Forum of Mathematics, Sigma (2017), Vol. 5, e1, 56 pages
doi:10.1017/fms.2016.29

ZETA ELEMENTS IN DEPTH 3 AND THE FUNDAMENTAL LIE ALGEBRA OF THE INFINITESIMAL TATE CURVE

FRANCIS BROWN
All Souls College, Oxford, Oxford OX1 4AL, UK;
email: francis.brown@all-souls.ox.ac.uk

Received 3 February 2016; accepted 8 August 2016

Abstract

This paper draws connections between the double shuffle equations and structure of associators; Hain and Matsumoto's universal mixed elliptic motives; 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 solve the double shuffle equations for multiple zeta values in depths two and three.

2010 Mathematics Subject Classification: 11M32 (primary); 11F67 (secondary)

1. Introduction

1.1. Motivation. A consequence of Belyi's theorem [2] is that the absolute Galois group of $\mathbb{Q}$ acts faithfully on the profinite completion of the fundamental group of $\mathbb{P}^1\backslash\{0, 1, \infty\}$, or in other words, the homomorphism

$$\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{Aut}(\widehat{\pi}_1(\mathbb{P}^1\backslash\{0, 1, \infty\}, \vec{1}_0)) \tag{1.1}$$

is injective. In his 'Esquisse d'un programme' ([22], $\frac{3}{4} - \frac{4}{5}$), Grothendieck suggests that this should give a way to 'parametrize' elements of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ by suitable elements of the profinite group on two generators.

Transposing this question to the prounipotent setting leads to a more attainable goal. The unipotent de Rham analogue of (1.1) is the statement that

$$G^{\mathrm{dR}}_{\mathcal{MT}(\mathbb{Z})} \longrightarrow \mathrm{Aut}(\pi_1^{\mathrm{dR}}(\mathbb{P}^1\backslash\{0, 1, \infty\}, \vec{1}_0)) \tag{1.2}$$

is injective [4], where $G^{\mathrm{dR}}_{\mathcal{MT}(\mathbb{Z})}$ is the de Rham motivic Galois group of the category of mixed Tate motives over $\mathbb{Z}$. Its graded Lie algebra is noncanonically generated by elements $\sigma_3, \sigma_5, \ldots$ in every odd degree $-3, -5, \ldots$. Furthermore, the de Rham fundamental groupoid of $\mathbb{P}^1\backslash\{0, 1, \infty\}$, and its group of inertia-preserving automorphisms, can be realized as formal power series in two noncommuting variables $\mathsf{x}_0, \mathsf{x}_1$. This gives a concrete version of Grothendieck's programme:

PROBLEM 1. Describe explicitly the images of elements σ_{2n+1} in $\mathbb{Q}\langle\langle \mathsf{x}_0, \mathsf{x}_1\rangle\rangle$.

A priori, the elements σ_{2n+1} are not canonical, since they are only well defined up to addition of commutators. For example, in degree -11, there is a two-dimensional space of possible generators $\mu\sigma_{11} + \lambda[\sigma_3, [\sigma_5, \sigma_3]]$, where $\mu, \lambda \in \mathbb{Q}$, with respect to some choices of σ_{2n+1}. Strangely enough, the proof [4] of the injectivity of (1.2) actually provides a canonical choice of generators σ^h_{2n+1}, but it seems very difficult to describe these explicitly, and their coefficients involve large prime factors. The proof in [4] also provides a canonical element $\tau^h \in \mathbb{Q}\langle\langle \mathsf{x}_0, \mathsf{x}_1\rangle\rangle$ in even degrees, which is a motivic version of a rational associator. The important problem of constructing an explicit rational associator was suggested by Drinfeld [13] in 1990 and is still open. The elements σ^h_{2n+1}, τ^h are related to the choice of the Hoffman–Lyndon basis for motivic multiple zeta values.

A better reformulation of problem 1 is therefore

PROBLEM 2. Give a canonical choice, and explicit construction of, elements σ_{2n+1} in $\mathbb{Q}\langle\langle x_0, x_1\rangle\rangle$ and a motivic rational associator τ in $\mathbb{Q}\langle\langle x_0, x_1\rangle\rangle$.

In this paper, we attempt to solve problem 2 using the depth filtration, which corresponds to the degree in the letter x_1. The heads of the elements σ_{2n+1} (respectively τ) in depths 1 and 2 (respectively depth 1) are canonical, but their tails are not. Using three different techniques (via double shuffle equations, the unipotent fundamental group of the punctured Tate curve, and the relative completion of $\mathrm{SL}_2(\mathbb{Z})$) we show, surprisingly, that there is an explicit way to write down canonical elements σ^c_{2n+1} (respectively τ^c) to the next order, namely depth 3 (respectively depth 2). Their coefficients involve products of Bernoulli numbers, which can be thought of as a higher-depth version of Euler's formula expressing even zeta values as multiples of powers of π. This raises the possibility that problem 2, which at first sight seems hopeless, may in fact have an explicit solution to all depths. Such a solution would give an algorithm to write any motivic multiple zeta value in terms of a basis, and in addition, would give an explicit representation of the motivic Galois group of $\mathcal{MT}(\mathbb{Z})$.

The general 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 by passing to genus one.

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

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

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

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 $\mathsf{x}_0, \mathsf{x}_1$ dual to loops around 0 and 1. Since ${}_1\Pi_1$ is the de Rham realization of a pro-object in the category of mixed Tate motives over $\mathbb{Z}$, it admits an action of the Tannakian fundamental group $G^{\mathrm{dR}}_{\mathcal{MT}(\mathbb{Z})}$. Denote the graded Lie algebra of the latter by

$$ \mathfrak{g}^{\mathfrak{m}} = \mathbb{L}(\sigma_3, \sigma_5, \ldots). \tag{1.3} $$

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

$$ i_0 : \mathfrak{g}^{\mathfrak{m}} \longrightarrow \mathrm{Der}^1\, \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1) \tag{1.4} $$

where $\mathrm{Der}^1\, \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ denotes derivations which send x_1 to 0. The subscript of i_0 refers to genus zero. Furthermore, we know that (1.4) factors through a morphism

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

where the second map sends $f \in \mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ to the derivation $\mathsf{x}_0 \mapsto [\mathsf{x}_0, f], \mathsf{x}_1 \mapsto 0$. The main result of [4] states that i, and hence i_0, is injective, and therefore enables us to expand elements $\sigma \in \mathfrak{g}^{\mathfrak{m}}$ in 'coordinates' x_0 and x_1. We wish to describe the $i(\sigma_{2n-1})$ as explicitly as possible. One way to do this is using the known relations which are satisfied by its image. Indeed, Racinet [35] showed 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 [17], are contained in $\mathfrak{dmr}_0$. The associator relations will not be used in this paper.

It is well known that

$$ i(\sigma_{2n+1}) = \mathrm{ad}(\mathsf{x}_0)^{2n}\mathsf{x}_1 + \ (\text{terms of degree } \geqslant 2 \text{ in } \mathsf{x}_1), \tag{1.5} $$

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

but little is known about the coefficients of $i(\sigma_{2n+1})$ of degrees $\geqslant 3$ in the x_1, and worse, they depend on the choice of generators σ_{2n+1}. In this paper, we show:

THEOREM 1.1.

(1) *There is a choice of generators* $\sigma^c_{2n+1} \in \mathfrak{g}^{\mathfrak{m}}$ *which is given by an explicit formula* (1.11) *modulo terms of degree* $\geqslant 5$ *in* x_1.

(2) *There is a rational associator* τ^c *which is given by an explicit formula modulo terms of degree* $\geqslant 4$ *in* x_1.

Statement (1) is surprising because a choice of generators σ_{2n+1} are *a priori* only well defined up to addition of triple commutators of σ_{2m+1}. The key point is that by passing to genus 1, we can fix these uniquely. A similar story holds for (2).

In the course of the proof of this theorem, we discover that it is more convenient to consider a different normalization for the σ_{2n+1} from the canonical normalization (1.5). For want of a better name, we shall call it the *heretical normalization*

$$\underline{\sigma}_{2n+1} \equiv \frac{B_{2n}}{(2n)!}\sigma_{2n+1} \text{ (mod terms of degree } \geqslant 2 \text{ in } \mathsf{x}_1), \tag{1.6}$$

where B_{2n} is the $2n$th Bernoulli number. Throughout this paper, objects which are normalized according to the heretical normalization will be underscored.

1.3. The fundamental group of the first-order Tate curve. Let $E^{\times}_{\partial/\partial q}$ denote the punctured fibre of the universal elliptic curve $\mathcal{M}_{1,2} \to \mathcal{M}_{1,1}$ over the tangential base point $\partial/\partial q$ 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 [25]) that its de Rham fundamental group

$$\mathcal{P} = \pi_1^{\mathrm{dR}}(E^{\times}_{\partial/\partial q}, \vec{1}_1) \tag{1.7}$$

where $\vec{1}_1$ is the tangent vector of length 1 with respect to a natural choice of holomorphic coordinate w on $E^{\times}_{\partial/\partial q}$, is the de Rham realization of a pro-object in the category of mixed Tate motives over $\mathbb{Z}$. Its weight filtration is denoted by M (for monodromy-weight), but it also possess an additional filtration W (the 'elliptic weight'), which coincides with the lower central series filtration. Its associated bigraded Lie algebra is the free Lie algebra on certain canonical generators a, b. Correspondingly, one obtains a morphism of Lie algebras

$$i_1 : \mathfrak{g}^{\mathfrak{m}} \longrightarrow \mathrm{Der}^{\Theta}\,\mathbb{L}(\mathsf{a}, \mathsf{b}) \subseteq \mathrm{Der}\,\mathbb{L}(\mathsf{a}, \mathsf{b}) \tag{1.8}$$

where Der^{Θ} denotes the subspace of derivations δ such that $\delta(\Theta) = 0$, where $\Theta = [\mathsf{a}, \mathsf{b}]$. We shall show as a consequence of [4] that (1.8) is injective. (If one

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

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, only the W-filtration, see [25].)

There exist distinguished elements $\varepsilon^{\vee}_{2n} \in \mathrm{Der}^{\Theta}\,\mathbb{L}(\mathsf{a},\mathsf{b})$ whose action on a is

$$\varepsilon^{\vee}_{2n}(\mathsf{a}) = \mathrm{ad}(\mathsf{a})^{2n}\mathsf{b} \quad \text{for } n \geqslant 1.$$

They were first studied by Tsunogai [32, 37] in a slightly different context and rediscovered in [11, 28]. The action of $\varepsilon^{\vee}_{2n}$ on b is determined by the condition $\varepsilon^{\vee}_{2n}\Theta = 0$ together with the fact that it is homogeneous of degree $2n$ in a, b. The derivations $\varepsilon^{\vee}_{2n}$ 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 the completion (with respect to the lower central series) $\mathbb{L}(\mathsf{a},\mathsf{b})^{\wedge}$ via the Lie algebra generated by the $\varepsilon^{\vee}_{2n}$ and their images $\mathrm{ad}(\varepsilon^{\vee}_0)^k\varepsilon^{\vee}_{2n}$ under the adjoint action of

$$\varepsilon^{\vee}_0 \in \mathrm{Der}^{\Theta}\,\mathbb{L}(\mathsf{a},\mathsf{b}) \quad \text{where } \varepsilon^{\vee}_0(\mathsf{a}) = \mathsf{b},\ \ \varepsilon^{\vee}_0(\mathsf{b}) = 0.$$

Denote the Lie subalgebra generated by the $\varepsilon^{\vee}_{2n}$, for all $n \geqslant 0$, by

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

It is the M, W-bigraded image of the universal monodromy [24]. The elements $\varepsilon^{\vee}_{2n}$ satisfy relations studied by Pollack [34]. The image of $\mathfrak{g}^{\mathfrak{m}}$ in $\mathrm{Der}^{\Theta}\,\mathbb{L}(\mathsf{a},\mathsf{b})$ under (1.8) is by no means contained in $\mathfrak{u}^{\mathrm{geom}}$, but in low degrees with respect to b, the $\varepsilon^{\vee}_{2n}$ give canonical 'coordinates' in which to write down the initial terms of elements $i_1(\sigma_{2n+1})$. Indeed, the motivic version of a formula due to Nakamura is

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

for all $n \geqslant 1$. We shall prove:

THEOREM 1.2. *Let $n \geqslant 2$. There exists a choice of elements σ^c_{2n+1} satisfying*

$$i_1(\underline{\sigma}^c_{2n+1}) \equiv \underline{\varepsilon}^{\vee}_{2n+2} + \sum_{a+b=n} \frac{1}{2b}[\underline{\varepsilon}^{\vee}_{2a+2}, [\underline{\varepsilon}^{\vee}_{2b+2}, \underline{\varepsilon}^{\vee}_0]] \ (\mathrm{mod}\ W_{-2n-5}) \tag{1.9}$$

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

This theorem is equivalent to an explicit formula for the $i_0(\sigma^c_{2n+1}) \in \mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$ modulo terms of degree $\geqslant 5$ in x_1. Note that the case σ^c_3 is exceptional. The sheer simplicity of formula (1.9) leads one to wonder if it can be extended further.

This theorem is equivalent to a formula for $i_0(\sigma^c_{2n+1})$ using an explicit morphism from the de Rham fundamental group of $\mathbb{P}^1\backslash\{0, 1, \infty\}$ to that of $E^{\times}_{\partial/\partial q}$ which was written down by Hain. Since i_0 and i_1 are compatible with this morphism, the expansion (1.9) of $i_1(\sigma^c_{2n+1})$ in terms of derivations $\varepsilon^{\vee}_{2k}$ implies an expansion for $i_0(\sigma^c_{2n+1})$ in terms of $\mathsf{x}_0, \mathsf{x}_1$ in genus 0.

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

1.4. Rankin–Selberg method. The third way of understanding the elements σ^c_{2n+1}, and the inspiration for this paper, came from the theory of iterated integrals of holomorphic modular forms for $SL_2(\mathbb{Z})$. The coefficients in equation (1.9) come from the computation [8] 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 how I found both the heretical normalizations (1.6), and the formula (1.9).

1.5. Further remarks. We discuss the methods used in this paper, and further applications to the double shuffle equations and Broadhurst–Kreimer conjecture.

1.5.1. Commutative power series and anatomy of associators. One tool which we use extensively is the method of commutative power series. It is closely related to Ecalle's theory of moulds [14, 15]. Let $\mathbb{L}(\mathsf{u},\mathsf{v})$ be the free bi-graded Lie algebra generated by two elements u,v. The degree in v will be called the depth-grading. For any $r\geqslant 1$, elements of depth r in the tensor algebra $T(\mathsf{u},\mathsf{v})$ can be represented as commutative polynomials in r variables

$$\begin{aligned} \rho : \mathrm{gr}^r_{\mathsf{v}}\, T(\mathsf{u},\mathsf{v}) &\longrightarrow \mathbb{Q}[x_1,\ldots,x_r] \quad r\geqslant 1 \\ \mathsf{u}^{i_0}\mathsf{v}\mathsf{u}^{i_1}\ldots\mathsf{v}\mathsf{u}^{i_r} &\mapsto x_1^{i_1}\ldots x_r^{i_r}. \end{aligned} \tag{1.10}$$

We apply this construction to $(\mathsf{u},\mathsf{v})=(\mathsf{x}_0,\mathsf{x}_1)$ and $(\mathsf{u},\mathsf{v})=(\mathsf{a},\mathsf{b})$, and their derivation algebras. We shall explain why, in certain situations, it is natural to rescale the morphism ρ by introducing polynomial denominators. In this manner, elements of $\mathrm{Der}^1\,\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$ and $\mathrm{Der}^{\Theta}\,\mathbb{L}(\mathsf{a},\mathsf{b})$ are uniquely 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 (in fact, several natural choices with different properties) if one allows poles:

THEOREM 1.3 [3]. *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 $\xi^{(r)}_{2n+1}\in\mathbb{Q}(x_1,\ldots,x_r)$ in weight $2n+1\geqslant 3$. Their components in depths $r=1,2$ are polynomials, but they have poles in depths $r\geqslant 3$. Furthermore, a new element emerges in weight -1 which we denote by $\xi^{(r)}_{-1}\in\mathbb{Q}(x_1,\ldots,x_r)$. The idea of [3] is to write the polynomial representation of zeta elements $\rho(i(\sigma_{2n-1}))$ in terms of the rational functions ξ_{2n+1}, for $n \geqslant -1$ ('anatomy'). It can be computed explicitly in low depth:

THEOREM 1.4. *If $\{\,,\}$ denotes the Ihara bracket, transposed and extended to rational functions via (1.10), then the canonical zeta elements up to depth 4 are given, in the heretical normalization, by the simple formula:*

$$\rho(i(\underline{\sigma}^c_{2n+1})) \equiv \underline{\xi}_{2n+1} + \sum_{a+b=n} \frac{1}{2b}\{\underline{\xi}_{2a+1}, \{\underline{\xi}_{2b+1}, \underline{\xi}_{-1}\}\} \ (\mathrm{mod}\ depth \geqslant 5). \quad (1.11)$$

Writing this formula in the canonical, as opposed to heretical normalizations, produces coefficients which are 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 Section 9.

