ZETA ELEMENTS IN DEPTH 3 AND THE FUNDAMENTAL LIE ALGEBRA OF A PUNCTURED ELLIPTIC CURVE

FRANCIS BROWN

ABSTRACT. This paper draws connections between the double shuffle equations and structure of associators; universal mixed elliptic motives as defined by Hain and Matsumoto; and the Rankin-Selberg method for modular forms for $\mathrm{SL}_2(\mathbb{Z})$. We write down explicit formulae for zeta elements σ_{2n-1} (generators of the Tannaka Lie algebra of the category of mixed Tate motives over $\mathbb{Z}$) in depths up to four, give applications to the Broadhurst-Kreimer conjecture, and completely solve the double shuffle equations for multiple zeta values in depths two and three.

1. INTRODUCTION

The theme of this paper is that certain constructions relating to the motivic fundamental group of the projective line minus 3 points, which are inherently ambiguous, can be explicitly determined, and simplified, by passing to genus one.

The main result can be viewed on the following three different levels.

1.1. The fundamental group of $\mathbb{P}^1 \backslash \{0, 1, \infty\}$.

The de Rham fundamental group

$$_1\Pi_1 = \pi_1^{dR}(\mathbb{P}^1 \backslash \{0, 1, \infty\}, \vec{1}_1)$$

of $\mathbb{P}^1 \backslash \{0, 1, \infty\}$ with tangential base point the unit tangent vector at 1 is a prounipotent affine group scheme over $\mathbb{Q}$. Its graded Lie algebra is the free Lie algebra $\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ on two generators $\mathsf{x}_0, \mathsf{x}_1$ corresponding to loops around 0 and 1. Since $_1\Pi_1$ is a (pro) object in the category of mixed Tate motives over $\mathbb{Z}$, it admits an action of the Tannakian fundamental group $G_{\mathcal{MT}(\mathbb{Z})}^{dR}$. Denote the graded Lie algebra of the latter by

$$(1.1) \qquad \qquad \mathfrak{g}^{\mathfrak{m}} = \mathbb{L}\langle \sigma_3, \sigma_5, \ldots \rangle .$$

It is the free graded Lie algebra generated by non-canonical elements σ_{2n+1} in degree $-2n-1$ for $n \geq 1$. We obtain a morphism of Lie algebras

$$(1.2) \qquad \qquad i_0 : \mathfrak{g}^{\mathfrak{m}} \longrightarrow \mathrm{Der}^1 \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$$

where $\mathrm{Der}^1 \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ denotes the set of derivations which send x_1 to 0. Furthermore, we know that (1.2) factors through a morphism

$$i : \mathfrak{g}^{\mathfrak{m}} \longrightarrow \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$$

and (1.2) maps $\sigma \in \mathfrak{g}^{\mathfrak{m}}$ to the derivation which sends $\mathsf{x}_0 \mapsto [i(\sigma), \mathsf{x}_0]$ and x_1 to 0. The main result of [3] states that i is injective, and therefore enables us to expand elements $\sigma \in \mathfrak{g}^{\mathfrak{m}}$ in 'coordinates' x_0 and x_1. It is an important problem to characterise the image of the map i and to describe the $i(\sigma_{2n-1})$ as explicitly as possible. It is known by the work of Racinet that the image of i is contained in the Lie algebra $\mathfrak{dmr}_0$ of solutions to the double shuffle equations. It is also contained in the space of solutions to Drinfeld's associator equations, which by a result of Furusho [21], are contained in $\mathfrak{dmr}_0$.

It is well-known that

$$(1.3) \qquad \qquad i(\sigma_{2n+1}) = \mathrm{ad}(\mathsf{x}_0)^{2n} \mathsf{x}_1 + \text{ terms of degree } \geq 2 \text{ in } \mathsf{x}_1 ,$$

1

This can be seen as follows. The Drinfeld associator is the formal power series which is the generating series of shuffle-regularised multiple zeta values,

$$Z = \sum_{w \in \{x_0, x_1\}^\times} \zeta(w) w \qquad \in \quad \mathbb{R}\langle\langle x_0, x_1 \rangle\rangle \; ;$$

and is easily computed explicitly to first order in x_1:

$$Z \equiv \sum_{n \geq 2} \zeta(n) \, \mathrm{ad}(x_0)^{n-1} x_1 \quad (\mathrm{mod} \ \text{terms of degree} \ \geq 2 \ \text{in} \ x_1) \; .$$

The coefficients of the odd zeta values in degree $2n+1$ are congruent to $i(\sigma_{2n+1})$, but very little is known in general about the coefficients of $i(\sigma_{2n+1})$ of higher degrees ≥ 3 in the x_1. In this paper, we show:

Theorem 1.1.   (1) *That there is a choice of generators $\sigma_{2n+1}^c \in \mathfrak{g}^{\mathfrak{m}}$ which is given by an explicit formula* (1.9) *modulo terms of degree ≥ 5 in x_1.*
   (2) *That there is a rational associator τ which is given by an explicit formula modulo terms of degree ≥ 4 in x_1. It computes the coefficients of the even zeta values (powers of π) in Z in this range.*

Statement (1) is surprising because the choice of generators σ_{2n+1} are *a priori* only well-defined up to addition of higher order commutators of σ_{2m+1}. The key point is that by passing to genus 1, we find canonical coordinates on $\mathfrak{g}^{\mathfrak{m}}$ which enable us to fix the commutators $[\sigma_{2a+1}, [\sigma_{2b+1}, \sigma_{2c+1}]]$ uniquely. A similar story holds for (2).

In so doing, we discover that it is more convenient to consider a different normalisation for the σ_{2n+1} from the canonical normalisation (1.3), which we call the *heretical normalisation*

$$(1.4) \qquad \underline{\sigma}_{2n+1} = \frac{B_{2n}}{(2n)!} \, \mathrm{ad}(x_0)^{2n} x_1 + \ \text{terms of degree} \ \geq 2 \ \text{in} \ x_1 \; ,$$

where B_{2n} is the $2n^{\mathrm{th}}$ Bernoulli number. Objects which are normalised according to the heretical normalisation will be underscored.

1.2. The fundamental group of the first-order Tate curve.
Let $E^\times_{\partial/\partial q}$ denote the fiber of the universal elliptic curve $\mathcal{M}_{1,2} \to \mathcal{M}_{1,1}$ over the tangential base point $\frac{\partial}{\partial q}$ on $\mathcal{M}_{1,1}$, where $\mathcal{M}_{g,n}$ denotes the moduli space of curves of genus g with n marked points. In a future paper with Hain, we shall show (as suggested in [27]) that its de Rham fundamental group

$$(1.5) \qquad\qquad\qquad \mathcal{P} = \pi_1^{dR}(E^\times_{\partial/\partial q}, \overrightarrow{1}_1)$$

where $\overrightarrow{1}_1$ is the tangent vector of length 1 with respect to the canonical holomorphic coordinate w on $E^\times_{\partial/\partial q}$, is the de Rham realisation of a pro-object in the category of mixed Tate motives over $\mathbb{Z}$. Its associated bigraded Lie algebra (bigraded for the weight W and relative monodromy-weight M filtrations) is the free Lie algebra on certain canonical generators a, b. Correspondingly[1], one obtains a morphism of Lie algebras

$$(1.6) \qquad\qquad\qquad i_1 : \mathfrak{g}^{\mathfrak{m}} \longrightarrow \mathrm{Der}^\Theta \, \mathbb{L}(a, b)$$

where Der^Θ denotes the set of derivations δ such that $\delta(\Theta) = 0$, where $\Theta = [a, b]$. We shall show as a consequence of [3] that (1.6) is injective.

---

[1]If one thinks of $\mathfrak{g}^{\mathfrak{m}}$ as being bigraded for M and W, with $W = M$, then the map i_1 respects the M-grading, but not the W-grading, see [27].

On the other hand, there are distinguished derivations $\varepsilon_{2n}^{\vee} \in \mathrm{Der}^{\Theta} \mathbb{L}(a,b)$ whose action on a are given by

$$\varepsilon_{2n}^{\vee} a = \mathrm{ad}(a)^{2n} b \qquad \text{for } n \geq 1 .$$

They were first studied by Nakamura [36] in a slightly different context and rediscovered in [11, 32]. The action of $\varepsilon_{2n}^{\vee}$ on b is determined by the condition $\varepsilon_{2n}^{\vee} \Theta = 0$ together with the fact that it is homogeneous of degree $2n$ in a, b. The derivations $\varepsilon_{2n}^{\vee}$ are 'geometric' in the sense that the relative completion of $\mathrm{SL}_2(\mathbb{Z}) = \pi_1(\mathcal{M}_{1,1}, \partial/\partial q)$ (or universal monodromy) acts on $\mathbb{L}(a,b)$ via the Lie algebra generated by the $\varepsilon_{2n}^{\vee}$ and their images $\mathrm{ad}(\varepsilon_0^{\vee})^k \varepsilon_{2n}^{\vee}$ under the adjoint action of

$$\varepsilon_0^{\vee} \in \mathrm{Der}^{\Theta} \mathbb{L}(a,b) \quad , \qquad \varepsilon_0^{\vee}(a) = b \quad \varepsilon_0^{\vee}(b) = 0 .$$

Denote the Lie subalgebra generated by the $\varepsilon_{2n}^{\vee}$, for all $n \geq 0$, by

$$\mathfrak{u}^{\mathrm{geom}} \subset \mathrm{Der}^{\Theta} \mathbb{L}(a,b) .$$

It is the bigraded image of the universal monodromy [26]. The elements $\varepsilon_{2n}^{\vee}$ satisfy many relations which were studied by Pollack [37]. The image of $\mathfrak{g}^{\mathfrak{m}}$ in $\mathrm{Der}^{\Theta} \mathbb{L}(a,b)$ under (1.6) is by no means contained in $\mathfrak{u}^{\mathrm{geom}}$, but in low degrees with respect to b, the $\varepsilon_{2n}^{\vee}$ give canonical 'coordinates' in which to write down the initial terms of elements $i_1(\sigma_{2n+1})$. The motivic version of a formula due to Nakamura in the ℓ-adic setting, is

$$i_1(\sigma_{2n+1}) \equiv \varepsilon_{2n+2}^{\vee} \quad (\mathrm{mod}\ W_{-2n-3})$$

for all $n \geq 1$. A more precise result can be obtained using the elements σ_{2n+1}^c.

Theorem 1.2. *Let $n \geq 2$. The elements σ_{2n+1}^c satisfy*

$$(1.7) \qquad i_1(\sigma_{2n+1}^c) \equiv \underline{\varepsilon}_{2n+2}^{\vee} + \sum_{a+b=n} \frac{1}{2b} \{\underline{\varepsilon}_{2a+2}^{\vee}, \{\underline{\varepsilon}_{2b+2}^{\vee}, \underline{\varepsilon}_0^{\vee}\}\} \qquad (\mathrm{mod}\ W_{-2n-5}) .$$

where the $\underline{\varepsilon}_{2n}^{\vee}$ are the heretical normalisations (3.4) of the $\varepsilon_{2n}^{\vee}$.

This theorem is equivalent to an explicit formula for the $i_0(\sigma_{2n+1}^c) \in \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ modulo terms of degree ≥ 5 in x_1, which is how it is proved here. This uses an explicit morphism from the de Rham fundamental group of $\mathbb{P}^1 \backslash \{0, 1, \infty\}$ to that of $E^{\times}$ which was written down by Hain. Since i_0 and i_1 are compatible with this morphism, the expansions of σ_{2n+1}^c under i_0 and i_1 can be related to each other explicitly.

1.3. **Rankin-Selberg method.** The third way of understanding the elements σ_{2n+1}^c, and the starting point for this paper, came from the theory of iterated integrals of holomorphic modular forms for $\mathrm{SL}_2(\mathbb{Z})$. The coefficients in equation (1.7) come from the computation [7] of the imaginary part of an iterated integral of two Eisenstein series using the Rankin-Selberg method. They turn out to be the coefficients of $\zeta(2n-1)$ in the convolution of two Eisenstein series of different weights, which are products of Bernoulli numbers. Equivalently, they are proportional to the coefficients in the odd period polynomials of Eisenstein series. This is the motivation for the heretical normalisations (1.4) and the source of the formula (1.7).

1.4. **Further remarks.** We discuss the methods used in this paper, and applications to the double shuffle equations and Broadhurst-Kreimer conjecture.

1.4.1. *Commutative power series and anatomy of associators.* One tool which we use extensively is the method of commutative power series and is closely related to Ecalle's theory of moulds [17, 18]. Let $\mathbb{L}(u,v)$ be the free graded Lie algebra generated by two elements u, v. It is graded for the degree in v. Elements of v-degree $r \geq 1$ in the tensor algebra $T(u,v)$ can be encoded by commutative polynomials

$$(1.8) \qquad \begin{aligned} \rho : \operatorname{gr}_v^r T(u,v) &\longrightarrow \mathbb{Q}[x_1, \ldots, x_r] \qquad r \geq 1 \\ u^{i_0} v u^{i_1} \ldots v u^{i_r} &\longmapsto x_1^{i_1} \ldots x_r^{i_r} \end{aligned}$$

We apply this construction to $(u,v) = (\mathsf{x}_0, \mathsf{x}_1)$ and $(u,v) = (a,b)$, and their derivation algebras. In certain contexts, we shall explain that it is natural to rescale the morphism ρ by introducing polynomial denominators. In this manner, elements of $\operatorname{Der}^1 \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ and $\operatorname{Der}^\Theta \mathbb{L}(a,b)$ are encoded by sequences of *rational functions* in $x_1, \ldots, x_r$. The double shuffle equations (defining equations for the Lie algebra $\mathfrak{dmr}_0$) can be translated into functional equations for commutative power series via the map ρ. A surprising discovery is that there exist canonical solutions if one allows poles:

Theorem 1.3. [2] *There exist explicit solutions to the double shuffle equations in the space of rational functions in all weights and all depths.*

There is a particular family of solutions we denote by $\xi_{2n+1}^{(r)} \in \mathbb{Q}(x_1, \ldots, x_r)$ in weight $2n + 1 \geq 3$. Their components in depths $r = 1, 2$ are polynomials, but they have poles in depths $r \geq 3$. Furthermore, a new element emerges in weight -1 which we denote by $\xi_{-1}^{(r)} \in \mathbb{Q}(x_1, \ldots, x_r)$. The idea of [2] is to expand the polynomial representation of Drinfeld elements $\rho(i(\sigma_{2n-1}))$ in terms of the ξ_{2n+1}. This 'anatomy' can be computed explicitly in low degrees:

Theorem 1.4. *If* { , } *denotes the Ihara bracket, extended to the setting of rational functions, then the canonical Drinfeld elements up to depth 4 are given, in the heretical normalisation, by the simple formula:*

$$(1.9) \quad \rho(i(\underline{\sigma}_{2n+1}^c)) \equiv \underline{\xi}_{2n+1} + \sum_{a+b=n} \frac{1}{2b} \{\underline{\xi}_{2a+1}, \{\underline{\xi}_{2b+1}, \underline{\xi}_{-1}\}\} \qquad (\mathrm{mod} \ depths \ \geq 5) \ .$$

If one were to write this formula in the canonical, as opposed to heretical normalisations, one would find coefficients given by products of Bernoulli numbers in the sum in the right-hand side. These coefficients are essentially the coefficients in the odd period polynomial of Eisenstein series, a fact which emerges from §9.

Theorem 1.9 is proved by combinatorial methods, and uses Goncharov's theorem that the solutions to the double shuffle equations in depth 3 are motivic. It makes no reference to the first-order Tate curve. It is more illuminating, however, to interpret this theorem by passing to genus 1. Via the Hain morphism 3.3, it turns out that the elements ξ_{2n+1} correspond in low depths to the derivations $\varepsilon_{2n+2}^\vee$, and this proves that theorems 1.4 and 1.2 are equivalent.

1.4.2. *Double shuffle equations.* Our elements σ_{2n+1}^c are explicit solutions to the double shuffle equations in depths ≤ 4 and odd weights. We also construct, in §7.1, an explicit solution τ in depths ≤ 3 and all even weights. Using a theorem due to Goncharov computing the dimension of the space of solutions to linearised double shuffle equations in depth 3 we deduce the

Theorem 1.5. *Every solution to the regularised double shuffle equations in depths ≤ 4 (odd weight) and depths ≤ 3 (even weight) can be expressed using the explicit elements σ_{2n+1}^c and the rational associator τ.*

This theorem can be applied to the semi-numerical algorithm described in [5] to decompose motivic multiple zeta values into a chosen basis using the motivic coaction. It involved a numerical computation of a regulator at each step. One application of the elements σ^c_{2n+1} and τ is to remove this transcendental step, leading to an exact and effective algorithm for proving motivic relations between multiple zeta values in low depths, and any weight. It replaces the need to store tables of multiple zeta values in this range of depths [10].

A further manifestation of the double shuffle equations occurs in genus 1. As mentioned above, we encode elements of $\mathrm{Der}^\Theta \mathbb{L}(a,b)$ by rational functions, by composing the morphism $\delta \mapsto \delta(a) : \mathrm{Der}^\Theta \mathbb{L}(a,b) \longrightarrow \mathbb{L}(a,b) \subset T(a,b)$ with the linear map

$$(1.10) \qquad \mathrm{gr}^r_b T(a,b) \quad \longrightarrow \quad \mathbb{Q}(x_1,\ldots,x_r)$$

$$a^{i_0} b a^{i_1} b \ldots b a^{i_r} \quad \mapsto \quad \frac{x_1^{i_1} \ldots x_r^{i_r}}{x_1(x_1-x_2)\ldots(x_{r-1}-x_r)x_r} \ .$$

In [2], we defined a bigraded Lie algebra $\mathfrak{pls}$ to be the space of solutions to the linearised double shuffle equations with poles at worst of the above form. We show that:

Proposition 1.6. *The Lie algebra of geometric derivations is contained, via* (1.10), *in the space solutions to the linearised double shuffle equations:* $\mathfrak{u}^{\mathrm{geom}} \subset \mathfrak{pls}$

Thus the linearised double shuffle equations occur naturally in the elliptic setting. It is natural to ask if $\mathfrak{u}^{\mathrm{geom}} = \mathfrak{pls}$, and easy to show [2] that this holds in depths ≤ 3. It follows from the previous proposition that the stuffle equations enable us to detect non-geometric derivations, i.e., elements in the quotient

$$(\mathrm{Der}^\Theta \mathbb{L}(a,b))/\mathfrak{u}^{\mathrm{geom}} \ .$$

1.4.3. *Depth 4 generators in the Broadhurst-Kreimer conjecture.* A further application of the canonical elements σ^c_{2n+1} is to the Broadhurst-Kreimer conjecture.

