# The Grothendieck-Teichmüller Group

Thomas Willwacher

February 27, 2014

# Contents

# Chapter 1

# Introduction

## Foreword

These are lecture notes (in progress) for a course held at ETH Zurich in fall 2012. The target audience are master students, advanced bachelor students, or doctoral students.

Prerequisites (... matter of discussion)

- category theory. groupoids. direct and inverse limit (co-limit and limit)
- topology / fundamental group
- groups, Hopf algebras
- commutative algebras, algebras and Lie algebras

## 1.1 Introduction

The Grothendieck-Teichmüller group is an important and somewhat mysterious object in algebra. It is important because it acts on a wide variety of other objects in many different fields of mathematics. It is mysterious because its structure and the relation to many of the objects it acts on is still unclear and a matter of ongoing research. In fact, there exist three different versions of the Grothendieck-Teichmüller group, a profinite version $\widehat{\mathsf{GT}}$, a pro-$l$ version $\mathrm{GT}_l$, and a pro-unipotent version GT. Roughly speaking, the former two versions are important in Galois theory and in the algebro-geometric context, while the latter version is "the correct" variant to be used in the homological algebra context. In this course we will study almost exclusively the pro-unipotent version, which is also the simplest of the three.

Historically however, the pro-finite version was invented first (by A. Grothendieck) and hence we start by giving a brief review.

### 1.1.1 Origins and the pro-finite version $\widehat{\mathsf{GT}}$

**Recollections from Galois theory**

Let $\mathbb{K}$ be a perfect field (e. g. of characteristic zero or finite). Let $\bar{\mathbb{K}}$ be the algebraic closure of $\mathbb{K}$. Then one defines the absolute Galois group of $\mathbb{K}$ to be

$$\mathrm{Gal}(\mathbb{K}) := \mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K}) := \mathrm{Aut}(\bar{\mathbb{K}}/\mathbb{K}) := \{\phi \mid \phi \text{ a field automorphisms of } \bar{\mathbb{K}}, \phi(x) = x \forall x \in \mathbb{K}\}.$$

For example $\mathrm{Gal}(\mathbb{R}) = \mathrm{Gal}(\mathbb{C}/\mathbb{R}) = \mathbb{Z}/2\mathbb{Z}$, the nontrivial element being complex conjugation. Of course, $\mathrm{Gal}(\mathbb{C}) = *$ trivially. The Artin-Schreier Theorem asserts that $*$ and $\mathbb{Z}_2$ are the only finite examples of absolute Galois groups.

**Remark 1.1.** Recall that a field extension $\mathbb{L}/\mathbb{K}$ is called algebraic if any element of $\mathbb{L}$ is a root of a polynomial with coefficients in $\mathbb{K}$. For example, $\mathbb{C}/\mathbb{Q}$ is not algebraic (because there exist transcendental numbers) and $\bar{\mathbb{Q}} \neq \mathbb{C}$.

In general $\mathrm{Gal}(\mathbb{K})$ has the structure of a pro-finite group.

**Definition 1.1.** *A topological group $G$ is* pro-finite *if it is homeomorphic to an inverse limit of finite groups $G \cong \lim_{\leftarrow} G_i$, where the finite groups $G_i$ are endowed with the discrete topology.*

**Remark 1.2.** Recall that elements of $\lim_{\leftarrow} G_i$ may be identified with collections of elements $g_i \in G_i$, such that for any arrow $G_i \to G_j$ in the inverse system $g_i \mapsto g_j$. The group structure is the obvious one. By definition there are maps $\pi_i : \lim_{\leftarrow} G_i \to G_i$. The topology on $\lim_{\leftarrow} G_i$ is the coarsest topology such that all $\pi_i$ are continuous, i. e., a basis for the topology s given by sets $\pi^{-1}(U_i)$ where $U_i \subset G_i$ is an arbitrary subset.

**Example 1.1.** Consider the inverse system

$$\mathbb{Z} \leftarrow \mathbb{Z}/p\mathbb{Z} \leftarrow \mathbb{Z}/p^2\mathbb{Z} \leftarrow \mathbb{Z}/p^3\mathbb{Z} \leftarrow \cdots .$$

Elements of $\mathbb{Z}_p := \lim_{\leftarrow} \mathbb{Z}/p^i\mathbb{Z}$ are called $p$-adic integers and may be identified with formal series of the form

$$\sum_{j=0}^{\infty} c_j p^j$$

where $c_j \in \mathbb{Z}/p\mathbb{Z}$.

**Example 1.2.** Consider the inverse system formed by groups $\mathbb{Z}/n\mathbb{Z}$, $n = 1, 2, 3, \dots$ with arrows $\mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ whenever $m \mid n$. The inverse limit may be seen to be

$$\hat{\mathbb{Z}} := \lim_{\leftarrow} \mathbb{Z}/n\mathbb{Z} = \{(a_n)_{n\in\mathbb{N}} \mid a_n \in \mathbb{Z}/n\mathbb{Z}; n \mid m \Rightarrow a_n \equiv a_m \mod n\} = \prod_{p \text{ prime}} \mathbb{Z}_p. \tag{1.1}$$

**Example 1.3.** Consider some (discrete) group $G$. Its pro-finite completion $\hat{G}$ is defined as

$$\hat{G} = \lim_{\leftarrow} G/G'$$

where the limit is over all normal subgroups $G' \subset G$ of finite index, i. e., such that $G/G'$ is finite. Exercise: Show that the notation is consistent with the one of the previous example.

Recall that a field extension $\mathbb{L}/\mathbb{K}$ of the perfect field $\mathbb{K}$ is a Galois extension if it is a smallest field extension over which some family of polynomials polynomial in $\mathbb{K}[X]$ splits into linear factors.[1] We will consider sub-extensions $\bar{\mathbb{K}}/\mathbb{L}/\mathbb{K}$, with $\mathbb{L}/\mathbb{K}$ finite Galois. In this case $\mathbb{L}$ is the smallest subfield of $\bar{\mathbb{K}}$ that contains all roots (in $\bar{\mathbb{K}}$) of the given polynomial(s). Clearly any element of $\mathrm{Gal}(\mathbb{K})$ fixes $\mathbb{L}$, but not necessarily pointwise. Similarly, if we have sub-extensions $\bar{\mathbb{K}}/\mathbb{L}'/\mathbb{L}/\mathbb{K}$ then $\mathrm{Gal}(\mathbb{L}'/\mathbb{K})(\mathbb{L}) = \mathbb{L}$. We may hence set up an inverse system with objects $\mathrm{Gal}(\mathbb{L}/\mathbb{K})$, one for each sub-extensions $\bar{\mathbb{K}}/\mathbb{L}/\mathbb{K}$, with $\mathbb{L}/\mathbb{K}$ finite Galois, and arrows $\mathrm{Gal}(\mathbb{L}'/\mathbb{K}) \to \mathrm{Gal}(\mathbb{L}/\mathbb{K})$ for each chain of extensions $\bar{\mathbb{K}}/\mathbb{L}'/\mathbb{L}/\mathbb{K}$.

**Lemma 1.1.** $\mathrm{Gal}(\mathbb{K}) = \lim_{\leftarrow} \mathrm{Gal}(\mathbb{L}/\mathbb{K})$ *as sets. In fact, one may endow* $\mathrm{Gal}(\mathbb{K})$ *with the topology from the right hand side. (This is then called Krull topology on* $\mathrm{Gal}(\mathbb{K})$*.)*

*Proof.* By what is said above there is a map $\mathrm{Gal}(\mathbb{K}) \to \mathrm{Gal}(\mathbb{L}/\mathbb{K})$ for each $\mathbb{L}$ and these maps are compatible with the arrows in the inverse system. Hence we have a map $f : \mathrm{Gal}(\mathbb{K}) \to \lim_{\leftarrow} \mathrm{Gal}(\mathbb{L}/\mathbb{K})$. We claim that this map is bijective. First note that every $x \in \bar{\mathbb{K}}$ is an element of some subfield $\mathbb{L}$ such that $\mathbb{L}/\mathbb{K}$ is (finite) Galois. Injectivity: Suppose $g \in \mathrm{Gal}(\mathbb{K})$ fixes pointwise all $\mathbb{L}$. Then it fixes all $x \in \bar{\mathbb{K}}$ and hence $g = id$. Surjectivity: Conversely, let $y = (g_{\mathbb{L}} \in \mathrm{Gal}(\mathbb{L}/\mathbb{K}))_{\mathbb{L}}$ be an element of the projective limit. Then define an element $g \in \mathrm{Gal}(\mathbb{K})$ by setting $g(x) := g_{\mathbb{L}}(x)$ where $\mathbb{L}$ is some sub-extension such that $x \in \mathbb{L}$. It is easy to see that $g$ is well-defined, is an element of $\mathrm{Gal}(\mathbb{K})$, and a preimage of $y$. $\square$

### Grothendieck's plan

One (or the most) important and mysterious example of an absolute Galois group is $\mathrm{Gal}(\mathbb{Q}) = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. The structure of this group is largely unknown. In fact, the author is not aware of any explicitly defined elements apart from complex conjugation and the identity. A. Grothendieck's idea [?] was to study $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ through its actions on objects that are easier to handle. (So ideally, find some simpler object $O$, with simpler to understand automorphism group $\mathrm{Aut}(O)$, and construct a map (ideally a bijective one) $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}(O)$.) Concretely, he proposed to study $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ through its (outer) action on algebraic fundamental groups of moduli spaces of curves.

[1] In general one requires the extension to be normal and separable. Here the extension is normal if $\mathbb{L}$ is the splitting field of some family of polynomials and separable if every irreducible polynomial has only distinct roots. That $\mathbb{K}$ is perfect means that every algebraic extension is separable.

**Digression: Action on algebraic fundamental groups**

Grothendieck defined a version of the fundamental group for any scheme $X$, the étale fundamentale group $\pi_1^{\text{ét}}(X)$.

The only properties we need are the following:

1. Let $X$ be a scheme over Spec($\mathbb{K}$) for some perfect field $\mathbb{K}$. Let $X_{\bar{K}} := X \times_{\text{Spec}(K)} \text{Spec}(\bar{K})$. Then there is an exact sequence
$$1 \to \pi_1^{\text{ét}}(X_{\bar{K}}) \to \pi_1^{\text{ét}}(X) \to \pi_1^{\text{ét}}(K) = \text{Gal}(\bar{\mathbb{K}}/\mathbb{K}) \to 1.$$

2. For $\mathbb{K} = \mathbb{Q}$ we have
$$\pi_1^{\text{ét}}(X_{\bar{K}}) = \pi_1^{\text{ét}}(X \times_{\text{Spec}(\mathbb{Q})} \text{Spec}(\mathbb{C})) = \hat{\pi}_1(X_{top})$$
where $\pi_1(X_{top})$ is the usual (topological) fundamental group and the hat denotes the pro-finite completion.

By the first property the adjoint action of $\pi_1^{\text{ét}}(X)$ on itself restricts to an action on $\pi_1^{\text{ét}}(X_{\bar{K}})$,
$$\pi_1^{\text{ét}}(X) \to \text{Aut}(\pi_1^{\text{ét}}(X_{\bar{K}})).$$
By definition $\pi_1^{\text{ét}}(X_{\bar{K}}) \subset \pi_1^{\text{ét}}(X)$ acts by inner automorphisms and we may pass to the quotient to obtain
$$\text{Gal}(\mathbb{K}) \to \text{Out}(\pi_1^{\text{ét}}(X))$$

**Grothendieck's plan (continued)**

In particular, consider the variety $X = \mathbb{P}^1 \setminus \{0, 1, \infty\}$, which is defined over $\mathbb{Q}$. Note that $\pi_1(X_{top}) = \mathsf{Free}(x, y)$ is the free group generated by two elements $x$ and $y$. Hence we obtain a morphism of groups
$$\phi : \text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \text{Out}(\widehat{\mathsf{Free}}(x, y)).$$
There is the following famous result.

**Theorem 1.1** (Belyi)**.** *The map $\phi$ is injective.*

Hence to study $\text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ we may study its image in $\text{Out}(\widehat{\mathsf{Free}}(x, y))$. It was realized by A. Grothendieck (but he did not write explicit formulas) that the image is contained in a subgroup characterised by some simple equations. Explicit formulas were written down by V. Drinfeld [**?**].

**The Grothendieck-Teichmüller group $\hat{\text{GT}}$**

Note first that for $a \in \hat{\mathbb{Z}}$ (say $a = (a_n)_n$ as in (1.1)) we may define elements $x^a \in \widehat{\mathsf{Free}}(x, y)$. To do this, it is sufficient to define for each finite index normal subgroup $G \subset \mathsf{Free}(x, y)$ the element $\pi_G(x^a)$ (in a way compatible with subgroup inclusion) where
$$\pi_G : \widehat{\mathsf{Free}}(x, y) \to \mathsf{Free}(x, y)/G.$$
Let $n$ be the order of the cyclic subgroup of the right hand side generated by $\pi_G(x)$. (I. e., $\pi_G(x)^n = id$.) We define $x^a$ by requiring $\pi_G(x^a) = \pi_G(x)^{a_n}$. Let us check that this is well defined. Let $H \subset G$ be a subgroup and suppose that the order of $\pi_H(x)$ is $m$. (Necessarily $n \mid m$.) Let
$$\pi_{G,H} : \mathsf{Free}(x, y)/H \to \mathsf{Free}(x, y)/G.$$
Then
$$\pi_{G,H}(\pi_H(x^a)) = \pi_{G,H}(\pi_H(x)^{a_n}) = \pi_{G,H}(\pi_H(x))^{a_n} = \pi_G(x)^{a_n} = \pi_G(x)^{a_m} = \pi_G(x^a).$$

**Definition 1.2** (V. Drinfeld)**.** *The (pro-finite version of the) Grothendieck-Teichmüller group is the subgroup $\widehat{\mathsf{GT}} \subset \text{Aut}(\widehat{\mathsf{Free}}(x, y))$ consisting of automorphisms $\phi$ that have the form*
$$\phi(x) = x^\lambda \qquad\qquad \phi(y) = f^{-1} y^\lambda f$$

*where $\lambda \in 1 + 2\hat{\mathbb{Z}}$ and $f \in \hat{\mathsf{Free}}(x,y)'$ satisfies*

$$f(y,x) = f(x,y)^{-1} \tag{1.2}$$
$$f(z,x)z^m f(y,z)y^m f(x,y)x^m = 1 \qquad (xyz = 1) \tag{1.3}$$
$$f(x_{12}, x_{23}x_{24})f(x_{13}x_{23}, x_{34}) = f(x_{23}, x_{34})f(x_{12}x_{13}, x_{24}x_{34})f(x_{12}, x_{23}) \tag{1.4}$$

*where the last equation takes place in the pro-finite completion of the pure braid group $\widehat{\mathsf{PB}}_4$, whose generators are $x_{ij}$, and where $m = (1-\lambda)/2$.*

**Remark 1.3.** One obtains similarly the pro-$l$ version $\mathrm{GT}_l$ of the Grothendieck-Teichmüller group (for any prime $l$) if one replaces all pro-finite completions by pro-$l$ completions. It comes with a map $\widehat{\mathsf{GT}} \to \mathrm{GT}_l$.

**Remark 1.4.** Note that elements of GT are pairs $(\lambda, f) \in \hat{\mathbb{Z}}^\times \times \widehat{\mathsf{Free}}(x,y)$ and that the latter object is also group. However, it is very important that the group structures are different.

**Theorem 1.2.** *The outer action of the absolute Galois group $\mathrm{Gal}(\bar{Q}/\mathbb{Q})$ factors uniquely through $\widehat{\mathsf{GT}}$, i. e., we have an injective map $\mathrm{Gal}(\bar{Q}/\mathbb{Q}) \to \widehat{\mathsf{GT}}$ fitting into a commutative diagram*

$$\begin{tikzcd}
& \widehat{\mathsf{GT}} \arrow[d] \\
\mathrm{Gal}(\mathbb{Q}) \arrow[r] \arrow[ur] & \mathrm{Out}(\widehat{\mathsf{Free}}(x,y))
\end{tikzcd}$$

One of the major problems in the field is to settle the following:

**Conjecture 1.1.** *The map $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \widehat{\mathsf{GT}}$ is an isomorphism.*

Several authors have added additional conditions to the above list, that are all satisfied by the image of $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ in $\widehat{\mathsf{GT}}$. However, the conjecture is still open.

**Remark 1.5.** Note that the pro-finite completion $\widehat{\mathsf{Free}}(x,y)$ is an unwieldy object. For example, any finite group with two generators will appear in the inverse system defining $\widehat{\mathsf{Free}}(x,y)$.

**Remark 1.6.** We simplified the discussion a little bit. Note that $\mathbb{P}^1 \setminus \{0,1,\infty\}$ may be identified with the moduli space $\mathcal{M}_{0,4}$ of curves of genus zero with four (distinguishable) marked points. The pro-finite completions of the fundamental groups of the moduli spaces $\mathcal{M}_{g,n}$ of arbitary genus $g$ and with arbitrarily many ($n$) marked points may be packaged into a "tower" of groups, the Teichmüller tower. The outer action of $\mathrm{Gal}(\mathbb{Q})$ extends to the full tower. (This is the origin of the second half of the name of GT.)

### 1.1.2 Actions of $\widehat{\mathsf{GT}}$

### 1.1.3 TODO

Concrete description for $\mathrm{GT}_l$?

Pure braid group?

## 1.2 The pro-unipotent version GT

One may define a pro-unipotent (to be defined later) analog of $\widehat{\mathsf{GT}}$.

**Definition 1.3.** *The (pro-unipotent version of the) Grothendieck-Teichmüller group $\mathrm{GT}(\mathbb{K})$ for $\mathbb{K}$ a field of characteristic 0, is the set of pairs $(\lambda, f) \in \mathbb{K}^\times \times \mathbb{K}\langle x,y\rangle$ such that*

$$\delta f = f \hat{\otimes} f \tag{1.5}$$
$$f(y,x) = f(x,y)^{-1} \tag{1.6}$$
$$f(z,x)z^m f(y,z)y^m f(x,y)x^m = 1 \tag{1.7}$$
$$f(x_{12}, x_{23}x_{24})f(x_{13}x_{23}, x_{34}) = f(x_{23}, x_{34})f(x_{12}x_{13}, x_{24}x_{34})f(x_{12}, x_{23}). \tag{1.8}$$

*Here the last equation takes place in the pro-unipotent completion of the pure braid group, $xyz = 1$ and $m = (\lambda - 1)/2$.*

*The group structure on* $\mathrm{GT}(\mathbb{K})$ *is defined by considering the pairs $(\lambda, f)$ as an automorphism of the pro-unipotent completion of the free group in two generators $x, y$ by setting*

$$x \mapsto x^\lambda \qquad\qquad y \mapsto f^{-1} y^\lambda f.$$

*Concretely, it is given by the equation*

$$(\lambda, f) \cdot (\lambda', f') = (\lambda\lambda', f F_f(f')).$$

**Remark 1.7.** In fact, there is a much simpler one-line definition of GT, but it requires knowledge of some further algebraic structures, so it is postponed a bit.

Note that there is a group homomorphism

$$\mathrm{GT}(\mathbb{K}) \to \mathbb{K}^\times$$
$$(\lambda, f) \to \lambda.$$

The kernel is denoted $\mathrm{GT}_1$ and is a pro-unipotent group.

Pro-unipotent groups are in general much easier to understand than pro-finite groups. In particular, we will see below that GT is isomorphic to a simpler group GRT, the "graded version of the Grothendieck-Teichmüller group". However, it is hard to to construct such isomorphisms, since one needs a Drinfeld associator for this task.

**Definition 1.4.** *Consider group-like elements $\Phi \in \mathbb{K}\langle\langle X, Y\rangle\rangle$. Consider the following set of equations, depending on $\mu \in \mathbb{K}$,*

$$\Phi(X, Y) = \Phi(Y, X)^{-1} \tag{1.9}$$
$$1 = e^{\frac{\mu}{2} Z} \Phi(X, Y) e^{\frac{\mu}{2} X} \Phi(Y, Z) e^{\frac{\mu}{2} Y} \Phi(Z, X) \tag{1.10}$$
$$\Phi(t_{12}, t_{23} + t_{24}) \Phi(t_{13} + t_{23}, t_{34}) = \Phi(t_{23}, t_{34}) \Phi(t_{12} + t_{13}, t_{24} + t_{34}) \Phi(t_{12}, t_{23}). \tag{1.11}$$

*Here for the middle equation $X + Y + Z = 0$ and the last equation takes values in the Drinfeld-Kohno Lie algebra $\mathfrak{t}_4$ with standard generators $t_{ij}$. The set of* Drinfeld associators DAss *is defined to be the set of pairs $(\mu, \Phi)$ solving these equations, with $\mu \neq 0$. The set of solutions $\Phi$ of these equations for $\mu = 0$ is called the* (graded version of the) Grothendieck-Teichmüller group $\mathrm{GRT}_1$. [2] *We will set*

$$\mathrm{GRT} := \mathbb{K}^\times \ltimes \mathrm{GRT}_1$$

*where the multiplicative group $\mathbb{K}^\times$ act on elements $\Phi \in \mathrm{GRT}_1$ by rescaling, i. e.,*

$$(\lambda \cdot \Phi)(X, Y) := \Phi(\lambda X, \lambda Y).$$

Later we will see the following results:

- There exist Drinfeld associators.
- The set of Drinfeld associators is a $\mathrm{GT} - \mathrm{GRT}$-torsor.
- In particular, it follows that $\mathrm{GT} \cong \mathrm{GRT}$.
- $\mathrm{GRT}_1$ and $\mathrm{GT}_1$ are pro-unipotent groups. We denote their Lie algebras by $\mathfrak{gt}_1$ and $\mathfrak{grt}_1$.
- $\mathfrak{grt}_1$ is graded and there are non-trivial element $\sigma_3, \sigma_5, \sigma_7, \cdots \in \mathfrak{grt}$ of degrees $3, 5, 7, \ldots$. In particular it follows that $\mathrm{GRT}_1$ and $\mathrm{GT}_1$ are infinite dimensional pro-affine varieties.

There is a famous conjecture about the structure of $\mathfrak{grt}$.

**Conjecture 1.2** (Deligne-Drinfeld-Ihara Conjecture)**.**

$$\mathfrak{grt} \cong \hat{\mathbb{F}}_{Lie}(\sigma_3, \sigma 5_, \cdots)$$

Unfortunately, we do not know how GRT/GT relate to their pro-finite cousin.

**Open Problem 1.1.** *Describe the precise relation between* $\widehat{\mathsf{GT}}$ *and* GT.

TODO: describe $\mathrm{GT}_l \to \mathrm{GT}(\mathbb{Q}_l)$???

[2] It carries a group structure that will be introduced later.

## 1.3 The role of GT and GRT in mathematics

The pro-unipotent version GT is a very important group in its own right and appears in many problems in a variety of fields in mathematics. Here we list a few:

### 1.3.1 Quantum groups

The original motivation of Drinfeld in defining the Grothendieck-Teichmüller group was to construct quasi-Hopf algebras. We will not discuss this, but we might discuss a similar result:

An important problem in quantum groups is the quantization of Lie bialgebras, i. e., the passage from a Lie bialgebra to a Hopf algebra, which is then considered to be the "quantum universal enveloping algebra" of the bialgebra. This quantization problem has been solved by P. Etingof and D. Kazhdan, and the essential ingredient of the solution is a Drinfeld associator.

### 1.3.2 Multiple zeta values

Zeta values are the numbers

$$\zeta(n) = \sum_{j\geq 1} \frac{1}{j^n}.$$

A result of Euler is that $\zeta(2n) = b_n\pi^{2n}$ for some $b_n \in \mathbb{Q}$. A conjecture (with no hope of settling it) is that the numbers $\pi, \zeta(3), \zeta(5), \dots$ are algebraically independent. If true, the algebra generated by the zeta values is "uninteresting" algebraically. However, currently one only knows that $\pi$ is transcendental and that $\zeta(3) \notin \mathbb{Q}$ (Apéry).

Multiple zeta values (MZVs) are the numbers

$$\zeta(n_1, n_2, \dots, n_k) = \sum_{j_1>j_2>\cdots>j_1\geq 1} \frac{1}{j_1^{n}j_2^{n_2}\cdots j_k^{n_k}}.$$

The multiple zeta values satisfy two families of combinatorial shuffle relation. These are referred to as double shuffle, or as shuffle and stuffle relations. Let us only consider one example of the stuffle relations here and leave the more detailed description and the shuffle relations for later.

**Exercise 1.1.**

$$\zeta(2,1)\zeta(4) = ... = \zeta(4,2,1) + \zeta(6,1) + \zeta(2,4,1) + \zeta(2,5) + \zeta(2,1,4)$$

It is conjectured that (regularized version of) these double shuffle relations generate all algebraic relations among the numbers $\zeta(n_1, n_2, \dots, n_k)$. This conjecture is equally hopeless as the previous. But one may at least ask what the algebraic structure of the algebra of MZV's is if the conjecture was true.

The relation to associators is as follows: For one particular associator (the Knizhnik-Zamolodchikov associator) all multiple zeta values appear as coefficients, more precisely the coefficient of

$$X^{n_1}YX^{n_2}Y\cdots X^{n_k}Y$$

is $\pm\zeta(n_1+1, n_2+1, \dots, n_k+1)$. Furthermore, it has been shown [?] that any associator gives a solution of the double shuffle relations. Conjecturally, the associators parameterize all solutions of the double shuffle relations. If this and the previous conjecture were true, then the algebra of (regularized) multiple zeta values could be identified with the algebra of functions on the set of associators (over $\mathbb{Q}$).

### 1.3.3 Lie theory: the KV problem

The Kashiwara-Vergne conjecture, proven by Alekseev and Meinrenken, is one of the few general Theorems about Lie algebras valid for all Lie algebras. It states (essentially) that there exists an automorphism $\phi$ of the free Lie algebra in two generators $\mathbb{F}_{Lie}(X,Y)$ that trivializes the Baker-Campbell-Hausdorff formula,

$$\phi(\mathrm{BCH}(X,Y)) = X + Y$$

and furthermore satisfies some suitable conditions. Any Drinfeld associator gives a solution of this KV problem and it is conjectured that the solutions thus found exhaust the space of solutions.

### 1.3.4 Knot invariants

A *knot invariant* is a function on the set of knots (i. e., on the space of embeddings $S^1 \to \mathbb{R}^3$) that is invariant under changing the knot by isotopy. Any such knot invariant may be extended to an invariant of singular knots, which are given by maps $S^1 \to \mathbb{R}^3$ with a finite number of self-crossings, by applying Vassiliev's "Skein relations"

.

A knot invariant is a Vassiliev invariant of type $\leq n$ if it vanishes on all singular knots with $> n$ self crossings. The Vassiliev invariants are at least as strong as the known knot polynomials (Jones, Alexander, HOMFLY etc.). Furthermore, one has the following conjecture

**Conjecture 1.3.** *Any two (non-isotopic) knots may be distinguished by a Vassiliev invariant.*

Given a Drinfeld associator one may construct a universal Vasiliev invariant, i. e., an invariant which is as strong as all Vassiliev invariants together. (This is the same as the one given by the Kontsevich integral.) Concretely, one may build an isotopy invariant map from the set of of singular knots (say $K$) to a certain algebra, the algebra of chord diagrams (say $\mathcal{A}$)

$$K \to \mathcal{A}$$

given any Drinfeld associator.

### 1.3.5 Relation to CFT

One of the two known explicit constructions of Drinfeld associators has its origin in Conformal Field Theory, and comes out of the Knizhnik-Zamolodchikov equations.

### 1.3.6 Deformation quantization and automorphisms of polyvector fields

In deformation quantization one studies the origins of quantum mechanics. Concretely, the basic question is whether for any Poisson manifold $M$ there is a non-commutative product $\star$ on the space of functions ("observables") $C^\infty(M)[[\hbar]]$. One can show that giving a universal solution of this problem (a universal star product) is equivalent to providing a Drinfeld associator.

### 1.3.7 Graph cohomology

The graph complex is a combinatorial complex build using graphs, the differential being vertex splitting. Computing its cohomology is a purely combinatorial problem. (Though this is essentially the stable Chevalley cohomology of the Lie algebras $\mathbb{K}[x_1, \dots, x_n, \xi_1, \dots, \xi_n]$ with Lie bracket determined by $[\xi_i, x_j] = \delta_{ij}$.) The zeroth cohomology of the graph complex may be shown to be isomorphic to $\mathfrak{grt}$.