Theorem 1.4 is proved by combinatorial methods, and uses Goncharov's theorem [20] enumerating the solutions to the double shuffle equations in depth 3. It makes no reference to the first-order Tate curve $E^{\times}_{\partial/\partial q}$. 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 rational function representations of the ξ_{2n+1} correspond in low depths to those of the derivations $\varepsilon^{\vee}_{2n+2}$, and enables us to deduce Theorem 1.2 from Theorem 1.4.

1.5.2. Double shuffle equations. Our elements σ^c_{2n+1} are explicit solutions to the double shuffle equations in depths $\leqslant 4$ and odd weights. We also construct, in Section 7.1, an explicit solution τ^c in depths $\leqslant 3$ and all even weights. Using Goncharov's theorem mentioned above, we deduce

THEOREM 1.5. *Every solution to the regularized double shuffle equations in depths $\leqslant 4$ (odd weight) and depths $\leqslant 3$ (even weight) can be expressed using the explicit elements σ^c_{2n+1} and the element τ^c.*

This theorem can be applied to the method of [5] for decomposing motivic multiple zeta values into a basis, which involved a numerical computation at each step. One application of the elements σ^c_{2n+1} and τ^c is to remove this transcendental step, leading to an exact algorithm for proving any motivic relation between multiple zeta values in depth $\leqslant 3$, and any weight. It replaces the need to store tables of multiple zeta values in this range [10].

A further manifestation of the double shuffle equations occurs in genus 1. As above, we encode elements of $\mathrm{Der}^{\Theta}\,\mathbb{L}(\mathsf{a},\mathsf{b})$ by rational functions, by composing

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

the morphism $\delta \mapsto \delta(\mathsf{a}) : \mathrm{Der}^{\Theta}\,\mathbb{L}(\mathsf{a},\mathsf{b}) \longrightarrow \mathbb{L}(\mathsf{a},\mathsf{b}) \subset T(\mathsf{a},\mathsf{b})$ with the linear map

$$\begin{aligned} \mathrm{gr}^r_{\mathsf{b}}\, T(\mathsf{a},\mathsf{b}) &\longrightarrow \mathbb{Q}(x_1,\ldots,x_r) \\ \mathsf{a}^{i_0}\mathsf{b}\mathsf{a}^{i_1}\mathsf{b}\ldots\mathsf{b}\mathsf{a}^{i_r} &\mapsto \frac{x_1^{i_1}\ldots x_r^{i_r}}{x_1(x_1-x_2)\ldots(x_{r-1}-x_r)x_r}. \end{aligned} \tag{1.12}$$

In [3], we defined a bigraded Lie algebra $\mathfrak{pls}$ to be the space of solutions to the linearized double shuffle equations with poles at worst of the above form.

PROPOSITION 1.6. *The Lie algebra of geometric derivations is contained, via (1.12), in the space of solutions to the linearized double shuffle equations:* $\mathfrak{u}^{\mathrm{geom}} \subset \mathfrak{pls}$.

Thus the linearized double shuffle equations arise naturally in the elliptic setting. An obvious question to ask is if $\mathfrak{u}^{\mathrm{geom}} = \mathfrak{pls}$. It is proved in depths $\leqslant 3$ in an appendix, where we also compute the generating function of dimensions for both $\mathfrak{u}^{\mathrm{geom}}$ and $\mathfrak{pls}$ in this range.

The previous proposition can be used to detect nongeometric derivations. Indeed, the stuffle relations give rise to an infinite family of functions

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

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

It is well known since Ihara and Takao [26] that there exist quadratic relations

$$\sum_{i,j} \lambda_{i,j}[\sigma_{2i+1},\sigma_{2j+1}] \equiv 0 \text{ (mod terms of degree } \geqslant 4 \text{ in } \mathsf{x}_1) \tag{1.13}$$

where $\lambda_{i,j} \in \mathbb{Q}$ are coefficients of period polynomials P of even, cuspidal $\mathrm{SL}_2(\mathbb{Z})$-cocycles. In [7], 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 nontrivial differential (conjecturally, the only nontrivial differential) in this spectral sequence. A motivic version of the Broadhurst–Kreimer conjecture provides an explicit presentation for $\mathrm{gr}_D\mathfrak{g}^{\mathfrak{m}}$ in terms of the σ^c_{2n+1} (see Section 8).

1.6. Aide-mémoire. There are several different filtrations at play in this paper. At a referee's request, and for the convenience of the reader, the corresponding degrees are summarized below.

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

For the de Rham Lie algebra $\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ of the fundamental group of $\mathbb{P}^1\backslash\{0, 1, \infty\}$, we define the Tate degree to be one half of the weight as a mixed Tate motive.

	$T =$ 'Tate' degree $= -L$	$D =$ 'Depth' degree
x_0	-1	0
x_1	-1	1

For the bigraded de Rham Lie algebra $\mathbb{L}(\mathsf{a}, \mathsf{b})$ in genus one:

	$M =$ 'Monodromy-weight'	$W =$'Elliptic weight'$= -L$	$B =$ b-degree
a	-2	-1	0
b	0	-1	1

The Hodge filtration will always be denoted by F, and the lower central series by L. The weight-grading in genus 0, which is double the Tate degree, is traditionally denoted by W. We shall never use this notation, since it corresponds to the M degree in genus 1, and W here will always denote the elliptic weight filtration.

Finally, we present a tableau of the main Lie algebras which will be defined and studied in this paper. The second column features the main objects of study: the motivic Lie algebra $\mathfrak{g}^{\mathfrak{m}}$, its depth-graded version $\mathfrak{d}$, and the geometric derivations $\mathfrak{u}^{\mathrm{geom}}$. The next column features Lie algebras of solutions to double shuffle equations (Racinet's double shuffle Lie algebra $\mathfrak{dmr}_0$, the linearized double shuffle algebra $\mathfrak{ls}$, and its version with poles $\mathfrak{pls}$). The right-hand column lists the ambient space of derivations on de Rham fundamental groups of curves. One of the main points in this paper is that the depth-graded motivic Lie algebra in genus 0 is very closely related to the B-graded geometric Lie algebra in genus 1. In the following table '$g = 0, 1$' denotes genus 0, 1, respectively.

	Protagonist		Double shuffle		Ambient space
$g = 0$	$\mathfrak{g}^{\mathfrak{m}}$	$\subseteq$	$\mathfrak{dmr}_0$	$\subseteq$	$\mathrm{Der}^1\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$
D-graded $g = 0$	$\mathfrak{d} := \mathrm{gr}_D\mathfrak{g}^{\mathfrak{m}}$	$\subseteq$	$\mathrm{gr}_D\mathfrak{dmr}_0 \quad \subseteq \quad \mathfrak{ls}$	$\subseteq$	$\mathrm{gr}_D\mathrm{Der}^1\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$
B-graded $g = 1$	$\mathfrak{u}^{\mathrm{geom}}$	$\subseteq$	$\mathfrak{pls}$	$\subseteq$	$\mathrm{gr}_B\mathrm{Der}^{\Theta}\mathbb{L}(\mathsf{a}, \mathsf{b})$

In the appendix, we introduce a new filtration $\mathfrak{R}_{\bullet}$ on $\mathfrak{pls}$ in terms of residues of rational functions such that $\mathfrak{ls} = \mathfrak{R}_0\mathfrak{pls}$. It would be interesting to interpret this filtration intrinsically on $\mathfrak{u}^{\mathrm{geom}}$ and verify that $\mathfrak{R}_0\mathfrak{u}^{\mathrm{geom}} \cong \mathfrak{d}$.

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

2. Reminders on the projective line minus 3 points

Background material can be found in [7, 12, 35].

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 T-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 $n \geqslant 2$. 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 $D^nT(\mathsf{x}_0, \mathsf{x}_1)$ spanned by words in $\geqslant n$ 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 [12]

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

is the de Rham realization 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 ${}_01_1$ from the tangential base point 1 at 0 to the tangential base point -1 at 1 gives a morphism of schemes

$$g \mapsto g._01_1 : G^{\mathrm{dR}} \longrightarrow {}_0\Pi_1. \tag{2.1}$$

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, its set of R-points ${}_0\Pi_1(R)$ 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 noncommuting variables. The Ihara group law is then given by the formula

$$\begin{aligned} \circ : {}_0\Pi_1 \times {}_0\Pi_1 &\longrightarrow {}_0\Pi_1 \\ F \circ G \;&=\; G(\mathsf{x}_0, F\mathsf{x}_1F^{-1})F. \end{aligned} \tag{2.2}$$

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

$$F \circ G = FG(F^{-1}\mathsf{x}_0F, \mathsf{x}_1).$$

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

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

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

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

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

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 $\mathfrak{g}^{\mathfrak{m}} = \mathrm{Lie}^{\mathrm{gr}}\, U^{\mathrm{dR}}$, where U^{dR} is the unipotent radical of $G^{\mathrm{dR}} = U^{\mathrm{dR}} \rtimes \mathbb{G}_m$. Equation (2.1) gives a morphism

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

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

$$ \begin{aligned} (\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1), \{\ ,\ \}) &\longrightarrow \mathrm{Der}^1\, \mathbb{L}(\mathsf{x}_0,\mathsf{x}_1) \\ \sigma &\mapsto \begin{cases} \mathsf{x}_0 \mapsto [\mathsf{x}_0, \sigma], \\ \mathsf{x}_1 \mapsto 0. \end{cases} \end{aligned} \tag{2.5} $$

THEOREM 2.1 [4]. *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 T-degree $-2n-1$, for all $n \geqslant 1$. Their images in $\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$ are the *zeta elements*

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

One also knows the coefficients of $(\mathsf{x}_0\mathsf{x}_1)^a\mathsf{x}_0(\mathsf{x}_0\mathsf{x}_1)^b$ in $i(\sigma_{2a+2b+1})$ by [4] and [39]. The elements σ_{2n+1} are not *a priori* canonical for $n \geqslant 5$. However, the Hoffman–Lyndon basis for motivic multiple zeta values allows one to define canonical choices σ^h_{2n+1} of generators [6]. Very little is known about these elements.

2.3. Linearized Ihara action and depth.

In [7], we considered the following linearized version of the Ihara action. For any $a_i \in \{\mathsf{x}_0,\mathsf{x}_1\}^\times$, let

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

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 $n \geqslant 0$, let

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

with the initial condition $a \,\underline{\circ}\, \mathsf{x}_0^n = \mathsf{x}_0^n \, a$, for $n \geqslant 0$.

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

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

It follows from this formula that the Ihara bracket is homogeneous for the D-degree (the weaker fact that it respects the depth filtration follows from the geometric interpretation of the depth filtration using the embedding $\mathbb{P}^1 \backslash \{0, 1, \infty\} \subset \mathbb{P}^1 \backslash \{0, \infty\}$ [12]). 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

$$A(a, b, c) = A(b, a, c) \tag{2.9}$$

for any $a, b, c \in T(\mathsf{x}_0, \mathsf{x}_1)$, which follows from the definitions, and implies that the linearized Ihara bracket (2.8) indeed 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 [35], Racinet defined a subspace

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

called the regularized double shuffle Lie algebra, which encodes these relations in terms of two Hopf algebra structures. He proved the

THEOREM 2.3 [35]. *The space $\mathfrak{dmr}_0$ is closed under the Ihara bracket $\{\, , \,\}$.*

Since the regularized 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

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

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

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 (Section 4), and the trick of introducing poles.

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

$$\mathfrak{d} = \mathrm{gr}^{\bullet}_D \mathfrak{g}^{\mathfrak{m}} \tag{2.10}$$

denote the associated graded Lie algebra. It is bigraded for weight and depth. The component of $\mathfrak{d}$ of depth d and T-degree $-n$ will be denoted by $\mathfrak{d}^d_n$. Let $\mathfrak{d}^d = \bigoplus_{n\geqslant 0} \mathfrak{d}^d_n$, denote the (infinite-dimensional) component in depth d, and let $\mathfrak{d}_n = \bigoplus_{d\geqslant 1} \mathfrak{d}^d_n$ denote the (finite-dimensional) component in T-degree $-n$.

The linearized double shuffle equations are a family of equations which were introduced in [18] and further studied in [7]. It follows from a variant of Racinet's theorem that their solutions, denoted by $\mathfrak{ls} \subset \mathrm{gr}^{\bullet}_D(\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1), \{\ ,\ \}_{\bullet})$, is a bigraded Lie algebra for the (graded) Ihara bracket $\{\ ,\ \}_{\bullet}$. One has $\mathrm{gr}^{\bullet}_D \mathfrak{dmr}_0 \subset \mathfrak{ls}$, and it is expected that equality holds.

A corollary of Theorem 2.1 and Racinet's theorem is

THEOREM 2.4 [7, Section 5]. $\mathfrak{d} \subset \mathfrak{ls}$.

The following theorem was proved by Tsumura [36]. See [7], Section 6.4 for a short proof. There are recent generalizations due to Glanois [19] and Panzer [33].

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

$$\mathfrak{ls}^d_n = 0 \quad \textit{if } n \equiv d+1 \ (\mathrm{mod}\ 2).$$

Combining the previous two theorems gives the

COROLLARY 2.6 (Depth-Parity theorem). *If* $n \equiv d+1 \ (\mathrm{mod}\ 2)$, *then* $\mathfrak{d}^d_n = 0$.

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

Background material for this section can be found in [23], [24], and [25].

3.1. Background. Let $E^{\times}_{\partial/\partial q}$ denote the punctured first-order Tate curve, which is the fibre 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

$$\mathcal{P} = \pi_1^{\mathrm{dR}}(E^{\times}_{\partial/\partial q}, \vec{1}_1) \tag{3.1}$$

where $\overrightarrow{1}_1$ is the tangent vector of length 1 with respect to a natural choice of holomorphic coordinate w on $E^{\times}_{\partial/\partial q}$, is, we claim, the de Rham realization of a pro-object in the category of mixed Tate motives over $\mathbb{Z}$. This will be proved in future joint work with Hain. Since its mixed Hodge structure is the limiting mixed Hodge structure of a variation, it comes equipped with a relative monodromy-weight filtration denoted by M, and a geometric weight filtration W. This data defines a universal mixed elliptic motive according to Hain and Matsumoto [25]. The associated M, W bigraded Lie algebra is the free Lie algebra $\mathbb{L}(H_{\mathrm{dR}})$ where

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

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

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

where $\Theta = [\mathsf{a},\mathsf{b}]$ and

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

This is because $\Theta = [\mathsf{a},\mathsf{b}]$ 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}, the prounipotent radical of G^{dR}.

