Special year in number theory at IMSc

Lectures on Multiple Zeta Values IMSC 2011

by

Michel Waldschmidt

Here are the notes of the lectures given in Chennai IMSc in April 2011 for the Special Year on Number Theory, based on notes taken during the lectures by Tapas Chatterjee. The author is also thankful to Sanoli Gun for the organization of this special year and for taking care of the publication of these notes.

1 Introduction to the results of F. Brown on Zagier's Conjecture

1.1 Zeta values: main Diophantine conjecture

Riemann zeta function

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

has been considered before Riemann by Euler for integer values of the variable s, both positive and negative ones. Among the many results he proved are

$$\zeta(2) = \frac{\pi^2}{6} \quad \text{and} \quad \frac{\zeta(2n)}{\zeta(2)^n} \in \mathbf{Q}$$

for any integer $n \geq 1$.

A quite ambitious goal is to determine the algebraic relations among the numbers

$$\pi,\ \zeta(3),\ \zeta(5), \ldots, \zeta(2n+1), \ldots$$

The expected answer is disappointingly simple: it is widely believed that there are no relations, which means that these numbers should be algebraically independent:

Conjecture 1. *For any $n \geq 0$ and any nonzero polynomial $P \in \mathbf{Z}[T_0, \ldots, T_n]$,*

$$P\bigl(\pi, \zeta(3), \zeta(5), \ldots, \zeta(2n+1)\bigr) \neq 0.$$

If true, this property would mean that there is no interesting algebraic structure.

There are very few results on the arithmetic nature of these numbers, even less on their independence: it is known that π is a transcendental numbers, hence, so are all $\zeta(2n)$, $n \geq 1$. It is also known that $\zeta(3)$ is irrational (Apéry, 1978), and that infinitely many $\zeta(2n+1)$ are irrational (further sharper results have been achieved by T. Rivoal and others – see [9, 14, 5]). But so far, it has not been disproved that all these numbers lie in the ring $\mathbf{Q}[\pi^2]$ (see the Open Problem 2).

1.2 Multizeta values: Zagier's conjecture

The situation changes drastically if we enlarge our set so as to include the so-called Multiple Zeta Values (MZV, also called *Polyzeta values, Euler-Zagier numbers* or *multiple harmonic series*):

$$\zeta(s_1,\dots,s_k)=\sum_{n_1>n_2>\cdots>n_k\geq 1}\frac{1}{n_1^{s_1}\cdots n_k^{s_k}}$$

which are defined for k, $s_1,\dots,s_k$ positive integers with $s_1\geq 2$. There are plenty of relations between them, providing a rich algebraic structure.

We will call k the *length* of the tuple $\underline{s}=(s_1,\dots,s_k)$ and $|s|:=s_1+\cdots+s_k$ the *weight* of this tuple. There are 2^{p-2} tuples $\underline{s}$ of weight p with $s_1\geq 2$ and $s_j\geq 1$ for $2\leq j\leq k$. The length k and the weight p are related by $k+1\leq p$.

One easily gets quadratic relations between MZV when one multiplies two such series: it is easy to express the product as a linear combination of MZV. We will study this phenomenon in detail, but we just give one easy example. Splitting the set of (n,m) with $n\geq 1$ and $m\geq 1$ into three disjoint subsets with respectively $n>m$, $m>n$ and $n=m$, we deduce, for $s\geq 2$ and $s'\geq 2$,

$$\sum_{n\geq 1}n^{-s}\sum_{m\geq 1}m^{-s'}=\sum_{n>m\geq 1}n^{-s}m^{-s'}+\sum_{m>n\geq 1}m^{-s'}n^{-s}+\sum_{n\geq 1}n^{-s-s'},$$

which is the so–called *Nielsen Reflexion Formula*:

$$\zeta(s)\zeta(s')=\zeta(s,s')+\zeta(s',s)+\zeta(s+s')$$

for $s\geq 2$ and $s'\geq 2$. For instance,

$$\zeta(2)^2=2\zeta(2,2)+\zeta(4).$$

Such expressions of the product of two zeta values as a linear combination of zeta values, arising from the product of two series, will be called "stuffle relations" (see §8.1).

They show that the $\mathbf{Q}$–vector space spanned by the multiple zeta values is in fact an algebra: a product of linear combinations of numbers of the form $\zeta(\underline{s})$ is again a linear combination of such numbers. As mentioned above, it has not been disproved that this algebra is $\mathbf{Q}[\pi^2]$. While looking at the result on the formal symbols representing the MZV, one should keep in mind that the following Open Problem 2 is not yet solved. Many results on these symbols are known, but almost nothing is known on the kernel of the corresponding specialization, which maps a symbol onto the corresponding real number.

We denote by $\mathfrak{Z}$ the $\mathbf{Q}$–vector space spanned by the numbers $\zeta(\underline{s})$. For $p\geq 2$, we denote by $\mathfrak{Z}_p$ the $\mathbf{Q}$-subspace of $\mathfrak{Z}$ spanned by the numbers $\zeta(\underline{s})$ with $|\underline{s}|=p$. For $k\geq 1$, we denote by $\mathcal{F}^k\mathfrak{Z}$ the $\mathbf{Q}$-subspace of $\mathfrak{Z}$ spanned by the numbers $\zeta(\underline{s})$ with $\underline{s}$ of length $\leq k$. Finally, for $p\geq k+1\geq 2$, we denote by $\mathcal{F}^k\mathfrak{Z}_p$ the $\mathbf{Q}$-subspace of $\mathfrak{Z}$ spanned by the numbers $\zeta(\underline{s})$ with $|\underline{s}|$ of weight p and length $\leq k$. The inclusion $\mathcal{F}^k\mathfrak{Z}_p\subset\mathcal{F}^k\mathfrak{Z}\cap\mathfrak{Z}_p$ is plain, that there is equality is only a conjecture. It is also conjectured but not proved that the weight defines a graduation on $\mathfrak{Z}$. It is a fact that the subspaces $\mathcal{F}^k\mathfrak{Z}$ define an increasing filtration of the algebra $\mathfrak{Z}$ (see §5.4), butit is not proved that this filtration is not the trivial one: for instance, it could happen that

$$\mathcal{F}^k\mathfrak{Z}_p = \mathfrak{Z}_p = \mathfrak{Z} = \mathbf{Q}[\pi^2]$$

for all $k \geq 1$ and $p \geq 2$. For $p \geq 2$, the space $\mathcal{F}^1\mathfrak{Z}_p$ has dimension 1, it is spanned by $\zeta(p)$. From Rivoal's result we know that $\mathfrak{Z}_p \neq \mathbf{Q}$ for infinitely many odd p.

Open Problem 2. *Is–it true that* $\mathfrak{Z} \neq \mathbf{Q}[\pi^2]$*?*

All know linear relations that express a multizeta $\zeta(\underline{s})$ as a linear combination of such numbers are homogeneous for the weight. The next conjecture (which is also still open) is that any linear relations among these numbers splits into homogeneous linear relations.

Conjecture 3. *The* $\mathbf{Q}$*–subspaces* $\mathfrak{Z}_p$ *of* $\mathbf{R}$ *are in direct sum:*

$$\bigoplus_{p\geq 2} \mathfrak{Z}_p \subset \mathbf{R}.$$

This is equivalent to saying that the weight defines a graduation (see §5.1) on the algebra $\mathfrak{Z}$. A very special case of Conjecture 3, which is open, is $\mathfrak{Z}_2 \cap \mathfrak{Z}_3 = \{0\}$, which means that the number $\zeta(3)/\pi^2$ should be irrational.

For $p \geq 1$, we denote by d_p the dimension over $\mathbf{Q}$ of $\mathfrak{Z}_p$ (with $d_1 = 0$); we also set $d_0 = 1$. It is clear that $d_p \geq 1$ for $p \geq 2$, because $\zeta(p)$ is not zero. For $p \geq 1$ and $k \geq 1$, we denote by $d_{p,k}$ the dimension of $\mathcal{F}^k\mathfrak{Z}_p/\mathcal{F}^{k-1}\mathfrak{Z}_p$ with $d_{p,1} = 1$ for $p \geq 1$. We also set $d_{0,0} = 1$ and $d_{0,k} = 0$ for $k \geq 1$, $d_{p,0} = 0$ for $p \geq 1$. We have, for all $p \geq 0$,

$$d_p = \sum_{k\geq 0} d_{p,k} \tag{4}$$

and $d_{p,k} = 0$ for $k \geq p \geq 1$.

We have $d_2 = 1$, since $\mathfrak{Z}_2$ is spanned by $\zeta(2)$. The relation $\zeta(2,1) = \zeta(3)$, which is again due to Euler, shows that $d_3 = 1$. Also the relations, essentially going back to Euler,

$$\zeta(3,1) = \frac{1}{4}\zeta(4), \quad \zeta(2,2) = \frac{3}{4}\zeta(4), \quad \zeta(2,1,1) = \zeta(4) = \frac{2}{5}\zeta(2)^2,$$

show that $d_4 = 1$. These are the only values of d_p which are known. It is not yet proved that there exists a $p \geq 5$ with $d_p \geq 2$. The upper bound $\zeta(5) \leq 2$ follows from the fact that there are 6 independent linear relations among the 8 numbers

$$\zeta(5),\ \zeta(4,1),\ \zeta(3,2),\ \zeta(3,1,1),\ \zeta(2,3),\ \zeta(2,2,1),\ \zeta(2,1,2),\ \zeta(2,1,1,1),$$

and $\mathfrak{Z}_5$ is the $\mathbf{Q}$-vector subspace of $\mathbf{R}$ spanned by $\zeta(2,3)$ and $\zeta(3,2)$:

$$\begin{aligned}
\zeta(5) &= \frac{4}{5}\zeta(3,2) + \frac{6}{5}\zeta(2,3) = \zeta(2,1,1,1),\\
\zeta(4,1) &= -\frac{1}{5}\zeta(3,2) + \frac{1}{5}\zeta(2,3) = \zeta(3,1,1),\\
\zeta(2,2,1) &= \zeta(3,2)\\
\zeta(2,1,2) &= \zeta(2,3).
\end{aligned}$$

The dimension of $\mathfrak{Z}_5$ is 2 if $\zeta(2,3)/\zeta(3,2)$ is irrational (which is conjectured, but not yet proved), and 1 otherwise.

Similarly, the $\mathbf{Q}$–space $\mathfrak{Z}_6$ has dimension ≤ 2, as it is spanned by $\zeta(2,2,2)$ and $\zeta(3,3)$:

$$\begin{aligned}
\zeta(6) &= \frac{16}{3}\zeta(2,2,2) = \zeta(2,1,1,1,1),\\
\zeta(5,1) &= \frac{4}{3}\zeta(2,2,2) - \zeta(3,3) = \zeta(3,1,1,1),\\
\zeta(4,2) &= -\frac{16}{9}\zeta(2,2,2) + 2\zeta(3,3) = \zeta(2,2,1,1),\\
\zeta(4,1,1) &= \frac{7}{3}\zeta(2,2,2) - 2\zeta(3,3),\\
\zeta(3,2,1) &= -\frac{59}{9}\zeta(2,2,2) + 6\zeta(3,3),\\
\zeta(2,4) &= \frac{59}{9}\zeta(2,2,2) - 2\zeta(3,3) = \zeta(2,1,2,1),\\
\zeta(2,3,1) &= \frac{34}{9}\zeta(2,2,2) - 3\zeta(3,3) = \zeta(3,1,2),\\
\zeta(2,1,2,1) &= \zeta(3,3).
\end{aligned}$$

Here is Zagier's conjecture on the dimension d_p of the $\mathbf{Q}$-vector space $\mathfrak{Z}_p$.

Conjecture 5 (Zagier)**.** *For $p \ge 3$, we have*

$$d_p = d_{p-2} + d_{p-3}.$$

Since $d_0 = 1$, $d_1 = 0$ and $d_2 = 1$, this conjecture can be written

$$\sum_{p\ge 0} d_p X^p = \frac{1}{1 - X^2 - X^3}.$$

It has been proved independently by Goncharov and Terasoma that the numbers defined by the recurrence relation of Zagier's Conjecture 5 with initial values $d_0 = 1$, $d_1 = 0$ provide upper bounds for the actual dimension d_p. This shows that there are plenty of linear relations among the numbers $\zeta(\underline{s})$. For each p from 2 to 11, we display the number of tuples $\underline{s}$ of length p which is 2^{p-2}, the number d_p given by Zagier's Conjecture 5 and the difference which is (a lower bound for) the number of linear relations among these numbers.

p	2	3	4	5	6	7	8	9	10	11
2^{p-2}	1	2	4	8	16	32	64	128	256	512
d_p	1	1	1	2	2	3	4	5	7	9
$2^{p-2} - d_p$	0	1	3	6	14	29	60	123	249	503

Since d_p grows like a constant multiple of r^p, where $r = 1.324\,717\,957\,244\,7\ldots$ is the real root of $x^3 - x - 1$, the difference $2^{p-2} - d_p$ is asymptotic to 2^{p-2}.

According to Zagier's Conjecture 5, a basis for $\mathfrak{Z}_p$ should be given as follows:
$p = 2,\quad d_2 = 1,\qquad \zeta(2);$
$p = 3,\quad d_3 = 1,\qquad \zeta(3);$
$p = 4,\quad d_4 = 1,\qquad \zeta(2,2);$
$p = 5,\quad d_5 = 2,\qquad \zeta(2,3),\ \zeta(3,2);$
$p = 6,\quad d_6 = 2,\qquad \zeta(2,2,2),\ \zeta(3,3);$
$p = 7,\quad d_7 = 3,\qquad \zeta(2,2,3),\ \zeta(2,3,2),\ \zeta(3,2,2);$
$p = 8,\quad d_8 = 4,\qquad \zeta(2,2,2,3),\ \zeta(2,3,3),\ \zeta(3,2,3),\ \zeta(3,3,2);$
$p = 9,\quad d_8 = 5,\qquad \zeta(2,2,2,3),\ \zeta(2,2,3,2),\ \zeta(2,3,2,2),\ \zeta(3,2,2,2)\zeta(3,3,3).$

For these small values of p, the dimension $d_{p,k}$ of $\mathcal{F}^k\mathfrak{Z}_p/\mathcal{F}^{k-1}\mathfrak{Z}_p$ is conjecturally given by the number of elements in the box (p,k) of the next figure, where conjectural generators of $\mathcal{F}^k\mathfrak{Z}_p/\mathcal{F}^{k-1}\mathfrak{Z}_p$ should be given by the classes of the following MZV:

$k \backslash p$	2	3	4	5	6	7	8	9
1	$\zeta(2)$	$\zeta(3)$	$\zeta(4)$	$\zeta(5)$	$\zeta(6)$	$\zeta(7)$	$\zeta(8)$	$\zeta(9)$
2				$\zeta(4,1)$	$\zeta(5,1)$	$\zeta(6,1)$ $\zeta(5,2)$	$\zeta(7,1)$ $\zeta(6,2)$	$\zeta(8,1)$ $\zeta(7,2)$ $\zeta(6,3)$
3							$\zeta(6,1,1)$	$\zeta(6,2,1)$
d_p	1	1	1	2	2	3	4	5

The displayed elements should all be linearly independent over $\mathbf{Q}$. The numerical computations have been performed online thanks to the computer program EZface [1].

1.3 Known results

A conjecture by M. Hoffman is that a basis of $\mathfrak{Z}_p$ over $\mathbf{Q}$ is given by the numbers $\zeta(s_1,\ldots,s_k)$, $s_1 + \cdots + s_k = p$, where each s_i is either 2 or 3: the dimension agrees with Conjecture 5.

As a side product of his main recent results, F. Brown [7, 8] obtains:

Theorem 6. *The numbers $\zeta(s_1, \ldots, s_k)$, with $k \geq 1$ and $s_j \in \{2, 3\}$, span the* $\mathbf{Q}$*–space* $\mathfrak{Z}$ *of multizeta.*

Hence, Conjecture 5 is equivalent to saying that the numbers $\zeta(s_1, \ldots, s_k)$, with $k \geq 1$ and $s_j \in \{2, 3\}$, occurring in Theorem 6 are $\mathbf{Q}$–linearly independent.

One of the auxiliary result which was needed by F. Brown is a formula which he conjectured and which has been established by D. Zagier (see §2).

1.4 Occurrences of powers of π^2 in the set of generators

The space $\mathcal{F}^1\mathfrak{Z}_{2n}$ is spanned by π^{2n} over $\mathbf{Q}$. The next Proposition 7 shows that the numbers π^{2n} occur in the set of generators $\zeta(\underline{s})$ with $s_j \in \{2, 3\}$ given by Theorem 6, by taking all s_j equal to 2.

For $s \geq 2$ and $n \geq 1$, we use the notation $\{s\}_a$ for a string with n elements all equal to s, that is $\{s\}_n = (s_1, \ldots, s_n)$ with $s_1 = \cdots = s_n = s$.

