数理解析研究所講究録
第 1549 巻 2007 年 47-63



**Derivation and double shuffle relations for multiple zeta values**
— joint work with M. Kaneko, D.Zagier.

九州大・数理学府　井原　健太郎 (Kentaro Ihara)

# 1 Introduction

The *multiple zeta value* (MZV for short) is a real number defined by

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

$$=\underbrace{\int_0^1\frac{dt}{t}\int_0^t\frac{dt}{t}\cdots\int_0^t\frac{dt}{t}}_{k_1-1}\int_0^t\frac{dt}{1-t}\cdots\underbrace{\int_0^t\frac{dt}{t}\cdots\int_0^t\frac{dt}{t}}_{k_n-1}\int_0^t\frac{dt}{1-t}. \tag{2}$$

where $\mathbf{k}=(k_1,k_2,\ldots,k_n)$ is an index set of positive integers with $k_1>1$. The condition $k_1>1$ ensures the convergence of the series and the integral. For the value $\zeta(k_1,k_2,\ldots,k_n)$, (strictly, for the index set $(k_1,k_2,\ldots,k_n)$) we call the number $n$ *depth* and $k=k_1+\cdots+k_n$ *weight.*

There are many linear and algebraic relations over $\mathbb{Q}$ among MZV's of the same weight, the simplest of which is $\zeta(3)=\zeta(2,1)$ found by Euler. To give a complete description of them is one of the main goal of the study of MZV's. From each representation (1) and (2), we can show that the product of two MZV's is written as a linear combination of MZV's with rational coefficients. Hence the $\mathbb{Q}$-vector space generated by MZV's is equipped with a $\mathbb{Q}$-algebra structure. In this report we investigate the structure of this $\mathbb{Q}$-algebra and give supplementaly explanations of the results in [1] and [2].

# 2 Double shuffle relations

To describe the multiplication rules of MZV's, we use an algebraic setup given by Hoffman in [5]. Let $\mathfrak{H}=\mathbb{Q}\langle x,y\rangle$ be the non-commutative polynomial algebra over $\mathbb{Q}$ 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 which sends the word $x^{k_1-1}yx^{k_2-1}y\cdots x^{k_n-1}y$ to the value $\zeta(k_1,k_2,\ldots,k_n)$ ("evaluation map"). The weight of $\zeta(k_1,k_2,\ldots,k_n)$ corresponds to the total degree of the word $x^{k_1-1}y\cdots x^{k_n-1}y$, and the depth $n$ the partial degree in $y$.

Put $z_k:=x^{k-1}y$, which corresponds to the Riemann zeta value $\zeta(k)$. Then the non-commutative algebra $\mathfrak{H}^1$ is freely generated by the set $\{z_k\mid k=1,2,3,\ldots\}$. Note that all $z_k$ are in $\mathfrak{H}^0$ except for $z_1=y$. We define the *harmonic product* $*$ on $\mathfrak{H}^1$ inductively by $1*w=w*1=w$ and

$$z_kw_1*z_lw_2=z_k(w_1*z_lw_2)+z_l(z_kw_1*w_2)+z_{k+l}(w_1*w_2), \tag{3}$$

where $k,l\geq 1$ and $w,w_1,w_2$ are any word in $\mathfrak{H}^1$, and extending by $\mathbb{Q}$-bilinearity. In [5], Hoffman showed that $\mathfrak{H}^1$ becomes an associative commutative algebra under the multiplication $*$ and $\mathfrak{H}^0$ a subalgebra. We will denote these algebras by $\mathfrak{H}^1_*$ and $\mathfrak{H}^0_*$ respectively.

Then the first multiplication law of MZV's can be stated that the map $Z$ is an algebra homomorphism with respect to the harmonic product $*$. For instance, the product $z_k * z_l = z_k z_l + z_l z_k + z_{k+l}$ corresponds to the identity $\zeta(k)\zeta(l) = \zeta(k,l) + \zeta(l,k) + \zeta(k+l)$.

The other commutative product $ш$, called *shuffle product* corresponding to the product of two integrals, is defined on all of $\mathfrak{H}$ inductively by $1 ш w = w ш 1 = w$ and

$$uw_1 ш vw_2 = u(w_1 ш vw_2) + v(uw_1 ш w_2), \tag{4}$$

where $w, w_1, w_2$ are any word in $\mathfrak{H}$ and $u, v \in \{x, y\}$, and again extending by $\mathbb{Q}$-bilinearity. Then the space $\mathfrak{H}$ make an associative commutative $\mathbb{Q}$-algebra ([11]) which we denote by $\mathfrak{H}_ш$. Obviously the subspaces $\mathfrak{H}^1$ and $\mathfrak{H}^0$ become subalgebras of $\mathfrak{H}_ш$, denoted by $\mathfrak{H}^1_ш$ and $\mathfrak{H}^0_ш$ respectively. By the standard shuffle product identity of iterated integrals, the evaluation map $Z$ is again an algebra homomorphism with respect to the multiplication $ш$.

Compareing the two products, we obtain the *double shuffle relations* (DSR for short) of MZVs:

$$Z(w_1 ш w_2) = Z(w_1 * w_2) \qquad (w_1, w_2 \in \mathfrak{H}^0). \tag{5}$$

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 get $4\zeta(3,1) = \zeta(4)$. However these double shuffle relations do not give the "all" relations. For instance, the relation $\zeta(3) = \zeta(2,1)$ can not be obtained from the double shuffle relations. Let $\mathcal{Z}_k$ be the $\mathbb{Q}$-vector space generated by all MZV's of weight $k$. Below is the table of the conjectural dimension $d_k$ of $\mathcal{Z}_k$ and the upper bounds of the $\dim\mathcal{Z}_k$ which are obtained by double shuffle relations. Therefore we need more larger class of relations to supply sufficiently many relations. In Section 4 we will show its extended version stated in [1].

| $k$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | $\cdots$ |
|---|---|---|---|---|---|---|---|---|---|
| $d_k$ | 1 | 1 | 1 | 2 | 2 | 3 | 4 | 5 | $\cdots$ |
| DSR | 1 | 2 | 3 | 6 | 9 | 16 | 24 | 36 | $\cdots$ |

# 3 Regularization

**Proposition 1** ([5],[11]) *For each product $\bullet = *$ or $ш$, we can regard $\mathfrak{H}^1_\bullet$ as a $\mathfrak{H}^0_\bullet$-algebra via the inclusion map $\mathfrak{H}^0_\bullet \to \mathfrak{H}^1_\bullet$. Then $\mathfrak{H}^1_\bullet$ is freely generated by the element $y$ over $\mathfrak{H}^0_\bullet$. In other words, for any $f \in \mathfrak{H}^1_\bullet$ there uniquely exist elements $f_0, \ldots, f_r \in \mathfrak{H}^0_\bullet$, ($f_r \neq 0$) such that*

$$f = f_0 + f_1 \bullet y + f_2 \bullet y^{\bullet 2} + \cdots + f_r \bullet y^{\bullet r}.$$

*Proof.* See [5] for the case $*$ and [11] for $ш$. ∎

**Definition 1** *For each product $\bullet = *$ or $ш$, we define two maps $Z^\bullet : \mathfrak{H}^1_\bullet \longrightarrow \mathbb{R}[T]$ which are uniquely characterized by the properties that they are algebra homomorphisms for $\bullet$ and both extend the evaluation map $Z : \mathfrak{H}^0 \to \mathbb{R}$ and send $y$ to $T$. In other words under the notaion in Proposition 1, we have*

$$Z^\bullet(f) = Z(f_0) + Z(f_1)T + Z(f_2)T^2 + \cdots + Z(f_r)T^r.$$

For example,

$$Z^*(yxy) = \zeta(2)T - \zeta(2,1) - \zeta(3), \quad Z^{ш}(yxy) = \zeta(2)T - 2\zeta(2,1).$$
$$Z^*(y^2xy) = \frac{1}{2}\zeta(2)T^2 - (\zeta(3)+\zeta(2,1))T + \frac{1}{2}\zeta(4) + \zeta(3,1) + \zeta(2,1,1),$$
$$Z^{ш}(y^2xy) = \frac{1}{2}\zeta(2)T^2 - 2\zeta(2,1)T + 3\zeta(2,1,1).$$

We introduce the following power series $A(u)$:

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

Note that the coefficient for $u^k$ of $A(u)$ is an element of weight $k$ in the $\mathbb{Q}$-algebra generated by Riemann zeta values. Define an $\mathbb{R}$-linear automorphism $\rho : \mathbb{R}[T] \to \mathbb{R}[T]$ by

$$\rho(e^{Tu}) = A(u)e^{Tu}. \tag{6}$$

For example, $\rho(T) = T$, $\rho(T^2) = T^2 + \zeta(2)$, and $\rho(T^3) = T^3 + 3\zeta(2)T - 2\zeta(3)$.

