Quasi-shuffle algebras and stuffle regularized multiple Eisenstein series

NUPACE guided independent study project, Nagoya University

Rasmus Peter Skadborg

January 23, 2024

Abstract

In this report, we state the definition and prove some basic properties of quasi-shuffle algebras, with the shuffle and stuffle products serving as our main examples. We give examples of applications to multiple zeta values (MZV), proving that the multiple zeta map gives a homomorphism from certain shuffle and stuffle algebras $\mathfrak{H}^0_{ш}$ and $\mathfrak{H}^0_*$ to the $\mathbb{Q}$-algebra $\mathcal{Z} \subset \mathbb{R}$ of multiple zeta values. From this, one obtains a family of $\mathbb{Q}$-linear relations among MZV, known as the finite double-shuffle relations. Further, these homomorphism properties give a natural way to extend MZV to shuffle and stuffle regularized MZV, and from this extension one obtains a larger family of linear relations among MZV known as the extended double-shuffle relations.

We also give applications to multiple Eisenstein series (MES) following [Bac4]. By using convolutions with respect to the stuffle-product, MES can be shown to have q-series whose coefficients are closely related to multiple zeta values and their q-analogues. As a byproduct of the proof of this fact, and using some properties of the multitangent functions studied by [Bou], one gets a natural way of extending MES to a stuffle-regularized version.

Contents

1 Introduction

The Riemann zeta function

$$\zeta(s) = \sum_{m>0} \frac{1}{m^s} \qquad (\mathrm{Re}(s) > 1)$$

is one of the most important objects in modern number theory. Famously, the zeroes of (the analytic continuation of) $\zeta(s)$ are related to the distribution of prime numbers. Of separate interest to the analytic aspects of the zeta function are its special values, such as $\zeta(k)$ for integers $k \geq 2$. These special numbers tend to show up in various places ranging from physics to the theory of modular forms. Euler famously proved that

$$\zeta(2k) = \frac{(-1)^{k-1} B_{2k}}{2(2k)!}(2\pi)^{2k},$$

where B_n are the Bernoulli numbers defined by the generating function

$$\sum_{n\geq 0} B_n \frac{x^n}{n!} = \frac{x}{e^x - 1} = 1 - \frac{1}{2}x + \frac{1}{6}\frac{x^2}{2!} - \frac{1}{30}\frac{x^4}{4!} + \frac{1}{42}\frac{x^6}{6!} + \dots$$

Thus for example, $\zeta(2) = \frac{\pi^2}{6}, \zeta(4) = \frac{\pi^4}{90}$ and $\zeta(6) = \frac{\pi^6}{945}$.

Since π is transcendental, it follows from Euler's formula that all even zeta values $\zeta(2k)$ are transcendental. By comparison, much less is known about the odd zeta values $\zeta(2k+1)$. Apéry proved in 1978 that $\zeta(3)$ is irrational, but no other particular odd zeta values are known to be irrational, let alone transcendental. However, it is conjectured that the numbers

$$1, \pi^2, \zeta(3), \zeta(5), \zeta(7), \dots$$

are algebraically independent over the rational numbers (though a proof of this conjecture seems far out of reach at the moment). In other words, we don't expect any algebraic combination of $\zeta(k_1), \dots, \zeta(k_n)$ to be rational when $k_1, \dots, k_n \geq 2$ are distinct integers, at most one of which is even. The simplest such (non-linear) algebraic combination is the product of two zeta-values:

$$\zeta(k_1)\zeta(k_2) = \sum_{m_1, m_2 > 0} \frac{1}{m_1^{k_1} m_2^{k_2}} = \left[\sum_{m_1 > m_2 > 0} + \sum_{m_2 > m_1 > 0} + \sum_{m_1 = m_2 > 0} \right] \frac{1}{m_1^{k_1} m_2^{k_2}}.$$

If one defines the double zeta-values $\zeta(k_1, k_2) := \sum_{m_1 > m_2 > 0} \frac{1}{m_1^{k_1} m_2^{k_2}}$, then the above gives the identity

$$\zeta(k_1)\zeta(k_2) = \zeta(k_1, k_2) + \zeta(k_2, k_1) + \zeta(k_1 + k_2). \tag{1}$$

In Section 2, we will define a certain $\mathbb{Q}$-algebra where the following equation resembling (1) holds:

$$z_{k_1} * z_{k_2} = z_{k_1} z_{k_2} + z_{k_2} z_{k_1} + z_{k_1 + k_2}.$$

The product $*$ is called the stuffle-product (or harmonic product) and is one of the main examples of a quasi-shuffle product, which are the subject of study in this report.

A different way of writing $\zeta(k_1)\zeta(k_2)$ comes from the partial fraction decomposition

$$\frac{1}{x^{k_1} y^{k_2}} = \sum_{j=1}^{k_1 + k_2 - 1} \left(\frac{\binom{j-1}{k_1 - 1}}{(x+y)^j y^{k_1 + k_2 - j}} + \frac{\binom{j-1}{k_2 - 1}}{(x+y)^j x^{k_1 + k_2 - j}} \right).$$

Noting that the $j = 1$ term vanishes when $k_1, k_2 \geq 2$, this gives

$$\begin{aligned}
\zeta(k_1)\zeta(k_2) &= \sum_{j=2}^{k_1+k_2-1} \sum_{m_1,m_2>0} \left(\frac{\binom{j-1}{k_1-1}}{(m_1+m_2)^j m_2^{k_1+k_2-j}} + \frac{\binom{j-1}{k_2-1}}{(m_1+m_2)^j m_1^{k_1+k_2-j}} \right) \\
&= \sum_{j=2}^{k_1+k_2-1} \left(\binom{j-1}{k_1-1} \sum_{n_1>n_2>0} \frac{1}{n_1^j n_2^{k_1+k_2-j}} + \binom{j-1}{k_2-1} \sum_{n_1>n_2>0} \frac{1}{n_1^j n_2^{k_1+k_2-j}} \right) \\
&= \sum_{j=2}^{k_1+k_2-1} \left(\binom{j-1}{k_1-1} + \binom{j-1}{k_2-1} \right) \zeta(j, k_1+k_2-j). \qquad (2)
\end{aligned}$$

Again, we will define a certain algebra in Section 2 where a similar equation holds, namely

$$x^{k_1-1}y \sqcup\!\sqcup\, x^{k_2-1}y = \sum_{j=1}^{k_1+k_2-1} \left(\binom{j-1}{k_1-1} + \binom{j-1}{k_2-1} \right) x^{j-1}yx^{k_1+k_2-j-1}y \qquad (k_1, k_2 \geq 1). \quad (3)$$

The product $\sqcup\!\sqcup$ is known as the shuffle product, and is another example of a quasi-shuffle product. Combining (1) and (3), one gets (a special case of) what is known as the double-shuffle relations:

$$\zeta(k_1,k_2)+\zeta(k_2,k_1)+\zeta(k_1+k_2) = \zeta(k_1)\zeta(k_2) = \sum_{j=2}^{k_1+k_2-1} \left(\binom{j-1}{k_1-1} + \binom{j-1}{k_2-1} \right) \zeta(j, k_1+k_2-j).$$

For example

$$\zeta(2,3) + \zeta(3,2) + \zeta(5) = \zeta(2)\zeta(3) = \zeta(2,3) + 3\zeta(3,2) + 6\zeta(4,1).$$

In Section 2, we will prove the double-shuffle relations in full generality using the quasi-shuffle products $*$ and $\sqcup\!\sqcup$. We will also use the theory of quasi-shuffle products to make sense of divergent multiple zeta values such as "$\zeta(1)$".

In Section 3, we will study several other families of objects that satisfy identities similar to (1), mainly multiple Eisenstein series, which are generalizations of Eisenstein series, the well-known modular forms. Using the theory of quasi-shuffle products built up in Section 2 (especially of the stuffle-product), we will be able to prove that the Fourier series expansions of multiple Eisenstein series are closely related to multiple zeta values. As a byproduct of the proof, we will find a "natural" way of making sense of divergent multiple Eisenstein series.

2 Quasi-shuffle products and multiple zeta values

Let $\Bbbk$ be a field, and let A (the "alphabet") be a countable set of symbols. We denote by $\Bbbk A$ the $\Bbbk$-vector space having a basis-element for each symbol in A, i.e.

$$\Bbbk A := \bigoplus_{a \in A} \Bbbk a.$$

Let $A^* = \{a_1 \cdots a_r \mid r \geq 0, a_1, \ldots, a_r \in A\}$ be the set of words in the alphabet A. For a word $w = a_1 \cdots a_r \in A^*$, we let $\ell(w) = r$ denote the length of w. We denote by $\Bbbk\langle A\rangle$ the **noncommutative polynomial algebra**, which as a vector space has A^* as a basis, i.e.

$$\Bbbk\langle A\rangle := \bigoplus_{w \in A^*} \Bbbk w.$$

This is a $\Bbbk$-algebra with the product given on basis-elements by concatenation of words. Note that the empty word is the unit for this algebra, and so we will simply denote it by 1, and we will identify $\Bbbk$ with the span of the empty word in $\Bbbk\langle A\rangle$.