Proposition 7. *For $s \geq 2$,*

$$\sum_{n \geq 0} \zeta\big(\{s\}_n\big) x^n = \prod_{j \geq 1} \left(1 + \frac{x}{j^s}\right) = \exp\left(\sum_{k \geq 1} \frac{(-1)^{k-1} x^k \zeta(sk)}{k}\right).$$

The proof will involve the infinite product

$$F_s(x) = \prod_{j \geq 1} \left(1 + \frac{x}{j^s}\right).$$

Proof. Expanding $F_s(x)$ as a series:

$$\begin{aligned} F_s(x) &= 1 + x \sum_{j \geq 1} \frac{1}{j^s} + x^2 \sum_{j_1 > j_2 \geq 1} \frac{1}{(j_1 j_2)^s} + \cdots \\ &= 1 + x\zeta(s) + x^2 \zeta(s, s) + \cdots \end{aligned}$$

yields the first equality in Proposition 7. For the second one, consider the logarithmic derivative of $F_s(x)$:

$$\begin{aligned} \frac{F_s'(x)}{F_s(x)} &= \sum_{j \geq 1} \frac{1}{j^s + x} = \sum_{j \geq 1} \frac{1}{j^s} \sum_{k \geq 1} (-1)^{k-1} \frac{x^{k-1}}{j^{s(k-1)}} \\ &= \sum_{k \geq 1} (-1)^{k-1} x^{k-1} \sum_{j \geq 1} \frac{1}{j^{sk}} = \sum_{k \geq 1} (-1)^{k-1} x^{k-1} \zeta(sk). \end{aligned}$$

Since $F_s(0) = \zeta\big(\{s\}_0\big) = 1$, Proposition 7 follows by integration. □

Lemma 8. *We have*

$$F_2(-z^2) = \frac{\sin(\pi z)}{\pi z}\cdot$$

Proof. Lemma 8 follows from the product expansion of the sine function (see for instance [18], Chap. 7, § 4.1, (27))

$$\sin z = z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2\pi^2}\right).$$

□

From Proposition 7 and Lemma 8 we deduce:

Corollary 9. *For any $n \geq 0$,*

$$\zeta(\{2\}_n) = \frac{\pi^{2n}}{(2n+1)!}\cdot$$

Proof. From the Taylor expansion of the sine function

$$\sin z = \sum_{k\geq 0} (-1)^k \frac{z^{2k+1}}{(2k+1)!}$$

and from Lemma 8, we infer

$$F_2(-z^2) = \frac{\sin(\pi z)}{\pi z} = \sum_{k\geq 0} (-1)^k \frac{\pi^{2k} z^{2k}}{(2k+1)!}\cdot$$

Corollary 9 now follows from Proposition 7. □

2 Zagier's contribution to Brown's proof

We explain the strategy of Zagier [21] for proving the relation which was needed and conjectured by F. Brown [7, 8] concerning the numbers

$$H(a,b) := \zeta(\{2\}_b 3 \{2\}_a)$$

for $a \geq 0$ and $b \geq 0$. In particular $H(0,0) = \zeta(3)$. Beware that the normalization used by F. Brown and by D. Zagier [7, 8, 21] is the opposite of ours; this is why a and b are in this reverse order here.

Another reference to this topic is the course [2] on MZV by J. Borwein and W. Zudilin at the University of Newcastle in 2011–2012.

Set also (cf. Corollary 9)

$$H(n) = \zeta(\{2\}_n) = \frac{\pi^{2n}}{(2n+1)!}$$

for $n \geq 0$, with $H(0) = 1$.

Consider the alphabet $\{2, 3\}$; give to the letter 2 the weight 2 and to the letter 3 the weight 3, so that the word $2^b \cdot 3 \cdot 2^a$ has weight $2a+2b+3$, while the word 2^n has weight $2n$. Give also the weight ℓ to $\zeta(\ell)$ for $\ell \geq 2$ (this is an abuse of language, since we do not know whether the weight defines a graduation – see Conjecture 3). Looking at homogeneous relations, one considers on the one side the numbers $H(a, b)$ and on the other side the numbers $\zeta(\ell)H(m)$ with $\ell \geq 2$ and $m \geq 0$, restricted to the relation $2a + 2b + 3 = \ell + 2m$. Notice that this implies that the number $\ell = 2a + 2b + 3 - 2m$ is odd.

Theorem 10 (Zagier, 2011). *Let a and b be non–negative integers. Set $k = 2a + 2b + 3$. Then there exist $a + b + 1$ rational integers $c_{m,r,a,b}$ with $m \geq 0$, $r \geq 1$, $m + r = a + b + 1$, such that*

$$H(a, b) = \sum_{m+r=a+b+1} c_{m,r,a,b} H(m)\zeta(2r + 1).$$

Conversely, given two integers r and m with $r \geq 1$ and $m \geq 0$, there exist $m+r$ rational numbers $c'_{m,r,a,b}$ with $a \geq 0$, $b \geq 0$, $a + b = m + r - 1$, such that

$$H(m)\zeta(2r + 1) = \sum_{a+b=m+r-1} c'_{m,r,a,b} H(a, b).$$

The integers $c_{m,r,a,b}$ are explicitly given:

$$c_{m,r,a,b} = 2(-1)^r \left[\binom{2r}{2a+2} - \left(1 - \frac{1}{2^{2r}}\right)\binom{2r}{2b+1} . \right]$$

It follows that the square matrices $\bigl(c_{m,r,a,b}\bigr)$ and $\bigl(c'_{m,r,a,b}\bigr)$, of size $a + b + 1 = m + r$, are inverse matrices.

We only give the sketch of proof of the first part of Theorem 10. Consider the generating series

$$F(x, y) = \sum_{a \geq 0} \sum_{b \geq 0} (-1)^{a+b+1} H(a, b) x^{2a+2} y^{2b+1}$$

and

$$\widehat{F}(x, y) = \sum_{a \geq 0} \sum_{b \geq 0} (-1)^{a+b+1} \widehat{H}(a, b) x^{2a+2} y^{2b+1},$$

where

$$\widehat{H}(a, b) = \sum_{m=0}^{a+b} c_{m,a,b} H(m)\zeta(k - 2m).$$

The first step relates $F(x, y)$ to a hypergeometric series ${}_3F_2$, namely $F(x, y)$ is the product of $(1/\pi)\sin(\pi y)$ by the z–derivative at $z = 0$ of the function

$${}_3F_2\left(\left. \begin{matrix} x,\ -x,\ z \\ 1+y,\ 1-y \end{matrix} \right| 1 \right).$$

The second step relates $\widehat{F}(x,y)$ to the digamma function $\psi(x) = \Gamma'(x)/\Gamma(x)$ (logarithmic derivative of Γ), namely $\widehat{F}(x,y)$ is a linear combination of fourteen functions of the form

$$\psi\left(1+\frac{u}{2}\right)\frac{\sin(\pi v)}{2\pi} \quad \text{with} \quad u \in \{\pm x \pm y,\ \pm 2x \pm 2y,\ \pm 2y\} \quad \text{and} \quad v \in \{x,y\}.$$

The third step is the proof that F and $\widehat{F}$ are both entire function on $\mathbf{C}\times\mathbf{C}$, they are bounded by a constant multiple of $e^{\pi X}\log X$, when $X = \max\{|x|,|y|\}$ tends to infinity, and also by a constant multiple of $e^{\pi|\Im(y)|}$, when $|y|$ tends to infinity while $x \in \mathbf{C}$ is fixed.

The fourth step is about the diagonal: for $z \in \mathbf{C}$, we have

$$F(z,z) = \widehat{F}(z,z).$$

Several equivalent explicit formula for this function are given. Let

$$A(z) := -\frac{\pi}{\sin(\pi z)}F(z,z).$$

Then

$$A(z) = \sum_{r=1}^{\infty}\zeta(2r+1)z^{2r} = \sum_{n=1}^{\infty}\frac{z^2}{n(n^2-z^2)}$$

and

$$A(z) = \frac{1}{2}\sum_{n=1}^{\infty}\left(\frac{1}{n-z}+\frac{1}{n+z}-\frac{2}{n}\right) = \psi(1) - \frac{1}{2}\big(\psi(1+z)+\psi(1-z)\big).$$

The fifth step shows that $F(n,y)$ and $\widehat{F}(n,y)$ are equal when $n \in \mathbf{N}$ and $y \in \mathbf{C}$, an explicit formula is given.
The sixth step shows that $F(x,k)$ and $\widehat{F}(x,k)$ are equal when $k \in \mathbf{N}$ and $x \in \mathbf{C}$, an explicit formula is given.
Now comes the conclusion, which rests on the next lemma:

Lemma 11. *An entire function $f : \mathbf{C} \to \mathbf{C}$ that vanishes at all rational integers and satisfies*

$$f(z) = O\left(e^{\pi|\Im(z)|}\right)$$

is a constant multiple of $\sin(\pi z)$.

A proof of this lemma, using a Theorem of Phragmén–Lindelöf, is given by Zagier (Lemma 2 in [21]), but he also notices that other references have been given subsequently to him, in particular by F. Gramain who pointed out that this lemma is known since the work of Pólya and Valiron; see [6]

3 Lyndon words: conjectural transcendence basis

D. Broadhurst considered the question of finding a transcendence basis for the algebra of MZV, he suggested that one should consider Lyndon words. We first give the definitions.

Here, we use the alphabet alphabet $A = \{a, b\}$. The set of words on A is denoted by A^* (see §6.2). The elements of A^* can be written

$$a^{n_1}b^{m_1}\cdots a^{n_k}b^{m_k}$$

with $k \geq 0$, $n_1 \geq 0$, $m_k \geq 0$ and $m_i \geq 1$ for $1 \leq i < k$, and with $n_i \geq 1$ for $2 \leq i \leq k$. We endowed A^* with the lexicographic order with $a < b$. A *Lyndon word* is a non–empty word $w \in A^*$ such that, for each decomposition $w = uv$ with $u \neq e$ and $v \neq e$, the inequality $w < v$ holds. Denote by $\mathcal{L}$ the set of Lyndon words. Examples of Lyndon words are a, b, ab^k $(k \geq 0)$, $a^\ell b$ $(\ell \geq 0)$, a^2b^2, a^2bab. Let us check, for instance, that a^2bab is a Lyndon word: this follows from the observation that a^2bab is smaller than any of

$$abab,\ bab,\ ab,\ b.$$

But a^2ba^2b is not a Lyndon word, since $a^2ba^2b > a^2b$.

Any Lyndon word other than b starts with a and any Lyndon word other than a ends with b.

Here are the 21 Lyndon words on the alphabet $\{a, b\}$ with weight ≤ 15, when a has weight 2 and b weight 3:

$$\begin{aligned} a &< a^6b < a^5b < a^4b < a^4b^2 < a^3b < a^3bab < a^3b^2 < a^3b^3 < a^2b < a^2bab \\ &< a^2bab^2 < a^2b^2 < a^2b^2ab < a^2b^3 < ab < abab^2 < ab^2 < ab^3 < ab^4 < b. \end{aligned}$$

We list them according to their weight $p = 2, \ldots, 15$, we display their number $N(p)$ and the corresponding multiple zeta values, where the word a^n is replaced by the tuple $\{2\}_n$ and b^n by $\{3\}_n$:

$p = 2$; $N(2) = 1$; a; $\zeta(2)$
$p = 3$; $N(3) = 1$; b; $\zeta(3)$
$p = 4$; $N(4) = 0$
$p = 5$; $N(5) = 1$; ab; $\zeta(2, 3)$
$p = 6$; $N(6) = 0$
$p = 7$; $N(7) = 1$; a^2b; $\zeta(2, 2, 3)$
$p = 8$; $N(8) = 1$; ab^2; $\zeta(2, 3, 3)$
$p = 9$; $N(9) = 1$; a^3b; $\zeta(2, 2, 2, 3)$
$p = 10$; $N(10) = 1$; a^2b^2; $\zeta(2, 2, 3, 3)$
$p = 11$; $N(11) = 2$; a^4b, ab^3; $\zeta(2, 2, 2, 2, 3)$, $\zeta(2, 3, 3, 3)$
$p = 12$; $N(12) = 2$; a^3b^2, a^2bab; $\zeta(2, 2, 2, 3, 3)$, $\zeta(2, 2, 3, 2, 3)$
$p = 13$; $N(13) = 3$; a^5b, a^2b^3, $abab^2$;
$\zeta(2, 2, 2, 2, 2, 3)$, $\zeta(2, 2, 3, 3, 3)$, $\zeta(2, 3, 2, 3, 3)$

$p = 14$; $N(14) = 3$; a^4b^2, a^3bab, ab^4;
$\zeta(2,2,2,2,3,3)$, $\zeta(2,2,2,3,2,3)$, $\zeta(2,3,3,3,3)$
$p = 15$; $N(15) = 4$; a^6b, a^3b^3, a^2bab^2, a^2b^2ab; ;
$\zeta(2,2,2,2,2,2,3)$, $\zeta(2,2,2,3,3,3)$, $\zeta(2,2,3,2,3,3)$, $\zeta(2,2,3,3,2,3)$.

Conjecture 12. *The set of multiple zeta values $\zeta(s_1,\dots,s_k)$, with $k \ge 1$ and $s_j \in \{2,3\}$ for $1 \le j \le k$, such that $s_1s_2\cdots s_k$ is a Lyndon word on the alphabet $\{2,3\}$, gives a transcendence basis of $\mathfrak{Z}$.*

The number $N(p)$ of elements of weight p in a transcendence basis of $\mathfrak{Z}$ should not depend on the choice of the transcendence basis, and it should be the number of Lyndon words of weight p on the alphabet $\{2,3\}$.

In the next section (Example 20), we will see that, for $p \ge 1$, we have

$$N(p) = \frac{1}{p}\sum_{\ell|p}\mu(p/\ell)P_\ell,$$

where $(P_\ell)_{\ell\ge1}$ is the linear recurrence sequence of integers defined by

$$P_\ell = P_{\ell-2} + P_{\ell-3} \quad \text{for } \ell \ge 4$$

with the initial conditions

$$P_1 = 0, \quad P_2 = 2, \quad P_3 = 3.$$

The sequence $N_0, N_1, \dots$ is

$$0,1,1,0,1,0,1,1,1,1,2,2,3,3,4,5,7,8,11,13,17,21,\dots$$

In Sloane [4] http://oeis.org/A113788, it is referred to as *Number of irreducible multiple zeta values at weight n.*

If one forgets about the weight of the words, one may list the Lyndon words according to the number of letters (which correspond to the length for MZV with $s_j \in \{2,3\}$), which yields a partial order on the words on the alphabet $\{a,b\}$, where there are 2^k words with k letters. Here are the Lyndon words with k letters on the alphabet $\{a,b\}$ for the first values of k, with their numbers L_k:
$k = 1$, $L_1 = 2$; a, b.
$k = 2$, $L_2 = 1$; ab.
$k = 3$, $L_3 = 2$; a^2b, ab^2.
$k = 4$, $L_4 = 3$, a^3b, a^2b^2, ab^3.
$k = 5$, $L_5 = 6$, a^4b, a^3b^2, a^2bab, a^2b^3, $abab^2$, ab^4.
$k = 6$, $L_6 = 9$, a^5b, a^4b^2, a^3bab, a^3b^3, a^2ba^2b, a^2bab^2, a^2b^2ab, $abab^3$, ab^5

The sequence $(L_k)_{k\ge1}$ starts with

$$2,1,2,3,6,9,18,30,56,99,186,335,630,1161,2182,4080,7710,14532,27594,\dots$$

See the reference [4] `A001037` to the `On-Line Encyclopedia of Integer Sequences` by Sloane.

The sequence $(L_1 + \cdots + L_n)_{n \geq 1}$ starts with

$2, 3, 5, 8, 14, 23, 41, 71, 127, 226, 412, 747, 1377, 2538, 4720, 8800, 16510, 31042, \ldots$

Also, $L_1 + \cdots + L_n$ is the number of irreducible polynomials over $\mathbf{F}_2$ of degree at most n — see [4] `A062692`.

4 Hilbert–Poincaré series

We will work with commutative algebras, namely polynomial algebras in variables having each a weight. A conjectural (recall the open problem 2) example is the algebra $\mathfrak{Z}$ of MZV which is a subalgebra of the real numbers. This algebra $\mathfrak{Z}$ is the image by specialization of an algebra of polynomials in infinitely many variables, with one variable of weight 2 (corresponding to $\zeta(2)$), one of weight 3 (corresponding to $\zeta(3)$), one of weight 5 (corresponding to $\zeta(3,2)$) and so on with a suitable number of variables for each odd weight p. According to Conjecture 12, the number of variables of weight p should be the number $N(p)$ of Lyndon words on the alphabet $\{2,3\}$ with 2 having weight 2 and 3 weight 3. In general, we will consider countably many variables with $N(p)$ variables of weight p for $p \geq 0$, with $N(0) = 0$. But later we will also consider non–commutative variables; for instance, the free algebra on the words $\{2,3\}$ will play a role.

5 Graduated algebras and Hilbert–Poincaré series

5.1 Graduations

We introduce basic definitions from algebra. A *graduation* on a ring A is a decomposition into a direct sum of additive subgroups

$$A = \bigoplus_{k \geq 0} A_k,$$

such that the multiplication $A \times A \to A$ which maps (a,b) onto the product ab maps $A_k \times A_h$ into A_{k+h} for all pairs (k,h) of non–negative integers. For us here, it will be sufficient to take for indices the non–negative integers, but we could more generally take a commutative additive monoid (see [17] Chap. X § 5). The elements in A_k are *homogeneous of weight (or degree) k*. Notice that A_0 is a subring of A and that each A_k is a A_0–module[1].

[1]According to this definition, 0 is homogeneous of weight k for all $k \geq 0$

Given a graduated ring A, a *graduation* on a A–module E is a decomposition into a direct sum of additive subgroups

$$E = \bigoplus_{k\geq 0} E_k,$$

such that $A_k E_n \subset E_{k+n}$. In particular each E_n is a A_0–module. The elements of E_n are *homogeneous of weight (or degree) n.*

A *graduated K–algebra* is a K–algebra A with a graduation as a ring $A = \bigoplus_{k\geq 0} A_k$ such that $KA_k \subset A_k$ for all $k \geq 0$ and $A_0 = K$, (see [17] Chap. XVI, § 6). If the dimension d_k of each A_k as a K–vector space is finite with $d_0 = 1$, the *Hilbert–Poincaré series* of the graduated algebra A is

$$\mathcal{H}_A(t) = \sum_{p\geq 0} d_p t^p.$$

If the K–algebra A is the tensor product $A' \otimes A''$ of two graded algebras A' and A'' over the field K, then A is graded with the generators of A_p as K–vector space being the elements $x' \otimes x''$, where x' runs over the generators of the homogeneous part A'_k of A' and where x'' runs over the generators of the homogeneous part A''_ℓ of A'', with $k + \ell = p$. Hence, the dimensions d_p, d'_k, d''_ℓ of the homogeneous subspaces of A, A' and A'' satisfy

$$d_p = \sum_{k+\ell=p} d'_k d''_\ell,$$

which means that the Hilbert–Poincaré series of A is the product of the Hilbert–Poincaré series of A' and A'':

$$\mathcal{H}_{A'\otimes A''}(t) = \mathcal{H}_{A'}(t)\mathcal{H}_{A''}(t).$$

5.2 Commutative polynomials algebras

Let

$$\bigl(N(1), N(2), \ldots, N(p), \ldots\bigr)$$

be a sequence of non–negative integers and let A denote the commutative K–algebra of polynomials with coefficients in K in the variables Z_{np} ($p \geq 1$, $1 \leq n \leq N(p)$). We endow the K-algebra A with the graduation for which each Z_{np} is homogeneous of weight p. We denote by d_p the dimension of the homogeneous space A_p over K.