### 1.3.8 Automorphisms of the $E_2$ operad

One may show that the set of Drinfeld associators is (essentially) in 1-1 correspondence to the set of formality morphisms of the (operad of chains of the) little disks operad, and object that is famous in topology, and also appears in many algebraic problems.

**Remark 1.8.** In this course we focus on studying GRT.

## 1.4 Tentative structure of the course

The course will be roughly divided into three parts.

1. In the first part, we will recall some standard notions and results from algebra, including
   - Associative and Lie algebras
   - Topological vector spaces
   - Categories and (co-)limit constructions
   - Operads

- (Pro-)unipotent groups and pro-unipotent completions

Experts may skip this part.

2. In the second part we will give a more compact definition of GT and GRT and show their basic properties stated above. In particular, we will discuss the construction of the Knizhnik-Zamolodchikov associator, and that of the Alekseev-Torossian associator. If time permits, we will also discuss F. Brown's recent result on the structure of $\mathfrak{grt}$.

3. In the third part we will discuss the role of GRT, or its conjectural role in mathematics. The lecturer's wish is to cover all topics mentioned above, but realistically speaking certain choices have to be made.

Other topics on the wish-list, a bit separate from the main storyline.

- Operads and homotopy algebras
- Rational Homotopy Theory
- Simplicial methods
- The free Lie algebra
- (maybe: Model categories, but this is not really important)

## 1.5 Literature

At the time of writing of these notes, there is unfortunately no complete and good textbook on the subject available. (Which is the main reason for the lecturer to write these notes.) A very good reference on the subject is Drinfeld's seminal paper [?]. Other introductory papers can be found on the web, though they mostly cover the pro-finite version $\widehat{\mathrm{GT}}$.

A book draft by B. Fresse is available online. He also uses the operadic approach to the definition of GT, and is hence quite close to this course.

# Chapter 2

# Prerequisites

## 2.1 Categories and limits

The standard reference here is [**?**].

**Definition 2.1.** *A category $\mathcal{C}$ is the following data:*

- *A collection of* objects $\mathrm{Ob}\mathcal{C}$*.*
- *A collection of* morphisms *(or arrows) $Mor(\mathcal{C})$.*
- *Two maps $dom, codom : Mor(\mathcal{C}) \to \mathrm{Ob}\mathcal{C}$. Notation: We will write $f : A \to B$ if $dom(f) = A$, $codom(f) = B$, and $\mathrm{Hom}_{\mathcal{C}}(A, B)$ for the collection of all such morphisms.*
- *For every three objects $A, B, C$, a map*

$$\circ : \mathrm{Hom}(B, C) \times \mathrm{Hom}(A, B) \to \mathrm{Hom}(A, C)$$

*(the* composition map*).*

*These data are required to satisfy*

- *Associativity. For all four objects $A, B, C, D$ and morphisms $f : A \to B$, $g : B \to C$, $h : C \to D$*

$$(h \circ g) \circ f = h \circ (g \circ f)$$

- *For every $A \in \mathrm{Ob}\mathcal{C}$, there is a distinguished morphism $id_A \in \mathrm{Hom}(A, A)$ such that for all morphisms $f : A \to B$,*

$$f \circ id_A = id_B \circ f.$$

*For $\mathcal{C}, \mathcal{D}$ categories a functor $F : \mathcal{C} \to \mathcal{D}$ is (i) a map $\mathrm{Ob}\mathcal{C} \to \mathrm{Ob}\mathcal{D}$ and (ii) a map $mor\mathcal{C} \to mor\mathcal{D}$ that is compatible with the above data (in the obvious sense).*

**Definition 2.2.** *A category $\mathcal{C}$ is called small if $\mathrm{Ob}\mathcal{C}$ is a set and locally small if for any pair of objects $A, B$, $\mathrm{Hom}_{\mathcal{C}}(A, B)$ is a set.*

**Definition 2.3.** *Let $\mathcal{C}, \mathcal{D}$ be categories. For two functors $F, G : \mathcal{C} \to \mathcal{D}$ a natural transformation $F \Rightarrow G$ is a map $alpha : \mathrm{Ob}\mathcal{C} \to mor\mathcal{D}$ $(A \mapsto \alpha_A)$ such that $\alpha_A : F(A) \to G(A)$ for all $A$ and for all morphisms $f : A \to B$ in $\mathcal{C}$ we have that*

$$\alpha_B \circ Ff = Gf \circ \alpha_A.$$

**Example 2.1.**

- The category Set of sets. Objects are the sets and morphisms are arbitrary maps between sets.
- The category Vect of ($\mathbb{K}$-)vector spaces.
- There is a functor $F : \mathrm{Vect} \to \mathrm{Set}$ forgetting the vector space structure.
- There is a functor $G : \mathrm{Set} \to \mathrm{Vect}$ assigning to a set $S$ the free vector space in that set. (It has a basis $\{e_s\}_{s \in S}$labelled by elements in the set.)

- There is a natural transformation $\alpha : GF \Rightarrow Id$ such that $\alpha_V : GF(V) \to V$ sends $e_v \to v$ for each $v \in V$.
- If $\mathcal{C}$ is a category then the opposite category $\mathcal{C}^{op}$ has the same objects and morphisms as $\mathcal{C}$, but the the domain (codomain) map of $\mathcal{C}^{op}$ is the codomain (domain) map of $\mathcal{C}$, and one flips arguments in the composition map.

Let us recall the notion of limits and colimits in a category.

**Definition 2.4.** *Let $J, \mathcal{C}$ be categories. A diagram of type $J$ in $\mathcal{C}$ is a functor $F : J \to \mathcal{C}$.*

*A* cone *for $F$ is an object $L$ of $\mathcal{C}$ together with maps $\pi_A : L \to F(A)$ for each object $A$ in* Ob$J$*, such that for every morphism $f : A \to B$ in $J$*

$$\pi_B \circ f = \pi_A.$$

*A (the) limit of $F$ is a cone $(L, \pi)$ such that for any other $(L', \pi')$ there is a unique morphism $f : L' \to L$ such that for all $A \in \mathrm{Ob}J$ we have*

$$\pi'_A = \pi_A \circ f.$$

*The colimit is the dual notion (reverse all arrows).*

**Example 2.2.**

1. Let $\mathcal{C}$ be a (small) category with only the identity morphisms. A functor $F : \mathcal{C} \to \mathcal{D}$ is given by picking a collection of objects in $\mathcal{D}$. Then the limit of $F$ is called the product of these objects and the colimit is called coproduct. For example, in $\mathcal{D} = \mathrm{Set}$ the product is the Cartesian product and the coproduct is the union.
2. Let $J = \emptyset$. Then there is a unique functor $J \to \mathcal{C}$ and a limit (colimit) is called terminal object (initial object) of $\mathcal{C}$.
3. If $J$ has the form $\cdot \rightrightarrows \cdot$, then the limit is called equalizer (and the colimit coequalizer).
4. If $J$ has the form $\cdot \to \cdot \leftarrow \cdot$, then the limit is called pullback.

The limits we will need have a very special form.

**Definition 2.5.** *A directed set is a set $J$ together with a transitive and reflexive relation $\leq$ (i.e., $A \leq B$ and $B \leq C$ implies $A \leq C$ and $A \leq A$) such that for any two objects $A, B \in J$ there is some object $C$ such that $A \leq C$ and $B \leq C$.*

*We will understand $J$ as a category, also denoted $J$, with objects $J$ and morphisms*

$$\mathrm{Hom}_J(A, B) = \begin{cases} \{*\} & \text{if } A \leq B \\ \emptyset & \text{otherwise.} \end{cases}$$

*If $\mathcal{C}$ is a category, than a limit of a diagram of type $J^{op}$, $F : J^{op} \to \mathcal{C}$ is* inverse limit*. Dually, the colimit of a diagram of type $J$ is called direct limit.*

We will use the notation $\lim_{\leftarrow}$ for the inverse limit and $\lim_{\rightarrow}$ for the direct limit.

**Example 2.3.** The canonical example is the following: Let $J$ be the directed set $\mathbb{N}$. Let $\mathcal{C} = \mathrm{Vect}$ be the category of vector spaces for concreteness. A functor $F : J^{op} \to \mathcal{C}$ is a collection of vector spaces $V_j$, $j = 1, 2, \ldots$ together with some arrows $f_j : V_j \to V_{j-1}$. The inverse limit of that functor

$$\lim_{\leftarrow} V_j$$

is the vector space $V \subset \prod_j V_j$, formed by the elements $(v_1, v_2, \ldots) \in \prod_j V_j$ that satisfy $f_j(v_j) = v_{j-1}$.

**Example 2.4.** If in the setting of the previous example the morphisms were going in the other direction, i. e., $f_j : V_j \to V_{j+1}$, then may form the direct limit (i.e., the colimit) of $F : \mathcal{C} \to \mathrm{Vect}$

$$\lim_{\rightarrow} V_j = (\bigoplus_j V_j)/\sim$$

where the equivalence relation is obtained by declaring that for all $j$ and for all $v \in V_j$

$$f_j(v) \sim v.$$

(In other words, one quotients by the sub-vector space spanned by $f_j(v) - v$ $j = 1, 2, \ldots, v \in V_j$.)

**Exercise 2.1.** Verify that the universal property holds in all cases above.

Without proof we state:

**Theorem 2.1.** *If $\mathcal{C}$ has equalizers and small products, then $\mathcal{C}$ has all small limits. (i.e., limits of diagrams of type $J$, $J$ small).*

## 2.2 Topological vector spaces

**Example 2.5.** Let $\mathbb{K}[X]$ be the space of polynomials in one indeterminate ($X$). Let $\mathbb{K}[[X]]$ be the space of power series in $X$. For example,

$$p = 1 + X + 2X^2 + 3X^3 + \cdots$$

is an element of $\mathbb{K}[[X]]$, but not of $\mathbb{K}[X]$. However,

$$p_n = 1 + X + 2X^2 + 3X^3 + \cdots + nX^n \in \mathbb{K}[X] \subset \mathbb{K}[[X]].$$

In this section we would like to introduce notation to make precise the following statements: (i) $p_n$ converges to $p$ as $n \to \infty$. (ii) $\mathbb{K}[[X]]$ is complete, and in fact the completion of $\mathbb{K}[X]$.

**Definition 2.6.** *A topological space is a set $X$ with subset $T$ of the power set of $X$ such that*

- $\emptyset, X \in T$.
- *For all $T' \subset T$, $\cup_{U \in T'} U \in T$.*
- *For all finite $T' \subset T$, $\cap_{U \in T'} U \in T$.*

*$T$ is called topology, elements of $T$ are called open sets, and complements of elements closed sets.*

*A continuous map between topological spaces $(X,T)$ and $(Y,R)$ is a map $f : X \to Y$ such that $f^{-1}U \in T$ for all $U \in R$. Such an $f$ is called homeomorphism if it is bijective and the inverse is also continuous.*

**Example 2.6.** For any set $X$ there are two "stupid" topologies, namely $T = \{\emptyset, X\}$, and the discrete topology $T = 2^X$.

**Definition 2.7.** *For $\mathbb{K}$ a topological field a topological $\mathbb{K}$-vector space is a $\mathbb{K}$-vector space endowed with a topology such that the scalar multiplication and the addition are continuous.*

In the "classical examples" $\mathbb{K} = \mathbb{R}, \mathbb{C}$ with the standard topology and the topology on the vector space is given by a metric, e. g., from a norm or inner product. Our examples however will look a bit different. For us, $\mathbb{K}$ is always endowed with the discrete topology. (For $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ for example this means that $\frac{1}{n}$ does not converge to 0.)

A topology $T$ on a set $X$ is always determined by giving a basis $B$ for the topology, i. e., a subset $B \subset T$ such that any open set $U \in T$ may be written as a union of sets in $B$. The topological space $X$ is called second countable if its topology has a countable basis.

An open neighborhood of a point $x \in X$ is an open set containing $x$. A neighborhood of a point $x \in X$ is a set containing an open neighborhood of $x$. A neighborhood basis at $x$ is a subset $B$ of the set of neighborhoods such that for any neighborhood $N$ of $x$ there is some $U \in B$ with $U \subset N$. $X$ is called first countable if every point has a countable neighborhood basis. Giving a neighborhood basis at every point determines the topology uniquely. (Define $U \subset X$ to be open iff for all $x \in U$ there is an element $U_x$ of the neighborhood basis at $x$ such that $U_x \subset U$.)

If $T$ is a topological vector space, it even suffices to specify a neighbourhood basis $B$ of $0 \in T$. For any other point $x$ we then declare $\{U + x \mid U \in B\}$ to be a neighborhood basis of $x$.

**Example 2.7.** For $\mathbb{K}[X]$ ($\mathbb{K}[[X]]$) a topology is given by declaring the sets $x^N\mathbb{K}[X]$ ($x^N\mathbb{K}[[X]]$) ($N = 0, 1, 2, \dots$) to be a neighbourhood basis of 0.

**Definition 2.8.** *A net in a topological space $X$ is a directed set $J$ together with a map $J \to X$. The net is called sequence if $J = \mathbb{N}$. We denote the net by $(x_j)_{j \in J}$, understanding that $j$ gets mapped to $x_j \in X$.*

*The net $(x_j)_{j \in J}$ converges to $x \in X$,if for any neighborhood $U$ of $x$ there is a $j \in J$ such that $x_k \in U$ for all $k \in J$ such that $j \leq k$. We also call $x$* limit point *of the net $(x_j)_{j \in J}$.*

**Exercise 2.2.** Check that for the topology of example 2.7 $\lim_{n\to\infty} p_n \to p$, where $p_n, p$ are as in example 2.5.

**Remark 2.1.** A subset $X' \subset X$ of a topological space $X$ is closed if for every net $(x_j)_{j\in J}$, with all $x_j \in X'$, and any limit point $x \in X$, we have $x \in X'$.

A map $f : X \to Y$ of topological spaces is continuous iff for any $x \in X$ and any net $(x_j)_j$ converging to $x$, $(f(x_j))_j$ converges to $f(x)$.

For first countable spaces one may replace "net" by "sequence".

**Remark 2.2.** In general, limits are not unique. However, if the topological space is Hausdorff, i.e., if any two distinct points have disjoint neighborhoods, limits are unique, if they exist. The converse also holds, since if any two neighborhoods $U \ni x \in X$ and $V \ni y \in X$ have non-empty intersection, then the net $(z_{U,V})$ where for every such neighborhoods we pick some $z_{U,V} \in U \cap V$ using the axiom of choice, converges to both $x$ and $y$.

We will write $\lim_j x_j := x$ if $x$ is the unique limit point of the net $x_j$.

**Definition 2.9.** *A Cauchy net in a topological vector space $X$ is a net $(x_j \in X)_j$ such that for all neighbourhoods $U$ of 0 there is an $N \in J$ such that $t_n - t_{n'} \in U$ for all $n, n' \geq N$.*

*A (first countable) topological space $X$ is called complete if every Cauchy net (sequence) in $X$ has a limit point.*

All important spaces in this course will be first countable.

**Exercise 2.3.** Show that $\mathbb{K}[[X]]$ is complete and $\mathbb{K}[X]$ is not.

**Exercise 2.4.** Show that the image of a Cauchy net under a continuous map is again Cauchy.

The most important example for us is the following.

**Example 2.8.** Let $V$ be a vector space that is an inverse limit of (ordinary) vector spaces,

$$V = \lim_{\leftarrow} V_\alpha.$$

Then we can endow $V$ with the *inverse limit topology*, which is defined by declaring the sets $\pi_\alpha^{-1}\{0\}$ (where $\pi_\alpha : V \to V_\alpha$ are the canonical projections) to be a neighbourhood basis for 0.

**Exercise 2.5.** Show that a topological vector space $V$ as in the previous example (i.e., equipped with the inverse limit topology) is complete.

**Solution 2.1.**

**Exercise 2.6.** Show that

$$K[[X]] \cong \lim_{\leftarrow} K[X]/x^N K[X]$$

The closure $\bar{U}$ of some subset $U$ of a topological space $X$ is the intersection of all closed sets containing $U$. $U \subset X$ is called dense if $\bar{U} = X$.

**Definition 2.10.** *A complete Hausdorff topological vector space $Y$ is called completion of another Hausdorff topological vector space $X$, if $X$ is homeomorphic to a dense subspace of $Y$.*

**Remark 2.3.** One may omit the Hausdorff-ness condition, but then the completion will be ill-behaved, e. g., it will not be unique.

**Example 2.9.** $K[[X]]$ is the completion of $K[x]$.

**Remark 2.4.** For $X$ Hausdorff, any completion of $X$ as above is the universal complete Hausdorff topological vector space $Y$ with a continuous map $X \to Y$, i. e., for any other Hausdorff complete topological vector space $Y'$ together with a continuous map $f : X \to Y'$, $f$ uniquely extends to a continuous map $\bar{f} : Y \to Y'$.

In particular, the completion is unique up to unique homeomorphism.

*Proof.* For any $y \in Y$ pick a net $(x_j)_{j\in J}$ converging to $y$. It is Cauchy automatically. Hence $f(x_j)$ is a Cauchy net in $Y'$, and hence has a (unique by Hausdorff-ness) limit $z$. We will define $\bar{f}(y) = z$.

$\bar{f}$ is well defined: Let $(x'_k)_{k\in K}$ be any other net converging to $y$. Then the net $(\tilde{x}_\alpha)_{\alpha\in K\times J\times\{1,2\}}$ with $\tilde{x}_{j,k,1} = x_j$ and $\tilde{x}_{j,k,2} = x'_k$ converges to $y$. Hence $f(\tilde{x}_\alpha)$ converges. Hence any sub-nets converge to the same limit point.

$\bar{f}$ is continuous: The same proof shows that $\bar{f}$ commutes with taking limits of convergent nets. □

Another ingredient we will need is the topological (projective) tensor product.

**Example 2.10.** Note that $\mathbb{K}[x]\otimes\mathbb{K}[y] \cong \mathbb{K}[x,y]$. However $\mathbb{K}[[x]]\otimes\mathbb{K}[[y]] \neq \mathbb{K}[[x,y]]$. We want to define a variant of the tensor product such that $\mathbb{K}[[x]]\hat{\otimes}\mathbb{K}[[y]] \cong \mathbb{K}[[x,y]]$.

**Definition 2.11.** *Let $V = \lim_\leftarrow V_\alpha$ and $W = \lim_\leftarrow W_\beta$ be as in example 2.8 (here $\alpha \in A, \beta \in B$). Then we will set*

$$V\hat{\otimes}W = \lim_{\leftarrow} V_\alpha \otimes W_\beta.$$

*We will call $V\hat{\otimes}W$ the projectively completed tensor product of $V$ and $W$.*

**Example 2.11.** For $V = W = K[[x]]$, we obtain

$$V\otimes W = \lim_{\leftarrow} K[x]/x^nK[x]\otimes K[y]/y^mK[y] = \lim_{\leftarrow} K[x,y]/((x^nK[x,y]+y^mK[x,y]) \cong K[[x,y]].$$

**Remark 2.5.** $V\hat{\otimes}W$ as in the definition above is the completion of $V\otimes W$, if we endow the latter space with the finest topology that makes $V\times W \to V\otimes W$ continuous.

To see this, first note that $V\otimes W \hookrightarrow V\hat{\otimes}W$ is a homeomorphism onto its image. Also it is clear that $V\hat{\otimes}W$ is complete. We still need to show that $V\otimes W$ is dense.

Reduction: It is sufficient to check that the image of $V\hat{\otimes}W$ in $V_\alpha\otimes W_\beta$ agrees with the image of $V\otimes W$, for all $\alpha,\beta$. Indeed, if this is true then for each $u\in V\hat{\otimes}W$ the net $(v_{\alpha,\beta})_{\alpha,\beta}$, where $v_{\alpha,\beta}\in V\otimes W$ is chosen such that $\pi_{alpha,\beta}(u - v_{\alpha,\beta}) = 0$, converges to $u$.

Let $u\in V\hat{\otimes}W$ and set $\pi_{alpha,\beta}(u) = \sum_j v_j\otimes w_j$. We may assume w.l.o.g. that the $v_j$ and the $w_j$ are linearly independent. Pick dual vectors $\tilde{v}_j$, $\tilde{w}_j$. Define $f_j$ to be the composition

$$V\hat{\otimes}W \to V_\alpha\otimes W \xrightarrow{\tilde{v}_j} W$$

and similarly define $g_j : V\hat{\otimes}W \to V$. Then we claim that $\pi_{alpha,\beta}(u - v_{\alpha,\beta}) = 0$, where $v_{\alpha,\beta} = \sum_j g_j(u)\otimes f_j(u)$. Indeed,

$$\pi_{alpha,\beta}v_{\alpha,\beta} = \sum_j \pi_\alpha g_j(u)\otimes\pi_\beta f_j(u) = \sum_j v_j\otimes w_j = \pi_{alpha,\beta}(u).$$

## 2.3 Associative algebras

**Definition 2.12.** *A* non-unital (associative) ($\mathbb{K}$-)algebra *is a $\mathbb{K}$-vector space $A$ together with a bilinear operation*

$$\mu\colon A\times A\to A$$

*such that $\forall x,y,z\in A$*

$$\mu(x,\mu(y,z)) = \mu(\mu(x,y),z) \qquad \textit{(associativity).}$$

*A (unital) associative algebra is a non-unital associative algebra $A$ together with a distinguished element $\mathbb{1}\in A$ such that $\forall x\in A$*

$$\mu(\mathbb{1},x) = \mu(x,\mathbb{1}) = x.$$

We will write (as usual) $xy = \mu(x,y)$ in the future.

**Definition 2.13.** *An ideal in the algebra $A$ is a sub-vectorspace $I\subset A$ closed under multiplication by elements of $A$. In this case $A/I$ is again an algebra.*

Furthermore, powers of the ideal $I^n$ are again ideals. One may put a topology on $A$ by declaring that $I^n$ form a neighborhood basis of 0.

**Definition 2.14.** *An augmentation is an algebra map $A\to\mathbb{K}$. The kernel is called the augmentation ideal.*

**Example 2.12.** $\mathbb{K}[x]$ is augmented, the augmentation being evaluation at $x = 0$. $\mathbb{K}[[x]]$ is the completion of $\mathbb{K}[x]$ with respect to (the topology generated by) the augmentation ideal.

## 2.4 Lie algebras

**Definition 2.15.** *A* Lie algebra *is a vector space* $\mathfrak{g}$ *together with a bilinear anti-symmetric operation (the Lie bracket)*

$$[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

*such that* $\forall x, y, z \in \mathfrak{g}$

$$[[x,y],z] + [[y,z],x] + [[z,x],y] = 0 \qquad \textit{(Jacobi identity).}$$

*A* Lie algebra morphism *between Lie algebras* $\mathfrak{g}$ *and* $\mathfrak{h}$ *is a map of vector spaces* $f : \mathfrak{g} \to \mathfrak{h}$ *that respects the Lie bracket, i. e.,* $f([x,y]) = [f(x), f(y)]\,\forall x, y \in \mathfrak{g}$.

**Example 2.13.**
- Any vector space is a Lie algebra with the trivial (zero) bracket. It is then called an Abelian Lie algebra.
- The Lie algebra $\mathfrak{su}(2) \cong \mathfrak{so}(3) \cong \mathbb{R}^3$ has bracket $[x,y] = x \times y$ where $\times$ is the cross product.

**Exercise 2.7.** Show that any associative algebra is a Lie algebra (with the commutator bracket). Show that any non-associative algebra whose associator is symmetric in the last two entries is a Lie algebra in the same way. Such an algebra is also called *pre-Lie algebra.* (Note: the associator is $A(x,y,z) := (xy)z - x(yz)$.)

For us, a important examples will be given by the free Lie algebras.

**Definition 2.16.** *Let* $S$ *be a set. Then the free Lie algebra is the (unique up to unique isomorphism) Lie algebra* $\mathbb{F}_{Lie}(S)$*, together with a map of sets* $S \to \mathbb{F}_{Lie}(S)$ *such that for each Lie algebra* $\mathfrak{g}$ *and map (of sets)* $S \to \mathfrak{g}$ *there is a unique Lie algebra morphism* $\mathbb{F}_{Lie}(S) \to \mathfrak{g}$ *such that the following diagram commutes*

$$\begin{tikzcd} S \arrow[d] \arrow[dr] & \\ \mathbb{F}_{Lie}(S) \arrow[r, dashed, "\exists!"] & \mathfrak{g} \end{tikzcd}$$

A *Lie word* in some symbols $X_1, \dots X_n$ is a formal bracketing of these symbols, e. g.,

$$[X_1, [X_2, X_1]].$$

The free Lie algebra $\mathbb{F}_{Lie}(S)$ is the vector space spanned by all Lie words in the symbol set $S$, modulo the subspace obtained by applying the antisymmetry and Jacobi identity. Unless $|S| \leq 1$ $\mathbb{F}_{Lie}(S)$ is infinite dimensional.

**Remark 2.6.** $\mathbb{F}_{Lie}(S)$ has a natural grading by the length of Lie words (i. e., the number of brackets used plus one),

$$\mathbb{F}_{Lie}(S) = \bigoplus_{n \geq 1} \mathbb{F}_{Lie}(S)_n$$

where $\mathbb{F}_{Lie}(S)_n$ is spanned by Lie words with $n-1$ brackets. Clearly

$$[\mathbb{F}_{Lie}(S)_n, \mathbb{F}_{Lie}(S)_m] \subset \mathbb{F}_{Lie}(S)_{n+m}.$$

**Example 2.14.** Consider $\mathbb{F}_{Lie}(X,Y) := \mathbb{F}_{Lie}(\{X,Y\})$. Bases for the first few graded subspaces are:

$$\begin{aligned} \mathbb{F}_{Lie}(X,Y)_1 &= \langle X, Y \rangle \\ \mathbb{F}_{Lie}(X,Y)_2 &= \langle [X,Y] \rangle \\ \mathbb{F}_{Lie}(X,Y)_3 &= \langle [X,[X,Y]], [Y,[Y,X]] \rangle \end{aligned}$$

**Exercise 2.8.** What is $\dim(\mathbb{F}_{Lie}(X,Y)_4)$?

**Definition 2.17.** *The universal enveloping algebra* $U(\mathfrak{g})$ *is the (unique up to unique isomorphism) algebra together with a map (of Lie algebras)* $\mathfrak{g} \to U(\mathfrak{g})$ *such that for any algebra* $A$ *together with a map* $\mathfrak{g} \to A$ *there is a unique map of algebras* $U(\mathfrak{g}) \to A$ *such that*

$$\begin{tikzcd} \mathfrak{g} \arrow[d] \arrow[dr] & \\ U(\mathfrak{g}) \arrow[r, dashed, "\exists!"] & A \end{tikzcd}$$

*commutes.*

Concretely, $U(\mathfrak{g}) = T\mathfrak{g}/I$ where $T\mathfrak{g}$ is the tensor algebra and $I$ is the two-sided ideal generates by the relations $x \otimes y - y \otimes x - [x, y]$.

**Example 2.15.** For $\mathfrak{g} = \mathbb{F}_{Lie}(S)$, $U(\mathfrak{g}) = \mathbb{K}\langle S\rangle$ is the free associative algebra in symbols $S$. A basis is given by words in symbols $S$.