The following theorem does originally to Zagier, and much work has been done by other writers Racinet, Goncharov, Minh, Petitot, Boutet de Monvel, Écalle,...

**Theorem 1** *We have*

$$Z^{ш} \equiv \rho \circ Z^* \quad on\ \mathfrak{H}^1.$$

*Proof.* (Sketch) For more detail see [1]. For each multiplication rule, we define two kinds of truncation of multiple zeta values: For $M > 0$ and index set $\mathbf{k} = (k_1, k_2, \ldots, k_n)$ (not necessarily $k_1 > 1$), set

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

If $k_1 > 1$ then $\zeta_M(\mathbf{k})$ converges to $\zeta(\mathbf{k})$ as $M \to \infty$. 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. With this fact and the classical formula $\zeta_M(1) = \sum_{M>m>0} 1/m = \log M + \gamma + O(M^{-1})$, we can show by induction that

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

where $Z^*_{\mathbf{k}}(T) := Z^*(z_{k_1} \cdots z_{k_n})$ is the associated polynomial defined in Definition 1.

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

$$\begin{aligned} Li_{\mathbf{k}}(t) &= \sum_{m_1>m_2>\cdots>m_n>0} \frac{t^{m_1}}{m_1^{k_1} m_2^{k_2} \cdots m_n^{k_n}} \\ &= \underbrace{\int_0^t \frac{dt}{t}\int_0^t \frac{dt}{t}\cdots\int_0^t \frac{dt}{t}}_{k_1-1}\int_0^t \frac{dt}{1-t}\cdots\underbrace{\int_0^t \frac{dt}{t}\cdots\int_0^t \frac{dt}{t}}_{k_n-1}\int_0^t \frac{dt}{1-t}. \end{aligned}$$

If $k_1 > 1$ then $Li_{\mathbf{k}}(1) = \zeta(\mathbf{k})$. We can write the product $Li_{\mathbf{k}}(t)Li_{\mathbf{k}'}(t)$ as a linear combination of $Li_{\mathbf{k}''}(t)$'s via the shuffle product identity of iterated integrals. When $k_1, k_1' > 1$, the formula specializes at $t = 1$ to that of the shuffle product of $\zeta(\mathbf{k})\zeta(\mathbf{k}')$. Together with $Li_1(t) = \log \frac{1}{1-t}$, we conclude by induction that

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

where $Z_{\mathbf{k}}{}^{ш}(T) := Z^{ш}(z_{k_1} \cdots z_{k_n})$ is the associated polynomial in Definition 1.

For any index set $\mathbf{k}$, we have

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

For any $P(T) \in \mathbb{R}[T]$ and $Q(T) := \rho(P(T))$, we can show the following behavior as $t \nearrow 1$:

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

for some $J > 0$. We omit the proof of this equation, (see [1]). This fact establishes the theorem. ∎

## 4 Extended double shuffle relations

In this section, we explain the meaning of Theorem 1 from the viewpoint of the algebra structure on $\mathfrak{H}$.

Let $\widehat{\mathfrak{H}} = \mathbb{Q}\langle\langle x, y \rangle\rangle$ be the algebra of non-commutative formal power series with $\mathbb{Q}$-coefficients. The algebra $\widehat{\mathfrak{H}}$ is complete with respect to the grading defined by $\deg x = \deg y = 1$ and then $\mathfrak{H}$ is a dense subalgebra of $\widehat{\mathfrak{H}}$. A *derivation* $d$ on $\mathfrak{H}$ (resp. $\widehat{\mathfrak{H}}$) is a $\mathbb{Q}$-linear (resp.+continious) map satisfing the derivation property for concatenation product: $d(uv) = d(u)v + ud(v)$ for any $u, v \in \mathfrak{H}$ (resp. $\in \widehat{\mathfrak{H}}$). The space of all derivations of $\widehat{\mathfrak{H}}$ form a Lie algebra, denoted by $\mathrm{Der}(\widehat{\mathfrak{H}})$, with usual commutater bracket: $[d, d'] := d \circ d' - d' \circ d$. On the other hand, the set of all algebra automorphisms of $\widehat{\mathfrak{H}}$ (with respect to the concatenation product) form a group, denoted by $\mathrm{Aut}(\widehat{\mathfrak{H}})$. Note that both derivations and autmorphisms on $\mathfrak{H}$ or $\widehat{\mathfrak{H}}$ are determined by the values on generators $x, y$. Let $\mathrm{Der}^+(\widehat{\mathfrak{H}})$ be the Lie subalgebra consisting of derivations which increase the degree, or equivalently which induce the zero derivation on the associated graded algebra $\mathrm{gr}(\widehat{\mathfrak{H}}) = \bigoplus \widehat{\mathfrak{H}}_k / \widehat{\mathfrak{H}}_{k+1}$, where $\widehat{\mathfrak{H}}_k$ is the subspace of $\widehat{\mathfrak{H}}$ generated by the words of degree $\geqslant k$. Let $\mathrm{Aut}^1(\widehat{\mathfrak{H}})$ be the subgroup of $\mathrm{Aut}(\widehat{\mathfrak{H}})$ consisting of automorphisms $\phi$ such that $\phi(x) - x$ and $\phi(y) - y$ belong to $\widehat{\mathfrak{H}}_2$, or equivalently which induce the identity automorphism on $\mathrm{gr}(\widehat{\mathfrak{H}})$.

In the discussion below it is usefull to keep in mind the following facts. There is a one to one correspondence between the Lie subalgebra $\mathrm{Der}^+(\widehat{\mathfrak{H}})$ and the subgroup $\mathrm{Aut}^1(\widehat{\mathfrak{H}})$

via the exponential and the logarithm maps; $\exp(d) = e^d = \sum_{m\geqslant 0} \frac{d^m}{m!}$, for $d \in \mathrm{Der}^+(\widehat{\mathfrak{H}})$, $\log(\phi) = -\sum_{m\geqslant 1} \frac{(1-\phi)^m}{m}$, for $\phi \in \mathrm{Aut}^1(\widehat{\mathfrak{H}})$.

**Proposition 2** ([1]) *Define the map $d : \mathfrak{H} \to \mathfrak{H}$ by $d(w) = y ш w - yw$. Then $d$ is a derivation and we have*

$$\exp(du)(w) = (1-yu)\left(\frac{1}{1-yu} ш w\right), \tag{7}$$

*where $u$ is a formal parameter.*

*Proof.* Using (4), we can show the derivation property of $d$ and $\frac{1}{m!}d^m(w) = y^m ш w - y(y^{m-1} ш w)$ by induction. Multiplying this by $u^m$ and summing over $m$ gives (7). ∎

The analogous result for $*$ product is as follows. See [1] for the proof. Recall $z_n = x^{n-1}y$.

**Proposition 3** ([1]) *For $n \geqslant 1$ the map $\delta_n : \mathfrak{H}^1 \to \mathfrak{H}^1$ defined by $\delta_n(w) := z_n * w - z_n w$ is a derivation and we have*

$$\exp\left(\sum_{n\geqslant 1} \frac{(-1)^{n-1}}{n} \delta_n u^n\right)(w) = (1-yu)\left(\frac{1}{1-yu} * w\right).$$

*These derivation $\delta_n$ extends to a derivation on all of $\mathfrak{H}$, with values on the generators given by $\delta_n(x) = 0$, $\delta_n(y) = (x+y)z_n$.*

**Proposition 4** ([1]) *We define two automorphisms by*

$$\Psi_u := \exp(du), \qquad \Phi_u := \exp\left(\sum_{n\geqslant 1} \frac{\delta_n}{n} u^n\right).$$

*Then the action on the generators is given by*

$$\Psi_u(x) = x(1-yu)^{-1}, \quad \Psi_u(y) = y(1-yu)^{-1}, \quad \Psi_u(z) = z(1-yu)^{-1},$$
$$\Phi_u(x) = x, \quad \Phi_u(y) = (1-zu)^{-1}y, \quad \Phi_u(z) = (1-zu)^{-1}z(1-xu),$$