Suppose $\Bbbk A$ is equipped with an associative, commutative, $\Bbbk$-bilinear product $\diamond$. This then induces a new product on $\Bbbk\langle A\rangle$ defined as follows:

Definition 2.1. Given a commutative, associative, $\Bbbk$-bilinear product $\diamond$ on $\Bbbk A$, we define the **quasi-shuffle product** $*_\diamond$ on $\Bbbk\langle A\rangle$ as follows: On the generators (words) it is given by

$$1 *_\diamond w = w *_\diamond 1 = w,$$

and

$$aw *_\diamond bv = a(w *_\diamond bv) + b(aw *_\diamond v) + (a \diamond b)(w *_\diamond v), \tag{4}$$

for arbitrary words $w, v \in A^*$ and symbols $a, b \in A$. $\alpha *_\diamond \beta$ is then given for general $\alpha, \beta \in \Bbbk\langle A\rangle$ by extending the above bilinearly.

Lemma 2.2. *$*_\diamond$ as defined above is commutative and associative whenever $\diamond$ is. Thus $(\Bbbk\langle A\rangle, *_\diamond)$ is a commutative $\Bbbk$-algebra.*

Proof. Let us start with commutativity. It suffices to show that $w *_\diamond v = v *_\diamond w$ for words $w, v \in A^*$. The proof is by induction in the total length of the words $\ell(w) + \ell(v)$. Note that the commutativity is obvious when either w or v is the empty word, so this gives us the base-case for free, and for the inductive step we may assume that $w = aw', v = bv'$ for some symbols $a, b \in A$ and words $w, v \in A^*$. Then using the inductive hypothesis, as well as the commutativity of $\diamond$, we get

$$\begin{aligned} w *_\diamond v - v *_\diamond w = \quad & a(w' *_\diamond bv') + b(aw' *_\diamond v') + (a \diamond b)(w' *_\diamond v') \\ & -b(v' *_\diamond aw') - a(bv' *_\diamond w') - (b \diamond a)(v' *_\diamond w') \\ = \quad & a(w' *_\diamond bv') + b(aw' *_\diamond v') + (a \diamond b)(w' *_\diamond v') \\ & -b(aw'' *_\diamond v') - a(w' *_\diamond bv') - (a \diamond b)(w' *_\diamond v') \\ = \quad & 0. \end{aligned}$$

The proof of associativity is similar, though the calculation gets slightly more cumbersome. We show $w *_\diamond (v *_\diamond u) = (w *_\diamond v) *_\diamond u$ for words $w, v, u \in A^*$ by induction in $\ell(w) + \ell(v) + \ell(u)$. Again this is trivial when either of the three words is empty, so we get the base case for free and can assume for the inductive step that $w = aw', v = bv', u = cu'$ for some symbols $a, b, c \in A$

and words $w, v, u \in A^*$. Then

$$\begin{aligned}
w *_\diamond (v *_\diamond u) - (w *_\diamond v) *_\diamond u = {} & aw' *_\diamond (b(v' *_\diamond cu') + c(bv' *_\diamond u') + (b \diamond c)(v' *_\diamond u')) \\
& -(a(w' *_\diamond bv') + b(aw' *_\diamond v') + (a \diamond b)(w' *_\diamond v')) *_\diamond cu' \\
= {} & a(w' *_\diamond b(v' *_\diamond cu')) + b(aw' *_\diamond (v' *_\diamond cu')) + (a \diamond b)(w' *_\diamond (v' *_\diamond cu')) \\
& +a(w' *_\diamond c(bv' *_\diamond u')) + c(aw' *_\diamond (bv' *_\diamond u')) + (a \diamond c)(w' *_\diamond (bv' *_\diamond u')) \\
& +a(w' *_\diamond (b \diamond c)(v' *_\diamond u')) + (b \diamond c)(aw' *_\diamond (v' *_\diamond u')) + (a \diamond (b \diamond c))(w' *_\diamond (v' *_\diamond u')) \\
& -a((w' *_\diamond bv') *_\diamond cu') - c(a(w' *_\diamond bv') *_\diamond u') - (a \diamond c)((w' *_\diamond bv') *_\diamond u') \\
& -b((aw' *_\diamond v') *_\diamond cu') - c(b(aw' *_\diamond v') *_\diamond u') - (b \diamond c)((aw' *_\diamond v') *_\diamond u') \\
& -(a \diamond b)((w' *_\diamond v') *_\diamond cu') - c((a \diamond b)(w' *_\diamond v') *_\diamond u') - ((a \diamond b) \diamond c)((w' *_\diamond v') *_\diamond u') \\
= {} & a(w' *_\diamond b(v' *_\diamond cu')) \\
& +a(w' *_\diamond c(bv' *_\diamond u')) + c(aw' *_\diamond (bv' *_\diamond u')) \\
& +a(w' *_\diamond (b \diamond c)(v' *_\diamond u')) \\
& -a((w' *_\diamond bv') *_\diamond cu') - c(a(w' *_\diamond bv') *_\diamond u') \\
& -c(b(aw' *_\diamond v') *_\diamond u') \\
& -c((a \diamond b)(w' *_\diamond v') *_\diamond u') \\
= {} & a(w' *_\diamond (bv' *_\diamond cu')) + c(aw' *_\diamond (bv' *_\diamond u')) \\
& -a((w' *_\diamond bv') *_\diamond cu') - c((aw' *_\diamond bv') *_\diamond u') \\
= {} & 0.
\end{aligned}$$

In the third equality above, we cancel several terms using the induction hypothesis along with the associativity of $\diamond$. The next equality is by recombining the $+a$-terms and the $-c$-terms using the recursive definition of $*_\diamond$. The cancellation in the final equality is then again by the inductive hypothesis. ☐

Example 2.3. Our main examples of quasi-shuffle products will be the following.

(i) Let $A = \{x, y\}$ and $\diamond = 0$. The associated quasi-shuffle product $*_\diamond$, which is typically denoted $ш$ is then given recursively by

$$1 ш w = w ш 1 = w, \qquad aw ш bv = a(w ш bv) + b(aw ш v),$$

for words $w, v \in A^*$ and symbols $a, b \in A$. One can think of a word $w \in A^*$ as a "deck of cards" with an x or a y written on each card. $w ш v$ is then the sum of all possible (riffle) shuffles of the two decks w and v. For this reason, $ш$ is also known as the **shuffle-product**. Quasi-shuffle products are then so named because they generalize the shuffle-product with the extra $\diamond$-term in the recursion. Note also that this card-shuffling interpretation of $ш$ gives a combinatorial proof of the equation (3).

(ii) Let $A = \{z_k \mid k \geq 1\}$, and define $\diamond$ on the symbols by $z_k \diamond z_\ell = z_{k+\ell}$. The associated quasi-shuffle product $*_\diamond$ is known as the **stuffle-product** and is simply denoted by $*$. Thus

$$w * 1 = 1 * w = w, \qquad z_k w * z_\ell v = z_k(w * z_\ell v) + z_\ell(z_k * v) + z_{k+\ell}(w * v),$$

for $k, \ell \geq 1$ and $w, v \in A^*$.

Let us remark that quasi-shuffle algebras are in fact bialgebras. Define a linear map $\Delta : \Bbbk\langle A\rangle \to \Bbbk\langle A\rangle \otimes \Bbbk\langle A\rangle$ (called the deconcatenation coproduct) on words $w \in A^*$ by

$$\Delta(w) = \sum_{uv=w} u \otimes v,$$

where the sum runs over all ways of writing w as the concatenation of two subwords u, v (either of which is allowed to be the empty word). This in turns the quasi-shuffle algebra $(\Bbbk\langle A\rangle, *_\diamond)$ into a bialgebra with coproduct Δ and counit

$$\varepsilon(w) = \begin{cases} 1 & w = 1, \\ 0 & \ell(w) \geq 1. \end{cases}$$

In fact, there is even a unique antipodal map turning $(\Bbbk\langle A\rangle, *_\diamond)$ into a Hopf-algebra (though we will not be making much use of this fact), see [Hof, HI] for details.