3.2. Derivations.

Define the B-filtration to be the decreasing filtration

$$B^r\mathbb{L}(H) = \{w \in \mathbb{L}(\mathsf{a},\mathsf{b}) : \deg_{\mathsf{b}} w \geqslant r\}$$

associated to the B-degree on $\mathbb{L}(\mathsf{a},\mathsf{b})$, which is defined to be the degree in b. It induces a decreasing filtration $B^{\bullet}$ on $\mathrm{Der}^{\Theta}\, \mathbb{L}(\mathsf{a},\mathsf{b})$, which is the filtration associated to the grading by B-degree. It satisfies $\mathrm{Der}^{\Theta}\, \mathbb{L}(\mathsf{a},\mathsf{b}) = B^{-1}\,\mathrm{Der}^{\Theta}\, \mathbb{L}(\mathsf{a},\mathsf{b})$. The subspace $B^0\,\mathrm{Der}^{\Theta}\, \mathbb{L}(\mathsf{a},\mathsf{b})$ is the space of derivations δ such that

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

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

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

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

which are uniquely determined by the property

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

The elements $\varepsilon^\vee_{2n+2}$ were first defined by Tsunogai [37]. The element $\varepsilon^\vee_2$ is central in $\mathrm{Der}^\Theta\,\mathbb{L}(\mathsf{a},\mathsf{b})$ and plays no role here. Let

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

be the Lie subalgebra spanned by the $\varepsilon^\vee_{2n+2}$, for $n \geqslant -1$. It contains $\varepsilon^\vee_0 = \mathsf{b}\partial/\partial\mathsf{a}$. One shows that the 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 $\mathbb{L}(\mathsf{a},\mathsf{b})$ via $\mathfrak{u}^{\mathrm{geom}}$ [24]. Quadratic relations between the elements $\varepsilon^\vee_{2n}$ predicted by Hain and Matsumoto were studied by Pollack in his thesis [34].

We define heretical normalizations of these derivations as follows. Let

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

where B_k denotes the kth Bernoulli number.

3.3. The Hain morphism. There is a natural morphism [23, Section 16–18]:

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

Using the work of Levin and Racinet, Hain computed this map in the de Rham realization [23], (18.1). On de Rham Lie algebras it is the continuous morphism

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

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

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

$$\widetilde{\sigma} \circ \phi = \phi \circ \sigma \tag{3.7}$$

which is equivalent to the equation $\widetilde{\sigma}\phi(\mathsf{x}_0) = \phi\sigma(\mathsf{x}_0)$. An element of $\mathrm{Der}^\Theta\,\mathbb{L}(\mathsf{a},\mathsf{b})^\wedge$ descends to an element of $\mathrm{Der}^1\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)^\wedge$ if and only if it

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

preserves the subspace $\phi(\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)^\wedge)$ in $\mathbb{L}(\mathsf{a},\mathsf{b})^\wedge$. Since (3.6) is injective, σ and $\widetilde{\sigma}$ determine each other uniquely. Furthermore, lifting derivations respects the Lie brackets: if $\sigma_1,\sigma_2\in \mathrm{Der}^1\,\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)^\wedge$ admit lifts $\widetilde{\sigma}_1,\widetilde{\sigma}_2\in \mathrm{Der}^\Theta\,\mathbb{L}(\mathsf{a},\mathsf{b})^\wedge$, 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 realization is [23], 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,

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

where i_0 was defined in (2.4).

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

Proof. The map i is injective by [4], and its image is contained in $D^1\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$. The map (2.5) restricts to an injective map on $D^1\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$, and hence their composition i_0 is injective. The morphism ϕ is injective (for example, because its associated graded (6.2) is injective). Since i_1 is the lift of i_0, it follows that i_1 is injective. □

3.4. *B*-filtration. We show that the B-filtration defined above on the bigraded Lie algebra of $\mathcal{P}$ is in fact induced by a natural filtration on the Lie algebra of $\mathcal{P}$, which we shall also call the B-filtration.

LEMMA 3.2. *Consider the filtration defined by the convolution of the Hodge filtration F and the lower central series filtration L:*

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

It induces a decreasing filtration on the Lie algebra of $\mathcal{P}$, which coincides with the B-filtration on its associated bigraded for M, W.

Proof. Note that $L^b=W_{-b}$, where W denotes the weight filtration. Since the filtrations F, W, and M can be split simultaneously [23], $F\star L$ induces a filtration on $\mathbb{L}(H)=\mathrm{gr}^W\mathrm{gr}^M\mathrm{Lie}\,\mathcal{P}$. We must check that it coincides with the B-filtration. To see this, note that $B^r\mathrm{gr}^n_L\mathbb{L}(H_{\mathrm{dR}})=F^{r-n}\mathrm{gr}^n_L\mathbb{L}(H_{\mathrm{dR}})$. Since $\mathrm{gr}^n_L\mathbb{L}(H_{\mathrm{dR}})$ consists of words of length n in a and b, and since

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

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

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

Note also that $2\deg_b + 2\deg_W = \deg_M$ in the bigraded Lie algebra of $\mathcal{P}$.

LEMMA 3.3. *The B-filtration on $\mathcal{P}$ is motivic, that is, it is stable under the image of the de Rham motivic Galois group $i_1(\mathfrak{g}^{\mathfrak{m}})$.*

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

We now show that the B-filtration gives one possible way to cut out the depth filtration on the image of $\mathfrak{g}^{\mathfrak{m}}$ inside $\operatorname{Der}^{\Theta} \mathbb{L}(\mathsf{a},\mathsf{b})$.

LEMMA 3.4. *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

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

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

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

LEMMA 3.5. *Let $\delta \in B^0 \operatorname{Der}^{\Theta} \mathbb{L}(\mathsf{a},\mathsf{b})$. The following are equivalent:*

(1) $\delta \in B^r \operatorname{Der}^{\Theta} \mathbb{L}(\mathsf{a},\mathsf{b})$;

(2) $\delta(\mathsf{a}) \in B^r\, \mathbb{L}(\mathsf{a},\mathsf{b})$;

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

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

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

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

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

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

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

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

PROPOSITION 3.6. *Let* $\sigma \in \mathrm{Der}^1\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ *which lifts to* $\widetilde{\sigma} \in B^0\,\mathrm{Der}^{\Theta}\,\mathbb{L}(\mathsf{a}, \mathsf{b})$. *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 $\widetilde{\sigma} \in B^r$. $\square$

COROLLARY 3.7. *The* B*-filtration cuts out the depth filtration on the image of* $\mathfrak{g}^{\mathfrak{m}}$*:*

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

Proof. Apply Lemma 3.3, the previous proposition, and $i_1\phi = \phi\, i_0$ (3.8). $\square$

By considering the symmetry $t \mapsto 1 - t$ on $\mathbb{P}^1\backslash\{0, 1, \infty\}$ (duality relation), one knows that an element $\sigma \in \mathfrak{g}^{\mathfrak{m}}$ of T-degree $-m$ is uniquely determined by its image in $D^1/D^{\lceil (m+1)/2\rceil}\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ under the map i. It follows from Theorem 3.1 that

COROLLARY 3.8. *An element* $\sigma \in \mathfrak{g}^{\mathfrak{m}}$ *of* T*-degree* $-m$ *is uniquely determined by the image of* $i_1(\sigma)$ *in* $B^1/B^{\lceil (m+1)/2\rceil}\mathrm{Der}^{\Theta}\mathbb{L}(\mathsf{a}, \mathsf{b})$.

REMARK 3.9. It follows that the 'tails' of the generators $i_1(\sigma)$ in the B-filtration, which could be infinitely long, are uniquely determined from their 'necks' $i_1(\sigma)$ (mod $B^{\lceil (m+1)/2\rceil}$). Furthermore, one can show (for instance, using the description [8, Section 9] of the group of automorphisms of a semidirect product) that, modulo B^m, $i_1(\sigma)$ is equivalent to an element of $\mathfrak{u}^{\mathrm{geom}}$, so can be expressed (nonuniquely) in terms of the geometric derivations $\varepsilon^{\vee}_{2n+2}$ for $n \geqslant -1$. It is by no means true, however, that the $i_1(\sigma)$ are contained in $\mathfrak{u}^{\mathrm{geom}}$.

For instance, a choice of zeta element of T-degree $-2n-1$ admits an expansion

$$i_1(\sigma_{2n+1}) \equiv \sum_{a_1,\ldots,a_r} \lambda^{(2n+1)}_{a_1,\ldots,a_r}\left[\varepsilon^{\vee}_{a_1}, \left[\varepsilon^{\vee}_{a_2}, \left[\ldots, \left[\varepsilon^{\vee}_{a_{r-1}}, \varepsilon^{\vee}_{a_r}\right]\cdots\right]\right.\right. \pmod{B^{2n+1}}$$

where $\lambda^{(2n+1)}_{a_1,\ldots,a_r} \in \mathbb{Q}$, and $a_i \geqslant 0$ are even. The nonuniqueness of this expansion corresponds to the fact that there exist relations between the generators $\varepsilon^{\vee}_{2n+2}$ in $\mathfrak{u}^{\mathrm{geom}}$ [34]. Nonetheless, an explicit formula for this geometric expansion for the elements $i_1(\sigma^c_{2n+1})$, for $n \geqslant 2$, will be determined modulo B^4 below.

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

4. Commutative power series

We discuss commutative power series representations for the Lie algebras $\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$, $\mathbb{L}(\mathsf{a},\mathsf{b})$ 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

$$\begin{aligned}\rho : \mathrm{gr}^r_D T(\mathsf{x}_0,\mathsf{x}_1) &\overset{\sim}{\longrightarrow} \mathbb{Q}[y_0,\ldots,y_r] \\ \mathsf{x}_0^{i_0}\mathsf{x}_1\mathsf{x}_0^{i_1}\ldots\mathsf{x}_1\mathsf{x}_0^{i_r} &\mapsto y_0^{i_0}\ldots y_r^{i_r},\end{aligned} \tag{4.1}$$

where we recall that the D-degree is the degree in x_1. Let us denote by

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

Since $\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$ is D-graded, (4.1) induces a map $\rho : \mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)\rightarrow P$. We have

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

Furthermore, by [7, Lemma 6.2], the image of $\mathrm{gr}^r_D\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$, for $r\geqslant 1$, is contained in the subspace of polynomials which are translation invariant:

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

Such a polynomial f is uniquely determined by its image in

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

We call this the reduced representation of a translation-invariant polynomial, and it applies equally well to translation-invariant rational functions. The variables x_n will be reserved for the reduced representations. Since $\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)\cong \mathrm{gr}^{\bullet}_D\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$ is graded with respect to the D-degree, there is an injection

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

One can describe the image of this map in terms of double shuffle equations (see below). The $\sigma^{(r)}$ will be called the depth r components of (the polynomial representation of) σ. The zeta elements satisfy $\sigma^{(1)}_{2n+1} = x_1^{2n}$, for $n\geqslant 1$ by (4.2).

4.2. Concatenation products. The concatenation of words in the alphabet x_0, x_1 defines a noncommutative multiplication law $f, g \mapsto f \cdot g$:

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

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

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

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

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

Note that equation (4.6) is equivalent to the condition that the linear map

$$w \mapsto f \underline{\circ} w - f \cdot w$$

is a derivation with respect to $\cdot$. There is another concatenation product, denoted by $\underline{\cdot}$, which comes from the stuffle Hopf algebra [3]. It will only be used once in this paper, so will not be discussed in any detail here.

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

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

whose r, s component is given explicitly by the formula

$$\begin{aligned} & f \underline{\circ} g\, (y_0, \ldots, y_{r+s}) \\ &\quad = \sum_{i=0}^{s} f(y_i, y_{i+1}, \ldots, y_{i+r}) g(y_0, \ldots, y_i, y_{i+r+1}, \ldots, y_{r+s}) \\ &\qquad + (-1)^{\deg f + r} \sum_{i=1}^{s} f(y_{i+r}, \ldots, y_{i+1}, y_i) g(y_0, \ldots, y_{i-1}, y_{i+r}, \ldots, y_{r+s}). \end{aligned} \tag{4.8}$$

This can be read off from equation (2.7). Antisymmetrizing gives a pairing

$$\{f, g\} = f \underline{\circ} g - g \underline{\circ} b$$

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

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 $f \,\underline{\circ}\, g$ and $\{f, g\}$.

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

$$\underline{\circ} : \mathbb{Q}(y_0, \ldots, y_r) \otimes_{\mathbb{Q}} \mathbb{Q}(y_0, \ldots, y_s) \longrightarrow \mathbb{Q}(y_0, \ldots, y_{r+s}),$$

and by restricting to translation-invariant rational functions, we obtain a map on their reduced versions $\underline{\circ} : \mathbb{Q}(x_1, \ldots, x_r) \otimes_{\mathbb{Q}} \mathbb{Q}(x_1, \ldots, x_s) \to \mathbb{Q}(x_1, \ldots, x_{r+s})$.

EXAMPLE 4.1. Let $r = s = 1$ and $f, g \in \mathbb{Q}(x_1)$. Then $\{f, g\} \in \mathbb{Q}(x_1, x_2)$ equals

$$f(x_1)g(x_2) - g(x_1)f(x_2) + f(x_2 - x_1)(g(x_1) - g(x_2)) + (f(x_2) - f(x_1))g(x_2 - x_1).$$

In the following sections we shall work in the graded vector space

$$RP = \bigoplus_{r \geqslant 0} \mathbb{Q}(y_0, \ldots, y_r)$$

and its translation-invariant subspace $\overline{RP} = \bigoplus_{r \geqslant 0} \mathbb{Q}(x_1, \ldots, x_r)$. The map $\underline{\circ}$ and its antisymmetrization $\{\,,\,\}$ define linear maps

$$\underline{\circ}, \{\,,\,\} : RP \otimes_{\mathbb{Q}} RP \longrightarrow RP \quad \text{and} \quad \underline{\circ}, \{\,,\,\} : \overline{RP} \otimes_{\mathbb{Q}} \overline{RP} \longrightarrow \overline{RP}$$

via the formulae $(\alpha \,\underline{\circ}\, \beta)^{(n)} = \sum_{i+j=n} \alpha^{(i)} \,\underline{\circ}\, \beta^{(j)}$ and $\{\alpha, \beta\} = \alpha \,\underline{\circ}\, \beta - \beta \,\underline{\circ}\, \alpha$. The equation (2.9) automatically extends to the setting of rational functions.

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

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

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

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

viewed inside RP. Using (4.1), define a linear map

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

where $\rho^{(r)}$ denotes the depth r component of ρ.

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

PROPOSITION 4.2. *The following diagram commutes:*

$$\begin{array}{ccccccc} & D^1\mathrm{Der}^1\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1) & \times & (\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1),[\ ,\ ]) & \longrightarrow & (\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1),[\ ,\ ]) \\ & \downarrow_{\rho'} & & \downarrow_{\rho} & & \downarrow_{\rho} \\ \odot: & P' & \times & P & \longrightarrow & RP \end{array}$$

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

$$f \odot g = f \,\underline{\circ}\, g - f \cdot g. \tag{4.11}$$

The composition of derivations is given by the linearized Ihara bracket:

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

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 the operator $\odot$, defined by formula (4.11), 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 $w \in D^1\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$, and $i = 0, 1$, we have

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

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

$$f(y_0, \ldots, y_r) \,\underline{\circ}\, y_0 = y_0 f(y_0, \ldots, y_r)$$
$$f(y_0, \ldots, y_r) \,\underline{\circ}\, 1 = f(y_0, \ldots, y_r)$$

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 a manipulation of (4.11), (4.6) and (2.9). $\square$

The following corollary is not essential for the remainder of this paper, but provides an interpretation of the map $f \mapsto \{x_1^{-1}, f\}$.