**Theorem 2.2** (Poincaré-Birkhoff-Witt)**.** *The symmetrization map $S\mathfrak{g} \to U(\mathfrak{g})$ is an isomorphism of vector spaces for any Lie algebra $\mathfrak{g}$.*

*Proof.* $T\mathfrak{g}$ is graded (by the length of words) and hence $U(\mathfrak{g}) = T\mathfrak{g}/I$ inherits a filtraton. The associated graded is $S\mathfrak{g}$. □

**Corollary 2.1.** *The dimension of $\mathbb{F}_{Lie}(S)_n$ where $|S| = N$ is*

$$\dim \mathbb{F}_{Lie}(S)_n = \frac{1}{n} \sum_{d|n} \mu(\frac{n}{d}) N^d$$

*where*

$$\mu(k) = \begin{cases} 1 & k \text{ square free with an even number of prime factors} \\ -1 & k \text{ square free with an odd number of prime factors} \\ 0 & \text{otherwise} \end{cases}$$

*is the Möbius function*

*Proof.* define the generating function $F(t) = \sum_n t^n \dim \mathbb{K}\langle S\rangle_n = \frac{1}{1-Nt}$ and abbreviate $a_n = \dim \mathbb{F}_{Lie}(S)_n$. By Poincaré-Birkhoff-Witt we have

$$\frac{1}{1 - Nt} = g(t) = \prod_n \frac{1}{(1 - t^n)^{a_n}}.$$

Taking logarithms and expanding in powers of $t$ we obtain

$$0 = \sum_n t^n \left( N^n - \sum_{d|n} a_d d \right).$$

Möbius inversion asserts that the solution of the equation

$$f(n) = \sum_{d|n} g(d)$$

is

$$g(n) = \sum_{d|n} \mu(\frac{n}{d}) g(d).$$

Inserting we are done. □

## 2.5 Hopf algebras

The universal enveloping algebra carries the structure of a *Hopf algebra*.

**Definition 2.18.** *A* bialgebra *is a vector space $A$ with an algebra and a coalgebra structure such that the multiplication and unit map (the comultiplication and counit) are maps of coalgebras (of algebras).*

**Definition 2.19.** *A* Hopf algebra *is a bialgebra $A$ with an additional operation $S : A \to A$ (called antipode) such that $\forall x \in A$*

$$S(x')x'' = x'S(x'') = \epsilon(x)\mathbb{1}.$$

Here we use the sumless Sweedler notation $\Delta x = x' \otimes x''$.

**Lemma 2.1.** *The antipode is an algebra anti-automorphism.*

*Proof.* Inserting $x = \mathbb{1}$ into the defining equation we see that $S(\mathbb{1}) = \mathbb{1}$. Next, for $x, y \in A$ given, let us compute

$$S(y)(S(x) = \epsilon(x'y')S(y'')S(x'') = S(x'y')x''y''S(y''')S(x''') = S(x'y')x''S(x''')\epsilon(y'') = S(x'y)x''S(x''') = S(x'y)\epsilon(x'') = S$$

□

**Example 2.16.** $U(\mathfrak{g})$ is a Hopf algebra with the coproduct being given on generators by $\Delta x = x \otimes 1 + 1 \otimes x$ for $x \in \mathfrak{g}$. The counit is defined such that $\epsilon(1) = 1$ and $\epsilon(x) = 0$ for $x \in \mathfrak{g}$. The antipode is $S(x) = -x$ for $x \in \mathfrak{g}$.

**Definition 2.20.** *An element $x \in H$ in a Hopf algebra $H$ is called* group-like *if $\Delta x = x \otimes x$ and is called* primitive *if $\Delta x = x \otimes 1 + 1 \otimes x$.*

**Exercise 2.9.** Show that the primitive elements in $U(\mathfrak{g})$ are exactly the elements of $\mathfrak{g} \subset U(\mathfrak{g})$ and that there is only one group-like element (namely 1).

## 2.6 Baker-Campbell-Hausdorff (BCH) formula

We define the complete free Lie algebra $\hat{\mathbb{F}}_{Lie}(S)$ as

$$\hat{\mathbb{F}}_{Lie}(S) = \prod_n \mathbb{F}_{Lie}(S)_n.$$

We define the algebra of (non-commutative) power series in symbols $S$ as $k\langle\langle S\rangle\rangle$. Clearly $\hat{\mathbb{F}}_{Lie}(S) \subset k\langle\langle S\rangle\rangle$. The Baker-Campbell-Hausdorff element $\mathrm{BCH} \in \hat{\mathbb{F}}_{Lie}(X, Y)$ is defined by the formula

$$\mathrm{BCH}(X,Y) = \log(e^X e^Y) = -\sum_{n\geq 1} \frac{1}{n}\left(1 - \sum_{k,l\geq 0} \frac{X^k Y^l}{k!l!}\right)^n = X + Y + \frac{1}{2}[X,Y] + \dots.$$

**Exercise 2.10.** Show that this is well-defined, i.e., that indeed $\mathrm{BCH}(X,Y) \in \hat{\mathbb{F}}_{Lie}(X,Y) \subset k\langle\langle X, Y\rangle\rangle$.

It is easy to see that

$$\begin{aligned}
\mathrm{BCH}(X,0) = \mathrm{BCH}(0,X) &= X \\
\mathrm{BCH}(X,-X) &= 1 \\
\mathrm{BCH}(\mathrm{BCH}(X,Y),Z) &= \mathrm{BCH}(X,\mathrm{BCH}(Y,Z)).
\end{aligned}$$

(use $(e^X e^Y)e^Z = e^X(e^Y e^Z)$).

## 2.7 (Pro-)nilpotent Lie algebras and (pro-)unipotent groups

Let $\mathfrak{g}$ be a Lie algebra. We define its lower central series to be the series of subalgebras

$$\mathfrak{g} = C_1 \supset C_2 \supset C_3 \supset \cdots$$

where recursively $C_{j+1} = [\mathfrak{g}, C_j]$. The Lie algebra $\mathfrak{g}$ is called *nilpotent* if the series terminates, i. e., if $C_{n+1} = 0$ for some $n$. The smallest such $n$ is called the nilpotence class of $\mathfrak{g}$.

**Example 2.17.** For $\mathfrak{g}$ semi-simple $\mathfrak{g} = C_1 = C_2 = C_3 = \cdots$. ($\Rightarrow$ not nilpotent)
Let $\mathfrak{g}$ be the Lie algebra of strictly upper triangular $n \times n$ matrices. Then $C_n = 0$. ($\Rightarrow$ nilpotent)

**Exercise 2.11.** The upper central series of $\mathfrak{g}$

$$0 = Z_0 \subset Z_1 = Z(\mathfrak{g}) \subset Z_2 \subset \cdots$$

is defined recursively through $Z_{i+1} = Z(\mathfrak{g}/Z_i)$ where $Z()$ denotes the center of a Lie algebra. Show that for nilpotent $\mathfrak{g}$ the length of the upper and lower central series agree.

**Exercise 2.12.** Verify that $[C_m, C_n] \subset C_{n+m}$.

**Proposition 2.1.** *Any finite dimensional nilpotent Lie algebra may be embedded into the Lie algebra of strictly upper triangular $n \times n$ matrices for some $n$.*

*Proof.* The descending filtration on $\mathfrak{g}$ by the $C_n$ induces a descending filtration on $U(\mathfrak{g})$. Concretely, $\mathcal{F}^p U(\mathfrak{g})$ is spanned by products $x_1 \dots x_r$, with $x_j \in C_{w_j}$ and $\sum_j w_j \geq p$. Define the left $\mathfrak{g}$ module $M := U(\mathfrak{g})/\mathcal{F}^{N+1}U(\mathfrak{g})$ where $N$ is the nilpotence class. Clearly $dim M < \infty$. Furthermore the action on $1 \in M$ embeds $\mathfrak{g}$ into $M$, hence the action is faithful, i. e., the map $\mathfrak{g} \to \mathfrak{gl}(M)$ is injective. Finally $M$ inherits the filtration $\mathcal{F}$ from $U(\mathfrak{g})$ and, by definition, $\mathfrak{g} \cdot \mathcal{F}^p M \subset \mathcal{F}^{p+1} M$. Hence, in a suitable basis, the image of $\mathfrak{g}$ in $\mathfrak{gl}(M)$ takes values in stricly upper triangular matrices. □

**Definition 2.21.** *A topological Lie algebra $\mathfrak{g}$ is pro-nilpotent if it is isomorphic to an inverse limit of finite dimensional nilpotent Lie algebras, i. e.,*

$$\mathfrak{g} = \lim_{\leftarrow} \mathfrak{g}_\alpha$$

*with each $\mathfrak{g}_\alpha$ finite dimensional, nilpotent, and considered equipped with the discrete topology. Morphisms of such Lie algebras are continuous morphisms of Lie algebras.*

**Remark 2.7.** Concretely, a basis of the topology is given by the open sets $\pi_\alpha^{-1}(\{0\})$, where $\pi_\alpha : \mathfrak{g} \to \mathfrak{g}_\alpha$ is the projection.

**Example 2.18.** Any nilpotent Lie algebra is trivially a pro-nilpotent Lie algebra. The complete free Lie algebra from the preceding section is pro-nilpotent, but not nilpotent.

**Lemma 2.2.** *A nilpotent Lie algebra $\mathfrak{g}$ is complete with respect to the topology given by the lower central series, i. e., the canonical map*

$$\mathfrak{g} \to \lim_{\leftarrow} \mathfrak{g}/C_n$$

*is an isomorphism of Lie algebras. (Not necessarily a homeomorphism, though.)*

*Proof.* We assume $\mathfrak{g} = \lim_{\leftarrow} \mathfrak{g}_\alpha$. An element of $\mathfrak{g}$ is hence a compatible collection of elements $x_\alpha \in \mathfrak{g}_\alpha$. An element of $\lim_{\leftarrow} \mathfrak{g}/C_n$ is a compatible collection of elements $x_n \in \mathfrak{g}/C_n$. Concretely, each $x_n$ is itself a compatible collection of elements $x_{n,\alpha} \in \mathfrak{g}_\alpha/C_{n,\alpha}$, where $C_{n,\alpha}$ are the terms of the lower central series for $\mathfrak{g}_\alpha$. The map in the Lemma sends $x = (x_\alpha)_\alpha$ to the quotients $(x_\alpha + C_{n,\alpha})_{\alpha,n}$.

Injectivity: If $x \mapsto 0$, i. e., $x_\alpha \in C_{n,\alpha}$ for all $n$ and $\alpha$, then by nilpotence of each $\mathfrak{g}_\alpha$ we have $x_\alpha = 0$ for each $\alpha$, hence $x = 0$.

Surjectivity: Given a compatible collection $y_{\alpha,n} \in \mathfrak{g}_\alpha/C_{n,\alpha}$, our task is to construct some compatible collection $x_\alpha \in \mathfrak{g}_\alpha$ in the preimage. Let $n_\alpha$ be the nilpotence class of $\mathfrak{g}_\alpha$. Then we set $x_\alpha = y_{\alpha,n_\alpha+1} \in \mathfrak{g}_\alpha$. □

**Remark 2.8.** The proof also shows that the lower-central-series induced topology is coarser than the true topology on $\mathfrak{g}$.

**Definition 2.22.** *For a nilpotent or pro-nilpotent Lie algebra $\mathfrak{g}$ we define the* exponential group $\mathrm{Exp}(\mathfrak{g})$ *to be the group with elements $\mathfrak{g}$ (as a set), with unit $0 \in \mathfrak{g}$, with inverse $x \mapsto -x$ for $x \in \mathfrak{g}$ and with composition*

$$x \cdot y := \mathrm{BCH}(x, y)$$

*for $x, y \in \mathfrak{g}$.*

Note that we are assuming, as always, that our ground field $\mathbb{K}$ has $char(K) = 0$. Otherwise $\mathrm{BCH}(x, y)$ does not make sense.

*Proof.* We need to check that this is well defined, i.e., that the formal series $\mathrm{BCH}(x, y)$ converges. But this follows from $\mathfrak{g} = \lim_{\leftarrow} \mathfrak{g}/C_n$. □

**Exercise 2.13.** Verify that $\mathrm{Exp}(\mathfrak{g})$ is indeed a group.

**Definition 2.23.** *For us, a unipotent group will be the exponential group of some finite dimensional nilpotent Lie algebra. For us, a pro-unipotent group will be the exponential group of a nilpotent Lie algebra. A morphism of such groups will be a morphism of the underlying (pro-)nilpotent Lie algebras.*

**Definition 2.24.** *Note that for $\mathfrak{g}$ finite dimensional, the group multiplication is given by polynomial formulas,*

**Remark 2.9.** Note that this is somewhat non-standard. For algebraic geometers a unipotent group is an algebraic group, isomorphic to a closed subgroup of the group $U_n$ of upper triangular $n \times n$ matrices with 1's on the diagonal, for some $n$. We have seen above that a finite dimensional nilpotent Lie algebra may be embedded into some $\mathfrak{u}_n (= Lie(U_n))$. Hence exponentiation of matrices gives an isomorphism of $\mathrm{Exp}(\mathfrak{g})$ with some closed subgroup $U_n$. However, we require $char(\mathbb{K}) = 0$, otherwise exponentiation or the BCH formula do not make sense.

The more standard notion of pro-unipotent group is that it is an inverse limit of unipotent groups. These groups are naturally pro-affine groups, and in particular pro-affine varieties and morphisms have to respect that structure. Concretely, in the standard approach one defines the allowed morphisms as morphisms of the (suitable defined) Hopf algebras of functions on these groups.

The pro-unipotent groups according to our definition are exactly the inverse limits of unipotent groups, however, we work only in characteristic zero.

### 2.7.1 Associated graded

For every Lie algebra the lower central series defines a filtration on the Lie algebra. (It is however not very useful unless the Lie-algebra is nilpotent or pro-nilpotent.) We may take the associated graded of this filtration, i. e., set

$$\mathrm{gr}^n \mathfrak{g} := C_n / C_{n+1}.$$

Since $[C_m, C_n] \subset C_{n+m}$ there is a natural way to put a Lie bracket on $\mathrm{gr}\mathfrak{g} = \oplus_n \mathrm{gr}^n \mathfrak{g}$ by declaring that the following diagram commutes:

$$\begin{tikzcd}
C_m \times C_n \arrow[r, "{[\,,\,]}"] \arrow[d] & C_{m+n} \arrow[d] \\
\mathrm{gr}^m \mathfrak{g} \times \mathrm{gr}^n \mathfrak{g} \arrow[r, "{[\,,\,]}"] & \mathrm{gr}^{m+n} \mathfrak{g}
\end{tikzcd}$$

We consider $\mathrm{gr}\mathfrak{g}$ as a Lie algebra with this bracket. It is a graded Lie algebra (of course).

**Example 2.19.** (Trivial examples) If $\mathfrak{g}$ is semi-simple, then $\mathrm{gr}\mathfrak{g} = 0$, so this construction is not so interesting. If $\mathfrak{g} = \mathbb{F}_{Lie}(S)$, then $\mathrm{gr}\mathfrak{g} = \mathfrak{g}$. (Note that $\mathfrak{g}$ was a graded from the start.) Similarly, for the upper triangular matrices $\mathfrak{g}$, $\mathrm{gr}\mathfrak{g} = \mathfrak{g}$. In this case $\mathrm{gr}^n \mathfrak{g}$ is spanned by matrices which are zero except on the $n$-th band above the diagonal.

**Example 2.20.** (Non-trivial examples) We will see some below.

## 2.8 Pro-unipotent algebras

**Definition 2.25.** *An augmented algebra is an algebra $A$ together with a morphism of algebras $\epsilon : A \to \mathbb{K}$.*

**Example 2.21.** $\mathbb{K}[X]$ is augmented, the augmentation being evaluation at $X = 0$.

For an augmented algebra $A$ the kernel of the augmentation morphism $I := \ker \epsilon$ is an ideal, the *augmentation ideal.*

**Definition 2.26.** *An augmented algebra $A$ is* unipotent *if the some power of the augmentation ideal vanishes, i. e., $I^{n+1} = 0$ for some $n$. The smallest such $n$ we call the* unipotence class *of $A$.*

**Example 2.22.** For $\mathbb{K}[X]$, the augmentation ideal is $X\mathbb{K}[X]$. $\mathbb{K}[X]$ is not unipotent. However $K[X]/\langle X^{n+1} \rangle$ is unipotent, and of unipotence class $n$.

**Definition 2.27.** *A topological augmented algebra $A$ is called* pro-unipotent *is it is (homeomorphic to) an inverse limit of finite dimensional unipotent algebras,*

$$A = \lim_{\leftarrow} A_\alpha.$$

*Here the right hand side is equipped with the inverse limit topology, after equipping each $A_\alpha$ with the discrete topology.*

Any unipotent algebra is a nilpotent Lie algebra and any pro-unipotent algebra is a pro-nilpotent Lie algebra.

**Definition 2.28.** *Let $L$ be a pro-nilpotent Lie algebra. The topological universal enveloping algebra $\hat{U}(L)$ of $L$ is the universal pro-unipotent algebra $A$, together with a map of pro-nilpotent Lie algebras $L \to A$, taking values in the augmentation ideal.*

Let $\mathfrak{g}$ be a finite dimensional nilpotent Lie algebra. Then one may set

$$\hat{U}(\mathfrak{g}) = \lim_{\leftarrow} U(\mathfrak{g})/I^n$$

i. e, take the completion of the algebraic Hopf algebra with respect to the augmentation ideal.

**Exercise 2.14.** Verify that $\hat{U}(\mathfrak{g})$ in this case is a pro-unipotent algebra, and that indeed the universal property is satisfied.

**Exercise 2.15.** Verify that $\hat{U}(\mathfrak{g})$ is a topological Hopf algebra. Show that the Lie algebra of primitive elements is exactly $L$, and that the group of group-like elements can be identified with $\mathrm{Exp}(\mathfrak{g})$. (Hint: The exponential and logarithm map exist and are inverse to each other.)

For a more general pro-nilpotent Lie algebra $\mathfrak{g} = \lim \mathfrak{g}_\alpha$ we may set

$$\hat{U}(\mathfrak{g}) := \lim_{\leftarrow} U(\mathfrak{g}_\alpha)/I_\alpha^n$$

where $I_\alpha$ is the augmentation ideal of $U(\mathfrak{g}_\alpha)$.

**Exercise 2.16.** Redo the previous two exercises in this case. I. e., verify that $\hat{U}(\mathfrak{g})$ satisfies the universal property, that it is a topological Hopf algebra, that the primitives are $\mathfrak{g}$ and that the group like elements are $\mathrm{Exp}(\mathfrak{g})$.

**Remark 2.10.** Compare this to the "usual" situation where we may associate (i) to a Lie algebra a Hopf algebra $U(\mathfrak{g})$, with no group-like elements and primitives $\mathfrak{g}$ of (ii) to a group $G$ a Hopf algebra $\mathbb{K}[G]$ with no primitives, but group like elements $G$. In the nilpotent case $\hat{U}(\mathfrak{g})$ unites these two Hopf algebras and may either be seen as a version of the universal enveloping algebra of $\mathfrak{g}$, or as a version of the group algebra $\mathbb{K}[\mathrm{Exp}(\mathfrak{g})]$.

**Exercise 2.17.** We saw before that $U(\mathbb{F}_{Lie}(S)) \cong \mathbb{K}\langle S\rangle$. Verify that $\hat{U}(\hat{\mathbb{F}}_{Lie}(S)) \cong \mathbb{K}\langle\langle S\rangle\rangle$, the topological Hopf algebra of formal power series in symbols from $S$.

## 2.9 Pro-unipotent (Malcev) completions of groups

**Definition 2.29.** *Let $G$ be a group. Then the pro-unipotent completion (or* Malcev *completion) of $G$, $\hat{G}$ is the universal pro-unipotent group with a morphism of groups $G \to \hat{G}$. The Lie algebra of $G$ is defined to be the pro-nilpotent Lie algebra underlying $\hat{G}$.*

**Remark 2.11.** More generally one may let $G$ be an affine algebraic group, and require that $G \to \hat{G}$ is pro-affine. (For unipotent $G$ then $\hat{G} = G$ trivially.) However, for us $G$ is always a discrete group, or in other words the underlying variety is a set of points.

Abstractly, one may construct $G \to \hat{G}$ as the inverse limit over arrows $G \to U$ with Zariski dense image and $U$ unipotent.

In this course, we will however use a more explicit construction, and restrict to finitely presented $G$ for simplicitly. I. e., $G$ is generated by some finite set $S$ (say $S = \{X_1, \dots, X_n\}$) and some finite set of relations $R$. Concretely, the most general relation looks like this

$$X_{j_1}^{\alpha_1} \cdots X_{j_k}^{\alpha_k} = 1.$$

Consider the (pro-nilpotent) Lie algebra of primitive elements $\mathfrak{g}_0 = \hat{\mathbb{F}}_{Lie}(x_1, \dots, x_n)$. We obtain relations in (i. e., elements of) $\mathfrak{g}_0$ by formally setting $X_j = \mathrm{Exp}(x_j)$ in the relations $R$. I. e., the relation above maps to the element

$$\mathrm{BCH}(\alpha_1 x_{j_1}, \mathrm{BCH}(\alpha_2 x_{j_2}, \dots, \mathrm{BCH}(\alpha_k x_{j_k}) \cdots) \in \mathfrak{g}_0.$$

Let us denote the relations thus obtained by $\{r_1, \dots r_N\}$. Let $I$ be the ideal generated by these relations. Also remember that $\mathfrak{g}_0 = \lim_{\leftarrow} \mathfrak{g}_0/C_l$ where $C_l$ are the terms of the lower central series. Then we set $\mathfrak{g} = \lim_{\leftarrow} \mathfrak{g}_0/(C_l + I)$. This is obviously a pro-nilpotent Lie algebra.

Furthermore, we claim there is a canonical map $G \to \mathrm{Exp}(\mathfrak{g})$, such that $X_j \mapsto \exp(x_j)$. Concretely, since $\mathrm{Exp}(\mathfrak{g}) = \lim_{\leftarrow} \mathrm{Exp}(\mathfrak{g}_0/(C_l + I))$ this means that the assignments $X_j \mapsto \exp(x_j + C_l + I)$ define (compatible) maps of groups $G \to \mathrm{Exp}(\mathfrak{g}_0/(C_l + I))$ for each $l$. For this one has to check that the images of the relations $R$ hold. But they are contained in $\exp(I)$ by construction so we are done.

**Proposition 2.2.** $\mathrm{Exp}(\mathfrak{g}) \cong \hat{G}$, *i. e.,* $\mathrm{Exp}(\mathfrak{g})$ *satisfies the universal property. (This also shows that the definition of* $\mathfrak{g}$ *is independent of the presentation of* $G$ *chosen, modulo unique isomorphism.)*

*Proof.* Suppose some other pro-unipotent group $\mathrm{Exp}(\tilde{\mathfrak{g}})$ with a map from $G$ is given, where

$$\tilde{\mathfrak{g}} = \lim_{\leftarrow} \tilde{\mathfrak{g}}_\alpha.$$

Then we need to show that there is a unique morphism $\mathfrak{g} \to \tilde{\mathfrak{g}}$ that makes the diagram

$$\begin{tikzcd} G \arrow[d] \arrow[dr] & \\ \mathrm{Exp}(\mathfrak{g}) \arrow[r] & \mathrm{Exp}(\tilde{\mathfrak{g}}) \end{tikzcd}$$

commute. This is equivalent to providing a compatible system of maps $\mathfrak{g} \to \tilde{\mathfrak{g}}_\alpha$ such that the diagrams

$$\begin{tikzcd} G \arrow[d] \arrow[dr] & \\ \mathrm{Exp}(\mathfrak{g}) \arrow[r] & \mathrm{Exp}(\tilde{\mathfrak{g}}_\alpha) \end{tikzcd}$$

commute. Let $\tilde{x}_j \in \tilde{\mathfrak{g}}$ be the unique elements such that $X_j \in G$ is mapped to $\exp(\tilde{x}_j)$. Clearly, our map to be constructed must send $x_j \mapsto \tilde{x}_j$. Since the Lie algebra generated by the $x_j$ is dense in $\mathfrak{g}$ this shows uniqueness. For existence, let $n_\alpha$ be the nilpotence class of $\mathfrak{g}_\alpha$ and define the map $\mathfrak{g} \to \tilde{\mathfrak{g}}_\alpha$ as the composition

$$\mathfrak{g} \to \mathfrak{g}_0/(C_{n_\alpha+1} + I) \to \tilde{\mathfrak{g}}_\alpha$$

where the right hand map is the unique one sending $x_j$ to $\tilde{x}_j$. This map exists since (i) $C_{n_\alpha+1}$ is sent to zero since the nilpotence class of $\tilde{\mathfrak{g}}_\alpha$ is $n_\alpha$ and (ii) $I$ is sent to zero since the existence of the map $G \to \tilde{\mathfrak{g}}_\alpha$ implies that the images of the relations $R$ must hold in $g_\alpha$. □

**Remark 2.12.** I do not know whether in general, for a pro-nilpotent Lie algebra $\mathfrak{g}$, and a closed ideal $I$, the map $\mathfrak{g}/I \to \lim_{\leftarrow} (g_\alpha/I)$ is an isomorphism (though it is surjective).

**Remark 2.13.** The kernel of the morphism $G \to \hat{G}$ is a normal subgroup which satisfies the following properties:

- If $x^n$ is in the kernel (for $n \geq 1$), so is $x$.
- The intersection of terms in the lower central series of $G$, $\cap_j C_j$ is in the kernel.

One may generalize the explicit description of the (topological) universal enveloping algebra of the free Lie algebra a bit, to also include the relations.

**Exercise 2.18.** Show that in the setting of this section

$$\hat{U}(\mathfrak{g}) \cong \lim_{\leftarrow} \mathbb{K}\langle\langle x_1, \dots, x_n \rangle\rangle / (\langle R_1, \dots, R_N \rangle + I^n)$$

where $I$ is the augmentation ideal in $U(\mathfrak{g}_0) = \mathbb{K}\langle\langle x_1, \dots, x_n \rangle\rangle$ and $R_1, \dots, R_N$ are the images of the relations therein. (These images are defined by replacing $X_j$ by $e^{x_j}$ in the relations.) This generalizes Exercise 2.17.

**Exercise 2.19.** Verify that $\hat{U}(\mathfrak{g}) \cong \lim_{\leftarrow} \mathbb{K}[G]/I^n$ where $\mathbb{K}[G]$ is the group algebra and $I$ is the augmentation ideal. Hence by Exercise **??** the group $\hat{G}$ may alternatively be described as the group-like elements of the limit on right hand side. (This is due to Quillen.)

**Example 2.23.** Consider the group $G = \mathbb{Z}/2\mathbb{Z}$. The group has one generator $X$ and one relation $X^2 = 1$. Setting $X = \exp(x)$ we obtain the relation $2x = 0$ an the Lie algebra, hence the Lie algebra of $G$ and the completion $\hat{G}$ are trivial.

**Example 2.24.** Show that the same assertion also holds for the permutation group $\mathbb{S}_n$ by the same reason (i. e., that torsion elements are mapped to 1 in $\hat{G}$).

**Remark 2.14.** The braid group on $n$ strands $B_n$ is the group with generators $\sigma_i$. $1 \leq i < n$ and relations

$$\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1} \qquad \text{for } 1 \leq i \leq n-2\sigma_i\sigma_j = \sigma_j\sigma_i \qquad \text{for } |i-j| \geq 2.$$

Elements may be considered as braids, see Figure **??**. There is a canonical map $B_n \to \mathbb{S}_n$. In fact, the right hand side is obtained by additionally imposing the relations $\sigma_i^2 = 1$. The pure braid group $\mathsf{PB}_n$ is the kernel of $B_n \to \mathbb{S}_n$. $\mathsf{PB}_n$ is presented as follows. The generators are

$$x_{ij} = \sigma_i^{-1}\sigma_{i+1}^{-1} \cdots \sigma_{j-2}^{-1}\sigma_{j-1}^2\sigma_{j-2} \cdots \sigma_{i+1}\sigma_i$$

for $1 \leq i < j \leq n$, see Figure **??**. The relations are

$$(x_{ij}, a_{ijk}) = (x_{ik}, a_{ijk}) = (x_{jk}, a_{ijk}) = 1$$
$$(x_{ij}, x_{kl}) = (x_{il}, x_{jk}) = 1$$
$$(x_{ik}, x_{ij}^{-1}x_{jl}x_{ij}) = 1$$

where $i < j < k < l$, $a_{ijk} := x_{ij}x_{ik}x_{jk}$ and $(\alpha, \beta) = \alpha\beta\alpha^{-1}\beta^{-1}$ denotes the commutator in the group.

**Exercise 2.20.** Verify that the map $\mathsf{PB}_3 \to \mathcal{F}_{\mathrm{Grp}}(X, Y)$ defined by

$$x_{12} \mapsto X \qquad x_{23} \mapsto Y \qquad x_{13} \mapsto X^{-1}Y^{-1}$$

is a group homomorphism. (I.e., check that the images of the pure braid group relations hold.)