It is well-known since Ihara and Takao [30] that there exist quadratic relations

$$(1.11) \qquad \sum_{i,j} \lambda_{i,j} \{\sigma_{2i+1}, \sigma_{2j+1}\} \equiv 0 \quad (\text{mod terms of degree } \geq 4 \text{ in } \mathsf{x}_1)$$

where $\lambda_{i,j} \in \mathbb{Q}$ are coefficients of period polynomials P of even, cuspidal $\mathrm{SL}_2(\mathbb{Z})$-cocycles. In [6], we reformulated the Broadhurst-Kreimer conjecture, which describes the dimensions of the space of multiple zeta values graded by the depth, in terms of the spectral sequence induced on $\mathfrak{g}^{\mathfrak{m}}$ by the depth filtration D. Using the elements σ^c_{2n+1} we can compute the first non-trivial differential (conjecturally, the only non-trivial differential) in this spectral sequence. A motivic version of the Broadhurst-Kreimer conjecture can thus be formulated by saying that $\mathrm{gr}_D \mathfrak{g}^{\mathfrak{m}}$ has the following presentation: it is generated by the classes $[\sigma_{2n+1}] \in \mathrm{gr}^1_D \mathfrak{g}^{\mathfrak{m}}$ for all $n \geq 1$, and certain elements $\mathfrak{c}(P) \in \mathrm{gr}^4_D \mathfrak{g}^{\mathfrak{m}}$, where $\mathfrak{c}$ can be expressed in terms of the canonical elements σ^c_{2n+1}, and it is subject only to the relations (1.11). For a precise statement, see §8. This formulation of the Broadhurst-Kreimer conjecture can be transposed to the elliptic setting and describes a presentation for a certain Lie subalgebra of $\mathfrak{u}^{\mathrm{geom}}$.

2. Reminders on the projective line minus 3 points

Background material can be found in [15], [38], [6].

2.1. Depth.

Let $\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ denote the free graded Lie algebra over $\mathbb{Q}$ on two generators $\mathsf{x}_0, \mathsf{x}_1$, where x_0 and x_1 have degree 1. The depth filtration $D^n \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ is the decreasing filtration such that $D^0 = \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ and

$$D^1 \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) = \ker(\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) \longrightarrow \mathbb{L}(\mathsf{x}_0))$$

where the map on the right sends x_1 to 0 and x_0 to x_0. It is defined by $D^n = [D^1, D^{n-1}]$ for all $n \geq 2$. It is the decreasing filtration associated to the D-degree, for which x_0 has D-degree 0 and x_1 has D-degree 1. Therefore $D^n \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ consists of $\mathbb{Q}$-linear combinations of Lie brackets of x_0 and x_1 with at least n x_1's.

The universal enveloping algebra of $\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ is the graded tensor algebra $T(\mathsf{x}_0, \mathsf{x}_1)$ on $\mathbb{Q}\mathsf{x}_0 \oplus \mathbb{Q}\mathsf{x}_1$. The D-degree is defined in the same manner on $T(\mathsf{x}_0, \mathsf{x}_1)$ and defines a decreasing filtration $D^n T(\mathsf{x}_0, \mathsf{x}_1)$ spanned by words in $\geq n$ x_1's. We shall embed $\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) \subset T(\mathsf{x}_0, \mathsf{x}_1)$; the embedding is compatible with the filtrations D.

2.2. Ihara bracket.

The de Rham fundamental groupoid [15]

$$_0\Pi_1 = \pi_1^{dR}(\mathbb{P}^1 \backslash \{0, 1, \infty\}, \overrightarrow{1}_0, -\overrightarrow{1}_1)$$

is the de Rham realisation of a mixed Tate motive over $\mathbb{Z}$, and admits an action of the de Rham motivic Galois group G^{dR}. The action on the trivial de Rham path $_0 1_1$ from the tangential base point 1 at 0 to the tangential base point -1 at 1 gives a morphism

$$(2.1) \qquad\qquad g \mapsto g._0 1_1 : G^{dR} \longrightarrow {}_0\Pi_1$$

of schemes. It becomes a morphism of groups if one equips $_0\Pi_1$ with the Ihara group law, which is denoted by $\circ$. If R is a commutative unitary algebra, the set of R-points of $_0\Pi_1$ is the set of invertible group-like (with respect to the completed coproduct for which $\mathsf{x}_0, \mathsf{x}_1$ are primitive) formal power series $R\langle\langle \mathsf{x}_0, \mathsf{x}_1 \rangle\rangle$ in two non-commuting variables. The Ihara group law is then given by the formula

$$(2.2) \qquad\qquad \begin{aligned} \circ : {}_0\Pi_1 \times {}_0\Pi_1 &\longrightarrow {}_0\Pi_1 \\ F \circ G &= G(\mathsf{x}_0, F\mathsf{x}_1 F^{-1})F \ . \end{aligned}$$

The expression on the right-hand side is also equal to

$$F \circ G = FG(F^{-1}\mathsf{x}_0 F, \mathsf{x}_1) \ .$$

Likewise, the de Rham fundamental group with tangential base point -1 at 1

$$_1\Pi_1 = \pi_1^{dR}(\mathbb{P}^1 \backslash \{0, 1, \infty\}, -\overrightarrow{1}_1)$$

admits an action of G^{dR}, which can be shown to factorise through the composition of the map (2.1) with the left-action

$$(2.3) \qquad\qquad \begin{aligned} \circ_1 : {}_0\Pi_1 \times {}_1\Pi_1 &\longrightarrow {}_1\Pi_1 \\ F \circ_1 H &= H(F^{-1}\mathsf{x}_0 F, \mathsf{x}_1) \ . \end{aligned}$$

Now pass to graded Lie algebras. The graded Lie algebras of both $_0\Pi_1$ and $_1\Pi_1$ can be identified with $\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$. Let $\mathfrak{g}^{\mathfrak{m}} = \mathrm{Lie}^{\mathrm{gr}} U^{dR}$, where U^{dR} is the unipotent radical of $G^{dR} = U^{dR} \rtimes \mathbb{G}_m$. Equation (2.1) gives a morphism

$$(2.4) \qquad\qquad i : \mathfrak{g}^{\mathfrak{m}} \longrightarrow (\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1), \{,\})$$

where $\{\,,\,\}$ is the Ihara bracket, for which we give a formula below. Let $\mathrm{Der}^1\,\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$ denote the Lie subalgebra of derivations $\delta \in \mathrm{Der}\,\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$ which satisfy $\delta(\mathsf{x}_1) = 0$. The left action (2.3) is given on the level of graded Lie algebras by

$$(2.5) \qquad (\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1),\{\,,\,\}) \quad \longrightarrow \quad \mathrm{Der}^1\,\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$$

$$\sigma \quad \mapsto \quad \begin{cases} \mathsf{x}_0 \mapsto [\mathsf{x}_0,\sigma] \\ \mathsf{x}_1 \mapsto 0 \end{cases} \; .$$

Theorem 2.1. [3] *The morphism* (2.4) *is injective.*

Because of this theorem, we can identify $\mathfrak{g}^{\mathfrak{m}}$ with its image in $\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$ via i. The graded Lie algebra $\mathfrak{g}^{\mathfrak{m}}$ is freely generated by elements σ_{2n+1} in degree $-2n-1$, for all $n \geq 1$. Their images in $\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$ are known as Drinfeld elements

$$(2.6) \qquad i(\sigma_{2n+1}) = \mathrm{ad}(\mathsf{x}_0)^{2n}\mathsf{x}_1 + \text{ terms of depth } \geq 2 \; .$$

They are not *a priori* canonical for $n \geq 5$. However, the Hoffman-Lyndon basis for motivic multiple zeta values allows one to define canonical choices σ_{2n+1}^{HL} for the σ_{2n+1} [4]. Very little is known about these elements, except for the coefficients of $(\mathsf{x}_0\mathsf{x}_1)^a\mathsf{x}_0(\mathsf{x}_0\mathsf{x}_1)^b$ in $\sigma_{2a+2b+1}$ (which are independent of the choice of σ_{2n+1}), which follows from [3] and [41].

2.3. **Linearized Ihara action and depth.**
In [6], we considered the following linearised version of the Ihara action. For any word $a \in \{\mathsf{x}_0,\mathsf{x}_1\}^{\times}$, let

$$(a_1\ldots a_n)^* = (-1)^n a_n \ldots a_1 \; .$$

Definition 2.2. Define a $\mathbb{Q}$-bilinear map

$$\underline{\circ} : T(\mathsf{x}_0,\mathsf{x}_1) \otimes_{\mathbb{Q}} T(\mathsf{x}_0,\mathsf{x}_1) \to T(\mathsf{x}_0,\mathsf{x}_1)$$

inductively as follows. For any words a,w in $\mathsf{x}_0,\mathsf{x}_1$, and for any integer $n \geq 0$, let

$$(2.7) \qquad a \underline{\circ} (\mathsf{x}_0^n\mathsf{x}_1 w) = \mathsf{x}_0^n a\mathsf{x}_1 w + \mathsf{x}_0^n\mathsf{x}_1 a^* w + \mathsf{x}_0^n\mathsf{x}_1(a \underline{\circ} w)$$

with the initial condition $a \underline{\circ} \mathsf{x}_0^n = \mathsf{x}_0^n a$.

The antisymmetrization of the map $\underline{\circ}$ restricts to the Ihara bracket on $\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$:

$$(2.8) \qquad \{f,g\} = f \underline{\circ} g - g \underline{\circ} f \qquad \text{for all } f,g \in \mathbb{L}(\mathsf{x}_0,\mathsf{x}_1) \; .$$

It follows from this formula that the Ihara bracket is homogeneous for the D-degree (the fact that it respects the depth filtration follows from the geometric interpretation of the depth filtration using the embedding $\mathbb{P}^1\backslash\{0,1,\infty\} \subset \mathbb{P}^1\backslash\{0,\infty\}$ [15]). If we embed $\mathfrak{g}^{\mathfrak{m}} \hookrightarrow \mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$ via (2.4) we can define the depth filtration $D\mathfrak{g}^{\mathfrak{m}}$ on $\mathfrak{g}^{\mathfrak{m}}$ to be the decreasing filtration induced by the depth filtration on $\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$.

For later use, remark that if $A(a,b,c) = a \underline{\circ} (b \underline{\circ} c) - (a \underline{\circ} b) \underline{\circ} c$, then

$$(2.9) \qquad\qquad A(a,b,c) = A(b,a,c)$$

for any $a,b,c \in T(\mathsf{x}_0,\mathsf{x}_1)$, which follows from the definitions, and implies that the linearised Ihara bracket satisfies the Jacobi identity.

2.4. **Double shuffle equations.**
The double shuffle equations are a family of equations satisfied by multiple zeta values which are well-adapted to the depth filtration. In his thesis [38], Racinet defined a subspace

$$\mathfrak{dmr}_0 \subset \mathbb{L}(\mathsf{x}_0,\mathsf{x}_1) \; ,$$

called the regularised double shuffle Lie algebra, which encodes these relations in terms of two Hopf algebra structures. The Lie algebra version of Racinet's theorem is:

Theorem 2.3. [38] *The space $\mathfrak{dmr}_0$ is closed under the Ihara bracket $\{\,,\}$.*

Since the regularised double shuffle equations hold for actual multiple zeta values, and are stable under the Ihara bracket, it follows that they are motivic. Combined with theorem 2.1 we deduce that there is an inclusion of Lie algebras

$$\mathfrak{g}^{\mathfrak{m}} \subset \mathfrak{dmr}_0 \subset \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) \ .$$

Therefore we can study elements σ_{2n+1} by attempting to solve the defining equations of $\mathfrak{dmr}_0$ in low depths. This will be achieved below using the language of commutative power series (§4).

2.5. **Depth-graded motivic Lie algebra.** The depth filtration induces a decreasing filtration $D^\bullet$ on G^{dR} and hence $\mathfrak{g}^{\mathfrak{m}}$ via the maps (2.1) and (2.4). Let

$$(2.10) \qquad\qquad\qquad \mathfrak{d} = \mathrm{gr}_D^\bullet \mathfrak{g}^{\mathfrak{m}}$$

denote the associated graded Lie algebra. It is bigraded for weight and depth. The component of $\mathfrak{d}$ of depth d and weight n will be denoted by $\mathfrak{d}_n^d$. Let $\mathfrak{d}^d = \bigoplus_n \mathfrak{d}_n^d$, denote the (infinite-dimensional) component in depth d, and let $\mathfrak{d}_n = \bigoplus_d \mathfrak{d}_n^d$ denote the (finite-dimensional) component in weight n.

The linearised double shuffle equations are a family of equations which were introduced in [22] and further studied in [6]. It follows from a variant of Racinet's theorem that their set of solutions, denoted by $\mathfrak{ls} \subset \mathrm{gr}_D^\bullet \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$, is a bigraded Lie algebra for the Ihara bracket $\{\,,\,\}$, and a corollary of theorem 2.1 and Racinet's theorem is

Theorem 2.4. *([6], §5)* $\mathfrak{d} \subset \mathfrak{ls}$.

The following theorem was first proved by Tsumura. See [6], §6.4 for a short proof.

Theorem 2.5. *(Depth-Parity theorem for double shuffle equations). We have*

$$\mathfrak{ls}_n^d = 0 \quad if \quad n \equiv d + 1 \pmod 2 \ .$$

Combining the previous two theorems gives the

Corollary 2.6. *(Depth-Parity theorem). If $n \equiv d + 1 \pmod 2$, then $\mathfrak{d}_n^d = 0$.*

3. The fundamental Lie algebra of the first-order Tate curve

The material in this section is an abridged account of results extracted from the papers [26], [25] and [27].

3.1. **Background.** Let $E_{\partial/\partial q}^\times$ denote the first order Tate curve, which is the fiber of the universal elliptic curve over $\mathcal{M}_{1,1}$ with respect to the tangential base point $\partial/\partial q$. Its de Rham fundamental group

$$(3.1) \qquad\qquad\qquad \mathcal{P} = \pi_1^{dR}(E_{\partial/\partial q}^\times, \overrightarrow{1}_1)$$

where $\overrightarrow{1}_1$ is the tangent vector of length 1 with respect to the canonical holomorphic coordinate w on $E_{\partial/\partial q}^\times$, is the de Rham realisation of a pro-object in the category of mixed Tate motives over $\mathbb{Z}$. Since its mixed Hodge structure is the limiting mixed Hodge structure of a variation, its relative monodromy weight filtration is denoted by M, and it comes equipped with a geometric weight filtration W. This data defines a universal mixed elliptic motive according to Hain and Matsumoto [27]. The associated M, W bigraded Lie algebra is the free Lie algebra $\mathbb{L}(H_{dR})$ where

$$H_{dR} = (H_{dR}^1(E_{\partial/\partial q}^\times; \mathbb{Q}))^\vee = \mathbb{Q}a \oplus \mathbb{Q}b \quad (= \mathbb{Q}(1) \oplus \mathbb{Q}(0))$$

which has two canonical de Rham generators a and b. It will be denoted by $\mathbb{L}(a, b)$. The generators a and b have (M, W) bidegrees $(-2, -1)$ and $(0, -1)$, respectively. The geometric weight filtration W coincides with the lower central series filtration on $\mathcal{P}$.

Since $\mathcal{P}$ is the de Rham realisation of a mixed Tate motive over $\mathbb{Z}$, it admits an action of the de Rham motivic Galois group G^{dR} of $\mathcal{MT}(\mathbb{Z})$ by the Tannakian formalism. Passing to Lie algebras gives a morphism

$$(3.2) \qquad i_1 : \mathfrak{g}^{\mathfrak{m}} \longrightarrow \mathrm{Der}^\Theta \mathbb{L}(a, b)$$

where $\Theta = [a, b]$ and

$$\mathrm{Der}^\Theta \mathbb{L}(a, b) = \{\delta \in \mathrm{Der}\, \mathbb{L}(a, b) : \delta(\Theta) = 0\} \ .$$

This is because $\Theta = [a, b]$ corresponds to the de Rham path which winds once around the puncture in $E^\times_{\partial/\partial q}$, and generates a copy of $\mathbb{Q}(1)$, which is fixed by U^{dR}.

3.2. **Derivations.** We only need to consider the subspace $B^0 \mathrm{Der}^\Theta \mathbb{L}(a, b)$ of derivations δ (see §3.4 for the definition of the B-filtration) such that

$$\delta(b) \in B^1\mathbb{L}(a, b) := \ker(\mathbb{L}(a, b) \to \mathbb{L}(a))$$

where the map on the right sends b to zero (it is the composition of the natural map $\mathbb{L}(H_{dR}) \to \mathbb{L}(H_{dR})^{ab} = H_{dR}$ followed by the projection $H_{dR} \to H_{dR}/F^0 H_{dR} = \mathbb{Q}a$). Such a derivation is uniquely determined by its value $\delta(a)$ since $[\delta(a), b] + [a, \delta(b)] = 0$, and the commutator of a is $a\,\mathbb{Q}$. Thus $\delta \mapsto \delta(a)$ gives an embedding $B^0 \mathrm{Der}^\Theta \mathbb{L}(a, b) \to \mathbb{L}(a, b)$.

For each $n \geq -1$, one shows that there exist elements

$$\varepsilon^\vee_{2n+2} \in B^0 \mathrm{Der}^\Theta \mathbb{L}(a, b) \subset \mathrm{Der}^\Theta \mathbb{L}(a, b)$$

which are uniquely determined by the property

$$\varepsilon^\vee_{2n+2}(a) = \mathrm{ad}(a)^{2n+2}(b) \ .$$

The elements $\varepsilon^\vee_{2n+2}$ were first defined by Nakamura in 1999 in the profinite setting. The element $\varepsilon^\vee_2$ is central in $\mathrm{Der}^\Theta \mathbb{L}(a, b)$ and plays no role here. Let us define

$$(3.3) \qquad \mathfrak{u}^{\mathrm{geom}} \subset \mathrm{Der}^\Theta \mathbb{L}(a, b)$$

to be the Lie subalgebra spanned by the $\varepsilon^\vee_{2n+2}$, for $n \geq -1$. One shows that the relative Malčev completion of $\pi_1(\mathcal{M}_{1,1}, \partial/\partial q) = \mathrm{SL}_2(\mathbb{Z})$ (relative to $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Q})$) acts on $\mathbb{L}(a, b)$ via $\mathfrak{u}^{\mathrm{geom}}$ [26]. Quadratic relations between the elements $\varepsilon^\vee_{2n}$ predicted by Hain and Matsumoto were studied by Pollack in his thesis [37].

We define heretical normalisations of these derivations as follows. Let

$$(3.4) \qquad \underline{\varepsilon}^\vee_0 = \frac{1}{12}\varepsilon^\vee_0 \qquad \text{and} \qquad \underline{\varepsilon}^\vee_{2n+2} = \frac{B_{2n}}{(2n)!}\varepsilon^\vee_{2n+2} \qquad \text{for } n \geq 1$$

where B_k denotes the kth Bernoulli number.