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*

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

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

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

Define

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

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

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

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

Denote also by ρ the linear map

$$\begin{aligned} \rho^{(r)} : \mathrm{gr}^r_B\, T(\mathsf{a}, \mathsf{b}) &\xrightarrow{\sim} \mathbb{Q}[y_0, \dots, y_r] \\ \mathsf{a}^{i_0}\mathsf{b}\mathsf{a}^{i_1}\dots\mathsf{b}\mathsf{a}^{i_r} &\mapsto y_0^{i_0}\dots y_r^{i_r}. \end{aligned} \tag{4.14}$$

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

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

Define a graded vector space

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

and a linear map

$$\begin{aligned} \ell : \mathbb{L}(\mathsf{a}, \mathsf{b}) &\longrightarrow Q \\ w &\mapsto \sum_r \ell_r^{-1}\rho^{(r)}(w). \end{aligned} \tag{4.16}$$

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

$$Q' = \bigoplus_{r \geqslant 1} Q_r \frac{1}{y_0 - y_r} \tag{4.17}$$

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

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

$$\begin{aligned} \ell' : D^{\Theta} &\longrightarrow Q' \\ \delta &\mapsto \sum_{r\geqslant 1} c_r^{-1}\rho^{(r)}\delta(\mathsf{a}). \end{aligned} \tag{4.18}$$

It is injective, since $\delta(\mathsf{a})$ uniquely determines $\delta \in D^{\Theta}$. Note that $Q, Q' \subset RP$.

PROPOSITION 4.4. *The following diagram commutes:*

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

where, as in (4.11), we have

$$f \circledast g = f \,\underline{\circ}\, g - f \cdot g. \tag{4.19}$$

Similarly, we have a commutative diagram

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

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

$$\ell'([\delta_1, \delta_2]) = \{\ell'(\delta_1), \ell'(\delta_2)\}. \tag{4.20}$$

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

Proof. The proof is similar to the proof of Proposition 4.2. We must check that $\ell'(\delta) \circledast \ell([\mathsf{a},\mathsf{b}]) = 0$ and $\ell'(\delta) \circledast \ell(\mathsf{a}) = \ell(\delta(\mathsf{a}))$. But $\ell([\mathsf{a},\mathsf{b}]) = (y_0 - y_1)/(y_0 - y_1) = 1$ and $\ell(\mathsf{a}) = y_0$, so this calculation is formally identical to the one in Proposition 4.2. The commutativity of the second diagram follows from the first, using the fact that an element in $B^1\mathrm{Der}^{\Theta}\mathbb{L}(\mathsf{a},\mathsf{b})$ is uniquely determined by its image in D^{Θ}, and the fact that $\circledast$ is a derivation for the shuffle concatenation product. $\square$

Derivations in D^{Θ} cannot be extended to derivations on $\mathbb{L}(\mathsf{a},\mathsf{b})$. But the two commutative diagrams in Proposition 4.4 show that, passing to rational function representations, namely Q', enables us to do precisely that.

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

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 [3]. We shall only require their restriction to depths $\leqslant 3$ and work with translation-invariant representations (4.3). Let

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

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

$$f^{(2)}(x_1, x_{12}) + f^{(2)}(x_2, x_{12}) = 0 \tag{4.21}$$
$$f^{(3)}(x_1, x_{12}, x_{123}) + f^{(3)}(x_2, x_{12}, x_{123}) + f^{(3)}(x_2, x_{23}, x_{123}) = 0.$$

Note that starting from depth four there will be several such equations in each depth. It is straightforward to show [3] that the solutions to the shuffle equations modulo products correspond, via (4.1) and (4.3), 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 (regularized versions of) the equations ($a, b, c \in \mathbb{N}$):

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

From depth four onwards, there are more than one such equation in each depth. By considering the series $Z^{(r)} = \sum_{n_1,\ldots,n_r \geqslant 0} \zeta_*(n_1, \ldots, n_r) x_1^{n_1-1} \ldots x_r^{n_r-1}$, where the subscript $*$ denotes the stuffle regularization, these equations translate into

$$f^{(2)}(x_1, x_2) + f^{(2)}(x_2, x_1) = \frac{f^{(1)}(x_1) - f^{(1)}(x_2)}{x_2 - x_1} \tag{4.22}$$

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

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 [3]. The Lie algebra $\mathfrak{dmr}_0$ is defined to be the sets of solutions to both shuffle and stuffle equations modulo products.

The *linearized double shuffle equations* [7] are the same sets of equations in which the right-hand sides are zero. The linearized shuffle equations are identical

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

to the ordinary shuffle equations, but the linearized stuffle equations are:

$$
\begin{aligned}
f^{(1)}(x_1) + f^{(1)}(-x_1) &= 0 \\
f^{(2)}(x_1, x_2) + f^{(2)}(x_2, x_1) &= 0 \\
f^{(3)}(x_1, x_2, x_3) + f^{(3)}(x_2, x_1, x_3) + f^{(3)}(x_2, x_3, x_1) &= 0.
\end{aligned} \tag{4.23}
$$

The linearized double shuffle equations are also closed under the Ihara bracket, by a (simpler) version of Racinet's theorem [35].

4.7. Geometric derivations and linearized double shuffle with poles. The Lie algebra $\mathfrak{u}^{\mathrm{geom}} \subset B^1\mathrm{Der}^{\Theta}\mathbb{L}(\mathsf{a},\mathsf{b})$ of geometric derivations was defined in (3.3). We shall identify $\mathfrak{u}^{\mathrm{geom}}$ with its associated B-graded. The definition of the linearized double shuffle equations (4.21) and (4.23) can be extended in the obvious way to rational functions.

DEFINITION 4.5. Define $\mathfrak{pls} \subset Q'$ to be the subspace of Q' of translation-invariant rational functions (4.17) whose reduced versions (image under (4.3)) satisfy the linearized double shuffle equations. It is bigraded by weight and depth.

Denote the reduced version of $\mathfrak{pls}$ by $\overline{\mathfrak{pls}}$. By a version of Racinet's theorem, $\mathfrak{pls}$ is also a Lie subalgebra of Q' for the linearized Ihara bracket $\{\, ,\}_\bullet$ (antisymmetrization of $\underline{\circ}$). The notation stands for 'polar linearized double shuffle' solutions. It is a bigraded Lie algebra in the category of $\mathfrak{sl}_2$-representations over $\mathbb{Q}$ (see appendix).

PROPOSITION 4.6. *The geometric derivations, in their rational function representation* (4.18)*, satisfy the linearized double shuffle equations:*

$$\ell'(\mathfrak{u}^{\mathrm{geom}}) \subset \mathfrak{pls}.$$

Proof. The images of the generators $\overline{\ell'}(\varepsilon^{\vee}_{2n+2}) = x_1^{2n}$ by (4.2) for $n \geqslant -1$. They are even and hence solutions to the linearized double shuffle equations. It follows from (4.20) that $\overline{\ell'}$ is a morphism of Lie algebras as $\mathfrak{pls}$ is closed under $\{\, ,\}_\bullet$. $\square$

A natural question to ask is whether the $\mathfrak{sl}_2$-equivariant map $\ell' : \mathfrak{u}^{\mathrm{geom}} \to \mathfrak{pls}$ is an isomorphism. It is true in depths $\leqslant 3$ (see appendix), and also in certain limits with respect to the residue filtration [3].

The previous proposition implies that the linearized stuffle equations define maps from the space of nongeometric derivations

$$B^1\mathrm{Der}^{\Theta}\,\mathbb{L}(\mathsf{a},\mathsf{b})/\mathfrak{u}^{\mathrm{geom}}$$

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

to spaces of rational functions. The map is as follows. For $\delta \in \mathrm{gr}^r_B \mathrm{Der}^{\Theta}\mathbb{L}(\mathsf{a}, \mathsf{b})$, compute the image of $\overline{\ell}'(\delta)$ under any stuffle equation of depth r. It defines an element in $\overline{RP}$, which is zero if $\delta \in \mathfrak{u}^{\mathrm{geom}}$, by the previous proposition. We shall show in Remark 5.8 that this map is nonzero and provides a new tool to prove that certain derivations are not geometric. It could in particular be used to study the image of $i_1(\mathfrak{g}^{\mathfrak{m}})$ in $B^1\mathrm{Der}^{\Theta}\,\mathbb{L}(\mathsf{a}, \mathsf{b})/\mathfrak{u}^{\mathrm{geom}}$.

5. Zeta 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 (4.21) and (4.22).

5.1. Polar solutions. The shape of the equations (4.22) suggests searching for solutions amongst the space of rational functions in x_i with $\mathbb{Q}$-coefficients. Let

$$s^{(1)} = \frac{1}{2\,x_1} \quad \text{and} \quad s^{(2)} = \frac{1}{6}\Big(\frac{1}{x_1x_2} + \frac{1}{x_2(x_1 - x_2)}\Big). \tag{5.1}$$

It is easy to verify that $(s^{(1)}, s^{(2)})$ is a solution to the double shuffle equations (4.21) and (4.22) in depths one and two.

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

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

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

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

which defines a polynomial in $\mathbb{Q}[x_1, x_2]$ whenever $n \geqslant 0$. On the other hand, $\xi^{(3)}_{2n+1}$ is a rational function in x_1, x_2, x_3 with nontrivial poles. When $n \geqslant 0$ 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 $\xi^{(2)}_1 = \xi^{(3)}_1 = 0$.

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

PROPOSITION 5.2. *Let $n \geqslant -1$. The elements*

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

satisfy the double shuffle equations modulo products (4.21), (4.22) *in depths* 2, 3.

Proof. This is a straightforward 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. □

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 (that is, whose depth i components $s^{(i)}$ are homogeneous rational functions of degree $-i$). This is discussed in [3], 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. It is given by $s^{(3)} = \{s^{(1)}, s^{(2)}\}_{\bullet}$.

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

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

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

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

Recall the well-known functional identity

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

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

$$\rho(\underline{\sigma}^c_{2n+1}) = \underline{\xi}_{2n+1} + \sum_{a+b=n} \frac{1}{2b}\{\underline{\xi}_{2a+1}, \{\underline{\xi}_{2b+1}, \underline{\xi}_{-1}\}\} \pmod{D^4} \tag{5.7}$$

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

$$\sigma^c_{2n+1} = \frac{(2n)!}{B_{2n}}\underline{\sigma}^c_{2n+1} \quad \text{for } n \geqslant 2$$

to be the canonical normalizations, and set $\sigma^c_3 = [\mathsf{x}_0, [\mathsf{x}_0, \mathsf{x}_1]] + [\mathsf{x}_1, [\mathsf{x}_0, \mathsf{x}_1]]$.

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

By Racinet's theorem, the space of solutions to the double shuffle equations is closed under the Ihara bracket, so for $n \geqslant 2$, the elements σ^c_{2n+1} are solutions to the double shuffle equations in depths $\leqslant 3$, by Proposition 5.2.

REMARK 5.5. The σ^c_{2n+1} have the canonical normalization. For $n \geqslant 1$,

$$(\sigma^c_{2n+1})^{(1)} = \xi^{(1)}_{2n+1} = x_1^{2n} \quad \text{and} \quad (\sigma^c_{2n+1})^{(2)} = \xi^{(2)}_{2n+1}$$

is given by (5.3). We have $(\sigma^c_3)^{(3)} = 0$ and for $n \geqslant 2$,

$$(\sigma^c_{2n+1})^{(3)} = \xi^{(3)}_{2n+1} + \sum_{a+b=n} \frac{B_{2a}B_{2b}}{B_{2n}} \binom{2n}{2a} \frac{1}{24\,b} \{x_1^{2a}, \{x_1^{2b}, x_1^{-2}\}_\bullet\}_\bullet. \tag{5.8}$$

The previous expression can be written more symmetrically in terms of lowest-weight vectors for the action of $\mathfrak{sl}_2$ (see appendix), namely:

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

On the other hand, compare the odd part of the period polynomial [27] for the Eisenstein series of weight $2n$, which is proportional to:

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

This is no accident, and follows from the computations in Section 5.3 as well as Section 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

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

In the notation of [3] it is $f \star g = f \circ g - g \,\underline{\cdot}\, f$, where $\underline{\cdot}$ is the 'stuffle concatenation'.

LEMMA 5.6. *For all $n, a, b \geqslant 1$,*

$$\operatorname{Res}_{x_3=0}\big(\xi^{(3)}_{2n+1}\big) = \tfrac{1}{12} x_1^{2n} \star x_1^{-1} \tag{5.9}$$

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

Proof. This is a straightforward computation and follows from the definitions. The second equation easily generalizes [3]. □

PROPOSITION 5.7. *The elements* $(\sigma^c_{2n+1})^{(3)}$ *have no poles, for all* $n \geqslant 2$.

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

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

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

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

where $b(x)$ was defined in (5.5). To compute this, 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 $n \geqslant 2$. The above argument proves that the $(\sigma^c_{2n+1})^{(3)}$ have no poles along $x_3 = 0$. Now we use the fact that σ^c_{2n+1} satisfies the double shuffle equations modulo products in depths two and three. Since $(\sigma^c_{2n+1})^{(i)}$ has no poles for $i = 1, 2$, the stuffle equation (4.22) implies that

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

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

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

By taking the residue of this expression at $x_2 = 0$, we deduce that $(\sigma^c_{2n+1})^{(3)}$ 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. $\square$

The last part of this argument can be generalized using the dihedral symmetry structure of the linearized double shuffle equations [7], Section 6.3.

REMARK 5.8. The element $(\sigma^c_3)^{(3)}$ does have poles. It is equal to $3z_3$, where

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

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

and corresponds to a lift to $\mathrm{Der}^{\Theta}\,\mathbb{L}(\mathsf{a},\mathsf{b})$ of the 'arithmetic image' of the element $i_1(\sigma_3)$ in $(\mathrm{Der}^{\Theta}\,\mathbb{L}(\mathsf{a},\mathsf{b}))/\mathfrak{u}^{\mathrm{geom}}$. The corresponding derivation was written down in [34]. Computing the stuffle equation (4.23) 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 nonzero, and shows, by Proposition 4.6, that z_3 is, as expected, nongeometric, that is, not in the image of $\mathfrak{u}^{\mathrm{geom}}$.

5.4. Zeta elements in depth three.

THEOREM 5.9. *The elements σ^c_{2n+1} are in the image of the map*

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

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

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

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

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

This is trivial for $d=1$, and follows from a computation of the dimensions of $\mathfrak{ls}^d_n$ obtained by Zagier [38] for $d=2$, and by Goncharov [21, Theorem 1.5] for $d=3$. □

REMARK 5.10. It follows from the depth-parity theorem that the elements σ^c_{2n+1} 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. Zeta elements in depth 3 via geometric derivations

Recall the notations from Section 3.

THEOREM 6.1. *For all $n\geqslant 2$, we have an explicit expansion*

$$i_1(\underline{\sigma}^c_{2n+1})\equiv\underline{\varepsilon}^{\vee}_{2n+2}+\sum_{a+b=n}\frac{1}{2b}\big[\underline{\varepsilon}^{\vee}_{2a+2},\big[\underline{\varepsilon}^{\vee}_{2b+2},\underline{\varepsilon}^{\vee}_{0}\big]\big]\ (\mathrm{mod}\ B^4).$$

For $n=1$, $i_1(\sigma^c_3)=\varepsilon^{\vee}_4+3z_3$ (mod B^4), where z_3 is defined in Remark 5.8.

