Compositio Math. 142 (2006) 307–338
doi:10.1112/S0010437X0500182X

# Derivation and double shuffle relations for multiple zeta values

Kentaro Ihara, Masanobu Kaneko and Don Zagier

Abstract

Derivation and extended double shuffle (EDS) relations for multiple zeta values (MZVs) are proved. Related algebraic structures of MZVs, as well as a 'linearized' version of EDS relations are also studied.

## Introduction

In recent years, there has been a considerable amount of interest in certain real numbers called *multiple zeta values* (MZVs). These numbers, first considered by Euler in a special case, have arisen in various contexts in geometry, knot theory, mathematical physics and arithmetical algebraic geometry. It is known that there are many linear relations over $\mathbb{Q}$ among the MZVs, but their exact structure remains quite mysterious.

The MZVs can be given both as sums (1.1) or as integrals (1.2). From each of these representations one finds that the product of two MZVs is a $\mathbb{Z}$-linear combination of MZVs, described by a so-called *shuffle product*, but the two expressions obtained are different. Their equality gives a large collection of relations among MZVs, which we call the *double shuffle relations*. These are not sufficient to imply all relations among MZVs, but it turns out that one can extend the double shuffle relations by allowing divergent sums and integrals in the definitions (roughly speaking, by adjoining a formal variable $T$ corresponding to the infinite sum $\sum 1/n$), and that these *extended double shuffle* (EDS) relations apparently suffice to describe the ring of MZVs completely. This observation, which was made by the third author a number of years ago and has been found independently by a number of other researchers in the field, is central to this paper. Our first goal (§§ 1, 2 and 3) is to explain the EDS relations in detail. This requires introducing a certain renormalization map whose definition, initially forced on us by the asymptotic properties of divergent multiple zeta sums and integrals, is later seen to have a purely algebraic meaning. This is carried out in §§ 4 and 5, in which we also prove the equivalence of a number of different versions of the basic conjecture on the sufficiency of the EDS relations. In the next two sections we prove a number of further algebraic properties of the ring of MZVs that can be deduced from the EDS relations. In particular, we introduce a number of derivations (and, by exponentiation, automorphisms) of the ring of formal MZVs and use them to give new, and in several cases conjecturally complete, sets of relations among MZVs. These identities contain previous results of Hoffman and Ohno as special cases. Finally, the last section of the paper contain a reformulation of the EDS relations as a problem of linear algebra and some general results concerning this problem.

Some of the results in this paper (in particular, in §§ 2 and 8 concerning the double shuffle relations and renormalization) originated in work that the third-named author carried out in 1988–1994 but never published. Since that time, much work has been done by other writers

Received 22 September 2004, accepted in final form 18 April 2005.
*2000 Mathematics Subject Classification* 11M41 (primary), 11M06 (secondary).
*Keywords:* multiple zeta value, double shuffle relation, regularization, derivation.

(Goncharov, Minh, Petitot, Boutet de Monvel, Écalle, Racinet, etc.; see the bibliography) and there is a considerable amount of overlap with their results. We nevertheless present a self-contained description of the work.

## 1. Double shuffle relations (convergent case)

The *multiple zeta value* (MZV) is defined by the convergent series

$$\zeta(\mathbf{k}) = \zeta(k_1, k_2, \ldots, k_n) = \sum_{m_1 > m_2 > \cdots > m_n > 0} \frac{1}{m_1^{k_1} m_2^{k_2} \cdots m_n^{k_n}}, \tag{1.1}$$

where $\mathbf{k} = (k_1, k_2, \ldots, k_n)$ is an *admissible index set* (ordered set of positive integers whose first element is strictly greater than 1). This value has an integral representation, known as the Drinfel'd integral [Dri90], [Zag94], as follows:

$$\zeta(k_1, k_2, \ldots, k_n) = \int \cdots \int_{1 > t_1 > t_2 > \cdots > t_k > 0} \omega_1(t_1)\omega_2(t_2) \cdots \omega_k(t_k), \tag{1.2}$$

where $k = k_1 + k_2 + \cdots + k_n$ is the weight and $\omega_i(t) = dt/(1-t)$ if $i \in \{k_1, k_1 + k_2, \ldots, k_1 + k_2 + \cdots + k_n\}$ and $\omega_i(t) = dt/t$ otherwise. There are many linear relations over $\mathbb{Q}$ among MZVs of the same weight. The main goal of the theory is to give as complete a description of them as possible.

The product of two MZVs is expressible as a sum of MZVs. We may see this by using either the defining series (1.1) or the integral representation (1.2) of $\zeta(\mathbf{k})$, but the multiplication rules obtained by the two methods are not the same; the equality of the products that they give will be our main tool for obtaining linear dependences among MZVs. To describe these multiplication rules, it is convenient to use the algebraic setup given by Hoffman [Hof97]. Let $\mathfrak{H} = \mathbb{Q}\langle x, y\rangle$ be the non-commutative polynomial algebra over the rationals in two indeterminates $x$ and $y$, and $\mathfrak{H}^1$ and $\mathfrak{H}^0$ its subalgebras $\mathbb{Q} + \mathfrak{H}y$ and $\mathbb{Q} + x\mathfrak{H}y$, respectively. Let $Z : \mathfrak{H}^0 \to \mathbb{R}$ be the $\mathbb{Q}$-linear map ('evaluation map') that assigns to each word (monomial) $u_1 u_2 \cdots u_k$ in $\mathfrak{H}^0$ the multiple integral

$$\int \cdots \int_{1 > t_1 > t_2 > \cdots > t_k > 0} \omega_{u_1}(t_1)\omega_{u_2}(t_2) \cdots \omega_{u_k}(t_k) \tag{1.3}$$

where $\omega_x(t) = dt/t$, $\omega_y(t) = dt/(1-t)$. We set $Z(1) = 1$. As the word $u_1 u_2 \cdots u_k$ is in $\mathfrak{H}^0$, we always have $\omega_{u_1}(t) = dt/t$ and $\omega_{u_k}(t) = dt/(1-t)$, so the integral converges. By the Drinfel'd integral representation (1.2), we have

$$Z(x^{k_1 - 1} y x^{k_2 - 1} y \cdots x^{k_n - 1} y) = \zeta(k_1, k_2, \ldots, k_n).$$

The *weight* $k = k_1 + k_2 + \cdots + k_n$ of $\zeta(k_1, k_2, \ldots, k_n)$ is the total degree of the corresponding monomial $x^{k_1 - 1} y x^{k_2 - 1} y \cdots x^{k_n - 1} y$, and the *depth* $n$ is the degree in $y$.

Let $z_k := x^{k-1} y$, which corresponds under $Z$ to the Riemann zeta value $\zeta(k)$. Then $\mathfrak{H}^1$ is freely generated by $z_k$ ($k = 1, 2, 3, \ldots$). We define the *harmonic product* $*$ on $\mathfrak{H}^1$ inductively by

$$1 * w = w * 1 = w,$$
$$z_k w_1 * z_l w_2 = z_k(w_1 * z_l w_2) + z_l(z_k w_1 * w_2) + z_{k+l}(w_1 * w_2),$$

for all $k, l \geqslant 1$ and any words $w, w_1, w_2 \in \mathfrak{H}^1$, and then extending by $\mathbb{Q}$-bilinearity. Equipped with this product, $\mathfrak{H}^1$ becomes a commutative algebra [Hof97] and $\mathfrak{H}^0$ a subalgebra. We denote these algebras by $\mathfrak{H}^1_*$ and $\mathfrak{H}^0_*$. The first multiplication law of MZVs can then be stated by saying that the evaluation map $Z : \mathfrak{H}^0 \to \mathbb{R}$ is an algebra homomorphism with respect to the multiplication $*$, i.e.

$$Z(w_1 * w_2) = Z(w_1)Z(w_2) \tag{1.4}$$



https://doi.org/10.1112/S0010437X0500182X Published online by Cambridge University Press

for all $w_1, w_2 \in \mathfrak{H}^0$. For instance, the harmonic product $z_k * z_l = z_k z_l + z_l z_k + z_{k+l}$ corresponds to the identity $\zeta(k)\zeta(l) = \zeta(k,l) + \zeta(l,k) + \zeta(k+l)$. Note that this multiplication rule corresponds simply to the formal multiplication and rearrangement of the terms of the sums (1.1), and would remain true if the numbers $m_i$ in these sums were to run over any other discrete subsets of $\mathbb{R}_+$, as long as the series converged absolutely.

The other commutative product ш, referred to as the *shuffle product*, corresponding to the product of two integrals in (1.2), is defined on all of $\mathfrak{H}$ inductively by setting

$$1 \sqcup\!\sqcup\, w = w \sqcup\!\sqcup\, 1 = w,$$
$$uw_1 \sqcup\!\sqcup\, vw_2 = u(w_1 \sqcup\!\sqcup\, vw_2) + v(uw_1 \sqcup\!\sqcup\, w_2),$$

for any words $w, w_1, w_2 \in \mathfrak{H}$ and $u, v \in \{x, y\}$, and again extending by $\mathbb{Q}$-bilinearity. This product gives $\mathfrak{H}$ the structure of a commutative $\mathbb{Q}$-algebra [Reu93], which we denote by $\mathfrak{H}_{\sqcup\!\sqcup}$. Obviously the subspaces $\mathfrak{H}^1$ and $\mathfrak{H}^0$ become subalgebras of $\mathfrak{H}_{\sqcup\!\sqcup}$, denoted by $\mathfrak{H}^1_{\sqcup\!\sqcup}$ and $\mathfrak{H}^0_{\sqcup\!\sqcup}$, respectively. By the standard shuffle product identity of iterated integrals, the evaluation map $Z$ is again an algebra homomorphism for the multiplication ш:

$$Z(w_1 \sqcup\!\sqcup\, w_2) = Z(w_1)Z(w_2). \tag{1.5}$$

Again, this rule is a formal consequence of the formula (1.3) and would hold for the values defined by these integrals if $\omega_x$ and $\omega_y$ were replaced by any other differential forms for which the integrals converged; it is only in the equality between the two multiplication rules that the specific definition of MZVs is important.

By equating (1.4) and (1.5), we get the *double shuffle relations* of MZV:

$$Z(w_1 \sqcup\!\sqcup\, w_2) = Z(w_1 * w_2) \quad (w_1, w_2 \in \mathfrak{H}^0). \tag{1.6}$$

The first example is

$$4\zeta(3,1) + 2\zeta(2,2) = 2\zeta(2,2) + \zeta(4) \quad (= \zeta(2)^2)$$

from which we deduce $4\zeta(3,1) = \zeta(4)$. These 'finite' double shuffle (FDS) relations, however, do not suffice to obtain 'all' relations. For instance, we have 1, 2 and 4 MZVs in weights 2, 3 and 4, respectively, but the relation above of weight 4 is obviously the only double shuffle relation in weight $\leqslant 4$, so that we are only able to reduce the dimensions to 1, 2, 3 rather than the correct 1, 1, 1. We therefore need a larger supply of relations. This is the object of the 'renormalization' procedure discussed in the next section.

## 2. Regularizations of MZVs

Proposition 1. *We have two algebra homomorphisms*

$$Z^* : \mathfrak{H}^1_* \longrightarrow \mathbb{R}[T] \quad \textit{and} \quad Z^{\sqcup\!\sqcup} : \mathfrak{H}^1_{\sqcup\!\sqcup} \longrightarrow \mathbb{R}[T]$$

*that are uniquely characterized by the properties that they both extend the evaluation map $Z : \mathfrak{H}^0 \to \mathbb{R}$ and send $y$ to $T$.*

*Proof.* This is clear from the isomorphisms $\mathfrak{H}^1_* \simeq \mathfrak{H}^0_*[y]$ and $\mathfrak{H}^1_{\sqcup\!\sqcup} \simeq \mathfrak{H}^0_{\sqcup\!\sqcup}[y]$ (see [Hof97] and [Reu93]) and the fact that the map $Z$ is an algebra homomorphism for both harmonic and shuffle products. □

For an index $\mathbf{k} = (k_1, \ldots, k_n)$ (not necessarily admissible, i.e. any ordered set of positive integers), the images under the maps $Z^*$ and $Z^{\sqcup\!\sqcup}$ of the corresponding word $x^{k_1-1}y \cdots x^{k_n-1}y$ are denoted by $Z^*_{\mathbf{k}}(T)$ and $Z^{\sqcup\!\sqcup}_{\mathbf{k}}(T)$, respectively. If $\mathbf{k}$ is admissible, we have $Z^*_{\mathbf{k}}(T) = Z^{\sqcup\!\sqcup}_{\mathbf{k}}(T) = \zeta(\mathbf{k})$. In general,

Table 1. Regularizations.

| $\mathbf{k}$ | $(1)$ | $(1,1)$ | $(1,2)$ |
|---|---|---|---|
| $Z^*_{\mathbf{k}}(T)$ | $T$ | $\frac{1}{2}T^2-\frac{1}{2}\zeta(2)$ | $\zeta(2)T-\zeta(2,1)-\zeta(3)$ |
| $Z^{ш}_{\mathbf{k}}(T)$ | $T$ | $\frac{1}{2}T^2$ | $\zeta(2)T-2\zeta(2,1)$ |

we see by induction on $s$ that, for $\mathbf{k}=(\underbrace{1,1,\dots,1}_{s},\mathbf{k}')$ with $\mathbf{k}'$ admissible and $s\geqslant 0$ we have

$$Z^*_{\mathbf{k}}(T)=\zeta(\mathbf{k}')\frac{T^s}{s!}+(\text{terms of lower degree in } T)$$

and, similarly,

$$Z^{ш}_{\mathbf{k}}(T)=\zeta(\mathbf{k}')\frac{T^s}{s!}+(\text{terms of lower degree in } T),$$

and also that the coefficients of $T^i$ in $Z^*_{\mathbf{k}}(T)$ and $Z^{ш}_{\mathbf{k}}(T)$ are $\mathbb{Q}$-linear combinations of MZVs of weight $k-i$ ($k=$ weight of $\mathbf{k}$). We give a few examples in Table 1.

To state the main renormalization formula, we introduce the following power series $A(u)$ with coefficients in the subring of $\mathbb{R}$ generated by Riemann zeta values:

$$A(u)=\exp\biggl(\sum_{n=2}^{\infty}\frac{(-1)^n}{n}\zeta(n)u^n\biggr). \tag{2.1}$$

As is easily seen from the standard Taylor expansion of $\log\Gamma(x)$ at $x=1$, this is the Taylor expansion of $e^{\gamma u}\Gamma(1+u)$ ($\gamma=$ Euler's constant) near $u=0$:

$$A(u)=e^{\gamma u}\Gamma(1+u)\quad(|u|<1).$$

Define an $\mathbb{R}$-linear map $\rho:\mathbb{R}[T]\to\mathbb{R}[T]$ by

$$\rho(e^{Tu})=A(u)e^{Tu}. \tag{2.2}$$

Equivalently, $\rho$ is determined by

$$\rho\biggl(\frac{T^l}{l!}\biggr)=\sum_{k=0}^{l}\gamma_k\frac{T^{l-k}}{(l-k)!}\quad(l=0,1,2,\dots)$$

and the $\mathbb{R}$-linearity, where the coefficients $\gamma_0=1,\gamma_1=0,\gamma_2=\zeta(2)/2,\dots$ are given by the generating function

$$A(u)=\sum_{k=0}^{\infty}\gamma_k u^k.$$

Note that, by (2.1), the coefficient $\gamma_k$ is a weighted homogeneous polynomial of degree $k$ with rational coefficients in the Riemann zeta values $\zeta(2),\zeta(3),\dots$ (with $\deg(\zeta(n))=n$).

Theorem 1. *For any index set* $\mathbf{k}$, *we have*

$$Z^{ш}_{\mathbf{k}}(T)=\rho(Z^*_{\mathbf{k}}(T)). \tag{2.3}$$

*Proof.* For $M>0$ and an index set $\mathbf{k}=(k_1,k_2,\dots,k_n)$, put

$$\zeta_M(k_1,k_2,\dots,k_n):=\sum_{M>m_1>m_2>\cdots>m_n>0}\frac{1}{m_1^{k_1}m_2^{k_2}\cdots m_n^{k_n}}.$$

If $\mathbf{k}$ is admissible, i.e. $k_1>1$, then $\zeta_M(\mathbf{k})$ converges to $\zeta(\mathbf{k})$ as $M\to\infty$. Note that we can write the product $\zeta_M(\mathbf{k})\zeta_M(\mathbf{k}')$ as a linear combination of $\zeta_M(\mathbf{k}'')$s by the same rule as in the case of harmonic product of the convergent MZVs. For instance, we have

$$\zeta_M(k)\zeta_M(k') = \zeta_M(k,k') + \zeta_M(k',k) + \zeta_M(k+k').$$

With this fact and the classical formula

$$\zeta_M(1) = 1 + \frac{1}{2} + \cdots + \frac{1}{M-1} = \log M + \gamma + \mathrm{O}\Big(\frac{1}{M}\Big),$$

we see by induction that for any index set $\mathbf{k}$ we have

$$\zeta_M(\mathbf{k}) = Z^*_{\mathbf{k}}(\log M + \gamma) + \mathrm{O}(M^{-1}\log^J M) \quad \text{for some } J \text{ as } M \to \infty,$$

where $Z^*_{\mathbf{k}}(T)$ is the associated polynomial defined in Proposition 1.