Lemma 13. *The Hilbert–Poincaré series of A is*

$$\mathcal{H}_A(t) = \prod_{p\geq 1} \frac{1}{(1-t^p)^{N(p)}}.$$

Proof. For $p \geq 1$, the K–vector space A_p of homogeneous elements of weight p has a basis consisting of monomials

$$\prod_{k=1}^{\infty} \prod_{n=1}^{N(k)} Z_{nk}^{h_{nk}},$$

where $\underline{h} = (h_{nk})_{\substack{k \geq 1 \\ 1 \leq n \leq N(k)}}$ runs over the set of tuples of non–negative integers satisfying

$$\sum_{k=1}^{\infty} \sum_{n=1}^{N(k)} k h_{nk} = p. \tag{14}$$

Notice that these tuples $\underline{h}$ have a *support*

$$\{(n,k)\ ;\ k \geq 1,\ 1 \leq n \leq N(k),\ h_{nk} \neq 0\}$$

which is finite, since $h_{n,k} = 0$ for $k > p$. The dimension d_p of the K–vector space A_p is the number of these tuples $\underline{h}$ (with $d_0 = 1$), and by definition we have

$$\mathcal{H}_A(t) = \sum_{p \geq 0} d_p t^p.$$

In the identity

$$\frac{1}{(1-z)^N} = \sum_{h_1 \geq 0} \cdots \sum_{h_N \geq 0} \prod_{n=1}^{N} z^{h_n}$$

we replace z by t^k and N by $N(k)$. We deduce

$$\prod_{k \geq 1} \frac{1}{(1-t^k)^{N(k)}} = \sum_{\underline{h}} \prod_{k=1}^{\infty} \prod_{n=1}^{N(k)} t^{k h_{nk}}.$$

The coefficient of t^p in the right hand side is the number of tuples $\underline{h} = (h_{nk})_{\substack{k \geq 1 \\ 1 \leq n \leq N(k)}}$ with $h_{nk} \geq 0$ satisfying (14); hence, it is nothing else than d_p. □

5.3 Examples

Example 15. In Lemma 13, take $N(p) = 0$ for $p \geq 2$ and write N instead of $N(1)$. Then A is the ring of polynomials $K[Z_1, \ldots, Z_N]$ with the standard graduation of the total degree (each variable Z_i, $i = 1, \ldots, N$, has weight 1). The Hilbert–Poincaré series is

$$\frac{1}{(1-t)^N} = \sum_{\ell \geq 0} \binom{N+\ell-1}{\ell} t^\ell.$$

If each variable Z_i has a weight other than 1 but all the same, say p, it suffices to replace t by t^p. For instance, the Hilbert–Poincaré series of the algebra of polynomials $K[Z]$ in one variable Z having weight 2 is $(1-t^2)^{-1}$.

Example 16. *More generally, if there are only finitely many variables, which means that there exists an integer $p_0 \geq 1$ such that $N(j) = 0$ for $j > p_0$, the same proof yields*

$$d_\ell = \prod_{\ell_1+2\ell_2+\cdots+j_0\ell_{k_0}=\ell} \prod_{j=1}^{k_0} \binom{N(j)+\ell_j-1}{\ell_j}.$$

Example 17. Denote by μ the Möbius function (see [15]— § 16.3):

$$\begin{cases} \mu(1) = 1, \\ \mu(p_1\cdots p_r) = (-1)^r \text{ if } p_1,\dots,p_r \text{ are distinct prime numbers distincts,} \\ \mu(n) = 0 \text{ if } n \text{ has a square factor} > 1. \end{cases}$$

Given a positive integer c, under the assumptions of Lemma 13*, the following conditions are equivalent:*

(i) *The Hilbert–Poincaré series of A is*

$$\mathcal{H}_A(t) = \frac{1}{1-ct}.$$

(ii) *For $p \geq 0$, we have $d_p = c^p$.*

(iii) *For $p \geq 1$, we have*

$$c^p = \sum_{n|p} nN(n) \quad \text{for all } p \geq 1.$$

(iv) *For $k \geq 1$, we have*

$$N(k) = \frac{1}{k}\sum_{n|k} \mu(k/n)c^n.$$

Proof. The equivalence between (i) and (ii) follows from the definition of $\mathcal{H}_A$ and the power series expansion

$$\frac{1}{1-ct} = \sum_{p\geq 0} c^p t^p.$$

The equivalence between (iii) and (iv) follows from Möbius inversion formula (see [17] Chap. II Ex. 12.c and Chap. V, Ex. 21; [15] § 16.4).

It remains to check the equivalence between (i) and (iii).The constant term of each of the developments of

$$\frac{1}{1-ct} \quad \text{and} \quad \prod_{k\geq 1} \frac{1}{(1-t^k)^{N(k)}}$$

into power series is 1; hence, the two series are the same if and only if their logarithmic derivatives are the same. The logarithmic derivative of $1/(1-ct)$ is

$$\frac{c}{1-ct} = \sum_{p\geq 1} c^p t^{p-1}.$$

The logarithmic derivative of $\prod_{k\geq 1}(1-t^k)^{-N(k)}$ is

$$\sum_{k\geq 1} \frac{kN(k)t^{k-1}}{1-t^k} = \sum_{p\geq 1} \left(\sum_{n|p} nN(n) \right) t^{p-1}. \tag{18}$$

□

Example 19. Let a and c be two positive integers. Define two sequences of integers $(\delta_p)_{p\geq 1}$ and $(P_\ell)_{\ell\geq 1}$ by

$$\begin{cases} \delta_p = 0 & \text{if } a \text{ does not divide } p, \\ \delta_p = c^{p/a} & \text{if } a \text{ divides } p \end{cases}$$

and

$$\begin{cases} P_\ell = 0 & a \text{ does not divide } \ell, \\ P_\ell = ac^{\ell/a} & \text{if } a \text{ divides } \ell. \end{cases}$$

Under the hypotheses of Lemma 13, the following properties are equivalent:

(i) The Hilbert–Poincaré series of A is

$$\mathcal{H}_A(t) = \frac{1}{1-ct^a}.$$

(ii) For any $p \geq 1$, we have $d_p = \delta_p$.

(iii) For any $\ell \geq 1$, we have

$$\sum_{n|\ell} nN(n) = P_\ell.$$

(iv) For any $k \geq 1$, we have

$$N(k) = \frac{1}{k} \sum_{\ell|k} \mu(k/\ell) P_\ell.$$

Proof. The definition of the numbers δ_p means

$$\frac{1}{1-ct^a} = \sum_{p\geq 0} \delta_p t^p,$$

while the definition of P_ℓ can be written

$$\sum_{\ell\geq 1} P_\ell t^{\ell-1} = \frac{cat^{a-1}}{1-ct^a},$$

where the right hand side is the logarithmic derivative of $1/(1-ct^a)$. Recall that the logarithmic derivative of $\prod_{k\geq 1}(1-t^k)^{-N(k)}$ is given by (18). This completes the proof. $\square$

Example 20. Let a and b be two positive integers with $a < b$. Define two sequences of integers $(\delta_p)_{p\geq 1}$ and $(P_\ell)_{\ell\geq 1}$ by the induction formulae

$$\delta_p = \delta_{p-a} + \delta_{p-b} \quad \text{for } p \geq b+1,$$

with initial conditions

$$\begin{cases} \delta_0 = 1, & \\ \delta_p = 0 & \text{for } 1 \leq p \leq b-1 \text{ if } a \text{ does not divide } p, \\ \delta_p = 1 & \text{for } a \leq p \leq b-1 \text{ if } a \text{ divides } p, \\ \delta_b = 1 & \text{if } a \text{ does not divide } b, \\ \delta_b = 2 & \text{if } a \text{ divides } b \end{cases}$$

and

$$P_\ell = P_{\ell-a} + P_{\ell-b} \quad \text{for } \ell \geq b+1,$$

with initial conditions

$$\begin{cases} P_\ell = 0 & \text{for } 1 \leq \ell < b \text{ if } a \text{ does not divides } \ell, \\ P_\ell = a & \text{for } a \leq \ell < b \text{ if } a \text{ divides } \ell, \\ P_b = b & \text{if } a \text{ does not divides } b, \\ P_b = a+b & \text{if } a \text{ divides } b. \end{cases}$$

Under the hypotheses of Lemma 13*, the following properties are equivalent:*

(i) The Hilbert–Poincaré series of A is

$$\mathcal{H}_A(t) = \frac{1}{1-t^a-t^b}\cdot$$

(ii) For any $p \geq 1$, we have $d_p = \delta_p$.

(iii) For any $\ell \geq 1$, we have

$$\sum_{n|\ell} nN(n) = P_\ell.$$

(iv) For any $k \geq 1$, we have

$$N(k) = \frac{1}{k} \sum_{\ell | k} \mu(k/\ell) P_\ell.$$

Proof. Condition (ii) means that the sequence $(d_p)_{p\geq 0}$ satisfies

$$(1 - t^a - t^b) \sum_{p\geq 0} d_p t^p = 1.$$

The equivalence between (i) and (ii) follows from the definition of d_p in condition (ii): the series

$$\mathcal{H}_A(t) = \sum_{p\geq 0} d_p t^p$$

satisfies

$$(1 - t^a - t^b)\mathcal{H}_A(t) = 1$$

if and only if the sequence $(d_p)_{p\geq 0}$ is the same as $(\delta_p)_{p\geq 0}$. The definition of the sequence $(P_\ell)_{\ell\geq 0}$ can be written

$$(1 - t^a - t^b) \sum_{\ell\geq 1} P_\ell t^{\ell-1} = at^{a-1} + bt^{b-1}.$$

The equivalence between (iii) and (iv) follows from Möbius inversion formula. It remains to check that conditions (i) and (iii) are equivalent. The logarithmic derivative of $1/(1 - t^a - t^b)$ is

$$\frac{at^{a-1} + bt^{b-1}}{1 - t^a - t^b} = \sum_{\ell\geq 1} P_\ell t^{\ell-1},$$

while the logarithmic derivative of $\prod_{k\geq 1} 1/(1 - t^k)^{N(k)}$ is given by (18). This completes the proof. $\square$

A first special case of example 20 is with $a = 1$ and $b = 2$: the sequence $(d_p)_{p\geq 1} = (1, 2, 3, 5 \dots)$ is the Fibonacci sequence $(F_n)_{n\geq 0}$ shifted by 1: $d_p = F_{p+1}$ for $p \geq 1$, while the sequence $(P_\ell)_{\ell\geq 1}$ is the sequence of Lucas numbers

$$1, 3, 4, 7, 11, 18, 29, 47, 76, 123, 199, 322, 521, 843, 1364, \dots$$

See the On-Line Encyclopedia of Integer Sequences by Sloane [4], `A000045` for the Fibonacci sequence and `A000032` for the Lucas sequence.

For the application to MZV, we are interested with the special case where $a = 2$, $b = 3$ in example 20. In this case the recurrence formula for the sequence $(P_\ell)_{\ell\geq 1}$ is $P_\ell = P_{\ell-2} + P_{\ell-3}$ and the initial conditions are $P_1 = 0$, $P_2 = 2$, $P_3 = 3$, so that, if we set $P_0 = 3$, then the sequence $(P_\ell)_{\ell\geq 0}$ is

$$3, 0, 2, 3, 2, 5, 5, 7, 10, 12, 17, 22, 29, 39, 51, 68, 90, 119, \dots$$

This is the so–called *Perrin sequence* or *Ondrej Such sequence* (see [4] `A001608`), defined by

$$P_\ell = P_{\ell-2} + P_{\ell-3} \quad \text{for } \ell \geq 3,$$

with the initial conditions

$$P_0 = 3, \quad P_1 = 0, \quad P_2 = 2.$$

The sequence $N(p)$ of the number of Lyndon words of weight p on the alphabet $\{2, 3\}$ satisfies, for $p \geq 1$,

$$N(p) = \frac{1}{p} \sum_{\ell | p} \mu(p/\ell) P_\ell.$$

The generating function of the sequence $(P_\ell)_{\ell \geq 1}$ is

$$\sum_{\ell \geq 1} P_\ell t^{\ell-1} = \frac{3 - t^2}{1 - t^2 - t^3}.$$

For $\ell > 9$, P_ℓ is the nearest integer to r^ℓ, with $r = 1.324\,717\,957\,244\,7\ldots$ the real root of $x^3 - x - 1$ (see [4] `A060006`), which has been also called the *silver number*, also the *plastic number*: this is the smallest *Pisot-Vijayaraghavan number*.

Example 21 (Words on the alphabet $\{f_3, f_5, \ldots, f_{2n+1} \ldots\}$)**.** Consider the alphabet $\{f_3, f_5, \ldots, f_{2n+1} \ldots\}$ with countably many letters, one for each odd weight. The free algebra on this alphabet (see §6.2) is the so–called *concatenation algebra* $\mathcal{C} := \mathbf{Q}\langle f_3, f_5, \ldots, f_{2n+1} \ldots\rangle$.

We get a word of weight p by concatenating a word of weight $p-(2k+1)$ with f_{2k+1}; in other terms, starting with a word of weight q having the last letter say f_{2k+1}, the prefix obtained by removing the last letter has weight $q - 2k - 1$. Hence, the number of words with weight p satisfies

$$d_p = d_{p-3} + d_{p-5} + \cdots$$

(a finite sum for each p,) with $d_0 = 1$ (the empty word), $d_1 = d_2 = 0$, $d_3 = 1$. The Hilbert–Poincaré series of $\mathcal{C}$

$$\mathcal{H}_\mathcal{C}(t) := \sum_{p \geq 0} d_p t^p$$

satisfies

$$\mathcal{H}_\mathcal{C}(t) = 1 - t^3 \mathcal{H}_\mathcal{C}(T) - t^5 \mathcal{H}_\mathcal{C}(t) - \cdots$$

Since

$$(1-t^2)(1-t^3-t^5-t^7-\cdots) = 1-t^3-t^5-t^7-\cdots-t^2+t^5+t^7+\cdots = 1-t^2-t^3$$

(telescoping series), we deduce

$$\mathcal{H}_\mathcal{C}(t) = \frac{1 - t^2}{1 - t^2 - t^3}.$$

Recall (example 15) that the Hilbert–Poincaré series of the commutative polynomial algebra $\mathbf{Q}[f_2]$ with f_2 a single variable of weight 2 is $1/(1-t^2)$. The algebra $\mathcal{C} \otimes_{\mathbf{Q}} \mathbf{Q}[f_2]$, which plays an important role in the theory of mixed Tate motives, can be viewed either as the free algebra on the alphabet $\{f_3, f_5, \dots, f_{2n+1} \dots\}$ over the commutative ring $\mathbf{Q}[f_2]$, or as the algebra $\mathcal{C}[f_2]$ of polynomials in the single variable f_2 with coefficients in $\mathcal{C}$. The Hilbert–Poincaré series of this algebra is the product

$$\mathcal{H}_{\mathcal{C}[f_2]}(t) = \frac{1-t^2}{1-t^2-t^3} \cdot \frac{1}{1-t^2} = \frac{1}{1-t^2-t^3},$$

which is conjectured to be also the Hilbert–Poincaré series of the algebra $\mathfrak{Z}$.

5.4 Filtrations

A *filtration* on a A-module E is an increasing or decreasing sequence of sub–A–modules

$$\{0\} = E_0 \subset E_1 \subset \cdots \subset E_n \subset \cdots$$

or

$$E = E_0 \supset E_1 \supset \cdots \supset E_n \supset \cdots$$

Sometimes one writes $\mathcal{F}^n(E)$ in place of E_n. For instance, if φ is an endomorphism of a A–module E, the sequence of kernels of the iterates

$$\{0\} \subset \ker \varphi \subset \ker \varphi^2 \subset \cdots \subset \ker \varphi^n \subset \cdots$$

is an increasing filtration on E, while the images of the iterates

$$E \supset \mathrm{Im}\varphi \supset \mathrm{Im}\varphi^2 \subset \cdots \supset \mathrm{Im}\varphi^n \supset \cdots$$

is a decreasing filtration on E.

A *filtration* on a ring A is an increasing or decreasing sequence of (abelian) additive subgoups

$$A = A_0 \supset A_1 \supset \cdots \supset A_n \supset \cdots$$

or

$$\{0\} = A_0 \subset A_1 \subset \cdots \subset A_n \subset \cdots$$

such that $A_n A_m \subset A_{n+m}$. In this case A_0 is a subring of A and each A_n is a A_0–module.

As an example, if $\mathfrak{A}$ is an ideal of A, a filtration on the ring A is given by the powers of $\mathfrak{A}$:

$$A = \mathfrak{A}^0 \supset \mathfrak{A}^1 \supset \cdots \supset \mathfrak{A}^n \supset \cdots$$

The *first graduated ring* associated with this filtration is

$$\bigoplus_{n \geq 0} \mathfrak{A}^n,$$

and the *second graduated ring* is

$$\bigoplus_{n\geq 0}\mathfrak{A}^n/\mathfrak{A}^{n+1}.$$

For instance, if $\mathfrak{A}$ is a *proper* ideal (that means distinct from $\{0\}$ and from A) and is principal, then the first graduated ring is isomorphic to the ring of polynomials $A[t]$ and the second to $(A/\mathfrak{A})[t]$.