2.1 Double-shuffle relations for multiple zeta-values

One of the main applications of quasi-shuffle products is to find relations among multiple zeta values (MZV). MZV are numbers of the form

$$\zeta(k_1, \dots, k_r) = \sum_{m_1 > \cdots > m_r > 0} \frac{1}{m_1^{k_1} \cdots m_r^{k_r}},$$

where $k_1, \dots, k_r$ are positive integers with $k_1 \geq 2$. The condition $k_1 \geq 2$ is to ensure convergence. Indeed, by comparing with the harmonic series, one sees that the sum above diverges when $k_1 = 1$, whereas for $k_1, \dots, k_r$ real with $k_j \geq 1$ and $k_1 \geq 1 + \varepsilon > 1$

$$\begin{aligned}
\left| \sum_{m_1 > \cdots > m_r > 0} \frac{1}{m_1^{k_1} \cdots m_r^{k_r}} \right| &\leq \sum_{m_1 > \cdots > m_r > 0} \frac{1}{m_1^{1+\varepsilon} m_2 \cdots m_r} \\
&= \sum_{m>0} \frac{1}{m^{1+\varepsilon}} \sum_{m > m_2 > \cdots > m_r > 0} \frac{1}{m_2 \cdots m_r} \\
&\leq \sum_{m>0} \frac{1}{m^{1+\varepsilon}} \sum_{\substack{m > m_2 > 0 \\ \vdots \\ m > m_r > 0}} \frac{1}{m_2 \cdots m_r} \\
&= \sum_{m>0} \frac{1}{m^{1+\varepsilon}} \left(\sum_{n=1}^{m-1} \frac{1}{n} \right)^{r-1}.
\end{aligned} \tag{5}$$

Upon comparing with an integral, one then sees that $\sum_{n=1}^{m-1} \frac{1}{n} = O(\log(m)) = O(m^{\frac{\varepsilon/2}{r-1}})$, so that the sum converges absolutely.

Let us introduce a bit of terminology. An **index** is a tuple $\mathbf{k} = (k_1, \dots, k_r)$ of positive integers, allowing the empty index $\mathbf{k} = \varnothing$. We say $\mathbf{k}$ is **admissible** if $\mathbf{k} = \varnothing$ or $k_1 \geq 2$, so that $\zeta(\mathbf{k})$ makes sense (by convention, $\zeta(\varnothing) = 1$). The **weight** of $\mathbf{k}$ is $\mathrm{wt}(\mathbf{k}) = k_1 + \cdots + k_r$, and the **depth** of $\mathbf{k}$ is $\mathrm{dep}(\mathbf{k}) = r$, with the convention that $\mathrm{wt}(\varnothing) = \mathrm{dep}(\varnothing) = 0$. We will study the vector spaces

$$\mathcal{Z} := \mathrm{Span}_{\mathbb{Q}}\{\zeta(\mathbf{k}) \mid \mathbf{k} \text{ admissible}\}$$

and

$$\mathcal{Z}_k := \mathrm{Span}_{\mathbb{Q}}\{\zeta(\mathbf{k}) \mid \mathbf{k} \text{ admissible with } \mathrm{wt}(\mathbf{k}) = k\}$$

for $k \geq 0$. It is clear by definition that $\mathcal{Z} = \sum_{k\geq 0} \mathcal{Z}_k$. One of the main conjectures about multiple zeta values is that in fact

$$\mathcal{Z} = \bigoplus_{k\geq 0} \mathcal{Z}_k,$$

i.e. that the space $\mathcal{Z}$ is graded by weight. Another way of phrasing this is that there should be no $\mathbb{Q}$-linear relations between multiple zeta values of different weight other than those that can be obtained as sums of weight-homogeneous relations. This is a rather strong conjecture; for example, since $\mathcal{Z}_0 = \mathbb{Q}$, it would immediately imply that all MZV of positive weight are irrational. In fact, it would imply that they are transcendental: We will show that $\mathcal{Z}$ is a $\mathbb{Q}$-algebra, with the product respecting the (conjectured) grading, so in particular, algebraic relations reduce to linear relations. On the other hand, there are plenty of weight-homogeneous relations among MZV. In this section, we will apply quasi-shuffle products to produce one family of relations, namely the (finite) **double-shuffle relations**.

Define the following three $\mathbb{Q}$-vector spaces $\mathfrak{H} \supset \mathfrak{H}^1 \supset \mathfrak{H}^0$:

$$\begin{aligned}
\mathfrak{H} &:= \mathbb{Q}\langle x, y\rangle, \\
\mathfrak{H}^1 &:= \mathbb{Q} + \mathfrak{H}y, \\
\mathfrak{H}^0 &:= \mathbb{Q} + x\mathfrak{H}y,
\end{aligned}$$

i.e. $\mathfrak{H}$ is generated by all words in the symbols x and y (including the empty word 1), $\mathfrak{H}^1$ is the subspace generated by the empty word and words ending in y, and $\mathfrak{H}^0$ is the subspace generated by the empty word and words beginning in x and ending in y. Note that $\mathfrak{H}$ can be equipped with the shuffle product $⧢$, and that $(\mathfrak{H}^1, ⧢)$ and $(\mathfrak{H}^0, ⧢)$ are easily seen to be subalgebras of $(\mathfrak{H}, ⧢)$. Note further that if we define $z_k = x^{k-1}y$ for $k \geq 1$, then $\mathfrak{H}^1 = \mathbb{Q}\Big\langle (z_k)_{k\geq 1}\Big\rangle$. We can then equip $\mathfrak{H}^1$ with the stuffle-product $*$, and again it is easy to see that $(\mathfrak{H}^0, *)$ is a subalgebra of $(\mathfrak{H}^1, *)$. We will usually write $\mathfrak{H}^0_{⧢}$, $\mathfrak{H}^0_*$ etc. to denote the quasi-shuffle algebras $(\mathfrak{H}^0, ⧢)$, $(\mathfrak{H}^0, *)$ etc.

An index $\mathbf{k} = (k_1, \ldots, k_r)$ may be identified with the generator $z_{\mathbf{k}} := z_{k_1} \cdots z_{k_r}$ in $\mathfrak{H}^1$. Note then that $z_{\mathbf{k}} \in \mathfrak{H}^0$ if and only if either $\mathbf{k} = \varnothing$ or $k_1 \geq 1$, i.e. if the index is admissible. It therefore makes sense to extend the multizeta map to a linear map $\zeta : \mathfrak{H}^0 \to \mathcal{Z}$ defined on the basis vectors by

$$\zeta(z_{\mathbf{k}}) := \zeta(\mathbf{k}).$$

The aim of this section is to show the following:

Theorem 2.4. *The map $\zeta : \mathfrak{H}^0 \to \mathcal{Z}$ is a $\mathbb{Q}$-algebra homomorphism, both with respect to the shuffle-product $⧢$, and with respect to the stuffle-product $*$.*

Technically, we haven't proven yet that $\mathcal{Z}$ is even a $\mathbb{Q}$-algebra (i.e. we don't know yet that the product of two MZV's is a $\mathbb{Q}$-linear combination of MZV's). A more proper formulation of Theorem 2.4 is then that ζ is a homomorphism from $\mathfrak{H}^0_{⧢}$ (resp. $\mathfrak{H}^0_*$) to $\mathbb{R}$. It then follows that the image $\zeta(\mathfrak{H}^0) = \mathcal{Z}$ is in fact a $\mathbb{Q}$-subalgebra of $\mathbb{R}$. Theorem 2.4 immediately gives us a family of linear relations among MZVs:

Corollary 2.5 (Finite double-shuffle relations)**.** *For any $w, v \in \mathfrak{H}^0$, we have*

$$\zeta(w ⧢ v - w * v) = 0.$$

Proof. By Theorem 2.4 we have

$$\zeta(w \sqcup\!\sqcup v - w * v) = \zeta(w \sqcup\!\sqcup v) - \zeta(w * v) = \zeta(w)\zeta(v) - \zeta(w)\zeta(v) = 0.$$

□

To prove Theorem 2.4, we will have to work on the larger space $\mathfrak{H}^1$ instead of $\mathfrak{H}^0$. To do this, we need some analogue of ζ that makes sense on this larger space. In the case of the stuffle-product, this will be the truncated multiple zeta function ζ_M (defined in (8)), and in the case of the shuffle-product it will be the multiple polylogarithm $\mathrm{Li}_{-}(z)$ (defined by (6) below). In both cases, we will prove that these give homomorphisms onto some $\mathbb{Q}$-algebra (respectively $\mathbb{Q}$ itself and the $\mathbb{Q}$-algebra $\mathcal{H}(D(0,1))$ of holomorphic functions on the open unit disk) that, when restricted to the subalgebra $\mathfrak{H}^0$, yield ζ in some limit (respectively $M \to \infty$ and $z \to 1$).