**Exercise 2.21.** Show that $\mathsf{PB}_3 \cong \mathcal{F}_{\mathrm{Grp}}(X, Y) \times \mathcal{F}_{\mathrm{Grp}}(C)$. *Hint: Define $X = x_{12}$, $Y = x_{23}$, $Z = x_{13}$ and let $C = XZY = ZYX = YXZ$ replace the generator $Z$ and show that the relations become $[C, X] = [C, Y] = 1$.*

**Exercise 2.22.** Verify that the map $\mathsf{PB}_4 \to \mathcal{F}_{\mathrm{Grp}}(X, Y)$ defined by

$$\begin{aligned} x_{12} &\mapsto X & x_{34} &\mapsto X \\ x_{23} &\mapsto Y & x_{14} &\mapsto Y \\ x_{13} &\mapsto X^{-1}Y^{-1} & x_{34} &\mapsto Y^{-1}X^{-1} \end{aligned}$$

is a group homomorphism. This exercise will be very important for us later.

**Example 2.25.** We are interested in the Lie algebra $\mathfrak{pb}_n$ of $\mathsf{PB}_n$. It is generated by elements $t_{ij}$, $1 \leq i < j \leq n$. (We will also set $t_{ji} = t_{ij}$ for convenience.) The relations are obtained from the relations on $\mathsf{PB}_n$ in the straightforward manner (set $x_{ij} = \exp(t_{ij})$) and are a bit unwieldy. (Exercise: Write them down.)

However, we do want to write down the relations for the associated graded Lie algebras $\mathfrak{t}_n = \mathrm{gr}\mathfrak{pb}_n$. They are (in this case) obtained by forgetting about all higher commutators in the relations of $\mathfrak{pb}_n$ and read.

$$\begin{aligned} [t_{ij}, t_{kl}] &= 0 & & i, j, k, l \text{ pairwise distinct} \\ [t_{ij}, t_{ik} + t_{kj}] &= 0 & & i, j, k \text{ pairwise distinct} \end{aligned}$$

**Exercise 2.23.** Show that the Malcev Lie algebra of $P_n$ is isomorphic to $\mathbb{F}_{Lie}(X)$ for all $n$.

**Solution 2.2.** Let $t_j$ be the generator of the Malcev Lie algebra corresponding to $\sigma_j (\cong \exp(t_j))$. Then the relations read

$$t_i - t_j = (\text{Lie words of length} \geq 2 \text{ in } t_i, t_j).$$

Hence inserting the formula into its right hand side we see that $t_i - t_j$ can be written as arbitrary high order bracket expressions. Hence necessarily $t_i = t_j$ in $\hat{P}_n$.

## 2.10 Groupoids

We may define a group as a category with one object in which all morphisms are invertible.

**Definition 2.30.** *A* groupoid *is a category in which all morphisms are invertible.*

For us, all groupoids we consider will in fact have a finite set of objects, and all objects are isomorphic. Concretely, there is hence the following structure: There is a set of objects. $\mathrm{Hom}(o, o)$ is a group for any object $o$. $\mathrm{Hom}(o, o')$ is a $\mathrm{Hom}(o', o')$-$\mathrm{Hom}(o, o)$-torsor.

**Definition 2.31.** *Let $G$ and $H$ be groups. A $G$-$H$ torsor $T$ is a set together with a left free and transitive action of $G$ and a commuting right free and transitive action of $H$.*

**Exercise 2.24.** Picking any element $* \in T$ uniquely defines a map $\phi : H \to G$ such that $*h = \phi(h)*$. Show that $\phi$ is an isomorphism of groups.

**Exercise 2.25.** Fix any isomorphism $\phi : H \to G$. Then $G$ becomes a $G$-$H$ torsor (say $G_\phi$) by defining the left and right actions as $g \cdot X := gX$, $X \cdot h := X\phi(h)$. Show that any $G$-$H$-torsor is isomorphic to a torsor of this form. Show further that two torsors $G_\phi$, $G_{\phi'}$ are isomorphic iff $\phi' \circ \phi^{-1}$ is inner, i. e., of the form $Ad_X : g \mapsto XgX^{-1}$ for some $X$ in $G$.

**Example 2.26.** Let $V, W$ be vector spaces. Then $\mathrm{End}(V, W)$ is a $GL(W)$-$GL(V)$-torsor.

**Example 2.27.** Any category can be made into a groupoid by dropping all non-invertible morphisms.

**Example 2.28.** Let $S$ be any set. We may consider it as a groupoid with one object for each element in $S$ and exactly one morphism between any pair of objects.

**Example 2.29.** (Action groupoid) Let $S$ be a set with an action of the group $G$. The *action groupoid* is the groupoid with objects $S$ and morphisms

$$\mathrm{Hom}(s, s') = \{g \in G \mid g \cdot s' = s\}.$$

The composition is the one from $G$.

**Example 2.30.** We define the groupoid of colored braids in $n$ strands $\mathsf{CoB}_n$ (notation stolen from B. Fresse) as the pair groupoid associated to the action of $B_n$ on $\mathbb{S}_n$ (the action is defined via the map $B_n \to \mathbb{S}_n$). See Figure **??**.

Let us examine closer the structure of a groupoid. Let $G, H, K$ be groups, let ${}_GT_H$ be a $G$-$H$-torsor and let ${}_HT_K$ be an $H$-$K$-torsor. We define ${}_GT_H \times_H {}_HT_K$ as ${}_GT_H \times {}_HT_K/\sim$, where the equivalence relation identifies $(a, h \cdot b) \sim (a \cdot h, b)$ for all $h \in H$. One checks that ${}_GT_H \times_H {}_HT_K$ is a $G$-$K$ torsor.

**Exercise 2.26.** Let $o, o', o''$ be three objects in a groupoid, let $G = \mathrm{Hom}(o, o)$, $H = \mathrm{Hom}(o', o')$, $K = \mathrm{Hom}(o'', o'')$, and let ${}_GT_H = \mathrm{Hom}(o', o)$and ${}_HT_K = \mathrm{Hom}(o'', o')$. Check that the composition map ${}_GT_H \times_H {}_HT_K \to \mathrm{Hom}(o''', o)$ factors through ${}_GT_H \times_H {}_HT_K$. (Use the associativity axiom.)

**Remark 2.15.** Showing that a $G$-$H$ torsor exists is sometimes a convenient way of showing that $G$ and $H$ are isomorphic, without constructing an explicit isomorphism.

## 2.11 Pro-unipotent groupoids

One may generalize our previous constructions to the groupoid setting.

**Definition 2.32.** *Let $G$, $H$ be pro-unipotent groups. Then a pro-unipotent $G$-$H$ torsor is a torsor of the form $G_\phi$ (cf. Exercise 2.25) for $\phi : H \to G$ an automorphism of pro-unipotent groups.*

A morphism of pro-unipotent torsors is the same as a morphism of the underlying torsors. Concretely, given morphisms of pro-unipotent groups $\alpha : G \to G'$ and $\beta : H \to H'$, then a morphism of torsors $G_\phi \to G'_{\phi'}$ is the same as an element $X \in G$ such that $\mathrm{Ad}_X \circ \phi' \circ \beta = \alpha \circ \phi$. (Such an element may not exist, in which case there are no morphisms of the torsors.)

Given a $G$-$H$ pro-unipotent torsor ${}_GT_H = G_\phi$ and an $H$-$K$ pro-unipotent torsor ${}_HT_K = H_{\phi'}$ we set

$${}_GT_H \times_H {}_HT_K = G_\phi \times_H H_{\phi'} := G_{\phi \circ \phi'}.$$

There is a canonical map

$$\begin{aligned} {}_GT_H \times {}_HT_K &\to {}_GT_H \times_H {}_HT_K \\ (g, h) &\mapsto g\phi(h). \end{aligned}$$

**Definition 2.33.** *A pro-unipotent groupoid is a groupoid such that*

- *All spaces* $\mathrm{Hom}(o,o)$ *are pro-unipotent groups.*
- *All spaces* $\mathrm{Hom}(o,o')$ *are either empty or pro-unipotent* $\mathrm{Hom}(o,o)$-$\mathrm{Hom}(o,o)$ *torsors.*
- *Suppose* $\mathrm{Hom}(o,o')$ *and* $\mathrm{Hom}(o',o'')$ *are non-empty. Then the composition maps* $\mathrm{Hom}(o,o') \times \mathrm{Hom}(o',o'') \to \mathrm{Hom}(o,o'')$ *factor as*

$$\mathrm{Hom}(o,o') \times \mathrm{Hom}(o',o'') \to \mathrm{Hom}(o,o') \times_{\mathrm{Hom}(o',o')} \mathrm{Hom}(o',o'') \to \mathrm{Hom}(o,o'')$$

*where the right hand arrow is a map of pro-unipotent torsors.*

*A morphism of pro-unipotent groupoids is defined in the obvious manner.*

**Definition 2.34.** *Let $G$ be a groupoid. The pro-unipotent completion $\hat{G}$ of $G$ is the universal pro-unipotent groupoid with a map of groupoids $G \to \hat{G}$.*

Let us construct $\hat{G}$. It has the same objects as $G$. $\mathrm{Hom}_{\hat{G}}(o,o)$ is defined to be the pro-unipotent completion of $\mathrm{Hom}_G(o,o)$ for each object $o$. In $G$, pick an element $f_{oo'} \in \mathrm{Hom}_G(o',o)$ for any isomorphic pair of objects $o \neq o'$. This fixes an isomorphism $\phi_{oo'}\mathrm{Hom}_G(o',o')\mathrm{Hom}_G(o,o)$. We define $\mathrm{Hom}_{\hat{G}}(o,o) := H_{\phi_{oo'}}$ where $H = \mathrm{Hom}_{\hat{G}}(o,o)$. Some composition morphisms are defined by the action on the torsors. The others define as follows:

$$\mathrm{Hom}_{\hat{G}}(o',o) \times \mathrm{Hom}_{\hat{G}}(o'',o') = H_{\phi_{oo'}} \times H'_{\phi_{o'o''}} \to H_{\phi_{oo'}} \times'_H H'_{\phi_{o'o''}} = H_{\phi_{oo'}\circ\phi_{o'o''}} \xrightarrow{X} H_{\phi_{oo''}}.$$

Here $H' = \mathrm{Hom}_{\hat{G}}(o',o')$ for brevity and $X = f_{oo''}^{-1} \circ f_{oo'} \circ f_{o'o''} \in H$. The canonical map of groupoids $G \to \hat{G}$ is defined as follows. It is the identity on the objects. The map $\mathrm{Hom}_G(o,o) \to \mathrm{Hom}_{\hat{G}}(o,o)$ is the canonical map contained in the definition of pro-unipotent completion. The maps $\mathrm{Hom}_G(o',o) \to \mathrm{Hom}_{\hat{G}}(o',o)$ is the composition

$$\mathrm{Hom}_G(o',o) \xrightarrow{\circ f_{oo'}^{-1}} \mathrm{Hom}_G(o,o) \to \mathrm{Hom}_{\hat{G}}(o,o) = H \to H_{\phi_{oo'}} = \mathrm{Hom}_{\hat{G}}(o',o).$$

**Exercise 2.27.** Verify the universal property for $\hat{G}$.

**Solution 2.3.** Assume the pro-unipotent groupoid $G'$ is given, with a map $G \to G'$. We have to show that the map factors through $G \to \hat{G}$. Let us construct $\hat{G} \to G'$. Since $G$ and $\hat{G}$ have the same objects it is clear how to map objects, and the map is uniquely defined. On spaces $H = \mathrm{Hom}_{\hat{G}}(o,o)$ the map is defined and unique by the universal property for the ordinary pro-unipotent completion. On spaces $\mathrm{Hom}_{\hat{G}}(o',o) = H_{\phi_{oo'}}$ we define the map uniquely by sending $\mathbb{1} \in H$ to the image of the chosen element $f_{oo'}$ in $G'$. This is the unique choice that makes the diagram inthe universal property commute. Hence uniqueness is shown. However, we still need to verify that the map thus defined is a functor, i. e., that it commutes with composition of morphisms. But that may be verified on the elements $\mathbb{1}$ in the torsors (the images of the $f_{oo'}$ in $G$), for all other elements it follows by equivariance. But on these elements it follows from the fact that $G \to G'$ is a functor using equivariance again.

**Example 2.31.** The most interesting example for us the pro-unipotent completion $\widehat{\mathsf{CoB}}_n$ of the groupoid of colored braids. Note that this completion is not trivial, the isomorphism group of each object being isomorphic to $\widehat{\mathsf{PB}}_n$.

## 2.12 (Braided) Monoidal Categories

Let us recall the following standard definitions from category theory.

**Definition 2.35.** *A monoidal category $(\mathcal{C}, \otimes, \mathbb{1})$ is a category together with*

1. *A binary operation (a functor)*
$$\mathcal{C} \times \mathcal{C} \to \mathcal{C}.$$
2. *An isomorphism of functors (i. e., an invertible natural transformation)*
$$\alpha : \cdot \otimes (\cdot \otimes \cdot) \Rightarrow (\cdot \otimes \cdot) \otimes \cdot$$
*It is called* associator.

*3. A distinguished object* $\mathbb{1} \in \mathrm{Ob}\mathcal{C}$.

*4. Isomorphisms of functors*

$$\lambda : \mathbb{1} \otimes \cdot \Rightarrow id$$
$$\rho : \cdot \otimes \mathbb{1} \Rightarrow id$$

*called* left and right unitor.

*These data are required to satisfy the following relations.*

$$\begin{tikzcd}
 & (A \otimes B) \otimes (C \otimes D) \arrow[dl] & \\
A \otimes (B \otimes (C \otimes D)) & & \\
 & & ((A \otimes B) \otimes C) \otimes D \arrow[uul] \arrow[ddl] \\
A \otimes ((B \otimes C) \otimes D)) \arrow[uu] & & \\
 & (A \otimes (B \otimes C)) \otimes D \arrow[ul] &
\end{tikzcd}$$

$$\begin{tikzcd}
(A \otimes I) \otimes B & A \otimes (I \otimes B) \\
A \otimes B &
\end{tikzcd}$$

$$\lambda_{\mathbb{1}} = \rho_{\mathbb{1}} : \mathbb{1} \otimes \mathbb{1} \to \mathbb{1}.$$

*A non-unital monoidal category is the subset of the above data not involving the unit.*

*Let* $\mathcal{C}, \mathcal{D}$ *be monoidal categories (respectively non-unital monoidal categories). We say that a functor* $F : \mathcal{C} \to \mathcal{D}$ *is* strict monoidal *if it preserves the above structures. Concretely, this means that* $F(\cdot \otimes_{\mathcal{C}} \cdot) = F(\cdot) \otimes_{\mathcal{D}} F(\cdot)$, $F(\mathbb{1}_{\mathcal{C}}) = \mathbb{1}_{\mathcal{D}}$ *(in the unital case) and that* $F$ *intertwines with the natural transformations* $\alpha$ *and (in the unital case)* $\rho$ *and lambda.*

**Example 2.32.** The category of sets Set is monoidal, with product the Cartesian product.

**Example 2.33.** The category of vector spaces Vect is monoidal with the usual tensor product. For given vector spaces $U$, $V$, $W$ the associator $\alpha_{U,V,W} : U \otimes (V \otimes W) \to (U \otimes V) \otimes W$ is defined by the obvious formula $u \otimes (v \otimes w) \mapsto (u \otimes v) \otimes w$. Alternatively, Vect can be made into a monoidal category by declaring the monoidal product to be the direct sum.

**Example 2.34.** The category of algebras is monoidal, with the tensor product being the tensor product of the underlying vector spaces. Note that the product on $A \otimes B$ for $A$, $B$ algebras is defined such that elements of $A$ and $B$ commute, and the unit is $1_A \otimes 1_B$.

**Example 2.35.** The category of Lie algebras is monoidal taking the direct sum of the underlying vector spaces as product. (Note: In the Lie algebra $\mathfrak{g} \otimes \mathfrak{h}$, $\mathfrak{g}$ commutes with $\mathfrak{h}$.)

It can be shown that the above axioms suffice to ensure that "any diagram built using only $\alpha$, $\rho$ and $\lambda$ and their inverses" commutes. However, like this the statement is strictly speaking not correct, since in the monoidal category in question identities additional identities betwen some tensor products may hold that make it possible to write down diagrams that do not commute. The way to state the Theorem correctly is the following. Let $M$ be the free magma generated by a single symbol $X$. Elements are just parenthesations of copies of $X$, like

$$X(X(XX)). \qquad \text{(Think: } X \otimes (X \otimes (X \otimes X)).\text{)}$$

Another name for $M$ is the set of planar binary trees. Let $M_{\mathbb{1}}$ be the free magma generated by symbols $X$ and $\mathbb{1}$, elements are just planar binary trees whose leafs are labeled by either $X$ or $\mathbb{1}$. We consider $M$ and $M_{\mathbb{1}}$ as categories with exactly one morphism between any pair of objects with the same number of $X$'s (and none between objects with different numbers of $X$'s). Clearly, $M_{\mathbb{1}}$ is monoidal (with the obvious monoidal structure) and $M$ is non-unital monoidal.

**Theorem 2.3** (MacLane Coherence Theorem). *Let $\mathcal{C}$ be a monoidal category (a non-unital monoidal category). For every $o \in \mathrm{Ob}\mathcal{C}$ there is a unique strict monoidal functor $F : M_{\mathbb{1}} \to \mathcal{C}$ (respectively $F : M \to \mathcal{C}$) such that $F(X) = o$.*

**Remark 2.16.** The non-trivial part here is not the (obvious) uniqueness, but the fact that the functor exists. In particular, note that all diagrams in $M\mathbb{1}$ trivially commute, and hence all images of these diagrams have to commute as well.

**Definition 2.36.** *A braided monoidal category is a monoidal category with an additional natural isomorphism*

$$\cdot \otimes ? \Rightarrow ? \otimes \cdot$$

*satisfying the following two conditions:*

$$\begin{tikzcd}
& A \otimes (B \otimes C) \arrow[r] \arrow[dl] & (B \otimes C) \otimes A & \\
(A \otimes B) \otimes C \arrow[dr] & & & B \otimes (C \otimes A) \arrow[ul] \\
& (B \otimes A) \otimes C \arrow[r] & B \otimes (A \otimes C) \arrow[ur] &
\end{tikzcd}
\qquad
\begin{tikzcd}
& (A \otimes B) \otimes C \arrow[r] \arrow[dl] & C \otimes (A \otimes B) & \\
A \otimes (B \otimes C) \arrow[dr] & & & (C \otimes A) \otimes B \arrow[ul] \\
& A \otimes (C \otimes B) \arrow[r] & (A \otimes C) \otimes B \arrow[ur] &
\end{tikzcd}$$

*A non-unital braided monoidal category is a non-unital monoidal category with the same additional datum as above.*

**Definition 2.37.** *A symmetric monoidal category is a braided monoidal category in which the the composition*

$$\cdot \otimes ? \Rightarrow ? \otimes \cdot \Rightarrow \cdot \otimes ?$$

*(i. e., the braiding applied twice) is the identity.*

**Example 2.36.** All the examples above are actually symmetric monoidal categories. We will soon see an example of a category that is monoidal but not symmetric.

**Example 2.37.** For the category of $\mathbb{Z}$-graded vector spaces the braiding $U \otimes V \to V \otimes U$ is defined as

$$u \otimes v \mapsto (-1)^{|u||v|} v \otimes u$$

for homogeneous elements $u \in U$, $v \in V$.

There is also a coherence theorem for braided monoidal categories. It says that the morphisms between two (arbitrarily bracketed) tensor products that can be written down using only the braidings, associators and unit are exactly given by the pure braid group, or the braid group if one assumes all involved objects identical. More precisely, similarly to $M$ above define a category $M^b$ with objects the free magma in one symbol $X$, and with sets of morphisms between any two objects of equal numbers $n$ of $X$'s the braid group $B_n$ in $n$ strands. Define similarly $M^b_{\mathbb{1}}$ to be the category with objects the free magma in symbols $\mathbb{1}, X$, and with sets of morphisms between any two objects of equal numbers $n$ of $X$'s the braid group $B_n$ in $n$ strands.

**Theorem 2.4.** *Let $\mathcal{C}$ be a baided monoidal category (a non-unital braided monoidal category). For every $o \in \mathrm{Ob}\mathcal{C}$ there is a unique strict braided monoidal functor $F : M_{\mathbb{1}} \to \mathcal{C}$ (respectively $F : M \to \mathcal{C}$) such that $F(X) = o$.*

### 2.12.1 Monoids

Fix a monoidal category $\mathcal{C}$. A *monoid* in $\mathcal{C}$ is an object $M \in \mathrm{Ob}\mathcal{C}$ together with morphisms

$$\begin{aligned} \mathbb{1} &\to M && \text{(unit)} \\ M \otimes M &\to M && \text{(product)} \end{aligned}$$

such that the following diagrams commute:

$$
\begin{tikzcd}
(M \otimes M) \otimes M \arrow[r] \arrow[d] & M \otimes (M \otimes M) \arrow[r] & M \otimes M \arrow[d] \\
M \otimes M \arrow[rr] & & M
\end{tikzcd}
$$

$$
\begin{tikzcd}
M \otimes \mathbb{1} \arrow[dr] & M \otimes M \arrow[l] \arrow[r] & \mathbb{1} \otimes M \arrow[dl] \\
& \hat{M} &
\end{tikzcd}
$$

**Example 2.38.** A monoid in Vect is an associative algebra.

# Chapter 3

# Operads

## 3.1 Motivation (what we want to achieve)

In algebra we deal with a variety of "algebraic structures" like associative algebras, commutative algebras, Lie algebras or modules over these objects. For each such type of algebraic structure, much of the theory is very similar. We would like to simplify the situation and move to a one step more abstract level. We would like to make a theory of all algebraic structures at once. Of course, the first step we have to take here is this:

**Task:** Give a precise definition of what is an "algebraic structure".

Operad theory attempts to do just that. There will be one operad for each type of algebraic structure (like the associative operad $Ass$, the Lie operad $Lie$ etc.). An algebraic structure on some object will then be a representation of an operad on that object. For example, a Lie algebra structure on some vector space $V$ is the same as a representation of the operad $Lie$ on $V$.

Concretely, an operad encodes the space of possible operations that may be applied to elements of some algebraic object. For example, if we are given $n$ elements $a_1, \dots, a_n$ of an associative algebra, we may multiply them in any one of $n!$ possible orders. Furthermore, we may take any linear combination of elements thus obtained. Hence the space of $n$-ary operations of the operad $Ass$ is $Ass(n) \cong \mathbb{K}^{n!}$.

## 3.2 Definition

**Definition 3.1.** *An $\mathbb{S}$-module $\mathcal{P}$ in a category $\mathcal{C}$ is a collection of objects $\mathcal{P}(n) \in \mathcal{C}$ together with right $\mathbb{S}_n$ actions for $n = 0, 1, 2, \dots$. The category of $\mathbb{S}$-modules in $\mathcal{C}$ is denoted by $\mathcal{C}^{\mathbb{S}}$.*

**Definition 3.2.** *An operad $\mathcal{P}$ is the following data:*

- *An $\mathbb{S}$ module in* Vect. *($\mathcal{P}(n)$ is called* spaces of $n$-ary operations*).*
- *A distinguished element $\mathbb{1} \in \mathcal{P}(1)$ (the* unit*).*
- *A family of morphisms*

$$\mu_{n,k_1,\dots,k_n} : \mathcal{P}(n) \otimes \mathcal{P}(k_1) \otimes \cdots \otimes \mathcal{P}(k_n) \to \mathcal{P}(k_1 + k_2 + \cdots + k_n)$$

  *(the* operadic compositions*).*

*These data are required to satisfy the following axioms:*

1. *(Equivariance) The compositions are equivariant with respect to the symmetric group actions.*
2. *(Unit axiom)*

$$\mu_{1,n}(\mathbb{1}, n) = x$$
$$\mu_{n,1,\dots,1}(x, \mathbb{1}, \dots, \mathbb{1}) = x$$

   *For each $n$ and each $x \in \mathcal{P}(n)$.*

3. *(Associativity)*

$$\mu_{n,k_1,\ldots,k_n}(x,\mu_{n,l_1^1,\ldots,l_n^1}(y_1,z_1^1,\ldots,z_l^1),\ldots,\mu_{n,l_1^1,\ldots,l_n^1}(y_1,z_1^1,\ldots,z_l^1))$$
$$= \mu_{n,k_1,\ldots,k_n}(\mu_{n,k_1,\ldots,k_n}(x,y_1,\ldots,y_n),z_1^1,\ldots,z_l^n)$$

In the obvious manner one defines morphisms of operads.

**Example 3.1.** The commutative operad $Com_1$ is defined by $Com_1(n) = \mathbb{K}$ for all $n$, with trivial $\mathbb{S}_n$ action. Let us fix a basis of all $Com_1(n)$, say $Com_1(n) = \mathbb{K}\cdot m_n$. The composition morphisms are defined by the formula

$$\mu_{n,k_1,\ldots,k_n}(m_n, m_{k_1},\ldots,m_{k_n}) = m_{k_1+\cdots+k_n}$$

The operadic unit is $\mathbb{1} = m_1$. There is also the sub-operad $Com \subset Com_1$ with $Com(0) = 0$, $Com(n) = Com_1(n)$ for $n \geq 1$.

**Example 3.2.** The associative operad $Ass_1$ is defined through $Ass_1(n) = \mathbb{K}[\mathbb{S}_n]$, where the right hand side denotes the group ring of $\mathbb{S}_n$ and the right $\mathbb{S}_n$-action is the canonical one. Let us fix as basis of these vector spaces the canonical one given by elements of $\mathbb{S}_n$. Then the compositions are defined such that

$$\mu_{n,k_1,\ldots,k_n}((12\cdots n),(1\cdots k_1),\ldots,(1\cdots k_n)) = (12\cdots N)$$

where $N = k_1 + \cdots + k_n$. All other cases are uniquely determined by the equivariance conditions. The unit is $\mathbb{1} = (1)$. We define $Ass \subset Ass_1$ by setting $Ass(0) = 0$, $Ass(n) = Ass_1(n)$, $n > 0$.

**Example 3.3.** Let $V$ be any vector space. Than the endomorphism operad $\mathrm{End}(V)$ is defined as follows.

$$\mathrm{End}(V)(n) = \mathrm{Hom}(V^{\otimes n}, V)$$

equipped with the natural right $\mathbb{S}_n$ module structure by permuting arguments. The compositions are the obvious ones

$$\mu_{n,k_1,\ldots,k_n}(f,g_1,\ldots,g_n)(x_1,\ldots,x_N) = f(g_1(x_1,\ldots,x_{k_1}),\ldots,g_n(x_{N-k_n+1},\ldots,x_N)).$$

The unit is the identity map in $\mathrm{End}(V)(1)$.

**Definition 3.3.** *Let $\mathcal{P}$ be an operad. A representation of $\mathcal{P}$ (or a $\mathcal{P}$-algebra structure) on some vector space $V$ is an operad map $\mathcal{P} \to \mathrm{End}(V)$.*

**Remark 3.1.** The operadic compositions are completely determined by specifying the reduced operadic compositions

$$\circ_j : \mathcal{P}(n) \otimes \mathcal{P}(m) \to \mathcal{P}(n+m-1)$$
$$a \circ_j b = \mu_{n,1,\ldots,1,m,1,\ldots,1}(\mathbb{1},\ldots,\mathbb{1},b,\mathbb{1},\ldots,\mathbb{1}).$$

**Exercise 3.1.** Verify that a representation of $\mathcal{P}$ is a collection of maps

$$\nu_n : \mathcal{P}(n) \otimes_{S_n} V^{\otimes n} \to V$$

satisfying the following conditions:

- $\nu_n(\mathbb{1}) \otimes x = x \forall x \in V$.
- $\nu_n(p,\nu_{k_1}(p_1,x_1,\ldots),\ldots,\nu_{k_n}(p_n,x_{\ldots},\ldots,x_N)) = \nu_n(\mu(p,p_1,\ldots,p_n),x_1,\ldots,x_N)$.