We come back to MZV. We have seen that the length k defines a filtration on the algebra $\mathfrak{Z}$ of multiple zeta values. Recall that $\mathcal{F}^k\mathfrak{Z}_p$ denotes the $\mathbf{Q}$-vector subspace of $\mathbf{R}$ spanned by the $\zeta(\underline{s})$ with $\underline{s}$ of weight p and length $\leq k$ and that $d_{p,k}$ is the dimension of $\mathcal{F}^k\mathfrak{Z}_p/\mathcal{F}^{k-1}\mathfrak{Z}_p$. The next conjecture is proposed by D. Broadhurst.

Conjecture 22.

$$\left(\sum_{p\geq 0}\sum_{k\geq 0} d_{p,k}X^pY^k\right)^{-1} = (1-X^2Y)\left(1-\frac{X^3Y}{1-X^2}+\frac{X^{12}Y^2(1-Y^2)}{(1-X^4)(1-X^6)}\right)\cdot$$

That Conjecture 22 implies Zagier's Conjecture 5 is easily seen by substituting $Y=1$ and using (4).

The left hand side in Conjecture 22 can be written as an infinite product: the next Lemma can be proved in the same way as Lemma 13.

Lemma 23. *Let $D(p,k)$ for $p\geq 0$ and $k\geq 1$ be non–negative integers. Then*

$$\prod_{p\geq 0}\prod_{k\geq 1}(1-X^pY^k)^{-D(p,k)} = \sum_{p\geq 0}\sum_{k\geq 1} d_{p,k}X^pY^k,$$

where $d_{p,k}$ is the number of tuples of non–negative integers of the form $\underline{h} = (h_{ij\ell})_{i\geq 0\ j\geq 1,\ 1\leq \ell\leq D(p,k)}$ satisfying

$$\sum_{i\geq 0}\sum_{j\geq 1}\sum_{n=1}^{D(p,k)} ih_{ij\ell} = p \quad \text{and} \quad \sum_{i\geq 0}\sum_{j\geq 1}\sum_{n=1}^{D(p,k)} jh_{ij\ell} = k.$$

If one believes Conjecture 12, a transcendence basis $\mathcal{T}$ of the field generated over $\mathbf{Q}$ by $\mathfrak{Z}$ should exist such that

$$D(p,k) = \mathrm{Card}\big(\mathcal{T}\cap\mathcal{F}^k\mathfrak{Z}_p\big)$$

is the number of Lyndon words on the alphabet $\{2,3\}$ with weight p and length k, so that

$$\prod_{p\geq 3}\prod_{k\geq 1}(1-X^pY^k)^{D(p,k)} = 1-\frac{X^3Y}{1-X^2}+\frac{X^{12}Y^2(1-Y^2)}{(1-X^4)(1-X^6)}\cdot$$

6 Words, non–commutative polynomials: free monoids and free algebras

6.1 Free structures as solution of universal problems

We will consider the categories of monoids, groups, abelian groups, vector spaces, commutative algebras, algebras, and, for each of them, we will consider the following *universal problem* of existence of an *initial object*:

Universal Problem 24. *Given a non–empty set X, does there exists an object $A(X)$ in this category and a map $\iota : X \to A(X)$ with the following property: for any object B in the category and any map $f : X \to B$, there exists a unique morphism $\underline{f} : A(X) \to B$ in this category for which the diagram*

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[dr, "\iota"] & B \\ & A(X) \arrow[u, "\underline{f}"'] \end{tikzcd}$$

commutes.

In each of these categories, the answer is yes, therefore, the solution $(A(X), i)$ is unique up to a unique isomorphism, meaning that if (A', ι') is another solution to this problem, then there is a unique isomorphism ι in the category between $A(X)$ and A' for which the diagram

$$\begin{tikzcd} X \arrow[r, "\iota'"] \arrow[dr, "\iota"] & A' \\ & A(X) \arrow[u, "\tilde{\iota}"'] \end{tikzcd}$$

commutes.

Our first example is the category of *monoid*, where the objects are pairs $(M, \cdot)$, where M is a non empty set and $\cdot$ a law

$$\begin{array}{ccc} M \times M & \longrightarrow & M \\ (a,b) & \longmapsto & ab \end{array}$$

which is associative with a neutral element e:

$$(ab)c = a(bc) \quad \text{and} \quad ae = ea = a$$

for all a, b and c in M, while the morphisms between two monoids M and M' are the maps $\varphi : M \to M'$ such that $\varphi(e) = e'$ and $\varphi(ab) = \varphi(a)\varphi(b)$.

When X is a non–empty set, we denote by $X^* = X^{(\mathbf{N})}$ the set of finite sequences of elements in X, including the empty sequence e. Write $x_1 \cdots x_p$ with $p \geq 0$ such a sequence (which is called a *word* on the *alphabet* X - the elements x_i in X are the *letters*). This set X^* is endowed with a monoid structure, the law on X^* is the *concatenation*:

$$(x_1 \cdots x_p)(x_{p+1} \cdots x_{p+q}) = x_1 \cdots x_{p+q},$$

which produces *the universal free monoid with basis* X. The neutral element is the *empty word* e. This monoid X^* with the canonical map $X \to X^*$, which maps a letter onto the corresponding word with a single letter, is the solution of the universal problem 24 in the category of monoids. The simplest case of a free monoid is when the given set X has a single element x, in which case the solution X^* of the universal problem 24 is the monoid $\{e, x, x^2, \dots, x^n \dots\}$ with the law $x^k x^\ell = x^{k+\ell}$ for k and ℓ non–negative integers. If one writes the law additively with the neutral element 0 and if we replace x by 1, this is a construction of the monoid $\mathbf{N} = \{0, 1, 2, \dots\}$ of the natural integers. We will study in §6.2 the free monoid X^* on a set $X := \{x_0, x_1\}$ with two elements as well as the free monoid Y^* on a countable set $Y := \{y_1, y_2, \dots\}$.

One can define a monoid by generators and relations: if X is the set of generators, we consider the free monoid X^* on X, we consider the equivalence relation, compatible with the concatenation, induced on X^* by the set of relations, and we take the quotient. For instance, one gets the solution of the universal problem 24 in the category of *commutative monoids* by taking the quotient of the free monoid on X by the equivalence relation induced by the relations $xy = yx$ for x and y in X. The *free commutative monoid* on a set X is the set $\mathbf{N}^{(X)}$ of maps $f : x \mapsto n_x$ from X to $\mathbf{N}$ with finite *support*: recall that the support of a map $f : X \to \mathbf{N}$ is the subset $\{x \in X \; ; \; f(x) \neq 0\}$ of X.

Another example is the construction of the *free group* on a set X: we consider the set Y which is the disjoint union of two copies of X. One can take, for instance, $Y = X_1 \cup X_2$, where $X_1 = X \times \{1\}$ and $X_2 = X \times \{2\}$. We consider the free monoid Y^* on Y and we take the equivalence relation on Y^* induced by $(x, 1)(x, 2) = e$ and $(x, 2)(x, 1) = e$ for all $x \in X$. A set of representative is given by the set $R(X)$ of so–called *reduced words*, which are the words $y_1 \cdots y_s$ in Y^* having no consecutive letters of the form $(x, 1)$, $(x, 2)$ nor of the form $(x, 2)$, $(x, 1)$ for some $x \in X$. One defines in a natural way a surjective map

$$\begin{array}{rccc} r: & Y^* & \longrightarrow & R(X) \\ & w & \longmapsto & r(w) \end{array}$$

by induction on the number ℓ of letters of the w's in Y^* as follows: if $w \in R(X)$, then $r(w) = w$; if $w \notin R(X)$, then there are two consecutive letters in the word w which are either of the form $(x, 1)$, $(x, 2)$ or of the form $(x, 2)(x, 1)$; we consider the word w' having $\ell - 2$ letters deduced from w by omitting such two consecutive letters (there may be several choices) and we use the induction hypothesis which allows us to define $r(w) = r(w')$. This maps is well defined (independent of the choices made) and allows one to define a law on the set of reduced words by $ww' = r(ww')$; this endows the set $R(X)$ with a structure of group. This is the *free group on* X.

Similarly, we get the *free abelian group on a set* X, which is just the group $\mathbf{Z}^{(X)}$. If X is finite with n elements, this is $\mathbf{Z}^n$.

The group defined by a set of generators X and a set of relations can be seen as the monoid constructed as follows: we consider the disjoint union Y of two copies of X, say X_1 and X_2 as above; the set of relations on X induces a set of relations on Y to which one adds the relations $(x,1)(x,2) = e$ and $(x,2)(x,1) = e$ for $x \in X$, and we take the quotient of Y^* by these relations.

In the category of *vector spaces over a field* K, the solution of the universal problem 24 associated with a set X is the *free vector space on* X, obtained by taking a set of variables e_x indexed by $x \in X$ and by considering the K–vector space E with basis $\{e_x\}_{x\in X}$. If X is finite with n elements, the free vector space over X has dimension n and is isomorphic to K^n. In general, the free vector space is $K^{(X)}$, which it the set of maps $X \to K$ with finite support, with the natural structure of K–vector space and with the natural injection $X \to K^{(X)}$ which maps $x \in X$ onto the characteristic function δ_x of $\{x\}$ (*Kronecker's symbol*):

$$\delta_x(y) = \begin{cases} 1 \text{ if } y = x, \\ 0 \text{ if } y \neq x, \end{cases} \qquad \text{for } y \in X.$$

In the category of *commutative algebras over a field* K, the solution of the universal problem 24 is the ring $K[\{T_x\}_{x\in X}]$ of polynomials in a set of variables indexed by X. One could take the elements of X as variables, and we will often do so, but if X has already some structure, it may be more convenient to introduce new letters (variables). For instance, when $X = \mathbf{N}$, it is better to introduce countably many variables $T_0, T_1, \dots$ in order to avoid confusion. Let us recall the construction of the *commutative polynomial algebra over a set* X, because the construction of the free algebra over X will be similar, only commutativity will not be there. Given the set of variables T_x with $x \in X$, an element in $K[\{T_x\}_{x\in X}]$ is a finite linear combination of *monomials*, where a *monomial* is a finite product

$$\prod_{x\in X} T_x^{n_x}.$$

Hence, to give a monomial is the same as to give a set $\underline{n} := \{n_x\}_{x\in X}$ of non–negative integers, all of which but a finite number are 0. We have seen that the set of such $\underline{n}$ is the underlying set of the free commutative monoid $\mathbf{N}^{(X)}$ on X. Hence, the K–vector space underlying the algebra of commutative polynomials $K[\{T_x\}_{x\in X}]$ is the free vector space on $\mathbf{N}^{(X)}$: to an element $\underline{n}$ in this set $\mathbf{N}^{(X)}$, we associate the coefficient $c_{\underline{n}}$ of the corresponding monomial, and an element in $K[\{T_x\}_{x\in X}]$ can be written in a unique way

$$\sum_{\underline{n}\in\mathbf{N}^{(X)}} c_{\underline{n}} \prod_{x\in X} T_x^{n_x}.$$

In the category of *algebras*, the solution of the universal problem 24 will be denoted by $K\langle X\rangle$. As a K–vector space, it is the free vector space over the free monoid X^*; hence, the elements of $K\langle X\rangle$ are the linear combinations of words with coefficients in K. The set underlying this space is the set $K^{(X^*)}$ of maps $S : X^* \to K$ having a finite support; for such a map, we denote by $(S|w)$ the image of $w \in X^*$ in K, so that the support is $\{w \in X^* \; ; \; (S|w) \neq 0\} \subset X^*$. We write also

$$S = \sum_{w \in X^*} (S|w)w.$$

On $K^{(X^*)}$, define an addition by

$$(S + T|w) = (S|w) + (T|w) \quad \text{for any } w \in X^* \tag{25}$$

and a multiplication [2] by

$$(ST|w) = \sum_{uv=w} (S|u)(T|v), \tag{26}$$

where, for each $w \in X^*$, the sum is over the (finite) set of (u, v) in $X^* \times X^*$ such that $uv = w$. Further, for $\lambda \in K$ and $S \in K^{(X^*)}$, define $\lambda S \in K^{(X^*)}$ by

$$(\lambda S|w) = \lambda(S|w) \quad \text{for any } w \in X^*. \tag{27}$$

With these laws, one checks that the set $K^{(X^*)}$ becomes a K-algebra, solution of the universal problem 24, which is denoted by $K\langle X\rangle$ and is called the *free algebra on* X.

This is a graded algebra, when elements of X are given weight 1: the weight of a word $x_1 \cdots x_p$ is p, and, for $p \geq 0$, the set $K\langle X\rangle_p$ of $S \in K\langle X\rangle$ for which

$$(S|w) = 0 \quad \text{if } w \in X^* \text{ has weight} \neq p$$

is the K–vector subspace whose basis is the set of words of length p. For $p = 0$, $K\langle X\rangle_0$ is the set Ke of *constant* polynomials λe, with $\lambda \in K$ – this is the K-subspace of dimension 1 spanned by e. For any $S \in K\langle X\rangle_p$ and $T \in K\langle X\rangle_q$, we have

$$ST \in K\langle X\rangle_{p+q}.$$

If X is finite with n elements, then, for each $p \geq 0$, there are n^p words of weight p; hence the dimension of $K\langle X\rangle_p$ over K is n^p, and the Hilbert–Poincaré series of the graded algebra $K\langle X\rangle$ is

$$\sum_{p \geq 0} t^p \dim_K K\langle X\rangle_p = \frac{1}{1 - nt} \cdot$$

For $n = 1$, this algebra $K\langle X\rangle$ is simply the (commutative) ring of polynomials $K[X]$ in one variable.

Remark. *Any group (resp. vector space, commutative algebra, algebra) is a quotient of a free group (resp. a free vector space, a free commutative algebra, a free algebra) by a normal subgroup (resp. a subspace, an ideal, a bilateral ideal). This is one more reason of the fundamental role of free structures!*

[2]Sometimes called *Cauchy product* - it is the usual multiplication, in opposition to the *Hadamard product*, where $(ST|w) = (S|w)(T|w)$.

6.2 The free algebra $\mathfrak{H}$ and its two subalgebras $\mathfrak{H}^1$ and $\mathfrak{H}^0$

Our first example of an alphabet was the trivial one with a single letter, the free monoid on a set with a single element 1 is just $\mathbf{N}$. The next example is when $X = \{x_0, x_1\}$ has two elements; in this case the algebra $K\langle x_0, x_1\rangle$ will be denoted by $\mathfrak{H}$. Each word w in X^* can be written $x_{\epsilon_1} \cdots x_{\epsilon_p}$, where each ϵ_i is either 0 or 1 and the integer p is the weight of w. The number of $i \in \{1, \dots p\}$ with $\epsilon_i = 1$ is called the *length* (or the *depth*) of w.

We will denote by X^*x_1 the set of word which end with x_1, and by $x_0X^*x_1$ the set of words which start with x_0 and end with x_1.

Consider a word w in X^*x_1. We write it $w = x_{\epsilon_1} \cdots x_{\epsilon_p}$ with $\epsilon_p = 1$. Let k be the number of occurrences of the letter x_1 in w. We have $p \geq 1$ and $k \geq 1$. We can write $w = x_0^{s_1-1}x_1x_0^{s_2-1}x_1 \cdots x_0^{s_k-1}x_1$ by defining $s_1 - 1$ as the number of occurrences of the letter x_0 before the first x_1 and, for $2 \leq j \leq k$, by defining $s_j - 1$ as the number of occurrences of the letter x_0 between the $(j-1)$-th and the j–th occurrence of x_1. This produces a sequence of non–negative integers $(s_1, \dots, s_k) \in \mathbf{N}^k$. Such a sequence $\underline{s} = (s_1, \dots, s_k)$ of positive integers with $k \geq 1$ is a called *composition* (the set of compositions is the union over $k \geq 1$ of these k–tuples $(s_1, \dots, s_k)$).

For $s \geq 1$, define $y_s = x_0^{s-1}x_1$:

$$y_1 = x_1, \;\; y_2 = x_0x_1, \;\; y_3 = x_0^2x_1, \;\; y_4 = x_0^3x_1, \;\; \dots$$

and let $Y = \{y_1, \; y_2, \; y_3, \dots\}$. For $\underline{s} = (s_1, \dots, s_k) \in \mathbf{N}^k$ with $s_j \geq 1$ $(1 \leq j \leq k)$, set $y_{\underline{s}} = y_{s_1} \cdots y_{s_k}$, so that

$$y_{\underline{s}} = x_0^{s_1-1}x_1 \cdots x_0^{s_k-1}x_1.$$

Lemma 28.

a) The set X^*x_1 is the same as the set of $y_{s_1} \cdots y_{s_k}$, where $(s_1, \dots, s_k)$ ranges over the finite sequences of positive integers with $k \geq 1$ and $s_j \geq 1$ for $1 \leq j \leq k$.

b) The free monoid Y^* on the set Y is $\{e\} \cup X^*x_1$.

c) The set $x_0X^*x_1$ is the set $\{y_{\underline{s}}\}$, where $\underline{s}$ is a composition having $s_1 \geq 2$.

The subalgebra of $\mathfrak{H}$ spanned by X^*x_1 is

$$\mathfrak{H}^1 = Ke + \mathfrak{H}x_1$$

and $\mathfrak{H}x_1$ is a left ideal of $\mathfrak{H}$. The algebra $\mathfrak{H}^1$ is the free algebra $K\langle Y\rangle$ on the set Y. We observe an interesting phenomenon, which does not occur in the commutative case: the free algebra $K\langle x_0, x_1\rangle$ on a set with only two elements contains as a subalgebra the free algebra $K\langle y_1, y_2, \dots\rangle$ on a set with countably many elements. Notice that, for each $n \geq 1$, this last algebra also contains as a subalgebra the free algebra on a set with n elements, namely $K\langle y_1, y_2, \dots, y_n\rangle$. From this point of view, it suffices to deal with only two variables! Any finite message can be encoded with an alphabet with only two letters. Also, we see that a naive definition of a dimension for such spaces, where $K\langle y_1,\dots,y_n\rangle$ would have dimension n, would be misleading.