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

In terms of the standard normalizations, this equation is equivalent to

$$i_1(\sigma^c_{2n+1}) \equiv \varepsilon^\vee_{2n+2} + \sum_{a+b=n} \frac{B_{2a}B_{2b}}{B_{2n}} \binom{2n}{2a} \frac{1}{24b} [\varepsilon^\vee_{2a+2}, [\varepsilon^\vee_{2b+2}, \varepsilon^\vee_0]] \pmod{B^4}.$$

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

$$\mathrm{lw}_{a,b} = \frac{1}{2b}[\underline{\varepsilon}^\vee_{2a+2}, [\underline{\varepsilon}^\vee_{2b+2}, \underline{\varepsilon}^\vee_0]] + \frac{1}{2a}[\underline{\varepsilon}^\vee_{2b+2}, [\underline{\varepsilon}^\vee_{2a+2}, \underline{\varepsilon}^\vee_0]].$$

These are lowest-weight vectors for the action of $\mathfrak{sl}_2$, that is, $\varepsilon_0(\mathrm{lw}_{a,b}) = 0$, where ε_0 is the derivation on $\mathbb{L}(\mathsf{a}, \mathsf{b})$ such that $\varepsilon_0(a) = 0$ and $\varepsilon(b) = a$. This yields a direct comparison with the period polynomials of Eisenstein series [8, Section 7.3].

The strategy for the proof is as follows. The elements $\varepsilon^\vee_{2n+2}$ 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 Section 4.1 and enlarge this space by introducing poles, then the elements $\varepsilon^\vee_{2n+2}$, 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': that is, by lifting the analogous result for derivations on the de Rham fundamental groupoid of $\mathbb{P}^1\backslash\{0, 1, \infty\}$ via the Hain morphism. 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. We shall apply the method of commutative power series Sections 4.1, 4.5 to both $\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ and $\mathbb{L}(\mathsf{a}, \mathsf{b})$.

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

$$\phi^r : \mathrm{gr}^\bullet_D\, T(\mathsf{x}_0, \mathsf{x}_1) \longrightarrow \mathrm{gr}^{\bullet+r}_B\, T(\mathsf{a}, \mathsf{b}).$$

Translating into rational functions Section 4, and using the fact that $\rho : \mathrm{gr}^\bullet_D T(\mathsf{x}_0, \mathsf{x}_1) \cong P$ is an isomorphism, we obtain a commutative diagram

$$\begin{array}{ccc} \mathrm{gr}^\bullet_D\, T(\mathsf{x}_0, \mathsf{x}_1) & \xrightarrow{\phi^r} & \mathrm{gr}^{\bullet+r}_B\, T(\mathsf{a}, \mathsf{b}) \\ \downarrow_\rho & & \downarrow_\ell \\ P & \xrightarrow{\phi_r} & Q. \end{array} \tag{6.1}$$

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):

$$\phi^0 = \mathrm{gr}\,\phi : \begin{cases} \mathsf{x}_0 \mapsto \mathsf{a}, \\ \mathsf{x}_1 \mapsto [\mathsf{a}, \mathsf{b}], \end{cases} \tag{6.2}$$

and via $T(\mathsf{x}_0,\mathsf{x}_1) = \mathrm{gr}_D T(\mathsf{x}_0,\mathsf{x}_1)$ and $T(\mathsf{a},\mathsf{b}) = \mathrm{gr}_B T(\mathsf{a},\mathsf{b})$, we have $\phi = \sum_{r\geqslant 0}\phi^r$. The idea of the following discussion is that although the map of Lie algebras $\phi^0+\phi^1+\phi^2 : \mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)\to\mathbb{L}(\mathsf{a},\mathsf{b})$ does not factorize through ϕ^0, its rational function representation $\phi_{\leqslant 2} = \phi_0+\phi_1+\phi_2$ will in fact factorize through ϕ_0:

$$\begin{array}{cccccc} & \mathbb{L}(\mathsf{x}_0,\mathsf{x}_1) & \xrightarrow{\phi^0} & \mathbb{L}(\mathsf{a},[\mathsf{a},\mathsf{b}]) & & \mathbb{L}(\mathsf{a},\mathsf{b}) \\ & \downarrow_{\rho} & & \downarrow_{\ell} & & \downarrow_{\ell} \\ \phi_{\leqslant 2}: & P & \xrightarrow{\phi_0} & Q & \xrightarrow{?} & Q \end{array}$$

After computing ϕ_0, we shall express the map ? using Proposition 4.4. It will be an exponential, with respect to $\circledast$, of the element $-s^{(1)}-s^{(2)}$, modulo depth $\geqslant 3$.

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 $\mathsf{x}_0\to\mathsf{a}$, $\mathsf{x}_1\to\mathsf{b}$, followed by the composition of r maps, where the kth map, for $1\leqslant k\leqslant r$, replaces the kth occurrence of the letter b in $\mathsf{a}^{i_0}\mathsf{b}\mathsf{a}^{i_1}\mathsf{b}\ldots\mathsf{a}^{i_{r-1}}\mathsf{b}\mathsf{a}^{i_r}$ with $\mathsf{ab}-\mathsf{ba}$. 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 $\phi(\mathsf{x}_0) = \mathsf{a}+\frac{1}{2}[\mathsf{a},\mathsf{b}]+\cdots$ that it is a composition of ϕ^0 (6.2), whose image consists of words in a, $[\mathsf{a},\mathsf{b}]$, followed by the derivation in D^{Θ} (Section 4.5) which sends $\mathsf{a}\mapsto\frac{1}{2}[\mathsf{a},\mathsf{b}]$ and $[\mathsf{a},\mathsf{b}]$ to zero. Note that the latter does not extend to an element of $\mathrm{Der}^{\Theta}\,\mathbb{L}(\mathsf{a},\mathsf{b})$. It is nonetheless represented, via Proposition 4.4, by

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

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

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

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

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

where

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

This holds from the definition of ϕ (which is a homomorphism, not a derivation) since $\phi^2 + \frac{1}{2}s^{(1)}\phi^1$ is the composition of ϕ^0 followed by the derivation $\mathbb{L}(\mathsf{a}, [\mathsf{a}, \mathsf{b}]) \to \mathbb{L}(\mathsf{a}, \mathsf{b})$ which sends $\mathsf{a} \mapsto \frac{1}{12}[\mathsf{b}, [\mathsf{b}, \mathsf{a}]]$ and $[\mathsf{a}, \mathsf{b}]$ 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 $\ell([\mathsf{a}, \mathsf{b}]) = -s^{(1)} \circledast y_0$,

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

This can also be read off Corollary 4.3 upon writing $\phi^1 = \frac{1}{2}\phi^0 \circ \partial_{\mathsf{x}_1}$.

6.2. Proof of Theorem 6.1. Recall that an element $\sigma \in \mathrm{Der}^1\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ lifts to $\widetilde{\sigma} \in \mathrm{Der}^1\mathbb{L}(\mathsf{a}, \mathsf{b})$ if and only if the following equation holds in $\mathbb{L}(\mathsf{a}, \mathsf{b})$:

$$\widetilde{\sigma}\phi(\mathsf{x}_0) = \phi\sigma(\mathsf{x}_0).$$

Finding an element $\sigma \in \mathrm{Der}^1\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ whose lift is $\varepsilon^\vee_{2n+2}$, is equivalent via (3.6), modulo terms of B-degree $\geqslant 4$, to the following equation

$$\varepsilon^\vee_{2n+2}\big(\mathsf{a} + \tfrac{1}{2}[\mathsf{a}, \mathsf{b}] + \tfrac{1}{12}[\mathsf{b}, [\mathsf{b}, \mathsf{a}]]\big) \equiv \phi(\sigma(\mathsf{x}_0)) \pmod{B^4}.$$

It is easy to verify that it has no solution $\sigma \in \mathrm{Der}^1\,\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$. Note that $\varepsilon^\vee_{2n+2}([\mathsf{a}, \mathsf{b}]) = 0$ 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 $\ell'(\varepsilon^\vee_{2n+2}) = (y_1 - y_0)^{2n}$, and $\rho(a) = y_0$, $\rho(\mathsf{x}_0) = y_0$, it is equivalent to

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

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

$$\underbrace{\chi \circledast y_0}_{\phi_0} - \underbrace{\{s^{(1)}, \chi\} \circledast y_0 + \tfrac{1}{2}\{s^{(1)}, \{s^{(1)}, \chi\}\} \circledast y_0}_{\phi_1} - \underbrace{s^{(2)} \circledast \big(\chi \circledast y_0\big)}_{\phi_2} \pmod{B^4}.$$

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

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 depths 1, 2, 3 are the equations:

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

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 (4.19). This proves that Theorem 6.1 is equivalent to Theorem 5.9.

It would be interesting to generalize this picture to all higher depths.

7. Explicit rational associator in depths $\leqslant 3$

In Section 5 we wrote down explicit solutions to the double shuffle equations modulo products, in odd weights and depths $\leqslant 3$. The goal of this paragraph is to discuss solutions to the full double shuffle equations with even weights.

7.1. Double shuffle equations.

The full double shuffle equations in depth two are given by the pair of equations:

$$\begin{aligned}f^{(2)}(x_1, x_1 + x_2) + f^{(2)}(x_2, x_1 + x_2) &= f^{(1)}(x_1) f^{(1)}(x_2) \\ f^{(2)}_*(x_1, x_2) + f^{(2)}_*(x_2, x_1) &= \frac{f^{(1)}_*(x_1) - f^{(1)}_*(x_2)}{x_2 - x_1} \\ &\quad + f^{(1)}_*(x_1) f^{(1)}_*(x_2)\end{aligned} \tag{7.1}$$

where $f^{(1)}, f^{(1)}_* \in \mathbb{Q}[[x_1]]$ and $f^{(2)}, f^{(2)}_* \in \mathbb{Q}[[x_1, x_2]]$ are formal power series in commuting variables. A power series f without a subscript will denote its shuffle-regularized version; a subscript $*$ will denote its stuffle-regularized version. They differ by a factor which is well understood [35]. Our normalizations will be such that

$$f^{(1)}_* = f^{(1)} \quad \text{and} \quad f^{(2)}_* = f^{(2)} + \tfrac{1}{48}.$$

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

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

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

$$\begin{aligned} & f_*^{(3)}(x_1, x_2, x_3) + f_*^{(3)}(x_2, x_1, x_3) + f_*^{(3)}(x_2, x_3, x_1) \\ &= \frac{f_*^{(2)}(x_2, x_1) - f_*^{(2)}(x_2, x_3)}{x_3 - x_1} + \frac{f_*^{(2)}(x_1, x_3) - f_*^{(2)}(x_2, x_3)}{x_2 - x_1} \\ &\quad + f_*^{(1)}(x_1) f_*^{(2)}(x_2, x_3), \end{aligned}$$

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

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

The general principle [3] of constructing solutions to these equations with poles and correcting with counterterms 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

$$b_2(x_1, x_2) = \tfrac{1}{3}\big(b_1(x_1)b_1(x_2) + b_1(x_2)b_1(x_1 - x_2)\big).$$

With these definitions, set

$$\begin{aligned} 2\gamma^{(1)} &= -b_1 \\ 4\gamma^{(2)} &= -b_2 + \tfrac{1}{2} b_1 \,\underline{\circ}\, b_1 \\ 8\gamma^{(3)} &= b_2 \,\underline{\circ}\, b_1 - \tfrac{1}{6} b_1 \,\underline{\circ}\, (b_1 \,\underline{\circ}\, b_1). \end{aligned}$$

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

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

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

$$\gamma^{(3)}(x_1, x_{12}, x_{123}) + \gamma^{(3)}(x_2, x_{12}, x_{123}) + \gamma^{(3)}(x_2, x_{23}, x_{123}) = \gamma^{(1)}(x_1)\gamma^{(2)}(x_2, x_{23})$$
$$\gamma^{(3)}_*(x_1, x_2, x_3) + \gamma^{(3)}_*(x_2, x_1, x_3) + \gamma^{(3)}_*(x_2, x_3, x_1) = \gamma^{(1)}(x_1)\gamma^{(2)}_*(x_2, x_3)$$

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 from (5.6).

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

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

They have poles in x_i. More precisely, the element $d_r \times \Theta^{(r)}$ is 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 $1 \leqslant r \leqslant 3$.

For $1 \leqslant r \leqslant 3$, 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 Section 7.1 with

$$\Phi^{(2)}_* = \Phi^{(2)} + \tfrac{1}{48} \quad \text{and} \quad \Phi^{(3)}_* = \Phi^{(3)} + \tfrac{1}{48}\Phi^{(1)}(x_1).$$

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

$$C = \sum_{n\geqslant 1} \frac{1}{2n} \{\underline{\xi}_{-1}, \underline{\xi}_{2n+1}\} \tag{7.3}$$

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

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

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

A straightforward residue computation along the lines of Section 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 Section 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 $\leqslant 3$.*

A similar construction holds in depth four, but there is *a priori* no canonical way to cancel the poles: one must subtract counterterms consisting of quadruple brackets in the ξ_{2n+1}'s, which involves some choices because of quadratic relations amongst them (see Section 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 $\exp_{\underline{\circ}} \sigma^c_{2n+1}$ to obtain all other rational solutions to the double shuffle equations in depths $\leqslant 3$.

COROLLARY 7.2. *Every rational solution s to the full double shuffle equations in depths $\leqslant 3$ can be written explicitly in the form*

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

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 $\leqslant 3$.

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

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

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

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

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

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

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 $\leqslant 4$ and odd weight to rational numbers given by

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

where $2n+1 = n_1+\cdots+n_r$. 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 involving the period map. However, the maps (7.4) and (7.5) can be used as an algebraic substitute for the period map. Thus we obtain an exact algorithm to decompose any multiple zeta value of depth $\leqslant 3$ (and depth $\leqslant 4$ in the case of odd weight) into a chosen basis of multiple zeta values of the same or smaller depth. Ideally, one would like to generalize this to all weights and depths.

8. Cuspidal elements and the Broadhurst–Kreimer conjecture

We can recast the version of the Broadhurst–Kreimer conjecture [9] formulated in [7] using the elements σ^c_{2n+1}, first in $\mathrm{gr}_D\,\mathrm{Der}^1\,\mathbb{L}(\mathsf{x}_0,\mathsf{x}_1)$ and then in the elliptic setting in $\mathrm{gr}_B\,\mathrm{Der}^{\Theta}\,\mathbb{L}(\mathsf{a},\mathsf{b})$.

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

$$\overline{\sigma}_{2n+1} \in \mathfrak{d}^1_{2n+1} \quad \text{for all } n \geqslant 1. \tag{8.1}$$

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 $n \geqslant 0$ and let $V_n = \bigoplus_{i+j=n}\mathbb{Q}x_1^ix_2^j$ 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

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

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 $V_n/(\mathbb{Q}x_1^n \oplus \mathbb{Q}x_2^n)$.