3.3. **The Hain morphism.** There is a morphism of fundamental groups [25], §16-18:

$$(3.5) \qquad \pi_1(\mathbb{P}^1\backslash\{0, 1, \infty\}, \overrightarrow{1_1}) \longrightarrow \pi_1(E^\times_{\partial/\partial q}, \partial/\partial w) \ .$$

Using the work of Levin and Racinet, Hain has computed this map in the de Rham realisation [25], (18.1). On de Rham Lie algebras it is the continuous morphism

$$(3.6) \qquad \begin{aligned} \phi : \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)^\wedge &\longrightarrow \mathbb{L}(a, b)^\wedge \\ \mathsf{x}_0 &\mapsto \frac{\mathrm{ad}(b)}{e^{\mathrm{ad}(b)} - 1}a = a - \frac{1}{2}[b, a] + \frac{1}{12}[b, [b, a]] + \dots \\ \mathsf{x}_1 &\mapsto [a, b] \end{aligned}$$

where $\wedge$ denotes completion with respect to the lower central series.

Let $\mathrm{Der}^1\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)^\wedge$, $\mathrm{Der}^\Theta\mathbb{L}(a,b)^\wedge$ denote the continuous derivations of completed Lie algebras which send, respectively, x_1 to 0 or $[a,b]$ to 0. We say that an element $\sigma \in \mathrm{Der}^1\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)^\wedge$ lifts to a derivation $\widetilde{\sigma} \in \mathrm{Der}^\Theta\mathbb{L}(a,b)^\wedge$ (and conversely, $\widetilde{\sigma}$ descends to the derivation σ) if

$$(3.7) \qquad \widetilde{\sigma} \circ \phi = \phi \circ \sigma \ .$$

which is equivalent to the equation $\widetilde{\sigma}\phi(\mathsf{x}_0) = \phi\sigma(\mathsf{x}_0)$. An element of $\mathrm{Der}^\Theta\mathbb{L}(a,b)^\wedge$ descends to an element of $\mathrm{Der}^1\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)^\wedge$ if and only if it preserves the subspace $\phi(\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)^\wedge)$ in $\mathbb{L}(a,b)^\wedge$. Since (3.6) is injective, σ and $\widetilde{\sigma}$ determine each other uniquely. Furthermore, lifting derivations preserves composition: if $\sigma_1,\sigma_2 \in \mathrm{Der}^1\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)^\wedge$ admit lifts $\widetilde{\sigma}_1,\widetilde{\sigma}_2 \in \mathrm{Der}^\Theta\mathbb{L}(a,b)^\wedge$, then $[\sigma_1,\sigma_2]$ admits the lift $[\widetilde{\sigma}_1,\widetilde{\sigma}_2]$.

Since (3.5) is geometric, it is compatible with the actions of G^{dR} on the de Rham fundamental groups (the case of the Hodge realisation is [25], Theorem 15.1). Thus ϕ commutes with the action of $\mathfrak{g}^{\mathfrak{m}}$, and the morphism (3.2) is the lift of the map $i_0 : \mathfrak{g}^{\mathfrak{m}} \to \mathrm{Der}^1\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$. In particular,

$$(3.8) \qquad i_1(\sigma)\big(\phi(\mathsf{x}_0)\big) = \phi(i_0(\sigma)(\mathsf{x}_0)) \qquad \text{for all } \sigma \in \mathfrak{g}^{\mathfrak{m}} \ ,$$

where i_0 was defined in (2.4). The injectivity of i [3] and ϕ implies the

Theorem 3.1. *The map $i_1 : \mathfrak{g}^{\mathfrak{m}} \to \mathrm{Der}^\Theta\mathbb{L}(a,b)$ is injective.*

3.4. B-**filtration.** One possible way to cut out the depth filtration on the image of $\mathfrak{g}^{\mathfrak{m}}$ inside $\mathrm{Der}^\Theta\mathbb{L}(a,b)$ is via the B-filtration on $\mathcal{P}$. As pointed out by Hain, it can be defined as the convolution of the Hodge filtration F and the lower central series filtration L:

$$B^r = (F \star L)^r = \sum_{a+b=r} F^a \cap L^b \ .$$

This induces a decreasing filtration on the Lie algebra of $\mathcal{P}$. Note that $L^b = W_{-b}$, where W denotes the weight filtration. Since the filtrations F, W, and M can be split simultaneously [25], B induces a filtration on $\mathbb{L}(H) = \mathrm{gr}^W\mathrm{gr}^M\mathrm{Lie}\,\mathcal{P}$, and

$$B^r\mathbb{L}(H) = \{w \in \mathbb{L}(a,b) : \deg_b w \geq r\} \ ,$$

is the filtration associated to the B-degree, where we define the B-degree on $\mathbb{L}(a,b)$ to be the degree in b. To see this, note that $B^r\mathrm{gr}_L^n\mathbb{L}(H_{dR}) = F^{n-r}\mathrm{gr}_L^n\mathbb{L}(H_{dR})$. Since $\mathrm{gr}_L^n\mathbb{L}(H_{dR})$ consists of words of length n in a and b, and since

$$\mathbb{Q}a \oplus \mathbb{Q}b = F^{-1}H_{dR} \supset F^0H_{dR} = \mathbb{Q}b \ ,$$

$F^{n-r}\mathrm{gr}_L^n\mathbb{L}(H_{dR})$ is spanned by words with at least r letter b's. Note also that

$$2\deg_b + 2\deg_W = \deg_M \ .$$

The B-filtration also induces a decreasing filtration $B^\bullet$ on $\mathrm{Der}^\Theta\mathbb{L}(a,b)$, such that $\mathrm{Der}^\Theta\mathbb{L}(a,b) = B^{-1}\mathrm{Der}^\Theta\mathbb{L}(a,b)$. It is the decreasing filtration associated to the grading induced by the B-degree. The subspace $B^0\mathrm{Der}^\Theta\mathbb{L}(a,b)$ is the space of derivations δ such that $\delta(b) \subset B^1\mathbb{L}(a,b)$, i.e., such that the coefficient of a in $\delta(b)$ is zero.

Lemma 3.2. *The B-filtration on $\mathcal{P}$ is motivic, i.e., it is stable under the image of the de Rham motivic Galois group $i_1(\mathfrak{g}^{\mathfrak{m}})$.*

Proof. We must verify that $i_1(\mathfrak{g}^{\mathfrak{m}}) \subset B^0\mathrm{Der}^\Theta\mathbb{L}(a,b)$. Since the group G^{dR} acts on $\mathbb{L}(a,b)^{ab} = H_{dR}$ through its quotient $\mathbb{G}_m$, and because $H_{dR} = \mathbb{Q} \oplus \mathbb{Q}(1)$ is a direct sum of pure Tate motives, the graded Lie algebra of its pro-unipotent radical $\mathfrak{g}^{\mathfrak{m}}$ acts trivially on $\mathbb{L}(H_{dR})^{ab}$. Therefore $i_1(\mathfrak{g}^{\mathfrak{m}})(b) \subset [\mathbb{L}(H_{dR}),\mathbb{L}(H_{dR})] \subset B^1\mathbb{L}(H_{dR})$. $\qquad\square$

Lemma 3.3. *Let $\alpha \in \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)^\wedge$. Then $\alpha \in D^r$ if and only if $\phi(\alpha) \in B^r$.*

Proof. The fact that $\phi\, D^r \subset B^r$ is clear from the definition (3.6). The converse follows from the fact that the associated graded morphism

$$\phi^0 : \mathrm{gr}_D^\bullet \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) \longrightarrow \mathrm{gr}_B^\bullet \mathbb{L}(a, b)$$

given by $\phi^0(\mathsf{x}_0) = a$ and $\phi^0(\mathsf{x}_1) = [a, b]$, is injective. $\qquad\square$

Observe that if $\delta \in \mathrm{Der}^1 \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ then $\delta \in D^r$ if and only if $\delta(\mathsf{x}_0) \in D^r$.

Lemma 3.4. *Let $\delta \in B^0 \mathrm{Der}^\Theta \mathbb{L}(a, b)$. The following are equivalent*

   (1) $\delta \in B^r \mathrm{Der}^\Theta \mathbb{L}(a, b)$
   (2) $\delta(a) \in B^r \mathbb{L}(a, b)$
   (3) $\delta(\phi(\mathsf{x}_0)) \in B^r \mathbb{L}(a, b)$.

Proof. Clearly (1) implies (2) and (3). Now suppose that (2) holds. We have

$$0 = \delta[a, b] = [a, \delta(b)] + [\delta(a), b]$$

which implies that $[a, \delta(b)] \in B^{r+1} \mathbb{L}(a, b)$. Since the coefficient of a in $\delta(b)$ vanishes, this implies that $\delta(b) \in B^{r+1}\mathbb{L}(a, b)$. Together with $\delta(a) \in B^r$ this implies (1).

Now suppose (3) holds. Write $\phi(\mathsf{x}_0) = a + w$ where $w \in B^1 \mathbb{L}(a, b)$, and by (3),

$$\delta(a) + \delta(w) \in B^r \ .$$

If we have shown that $\delta(a) \in B^i$, then by (2) $\Rightarrow$ (1), we have $\delta \in B^i$ and hence $\delta(w) \in B^{i+1}$. The previous equation then implies $\delta(a) \in B^{\min(i+1, r)}$. Starting with $i = 0$, repeat this argument to deduce that $\delta \in B^r$ which proves (3) $\Rightarrow$ (1). $\qquad\square$

Proposition 3.5. *Let $\sigma \in \mathrm{Der}^1 \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ which lifts to an element $\widetilde{\sigma} \in B^0 \mathrm{Der}^\Theta \mathbb{L}(a, b)$. Then $\sigma \in D^r$ if and only if $\widetilde{\sigma} \in B^r$.*

Proof. We have $\phi(\sigma(\mathsf{x}_0)) = \widetilde{\sigma}(\phi(\mathsf{x}_0))$. Now apply the previous two lemmas to deduce that $\sigma \in D^r$ if and only if $\widetilde{\sigma} \in B^r$. $\qquad\square$

Corollary 3.6. *The B-filtration cuts out the depth filtration on the image of $\mathfrak{g}^\mathfrak{m}$:*

$$(3.9) \qquad\qquad\qquad B^r \cap i_1(\mathfrak{g}^\mathfrak{m}) = i_1(D^r \mathfrak{g}^\mathfrak{m}) \ .$$

Proof. Apply lemma 3.2, the previous proposition, and $i_1 \phi = \phi\, i_0$ (3.8). $\qquad\square$

By considering the symmetry $t \mapsto 1 - t$ on $\mathbb{P}^1 \backslash \{0, 1, \infty\}$, one knows that an element $\sigma \in \mathfrak{g}^\mathfrak{m}$ of degree $-m$ is uniquely determined by its image in $D^1/D^{\lfloor \frac{m}{2} \rfloor} \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ (duality relation). It follows from the injectivity of ϕ and i that:

Corollary 3.7. *An element $\sigma \in \mathfrak{g}^\mathfrak{m}$ of degree $-m$ is uniquely determined by the image of $i_1(\sigma)$ in $B^1/B^{\lfloor \frac{m}{2} \rfloor} \mathbb{L}(a, b)$.*

Remark 3.8. Thus the 'tails' of the elements $i_1(\sigma)$ in the B-filtration, which may be infinitely long, are uniquely determined from their 'heads' $i_1(\sigma)$ (mod $B^{\lfloor \frac{m}{2} \rfloor}$). Furthermore, one can show (for instance, using the description [7], §9 of the group of automorphisms of a semi-direct product) that, modulo B^m, $i_1(\sigma)$ is equivalent to an element of $\mathfrak{u}^{\mathrm{geom}}$, so can be expressed (non-uniquely) in terms of the geometric derivations $\varepsilon_{2n+2}^\vee$ for $n \geq -1$. For instance, a choice of Drinfeld element of degree $-2n - 1$ admits an expansion

$$i_1(\sigma_{2n+1}) \equiv \sum_{a_1, \ldots, a_r \in 2\mathbb{N}} \lambda_{a_1, \ldots, a_r}^{(2n+1)} \{\varepsilon_{a_1}^\vee, \{\varepsilon_{a_2}^\vee, \{\ldots, \{\varepsilon_{a_{r-1}}^\vee, \varepsilon_{a_r}^\vee\} \cdots\} \quad (\mathrm{mod}\ B^{2n+1})$$

where $\lambda_{a_1,\ldots,a_r}^{(2n+1)} \in \mathbb{Q}$. The non-uniqueness of this expansion corresponds to the fact that there exist relations between the generators $\varepsilon_{2n+2}^\vee$ in $\mathfrak{u}^{\mathrm{geom}}$ [37]. Nonetheless, this geometric expansion or 'anatomy' of the elements $i_1(\sigma_{2n+1}^c)$, for $n \geq 2$, will be determined explicitly modulo B^4 below.

4. Commutative power series

We discuss commutative power series representations for the Lie algebras $\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$, $\mathbb{L}(a, b)$ and describe the composition laws for their derivation algebras.

4.1. Commutative power series.
The method of commutative power series is based on the observation that there is an isomorphism of $\mathbb{Q}$-vector spaces, for $r \geq 0$

$$(4.1) \qquad \begin{aligned} \rho \quad : \quad \mathrm{gr}_D^r T(\mathsf{x}_0, \mathsf{x}_1) \quad &\xrightarrow{\sim} \quad \mathbb{Q}[y_0, \ldots, y_r] \\ \mathsf{x}_0^{i_0} \mathsf{x}_1 \mathsf{x}_0^{i_1} \ldots \mathsf{x}_1 \mathsf{x}_0^{i_r} \quad &\mapsto \quad y_0^{i_0} \ldots y_r^{i_r} \end{aligned}$$

Let us denote by

$$P = \bigoplus_{r \geq 0} \mathbb{Q}[y_0, \ldots, y_r] \ .$$

Since $\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ is D-graded, (4.1) induces a map $\rho : \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) \to P$. One checks that

$$(4.2) \qquad \rho(\mathrm{ad}(\mathsf{x}_0)^{2n}(\mathsf{x}_1)) = (y_0 - y_1)^{2n} \ .$$

Furthermore, one shows ([6], lemma 6.2) that the image of $\mathrm{gr}_D^r \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$, for $r \geq 1$, is contained in the subspace of polynomials which are translation invariant:

$$f(y_0, \ldots, y_r) = f(y_0 + \lambda, \ldots, y_r + \lambda) \quad \text{for all} \quad \lambda \in \mathbb{Q} \ .$$

Such a polynomial f is uniquely determined by its image in

$$(4.3) \qquad \begin{aligned} \mathbb{Q}[y_0, \ldots, y_r] \quad &\longrightarrow \quad \mathbb{Q}[x_1, \ldots, x_r] \\ f(y_0, \ldots, y_r) \quad &\mapsto \quad \overline{f}(x_1, \ldots, x_r) = f(0, x_1, \ldots, x_r) \ . \end{aligned}$$

We sometimes call this the reduced representation of a translation-invariant polynomial, and it applies equally well to translation-invariant rational functions. In this way, since $\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) \cong \mathrm{gr}_D^\bullet \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ is graded with respect to the D-degree, every element of $D^1 \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ can be uniquely written

$$(4.4) \qquad \begin{aligned} \overline{\rho} \quad : D^1 \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) \quad &\longrightarrow \quad \bigoplus_{r \geq 1} \mathbb{Q}[x_1, \ldots, x_r] \\ \sigma \quad &\mapsto \quad \sum_{r \geq 1} \sigma^{(r)} \end{aligned}$$

where $\sigma^{(r)}$ denotes the polynomial representation of the component in depth r of σ. The Drinfeld elements satisfy $\sigma_{2n+1}^{(1)} = x_1^{2n}$, for all $n \geq 1$ by (4.2).

4.2. Concatenation products.
The concatenation of words in the alphabet $\mathsf{x}_0, \mathsf{x}_1$ defines a non-commutative multiplication law $f, g \mapsto f \cdot g$:

$$\mathrm{gr}_D^r \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) \times \mathrm{gr}_D^s \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) \longrightarrow \mathrm{gr}_D^{r+s} \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) \ .$$

On the level of commutative power series, it is the operation

$$(4.5) \qquad \begin{aligned} \mathbb{Q}[y_0, \ldots, y_r] \otimes \mathbb{Q}[y_0, \ldots, y_{r+s}] \quad &\longrightarrow \quad \mathbb{Q}[y_0, \ldots, y_{r+s}] \\ f(y_0, \ldots, y_r) \cdot g(y_0, \ldots, y_{r+s}) \quad &= \quad f(y_0, \ldots, y_r) g(y_r, \ldots, y_{r+s}) \end{aligned}$$

It follows from the definition of the linearized Ihara action (2.7) that

$$(4.6) \qquad f \underline{\circ} (g \cdot h) = (f \underline{\circ} g) \cdot h + g \cdot (f \underline{\circ} h) - g \cdot f \cdot h \ .$$

Note that equation (4.6) is equivalent to the condition that the linear map
$$w \mapsto f \mathbin{\underline{\circ}} w - w \cdot f$$
is a derivation with respect to $\cdot$. There is another concatenation product, denoted by $\underline{\cdot}$, which comes from the stuffle Hopf algebra [2]. It will only be used once in this paper, so will not be discussed.

4.3. **Linearized Ihara action.** The operator $\underline{\circ} : T(\mathsf{x}_0, \mathsf{x}_1) \otimes_{\mathbb{Q}} T(\mathsf{x}_0, \mathsf{x}_1) \to T(\mathsf{x}_0, \mathsf{x}_1)$ is homogeneous for the D-degree, and therefore defines a map

$$(4.7) \qquad \begin{aligned} \underline{\circ} : \mathbb{Q}[y_0, \ldots, y_r] \otimes_{\mathbb{Q}} \mathbb{Q}[y_0, \ldots, y_s] &\longrightarrow \mathbb{Q}[y_0, \ldots, y_{r+s}] \\ f(y_0, \ldots, y_r) \otimes g(y_0, \ldots, y_s) &\mapsto f \mathbin{\underline{\circ}} g \, (y_0, \ldots, y_{r+s}) \end{aligned}$$

whose r, s component is given explicitly by the formula

$$(4.8) \quad f \mathbin{\underline{\circ}} g \, (y_0, \ldots, y_{r+s}) = \sum_{i=0}^{s} f(y_i, y_{i+1}, \ldots, y_{i+r}) g(y_0, \ldots, y_i, y_{i+r+1}, \ldots, y_{r+s}) +$$

$$(-1)^{\deg f + r} \sum_{i=1}^{s} f(y_{i+r}, \ldots, y_{i+1}, y_i) g(y_0, \ldots, y_{i-1}, y_{i+r}, \ldots, y_{r+s}) \ .$$

This can be read off from equation (2.7). Antisymmetrizing gives a pairing
$$\{f, g\} = f \mathbin{\underline{\circ}} g - g \mathbin{\underline{\circ}} b$$
which coincides with the Ihara bracket on the image of $\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$. Clearly, if f, g are both translation invariant, then so too are $f \mathbin{\underline{\circ}} g$ and $\{f, g\}$.