A word in $x_0X^*x_1$ is called *convergent*. The reason is that one defines a map on the convergent words

$$\widehat{\zeta}: x_0X^*x_1 \to \mathbf{R}$$

by setting $\widehat{\zeta}(y_{\underline{s}})=\zeta(s)$. Also the subalgebra of $\mathfrak{H}^1$ spanned by $x_0X^*x_1$ is

$$\mathfrak{H}^0 = Ke + x_0\mathfrak{H}x_1.$$

One extends the map $\widehat{\zeta}: x_0X^*x_1 \to \mathbf{R}$ by K–linearity and obtain a map $\widehat{\zeta}:\mathfrak{H}^0\to\mathbf{R}$ such that $\widehat{\zeta}(e)=1$.

On $\mathfrak{H}^0$ there is a structure of non–commutative algebra, given by the concatenation – however, the map $\widehat{\zeta}$ has no good property for this structure. The concatenation of $y_2=x_1x_0$ and $y_3=x_0^2x_1$ is $y_2y_3=y_{(2,3)}=x_0x_1x_0^2x_1$, while $\zeta(2)\zeta(3)\neq\zeta(2,3)$ and $\zeta(2)\zeta(3)\neq\zeta(2,3)$; indeed according to [1] we have $\zeta(2)=1.644\dots$, $\zeta(3)=1.202\dots$

$$\zeta(2)\zeta(3)=1.977\dots\quad \zeta(2,3)=0.711\dots\quad \zeta(3,2)=0.228\dots$$

so

$$\widehat{\zeta}(y_2)\widehat{\zeta}(y_3)\neq\widehat{\zeta}(y_2y_3).$$

As we have seen in §3, it is expected that $\zeta(2)$, $\zeta(3)$ and $\zeta(2,3)$ are algebraically independent.

However, there are other structures of algebras on $\mathfrak{H}^0$, in particular two commutative algebra structures ш (*shuffle* – see §7.3) – and $\star$ (*stuffle* or *harmonic law* – see §8.1 –) for which $\widehat{\zeta}$ will become an algebra homomorphism: for w and w' in $\mathfrak{H}^0$,

$$\widehat{\zeta}(w\text{ш}w')=\zeta(w)\zeta(w')\quad\text{and}\quad\widehat{\zeta}(w\star w')=\zeta(w)\zeta(w').$$

7 MZV as integrals and the Shuffle Product

7.1 Zeta values as integrals

We first check

$$\zeta(2)=\int_{1>t_1>t_2>0}\frac{dt_1}{t_1}\cdot\frac{dt_2}{1-t_2}\cdot \tag{29}$$

For t_2 in the interval $0<t_2<1$, we expand $1/(1-t_2)$ in power series; next we integrate over the interval $[0,t_1]$, where t_1 is in the interval $0<t_1<1$, so that the integration terms by terms is licit:

$$\frac{1}{1-t_2}=\sum_{n\ge1}t_2^{n-1},\qquad \int_0^{t_1}\frac{dt_2}{1-t_2}=\sum_{n\ge1}\int_0^{t_1}t_2^{n-1}dt_2=\sum_{n\ge1}\frac{t_1^n}{n}\cdot$$

Hence, the integral in the right hand side of (29) is

$$\int_0^1 \frac{dt_1}{t_1} \int_0^{t_1} \frac{dt_2}{1-t_2} = \sum_{n\geq 1} \frac{1}{n} \int_0^1 t_1^{n-1} dt_1 = \sum_{n\geq 1} \frac{1}{n^2} = \zeta(2).$$

In the same way, one checks

$$\zeta(3) = \int_{1>t_1>t_2 t_3>0} \frac{dt_1}{t_1} \cdot \frac{dt_2}{t_2} \cdot \frac{dt_3}{1-t_3} \cdot$$

We continue by induction and give an integral formula for $\zeta(s)$ when $s \geq 2$ is an integer, and, more generally, for $\zeta(\underline{s})$, when $\underline{s} = (s_1, \ldots, s_k)$ is a composition. To state the result, it will be convenient to introduce a definition: we set

$$\omega_0(t) = \frac{dt}{t} \quad \text{and} \quad \omega_1(t) = \frac{dt}{1-t} \cdot$$

For $p \geq 1$ and $\epsilon_1, \ldots, \epsilon_p$ in $\{0, 1\}$, we define

$$\omega_{\epsilon_1} \cdots \omega_{\epsilon_p} = \omega_{\epsilon_1}(t_1) \cdots \omega_{\epsilon_p}(t_p).$$

We will integrate this differential form on the simplex

$$\Delta_p = \{(t_1, \ldots, t_p) \in \mathbf{R}^p \; ; \; 1 > t_1 > \cdots > t_p > 0\}$$

The two previous formulae are

$$\zeta(2) = \int_{\Delta_2} \omega_0 \omega_1 \quad \text{and} \quad \zeta(3) = \int_{\Delta_3} \omega_0^2 \omega_1.$$

The integral formula for zeta values that follows by induction is

$$\zeta(s) = \int_{\Delta_s} \omega_0^{s-1} \omega_1 \quad \text{for} \quad s \geq 2.$$

We extend this formula to multiple zeta values as follows. Firstly, for $s \geq 1$, we define $\omega_s = \omega_0^{s-1}\omega_1$, which matches the previous definition when $s = 1$ and produces, for $s \geq 2$,

$$\zeta(s) = \int_{\Delta_s} \omega_s.$$

Next, for $\underline{s} = (s_1, \ldots, s_k)$, we define $\omega_{\underline{s}} = \omega_{s_1} \cdots \omega_{s_k}$.

Proposition 30. *Assume $s_1 \geq 2$. Let $p = s_1 + \cdots + s_k$. Then*

$$\zeta(\underline{s}) = \widehat{\zeta}(y_{\underline{s}}) = \int_{\Delta_p} \omega_{\underline{s}}.$$

The proof is by induction on p. For this induction argument, it is convenient to introduce the *multiple polylogarithm functions in one variable*:

$$\mathrm{Li}_{\underline{s}}(z) = \sum_{n_1 > n_2 > \cdots > n_k \geq 1} \frac{z^{n_1}}{n_1^{s_1} \cdots n_k^{s_k}},$$

for $\underline{s} = (s_1, \dots, s_k)$ with $s_j \geq 1$ and for $z \in \mathbf{C}$ with $|z| < 1$. Notice that $\mathrm{L}i(z)$ is defined also at $z = 1$ when $s_1 \geq 2$, where it takes the value $\mathrm{Li}_{\underline{s}}(1) = \zeta(\underline{s})$. One checks, by induction on the weight p, for $0 < z < 1$,

$$\mathrm{Li}_{\underline{s}}(z) = \int_{\Delta_p(z)} \omega_{\underline{s}},$$

where $\Delta_p(z)$ is the simplex

$$\Delta_p(z) = \{(t_1, \dots, t_p) \in \mathbf{R}^p \; ; \; z > t_1 > \cdots > t_p > 0\}.$$

We now consider products of such integrals. Using (29), write $\zeta(2)^2$ as product of two integrals

$$\zeta(2)^2 = \int_{\substack{1 > t_1 > t_2 > 0 \\ 1 > u_1 > u_2 > 0}} \frac{dt_1}{t_1} \cdot \frac{dt_2}{1 - t_2} \cdot \frac{du_1}{u_1} \cdot \frac{du_2}{1 - u_2}.$$

We decompose the domain

$$1 > t_1 > t_2 > 0, \quad 1 > u_1 > u_2 > 0$$

into six disjoint domains (and further subsets of zero dimension) obtained by "shuffling" (t_1, t_2) with (u_1, u_2):

$$\begin{array}{ll} 1 > t_1 > t_2 > u_1 > u_2 > 0, & 1 > t_1 > u_1 > t_2 > u_2 > 0, \\ 1 > u_1 > t_1 > t_2 > u_2 > 0, & 1 > t_1 > u_1 > u_2 > t_2 > 0, \\ 1 > u_1 > t_1 > u_2 > t_2 > 0, & 1 > u_1 > u_2 > t_1 > t_2 > 0, \end{array}$$

Each of the six simplices have either t_1 or u_1 as the largest variable (corresponding to $\omega_0(t) = 1/t$) and u_2 or t_2 as the lowest (corresponding to $\omega_1(t) = dt/(1 - t)$). The integrals of $\omega_0^2\omega_1^2$ produce $\zeta(3, 1)$, there are 4 of them, the integrals of $\omega_0\omega_1\omega_0\omega_1$ produce $\zeta(2, 2)$, there are 2 of them. From Proposition 30, we deduce

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

This is a typical example of a "shuffle relation":

$$\omega_0\omega_1 ⧢ \omega_0\omega_1 = 4\omega_0^2\omega_1^2 + 2\omega_0\omega_1\omega_0\omega_1.$$

7.2 Chen's integrals

Chen iterated integrals are defined by induction as follows. Let $\varphi_1, \dots, \varphi_p$ be holomorphic differential forms on a simply connected open subset D of the complex plane and let x and y be two elements in D. Define, as usual, $\int_x^y \varphi_1$ as the value, at y, of the primitive of φ_1 which vanishes at x. Next, by induction on p, define

$$\int_x^y \varphi_1 \cdots \varphi_p = \int_x^y \varphi_1(t) \int_x^t \varphi_2 \cdots \varphi_p.$$

By means of a change of variables

$$t \longmapsto x + t(y - x),$$

one can assume that $x = 0$, $y = 1$ and that D contains the real segment $[0, 1]$. In this case the integral is

$$\int_0^1 \varphi_1 \cdots \varphi_p = \int_{\Delta_p} \varphi_1(t_1)\varphi_2(t_2) \cdots \varphi_p(t_p),$$

where the domain of integration Δ_p is the simplex of $\mathbf{R}^p$ defined by

$$\Delta_p = \bigl\{(t_1, \dots, t_p) \in \mathbf{R}^p\ ,\ 1 > t_1 > \cdots > t_p > 0\bigr\}.$$

In our applications, the open set D will be the open disk $|z - (1/2)| < 1/2$, the differential forms will be either dt/t or $dt/(1 - t)$; so one needs to take care of the fact that the limit points 0 and 1 of the integrals are not in D. One way to overcome this difficulty is by integrating from ϵ_1 to $1 - \epsilon_2$ with $\epsilon_1 > 0$, $\epsilon_2 > 0$ and $\epsilon_1 + \epsilon_2 < 1$, and by letting ϵ_1 and ϵ_2 tend to 0. Here, we just ignore these convergence questions by restricting our discussion to the convergent words and to the algebra $\mathfrak{H}^0$ they generate.

The product of two integrals is a Chen integral, and more generally the product of two Chen integrals is a Chen integral. This is where the shuffle comes in. We consider a special case: the product of the two integrals

$$\int_{1>t_1>t_2>0} \varphi_1(t_1)\varphi_2(t_2) \int_{1>t_3>0} \varphi_3(t_3)$$

is the sum of three integrals

$$\int_{1>t_1>t_2>t_3>0} \varphi_1(t_1)\varphi_2(t_2)\varphi_3(t_3),$$

$$\int_{1>t_1>t_3>t_2>0} \varphi_1(t_1)\varphi_3(t_3)\varphi_2(t_2)$$

and

$$\int_{1>t_3>t_1>t_2>0} \varphi_3(t_3)\varphi_1(t_1)\varphi_2(t_2).$$

Consider, for instance, the third integral: we write it as

$$\int_{1>t_{\tau(1)}>t_{\tau(2)}>t_{\tau(3)}>0} \varphi_{\tau(1)}(t_{\tau(1)})\,\varphi_{\tau(2)}(t_{\tau(2)})\,\varphi_{\tau(3)}(t_{\tau(3)}).$$

The permutation τ of $\{1,2,3\}$ is $\tau(1)=3$, $\tau(2)=1$, $\tau(3)=2$, and it is one of the three permutations of $\mathfrak{S}_3$ which is of the form σ^{-1}, where $\sigma(1)<\sigma(2)$.

The shuffle ш will be defined in §7.3 so that the next lemma holds:

Lemma 31. *Let $\varphi_1,\ldots,\varphi_{p+q}$ be differential forms with $p\ge 0$ and $q\ge 0$. Then*

$$\int_0^1 \varphi_1\cdots\varphi_p \int_0^1 \varphi_{p+1}\cdots\varphi_{p+q} = \int_0^1 \varphi_1\cdots\varphi_p ш \varphi_{p+1}\cdots\varphi_{p+q}.$$

Proof. Define $\Delta'_{p,q}$ as the subset of $\Delta_p\times\Delta_q$ of those elements $(z_1,\ldots,z_{p+q})$ for which we have $z_i\ne z_j$ for $1\le i\le p<j\le p+q$. Hence,

$$\int_0^1 \varphi_1\cdots\varphi_p \int_0^1 \varphi_{p+1}\cdots\varphi_{p+q} = \int_{\Delta_p\times\Delta_q} \varphi_1\cdots\varphi_{p+q} = \int_{\Delta'_{p,q}} \varphi_1\cdots\varphi_{p+q},$$

where $\Delta'_{p,q}$ is the disjoint union of the subsets $\Delta^\sigma_{p,q}$ defined by

$$\Delta^\sigma_{p,q} = \bigl\{(t_1,\ldots,t_{p+q})\ ;\ 1>t_{\sigma^{-1}(1)}>\cdots>t_{\sigma^{-1}(p+q)}>0\bigr\},$$

for σ running over the set $\mathfrak{S}_{p,q}$ of permutations of $\{1,\ldots,p+q\}$ satisfying

$$\sigma(1)<\sigma(2)<\cdots<\sigma(p) \quad\text{and}\quad \sigma(p+1)<\sigma(p+2)<\cdots<\sigma(p+q).$$

Hence,

$$\int_{\Delta'_{p,q}} \varphi_1\cdots\varphi_{p+q} = \sum_{\sigma\in\mathfrak{S}_{p,q}} \int_0^1 \varphi_{\sigma^{-1}(1)}\cdots\varphi_{\sigma^{-1}(p+q)}.$$

Lemma 31 follows, provided that we define the shuffle ш so that

$$\sum_{\sigma\in\mathfrak{S}_{p,q}} \varphi_{\sigma^{-1}(1)}\cdots\varphi_{\sigma^{-1}(p+q)} = \varphi_1\cdots\varphi_p ш \varphi_{p+1}\cdots\varphi_{p+q}.$$

□

7.3 The shuffle ш and the shuffle Algebra $\mathfrak{H}_ш$

Let X be a set and K a field. On $K\langle X\rangle$ we define the *shuffle* product as follows. On the words, the map ш $: X^*\times X^*\to\mathfrak{H}$ is defined by the formula

$$(x_1\cdots x_p)ш(x_{p+1}\cdots x_{p+q}) = \sum_{\sigma\in\mathfrak{S}_{p,q}} x_{\sigma^{-1}(1)}\cdots x_{\sigma^{-1}(p+q)},$$

where $\mathfrak{S}_{p,q}$ denotes the set of permutation σ on $\{1,\ldots,p+q\}$ satisfying

$$\sigma(1)<\sigma(2)<\cdots<\sigma(p) \quad\text{and}\quad \sigma(p+1)<\sigma(p+2)<\cdots<\sigma(p+q).$$

This set $\mathfrak{S}_{p,q}$ has $(p+q)!/p!q!$ elements; if $(p,q) \neq (0,0)$, it is the disjoint union of two subsets, the first one with $(p-1+q)!/(p-1)!q!$ elements consists of those σ for which $\sigma(1)=1$, and the second one with $(p+q-1)!/p!(q-1)!$ elements consists of those σ for which $\sigma(p+1)=1$.

Write $y_i = x_{\sigma^{-1}(i)}$, so that $x_j = y_{\sigma(j)}$ for $1 \le i \le p+q$ and $1 \le j \le p+q$. The letters $x_1,\dots,x_{p+q}$ and $y_1,\dots,y_{p+q}$ are the same, only the order may differ. However, $x_1,\dots,x_p$ (which is the same as $y_{\sigma(1)},\dots,y_{\sigma(p)}$) occur in this order in $y_1,\dots,y_{p+q}$, and so do $x_{p+1},\dots,x_{p+q}$ (which is the same as $y_{\sigma(p+1)},\dots,y_{\sigma(p+q)}$).

Accordingly, the previous definition of $ш : X^* \times X^* \to \mathfrak{H}$ is equivalent to the following inductive one:

$$e ш w = w ш e = w \quad \text{for any } w \in X^*,$$

and

$$(xu) ш (yv) = x\bigl(u ш (yv)\bigr) + y\bigl((xu) ш v\bigr)$$

for x and y in X (letters), u and v in X^* (words).

Examples. *For k and ℓ non-negative integers and $x \in X$,*

$$x^k ш x^\ell = \frac{(k+\ell)!}{k!\ell!} x^{k+\ell}.$$

From

$$\mathfrak{S}_{2,2} = \bigl\{(1)\ ;\ (2,3)\ ;\ (2,4,3)\ ;\ (1,2,3)\ ;\ (1,2,4,3)\ ;\ (1,3)(2,4)\bigr\}$$

one deduces

$$x_1x_2 ш x_3x_4 = x_1x_2x_3x_4 + x_1x_3x_2x_4 + x_1x_3x_4x_2 + x_3x_1x_2x_4 + x_3x_1x_4x_2 + x_3x_4x_1x_2,$$

hence,

$$x_0x_1 ш x_0x_1 = 2x_0x_1x_0x_1 + 4x_0^2x_1^2.$$

In the same way the relation

$$x_0x_1 ш x_0^2x_1 = x_0x_1x_0^2x_1 + 3x_0^2x_1x_0x_1 + 6x_0^3x_1^2$$

is easily checked by computing more generally $x_0x_1 ш x_2x_3x_4$ as a sum of $6!/(2!3!) = 10$ *terms.*

Notice that the shuffle product of two words is most often not a word but a polynomial in $K\langle X\rangle$. We extend the definition of $ш : X^* \times X^* \to \mathfrak{H}$ to $ш : \mathfrak{H} \times \mathfrak{H} \to \mathfrak{H}$ by distributivity with respect to addition:

$$\sum_{u\in X^*} (S|u)u \ ш \sum_{v\in X^*} (T|v)v = \sum_{u\in X^*}\sum_{v\in X^*} (S|u)(T|v) u ш v.$$

One checks that the shuffle ш endows $K\langle X\rangle$ with a structure of commutative K-algebra.