*where we put $z = x+y$. In particular, both automorphisms $\Psi_u$ and $\Phi_u$ preserve $\mathfrak{H}^0$.*

*Proof.* By induction, we can check $\frac{1}{m!}d^m(x) = xy^m$, and $\frac{1}{m!}d^m(y) = y^{m+1}$, which gives the result for $\Psi_u$. For the $\Phi_u$, see [1]. ∎

**Definition 2** *Let $\Delta_u$ be the automorphism of $\mathfrak{H}$ defined by $\Delta_u = \Psi_u \circ \Phi_u^{-1}$. The images of the generators $x$ and $y$ of $\Delta_u$ are given by*

$$\Delta_u(x) = x(1-yu)^{-1}, \quad \Delta_u(y) = (1-zu)(1-yu)^{-1}y, \quad \Delta_u(z) = z.$$

*where $z = x+y$.*

**Definition 3** *For each product* $\bullet = *$ *or* $ш$, *we define algebra homomorphisms* $\mathrm{reg}_\bullet : \mathfrak{H}^1_\bullet \longrightarrow \mathfrak{H}^0_\bullet$ *which is uniquely characterized by the properties that it is identity on* $\mathfrak{H}^0$ *and sends* $y$ *to* $0$. *Specifically* $\mathrm{reg}_\bullet(f) := f_0$ *for* $f \in \mathfrak{H}^1$, *where* $f_0$ *is the element given in Proposition 1.*

By Definition 1 and Definition 3, for each $\bullet = *$ or $ш$ it clearly holds that

$$Z \circ \mathrm{reg}_\bullet(f) = Z^\bullet(f)\ \big|_{T=0} \tag{8}$$

for all $f \in \mathfrak{H}^1$.

**Theorem 2 (Extended double shuffle relations)** ([1]) *The following statements are true and equivalent:*

**(i)** $Z^{ш} - \rho \circ Z^* \equiv 0$ *on* $\mathfrak{H}^1$,

**(ii)** $Z \circ (\Delta_u - 1) \equiv 0$ *on* $\mathfrak{H}^0$,

**(iii)** $Z[\ \mathrm{reg}_{ш}(w_1 ш\, w_0 - w_1 * w_0)\ ] = 0$ *for* $w_1 \in \mathfrak{H}^1$, $w_0 \in \mathfrak{H}^0$,

**(iv)** $Z[\ \mathrm{reg}_*(w_1 ш\, w_0 - w_1 * w_0)\ ] = 0$ *for* $w_1 \in \mathfrak{H}^1$, $w_0 \in \mathfrak{H}^0$.

*We call this equivalent class of relations of MZV's "extended double shuffle relations".*

**Conjecture 1** ([1]) *The extended double shuffle relations give the all relations among MZV's.*

**Lemma 1** *We have*

$$\exp_{ш}(yu) = \frac{1}{1-yu} = \exp_*\Big(\sum_{n\geqslant 1}\frac{(-1)^{n-1}}{n} z_n u^n\Big).$$

*where* $\exp_\bullet(f) = \sum_{n\geqslant 0}\frac{1}{n!}f^{\bullet n}$ *for* $f \in \mathfrak{H}^1$.

*Proof.* The first equation is direct from $y^{ш n} = n! y^n$. For second equation, see [1]. ∎

*Proof of Theorem 2.* (Sketch) In Proposition 2, replace $w$ by $\Delta_{-u}(w_0)$ and divide both sides by $1 - yu$, and use the lemma,

$$\frac{1}{1-yu}\Phi^{-1}_{-u}(w_0) = \frac{1}{1-yu} ш\, \Delta_{-u}(w_0) = \exp_{ш}(yu) ш\, \Delta_{-u}(w_0), \tag{9}$$

for $w_0 \in \mathfrak{H}^0$. On the other hand, use Proposition 3 and the lemma in the same way, we have

$$\frac{1}{1-yu}\Phi^{-1}_{-u}(w_0) = \frac{1}{1-yu} * w_0 = \exp_*\Big(\sum_{n\geqslant 1}\frac{(-1)^{n-1}}{n} z_n u^n\Big) * w_0. \tag{10}$$

Apply $Z^{ш}$ and $\rho \circ Z^*$ to (9) and (10) respectively, we have

$$Z^{ш}\Big(\frac{1}{1-yu}\Phi^{-1}_{-u}(w_0)\Big) = Z(\Delta_{-u}(w_0))e^{Tu}, \tag{11}$$

$$\rho \circ Z^*\Big(\frac{1}{1-yu}\Phi^{-1}_{-u}(w_0)\Big) = \rho\big(Z(w_0)e^{Tu}A(u)^{-1}\big) = Z(w_0)e^{Tu}. \tag{12}$$

Since $\Phi_{-u}^{-1}$ acts as an automorphism of $\mathfrak{H}^0$ and since the elements $\frac{1}{1-yu}\mathfrak{H}^0$ span $\mathfrak{H}^1$, the equations (11) and (12) ensures the equivalence between (i) and (ii). Next we show that (iii)$\Rightarrow$(ii).

$$\begin{aligned}
&\mathrm{reg}_{Ш}\Big(\frac{1}{1-yu}ш\, w_0 - \frac{1}{1-yu} * w_0\Big)\\
&= \mathrm{reg}_{Ш}\Big(\frac{1}{1-yu}ш\, w_0 - \frac{1}{1-yu}ш\, \Delta_{-u}(w_0)\Big)\\
&= \mathrm{reg}_{Ш}\Big(\frac{1}{1-yu}\Big)(1-\Delta_{-u})(w_0) = (1-\Delta_{-u})(w_0),
\end{aligned}$$

where we used (9), (10) for the first equation and used the fact $\mathrm{reg}_{Ш}(1-yu)^{-1} = 1$ for the last equation, which follows from $\mathrm{reg}_{Ш}(y^m) = 0$, $(m \geqslant 1)$. Taking $\frac{1}{1-yu}$ for $w_1$ in (iii), the above equation shows that (iii)$\Rightarrow$ (ii). By the same arguments we can show (iv)$\Rightarrow$ (ii). For (i) $\Rightarrow$ (iii), multiply $Z(w_0) \in \mathbb{R}$ on both sides of $Z^{Ш}(w_1) = \rho(Z^*(w_1))$ and use the $\mathbb{R}$-linearity of $\rho$ to get $Z^{Ш}(w_1 ш\, w_0) = \rho(Z^*(w_1 * w_0))$. Using (i) on the right, we obtain $Z^{Ш}(w_1 ш\, w_0 - w_1 * w_0) = 0$. From (8), comparing the constant term of this equation, which shows (iii). The implication (i)$\Rightarrow$(iv) is proved similarly. ∎

## 5 Derivation and Ohno's relations

**Theorem 3 (Derivation relations, [1])** *For $n \geqslant 1$, let $\partial_n$ be the derivation on $\mathfrak{H}$ defined by the following action on generators:*

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

*Then $\partial_n$ can be ristricted to a derivation on $\mathfrak{H}^0$ and we have $Z[\ \partial_n(\mathfrak{H}^0)\ ] = 0$.*

Define a space of linear endomorphisms on $\mathfrak{H}^0$ by

$$\mathcal{N} = \{\ \varphi \in \mathrm{End}_{\mathbb{Q}}(\mathfrak{H}^0)\ |\ Z[\ \varphi(\mathfrak{H}^0)\ ] = 0\ \}$$

Note that the space $\mathcal{N}$ is a right ideal of $\mathrm{End}_{\mathbb{Q}}(\mathfrak{H}^0)$ under the composition of endomorphisms. Then the derivation relations can be restated as $\partial_n \in \mathcal{N}$. For later use, we review several relations of MZV's.