Example 4.1. When $r = s = 1$ and $f, g \in \mathbb{Q}[y_0, y_1]$ are translation-invariant, one verifies that the image of $\{f, g\} \in \mathbb{Q}[y_0, y_1, y_2]$ in $\mathbb{Q}[x_1, x_2]$ is given by

$$\overline{f}(x_1)\overline{g}(x_2) - \overline{g}(x_1)\overline{f}(x_2) + \overline{f}(x_2 - x_1)(\overline{g}(x_1) - \overline{g}(x_2)) + (\overline{f}(x_2) - \overline{f}(x_1))\overline{g}(x_2 - x_1) \ .$$

The Ihara action (4.8) extends to rational functions by the identical formula.

4.4. **Derivations on $\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ and power series.** Consider the isomorphism

$$(4.9) \qquad\qquad \begin{aligned} \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) &\xrightarrow{\sim} \mathrm{Der}^1 \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) \\ w &\mapsto \delta_w \end{aligned}$$

where δ_w denotes the derivation $\delta_w(\mathsf{x}_0) = w$, $\delta_w(\mathsf{x}_1) = 0$. This isomorphism respects the D-grading on both sides. Let

$$P' = \bigoplus_{r \geq 1} \mathbb{Q}[y_0, \ldots, y_r] \frac{1}{y_0 - y_r} \ .$$

Using (4.1), define a linear map

$$(4.10) \qquad\qquad \begin{aligned} \rho' : D^1 \mathrm{Der}^1 \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) &\longrightarrow P' \\ \delta_w &\mapsto \sum_{r \geq 1} \rho^r(w) \frac{1}{y_0 - y_r} \ . \end{aligned}$$

Proposition 4.2. *The following diagram commutes:*

$$\begin{array}{ccccc} D^1 \mathrm{Der}^1 \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) & \times & \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) & \longrightarrow & \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) \\ \downarrow{\scriptstyle\rho'} & & \downarrow{\scriptstyle\rho} & & \downarrow{\scriptstyle\rho} \\ \odot : \qquad P' & \times & P & \longrightarrow & P \end{array}$$

where for $f \in P'$ and $g \in P$,

$$(4.11) \qquad\qquad\qquad f \odot g = f \mathbin{\underline{\circ}} g - f \cdot g \ .$$

Furthermore, the composition of derivations is given by the linearised Ihara bracket:

(4.12) $$\rho'([\delta_w, \delta_v]) = \{\rho'(\delta_w), \rho'(\delta_v)\} \ .$$

Thus $\rho' : D^1\mathrm{Der}^1\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) \to (P', \{,\})$ is a morphism of Lie algebras.

Proof. The shuffle distributivity law (4.6), and the remark which follows, implies that $\odot$ is a derivation: $f \odot (g \cdot h) = (f \odot g) \cdot h + g \cdot (f \odot h)$, for all $f \in P'$, $g, h \in P$. It therefore suffices to show that for all $w \in \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$, and $i = 0, 1$, we have

$$\rho'(\delta_w) \odot \rho(\mathsf{x}_i) = \rho(\delta_w(\mathsf{x}_i)) \ .$$

Without loss of generality, assume that w is of D-degree $r \geq 1$, and write $f = \rho'(w)$. Since $w \in \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$, we have $w + w^* = 0$ and hence

$$f(y_0, \ldots, y_r) + (-1)^{\deg f + r} f(y_r, \ldots, y_0) = 0 \ .$$

We check that $\rho(\mathsf{x}_0) = y_0 \in \mathbb{Q}[y_0]$ and $\rho(\mathsf{x}_1) = 1 \in \mathbb{Q}[y_0, y_1]$, and verify that

$$\begin{aligned} f(y_0, \ldots, y_r) \, \underline{\circ} \, y_0 &= y_0 f(y_0, \ldots, y_r) \\ f(y_0, \ldots, y_r) \, \underline{\circ} \, 1 &= f(y_0, \ldots, y_r) \end{aligned}$$

directly from the definition (4.8), applied in the cases $s = 0, g = y_0$ and $s = 1, g = 1$ respectively. Via (4.11), these equations imply that $f \odot y_0 = (y_0 - y_r)f$ and $f \odot 1 = 0$. This proves that $\rho'(\delta_w) \odot \rho(\mathsf{x}_0) = \rho(w)$ and $\rho'(\delta_w) \odot \rho(\mathsf{x}_1) = 0$ as required. For the last part, use the fact that $\mathrm{Der}^1\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ acts faithfully on $\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ and the identity

$$f \odot (g \odot h) - g \odot (f \odot h) = \{f, g\} \odot h$$

which follows from (4.11), (4.6) and (2.9). $\qquad\qquad\Box$

The following corollary is not required for the remainder of this paper.

Corollary 4.3. *Let $w \in \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ of D-degree r, and write $f = \overline{\rho}^r(w)$. Then*

(4.13) $$\overline{\rho}^{(r+1)}(\delta_{\mathsf{x}_1} w) = \{f x_r^{-1}, x_1^{-1}\} x_{r+1} \quad \in \quad \mathbb{Q}[x_1, \ldots, x_{r+1}]$$

Proof. Use (4.9) and (4.10). Apply (4.12) to $[\delta_{\mathsf{x}_1}, \delta_w] = \delta_{\delta_{\mathsf{x}_1} w}$ to give

$$\rho'(\delta_{\delta_{\mathsf{x}_1} w}) = \{\rho'(\delta_{\mathsf{x}_1}), \rho'(\delta_w)\} \ .$$

The left-hand side is $\frac{\rho^{(r+1)}(\delta_{\mathsf{x}_1} w)}{(y_0 - y_{r+1})}$, the right-hand side is $\{\frac{1}{y_0 - y_1}, \frac{\rho^{(r)}(w)}{y_0 - y_r}\}$. Then pass to the reduced representation $(y_0, y_1, \ldots, y_r) \mapsto (0, x_1, \ldots, x_r)$. $\qquad\Box$

4.5. **Derivations on $\mathbb{L}(a, b)$ and power series.** Define

$$D^{\Theta} = \mathrm{Der}^{\Theta}(\mathbb{L}(a, [a, b]), B^1\mathbb{L}(a, b))$$

to be the vector space of linear maps $\delta : \mathbb{L}(a, [a, b]) \to B^1\mathbb{L}(a, b)$ satisfying $\delta[p, q] = [\delta(p), q] + [p, \delta(q)]$ and $\delta([a, b]) = 0$. Such a $\delta \in D^{\Theta}$ is uniquely determined by the element $\delta(a) \in B^1\mathbb{L}(a, b)$. Furthermore, there is an injective map

$$B^1\mathrm{Der}^{\Theta}(\mathbb{L}(a, b)) \longrightarrow D^{\Theta}$$

obtained by restricting to the Lie subalgebra $\mathbb{L}(a, [a, b]) \subset \mathbb{L}(a, b)$.

Denote also by ρ the linear map

(4.14) $$\begin{aligned} \rho^{(r)} \quad &: \quad \mathrm{gr}_B^r T(a, b) \quad \overset{\sim}{\longrightarrow} \quad \mathbb{Q}[y_0, \ldots, y_r] \\ &\quad\ a^{i_0} b a^{i_1} \ldots b a^{i_r} \quad \mapsto \quad y_0^{i_0} \ldots y_r^{i_r} \ . \end{aligned}$$

Let $\ell_0 = 1$ and for $r \geq 1$, set

(4.15) $$\ell_r = (y_0 - y_1)(y_1 - y_2) \ldots (y_{r-1} - y_r) \ .$$

Define a graded vector space

$$Q = \bigoplus_{r \geq 0} \ell_r^{-1} \mathbb{Q}[y_0, \ldots, y_r]$$

and a linear map

(4.16)
$$\ell : \mathbb{L}(a, b) \longrightarrow Q$$
$$w \longmapsto \sum_r \ell_r^{-1} \rho^{(r)}(w)$$

Since $\ell_r \cdot \ell_s = \ell_{r+s}$, Q is an algebra for the shuffle concatenation (4.5). Define

$$Q' = \bigoplus_{r \geq 1} Q_r \frac{1}{y_0 - y_r}$$

and setting $c_r = \ell_r(y_0 - y_r)$ (c for 'cyclic') consider the linear map

(4.17)
$$\ell' : D^\Theta \longrightarrow Q'$$
$$\delta \longmapsto \sum_{r \geq 1} c_r^{-1} \rho^{(r)} \delta(a) \ .$$

It is injective, since $\delta(a)$ uniquely determines $\delta \in D^\Theta$.

Proposition 4.4. *The following diagram commutes:*

$$
\begin{array}{ccccc}
D^\Theta & \times & \mathbb{L}(a, [a, b]) & \longrightarrow & \mathbb{L}(a, b) \\
\downarrow{\scriptstyle \ell'} & & \downarrow{\scriptstyle \ell} & & \downarrow{\scriptstyle \ell} \\
\circledast : \quad Q' & \times & Q & \longrightarrow & Q
\end{array}
$$

where, as in (4.11), we have

(4.18)
$$f \circledast g = f \circ g - f \cdot g \ .$$

Restricting to the subspace $B^1 \mathrm{Der}^\Theta \mathbb{L}(a, b) \subset D^\Theta$ we obtain a commutative diagram

$$
\begin{array}{ccccc}
B^1 \mathrm{Der}^\Theta \mathbb{L}(a, b) & \times & \mathbb{L}(a, b) & \longrightarrow & \mathbb{L}(a, b) \\
\downarrow{\scriptstyle \ell'} & & \downarrow{\scriptstyle \ell} & & \downarrow{\scriptstyle \ell} \\
\circledast : \quad Q' & \times & Q & \longrightarrow & Q
\end{array}
$$

Furthermore, we have the identity for all $\delta_1, \delta_2 \in B^1 \mathrm{Der}^\Theta \mathbb{L}(a, b)$:

(4.19)
$$\ell'([\delta_1, \delta_2]) = \{\ell'(\delta_1), \ell'(\delta_2)\} \ .$$

Thus $\ell' : B^1 \mathrm{Der}^\Theta \mathbb{L}(a, b) \to (Q', \{,\})$ is a Lie algebra homomorphism.

Proof. The proof is similar to the proof of proposition 4.2, except that we must check that $\ell'(\delta) \circledast \ell([a, b]) = 0$ and $\ell'(\delta) \circledast \ell(a) = \ell(\delta(a))$. But $\ell([a, b]) = \frac{y_0 - y_1}{y_0 - y_1} = 1$ and $\ell(a) = y_0$, so this calculation is formally identical to the one in proposition 4.2. $\qquad \square$

4.6. **Double shuffle equations.** The equations defining $\mathfrak{dmr}_0$ can be spelt out explicitly and translated via (4.1) into the language of commutative power series [2]. We shall only require their restriction to depths ≤ 3 and work with translation-invariant representations (4.3). Let

$$(f^{(1)}, f^{(2)}, f^{(3)}) \quad \in \quad \mathbb{Q}[x_1] \oplus \mathbb{Q}[x_1, x_2] \oplus \mathbb{Q}[x_1, x_2, x_3]$$

Writing x_{ij} for $x_i + x_j$ and x_{ijk} for $x_i + x_j + x_k$, the shuffle equations modulo products in depths 2 and 3 are given by

(4.20)
$$f^{(2)}(x_1, x_{12}) + f^{(2)}(x_2, x_{12}) = 0$$
$$f^{(3)}(x_1, x_{12}, x_{123}) + f^{(3)}(x_2, x_{12}, x_{123}) + f^{(3)}(x_2, x_{23}, x_{123}) = 0 \ .$$

The solutions to the shuffle equations modulo products correspond, via (4.1), to the image of $\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ inside $T(\mathsf{x}_0, \mathsf{x}_1)$. The stuffle equations modulo products, in depths 2 and 3, correspond to the (regularised versions of) the equations ($a, b, c \in \mathbb{N}$):

$$\zeta(a, b) + \zeta(b, a) + \zeta(a + b) \equiv 0 \quad (\text{mod products})$$

$$\zeta(a, b, c) + \zeta(b, a, c) + \zeta(b, c, a) + \zeta(a + b, c) + \zeta(a, b + c) \equiv 0 \quad (\text{mod products}) \,.$$

By considering the series $Z^{(r)} = \sum_{n_1,\ldots,n_r \geq 0} \zeta_*(n_1, \ldots, n_r) x_1^{n_1 - 1} \ldots x_r^{n_r - 1}$, where the subscript $*$ denotes the stuffle regularisation, the previous equations translate into

$$(4.21) \qquad f^{(2)}(x_1, x_2) + f^{(2)}(x_2, x_1) \quad = \quad \frac{f^{(1)}(x_1) - f^{(1)}(x_2)}{x_2 - x_1}$$

$$f^{(3)}(x_1, x_2, x_3) + f^{(3)}(x_2, x_1, x_3) + f^{(3)}(x_2, x_3, x_1) \quad = $$
$$\frac{f^{(2)}(x_2, x_1) - f^{(2)}(x_2, x_3)}{x_3 - x_1} + \frac{f^{(2)}(x_1, x_3) - f^{(2)}(x_2, x_3)}{x_2 - x_1}$$

Note that the right-hand sides of the equations are in fact polynomials. These equations extend to an infinite family of equations in every depth [2]. The Lie algebra $\mathfrak{dmr}_0$ is defined to be the sets of solutions to both shuffle and stuffle equations.

The linearised double shuffle equations [6] are the same sets of equations in which the right-hand sides are zero. The linearised shuffle equations are identical to the ordinary shuffle equations, but the linearised stuffle equations are:

$$(4.22) \qquad\qquad\qquad f^{(1)}(x_1) + f^{(1)}(-x_1) \quad = \quad 0$$
$$f^{(2)}(x_1, x_2) + f^{(2)}(x_2, x_1) \quad = \quad 0$$
$$f^{(3)}(x_1, x_2, x_3) + f^{(3)}(x_2, x_1, x_3) + f^{(3)}(x_2, x_3, x_1) \quad = \quad 0$$

The linearised double shuffle equations are also closed under the Ihara bracket, by a (simpler) version of Racinet's theorem.

4.7. Geometric derivations and linearised double shuffle equations with poles.
The Lie algebra $\mathfrak{u}^{\text{geom}} \subset B^1 \text{Der}^\Theta \mathbb{L}(a, b)$ of geometric derivations was defined in (3.3). The definition of the linearised double shuffle equations (4.20) and (4.22) can be extended in the obvious way to rational functions.

Definition 4.5. Define $\mathfrak{pls} \subset Q'$ to be the subspace of Q' satisfying the linearised double shuffle equations. It is bigraded by weight and depth.

By a version of Racinet's theorem, $\mathfrak{pls}$ is also a Lie subalgebra of Q' for the linearised Ihara bracket $\{\,,\}$. The notation stands for 'polar linearised double shuffle' solutions. It is a bigraded Lie algebra in the category of $\mathfrak{sl}_2$-representations over $\mathbb{Q}$.

Proposition 4.6. *The geometric derivations, in their reduced rational function representation* (4.17)*, satisfy the linearised double shuffle equations:*

$$\overline{\ell'}(\mathfrak{u}^{\text{geom}}) \subset \mathfrak{pls} \,.$$

Proof. The images of the generators $\overline{\ell'}(\varepsilon_{2n+2}^\vee) = x_1^{2n}$ by (4.2) for $n \geq -1$. They are even and hence solutions to the linearised double shuffle equations. It follows from (4.19), that $\overline{\ell'}$ is a morphism of Lie algebras since $\mathfrak{pls}$ is closed under $\{\,,\}$. $\square$

A natural question to ask is whether $\mathfrak{pls} = \mathfrak{u}^{\text{geom}}$. It is true in depths ≤ 3 and in certain limits [2].

The previous proposition implies that the stuffle equations define maps from the space of non-geometric derivations

$$\mathrm{Der}^\Theta \, \mathbb{L}(a,b)/\mathfrak{u}^{\mathrm{geom}}$$

to spaces of rational functions. We shall show in remark 5.8 that this map is non-zero and provides a tool to prove that certain derivations are not geometric. This can in particular be applied to the image of $i_1(\mathfrak{g}^\mathfrak{m})$ in $\mathrm{Der}^\Theta \, \mathbb{L}(a,b)/\mathfrak{u}^{\mathrm{geom}}$.

5. Drinfeld elements in depth 3 via anatomical construction

We wish to write down elements

$$\sigma^c_{2n+1} \in D^1\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)/D^4\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$$

by exhibiting explicit polynomials

$$(\sigma^{(1)}_{2n+1}, \sigma^{(2)}_{2n+1}, \sigma^{(3)}_{2n+1}) \in \mathbb{Q}[x_1] \oplus \mathbb{Q}[x_1,x_2] \oplus \mathbb{Q}[x_1,x_2,x_3]$$

which are solutions to the equations (4.20) and (4.21).

5.1. **Polar solutions.** The shape of the equations (4.21) suggests searching for solutions amongst the space of rational functions in x_i with $\mathbb{Q}$-coefficients. Let

$$(5.1) \qquad s^{(1)} = \frac{1}{2\,x_1} \qquad \text{and} \qquad s^{(2)} = \frac{1}{12}\left(\frac{1}{x_1 x_2} + \frac{1}{x_2(x_1-x_2)}\right).$$

It is easy to verify that $(s^{(1)}, s^{(2)})$ is a solution to the double shuffle equations (4.20) and (4.21) in depths one and two.

Definition 5.1. For $n \geq -1$, define rational functions in x_1, x_2, x_3 by

$$(5.2) \qquad \begin{aligned} \xi^{(1)}_{2n+1} &= x_1^{2n} \\ \xi^{(2)}_{2n+1} &= \{s^{(1)}, x_1^{2n}\} \\ \xi^{(3)}_{2n+1} &= \{s^{(2)}, x_1^{2n}\} + \frac{1}{2}\{s^{(1)}, \{s^{(1)}, x_1^{2n}\}\} \end{aligned}$$

where curly brackets denote the linearised Ihara bracket. Explicitly, we have

$$(5.3) \qquad \xi^{(2)}_{2n+1} = \frac{x_2^{2n} - (x_2-x_1)^{2n}}{x_1} + \frac{x_1^{2n} - x_2^{2n}}{x_2 - x_1} + \frac{(x_2-x_1)^{2n} - x_1^{2n}}{x_2}$$

which defines a polynomial in $\mathbb{Q}[x_1,x_2]$ whenever $n \geq 0$. On the other hand, $\xi^{(3)}_{2n+1}$ is a rational function in x_1, x_2, x_3 with non-trivial poles. When $n \geq 0$ it has at most simple poles along $x_1 = 0, x_3 = 0, x_1 = x_2$ and $x_2 = x_3$.