From now on we consider only the special case $X = \{x_0, x_1\}$. The set $\mathfrak{H} = K\langle X\rangle$ with the shuffle law ш is a commutative algebra which will be denoted by $\mathfrak{H}_{ш}$. Since $\mathfrak{H}^1$ as well as $\mathfrak{H}^0$ are stable under ш, they define subalgebras

$$\mathfrak{H}^0_{ш} \subset \mathfrak{H}^1_{ш} \subset \mathfrak{H}_{ш}.$$

Since, for $\underline{s} = (s_1, \ldots, s_k)$ and $\underline{s}' = (s'_1, \ldots, s'_k)$, with $s_j \geq 1$, $s'_j \geq 1$, $s'_1 \geq 2$ and $s_1 \geq 2$, we have

$$\widehat{\zeta}(y_{\underline{s}})\widehat{\zeta}(y_{\underline{s}'}) = \widehat{\zeta}(y_{\underline{s}} ш y_{\underline{s}'}),$$

we deduce:

Theorem 32. *The map*

$$\widehat{\zeta} : \mathfrak{H}^0_{ш} \longrightarrow \mathbf{R}$$

is a homomorphism of commutative algebras.

8 Product of series and the harmonic algebra

8.1 The stuffle $\star$ and the harmonic algebra $\mathfrak{H}_\star$

There is another shuffle-like law on $\mathfrak{H}$, called the *harmonic product* by M. Hoffman and *stuffle* by other authors, denoted with a star, which also gives rise to subalgebras

$$\mathfrak{H}^0_\star \subset \mathfrak{H}^1_\star \subset \mathfrak{H}_\star.$$

The starting point is the observation that the product of two multizeta series is a linear combination of multizeta series. Indeed, the cartesian product

$$\big\{(n_1, \ldots, n_k)\ ;\ n_1 > \cdots > n_k\big\} \times \big\{(n'_1, \ldots, n'_{k'})\ ;\ n'_1 > \cdots > n'_{k'}\big\}$$

breaks into a disjoint union of subsets of the form

$$\big\{(n''_1, \ldots, n''_{k''})\ ;\ n''_1 > \cdots > n''_{k''}\big\}$$

with each k'' satisfying $\max\{k, k'\} \leq k'' \leq k + k'$. The simplest example is $\zeta(2)^2 = 2\zeta(2,2) + \zeta(4)$, a special case of Nielsen Reflexion Formula already seen in §1.2.

We write this as follows:

$$y_{\underline{s}} \star y_{\underline{s}'} = \sum_{\underline{s}''} y_{\underline{s}''}, \tag{33}$$

where $\underline{s}''$ runs over the tuples $(s''_1, \ldots, s''_{k''})$ obtained from $\underline{s} = (s_1, \ldots, s_k)$ and $\underline{s}' = (s'_1, \ldots, s'_{k'})$ by inserting, in all possible ways, some 0 in the string $(s_1, \ldots, s_k)$ as well as in the string $(s'_1, \ldots, s'_{k'})$ (including in front and at the end), so that the new strings have the same length k'', with $\max\{k, k'\} \leq k'' \leq k + k'$, and by adding the two sequences term by term. Here is an example:

$$\begin{array}{cccccccc}
\underline{s} & s_1 & s_2 & 0 & s_3 & s_4 & \cdots & 0 \\
\underline{s}' & 0 & s'_1 & s'_2 & 0 & s'_3 & \cdots & s'_{k'} \\
\underline{s}'' & s_1 & s_2 + s'_1 & s'_2 & s_3 & s_4 + s'_3 & \cdots & s'_{k'}.
\end{array}$$

Notice that the weight of the last string (sum of the s''_j) is the sum of the weight of $\underline{s}$ and the weight of $\underline{s}'$.

More precisely, the law $\star$ on $\mathfrak{H}$ is defined as follows. First on X^*, the map $\star : X^* \times X^* \to \mathfrak{H}$ is defined by induction, starting from

$$x_0^n \star w = w \star x_0^n = w x_0^n$$

for any $w \in X^*$ and any $n \geq 0$ (for $n = 0$, it means $e \star w = w \star e = w$ for all $w \in X^*$), and then

$$(y_s u) \star (y_t v) = y_s\big(u \star (y_t v)\big) + y_t\big((y_s u) \star v\big) + y_{s+t}(u \star v)$$

for u and v in X^*, s and t positive integers.

We will not use so many parentheses later: in a formula where there are both concatenation products and either shuffle of star products, we agree that concatenation is always performed first, unless parentheses impose another priority:

$$y_s u \star y_t v = y_s(u \star y_t v) + y_t(y_s u \star v) + y_{s+t}(u \star v)$$

Again, this law is extended to all of $\mathfrak{H}$ by distributivity with respect to addition:

$$\sum_{u \in X^*} (S|u)u \quad \star \sum_{v \in X^*} (T|v)v = \sum_{u \in X^*} \sum_{v \in X^*} (S|u)(T|v) u \star v.$$

Remark. *From the definition (by induction on the length of uv) one deduces*

$$(u x_0^m) \star (v x_0^n) = (u \star v) x_0^{m+n}$$

for $m \geq 0$, u and v in X^*.

Example.

$$y_s^{\star 3} = y_s \star y_s \star y_s = 6 y_s^3 + 3 y_s y_{2s} + 3 y_{2s} y_2 + y_{3s}.$$

The additive group $K\langle X\rangle$ endowed with the harmonic law $\star$ is a commutative algebra which will be denoted by $\mathfrak{H}_\star$. Since $\mathfrak{H}^1$ as well as $\mathfrak{H}^0$ are stable under $\star$, they define subalgebras

$$\mathfrak{H}_\star^0 \subset \mathfrak{H}_\star^1 \subset \mathfrak{H}_\star.$$

Since, for $\underline{s} = (s_1, \ldots, s_k)$ and $\underline{s}' = (s'_1, \ldots, s'_k)$, with $s_j \geq 1$, $s'_j \geq 1$, $s'_1 \geq 2$ and $s_1 \geq 2$, we have

$$\widehat{\zeta}(y_{\underline{s}}) \widehat{\zeta}(y_{\underline{s}'}) = \widehat{\zeta}(y_{\underline{s}} \star y_{\underline{s}'}),$$

we deduce:

Theorem 34. *The map*

$$\widehat{\zeta} : \mathfrak{H}_\star^0 \longrightarrow \mathbf{R}$$

is a homomorphism of commutative algebras.

8.2 Regularized double shuffle relations

As a consequence of theorems 32 and 34, the kernel of $\widehat{\zeta}$ contains all elements $w ш w' - w \star w'$ for w and w' in $\mathfrak{H}^0$: indeed

$$\widehat{\zeta}(w ш w') = \widehat{\zeta}(w)\widehat{\zeta}(w') = \widehat{\zeta}(w \star w'), \quad \text{hence,} \quad \widehat{\zeta}(w ш w' - w \star w') = 0.$$

However, the relation $\zeta(2,1) = \zeta(3)$ (due to Euler) is not a consequence of these relations, but one may derive it in a formal way as follows.

Consider

$$y_1 ш y_2 = x_1 ш x_0 x_1 = 2x_0^2 x_1 + x_1 x_0 x_1 = 2y_2 y_1 + y_1 y_2$$

and

$$y_1 \star y_2 = y_1 y_2 + y_2 y_1 + y_3.$$

They are not in $\mathfrak{H}^0$, but their difference

$$y_1 ш y_2 - y_1 \star y_2 = y_2 y_1 - y_3$$

is in $\mathfrak{H}^0$, and Euler's relation says that this difference is in the kernel of $\widehat{\zeta}$. This is the simplest example of the so–called *Regularized double shuffle relations.*

9 The structure of the shuffle and harmonic algebras

The shuffle and harmonic algebras are polynomial algebras in the Lyndon words.

Consider the lexicographic order on X^* with $x_0 < x_1$ and denote by $\mathcal{L}$ the set of Lyndon words (see §3). We have seen that x_0 is the only Lyndon word which is not in $\mathfrak{H}^1_ш$, while x_0 and x_1 are the only Lyndon words which are not in $\mathfrak{H}^0_ш$.

For instance, there are $1 + 2 + 2^2 + 2^3 = 15$ words of weight ≤ 3 (where x_0 and x_1 have weight 1). Since $L_1 + L_2 + L_3 = 5$, among them 5 are Lyndon words:

$$x_0 < x_0^2 x_1 < x_0 x_1 < x_0 x_1^2 < x_1.$$

Accordingly, we introduce 5 variables

$$T_{10},\ T_{21},\ T_{11},\ T_{12},\ T_{01},$$

where T_{ij} corresponds to $x_0^i x_1^j$.

9.1 Shuffle

The structure of the commutative algebra $\mathfrak{H}_ш$ is given by *Radford Theorem.*

Theorem 35. *The three shuffle algebras are (commutative) polynomial algebras*

$$\mathfrak{H}_ш = K[\mathcal{L}]_ш, \quad \mathfrak{H}^1_ш = K\big[\mathcal{L} \setminus \{x_0\}\big]_ш \quad \textit{and} \quad \mathfrak{H}^0_ш = K\big[\mathcal{L} \setminus \{x_0, x_1\}\big]_ш.$$

We write the 10 non-Lyndon words of weight ≤ 3 as polynomials in these Lyndon words as follows:

$$e = e,$$

$$x_0^2 = \frac{1}{2} x_0 ш x_0 = \frac{1}{2} T_{10}^2,$$

$$x_0^3 = \frac{1}{3} x_0 ш x_0 ш x_0 = \frac{1}{3} T_{10}^3,$$

$$x_0 x_1 x_0 = x_0 ш x_0 x_1 - 2x_0^2 x_1 = T_{10} T_{11} - 2T_{21},$$

$$x_1 x_0 = x_0 ш x_1 - x_0 x_1 = T_{10} T_{01} - T_{11},$$

$$x_1 x_0^2 = \frac{1}{2} x_0 ш x_0 ш x_1 - x_0 ш x_0 x_1 + x_0^2 x_1 = \frac{1}{2} T_{10}^2 T_{01} - T_{10} T_{11} + T_{21},$$

$$x_1 x_0 x_1 = x_0 x_1 ш x_1 - 2x_0 x_1^2 = T_{11} T_{01} - 2T_{12},$$

$$x_1^2 = \frac{1}{2} x_1 ш x_1 = \frac{1}{2} T_{01}^2,$$

$$x_1^2 x_0 = \frac{1}{2} x_0 ш x_1 ш x_1 - x_0 x_1 ш x_1 + x_0 x_1^2 = \frac{1}{2} T_{10} T_{01}^2 - T_{11} T_{01} + T_{12},$$

$$x_1^3 = \frac{1}{3} x_1 ш x_1 ш x_1 = \frac{1}{3} T_{01}^3.$$

Corollary 36. *We have*

$$\mathfrak{H}_ш = \mathfrak{H}_ш^1[x_0]_ш = \mathfrak{H}_ш^0[x_0, x_1]_ш \quad \textit{and} \quad \mathfrak{H}_ш^1 = \mathfrak{H}_ш^0[x_1]_ш.$$

9.2 Harmonic algebra

Hoffman's Theorem gives the structure of the harmonic algebra $\mathfrak{H}_\star$:

Theorem 37. *The harmonic algebras are polynomial algebras on Lyndon words:*

$$\mathfrak{H}_\star = K[\mathcal{L}]_\star, \quad \mathfrak{H}_\star^0 = K\big[\mathcal{L} \setminus \{x_0, x_1\}\big]_\star \quad \textit{and} \quad \mathfrak{H}_\star^1 = K\big[\mathcal{L} \setminus \{x_0, x_1\}\big]_\star.$$

For instance, the 10 non-Lyndon words of weight ≤ 3 are polynomials in the

5 Lyndon words as follows:

$$e = e,$$

$$x_0^2 = x_0 \star x_0 = T_{10}^2,$$

$$x_0^3 = x_0 \star x_0 \star x_0 = T_{10}^3,$$

$$x_0 x_1 x_0 = x_0 \star x_0 x_1 = T_{10} T_{11},$$

$$x_1 x_0 = x_0 \star x_1 = T_{10} T_{01},$$

$$x_1 x_0^2 = x_0 \star x_0 \star x_1 = T_{10}^2 T_{01},$$

$$x_1 x_0 x_1 = x_0 x_1 \star x_1 - x_0^2 x_1 - x_0 x_1^2 = T_{11} T_{01} - T_{21} - T_{12},$$

$$x_1^2 = \tfrac{1}{2} x_1 \star x_1 - \tfrac{1}{2} x_0 x_1 = \tfrac{1}{2} T_{01}^2 - \tfrac{1}{2} T_{11},$$

$$x_1^2 x_0 = \tfrac{1}{2} x_0 \star x_1 \star x_1 - \tfrac{1}{2} x_0 \star x_0 x_1 = \tfrac{1}{2} T_{10} T_{01}^2 - \tfrac{1}{2} T_{10} T_{11},$$

$$x_1^3 = \tfrac{1}{6} x_1 \star x_1 \star x_1 - \tfrac{1}{2} x_0 x_1 \star x_1 + \tfrac{1}{3} x_0^2 x_1 = \tfrac{1}{6} T_{01}^3 - \tfrac{1}{2} T_{11} T_{01} + \tfrac{1}{3} T_{21}.$$

In the same way as Corollary 36 follows from Theorem 35, we deduce from Theorem 37:

Corollary 38. *We have*

$$\mathfrak{H}_\star = \mathfrak{H}_\star^1[x_0]_\star = \mathfrak{H}_\star^0[x_0, x_1]_\star \quad \textit{and} \quad \mathfrak{H}_\star^1 = \mathfrak{H}_\star^0[x_1]_\star.$$

Remark. Consider the diagram

$$\begin{array}{ccc} \mathfrak{H}_{\text{ш}} & \longrightarrow & K[\mathcal{L}]_{\text{ш}} \\ \Big\downarrow f & & \Big\downarrow g \\ \mathfrak{H}_\star & \longrightarrow & K[\mathcal{L}]_\star \end{array}$$

The horizontal maps are just the identitication of $\mathfrak{H}_{\text{ш}}$ with $K[\mathcal{L}]_{\text{ш}}$ and of $\mathfrak{H}_\star$ with $K[\mathcal{L}]_\star$. The vertical map f is the identity map on $\mathfrak{H}$, since the algebras $\mathfrak{H}_{\text{ш}}$ and $\mathfrak{H}_\star$ have the same underlying set $\mathfrak{H}$ (only the law differs). But the map g which makes the diagram commute is not a morphism of algebras: it maps each Lyndon word onto itself; but consider, for instance, the image of the word x_0^2. As a polynomial in $K[\mathcal{L}]_{\text{ш}}$, we have

$$x_0^2 = \frac{1}{2} x_0 \text{ш} x_0 = \frac{1}{2} x_0^{\text{ш}2},$$

but as a polynomial in $K[\mathcal{L}]_\star$, we have

$$x_0^2 = x_0 \star x_0 = x_0^{\star 2},$$

hence, $g(T_{10}^2) = 2T_{10}^2$.

10 Regularized double shuffle relations

10.1 Hoffman standard relations

The relation $\zeta(2,1)=\zeta(3)$ is the easiest example of a whole class of linear relations among MZV.

For any word $w\in X^*$, each of $x_1\star x_0wx_1$ and $x_1ш x_0wx_1$ is the sum of $x_1x_0wx_1$ with other words in $x_0X^*x_1$ (i.e. convergent words):

$$x_1ш w - x_1w \in \mathfrak{H}^0, \quad x_1\star w - x_1w\in\mathfrak{H}^0.$$

Hence,

$$x_1ш w - x_1\star w\in\mathfrak{H}^0.$$

It turns out that these elements $x_1ш w - x_1\star w$ are in the kernel of $\widehat{\zeta}:\mathfrak{H}^0\to\mathbf{R}$.

Proposition 39. *For $w\in\mathfrak{H}^0$,*

$$\widehat{\zeta}(x_1\star w - x_1ш w)=0.$$

Since $x_1ш e = x_1\star e = e$, these relations can be written in an equivalent way

$$\widehat{\zeta}(x_1\star y_{\underline{s}} - x_1ш y_{\underline{s}})=0 \qquad \text{whenever } s_1\geq 2.$$

They are called *Hoffman standard relations* between multiple zeta values.

Example. *Writing*

$$x_1\star(x_0^{s-1}x_1)=y_1\star y_s = y_1y_s+y_sy_1+y_{s+1}$$

and

$$x_1ш(x_0^{s-1}x_1)=x_1x_0^{s-1}x_1+x_0(x_1ш x_0^{s-2}x_1)=\sum_{\nu=1}^{s}y_\nu y_{s+1-\nu}+y_sy_1,$$

we deduce

$$x_1\star(x_0^{s-1}x_1)-x_1ш(x_0^{s-1}x_1)=x_0^sx_1-\sum_{\nu=1}^{s}x_0^{\nu-1}x_1x_0^{s-\nu}x_1,$$

hence,

$$\zeta(p)=\sum_{\substack{s+s'=p\\ s\geq 2,\ s'\geq 1}}\zeta(s,s').$$

Remark. *As we have seen, the word*

$$x_1\star x_1 - x_1ш x_1 = x_0x_1$$

is convergent, but

$$\widehat{\zeta}(x_1\star x_1 - x_1ш x_1)=\zeta(2)\neq 0.$$

On the other hand, a word like

$$x_1^2 \star x_0x_1 - x_1^2 \shuffle x_0x_1 = x_1x_0^2x_1 + x_0^2x_1^2 - x_1x_0x_1^2 - 2x_0x_1^3$$

is not convergent; also the "convergent part" of this word in the concatenation algebra $\mathfrak{H}$, namely $x_0^2x_1^2 - 2x_0x_1^3 = y_3y_1 - 2y_2y_1^2$, does not belong to the kernel of $\widehat{\zeta}$:

$$\zeta(3,1) = \frac{1}{4}\zeta(4), \quad \zeta(2,1,1) = \zeta(4).$$

We will see (theorem 42) that the convergent part in the shuffle algebra $\mathfrak{H}_{\shuffle}$ of a word of the form $w \shuffle w_0 - w \star w_0$, for $w \in \mathfrak{H}^1$ and $w_0 \in \mathfrak{H}^0$, is always in the kernel of $\widehat{\zeta}$.