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

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

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

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 $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ in $\varGamma$. Consider the following linear map from right $\varGamma$ group cochains [8, Section 2.3] to polynomials

$$f \mapsto \pi(f(S)) : Z^1_{\mathrm{cusp}}(\varGamma; V_{2n}) \longrightarrow V'_{2n} \tag{8.2}$$

where $Z^1_{\mathrm{cusp}}(\varGamma; V_n) \subset Z^1(\varGamma; V_n)$ 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

$$H^1_{\mathrm{cusp}}(\varGamma; V_{2n}) \xrightarrow{\sim} S_{2n+2}$$

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

$$H^1_{\mathrm{cusp}}(\varGamma; V_{2n})^+ \xrightarrow{\sim} S^+_{2n+2} \tag{8.3}$$

where the $+$ on the left-hand factor denotes invariants with respect to the action of the real Frobenius involution [8, Sections 5.4, 7.4]. The Eichler–Shimura theorem states in particular that integration defines an isomorphism:

$$S_{2n}(\varGamma) \xrightarrow{\sim} H^1_{\mathrm{cusp}}(\varGamma; V_{2n-2})^+ \otimes \mathbb{R}$$

where $S_{2n}(\varGamma)$ 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 (twice the T-degree) 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, that is,

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

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

is an exact sequence. Now embed $\mathfrak{g}^{\mathfrak{m}}$ in $\mathbb{L}(\mathsf{x}_0, \mathsf{x}_1)$ via (2.4). It follows that $\mathfrak{d} = \mathrm{gr}_D \mathfrak{g}^{\mathfrak{m}}$ 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 Section 4.1, we have a canonical isomorphism

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

We can thus identify $\mathfrak{d}^1 \otimes \mathfrak{d}^1 = x_1^2\mathbb{Q}[x_1^2] \otimes x_1^2\mathbb{Q}[x_1^2] \cong x_1^2x_2^2\mathbb{Q}[x_1^2, x_2^2]$, 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*

$$K_{2n} \xrightarrow{\sim} S^+_{2n}. \tag{8.4}$$

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

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

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

These quadratic relations appear in several contexts:

COROLLARY 8.4. *The elements $\varepsilon^\vee_{2n+2}$, ξ_{2n+1}, $\overline{\sigma}_{2n+1}$ and x_1^{2n} satisfy the identical quadratic relations. In other words, if $f_\bullet \in \{\varepsilon_{\bullet+2}, \xi_{\bullet+1}, \overline{\sigma}_{\bullet+1}\}$, then*

$$\sum \lambda_{ij}\{f_{2i}, f_{2j}\}_\bullet = 0 \Longleftrightarrow \sum \lambda_{ij} x_1^{2i} \wedge x_1^{2j} \in K.$$

Proof. The reduced polynomial representations of $\varepsilon^\vee_{2n+2}$ 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 either follows from their definition, because they are obtained from the x_1^{2n} via the Ihara bracket, or from the computations of Section 6 relating them to the $\varepsilon^\vee_{2n+2}$. $\square$

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 $x_1^2x_2^2(x_1^2 - x_2^2)^3 = x_1^8x_2^2 - 3x_1^6x_2^4 + 3x_1^4x_2^6 - x_1^2x_2^8$. It corresponds to the relations

$$\begin{aligned} [\overline{\sigma}_3, \overline{\sigma}_9] - 3[\overline{\sigma}_5, \overline{\sigma}_7] = 0, &\quad \{x_1^2, x_1^8\}_\bullet - 3\{x_1^4, x_1^6\}_\bullet = 0. \\ [\varepsilon^\vee_4, \varepsilon^\vee_{10}] - 3[\varepsilon^\vee_6, \varepsilon^\vee_8] = 0, &\quad \{\xi_3, \xi_9\} - 3\{\xi_5, \xi_7\} = 0. \end{aligned}$$

8.3. Cuspidal generators in depth 4. As explained in [7], 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 on K since the image of $\wedge^3\mathfrak{d}$ is in depth $\geqslant 3$. Composing with this map gives a linear map $d : K \to (\mathfrak{d}^4)^{ab}$ as we explain presently, and the canonical zeta elements defined in Section 5 give a means to compute it explicitly. To see this, the elements σ^c_{2n+1} can be interpreted as a linear map

$$\begin{aligned} \sigma^c : \mathfrak{d}^1 &\longrightarrow D^1\mathfrak{g}^{\mathfrak{m}}/D^4\mathfrak{g}^{\mathfrak{m}} \\ \overline{\sigma}_{2n+1} &\mapsto \sigma^c_{2n+1} \end{aligned}$$

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

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

The subspace K maps via (8.5) to $D^3/D^5\mathfrak{g}^{\mathfrak{m}}$, 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. Therefore, the restriction of (8.5) to K gives a linear map

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

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

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

These maps are closely related to the discussion in [31, Section 9].

THEOREM 8.5. *Let $P(x_1, x_2) = \sum_{i+j=n} \lambda_{i,j} x_1^{2i} x_2^{2j}$ be an even period polynomial of degree $2n$, where $\lambda_{i,j} = -\lambda_{j,i}$. Thus $P \in K$. It gives rise to a relation of the form*

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

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

$$\begin{aligned} \rho^{(4)}(\mathfrak{c}(P)) = &\sum_{i+a+b=n} \lambda_{i,a+b} \frac{B_{2a}B_{2b}}{B_{2a+2b}} \binom{2a+2b}{2a} \frac{1}{24b} \{x_1^{2i}, \{x_1^{2a}, \{x_1^{2b}, x_1^{-2}\}_\bullet\}_\bullet\}_\bullet \\ &- 3\lambda_{n-1,1}\{x_1^{2n-2}, z_3\}_\bullet \end{aligned} \tag{8.7}$$

where z_3 was defined in Remark 5.8.

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

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

Now substitute the expressions (5.8) for σ^c_{2j+1} 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

$$\Big(\sum_{i<j} \lambda_{i,j} \{\xi_{2i+1}, \xi_{2j+1}\}\Big)^{(r)} = 0 \quad \text{for } 1 \leqslant r \leqslant 3,$$

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

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

CONJECTURE 1 (Broadhurst–Kreimer: compare with [7, Section 9]).

$$\begin{aligned} H_1(\mathfrak{d}; \mathbb{Q}) &\cong \bigoplus_{n\geqslant 1} \overline{\sigma}_{2n+1}\mathbb{Q} \oplus \mathfrak{c}(K) \\ H_2(\mathfrak{d}; \mathbb{Q}) &\cong K \\ H_i(\mathfrak{d}; \mathbb{Q}) &= 0 \quad \textit{for all } i \geqslant 3. \end{aligned} \tag{8.8}$$

Thus $\mathfrak{d}$ admits the following conjectural presentation. It should have

(i) Generators: the $\overline{\sigma}_{2n+1}$ in depth 1 for $n \geqslant 1$, and $\mathfrak{c}(K)$ in depth 4.

(ii) Relations: the quadratic relations of Section 8.2.

REMARK 8.6. As noted in [16], $H_3(\mathfrak{d}; \mathbb{Q}) = 0$ implies that $H_i(\mathfrak{d}; \mathbb{Q}) = 0$ for all $i \geqslant 3$. In fact, for any pronilpotent 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 $n \geqslant i$. 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 $\mathrm{Ext}^n(k, M)$ in the category of $U\mathfrak{g}$-modules, and use the well-known fact that if $\mathrm{Ext}^i(k, M)$ vanishes for all M then it also vanishes for all $n \geqslant i$.

The conjecture given in [7] involved certain exceptional generators denoted $\mathbf{e}_f$, for $f \in P$, in the depth 4 component of the larger Lie algebra $\mathrm{gr}^4_D\mathfrak{dmr}$ of double shuffle equations. It is not known if they are in the image of $\mathfrak{d}^4$. Thus the formulation (8.8) eliminates part of the conjecture given in [7].

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

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 $\mathrm{Der}^{\Theta}\,\mathbb{L}(\mathsf{a},\mathsf{b})$ which are $\mathfrak{sl}_2$-invariant and well-defined modulo $(\mathfrak{u}^{\mathrm{geom}})^{\mathfrak{sl}_2}$. It follows from [8, Theorem 10.1] that their action on the derivations $\varepsilon^{\vee}_{2k+2}$ are known explicitly modulo Lie brackets involving at least three $\varepsilon^{\vee}_{2n+2}$, with $n \geqslant 0$. 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 $n \geqslant 1$.

REMARK 8.7. In [7] we defined an injective linear map

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

from the space P of even period polynomials to the space of solutions $\mathfrak{ls}^4$ to the linearized 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 are the images of 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:

$$\bar{\mathbf{e}}(x_1^{2n} - x_2^{2n}) + \{z_3, x_1^{2n-2}\}_{\bullet} = 0. \tag{8.9}$$

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 $\mathrm{Der}^{\Theta}\,\mathbb{L}(\mathsf{a},\mathsf{b})$ as follows. Recall that the map $i_1 : \mathfrak{g}^{\mathfrak{m}} \to \mathrm{Der}^{\Theta}\,\mathbb{L}(\mathsf{a},\mathsf{b})$ (3.2) is injective by Theorem 3.1. Since B cuts out the depth filtration on the image $i_1(\mathfrak{g}^{\mathfrak{m}})$ (Corollary 3.7) we obtain an injective morphism

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

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

More precisely, consider linear map

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

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

$$\sum_{i<j} \lambda_{i,j} [\underline{\varepsilon}^{\vee}_{2i+2}, \underline{\varepsilon}^{\vee}_{2j+2}] = 0.$$

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

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

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

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

$$\mathfrak{c}(P) = \sum_{j} \lambda_{1,j} \big[z_3, \underline{\varepsilon}^{\vee}_{2j+2}\big] + \sum_{i,a,b} \lambda_{i,a+b} \frac{1}{2b} \big[\underline{\varepsilon}^{\vee}_{2i+2}, \big[\underline{\varepsilon}^{\vee}_{2a+2}, \big[\underline{\varepsilon}^{\vee}_{2b+2}, \underline{\varepsilon}^{\vee}_{0}\big]\big]\big].$$

The Broadhurst–Kreimer conjecture suggests the following:

CONJECTURE 2 (Elliptic geometric Broadhurst–Kreimer conjecture).

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

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

8.6. Some related problems.

(1) Show that the map $\mathfrak{c}(P) : K \to (\mathfrak{d}^4)^{\mathrm{ab}}$ is injective.

(2) Relate the elements $\mathfrak{c}(P)$ to the exceptional elements $\mathbf{e}_f$ defined in [7].

(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 $\leqslant 4$.

9. Some motivation from the relative completion of $\mathbf{SL_2}(\mathbb{Z})$

I shall very briefly sketch how I arrived at formula (1.11) and (1.9) by considering double integrals of Eisenstein series. This explains why the coefficients for the explicit formula for the σ^c_{2n+1} involve the odd period polynomials of Eisenstein series. The arguments which follow require more substantial technical background.

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

9.1. Overview. Denote the Hecke-normalized Eisenstein series by

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

where $2k \geqslant 4$ and $\sigma_k(n)$ denotes the divisor function. For any modular form $f(\tau)$ of weight $2k \geqslant 4$ for $\mathrm{SL}_2(\mathbb{Z})$ we shall write (see [8] for further details):

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

where $q = \exp(2i\pi\tau)$. It is to be viewed as a global section of $V_{2k-2} \otimes \Omega^1_{\mathbb{H}}$ over the upper-half plane $\mathbb{H}$. In [8, Section 5] we defined regularized iterated integrals of Eisenstein series between cusps, building on [29, 30]. Consider the double integrals:

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

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 $k \geqslant 0$, there is an explicit morphism of SL_2-representations [8, Section 2.4]

$$\partial^k : V_{2m} \otimes_{\mathbb{Q}} V_{2n} \longrightarrow V_{2m+2n-2k}.$$

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 [8, Section 11], computes this polynomial as the Petersen inner product of two (real analytic) modular forms. The part we are interested in, by the unfolding method, corresponds to the convolution of two Eisenstein series, and yields a certain multiple of an odd zeta value. The ratios of the coefficients are the odd period polynomials of Eisenstein series.

9.2. Precise statement. All the notation in this section is borrowed from [8], Section 11. Let $k \geqslant 1$ be odd, $a, b \geqslant 2$, and $w = 2a + 2b - 2k - 2$. Set

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

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

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

where v_{2k} was defined in [8], (10.7). Then I claim that $\widetilde{I}^k_{2a,2b}$ is cocycle for $\mathrm{SL}_2(\mathbb{Z})$, and cuspidal for $k < 2\min\{a, b\} - 2$. For such k,

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

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

where e^0_w the rational Eisenstein cocycle defined in [8, Section 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 [8] with minor modifications to account for divergences.

9.3. Zeta elements from periods. Let $G^{B/\mathrm{dR}}_{1,1}$ denote the Betti and de Rham versions of the completion [24] of $\pi_1(\mathcal{M}_{1,1}, \partial/\partial q) = \mathrm{SL}_2(\mathbb{Z})$ relative to its inclusion into $\mathrm{SL}_2(\mathbb{Q})$, and $U^{B/\mathrm{dR}}_{1,1}$ their unipotent radicals. They are affine group schemes over $\mathbb{Q}$, in a Tannakian category of realizations $\mathcal{H}$ [6]. The de Rham Tannaka group $G^{\mathrm{dR}}_{\mathcal{H}}$ of $\mathcal{H}$ acts on $G^{\mathrm{dR}}_{1,1}$. The Lie algebra of $G^{\mathrm{dR}}_{\mathcal{H}}$ contains zeta elements σ_{2n+1} which are dual to the zeta values $\zeta^{\mathfrak{m}}(2n+1)$ in the ring of periods of $\mathcal{H}$. They are only well defined up to commutators.

The (M,W-bigraded) Lie algebra $\mathfrak{u}^{\mathrm{dR}}_{1,1}$ of $U^{\mathrm{dR}}_{1,1}$ is free, generated by symbols

$$\mathsf{e}_{2n+2} \otimes V^{\mathrm{dR}}_{2n} \quad \text{and} \quad \mathsf{m}_f \otimes V^{\mathrm{dR}}_{2n}$$

corresponding to Eisenstein series and a $\mathbb{Q}$-basis for the generalized Hecke-eigenspaces of cusp forms f of weight $2n+2$. Here V^{dR}_{2n} is the SL_2-module isomorphic to homogeneous polynomials over $\mathbb{Q}$ in two variables of degree $2n$.

Let us choose a basis for the (motivic) periods of regularized iterated integrals of Eisenstein series of length $\leqslant 2$ along the path S (that is, from 0 to ∞). Via the theory in [8], we can view the 'geometric part' of the σ_{2n+1} as elements of $\mathfrak{u}^{\mathrm{dR}}_{1,1}$. Taking the coefficient of $\zeta^{\mathfrak{m}}(2n+1)$, for $n \geqslant 1$, enables us to compute the image of a choice of zeta elements σ_{2n+1} in $\mathfrak{u}^{\mathrm{dR}}_{1,1}$. We can read off the coefficients of Lie brackets of length two in the e_{2n+2}, from the period computations of Section 9.2.

The group $G^{\mathrm{dR}}_{1,1}$ acts on $\mathcal{P}$, the de Rham fundamental group of the punctured curve $E^{\times}_{\partial/\partial q}$ [24, 25]. We obtain a morphism of (M, W)-bigraded Lie algebras

$$\mathfrak{u}^{\mathrm{dR}}_{1,1} \longrightarrow \mathrm{Der}\,\mathbb{L}(\mathsf{a}, \mathsf{b}).$$

It sends the generators m_f to zero, and the lowest-weight vectors in $\mathsf{e}_{2n+2} \otimes V^{\mathrm{dR}}_{2n}$ to the derivations $\varepsilon^{\vee}_{2n+2}$. The images of our σ_{2n+1} provide a choice of zeta elements in $\mathrm{Der}\,\mathbb{L}(\mathsf{a}, \mathsf{b})$, for which we know the coefficients of $[\varepsilon^{\vee}_{2a+2}, [\varepsilon^{\vee}_{2b+2}, \varepsilon^{\vee}_0]]$, corresponding to the case $k = 1$ in Section 9.2. The cases $k \geqslant 2$ also provide the coefficients of Lie brackets involving two $\varepsilon^{\vee}_{2n+2}$, where $n \geqslant 0$, and several $\varepsilon^{\vee}_0$.