One checks that the case $n = 0$ is trivial: $\xi^{(1)}_1 = 1$ and $\xi^{(2)}_1 = \xi^{(3)}_1 = 0$.

Proposition 5.2. *Let $n \geq -1$. The elements*

$$\xi_{2n+1} = (\xi^{(1)}_{2n+1}, \xi^{(2)}_{2n+1}, \xi^{(3)}_{2n+1})$$

satisfy the double shuffle equations modulo products (4.20) *and* (4.21) *in depths* $2, 3$.

Proof. This is a finite computation and only uses the fact that x_1^{2n} is an even function. It also follows from the fact that $(s^{(1)}, s^{(2)})$, and x_1^{2n} are solutions to the double shuffle equations via a version of Racinet's theorem for rational functions. $\qquad \square$

Remark 5.3. The elements ξ_{2n+1} can be extended to all higher depths by the equation

$$\xi_{2n+1} = \exp(\mathrm{ad}(s))x_1^{2n}$$

to define solutions to the full set of double shuffle equations with poles, where s is one of (many possible) solutions to the polar double shuffle equations in weight 0. This is discussed in [2], and implies the previous proposition. The component of s in depth 3 is unique by an extension of the depth-parity theorem for double shuffle equations (theorem 2.5) to the case of rational functions and given by $s^{(3)} = \frac{1}{2}\{s^{(1)}, s^{(2)}\}$.

5.2. **Definition of canonical elements.** It is convenient to define heretical normalisations of the elements ξ as follows. Let

(5.4) $$\underline{\xi}_{-1} = \frac{1}{12}\,\xi_{-1} \qquad \text{and} \qquad \underline{\xi}_{2n+1} = \frac{B_{2n}}{(2n)!}\,\xi_{2n+1} \quad \text{for } n \geq 0$$

where B_{2n} is the $2n^{\mathrm{th}}$ Bernoulli number. Set

(5.5) $$b(x) = \frac{1}{e^x - 1} + \frac{1}{2}$$

Recall the well-known functional identity

(5.6) $$b(x_1)b(x_2) - b(x_1)b(x_2 - x_1) + b(x_2)b(x_2 - x_1) = \frac{1}{4}\;.$$

Definition 5.4. Let $n \geq 2$. Define elements $\underline{\sigma}_{2n+1}^c \in \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)/D^4\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ by

(5.7) $$\rho(\underline{\sigma}_{2n+1}^c) = \underline{\xi}_{2n+1} + \sum_{a+b=n} \frac{1}{2b}\{\underline{\xi}_{2a+1}, \{\underline{\xi}_{2b+1}, \underline{\xi}_{-1}\}\} \quad (\mathrm{mod}\ D^4)$$

where the sum is over $a, b \geq 1$. Definition (5.7) makes sense, since we shall prove in the next paragraph that the right-hand side has no poles. Define

$$\sigma_{2n+1}^c = \frac{(2n)!}{B_{2n}}\underline{\sigma}_{2n+1}^c \qquad \text{for } n \geq 2$$

to be the canonical normalisations, and set $\sigma_3^c = [\mathsf{x}_0, [\mathsf{x}_0, \mathsf{x}_1]] + [\mathsf{x}_1, [\mathsf{x}_0, \mathsf{x}_1]]$.

Since by Racinet's theorem, the space of solutions to the double shuffle equations is closed under the Ihara bracket, this means that for $n \geq 2$ the elements σ_{2n+1}^c are solutions to the double shuffle equations in depths ≤ 3.

Remark 5.5. The σ_{2n+1}^c have the canonical normalisation. For $n \geq 1$,

$$(\sigma_{2n+1}^c)^{(1)} = \xi_{2n+1}^{(1)} = x_1^{2n} \qquad \text{and} \qquad (\sigma_{2n+1}^c)^{(2)} = \xi_{2n+1}^{(2)}$$

is given by (5.3). We have $(\sigma_3^c)^{(3)} = 0$ and for $n \geq 2$,

(5.8) $$(\sigma_{2n+1}^c)^{(3)} = \xi_{2n+1}^{(3)} + \sum_{a+b=n} \frac{B_{2a}B_{2b}}{B_{2n}}\binom{2n}{2a}\frac{1}{24\,b}\{x_1^{2a}, \{x_1^{2b}, x_1^{-2}\}\}\;.$$

Due to the obvious symmetry in a and b, the previous expression can also be written in terms of lowest weight vectors for the action [2] of $\mathfrak{sl}_2$, namely:

$$\frac{1}{2b}\{x_1^{2a}, \{x_1^{2b}, x_1^{-2}\}\} + \frac{1}{2a}\{x_1^{2b}, \{x_1^{2a}, x_1^{-2}\}\}\;.$$

On the other hand, compare the odd part of the period polynomial [31] for the Eisenstein series of weight $2n$, which is proportional to:

$$\sum_{a+b=n, a,b \geq 1} \binom{2n}{2a} B_{2a}B_{2b}X^{2a-1}Y^{2b-1} \quad \in \quad \mathbb{Q}[X, Y]\;.$$

This is no accident, and follows from the computations in §5.3 as well as §9.

5.3. **Cancellation of poles.** We show that (5.8) has no poles. We need the following notation. Given two even functions f, g of one variable, define a new function

$$(f \star g)(x_1, x_2) = f(x_1)g(x_2) - f(x_2-x_1)g(x_2) + f(x_2-x_1)g(x_1) - f(x_2)g(x_1) \ .$$

In the notation of [2] it is $f \star g = f \circ g - g : f$.

Lemma 5.6. *For all $n, a, b \geq 1$,*

$$(5.9) \qquad \mathrm{Res}_{x_3=0}\big(\xi_{2n+1}^{(3)}\big) \ = \ \frac{1}{12}x_1^{2n} \star x_1^{-1}$$

$$\mathrm{Res}_{x_3=0}\big(\{x_1^{2a}, \{x_1^{2b}, x_1^{-2}\}\}\big) \ = \ 2b\, x_1^{2a} \star x_1^{2b-1} \ .$$

Proof. This is a straightforward computation and follows from the definitions. See also [2] for a generalisation of the second equation. $\qquad\square$

Proposition 5.7. *The elements $(\sigma_{2n+1}^c)^{(3)}$ have no poles, for all $n \geq 2$.*

Proof. Since $\mathfrak{pls}$ (definition 4.5) is a Lie algebra for the Ihara bracket and is contained in Q', the element $\{x_1^{2a}, \{x_1^{2b}, x_1^{-2}\}\}$ is in $\mathfrak{pls}$ has at most simple poles along $x_1 = 0, x_2 = x_1, x_3 = x_2$ and $x_3 = 0$, whenever $a, b \geq -1$. We first check that the residue of $(\sigma_{2n+1}^c)^{(3)}$ along $x_3 = 0$ vanishes for $n \geq 2$. It is given via (5.8) by

$$\frac{B_{2n}}{(2n)!}\mathrm{Res}_{x_3=0}\big(\xi_{2n+1}^{(3)}\big) + \frac{1}{12}\sum_{a+b=n}\frac{B_{2a}}{(2a)!}\frac{B_{2b}}{(2b)!}\frac{1}{2b}\mathrm{Res}_{x_3=0}\big(\{x_1^{2a}, \{x_1^{2b}, x_1^{-2}\}\}\big) \ .$$

Pass to generating series and substitute (5.9) into the previous expression to give

$$12\sum_{n\geq 1}\mathrm{Res}_{x_3=0}\,\underline{\sigma}_{2n+1}^{(3)} \ = \ (xb(x) - 1) \star x^{-1} + (xb(x) - 1) \star (b(x) - x^{-1})$$

$$= \ (xb(x) - 1) \star b(x) \ ,$$

where $b(x)$ was defined in (5.5). To compute this quantity, observe that $1 \star f = 0$ and

$$(xf \star f)(x_1, x_2) = (x_1 - x_2)\big(f(x_1)f(x_2) - f(x_1)f(x_2-x_1) + f(x_2)f(x_2-x_1)\big) \ .$$

for any even function f. Substituting for $f = b(x)$ and using (5.6), we deduce that $(xb(x) - 1) \star b(x) = \frac{1}{4}(x_1 - x_2)$. Now let $n \geq 2$. The above argument proves that the $(\sigma_{2n+1}^c)^{(3)}$ have no poles along $x_3 = 0$. Now we use the fact that σ_{2n+1}^c satisfies the double shuffle equations modulo products in depths two and three. Since $(\sigma_{2n+1}^c)^{(i)}$ has no poles for $i = 1, 2$, the stuffle equation (4.21) implies that

$$(5.10) \qquad (\sigma_{2n+1}^c)^{(3)}(x_1, x_2, x_3) - (\sigma_{2n+1}^c)^{(3)}(x_3, x_2, x_1) \in \mathbb{Q}[x_1, x_2, x_3] \ .$$

It follows that its residue at $x_1 = 0$ also vanishes. The shuffle equation is

$$(\sigma_{2n+1}^c)^{(3)}(x_1, x_{12}, x_{123}) + (\sigma_{2n+1}^c)^{(3)}(x_2, x_{12}, x_{123}) + (\sigma_{2n+1}^c)^{(3)}(x_2, x_{23}, x_{123}) = 0 \ .$$

By taking the residue of this expression at $x_2 = 0$, we deduce that $(\sigma_{2n+1}^c)^{(3)}$ has no pole along $x_2 = x_1$. Finally by (5.10), this implies that it has no pole along $x_2 = x_3$ either. $\qquad\square$

The last part of this argument is completely general and follows from the dihedral symmetry structure of the linearised double shuffle equations [6], §6.3.

Remark 5.8. The element $(\sigma_3^c)^{(3)}$ does have poles. It is equal to $3z_3$, where

$$z_3 = \frac{4}{3} + \frac{x_1}{x_3 - x_2} + \frac{x_3}{x_1 - x_2} + \frac{x_3 - x_2}{x_1} + \frac{x_1 - x_2}{x_3}$$

and corresponds to a lift to $\mathrm{Der}^\Theta \mathbb{L}(a,b)$ of the 'arithmetic image' of the element $i_1(\sigma_3)$ in $(\mathrm{Der}^\Theta \mathbb{L}(a,b))/\mathfrak{u}^{\mathrm{geom}}$. The corresponding derivation was written down in [37]. Computing the stuffle equation (4.22) gives

$$z_3(x_1, x_2, x_3) + z_3(x_2, x_1, x_3) + z_3(x_2, x_3, x_1) = 4$$

which is non-zero, and proves, by proposition 4.6, that z_3 is non-geometric, i.e., not an element of $\mathfrak{u}^{\mathrm{geom}}$.

5.4. Drinfeld elements in depth three.

Theorem 5.9. *The elements σ_{2n+1}^c are in the image of the map*

$$i_0 \quad : \quad \mathfrak{g}^{\mathfrak{m}}/D^4\mathfrak{g}^{\mathfrak{m}} \longrightarrow \mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)/D^4\mathbb{L}(x_0,x_1) \ .$$

Proof. The elements σ_{2n+1}^c satisfy the double shuffle equations so lie in $D^1/D^4\mathfrak{dmr}_0$. The theorem follows immediately from the fact that

$$i: \quad D^1/D^4\mathfrak{g}^{\mathfrak{m}} \longrightarrow D^1/D^4\mathfrak{dmr}_0$$

is an isomorphism. This is equivalent to the statement that i induces an isomorphism on each depth-graded piece

$$i: \quad \mathfrak{d}^d \cong \mathfrak{ls}^d \qquad \text{for } d \leq 3 \ .$$

This is trivial for $d = 1$, and follows from a computation of the dimensions of $\mathfrak{ls}_n^d$ obtained by Zagier [40] for $d = 2$, and by Goncharov ([24], theorem 1.5) for $d = 3$. $\square$

Remark 5.10. It follows from the depth-parity theorem that the elements σ_{2n+1}^c are uniquely determined in depth 4 also (but not in depth 5). A closed formula for these elements can be deduced from remark 5.3.

6. DRINFELD ELEMENTS IN DEPTH 3 VIA GEOMETRIC DERIVATIONS

Recall the notations from §3.

Theorem 6.1. *For all $n \geq 2$, we have an explicit expansion*

$$i_1(\underline{\sigma}_{2n+1}^c) \equiv \varepsilon_{2n+2}^\vee + \sum_{a+b=n} \frac{1}{2b}\{\varepsilon_{2a+2}^\vee, \{\varepsilon_{2b+2}^\vee, \varepsilon_0^\vee\}\} \pmod{B^4}$$

For $n = 1$, we have $i_1(\sigma_3^c) = \varepsilon_4^\vee + z_3 \pmod{B^4}$, where z_3 is defined below.

In terms of the standard normalisations, the previous equation is equivalent to

$$i_1(\sigma_{2n+1}^c) \equiv \varepsilon_{2n+2}^\vee + \sum_{a+b=n} \frac{B_{2a}B_{2b}}{B_{2n}}\binom{2n}{2a}\frac{1}{24b}\{\varepsilon_{2a+2}^\vee, \{\varepsilon_{2b+2}^\vee, \varepsilon_0^\vee\}\} \pmod{B^4}$$

One can also write the right-hand side symmetrically using elements

$$\mathrm{lw}_{a,b} = \frac{1}{2b}\{\varepsilon_{2a+2}^\vee, \{\varepsilon_{2b+2}^\vee, \varepsilon_0^\vee\}\} + \frac{1}{2a}\{\varepsilon_{2b+2}^\vee, \{\varepsilon_{2a+2}^\vee, \varepsilon_0^\vee\}\} \ .$$

These are lowest weight vectors for the action of $\mathfrak{sl}_2$, i.e., $\varepsilon_0 \mathrm{lw}_{a,b} = 0$, where ε_0 is the derivation on $\mathbb{L}(a,b)$ such that $\varepsilon_0(a) = 0$ and $\varepsilon(b) = a$. This yields a direct comparison with the period polynomials of Eisenstein series.

The strategy for the proof is as follows. The elements $\varepsilon_{2n+2}^\vee$ do not preserve the image of $\mathrm{Der}^1 \mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)^\wedge$ under (3.6) and do not descend to derivations on $\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)^\wedge$. However, if we pass to commutative power series representations via §4.1 and enlarge them by introducing poles, then the elements $\varepsilon_{2n+2}^\vee$, considered modulo B^4, descend to the elements ξ_{2n+1} defined in the previous section. Theorem 6.1 is then equivalent to theorem 5.9 via definition 5.4.

The proof of theorem 6.1 given here is from the 'bottom up': i.e. by lifting the analogous theorem on $\mathbb{P}^1\backslash\{0, 1, \infty\}$. A different way to prove theorem 5.9 from the 'top down' via $\mathcal{M}_{1,1}$, is sketched in the final section of the paper.

6.1. **Hain homomorphism in low depth.** Throughout this paragraph, we apply the method of commutative power series §4.1, §4.5 to both $\mathbb{L}(x_0, x_1)$ and $\mathbb{L}(a, b)$.

We wish to consider the depth r components of the Hain morphism §3.3:

$$\phi^r : \mathrm{gr}_D^\bullet T(x_0, x_1) \longrightarrow \mathrm{gr}_B^{\bullet+r} T(a, b) .$$

Translating into rational functions §4, and using the fact that $\rho : \mathrm{gr}_D^\bullet T(x_0, x_1) \cong P$ is an isomorphism, we obtain a commutative diagram

(6.1)
$$\begin{array}{ccc} \mathrm{gr}_D^\bullet T(x_0, x_1) & \xrightarrow{\phi^r} & \mathrm{gr}_B^{\bullet+r} T(a, b) \\ \downarrow\rho & & \downarrow\ell \\ P & \xrightarrow{\phi_r} & Q . \end{array}$$

The map along the top is denoted by a superscript, the one along the bottom by a subscript. The map ϕ^0 is simply the associated graded of (3.6):

(6.2)
$$\phi^0 = \mathrm{gr}\,\phi : \begin{cases} x_0 \mapsto a \\ x_1 \mapsto [a, b] \end{cases}$$

and via $T(x_0, x_1) = \mathrm{gr}_D T(x_0, x_1)$ and $T(a, b) = \mathrm{gr}_B T(a, b)$, we have $\phi = \sum_{r\geq 0} \phi^r$. The idea of the following discussion is to factorise the map ϕ as a composition

$$\mathbb{L}(x_0, x_1) \xrightarrow{\phi^0} \mathbb{L}(a, [a, b]) \xrightarrow{p} \mathbb{L}(a, b)^\wedge$$

where $p : \mathbb{L}(a, [a, b]) \to \mathbb{L}(a, b)^\wedge$ is a continuous map which satisfies $p([a, b]) = 0$, and to express p, in low degrees, via its rational function representation.

Lemma 6.2. *The map $\phi_0 : P \to Q$ is the inclusion $P \subset Q$.*

Proof. We show that $\phi_0 : \mathbb{Q}[y_0, y_1, \ldots, y_r] \to \mathbb{Q}[y_0, y_1, \ldots, y_r]$ is multiplication by the element ℓ_r of (4.15). To see this, (6.2) is the map $x_0 \to a, x_1 \to b$, followed by the composition of r maps, where the kth map, for $1 \leq k \leq r$, replaces the kth occurrence of the letter b in $a^{i_0}ba^{i_1}b \ldots a^{i_{r-1}}ba^{i_r}$ with $ab-ba$. On commutative power series (4.14), this is multiplication by $y_{k-1} - y_k$. $\square$

We next determine ϕ_r for $r = 1, 2$. In degree $r = 1$, it follows from the definition (3.6) of $\phi(x_0) = a + \frac{1}{2}[a, b] + \ldots$ that it is a composition of ϕ^0 (6.2), whose image consists of words in $a, [a, b]$, followed by the derivation in D^Θ (§4.5) which sends $a \mapsto \frac{1}{2}[a, b]$ and $[a, b]$ to zero. Note that the latter does not extend to an element of $\mathrm{Der}^\Theta \mathbb{L}(a, b)$. It is nonetheless represented, via proposition 4.4, by

$$-s^{(1)} = \frac{1}{2}\ell'([a, b]) = \frac{1}{2(y_0 - y_1)} \in Q' ,$$

whose reduced representation is minus (5.1). Therefore if $f \in P$, we have

(6.3)
$$\phi_1(f) = -s^{(1)} \circledast \phi_0(f) .$$

Similarly, in degree $r = 2$, we have for $f \in P$,

(6.4)
$$\phi_2(f) = \frac{1}{2}s^{(1)} \circledast (s^{(1)} \circledast \phi_0(f)) - s^{(2)} \circledast \phi_0(f) ,$$

where

$$-s^{(2)} = \frac{1}{12}\ell'([b, [b, a]]) = \frac{1}{12}\frac{y_0 - 2y_1 + y_2}{(y_0 - y_1)(y_1 - y_2)(y_0 - y_2)} \in Q' .$$

This holds from the definition of ϕ, since $\phi^2 - \frac{1}{2}\phi^1\phi^1$ is the composition of ϕ^0 followed by the derivation $\mathbb{L}(a, [a, b]) \to \mathbb{L}(a, b)$ which sends $a \mapsto \frac{1}{12}[b, [b, a]]$ and $[a, b]$ to zero. By proposition 4.4 the latter corresponds to the action of $s^{(2)}$.