Let us start with the shuffle-product. For any index $\mathbf{k} = (k_1, \dots, k_r)$ (not necessarily admissible), define the **multiple polylogarithm** $\mathrm{Li}_{\mathbf{k}}$ by

$$\mathrm{Li}_{\mathbf{k}}(z) := \sum_{m_1 > \cdots > m_r > 0} \frac{z^{m_1}}{m_1^{k_1} \cdots m_r^{k_r}}, \qquad |z| < 1, \tag{6}$$

for $r \geq 1$ and

$$\mathrm{Li}_{\varnothing} := 1.$$

The power-series above does indeed have radius of convergence 1 (or ∞ in the case of $\mathrm{Li}_{\varnothing}$), so $\mathrm{Li}_{\mathbf{k}}$ defines a holomomorphic function on the open unit disk $D(0,1)$. Note that when $\mathbf{k}$ is admissible, the power-series is in fact convergent at $z = 1$, and we have $\lim_{z \to 1} \mathrm{Li}_{\mathbf{k}}(z) = \zeta(k)$.

If we see $\mathrm{Li}_{\mathbf{k}}(z)$ as a function of $\mathbf{k}$, then that gives a map from the set of indices to the space $\mathcal{H}(D(0,1))$ of holomorphic functions on the unit disk. Analogously to ζ, we can extend this to a linear map on $\mathfrak{H}^1$:

$$\begin{aligned} \mathrm{Li}_{-} : \mathfrak{H}^1 &\to \mathcal{H}(D(0,1)) \\ z_{\mathbf{k}} &\mapsto \mathrm{Li}_{\mathbf{k}}. \end{aligned}$$

Our goal is to show that this is a homomorphism with respect to the shuffle product. The key to proving this is that the multiple polylogarithm satisfies a certain differential equation:

Lemma 2.6. *For any symbol $a \in \{x, y\}$ and any word $w \in \mathfrak{H}^1$, we have*

$$\frac{d}{dz}\mathrm{Li}_{aw}(z) = f_a(z)\mathrm{Li}_w(z),$$

where

$$f_x(z) = \frac{1}{z}, \qquad f_y(z) = \frac{1}{1-z}.$$

Proof. Suppose $aw = z_{k_1} \cdots z_{k_r}$. Then

$$\frac{d}{dz}\mathrm{Li}_{aw}(z) = \sum_{m_1 > \cdots > m_r > 0} \frac{m_1 z^{m_1 - 1}}{m_1^{k_1} \cdots m_r^{k_r}} = \sum_{m_1 > \cdots > m_r > 0} \frac{z^{m_1 - 1}}{m_1^{k_1 - 1} m_2^{k_2} \cdots m_r^{k_r}}. \tag{7}$$

There are then two cases to check:

Case 1: Suppose $a = x$. This means that $k_1 \geq 2$, so $w = z_{k_1 - 1} z_{k_2} \cdots z_{k_r}$. But then (7) says exactly that

$$\frac{d}{dz}\mathrm{Li}_{aw}(z) = \frac{1}{z}\mathrm{Li}_{z_{k_1-1}z_{k_2}\cdots z_{k_r}}(z) = \frac{1}{z}\mathrm{Li}_w(z),$$

as claimed

Case 2: Suppose $a = y$. This means that $k_1 = 1$, so $w = z_{k_2} \cdots z_{k_r}$. But then we can rewrite (7) to

$$\begin{aligned}
\frac{d}{dz}\mathrm{Li}_{aw}(z) &= \sum_{m_2 > \cdots > m_r > 0} \left(\frac{1}{m_2^{k_2} \cdots m_r^{k_r}} \sum_{m_1 > m_2} z^{m_1 - 1} \right) \\
&= \sum_{m_2 > \cdots > m_r > 0} \frac{1}{m_2^{k_2} \cdots m_r^{k_r}} \cdot \frac{z^{m_2}}{1 - z} \\
&= \frac{1}{1 - z} \mathrm{Li}_{z_{k_2} \cdots z_{k_r}}(z) = \frac{1}{1 - z} \mathrm{Li}_w(z),
\end{aligned}$$

where we used the fact that the geometric series $\sum_{m_1 > m_2} z^{m_1 - 1}$ with initial term z^{m_2} and common ratio z converges to $\frac{z^{m_2}}{1-z}$ for $|z| < 1$.

□

Proposition 2.7. $\mathrm{Li}_{\_} : \mathfrak{H}^1_{ш} \to \mathcal{H}(D(0,1))$ *is a homomorphism of* $\mathbb{Q}$*-algebras.*

Proof. We prove that $\mathrm{Li}_{w ш v}(z) = \mathrm{Li}_w(z)\mathrm{Li}_v(z)$ for words $w, v \in \mathfrak{H}^1$ by induction in $\ell(w) + \ell(v)$. As usual, this is trivial when either word is empty, so we get the base case for free, and can assume for the induction step that $w = aw'$ and $v = bv'$ for some symbols $a, b \in x, y$ and words $w', v' \in \mathfrak{H}^1$. Then using Lemma 2.6, the induction hypothesis, and then Lemma 2.6 again, we get

$$\begin{aligned}
\frac{d}{dz}\mathrm{Li}_w(z)\mathrm{Li}_v(z) &= f_a(z)\mathrm{Li}_{w'}(z)\mathrm{Li}_{bv'}(z) + f_b(z)\mathrm{Li}_{aw'}(z)\mathrm{Li}_{v'}(z) \\
&= f_a(z)\mathrm{Li}_{w' ш bv'}(z) + f_b(z)\mathrm{Li}_{aw' ш v'}(z) \\
&= \frac{d}{dz}\left(\mathrm{Li}_{a(w' ш bv')} + \mathrm{Li}_{b(aw' ш v')}(z)\right) \\
&= \frac{d}{dz}(\mathrm{Li}_{a(w' ш bv') + b(aw' ш v')}(z) \\
&= \frac{d}{dz}\mathrm{Li}_{w ш v}(z).
\end{aligned}$$

Thus $\mathrm{Li}_w(z)\mathrm{Li}_v(z) - \mathrm{Li}_{w ш v}(z)$ is constant, and since

$$\mathrm{Li}_w(0) = \begin{cases} 1 & w = 1, \\ 0 & \ell(w) \geq 1, \end{cases}$$

and similarly for Li_v and $\mathrm{Li}_{w ш v}$, we see that this constant is zero. □

From this, we get the first half of Theorem 2.4:

Corollary 2.8. $\zeta : \mathfrak{H}^0_{ш} \to \mathcal{Z}$ *is a homomorphism of* $\mathbb{Q}$*-algebras.*

Proof. We have $\mathrm{Li}_{w ш v}(z) = \mathrm{Li}_w(z)\mathrm{Li}_v(z)$ for any $w, v \in \mathfrak{H}^1$, so this holds in particular if we restrict to $w, v \in \mathfrak{H}^0$. Taking the limit $z \to 1$ then yields $\zeta(w ш v) = \zeta(w)\zeta(v)$. □

Let us now turn to the stuffle product. Given a positive integer M, we define the **Truncated Multiple Zeta Value** $\zeta_M(\mathbf{k})$ for any index $\mathbf{k} = (k_1, \ldots, k_r)$ by

$$\zeta_M(\mathbf{k}) = \sum_{M > m_1 > \cdots > m_r > 0} \frac{1}{m_1^{k_1} \cdots m_r^{k_r}} \in \mathbb{Q} \tag{8}$$

for $r \geq 1$, and

$$\zeta_M(\varnothing) = 1.$$

As usual, we can extend ζ_M to a linear map $\zeta_M : \mathfrak{H}^1 \to \mathbb{Q}$. Our goal is then to show that this is a homomorphism with respect to the stuffle product, which upon restricting to $\mathfrak{H}^0$ and taking the limit $M \to \infty$ implies that ζ is a homomorphism from $\mathfrak{H}^0_*$ to $\mathcal{Z}$. It is actually quite straightforward to show this by induction in M. However, we will here take a different, more abstract approach, which has the advantage of being applicable in many other settings (for example, we will make extensive use of it in Section 3). The idea will be to use the following simple fact:

Lemma 2.9. *Let A be a $\Bbbk$-bialgebra with coproduct $\Delta : A \to A \otimes A$, and let B be a commutative $\Bbbk$-algebra with product $m : B \otimes B$. If $f, g : A \to B$ are homomorphisms of $\Bbbk$-algebras, then so is their convolution*

$$f \star g = m \circ (f \otimes g) \circ \Delta.$$

Proof. This simply follows from the fact that $f \star g$ is a composition of homomorphisms: $\Delta : A \to A \otimes A$ is a homomorphism because it is the coproduct of the bialgebra A, $f \otimes g : A \otimes A \to B \otimes B$ is a homomorphism because f and g are homomorphisms, and $m : B \otimes B \to B$ is a homomorphism because B is commutative: For any $a, b, c, d \in B$ we have

$$m((a \otimes b)(c \otimes d)) = m(ac \otimes bd) = acbd = abcd = m(ab \otimes cd) = m((a \otimes c)(b \otimes d)).$$

□

It thus suffices to realize ζ_M as a convolution of homomorphisms. We can produce these homomorphisms using the following lemma:

Lemma 2.10. *Suppose $\tilde{f} : (\Bbbk A, \diamond) \to B$ is a homomorphism from the semialgebra of symbols to some $\Bbbk$-algebra B. Then the linear map $f : (\Bbbk\langle A \rangle, *_\diamond) \to B$ given on a word $a_1 \cdots a_r \in A^*$ by*

$$f(a_1 \cdots a_r) = \begin{cases} 1 & r = 0, \\ \tilde{f}(a_1) & r = 1, \\ 0 & r \geq 2, \end{cases}$$

is a homomorphism of $\Bbbk$-algebras.