Next, for $\mathbf{k} = (k_1, k_2, \ldots, k_n)$ and $0 < t < 1$, put

$$Li_{\mathbf{k}}(t) = \int\cdots\int_{t>t_1>t_2>\cdots>t_k>0} \omega_1(t_1)\omega_2(t_2)\cdots\omega_k(t_k),$$

where $k = k_1 + k_2 + \cdots + k_n$ and $\omega_i(t) = dt/(1-t)$ if $i \in \{k_1, k_1+k_2, \ldots, k_1+k_2+\cdots+k_n\}$ and $\omega_i(t) = dt/t$ otherwise. Iterated integration shows that

$$Li_{\mathbf{k}}(t) = \sum_{m_1>m_2>\cdots>m_n>0} \frac{t^{m_1}}{m_1^{k_1}m_2^{k_2}\cdots m_n^{k_n}}. \tag{2.4}$$

The product $Li_{\mathbf{k}}(t)Li_{\mathbf{k}'}(t)$ is a linear combination of $Li_{\mathbf{k}''}(t)$s via the shuffle product identity of iterated integrals, and the formula specializes at $t = 1$ to that (with the shuffle product ш) of $\zeta(\mathbf{k})\zeta(\mathbf{k}')$ if $\mathbf{k}$ and $\mathbf{k}'$ are admissible. Together with $Li_1(t) = \log(1/1-t)$, we conclude by induction that, for each index set $\mathbf{k}$, we have

$$Li_{\mathbf{k}}(t) = Z^{ш}_{\mathbf{k}}\Big(\log\frac{1}{1-t}\Big) + \mathrm{O}\Big((1-t)\log^J\Big(\frac{1}{1-t}\Big)\Big) \quad \text{for some } J \text{ as } t \nearrow 1.$$

We shall compare the behaviors of $\zeta_M(\mathbf{k})$ and $Li_{\mathbf{k}}(t)$. For that we start with (2.4) to obtain

$$\begin{aligned} Li_{\mathbf{k}}(t) &= \sum_{m_1>m_2>\cdots>m_n>0} \frac{t^{m_1}}{m_1^{k_1}m_2^{k_2}\cdots m_n^{k_n}} \\ &= \sum_{m=1}^{\infty}\Big(\sum_{m>m_2>\cdots>m_n>0}\frac{1}{m^{k_1}m_2^{k_2}\cdots m_n^{k_n}}\Big)t^m \\ &= \sum_{m=1}^{\infty}(\zeta_{m+1}(\mathbf{k}) - \zeta_m(\mathbf{k}))t^m \\ &= (1-t)\sum_{m=1}^{\infty}\zeta_m(\mathbf{k})t^{m-1}. \end{aligned}$$

To deduce the theorem from this, we use the following lemma.

Lemma 1.

(i) *Let $P(T) \in \mathbb{R}[T]$ and $Q(T) = \rho(P(T))$. Then*

$$\sum_{m=1}^{\infty} P(\log m + \gamma)t^{m-1} = \frac{1}{1-t}Q\Big(\log\frac{1}{1-t}\Big) + O\Big(\log^J\Big(\frac{1}{1-t}\Big)\Big)$$

*for some $J$ $(= \deg P - 1)$ as $t \nearrow 1$.*

(ii) *For $l \geqslant 0$, we have*

$$\sum_{m=1}^{\infty} \frac{\log^l m}{m} t^{m-1} = O\biggl(\log^{l+1}\biggl(\frac{1}{1-t}\biggr)\biggr) \quad \text{as } t \nearrow 1.$$

As $\zeta_m(\mathbf{k}) = Z^*_{\mathbf{k}}(\log m + \gamma) + \mathrm{O}(m^{-1}\log^J m)$, Lemma 1 gives

$$(1-t)\sum_{m=1}^{\infty} \zeta_m(\mathbf{k}) t^{m-1} = Q\biggl(\log\frac{1}{1-t}\biggr) + \mathrm{O}\biggl((1-t)\log^{J+1}\biggl(\frac{1}{1-t}\biggr)\biggr)$$

with $Q(T) = \rho(Z^*_{\mathbf{k}}(T))$, so we conclude that $Z^{ш}_{\mathbf{k}}(T) = \rho(Z^*_{\mathbf{k}}(T))$. □

*Proof of Lemma 1.* We first prove part (ii). For $l = 0$, the left-hand side is $(1/t)\log(1/1-t)$ which is clearly $\mathrm{O}(\log(1/(1-t)))$ as $t \to 1$. We now proceed by induction on $l$. We have

$$\log^{l+1} m \leqslant C_l \sum_{n=1}^{m} \frac{\log^l n}{n} \quad (m \geqslant 1, l \geqslant 0)$$

for some constant $C_l$ independent of $m$. (This is easily seen by comparing the sum with the corresponding integral $\int_1^m (\log^l x/x)\,dx$.) Hence, for $t < 1$ we obtain

$$\begin{aligned} \sum_{m=1}^{\infty} \frac{\log^{l+1} m}{m} t^{m-1} &\leqslant C_l \sum_{m=1}^{\infty} \frac{t^{m-1}}{m} \sum_{n=1}^{m} \frac{\log^l n}{n} \\ &= C_l \sum_{n=1}^{\infty} \frac{\log^l n}{n} t^{n-1} \sum_{r=1}^{\infty} \frac{t^{r-1}}{r+n-1} \\ &< C_l \biggl(\sum_{n=1}^{\infty} \frac{\log^l n}{n} t^{n-1}\biggr)\biggl(\frac{1}{t}\log\frac{1}{1-t}\biggr). \end{aligned}$$

The estimate in part (ii) now follows for all $l$ by induction.

For part (i), it is enough by linearity to prove the identity for $P(T) = (T-\gamma)^l$. Put $Q(T) = \rho((T-\gamma)^l)$. Then

$$Q(T) = \frac{d^l}{du^l}[A(u)e^{(T-\gamma)u}]_{u=0} = \frac{d^l}{du^l}[\Gamma(1+u)e^{Tu}]_{u=0},$$

and hence

$$\begin{aligned} \frac{1}{1-t} Q\biggl(\log\frac{1}{1-t}\biggr) &= \frac{d^l}{du^l}\biggl[\frac{\Gamma(1+u)}{(1-t)^{1+u}}\biggr]_{u=0} \\ &= \frac{d^l}{du^l}\biggl[\sum_{m=1}^{\infty} \frac{\Gamma(m+u)}{\Gamma(m)} t^{m-1}\biggr]_{u=0} \quad \text{(binomial theorem)} \\ &= \sum_{m=1}^{\infty} \frac{\Gamma^{(l)}(m)}{\Gamma(m)} t^{m-1}. \end{aligned}$$

From the standard integral representation

$$\log\Gamma(x) = \biggl(x - \frac{1}{2}\biggr)\log x - x + C - \int_0^{\infty} \frac{t - [t] - 1/2}{x+t}\,dt \quad (x > 0)$$

we see by induction on $l$ that

$$\frac{\Gamma^{(l)}(m)}{\Gamma(m)} = \log^l m + \mathrm{O}\biggl(\frac{\log^{l-1} m}{m}\biggr) \quad (m \to \infty)$$

for all $l \geqslant 0$, and from this and Lemma 1(ii) we deduce that

$$\begin{aligned}\sum_{m=1}^{\infty} \frac{\Gamma^{(l)}(m)}{\Gamma(m)} t^{m-1} &= \sum_{m=1}^{\infty} \log^l m \ t^{m-1} + \mathrm{O}\biggl(\log^l \frac{1}{1-t}\biggr) \\ &= \sum_{m=1}^{\infty} P(\log m + \gamma) t^{m-1} + \mathrm{O}\biggl(\log^l \frac{1}{1-t}\biggr).\end{aligned}$$

This completes the proof. □

As an example of the theorem, by comparing the two entries for $\mathbf{k} = (1,2)$ in the table at the beginning of the section, we find

$$\zeta(2)T - 2\zeta(2,1) = \zeta(2)T - \zeta(2,1) - \zeta(3),$$

the left-hand side being $Z^{ш}_{\mathbf{k}}(T)$ and the right $\rho(Z^{*}_{\mathbf{k}}(T))$. (Note that $\rho(T) = T$.) This equality gives Euler's formula

$$\zeta(2,1) = \zeta(3)$$

and shows that the space of weight-3 MZVs is one-dimensional, whereas using only the finite shuffle relations as in § 1 we were not able to reduce the dimension below 2. Similarly, by comparing the two sides of (2.3) for $\mathbf{k} = (1,3), (1,2,1)$, we find

$$\begin{aligned}\zeta(3)T - 2\zeta(3,1) - \zeta(2,2) &= \zeta(3)T - \zeta(4) - \zeta(3,1), \\ \zeta(2,1)T - 3\zeta(2,1,1) &= \zeta(2,1)T - \zeta(3,1) - \zeta(2,2) - 2\zeta(2,1,1)\end{aligned}$$

and, hence, the identities

$$\zeta(4) = \zeta(3,1) + \zeta(2,2) = \zeta(2,1,1),$$

which, together with the formula $4\zeta(3,1) = \zeta(4)$ obtained in § 1 as a FDS relation, reduce the dimension of weight-4 MZVs to 1.

In each of these examples, the degree with respect to $T$ was at most 1 and hence the effect of the automorphism $\rho$, which acts as the identity on the subspace $\mathbb{R} + \mathbb{R}T$ of $\mathbb{R}[T]$, was not visible. As an example involving higher powers of $T$, take $\mathbf{k} = (1,1,2)$ in (2.3). Then

$$\begin{aligned}Z^{ш}_{1,1,2}(T) &= \tfrac{1}{2}\zeta(2)T^2 - 2\zeta(2,1)T + 3\zeta(2,1,1), \\ Z^{*}_{1,1,2}(T) &= \tfrac{1}{2}\zeta(2)T^2 - (\zeta(3) + \zeta(2,1))T + \tfrac{1}{2}\zeta(4) + \zeta(3,1) + \zeta(2,1,1),\end{aligned}$$

so from $Z^{ш}_{1,1,2}(T) = \rho(Z^{*}_{1,1,2}(T))$ and $\rho(T^2) = T^2 + \zeta(2)$ we find (again) $\zeta(2,1) = \zeta(3)$ and also

$$3\zeta(2,1,1) = \tfrac{1}{2}\zeta(2)^2 + \tfrac{1}{2}\zeta(4) + \zeta(3,1) + \zeta(2,1,1).$$

This latter relation is a consequence of the relations obtained above, which already reduced the dimension of the MZV space in this weight to 1, but at the same time we see that here $Z^{ш}_{1,1,2}(T) \neq Z^{*}_{1,1,2}(T)$ and hence that the presence of the automorphism $\rho$ in (2.3) is really necessary to make this identity correct.

These examples show that in weight up to 4, Theorem 1, together with the homomorphism properties of the evaluation maps $Z^{ш}$ and $Z^{*}$, suffices to give all relations of MZVs, and that this still remains true even if we restrict our attention to index sets $\mathbf{k}$ with at most one leading 1. It is conjectured that both of these statements remain true in all weights. These conjectures will be formulated more precisely in the next section in the language of the algebra $\mathfrak{H}$.



https://doi.org/10.1112/S0010437X0500182X Published online by Cambridge University Press

## 3. EDS relations

Denote by $\mathrm{reg}_{ш}^T$ (respectively $\mathrm{reg}_*^T$) the map (actually an isomorphism) $\mathfrak{H}_{ш}^1 \to \mathfrak{H}_{ш}^0[T]$ (respectively $\mathfrak{H}_*^1 \to \mathfrak{H}_*^0[T]$) defined by the properties that it is the identity on $\mathfrak{H}^0$, maps $y$ to $T$, and is an algebra homomorphism. The maps $\mathfrak{H}_{ш}^1 \to \mathfrak{H}_{ш}^0$ and $\mathfrak{H}_*^1 \to \mathfrak{H}_*^0$ obtained by specializing $\mathrm{reg}_{ш}^T$ and $\mathrm{reg}_*^T$ to $T=0$ will be denoted by $\mathrm{reg}_{ш}$ and $\mathrm{reg}_*$, respectively. If $R$ is a commutative $\mathbb{Q}$-algebra with 1 and $Z_R$ is any map from $\mathfrak{H}^0$ to $R$ which is a homomorphism with respect to *both* multiplications ш and $*$, i.e.

$$Z_R(w_1 ш w_2) = Z_R(w_1 * w_2) = Z_R(w_1)Z_R(w_2), \tag{3.1}$$

we say that $Z_R$ has the 'finite double shuffle (FDS)' property. We can then extend $Z_R$ to maps $Z_R^{ш} : \mathfrak{H}^1 \to R[T]$ and $Z_R^* : \mathfrak{H}^1 \to R[T]$ in the same way (namely, they agree with $Z_R$ on $\mathfrak{H}^0$, send $y$ to $T$, and are homomorphisms with respect to ш or $*$), or equivalently, we can define $Z_R^{ш}$ and $Z_R^*$ as the composites of the maps $\mathrm{reg}_{ш}^T$ and $\mathrm{reg}_*^T$ with the map $Z_R \otimes 1 : \mathfrak{H}^0[T] = \mathfrak{H}^0 \otimes \mathbb{Q}[T] \to R \otimes \mathbb{Q}[T] = R[T]$. Finally, we define an $R$-module automorphism $\rho_R$ of $R[T]$, generalizing the map $\rho$ in Theorem 1, by the formula

$$\rho_R(e^{Tu}) = A_R(u)e^{Tu} \tag{3.2}$$

(together with the requirement of $R$-linearity), where $A_R(u)$ is the power series defined, in analogy with (2.1), by

$$A_R(u) = \exp\biggl(\sum_{n=2}^{\infty} \frac{(-1)^n}{n} Z_R(x^{n-1}y)u^n\biggr) \in R[[u]]. \tag{3.3}$$

Theorem 2. *Let $(R, Z_R)$ be as above with the FDS property. Then the following properties are equivalent:*

(i) $(Z_R^{ш} - \rho_R \circ Z_R^*)(w) = 0$ *for all* $w \in \mathfrak{H}^1$;
(ii) $(Z_R^{ш} - \rho_R \circ Z_R^*)(w)|_{T=0} = 0$ *for all* $w \in \mathfrak{H}^1$;
(iii) $Z_R^{ш}(w_1 ш w_0 - w_1 * w_0) = 0$ *for all* $w_1 \in \mathfrak{H}^1$ *and all* $w_0 \in \mathfrak{H}^0$;
(iii′) $Z_R^*(w_1 ш w_0 - w_1 * w_0) = 0$ *for all* $w_1 \in \mathfrak{H}^1$ *and all* $w_0 \in \mathfrak{H}^0$;
(iv) $Z_R(\mathrm{reg}_{ш}(w_1 ш w_0 - w_1 * w_0)) = 0$ *for all* $w_1 \in \mathfrak{H}^1$ *and all* $w_0 \in \mathfrak{H}^0$;
(iv′) $Z_R(\mathrm{reg}_*(w_1 ш w_0 - w_1 * w_0)) = 0$ *for all* $w_1 \in \mathfrak{H}^1$ *and all* $w_0 \in \mathfrak{H}^0$;
(v) $Z_R(\mathrm{reg}_{ш}(y^m * w_0)) = 0$ *for all* $m \geqslant 1$ *and all* $w_0 \in \mathfrak{H}^0$;
(v′) $Z_R(\mathrm{reg}_*(y^m ш w_0 - y^m * w_0)) = 0$ *for all* $m \geqslant 1$ *and all* $w_0 \in \mathfrak{H}^0$.