Acknowledgements

This paper was written at the IAS and IHES, partially supported by NSF grant DMS-1128155 and ERC grant 257638. Its origin was an attempt to interpret the

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

algebraic structures in the notes [3] geometrically. This paper owes a great deal to Richard Hain, who kindly explained his recent papers [24, 25] to me and made many suggestions. Many thanks also to H. Bachmann, Ding Ma, N. Matthes, and K. Tasaka for corrections.

Appendix

Recall that $\mathfrak{u}^{\text{geom}} \subset \text{gr}_B^{\geqslant 1}\text{Der}^{\Theta}\mathbb{L}(\mathsf{a},\mathsf{b})$ is the Lie algebra generated by the set of derivations $\varepsilon_{2n}^{\vee}$, for $n \geqslant 0$. It is bigraded for the weight W and monodromy-weight M filtrations. We defined in Section 4.5 an injective morphism of Lie algebras

$$\ell' : \mathfrak{u}^{\text{geom}} \longrightarrow \mathfrak{pls}. \tag{A.1}$$

Throughout this section we identify $\mathfrak{pls} \cong \overline{\mathfrak{pls}}$. The latter is the space of solutions to the linearized double shuffle equations with poles. Let $\mathfrak{pls}^k$ denote the component of $\mathfrak{pls}$ in depth k. The following theorem is extracted from [3].

THEOREM A.1. *The morphism* (A.1) *is an isomorphism in B-degrees $\leqslant 3$, that is,*

$$\text{gr}_k^B \mathfrak{u}^{\text{geom}} \xrightarrow{\sim} \mathfrak{pls}^k$$

for $k = 1, 2, 3$. Denote the Poincaré series for $\text{gr}_k^B \mathfrak{u}^{\text{geom}}$ by

$$u_k(s) = \sum_{n\in\mathbb{Z}} s^n \dim_{\mathbb{Q}}(\text{gr}_k^B \text{gr}_{2n}^M \mathfrak{u}^{\text{geom}}).$$

Then

$$u_1(s) = \frac{s^{-1}}{1-s^2}, \quad u_2(s) = \frac{s^2}{(1-s^2)(1-s^6)}, \quad u_3(s) = \frac{s}{(1-s^2)(1-s^4)(1-s^6)}.$$

The spaces $\mathfrak{pls}^k$ can be further broken down into smaller spaces using an action of $\mathfrak{sl}_2$, and a new filtration we call the residue. The proof will give, more precisely, dimensions and in fact generators for each graded piece.

A.1. Description of $\mathfrak{pls}$ in depths $\leqslant$ 3. The space $\overline{\mathfrak{pls}}^1 \subset \mathbb{Q}(x_1)$ is the graded vector space of even rational functions $f \in \mathbb{Q}(x_1^2)$ such that $x_1^2 f(x_1)$ is a polynomial. By the definitions in Section 4.5, and the formula for $\varepsilon_{2n}^{\vee}(a)$, we have

$$\overline{\ell'} : \varepsilon_{2n}^{\vee} \mapsto x_1^{2n-2} \quad \text{for } n \geqslant 0.$$

Therefore, in particular

$$\bigoplus_{n\geqslant 0} \varepsilon_{2n}^{\vee}\mathbb{Q} \cong \mathfrak{pls}^1.$$

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

Now $\overline{\mathfrak{pls}}^2 \subset \mathbb{Q}(x_1, x_2)$ is the graded vector space of rational functions f such that

$$f(x_1, x_2)x_1(x_2 - x_1)x_2 \in \mathbb{Q}[x_1, x_2]$$

which also satisfy the linearized double shuffle equations

$$\begin{gathered} f(x_1, x_1 + x_2) + f(x_2, x_1 + x_2) = 0 \\ f(x_1, x_2) + f(x_2, x_1) = 0, \end{gathered}$$

and $\overline{\mathfrak{pls}}^3 \subset \mathbb{Q}(x_1, x_2, x_3)$ is the graded vector space of rational functions f with

$$f(x_1, x_2, x_3)x_1(x_2 - x_1)(x_3 - x_2)x_3 \in \mathbb{Q}[x_1, x_2, x_3],$$

which satisfy the linearized double shuffle equations

$$\begin{gathered} f(x_1, x_{12}, x_{123}) + f(x_2, x_{12}, x_{123}) + f(x_2, x_{23}, x_{123}) = 0 \\ f(x_1, x_2, x_3) + f(x_2, x_1, x_3) + f(x_2, x_3, x_1) = 0, \end{gathered} \tag{A.2}$$

where $x_{ij} = x_i + x_j$ and $x_{ijk} = x_i + x_j + x_k$. The usual T-grading for multiple zeta values corresponds to one half of the M-grading. It is given by the degree of the polynomials plus the number of variables. Thus x_1^{2n} has T-degree $2n + 1$ and $x_1^{2n}x_2^{2m}$ has T-degree $2n + 2m + 2$. We give a complete description of $\mathfrak{pls}^{\leqslant 3}$.

A.2. The residue filtration. An element $f \in \overline{\mathfrak{pls}}^k$ is a rational function in $x_1, \ldots, x_k$ with at most simple poles along $x_i = x_{i+1}$ for $1 \leqslant i \leqslant k - 1$ and $x_1 = 0$, $x_k = 0$ (the latter coincide when $k = 1$, giving a possible double pole at $x_1 = 0$ in this case) and satisfying the linearized double shuffle equations.

For any such function $f \in \overline{\mathfrak{pls}}^k$, define the *residue*

$$R_k f = \mathrm{Res}_{x_k=0} f \in \mathbb{Q}(x_1, \ldots, x_{k-1}) \quad \text{when } k \geqslant 2,$$

and in the case $k = 1$, let $R_1 f = \mathrm{Res}_{x_1=0}(x_1 f)$. The rational function $R_k f$ depends on $x_1, \ldots, x_{k-1}$ and has the same pole structure described above. It is not an element of $\overline{\mathfrak{pls}}^{k-1}$ in general, however. Define the residue filtration

$$\mathfrak{R}_m \overline{\mathfrak{pls}}^k = \{f \in \overline{\mathfrak{pls}}^k \text{ such that } R_{k-m} R_{k-m+1} \ldots R_k f = 0\},$$

for all $m \geqslant 0$, and set $\mathfrak{R}_{-1}\mathfrak{pls} = 0$. (One can show [3] that the linearized stuffle equations are stable under the residue R_k, but the linearized shuffle equations are not.) Although we shall not need it, a theorem proved in [3] states that this filtration is compatible with the Lie algebra structure:

$$\{\mathfrak{R}_p \mathfrak{pls}, \mathfrak{R}_q \mathfrak{pls}\}_{\bullet} \subset \mathfrak{R}_{p+q} \mathfrak{pls},$$

and furthermore, using the dihedral symmetry of the linearized double shuffle equations [7, Section 6.5], that $\mathfrak{R}_0\mathfrak{pls}$ is the Lie subalgebra of $\mathfrak{pls}$ consisting of rational functions with no poles at all (that is, polynomials satisfying the linearized double shuffle equations). It is denoted by

$$\mathfrak{ls} = \mathfrak{R}_0\mathfrak{pls}$$

and was studied in [7]. The generating series of its dimensions are predicted by a version of the Broadhurst–Kreimer conjecture, proven in depths $\leqslant 3$. Since the residue maps R_i are homogeneous for the M-weight, the Lie algebra $\mathrm{gr}^{\mathfrak{R}}\mathfrak{pls}$ is bigraded for the M-weight and $\mathfrak{R}$-grading. It follows from the definitions that

$$\mathfrak{R}_0\overline{\mathfrak{pls}}^1 = \bigoplus_{n\geqslant 0}\mathbb{Q}x^{2n} \quad \text{and} \quad \mathrm{gr}_1^{\mathfrak{R}}\overline{\mathfrak{pls}}^1 \cong \mathbb{Q}x_1^{-2}. \tag{A.3}$$

We have $\mathfrak{R}_{k-1}\overline{\mathfrak{pls}}^k = \overline{\mathfrak{pls}}^k$ for $k \geqslant 2$, since if $R_2 \ldots R_k f \in \mathbb{Q}(x_1)$ had a pole in x_1 then it would be a multiple of x_1^{-2} and this forces f to be of M weight -2, and hence a rational multiple of $(x_1(x_2-x_1)\ldots(x_k-x_{k-1})x_k)^{-1}$. One easily shows that the latter does not satisfy the requisite equations to be in $\overline{\mathfrak{pls}}^k$.

In summary, in depths two and three we have

$$\begin{aligned}\mathrm{gr}_{\bullet}^{\mathfrak{R}}\mathfrak{pls}^2 &= \mathrm{gr}_0^{\mathfrak{R}}\mathfrak{pls}^2 \oplus \mathrm{gr}_1^{\mathfrak{R}}\mathfrak{pls}^2 \\ \mathrm{gr}_{\bullet}^{\mathfrak{R}}\mathfrak{pls}^3 &= \mathrm{gr}_0^{\mathfrak{R}}\mathfrak{pls}^3 \oplus \mathrm{gr}_1^{\mathfrak{R}}\mathfrak{pls}^3 \oplus \mathrm{gr}_2^{\mathfrak{R}}\mathfrak{pls}^3.\end{aligned} \tag{A.4}$$

We shall compute the generating series for the dimensions of these five pieces.

A.3. $\mathfrak{sl}_2$-algebra structure. The Lie algebra $\mathfrak{u}^{\mathrm{geom}}$ admits an action of $\mathfrak{sl}_2$ which is compatible with the Lie algebra structure. Similarly, one can show [3], that there is also an action of $\mathfrak{sl}_2$ on $\mathfrak{pls}$ which is compatible with (A.1).

DEFINITION A.2. Define operators $\mathfrak{e} : \mathfrak{pls}^k \to \mathfrak{pls}^{k+1}$ and $\mathfrak{f} : \mathfrak{pls}^r \to \mathfrak{pls}^{r-1}$ on the level of reduced rational function representations by the formulae

$$\mathfrak{e}(f) = \{x_1^{-2}, f\}_{\bullet} \tag{A.5}$$

$$\mathfrak{f}(f) = \sum_{i=1}^{r-1} x_i \,\mathrm{Res}_{z=x_i} f(x_1, \ldots, x_i, z, x_{i+1}, \ldots, x_{r-1}). \tag{A.6}$$

One can prove [3] that these operators indeed preserve $\mathfrak{pls}$, generate a copy of $\mathfrak{sl}_2$, and are derivations for the Lie bracket. There is a simple general formula [3] for $\mathfrak{f}$ on $\mathfrak{pls}^k$ in terms of the residue operator R_k and a certain cyclic symmetry. We only need the case of depth 3. If $g \in \overline{\mathfrak{pls}}^3$, then

$$\mathfrak{f}(g(x_1,x_2,x_3)) = x_1g_3(x_2-x_1,-x_1) + x_2g_3(-x_2,x_1-x_2), \tag{A.7}$$

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

where $g_3 = \mathrm{Res}_{x_3=0}\, g$. One can check that $\mathfrak{e}(\mathfrak{R}_k) \subset \mathfrak{R}_{k+1}$, and that the composition $\mathfrak{fe} : \mathfrak{ls}^k \to \mathfrak{R}_1\mathfrak{p}\mathfrak{ls}^{k+1}$ is multiplication by the polynomial degree. In particular, we shall need the fact that $\mathfrak{e} : \mathfrak{ls}^2 \to \mathfrak{R}_1\mathfrak{p}\mathfrak{ls}^3$ is injective.

Here we only need the case of depths 2, and 3, and all of the claims made above can be checked by elementary manipulations on rational functions.

A.4. Recap on $\mathfrak{ls}$ in depths 2, 3. As a consequence of the work of Zagier and Goncharov, who computed the dimensions of the weight-graded pieces of $\mathfrak{ls}^2$ and $\mathfrak{ls}^3$ respectively (see also [1]), one can prove that there are exact sequences:

$$0 \longrightarrow S \longrightarrow \bigwedge^2 \mathfrak{ls}^1 \longrightarrow \mathfrak{ls}^2 \longrightarrow 0 \tag{A.8}$$
$$0 \longrightarrow S \otimes \mathfrak{ls}^1 \longrightarrow \mathrm{Lie}_3(\mathfrak{ls}^1) \longrightarrow \mathfrak{ls}^3 \longrightarrow 0.$$

For further details, see [7, equations (7.8) and (7.10)]. The third map in each line is given by the Lie bracket on $\mathfrak{ls}$. Here S is the space of solutions to the period polynomial equations, and is isomorphic to the graded vector space of cuspidal cocycles for $\mathrm{SL}_2(\mathbb{Z})$.

There are two consequences: every element of $\mathfrak{ls}^2$, $\mathfrak{ls}^3$ is generated by Lie brackets of elements in $\mathfrak{ls}^1$, and so the morphism of Lie algebras

$$\ell' : \mathfrak{u}^{\mathrm{geom},+} \longrightarrow \mathfrak{ls}$$

is surjective in depths $\leqslant 3$, where $\mathfrak{u}^{\mathrm{geom},+}$ is the Lie subalgebra of $\mathfrak{u}^{\mathrm{geom}}$ generated by the $\varepsilon^{\vee}_{2n}$ for $n \geqslant 1$. The second consequence is that it gives formulae for the generating series of dimensions with respect to the M-weight. If

$$d_k(s) = \sum_{n\geqslant 0} s^n \dim(\mathfrak{ls}^k_n) \quad \text{and} \quad S(s) = \sum_n s^n \dim(S_n) = \frac{s^{12}}{(1-s^4)(1-s^6)}$$

then it follows from (A.3) that $d_1(s) = s^3/(1-s)$, and the first line of (A.8) gives

$$d_2(s) = \frac{1}{2}\big(d_1(s^2) - d_1(s)\big) - S(s) = \frac{s^8}{(1-s^2)(1-s^6)}. \tag{A.9}$$

Remark that the smallest element in $\mathfrak{ls}^2$ is of degree 8 and is given explicitly by

$$\{x_1^2, x_1^4\}_{\bullet} = x_1x_2(x_1-x_2)(2x_1-x_2)(x_1-2x_2)(x_2+x_1) \tag{A.10}$$

which will play a role later on.

The second short exact sequence gives, after a similar calculation,

$$d_3(s) = \frac{s^{11}(1+s^2-s^4)}{(1-s^2)(1-s^4)(1-s^6)}. \tag{A.11}$$

The coefficient of s^{2n+1}, for $n \geqslant 2$, is given simply by the sequence $\lfloor ((n-1)^2-1)/12 \rfloor$, which is essentially the statement of Goncharov's theorem for $\dim_{\mathbb{Q}} \mathfrak{ls}^3_n$ [21].

Thus the parts $\mathfrak{ls}^2$, $\mathfrak{ls}^3$ of $\mathfrak{pls}^2$, $\mathfrak{pls}^3$ are generated respectively by

$$\{x_1^{2a}, x_1^{2b}\}_{\bullet} \quad \text{and} \quad \{x_1^{2a}, \{x_1^{2b}, x_1^{2c}\}_{\bullet}\}_{\bullet} \quad \text{where } a, b, c \geqslant 1. \tag{A.12}$$

A.5. The trivial piece $\mathrm{gr}^{\mathfrak{R}}_{k-1}\mathfrak{p}\mathfrak{ls}^k$.