Proof. It suffices to verify that $f(w *_\diamond v) = f(w) f(v)$ for words $w, v \in A^*$. This is obvious in the case when either word is empty. It is also clear in the case when either word has length at least 2: Note that when we expand $w *_\diamond v$ using the recursive definition (4), all the terms that we get will have length at least $\max\{\ell(w), \ell(v)\} \geq 2$, so $f(w *_\diamond v) = 0 = f(w) f(v)$. This leaves only the case when w and v are both single symbols, say $w = a$ and $v = b$. Then

$$f(w *_\diamond v) = f(wv) + f(vw) + f(w *_\diamond v) = 0 + 0 + \tilde{f}(w *_\diamond v) = \tilde{f}(w)\tilde{f}(v) = f(w) f(v).$$

□

Proposition 2.11. *For every $M > 0$, $\zeta_M : \mathfrak{H}^1_* \to \mathbb{Q}$ is a homomorphism of $\mathbb{Q}$-algebras.*

In fact, we will prove something more general:

Lemma 2.12. *Let B be a commutative $\Bbbk$-algebra, and let $\mathcal{N}$ be a finite set equipped with a total order $\succ$. Suppose we are given a family of homomorphisms $\tilde{f}_m : (\Bbbk\langle A\rangle, \diamond) \to B$ for $m \in \mathcal{N}$. Then the linear map $F_{\mathcal{N}} : \Bbbk\langle A\rangle \to B$ given on words $a_1 \cdots a_r \in A^*$ by*

$$F_{\mathcal{N}}(a_1 \cdots a_r) = \sum_{\substack{m_1,\ldots,m_r \in \mathcal{N} \\ m_1 \succ \cdots \succ m_r}} \tilde{f}_{m_1}(a_1) \cdots \tilde{f}_{m_r}(a_r) \tag{9}$$

for $r \geq 1$, and

$$F_{\mathcal{N}}(1) = 1$$

is a homomorphism from $(\Bbbk\langle a\rangle, *_\diamond)$ to B.

Proof. Note that in the case $\mathcal{N} = \emptyset$, we have

$$F_{\emptyset}(w) = \begin{cases} 1 & w = 1, \\ 0 & \ell(w) \geq 1, \end{cases}$$

which is clearly a homomorphism. Let us therefore restrict ourselves to the cases where $\mathcal{N} \neq \emptyset$.

The idea for the proof is to realize $F_{\mathcal{N}}$ as a convolution of homomorphisms. For every $m \in \mathcal{N}$, define a linear map $f_m : \mathfrak{H}^1_* \to \mathbb{Q}$ on words by

$$f(a_1 \cdots a_r) = \begin{cases} 1 & r = 0, \\ \tilde{f}(a_1) & r = 1, \\ 0 & r \geq 2. \end{cases}$$

By Lemma 2.10, f_m is a homomorphism for every m. Therefore, by Lemma 2.9, so is the map

$$\bigstar_{m \in \mathcal{N}} f_m,$$

where the convolution is to be taken in decreasing order, i.e. if $n_1 \succ \cdots \succ n_M$ are all the elements of $\mathcal{N}$, then $\bigstar_{m\in\mathcal{N}} f_m = f_{n_1} \star \cdots \star f_{n_M}$. The claim is now that in fact $F_{\mathcal{N}} = \bigstar_{m\in\mathcal{N}} f_m$, from which the proposition immediately follows (note that $F_\emptyset$ is in fact equal to the unit $u \circ \varepsilon$ for the convolution product $\star$, so we can also think of $F_\emptyset$ as the "empty convolutoin"). Let $w = a_1 \cdots a_r$. Then

$$\left(\bigstar_{m\in\mathcal{N}} f_m\right)(w) = \sum_{w_1 \cdots w_M = w} f_{n_1}(w_1) \cdots f_{n_M}(w_M),$$

where $n_1 \succ \cdots \succ n_M$ are all the elements of $\mathcal{N}$ in decreasing order, and the sum runs over all ways of splitting w into M subwords, allowing the empty word. Consider some term $f_{n_1}(w_1) \cdots f_{n_M}(w_M)$ in this sum. If $\ell(w_j) \geq 2$ for some j, then $f_{n_j}(w_j) = 0$, so the whole term vanishes. We may therefore restrict the sum to only run over ways of splitting w up into M subwords, each of which is either empty or a single symbol. The number of single symbol words in each term must be r, and said single symbol words must be $w_{m_1} = a_1, \ldots, w_{m_r} = a_r$ for some $m_1 \succ \cdots \succ m_r$. Any remaining empty words just contribute factors of 1 to the term, and can thus be disregarded. We therefore have

$$\begin{aligned}\left(\bigstar_{m\in\mathcal{N}} f_m\right)(w) &= \sum_{\substack{m_1,\ldots,m_r \in \mathcal{N} \\ m_1 \succ \cdots \succ m_r}} f_{m_1}(a_1) \cdots f_{m_r}(a_r) \\ &= \sum_{\substack{m_1,\ldots,m_r \in \mathcal{N} \\ m_1 \succ \cdots \succ m_r}} \tilde{f}_{m_1}(a_1) \cdots \tilde{f}_{m_r}(a_r) \\ &= F_{\mathcal{N}}(w),\end{aligned}$$

as claimed, which finishes the proof. $\square$

Note that Proposition 2.11 follows from Lemma 2.12 by taking $\mathcal{N} = \{1, \ldots, M-1\}$ (equipped with its usual ordering), and $\tilde{f}_m(z_k) = \frac{1}{m^k}$, which is indeed a homomorphism for $\diamond$, since $\tilde{f}_m(z_k)\tilde{f}_m(z_\ell) = \frac{1}{m^k}\frac{1}{m^\ell} = \frac{1}{m^{k+\ell}} = \tilde{f}_m(z_{k+\ell}) = \tilde{f}_m(z_k \diamond z_\ell)$. As before, we can restrict ζ_M to $\mathfrak{H}^0$, which upon taking the limit $M \to \infty$ yields the second part of Theorem 2.4:

Corollary 2.13. *$\zeta : \mathfrak{H}^0_* \to \mathcal{Z}$ is a homomorphism of $\mathbb{Q}$-algebras.* $\square$

With this, we have finally proven Theorem 2.4 and Corollary 2.5.

2.2 Regularization and the extended double-shuffle relations

The finite double shuffle relations $\zeta(w \sqcup\!\sqcup v - w * v) = 0$ do not give us all relations among multiple zeta values. For one thing, if w, v are nonempty words in $\mathfrak{H}^0$, then they each represent indices of weight at least 2, so there is no way to achieve any relations in weight 3 such as $\zeta(2,1) = \zeta(3)$. On the other hand, if we forget about the fact that $z_1 \notin \mathfrak{H}^0$, we could simply try to compute

$$z_1 \sqcup\!\sqcup z_2 - z_1 * z_2 = z_2 z_1 - z_3,$$

and noting that the right-hand side is in $\mathfrak{H}^0$ we "conclude" that $\zeta(z_2 z_1 - z_3) = 0$. In this section, we shall make sense of calculations like this. The key will be to extend the map ζ from $\mathfrak{H}^0$ to $\mathfrak{H}^1$ in a way that preserves its homomorphism property with respect to the stuffle and shuffle product. This will yield two different homomorphisms $\zeta^{\sqcup\!\sqcup}$ and ζ^* with respect to $\sqcup\!\sqcup$ and $*$ respectively, but somewhat miraculously, we still have $\zeta^{\sqcup\!\sqcup}(w \sqcup\!\sqcup v - w * v) = \zeta^*(w \sqcup\!\sqcup v - w * v) = 0$ whenever $w \in \mathfrak{H}^0$ and $v \in \mathfrak{H}^1$. The exposition given here follows [Bac2, section 2.3-2.4].