The implications (i) ⇒ (ii), (iii) ⇒ (iv) and (iii′) ⇒ (iv′) are obvious (note that $Z_R \circ \mathrm{reg}_{ш}$ (respectively $Z_R \circ \mathrm{reg}_*$) is the specialization $Z_R^{ш}|_{T=0}$ (respectively $Z_R^*|_{T=0}$)). For (i) ⇒ (iii), multiply $Z_R(w_0)$ ($\in R$) on both sides of $Z_R^{ш}(w_1) = \rho_R(Z_R^*(w_1))$ and use the $R$-linearity of $\rho_R$ to get $Z_R^{ш}(w_1 ш w_0) = \rho_R(Z_R^*(w_1 * w_0))$. Using (i) on the right, we obtain (iii). The implication (i) ⇒ (iii′) is proved similarly (multiply $Z_R(w_0)$ on both sides of $\rho_R^{-1}(Z_R^{ш}(w_1)) = Z_R^*(w_1)$). By the same arguments we can show (ii) ⇒ (iv) and (ii) ⇒ (iv′). The properties (v) and (v′) are obtained, respectively, from (iv) and (iv′) by setting $w_1 = y^m$ and noting (for (iv)) $\mathrm{reg}_{ш}(y^m ш w) = 0$ which follows from $\mathrm{reg}_{ш}(y^m) = 0$ (since $y^m = y^{шm}/m!$). We have thus shown the following implications.

$$\begin{array}{ccccc} \text{(i)} & \Longrightarrow & \text{(ii)} & & \\ \Downarrow & & \Downarrow & & \\ \text{(iii)} & \Rightarrow & \text{(iv)} & \Rightarrow & \text{(v)} \end{array} \qquad \text{and} \qquad \begin{array}{ccccc} \text{(i)} & \Longrightarrow & \text{(ii)} & & \\ \Downarrow & & \Downarrow & & \\ \text{(iii}'\text{)} & \Rightarrow & \text{(iv}'\text{)} & \Rightarrow & \text{(v}'\text{)} \end{array}$$

Note that these implications are formal and do not depend on the precise definition of the algebraic renormalization map $\rho_R$ (i.e., they would remain true if $A_R(u)$ in (3.2) were replaced by any other power series). The real content of the theorem is the implications (v) ⇒ (i) and (v′) ⇒ (i), which we will prove in § 5 after some algebraic preliminaries.



https://doi.org/10.1112/S0010437X0500182X Published online by Cambridge University Press

Definition 1. The $\mathbb{Q}$-algebra $R$ and the map $Z_R : \mathfrak{H}^0 \to R$ with the FDS property have the EDS property if the eight equivalent properties of Theorem 2 are satisfied.

The content of Theorem 1 is now precisely that $(\mathbb{R}, Z)$ satisfies the EDS property, in the form (i) of Theorem 2, and hence also in the forms (ii)–(v) and (iii′)–(v′). In particular, we have

$$Z(\mathrm{reg}_{ш}(w_1 ш w_0 - w_1 * w_0)) = 0 \quad (\forall w_1 \in \mathfrak{H}^1,\ \forall w_0 \in \mathfrak{H}^0) \tag{3.4}$$

and

$$Z(\mathrm{reg}_{ш}(y^m * w_0)) = 0 \quad (\forall m \geqslant 1,\ \forall w_0 \in \mathfrak{H}^0). \tag{3.5}$$

The main conjecture about multiple zeta values is that the relations (3.4) suffice to give *all* linear relations over $\mathbb{Q}$ among MZVs. To state this formally, we introduce the universal EDS ring, as follows.

If $(R, Z_R)$ has the EDS property and $\varphi : R \to R'$ is a $\mathbb{Q}$-algebra homomorphism, then $(R', \varphi \circ Z_R)$ also has the EDS property. Clearly, there exists a universal algebra $R_{\mathrm{EDS}}$, namely the quotient of $\mathfrak{H}^0$ divided by the necessary relations, and a map $\varphi_R : R_{\mathrm{EDS}} \to R$ for any $(R, Z_R)$ with the EDS property that makes the diagram

$$\begin{tikzcd} \mathfrak{H}^0 \arrow[r] \arrow[dr, "Z_R"'] & R_{\mathrm{EDS}} \arrow[d, "\varphi_R"] \\ & R \end{tikzcd}$$

commute. Now for $(R, Z_R) = (\mathbb{R}, Z)$ we formulate the following.

Conjecture 1. The map $\varphi_{\mathbb{R}}$ is injective, i.e. the algebra of multiple zeta values is isomorphic to $R_{\mathrm{EDS}}$.

We briefly discuss here several possible different versions of the conjecture and their experimental status. We can consider statements of various strength, namely:

(1) the FDS and EDS relations suffice to give all relations among MZVs;

(2) the FDS relations alone suffice;

(3) only FDS and EDS with $\zeta(1)$ suffice;

(4) only double shuffle relations with $\zeta(n)$ $(n = 1, 2, 3, \dots)$ against finite zetas suffice.

(Here (3) means that we only need FDS and the formula (v) of Theorem 2 with $m = 1$, and (4) means that we need only Theorem 2(iv) with $w_1 = z_n = x^{n-1}y$.) Statement (1) is just the conjecture stated above. We checked up to $k = 13$ that the FDS and EDS relations suffice to give all expected relations, i.e. to reduce the dimension to the numerically and theoretically predicted value. (This is all one can hope to do, because actually proving the linear independence over $\mathbb{Q}$ of MZVs is out of reach.) The more optimistic statement (2) is wrong, as the weight 3 and weight 4 examples in § 2 showed. The intermediate statement (3) says that it is enough to use Theorem 1 for index sets $\mathbf{k}$ for which $Z^{ш}_{\mathbf{k}}(T)$ is at most linear in $T$ (in which case the map '$\rho$' is not needed in the formula (2.3)). The examples in § 2 verified this up to weight 4, and it has been checked up to weight 16 by Minh, Petitot *et al.* in Lille. (This of course also verifies the weaker conjecture (1) up to this weight.) Finally, statement (4), a different and particularly simple-looking strengthening of (1), holds up to weight 12 but fails at weight 13, where the relations in (4) suffice only to reduce the dimension of the space of MZVs to 17 instead of the value 16, which is what is obtained in this weight if one uses all the double shuffle relations.



https://doi.org/10.1112/S0010437X0500182X Published online by Cambridge University Press

## 4. Algebraic formulas

Recall that the algebra $\mathfrak{H}^1$ (with concatenation product) is free on the generators $z_k = x^{k-1}y$ ($k \geqslant 1$). Denote by $\mathfrak{z}$ the $\mathbb{Q}$-linear span of the $z_k$[1].

Proposition 2.

(i) *For $z \in \mathfrak{z}$ the map $\delta_z : \mathfrak{H}^1 \to \mathfrak{H}^1$ defined by*

$$\delta_z(w) := z * w - zw \quad (z \in \mathfrak{z}, w \in \mathfrak{H}^1) \tag{4.1}$$

*is a derivation, and these derivations all commute.*

(ii) *The above derivation $\delta_z$ extends to a derivation on all of $\mathfrak{H}$, with values on the generators given by*

$$\delta_z(x) = 0, \quad \delta_z(y) = (x + y)z \quad (z \in \mathfrak{z}). \tag{4.2}$$

*In particular, $\delta_z$ preserves $\mathfrak{H}^0$.*

*Proof.* (i) It is almost immediate from the definition of $*$ that we have

$$z * (ww') = (z * w)w' + w(z * w') - wzw' \tag{4.3}$$

for all $w, w' \in \mathfrak{H}^1$, and this is equivalent to the derivation property. Now if $z, z' \in \mathfrak{z}$ and $w \in \mathfrak{H}^1$ then from (4.1) and (4.3) and the associativity of $*$ we have

$$\delta_z(\delta_{z'}(w)) = z * (z' * w - z'w) - z\delta_{z'}(w) = \delta_{z*z'}(w) - z'\delta_z(w) - z\delta_{z'}(w),$$

which by virtue of the commutativity of $*$ is symmetric in $z$ and $z'$.

(ii) Define a derivation $\delta'_z$ on $\mathfrak{H}$ by the formulas (4.2). Then, for $k \geqslant 1$, we have

$$\delta'_z(z_k) = \delta'_z(x^{k-1}y) = x^{k-1}(x + y)z = z_k * z - zz_k,$$

so $\delta'_z$ agrees with $\delta_z$ on the generators of $\mathfrak{H}^1$ and hence on all of $\mathfrak{H}^1$. □

Proposition 3. *The vector space $\mathfrak{z}$ becomes a commutative and associative algebra with respect to the multiplication $\circ$ defined by*

$$z * z' = zz' + z'z + z \circ z' \quad (z, z' \in \mathfrak{z}). \tag{4.4}$$

*Proof.* We find immediately that $z_k \circ z_l = z_{k+l}$, from which these properties follow. Note that the map

$$\gamma : X\,\mathbb{Q}[[X]] \to \mathfrak{z}, \quad \gamma(X^k) = z_k \quad (k = 1, 2, \dots) \tag{4.5}$$

is an algebra isomorphism for $\circ$. We will use it occasionally later. □

We now have three different associative and commutative multiplications: the shuffle product ш (defined on all of $\mathfrak{H}$), the harmonic product $*$ (defined on $\mathfrak{H}^1 \subset \mathfrak{H}$) and the 'circle' product $\circ$ (defined on $\mathfrak{z} \subset \mathfrak{H}^1$). We denote by $\exp_{ш}$, $\exp_*$ and $\exp_\circ$ the corresponding exponential maps. The next two propositions describe some of their properties.

Proposition 4. *For $z \in \mathfrak{z}$ we have*

$$\exp_*(z) = (2 - \exp_\circ(z))^{-1}.$$

*(The inverse on the right is with respect to the concatenation product.)*

[1]We will not distinguish between direct sums and direct products, using the same letters $\mathfrak{H}^1$ and $\mathfrak{z}$ also for the completions and freely allowing infinite sums (of elements with grading going to infinity); this will not lead to any problems since $\mathfrak{H}^1$ is finite-dimensional in each weight.

*Proof.* Define a power series $f(u) \in \mathfrak{z}[[u]]$ by

$$f(u) = \exp_\circ(zu) - 1 = zu + z \circ z\frac{u^2}{2} + \cdots$$

(here the map $\exp_\circ$ has been extended to $\mathfrak{z}[[u]]$ in the obvious way, in accordance with the footnote above). Then $f'(u) = z \circ (1 + f(u))$, so

$$\begin{aligned} z * \frac{1}{1-f(u)} &= z * \sum_{n\geqslant 0} f(u)^n = \sum_{\alpha,\beta\geqslant 0} f(u)^\alpha z f(u)^\beta + \sum_{\alpha,\beta\geqslant 0} f(u)^\alpha (z \circ f(u)) f(u)^\beta \\ &= \sum_{\alpha,\beta\geqslant 0} f(u)^\alpha f'(u) f(u)^\beta = \frac{d}{du}\biggl(\sum_{n\geqslant 0} f(u)^n\biggr) = \frac{d}{du}\biggl(\frac{1}{1-f(u)}\biggr). \end{aligned}$$

For the second equality we have used the identity

$$z * w_1 w_2 \cdots w_n = \sum_{i=0}^{n} w_1 \cdots w_i z w_{i+1} \cdots w_n + \sum_{i=1}^{n} w_1 \cdots w_{i-1}(z \circ w_i) w_{i+1} \cdots w_n$$

for $z, w_i \in \mathfrak{z}$, which follows from the derivation property of $\delta_z$ and the formula $\delta_z(w) = wz + z \circ w (z, w \in \mathfrak{z})$. It follows that the function $F(u) := (1 - f(u))^{-1}$ satisfies $F'(u) = z * F(u)$ and $F(0) = 1$, so $F(u) = \exp_*(zu)$. □

COROLLARY 1. *For all $z \in \mathfrak{z}$ we have*

$$\exp_*(\log_\circ(1+z)) = \frac{1}{1-z}. \tag{4.6}$$

As an example, putting $z = z_k$ in (4.6) and applying the evaluation map $Z$, we obtain

COROLLARY 2. *For $k \geqslant 2$, we have the identity*

$$\exp\biggl(\sum_{n=1}^{\infty} (-1)^{n-1} \zeta(nk) \frac{u^n}{n}\biggr) = 1 + \sum_{n=1}^{\infty} \zeta(\underbrace{k, k, \ldots, k}_{n}) u^n.$$

PROPOSITION 5. *For $z, z' \in \mathfrak{z}$, $w \in \mathfrak{H}^1$ we have the identities*

$$\exp(\delta_z)(z') = (\exp_\circ(z) \circ z') \exp_*(z), \tag{4.7}$$

$$\exp(\delta_z)(w) = (\exp_*(z))^{-1}(\exp_*(z) * w). \tag{4.8}$$

*Proof.* Using the derivation property of $\delta_z$ and (4.4), we find

$$\begin{aligned} \delta_z(z'w) &= \delta_z(z')w + z'\delta_z(w) = (z \circ z' + z'z)w + z'(z * w - zw) \\ &= (z \circ z')w + z'(z * w) \end{aligned}$$

for $z, z' \in \mathfrak{z}$, $w \in \mathfrak{H}^1$. Now replacing $z'$ and $w$ by $z^{\circ\alpha} \circ z'$ and $z^{*\beta} * w$ (which again belong to $\mathfrak{z}$ and $\mathfrak{H}^1$, respectively) we obtain

$$\delta_z((z^{\circ\alpha} \circ z')(z^{*\beta} * w)) = (z^{\circ(\alpha+1)} \circ z')(z^{*\beta} * w) + (z^{\circ\alpha} \circ z')(z^{*(\beta+1)} * w)$$

and, hence, by induction

$$\delta_z^n(z'w) = \sum_{\alpha+\beta=n} \binom{n}{\alpha} (z^{\circ\alpha} \circ z')(z^{*\beta} * w)$$

or, dividing by $n!$ and summing over $n$,

$$\exp(\delta_z)(z'w) = (\exp_\circ(z) \circ z')(\exp_*(z) * w). \tag{4.9}$$

Setting $w = 1$ in (4.9) gives (4.7). Observe that if $\delta$ is a derivation which increases the weight, then by Leibniz's rule, $\exp(\delta)$ is a well-defined automorphism. With this, dividing (4.9) by (4.7) gives (4.8). □

Combining the last two propositions, we obtain the following proposition.

Proposition 6.

(i) *For* $z \in \mathfrak{z}$ *define* $\Phi_z : \mathfrak{H}^1 \to \mathfrak{H}^1$ *by*

$$\Phi_z(w) := (1-z)\left(\frac{1}{1-z} * w\right) \quad (z \in \mathfrak{z}, w \in \mathfrak{H}^1). \tag{4.10}$$

*Then* $\Phi_z$ *is an automorphism of* $\mathfrak{H}^1$ *and we have the identity*

$$\Phi_z(w) = \exp(\delta_t)(w), \quad \textit{where } t = \log_\circ(1+z) \in \mathfrak{z}. \tag{4.11}$$

*The collection of all* $\Phi_z$ $(z \in \mathfrak{z})$ *forms a commutative subgroup of* $\mathrm{Aut}(\mathfrak{H}^1)$*, with* $\Phi_z\Phi_{z'} = \Phi_{z+z'+z\circ z'}$ *for* $z, z' \in \mathfrak{z}$*. Equivalently, the map* $1 + X\,\mathbb{Q}[[X]] \to \mathrm{Aut}(\mathfrak{H}^1)$ *mapping* $1 + f(X)$ *to* $\Phi_{\gamma(f)}$*, with* $\gamma$ *defined by (4.5), is a homomorphism of groups.*

(ii) *The automorphism* $\Phi_z$ *of* $\mathfrak{H}^1$ *extends to an automorphism of all of* $\mathfrak{H}$*, with* $\Phi_z(x) = x$ *and* $\Phi_z(x+y) = (x+y)(1-z)^{-1}$*. In particular,* $\Phi_z$ *also induces an automorphism of* $\mathfrak{H}^0$*.*

*Proof.* (i) From (4.6) we have $\exp_*(t) = (1-z)^{-1}$, and substituting this into (4.8) (with $z$ replaced by $t$) immediately gives (4.11). The fact that $\Phi_z$ is an automorphism now follows from Proposition 2(i) and the fact that the exponential of a derivation is an automorphism, as noted before, and the last statements of the proposition follow immediately from formula (4.11), since for $t = \log_\circ(1+z)$ and $t' = \log_\circ(1+z')$ we have $\delta_t + \delta_{t'} = \delta_{t+t'}$ and $t + t' = \log_\circ(1 + z + z' + z \circ z')$ by the commutativity of $\circ$.

(ii) If we iterate $\delta_z$, then we find $\delta_z^n(x) = 0$ and $\delta_z^n(y) = (x+y)z^{*n}$ for all $n \geqslant 1$. (This is clear by induction, as $\delta_z((x+y)w) = (x+y)(zw + \delta_z(w)) = (x+y)(z * w)$ for any $w \in \mathfrak{H}^1$.) It follows that $\exp(\delta_z)(x) = x$, $\exp(\delta_z)(x+y) = (x+y)\exp_*(z)$. Combining this with formulas (4.11) and (4.6) we obtain the assertion. □

*Remark.* Written out, the fact that $\Phi_z$ is a homomorphism says that

$$(z^n) * (ww') = \sum_{\alpha+\beta=n} (z^\alpha * w)(z^\beta * w') - \sum_{\alpha+\beta=n-1} (z^\alpha * w)z(z^\beta * w')$$

for all $n \geqslant 0$, a generalization of (4.3) that can also be proved directly by a tedious induction using (4.3), (4.4) and the commutativity and associativity of $*$ and $\circ$.

The analogous (but simpler) result for the shuffle product is as follows.