The iterated residue defines an injection

$$R_2 \ldots R_k : \mathrm{gr}^{\mathfrak{R}}_{k-1}\overline{\mathfrak{pls}}^k \longrightarrow \mathbb{Q}(x_1). \tag{A.13}$$

It is easy to show that $\mathrm{gr}^{\mathfrak{R}}_{k-1}\mathfrak{pls}^k$ is generated by $\mathfrak{e}^{k-1}\mathfrak{ls}_1$, that is, the elements

$$\{x_1^{-2}, \ldots \{x_1^{-2}, x_1^{2n-2}\}_{\bullet} \cdots\}_{\bullet},$$

with $k-1$ terms x_1^{-2}. They are the images of the elements $(\mathrm{ad}(\varepsilon_0^{\vee}))^k \varepsilon_{2n}^{\vee} \in \mathfrak{u}^{\mathrm{geom}}$ under (A.1). Note that for reasons to do with the $\mathfrak{sl}_2$-weights, this vanishes if $k \geqslant 2n-1$. It follows that if we write

$$t_k(s) = \sum_{n \geqslant 0} s^n \dim(\mathrm{gr}^{\mathfrak{R}}_{k-1}\mathfrak{pls}^k_n)$$

then for $k \geqslant 1$,

$$t_{2k}(s) = \frac{s^2}{1-s^2} \quad \text{and} \quad t_{2k+1}(s) = \frac{s}{1-s^2}.$$

We immediately deduce the enumeration of $\mathfrak{pls}$ in depth 2, from (A.4).

COROLLARY A.3. *The generating series for the dimensions* $p_2(s) = \sum_n s^n \dim(\mathfrak{pls}^2_n)$ *is given by the formula*

$$p_2(s) = \frac{s^2}{(1-s^2)(1-s^6)}.$$

The coefficient of s^{2n} *is* $\lfloor (n+2)/3 \rfloor$.

Proof. Since $\mathfrak{pls}^2 = \mathrm{gr}^{\mathfrak{R}}_0\mathfrak{pls}^2 \oplus \mathrm{gr}^{\mathfrak{R}}_1\mathfrak{pls}^2 = \mathfrak{ls}^2 \oplus \mathrm{gr}^{\mathfrak{R}}_1\mathfrak{pls}^2$, its Poincaré series is given by $p_2(s) = d_2(s) + t_2(s)$. $\square$

It follows that the pieces $\mathrm{gr}^{\mathfrak{R}}_1\overline{\mathfrak{pls}}^2$, $\mathrm{gr}^{\mathfrak{R}}_2\overline{\mathfrak{pls}}^3$ are generated respectively by

$$\{x_1^{-2}, x_1^{2n}\}_{\bullet} \quad \text{and} \quad \{x_1^{-2}, \{x_1^{-2}, x_1^{2n}\}_{\bullet}\}_{\bullet} \quad \text{where } n \geqslant 1.$$

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

A.6. The final piece $\mathrm{gr}_1^{\mathfrak{N}}\mathfrak{pls}^3$. Using the $\mathfrak{sl}_2$ action, we deduce that there is a direct sum decomposition

$$\mathrm{gr}_1^{\mathfrak{N}}\mathfrak{pls}^3 = \mathfrak{e}(\mathfrak{ls}^2) \oplus H$$

where $H = \ker(\mathfrak{f})/\mathfrak{N}_0$ is the space of lowest-weight vectors. Since $\mathfrak{e}$ is injective on $\mathfrak{ls}^2$, it suffices to compute H.

LEMMA A.4. *There is an exact sequence of graded vector spaces*

$$0 \longrightarrow H(-5) \longrightarrow \mathfrak{ls}^2 \longrightarrow \mathbb{Q}(-8) \longrightarrow 0$$

where the Tate twist $(-n)$ denotes a shift in M-weight of $+2n$. Thus the copy of $\mathbb{Q}$ in the right-hand factor sits in degree 8, that is, M-weight 16.

Proof. Let $f \in \mathfrak{pls}^3$ satisfying $\mathfrak{f}(f) = 0$. Let $g = R_3 f$ denote its residue along $x_3 = 0$. By taking the residue along $x_3 = 0$ of the first equation of (A.2), we get

$$g(x_1, x_2) + g(x_2, x_1) = 0 \tag{A.14}$$

since f has no poles along $x_2 = 0$. Now from the formula for $\mathfrak{f}(f)$ it also satisfies

$$x_1 g(x_2 - x_1, -x_1) + x_2 g(-x_2, x_1 - x_2) = 0. \tag{A.15}$$

Inspired by (A.10), define

$$h(x_1, x_2) = x_1 x_2 (2x_1 - x_2)(x_1 - 2x_2)(x_1 + x_2) \times g(x_1, x_2).$$

Then h is homogeneous of even degree in x_1, x_2. A trivial calculation shows that (A.14) and (A.15) imply that h satisfies the pair of equations

$$h(x_1, x_2) + h(x_2, x_1) = 0$$
$$h(x_2 - x_1, -x_1) - h(-x_2, x_1 - x_2) = 0$$

which is equivalent, by replacing x_1 with $-x_1$ and noting that h is of even degree, to the linearized double shuffle equations. So we have constructed an injection

$$H \longrightarrow \mathfrak{ls}^2$$

of degree 5. Now we want a lower bound on H. For this, consider elements

$$\ell_{a,b} = \frac{1}{2b}\{x_1^{2a}, \{x_1^{-2}, x_1^{2b}\}_\bullet\}_\bullet + \frac{1}{2a}\{x_1^{2b}, \{x_1^{-2}, x_1^{2a}\}_\bullet\}_\bullet, \quad \text{for } a, b \geqslant 1. \tag{A.16}$$

These elements are in $\mathfrak{R}_1\overline{\mathfrak{pls}_3}$ and in the image of (A.1). They are the images of lowest-weight vectors in $\mathfrak{u}^{\mathrm{geom}}$, and therefore satisfy $\mathfrak{f}(\ell_{a,b}) = 0$, a fact which can also be checked directly. Since $\ell_{a,b} = \ell_{b,a}$, a linear combination $\sum_{a,b\geqslant 1} c_{a,b}\ell_{a,b}$ can be represented by a symmetric polynomial in two variables

$$P(x_1, x_2) = \sum_{a,b\geqslant 1} c_{a,b}\, x_1^{2a-1} x_2^{2b-1}.$$

Equivalently, let V be the graded vector space of symmetric polynomials in two variables x_1, x_2 of odd degree in x_1 and x_2, and consider the linear map

$$\begin{aligned} \ell : V &\longrightarrow H \\ x_1^{2a-1}x_2^{2b-1} &\mapsto \ell_{a,b}. \end{aligned}$$

Using the definitions or (5.9), we have

$$\begin{aligned} &R_3\, \frac{1}{2b}\{x_1^{2a}, \{x_1^{-2}, x_1^{2b}\}_\bullet\}_\bullet \\ &\quad = x_1 r(x_1, x_2) - x_2 r(x_2, x_1) + (x_2 - x_1)(r(x_2 - x_1, x_1) - r(x_2 - x_1, x_2)), \end{aligned}$$

where $r = x_1^{2a-1}x_2^{2b-1}$. Using the fact that P is symmetric in x_1, x_2, it follows that the kernel of $R_3\ell \subset V$ is the set of polynomials satisfying

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

The solutions to this equation are precisely the defining equations for the space of odd period polynomials. As is well known, the generating series for its dimensions are again given by the generating series $S(s)$ of dimensions of the space of cusp forms, which gives us a lower bound for the dimensions of H. This lower bound coincides, via (A.8), with the upper bound on the dimension coming from $\mathfrak{ls}^2$ except for the first term. This gives the exact sequence. $\square$

The kernel of the map ℓ was first computed by Pollack [34] by a different method. He showed that the relations between the $\ell_{a,b}$ (or rather their versions in $\mathfrak{u}^{\mathrm{geom}}$) are exactly given by odd period polynomials.

REMARK A.5. In the course of the proof we have a given a new interpretation of the double shuffle space $\mathfrak{ls}^2$ as the set of symmetric odd polynomials modulo the *odd* period polynomial relations.

In particular, the spaces $\mathbf{e}(\mathfrak{ls}^2)$ and H are, respectively, generated by

$$\{x_1^{-2}, \{x_1^{2m}, x_1^{2n}\}_\bullet\}_\bullet \quad \text{where } m, n \geqslant 1, \text{ and } \ell_{a,b}.$$

https://doi.org/10.1017/fms.2016.29 Published online by Cambridge University Press

As a consequence of the previous lemma, we have

COROLLARY A.6. *The Poincaré series for* $\mathrm{gr}^{\mathfrak{R}_1}\mathfrak{pls}^3 = H \oplus \mathfrak{e}(\mathfrak{ls}^2)$ *is*

$$\frac{d_2(s)-s^8}{s^5}+\frac{d_2(s)}{s}$$

In conclusion, the Poincaré series of $\mathrm{gr}^{\mathfrak{R}}\mathfrak{pls}^3 = \mathfrak{ls}^3 \oplus \mathrm{gr}_1^{\mathfrak{R}}\mathfrak{pls}^3 \oplus \mathrm{gr}_2^{\mathfrak{R}}\mathfrak{pls}^3$ is

$$u_3(s) = d_3(s) + \frac{d_2(s)-s^8}{s^5}+\frac{d_2(s)}{s}+\frac{s}{1-s^2}$$

which, when simplified, gives the formula stated in the theorem. The surjectivity of ℓ' follows from our explicit description of generators on each piece $\mathrm{gr}_i^{\mathfrak{R}}\mathfrak{pls}^k$.

References

[1] N. Arbesfeld and B. Enriquez, 'On a lower central series filtration of the Grothendieck–Teichmüller Lie algebra grt_1', *Mosc. Math. J.* **15**(2) (2015), 205–256, 403.

[2] G. V. Belyi, 'On Galois extensions of a maximal cyclotomic field', *Math. USSR Izvestiya* **14** (1980), 247–256.

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

[4] F. Brown, 'Mixed Tate motives over $\mathbb{Z}$', *Ann. of Math.* (2) **175**(1) (2012), 949–976.

[5] F. Brown, 'On the decomposition of motivic multiple zeta values', in *Galois–Teichmüller Theory and Arithmetic Geometry*, Advanced Studies in Pure Mathematics, 63 (Math. Soc. Japan, Tokyo, 2012), 31–58.

[6] F. Brown, 'Motivic periods and the projective line minus 3 points', in *Proceedings of the ICM*, 2014, arXiv:1407.5165.

[7] F. Brown, 'Depth-graded motivic multiple zeta values', Preprint, 2013, arXiv:1301.3053.

[8] F. Brown, 'Multiple modular values for $\mathrm{SL}_2(\mathbb{Z})$', Preprint, 2014, arXiv:1407.5167.

[9] D. Broadhurst and D. Kreimer, 'Association of multiple zeta values with positive knots via Feynman diagrams up to 9 loops', *Phys. Lett. B* **393**(3–4) (1997), 403–412.

[10] J. Blümlein, D. J. Broadhurst and J. A. M. Vermaseren, 'The multiple zeta value data mine', *Comput. Phys. Comm.* **181** (2010), 582–625.

[11] D. Calaque, B. Enriquez and P. Etingof, 'Universal KZB equations: the elliptic case', in *Algebra, Arithmetic, and Geometry: in honor of Yu. I. Manin.* Vol. I, Progress in Mathematics, 269 (Birkhäuser Boston, Inc., Boston, MA, 2009), 165–266.

[12] P. Deligne and A. B. Goncharov, 'Groupes fondamentaux motiviques de Tate mixte', *Ann. Sci. Éc. Norm. Supér.* (4) **38** (2005), 1–56.

[13] 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**(4) (1990), 149–181.

[14] J. Ecalle, 'ARI/GARI, la dimorphie et l'arithmétique des multizêtas: un premier bilan', *J. Théor. Nombres Bordeaux* **15**(2) (2003), 411–478.

[15] J. Ecalle, 'Multizetas, perinomal numbers, arithmetical dimorphy, and ARI/GARI', *Ann. Fac. Sci. Toulouse Math.* (6) **13**(4) (2004), 683–708.

[16] B. Enriques and P. Lochak, 'Homology of depth-graded motivic Lie algebras and koszulity', Preprint, 2014, arXiv:1407.4060.
[17] H. Furusho, 'Double shuffle relation for associators', *Ann. of Math.* (2) **174**(1) (2011), 341–360.
[18] H. Gangl, M. Kaneko and D. Zagier, 'Double zeta values and modular forms', in *Automorphic Forms and Zeta Functions* (World Scientific Publishing, Hackensack, NJ, 2006), 71–106.
[19] C. Glanois, 'Periods of the motivic fundamental groupoid of $\mathbb{P}^1\backslash\{0, \mu_N, \infty\}$, PhD Thesis, Université Pierre et Marie Curie, 2016.
[20] A. B. Goncharov, 'The dihedral Lie algebras and Galois symmetries of $\pi_1^{(l)}(\mathbb{P}^1 - (\{0, \infty\} \cup \mu_N))$', *Duke Math. J.* **110**(3) (2001), 397–487. (English summary).
[21] A. B. Goncharov, 'Multiple polylogarithms, cyclotomy and modular complexes', *Math. Res. Lett.* **5**(4) (1998), 497–516.
[22] A. Grothendieck, 'Esquisse d'un programme', 1984. http://matematicas.unex.es/~navarro/res/esquissefr.pdf.
[23] R. Hain, 'Notes on the Universal Elliptic KZB Equation', Preprint, 2013, arXiv:1309.0580.
[24] R. Hain, 'The Hodge–de Rham theory of modular groups', in *Recent Advances in Hodge Theory*, London Mathematical Society Lecture Note Series, 427 (Cambridge University Press, Cambridge, 2016), 422–514.
[25] R. Hain and M. Matsumoto, 'Universal mixed elliptic motives', Preprint, 2015, arXiv:1512.03975.
[26] Y. Ihara and N. Takao, Seminar talk (May 1993).
[27] W. Kohnen and D. Zagier, 'Modular forms with rational periods', in *Modular Forms (Durham, 1983)*, Ellis Horwood Ser. Math. Appl.: Statist. Oper. Res. (Horwood, Chichester, 1984), 197–249.
[28] A. Levin and G. Racinet, 'Towards multiple elliptic polylogarithms', Preprint, 2007, arXiv:math/0703237.
[29] Y. Manin, 'Iterated Shimura integrals', *Mosc. Math. J.* **5** (2005), 869–881.
[30] Y. Manin, 'Iterated integrals of modular forms and noncommutative modular symbols', in *Algebraic Geometry and Number Theory*, Progress in Mathematics, 253 (Birkhäuser Boston, Boston, MA, 2006), 565–597.
[31] W. McCallum and R. Sharifi, 'A cup product in the Galois cohomology of number fields', *Duke Math. J.* **120**(2) (2003), 269–310.
[32] H. Nakamura, 'Tangential base points and Eisenstein power series', in *Aspects of Galois Theory (Gainesville, FL, 1996)*, London Mathematical Society Lecture Note Series, 256 (Cambridge University Press, Cambridge, 1999), 202–217.
[33] E. Panzer, 'The parity theorem for multiple polylogarithms', *J. Number Theory* **172** (2017), 93–113.
[34] A. Pollack, 'Relations between derivations arising from modular forms', Undergraduate Thesis, Duke University, 2009.
[35] G. Racinet, 'Doubles mélanges des polylogarithmes multiples aux racines de l'unité', *Publ. Math. Inst. Hautes Études Sci.* **95** (2002), 185–231.
[36] H. Tsumura, 'Combinatorial relations for Euler–Zagier sums', *Acta Arith.* **111** (2004), 27–42.
[37] H. Tsunogai, 'On some derivations of Lie algebras related to Galois representations', *Publ. Res. Inst. Math. Sci.* **31** (1995), 113–134.
[38] D. Zagier, 'Periods of modular forms, traces of Hecke operators, and multiple zeta values', *RIMS Kokyuroku* **843** (1993), 162–170.
[39] D. B. Zagier, 'Evaluation of the multiple zeta values $\zeta(2, \ldots, 2, 3, 2, \ldots, 2)$', *Ann. of Math.* (2) **175**(2) (2012), 977–1000.