Let us first fix some notation. Let $Q = \Bbbk\langle A \rangle$ be a quasi-shuffle algebra with respect to the quasi-shuffle product $*_\diamond$, and let L, R be subsets of the alphabet A. We will denote by Q_L^R the subspace

$$Q_L^R := \Bbbk + \mathrm{Span}_\Bbbk(\{a_1 \cdots a_r \mid a_1, \ldots, a_r \in A, a_1 \in L, a_r \in R\})$$

of Q, which is generated by the empty word, as well as words starting with a symbol from L and ending with a symbol from R. Note that if $\Bbbk L$ and $\Bbbk R$ are closed under $\diamond$, then Q_L^R is a subalgebra of Q.

Theorem 2.14. *Let $L, R \subset A$ and $a \in A$ be such that $\Bbbk L$ and $\Bbbk R$ are closed under $\diamond$, and such that $A \diamond A \setminus \{A\} \subseteq \Bbbk A \setminus \{a\}$.*

(i) If $a \in R$, then we have an isomorphism of $\Bbbk$-algebras

$$\begin{aligned} Q^R_{A\setminus\{a\}}[X] &\xrightarrow{\sim} Q^R_A \\ \sum_{j=0}^n \alpha_j X^j &\mapsto \sum_{j=0}^n \alpha_j *_\diamond a^{*_\diamond j}. \end{aligned}$$

(ii) If $a \in L$, then we have an isomorphism of $\Bbbk$-algebras

$$\begin{aligned} Q_L^{A\setminus\{a\}}[X] &\xrightarrow{\sim} Q_L^A \\ \sum_{j=0}^n \alpha_j X^j &\mapsto \sum_{j=0}^n \alpha_j *_\diamond a^{*_\diamond j}. \end{aligned}$$

Proof. We only prove (i), as (ii) is similar. Let $F : Q^R_{A\setminus\{a\}}[X] \to Q^R_A$ be the map sending $\sum_{j=0}^n \alpha_j X^j$ to $\sum_{j=0}^n \alpha_j *_\diamond a^{*_\diamond j}$. This is clearly a homomorphism, so we only need to show that it is bijective.

For surjectivity, let w be a word in Q^R_A, so we can write w on the form $a^n w'$ with $w' \in Q^R_{A\setminus\{a\}}$. We shal prove by induction in n that w is in the image of F. The base case $n = 0$ is trivial, as then $w = w' = F(w')$. For the induction step, we let $n > 0$ and note that

$$a *_\diamond a^{n-1}w' = na^n w' + (\text{terms starting with} < n\ a\text{'s}). \tag{10}$$

The n copies of the word $a^n w'$ comes from the n ways the a can be inserted into initial string of a's in w. Inserting a later in the word yields words starting with $n-1$ a's (since the first letter of w' is not a). The terms coming from "diamonding" a with one of the first $n-1$ a's, give linear combinations of words starting with $\leq n-1$ a's for the same reason, and similarly diamonding a with a letter of w' other than the first one give linear combinations of words starting with $n-1$ a's. This leaves the terms coming from diamonding a with the first letter of w. Since this letter is an element of $A \setminus \{a\}$, and we assumed $A \diamond (A \setminus \{a\}) \subseteq \Bbbk A \setminus \{a\}$, this also yields terms starting with $n-1$ a's, thus proving (10).

By induction, all terms on the right-hand side of (10) except $na^n w'$ are in the image of F. Also by induction, $a^{n-1}w'$ is in the image of F, say $a^{n-1}w' = F(v)$. Then $F(Xv) = a *_\diamond a^{n-1}w'$, so the left-hand side of (10) is also in the image of F. Upon rearranging, this shows that $a^n w'$ is in the image of F, which finishes the proof that F is surjective.

For the injectivity, we assume that $F\left(\sum_{j=0}^n \alpha_j X^j\right) = 0$. We will show, by induction in n, that this implies that $\alpha_0 = \cdots = \alpha_n = 0$. From (10), it follows that

$$\alpha_j *_\diamond a^{*_\diamond j} = j! a^j \alpha_j + (\text{terms starting with} < n\ a\text{'s}).$$

In particular, the only term in $F\left(\sum_{j=0}^n \alpha_j X^j\right) = 0$ that can contribute words starting with n a's is $\alpha_n *_\diamond a^{*_\diamond n}$, so we must have $n! a^n \alpha_n = 0$, which implies $\alpha_n = 0$. Thus

$$0 = F\left(\sum_{j=0}^{n} \alpha_j X^j\right) = F\left(\sum_{j=0}^{n-1} \alpha_j X^j\right),$$

which inductively implies $\alpha_0 = \cdots = \alpha_{n-1} = 0$. $\square$

As special cases, we have

Corollary 2.15. *We have isomorphisms*

$$\begin{aligned} \mathfrak{H}^0_*[X] &\xrightarrow{\sim} \mathfrak{H}^1_* \\ \sum_{j=0}^n \alpha_j X^j &\mapsto \sum_{j=0}^n \alpha_j * z_1^{*j}, \end{aligned} \qquad \begin{aligned} \mathfrak{H}^0_{\sqcup\!\sqcup}[X] &\xrightarrow{\sim} \mathfrak{H}^1_{\sqcup\!\sqcup} \\ \sum_{j=0}^n \alpha_j X^j &\mapsto \sum_{j=0}^n \alpha_j \sqcup\!\sqcup\, y^{\sqcup\!\sqcup j}. \end{aligned}$$

Proof. This follows from Theorem 2.14(i) by taking respectively $A = R = \{z_k \mid k \geq 1\}, a = z_1$ and $A = R = \{x, y\}, a = y$. $\square$

From this, we get a general way of extending any shuffle/stuffle-homomorphism of $\mathfrak{H}^0$ to a shuffle/stuffle-homomorphism of $\mathfrak{H}^1$:

Lemma 2.16. *Let $\bullet \in \{*, ш\}$, and let $F : \mathfrak{H}^0_\bullet \to B$ be an algebra homomorphism for some $\Bbbk$-algebra B. Then for any $\beta \in B$, there exists a unique homomorphism $F^\bullet : \mathfrak{H}^1_\bullet \to B$ such that $F^\bullet|_{\mathfrak{H}^0} = F$ and $F^\bullet(z_1) = \beta$.*

Proof. This follows from the universal property of the polynomial ring: Any homomorphism $F : \mathfrak{H}^0_\bullet \to B$ can be extended uniquely to $\mathfrak{H}^0_\bullet[X]$ by choosing where X goes, so by Corollary 2.15, it can be extended uniquely to $\mathfrak{H}^1_\bullet$ by choosing where $z_1 = y$ goes. □

Note that the extension of F depends on the choice of $F^\bullet(z_1)$. However, one can avoid making this choice by considering F as a homomorphism $\mathfrak{H}^0_\bullet \to B \hookrightarrow B[X]$ and then setting $F^\bullet(z_1) = X$. One can then always make the choice later by evaluating at $X = \beta$ for some $\beta \in B$. We use this to define extensions ζ^* and $\zeta^ш$ of ζ in the most general way.

Definition 2.17. Let $\bullet \in \{*, ш\}$. Using Lemma 2.16, we define a homomorphism $\zeta^\bullet : \mathfrak{H}^1 \to \mathcal{Z}[X]$ by $\zeta^\bullet(w; X) = \zeta(w)$ when $w \in \mathfrak{H}^0$, and $\zeta^\bullet(z_1; X) = X$. We use the notation $\zeta^\bullet(w) := \zeta^\bullet(w; 0)$ to denote its evaluation at $X = 0$.

In general, $\zeta^*(w; X)$ and $\zeta^ш(w; X)$ differ. They are however related in the following way:

Theorem 2.18 ([IKZ, Theorem 1])**.** *There exists a bijective $\mathbb{R}$-linear map $\rho : \mathbb{R}[X] \to \mathbb{R}[X]$ such that*

$$\zeta^ш(\mathbf{k}; X) = \rho(\zeta^*(\mathbf{k}; X))$$

for all indices $\mathbf{k}$*.*

We shall not prove this theorem here. The map ρ is defined by the equation

$$\rho(\exp(uX)) = A(u)\exp(uX),$$

where

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

Taking for granted the existence of ρ, we now get a large family of relations, called **extended double-shuffle relations** among multiple zeta values.