Proposition 7. *Define the map* $d : \mathfrak{H} \to \mathfrak{H}$ *by* $d(w) = y ш w - yw$*. Then* $d$ *is a derivation and we have*

$$\exp(du)(w) = (1-yu)\left(\frac{1}{1-yu} ш w\right) \quad (w \in \mathfrak{H}). \tag{4.12}$$

*(Here,* $u$ *is a formal parameter.) In particular, we have*

$$\exp(du)(x) = x\frac{1}{1-yu} \quad \textit{and} \quad \exp(du)(y) = y\frac{1}{1-yu}. \tag{4.13}$$

*Proof.* That the $d$ is a derivation on $\mathfrak{H}$ is easily checked. For (4.12), we show by induction the identity

$$\frac{1}{m!}\,d^m(w) = y^m ш w - y(y^{m-1} ш w) \quad (m \geqslant 1, w \in \mathfrak{H}). \tag{4.14}$$



https://doi.org/10.1112/S0010437X0500182X Published online by Cambridge University Press

The case $m=1$ is the definition of $d$. Assuming the identity for $m$, we have

$$\begin{aligned}\frac{1}{(m+1)!}\,d^{m+1}(w) &= \frac{1}{m+1}d(y^m ш w - y(y^{m-1} ш w))\\ &= \frac{1}{m+1}[y ш (y^m ш w - y(y^{m-1} ш w)) - y(y^m ш w - y(y^{m-1} ш w))]\\ &= \frac{1}{m+1}[(m+1)y^{m+1} ш w - y^2(y^{m-1} ш w) - y(y ш y^{m-1} ш w)\\ &\quad - y(y^m ш w) + y^2(y^{m-1} ш w)]\\ &= y^{m+1} ш w - y(y^m ш w).\end{aligned}$$

Multiplying (4.14) by $u^m$ and summing over $m$ gives (4.12). Putting $w=x$ (respectively $w=y$) in (4.14), we have $d^m(x)/m! = xy^m$ (respectively $d^m(y)/m! = y^{m+1}$), which gives (4.13). $\square$

Corollary 3. *Let $\Delta_u$ be the automorphism of $\mathfrak{H}$ defined by*

$$\Delta_u = \exp(-du)\circ\Phi_{yu}. \tag{4.15}$$

*(Here $\circ$ denotes composition.) Then for $w\in\mathfrak{H}^1$ we have*

$$\frac{1}{1-yu} * w = \frac{1}{1-yu} ш \Delta_u(w). \tag{4.16}$$

*In particular, for $w_0\in\mathfrak{H}^0$ we have*

$$\mathrm{reg}_ш\left(\frac{1}{1-yu} * w_0\right) = \Delta_u(w_0). \tag{4.17}$$

*The images of the generators $x$ and $y$ of $\mathfrak{H}$ under $\Delta_u$ are given by*

$$\Delta_u(x) = x(1+yu)^{-1},\quad \Delta_u(y) = y + x(1+yu)^{-1}yu. \tag{4.18}$$

*Proof.* The first identity directly follows from (4.12) by replacing $w$ with $\Delta_u(w)$ and dividing both sides on the left by $1-yu$. To get the second, we put $w=w_0$ in the first and take $\mathrm{reg}_ш$ of both sides, noting that $\mathrm{reg}_ш(y^m)=0$ for $m\geqslant 1$. For the images of $x$ and $y$, we use Proposition 6(ii) and (4.13) to obtain

$$\Delta_u(x) = \exp(-du)(\Phi_{yu}(x)) = \exp(-du)(x) = x(1+yu)^{-1}$$

and

$$\begin{aligned}\Delta_u(x+y) &= (\exp(-du)\circ\Phi_{yu})(x+y) = \exp(-du)((x+y)(1-yu)^{-1})\\ &= x+y,\end{aligned}$$

and so

$$\Delta_u(y) = \Delta_u(x+y) - \Delta_u(x) = y + x(1+yu)^{-1}yu. \qquad \square$$

## 5. Regularization formulas

In this section we give various algebraic formulas for the regularization maps $\mathrm{reg}_ш^T$ and $\mathrm{reg}_ш$ (and also $\mathrm{reg}_*^T$ and $\mathrm{reg}_*$). Using these, we complete the proof of Theorem 2. We also apply these formulas to show that the well-known sum formula for MZVs is a formal consequence of the EDS relations.

Proposition 8. *For $w_0 = xw_0'\in\mathfrak{H}^0$ we have the regularization formula*

$$\mathrm{reg}_ш^T\left(\frac{1}{1-yu}w_0\right) = \exp(-du)(w_0)e^{Tu} = x\left(\frac{1}{1+yu} ш w_0'\right)e^{Tu}. \tag{5.1}$$



https://doi.org/10.1112/S0010437X0500182X Published online by Cambridge University Press

*In particular, for all $m \geqslant 0$ we have*

$$\operatorname{reg}_{ш}(y^m w_0) = \frac{(-1)^m}{m!}\, d^m(w_0) = (-1)^m\, x\,(y^m \,ш\, w_0'). \tag{5.2}$$

*Proof.* Putting $w = \exp(-du)(w_0)$ in Proposition 7 and multiplying $(1-yu)^{-1}$ from the left, we have (note the obvious identity $(1-yu)^{-1} = \exp_{ш}(yu)$)

$$\begin{aligned}\frac{1}{1-yu} w_0 &= \frac{1}{1-yu} \,ш\, \exp(-du)(w_0)\\ &= \exp_{ш}(yu) \,ш\, \exp(-du)(w_0).\end{aligned} \tag{5.3}$$

Taking $\operatorname{reg}_{ш}^T$ of this gives the first equality of (5.1). For the second, use the first equality in (4.13) with $u$ replaced by $-u$ and Proposition 7 applied to $w = xw_0'$ to get

$$\begin{aligned}\exp(-du)(w_0) &= \exp(-du)(x)\exp(-du)(w_0')\\ &= \left(x\frac{1}{1+yu}\right)\left((1+yu)\left(\frac{1}{1+yu} \,ш\, w_0'\right)\right)\\ &= x\left(\frac{1}{1+yu} \,ш\, w_0'\right).\end{aligned}$$

This proves (5.1). Comparing the coefficients of $u^m$ on both sides of this equation, we can rewrite it more explicitly as

$$\operatorname{reg}_{ш}^T(y^m w_0) = \frac{1}{m!}\sum_{l=0}^{m}(-1)^l\binom{m}{l}\, d^l(w_0)T^{m-l} = \sum_{l=0}^{m}(-1)^l x(y^l \,ш\, w_0')\frac{T^{m-l}}{(m-l)!}.$$

Equation (5.2) is the special case $T = 0$. ☐

Using this result, we can now complete the proof begun in § 3.

*Proof of Theorem 2.* We only have to prove the implications (v) $\Rightarrow$ (i) and (v$'$) $\Rightarrow$ (i), as the other parts of the theorem were proved after its statement in § 3. Noting $\Phi_{yu}(w_0) \in \mathfrak{H}^0$ if $w_0 \in \mathfrak{H}^0$, replace $w_0$ in Proposition 8 by $\Phi_{yu}(w_0)$ to get

$$\operatorname{reg}_{ш}^T\left(\frac{1}{1-yu}\Phi_{yu}(w_0)\right) = (\exp(-du)\circ\Phi_{yu})(w_0)e^{Tu} = \Delta_u(w_0)e^{Tu}$$

and thus

$$Z_R^{ш}\left(\frac{1}{1-yu}\Phi_{yu}(w_0)\right) = Z_R(\Delta_u(w_0))e^{Tu}. \tag{5.4}$$

On the other hand, the definition of $\Phi_{yu}$ and (4.6) give

$$\frac{1}{1-yu}\Phi_{yu}(w_0) = \frac{1}{1-yu} * w_0 = \exp_*(\log_\circ(1+yu)) * w_0. \tag{5.5}$$

Now we observe that

$$\begin{aligned}Z_R^*(\exp_*(\log_\circ(1+yu))) &= Z_R^*\left(\exp_*\left(\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n} z_n\, u^n\right)\right)\\ &= \exp\left(Tu - \sum_{n=2}^{\infty}\frac{(-1)^n}{n}Z_R(z_n)u^n\right) = A_R(u)^{-1}e^{Tu}\end{aligned} \tag{5.6}$$

with $A_R(u)$ as in (3.3). (This is the only point where the specific definition of the power series $A_R$ is used.) Hence, applying $\rho_R \circ Z_R^*$ to (5.5) yields

$$\rho_R \circ Z_R^*\left(\frac{1}{1-yu}\Phi_{yu}(w_0)\right) = e^{Tu}Z_R(w_0). \tag{5.7}$$



https://doi.org/10.1112/S0010437X0500182X Published online by Cambridge University Press

From (5.4) and (5.7) we get

$$\begin{aligned}(Z_R^{ш} - \rho_R \circ Z_R^{*})\left(\frac{1}{1-yu}\Phi_{yu}(w_0)\right) &= Z_R((\Delta_u - 1)(w_0))e^{Tu} \\ &= \left(\sum_{m=1}^{\infty} Z_R(\mathrm{reg}_{ш}(y^m * w_0))u^m\right)e^{Tu} \quad \text{(by (4.16))}.\end{aligned}$$

As $\Phi_{yu}$ acts as an automorphism of $\mathfrak{H}^0$ and as the components of $(1/1-yu)\mathfrak{H}^0$ span $\mathfrak{H}^1$, this shows the equivalence of statements (i) and (v) of Theorem 2. The equivalence of statements (i) and (v′) is shown in a similar manner. Applying $Z_R \circ \mathrm{reg}_*$ to (4.6) (with $z = yu$) and using (5.6) with $T = 0$, we have

$$Z_R\left(\mathrm{reg}_*\left(\frac{1}{1-yu}\right)\right) = A_R(u)^{-1}. \tag{5.8}$$

Now replace $w_0$ in (5.4) and (5.7) by $\Delta_u^{-1}(w_0)$ and take the difference to get

$$(Z_R^{ш} - \rho_R \circ Z_R^{*})\left(\frac{1}{1-yu}\exp(du)(w_0)\right) = Z_R(w_0 - \Delta_u^{-1}(w_0))e^{Tu}.$$

Multiplying $A_R(u)^{-1}$ on both sides of this and using (5.8), we obtain

$$\begin{aligned}&A_R(u)^{-1}(Z_R^{ш} - \rho_R \circ Z_R^{*})\left(\frac{1}{1-yu}\exp(du)(w_0)\right) \\ &= Z_R\left(\mathrm{reg}_*\left(\frac{1}{1-yu}\right) * (w_0 - \Delta_u^{-1}(w_0))\right)e^{Tu} \\ &= Z_R\left(\mathrm{reg}_*\left(\frac{1}{1-yu} * w_0 - \frac{1}{1-yu} * \Delta_u^{-1}(w_0)\right)\right)e^{Tu} \\ &= Z_R\left(\mathrm{reg}_*\left(\frac{1}{1-yu} * w_0 - \frac{1}{1-yu} ш\, w_0\right)\right)e^{Tu} \quad \text{(by (4.16))}.\end{aligned}$$

As before, this gives the equivalence of statements (i) and (v′) of Theorem 2. □

As a second application of Proposition 8, we show that the 'sum formula' for MZVs [Gra97], which states that the sum of all MZVs of fixed weight and depth is equal to the Riemann zeta value of that weight, is a consequence of the EDS relations.

PROPOSITION 9. *Denote by $S(k, m)$ the sum of all monomials in $\mathfrak{H}^0$ of weight $k$ and depth $m$. For any $k$ and $m$ with $k > m + 1 \geqslant 2$, we have*

$$(-1)^m \mathrm{reg}_{ш}(y^m * x^{k-m-1}y) = S(k, m+1) - S(k, m).$$

COROLLARY 4. *If $(R, Z_R : \mathfrak{H}^0 \to R)$ has the EDS property, then $Z_R(S(k, m)) = Z_R(x^{k-1}y)$ for $0 < m < k$. In particular, the sum of all MZVs of weight $k$ and depth $m$ is equal to $\zeta(k)$ for each value $m = 1, 2, \ldots, k-1$.*

*Proof.* Apply Theorem 2(v) to the statement of the proposition. □

*Proof of Proposition 9.* The harmonic product $y^m * x^{k-m-1}y$, which corresponds to the product $\zeta(\underbrace{1, 1, \ldots, 1}_{m})\zeta(k-m)$ of MZVs, is easily computed as

$$y^m * x^{k-m-1}y = \sum_{i=0}^{m} y^i x^{k-m-1} y^{m+1-i} + \sum_{j=0}^{m-1} y^j x^{k-m} y^{m-j}.$$

By (5.2), we then obtain

$$\begin{aligned}
\operatorname{reg}_{ш}(y^m * x^{k-m-1}y) &= \sum_{i=0}^{m}(-1)^i x(y^i ш x^{k-m-2}y^{m+1-i}) + \sum_{j=0}^{m-1}(-1)^j x(y^j ш x^{k-m-1}y^{m-j}) \\
&= x^{k-m-1}y^{m+1} + \sum_{i=1}^{m}(-1)^i x\{(y^i ш x^{k-m-2}y^{m-i})y + (y^{i-1} ш x^{k-m-2}y^{m+1-i})y\} \\
&\quad + x^{k-m}y^m + \sum_{j=1}^{m-1}(-1)^j x\{(y^j ш x^{k-m-1}y^{m-1-j})y + (y^{j-1} ш x^{k-m-1}y^{m-j})y\} \\
&= \sum_{i=0}^{m}(-1)^i x(y^i ш x^{k-m-2}y^{m-i})y + \sum_{i=0}^{m-1}(-1)^{i+1} x(y^i ш x^{k-m-2}y^{m-i})y \\
&\quad + \sum_{j=0}^{m-1}(-1)^j x(y^j ш x^{k-m-1}y^{m-1-j})y + \sum_{j=0}^{m-2}(-1)^{j+1} x(y^j ш x^{k-m-1}y^{m-1-j})y \\
&= (-1)^m x(y^m ш x^{k-m-2})y + (-1)^{m-1}x(y^{m-1} ш x^{k-m-1})y \\
&= (-1)^m(S(k,m+1) - S(k,m)).
\end{aligned}$$

An alternative way of deducing this identity is by making use of the automorphism $\Delta_u$ (note the relation (4.17)) and the formula (4.18), as follows:

$$\begin{aligned}
(-1)^m \operatorname{reg}_{ш}(y^m * x^{k-m-1}y) &= \text{the degree } k \text{ component of } \Delta_{-1}(x^{k-m-1}y) \\
&= \text{the degree } k \text{ component of } \left(x\frac{1}{1-y}\right)^{k-m-1}\left(y - x\frac{y}{1-y}\right) \\
&= \text{the degree } k \text{ component of } \left(x\frac{1}{1-y}\right)^{k-m-1} y - \left(x\frac{1}{1-y}\right)^{k-m} y \\
&= S(k,m+1) - S(k,m). \qquad \square
\end{aligned}$$

We end this section with a collection of formulas for the regularization map $\operatorname{reg}_{ш}^T$ and its harmonic analog $\operatorname{reg}_*^T$.