**Proposition 5 (Duality)** *Let $\tau : \mathfrak{H} \to \mathfrak{H}$ be the involutive anti-automorphism which interchanges $x$ and $y$: $\tau(x) = y$, $\tau(y) = x$; $\tau(uv) = \tau(v)\tau(u)$ for $u, v \in \mathfrak{H}$. Then $1 - \tau \in \mathcal{N}$, where 1 denotes the identity map on $\mathfrak{H}^0$.*

**Proposition 6 (Ohno's relations, [9])** *For $l \geqslant 0$, let $\sigma_l : \mathfrak{H}^0 \to \mathfrak{H}^0$ be the $\mathbb{Q}$-linear map defined by*

$$\sigma_l(z_{k_1} z_{k_2} \cdots z_{k_n}) = \sum_{\substack{e_1+e_2+\cdots+e_n = l\\ e_i \geq 0}} z_{k_1+e_1} z_{k_2+e_2} \cdots z_{k_n+e_n}.$$

*Then $\sigma_l - \sigma_l \tau \in \mathcal{N}$.*

For any endomorphism $\varphi \in \mathrm{End}_{\mathbb{Q}}(\mathfrak{H}^0)$, put $\overline{\varphi} := \tau\varphi\tau$. If $\varphi$ is a derivation or an automorphism, then so is $\overline{\varphi}$. Since $\tau^2 = 1$, it holds

$$\sigma_l - \overline{\sigma_l} = (\sigma_l - \sigma_l\tau) - (1-\tau)\overline{\sigma_l}. \tag{13}$$

Since $\mathcal{N}$ is a right ideal, Proposition 5, 6 imply $\sigma_l - \overline{\sigma_l} \in \mathcal{N}$. We call these relations *weak Ohno's relations.* Indeed, from (13) the Ohno's relations is deduced from its weak version and duality.

We give a table of all derivations which have been defined above. Here $z$ denotes $x+y$.

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

Define the derivations on $\widehat{\mathfrak{H}}$ as follows.

$$\delta_u = \sum_{n\geqslant 1} \frac{\delta_n}{n} u^n, \quad \partial_u = \sum_{n\geqslant 1} \frac{\partial_n}{n} u^n, \quad D_u = \sum_{n\geqslant 1} \frac{D_n}{n} u^n, \quad \overline{D}_u = \sum_{n\geqslant 1} \frac{\overline{D}_n}{n} u^n.$$

**Theorem 4** ([1]) *We have following equations among the corresponding automorphisms:*

$$\Delta_u := \exp(du)\exp(-\delta_u) = \exp(\partial_u) = \exp(\overline{D_u})\exp(-D_u).$$

*Proof.* It is enough to show that the images of generators for each automorphism coincides with each other. From the definition of $D_n$, we have $D_u^n(x) = 0$ and $D_u^n(y) = \big(-\log(1-xu)\big)^n y$ for $n \geqslant 1$. Hence this implies

$$\begin{aligned} \exp(D_u)(x) &= x, & \exp(D_u)(y) &= (1-xu)^{-1}y, \\ \exp(-D_u)(x) &= x, & \exp(-D_u)(y) &= (1-xu)y. \end{aligned} \tag{14}$$

Consider the dual of (14), then we have

$$\exp(\overline{D_u})(y) = y, \qquad \exp(\overline{D_u})(x) = x(1-yu)^{-1}.$$

Therefore we have

$$\begin{aligned} &\exp(\overline{D_u})\big(\exp(-D_u)(x)\big) = \exp(\overline{D_u})(x) = x(1-yu)^{-1}, \\ &\exp(\overline{D_u})\big(\exp(-D_u)(y)\big) = \exp(\overline{D_u})\big((1-xu)y\big) = \big(1-x(1-yu)^{-1}u\big)y = (1-zu)(1-yu)^{-1}y, \end{aligned}$$

which coincides with that of $\Delta_u$ in Definition 2. For $\exp(\partial_u)$, it will be shown in a corollary of Theorem 5 in the next section. ∎

As a consequence of the theorem, we find a connection among the regularization, derivation relations and Ohno's relations:

**Corollary 1** ([1]) *The following three statements are true and equivalent:*

(i) **(Regularization)** $\Delta_u - 1 \in \mathcal{N}$ ,

(ii) **(Derivation relations)** $\exp(\partial_u) - 1 \in \mathcal{N}$ ,

(iii) **(Weak Ohno's relations)** $\exp(\overline{D}_u) - \exp(D_u) \in \mathcal{N}$ .

Before the proof, we give the table of the upper bounds of the $\dim \mathcal{Z}_k$ which are obtained by derivation relation and (weak) Ohno's relations.

| $k$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | $\cdots$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $d_k$ | 1 | 1 | 1 | 2 | 2 | 3 | 4 | 5 | 7 | 9 | 12 | $\cdots$ |
| Der. rel. | 1 | 1 | 2 | 3 | 6 | 10 | 20 | 38 | 75 | 147 | 305 | $\cdots$ |
| Weak Ohno rel. | 1 | 1 | 2 | 3 | 6 | 10 | 20 | 38 | 75 | 147 | 305 | $\cdots$ |
| Ohno rel. | 1 | 1 | 2 | 3 | 6 | 9 | 18 | 30 | 57 | 99 | 192 | $\cdots$ |

*Proof.* Since we have already shown (i) in Theorem 2, it is enough to prove the equivalence. The equivalence between (i) and (ii) is directly deduced from Theorem 4. Multiply $e_u^D$ from the right to $e_u^\partial - 1 = e^{\overline{D}_u}e^{-D_u} - 1$, then (iii) is deduced from (ii). The reverse direction is same augument. The reason to put the tag 'weak Ohno's relations' is as folows: Since $e^{D_u}(x) = x$ and $e^{D_u}(y) = (1-xu)^{-1}y$, we have

$$\begin{aligned}\exp(D_u)(x^{k_1-1}y\cdots x^{k_n-1}y) &= x^{k_1-1}(1-xu)^{-1}y\cdots x^{k_n-1}(1-xu)^{-1}y\\ &= \sum_{l\geqslant 0}\sum_{e_1+\cdots+e_n} x^{k_1+e_1-1}y\cdots x^{k_n+e_n-1}yu^l = \sum_{l\geqslant 0}\sigma_l(x^{k_1-1}y\cdots x^{k_n-1}y)u^l.\end{aligned}$$

Hence we have $\exp(D_u) = \sum_{l\geqslant 0}\sigma_l u^l$, and $\exp(\overline{D_u}) = \sum_{l\geqslant 0}\overline{\sigma_l}u^l$.

Therefore $e^{\overline{D}_u} - e^{D_u} \in \mathcal{N}$ is equivalent to the weak Ohno's relations $\sigma_l - \overline{\sigma_l} \in \mathcal{N}$ $(l \geqslant 0)$.

∎

## 6 Derivations and automorphisms

Following [2], we discuss the derivations and automorphisms more generally. In this section we define a family of derivations which generalize $\{D_n\}$, $\{\overline{D}_n\}$, $\{\delta_n\}$, $\{\overline{\delta_n}\}$ and $\{\partial_n\}$ in previous section and discuss the corresponding automorphisms via exponential map.

Let $\{a,b\}$ be an arbitrary set of (topological) generators of $\widehat{\mathfrak{H}}$, for example $a$ and $b$ are both linear combinations of $x$ and $y$ which are not proportional. In general, the generators $a$ and $b$ need not be of degree 1 homogeneous elements. We will fix such $\{a,b\}$. In this section we use the letter $D_n$ to express $D_n^{(\alpha,\beta,\gamma,\delta)}$ defined below, unlike the previous section.

**Definition 4** ([2], [1]) *For all $n > 0$ and elements $\alpha,\beta,\gamma,\delta$ in $\mathbb{Q}$, define the derivations $D_n = D_n^{(\alpha,\beta,\gamma,\delta)}$ by*

$$D_n(a) = 0, \qquad D_n(b) = \alpha a^{n+1} + \beta a^n b + \gamma b a^n + \delta b a^{n-1} b,$$

*which are clearly in* $\mathrm{Der}^+(\widehat{\mathfrak{H}})$.

**Proposition 7** ([2], [1]) *Fix the elements $\alpha,\beta,\gamma,\delta \in \mathbb{Q}$, then the sequence of derivations $\{D_n = D_n^{(\alpha,\beta,\gamma,\delta)} \mid n \geqslant 1\}$ commute with each other: $[D_m^{(\alpha,\beta,\gamma,\delta)}, D_n^{(\alpha,\beta,\gamma,\delta)}] = 0$ for all $m,n \geqslant 1$.*

*Proof.* Clearly the $[D_m, D_n]$ is also a derivation on $\widehat{\mathfrak{H}}$. One can check easily the images of $a$ and $b$ are both 0. ∎

To consider any linear combination of $D_n$'s, we use the notation $D_f$ which was introduced in [1]:

**Definition 5** *Let $f(X) = \sum_{n\geqslant 1} c_n X^n \in X\mathbb{Q}[[X]]$ be a formal power series in one indeterminate $X$ without constant term. We define the derivation $D_f \in \mathrm{Der}^+(\widehat{\mathfrak{H}})$ by $D_f = \sum_{n\geqslant 1} c_n D_n$.*

The action on generators $\{a, b\}$ is given by $D_f(a) = 0$ and

$$D_f(b) = \alpha f(a)a + \beta f(a)b + \gamma b f(a) + \delta b\frac{f(a)}{a}b = f(a)u + b\frac{f(a)}{a}v$$

where $u = \alpha a + \beta b$ and $v = \gamma a + \delta b$. The element $\frac{f(a)}{a} \in \widehat{\mathfrak{H}}$ is given by substituting $a$ for $X$ in the power series $\frac{f(X)}{X} \in \mathbb{Q}[[X]]$.