Theorem 2.19 (Extended double-shuffle relations)**.** *For $w \in \mathfrak{H}^1$ and $v \in \mathfrak{H}^0$, we have*

$$\zeta^ш(w ш v - w * v; X) = \zeta^*(w ш v - w * v; X) = 0.$$

Proof. Using Theorem 2.18, the linearity of ρ, and the fact that $\zeta^ш(v; X) = \zeta^*(v; X) = \zeta(v)$, we get

$$\begin{aligned}\zeta^ш(w ш v; X) &= \zeta^ш(w; X)\zeta(v) = \rho(\zeta^*(w; X))\zeta(v)\\ &= \rho(\zeta^*(w; X)\zeta(v)) = \rho(\zeta^*(w * v; X)) = \zeta^ш(w * v; X),\end{aligned}$$

so $\zeta^ш(w ш v - w * v; X) = 0$. Applying ρ^{-1} gives $\zeta^*(w ш v - w * v; X) = 0$. □

It is conjectured that the extended double-shuffle relations generate all linear relations among multiple zeta values. More precisely, for $\bullet \in \{ш, *\}$, let $\mathrm{reg}^X_\bullet : \mathfrak{H}^1_\bullet \xrightarrow{\sim} \mathfrak{H}^0_\bullet[X]$ be the inverse of the isomorphism from Corollary 2.15, and let $\mathrm{reg}_\bullet = \mathrm{reg}^0_\bullet : \mathfrak{H}_1 \to \mathfrak{H}_0$ be the composition of $\mathrm{reg}^X_\bullet$ with the evaluation map at $X = 0$. Then it is conjectured that

$$\ker(\zeta) = \mathrm{Span}_{\mathbb{Q}}\{\mathrm{reg}_ш(w ш v - w * v) \mid w \in \mathfrak{H}^1, v \in \mathfrak{H}^0\} = \mathrm{Span}_{\mathbb{Q}}\{\mathrm{reg}_*(w ш v - w * v) \mid w \in \mathfrak{H}^1, v \in \mathfrak{H}^0\}.$$

This would for example immediately imply that the $\mathbb{Q}$-algebra $\mathcal{Z}$ is graded by weight, since the shuffle and stuffle products are weight-homogeneous.

3 Multiple Eisenstein series and their regularizations

In this section, we will use the algebraic machinery we have built up in the previous section to study multiple Eisenstein series (MES). Before we define what these are, let us briefly discuss (single) Eisenstein series, of which MES are a generalization.

Let $\mathbb{H} = \{\tau \in \mathbb{C} \mid \operatorname{Im}(\tau) > 0\}$ be the upper half-plane. For an even integer $k \geq 4$, the Eisenstein series $G_k \in \mathcal{H}(\mathbb{H})$ is given by

$$G_k(\tau) = \frac{1}{2} \sum_{\substack{\lambda \in \mathbb{Z}\tau + \mathbb{Z} \\ \lambda \neq 0}} \frac{1}{\lambda^k} = \frac{1}{2} \sum_{\substack{m,n \in \mathbb{Z} \\ (m,n) \neq (0,0)}} \frac{1}{(mz+n)^k}. \tag{11}$$

It is not hard to check that this sum converges absolutely and locally uniformly whenever $k > 2$, so that the limit function is holomorphic, see e.g. [BGHZ, p. 14]. The main reason for excluding odd k is simply that it is not very interesting: When k is odd each λ-term in the sum is cancelled by the $-\lambda$-term, so the whole sum is zero (in a moment, we will give a different definition under which G_k does not vanish for odd k). When k is even, λ and $-\lambda$ contribute the same to the sum, which is the reason for the normalizing factor of $\frac{1}{2}$. Let us remark that we can still make sense of the non-absolutely convergent case $k = 2$, if we agree to define

$$G_2(\tau) := \lim_{M \to \infty} \lim_{N \to \infty} \frac{1}{2} \sum_{\substack{\lambda \in \mathbb{Z}_M \tau + \mathbb{Z}_N \\ \lambda \neq 0}} \frac{1}{\lambda^2},$$

where $\mathbb{Z}_M := \{-(M-1), \ldots, M-1\}$.

The reason why Eisenstein series are interesting is that they provide the main examples of modular forms.

Definition 3.1. Let k be an integer. A **modular form** of weight k (and level 1) is a holomorphic function $f \in \mathcal{H}(\mathbb{H})$ such that

(i) $f\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^k f(\tau)$ for $\begin{pmatrix} a & b \\ c & c \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z})$, and

(ii) $f(\tau)$ is bounded as $\tau \to i\infty$.

In the case of G_k where $k \geq 4$ is even, condition (i) follows by rearranging the absolutely convergent sum (11): For $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z})$

$$\begin{aligned} G_k\left(\frac{a\tau+b}{c\tau+d}\right) &= \sum_{\substack{m,n \in \mathbb{Z} \\ (m,n) \neq (0,0)}} \frac{1}{(m\frac{a\tau+b}{c\tau+d}+n)^k} \\ &= (c\tau+d)^k \sum_{\substack{m,n \in \mathbb{Z} \\ (m,n) \neq (0,0)}} \frac{1}{((am+cn)\tau+(bm+dn))^k} \\ &= (c\tau+d)^k G_k(\tau), \end{aligned}$$

since the map $(m,n) \mapsto (am+cn, bm+dn)$ maps $\mathbb{Z}^2 \setminus \{(0,0)\}$ bijectively to itself for $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z})$. In the case of G_2, we cannot rearrange the order of summation freely, so the above argument does not work. Indeed, G_2 is not a modular form, though it turns out to have the quasi-modular property

$$G_2\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^2 G_2(\tau) - \pi i c(c\tau+d),$$

see [BGHZ, Prop. 6]. If one does attempt to change the order of summation, one then gets

$$\begin{aligned}\lim_{N\to\infty}\lim_{M\to\infty}\frac{1}{2}\sum_{\substack{m\in\mathbb{Z}_M, n\in\mathbb{Z}_N\\(m,n)\neq 0}}\frac{1}{(m\tau+n)^2} &= \lim_{N\to\infty}\lim_{M\to\infty}\frac{1}{2\tau^2}\sum_{\substack{m\in\mathbb{Z}_M, n\in\mathbb{Z}_N\\(m,n)\neq 0}}\frac{1}{(m+n\tau^{-1})^2}\\ &= \lim_{N\to\infty}\lim_{M\to\infty}\frac{1}{2\tau^2}\sum_{\substack{m\in\mathbb{Z}_M, n\in\mathbb{Z}_N\\(m,n)\neq 0}}\frac{1}{(m-n\tau^{-1})^2}\\ &= \tau^{-2}G_2\left(\frac{-1}{\tau}\right) = G_2(\tau) - \frac{\pi i}{\tau}.\end{aligned}$$

As for condition (ii), note that the matrix $\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z})$ corresponds to the Möbius transformation $\tau \to \tau+1$, so by condition (i), a modular form is 1-periodic in the real direction. It can thus be expressed as a Fourier series $f(\tau) = \sum_{n\in\mathbb{Z}} a_n q^n$, where $q = e^{2\pi i\tau}$ (this notation is standard, and will be used throughout this chapter). The condition that $f(\tau)$ is bounded as $\tau \to i\infty$ is then equivalent to $a_n = 0$ for all $n < 0$; in other words that f has a Fourier series of the form

$$f(\tau) = \sum_{n\geq 0} a_n q^n.$$

In the case of Eisenstein series, we have the following remarkable result:

Theorem 3.2. *For $k \geq 2$ even,*

$$G_k(\tau) = \zeta(k) + \frac{(-2\pi i)^k}{(k-1)!}\sum_{n>0}\sigma_{k-1}(n)q^n,$$

where ζ is the Riemann zeta function, and σ_{k-1} is the sum-of-divisors function given by

$$\sigma_{k-1}(n) = \sum_{d\,|\,n} d^{k-1}.$$

The proof relies on the following formula:

Proposition 3.3 (Lipschitz‘s formula)**.** *For $k \geq 2$, and $\tau \in \mathbb{C}\setminus\mathbb{Z}$*

$$\sum_{n\in\mathbb{Z}}\frac{1}{(\tau+n)^k} = \frac{(-2\pi i)^k}{(k-1)!}\sum_{r>0} r^{k-1}q^r.$$

Proof. We start with the well-known formula

$$\sum_{n\in\mathbb{Z}}\frac{1}{\tau+n} = \frac{\pi}{\tan\pi\tau}, \qquad \tau\in\mathbb{C}\setminus\mathbb{Z},$$

where the conditionally convergent series on the left-hand side should be interpreted as $\lim_{N\to\infty}\sum_{n\in\mathbb{Z}_N}$. Note that

$$\begin{aligned}\frac{\pi}{\tan \pi\tau} &= \pi\frac{(e^{\pi i\tau}+e^{-\pi i\tau})/2}{(e^{\pi i\tau}-e^{-\pi i\tau})/2i} = -\pi i\left(\frac{1+e^{2\pi i\tau}}{1-e^{2\pi i\tau}}\right)\\ &= -\pi i\left(\frac{1+q}{1-q}\right) = -\pi i\left(1+\frac{2q}{1-q}\right) = -\pi i - 2\pi i\sum_{r>0} q^r. \end{aligned} \tag{12}$$