Proposition 10. *For $w_0 \in \mathfrak{H}^0$ we have*

$$\begin{aligned}
\operatorname{reg}_{ш}^T\left(\frac{1}{1-yu}w_0\right) &= \operatorname{reg}_{ш}\left(\frac{1}{1-yu}w_0\right)e^{Tu} \\
&= \left(\exp_{ш}(-yu) ш \left(\frac{1}{1-yu}w_0\right)\right)e^{Tu}
\end{aligned} \tag{5.9}$$

*and*

$$\begin{aligned}
\operatorname{reg}_*^T\left(\frac{1}{1-yu}w_0\right) &= \operatorname{reg}_*\left(\frac{1}{1-yu}w_0\right)e^{Tu} \\
&= \left(\exp_*(-yu) * \left(\frac{1}{1-yu}w_0\right)\right)e^{Tu}.
\end{aligned} \tag{5.10}$$

Corollary 5 (Explicit regularization formula). *Let $w \in \mathfrak{H}^1$ and write $w = y^m w_0$ with $m \geqslant 0$ and $w_0 \in \mathfrak{H}^0$. Then*

$$\operatorname{reg}_{ш}(w) = \sum_{i=0}^{m}(-1)^i y^i ш y^{m-i}w_0, \quad \operatorname{reg}_*(w) = \sum_{i=0}^{m}\frac{(-1)^i}{i!}y^{*i} * y^{m-i}w_0,$$

*and, conversely,*

$$w = \sum_{i=0}^{m} \mathrm{reg}_{ш}(y^{m-i}w_0) \, ш \, y^i, \quad w = \sum_{i=0}^{m} \frac{1}{i!}\mathrm{reg}_*(y^{m-i}w_0) * y^{*i}.$$

*Proof.* Equation (5.9) is essentially contained in Proposition 8 and its proof, as (5.1) and (5.3) give

$$\mathrm{reg}_{ш}\left(\frac{1}{1-yu}w_0\right) = \exp(-du)(w_0) = \exp_{ш}(-yu) \, ш \left(\frac{1}{1-yu}w_0\right).$$

For (5.10), we first combine (4.10) and (4.6) to obtain

$$\begin{aligned}\frac{1}{1-yu}w_0 &= \frac{1}{1-yu} * \Phi_{yu}^{-1}(w_0) = \exp_*(\log_\circ(1+yu)) * \Phi_{yu}^{-1}(w_0) \\ &= \exp_*(yu) * [\exp_*(\log_\circ(1+yu) - yu) * \Phi_{yu}^{-1}(w_0)]. \end{aligned} \tag{5.11}$$

As the expression in square brackets belongs to $\mathfrak{H}^0[[u]]$ by virtue of Proposition 6(ii), we have

$$\begin{aligned}\mathrm{reg}_*\left(\frac{1}{1-yu}w_0\right) &= \exp_*(\log_\circ(1+yu) - yu) * \Phi_{yu}^{-1}(w_0) \\ &= \exp_*(-yu) * \frac{1}{1-yu}w_0.\end{aligned}$$

From this, (5.10) follows. Finally, the corollary is obtained by specializing (5.9) and (5.10) to $T = 0$ and comparing the coefficients of $u^m$ on both sides either before or after $\bullet$-multiplying both sides by $\exp_\bullet(yu)$, where $\bullet = ш$ or $*$, and noting that $y^{шi} = i!y^i$. □

## 6. Derivation and double shuffle relations

Let $\mathrm{Der}(\mathfrak{H})$ be the Lie algebra of derivations of $\mathfrak{H}$ (with respect to the concatenation product, Lie algebra structure being defined by $[\partial, \partial'] := \partial\partial' - \partial'\partial$, as usual). Clearly, an element of $\mathrm{Der}(\mathfrak{H})$ is uniquely determined by the images of $x$ and $y$. Examples of elements of $\mathrm{Der}(\mathfrak{H})$ include the maps $\delta_z$ ($z \in \mathfrak{z}$) and $d$ introduced in Propositions 2 and 7. Let $\tau : \mathfrak{H} \to \mathfrak{H}$ be the involutory anti-automorphism that interchanges $x$ and $y$. If $\partial \in \mathrm{Der}(\mathfrak{H})$, then $\overline{\partial} := \tau\partial\tau$ is also an element of $\mathrm{Der}(\mathfrak{H})$. The involution $\tau$ preserves $\mathfrak{H}^0$ and the standard duality theorem for MZVs can be stated as $Z((1-\tau)(w_0)) = 0$ for any $w_0 \in \mathfrak{H}^0$.

For each integer $n \geqslant 1$, define the derivations $\partial_n$ and $D_n$ in $\mathrm{Der}(\mathfrak{H})$ by

$$\partial_n(x) = x(x+y)^{n-1}y, \quad \partial_n(y) = -x(x+y)^{n-1}y,$$
$$D_n(x) = 0, \quad D_n(y) = x^n y.$$

Each of these derivations preserves $\mathfrak{H}^1$ and $\mathfrak{H}^0$. As is easily seen by checking the images of the generators $x$ and $y$, the derivations in each of the three families $\{\partial_n\}$, $\{D_n\}$ and $\{\overline{D}_n\}$ commute with one another: $[\partial_m, \partial_n] = [D_m, D_n] = [\overline{D}_m, \overline{D}_n] = 0$ for any $m, n \geqslant 1$. For each $m \geqslant 0$, define the linear maps $\sigma_m$ and $\overline{\sigma}_m : \mathfrak{H} \to \mathfrak{H}$ as homogeneous components of degree $m$ of the homomorphisms

$$\sigma := \exp\left(\sum_{n=1}^{\infty} \frac{D_n}{n}\right), \quad \overline{\sigma} := \exp\left(\sum_{n=1}^{\infty} \frac{\overline{D}_n}{n}\right)(= \tau\sigma\tau),$$

so that

$$\sigma = \exp\left(\sum_{n=1}^{\infty} \frac{D_n}{n}\right) = \sum_{m=0}^{\infty} \sigma_m, \quad \overline{\sigma} = \exp\left(\sum_{n=1}^{\infty} \frac{\overline{D}_n}{n}\right) = \sum_{m=0}^{\infty} \overline{\sigma}_m.$$

We can use these endomorphisms to give two further collections of relations that are equivalent to the EDS relations.



https://doi.org/10.1112/S0010437X0500182X Published online by Cambridge University Press

Theorem 3. *Assume that $R$, $Z_R : \mathfrak{H}^0 \to R$ satisfy the FDS. Then the following three properties are equivalent:*

(i) *$(R, Z_R)$ satisfies the EDS;*

(ii) *$Z_R(\partial_n(w_0)) = 0$ for all $w_0 \in \mathfrak{H}^0$ and all $n$;*

(iii) *$Z_R((\sigma_m - \overline{\sigma}_m)(w_0)) = 0$ for all $w_0 \in \mathfrak{H}^0$ and all $m$.*

Combining this result (for $R = \mathbb{R}$) with Theorem 1, we obtain the following 'derivation relations' for MZVs.

Corollary 6. *One has $Z(\partial_n(w_0)) = 0$ for all $n \geqslant 1$ and all $w_0 \in \mathfrak{H}^0$.*

This corollary is a generalization of Hoffman's relation [Hof92] which is equivalent to the case $n = 1$. An alternative proof is given in [HO03].

Part (iii) of the theorem enables us to understand Ohno's relation [Ohn99] in light of the EDS relations. To see this we describe the map $\sigma_m$ more concretely as follows. Put

$$D = \sum_{n=1}^{\infty} \frac{D_n}{n}.$$

As $D(x) = 0$ and $D(y) = (x + x^2/2 + x^3/3 + \cdots)y = (-\log(1-x))y$, we have $D^n(x) = 0, D^n(y) = (-\log(1-x))^n y$ and hence $\sigma(x) = x$ and $\sigma(y) = \sum_{n=0}^{\infty}(1/n!)(-\log(1-x))^n y = (1-x)^{-1}y$. From this, we have

$$\begin{aligned}\sigma(x^{k_1-1}yx^{k_2-1}y\cdots x^{k_n-1}y) &= x^{k_1-1}(1-x)^{-1}yx^{k_2-1}(1-x)^{-1}y\cdots x^{k_n-1}(1-x)^{-1}y\\ &= \sum_{m=0}^{\infty}\sum_{\substack{e_1+e_2+\cdots+e_n=m\\ e_i\geqslant 0}} x^{k_1+e_1-1}yx^{k_2+e_2-1}y\cdots x^{k_n+e_n-1}y,\end{aligned}$$

namely,

$$\sigma_m(x^{k_1-1}yx^{k_2-1}y\cdots x^{k_n-1}y) = \sum_{\substack{e_1+e_2+\cdots+e_n=m\\ e_i\geqslant 0}} x^{k_1+e_1-1}yx^{k_2+e_2-1}y\cdots x^{k_n+e_n-1}y.$$

Ohno's relation [Ohn99] then states that $Z((\sigma_m - \sigma_m\tau)(w_0)) = 0$ for all $m \geqslant 1$, $w_0 \in \mathfrak{H}^0$. Under duality, this follows from Theorem 3(iii) and Theorem 1.

The crucial identities needed for the proof of Theorem 3 are given in the following result, which will be proved in a more general form in the next section.

Theorem 4.

(i) *The automorphism $\Delta_u$ defined in (4.15) is given in terms of the derivations $\partial_n$ by*

$$\Delta_u = \exp\biggl(\sum_{n=1}^{\infty}(-1)^n\frac{\partial_n}{n}u^n\biggr). \tag{6.1}$$

(ii) *Set $\partial = \sum_{n=1}^{\infty}\dfrac{\partial_n}{n}$. Then we have*

$$\exp(\partial) = \overline{\sigma}\sigma^{-1}. \tag{6.2}$$

*Proof of Theorem 3.* The equivalence of Theorem 3(ii) and Theorem 2(v) follows from (6.1) because $(\Delta_u - 1)(w_0) = \sum_{m=1}^{\infty}\mathrm{reg}_{ш}(y^m * w_0)u^m$ by (4.17). By (6.2), we have $\partial = \log(1 - (\sigma - \overline{\sigma})\sigma^{-1})$ and $\sigma - \overline{\sigma} = (1 - \exp(\partial))\sigma$. This gives the equivalence (ii) $\Leftrightarrow$ (iii) of Theorem 3. $\square$



https://doi.org/10.1112/S0010437X0500182X Published online by Cambridge University Press

## 7. Various derivations and automorphisms of $\mathfrak{H}$

In this section we study the various derivations and automorphisms being considered in more detail and more systematically, and give a proof of Theorem 4.

Set $\delta_n := \delta_{z_n}$. For the convenience of the reader, in Table 2 we give the various derivations that have been introduced (in Propositions 2, 7 and § 6). In this table and for the rest of this section, unlike the previous sections, $z$ denotes $x + y$.

Table 2. Various derivations.

| | $d$ | $\overline{d}$ | $D_n$ | $\overline{D_n}$ | $\delta_n$ | $\overline{\delta_n}$ | $\partial_n = -\overline{\partial_n}$ |
|---|---|---|---|---|---|---|---|
| $x$ | $xy$ | $x^2$ | $0$ | $xy^n$ | $0$ | $xy^{n-1}z$ | $xz^{n-1}y$ |
| $y$ | $y^2$ | $xy$ | $x^ny$ | $0$ | $zx^{n-1}y$ | $0$ | $-xz^{n-1}y$ |
| $z$ | $zy$ | $xz$ | $x^ny$ | $xy^n$ | $zx^{n-1}y$ | $xy^{n-1}z$ | $0$ |

Note that each of these derivations belongs not merely to $\mathrm{Der}(\mathfrak{H})$ = Lie algebra of all derivations of $\mathfrak{H}$, but to the subspace $\mathrm{Der}^+(\mathfrak{H})$ consisting of derivations that increase the weight or, equivalently, that induce the zero derivation on the associated graded space $\mathrm{Gr}(\mathfrak{H}) = \bigoplus(\mathfrak{H}^{(k)}/\mathfrak{H}^{(k+1)})$ of $\mathfrak{H}$, where $\mathfrak{H}^{(k)}$ denotes the subspace generated by all monomials in $x$ and $y$ of weight $\geqslant k$. The space $\mathrm{Der}^+(\mathfrak{H})$ is isomorphic via the exponential map to the subgroup $\mathrm{Aut}^1(\mathfrak{H})$ of $\mathrm{Aut}(\mathfrak{H})$ consisting of all automorphisms that induce the identity automorphism on $\mathrm{Gr}(\mathfrak{H})$ (i.e. automorphisms $\phi$ of $\mathfrak{H}$ such that $\phi(x) - x$ and $\phi(y) - y$ contain only monomials of weight $\geqslant 2$). We want to understand the automorphisms corresponding to all of the above derivations. In particular, the fact that the derivations in each of the four families $\{D_n\}$, $\{\overline{D_n}\}$, $\{\partial_n\}$ and $\{\delta_n\}$ commute with one another is more naturally interpreted in terms of the commutation of the corresponding automorphisms, as the set of automorphisms is closed under composition and the set of derivations is not. As a first step, we generalize the notation slightly. We define $D, \overline{D}, \partial, \delta : X\mathbb{Q}[[X]] \to \mathrm{Der}^+(\mathfrak{H})$ as the $\mathbb{Q}$-linear maps sending $X^n$ $(n \geqslant 1)$ to $D_n$, $\overline{D_n}$, $\partial_n$ and $\delta_n$, respectively, or alternatively as the maps sending $f(X) \in X\mathbb{Q}[[X]]$ to the derivations $D_f, \overline{D}_f, \partial_f, \delta_f$ defined on generators by

$$\begin{aligned} &D_f(x) = 0, \quad D_f(y) = f(x)y; \quad \overline{D}_f(y) = 0, \quad \overline{D}_f(x) = xf(y); \\ &\partial_f(z) = 0, \quad \partial_f(x) = x\frac{f(z)}{z}y; \quad \delta_f(x) = 0, \quad \delta_f(z) = z\frac{f(x)}{x}y. \end{aligned} \tag{7.1}$$

(Note that $\delta_f$ is just $\delta_{\gamma(f)}$, with $\gamma$ as in (4.5), and that $\overline{D}_f = \overline{D_f}$.) The corresponding automorphisms are described by the following proposition.

Proposition 11. *For $h(X) \in 1 + X\mathbb{Q}[[X]]$ let $\sigma_h, \overline{\sigma}_h, \Delta_h, \Psi_h \in \mathrm{Aut}^1(\mathfrak{H})$ be the automorphisms defined by the following action on generators:*

$$\sigma_h(x) = x, \quad \sigma_h(y) = h(x)y, \tag{7.2}$$

$$\overline{\sigma}_h(y) = y, \quad \overline{\sigma}_h(x) = xh(y), \tag{7.3}$$

$$\Delta_h(z) = z, \quad \Delta_h(x) = x\left(1 + \frac{h(z)-1}{z}y\right)^{-1}, \tag{7.4}$$

$$\Psi_h(x) = x, \quad \Psi_h(z) = z\left(1 - \frac{h(x)-1}{x}y\right)^{-1}. \tag{7.5}$$

*Then each of the maps $h \mapsto \sigma_h, \overline{\sigma}_h, \Delta_h, \Psi_h$ is a homomorphism from $1 + X\mathbb{Q}[[X]]$ to $\mathrm{Aut}(\mathfrak{H})$, and they are related to the derivations $D_f$, $\overline{D}_f$, $\partial_f$ and $\delta_f$ by*

$$\sigma_h = \exp(D_f), \quad \overline{\sigma}_h = \exp(\overline{D}_f), \quad \Delta_h = \exp(-\partial_f), \quad \Psi_h = \exp(\delta_f) \tag{7.6}$$

*for all $h \in 1 + X\mathbb{Q}[[X]]$, where $f = \log(h) \in X\mathbb{Q}[[X]]$.*

*Proof.* Let $g(X)$ and $h(X)$ be two elements of $1 + X\mathbb{Q}[[X]]$. It is easy to verify the equations $\sigma_{gh} = \sigma_g\sigma_h$, $\overline{\sigma}_{gh} = \overline{\sigma}_g\overline{\sigma}_h$, and $\sigma_h = \exp(D_f)$, $\overline{\sigma}_h = \exp(\overline{D}_f)$ with $f = \log(h)$. If we use the map $\gamma$ in (4.5) then $\gamma(X^k) = x^{k-1}y$ implies that $\gamma(h(X) - 1) = ((h(x) - 1)/x)\, y$ and, hence, we have $\Psi_h = \Phi_{\gamma(h-1)}$. The proposition for $\Psi_h$ then follows from Proposition 6(i), and this also implies the statement for $\Delta_h$ because of the identities $\Delta_h = \varepsilon\Psi_h\varepsilon$ and $-\partial_f = \varepsilon\delta_f\varepsilon$, where $\varepsilon$ is the involution of $\mathfrak{H}$ interchanging $x$ and $z$ and sending $y$ to $-y$. □

We can now relate the above derivations and automorphisms to those previously studied and give the proof of Theorem 4. First, note that the automorphism $\Delta_{1+uX}$ is nothing other than the automorphism $\Delta_u$ in Theorem 4(i) defined by $\Delta_u = \exp(-du) \circ \Phi_{yu}$ in § 4, as is seen by comparing (4.18) and (7.4). The formula of Theorem 4(i) then follows from the third equation of (7.6) with $h = 1+uX$ and $f = \log(1+uX)$. On the other hand, the automorphisms denoted $\sigma$ and $\overline{\sigma}$ in § 6 are $\sigma_{1-X}^{-1}$ and $\overline{\sigma}_{1-X}^{-1}$ in our notation. Part (ii) of Theorem 4 then follows from the special case $u = -1$ of part (i) and the formula $\Delta_{1+uX} = \overline{\sigma}_{1+uX}^{-1} \cdot \sigma_{1+uX}$, whose proof is immediate from the definitions (7.2)–(7.4) by checking the action on generators.

In the rest of the section, we show that the Lie algebra generated by the four derivations $d, \overline{d}, \overline{d} - D_1$, and $d - \overline{D_1}$ contain all derivations considered so far, and discuss some properties of this Lie algebra.