Next, we give the automorphism corresponding to $D_f$ via the exponential map.

**Definition 6** ([2]) *Let $h(X) \in 1 + X\mathbb{Q}[[X]]$ be a power series with constant term 1. We define an automorphism $\Delta_h$ as follows: Denote by $\varepsilon$ and $\varepsilon'$ the two roots of the quadratic equation $T^2 + (\beta+\gamma)T + \alpha\delta = 0$ and put $\omega = \varepsilon - \varepsilon'$. The elements $\varepsilon, \varepsilon'$ and $\omega$ belong to a quadratic extensiton $K$ of $\mathbb{Q}$, but the elements $\varepsilon + \varepsilon' = -(\beta+\gamma)$ and $\varepsilon\varepsilon' = \alpha\delta$ are in $\mathbb{Q}$.*

*Let $\Delta_h \in \mathrm{Aut}^1(\widehat{\mathfrak{H}})$ be the automorphism defined by the following action on generators: $\Delta_h(a) = a$ and*

$$\Delta_h(b) = h(a)^{\beta+\varepsilon}\Big[b + \frac{h(a)^{-\omega}-1}{-\omega}(\alpha a - \varepsilon b)\Big] \times \Big[1 + \frac{h(a)^{\omega}-1}{\omega a}(\varepsilon a - \delta b)\Big]^{-1} h(a)^{\gamma+\varepsilon} \tag{15}$$

$$= h(a)^{\beta}\Big[(h(a)^{\varepsilon} - h(a)^{\varepsilon'})\alpha a - (\varepsilon' h(a)^{\varepsilon} - \varepsilon h(a)^{\varepsilon'})b\Big]$$

$$\times \Big[(\varepsilon h(a)^{\varepsilon} - \varepsilon' h(a)^{\varepsilon'}) - \frac{h(a)^{\varepsilon} - h(a)^{\varepsilon'}}{a}\delta b\Big]^{-1} h(a)^{-\beta} \tag{16}$$

*where $h(a)^{\lambda} = \exp(\lambda \log h(a))$ for any $\lambda \in K$, and the quotients $(h(a)^{\omega}-1)/\omega a$ and $(h(a)^{\varepsilon} - h(a)^{\varepsilon'})/a$ define the elements of $K\langle\langle x, y\rangle\rangle$, since each numerator has no constant term, one can divide it by $a$. In the case $\omega = 0$, we regard the elements $(h(a)^{-\omega}-1)/(-\omega)$ and $(h(a)^{\omega}-1)/\omega a$ as $\log h(a)$ and $(\log h(a))/a$ respectively. Since the expression (16) is symmetric in $\varepsilon$ and $\varepsilon'$, it defines an element of $\widehat{\mathfrak{H}}$.*

First we check the expression (15) equals (16):

$$\begin{aligned} A_h := {}& h(a)^{\beta+\varepsilon}\Big[b + \frac{h(a)^{-\omega}-1}{-\omega}(\alpha a - \varepsilon b)\Big] \\ = {}& h(a)^{\beta}\Big[\frac{h(a)^{\varepsilon} - h(a)^{\varepsilon'}}{\omega}\alpha a + \big(h(a)^{\varepsilon} - \frac{\varepsilon(h(a)^{\varepsilon} - h(a)^{\varepsilon'})}{\omega}\big)b\Big] \\ = {}& h(a)^{\beta}\Big[\frac{h(a)^{\varepsilon} - h(a)^{\varepsilon'}}{\omega}\alpha a - \frac{\varepsilon' h(a)^{\varepsilon} - \varepsilon h(a)^{\varepsilon'}}{\omega}b\Big]. \end{aligned} \tag{17}$$

On the other hand, we have

$$
\begin{aligned}
B_h^{-1} :&= \Big[1+\frac{h(a)^{\omega}-1}{\omega a}(\varepsilon a-\delta b)\Big]^{-1}h(a)^{\gamma+\varepsilon} = \Big[1+\frac{h(a)^{\omega}-1}{\omega a}(\varepsilon a-\delta b)\Big]^{-1}h(a)^{-(\beta+\varepsilon')}\\
&= \Big[h(a)^{\varepsilon'}+\frac{\varepsilon(h(a)^{\varepsilon}-h(a)^{\varepsilon'})}{\omega}-\frac{h(a)^{\varepsilon}-h(a)^{\varepsilon'}}{\omega a}\delta b\Big]^{-1}h(a)^{-\beta}\\
&= \Big[\frac{\varepsilon h(a)^{\varepsilon}-\varepsilon' h(a)^{\varepsilon'}}{\omega}-\frac{h(a)^{\varepsilon}-h(a)^{\varepsilon'}}{\omega a}\delta b\Big]^{-1}h(a)^{-\beta}. \qquad (18)
\end{aligned}
$$

Thus we have shown that (15)=(16). Since the equation (18) defines an invertible element of $\widehat{\mathfrak{H}}$, we denote the inverse by $B_h$. Hence we have $\Delta_h(b) = A_h B_h^{-1}$.

**Theorem 5** ([2]) *For any $f(X) \in X\mathbb{Q}[[X]]$, set $h(X) = e^{f(X)} \in 1 + X\mathbb{Q}[[X]]$. Then we have*

$$\Delta_h = \exp(D_f). \qquad (19)$$

*Proof.* For the derivation $D_f$ we can consider the 1-dimensional commutative Lie subalgebra $\{tD_f = D_{tf}\}$ spaned by $D_f$. Then the image of the Lie algebra under the exponential map forms a 1-parameter subgroup $\{e^{tD_f} = e^{D_{tf}}\}$ of $\mathrm{Aut}^1(\widehat{\mathfrak{H}})$. The tangent vector along the path at the unit (identity automorphism on $\widehat{\mathfrak{H}}$) corresponds to $\log(e^{D_f}) = D_f$. Therefore it is enough to show that (i) $\frac{d}{dt}\Delta_{h^t}|_{t=0} = D_f$, and (ii) $\Delta_{gh} = \Delta_g\Delta_h$ for $g, h \in 1 + X\mathbb{Q}[[X]]$, i.e., the map $h \mapsto \Delta_h$ is a group homomorphism.

For (i), from the definition of $D_f$ and $\Delta_h$ it is clear that $\frac{d}{dt}\Delta_{h^t}(a)|_{t=0} = D_f(a) = 0$. Next we have from (15)

$$\Delta_{h^t}(b) = h^{(\beta+\varepsilon)t}\Big[b+\frac{h^{-\omega t}-1}{-\omega}(\alpha a-\varepsilon b)\Big]\Big[1+\frac{h^{\omega t}-1}{\omega a}(\varepsilon a-\delta b)\Big]^{-1}h^{(\gamma+\varepsilon)t},$$

where we write $h$ for $h(a)$ for simplicity. By using the formula $\frac{d}{dt}h^{\lambda t}|_{t=0} = \frac{d}{dt}e^{\lambda t f(a)}|_{t=0} = \lambda f(a)$ for $\lambda \in K$, we have

$$
\begin{aligned}
\frac{d}{dt}\Delta_{h^t}(b)\big|_{t=0} &= (\beta+\varepsilon)f(a)b + f(a)(\alpha a-\varepsilon b) - b\frac{f(a)}{a}(\varepsilon a-\delta b) + b(\gamma+\varepsilon)f(a)\\
&= \alpha f(a)a + \beta f(a)b + \gamma b f(a) + \delta b\frac{f(a)}{a}b = f(a)u + b\frac{f(a)}{a}v.
\end{aligned}
$$

This coincides with the expression in Definition 5. For the proof of (ii) we need the following lemma, which is proved in [2].