Hoffman's operator $d_1 : \mathfrak{H} \to \mathfrak{H}$ is defined by

$$\delta(w) = x_1 \star w - x_1 \shuffle w.$$

For $w \in \mathfrak{H}^0$, we have $d_1(w) \in \mathfrak{H}^0$ and the Hoffman standard relations (39) mean that it satisfies $d_1(\mathfrak{H}^0) \subset \ker \widehat{\zeta}$.

10.2 Ihara–Kaneko–Zagier

There are further similar linear relations between MZV arising from *regularized double shuffle relations* (see [16]).

Recall Corollaries 36 and 38 of Radford and Hoffman concerning the structures of the algebras $\mathfrak{H}_{\shuffle}$ and $\mathfrak{H}_\star$ respectively (we take here for ground field K the field $\mathbf{R}$ of real numbers). From

$$\mathfrak{H}^1_{\shuffle} = \mathfrak{H}^0_{\shuffle}[x_1]_{\shuffle} \quad \text{and} \quad \mathfrak{H}^1_\star = \mathfrak{H}^0_\star[x_1]_\star$$

we deduce that there are two uniquely determined algebra morphisms

$$\widehat{Z}_{\shuffle} : \mathfrak{H}^1_\star \longrightarrow \mathbf{R}[X] \quad \text{and} \quad \widehat{Z}_\star : \mathfrak{H}^1_{\shuffle} \longrightarrow \mathbf{R}[T]$$

which extend $\widehat{\zeta}$ and map x_1 to X and T respectively: for $a_i \in \mathfrak{H}^0$, According to [11], the next result is due to Boutet de Monvel and Zagier (see also [16])

$$\widehat{Z}_{\shuffle}\left(\sum_i a_i \shuffle x_1^{\shuffle i}\right) = \sum_i \widehat{\zeta}(a_i)X^i \quad \text{and} \quad \widehat{Z}_\star\left(\sum_i a_i \star x_1^{\star i}\right) = \sum_i \widehat{\zeta}(a_i)T^i.$$

Proposition 40. *There is a $\mathbf{R}$-linear isomorphism $\varrho : \mathbf{R}[T] \to \mathbf{R}[X]$ which makes commutative the following diagram:*

$$\begin{tikzcd} & \mathbf{R}[X] \\ \mathfrak{H}^1 \arrow[ur, "\widehat{Z}_{\shuffle}"] \arrow[dr, "\widehat{Z}_\star"'] & \\ & \mathbf{R}[T] \arrow[uu, "\varrho"'] \end{tikzcd}$$

The kernel of $\widehat{\zeta}$ is a subset of $\mathfrak{H}^0$ which is an ideal of the algebra $\mathfrak{H}^0_{ш}$ and also of the algebra $\mathfrak{H}^0_\star$. One can deduce from the existence of a bijective map ϱ as in proposition 40 that $\ker\widehat{\zeta}$ generates the same ideal in both algebras $\mathfrak{H}^1_{ш}$ and $\mathfrak{H}^1_\star$.

Explicit formulae for ϱ and its inverse ϱ^{-1} are given by means of the generating series

$$\sum_{\ell\geq 0}\varrho(T^\ell)\frac{t^\ell}{\ell!} = \exp\left(Xt + \sum_{n=2}^{\infty}(-1)^n\frac{\zeta(n)}{n}t^n\right) \tag{41}$$

and

$$\sum_{\ell\geq 0}\varrho^{-1}(X^\ell)\frac{t^\ell}{\ell!} = \exp\left(Tt - \sum_{n=2}^{\infty}(-1)^n\frac{\zeta(n)}{n}t^n\right).$$

For instance,

$$\varrho(T^0)=1,\quad \varrho(T)=X,\quad \varrho(T^2)=X^2+\zeta(2),\quad \varrho(T^3)=X^3+3\zeta(2)X-2\zeta(3),$$

$$\varrho(T^4)=X^4+6\zeta(2)X^2-8\zeta(3)X+\frac{27}{2}\zeta(4).$$

Compare the right hand side in (41) with the formula giving the expansion of the logarithm of Euler Gamma function:

$$\Gamma(1+t)=\exp\left(-\gamma t+\sum_{n=2}^{\infty}(-1)^n\frac{\zeta(n)}{n}t^n\right).$$

The map ϱ may be seen as the differential operator of infinite order

$$\exp\left(\sum_{n=2}^{\infty}(-1)^n\frac{\zeta(n)}{n}\left(\frac{\partial}{\partial T}\right)^n\right)$$

(consider the image of e^{tT}).

10.3 Shuffle regularization of the divergent multiple zeta values, following Ihara, Kaneko and Zagier [16]

Recall that $\mathfrak{H}_{ш}=\mathfrak{H}^0[x_0,x_1]_{ш}$. Denote by $\mathrm{reg}_{ш}$ the $\mathbf{Q}$-linear map $\mathfrak{H}\to\mathfrak{H}^0$ which maps $w\in\mathfrak{H}$ onto its constant term, when w is written as a polynomial in x_0,x_1 in the shuffle algebra $\mathfrak{H}^0[x_0,x_1]_{ш}$. Then $\mathrm{reg}_{ш}$ is a morphism of algebras $\mathfrak{H}_{ш}\to\mathfrak{H}^0_{ш}$. Clearly, for $w\in\mathfrak{H}^0$, we have

$$\mathrm{reg}_{ш}(w)=w.$$

Here are the *regularized double shuffle relations* [16].

Theorem 42. *For $w\in\mathfrak{H}^1$ and $w_0\in\mathfrak{H}^0$,*

$$\mathrm{reg}_{ш}(w ш w_0 - w\star w_0)\in\ker\widehat{\zeta}.$$

Define the *shuffle regularized extension of* $\widehat{\zeta} : \mathfrak{H}^0 \to \mathbf{R}$ as the map $\widehat{\zeta}_{ш} : \mathfrak{H} \to \mathbf{R}$ defined by

$$\widehat{\zeta}_{ш} = \widehat{\zeta} \circ \mathrm{reg}_{ш}.$$

Hence, $\widehat{\zeta}_{ш}$ is nothing else than the composite of $\widehat{Z}_{ш}$ with the specialization map $\mathbf{R}[X] \to \mathbf{R}$ which sends X to 0.

With this definition of $\widehat{\zeta}_{ш}$, Theorem 42 can be written

$$\widehat{\zeta}_{ш}(w ш w_0 - w \star w_0) = 0$$

for any $w \in \mathfrak{H}^1$ and $w_0 \in \mathfrak{H}^0$.

Define the map $D_{ш} : \mathfrak{H} \longrightarrow \mathfrak{H}$ by $D_{ш}(e) = 0$ and

$$D_{ш}(x_{\epsilon_1} x_{\epsilon_2} \cdots x_{\epsilon_p}) = \begin{cases} 0 & \text{if } \epsilon_1 = 0, \\ x_{\epsilon_2} \cdots x_{\epsilon_p} & \text{if } \epsilon_1 = 1. \end{cases}$$

For instance, for $m \geq 0$ and for $w_0 \in \mathfrak{H}^0$,

$$D^i_{ш}(x_1^m w_0) = \begin{cases} x_1^{m-i} w_0 & \text{for } 0 \leq i \leq m, \\ 0 & \text{for } i > m. \end{cases}$$

One checks that $D_{ш}$ is a derivation on the algebra $\mathfrak{H}_{ш}$. Its kernel contains x_0^m and $\mathfrak{H}^0_{ш}$. There is a Taylor expansion for the elements of $\mathfrak{H}^1_{ш}$:

$$u = \sum_{i \geq 0} \frac{1}{i!} \mathrm{reg}_{ш}(D^i_{ш} u) ш y_1^{ш i}, \quad \text{hence,} \quad \widehat{Z}_{ш}(u) = \sum_{i \geq 0} \frac{1}{i!} \mathrm{reg}_{ш}(D^i_{ш}(u)) X^i,$$

and

$$\mathrm{reg}_{ш}(u) = \sum_{i \geq 0} \frac{(-1)^i}{i!} y_1^{ш i} ш D^i_{ш} u.$$

For $w_0 \in \mathfrak{H}^0$ with $w_0 = x_0 w$ (with $w \in \mathfrak{H}^1$) and for $m \geq 0$, we have

$$\mathrm{reg}_{ш}(y_1^m w_0) = (-1)^m x_0 (w ш y_1^m).$$

10.4 Harmonic regularization of the divergent multiple zeta values

There is also a *harmonic regularized extension of* $\widehat{\zeta}$ by means of the star product. Recall that, according to Hoffman's Corollary 38, $\mathfrak{H}_\star = \mathfrak{H}^0_\star[x_0, x_1]_\star$. Denote by $\mathrm{reg}_\star$ the $\mathbf{Q}$-linear map $\mathfrak{H} \to \mathfrak{H}^0$ which maps $w \in \mathfrak{H}$ onto its constant term, when w is written as a polynomial in x_0, x_1 in the harmonic algebra $\mathfrak{H}^0_\star[x_0, x_1]_\star$. Then, $\mathrm{reg}_\star$ is a morphism of algebras $\mathfrak{H}_\star \to \mathfrak{H}^0_\star$. Clearly, for $w \in \mathfrak{H}^0$, we have

$$\mathrm{reg}_\star(w) = w.$$

The map $D_\star : \mathfrak{H}^1 \longrightarrow \mathfrak{H}^1$ defined by $D_\star(e) = 0$ and

$$D_\star(y_{s_1} y_{s_2} \cdots y_{s_k}) = \begin{cases} 0 & \text{if } s_1 = 1, \\ y_{s_2} \cdots y_{s_k} & \text{if } s_1 \geq 2, \end{cases}$$

is a derivation on the algebra $\mathfrak{H}^1_\star$ with kernel $\mathfrak{H}^0_\star$; there is a Taylor expansion for the elements of $\mathfrak{H}^1_\star$:

$$u = \sum_{i \geq 0} \frac{1}{i!} y_1^{\star i} \star \operatorname{reg}_\star(D_\star^i u), \quad \text{hence,} \quad \widehat{Z}_\star(u) = \sum_{i \geq 0} \frac{1}{i!} \operatorname{reg}_\star(D_\star^i u) T^i,$$

and

$$\operatorname{reg}_\star(u) = \sum_{i \geq 0} \frac{(-1)^i}{i!} y_1^{\star i} \star (D_\star^i u).$$

For $w_0 \in \mathfrak{H}^0$ and for $m \geq 0$, we have

$$\operatorname{reg}_\star(y_1^m w_0) = \sum_{i=0}^{m} \frac{(-1)^i}{i!} (y_1^{m-i} w_0) \star y_1^i.$$

11 The Zagier–Broadhurst formula

Theorem 43. *For any $n \geq 1$,*

$$\zeta\big(\{3,1\}_n\big) = \frac{1}{2^{2n}} \zeta\big(\{4\}_n\big).$$

This formula was originally conjectured by D. Zagier and first proved by D. Broadhurst.

The right hand side is known to be

$$\frac{1}{2n+1} \zeta\big(\{2\}_{2n}\big) = 2 \cdot \frac{\pi^{4n}}{(4n+2)!}.$$

These relations can be written

$$\widehat{\zeta}(y_{3,1}^n) = \frac{1}{2^{2n}} \widehat{\zeta}(y_4^n) = \frac{1}{2n+1} \widehat{\zeta}(y_2^{2n}) = 2 \cdot \frac{\pi^{4n}}{(4n+2)!}.$$

11.1 Rational power series

We will use power series without formal justification; the necessary bases for that will be given in § 11.4 below.

We introduce a new map, denoted again with a star * in the exponent[3], defined on the set of series S in $\widehat{\mathfrak{H}}$ which satisfy $(S|e) = 0$, with values in $\widehat{\mathfrak{H}}$, by

$$S^\star = \sum_{n \geq 0} S^n = e + S + S^2 + \cdots \tag{44}$$

[3] There should be no confusion with the notation X^* for the set of words, nor with the star in the exponent for the dual, nor with the harmonic product $\star$ of §8.

The fact that the right hand side of (44) is well defined is a consequence of the assumption $(S|e) = 0$. Notice that $S^\star$ is the unique solution in $\widehat{\mathfrak{H}}$ to the equation

$$(1 - S)S^\star = e,$$

and it is also the unique solution to the equation

$$S^\star(1 - S) = e.$$

A *rational* series is a series in $\widehat{\mathfrak{H}}$ which is obtained by starting with a finite number of letters (this is a restriction only in case where X is infinite) and using only a finite number of *rational operations*, namely addition (25), product (26), multiplication (27) by an element in K and the star (44). The set of rational series over K is a field $\mathrm{Rat}_K(X)$.

For instance, for $x \in X$, the series

$$e + x^2 + x^4 + \cdots + x^{2n} + \cdots = x^\star(-x)^\star$$

is rational, and, for $m \geq 1$, so is the series

$$\sum_{p\geq 0} \varphi_m(p)x^p = (mx)^\star,$$

when $\varphi_m(p) = m^p$. Notice that $\varphi_m(p)$ is also the number of words of weight p on the alphabet with m letters. Series like

$$\sum_{p\geq 0} x^p/p, \quad \sum_{p\geq 0} x^p/p!, \quad \sum_{p\geq 0} x^{2^p}$$

are not rational: if X has a single element, say x, one can prove that *rational series can be identified with elements in $K(x)$ with no poles at $x = 0$.*

For a series S without constant term, i.e. such that $(S|e) = 0$, one defines

$$\exp(S) = \sum_{n=0}^{\infty} \frac{S^n}{n!}\cdot$$

It is easy to check that if T satisfies $(T|e) = 0$, then the series

$$S = \sum_{n=1}^{\infty} \frac{T^n}{n}$$

is well defined and has

$$\exp(S) = T^\star.$$

11.2 Syntaxic identities, following Hoang Ngoc Minh and Petitot

Here, we assume that the next *syntaxic identity* (due to Minh and Petitot) holds. We will prove it in §11.3. The star has been defined in (44) and we use rational series which belong to the algebra of power series considered in §11.4.2.

Lemma 45. *The following identity holds:*

$$(x_0x_1)^\star ш (-x_0x_1)^\star = (-4x_0^2x_1^2)^\star.$$

Taking this identity for granted, we complete the proof of Theorem 43.

Proof of Theorem 43. We will introduce generating series; we work in the ring $\mathbf{R}[[t]]$ of formal power series, where t is a variable. The integration will be according to Chen integrals with respect to ω_0 and ω_1 (not with respect to t!).

Consider the left hand side of the formula of Lemma 45. Integrating between 0 and 1, we deduce, for $k \geq 0$,

$$\int_0^1 (\omega_0\omega_1)^k = \widehat{\zeta}(y_2^k) = \zeta\big(\{2\}_k\big), \quad \int_0^1 (\omega_0^2\omega_1^2)^k = \widehat{\zeta}((y_3y_1)^k) = \zeta\big(\{3,1\}_k\big)$$

and

$$\int_0^1 (\omega_0^3\omega_1)^k = \widehat{\zeta}(y_4^k) = \zeta\big(\{4\}_k\big).$$

We replace x_0 by $t\omega_0$, x_1 by $t\omega_1$. We use Proposition 7

$$\prod_{n\geq 1}\left(1+\frac{t}{n^s}\right) = \sum_{k\geq 0}\zeta\big(\{s\}_k\big)t^k = \sum_{k\geq 0}\widehat{\zeta}\big((y_s)^k\big)t^k$$

with $s=2$ and $s=4$. Replacing t by t^2, $-t^2$ and $-t^4$ respectively, we deduce

$$\sum_{k=0}^\infty t^{2k}\zeta\big(\{2\}_k\big) = \prod_{n\geq 1}\left(1+\frac{t^2}{n^2}\right), \quad \sum_{k=0}^\infty (-t^2)^k\zeta\big(\{2\}_k\big) = \prod_{n\geq 1}\left(1-\frac{t^2}{n^2}\right)$$

and

$$\sum_{k=0}^\infty (-t^4)^k\zeta\big(\{4\}_k\big) = \prod_{n\geq 1}\left(1-\frac{t^4}{n^4}\right).$$

On the other hand, if we replace x_0 by $t\omega_0$ and x_1 by $t\omega_1$ in the left hand side of the formula in Lemma 45, we get

$$\left(\sum_{k=0}^\infty t^{2k}(\omega_0\omega_1)^k\right) ш \left(\sum_{k=0}^\infty (-t^2)^k(\omega_0\omega_1)^k\right) = \sum_{k=0}^\infty (-4t^4)^k(\omega_0^3\omega_1)^k.$$

Thanks to the compatibility of the shuffle product with Chen integrals (Lemma 31), one deduces

$$\left(\int_0^1 \sum_{k=0}^\infty t^{2k}(\omega_0\omega_1)^k\right)\left(\int_0^1 \sum_{k=0}^\infty (-t^2)^k(\omega_0\omega_1)^k\right) = \int_0^1 \sum_{k=0}^\infty (-4t^4)^k(\omega_0^2\omega_1^2)^k.$$

Hence, Theorem 43 follows from

$$\prod_{n\geq 1}\left(1+\frac{t^2}{n^2}\right)\prod_{n\geq 1}\left(1-\frac{t^2}{n^2}\right) = \prod_{n\geq 1}\left(1-\frac{t^4}{n^4}\right).$$

$\square$

In §11.4.3, we will use the same arguments to prove another syntaxic identity, where the star in the exponent is the rational map introduced in §11.1, while the operator $\star$ is the harmonic product of §8.