Proposition 12. *For all $n \geqslant 1$ we have the commutation formulas*

$$[\overline{d}, \delta_n] = n\delta_{n+1}, \quad [d, \overline{\delta_n}] = n\overline{\delta_{n+1}}, \tag{7.7}$$

$$[\overline{d}, D_n] = nD_{n+1}, \quad [d, \overline{D_n}] = n\overline{D_{n+1}}. \tag{7.8}$$

*Proof.* Since in each case both sides of the equation to be proved are derivations on all of $\mathfrak{H}$, it suffices to show that they agree on the generators $x$ and $y$. We have

$$[\overline{d}, \delta_n](x) = \overline{d}(0) - \delta_n(x^2) = 0 = n\delta_{n+1}(x)$$

and

$$\begin{aligned}[\overline{d}, \delta_n](y) &= \overline{d}(zx^{n-1}y) - \delta_n(xy)\\ &= xzx^{n-1}y + (n-1)zx^n y + zx^n y - xzx^{n-1}y\\ &= nzx^n y = n\delta_{n+1}(y).\end{aligned}$$

The proofs of the other three identities are similar and will be omitted. □

Theorem 5. *Let $\mathfrak{L}_4$ be the Lie subalgebra of* $\mathrm{Der}(\mathfrak{H})$ *generated by the four elements $d$, $\overline{d}$, $D = \overline{d} - D_1$ and $\overline{D} = d - \overline{D_1}$, and $\mathfrak{L}_3$ the Lie subalgebra of $\mathfrak{L}_4$ generated by $d$, $\overline{d}$, and $D - \overline{D}$. Then:*

(i) $\mathfrak{L}_3 = \overline{\mathfrak{L}_3}$, $\mathfrak{L}_4 = \overline{\mathfrak{L}_4}$;

(ii) $\mathfrak{L}_3$ *contains* $\partial_n$, $\delta_n$, $\overline{\delta_n}$ *for all* $n \geqslant 1$;

(iii) $\mathfrak{L}_4$ *also contains* $D_n$ *and* $\overline{D_n}$ *for all* $n \geqslant 1$.

*Proof.* Property (i) is clear from the definition. We have

$$\delta_1 = \overline{d} - D + \overline{D} \tag{7.9}$$

(both sides are derivations sending $x$ to 0 and $y$ to $zy$) and $D_1 = \overline{d} - D$ (by the definition of $D$), so $\delta_1 \in \mathfrak{L}_3, D_1 \in \mathfrak{L}_4$, and also $\overline{\delta_1} \in \mathfrak{L}_3, \overline{D_1} \in \mathfrak{L}_4$ by (i). Proposition 12 and induction over $n$ then show that the derivations $\delta_n$, $\overline{\delta_n}$ belong to $\mathfrak{L}_3$ and $D_n$, $\overline{D_n}$ to $\mathfrak{L}_4$ for all $n \geqslant 1$. Finally, combining (6.1), (4.15) and (4.11) we have

$$\exp\biggl(\sum_{n=1}^{\infty} \frac{\partial_n}{n} u^n\biggr) = \Delta_{-u} = \exp(du)\exp(\delta_{\log_\circ(1-yu)}) = \exp(du)\exp\biggl(-\sum_{n=1}^{\infty}\frac{\delta_n}{n}u^n\biggr). \tag{7.10}$$



https://doi.org/10.1112/S0010437X0500182X Published online by Cambridge University Press

From this, we have $\partial_n \in \mathfrak{L}_3$ for all $n$ by the Baker–Campbell–Hausdorff formula. $\square$

A more explicit formula for $\partial_n$ is obtained as follows. We differentiate (7.10) with respect to $u$ to get

$$\biggl(\sum_{n=1}^{\infty} \partial_n u^{n-1}\biggr) \exp\biggl(\sum_{n=1}^{\infty} \frac{\partial_n}{n} u^n\biggr) = d \exp(du) \exp\biggl(-\sum_{n=1}^{\infty} \frac{\delta_n}{n} u^n\biggr)$$
$$- \exp(du)\biggl(\sum_{n=1}^{\infty} \delta_n u^{n-1}\biggr) \exp\biggl(-\sum_{n=1}^{\infty} \frac{\delta_n}{n} u^n\biggr)$$

and hence

$$\sum_{n=1}^{\infty} \partial_n u^{n-1} = d - \exp(du)\biggl(\sum_{n=1}^{\infty} \delta_n u^{n-1}\biggr) \exp(-du)$$
$$= d - \exp\Bigl(\mathrm{ad}(d)u\Bigr)\biggl(\sum_{n=1}^{\infty} \delta_n u^{n-1}\biggr).$$

Here, by (7.7), we have

$$\sum_{n=1}^{\infty} \delta_n = \sum_{n=0}^{\infty} \frac{\mathrm{ad}(\overline{d})^n}{n!}(\delta_1) = \exp(\mathrm{ad}(\overline{d}))(\delta_1).$$

Hence, we obtain

$$\sum_{n=1}^{\infty} \partial_n = d - \exp(\mathrm{ad}(d)) \exp(\mathrm{ad}(\overline{d}))(\delta_1).$$

Written out, this gives an explicit description of $\partial_n$ in terms of the generators of $\mathfrak{L}_3$:

$$\partial_n = \begin{cases} d - \overline{d} + D - \overline{D} & \text{if } n = 1, \\ \displaystyle\sum_{i+j=n-1} \frac{\mathrm{ad}(d)^i}{i!} \frac{\mathrm{ad}(\overline{d})^j}{j!}(D - \overline{D} - \overline{d}) & \text{if } n \geqslant 2. \end{cases}$$

The following proposition gives further relations in the Lie algebra $\mathfrak{L}_4$. As we do not use this proposition later (except that the relations do contribute to reduce the dimensions of $\mathfrak{L}_4^{(n)}$ as in the table below), we omit the proof.

Proposition 13.

(i) *The two derivations $D$ and $\overline{D}$ commute.*

(ii) *We have*

$$\frac{\mathrm{ad}(\overline{d})^n}{n!}(d) = \biggl[\mathrm{ad}(d), \frac{\mathrm{ad}(\overline{d})^{n-1}}{(n-1)!}\biggr](D) \quad (n > 1),$$
$$\frac{\mathrm{ad}(d)^n}{n!}(\overline{d}) = \biggl[\mathrm{ad}(\overline{d}), \frac{\mathrm{ad}(d)^{n-1}}{(n-1)!}\biggr](\overline{D}) \quad (n > 1).$$

Below we give a short table of the dimensions of the graded piece $\mathfrak{L}_3^{(n)}$ and $\mathfrak{L}_4^{(n)}$ of $\mathfrak{L}_3$ and $\mathfrak{L}_4$ for $n \leqslant 8$ and, for comparison, the same data for the free Lie algebras $\mathfrak{F}_3$ and $\mathfrak{F}_4$ on 3 and 4 generators, respectively.



https://doi.org/10.1112/S0010437X0500182X Published online by Cambridge University Press

| $n$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| $\mathfrak{L}_3^{(n)}$ | 3 | 3 | 5 | 8 | 14 | 23 | 43 | 75 |
| $\mathfrak{F}_3^{(n)}$ | 3 | 3 | 8 | 18 | 48 | 116 | 312 | 810 |
| $\mathfrak{L}_4^{(n)}$ | 4 | 5 | 8 | 13 | 22 | 37 | 66 | 115 |
| $\mathfrak{F}_4^{(n)}$ | 4 | 6 | 20 | 60 | 204 | 670 | 2340 | 8160 |

We end this section by presenting some results which seem interesting to us, even though there is no obvious application to the structure of the double shuffle relations. Each of the four families $\{\delta_n\}$, $\{\overline{\delta_n}\}$, $\{D_n\}$, and $\{\overline{D_n}\}$ has the property that their elements mutually commute and, by Proposition 12, that they have the form $\mathrm{ad}(X)^n(Y)$ for some $X, Y \in \mathrm{Der}(\mathfrak{H})$. We give two results, one generalizing and one reinterpreting this phenomenon.

Proposition 14. *Let $a, b$ be a basis of the two-dimensional space spanned by $x$ and $y$, and $\delta$ a derivation that sends $a$ to 0 and $b$ to any element of the form $\kappa b^2 + \lambda ba + \mu ab + \nu a^2$. Suppose that $\theta$ is another derivation sending $a$ to $a^2$ and $b$ to the sum of $ab$ and any linear combination of $[a, b]$ and $\delta(b)$. Then the derivations $\delta'_n$ defined by $\delta'_n = \mathrm{ad}(\theta)^{n-1}(\delta)/(n-1)!$ are given on generators by $\delta'_n(a) = 0$, $\delta'_n(b) = \kappa ba^{n-1}b + \lambda ba^n + \mu a^n b + \nu a^{n+1}$ and all commute with one another.*

The proof is routine and is omitted.

If we set $a = z (= x+y)$, $b = y$, $\delta(b) = b^2 - ab (= -xy)$ and $\theta(b) = ab - [a, b]/2$ in the proposition, then we have $\delta = \partial_1$ and $\delta'_n = \partial_n$, thereby proving not only that the elements $\{\partial_n\}$ mutually commute, as remarked at the beginning of § 7, but also that they have the form $\mathrm{ad}(\theta)^{n-1}(\partial_1)/(n-1)!$. Note, however, that the derivation $\theta$ in this case (defined by $\theta(x) = (xz+zx)/2, \theta(y) = (yz+zy)/2$) does not belong to the Lie algebra $\mathfrak{L}_4$.

Definition 2. Two elements $X$ and $Y$ in a Lie algebra over a field of characteristic 0 *semi-commute* if the power series $\exp(tX)\exp(-tY)$ in (the completion of) the universal enveloping algebra commute with each other for different $t$, i.e. if all of the coefficients of this power series commute with each other. Clearly, if $X$ and $Y$ commute, then they semi-commute.

Proposition 15. *Let $X$ and $\delta$ be two elements in a Lie algebra over a field $K$ of characteristic 0. Then the following three statements are equivalent.*

(i) *All of the elements $\mathrm{ad}(X)^n(\delta)$ $(n = 0, 1, 2, \dots)$ commute with one another.*

(ii) *The elements $X$ and $X - \delta$ semi-commute.*

(iii) *Any two elements of $X + K\delta$ semi-commute.*

*Proof.* We put $\delta_n := \mathrm{ad}(X)^{n-1}(\delta)/(n-1)!$ $(n = 1, 2, 3, \dots)$. Assume statement (i). Then we have

$$\exp\left(\sum_{n=1}^{\infty} \frac{\delta_n}{n} t^n\right) = \exp(tX)\exp(-t(X-\delta)),$$

as both sides satisfy the differential equation $f'(t) = \exp(tX)\delta\exp(-tX)f(t)$ and $f(0) = 1$. This implies statement (ii). Conversely, assume statement (ii) and set

$$f(t) = \exp(tX)\exp(-t(X-\delta)).$$

We calculate the derivative:

$$f'(t) = \exp(tX)\delta\exp(-(X-\delta)) = Z(t)f(t),$$

where

$$Z(t) = \exp(tX)\delta\exp(-tX) = \exp(\mathrm{ad}(X)t)(\delta) = \sum_{n=1}^{\infty} \delta_n t^{n-1}.$$

We want to show that if $f(t)$ and $f(u)$ commute for all $t$ and $u$, then $Z(t)$ and $Z(u)$ commute for all $t$ and $u$ (i.e. all $\delta_n$ commute). Differentiating the equation $f(t)f(u) = f(u)f(t)$ with respect to $u$, we find $f(t)Z(u) = Z(u)f(t)$. Differentiating this with respect to $t$, we obtain $Z(t)Z(u) = Z(u)Z(t)$. Thus we have shown the equivalence of (i) and (ii). However, then (i) and (iii) are also equivalent, as (i) is invariant under $\delta \to -c\delta$ for any $c \in K^*$. $\square$

*Remark.* (i) If $X$ and $Y$ are two elements of a Lie algebra that semi-commute with each other, then the commutation identity is

$$e^{uX}e^{-(u+v)Y}e^{vX} = e^{-vY}e^{(u+v)X}e^{-uY}$$

or, more symmetrically,

$$e^{aX}e^{bY}e^{cX}e^{aY}e^{bX}e^{cY} = 1 \quad \text{if } a+b+c=0,$$

which is somewhat reminiscent of the so-called Yang–Baxter equation.

(ii) Let $L_0$ be the quotient of the free Lie algebra on two letters $X$ and $\delta$ by the relation that all $\mathrm{ad}(X)^n(\delta)$ commute or, equivalently, setting $\delta_n = \mathrm{ad}(X)^{n-1}(\delta)/(n-1)!$, the Lie algebra with basis $(X, \delta_1, \delta_2 \ldots)$ and brackets given by $[X, \delta_n] = n\delta_{n+1}$, $[\delta_m, \delta_n] = 0$. As this is a graded Lie algebra (with $X$ and $\delta$ of weight 1, and $\delta_n$ of weight $n$), it extends naturally to a slightly larger Lie algebra $L_1 = L_0 + K \cdot H$ where $[H, X] = X$ and $[H, \delta_n] = n\delta_n$. Proposition 12 then states that there are several natural copies of $L_1$ in the endomorphisms of $\mathfrak{H}$, and Proposition 15 gives a nice interpretation of embeddings of $L_1$ into any Lie algebra in terms of semi-commuting elements. It is therefore of interest to note that the same Lie algebra $L_1$ has occurred in the work of Connes and Moscovici in connection with questions about cyclic homology, foliations and, more recently, the so-called 'Rankin–Cohen brackets' in the theory of modular forms [CM04a], [CM04b]. However, while our copies of $L_1$ act on $\mathfrak{H}$ as derivations, and hence can be thought of as Hopf algebra representations of the standard Hopf algebra associated with the Lie algebra $L_1$ (i.e. the Hopf algebra whose underlying algebra is the universal enveloping algebra of $L_1$ and whose coproduct is defined by requiring the generators $X$, $\delta$ and $H$, and hence all elements of $L_1$, to be primitive), Connes and Moscovici consider a 'twisted' Hopf algebra structure that has the same underlying algebra and in which $H$ and $\delta$ are still primitive, but with $\Delta(X) = X \otimes 1 + \delta \otimes H + 1 \otimes X$. It would be very interesting to discover whether there is any deeper reason for the occurrence of the same Lie algebra $L_1$ in these very different contexts, and whether the twisted Hopf structure considered by Connes and Moscovici plays any role in the context of MZVs.

## 8. Linearized double shuffle relations

In this section, we fix the depth $n$ and look at the (extended) double shuffle relation modulo elements of lower depth and products. This amounts to 'linearizing' the double shuffle relation and reduces the problem of finding an upper bound (and conjecturally exact value) for the number of generators of the $\mathbb{Q}$-algebra of MZVs of given weight $k$ and depth $n$ to the solution of an elementary, but hard, problem of linear algebra. Some consequences of this reduction for general $n$ are given at the end of this section. A more detailed discussion of the special cases $n = 2$ and 3 as well as some general results will be given in a subsequent paper.

Let $\mathcal{Z} = \bigoplus_{k \geqslant 0} \mathcal{Z}_k$ be the graded algebra generated by MZVs over $\mathbb{Q}$, where $\mathcal{Z}_k$ is the $\mathbb{Q}$-vector space generated by MZVs of weight $k$. The space $\mathcal{Z}_k$ has a natural filtration $\mathcal{Z}_k = \bigcup_{n \geqslant 0} \mathcal{Z}_k^{(n)}$, where $\mathcal{Z}_k^{(n)}$ is the $\mathbb{Q}$-vector space spanned by MZVs of weight $k$ and depth $\leqslant n$, and setting $\mathcal{Z}^{(n)} = \bigoplus_{k\geqslant 0} \mathcal{Z}_k^{(n)}$ gives a corresponding filtration $\mathcal{Z} = \bigcup_{n\geqslant 0} \mathcal{Z}^{(n)}$ on the whole algebra $\mathcal{Z}$. Let $\mathcal{I} = \bigoplus_{k\geqslant 1} \mathcal{Z}_k$ be the augumentation ideal of $\mathcal{Z}$ and $\mathcal{I}^2$ its square. The quotient space $\mathcal{T} = \mathcal{I}/\mathcal{I}^2$ also inherits the grading and filtration. All products of MZVs become trivial in $\mathcal{T}$, so the dimension of $\mathcal{T}_k$, the weight $k$ component of $\mathcal{T}$, coincides with the number $D_k$ of algebra generators of $\mathcal{Z}$ in weight $k$. We can also consider the bigraded vector space $\mathcal{M}$ associated with the graded filtered space $\mathcal{T}$:

$$\mathcal{M} = \bigoplus_{k,\, n\geqslant 1} \mathcal{M}_k^{(n)}, \quad \mathcal{M}_k^{(n)} = \mathcal{T}_k^{(n)}/\mathcal{T}_k^{(n-1)} \simeq \mathcal{Z}_k^{(n)}/(\mathcal{Z}_k^{(n-1)} + \mathcal{Z}_k^{(n)} \cap \mathcal{I}^2).$$

Then clearly the dimension $D_{k,n}$ of $\mathcal{M}_k^{(n)}$ equals the number of algebra generators of $\mathcal{Z}$ of weight $k$ and depth $n$, and we have $D_k = \sum_{n=1}^{k-1} D_{k,n}$.

*Remark.* There is a conjectural formula giving these dimensions $D_{k,n}$, due to Broadhurst and Kreimer [BK97]. As it is too beautiful to omit, but we did not want to interrupt the text here, we have reproduced this formula in the Appendix.