**Lemma 2** *For any $g, h \in 1 + X\mathbb{Q}[[X]]$, we obtain*

$$\Delta_g(A_h)B_g = A_{gh}, \qquad \Delta_g(B_h)B_g = B_{gh} \qquad (20)$$

*where $A_h, B_h$ are the elements defined above.*

Using this lemma we can prove (ii): $\Delta_g(\Delta_h(a)) = a = \Delta_{gh}(a)$ and

$$\Delta_g(\Delta_h(b)) = \Delta_g(A_h B_h^{-1}) = (A_{gh}B_g^{-1})(B_{gh}B_g^{-1})^{-1} = A_{gh}B_{gh}^{-1} = \Delta_{gh}(b).$$

∎

The following theorem is a special case of Theorem 5, but is worth stating separately because of the conciseness of the expression.

**Theorem 6** ([2]) *Suppose that $\alpha, \beta, \gamma, \delta \in \mathbb{Q}$ satisfy $\alpha\delta - \beta\gamma = 0$. Then the derivation $D_f$ is defined by the images $D_f(a) = 0$, $D_f(b) = w\frac{f(a)}{a}w'$ on the generators for some $w, w' \in \mathbb{Q}a + \mathbb{Q}b$ and the automorphism $\Delta_h = \exp(D_f)$ for $h = e^f$ sends the generators to*

$$\Delta_h(a) = a, \quad \Delta_h(b) = \left[b + \frac{h(a)^{\beta-\gamma} - 1}{\beta - \gamma}u\right]\left[1 - \frac{h(a)^{\beta-\gamma} - 1}{(\beta - \gamma)a}v\right]^{-1}, \tag{21}$$

*where $u = \alpha a + \beta b$, $v = \gamma a + \delta b$.*

**Corollary 2** *We have*

$$\exp(\partial_u)(x) = x(1 - yu)^{-1}, \quad \exp(\partial_u)(z) = z,$$

*where $\partial_u$ is a derivation defined in Section 5.*

*Proof.* Take the generator $\{a, b\}$ as $\{z = x + y, x\}$ and $(\alpha, \beta, \gamma, \delta)$ as $(0, 0, 1, -1)$, which satisfies the assumption of Theorem 6. Moreover put $f(X) = -\log(1 - uX)$ for parameter $u$, then $D_f^{(\alpha,\beta,\gamma,\delta)} = \partial_u$. In this case we obtain $\exp(\partial_u)(x) = x(1 - yu)^{-1}$, and $\exp(\partial_u)(z) = z$ from the theorem. ∎

## 7 Linearized double shuffle relations

In this section we estimate the number of generators of the algebra of MZV's of given weight $k$ and depth $n$ by considering the extended double shuffle relation modulo elements of lower depth and products.

Let $\mathcal{Z} = \bigoplus_{k\geqslant 0} \mathcal{Z}_k$ be the graded algebra generated by all MZV's over $\mathbb{Q}$, where $\mathcal{Z}_k$ is the $\mathbb{Q}$-vector space generated by MZV's 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 MZV's of weight $k$ and depth $\leqslant n$. Thus $\mathcal{Z}^{(n)} = \bigoplus_{k\geqslant 0} \mathcal{Z}_k^{(n)}$ gives a corresponding filtration $\mathcal{Z} = \bigcup_{n\geqslant 0} \mathcal{Z}^{(n)}$ on the 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 ideal. The grading and filtration are induced to the cotanjent space $\mathcal{T} = \mathcal{I}/\mathcal{I}^2$. The dimension of the space $\mathcal{T}_k$, the weight $k$ component of $\mathcal{T}$, coincides with the minimum number $D_k$ of algebra generators of $\mathcal{Z}$ in weight $k$. We can consider the bigraded vector space $\mathcal{M} = \mathrm{gr}(\mathcal{T})$ associated to 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 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 $D_k = \sum_{n=1}^{k-1} D_{k,n}$. There is a conjectural formula giving these dimensions $D_{k,n}$, due to Broadhurst and Kreimer.

**Conjecture 2** ([3]) *The number $D_{k,n}$ of algebra generaters of weight $k$ and depth $n$ are given by*

$$\prod_{\substack{k\geqslant 2\\ n\geqslant 1}} (1 - x^k y^n)^{-D_{k,n}} = \frac{1}{1 - x^2}\left(1 - \frac{x^3 y}{1 - x^2} + \frac{x^{12}y^2(1 - y^2)}{(1 - x^4)(1 - x^6)}\right)^{-1}.$$

Following is the table of this conjectural values of $D_{k,n}$.

| $k$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | $\cdots$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $D_k$ | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 2 | 2 | 3 | 3 | 4 | 5 | 7 | 8 | 11 | $\cdots$ |
| $D_{k,1}$ | | 1 | 1 | | 1 | | 1 | | 1 | | 1 | | 1 | | 1 | | 1 | | 1 | $\cdots$ |
| $D_{k,2}$ | | | | | | | | 1 | | 1 | | 1 | | 2 | | 2 | | 2 | | $\cdots$ |
| $D_{k,3}$ | | | | | | | | | | | 1 | | 2 | | 2 | | 4 | | 5 | $\cdots$ |
| $D_{k,4}$ | | | | | | | | | | | | 1 | | 1 | | 3 | | 5 | | $\cdots$ |
| $D_{k,5}$ | | | | | | | | | | | | | | | 1 | | 2 | | 5 | $\cdots$ |
| $D_{k,6}$ | | | | | | | | | | | | | | | | | | 1 | | $\cdots$ |

In [1], certain vector space $DS_n(d)$ was intrduced for each $n, d > 0$ whose dimension gives an upper bound of the numbers $D_{n+d,n}$. In this section, we summarize a result in [1] and estimate the dimensions of $DS_n(d)$ for small $n$. As a consequence, we obtain a non-trivial upper bound of $D_{k,n}$ for small $n$.

Let $\mathfrak{S}_n$ be the symmetric group of degree $n$ and $\mathbb{Z}[\mathfrak{S}_n]$ its group algebra. We denote $\mathbb{Q}[x_1, \dots, x_n]$ the space of (commutative) polynomials in $n$ variables with rational coefficients and by $\mathbb{Q}[x_1, \dots, x_n]_{(d)}$ its subspace of homogeneous polynomials of degree $d$. The group $\mathfrak{S}_n$ acts on these spaces by permutation of variables: $(f|\sigma)(x_1, \dots, x_n) = f(x_{\sigma^{-1}(1)}, \dots, x_{\sigma^{-1}(n)})$. For any $\sigma$ and $\tau$ in $\mathfrak{S}_n$, it holds $f|(\sigma\tau) = (f|\sigma)|\tau$. We extend the action $\mathbb{Z}$-linearly to an action of $\mathbb{Z}[\mathfrak{S}_n]$.

Define the *double shuffle subspace* $DS_n$ of $\mathbb{Q}[x_1, \dots, x_n]$ as follws: For each integer $l$ with $1 \leqslant l < n$, define the *l-th shuffle element* by $sh_l = \sum \sigma \in \mathbb{Z}[\mathfrak{S}_n]$, where the sum runs over the element $\sigma \in \mathfrak{S}_n$ satisfying $\sigma(1) < \cdots < \sigma(l)$ and $\sigma(l+1) < \cdots < \sigma(n)$. Then

$$DS_n = \{f \in \mathbb{Q}[x_1, \dots, x_n] \mid f|sh_l = f^{\sharp}|sh_l = 0 \text{ for } 1 \leqslant l < n\} \tag{22}$$

where for any polynomial $f \in \mathbb{Q}[x_1, \dots, x_n]$, we put

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

We write $DS_n(d)$ for the degree $d$ part of $DS_n$. For example, the case $n = 2$ is

$$DS_2 = \left\{ f \in \mathbb{Q}[x_1, x_2] \;\middle|\; \begin{matrix} f(x_1, x_2) + f(x_2, x_1) = 0, \\ f(x_1 + x_2, x_1) + f(x_1 + x_2, x_2) = 0 \end{matrix} \right\}.$$

**Theorem 7** ([1]) *For all $k > n > 0$, we have*

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

It is conjectured that $D_{k,n} = \dim DS_n(k - n)$ for $n > 1$.

For example, the case $n = 2$ and $d = 6$, the space $DS_2(6)$ is spaned by a single polynomial: $DS_2(6) = \langle\ 2x_1^5 x_2 - 2x_1 x_2^5 - 5x_1^4 x_2^2 + 5x_1^2 x_2^4\ \rangle_{\mathbb{Q}}$. By the therem, $D_{8,2} \leqslant 1$ is deduced. We will sketch a proof of the theorem after some preliminaries. A part of the proof is different from the original one given in [1].