## 3.3 A slightly better description of operad

Elements of $\mathcal{P}(n)$ of an operad $\mathcal{P}$ may be thought of as operation with $n$ inputs labelled by 1 to $n$. It is often more convenient to use a slightly different (but equivalent) definition of operad, in which "the inputs are labelled by" some arbitrary set $S$.

Let $\mathrm{Set}_f$ be the category with objects the finite sets and morphisms the bijections.

**Definition 3.4.** *An operad $\mathcal{P}$ is a functor* $\mathrm{Set}_f \vec{\ } \mathrm{Vect}$, *together with*

- *A distinguished object* $\mathbb{1}_S \in \mathcal{P}(S)$ *for each one element set* $S$*.*
- *Composition morphisms: For all finite sets* $S, T$ *and* $s$ *in* $S$ *a morphism*

$$\mathcal{P}(S) \otimes_s \mathcal{P}(T) \to \mathcal{P}((S \setminus \{s\} \sqcup T).$$

*These data have to satisfy the following conditions:*

- *Naturality: For every bijection* $f : S \to S'$ *we have* $\mathcal{P}(f)(\mathbb{1}_{S'}) = \mathbb{1}_S$*.*[1] *Similarly, for all bijections* $f : S \to S'$, $g : T \to T'$

$$\begin{tikzcd} \mathcal{P}(S') \otimes_{f(s)} \mathcal{P}(T') \arrow[r] \arrow[d] & \mathcal{P}((S \setminus \{s\} \sqcup T) \arrow[d] \\ \mathcal{P}(S) \otimes_s \mathcal{P}(T) \arrow[r] & \mathcal{P}((S \setminus \{s\} \sqcup T) \end{tikzcd}$$

  *commutes.*
- *Unit axiom: Operadic composition with the operadic units are the identity.*
- *Associativity: ...*

We may equivalently define an $\mathbb{S}$-module to be a functor $\mathrm{Set}_f^{\to}\mathrm{Vect}$.

## 3.4 Operads in other categories

Looking at the definition of operad, we see that it may be copied word by word for any symmetric monoidal category $\mathcal{C}$ having finite limits and colimits replacing Vect. We will call the objects thus obtained *operads in* $\mathcal{C}$.

We may also define the notion of representation of an operad in $\mathcal{C}$ by copying the definition of Exercise 3.1 above.

## 3.5 The free operad

There is an obvious forgetful functor

$$\mathrm{Op}_{\mathcal{C}} \to \mathcal{C}^{\mathbb{S}}.$$

It has a left adjoint, the free operad functor, which assigns to every $\mathbb{S}$-module $\mathcal{A}$ some operad $\mathsf{Free}_{\mathcal{A}}$.

The explicit construction for $\mathcal{C} = \mathrm{Vect}$ is as follows:

1. Define the auxiliary space $V(n)$ to be the space spanned by all formal functional expressions one may write down using elements of $\mathcal{A}(k)$ ($k = 0, 1, \dots$) and formal variables $X_1, \dots, X_n$, where elements of $\mathcal{A}(k)$ are considered as $k$-ary multilinear functions and each $X_j$ occurs exactly once. Example: $f_4(f_3(X_2, X_5, X_1), X_4, X_6, f_0)$ is allowed where $f_j \in \mathcal{A}(j)$. Furthermore $f_2(f_0, 2f_0) = 2f_2(f_0, f_0)$ etc.
2. We identify $f_n(A_1, \dots, A_n) \sim (f_n\sigma)(A_{\sigma(1)}, \dots, A_{\sigma(n)})$ for all permutations $\sigma \in \mathbb{S}_n$ everywhere in expressions.
3. The right $\mathbb{S}_n$ action on $\mathsf{Free}_{\mathcal{A}}(n)$ is given by permuting indices of the $X_j$.
4. The operadic unit is given by the expression $X_1$.
5. The operadic composition is defined in the obvious way.

For other categories $\mathcal{C}$ the construction is analogous.

[1] This may be rephrased as saying that $\mathbb{1}$ is a natural transformation from the operad $*$ with only the identity operation to $\mathcal{P}$.

**Example 3.4.** Let $\mathcal{C} = \mathrm{Set}$ and let the $\mathbb{S}$-module $\mathcal{A}$ be defined by setting

$$\mathcal{A}(n) = \begin{cases} \mathbb{S}_2 & \text{for } n = 2 \\ \emptyset & \text{otherwise} \end{cases}$$

Let $PaP = \mathsf{Free}_{\mathcal{A}}$. Algebras over this operad are called magmas, and are just sets with a binary operation not required to satisfy any axioms. Elements of $PaP(n)$ have a very concrete description and may be understood as parenthesized permutations (hence the notation). For example

$$PaP(2) = \{(12), (21)\}$$
$$PaP(3) = \{((12)3), (1(23)), ((21)3), (2(13)), \text{etc.}\}$$

**Exercise 3.2.** Show that $|PaP(n)| = n!C_n$ where $C_n = \frac{1}{n+1}\binom{2n}{n}$ is the $n$-th Catalan number.

**Example 3.5.** We may add take a quotient of $PaP$ by the equivalence relation

$$\mu \circ_1 \mu = \mu \circ_2 \mu.$$

Algebras over the resulting operad $Ass_{\mathrm{Set}}$ are associative algebra objects in Set. Concretely

$$Ass_{\mathrm{Set}}(n) = \mathbb{S}_n.$$

**Example 3.6.** There is a functor $\mathsf{Fin} \to \mathcal{C}$ into any monoidal category $\mathcal{C}$ sending a finite set $S$ to $\coprod_{s\in S} \mathbb{1}$. Hence from an operad in $\mathsf{Fin}$ we may obtain an operad in $\mathcal{C}$. For example, for $Ass_{\mathrm{Set}}$ the resulting operad will govern algebra objects in $\mathcal{C}$. (discuss unital/non-unital)

### 3.5.1 More free operads

Let us start by recalling several familiar definitions.

**Example 3.7.** Let $S$ be a set. The free group $\mathcal{F}_{\mathrm{Grp}}(S)$ in $S$ may be identified with the set of (possibly empty) strings in letters $S \sqcup S$, where a symbol $s$ from the second copy of $S$ is written $s^{-1}$ to distinguish it from one in the first, modulo relations $ss^{-1} = s^{-1}s = 1$. The product is juxtaposition, the inverse is the obvious map. There is an (equally obvious) map of sets $S \to \mathcal{F}_{\mathrm{Grp}}(S)$. It satisfies a universal property: Let $G$ be another group. Then any map of sets $S \to G$ factors uniquely as

$$S \to \mathcal{F}_{\mathrm{Grp}}(S) \to G$$

where the right hand arrow is a group homomorphism. Another way to say this is that the functor $\mathcal{F}_{\mathrm{Grp}}(\cdot) : \mathrm{Set} \to \mathsf{Grp}$ is left adjoint to the forgetful functor $\mathsf{Grp} \to \mathrm{Set}$.

**Example 3.8.** Let $S$ be a set. The free algebra in $S$ is the vector space spanned by (possibly empty) strings with letters in $S$. Exercise: Write down the universal property.

**Example 3.9.** To generate a groupoid it is not enough to provide a generating (morphism) set. One first has to provide a set of objects $O$, and then a set of morphisms $M$, together with two maps $M \rightrightarrows O$ (the source and target maps). Then the free groupoid in $M \rightrightarrows O$ is a category such that $\mathrm{Hom}(o, o')$ is given by all *admissible* strings in letters $M \sqcup M$, with relations $mm^{-1} = m^{-1}m = 1$. Here a string is admissible if the source of any letter equals the target of the next, the first letter has target $o'$, and the last letter has source $o$. The free groupoid functor is left adjoint to the forgetful functor

$$\mathsf{Groupoids} \to (\mathrm{Set} \rightrightarrows \mathrm{Set}).$$

Exercise: Write down the universal property explicitly.

**Exercise 3.3.** Recall from Example **??** that the braid group $B_n$ was generated by the pair transpositions $b_i$ under the relations (**??**), and that the colored braid groupoid $\mathsf{CoB}_n$ is the action groupoid of the action of $B_n$ on $\mathbb{S}_n$. Concretely, morphisms are pairs $(b, \sigma)$, with $b \in B_n$ and $\sigma \in \mathbb{S}_n$. Show that the groupoid $\mathsf{CoB}_n$ is generated by morphisms $(b_i, \sigma)$ under relations $(R, \sigma)$, where $R$ stands for the relations (**??**).

Let us move on to more elaborate examples, involving operads.

**Example 3.10.** Consider the forgetful functor

$$(\text{Operads in algebras}) \to \mathrm{Set}^{\mathbb{S}}.$$

To describe its left adjoint, let us introduce the following notation. Let an $\mathbb{S}$-module $S$ in Set be given, and let us define another (bigger) $\mathbb{S}$-module $\tilde{S}$. $\tilde{S}(N)$ has as elements (equivalence classes of) tuples

$$s^{I_1,\dots,I_n} := (s; I_1, \dots, I_n)$$

where $s \in S(n)$, for $N \geq n$ and any $n$ disjoint subsets $I_1, \dots I_n \subset \{1, \dots, N\}$. We consider two such tuples identical if they can be mapped onto each other using the symmetric group action, i.e.,

$$(s \cdot \sigma; I_1, \dots, I_n) \equiv (s; I_{\sigma(1)}, \dots, I_{\sigma(n)}).$$

for some $\sigma \in \mathbb{S}_n$. TODO: inverse?

The action of the symmetric group $\mathbb{S}_N$ on such tuples is defined in the obvious manner by changing the elements of the sets $I_1, \dots, I_n$.

Then the free operad in algebras generated by $S$, say $\mathcal{P}$, is defined such that $\mathcal{P}(N)$ is the algebra generated by $\tilde{S}(N)$, modulo relations

$$\left[s^{I_1,\dots,I_n}, s^{J_1,\dots,J_k}\right] = 0 \qquad \text{if } (I_1 \sqcup \cdots \sqcup I_n) \cap (J_1 \sqcup \cdots \sqcup J_k) = \emptyset \text{ or } \exists i : J_1 \sqcup \cdots \sqcup J_k \subset I_i. \tag{3.1}$$

Here the symbols $s^{I_1,\dots,I_n}$ should be thought of as operadic compositions of (the image of) $s$ with units, e. g., for $N = 8$

$$s^{23,456,7} = \mu(1_8, \mu(s, 1_2, 1_3, 1_1), 1_1)$$

where $\mu$ is the operadic composition as in eqn. (**??**), and $1_n \in \mathcal{P}(n)$ is the unit. (Here we abbreviate $s^{23,456,7} = s^{\{2,3\},\{4,5\},\{7\}}$.)

To define the operad structure it suffices to specify the compositions

$$\mathcal{P}(n) \otimes \mathcal{P}(k_1) \otimes \cdots \otimes \mathcal{P}(k_n) \to \mathcal{P}(k_1 + \cdots + k_n)$$

on generators, i. e., we may pick a unit in all but one factor in the tensor product. But given our interpretation of the symbols $s^{\cdots}$ it is then clear how to define the composition. The commutativity relations (3.1) are a translation of the condition that the factors in the tensor product above must commute.

Universality of his construction follows since we have not used any relation except those coming from the axioms of an operad in algebras.

**Example 3.11.** Similarly, the free operad in Lie algebras $\mathcal{P}$ generated by $S \in \mathrm{Set}^{\mathbb{S}}$ may be defined. $\mathcal{P}(N)$ is the Lie algebra generated by $\tilde{S}(N)$, with relations (3.1).

**Example 3.12.** Let $S$ be the $\mathbb{S}$-module in Set such that $S(2) = \{t\}$ and $S(n) = \emptyset$ for $n \neq 2$. Then the Drinfeld Kohno operad in Lie algebras $\mathfrak{t}$ from Example **??** is the free operad in Lie algebras generated by $S$ modulo the single relation

$$t^{1,23} = t^{1,2} + t^{1,3}.$$

(Here we abbreviate $t^{1,23} := t^{\{1\},\{2,3\}}$ etc.)

**Example 3.13.** Free operads in groupoids are a bit more tricky to define. Clearly there is a forgetful functor

$$(\text{Operads in groupoids}) \to (\mathrm{Set}^{\mathbb{S}} \rightrightarrows \mathrm{Set}^{\mathbb{S}}).$$

We want to construct its left adjoint. For concreteness, fix an element $M \rightrightarrows O$ on the right hand side, and call our free operad in groupoids to be constructed $\mathcal{P}$. We set $\mathrm{Ob}\mathcal{P} = \mathcal{F}_{\mathrm{Op}}(O)$. Let us next construct generating sets for morphisms. Analogously to Example 3.10 above, let us define an $\mathbb{S}$-module $\tilde{M}$ in sets as follows. Elements of $\tilde{M}(N)$ are equivalence classes of tuples

$$(m, \alpha_0, \alpha_1, \dots, \alpha_n)$$

where $m \in M(n)$, $n < N$, $\alpha_0 \in \mathcal{F}_{\mathrm{Op}}(O)(\{0\} \cup I_0)$, $\alpha_j \in \mathcal{F}_{\mathrm{Op}}(O)(I_j)$ $(j = 1, \dots, n)$, and where we fixed an underlying partition $[N] = I_0 \sqcup \cdots \sqcup I_n$. There is an equivalence relation as follows

$$(m \cdot \sigma, \alpha_0, \alpha_1, \dots, \alpha_n) \equiv (m, \alpha_0, \alpha_{\sigma(1)}, \dots, \alpha_{\sigma(n)})$$

for $\sigma \in \mathbb{S}_n$ (TODO: inverse?). Again a tuple as above shall be understood as a suitable composition of $m$ (or rather the image of $m$ in $\mathcal{P}$) with identity morphisms. there are two maps

$$\tilde{M} \rightrightarrows \mathcal{F}_{\mathrm{Op}}(O)$$

by mapping $(m, \alpha_0, \alpha_1, \dots, \alpha_n)$ to the compositions

$$\alpha_0 \circ_0 (s(m) \circ (\alpha_1, \dots, \alpha_n)) \in \mathcal{F}_{\mathrm{Op}}(O)(N)$$
$$\alpha_0 \circ_0 (t(m) \circ (\alpha_1, \dots, \alpha_n)) \in \mathcal{F}_{\mathrm{Op}}(O)(N).$$

We define the $\mathcal{P}(N)$ to be the groupoid generated by $\tilde{M}(N) \rightrightarrows \mathcal{F}_{\mathrm{Op}}(O)(N) = P(N)$, with certain commutativity conditions resembling (3.1). They are a bit cumbersome to write down, and we have encoded them in the pictures **??**.

**Example 3.14.** Note that in the previous example the generating morphisms $M$ came with two maps to $O$. Later we will like to have the additional freedom to specify generating morphism between objects in $\mathcal{F}_{\mathrm{Op}}(O)$, i. e., as initial data we are given the $O \in \mathrm{Set}^{\mathbb{S}}$ and $M \in \mathrm{Set}^{\mathbb{S}}$ with arrows $M \rightrightarrows \mathcal{F}_{\mathrm{Op}}(O)$. We may still talk about an operad in groupoids generated by these data, and the construction can be extracted from that in the previous example as follows:

- For every element $x \in \mathcal{F}_{\mathrm{Op}}(O)(n)$ in the image of $M(n) \rightrightarrows \mathcal{F}_{\mathrm{Op}}(O)(n)$ add an extra element $o_x$ to the generating set of objects $O(n)$. Call the resulting $\mathbb{S}$-module $O'$.
- Form the free operad $\mathcal{P}$ in groupoids generated by $\mathcal{M} \rightrightarrows O'$.
- Take its quotient modulo the relation $x = o_x$ in $\mathcal{P}(n)$. This quotient is be the operad we look for.

**Exercise 3.4.** Show that the operad in groupoids $\mathsf{CoB}$ of colored braids is generated (over the base $Ass_{\mathrm{Set}}$) by a single generator $b^{1,2} : 12 \to 21$, and relations

$$b^{1,23} = b^{13} b^{12}$$
$$b^{12,3} = b^{13} b^{23}$$

TODO: check whether both are necessary.

## 3.6 Operads governing monoidal categories

The goal of this section is to define and examine the operads governing monoidal and braided monoidal categories. In fact, we will restrict to the non-unital versions for simplicity.

Recall the operad (in Set) $PaP$ from example **??**. Clearly there is a map of operads $PaP \to Ass_{\mathrm{Set}}$. We define the operad in groupoids Mon to be the fiber pair (operad in) groupoid(s) of $PaP \to Ass_{\mathrm{Set}}$. So ObMon $= PaP$, and between two objects of Mon$(n)$ there is exactly one morphism if they encode the same permutation after forgetting the parenthesation, and no morphisms otherwise.

The justification of the name Mon comes from the following Lemma.

**Lemma 3.1.** *Consider* Mon *as an operad in categories. Then* Mon*-algebras are the same as (small) non-unital monoidal categories.*

*Proof.* We have to show two things: every Mon-algebra is a non-unital monoidal category and every such category is a Mon-algebra. Let us start with the latter statement, so assume a $\mathcal{C}$ is a non-unital monoidal category. To construct the Mon-algebra structure we have to provide functors

$$\rho_n : \mathrm{Mon}(n) \times \mathcal{C} \times \cdots \times \mathcal{C} \to \mathcal{C} \tag{3.2}$$

in a way compatible with the composition in Mon. On objects, it suffices to define the functor on the generator (12) of ObMon, and then

$$(12, A, B) \to A \otimes B$$

for $A, B \in \mathcal{C}$. For all other objects in ObMon the map is defined by the operadic compositions. For example

$$((13)2, A, B, C) \to (A \otimes C) \otimes B.$$

More generally, denote the $n-1$-fold tensor product functor bracketed according to $o \in \mathrm{ObMon}$, by $\otimes_o$. Then (3.2) is defined on objects as

$$(o, A_1, \dots, A_n) \mapsto \otimes_o(A_1, \dots, A_n).$$

Next we need to define (3.2) on morphisms. Recall that there is one morphism between every pair of objects in Mon that correspond to the same permutation. Fix two such objects $o, o' \in \mathrm{Ob}Mon(n)$, and let $a : o' \to o$ be the unique morphism between them. We may use the associator $\alpha$ from $\mathcal{C}$ to build a natural transformation $\alpha_{oo'} : \otimes_{o'} \Rightarrow \otimes_o$. We define (3.2) on morphisms as

$$(a, f_1, \dots, f_n) \mapsto \otimes_o(f_1, \dots, f_n) \circ (\alpha_{oo'})_{\otimes_{o'}(A'_1,\dots,A'_n)} = (\alpha_{oo'})_{\otimes_{o'}(A_1,\dots,A_n)} \circ \otimes_{o'}(f_1, \dots, f_n)$$

where $f_j \in \mathrm{Hom}_{\mathcal{C}}(A'_j, A_j)$ for $j = 1, \dots, n$. We claim that these assignments are a functor, i.e., preserve the identity morphisms and commute with compositions. The first fact is easy (since $\alpha_{oo} = id$), let us only show the second. Because the $\alpha$ are natural transformations, this amounts to verifying that $\alpha_{oo'} \circ \alpha_{o'o''} = \alpha_{oo''}$. However, note that both sides are natural transformations $\otimes_{o''} \to \otimes_o$ build (formally) using the associator $\alpha$ and hence MacLane's coherence Theorem says that both are identical.

To see that these formulas define a Mon algebra structure, we have to verify that the two double compositions

$$\mathrm{Mon}(n) \times Mon(k_1^1, \dots, k_{r_1}^1) \times \cdots Mon(k_1^n, \dots, k_{r_n}^n) \times \mathcal{C} \times \cdots \times \mathcal{C} \rightrightarrows \mathcal{C}$$

agree. On objects, they agree by construction since $\mathrm{ObMon} = PaP$ is a free operad. On morphisms this again reduces to checking that two natural transformations between iterated tensor product functors that are build using $\alpha$ agree. But this again is guaranteed by the MacLane coherence Theorem.

Let us also show the other direction, namely that a Mon-algebra $\mathcal{C}$ is a monoidal category. To do that, we have to define the tensor product and the associator $\alpha$, and check that the pentagon equation is satisfied. First, we define the bifunctor

$$\cdot \otimes \cdot := \rho_2((12), \cdot, \cdot)$$

where $\rho_2$ is the Mon-algebra structure as in (3.2). Secondly, we define for $A, B, C \in \mathrm{Ob}\mathcal{C}$

$$\alpha_{A,B,C} : (A \otimes B) \otimes C \to A \otimes (B \otimes C)$$
$$\alpha_{A,B,C} = \rho_3(f, id_A, id_B, id_C)$$

where $f$ is the unique morphism in $\mathrm{Hom}_{\mathrm{Mon}(3)}((12)3, 1(23))$. Since $\rho_3$ is a map of categories it follows that $\alpha$ is natural. It remains to check commutativity of the pentagon diagram. However, the pentagon diagram is the image of a diagram in Mon(4), which commutes since all diagrams in Mon(4) commute. Hence the pentagon equation is satisfied. □

Also, the MacLane coherence Theorem may be translated into operadic language.

**Proposition 3.1** (Variant of MacLane's coherence Theorem)**.** *The operad* Mon *is (isomorphic to) the (reduced) operad in groupoids with objects PaP generated by (see Example* **??***) a single morphism*

$$\alpha^{1,2,3} : (12)3 \to 1(23)$$

*with relations the pentagon relation*

$$\alpha^{1,2,34} \circ \alpha^{12,3,4} = \alpha^{2,3,4} \circ \alpha^{1,23,4} \circ \alpha^{1,2,3}. \tag{3.3}$$

The following prof is essentially a copy of MacLane's proof of his coherence Theorem.

*Proof.* Temporarily call the operad in groupoids generated by the above relation $\mathcal{P}$. Our goal is to show $\mathrm{Mon} = \mathcal{P}$. Certainly this is true on objects by definition. Next we need to show that the spaces of morphisms agree. Since $\alpha$ only maps between parenthesized permutations with the same underlying permutation, $\mathrm{Hom}_{\mathcal{P}}(o', o) = \emptyset$ if $o$ and $o'$ have different underlying permutations. We have to show that otherwise $|\mathrm{Hom}_{\mathcal{P}}(o', o)| = 1$. By $\mathbb{S}_n$ equivariance we may in fact assume that the underlying permutation is the trivial one and consider parenthesations of the trivial permutation henceforth.

We consider the sub-operad in categories $\mathcal{P}'' \subset \mathcal{P}$ generated by the same generators and relations, but only as an operad in categories, not as an operad in groupoids. In other words, in $\mathcal{P}''$ the morphism

$\alpha$ is not invertible, there is no $\alpha^{-1}$. In yet other words, one may only associate to the right, never to the left. It is easy to see that in fact $\mathcal{P}'''(n)$ is a poset, where $o' \geq o$ iff there exists a morphism $o' \to o$. Let $\mathcal{P}'(n)$ be the full subcategory (and sub-poset) formed by parenthesized permutations whose underlying permutation is the identity. It has a unique lowest element, namely $o_n := 1(2(3(4(\cdots n)\cdots)$.

Reduction 1: To show the Proposition, it suffices to check that for each $n$ and each $o' \in \mathcal{P}'(n)$ there is a unique arrow $o' \to o_n$ in $\mathcal{P}'(n)$. (In category theorists' slang: $o_n$ is a *final object.*)

Proof of Reduction 1: Denote by $f_o : o \to o_n$ the (unique by assumption) arrow in $\mathcal{P}'(n)$. Let $f : o' \to o$ be some arrow in $\mathcal{P}(n)$ (where $o, o' \in \mathrm{Ob}\mathcal{P}'(n)$). We claim (Claim 1) that $f_{o'} = f_o \circ f$. If we can show the Claim 1 then Reduction 1 follows since then $f = f_o^{-1} \circ f_{o'}$ irrespective of the $f$ chosen. Also, Claim 1 is true by assumption as long as $f$ is in $\mathcal{P}'(n)$. But any morphism $f : o' \to o$ is a composition of morphisms in $\mathcal{P}'(n)$ and their inverses, say

$$f = g_1 \circ g_2^{-1} \circ g_3 \circ \cdots \circ g_n$$

where each $g_j : o'_j \to o_j$ is a morphism in $\mathcal{P}'(n)$. (Here $o_1 = o$, $o'_1 = o'_2$, etc.) But inserting $g_j = f_{o_j}^{-1} \circ f_{o'_j}$ we obtain

$$f = f_o^{-1} \circ f_{o'_1} \circ f_{o'_1}^{-1} \circ f_{o_2} \circ f_{o_2}^{-1} \circ f_{o'_3} \circ \cdots \circ f_{o'} = f_o^{-1} \circ f_{o'}.$$

Hence Reduction 1 is shown, and we can focus on showing the assumption therein.

We do this by a two-fold induction: The outer induction is on $n$. We start it by noticing that for $n \leq 2$ there is nothing to be shown and that for $n = 3$ the statement is trivial. The inner induction is done by the partial order on $\mathcal{P}'(n)$. For the lowest object (i. e., for $o_n$) the statement is trivial. Suppose that $o \in \mathrm{Ob}\mathcal{P}'(n)$ is given and we want to show that there is a unique arrow $o \to o_n$ in $P'(n)$. By the induction hypothesis we may assume this is true for all $o' < o$. It is also clear that there is an arrow, we only have to show uniqueness. So suppose there are two arrows that factor as

$$\begin{tikzcd} & o \arrow[dl] \arrow[dr] & \\ o' \arrow[dr] & & o'' \arrow[dl] \\ & o_n & \end{tikzcd}$$

where we can assume that $o \to o'$ and $o \to o''$ are obtained by a single application of the associator $\alpha$. By the induction hypothesis we are done if we can show that the diagram may be completed to a diagram

$$\begin{tikzcd} & o \arrow[dl] \arrow[dr] & \\ o' \arrow[dr] \arrow[ddr] & & o'' \arrow[dl] \arrow[ddl] \\ & o''' \arrow[d] & \\ & o_n & \end{tikzcd}$$

in such a way that the upper diamond commutes. (Note that by the induction hypothesis all other triangles commute.) Now suppose $o = (a)(b)$. For each of the maps $f : o \to o'$, $g : o \to o''$ there are 3 choices:

1. The associator acts within $a$.
2. The associator acts within $b$.
3. The associator acts on the top level, i.e., $a = (a')(a'')$ and the associator acts as $(a'a'')b \to a'(a''b)$.

So there are $3^2 = 9$ cases to consider. If both $f$ and $g$ act on the top level, $f = g$ and we are done. If both $f$ and $g$ act on $a$ we may set $o''' = (o_{n'})(b)$, where $n'$ is the arity of $a$. The diamond is then obtained from a diamond in lower $n$ by the operad maps, and by the (outer) induction hypothesis it commutes. The analogous argument holds when both $f$ and $g$ act on $b$. Next suppose one of $f$ and $g$ acts on $a$ and the other on $b$, say $f : (a)(b) \to (a')(b)$ and $g : (a)(b) \to (a)(b')$. Concretely, this means that $f$ and $g$ are obtained from some $f' : a \to a'$ and $g' : b \to b'$ through the operadic compositions as

$$f = \mu((12), f', id) \qquad\qquad g = \mu((12), id, g').$$

In this case we may set $o''' = (a')(b')$, and let the morphism $o' \to o'''$ be $\mu((12), id, g')$ and $o'' \to o'''$ be $\mu((12), f', id)$. Then the diamond commutes by the operad axioms. (Concretely, both compositions equal $\mu((12), f', g')$.)

Next assume that one of $f$, $g$, say $g$, acts on $b$, and the other ($f$) acts on the top level. It means that $a = (a')(a'')$, $f : (a'a'')b \to a'(a''b)$ and $g : b \to b'$. In this case we may set $o''' = a'(a''b')$. A similar argument as before constructs the arrows to $o'''$ and shows that the diamond commutes.

Note that we have not used the pentagon axiom so far. It is needed for the final case, namely that $f$ acts on $a$ and $g$ acts on the top level, $g : (a'a'')b \to a'(a''b)$. Here one has to distinguish 3 subcases:

1. $f$ acts within $a'$.
2. $f$ acts within $a''$.
3. $a' = cc'$ and $f$ act as $((cc')a'')b \to (c(c'a''))b$.