Remark 6.3. Because elements of Q' act trivially on $\ell([a, b]) = -s^{(1)} \circledast y_0$, we have

$$(6.5) \qquad\qquad s^{(1)} \circledast (f \circledast y_0) = \{s^{(1)}, f\} \circledast y_0 \quad , \quad \text{for all } f \in Q' \ .$$

This can also be read off corollary 4.3 upon writing $\phi^1 = \frac{1}{2}\phi^0 \circ \partial_{x_1}$.

6.2. **Proof of theorem 6.1.** Recall that an element $\sigma \in \mathrm{Der}^1\mathbb{L}(x_0, x_1)$ lifts to $\widetilde{\sigma} \in \mathrm{Der}^1\mathbb{L}(a, b)$ if and only if the following equation holds in $\mathbb{L}(a, b)$:

$$\widetilde{\sigma}\phi(x_0) = \phi\sigma(x_0)$$

Finding an element $\sigma \in \mathrm{Der}^1\mathbb{L}(x_0, x_1)$ whose lift is $\varepsilon_{2n+2}^\vee$, is equivalent via (3.6), modulo terms of B-degree ≥ 4, to the following equation

$$\varepsilon_{2n+2}^\vee\big(a + \frac{1}{2}[a, b] + \frac{1}{12}[b, [b, a]]\big) \equiv \phi(\sigma(x_0)) \pmod{B^4} \ .$$

It is easy to verify that it has no solution $\sigma \in \mathrm{Der}^1\mathbb{L}(x_0, x_1)$. Note that $\varepsilon_{2n+2}^\vee([a, b]) = 0$ so the middle term on the left-hand side can be dropped. We can pass to rational function representations via propositions 4.2 and 4.4, and view the previous equation in Q. Since $\ell'(\varepsilon_{2n+2}^\vee) = (y_1 - y_0)^{2n}$, and $\rho(a) = y_0$, $\rho(x_0) = y_0$, it is equivalent to

$$(6.6) \qquad (y_1 - y_0)^{2n} \circledast \big((1 - s^{(2)}) \circledast y_0\big) \equiv \phi(\rho'(\sigma) \odot y_0) \pmod{B^4}$$

It has no solutions $\rho'(\sigma) \in P'$. Now observe that

$$\phi_0(\rho'(\sigma) \odot y_0) = \phi_0(\rho'(\sigma)) \circledast y_0$$

since the formulae for $\odot$ and $\circledast$ (propositions 4.2 and 4.4) are formally identical and ϕ_0 is the identity. Let us write χ instead of $\phi_0(\rho'(\sigma))$ and try to solve (6.6) for $\chi \in Q'$. By (6.3), (6.4), the right-hand side of (6.6) is equal, after expanding $\phi \equiv \phi^0 + \phi^1 + \phi^2$ mod B^4 and applying (6.5), to

$$\chi \circledast y_0 - \{s^{(1)}, \chi\} \circledast y_0 + \frac{1}{2}\{s^{(1)}, \{s^{(1)}, \chi\}\} \circledast y_0 - s^{(2)} \circledast (\chi \circledast y_0) \pmod{B^4} \ .$$

Using the fact (proposition 4.4) that the Lie bracket $\{\ ,\ \}$ is the antisymmetrization of $\circledast$, and that the action of Q' on $y_0 \in Q$ is faithful, we deduce that the components of equation (6.6) in degrees $1, 2, 3$ are the equations:

$$
\begin{aligned}
(y_1 - y_0)^{2n} &= \chi^{(1)} \\
0 &= \chi^{(2)} - \{s^{(1)}, \chi^{(1)}\} \\
-(y_1 - y_0)^{2n} \circledast s^{(2)} &= \chi^{(3)} - \{s^{(1)}, \chi^{(2)}\} + \frac{1}{2}\{s^{(1)}, \{s^{(1)}, \chi^{(1)}\}\} - s^{(2)} \circledast \chi^{(1)} \ .
\end{aligned}
$$

These three equations are equivalent to the definition of the elements χ_{2n+1} after passing to reduced versions $(y_0, y_1, y_2) \mapsto (0, x_1, x_2)$ and using definition (4.18).

7. Explicit rational associator in depths ≤ 3

In section §5 we wrote down explicit solutions to the double shuffle equations modulo products, in odd weights and depths ≤ 3. The goal of this paragraph is to discuss solutions to the full double shuffle equations with even weights in the same range.

7.1. **Double shuffle equations.** The full double shuffle equations in depth two are given by the pair of equations:

$$f^{(2)}(x_1, x_1 + x_2) + f^{(2)}(x_2, x_1 + x_2) \quad = \quad f^{(1)}(x_1)f^{(1)}(x_2)$$

$$(7.1) \qquad f_\star^{(2)}(x_1, x_2) + f_\star^{(2)}(x_2, x_1) \quad = \quad \frac{f_\star^{(1)}(x_1) - f_\star^{(1)}(x_2)}{x_2 - x_1} + f_\star^{(1)}(x_1)f_\star^{(1)}(x_2)$$

where $f^{(1)}, f_\star^{(1)} \in \mathbb{Q}[[x_1]]$ and $f^{(2)}, f_\star^{(2)} \in \mathbb{Q}[[x_1, x_2]]$ are formal power series in commuting variables. A power series f without a subscript will denote its shuffle-regularised version; a subscript $\star$ will denote its stuffle-regularised version. They differ by a factor which is well-understood [38]. Our normalisations will be such that

$$f_\star^{(1)} = f^{(1)} \quad \text{and} \quad f_\star^{(2)} = f^{(2)} + \frac{1}{48}$$

One can easily convince oneself that the second equation of (7.1) is the direct translation of the stuffle product formula

$$\zeta(m, n) + \zeta(n, m) + \zeta(m + n) = \zeta(m)\zeta(n) .$$

Note that, in contrast to the double shuffle equations modulo products, the right-hand term in the previous equation means that we must consider all weights simultaneously. The shuffle equation in depth three takes the form

$$f^{(3)}(x_1, x_{12}, x_{123}) + f^{(3)}(x_2, x_{12}, x_{123}) + f^{(3)}(x_2, x_{23}, x_{123}) = f^{(1)}(x_1)f^{(2)}(x_2, x_{23})$$

and the stuffle equation takes the form

$$f_\star^{(3)}(x_1, x_2, x_3) + f_\star^{(3)}(x_2, x_1, x_3) + f_\star^{(3)}(x_2, x_3, x_1) \quad =$$

$$\frac{f_\star^{(2)}(x_2, x_1) - f_\star^{(2)}(x_2, x_3)}{x_3 - x_1} + \frac{f_\star^{(2)}(x_1, x_3) - f_\star^{(2)}(x_2, x_3)}{x_2 - x_1} + f_\star^{(1)}(x_1)f_\star^{(2)}(x_2, x_3) .$$

where in this case the comparison between the two regularisations is given by

$$f^{(3)}(x_1, x_2, x_3) = f_\star^{(3)}(x_1, x_2, x_3) + \frac{1}{96}\left(b(x_1) - \frac{1}{x_1}\right)$$

The general principle [2] of constructing solutions to these equations with poles and correcting with counter terms also holds in this situation. The full double shuffle equations are inhomogeneous in two different ways: there are several linear terms of lower depths and a single term consisting of products of elements of lower depth. The strategy is to construct solutions γ to the equations in which lower depth terms are omitted, but with all product terms retained, and to use the element (5.1) to convert these solutions into polar solutions to the full equations. The polar parts are then subtracted using counterterms involving the elements ξ_{2n+1} constructed before.

7.2. **Polar solutions.** Recall that $b_1(x) = b(x)$ (5.5) is a generating series for Bernoulli numbers whose Laurent series is $x^{-1} + O(x)$. Thinking of $b_1(x)$ as a deformation of the rational function x^{-1}, leads us to introduce, following (5.1), the function

$$b_2(x_1, x_2) = \frac{1}{3}\left(b_1(x_1)b_1(x_2) + b_1(x_2)b_1(x_1 - x_2)\right) .$$

With these definitions, set

$$2\gamma^{(1)} \quad = \quad -b_1$$
$$4\gamma^{(2)} \quad = \quad -b_2 + \tfrac{1}{2}b_1 \underset{\circ}{\circ} b_1$$
$$8\gamma^{(3)} \quad = \quad -b_2 \underset{\circ}{\circ} b_1 + \tfrac{1}{6}b_1 \underset{\circ}{\circ} (b_1 \underset{\circ}{\circ} b_1)$$

The element $\gamma^{(2)}$, for example, solves the semi-homogeneous equations

$$(7.2) \qquad \begin{aligned} \gamma^{(2)}(x_1, x_1 + x_2) + \gamma^{(2)}(x_2, x_1 + x_2) &= \gamma^{(1)}(x_1)\gamma^{(1)}(x_2) \\ \gamma_*^{(2)}(x_1, x_2) + \gamma_*^{(2)}(x_2, x_1) &= \gamma^{(1)}(x_1)\gamma^{(1)}(x_2) \end{aligned}$$

where $\gamma_*^{(2)} = \gamma^{(2)} + \frac{1}{48}$, and $\gamma^{(3)}$ satisfies the equations

$$\begin{aligned} \gamma^{(3)}(x_1, x_{12}, x_{123}) + \gamma^{(3)}(x_2, x_{12}, x_{123}) + \gamma^{(3)}(x_2, x_{23}, x_{123}) &= \gamma^{(1)}(x_1)\gamma^{(2)}(x_2, x_{23}) \\ \gamma_*^{(3)}(x_1, x_2, x_3) + \gamma_*^{(3)}(x_2, x_1, x_3) + \gamma_*^{(3)}(x_2, x_3, x_1) &= \gamma^{(1)}(x_1)\gamma_*^{(2)}(x_2, x_3) \end{aligned}$$

where $\gamma_*^{(3)}(x_1, x_2, x_3) = \gamma^{(3)}(x_1, x_2, x_3) + \frac{1}{48}\gamma^{(1)}(x_1)$. This follows easily from (5.6).

New series Θ are now defined by twisting on the left by the elements (5.1):

$$\begin{aligned} \Theta^{(1)} &= \gamma^{(1)} \\ \Theta^{(2)} &= \gamma^{(2)} + s^{(1)} \underline{\circ} \gamma^{(1)} \\ \Theta^{(3)} &= \gamma^{(3)} + s^{(1)} \underline{\circ} \gamma^{(2)} + s^{(2)} \underline{\circ} \gamma^{(1)} + \tfrac{1}{2} s^{(1)} \underline{\circ} (s^{(1)} \underline{\circ} \gamma^{(1)}) \end{aligned}$$

They have poles in x_i. More precisely, the element $d_r \Theta^{(r)}$ can be viewed as a formal power series in $\mathbb{Q}[[x_1, \ldots, x_r]]$, where $d_r = x_1 \ldots x_r \prod_{i<j}(x_i - x_j)$, for $1 \le r \le 3$.

For $1 \le r \le 3$, we can write

$$\Theta^{(r)} = p_r + \Phi^{(r)} \ ,$$

where p_r is a homogeneous rational function in $x_1, \ldots, x_r$ of degree $-r$, and $\Phi^{(r)}$ is a power series in homogeneous rational functions of degrees $> 1 - r$. With these definitions, one verifies that the truncated elements $\Phi^{(r)}$ are polar solutions to the full double shuffle equations §7.1 with

$$\Phi_*^{(2)} = \Phi^{(2)} + \frac{1}{48} \qquad \text{and} \qquad \Phi_*^{(3)} = \Phi^{(3)} + \frac{1}{48}\Phi^{(1)}(x_1) \ .$$

It remains to remove the poles from the $\Phi^{(r)}$ to obtain bona fide polynomial solutions to the double shuffle equations with no polar terms.

7.3. **Subtraction of counterterms.** Let us define a formal power series by

$$(7.3) \qquad C = \sum_{n \ge 1} \frac{1}{2n} \{\underline{\xi}_{-1}, \underline{\xi}_{2n+1}\}$$

where the elements $\underline{\xi}_{2n+1}$ were defined in §5.1. Its definition was only given in depths $1, 2, 3$. Using this element to provide counter terms, we can finally write down a canonical element τ in depths $1, 2, 3$ as follows:

$$\begin{aligned} \tau^{(1)} &= \Phi^{(1)} \\ \tau^{(2)} &= \Phi^{(2)} + C^{(2)} \\ \tau^{(3)} &= \Phi^{(3)} + C^{(2)} \underline{\circ} \Phi^{(1)} + C^{(3)} \end{aligned}$$

A straightforward residue computation along the lines of §5.3 suffices to show that the elements $\tau^{(i)}$, where $i = 1, 2, 3$ have no poles, and therefore lie in $\mathbb{Q}[[x_1, \ldots, x_i]]$. We omit the details. Note that by the depth-parity theorem, the element $\tau^{(3)}$ is uniquely determined from $\tau^{(2)}$. By a version of §4.1, the coefficients of $\tau^{(i)}$ correspond to words in $\mathsf{x}_0, \mathsf{x}_1$, and taking the limit defines a unique element

$$\tau \in \mathbb{Q}\langle\langle \mathsf{x}_0, \mathsf{x}_1 \rangle\rangle / D^4 \mathbb{Q}\langle\langle e_0, e_1 \rangle\rangle \ .$$

Theorem 7.1. *The element τ is an explicit (shuffle-regularized) solution to the full double shuffle equations in depths ≤ 3.*

A similar construction holds in depth four, but there is *a priori* no canonical way to cancel the poles: one must subtract counter-terms consisting of quadruple brackets in the ξ_{2n+1}'s, which involves some choices because of quadratic relations amongst them (see §8.2). It is an interesting question to ask if the element τ defined above can in fact be extended to an explicit associator in higher depths.

Since the solutions to the full double shuffle equations is a torsor under the left action of the prounipotent algebraic group DMR_0 whose Lie algebra is $\mathfrak{dmr}_0$, we can twist the elements $\tau^{(i)}$ on the left with our canonical elements $\exp_\circ \sigma^c_{2n+1}$ to obtain all other rational solutions to the double shuffle equations in depths ≤ 3.

Corollary 7.2. *Every rational solution s to the full double shuffle equations in depths ≤ 3 can be written explicitly in the form*

$$s \equiv \exp_\circ(g) \circ \tau \pmod{D^4}$$

where $g \in (\mathfrak{g}^\mathfrak{m}/D^4\mathfrak{g}^\mathfrak{m})^\wedge$ is an (infinite) linear combination of commutators in the canonical elements σ^c_{2n+1} of length ≤ 3.

Note that the element g in the corollary is not unique because of quadratic relations §8.2 in $\mathfrak{g}^\mathfrak{m}/D^4\mathfrak{g}^\mathfrak{m}$.

7.4. **Remarks.** The elements $\tau^{(i)}$ for $i \leq 3$ define a homomorphism from motivic multiple zeta values in depths ≤ 3 and even weight to rational numbers, given by

(7.4) $$\tau^{(r)}\zeta^\mathfrak{m}(n_1, \ldots, n_r) = \text{ coeff. of } x_1^{n_1-1} \ldots x_r^{n_r-1} \text{ in } \tau^{(r)} .$$

They respect all the relations between motivic multiple zeta values and satisfy

$$\tau^{(1)}\zeta^\mathfrak{m}(2n) = \zeta(2n)/(2\pi i)^{2n} \quad \in \mathbb{Q} .$$

Likewise, the canonical elements $\sigma^c_{2n+1} \in \mathfrak{g}^\mathfrak{m}/D^4\mathfrak{g}^\mathfrak{m}$ define a map from motivic multiple zeta values in depth ≤ 4 and odd weight to rational numbers given by

(7.5) $$\sigma^{(r)}_{2n+1}\zeta^\mathfrak{m}(n_1, \ldots, n_r) = \text{ coeff. of } x_1^{n_1-1} \ldots x_r^{n_r-1} \text{ in } \sigma^{(r)}$$

where $2n+1 = n_1 + \ldots + n_r$. The maps (7.5) annihilate products, respect all relations between motivic multiple zeta values (modulo products) and satisfy

$$\sigma^{(1)}_{2n+1}\zeta^\mathfrak{m}(2n+1) = 1 .$$

In [5], a method was described to decompose any motivic multiple zeta value (and hence, by taking the period, any actual multiple zeta value) into a chosen basis of motivic multiple zeta values using the motivic coaction. The method is not an algorithm because it requires a transcendental computation at each step using the period map. However, the maps (7.4) and (7.5) can be used as a substitute for the period map. Thus we obtain as a corollary an exact algorithm to decompose any multiple zeta value of depth ≤ 3 (and depth ≤ 4 in the case of odd weight) into a chosen basis of multiple zeta values of the same or smaller depth.

8. Cuspidal elements and the Broadhurst-Kreimer conjecture

We can recast the version of the Broadhurst-Kreimer conjecture stated in [6] using the σ^c_{2n+1}, first in $\mathrm{gr}_D \mathrm{Der}^1 \mathbb{L}(\mathsf{x}_0, x_1)$ and then in the elliptic setting in $\mathrm{gr}_B \mathrm{Der}^\Theta \mathbb{L}(a, b)$.

We seek a conjectural presentation for $\mathfrak{d}$. The first set of obvious generators are the images of the Drinfeld elements $\sigma_{2n+1} \in D^1\mathfrak{g}^\mathfrak{m}$ in the associated graded $\mathfrak{d}^\bullet = \mathrm{gr}^\bullet_D\mathfrak{g}^\mathfrak{m}$:

(8.1) $$\overline{\sigma}_{2n+1} \in \mathfrak{d}^1_{2n+1} \qquad \text{for all } n \geq 1 .$$

They are well-defined (independent of the choice of σ_{2n+1}). They satisfy quadratic relations which can be described in terms of period polynomials.