Recall that $\mathfrak{H}^1$ is generated by $z_k = x^{k-1}y$, $(k \geqslant 1)$. For a fixed $n$, consider the generating function

$$F_n(x_1, \dots, x_n) = \sum_{\mathbf{k}} z_{k_1} \cdots z_{k_n} x_1^{k_1 - 1} \cdots x_n^{k_n - 1} \in \mathfrak{H}^1[[x_1, \dots, x_n]]$$

where the sum runs over all index sets $\mathbf{k} = (k_1, \ldots, k_n)$ (allowing $k_1 = 1$) of depth $n$, and $\mathfrak{H}^1[[x_1, \ldots, x_n]]$ is the algebra of formal power series in $n$ variables with coefficients ring $\mathfrak{H}^1$. For each product $\bullet$, ($\bullet = \cdot$, $*$ or ш), we can consider the algebra structure on $\mathfrak{H}^1[[x_1, \ldots, x_n]]$ which is isomorphic to the tensor algebra $\mathfrak{H}^1_\bullet \widehat{\otimes} \mathbb{Q}[[x_1, \ldots, x_n]]$. For $n \geqslant 2$ and $1 \leqslant l < n$, we have easily

$$F_l(x_1, \ldots, x_l) \cdot F_{n-l}(x_{l+1}, \ldots, x_n) = F_n(x_1, \ldots, x_n).$$

**Proposition 8** *For any $n \geqslant 2$ and $1 \leqslant l < n$, we have*

(i) $F_l(x_1, \ldots, x_l) * F_{n-l}(x_{l+1}, \ldots, x_n)$

$$\begin{aligned} &= F_1(x_1) \cdot \big(F_{l-1}(x_2, \ldots, x_l) * F_{n-l}(x_{l+1}, \ldots, x_n)\big) \\ &\quad + F_1(x_{l+1}) \cdot \big(F_l(x_1, \ldots, x_l) * F_{n-l-1}(x_{l+2}, \ldots, x_n)\big) \\ &\quad + \frac{F_1(x_1) - F_1(x_{l+1})}{x_1 - x_{l+1}} \cdot \big(F_{l-1}(x_2, \ldots, x_l) * F_{n-l-1}(x_{l+2}, \ldots, x_n)\big). \end{aligned}$$

(ii) $F_l(x_1, \ldots, x_l) ш F_{n-l}(x_{l+1}, \ldots, x_n)$

$$\begin{aligned} &= F_1(x_1 + x_{l+1}) \cdot \big(F_{l-1}(x_2, \ldots, x_l) ш F_{n-l}(x_{l+1}, \ldots, x_n)\big) \\ &\quad + F_1(x_1 + x_{l+1}) \cdot \big(F_l(x_1, \ldots, x_l) ш F_{n-l-1}(x_{l+2}, \ldots, x_n)\big). \end{aligned}$$

*Proof.* From (3) and (4), it is enough to show the case $n = 2$ and $l = 1$. For (i), we have

$$\begin{aligned} F_1(x_1) * F_1(x_2) &= \Big(\sum z_{k_1} x_1^{k_1-1}\Big) * \Big(\sum z_{k_2} x_2^{k_2-1}\Big) = \sum z_{k_1} * z_{k_2} x_1^{k_1-1} x_2^{k_2-1} \\ &= \sum (z_{k_1} z_{k_2} + z_{k_2} z_{k_1} + z_{k_1+k_2}) x_1^{k_1-1} x_2^{k_2-1} \\ &= \sum z_{k_1} z_{k_2} x_1^{k_1-1} x_2^{k_2-1} + \sum z_{k_2} z_{k_1} x_1^{k_1-1} x_2^{k_2-1} + \sum z_{k_1+k_2} x_1^{k_1-1} x_2^{k_2-1} \\ &= F_2(x_1, x_2) + F_2(x_2, x_1) + \frac{F_1(x_1) - F_1(x_2)}{x_1 - x_2}. \end{aligned}$$

For (ii), we use $F_1(x_i) = \sum_{k_1 \geqslant 1} x^{k_1-1} y x_i^{k_1-1} = (1 - x x_i)^{-1} y = y + x x_i F_1(x_i)$, for $i = 1, 2$, and (4), then

$$\begin{aligned} F_1(x_1) ш F_1(x_2) &= \big(y + x x_1 F_1(x_1)\big) ш \big(y + x x_2 F_1(x_2)\big) \\ &= y ш y + y ш x\big(x_1 F_1(x_1) + x_2 F_1(x_2)\big) + x x_1 F_1(x_1) ш x x_2 F_1(x_2) \\ &= y ш y + y x\big(x_1 F_1(x_1) + x_2 F_1(x_2)\big) + x\Big(y ш \big(x_1 F_1(x_1) + x_2 F_1(x_2)\big)\Big) \\ &\quad + x\big(x_1 F_1(x_1) + x x_2 F_1(x_2)\big) + x\big(x x_1 F_1(x_1) + x_2 F_1(x_2)\big) \\ &= y ш y + y\big(F_1(x_1) - y + F_1(x_2) - y\big) + x\Big(y ш \big(x_1 F_1(x_1) + x_2 F_1(x_2)\big)\Big) \\ &\quad + x\Big(x_1 F_1(x_1) ш \big(F_1(x_2) - y\big)\Big) + x\Big(\big(F_1(x_1) - y\big) ш x_2 F_1(x_2)\Big) \\ &= y\big(F_1(x_1) + F_1(x_2)\big) + x(x_1 + x_2)\big(F_1(x_1) ш F_1(x_2)\big). \end{aligned}$$

Therefore

$$\begin{aligned} F_1(x_1) ш F_1(x_2) &= \big(1 - x(x_1 + x_2)\big)^{-1} y\big(F_1(x_1) + F_1(x_2)\big) \\ &= F_1(x_1 + x_2)\big(F_1(x_1) + F_1(x_2)\big) = F_2(x_1 + x_2, x_1) + F_2(x_1 + x_2, x_2). \end{aligned}$$

∎

We can define a filtered graded structure on $\mathfrak{H}^1_\bullet$. The grading and filtration are defined by the total degree and partial degree in $y$ respectively. The space $\mathfrak{H}^0_\bullet$ is a filterd graded subalgebra of $\mathfrak{H}^1_\bullet$ for each $\bullet = *$ or ш. Then the both evaluation map $Z : \mathfrak{H}^0 \longrightarrow \mathcal{Z}$ and regualization map $\mathrm{reg}_\bullet : \mathfrak{H}^1_\bullet \longrightarrow \mathfrak{H}^0_\bullet$ are morphisms preserving the grading and filtration.

Let $\iota^{(n)}_k : \mathcal{Z}^{(n)}_k \longrightarrow \mathcal{M}^{(n)}_k$ be the natural surjection and $\iota^{(n)} : \mathcal{Z}^{(n)} \longrightarrow \mathcal{M}^{(n)}$ be its direct sum : $\iota^{(n)} = \bigoplus_k \iota^{(n)}_k$. For each product $\bullet = *$ or ш, consider the composition map

$$\iota^{(n)} \circ Z \circ \mathrm{reg}_\bullet : \mathfrak{H}^{1,(n)}_\bullet \longrightarrow \mathfrak{H}^{0,(n)}_\bullet \longrightarrow \mathcal{Z}^{(n)} \longrightarrow \mathcal{M}^{(n)},$$

where $\mathfrak{H}^{1,(n)}_\bullet$ is the $n$-th filtered subspace of $\mathfrak{H}^1_\bullet$, namely which is generated by the words whose partial degree in $y$ are less than or equal to $n$. By the definition of $\mathcal{M}$, the image of the subspace $\mathfrak{H}^{1,(n-1)}_\bullet$ in $\mathcal{M}^{(n)}$ via this composition map is $\{0\}$. Furthermore, the image of the product $f \bullet f' \in \mathfrak{H}^{1,(n)}_\bullet$ in $\mathcal{M}^{(n)}$ is also $\{0\}$ for $f \in \mathfrak{H}^{1,(l)}_\bullet$ and $f' \in \mathfrak{H}^{1,(n-l)}_\bullet$.