Our object is to introduce certain vector spaces $DSh_n(d)$ $(n, d > 0)$ whose dimensions give upper bounds for (and are conjecturally equal to) the numbers $D_{n+d,n}$, and then to discuss the calculation of these spaces. Let $\mathfrak{S}_n$ denote the symmetric group of order $n$ and $\mathcal{R} = \mathcal{R}_n = \mathbb{Z}[\mathfrak{S}_n]$ its group ring. We denote by $V_n$ the space $\mathbb{Q}[x_1, \dots, x_n]$ of polynomials in $n$ variables with rational coefficients and by $V_n(d)$ its subspace of homogeneous polynomials of degree $d$. The group $\mathfrak{S}_n$ acts on these spaces in a natural way by $(f|\sigma)(x_1, \dots, x_n) = f(x_{\sigma^{-1}(1)}, \dots, x_{\sigma^{-1}(n)})$. Note that this is a right action, i.e. $f|(\sigma\tau) = (f|\sigma)|\tau$ for all $\sigma$ and $\tau$ in $\mathfrak{S}_n$. This action extends to an action of $\mathcal{R}$ in the standard way by $f|(\sum a_i\sigma_i) = \sum a_i(f|\sigma_i)$.

For each integer $l$ with $1 \leqslant l \leqslant n-1$ we define the $l$th *shuffle element* $sh_l = sh_l^{(n)}$ in the group ring $\mathcal{R}$ by

$$sh_l^{(n)} = \sum_{\substack{\sigma\in\mathfrak{S}_n \\ \sigma(1)<\cdots<\sigma(l) \\ \sigma(l+1)<\cdots<\sigma(n)}} \sigma.$$

We denote by $\mathfrak{I}^{(n)} = sh_1\mathcal{R} + \cdots + sh_{n-1}\mathcal{R}$ the right ideal in $\mathcal{R}$ generated by all of these elements and, for any (right) representation $V$ of $\mathfrak{S}_n$, define the 'shuffle subspace' $Sh(V)$ of $V$ by

$$Sh(V) = \operatorname{Ker}(\mathfrak{I}^{(n)}, V) = \bigcap_{l=1}^{n-1} \operatorname{Ker}(sh_l^{(n)}, V),$$

the space of elements of $V$ annihilated by the ideal $\mathfrak{I}^{(n)}$. In particular, we have the *shuffle space* $Sh_n := Sh(V_n) \subset V_n$ and its homogeneous part of degree $d$, $Sh_n(d) := Sh(V_n(d)) \subset V_n(d)$.

We remark that the dimension of $Sh(V)$ can be computed for any $\mathfrak{S}_n$-module $V$ in a nice way using the theory of representation of finite groups. The following proposition gives explicit formulas for this dimension.

Proposition 16. *Let $V$ be any representation of $\mathfrak{S}_n$ and let $C \in \mathfrak{S}_n$ be an element of order $n$ (i.e. a cyclic permutation of $1, \dots, n$). Then we have the two formulas*

$$\dim Sh(V) = \frac{1}{n} \sum_{m|n} \mu\left(\frac{n}{m}\right) \operatorname{tr}(C^m, V), \tag{8.1}$$

*and*

$$\dim Sh(V) = \frac{1}{\varphi(n)} \sum_{m|n} \mu\left(\frac{n}{m}\right) \dim(V^{C^m}), \tag{8.2}$$

*where $\mu$ and $\varphi$ denote the Möbius function and the Euler function respectively. In particular,*

$$\dim Sh_n(d) = \frac{1}{n} \sum_{e|(n,d)} \mu(e) \binom{n/e + d/e - 1}{d/e}. \tag{8.3}$$

Actually, we proved this in the reverse direction, first proving the special case (8.3) by identifying the space $Sh_n(d)$ with the graded part of degree $d$ and weight $n+d$ of the free Lie algebra on infinitely many generators of weights $1, 2, \dots$, then noting that (8.1) coincides with (8.3) for $V = V_n(d)$, and that this suffices for the general case because the representations $V_n(d)(d = 0, 1, 2, \dots)$ span the Grothendieck group of representations of $\mathfrak{S}_n$ over $\mathbb{Q}$; and finally deducing (8.2) from (8.1) by using the formula $\dim(V^G) = |G|^{-1} \sum_{g \in G} \mathrm{tr}(g, V)$ which holds for any representation $V$ of a group $G$. In any case we omit the details of the proofs as we will not use any of these formulas and there must exist a simpler proof anyway, undoubtedly already in the literature.

The space we are actually interested in, however, is a subspace of $Sh_n(d)$ that is much harder to compute. To define it, we first extend the action of $\mathfrak{S}_n$ on $V_n$ to an action of $\Gamma_n = GL_n(\mathbb{Z})$ on $V_n$ by setting

$$(f|S)(x_1, \dots, x_n) := f((x_1, \dots, x_n) \cdot S^{-1}) \quad (S \in \Gamma_n),$$

which agrees with the previous definition for $\mathfrak{S}_n$ if the elements of the symmetric group are identified with permutation matrices in $\Gamma_n$ in the usual way. In particular, we will be interested in the element $P$ of $\Gamma_n$ given by

$$P = \begin{pmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 1 & & & \\ 1 & 1 & & \\ \vdots & & \ddots & \\ 1 & 1 & \dots & 1 \end{pmatrix}.$$

We now define the *double shuffle subspace* $DSh_n$ of $V_n$ as the intersection

$$\begin{aligned} DSh_n &= Sh_n \cap Sh_n | P^{-1} \\ &= \{ f \in V_n \mid f|sh_l = f^\sharp | sh_l = 0 \text{ for } l = 1, \dots, n-1 \}, \end{aligned}$$

where for any polynomial $f \in V_n$ we have set $f^\sharp = f|P$ or explicitly

$$f^\sharp(x_1, \dots, x_n) = f(x_1 + x_2 + \dots + x_n, x_2 + \dots + x_n, \dots, x_{n-1} + x_n, x_n).$$

For an element $S$ in $\mathbb{Z}[GL_n(\mathbb{Z})]$, we define $S^\sharp := PS\,P^{-1}$. The condition $f^\sharp | sh_l = 0$ is then equivalent to $f|sh_l{}^\sharp = 0$. We write $DSh_n(d)$ for the degree $d$ part of $DSh_n$.

As already mentioned, the double shuffle space $DSh_n(d)$ is much harder to compute than the single shuffle space $Sh_n(d)$, the reason being that it is defined in terms of the action of an infinite rather than a finite group and therefore cannot be computed using the theory of characters of finite groups. Some information about the dimension of $DSh_n(d)$ can be obtained by looking at finite subgroups or quotients of the subgroup of $\Gamma_n$ generated by $\mathfrak{S}_n$ and $P$ and applying the theory of characters. A few results of this nature, both for general $n$ and for the special cases $n = 2$ and 3, will be discussed at the end of the section. Before doing this, we prove a theorem that states that the generating function of MZVs belong to the double shuffle space and, as a consequence, bounds the number of linearly independent MZVs in terms of the dimensions of the vector spaces $DSh_n(d)$.

Let $\mathbf{k}$ be any (not necessarily admissible) index set. Recall that the 'regularized MZVs' $Z^*_{\mathbf{k}}(0)$ and $Z^{ш}_{\mathbf{k}}(0)$ are, respectively, the constant terms of the polynomials $Z^*_{\mathbf{k}}(T)$ and $Z^{ш}_{\mathbf{k}}(T)$ defined in § 2.



https://doi.org/10.1112/S0010437X0500182X Published online by Cambridge University Press

For a fixed $n$, consider the two generating functions

$$F_n^*(x_1,\dots,x_n)=\sum_{\mathbf{k}} Z_{\mathbf{k}}^*(0)x_1^{k_1-1}\cdots x_n^{k_n-1}$$

and

$$F_n^{ш}(x_1,\dots,x_n)=\sum_{\mathbf{k}} Z_{\mathbf{k}}^{ш}(0)x_1^{k_1-1}\cdots x_n^{k_n-1}$$

in $\mathcal{Z}[x_1,\dots,x_n]$, where both the sums run over all index sets $\mathbf{k}=(k_1,\dots,k_n)$ (allowing $k_1=1$) of depth $n$. We regard the coefficients as elements in $\mathcal{Z}^{(n)}$ and consider their images in $\mathcal{M}$, and look at the images of $F_n^*$ and $F_n^{ш}$ in $\widetilde{\mathcal{M}}[x_1,\dots,x_n]=\widetilde{\mathcal{M}}\otimes_{\mathbb{Q}} V_n$, where $\widetilde{\mathcal{M}}=\mathbb{Q}\oplus\mathcal{M}$ as a $\mathbb{Q}$-vector space and is regarded as a $\mathbb{Q}$-algebra in a trivial way (only $\mathbb{Q}$-multiplication is non-zero).

Theorem 6. *The two polynomials $F_n^*$ and $F_n^{ш}$ agree as elements of $\widetilde{\mathcal{M}}\otimes V_n$ and belong to the subspace $\widetilde{\mathcal{M}}\otimes DSh_n$.*

Corollary 7. *For all $k>n>0$, we have the upper bound*

$$D_{k,n}\leqslant \dim_{\mathbb{Q}} DSh_n(k-n).$$

*Proof.* The statement of the theorem follows from the following three assertions:

(i) $(F_n^*|sh_l)(x_1,\dots,x_n)=0$ in $\widetilde{\mathcal{M}}\otimes V_n$ for $1\leqslant l\leqslant n-1$;

(ii) $((F_n^{ш})^\sharp|sh_l)(x_1,\dots,x_n)=0$ in $\widetilde{\mathcal{M}}\otimes V_n$ for $1\leqslant l\leqslant n-1$;

(iii) $F_n^*(x_1,\dots,x_n)=F_n^{ш}(x_1,\dots,x_n)$ in $\widetilde{\mathcal{M}}\otimes V_n$.

(i) With an obvious notational convention, we know from Proposition 1 that the values $Z_{\mathbf{k}}^*(0)$ satisfy $Z_{\mathbf{k}}^*(0)Z_{\mathbf{k}'}^*(0)=Z_{\mathbf{k}*\mathbf{k}'}^*(0)$ for any index sets $\mathbf{k}$ and $\mathbf{k}'$. From this we easily have $0=F_l^*(x_1,\dots,x_l)F_{n-l}^*(x_{l+1},\dots,x_n)=(F_n^*\big|sh_l)(x_1,\dots,x_n)$ in $\widetilde{\mathcal{M}}[x_1,\dots,x_n]$. This proves (i).

(ii) Here we use the iterated integral expression. For small $\varepsilon>0$, put

$$\zeta_\varepsilon(k_1,k_2,\dots,k_n)=\int\cdots\int_{1-\varepsilon>t_1>t_2>\cdots>t_k>0}\omega_1(t_1)\omega_2(t_2)\cdots\omega_k(t_k),\tag{8.4}$$

where the integrand is the same as in (1.2). First we have

$$\int\cdots\int_{a>x_1>\cdots>x_r>b}\frac{dx_1}{x_1}\cdots\frac{dx_r}{x_r}=\frac{1}{r!}\left(\log\frac{a}{b}\right)^r$$

for any $a>b>0$, as the integral would be the same if $x_1,\dots,x_r$ were ordered in any other way and the integral over unordered $r$-tuples $a>x_1,\dots,x_r>b$ is simply the $r$th power of $\int_b^a(dx/x)$. Applying this to (8.4) with $(a,b,r)=(1-\varepsilon,u_1,k_1-1)$, $(u_1,u_2,k_2-1)$,..., $(u_{n-1},u_n,k_n-1)$, where $u_1=t_{k_1},u_2=t_{k_1+k_2},\dots,u_n=t_{k_1+k_2+\cdots+k_n}=t_k$, we find

$$\begin{aligned}\zeta_\varepsilon(k_1,k_2,\dots,k_n)&=\frac{1}{\prod_{i=1}^n(k_i-1)!}\\&\quad\times\int\cdots\int_{1-\varepsilon>u_1>\cdots>u_n>0}\left(\log\frac{1-\varepsilon}{u_1}\right)^{k_1-1}\frac{du_1}{1-u_1}\left(\log\frac{u_1}{u_2}\right)^{k_2-1}\frac{du_2}{1-u_2}\cdots\\&\quad\cdots\left(\log\frac{u_{n-1}}{u_n}\right)^{k_n-1}\frac{du_n}{1-u_n}.\end{aligned}$$



https://doi.org/10.1112/S0010437X0500182X Published online by Cambridge University Press

From this, we obtain for the corresponding generating function $F^{ш}_{n,\varepsilon}$:

$$F^{ш}_{n,\varepsilon}(x_1,\dots,x_n) := \sum_{k_1,\dots,k_n\geqslant 1} \zeta_\varepsilon(k_1,\dots,k_n)x_1^{k_1-1}\cdots x_n^{k_n-1}$$
$$= (1-\varepsilon)^{x_1} \int\cdots\int_{1-\varepsilon>u_1>\cdots>u_k>0} \frac{u_1^{-x_1+x_2}}{1-u_1}\,du_1 \frac{u_2^{-x_2+x_3}}{1-u_2}\,du_2\cdots\frac{u_n^{-x_n}}{1-u_n}\,du_n$$

and, hence,

$$(F^{ш}_{n,\varepsilon})^\sharp(x_1,\dots,x_n) = F^{ш}_{n,\varepsilon}(x_1+x_2+\cdots+x_n, x_2+\cdots+x_n,\dots,x_n)$$
$$= (1-\varepsilon)^{x_1+\cdots+x_n} \int\cdots\int_{1-\varepsilon>u_1>\cdots>u_k>0} \frac{u_1^{-x_1}}{1-u_1}\,du_1 \frac{u_2^{-x_2}}{1-u_2}\,du_2\cdots\frac{u_n^{-x_n}}{1-u_n}\,du_n.$$

Now, it is clear that $(F^{ш}_{n,\varepsilon})^\sharp(x_1,\dots,x_n)$ satisfies the shuffle relation

$$(F^{ш}_{l,\varepsilon})^\sharp(x_1,\dots,x_l)(F^{ш}_{n-l,\varepsilon})^\sharp(x_{l+1},\dots,x_n) = \big((F^{ш}_{n,\varepsilon})^\sharp\big|sh_l\big)(x_1,\dots,x_n)$$

for any $\varepsilon$ and $1\leqslant l\leqslant n-1$. From this and the definition of the polynomial $Z^{ш}_{\mathbf{k}}(T)$, we obtain the desired assertion.

(iii) This follows from our fundamental relation (2.3) and the fact that the coefficient of $\rho(T^i)$ is contained in the ring generated by the Riemann zeta values (MZVs of depth 1). $\square$

To prove the corollary, let $\{f_i\}_{i=1,\dots,r}$ ($r=\dim DSh_n(d)$) be a basis of $DSh_n(d)$ over $\mathbb{Q}$ and write each $f_i(x_1,\dots,x_n)$ as $\sum_{\mathbf{k}} c_i(\mathbf{k})x_1^{k_1-1}\cdots x_n^{k_n-1}$, where $\mathbf{k}=(k_1,\dots,k_n)$ runs over the index sets of depth $n$ and weight $k$. Then the $r\times\binom{n+d-1}{d}$ matrix $\{c_i(\mathbf{k})\}_{i,\mathbf{k}}$ has rank $r$, so all of its columns can be expressed as $\mathbb{Q}$-linear combinations of $r$ of them, say, those labeled by the index sets $\mathbf{k}_1,\dots,\mathbf{k}_r$. This implies that the $\mathbf{k}$th coefficient of any polynomial $F\in A\otimes DSh_n(d)$, for any index set $\mathbf{k}$ of depth $n$ and weight $k$ and any $\mathbb{Q}$-algebra $A$, is a rational linear combination (with coefficients not depending on $F$) of its $\mathbf{k}_1$th$,\dots,\mathbf{k}_r$th coefficients. Applying this to the polynomial $F$ defined as the homogeneous component of degree $d=k-n$ of $F^*_n(x_1,\dots,x_n)=F^{ш}_n(x_1,\dots,x_n)$, which belongs to $\widetilde{\mathcal{M}}\otimes DSh_n(d)$ by the theorem, we see that the image in $\mathcal{M}_k^{(n)}$ of each MZV $\zeta(\mathbf{k})$ of weight $k$ and depth $n$ is a rational linear combination of the $r$ MZVs $\zeta(\mathbf{k}_1),\dots,\zeta(\mathbf{k}_r)$, so $\dim_{\mathbb{Q}}\mathcal{M}_k^{(n)}\leqslant r$. $\square$

In the remainder of this section we give some estimates of the spaces $DSh_n(d)$ and corollaries for MZVs.