Lemma 46. *The following identity holds:*

$$y_2^\star \star (-y_2)^\star = (-y_4)^\star.$$

11.3 Shuffle product and automata

There is a description of the shuffle product in terms of automata due to Schutzenberger. Here is a sketch of proof of Lemma 45.

Sketch of proof of Lemma 45. To a series $S^\star$, one associates an automaton, with the following property: the sum of paths going out from the entry vertex is S. As an example the series associated to

$$\begin{tikzcd} {} \arrow[r, Rightarrow, shift right] & \boxed{1} \arrow[l, Rightarrow, shift right] \arrow[r, shift right, "x_0"'] & \boxed{2} \arrow[l, shift right, "x_1"'] \end{tikzcd} \tag{47}$$

is

$$S_1 = e + x_0x_1 + (x_0x_1)^2 + \cdots + (x_0x_1)^n + \cdots = (x_0x_1)^\star$$

and similarly, the series associated to

$$\begin{tikzcd} {} \arrow[r, Rightarrow, shift right] & \boxed{A} \arrow[l, Rightarrow, shift right] \arrow[r, shift right, "-x_0"'] & \boxed{B} \arrow[l, shift right, "x_1"'] \end{tikzcd} \tag{48}$$

is

$$S_A = e - x_0x_1 + (x_0x_1)^2 + \cdots + (-x_0x_1)^n + \cdots = (-x_0x_1)^\star.$$

The cartesian product of these two automata is the following:

$$\begin{tikzcd} {} \arrow[r, Rightarrow, shift right] & \boxed{1A} \arrow[l, Rightarrow, shift right] \arrow[r, shift right, "x_0"'] \arrow[d, shift right, "-x_0"'] & \boxed{2A} \arrow[l, shift right, "x_1"'] \arrow[d, shift right, "-x_0"'] \\ & \boxed{1B} \arrow[r, shift right, "x_0"'] \arrow[u, shift right, "x_1"'] & \boxed{2B} \arrow[l, shift right, "x_1"'] \arrow[u, shift right, "x_1"'] \end{tikzcd} \tag{49}$$

Let S_{1A} be the series associated with this automaton (49). One computes it by solving a system of linear (noncommutative) equations as follows. Define also S_{1B}, S_{2A} and S_{2B} as the series associated with the paths going out from the corresponding vertex. Then

$$\begin{aligned} S_{1A} &= e - x_0S_{1B} + x_0S_{2A}, \\ S_{1B} &= x_1S_{1A} + x_0S_{2B}, \\ S_{2A} &= x_1S_{1A} - x_0S_{2B}, \\ S_{2B} &= x_1S_{1B} + x_1S_{2A}. \end{aligned}$$

The rule is as follows: if Σ is the sum associated with a vertex (also denoted by Σ) with oriented edges $\xi_i : \Sigma \to \Sigma_i \quad (1 \le i \le m)$, then

$$\Sigma = x_1\Sigma_1 + \cdots + x_m\Sigma_m,$$

and $x_i\Sigma_i$ is replaced by e for the entry vertex.

In the present situation, one deduces

$$S_{1A} = e - x_0(S_{1B} - S_{2A}), \quad S_{1B} - S_{2A} = -2x_0S_{2B},$$

$$S_{2B} = x_1(S_{1B} + S_{2A}), \quad S_{1B} + S_{2A} = 2x_1S_{1A}$$

and therefore,

$$S_{1A} = e + 4x_0^2x_1^2S_{1A},$$

which completes the proof of Lemma 45, since the series associated with the automaton (49) is the shuffle product of the series associated with the automata (47) and (48).

A proof that the cartesian product of two automata recognizes the shuffle of the two languages which are recognized by each factor can be found in [13], pp. 19–20.

□

11.4 Formal power series, rational series, symmetric and quasi-symmetric series

11.4.1 Dual

Let K be a field and X a set. Denote by E_X the free vector space on X (see §6.1). By definition of E_X as a solution of the universal problem 24, there is a one to one map j between the set of linear maps $E_X \to K$ (which is the *dual* of the vector space E_X which we denote by E^*) and the set K^X of maps $X \to K$. Notice that there is no restriction on the support of these maps (such a condition on the support makes a difference only when X is infinite: for a finite set X, we have $K^X = K^{(X)}$). There is a natural structure of K–vector space on the dual E^* and there is also a natural structure of K–vector space on K^X, and the bijective map j is an isomorphism of K–vector spaces.

Consider next the free commutative algebra $K[X]$ on X (see §6.1). By definition of $K[X]$ as a solution of universal problem 24, there is a one to one map ι between the set of morphisms of algebras $K[X] \to K$ (which is also called the *dual* of the algebra $K[X]$) and the set K^X of maps $X \to K$. There is a natural structure of K–algebra on the dual and there is also a natural structure of K–algebra on K^X, and the bijective map ι is an isomorphism of K–algebras.

Consider a non–empty set X. According to the definition of $K\langle X\rangle$ as a solution of a universal problem 24, for each K-algebra A, the map $f \to \underline{f}$ defines a bijection between A^X and the set of morphisms of K-algebras $K\langle X\rangle \to A$. This is one dual of $\mathfrak{H}$, considered as an algebra, but this is not the one we are going to consider: we are interested with the dual $\widehat{\mathfrak{H}}$ of $\mathfrak{H}$ as a K–vector space.

11.4.2 The Algebra $\widehat{\mathfrak{H}} = K\langle\langle X\rangle\rangle$ of formal power series

Let X be a non–empty set. We are interested in the dual of the free K–algebra $K\langle X\rangle$ as a K–vector space, namely the K–vector space $\mathrm{Hom}_K\big(K\langle X\rangle, K\big)$ of K-linear maps $K\langle X\rangle \to K$. We will see that there is a natural structure of algebra on this dual, which is the algebra $\widehat{\mathfrak{H}} = K\langle\langle X\rangle\rangle$ *of formal power series on* X.

The underlying set of the algebra $K\langle\langle X\rangle\rangle$ is the set K^{X^*} of maps $X^* \to K$. Here, there is no restriction on the support. For such a map S, write $(S|w)$ for the image of $w \in X^*$ in K and write also

$$S = \sum_{w\in X^*} (S|w)w.$$

On this set K^{X^*}, the addition is defined by (25) and the multiplication is again Cauchy product (26). Further, for $\lambda \in K$ and $S \in K^{(X^*)}$, define $\lambda S \in K^{X^*}$ by (27). With these laws, one checks that the set K^{X^*} becomes a K-algebra which we denote by either $K\langle\langle X\rangle\rangle$ or $\widehat{\mathfrak{H}}$.

To a formal power series S we associate a K-linear map:

$$\begin{array}{ccc} K\langle X\rangle & \longrightarrow & K \\ P & \longmapsto & \displaystyle\sum_{w\in X^*} (S|w)(P|w). \end{array}$$

Notice that the sum is finite, since $P \in K\langle X\rangle$ has finite support.

Since X^* is a basis of the K-vector space $K\langle X\rangle$, a linear map $f \in \mathrm{Hom}_K\big(K\langle X\rangle, K\big)$ is uniquely determined by its values $(f|w)$ on the set X^*. Hence, the map

$$\begin{array}{ccc} \mathrm{Hom}_K\big(K\langle X\rangle, K\big) & \longrightarrow & \widehat{\mathfrak{H}} \\ f & \longmapsto & \sum_{w\in X^*} (f|w)w \end{array}$$

is an isomorphism of vector spaces between the dual [4] $\mathrm{Hom}_K\big(K\langle X\rangle, K\big)$ of $\mathfrak{H} = K\langle X\rangle$ and $\widehat{\mathfrak{H}}$.

11.4.3 Symmetric Series, Quasi-Symmetric Series and Harmonic product

Denote by $\underline{t} = (t_1, t_2, \dots)$ a sequence of commutative variables. To $\underline{s} = (s_1, \dots, s_k)$, where each s_j is an integer ≥ 1, associate the series

$$\mathrm{S}_{\underline{s}}(\underline{t}) = \sum_{\substack{n_1\geq 1,\dots,n_k\geq 1\\ n_1,\dots,n_k \text{pairwise distinct}}} t_{n_1}^{s_1}\cdots t_{n_k}^{s_k}.$$

[4]This is the classical dual; there are other notions of dual, in particular the "graduate dual", which in the present case is isomorphic to $\mathfrak{H}$, and the "restricted dual", which is the field $\mathrm{Rat}_K(X)$ of series which are "rational" (see § 1.3).

The space of power series spanned by these $\mathrm{S}_{\underline{s}}$ is denoted by Sym and its elements are called *symmetric series*. A basis of Sym is given by the series $\mathrm{S}_{\underline{s}}$ with $s_1 \geq s_2 \geq \cdots \geq s_k$ and $k \geq 0$.

A *quasi-symmetric series* is an element of the algebra QSym spanned by the series

$$\mathrm{QS}_{\underline{s}}(\underline{t}) = \sum_{n_1 > \cdots > n_k \geq 1} t_{n_1}^{s_1} \cdots t_{n_k}^{s_k},$$

where $\underline{s}$ ranges over the set of tuples $(s_1, \ldots, s_k)$ with $k \geq 0$ and $s_j \geq 1$ for $1 \leq j \leq k$. Notice that, for $\underline{s} = (s_1, \ldots, s_k)$ of length k,

$$\mathrm{S}_{\underline{s}} = \sum_{\tau \in \mathfrak{S}_k} \mathrm{QS}_{\underline{s}^\tau},$$

where $\mathfrak{S}_k$ is the symmetric group on k elements and $\underline{s}^\tau = (s_{\tau(1)}, \ldots, s_{\tau(k)})$. Hence, any symmetric series is also quasi-symmetric. Therefore, Sym is a subalgebra of QSym.

Notice that these algebras are commutative.

Proposition 50. *The K-linear map $\phi : \mathfrak{H}^1 \to \mathrm{QSym}$ defined by $y_{\underline{s}} \mapsto \mathrm{QS}_{\underline{s}}$ is an isomorphism of K-algebras from the harmonic algebra $\mathfrak{H}^1_\star$ to* QSym.

In other terms, we can write (33) as follows:

$$\mathrm{QS}_{\underline{s}}(\underline{t})\ \mathrm{QS}_{\underline{s}'}(\underline{t}) = \sum_{\underline{s}''} \mathrm{QS}_{\underline{s}''}(\underline{t}),$$

which means

$$\sum_{n_1 > \cdots > n_k \geq 1} t_{n_1}^{s_1} \cdots t_{n_k}^{s_k} \sum_{n'_1 > \cdots > n'_k \geq 1} t_{n'_1}^{s'_1} \cdots t_{n'_k}^{s'_k} = \sum_{\underline{s}''} \sum_{n''_1 > \cdots > n''_k \geq 1} t_{n''_1}^{s''_1} \cdots t_{n''_k}^{s''_k},$$

where $\underline{s}''$ is the same as for the definition of the harmonic product in §8. The star (stuffle) law gives an explicit way of writing the product of two quasi-symmetric series as a sum of quasi-symmetric series.

Let QSym^0 be the subspace of QSym spanned by the $\mathrm{QS}_{\underline{s}}(\underline{t})$ for which $s_1 \geq 2$. The restriction of ϕ to $\mathfrak{H}^0$ gives an isomorphism of K-algebra from $\mathfrak{H}^0$ to QSym^0. The specialization $t_n \to 1/n$ for $n \geq 1$, restricted to QSym^0, maps $\mathrm{QS}_{\underline{s}}$ onto $\zeta(\underline{s})$. Hence, we have a commutative diagram:

$$\begin{array}{ccc}
\mathfrak{H} & & \\
\bigcup & & \\
\mathfrak{H}^1 & \xrightarrow[\sim]{\phi} & \mathrm{QSym} \\
\bigcup & & \bigcup \\
\mathfrak{H}^0 & \xrightarrow{\sim} & \mathrm{QSym}^0 \\
\hat{\zeta}\Big\downarrow & \swarrow & \\
\mathbf{R} & &
\end{array}
\qquad\qquad
\begin{array}{ccc}
y_{\underline{s}} & \longmapsto & \mathrm{QS}_{\underline{s}}(\underline{t}) \\
\downarrow & \swarrow & \\
\zeta(\underline{s}) & &
\end{array}$$

Proof of Lemma 46. From the definition of ϕ in Proposition 50, we have

$$\phi(y_2^\star) = \sum_{k=0}^{\infty} \sum_{n_1 > \cdots > n_k \geq 1} t_{n_1}^2 \cdots t_{n_k}^2,$$

$$\phi\bigl((-y_2)^\star\bigr) = \sum_{k=0}^{\infty} (-1)^k \sum_{n_1 > \cdots > n_k \geq 1} t_{n_1}^2 \cdots t_{n_k}^2$$

and

$$\phi\bigl((-y_4)^\star\bigr) = (-1)^k \sum_{n_1 > \cdots > n_k \geq 1} t_{n_1}^4 \cdots t_{n_k}^4.$$

Hence, from the identity

$$\prod_{n=1}^{\infty} (1 + t_n t) = \sum_{k=0}^{\infty} t^k \sum_{n_1 > \cdots > n_k \geq 1} t_{n_1} \cdots t_{n_k},$$

one deduces

$$\phi(y_2^\star) = \prod_{n=1}^{\infty} (1+t_n^2), \quad \phi\bigl((-y_2)^\star\bigr) = \prod_{n=1}^{\infty} (1-t_n^2) \quad \text{and} \quad \phi\bigl((-y_4)^\star\bigr) = \prod_{n=1}^{\infty} (1-t_n^4),$$

which implies Lemma 46.

□

We give a short list of references, starting with internet web sites.

References

[1] J. Borwein, P. Lisonek and P. Irvine, *EZ-Face*, online computer program.
`http://oldweb.cecm.sfu.ca/projects/EZFace/`

[2] J. Borwein and W. Zudilin, *Multiple Zeta Values*, Math Honours, 2011.
`http://carma.newcastle.edu.au/MZVs/mzv.pdf`

[3] Michael Hoffman, *Multiple Zeta Values and Euler Sums*
`http://www.usna.edu/Users/math/meh/biblio.html`

[4] N.J. Sloane, *The On-Line Encyclopedia of Integer Sequences.*
`http://oeis.org/`.

[5] Wadim Zudilin, *Zeta values on the Web*
`http://wain.mi.ras.ru/zw/index.html`

[6] Boas, R. P. Jr., *Entire functions.* New York: Academic Press Inc., 1954.

[7] Brown, F., *On the decomposition of motivic multiple zeta values.* `arXiv:1102.1310`

[8] — , *Mixed Tate motives over* **Z**. Annals of Mathematics **175** (2012), 949–976. `arXiv:1102.1312`

[9] — , *Irrationality proofs for zeta values, moduli spaces and dinner parties.* `arXiv:1412.6508`

[10] J. I. Burgos Gil and J. Fresán, *Multiple zeta values: from numbers to motives.* Clay Mathematics Proceedings, to appear.

[11] P. Cartier, *Fonctions polylogarithmes, nombres polyzetas et groupes pro–unipotents*, Séminaire Bourbaki t. **42** 53è année, 2000–2001, Exposé No°885, Mars 2001. Astérisque No.**282** (2002), 137–173. `http://www.numdam.org/numdam-bin/fitem?id=SB_2000-2001__43__137_0`

[12] Dummit, D. S. and Foote, R. M, *Abstract algebra.* Hoboken, NJ: John Wiley & Sons Inc., third ed., 2004.

[13] Eilenberg, S., *Automata, Languages and Machines*, vol. A, Academic Press, 1974.

[14] S. Fischler – *Irrationalité de valeurs de zêta, (d'après Apéry, Rivoal, ...)*, Sém. Bourbaki 2002-2003, N° 910 (Novembre 2002). Astérisque **294** (2004), 27-62.
http://www.math.u-psud.fr/∼fischler/publi.html

[15] Hardy, G. H. and Wright, E. M., *An introduction to the theory of numbers.* Oxford: Oxford University Press, sixth ed., 2008. Revised by D. R. Heath-Brown and J. H. Silverman, With a foreword by Andrew Wiles.

[16] K. Ihara, M. Kaneko and D. Zagier , *Derivation and double shuffle relations for multiple zeta values.* Compositio Math. **142** (2006) 307–338.

[17] Lang, S., *Algebra*, vol. 211 of *Graduate Texts in Mathematics.* New York: Springer-Verlag, third ed., 2002.

[18] Serre, J.-P., *A course in arithmetic.* New York: Springer-Verlag, 1973. Translated from the French, Graduate Texts in Mathematics, No. 7.

[19] Waldschmidt, M., *Valeurs zêta multiples. Une introduction*, J. Théor. Nombres Bordeaux 12 (2000), 581–595. `http://www.numdam.org/numdam-bin/fitem?id=JTNB_2000__12_2_581_0`

[20] — , *Multiple polylogarithms: an introduction*, in Number Theory and Discrete Mathematics (Chandigarh, 2000), A. K. Agarwal et. al. (eds.), Birkhäuser, Basel (2002), 1–12.
`http://hal.archives-ouvertes.fr/hal-00416166/fr/`

[21] Zagier, D., *Evaluation of the multiple zeta value* $\zeta(2,\dots,2,3,2,\dots,2)$, Annals of Mathematics **175** (2012), 977–1000.
http://dx.doi.org/10.4007/annals.2012.175.2.11

[22] Zhonghua Li, *Another proof of Zagier's evaluation formula of the multiple zeta values* $\zeta(2,\dots,2,3,2,\dots,2)$.
http://arxiv.org/abs/1204.2060

[23] Zudilin, W., *Algebraic relations for multiple zeta values*, Russian Math. Surveys **58**:1 (2003), 1—29.

This text can be downloaded on the internet at URL
http://www.math.jussieu.fr/~miw/articles/pdf/MZV2011IMSc.pdf

Michel WALDSCHMIDT
Université Pierre et Marie Curie-Paris 6
Institut de Mathématiques de Jussieu IMJ UMR 7586
Théorie des Nombres Case Courrier 247
4 Place Jussieu
F–75252 Paris Cedex 05 France
e-mail: miw@math.jussieu.fr
URL: http://www.math.jussieu.fr/~miw/