8.1. **Reminders on period polynomials.** Let $n \geq 0$ and let $V_n = \bigoplus_{i+j=n} \mathbb{Q} x_1^i x_2^j$ denote the vector space of homogeneous polynomials of degree n. It is equipped with the right action of $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ given by the formula

$$P(x_1, x_2)|_\gamma = P(ax_1 + bx_2, cx_1 + dx_2) \quad \text{if} \quad \gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma \, , P \in V_n \ .$$

Let $V_n' \subset V_n$ denote the subspace of polynomials which vanish at $x_1 = 0$ and $x_2 = 0$. It is naturally isomorphic to the vector space quotient $V_n/(\mathbb{Q}x_1^n \oplus \mathbb{Q}x_2^n)$.

Definition 8.1. Let $n \geq 1$ and let $S_{2n} \subset V_{2n}'$ denote the vector space of homogeneous polynomials $P(x_1, x_2)$ of degree $2n$ satisfying $P(x_1, 0) = P(0, x_2) = 0$ and

$$P(x_1, x_2) + P(x_2, x_1) = 0 \quad , \quad P(x_1, x_2) + P(x_1 - x_2, x_1) + P(-x_2, x_1 - x_2) = 0 \ .$$

The subspace $S_{2n}^+ \subset S_{2n}$ consisting of polynomials which are of even degree in both x_1 and x_2 is called the space of even (cuspidal) period polynomials.

Remark 8.2. Denote the standard elements $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ in Γ. Consider the following linear map from right Γ group cochains ([7], §2.3) to polynomials

$$(8.2) \qquad\qquad f \mapsto \pi(f(S)) : Z_{\mathrm{cusp}}^1(\Gamma; V_{2n}) \longrightarrow V_{2n}'$$

where $Z_{\mathrm{cusp}}^1(\Gamma; V_n) \subset Z^1(\Gamma; V_n)$ is the subgroup of cochains f such that $f(T) = 0$, and $\pi : V_n \to V_n'$ is the projection. It is well-known that this induces an isomorphism

$$H_{\mathrm{cusp}}^1(\Gamma; V_{2n}) \xrightarrow{\sim} S_{2n}$$

where $H_{\mathrm{cusp}}^1(\Gamma; V_{2n}) = \ker(H^1(\Gamma; V_{2n}) \to H^1(\Gamma_\infty; V_{2n}))$, and $\Gamma_\infty \leq \Gamma$ is the subgroup generated by $-1, T$. This in turn induces an isomorphism

$$(8.3) \qquad\qquad H_{\mathrm{cusp}}^1(\Gamma; V_{2n})^+ \xrightarrow{\sim} S_{2n}^+$$

where the $+$ on the left-hand factor denotes invariants with respect to the action of the real Frobenius involution ([7] §5.4, §7.4). The Eichler-Shimura theorem states in particular that the integration maps gives an isomorphism:

$$S_{2n}(\Gamma) \xrightarrow{\sim} H_{\mathrm{cusp}}^1(\Gamma; V_{2n+2})^+ \otimes \mathbb{R}$$

where S_{2n} denotes the space of cuspidal modular forms of weight $2n$.

8.2. **Quadratic relations.** Define a vector space K

$$K = \ker(\{\,,\} : \mathfrak{d}^1 \wedge \mathfrak{d}^1 \to \mathfrak{d}^2)$$

to be the kernel of the Ihara bracket. It is weight-graded in even degrees $K = \bigoplus_n K_{2n}$. Since $\mathfrak{g}^{\mathfrak{m}} \subset D^1\mathfrak{g}^{\mathfrak{m}}$ is generated in depth 1, $D^1 \wedge D^1 \to D^2$ is surjective, and hence $\mathfrak{d}^1 \wedge \mathfrak{d}^1 \to \mathfrak{d}^2$ is surjective, i.e.,

$$0 \longrightarrow K \longrightarrow \mathfrak{d}^1 \wedge \mathfrak{d}^1 \xrightarrow{\{\,,\}} \mathfrak{d}^2 \longrightarrow 0$$

is an exact sequence. Now embed $\mathfrak{g}^{\mathfrak{m}}$ in $\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ via (2.4), and therefore $\mathfrak{d} = \mathrm{gr}_D \mathfrak{g}^{\mathfrak{m}}$ is also embedded in $\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ since the latter is graded for the D-degree. By passing to reduced polynomial representations §4.1, we have a canonical isomorphism

$$\begin{aligned} \mathfrak{d}^1 &= x_1^2 \mathbb{Q}[x_1^2] \\ \overline{\sigma}_{2n+1} &\mapsto x_1^{2n} \qquad\qquad \text{for } n \geq 1. \end{aligned}$$

We can thus identify $\mathfrak{d}^1 \otimes \mathfrak{d}^1 = x_1^2 \mathbb{Q}[x_1^2] \otimes x_1^2 \mathbb{Q}[x_1^2] \cong x_1^2 x_2^2 \mathbb{Q}[x_1^2, x_2^2]$, and hence view elements of $\mathfrak{d}^1 \wedge \mathfrak{d}^1 \subset \mathfrak{d}^1 \otimes \mathfrak{d}^1$ as antisymmetric polynomials in x_1^2, x_2^2.

Lemma 8.3. *The polynomial representation gives an isomorphism*

$$(8.4) \qquad\qquad\qquad K_{2n} \xrightarrow{\sim} S_{2n}^+ \ .$$

Proof. This is immediate from the formula for $\underline{\circ}$ given in §2.3 (example 4.1)

$$\{x_1^{2a}, x_1^{2b}\} = P(x_1, x_2) + P(x_2 - x_1, x_1) + P(-x_2, x_1 - x_2)$$

where $P(x_1, x_2) = x_1^{2a}x_2^{2b} - x_2^{2a}x_1^{2b}$ and $a, b \in \mathbb{N}$. $\qquad\qquad\square$

These quadratic relations appear in several contexts:

Corollary 8.4. *The elements $\varepsilon_{2n+2}^{\vee}$, ξ_{2n+1}, $\overline{\sigma}_{2n+1}$ and x_1^{2n} all satisfy the same quadratic relations in K_{2n}.*

Proof. The polynomial representations of $\varepsilon_{2n+2}^{\vee}$ and $\overline{\sigma}_{2n+1}$ via propositions 4.2 and 4.4 are both x_1^{2n}, and the Lie brackets correspond to the Ihara bracket $\{,\}$. Therefore they satisfy the identical quadratic relations. For the elements ξ_{2n+1}, this follows from their definition, because they are obtained from the x_1^{2n} via the Ihara bracket, or from the computations of §6 relating them to the $\varepsilon_{2n+2}^{\vee}$. $\qquad\square$

The existence of such quadratic relations was first observed by Ihara-Takao and has been reproved in many ways since. The smallest example of a period polynomial is the element $x_1^2 x_2^2 (x_1^2 - x_2^2)^3 = x_1^8 x_2^2 - 3x_1^6 x_2^4 + 3x_1^4 x_2^6 - x_1^2 x_2^8$. It corresponds to the relations

$$\{x_1^2, x_1^8\} - 3\{x_1^4, x_1^6\} = 0 \qquad , \qquad \{\overline{\sigma}_3, \overline{\sigma}_9\} - 3\{\overline{\sigma}_5, \overline{\sigma}_7\} = 0 \ .$$
$$\{\varepsilon_4^{\vee}, \varepsilon_{10}^{\vee}\} - 3\{\varepsilon_6^{\vee}, \varepsilon_8^{\vee}\} = 0 \qquad , \qquad \{\xi_3, \xi_9\} - 3\{\xi_5, \xi_7\} = 0 \ .$$

8.3. **Cuspidal generators in depth 4.** As explained in [6], the depth filtration on $\mathfrak{g}^{\mathfrak{m}}$ gives rise to a spectral sequence and in particular a differential

$$d : H_2(\mathfrak{d}) \longrightarrow H_1(\mathfrak{d}) \ .$$

Since $H_2(\mathfrak{d}) = \ker(\wedge^2 \mathfrak{d} \to \mathfrak{d})/\wedge^3 \mathfrak{d}$, there is a natural map $K \to H_2(\mathfrak{d})$. It is in fact injective since the image of $\wedge^3 \mathfrak{d}$ is in depth ≥ 3. Composing with this map gives a linear map $d : K \to (\mathfrak{d}^4)^{ab}$ as we explain presently, and the canonical Drinfeld elements defined in §5 give a means to compute it explicitly. To see this, the elements σ_{2n+1}^c can be interpreted as a linear map

$$\sigma^c : \mathfrak{d}^1 \quad \longrightarrow \quad D^1\mathfrak{g}^{\mathfrak{m}}/D^4\mathfrak{g}^{\mathfrak{m}}$$
$$\overline{\sigma}_{2n+1} \quad \mapsto \quad \sigma_{2n+1}^c$$

which splits the natural map $D^1/D^4\mathfrak{g}^{\mathfrak{m}} \to D^1/D^2\mathfrak{g}^{\mathfrak{m}} = \mathfrak{d}^1$. Consider

(8.5) $$\mathfrak{d}^1 \wedge \mathfrak{d}^1 \xrightarrow{\sigma^c \wedge \sigma^c} D^1/D^4\mathfrak{g}^{\mathfrak{m}} \wedge D^1/D^4\mathfrak{g}^{\mathfrak{m}} \xrightarrow{\{\,,\,\}} D^2/D^5\mathfrak{g}^{\mathfrak{m}} \ .$$

The subspace K maps into $D^3/D^5\mathfrak{g}^{\mathfrak{m}}$, since its image in $D^2/D^3 = \mathfrak{d}^2$ is zero. Since K has even weights, the depth-parity theorem 2.6 implies that $D^3/D^4\mathfrak{g}^{\mathfrak{m}} = \mathfrak{d}^3$ vanishes in even weights, and the restriction of (8.5) to K gives a linear map

(8.6) $$\mathfrak{c} : K \longrightarrow D^4/D^5\mathfrak{g}^{\mathfrak{m}} = \mathfrak{d}^4 \ .$$

The letter $\mathfrak{c}$ was chosen to stand for 'cuspidal', for the following reason. Its weight-graded components by $\mathfrak{c}_{2n}$ can be viewed, via (8.4), as linear maps

$$\mathfrak{c}_{2n} : H_{\text{cusp}}^1(\Gamma, V_{2n})^+ \longrightarrow \mathfrak{d}_{2n}^4 \ .$$

Theorem 8.5. *Let $P(x_1, x_2) = \sum_{i,j} \lambda_{i,j} x_1^{2i} x_2^{2j}$ be in K, where $\lambda_{i,j} = -\lambda_{j,i}$. It gives rise to a relation of the form*

$$\sum_{i<j} \lambda_{i,j}\{\overline{\sigma}_{2i+1}, \overline{\sigma}_{2j+1}\} = 0 \quad in \quad \mathfrak{d}^2 \ .$$

Then the image of the element $\mathfrak{c}(P) \in \mathfrak{d}^4$ *in* $\mathbb{Q}[x_1, x_2, x_3, x_4]$ *is*

$$(8.7) \quad \rho^{(4)}(\mathfrak{c}(P)) = \sum_{i,a,b} \lambda_{i,a+b} \frac{B_{2a}B_{2b}}{B_{2a+2b}} \binom{2a+2b}{2a} \frac{1}{24b} \{x_1^{2i}, \{x_1^{2a}, \{x_1^{2b}, x_1^{-2}\}\}\}$$
$$- 3 \sum_i \lambda_{i,2} \{x_1^{2i}, z_3\}$$

where z_3 *was defined in remark* 5.8.

Proof. The element $\mathfrak{c}(P)$ is by definition

$$\mathfrak{c}(P) = \sum_{i<j} \lambda_{i,j} \{\sigma_{2i+1}^c, \sigma_{2j+1}^c\} \pmod{D^5}$$

Now substitute the expressions (5.8) for σ_{2j+1}^c in terms of the polar elements ξ_{2a+1} (work in $\mathbb{Q}(x_1, \ldots, x_4)$). By corollary 8.4, the ξ_{2a+1} satisfy the relations

$$\left(\sum_{i<j} \lambda_{i,j} \{\xi_{2i+1}, \xi_{2j+1}\} \right)^{(r)} = 0$$

for $1 \le r \le 3$, where a superscript (r) denotes the depth r component. The theorem follows from formula (5.8), together with the definition of the element $z_3 = -\frac{1}{3}\xi_3^{(3)}$. $\quad\square$

If one believes the Broadhurst-Kreimer conjecture, one is led to the following

Conjecture 1. (Broadhurst-Kreimer: compare with [6], §9)

$$(8.8) \qquad\qquad H_1(\mathfrak{d}; \mathbb{Q}) \cong \bigoplus_{n \ge 1} \overline{\sigma}_{2n+1} \mathbb{Q} \oplus \mathfrak{c}(K)$$
$$H_2(\mathfrak{d}; \mathbb{Q}) \cong K$$
$$H_i(\mathfrak{d}; \mathbb{Q}) = 0 \quad \text{for all} \quad i \ge 3 .$$

Thus $\mathfrak{d}$ admits the following conjectural presentation. It has generators the $\overline{\sigma}_{2n+1}$ in depth 1 for $n \ge 1$ together with cuspidal elements $\mathfrak{c}(K)$ in depth 4. The only relations are the quadratic relations of §8.2.

Remark 8.6. As noted in [20], $H_3(\mathfrak{d}; \mathbb{Q}) = 0$ implies that $H_i(\mathfrak{d}; \mathbb{Q}) = 0$ for all $i \ge 3$. In fact, for any pro-nilpotent Lie algebra $\mathfrak{g}$ over a field k of characteristic zero, $H_i(\mathfrak{g}, k) = 0$ implies that $H_n(\mathfrak{g}, k) = 0$ for all $n \ge i$. To see this, note that since $\mathfrak{g}$ is a projective limit of finite-dimensional nilpotent Lie algebras, and (co)homology commutes with limits, we can assume $\mathfrak{g}$ nilpotent and $H^i(\mathfrak{g}, k) = 0$. Every finite-dimensional $\mathfrak{g}$-module M has an increasing filtration by submodules $M_m \subset M$ such that the associated graded is a trivial module. By the long exact cohomology sequence and induction on m, $H^i(\mathfrak{g}; M) = 0$ for all such M. Now interpret $H^n(\mathfrak{g}; M)$ as the Ext group $\text{Ext}^n(k, M)$ in the category of $U\mathfrak{g}$-modules, and use the well-known fact that if $\text{Ext}^i(k, M)$ vanishes for all M then it also vanishes for all $n \ge i$.

The conjecture given in [6] involved certain exceptional generators denoted $\mathbf{e}_f$, for $f \in P$, in the depth 4 component of the larger Lie algebra $\text{gr}_D^4 \mathfrak{dmr}$ of double shuffle equations. It is not known if they are in the image of $\mathfrak{d}^4$. Thus the formulation (8.10) eliminates part of the conjecture given in [6].

8.4. **Remarks on the role of** z_3**.** The element z_3 is the first of a sequence z_{2n+1} of derivations in $\mathrm{Der}^\Theta \mathbb{L}(a,b)$ which are $\mathfrak{sl}_2$-invariant and well-defined modulo $(\mathfrak{u}^{\mathrm{geom}})^{\mathfrak{sl}_2}$. It follows from theorem 10.1 in [7] that their action on the derivations $\varepsilon_{2k+2}^\vee$ are known explicitly modulo Lie brackets involving at least three $\varepsilon_{2n+2}^\vee$, with $n \geq 0$. It is possible that this computation can be extended to the next order, which in particular would give a formula for $\{z_3, x_1^{2n}\}$ for all $n \geq 1$.

Remark 8.7. In [6] we defined an injective linear map

$$\overline{\mathbf{e}} : P \longrightarrow \mathfrak{ls}^4$$

from the space P of even period polynomials to the space of solutions $\mathfrak{ls}^4$ to the linearised double shuffle equations in depth 4. It only depends on the functional equations satisfied by elements of P. It is natural to extend this linear map to the polynomials $x_1^{2n} - x_2^{2n} \in V_{2n}$, which correspond to coboundaries under the morphism (8.2). Since they satisfy the same functional equations as elements of P, they define elements of $\mathfrak{pls}^4$ which have poles. One easily verifies from the definitions that:

$$(8.9) \qquad \overline{\mathbf{e}}(x_1^{2n} - x_2^{2n}) + \{z_3, x_1^{2n-2}\} = 0 \ .$$

This gives a different interpretation of the role of z_3 in formula (8.7).

8.5. **Elliptic interpretation of the Broadhurst-Kreimer conjecture.** We can transpose the previous conjecture into the Lie algebra $\mathrm{Der}^\Theta \mathbb{L}(a,b)$ as follows. Recall that the map $i_1 : \mathfrak{g}^{\mathrm{m}} \to \mathrm{Der}^\Theta \mathbb{L}(a,b)$ (3.2) is injective. Since B cuts out the depth filtration on the image $i_1(\mathfrak{g}^{\mathrm{m}})$ (corollary 3.6) we obtain an injective morphism

$$i_1 : \mathfrak{d} \to \mathrm{gr}_B \mathrm{Der}^\Theta \mathbb{L}(a,b) \ .$$

We wish to describe the conjectural generators in B-degrees 1 and 4. For simplicity, we shall use the heretical normalisations $\underline{\varepsilon}_{2n}$ to simplify the statement. This has the side-effect of rescaling the period polynomial relations.

More precisely, consider linear map

$$\begin{aligned} \mathbb{Q}[x_1^2, x_2^2] &\longrightarrow \ \mathbb{Q}[x_1^2, x_2^2] \\ x_i^{2n} &\longmapsto \ \frac{(2n)!}{B_{2n}} x_i^{2n} \quad \text{for } n \geq 1 \end{aligned}$$

and let $\underline{K}$ denote the image of K. Lemma 8.3 and corollary 8.4 imply that $P = \sum_{i,j} \lambda_{i,j} x_1^{2i} x_2^{2j} \in \underline{K}$ where $\lambda_{i,j} + \lambda_{j,i} = 0$, if and only if

$$\sum_{i<j} \lambda_{i,j} \{\varepsilon_{2i+2}^\vee, \varepsilon_{2j+2}^\vee\} = 0 \ .$$

Define, for all $P \in \underline{K}$, elements

$$\mathfrak{c}(P) = \sum_{i<j} \lambda_{i,j} \{\underline{\sigma}_{2i+1}^c, \underline{\sigma}_{2j+1}^c\} \in \mathrm{gr}_B^4 \mathrm{Der}^\Theta \mathbb{L}(a,b) \ ,$$

and let $z_3 \in \mathrm{gr}_B^3 \mathrm{Der}^\Theta \mathbb{L}(a,b)$ denote the derivation such that $\ell'(z_3)$ is the element (5.8).