Theorem 7. *If $d$ is odd, then $DSh_n(d)=\{0\}$ for every $n>0$.*

Corollary 8 (Parity result). *If $k\not\equiv n \bmod 2$, then $D_{k,n}=0$. In other words, any MZV of weight $k$ and depth $n$ with $k$ and $n$ of opposite parity is a linear combination of MZVs of smaller depth and products of MZVs of lower weight.*

This result, of which a different proof was given by Tsumura [Tsu04], generalizes classical results of Euler for $n=1$ (that $\zeta(2j)$ is expressible as a product of smaller zeta values for $j\geqslant 2$) and $n=2$ (that all double zeta values of odd weight can be expressed in terms of Riemann zeta values.)

To prove Theorem 7, we will construct a space $ShC_n(d)$ containing $DSh_n(d)$ and show that it is zero when $d$ is odd. This larger space will also be of interest for $d$ even, as it can be computed in terms of representation of finite groups and hence, in principle, gives a non-trivial upper bound for $\dim DSh_n(d)$ for all $n$ and $d$. This subject will be treated in a later paper.

To define the space $ShC_n(d)$, we introduce an action of $\mathfrak{S}_{n+1}$ on the space $\mathbb{Q}[x_1,\dots,x_n]$ by identifying this with the space $\mathbb{Q}[y_1,\dots,y_{n+1}]/(y_1+\cdots+y_{n+1})$ via the obvious isomorphism

$$\varphi:\mathbb{Q}[x_1,\dots,x_n]\simeq\mathbb{Q}[y_1,\dots,y_{n+1}]/(y_1+\cdots+y_{n+1})$$

given by

$$\varphi(x_i) = y_i \quad (1 \leqslant i \leqslant n).$$

The action of $\mathfrak{S}_{n+1}$ is then given by

$$(F|\sigma)(y_1, \ldots, y_{n+1}) := F(y_{\sigma^{-1}(1)}, \ldots, y_{\sigma^{-1}(n+1)}) \quad (\sigma \in \mathfrak{S}_{n+1}).$$

Identifying $\mathfrak{S}_n$ as the subgroup of $\mathfrak{S}_{n+1}$ consisting of elements that fix $n+1$, we can also regard the space $\mathbb{Q}[y_1, \ldots, y_{n+1}]/(y_1 + \cdots + y_{n+1})$ as an $\mathfrak{S}_n$-module; the map $\varphi$ is then $\mathfrak{S}_n$-equivariant. As before, the actions of $\mathfrak{S}_{n+1}$ extends naturally to an action of $\mathbb{Z}[\mathfrak{S}_{n+1}] \supset \mathbb{Z}[\mathfrak{S}_n]$.

Let

$$C_{n+1} = \begin{pmatrix} 1 & \cdots & n & n+1 \\ 2 & \cdots & n+1 & 1 \end{pmatrix} \in \mathfrak{S}_{n+1}$$

be the cyclic permutation and $\varepsilon$ the involution $(y_1, \ldots, y_{n+1}) \mapsto (-y_1, \ldots, -y_{n+1})$. Theorem 7 is a consequence of the following proposition.

Proposition 17. *For $n, d \geqslant 1$, define*

$$ShC_n(d) := \{f \in V_n(d) \mid f^\sharp | sh_l = 0 \ (1 \leqslant l \leqslant n-1), \varphi(f^\sharp)|C_{n+1} = \varepsilon\varphi(f^\sharp)\}.$$

*Then:*

(i) $DSh_n(d) \subset ShC_n(d)$;

(ii) $ShC_n(d) = \{0\}$ *if $d$ is odd.*

*Proof.* We first prove statement (ii). For this, we use the easily checked identity

$$1 + sh_1^{(n)} C_{n+1} = C_{n+1}(1 + sh_1^{(n)}\tau) \tag{8.5}$$

in $\mathbb{Z}[\mathfrak{S}_{n+1}]$, where $\tau$ is the transposition $1 \leftrightarrow n+1$ and where the shuffle element $sh_1^{(n)} \in \mathbb{Z}[\mathfrak{S}_n]$ is viewed as an element in $\mathbb{Z}[\mathfrak{S}_{n+1}]$ in the way described above.

For $f \in ShC_n(d)$, put $F = \varphi(f^\sharp)$. Applying (8.5) to $F$ and using the conditions $F|sh_1^{(n)} = 0$ and $F|C_{n+1} = \varepsilon F$, we obtain

$$F = F|C_{n+1}(1 + sh_1^{(n)}\tau) = \varepsilon F|(1 + sh_1^{(n)}\tau) = \varepsilon F.$$

If $d$ is odd, this gives $F = 0$ and hence $f = 0$.

To prove statement (i), we need a lemma.

Lemma 2. *For $0 \leqslant l \leqslant n$, let*

$$T_l = \begin{pmatrix} 1 & \cdots & l & l+1 & \cdots & n \\ 1 & \cdots & l & n & \cdots & l+1 \end{pmatrix} \in \mathfrak{S}_n.$$

*Then we have the relation*

$$\sum_{l=1}^{n-1} (-1)^{l-1} sh_l \, T_l = T_0 + (-1)^n,$$

*or, since $sh_0 = sh_n = \mathrm{id}$, simply $\sum_{l=0}^{n} (-1)^l sh_l T_l = 0$.*

*Proof.* For $1 \leqslant l \leqslant n-1$, we have

$$sh_l \, T_l = \sum_{\substack{\sigma \in \mathfrak{S}_n \\ \sigma(1) < \cdots < \sigma(l) \\ \sigma(l+1) > \cdots > \sigma(n)}} \sigma = R_l + R_{l+1},$$

where

$$R_i = \sum_{\substack{\sigma\in\mathfrak{S}_n \\ \sigma(1)<\cdots<\sigma(i)>\cdots>\sigma(n)}} \sigma \in \mathbb{Z}[\mathfrak{S}_n] \quad (i=1,\dots,n).$$

Hence,

$$\sum_{l=1}^{n-1}(-1)^{l-1} sh_l\, T_l = R_1 + (-1)^n R_n = T_0 + (-1)^n. \qquad \square$$

By Lemma 2, we have

$$f|T_0 = f|T_0^\sharp = (-1)^{n-1} f \quad \text{for} \quad f \in DSh_n(d). \tag{8.6}$$

Let

$$T' = \begin{pmatrix} 1 & \cdots & n-1 & n & n+1 \\ n-1 & \cdots & 1 & n+1 & n \end{pmatrix} \in \mathfrak{S}_{n+1}.$$

It is straightforward to check the relations

$$T_0T' = C_{n+1} \quad (\text{in } \mathfrak{S}_{n+1}), \quad \varphi(f|P)|T' = \varepsilon\varphi(f|T_0P). \tag{8.7}$$

Now suppose that $f \in DSh_n(d)$. We only need to check that the second condition in the definition of $ShC_n(d)$ is satisfied. With (8.6) and (8.7), we have

$$\begin{aligned}\varphi(f^\sharp)|C_{n+1} &= \varphi(f^\sharp)|T_0T' = \varphi(f|P)|T_0T' = \varphi(f|PT_0)|T' \\ &= \varphi(f|T_0^\sharp P)T' = \varepsilon\varphi(f|T_0^\sharp T_0 P) = \varepsilon\varphi(f^\sharp).\end{aligned}$$

This completes the proof of Proposition 17 and hence of Theorem 7. The corollary then follows immediately by virtue of the corollary to Theorem 6. $\square$

One can sometimes get upper and lower bounds for $DSh_n(d)$ using representation theory of finite groups. In particular, we can prove the following proposition[2] and also some results for general $n$.

PROPOSITION 18. *Assume $d$ is even. Then we have:*

(i) $\dim_{\mathbb{Q}} DSh_2(d) = \left[\dfrac{d}{6}\right]$;

(ii) $\left[\dfrac{d^2-1}{48}\right] \leqslant \dim_{\mathbb{Q}} DSh_3(d) \leqslant \left[\dfrac{(d+3)^2}{24}\right]$.

## Appendix. The conjectural value of $D_{k,n}$

In this appendix we describe the conjectural formula for $D_{k,n}$ due to Broadhurst and Kreimer [BK97].

Let $d_{k,n} = \dim(\mathcal{Z}_k^{(n)}/\mathcal{Z}_k^{(n-1)})$ be the 'number of $\mathbb{Q}$-linearly independent MZVs of weight $k$ and depth exactly $n$'. Denote by $\mathcal{D}(x,y) = \sum_{k,n} d_{k,n}\, x^k y^n$ the generating function of the numbers $d_{k,n}$ and by $\mathcal{D}^0(x,y)$ the corresponding generating function for the 'primitive part' $\mathcal{Z}^0$, where $\mathcal{Z} = \mathcal{Z}^0 \otimes_{\mathbb{Q}} \mathbb{Q}[\pi^2]$. Then

$$\mathcal{D}(x,y) = \left(1 + \frac{x^2}{1-x^2}y\right)\mathcal{D}^0(x,y)$$

[2] The referee pointed out that the dimensions of spaces essentially equivalent to our $DSh_2(d)$ and $DSh_3(d)$ have been calculated (in a quite different way) in [Gon01a, pp. 474–475]. The space $\mathcal{D}_{d+n,n}$ there is isomorphic to our $DSh_n(d)$. According to [Gon01a], $\dim_{\mathbb{Q}} DSh_3(d) = [(d^2-1)/48]$. We thank the referee for this indication.

and the numbers $D_{k,n}$ are given in terms of $\mathcal{D}^0(x,y)$ by the product expansion formula

$$\prod_{k,\,n}\left(\frac{1}{1-x^k y^n}\right)^{D_{k,n}} = \frac{1}{1-x^2y}\mathcal{D}^0(x,y)$$

or, more explicitly, by the formulas

$$D_{k,n} = \sum_{d|(k,n)} \frac{\mu(d)}{d}\cdot \text{coefficient of } x^{k/d}y^{n/d} \text{ in } \log\mathcal{D}^0(x,y),$$

where $\mu(d)$ denotes the Möbius function. Knowing the numbers $D_{k,n}$ is therefore equivalent to knowing the function $\mathcal{D}(x,y)$ or $\mathcal{D}^0(x,y)$. The Broadhurst–Kreimer conjecture states that the power series $\mathcal{D}^0(x,y)$ is given by

$$\mathcal{D}^0(x,y) = \frac{1}{1-\mathbf{O}y \;+\; \mathbf{S}y^2 \;-\; \mathbf{S}y^4},$$

where

$$\mathbf{O} = \sum_{k>1,\ k \text{ odd}} x^k = \frac{x^3}{1-x^2}$$

corresponds to the odd zeta values $\zeta(3)$, $\zeta(5), \ldots$ and

$$\mathbf{S} = \sum_{k>0} \dim S_k(SL_2(\mathbb{Z}))x^k = \frac{x^{12}}{(1-x^4)(1-x^6)}$$

is the generating series of the dimension of the graded vector space of cusp forms on the full modular group. The conjectural value for the full power series $\mathcal{D}(x,y)$ is then given by the rational function

$$\mathcal{D}(x,y) = \frac{1+\mathbf{E}y}{1-\mathbf{O}y+\mathbf{S}y^2-\mathbf{S}y^4}$$

in $x$ and $y$, where

$$\mathbf{E} = \sum_{k>0,\ k \text{ even}} x^k = \frac{x^2}{1-x^2}$$

now counts the *even* zeta values $\zeta(2)$, $\zeta(4), \ldots$ . In particular, for the full number $D_k = \sum_n D_{k,n}$ of weight $k$ generators of the MZV algebra, this would imply

$$\prod_{k\geqslant 2}\left(\frac{1}{1-x^k}\right)^{D_k} = \mathcal{D}(x,1) = \frac{1}{1-x^2-x^3},$$

in accordance with third-named author's conjecture on the value of $\dim \mathcal{Z}_k$. This conjecture is still open, but it has been shown by Terasoma [Ter02] and independently by Goncharov [Gon01c] that the coefficient of $x^k$ in $(1-x^2-x^3)^{-1}$ is an upper bound for $\dim \mathcal{Z}_k$.

The Broadhurst–Kreimer conjecture implies that $D_{k,n} = 0$ unless $k \geqslant 3n$ and $k$ is congruent to $n$ modulo 2, and gives Table 3 for $n \leqslant 6$ and $k-3n \leqslant 14$.

It also implies the formulas

$$D_{k,1} = 1, \quad D_{k,2} = \left[\frac{d}{6}\right], \quad D_{k,3} = \left[\frac{d^2-1}{48}\right]$$

for $1 \leqslant n \leqslant 3$ and $k-n = d \equiv 0 \pmod 2$. Each of these three formulas is known to give an upper bound for the true dimension (cf. [Zag93] for $n=2$ and Goncharov [Gon01b] for $n=3$; see also the footnote before Proposition 18 of § 8).

Table 3. Values of $D_{k,n}$.

| $n$ \ $k-3n$ | 0 | 2 | 4 | 6 | 8 | 10 | 12 | 14 |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 0 | 1 | 1 | 1 | 2 | 2 | 2 | 3 |
| 3 | 0 | 1 | 2 | 2 | 4 | 5 | 6 | 8 |
| 4 | 1 | 1 | 3 | 5 | 7 | 11 | 16 | 20 |
| 5 | 1 | 2 | 5 | 9 | 15 | 23 | 36 | 50 |
| 6 | 1 | 3 | 7 | 14 | 27 | 45 | 73 | 113 |

## References


Bou00 L. Boutet de Monvel, *Remarques sur les séries logarithmiques divergentes*, in *Polylogarithmes et conjecture de Deligne–Ihara*, Luminy, April 2000 (CIRM, Marseille, 2000).

BK97 D. J. Broadhurst and D. Kreimer, *Association of multiple zeta values with positive knots via Feynman diagrams up to* 9 *loops*, Phys. Lett. B **393** (1997), 403–412.

CM04a A. Connes and H. Moscovici, *Modular Hecke algebras and their Hopf symmetry*, Mosc. Math. J. **4** (2004), 67–109.

CM04b A. Connes and H. Moscovici, *Rankin–Cohen brackets and the Hopf algebra of transverse geometry*, Mosc. Math. J. **4** (2004), 111–130.

Dri90 V. G. Drinfel'd, *On quasitriangular quasi-Hopf algebras and on a group that is closely connected with* $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, Algebra i Analiz **2** (1990), 149–181 (in Russian). (English translation in Leningrad Math. J. **2**(4) (1991), 829–860).

Gon01a 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** (2001), 397–487.

Gon01b A. B. Goncharov, *Multiple* $\zeta$*-values, Galois groups and geometry of modular varieties*, Progr. Math. **201** (2001), 361–392.

Gon01c A. B. Goncharov, *Multiple polylogarithms and mixed Tate motives*, Preprint (2001), math.AG/0103059.

Gra97 A. Granville, *A decomposition of Riemann's zeta-function*, in *Analytic number theory*, London Mathematical Society Lecture Note Series, vol. 247 (Cambridge University Press, Cambridge 1997), 95–101.

Hof92 M. Hoffman, *Multiple harmonic series*, Pacific J. Math. **152** (1992), 275–290.

Hof97 M. Hoffman, *The algebra of multiple harmonic series*, J. Algebra **194** (1997), 477–495.

HO03 M. Hoffman and Y. Ohno, *Relations of multiple zeta values and their algebraic expression*, J. Algebra **262** (2003), 332–347.

MPH00 H. N. Minh, M. Petitot and J. V. D. Hoeven, *Shuffle algebra and polylogarithms*, Discrete Math. **225** (2000), 217–230.

Ohn99 Y. Ohno, *A generalization of the duality and sum formulas on the multiple zeta values*, J. Number Theory **74** (1999), 39–43.

Rac02 G. Racinet, *Doubles mélanges des polylogarithmes multiples aux racines de l'unité*, Publ. Math. Inst. Hautes Études Sci. **95** (2002), 185–231.

Reu93 C. Reutenauer, *Free Lie algebras* (Oxford Science Publications, Oxford, 1993).

Ter02 T. Terasoma, *Mixed Tate motives and multiple zeta values*, Invent. Math. **149** (2002), 339–369.

Tsu04 H. Tsumura, *Combinatorial relations for Euler–Zagier sums*, Acta Arith. **111** (2004), 27–42.

Zag93 D. Zagier, *Periods of modular forms, trace of Hecke operators, and multiple zeta values*, RIMS Kokyuroku **843** (1993), 162–170.

Zag94 D. Zagier, *Values of zeta functions and their applications*, in *First European Congress of Mathematics, Vol. II*, Progr. Math. **120** (1994), 497–512.

Kentaro Ihara k-ihara@math.nagoya-u.ac.jp
Graduate School of Mathematics, Nagoya University, Chikusa-ku, Nagoya 464-8602, Japan

Masanobu Kaneko mkaneko@math.kyushu-u.ac.jp
Graduate School of Mathematics, Kyushu University 33, Fukuoka 812-8581, Japan

Don Zagier zagier@mpim-bonn.mpg.de
Max-Planck-Institut für Mathematik, Vivatsgasse 7, D-53111 Bonn, Germany
and
Collège de France, 3 rue d'Ulm, F-75005 Paris, France