**Lemma 3** ([1]) *We have a following equation in $\mathcal{M}^{(n)}\widehat{\otimes}\mathbb{Q}[[x_1,\ldots,x_n]]$*

$$\iota^{(n)} \circ Z \circ \mathrm{reg}_*\big(F_n(x_1,\ldots,x_n)\big) = \iota^{(n)} \circ Z \circ \mathrm{reg}_{ш}\big(F_n(x_1,\ldots,x_n)\big),$$

*where the composition maps acts on the coefficient part.*

*Proof.* From (8), the gap between $Z \circ \mathrm{reg}_*$ and $Z \circ \mathrm{reg}_{ш}$ is given by the map $\rho$ defined in (6). The lemma follows from Theorem 1 and the fact that the coefficient of $\rho(T^i)$ is contained in the algebra generated by Riemann zeta values i.e., MZV's of depth 1. ∎

Let $\widetilde{\mathcal{M}}$ be the bigraded $\mathbb{Q}$-algebra associated to the filterd graded algebra $\mathcal{Z}/\mathcal{I}^2$. As a $\mathbb{Q}$-vector space $\widetilde{\mathcal{M}} = \mathbb{Q} \oplus \mathcal{M}$, here $\mathbb{Q}$ is regarded as the $(0,0)$-degree component of $\widetilde{\mathcal{M}}$. In the following, we think $\mathcal{M}$ as a subspace of $\widetilde{\mathcal{M}}$. Consider $\widetilde{\mathcal{M}}[[x_1,\ldots,x_n]]$ the algebra of power series with $\widetilde{\mathcal{M}}$ coefficients and extend the $\mathbb{Z}[\mathfrak{S}_n]$-action to $\widetilde{\mathcal{M}}[[x_1,\ldots,x_n]]$ in the obvious way.

**Definition 7** *Define a power series in $\widetilde{\mathcal{M}}[[x_1,\ldots,x_n]]$ by*

$$\overline{F_n}(x_1,\ldots,x_n) := \iota^{(n)} \circ Z \circ \mathrm{reg}_*\big(F_n(x_1,\ldots,x_n)\big) = \iota^{(n)} \circ Z \circ \mathrm{reg}_{ш}\big(F_n(x_1,\ldots,x_n)\big).$$

**Proposition 9** ([1]) *For $1 \leqslant l < n$, we have*

$$(\overline{F_n}|sh_l)(x_1,\ldots,x_n) = (\overline{F_n}^\sharp|sh_l)(x_1,\ldots,x_n) = 0.$$

*Hence the polynomial $\overline{F_n}(d)$, the homogeneous degree d part of $\overline{F_n}$ , is in $\widetilde{\mathcal{M}} \otimes DS_n(d)$.*

*Proof.* For each product $\bullet = *$ or ш, apply $\iota^{(n)} \circ Z \circ \mathrm{reg}_\bullet$ to Proposition 8, then

$$\begin{aligned} 0 = \overline{F_1}(x_1)\big(\overline{F_{l-1}}(x_2,\ldots,x_l)\overline{F_{n-l}}(x_{l+1},\ldots,x_n)\big) \\ + \overline{F_1}(x_{l+1})\big(\overline{F_l}(x_1,\ldots,x_l)\overline{F_{n-l-1}}(x_{l+2},\ldots,x_n)\big) = (\overline{F_n}|sh_l)(x_1,\ldots,x_n). \end{aligned}$$

and

$$\begin{aligned} 0 = \overline{F_1}(x_1 + x_{l+1},) \cdot \big(\overline{F_{l-1}}(x_2,\ldots,x_l)\overline{F_{n-l}}(x_{l+1},\ldots,x_n)\big) \\ + \overline{F_1}(x_1 + x_{l+1}) \cdot \big(\overline{F_l}(x_1,\ldots,x_l)\overline{F_{n-l-1}}(x_{l+2},\ldots,x_n)\big) = (\overline{F_n}^\sharp|sh_l)(x_1,\ldots,x_n). \end{aligned}$$

Thus we conclude the proof. ∎

*Proof of Theorem 7.* As a corollary of Proposition 9, we can show that the dimension of the $\mathbb{Q}$-vector subspace of $\mathcal{M}_k^{(n)}$ spaned by the coefficients of $\overline{F_n}(k-n)$ is less than or equal to the dimension of $DS_n(k-n)$. Since images of all MZV's of weight $k$ and depth $n$ in $\mathcal{M}_k^{(n)}$ are appered as the coefficients of $\overline{F_n}(k-n)$, we have $\dim \mathcal{M}_k^{(n)} \leqslant \dim DS_n(k-n)$, which proves the theorem. ∎

In the rest of this section we give some estimates of the space $DS_n(d)$.
Let $T_n = \begin{pmatrix} 1 & 2 & \cdots & n \\ n & n-1 & \cdots & 1 \end{pmatrix} \in \mathfrak{S}_n$. For $n, d \geqslant 1$, define the space

$$W_n(d) := \{f \in \mathbb{Q}[x_1, \ldots, x_n]_{(d)} \mid f^\sharp|sh_l = 0 \ (1 \leqslant l < n),\ f|T_n = (-1)^{n-1} f\}.$$

**Proposition 10** ([1]) *We have* (i) $DS_n(d) \subset W_n(d)$, (ii) $W_n(d) = \{0\}$ *if $d$ is odd.*

*Proof.* Omitted. The space $W_n(d)$ is equal to the space '$ShC_n(d)$' in [1].

**Corollary 3 (Parity result)** *If $d$ is odd, then $DS_n(d) = \{0\}$ for every $n > 0$. Conseqently $D_{k,n} = 0$ if $k \not\equiv n \bmod 2$.*

This result was proved independently by Tsumura [13] by a different method.

For small $n$, we can compute explicitly the dimension of the space $W_n(d)$, which gives a non-trivial upper bound of the number $D_{n+d,n}$.

**Proposition 11** ([6]) *Let $E_n(t) = \sum_{d \geqslant 0} \dim W_n(d) t^d$ be the Poincaré series of the spaces $W_n(d)$. Then,*

(i) $E_2(t) = \dfrac{t^6}{(1-t^2)(1-t^6)}$,

(ii) $E_3(t) = \dfrac{t^2}{(1-t^2)^2(1-t^6)}$,

(iii) $E_4(t) = \dfrac{t^4(1+t^4)}{(1-t^2)^3(1-t^{10})}$,

(iv) $E_5(t) = \dfrac{t^2(1+t^2+4t^4+2t^6+5t^8+4t^{10}+4t^{12}+t^{14}+2t^{16})}{(1-t^2)^2(1-t^6)^2(1-t^{10})}$.

We give the table of $\dim W_n(k-n)$ up to $n \leqslant 5$ and $k \leqslant 19$ as follows.

| $n \backslash k$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | ⋯ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | 1 | 1 | | 1 | | 1 | | 1 | | 1 | | 1 | | 1 | | 1 | | 1 | ⋯ |
| 2 | | | | | | | | 1 | | 1 | | 1 | | 2 | | 2 | | 2 | | ⋯ |
| 3 | | | | | 1 | | 2 | | 3 | | 5 | | 7 | | 9 | | 12 | | 15 | ⋯ |
| 4 | | | | | | | | 1 | | 3 | | 7 | | 13 | | 21 | | 32 | | ⋯ |
| 5 | | | | | | | 1 | | 3 | | 9 | | 19 | | 36 | | 66 | | 108 | ⋯ |

**Acknowledgements** I wish to express my appriciation to Prof. Yasuo Ohno for giving the opportunity to talk.

## References


[1] K. Ihara, M. Kaneko, D. Zagier, *Derivations and double shuffle relations for multiple zeta values*, preprint (2004), Max-Planck-Institut für Mathematik preprint series 2004-100.

[2] K. Ihara, *Derivations and automorphisms on the algebra of non-commutative power series*, Math. J. of Okayama university, to appear.

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

[4] A. B. Goncharov, *Multiple $\zeta$-values, Galois groups and geometry of modular varieties*, Progress in Math. **201** (2001), 361–392

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

[6] K. Ihara, M. Kaneko, D. Zagier, *The double shuffle vector space*, planned.

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

[8] L. Boutet de Monvel, *Remarques sur les séries logarithmiques divergentes*, lecture at the workshop "Polylogarithmes et conjecture de Deligne-Ihara", C.I.R.M. (Luminy) (2000).

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

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

[11] C. Reutenauer, *Free Lie Algebras*, Oxford Science Publications (1993).

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

[13] H. Tsumura, *Combinatorial relations for Euler-Zagier sums* , Acta Arith. **111.1** (2004), 27–42.

[14] D. Zagier, *Values of zeta functions and their applications*, Progress in Math. **120** (1994), 497–512.