The first two cases are handled as before. For the last case we use the pentagon identity to complete the diamond, i. e., we set $o''' = c(c'(a''b))$. The morphism $o'' \to o'''$ is set to equal one edge of the pentagon, and the morphism $o' \to o'''$ is the composition of the remaining yet unused two. Commutativity of the pentagon yields commutativity of the diamond and hence we are done. □

Let us turn to (non-unital) braided monoidal categories. Recall that the colored braids $\mathsf{CoB}$ form an operad in groupoids, with base $Ass_{\mathrm{Set}}$. We build the operad $PaB$ ("parenthesized braids") as the base change of $\mathsf{CoB}$ over $PaP \to Ass_{\mathrm{Set}}$, cf. Example **??**. Concretely, objects in $PaB(n)$ are elements of $PaP(n)$ (parenthesized permutations). Morphisms between two such objects are braids which take one permutation into the other.

**Lemma 3.2.** *Consider $PaB$ as an operad in categories. Then $PaB$-algebras are exactly (small) non-unital braided monoidal categories.*

*Proof.* The proof is similar to the one above, using the braided version of the MacLane Coherence Theorem. □

**Proposition 3.2.** *The operad in groupoids $PaB$ is the operad in groupoids over the base $PaP$ generated by two morphisms*

$$\alpha^{1,2,3} : (12)3 \mapsto 1(23)$$
$$\gamma^{1,2} : 12 \mapsto 21$$

*with relations the pentagon relation* (**??**) *and additionally the two hexagon relations*

$$\gamma^{1,23} = (\alpha^{2,3,1})^{-1}\gamma^{1,3}\alpha^{2,1,3}\gamma^{1,2}(\alpha^{1,2,3})^{-1}$$
$$\gamma^{23,1} = \alpha^{1,2,3}\gamma^{2,1}(\alpha^{2,1,3})^{-1}\gamma^{3,1}\alpha^{2,3,1}$$

*Proof.* Temporarily denote by $\mathcal{P}$ the operad in groupoids generated by the above generators and relations. There is clearly a map of operads in groupoids $\mathcal{P} \to PaB$ and we want to show this is an isomorphism. It actually suffices to show that $\mathrm{Hom}_{\mathcal{P}(n)}(o_n, o_n) \to \mathrm{Hom}_{PaB(n)}(o_n, o_n)$ is an isomorphism for all $n$, where $o_n$ is as in the proof of Proposition **??**. To see that it is a bijection, it suffices to check that the full subgroupoid with objects those parenthesized permutations with parenthesation $\cdot(\cdot(\cdot(\cdots)))))$ is isomorphic to $\mathsf{CoB}(n)$. Recall the description of $\mathsf{CoB}(n)$ by generators and relations from Exercise **??**. We construct a map $\mathsf{CoB}(n) \to \mathcal{P}(n)$ by sending each object (a permutation $\sigma$) to the parenthesized permutation obtained by endowing $\sigma$ with the aforementioned (rightmost) parenthesation. We define the map on morphisms by sending the generator $(b_i, \sigma)$ to $f \circ \gamma^{\sigma(i),\sigma(i+1)} \circ f'$, where $f$ and $f'$ are morphisms in $\mathrm{Mon}(n) \subset \mathcal{P}(n)$ restoring the parenthesation. (They are unique by Proposition **??**.) To show that this defines a map of groupoids $\mathsf{CoB}(n) \to \mathcal{P}(n)$ one has to verify the braid relations. This is done as in Exercise **??**, where on uses Proposition **??** to remove occurring associators, see Figure **??** for an illustration. One easily checks that the composition $\mathsf{CoB}(n) \to \mathcal{P}(n) \to PaB(n)$ is the usual full inclusion of $\mathsf{CoB}(n)$ in $PaB(n)$ and hence we are done. □

**Example 3.15.** The operad *Lie* of Lie algebras.

**Exercise 3.5.** Define the operad $\mathsf{Grp}$ whose algebras are groups. Show that

$$|\mathsf{Grp}(n)| = 2^n n!$$

## 3.7 Little $n$-cubes operads

Historically, the first studied examples of operads were in fact in the category $\mathcal{T}\wr_{\surd}$ of topological spaces.

Let

$$\mathsf{LD}_n(N) = \{\text{space of squarilinear embeddings } (0,1)^n \sqcup \cdots \sqcup (0,1)^n \to (0,1)^n\}.$$

Here "squarilinear" means that the map on each cube must have the form

$$(x_1, \dots, x_n) \mapsto \lambda \cdot ((x_1 + a_1, \dots, x_n + a_n))$$

for some real numbers $\lambda, a_1, \dots, a_n$. A point in $\mathsf{LD}_1(3)$ and one in $\mathsf{LD}_2(3)$ is depicted in Figure **??**. There are natural compositions of morphisms depicted in Figure **??** which make the collection of spaces $\mathsf{LD}_n(N)$ into an operad $\mathsf{LD}_n$.

These operads received a lot of interest from topologist because of a Theorem proved by May saying roughly that "algebras over $\mathsf{LD}_n$ are the same as $n$-fold loop spaces". More concretely:

**Theorem 3.1.** *TODO*

### 3.7.1 A variant: operad of configuration spaces $\mathsf{FM}_n$

Often it is convenient to make the "big cube" in the little cubes operad of infinite size and the little cubes of zero size. The resulting object would then be called the configuration space of points in $\mathbb{R}^n$. Unfortunately, as it stands these configuration spaces do not form an operad. Here is one way to repair this defect.

### 3.7.2 Homotopy type of $\mathsf{LD}_n$

There are fibrations

$$\mathrm{Conf}_1(\mathbb{R}^n \setminus \{p_2, \dots, p_N\}) \to \mathsf{LD}_n(N) \to \mathsf{LD}_n(N-1).$$

Here the right hand map forgets the position of the first point. The left hand side (i.e., the fiber) is the configuration space of one point in the $\mathbb{R}^n$ with $N-1$ punctures. The fiber may be contracted to an $N-1$ fold wedge product of $n-1$-spheres. In principal, one may use the resulting long exact sequence of homotopy groups to compute the homotopy groups of $\mathsf{LD}_n(N)$ for all $n$ and $N$. However, this involves knowing the homotopy groups of spheres which are very difficult to compute in general.

For $n = 2$ the answer is simple. The $i$-th homotopy groups of the fiber (a wedge of circles) is trivial for $i \geq 2$. It then follows by induction from the long exact sequence

$$\cdots \to \pi^i(S^1 \wedge \cdots \wedge S^1) \to \pi^i(\mathsf{LD}_2(N)) \to \pi^i(\mathsf{LD}_2(N-1)) \to \pi^{i-1}(S^1 \wedge \cdots \wedge S^1) \to \cdots$$

that the same holds for all $\mathsf{LD}_2(N)$. Hence the $\mathsf{LD}_2(N)$ are in fact $K(\pi, 1)$ spaces. Here the fundamental group $\pi$ is the pure braid group $\mathsf{PB}_N$. From the exact sequence we can in fact read off an interesting property of the pure braid groups. Since we know that $\pi^1(S^1 \wedge \cdots \wedge S^1)$ is a free group in $n$ generators, we obtain the exact sequence

$$1 \to \mathcal{F}_{\mathrm{Grp}}(X_2, \dots, X_N) \to \mathsf{PB}_N \to \mathsf{PB}_{N-1} \to 1.$$

In fact, this sequence is split; there is a natural map $\mathsf{PB}_{N-1} \to \mathsf{PB}_N$. Let us summarize these findings:

**Proposition 3.3.** $\mathsf{LD}_2(N)$ *(where $N \geq 1$) are $K(\pi, 1)$ spaces, i.e.,*

$$\pi^i(\mathsf{LD}_2(N)) = \begin{cases} \mathsf{PB}_N & \text{for } i = 1 \\ 1 & \text{otherwise.} \end{cases}$$

*Furthermore the pure braid groups may be written recursively as semidirect products*

$$\mathsf{PB}_N \cong \mathsf{PB}_{N-1} \ltimes \mathcal{F}_{\mathrm{Grp}}(X_2, \dots, X_N)$$

*where $\mathsf{PB}_{N-1} \subset \mathsf{PB}_N$.*

The subgroup $\mathsf{PB}_{N-1} \subset \mathsf{PB}_N$ may be understood as the subgroup of braids “leaving the first strand alone”, while $X_j = x_{1j}$ may be understood as the braid obtained by wrapping the first strand around the $j$-th..

### 3.7.3 (Co-)Homology of $\mathsf{LD}_n$

**Definition 3.5.** *Let* $\mathsf{Fin}$ *be the groupoid of finite sets with bijections as morphisms. An* $\mathbb{S}$*-module in some category* $\mathcal{C}$ *is a contravariant functor* $\mathsf{Fin} \to \mathcal{C}$*. We denote the category of* $\mathbb{S}$*-modules in* $\mathcal{C}$ *by* $\mathcal{C}_{\mathbb{S}}$

Another name for $\mathbb{S}$-modules is symmetric sequences. Of course, the $\mathbb{S}$-module $\mathcal{P}$ is completely determined by giving for each $n = 0, 1, 2, \dots$ an object $\mathcal{P}(n) = \mathcal{P}(\{1, 2, \dots, n\})$ of $\mathcal{C}$ together with a right action of the symmetric group $\mathbb{S}_n$.

**Example 3.16.** The $\mathbb{S}$-module (in $\mathfrak{K}$-vector spaces) *Com* is defined by setting $Com(n) = \mathbb{R}$ for each $n = 1, 2, \dots$.

Next suppose that the category $\mathcal{C}$ is in fact symmetric monoidal, for example $\mathcal{C}$ is the category of $\mathbb{K}$-vector spaces. We also assume quietly that $\mathcal{C}$ has all small limits and colimits. Then we may equip $\mathcal{C}_{\mathbb{S}}$ with a monoidal product $\boxtimes$.

**Definition 3.6.** *Let* $\mathcal{P}, \mathcal{Q} \in \mathcal{C}_{\mathbb{S}}$*. Then we define* $\mathcal{P} \boxtimes \mathcal{Q} \in \mathcal{C}_{\mathbb{S}}$ *as follows. For an object (i. e., a finite set)* $S \in \mathrm{Ob}(\mathcal{C})$ *we set*

$$(\mathcal{P} \boxtimes \mathcal{Q})(S) = \coprod$$

**Remark 3.2.** The product $\boxtimes$ is called plethysm. (...in rep. th.)

**Definition 3.7.** *An operad* $\mathcal{P}$ *in some symmetric monoidal category* $\mathcal{C}$ *is monoid in* $(\mathcal{C}_{\mathbb{S}}, \boxtimes)$*.*

# Chapter 4

# The Grothendieck-Teichmüller group

## 4.1 Another look at the definition of GT

**Definition 4.1.** *The Grothendieck-Teichmüller group* GT *is the group of automorphisms of the operad in pro-unipotent groupoids* $\widehat{PaB}$ *which are the identity on objects,*

$$\mathrm{GT} := \mathrm{Aut}(\widehat{PaB}).$$

**Remark 4.1.** We restrict to automorphism which are the identity on objects to adhere to custom conventions. The full automorphism group is in our case $\mathbb{Z}_2 \ltimes \mathrm{Aut}(\widehat{PaB})$, where the non-trivial morphism in $\mathbb{Z}_2$ rotates a braid by 180 degrees. It corresponds to the fact that for a braided monoidal category the opposite of the monoidal product is also a monoidal product.

**Proposition 4.1.** *Elements of* GT *are in one-to one correspondence with pairs* $(\lambda, f) \in \mathbb{K}^\times \times \hat{\mathcal{F}}_{\mathrm{Grp}}(X, Y)$*, that satisfy the following equations*

*Proof.* By the universal property of the pro-finite completion (Definition **??**) we know that endomorphisms of $\widehat{PaB}$ are in one-to-one correspondence with maps of operads in groupoids

$$\phi : PaB \to \widehat{PaB}.$$

By Proposition **??** we however know that any such morphism is uniquely determined by specifying the images of the generators $\gamma$ and $\alpha$. Note that $\gamma^{1,2} \in \mathrm{Hom}_{\widehat{PaB}(2)}(12, 21) \cong \widehat{\mathsf{PB}}_2\gamma^{1,2}$. Since $\widehat{PaB}(2)$ is the pro-unipotent completion of a free group in one generator, there must be $\lambda \in \mathbb{K}$ such that

$$\phi(\gamma) = (s_{1,2})^m \gamma^{1,2}$$

where $m = (\lambda - 1)/2$.[1] In order for $\phi$ to be invertible, we must necessarily have $\lambda \in \mathbb{K}^\times$. Let us denote $\phi(\gamma) =: \tilde{f}\gamma$, where $\tilde{f} \in \widehat{\mathsf{PB}}_3$. Recalling the structure of $\mathsf{PB}_3$ from Exercise 2.21, we must have

$$\tilde{f} = f(x_{12}, x_{23})C^\mu$$

where $C$ is the central element and $\mu \in \mathbb{K}$. In order for the data $(\lambda, \tilde{f})$ to generate a morphism $PaB \to \widehat{PaB}$ it is necessary and sufficient that the images of the relations (see Proposition **??**) hold in $\widehat{PaB}$, i. e., that we have

$$\gamma^{1,23} x_{1,23}^m = (\tilde{f}^{2,3,1})^{-1} \gamma^{1,3} \tilde{f}^{2,1,3} \gamma^{1,2} (\tilde{f}^{1,2,3})^{-1} \tag{4.1}$$

$$\gamma^{12,3} x_{12,3}^m = \tilde{f}^{3,1,2} \gamma^{1,3} (f^{1,3,2})^{-1} \gamma^{2,3} f^{1,2,3} \tag{4.2}$$

$$\tilde{f}^{1,2,34} \tilde{f}^{12,3,4} = \tilde{f}^{2,3,4} \tilde{f}^{1,23,4} \tilde{f}^{1,2,3}. \tag{4.3}$$

[1] The apparently strange an unnecessary change of variables from $\lambda$ to $m$ (or vice versa) is made such that if one composes two such automorphisms (say with parameters $\lambda$, $\lambda'$) the parameters just multiply, i. e., the new parameter is $\tilde{\lambda} = \lambda\lambda'$. Exercise: Verify this, using that $s_{12} = \gamma^{2,1}\gamma^{1,2}$.

Let us rewrite the first two equations (TODO: INTRODUCE NOTATION).

$$\sigma_2\sigma_1(Cx_{23}^{-1})^m = f^{-1}\sigma_2 x_{23}^m f \sigma_1 x_{12}^m f^{-1} C^{-\mu}$$
$$\sigma_1\sigma_2(Cx_{12}^{-1})^m = f\sigma_1 x_{12}^m f^{-1}\sigma_2 x_{23}^m f C^{\mu}.$$

Here we used that $x_{1,23} = Cx_{23}^{-1}$ and $x_{12,3} = Cx_{12}^{-1}$ and abbreviated $f = f(x_{12}, x_{23})$. (Note also, that by definition the central element $C$ commutes with everything.) We multiply both sides of the first equation by $(\sigma_2\sigma_1)^{-1}$ from the left, and we multiply both sides of the second equation by $\sigma_1^{-1}$ from the left and by $\sigma_2^{-1}$ from the right.

$$(Cx_{23}^{-1})^m = \sigma_1^{-1}\sigma_2^{-1} f^{-1}\sigma_2\sigma_1(\sigma_1^{-1}x_{23}\sigma_1)^m(\sigma_1^{-1}f\sigma_1)x_{12}^m f^{-1}C^{-\mu}$$
$$(Cx_{13}^{-1})^m = (\sigma_1^{-1}f\sigma_1)x_{12}^m f^{-1}x_{23}^m(\sigma_2 f\sigma_2^{-1})C^{\mu}.$$

Now by Exercise 4.1 we have

$$\sigma_1^{-1}f\sigma_1 = f(x_{12}, x_{13}) \qquad \sigma_2 f\sigma_2^{-1} = f(x_{13}, x_{23})$$
$$(\sigma_2\sigma_1)^{-1}f\sigma_2\sigma_1 = f(x_{23}, x_{13}).$$

Inserting this into the previous equation and abbreviating $X = x_{12}$, $Y = x_{13}$, $Z = x_23$ we obtain (after moving all factors to the right hand sides)

$$1 = Z^m f(Z,Y)^{-1}Y^m f(X,Y)X^m f(X,Z)^{-1}C^{-\mu-m} \quad = Y^m f(X,Y)X^m f(X,Z)^{-1}Z^m f(Z,Y)^{-1}C^{-\mu-m}$$
$$1 = Y^m f(X,Y)X^m f(X,Z)^{-1}Z^m f(Y,Z)C^{\mu-m}.$$

Note that this is an equation in $\widehat{\mathsf{PB}}_3$ and by Exercise 2.21 the $X, Y, Z, C$ are free variables, except for the single relation $C = XYZ$. Equating both 1's one cancelling terms on the left we obtain

$$f(Z,Y)^{-1}C^{-\mu} = f(Y,Z)C^{\mu}.$$

It follows that $\mu = 0$ and that $f(Z,Y)^{-1} = f(Y,Z)$, i.e., the antisymmetry equation. Inserting this back in we obtain the remaining equation

$$1 = Y^m f(X,Y)X^m f(Z,X)Z^m f(Y,Z)C^{-m} = Y^m f(X,Y)X^m f(C\tilde{Z},X)\tilde{Z}^m f(Y,C\tilde{Z})$$

where we defined $\tilde{Z} = C^{-1}Z$. Note that there are always $\alpha, \beta \in \mathbb{K}$ so that $f(CX,Y) = C^{\alpha}f(X,Y)$ and $f(X,CY) = C^{\beta}f(X,Y)$. Hence from the antisymmetry relation we see that $\alpha+\beta = 0$ and the remaining hexagon equation becomes equivalent to

$$1 = Y^m f(X,Y)X^m f(\tilde{Z},X)\tilde{Z}^m f(Y,\tilde{Z})$$

where $XY\tilde{Z} = 1$.

Finally consider the pentagon equation. Using that

$$f^{1,2,3} = f(x_{12}, x_{23}) \qquad f^{2,3,4} = f(x_{23}, x_{34})$$
$$f^{12,3,4} = f(x_{12,3}, x_{34}) = f(x_{13}x_{23}, x_{34}) \qquad f^{1,2,34} = f(x_{12}, x_{2,34}) = f(x_{12}, x_{34}x_{24})$$
$$f^{1,23,4} = f(x_{1,23}, x_{23,4}) = f(x_{12}x_{13}, x_{23}x_{34})$$

the equation becomes

$$f(x_{12}, x_{34}x_{24})f(x_{13}x_{23}, x_{34}) = f(x_{23}, x_{34})f(x_{12}x_{13}, x_{23}x_{34})f(x_{12}, x_{23}).$$

□

**Exercise 4.1.** Let $x_{12} = \sigma_1^2$, $x_{23} = \sigma_2^2$ and $x_{13} = \sigma_1^{-1}\sigma_2^2\sigma_1$. Verify that

$$\sigma_1^{-1}x_{23}\sigma_1 = x_{13} \qquad \sigma_2 x_{12}\sigma_2^{-1} = x_{13}$$
$$(\sigma_2\sigma_1)^{-1}x_{12}\sigma_2\sigma_1 = x_{23} \qquad (\sigma_2\sigma_1)^{-1}x_{23}\sigma_2\sigma_1 = x_{13}$$

**Remark 4.2.** In fact, it was shown by H. Furusho that the antisymmetry equation follows from the pentagon equation. To see that, apply the group morphism $\widehat{\mathsf{PB}}_4 \to \hat{\mathcal{F}}_{\mathrm{Grp}}(X,Y)$ worked out in Exercise 2.22 to both sides of the pentagon equation. Note that under that map

$$f^{1,2,3} = f(S_{12}, S_{23}) \mapsto f(X,Y) \qquad f^{2,3,4} = f(S_{23}, S_{34}) \mapsto f(Y,X)$$
$$f^{12,3,4} = f(S_{13}S_{23}, S_{34}) \mapsto f(X^{-1}Y^{-1}Y, X) = 1 \qquad f^{1,2,34} = f(S_{12}, S_{23}S_{24}) \mapsto f(X, YY^{-1}X^{-1}) = 1$$
$$f^{1,23,4} = f(S_{12}S_{13}, S_{24}S_{34}) \mapsto f(XX^{-1}Y^{-1}, Y^{-1}X^{-1}X) = 1.$$

What remains from the pentagon equation is the statement

$$1 = f(Y,X)f(X,Y)$$

i. e., antisymmetry.

## 4.2 Drinfeld associators

Recall from section ?? that we have a functor gr that sends a pro-unipotent Lie algebra to its associated graded under the lower-central series filtration. Recall that the associated graded of the pure braid Lie algebra $\mathfrak{pb}_n$ is the Drinfeld-Kohno Lie algebra $\mathfrak{t}_n$. The $\mathfrak{t}_n$ assemble to form an operad of Lie algebras $\mathfrak{t}$. It follows that $T := \mathrm{Exp}(\mathfrak{t})$ is an operad of pro-unipotent groups. We may consider it as an operad of pro-unipotent groupoids with all groupoids having one object. Recall also that $PaP$ was the operad in sets governing monoids. Its pair groupoid is an operad in groupoids $\widetilde{PaP}$. We define the operad in pro-unipotent groupoids

$$\widehat{GPaCD} := T \times \widetilde{PaP}.$$

The group of endomorphisms of any object in $\widehat{GPaCD}(n)$ is identified with

$$\mathrm{Exp}(\mathfrak{t}_n) = \mathrm{Gr}(\hat{U}\mathfrak{t}_n).$$

where $\mathfrak{t}_n$ is the Drinfeld-Kohno algebra. The topological Hopf algebra $\hat{U}\mathfrak{t}_n$ appears in knot theory and is called the algebra of chord diagrams. Elements may be identified with (possibly infinite) linear combinations of diagrams with $n$ strands and some "chords" drawn in between, see Figure ??. The generator $t_{ij}$ corresponds to a chord between strand $i$ and $j$. Consequently, the operad in topological Hopf algebras $\hat{U}\mathfrak{t}_n \times \widetilde{PaP}$ is often denoted by $\widehat{PaCD}$ ("parenthesized chord diagrams").

The (graded version of the) Grothendieck-Teichmüller group GRT is the group of automorphism of $\widehat{GPaCD}$ which are the identity on objects, i. e.,

$$\mathrm{GRT} := \mathrm{Aut}(\widehat{GPaCD})(\cong \mathrm{Aut}(\widehat{PaCD})).$$

The set Drinfeld associators DAss is the set of isomorphisms which are the identity on objects

$$\widehat{PaB} \to \widehat{GPaCD} = \mathrm{Gr}\widehat{PaCD}.$$

We will denote the "associator" element in $\widehat{GPaCD}(3)$ by $A : (12)3 \to 1(23)$ and the "braid" or rather transposition element in $\widehat{GPaCD}(2)$ by $X : 12 \to 21$. It is important to note that in this case $X^{2,1}X^{1,2} = id$, while for the braiding $\gamma$ we had before $\gamma^{2,1}\gamma^{1,2} \neq id$.

**Proposition 4.2** (Symmetric Monoidal Coherence Theorem). *The operad in groupoids $\widetilde{PaP}$ has the following description in terms of generators and relations.*

- *On objects, it is generated by one operation $12 \in \mathrm{Ob}\widetilde{PaP}(2)$ and no relations.*
- *On morphisms it is generated by one operation $A : (12)3 \to 1(23)$ and one operation $X : 12 \to 21$. The relations are the same relations as for PaB (two hexagon relations and the pentagon relation), plus the additional relation that*

$$X^{2,1} = (X^{1,2})^{-1}$$

The proof is essentially identical to the one of Proposition ??, with the role of the braid group replaced by the permutation group.

**Proposition 4.3.** *A Drinfeld associator is the same data as a pair* $(\mu, \Phi) \in \mathbb{K}^\times \times \mathbb{K}\langle\langle x, y\rangle\rangle$ *such that* $\Phi$ *is group-like (i.e.,* $\Delta\Phi = \Phi\hat{\otimes}\Phi$*) and furthermore*

$$\Phi(x, y) = \Phi(x, y)^{-1} \tag{4.4}$$

$$1 = e^{\frac{\mu}{2}z}\Phi(x, y)e^{\frac{\mu}{2}x}\Phi(y, z)e^{\frac{\mu}{2}y}\Phi(z, x) \tag{4.5}$$

$$\Phi(t_{12}, t_{23} + t_{24})\Phi(t_{13} + t_{23}, t_{34}) = \Phi(t_{23}, t_{34})\Phi(t_{12} + t_{13}, t_{24} + t_{34})\Phi(t_{12}, t_{23}). \tag{4.6}$$

*Here for the middle equation* $x + y + z = 0$ *and the last equation takes values in the universal enveloping algebra of the Drinfeld-Kohno Lie algebra* $\mathfrak{t}_4$ *with standard generators* $t_{ij}$*. The set of* Drinfeld associators DAss *is defined to be the set of pairs* $(\mu, \Phi)$ *solving these equations, with* $\mu \neq 0$*.*

We will denote be $\mathrm{DAss}_\mu$ the set of associators of the form $(\mu, \Phi)$.

*Proof.* As in the proof of Proposition 4.1, the map is uniquely determined by the image of the two generators. The generator $\gamma$ must be mapped to

$$e^{\mu t_{12}/2}X$$

for some number $\mu \in \mathbb{K}$. For the map to be invertible, we must have $\mu \neq 0$. The associator must be mapped to

$$\Phi(t_{12}, t_{23})e^{\lambda c}A$$

for some group-like element $\Phi(x, y) \in \mathbb{K}\langle x, y\rangle$ and some number $\lambda$, where $c = t_{12} + t_{23} + t_{13}$ is the central element. An almost identical derivation as that leading to Proposition 4.1 then shows that $\lambda = 0$ and that relations (4.4)-(4.6) hold.

Conversely, suppose that the data $(\mu, \Phi)$ are given and satisfy (4.4)-(4.6). We obtain a morphism $\widehat{PaB} \to \widehat{GPaCD}$. We have to show that it is bijective. It suffices to check that for each $n$ and fixed object in $PaB(n)$, say $o = 1(2(\cdots n)))$, the morphism restricts to an isomorphism of the endomorphism groups of $o$. However, these groups are $\widehat{\mathsf{PB}}_n$ and $T_n$ respectively. Equivalently, we have to check that the underlying maps $\mathfrak{pb}_n \to \mathfrak{t}_n$ are isomorphisms. One checks that $\log x_{ij} \mapsto \mu t_{ij} + (\dots)$ where $\dots$ are higher order terms. Since $\mu \neq 0$ this map is clearly surjective, since all generators are in the image. We nee to verify injectivity. Call the kernel of the above map $K$. By completeness of $\mathfrak{pb}_n$ it suffices to check that

$$K \subset C_j$$

for all $j$ where $C_j$ is the $j$-th term of the lower central series. Also recall that the relations on $\mathfrak{t}_n$ are just the leading terms (with respect to the number of commutators) of the relations on $\mathfrak{pb}_n$. So suppose $F(\{x_{ij}\}) \in K \cap C_j$ for some Lie series $F$. Then $F(\{t_{ij}\}) = 0$. But since the relations on $\mathfrak{t}_n$ are just the leading terms of the relations on $\mathfrak{pb}_n$ we have that in fact $F(\{x_{ij}\}) \in C_{j+1}$. This shows the claim and hence bijectivity. □

**Remark 4.3.** Note that a Drinfeld associator in particular determines isomorphisms $\mathfrak{pb}_n \to \mathfrak{t}_n$.

**Proposition 4.4.** *The Grothendieck-Teichmüller group* GRT *has the structure*

$$\mathrm{GRT} \cong \mathbb{K}^\times \ltimes \mathrm{GRT}_1$$

*where* $\mathrm{GRT}_1$ *is a pro-unipotent group whose elements may be identified with solutions* $\Phi$ *of equations* (4.4)-(4.6) *for* $\mu = 0$*.* $\mathbb{K}^\times$ *acts on such solutions by rescaling, i. e.,*

$$(\lambda \cdot \Phi)(x, y) := \Phi(\lambda x, \lambda y).$$

Concretely, $\mathrm{GRT}_1 = \mathrm{Exp}(\mathfrak{grt}_1)$ where the pro-nilpotent Lie algebra $\mathfrak{grt}_1$ may be identified with series $\psi \in \hat{\mathbb{F}}_{Lie}(x, y)$ that satisfy

$$\psi(x, y) = -\psi(y, x) \tag{4.7}$$

$$\psi(x, y) + \psi(y, z) + \psi(z, x) = 0 \tag{4.8}$$

$$\psi(t_{12}, t_{23}) - \psi(t_{12}, t_{23} + t_{24}) + \psi(t_{12} + t_{13}, t_{24} + t_{34}) - \psi(t_{13} + t_{23}, t_{34}) + \psi(t_{23}, t_{34}) = 0 \tag{4.9}$$

where $x + y + z = 0$ and the last equation takes place in $\mathfrak{t}_4$. Note also that $\mathfrak{grt}_1$ is graded.