Theorem 8.8. *For any $P \in \underline{K}$,*

$$\mathfrak{c}(P) = \sum_{j} \lambda_{2,j} \{z_3, \underline{\varepsilon}_{2j+2}^\vee\} + \sum_{i,a,b} \lambda_{i,a+b} \frac{1}{2b} \{\underline{\varepsilon}_{2i+2}^\vee, \{\underline{\varepsilon}_{2a+2}^\vee, \{\underline{\varepsilon}_{2b+2}^\vee, \underline{\varepsilon}_0^\vee\}\}\} \ .$$

The Broadhurst-Kreimer conjecture suggests the following:

Conjecture 2. (Elliptic (geometric) Broadhurst-Kreimer conjecture)

$$(8.10) \qquad \begin{aligned} H_1(\mathfrak{o}, \mathbb{Q}) &\cong \bigoplus_{n \geq 1} \underline{\varepsilon}_{2n+2}^{\vee} \mathbb{Q} \oplus \mathfrak{c}(\underline{K}) \\ H_2(\mathfrak{o}, \mathbb{Q}) &\cong \underline{K} \\ H_i(\mathfrak{o}, \mathbb{Q}) &= 0 \quad \text{for all} \quad i \geq 3 \;. \end{aligned}$$

Note that $\underline{K}$ in the above is interpreted as the space of quadratic relations between the $\underline{\varepsilon}_{2n+2}^{\vee}$, for $n \geq 1$. This conjecture is equivalent to conjecture (1) by §6.

8.6. **Some related problems.**

(1) Show that the map $\mathfrak{c}(P) : K \to (\mathfrak{o}^4)^{\mathrm{ab}}$ is injective.
(2) Relate the elements $\mathfrak{c}(P)$ to the exceptional elements $\mathbf{e}_f$ defined in [6].
(3) Construct a basis of the space of motivic periods of $\mathcal{MT}(\mathbb{Z})$ of motivic depth 2 and even weight out of motivic multiple zeta values of depth ≤ 4, using the formula for the σ_{2n+1}^c modulo depth 4.

9. Some motivation from the relative completion of $\mathrm{SL}_2(\mathbb{Z})$

I shall very briefly sketch how I arrived at formula (1.9) and (1.7) by considering double integrals of Eisenstein series. This explains why the coefficients for the explicit formula for the σ_{2n+1}^c involve the odd period polynomials of Eisenstein series.

9.1. Denote the Hecke-normalised Eisenstein series of weight $2k \geq 4$ by

$$E_{2k}(q) = -\frac{B_{2k}}{4k} + \sum_{n \geq 0} \sigma_{2k-1}(n) q^n \;,$$

where $\sigma_k(n)$ denotes the divisor function. For any modular form $f(\tau)$ of weight $2k \geq 4$ for $\mathrm{SL}_2(\mathbb{Z})$ we shall write (see [7] for further details):

$$(9.1) \qquad \underline{f}(\tau) = (2\pi i)^{2k-1} f(\tau)(X - \tau Y)^{2k-2} d\tau$$

where $q = \exp(2i\pi\tau)$. It is to be viewed as a global section of $V_{2k-2} \otimes \Omega_{\mathbb{H}}^1$ over the upper-half plane $\mathbb{H}$. In [7], §5, we defined regularised iterated integrals of Eisenstein series between cusps. Consider the double integrals:

$$(9.2) \qquad \int_0^{\infty} \underline{E}_{2m+2}(\tau) \underline{E}_{2n+2}(\tau) \quad \in \quad V_{2m} \otimes V_{2n} \otimes \mathbb{C}$$

along the geodesic path from 0 to ∞ (suitably interpreted as the path S from $\overrightarrow{1}_{\infty}$, the unit tangential base point at the cusp to itself). For each $k \geq 0$, there is a canonical morphism of SL_2-representations ([7], §2.4)

$$\partial^k : V_{2m} \otimes V_{2n} \longrightarrow V_{2m+2n-2k} \;.$$

In this way the imaginary part of the image of (9.2) under ∂^1 defines a homogeneous polynomial in $\mathbb{R}[X, Y]$ of degree $2m + 2n - 2$ whose coefficients can be described explicitly. The method described in [7], §11, computes this polynomial as the Petersen inner product of two (real analytic) modular forms. The part we are interested, by the unfolding method, corresponds to the convolution of two Eisenstein series, and yields a certain multiple of an odd zeta value. One knows that the ratios of these coefficients are the odd period polynomials of Eisenstein series.

First of all, we give the precise technical statement about the periods of double Eisenstein series, and then explain how this relates to the image of $\mathfrak{g}^{\mathbf{m}}$ in $\mathrm{Der}^{\Theta} \mathbb{L}(a, b)$.

9.2. Precise statement. All the notation in this section is borrowed from [7], §11. Let $k \geq 1$ be odd, $a, b \geq 2$, and $w = 2a + 2b - 2k - 2$. Set

$$(9.3) \qquad \widetilde{I}^k_{2a,2b} = I^k_{2a,2b} + \delta^0 \partial^k(\overline{v}_{2a} \cup \overline{b}_{2b} - \overline{b}_{2a} \cup \overline{v}_{2b})$$

where δ^0 is the boundary for 0-cochains, and for all $k \geq 2$,

$$\overline{v}_{2k} = (2\pi i)^{2k-1} v_{2k}$$

where v_{2k} was defined in [7], (10.7). Then I claim that $\widetilde{I}^k_{2a,2b}$ is cocycle for $\mathrm{SL}_2(\mathbb{Z})$ which vanishes on $T \in \mathrm{SL}_2(\mathbb{Z})$. Then it satisfies

$$(9.4) \qquad \{i\, \widetilde{I}^k_{2a,2b}, e^0_w\} = 6(2\pi i)^{-w-1} C^k_{a,b}\, \zeta(k+1)\zeta(2a-k-1)\zeta(2b-k-1)\,\zeta(k+w)$$

where e^0_w the rational Eisenstein cocycle defined in [7] §7.3 and

$$C^k_{a,b} = k!(2a-2)!(2b-2)!(k+w-1)!.$$

The equation (9.4) can be written, using Euler's formula, in terms of a product of three bernoulli numbers and a single odd zeta value. Note that we only require the case $k = 1$ here. The proof is essentially the same as in [7] with modifications to account for divergences.

9.3. Putting the pieces together. We sketch the main ingredients. Let G^{rel}_B denote the completion of $\mathrm{SL}_2(\mathbb{Z})$ relative to its inclusion into $\mathrm{SL}_2(\mathbb{Q})$. It acts on $\mathcal{P}$ (3.1) through a quotient, called the 'Eisenstein quotient' G^{eis}_B by Hain and Matsumoto [26, 27]. Let us suppose for simplicity[2] that its affine ring is an ind-object of the category of mixed Tate motives over $\mathbb{Z}$. There is a natural map

$$\mathrm{SL}_2(\mathbb{Z}) \to G^{\mathrm{rel}}_B(\mathbb{Q}) \to G^{\mathrm{eis}}_B(\mathbb{Q})$$

and hence a morphism, for any $\gamma \in \mathrm{SL}_2(\mathbb{Z})$,

$$(9.5) \qquad \begin{array}{ccc} \mathrm{Isom}_{\mathcal{MT}(\mathbb{Z})}(\omega_{dR}, \omega_B) & \longrightarrow & G^{\mathrm{eis}}_{dR} \\ \phi & \mapsto & \phi(\gamma) \end{array}$$

where G^{eis}_{dR} denotes the $\mathbb{Q}$-affine group scheme underlying $G^{\mathrm{eis}}_B(\mathbb{C})$ which coincides with the $\mathbb{Q}$-de Rham structure of[25]. The map (9.5) defines a canonical homomorphism $\mathrm{SL}_2(\mathbb{Z}) \to G^{\mathrm{eis}}_{dR}(P^{\mathfrak{m}})$ where $P^{\mathfrak{m}} = \mathcal{O}(\mathrm{Isom}_{\mathcal{MT}(\mathbb{Z})}(\omega_{dR}, \omega_B))$ is the ring of motivic periods of $\mathcal{MT}(\mathbb{Z})$. Now Hain and Matsumoto show [27] that there is a splitting

$$(9.6) \qquad G^{\mathrm{eis}}_{dR} \cong U^{\mathrm{eis}}_{dR} \rtimes \mathrm{SL}_2$$

so the composition $\mathrm{SL}_2(\mathbb{Z}) \to G^{\mathrm{eis}}_{dR}(P^{\mathfrak{m}}) \to U^{\mathrm{eis}}_{dR}(P^{\mathfrak{m}})$ defines a cocycle

$$C^{\mathfrak{m}} \in Z^1(\mathrm{SL}_2(\mathbb{Z}), U^{\mathrm{eis}}_{dR}(P^{\mathfrak{m}}))\ .$$

Its period $\mathrm{per}(C^{\mathfrak{m}}) \in Z^1(\mathrm{SL}_2(\mathbb{Z}), U^{\mathrm{eis}}_{dR}(\mathbb{C}))$ is the image of the 'canonical cocycle' $\mathcal{C}$ defined in [7] in the group $U^{\mathrm{eis}}(\mathbb{C})$. Its coefficients are given by certain linear combinations of regularised iterated integrals of Eisenstein series.

There is a canonical isomorphism $\mathbb{Q}(n)_{dR} \xrightarrow{\sim} \mathbb{Q}(n)_B$ given by the comparison isomorphism $\mathrm{comp}_{B,dR}$ scaled by a suitable power of $2\pi i$. Since $\mathcal{O}(\mathrm{SL}_2)$ is a direct sum of pure Tate motives, we have in particular an isomorphism $M_0\mathcal{O}(\mathrm{SL}_2)_{dR} \xrightarrow{\sim} M_0\mathcal{O}(\mathrm{SL}_2)_B$ of $\mathbb{Q}$-vector spaces, where M is the weight filtration. Now consider the map

$$\mathcal{O}(G^{\mathrm{eis}}_{dR}) \to M_0\mathcal{O}(G^{\mathrm{eis}}_{dR}) = M_0(\mathcal{O}(\mathrm{SL}_2))_{dR} \cong M_0(\mathcal{O}(\mathrm{SL}_2))_B \to \mathcal{O}(G^{\mathrm{eis}}_B)$$

where the first map is the natural map to $\mathcal{O}(G^{\mathrm{eis}}_{dR})/F^1$, and the second follows from (9.6), Dually, this defines a linear map $\mathcal{O}(G^{\mathrm{eis}}_B)^\vee \to \mathcal{O}(G^{\mathrm{eis}}_{dR})^\vee$, and hence a map $\mathrm{SL}_2(\mathbb{Z}) \to \mathcal{O}(G^{\mathrm{eis}}_{dR})^\vee$. One verifies that the element $STS^{-1} \in \mathrm{SL}_2(\mathbb{Z})$, being lower

_____

[2]This is not strictly required for the following argument

triangular, maps to the element $STS^{-1} \in \mathcal{O}(\mathrm{SL}_2)_{dR}^{\vee} \subset \mathcal{O}(G_{dR}^{\mathrm{eis}})^{\vee}$. By applying this construction to the coefficients of $C_{STS^{-1}}^{\mathfrak{m}}$, and sending $U_{dR}^{\mathrm{eis}} \to \mathrm{Aut}(\mathcal{P})$ we obtain an element

$$C_{STS^{-1}}^{dR} \in \mathrm{Aut}(\mathcal{P})(\mathcal{O}(G^{dR}))$$

in the $\mathcal{O}(G^{dR})$-points of $\mathrm{Aut}(\mathcal{P})$. On the other hand, the structure of mixed Tate motive over $\mathbb{Z}$ on $\mathcal{P}$ defines a morphism

$$I_1 : G^{dR} \longrightarrow \mathrm{Aut}(\mathcal{P}) \ .$$

Unravelling all the definitions, one verifies that $C_{STS^{-1}}^{dR} = I_1^{STS^{-1}} I_1^{-1}$. Taking a logarithm and applying the Baker-Campbell-Hausdorff formula gives a relation between i_1 and $C_{STS^{-1}}^{dR}$. The latter is determined from C_S^{dR} and C_T^{dR} by the cocycle relations. From this, and knowledge of C_T, one can read off the first few coefficients in the expansion of $i_1(\sigma_{2n+1})$, for $n \geq 2$ from the coefficients of $\zeta^{dR}(2n+1)$ in C_S^{dR}. This is the mechanism by which information about the map i_1 can be computed from double Eisenstein integrals. A different approach would be via a motivic version of the analytic arguments of [19], which give a relation between the Drinfeld associator and the image of $C_{STS^{-1}}$ in U_{dR}^{eis}. Further details will be given in a joint work with Hain.

References

[1] **N. Arbesfeld, B. Enriquez**: 'On a lower central series filtration of the Grothendieck-Teichmüller Lie algebra grt_1', `arXiv:1406.0675` (2014).

[2] **F. Brown**: *Anatomy of an associator*, notes (see also `http://www.ihes.fr/~brown/AnatomyBeamer.pdf`).

[3] **F. Brown**: *Mixed Tate motives over* $\mathbb{Z}$, Annals of Math., volume 175, no. 1, 949-976, (2012).

[4] **F. Brown**: *Motivic periods and the projective line minus 3 points*, proceedings of the ICM (2014), `arXiv:1407.5165`

[5] **F. Brown**: *Decomposition of motivic multiple zeta values*, 'Galois-Teichmuller theory and Arithmetic Geometry', Adv. Stud. Pure Math., 63, (2012).

[6] **F. Brown**: *Depth-graded motivic multiple zeta values*, `arXiv:1301.3053`, (2013)

[7] **F. Brown**: *Multiple modular values*, `arXiv:1407.5167`, (2014)

[8] **D. Broadhurst, D. Kreimer** : *Association of multiple zeta values with positive knots via Feynman diagrams up to 9 loops*, Phys. Lett. B 393 (1997), no. 3-4, 403-412.

[9] **F. Brown, A. Levin**: *Multiple Elliptic Polylogarithms*, `arXiv:1110.6917`, (2011)

[10] **J. Blümlein, D.J. Broadhurst, J.A.M. Vermaseren**: *The Multiple Zeta Value Data Mine*, Comput. Phys. Commun. 181, 582-625, (2010).

[11] **D. Calaque, B. Enriquez, P. Etingof**: *Universal KZB equations: the elliptic case*, Algebra, arithmetic, and geometry: in honor of Yu. I. Manin. Vol. I, 165266, Progr. Math., 269, (2009).

[12] **P. Cartier**: *Fonctions polylogarithmes, nombres polyzêtas et groupes pro-unipotents*, Séminaire Bourbaki, Exp. No. 885, Astérisque No. 282 (2002), 137-173.

[13] **P. Deligne**: *Multizêtas*, Séminaire Bourbaki (2012).

[14] **P. Deligne**: *Le groupe fondamental unipotent motivique de* $\mathbb{G}_m - \mu_N$, *pour* $N = 2, 3, 4, 6$ *ou* 8, Publ. Math. Inst. Hautes Études Sci. 101 (2010).

[15] **P. Deligne, A. B. Goncharov**: *Groupes fondamentaux motiviques de Tate mixte*, Ann. Sci. École Norm. Sup. 38 (2005), 1–56.

[16] **V. Drinfeld**: *On quasi-triangular quasi-Hopf algebras and some group closely related with* $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, Algebra i Analiz 2 (1990), no. 4, 149-181.

[17] **J. Ecalle**: *ARI/GARI, la dimorphie et l'arithmétique des multizêtas: un premier bilan*, J. Théor. Nombres Bordeaux 15 (2003), no. 2, 411-478.

[18] **J. Ecalle**: *Multizetas, perinomal numbers, arithmetical dimorphy, and ARI/GARI*, Ann. Fac. Sci. Toulouse Math. (6) 13 (2004), no. 4, 683-708.

[19] **B. Enriques**: *Elliptic associators*, Selecta Math. (N.S.) 20 (2014), no. 2, 491584.

[20] **B. Enriques, P. Lochak**: *Homology of depth-graded motivic Lie algebras and koszulity*, arXiv:1407.4060

[21] **H. Furusho**: *Double shuffle relation for associators*, Ann. of Math. (2) 174 (2011), no. 1, 341-360.

[22] **H. Gangl, M. Kaneko, D. Zagier**: *Double zeta values and modular forms*, Automorphic forms and zeta functions, 71-106, World Sci. Publ., Hackensack, NJ, 2006.

[23] **A. B. Goncharov**: *Galois symmetries of fundamental groupoids and noncommutative geometry*, Duke Math. J.128 (2005), 209-284.

[24] **Goncharov, A. B.**:*Multiple polylogarithms, cyclotomy and modular complexes*, Math. Res. Lett. 5 (1998), no. 4, 497-516.

[25] **R. Hain** : *Notes on the Universal Elliptic KZB Equation*, arXiv:1309.0580, (2013)

[26] **R. Hain** : *The Hodge-de Rham theory of modular groups*, arXiv:1403.6443, (2014)

[27] **R. Hain, M. Matsumoto**: *Universal Mixed Elliptic Motives*, notes.

[28] **Y. Ihara**: *Some arithmetic aspects of Galois actions on the pro-p fundamental group of* $\mathbb{P}^1\backslash\{0,1,\infty\}$, Proceedings Symp. in Pure Math. 70 (2002), 247-273.

[29] **K. Ihara, M. Kaneko, D. Zagier**: *Derivation and double shuffle relations for multiple zeta values*, Compos. Math, 142 (2006) 307-338.

[30] **Y. Ihara, N. Takao**: seminar talk (May 1993)

[31] **W. Kohnen, D. Zagier**: *Modular forms with rational periods*, Modular forms (Durham, 1983), 197-249, Ellis Horwood 1984.

[32] **A. Levin, G. Racinet**: *Towards multiple elliptic polylogarithms*, arXiv:math/0703237 (2007).

[33] **Y. Manin**: *Iterated integrals of modular forms and non-commutative modular symbols*, Algebraic geometry and number theory, 565-597, Prog. Math. 253 (2006).

[34] **Y. Manin**: *Iterated Shimura integrals*, Moscow Math. J. 5 (2005), 869-881

[35] **W. McCallum, R. Sharifi**: *A cup product in the Galois cohomology of number fields*, Duke Math. J. Volume 120, Number 2 (2003), 269-310.

[36] **H. Nakamura**, Tangential base points and Eisenstein power series, Aspects of Galois theory 202-217, London Math. Soc. Lecture Note Ser., 256, CUP (1999).

[37] **A. Pollack**: *Relations between derivations arising from modular forms*, undergraduate thesis, Duke University, (2009)

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

[39] **R. Sharifi**: *Relationships between conjectures on the structure of pro-p Galois groups unramified outside p*, Proceedings Symp. in Pure Math. 70 (2002), 275-284.

[40] **D. Zagier**: *Periods of modular forms, traces of Hecke operators, and multiple zeta values*, RIMS Kokyuroku 843 (1993), 162-170.

[41] **D. B. Zagier**, *Evaluation of the multiple zeta values* $\zeta(2,\ldots,2,3,2,\ldots,2)$,