*Proof.* Any automorphism of $\widehat{GPaCD}$ (which is the identity on objects) is determined by its action on $T$ and on $\widetilde{P}aP$. But since $\mathfrak{t}$ is generated by one generator $t_{12}$ and $\widetilde{P}aP$ is generated by $A$ and $X$, any automorphism is determined by the images of these objects. Clearly $t_{12}$ can only be sent to a non-zero multiple $\lambda t_{12}$. The map sending an isomorphism to this $\lambda$ is a character of GRT. Conversely, since $\mathfrak{t}$ is graded, rescaling $t_{12}$ by $\lambda \neq 0$ yields an automorphism of $\mathfrak{t}$ and hence of $\widehat{GPaCD}$ (so in particular $\mathbb{K}^\times \subset \mathrm{GRT}$). Hence $\mathrm{GRT} \cong \mathbb{K}^\times \rtimes \mathrm{GRT}_1$ where $\mathrm{GRT}_1 := \ker(\mathrm{GRT} \to \mathbb{K}^\times)$.

An element of $\mathrm{GRT}_1$ is determined by its action on $A$ and $X$. In general $X$ may be sent to $Xe^{\mu t_{12}}$ for some $\mu \in \mathbb{K}$. But by the relation $X^{2,1}X^{1,2} = id$ we must have $\mu = 0$. The associator may be sent to

$$A\Phi(t_{12}, t_{23})e^{\mu' c}$$

as before. Also as before, it must satisfy the hexagon and pentagon relations (for $\mu = 0$) and $\mu' = 0$. Conversely, given a $\Phi$ that satisfies (4.4)-(4.6) we obtain an endomorphism of $\widehat{GPaCD}$. We still need to show that this endomorphism is invertible. It suffices to check that the map on endomorphisms of some (and hence any) object of $\widehat{GPaCD}(n)$ is an isomorphism for each $n$. But these endomorphisms are isomorphic to $T_n = \mathrm{Exp}(\mathfrak{t}_n)$, and the morphism is easily checked to have the form

$$t_{ij} \mapsto t_{ij} + (\text{commutators})$$

which is invertible.

Exercise: Verify that $\mathrm{GRT}_1$ is a pro-unipotent group (e.g., check that the logarithm exists). $\square$

**Remark 4.4.** Note that the composition of two elements $\Phi(X, Y)$, $\Phi'(X, Y)$ is not the "naive" product, but given by the formula

$$(\Phi \cdot \Phi')(X, Y) = \Phi(X, Y)\Phi'(X, \Phi^{-1}Y\Phi).$$

The action of $\mathrm{GRT}_1$ on associators is given by the same formula, just interpreting $\Phi'$ as an associator.

Similarly, the bracket on $\mathfrak{grt}_1$ is not just the ordinary bracket of Lie series, but the "Poisson bracket"

$$\{\psi, \psi'\}(x, y) = [\psi(x, y), \psi'(x, y)] + D_\psi \psi'(x, y) - D_{\psi'}\psi(x, y)$$

where $D_\psi$ is the derivation of the free Lie algebra sending $x$ to $x$ and $y$ to $[y, \psi]$.

# Chapter 5

# The Knizhnik-Zamolodchikov associator

## 5.1 Introduction

The purpose of the present chapter is to prove the following Theorem.

**Theorem 5.1.** *The set of Drinfeld associators is not empty.*

In view of definition **??** we immediately obtain the following Corollary.

**Corollary 5.1.** *The set of Drinfeld associators* DAss *is a* GRT-GT *torsor.*

In other words, GRT and GT act freely transitively on DAss and the actions commute.

We will show Theorem 5.1 by a explicitly constructing one associator, the Knizhnik-Zamolodchikov associator $\Phi_{KZ}$. Concretely, this associator will have the following important property.

**Lemma 5.1.** *The coefficient of $X^{2p}Y$ in the Knizhnik-Zamolodchikov associator $\Phi_{KZ}(X,Y)$ (to be constructed below) is non-zero for $p = 1, 2, 3, \ldots$.*

**Remark 5.1.** Note that there is a special automorphism of $PaB$ (i.e., an element $\phi \in \mathrm{GT}$) which sends a braid to its mirror image "at the real axis". (On the braid group it sends $\sigma_i$ to $\sigma_i^{-1}$.) Its action on associators sends a $\mu$-associator to a $-\mu$-associator. Similarly, $\phi' = -1 \in \mathbb{K}^\times \subset \mathrm{GRT}$ sends a $-\mu$-associator to a $\mu$-associator, and hence the combined action of $\phi$ and $\phi'$ is an automorphism (even an involution) of the set of $\mu$-associators $\mathrm{DAss}_\mu$. Concretely, this action sends

$$\Phi(X,Y) \mapsto \Phi(-X,-Y).$$

Hence we see that there exists a second associator

$$\Phi_{\overline{\mathrm{KZ}}}(X,Y) := \Phi_{KZ}(-X,-Y) \neq \Phi_{KZ}(X,Y).$$

Now by the transitivity of the action of $\mathrm{GRT}_1$ on $\mathrm{DAss}_1$ there is a unique element $g \in \mathrm{GRT}_1$ such that

$$g \cdot \Phi_{KZ} = \Phi_{\overline{\mathrm{KZ}}}.$$

Since $\mathrm{GRT}_1$ is pro-unipotent, there is a unique $x \in \mathfrak{grt}_1$ such that $g = \exp(x)$. Since $\mathfrak{grt}_1$ is graded, we may write

$$x = \sum_{j=1}^{\infty} \sigma_j$$

where $\sigma_j$ is the degree $j$ part. Unravelling the action of $\mathfrak{grt}_1$ on $\mathrm{DAss}_1$, $\sigma_j$ is in particular "responsible for" changing the coefficient of $x^{j-1}y$ in the associator. Since for $j \geq 3$ odd these coefficients are different in $\Phi_{KZ}$ and $\Phi_{\overline{\mathrm{KZ}}}$, we find that $\sigma_j \neq 0$ for $j \geq 3$ odd. Summarizing, one obtains the following result.

**Corollary 5.2.** *The Grothendieck-Teichmller Lie algebra is inifnite dimensional and contains non-zero elements $\sigma_3, \sigma_5, \sigma_7, \ldots$ in degrees $3, 5, 7, \ldots$.*

In fact, we will see below that all other $\sigma_j$ are 0. The famous Deligne-Drinfeld-Ihara conjecture **??** states that $\mathfrak{grt}_1$ is in fact equal to the (completed) free Lie algebra generated by $\sigma_{2p+1}$, $p = 1, 2, 3, \ldots$. F. Brown has shown one half of this result, and we will present his proof in Chapter 7.

## 5.2 The Knizhnik-Zamolodchikov equation

The Knizhnik-Zamolodchikov equations originally are a set of differential equation satisied by the correlation functions of conformal field theory. In the simplest form, the Knizhnik-Zamolodchikov equation is the differential equation

$$\frac{d}{dz}u - \frac{1}{2\pi i}\left(\frac{X}{z} + \frac{Y}{1-z}\right)u = 0. \tag{5.1}$$

Here $z$ varies over $\mathbb{C}\setminus\{0,1\}$ and the unknown function $u$ takes values in $\mathbb{K}\langle\langle X, Y\rangle\rangle$. The equation asserts flatness of $u$ with respect to the holomorphic flat connection

$$d - \frac{1}{2\pi i}\left(\frac{X}{z} + \frac{Y}{1-z}\right)dz$$

with values in $\hat{\mathbb{F}}_{Lie}(X, Y)$.

There is also an extended form of this connection, also called the Knizhnik-Zamolodchikov connection. Namely, define a connection $\nabla$ with values in $\mathfrak{t}_n$ on the configuration space of $n$ points in $\mathbb{C}$

$$\mathrm{Conf}_n(\mathbb{C}) = \{z_1, \ldots, z_n \in \mathbb{C} \mid z_i \neq z_j \forall i \neq j\}$$

by the formula

$$\nabla = d - \frac{1}{2\pi i}\sum_{i<j}\frac{t_{ij}d(z_i - z_j)}{z_i - z_j}. \tag{5.2}$$

**Proposition 5.1.** *$\nabla$ is flat, i.e., $\nabla^2 = 0$. $\nabla$ is translation invariant, i. e.,*

$$[L_\tau, \nabla] = 0$$

*where $\tau$ is the vector field generating translation and $L$ denotes the Lie derivative. Furthermore $\nabla$ is also scale invariant*

$$[L_\sigma, \nabla] = 0$$

*where $\sigma$ is the scaling vector field. However, while $[\iota_\tau, \nabla] - L_\tau = 0$,*

$$[\iota_\sigma, \nabla] - L_\sigma = -\frac{1}{2\pi i}C$$

*where $C = \sum_{i<j} t_{ij}$ is the central element of $\mathfrak{t}_n$. It follows that $\nabla$ descend to the quotient of $\mathrm{Conf}_n(\mathbb{C})$ under translations, but it does not descend (readily) to the quotient under scaling.*

*Proof.* Compute:

$$\begin{aligned}
-4\pi^2\nabla^2 &= \sum_{\substack{i<j\\k<l}}\frac{[t_{ij}, t_{kl}]d(z_i - z_j)d(z_k - z_l)}{(z_i - z_j)(z_k - z_l)}\\
&= \sum_{i\neq j} dz_i dz_j \left(\sum_{k\neq i, l\neq j}\frac{[t_{ik}, t_{jl}]}{(z_i - z_k)(z_j - z_l)}\right)\\
&= \sum_{i\neq j} dz_i dz_j \left(\sum_{l\neq j}\frac{[t_{ij}, t_{jl}]}{(z_i - z_j)(z_j - z_l)} + \sum_{k\neq i}\frac{[t_{ik}, t_{ji}]}{(z_i - z_k)(z_j - z_i)}\right)\\
&= \sum_{i\neq j} dz_i dz_j \left(\sum_{k\neq i,j}\frac{-[t_{ik}, t_{jk}]}{(z_i - z_j)(z_j - z_k)} + \sum_{k\neq i,j}\frac{-[t_{ik}, t_{jk}]}{(z_i - z_k)(z_j - z_i)}\right)\\
&= -\sum_{i\neq j} dz_i dz_j \sum_{k\neq i,j}\frac{[t_{ik}, t_{jk}]}{(z_i - z_k)(z_j - z_k)} = 0.
\end{aligned}$$

The remainder of the assertions is trivial to verify. □

## 5.3 The Knizhnik-Zamolodchikov associator

Let $\Phi_\epsilon(x,y)$ be the parallel transport of the Knizhnik-Zamolodchikov connection betwen the points $z=\epsilon$ and $z=1-\epsilon$ in $\mathbb{C}\setminus\{0,1\}$. Concretely, it is given as a path ordered exponential

$$\Phi_\epsilon(x,y)=\mathrm{Pexp}\left(\frac{1}{2\pi i}\int_\epsilon^{1-\epsilon}\left(\frac{X}{z}+\frac{Y}{z-1}\right)dz.\right)$$

Even more concretely, let $x_0=x$, $x_1=y$, $z_0=0$ and $z_1=1$. Then the coefficient of some word

$$w=x_{j_1}\cdots x_{j_n}\in\mathbb{K}\langle\langle x_0,x_1\rangle\rangle,$$

in $\Phi_\epsilon(x_0,x_1)$ where $j_1,\dots,j_n\in\{0,1\}$ is

$$c_w=\frac{1}{(2\pi i)^n}\int_\epsilon^{1-\epsilon}\frac{dt_1}{t_1-z_{j_1}}\int_\epsilon^{t_1}\frac{dt_2}{t_2-z_{j_2}}\cdots\int_\epsilon^{t_{n-1}}\frac{dt_n}{t_n-z_{j_n}}.$$

These integrals may all be computed explicitly by the following recipe:

1. Expand each occurrence of $\frac{1}{t-1}$ as a power series $-\sum_{j\geq 0}t^j$.
2. Use the integral formulas

$$\int\frac{\log^n(t)dt}{t}=\frac{1}{n+1}\log^{n+1}(t)$$
$$\int\log^n(t)t^m dt=\sum_{j=0}^{n}\frac{(-1)^j}{(m+1)^j}\frac{n!}{(n-j)!}\log^{n-j}(t)t^{m+1}.$$

**Example 5.1.** Let us consider $w=001$. Let us also disregrad terms that tend to 0 when $\epsilon\to 0$ for simplicity. Then the innermost of the three integrals in $c_w$ yields

$$-\sum_{j\geq 0}\frac{t_2^{j+1}}{j+1}.$$

The next integral yields

$$-\sum_{j\geq 0}\frac{t_1^{j+1}}{(j+1)^2}.$$

The last integral transforms this to

$$-\frac{1}{(2\pi i)^3}\sum_{j\geq 0}\frac{(1-\epsilon)^{j+1}}{(j+1)^3}\to-\frac{1}{(2\pi i)^3}\zeta(3).$$

**Example 5.2.** To see what happens in a singular case, consider $w=010$. The first integral produces

$$\log t_2-\log\epsilon.$$

The next integral produces

$$\sum_j\frac{t_1^{j+1}}{j+1}\log(t_1/\epsilon)-\sum_j\frac{t_1^{j+1}}{(j+1)^2}-\sum_j\frac{\epsilon^{j+1}}{j+1}\log\epsilon-\sum_j\frac{\epsilon^{j+1}}{(j+1)^2}.$$

Omitting the terms that approach 0 as $\epsilon\to 0$ we obtain

$$-\sum_j\frac{(1-\epsilon)^{j+1}}{(j+1)^3}-\sum_j\frac{(1-\epsilon)^{j+1}}{(j+1)^3}-\sum_j\frac{(1-\epsilon)^{j+1}}{(j+1)^2}\log\epsilon\sim-2\zeta(3)-\zeta(2)\log\epsilon.$$

This expression diverges logarithmically with $\epsilon$.

**Lemma 5.2.** *For each word $w$ $c_w(\epsilon)$ is a polynomial in polylogarithm functions $\mathrm{Li}_n(\epsilon)$ and $\log\epsilon$. Moreover, if $w = x_0^{n_1-1}x_1x_0^{n_2-1}\cdots x_1x_0^{n_k-1}x_1$ with $n_1 \geq 2$ then*

$$\lim_{\epsilon\to 0} c_w(\epsilon) = \frac{(-1)^k}{(2\pi i)^n}\zeta(n_1,\dots,n_k).$$

*Proof.* □

In particular, it follows that $\Phi_\epsilon$ has an asymptotic expansion which is degree-wise polynomial in $\log\epsilon$. We denote this (degree-wise) polynomial by

$$\Phi_\sim(\log\epsilon).$$

(Concretely, $\Phi_\epsilon - \Phi_\sim(\log\epsilon) \to 0$ as $\epsilon \to 0$.)

**Definition 5.1.** *The Knizhnik-Zamolodchikov associator is defined to be*

$$\Phi_{KZ} := \Phi_\sim(0).$$

*In other words, we regularize $\Phi_\epsilon$ by formally set $\log\epsilon = 0$.*

**Remark 5.2.** The Knizhnik-Zamolodchikov associator may equivalently be defined in the following ways:

- As
$$\Phi_{KZ}(X,Y) := \lim_{\epsilon\to 0}\epsilon^Y\Phi_\epsilon(X,Y)\epsilon^{-X}.$$
- Let $u_0(z), u_1(z)$ be solutions of (5.1) such that $u_0(z) \sim z^X$ as $z \to 0$ and $u_1(z) \sim (1-z)^Y$ as $z \to 1$. Then we set $\Phi_{KZ} := u_1^{-1}u_0$ (which is independent of $z$).

*Proof.* Step 1: Show that such solution $u_0$, $u_1$ exists (insert the Ansatz $u_0 = \tilde{u}z^X$ into the KZ equation).

Step 2: Given $u_0$ and $u_1$ we may express

$$\Phi_\epsilon = u_1(1-\epsilon)u_1^{-1}u_0u_0(\epsilon)^{-1}.$$

Hence the limit

$$\lim_{\epsilon\to 0}\epsilon^Y\Phi_\epsilon(X,Y)\epsilon^{-X} = \lim_{\epsilon\to 0}(\epsilon^Y u_1(1-\epsilon))u_1^{-1}u_0(u_0(\epsilon)^{-1}\epsilon^{-X}) = u_1^{-1}u_0$$

exists. Hence the asymptotic expansion of $\Phi_\epsilon$ is $\epsilon^{-Y}u_1^{-1}u_0\epsilon^X$ an the result follows. □

**Theorem 5.2.** *$\Phi_{KZ}$ is a Drinfeld associator, i. e., it satisfies (4.4)-(4.6) for $\mu = 1$.*

The antisymmetry equation (4.4) is easy to see. Since $\Phi_\epsilon(x,y) = \Phi_\epsilon(y,x)^{-1}$ by reflection symmetry, the same equation holds for the asymptotic expansion and for the regularization. For the hexagon equation (4.5) consider the version of the Knizhnik-Zamolodchikov connection with values in $\mathfrak{t}_3$ defined in equation (5.2) (for $n = 3$). Consider the parallel transport around the closed loop shown in Figure **??**. Since the loop is contractible the parallel transport is 1 by flatness of the connection. Since the path is composed of 6 segments, we obtain an equation with 6 terms.

$$1 = U_{3,12}\Phi_\epsilon^{3,1,2}U_{13}(\Phi_\epsilon^{1,3,2})^{-1}U_{23}\Phi_\epsilon^{1,2,3}$$

where $U_{23}, U_{13}, U_{3,12}$ denotes the parallel transport along the three semicircle path segments. It is not hard to see that the limits of these transports as $\epsilon \to 0$ exist are are

$$U_{23} \to e^{t_{23}/2} \qquad U_{13} \to e^{t_{13}/2} \qquad U_{3,12} \to e^{-(t_{13}+t_{23})/2}.$$

Using the asymptotic expansions for $\Phi_\epsilon$ we see that the right hand side of the above equation has an asymptotic expansion that is a polynomial in $\epsilon$ in each degree. In particular we hence learn that the equation must hold on the constant term, i. e., when setting $\log\epsilon = 0$. Using furthermore that $t_3 \cong \hat{\mathbb{F}}_{Lie}(t_{12}, t_{23}) \oplus \mathbb{K}C$ where $C = t_{12} + t_{23} + t_{13}$ is the central element, we obtain the hexagon equation.

The pentagon equation is shown in a similar way, considering the parallel transport in $\mathrm{Conf}_4(\mathbb{C})$ along the path depicted in Figure **??**.

# Chapter 6

# Associators and double shuffle relations

## 6.1 Multiple zeta values and double shuffle relations

Multiple zeta values are the numbers

$$\zeta(n_1, \ldots, n_k) := \sum_{j_1 > j_2 > \cdots > j_k \geq 1} \frac{1}{j_1^{n_1} j_2^{n_2} \cdots j_k^{n_k}}$$

defined as long as $n_1 \geq 2$. We introduce the following notation. Let $w = n_1 \cdots n_k$ be a word with letters from the alphabet $\mathbb{N}$. We call it admissible if $n_1 \geq 2$ (or $k = 0$). Let $A_{adm}$ be the vector space of formal linear combinations of all admissible words. We define the map $\zeta : A_{adm} \to \mathbb{R}$ by setting

$$\zeta(w) = \zeta(n_1, \ldots, n_k)$$

on words $w$ and extending linearly. Furthermore $\zeta(\emptyset) := 1$ by convention. Define the associative commutative *stuffle product* $\sqcup\!\sqcup$ on words as

$$nw \sqcup\!\sqcup\, n'w' = n(w \sqcup\!\sqcup\, n'w') + (n + n')(w \sqcup\!\sqcup\, w') + n'(nw \sqcup\!\sqcup\, w')$$

where $n, n' \in \mathbb{N}$ are symbols (numbers) and $w, w'$ are words, the tails of the words $nw$, $n'w'$.

**Example 6.1.**

$$23 \sqcup\!\sqcup\, 5 = 523 + 73 + 243 + 28 + 235$$

**Proposition 6.1.** *$\zeta$ is an algebra morphism $(A_{adm}, \sqcup\!\sqcup) \to \mathbb{R}$.*

*Proof.* Straightforward. □

This means that the multizeta values satisfy the stuffle relations,

$$\zeta(w)\zeta(w') = \zeta(w \sqcup\!\sqcup\, w').$$

The multizeta values satisfy another class of relations, the *shuffle relations*. Consider words $w$ in symbols $\{0, 1\}$. We call such a word admissible if it is empty or if it starts with 0 ends in 1. So each (non-empty) admissible word has the form

$$w = 0^{n_1 - 1} 1 0^{n_2 - 1} 1 \cdot 0^{n_k - 1} 1$$

with $n_1 \geq 2$, $n_2, \ldots n_k \geq 1$. Let $B_{adm}$ be the vector space spanned by such admissible "binary" words. We may define a linear function $\zeta : B_{adm} \to \mathbb{R}$ be letting

$$\zeta(w) = \zeta(n_1, \ldots, n_k)$$

and extending by linearity, where $w$ is as above. Again we set $\zeta(\emptyset) := 1$. On $B$ we may define an associative and commutative product, the shuffle product $*$ recursively by

$$\alpha w * \alpha' w' := \alpha(w * \alpha' w') + \alpha'(\alpha w * w').$$

**Proposition 6.2.** *$\zeta : (B_{adm}, *) \to \mathbb{R}$ is an algebra map.*

The proof is straightforward (and left to the reader) using Kontsevich's integral representation of the multizeta values

$$\zeta(a_1 \cdots a_n) = \int_0^1 \frac{dt_1}{f_{a_1}(t)} \int_0^{t_1} \frac{dt_2}{f_{a_2}(t)} \cdots \int_0^{t_{n-1}} \frac{dt_n}{f_{a_n}(t)}.$$

where $f_0(t) = t$ and $f_1(t) = 1 - t$.

The shuffle and stuffle relations thus obtained on multi zeta values are together called the "double shuffle relations". Many algebraic relations between multizeta values may be derived using these double shuffle relations. But not all, the simplest example being Euler's famous identity

$$\zeta(2,1) = \zeta(3).$$

In the nect section, we extend the double shuffle relations to "regularized" double shuffle relations. Euler's identity and conjecturally all algebraic identities between multi zeta values will follow from these extended relations.

## 6.2 Regularized double shuffle relations

Let $(A, \sqcup\!\sqcup) \supset (A_{adm}, \sqcup\!\sqcup)$ be the algebra spanned by all words in symbols $\mathbb{N}$ (not just the admissible ones). Similarly let $(B, *) \supset (B_{adm}, *)$ the shuffle algebra of all words in letters $\{0, 1\}$.

**Proposition 6.3.** *The map $\zeta : (A_{adm}, \sqcup\!\sqcup) \to \mathbb{R}$ uniquely extends to a map of algebras $\zeta : (A, \sqcup\!\sqcup) \to \mathbb{R}$ such that $\zeta(1) = 0$.*

*Similarly, the map $\zeta : (B_{adm}, *) \to \mathbb{R}$ uniquely extends to a map of algebras $\zeta : (B, *) \to \mathbb{R}$ such that $\zeta(1) = \zeta(0) = 0$.*

In other words we regularize the "ill-defined" multiple zeta values in two different ways. In the first case, one talks about stuffle regularized multiple zeta values, while in the second case one talks about shuffle regularized multiple zeta values.

# Chapter 7

# F. Brown's Theorem

## 7.1 (Pre-)dual of the universal enveloping algebra

Let $\mathfrak{g}$ be a pro-nilpotent Lie algebra. Recall that $\hat{U}\mathfrak{g}$ is the completed universal enveloping algebra and $\mathrm{Exp}(\mathfrak{g}) =: G$ is the exponential group, which may be identified with $\mathfrak{g}$ as a set or with the group-like elements of $\hat{U}\mathfrak{g}$ as a group.

Below we want to dualize the Hopf algebra $\hat{U}\mathfrak{g}$. Define the Hopf algebra $\mathcal{O}(G)$ (think: functions on $G$) as the Hopf algebra topologically dual to $\hat{U}\mathfrak{g}$. Concretely, suppose $\mathfrak{g} = \lim_{\leftarrow} \mathfrak{g}_\alpha$. Then we set

$$\mathcal{O}(G) = \lim_{\rightarrow} \mathcal{O}(\mathrm{Exp}(\mathfrak{g}_\alpha))$$

where $\mathcal{O}(\mathrm{Exp}(\mathfrak{g}_\alpha))$ is the universal enveloping coalgebra of the Lie coalgebra $\mathfrak{g}_\alpha^*$.

**Example 7.1.** Let $\mathfrak{g} = \hat{\mathbb{F}}_{Lie}(x, y)$ then $\hat{U}\mathfrak{g} = \mathbb{K}\langle\langle x, y\rangle\rangle$ while $\mathcal{O}(G) \cong \mathbb{K}\langle a, b\rangle$, where $a, b$ are dual to $x, y$. The product on $\mathcal{O}(G)$ is the shuffle product of words, and is commutative. The coproduct is the deconcatenation, and is not co-commutative.

There is a natural pairing $\hat{U}\mathfrak{g} \times \mathcal{O}(G) \to \mathbb{K}$.

## 7.2 Preliminaries on actions and coactions

**Lemma 7.1.** *Let $G = \mathrm{Exp}(\mathfrak{g})$ be a pro-unipotent group acting on $X$ as above. Then a function $f \in \mathcal{O}(X)$ is $G$-invariant iff*

$$0 = \Delta f \in \mathfrak{g}^\vee \otimes \mathcal{O}(X).$$

*Proof.* "$\Rightarrow$": $G$-invariance means that $f'(g)f''(x) = f(x)$ for all $g \in G, x \in X$. If we let $g = \exp(t\Psi)$ for $\Psi \in \mathfrak{g}$ and take the $t^1$-coefficient, the result follows.

"$\Leftarrow$": Conversely, suppose the infinitesimal coaction vanishes. Let $x \in X$ and $\Psi \in \mathfrak{g}$ be arbitrary. Consider the function $p(t) := f(e^{t\Psi}x) - f(x)$. Clearly $p(0) = 0$, and $p(t)$ is a polynomial in $t$. By the assumption its derivative vanishes. Hence $p(t) \equiv 0$. ☐

## 7.3 The Poisson bracket on the free Lie algebra

For $f \in \hat{\mathbb{F}}_{Lie}(x, y)$ define the derivation $D_f$ of $\hat{\mathbb{F}}_{Lie}(x, y)$ which on generators is given by

$$D_f x = x \qquad\qquad D_f y = [y, f]\,.$$

**Lemma 7.2.** *For all $f, g \in \hat{\mathbb{F}}_{Lie}(x, y)$:*

$$[D_f, D_g] = D_{[f,g]} + D_{D_f g} - D_{D_g f}$$

*Proof.* Exercise. ☐

**Corollary 7.1.** *The* Poisson bracket

$$\{,\} : \hat{\mathbb{F}}_{Lie}(x,y) \times \hat{\mathbb{F}}_{Lie}(x,y) \to \hat{\mathbb{F}}_{Lie}(x,y)$$
$$\{f,g\} = [f,g] + D_f g - D_g f$$

*defines a Lie algebra structure on* $\hat{\mathbb{F}}_{Lie}(x,y)$.

*Proof.* Bilinearity and antisymmetry are clear. It suffices to check the Jacobi identity.

$$\begin{aligned}\{f,\{g,h\}\} + (cyc) &= [f,[g,h]] + [f, D_g h - D_h g] + D_f([g,h]) + D_f(D_g h - D_h g) - D_{[g,h]} f - D_{D_g h - D_h g} f + (cyc) \\ &= 0 + [f, D_g h] - [h, D_g f] + D_g([h,f]) + D_f D_g h - D_g D_f h - D_{[f,g]} h - D_{D_f g} h + D_{D_g f} h + (cyc) = 0.\end{aligned}$$

Here $(cyc)$ stands for the other cyclic permutations of terms and for the last equality we used that $D_{\cdot}$ is a derivation and Lemma 7.2. □

**Remark 7.1.** $\mathfrak{grt}_1$ is a Lie subalgebra of $(\hat{\mathbb{F}}_{Lie}(x,y), \{,\})$.

**Remark 7.2.** Note that the Poisson bracket is a bit peculiar. In particular $\{x,y\} = [x,y] + [y,x] = 0$, and hence $(\hat{\mathbb{F}}_{Lie}(x,y), \{,\})$ is not generated by $x$ and $y$, and the map $(\hat{\mathbb{F}}_{Lie}(x,y), [,]) \to (\hat{\mathbb{F}}_{Lie}(x,y), \{,\})$ which exists by the universal property is quite trivial.

It is interesting to work out the universal enveloping algebra of $\mathfrak{grt}_1$. We claim that this (topological) Hopf algebra is $\mathbb{K}\langle\langle x,y\rangle\rangle$ with the following (non-standard) Hopf algebra structure:

- The coproduct is the standard coproduct determined by the equations

$$\Delta x = x \otimes 1 + 1 \otimes x \qquad\qquad \Delta y = y \otimes 1 + 1 \otimes y.$$

- The unit and counit are the standard ones.
- The product is non-standard, and is given by the following formula.

$$\begin{aligned}F \cdot X^{k_0} Y X^{k_1} Y X^{k_2} \cdots X^{k_{n-1}} Y X^{k_n} = F^{(0)} X^{k_0} (SF^{(1)}) Y F^{(2)} X^{k_1} (SF^{(3)}) Y F^{(4)} X^{k_2} \cdots X^{k_{n-1}} \\ (SF^{(2n-1)}) Y F^{(2n)} X^{k_n}\end{aligned}$$

  where $S$ is the standard antipode and

$$\Delta^{2n} F =: \sum F^{(0)} \otimes \cdots \otimes F^{(2n)}.$$

- The antipode $\tilde{S}$ is not the standard one, but instead determined recursively by the formulas

$$\mu(1 \otimes \tilde{S})\Delta = \mu(1 \otimes \tilde{S})\Delta = 1\epsilon.$$

**Remark 7.3.** For a group-like element $F$ and an arbitrary $G$ the product formula becomes

$$F(X,Y) \cdot G(X,Y) = F(X,Y) G(X, F^{-1} Y F)$$

where $F^{-1} = SF$ is the inverse of $F$ with respect to the usual product.

**Proposition 7.1.** *The operations just described endow* $\mathbb{K}\langle\langle x,y\rangle\rangle$ *with a Hopf algebra structure. It is the universal enveloping algebra of* $(\hat{\mathbb{F}}_{Lie}(x,y), \{,\})$.

*Proof.* By construction of the antipode, we only have to verify that the operations define a bialgebra structure. It is clear that the operations $(\Delta, \epsilon)$ describe a coalgebra structure. It is also clear that 1 is indeed a left and right unit for the product. Next we have to show the associativity of the product (i.e., $F \cdot (G \cdot H) = (F \cdot G) \cdot H$). Since the linear combinations of group-like elements are dense in $\mathbb{K}\langle\langle x,y\rangle\rangle$ and the product is continuous we may assume the $F, G, H$ are group-like. Then by the formula from the previous remark

$$F \cdot (G \cdot H) = F(X,Y) G(X, F^{-1} Y F) H(X, G_F^{-1} F^{-1} Y F G_F)$$
$$(F \cdot G) \cdot H = F(X,Y) G(X, F^{-1} Y F) H(X, G_F^{-1} F^{-1} Y F G_F)$$

where we used the notation $G_F = G(X, F^-1YF)$. Hence associativity holds. Next we have to show the compatibility of product and coproduct, i.e., that $(\Delta G) \cdot (\Delta H) = \Delta(GH)$ and that $\Delta$ and $\cdot$ preserve unit and counit. The second statement is obvious. For the first statement we may again assume $G, H$ group-like. Compute

$$\begin{aligned}\Delta(GH) &= \Delta(F(X,Y)G(X,F^{-1}YF)) = (F(X,Y)G(X,F^{-1}YF)) \otimes (F(X,Y)G(X,F^{-1}YF)) \\ &= (F \otimes F) \cdot (G \otimes G) = (\Delta G) \cdot (\Delta H).\end{aligned}$$

Finally let us check that $A := \mathbb{K}\langle\langle x, y\rangle\rangle$ is indeed (isomorphic to) the universal enveloping algebra $U\tilde{\mathfrak{g}}$ of $\tilde{\mathfrak{g}} := (\hat{\mathbb{F}}_{Lie}(x,y), \{,\})$. We claim first that $A$ is generated by $\tilde{\mathfrak{g}}$, i. e., by the primitive elements. Indeed, the linear combinations of group-like elements are dense in $A$, any group-like element is the exponential of its logarithm, and the logarithm is primitive, hence the primitive elements generate a dense subspace.

Next note that by the universal property there is a continuous map $U\tilde{\mathfrak{g}} \to A$. It preserves the grading. Furthermore, in each degree both sides have the same (finite) dimension. Hence it suffices to show that the map is surjective in each degree to conclude that is is bijective. However, by the construction the generators $\tilde{\mathfrak{g}}$ of $A$ are in the image, and hence surjectivity holds.

Note furthermore that the map $U\tilde{\mathfrak{g}} \to A$ is also a morphism of bialgebras, and hence also of Hopf algebras. □

**Remark 7.4.** Note that the exponential and logarithm in $A = \mathbb{K}\langle\langle x, y\rangle\rangle$ (with the "Poisson" product are not the same as the exponential and logarithm in $\mathbb{K}\langle\langle x, y\rangle\rangle$ with the ordinary (concatenation) product.

**Remark 7.5.** Note that both $\mathrm{GRT}_1$ and DAss are naturally subsets of $\tilde{G} := \mathrm{Exp}(\tilde{\mathfrak{g}})$, and the group composition and the action are both the group composition in $\tilde{G}$.

Below a very important role will be played by the Hopf algebra of functions $\mathcal{O}(\tilde{G}) = \mathbb{K}\langle\langle x, y\rangle\rangle^\vee = \mathbb{K}\langle a, b\rangle$. The Hopf algebra structure is just the dual of the one on $\mathbb{K}\langle\langle x, y\rangle\rangle$, but let us nevertheless write it down explicitly. To do that, and to be more consistent with the literature, let us introduce the notation $I(a_1, \dots, a_n)$ for elemenrts of $\mathbb{K}\langle a, b\rangle$. Here each $a_j \in \{0, 1\}$ and the element of $\mathbb{K}[a, b]$ is obtained by replacing each occurrence of 0 by $a$ and each occurrence of 1 by $b$, and form a word by concatenating the symbols. So, for example

$$I(0, 0, 1) := aab \in \mathbb{K}\langle a, b\rangle.$$

Furthermore, to simplify the formula for the coproduct below, we will also introduce the notatiom

$$I_{a_0,a_{n+1}}(a_1, \dots, a_n) := \begin{cases} 1 & \text{if } n = 0 \text{ TODO: check if omissible} \\ I(a_1, \dots, a_n) & \text{if } a_0 = 0 \text{ and } a_1 = 1 \\ (-1)^n I(a_n, \dots, a_1) & \text{if } a_0 = 1 \text{ and } a_1 = 0 \\ 0 & \text{if } a_0 = a_1 \text{ and } n \geq 1. \end{cases} \tag{7.1}$$

Note that the formula for the third case is the (standard) antipode applied to the word $I(a_1, \dots, a_n)$.

- The product is the standard shuffle product. It is commutative.
- The unit and counit are the standard ones.
- The coproduct is non-standard, and is given by the following formula.

$$\Delta I(a_1, \dots, a_n) = \sum_k \sum_{0=i_0<i_1<\cdots<i_{k+1}=n+1} \left( \prod_{j=0}^{k} I_{a_{i_j},a_{i_{j+1}}}(a_{i_j+1}, \dots, a_{i_{j+1}-1}) \right) \otimes I(a_{i_1}, \dots, a_{i_k})$$

  Here we set $a_0 := 0$ and $a_{rn+1} = 1$ for notational simplicity.
- As above, the antipode $\tilde{S}$ is determined by the bialgebra structure.

We will denote the Hopf algebra $\mathbb{K}\langle a, b\rangle$ (with this non-standard Hopf algebra structure) by $\mathcal{A}_0$.

**Example 7.2.**

**Remark 7.6.** Note that in particular any Drinfeld associator and any $\mathrm{GRT}_1$ element give points in $\mathrm{Spec}(\mathcal{A}_0)$, i. e., algebra maps $\mathcal{A}_0 \to \mathbb{K}$ by evaluation.

## 7.4 The orbit of $\Phi_{KZ}$

Let $\mathfrak{f}_1 = \hat{\mathbb{F}}_{Lie}(\sigma_3, \sigma_5, \dots)$, with $\sigma_3$ in degree 3 etc., and let $F_1 = \mathrm{Exp}(\mathfrak{f}) \cong \hat{\mathcal{F}}_{\mathrm{Grp}}(\Sigma_3, \Sigma_5, \dots)$. As we saw in section **??**, we have a map $\mathfrak{f}_1 \to \mathfrak{grt}_1$ and hence a map $F_1 \to \mathrm{GRT}_1$. Remember that our goal here is to show that this map is an injection. The algebra of functions on $F_1$ is defined to be $\mathcal{O}(F_1) = \mathbb{K}\langle s_3, s_5, \dots \rangle$ with the usual commutative Hopf algebra structure. Concretely, the product is the shuffle product and the coproduct is deconcatenation.

Furthermore, define $F = F_1 \rtimes \mathbb{K}^\times$, where $\lambda \mathbb{K}^\times$ acts by scaling $\sigma_{2j+1}$ by $\lambda^{2j+1}$, i.e., by the grading. $F$ is not pro-unipotent, because of the factor $\mathbb{K}^\times$. We have a map $F \to \mathrm{GRT}$. Hence we have an action of $F$ on DAss.

We will from now on assume $\mathbb{K} = \mathbb{C}$. We will consider the orbit of the Knizhnik-Zamolodchikov associator $\Phi_{KZ}$ under the action of $F$.

$$\mathcal{Y} := \overline{\mathrm{GRT} \cdot \Phi_{KZ}}.$$

The space of functions on $\mathcal{Y}$, $\mathcal{O}(\mathcal{Y})$ is defined to be

$$\mathcal{O}(\mathcal{Y}) = \mathcal{A}_0 / J$$

where $J := \{f \in \mathcal{A}_0 \mid f(g \cdot \Phi_{KZ}) = 0 \forall g \in \mathrm{GRT}\}$. Note that $J$ is a graded ideal (because the rescaling operation is in GRT), so that $\mathcal{O}(\mathcal{Y})$ inherits the grading from $\mathcal{A}_0$. We call the piece of degree $N$ $\mathcal{O}(\mathcal{Y})_N$. Furthermore, since $F_1$ acts on $\mathcal{Y}$, we obtain a caction of $\mathcal{O}(F_1)$ on $\mathcal{O}(\mathcal{Y})$. Concretely, the formula for this coaction is obtained by computing the coproduct in $\mathcal{A}_0$ (by formula (**??**)) and then projecting the first factor to $\mathcal{O}(F_1)$.

**Remark 7.7.** We saw in section **??** that all Drinfeld associators satisfy the regularized double shuffle relations. Hence these relations are contained in $J$.

Let $\tilde{F}_1$ be the image of $F_1$ in $\mathrm{GRT}_1$. It is a graded (hence closed), pro-unipotent subgroup of $\mathrm{GRT}_1$. Our goal is to show that $\tilde{F}_1 \cong F_1$.

**Lemma 7.3.** *There is an isomorphism of $\mathcal{O}(\tilde{F}_1)$ comodules*

$$\mathcal{O}(\mathcal{Y}) \cong \mathcal{O}(\tilde{F}_1) \times \mathbb{K}[\zeta^f(2)]$$

*where $\zeta^f(2)$ here is considered as some symbol and $\mathbb{K}[\zeta^f(2)]$ is considered as a trivial comodule. (I.e., $\Delta\zeta^f(2) = 1 \otimes \zeta^f(2)$.)*

The morphism is, however, non-canonical and does not preserve the grading.

*Proof.* Act with GRT on an even associator... □

**Remark 7.8.** We will use the notation $\zeta^f(n_1, \dots, n_k) \in \mathcal{O}(\mathcal{Y})$ to denote the image of the element

$$I(10^{\{n_1-1\}}10^{\{n_1-1\}}1 \dots 10^{\{n_k-1\}}) \in \mathcal{A}_0$$

and more generally $\zeta^f_m(n_1, \dots, n_k) \in \mathcal{O}(\mathcal{Y})$ to denote the image of the element

$$I(0^m 10^{\{n_1-1\}}10^{\{n_1-1\}}1 \dots 10^{\{n_k-1\}}) \in \mathcal{A}_0$$

The result we want to show here is the following.

**Theorem 7.1** (F. Brown)**.** *The elements $\zeta^f(r_1, \dots, r_n) \in \mathcal{O}(\mathcal{Y})$, $n = 0, 1, \dots$ and $r_j \in \{2, 3\}$ form a vector space basis of $\mathcal{O}(\mathcal{Y})$.*

The proof will occupy us for some while. However, for now note the following Corollary.

**Corollary 7.2.** *The map $F_1 \to \mathrm{GRT}_1$ is injective.*

*Proof.* We may equivalently show that $\mathcal{O}(\mathrm{GRT}_1) \to \mathcal{O}(F_1)$ is surjective. But the latter map factors as

$$\mathcal{O}(\mathrm{GRT}_1) \to \mathcal{O}(\tilde{F}_1) \to \mathcal{O}(F_1)$$

by definition of $\tilde{F}_1$, with the left hand map being a surjection and the right hand map being an injection. However, from the Theorem and the Lemma it follows that $\mathcal{O}(\tilde{F}_1)$ has in each degree the same dimension as $\mathcal{O}(F_1)$. Hence the right hand map must also be surjective, and we are done. □

For now, let us introduce some notation. Let the map $D_{2r+1}$ be the composition

$$D_{2r+1}\colon \mathcal{O}(\mathcal{Y}) \to \mathcal{O}(\tilde{F}_1)\otimes\mathcal{O}(\mathcal{Y}) \to \mathcal{O}(F_1)\otimes\mathcal{O}(\mathcal{Y}) \to \mathfrak{f}_1^\vee\otimes\mathcal{O}(\mathcal{Y}) \xrightarrow{\mathrm{ev}_{\sigma'_{2r+1}}} \mathcal{O}(\mathcal{Y}).$$

The composition of the first three arrows is just the infinitesimal coaction of $\mathfrak{f}_1^\vee$. The last arrow is the projection of the first factor to the cogenerator of degree $2r+1$, $\sigma'_{2r+1}\propto\sigma_{2r+1}$, normalized in such a way that $D_{2r+1}\zeta^f(2r+1)=1$.

**Lemma 7.4.** *Let $x\in\mathcal{O}(\mathcal{Y})_N$ be of degree $N$. Suppose that $D_{2r+1}x=0$ for all $2r+1<N$. Then $x$ is a multiple of $\zeta^f(N)$.*

**Remark 7.9.** Let $\mathfrak{g}$ be a Lie algebra and $U\mathfrak{g}\supset\mathfrak{g}$ be the universal enveloping algebra. Then the coproduct on $x\in\mathfrak{g}$ is $\Delta x = 1\otimes x + x\otimes 1$, which is send to zero modulo the units. The dual statement for Lie coalgebras $C$ is that the projection $UC\to C$ annihilates products $ab$, which are non-trivial, i.e., $\epsilon(a)=\epsilon(b)=0$. This has two nice consequences for us, regarding the computation of $D_{2r+1}$:

- When computing $D_{2r+1}$, we may omit the product terms from the formula for the coaction (**??**).
- A priori the projection $\mathfrak{f}_1^\vee\to\mathbb{K}$ to the cogenerators is hard to compute, because we do not know the explicit formula for $\sigma_{2r+1}$. In practice, we will solve this by showing identities of multi-zeta values of the form
$$\zeta^f(XXX) = c\zeta^f(2r+1) + \text{products}.$$
In this case the projection just picks out the constant $c$. In fact, we need two such identities, given in the next section.

## 7.5 Some multizeta identities in $\mathcal{O}(\mathcal{Y})$

Two identities about formal multi zeta values are needed. In the first Lemma, note that $\zeta^f(2^k)\propto\zeta^f(2)^k$ by Exercise **??**.

**Lemma 7.5.**
$$\zeta_1^f(2^{\{n\}}) = 2\sum_{i=1}^{n}(-1)^i\zeta^f(2i+1)\zeta^f(2^{n-i})$$

*Proof.* Apply the stuffle relations

$$\zeta^f(2i+1)\zeta^f(2^k) = \sum_{j=0}^{k}\zeta^f(2^j(2i+1)2^{k-j}) + \sum_{j=0}^{k-1}\zeta^f(2^j(2i+3)2^{k-j-1}).$$

The terms on the right cancel telescopically and one is left with the equation

$$\zeta_1^f(12^{\{n\}}) = -2\sum_{j=0}^{n-1}\zeta^f(2^j32^{n-1-j})$$

which is a (regularized) shuffle product relation. □

The second is the formal version of an identity of D. Zagier. (Zagier's result may be recovered by omitting all superscripts $f$, or equivalently by evaluating on the Knizhnik-Zamolodchikov associator.)

**Theorem 7.2.**
$$\zeta^f(2^a32^b) = 2\sum_{r=1}^{a+b+1}(-1)^r(A^r_{a,b}-B^r_{a,b})\zeta^f(2r+1)\zeta^f(2^{a+b+1-r})$$

*where*

$$A^r_{a,b} = \binom{2r}{2a+2} \qquad B^r_{a,b} = (1-2^{-2r})\binom{2r}{2b+1}.$$

*Idea of proof.* By induction on $a+b$, using Lemma 7.4. Apply $D_{2r+1}$ to the difference $\Theta$ of both sides. $D_{2r+1}\zeta^f(2^a32^b)$ may be computed, the result being expressed as projections of $\zeta^f(2^\alpha32^\beta)$ for $\alpha+\beta<a+b$ to the cogenerators. These cases are handled by the induction hypothesis. Hence one finds that $\Theta\propto\zeta^f(2a+2b+3)$ by Lemma 7.4. The missing constant is read off from D. Zagier's non-formal version of the Theorem. □

## 7.6 F. Brown's proof of Theorem 7.1

Temporarily let $V_N$ be the free vector space spanned by words in symbols $\{2,3\}$ of weight (i. e., digit sum) $N$.[1] To show 7.1 it is sufficient to show that the obvious maps

$$V_N \to \mathcal{O}(\mathcal{Y})_N$$

are injective for all $N$. Then these maps are actually bijective since we know that $\dim(\mathcal{O}(\mathcal{Y})_N) \leq \dim(V_N)$.

We will show this statement by induction. For $N = 0$, both spaces are 1 dimensional and the statement is trivial. Assume that the statement has been shown up to $N - 1$, and our task is to show it for $N$.

Put a grading (the *level grading*) on the vector spaces $V_n$ by the number of 3's in words. It induces a filtration on the image $W_N$ of $V_N$ in $\mathcal{O}(\mathcal{Y})_N$. Concretely, $\mathcal{F}^p W_N$ is the vector space spanned by all $\zeta^f(r_1, \dots, r_k)$ with $r_j \in \{2,3\}$, $\sum_j r_j = N$, and with at most $p$ of the $r_j$'s being 3's. To reach our goal it clearly suffices to show that

$$\mathrm{gr}^p V_N \to \mathrm{gr}^p W_N$$

is injective for all $p$.

There is one case we can settle right away, namely the case $p = 0$. If $N$ is odd, then $\mathrm{gr}^0 V_N = 0$ and we are done. If $N$ is even, then $\mathrm{gr}^0 V_N$ is one dimensional, and it is easily checked that the image is non-zero. (For example since the non-formal zeta value $\zeta(2, 2, \dots, 2) = \frac{\pi^n}{(2n+1)!} \neq 0$ where $N = 2n$.)

**Exercise 7.1.** Show that in fact $\zeta^f(2^{\{n\}})$ is a multiple of $\zeta^f(2)^n$ by using Lemma 7.4.

Hence we are left with checking injectivity for $p \geq 1$.

**Lemma 7.6.**

$$D_{2r+1}\mathcal{F}^p W_N \subset \mathcal{F}^{p-1} W_{N-2r-1}$$

*Proof.* We have to cut out a subsequence of length $2r+1$ of a sequence of 01's and 001's. Furthermore, to get a nonzero contribution, the symbols just to the right and left of the subsequence have to be unequal. Schematically

$$\cdots 0XXXXX1 \cdots$$

or

$$\cdots 1XXXXX0 \cdots .$$

(The beginning of the sequence counts as 0, the end as 1.) The possible neighbourhoods of the cut region are: $(1)00XXX(1)$, $(1)0XXX(1)$, $1XXX01$, $1XXX001$. In all cases, after cutting there remains a series of 2s and 3s, with at most the same number of 3s that were there before. Furthermore, since we cut out an odd length piece, the number of 3s is in fact reduced by at least 1. $\square$

Hence let us consider the composition

$$\partial_{N,p} : \mathrm{gr}^p V_N \to \mathrm{gr}^p W_N \to \bigoplus_{3 \leq 2r+1 \leq N} \mathrm{gr}^{p-1} W_{N-2r-1}$$

where the the right hand aroow is the sum of maps induced by $D_{2r+1}$. here we also used the previous Lemma. Of course, it suffices to show that $\partial_{N,p}$ is injective. Also note that by the induction hypothesis, we know a basis for the right hand side.

The very remarkable insight of F. Brown is now that the (square) matrix of $\partial_{N,p}$ with repect to the bases given may be explicitly computed (using the formal version of Zagier's Theorem **??**), and that invertibility of this matrix follows by considering arithmetic properties of its entries.

### 7.6.1 Formula for the matrix of $\partial_{N,p}$

All entries of the matrix of $\partial_{N,p}$ will be rational. Modulo $2\mathbb{Z}$, F. Brown showed that the operators $\partial_{N,p}$ acts just by deconcatenation.

[1] By definition, the empty word has digit sum 0.

**Theorem 7.3** (F. Brown)**.** *Let $w$ be a word in symbols $\{2,3\}$ of weight $N$ and level $p$. Then*

$$\partial_{N,p}\zeta^f(w) = \sum_{\substack{v:uv=w \\ \deg_3(v)=1}} c_v \zeta^f(u) \quad (\textit{mod } I) \tag{7.2}$$

*where $\deg_3(v) = 1$ shall mean that the word $v$ contains only one 3, i. e. $v = 2^a 3 2^b$ for some numbers $a, b$, and the set $I$ is composed of even integer multiples of $\zeta^f(\cdots)$. (In other words, we are giving the matrix elements of $\partial_{N,p}$ only modulo $2\mathbb{Z}$.)*

*Furthermore the constant*

$$c_v = 2(-1)^{a+b+1}\left(A_{a,b}^{a+b+1} - B_{a,b}^{a+b+1}\right) = 2(-1)^{a+b+1}\left(\binom{2a+2b+2}{2a+2} - (1-2^{-2a-2b-2})\binom{2a+2b+2}{2b+1}\right). \tag{7.3}$$

*are the constants appearing in Theorem 7.2. The "(mod $2\mathbb{Z}$)" shall indicate that all numbers are given only modulo addition of even integers, which are irrelevant for invertibility as we will see below.*

*Proof.* One needs to consider 4 non-trivial cases, according to how the cut out subsequence sits inside the sequence $w$, so that the symbols to the left and right of the subsequence are distinct. They the length of the subsequence we cut out is $2r + 1$.

1. The subsequence starts and ends in 0. In that case it has the form $0(10)^r$. But this contributes an even integer to the matrix element by Lemma **??**.

2. After cutting out the subsequence, the remaining sequence contains less than $p - 1$ 3s. Then there is no contribution by definition of $\partial_{N,p}$.

3. We cut out the subsequence in the middle (not at the end) of $w$. Then there are always two ways to cut:

   $$0(101\cdots1001\cdots0)10$$

   or

   $$01(01\cdots1001\cdots01)0.$$

   The first way contributes a coefficient $c_{2^a32^b}$ and the second one a coefficient $(-1)^{2r+1}c_{2^b32^a}$ by Theorem 7.2. Hence their sum is

   $$c_{2^a32^b} - c_{2^b32^a} = 2(-1)^{a+b+1}\left(\binom{2a+2b+2}{2a+2} - \binom{2a+2b+2}{2b+2}\right) \in 2\mathbb{Z}.$$

4. If the subsequence is a tail of $w$, the contribution is as stated in the Theorem, again by Theorem 7.2. (Note that in contrast to the previous case, there is now no "partner subsequence" that could make the contribution even.)

□

### 7.6.2 End of the proof

To finish the proof it suffices to check that the determinant of (the matrix of) $\partial_{N,p}$ is non-zero. We write $\det(\partial_{N,p})$ abusively, assuming that the unimportant overall sign has been fixed, e. g. by putting some order on the basis elements. Clearly $\det(\partial_{N,p}) \in \mathbb{Q} \subset \mathbb{Q}_2$, where $\mathbb{Q}_2$ are the 2-adic numbers. The 2-adic expansion of $\det(\partial_{N,p})$ will have the form

$$\sum_{j \geq -q} c_j 2^j$$

where $c_j \in \{0, 1\}$. For us it suffices to show that one $c_j \neq$. We will try to take $j$ smallest.

Let $M_{w,u}$ be the entry of the matrix of $\det(\partial_{N,p})$ between words $w$ and $v$. According to Theorem 7.3 the only $M_{w,u}$ of non-positive valuation are those for which there are $a$, $b$ such that $w = u2^a32^b$, and in this case $M_{w,u} = c_{2^a32^b}$ (sf. (7.2)). Hence in this case

$$\begin{aligned}\mathrm{val}_2(M_{w,u}) = \mathrm{val}_2(c_{2^a32^b}) &= \mathrm{val}_2\left(2^{-2a-2b-1}\binom{2a+2b+2}{2b+1}\right) \\ &= -2a-2b-1+\mathrm{val}_2\binom{2a+2b+2}{2b+1} = -2a-2b+\mathrm{val}_2(a+b+1)+\mathrm{val}_2\binom{2a+2b+1}{2b}.\end{aligned}$$

Considering the column $M_{\cdot,u}$ for fixed $u$, only pairs $a, b$ with $a+b =: n$ fixed occur. Among these $a, b$, the minimum of $\mathrm{val}_2 \binom{2a+2b+1}{2b} \geq 0$ is obtained at $a = 0$, $b = n$, in which case the valuation becomes 0. This minimum is however also obtained at other values (e. g., $b = 0$).

Let us put the reverse ("middle-eastern") lexicographic ordering on the the words $w$ of weight $N \geq 1$ and level $p$. For each $w$ we assign one possible $u$ as above, namely the unique $u$ such that $w = 32 \cdots 2$. This puts an ordering on the $u$'s.

**Proposition 7.2.** *Ordering the words $w$ and $u$ as said. Then the matrix $M_{w,u}$ has the following form:*

- *All entries above the diagonal have positive valuation.*
- *The lowest valuation in each column is attained on the diagonal.*

*In particular the determinant is non-zero, i.e., $\partial_{N,p}$ is invertible.*

*Proof.* To see the first statements, consider some fixed $u$. The highest $w$ that would yield an $M_{w,u}$ of non-positive-valuation is $w = u32 \cdots 2$, and we saw above that in this case the valuation becomes minimal.

To see the final statement, note that the lowest coefficient in the 2-adic expansion is solely contributed by the term in the expansion of the determinant coming from the product of the diagonal elements. $\square$

Since $\partial_{N,p}$ is invertible it follows that the maps $\mathrm{gr}^p V_N \to \mathrm{gr}^p W_N$ are injective, hence the maps $V_N \to W_N$ are injective, and hence isomorphisms. Hence Theorem 7.1 follows. $\square$

# Bibliography