Now differentiate both sides of the equation

$$\sum_{n\in\mathbb{Z}}\frac{1}{\tau+n} = -2\pi i\sum_{r>0} q^r$$

$k-1$ times with respect to τ to get

$$(-1)^{k-1}(k-1)!\sum_{n\in\mathbb{Z}}\frac{1}{(\tau+n)^k} = (-2\pi i)(2\pi i)^{k-1}\sum_{r>0} r^{k-1}q^r$$

(note that $\frac{d}{d\tau}q^r = \frac{d}{d\tau}e^{2\pi i r\tau} = 2\pi i r q^r$). Dividing by $(-1)^{k-1}(k-1)!$ on both sides gives the claimed formula. $\square$

Proof of Theorem 3.2. We can split up the sum (11) into terms with $m=0$ and with $m\neq 0$, taking care to take the limits in the correct order, so that our calculations remain valid for G_2:

$$G_k(\tau) = \frac{1}{2}\sum_{n\neq 0}\frac{1}{n^k} + \frac{1}{2}\sum_{m\neq 0}\sum_{n\in\mathbb{Z}}\frac{1}{(m\tau+n)^k}.$$

Since k is even, the first sum is

$$\frac{1}{2}\sum_{n\neq 0}\frac{1}{n^k} = \sum_{n>0}\frac{1}{n^k} = \zeta(k). \tag{13}$$

As for the second sum, we use again the fact that k is even, as well as Lipschitz formula (with $m\tau$ substituted for τ) to see that

$$\begin{aligned}\frac{1}{2}\sum_{m\neq 0}\sum_{n\in\mathbb{Z}}\frac{1}{(m\tau+n)^k} &= \sum_{m>0}\sum_{n\in\mathbb{Z}}\frac{1}{(m\tau+n)^k}\\ &= \frac{(-2\pi i)^k}{(k-1)!}\sum_{m>0}\sum_{r>0} r^{k-1}q^{mr}\\ &= \frac{(-2\pi i)^k}{(k-1)!}\sum_{n>0}\sum_{mr=n} r^{k-1}q^{n}\\ &= \frac{(-2\pi i)^k}{(k-1)!}\sum_{n>0}\sigma_{k-1}(n)q^n. \end{aligned} \tag{14}$$

$\square$

Theorem 3.2 gives one of the main motivations for multiple Eisenstein series: The regular Eisenstein series G_k have $\zeta(k)$ as the constant term in its q-series, so one might hope that a suitably defined "multiple" version of Eisenstein series $G(\mathbf{k};\tau)$ will have a q-series with the MZV

$\zeta(\mathbf{k})$ as its constant term. We will show that this is indeed the case; in fact all the Fourier coefficients turn out to be certain $\mathbb{Q}[2\pi i]$-linear combinations of multiple zeta values.

The idea for the definition of multiple Eisenstein series is to replace the sum (11) over λ with an ordered sum over $\lambda_1 \succ \cdots \succ \lambda_r \succ 0$ for a suitable total order on the lattice $\mathbb{Z}\tau + \mathbb{Z}$.

Definition 3.4. Let $\tau \in \mathbb{H}$. We define a relation $\succ$ the lattice $\mathbb{Z}\tau + \mathbb{Z}$ as follows:

$$m_1\tau + n_1\tau > m_2\tau + n_2 \Leftrightarrow (m_1 > m_2) \vee (m_1 = m_2 \wedge n_1 > n_2).$$

Note that this is indeed a total ordering: It is the lexicographical ordering of $\mathbb{Z}\tau + \mathbb{Z} = \{m\tau + n \mid m, n \in \mathbb{Z}\}$ with m having precedence over n. Note also that this makes $\mathbb{Z}_M\tau + \mathbb{Z}_N$ into a finite totally ordered set, which will allow us to make use of Lemma 2.12.

Definition 3.5. Let $\mathbf{k} = (k_1, \ldots, k_r)$ be any index. For $M, N > 0$, we define the **truncated multiple Eisenstein series** $G_{M,N}(\mathbf{k}; \tau) \in \mathcal{H}(\mathbb{H})$ by

$$G_{M,N}(\mathbf{k}; \tau) = \sum_{\substack{\lambda_1 \succ \cdots \succ \lambda_r \succ 0 \\ \lambda_j \in \mathbb{Z}_M\tau + \mathbb{Z}_N}} \frac{1}{\lambda_1^{k_1} \cdots \lambda_r^{k_r}},$$

and $G_{M,N}(\varnothing; \tau) = 1$. If $k_1, \ldots, k_r \geq 2$, we define the **multiple Eisenstein series** $G(\mathbf{k}; \tau) \in \mathcal{H}(\mathbb{H})$ by

$$G(\mathbf{k}; \tau) = \lim_{M \to \infty} \lim_{N \to \infty} G_{M,N}(\mathbf{k}; \tau).$$

Note that if $\mathbf{k} = (k)$ is of depth 1 with k even, then this agrees with the regular Eisenstein series (instead of summing over all nonzero lattice points and dividing by 2, we sum only the "positive" lattice points, namely those with $\lambda \succ 0$). However, when k is odd the lattice point $\lambda \succ 0$ is not cancelled by $-\lambda$ anymore, since $0 \succ -\lambda$, so the latter does not appear in the sum. Thus, $G(k; \tau)$ is not identically 0 for k odd, unlike G_k. In fact $G(k; \tau)$ satisfies Theorem 3.2 for all $k \geq 2$, since the calculations (13) and (14) (skipping the first equality of both) remain valid when k is odd.

We extend $G_{M,N}$ to linear maps $\mathfrak{H}^1 \to \mathcal{H}(\mathbb{H})$ in the usual way. Letting

$$\mathfrak{H}^2 := \mathbb{Q}\langle (z_k)_{k \geq 2} \rangle,$$

we can also extend G to a linear map $\mathfrak{H}^2 \to \mathcal{H}(\mathbb{H})$. Note that $\mathfrak{H}^2$ is in fact closed under the stuffle-product and thus a subalgebra of $\mathfrak{H}^0_*$ (which is in turn a subalgebra of $\mathfrak{H}^1_*$). Applying Lemma 2.12 with $\mathcal{N} = \mathbb{Z}_M\tau + \mathbb{Z}_N$ and $\tilde{f}_\lambda(z_k) = \frac{1}{\lambda^k}$ we see that the $G_{M,N}$ are homomorphisms with respect to the stuffle product, so upon restricting to $\mathfrak{H}^2_*$ and taking limits, we see that G is a homomorphism $\mathfrak{H}^2_* \to \mathcal{H}(\mathbb{H})$.

To state our goals for the remainder of this section, we need to define a couple of auxiliary functions.

Definition 3.6. For an index $\mathbf{k} = (k_1, \ldots, k_r)$ we define

(i)
$$g(\mathbf{k}; \tau) = \sum_{\substack{m_1 > \cdots > m_r \\ n_1, \ldots, n_r > 0}} \frac{n^{k_1 - 1}}{(k_1 - 1)!} \cdots \frac{n^{k_r - 1}}{(k_r - 1)!} q^{m_1 n_1 + \cdots + m_r n_r},$$

(ii)
$$\hat{g}(\mathbf{k}; \tau) = (-2\pi i)^{k_1 + \cdots + k_r} g(\mathbf{k}; \tau).$$

The g-functions can be seen as q-analogues of multiple zeta values, in the sense that as functions of q they satisfy

$$\lim_{q\to 1}(1-q)^{k_1+\cdots+k_r} g(\mathbf{k}) = \zeta(\mathbf{k})$$

when $\mathbf{k}$ is an admissible index, see [BK, Prop. 6.4].

Our first goal, which we will achieve in Section 3.2, is to prove the following

Theorem 3.7 ([GKZ] in depth 2, [Bac3] in general)**.** *For* $\mathbf{k} = k_1, \ldots, k_r \geq 2$, $G(\mathbf{k};\tau)$ *has a* q*-series of the form*

$$G(\mathbf{k};\tau) = \zeta(\mathbf{k}) + \sum_{n>0} a_{\mathbf{k}}(n) q^n,$$

where $a_{\mathbf{k}}(n) \in \mathcal{Z}[\pi i]$*. More precisely,* $G(\mathbf{k};\tau)$ *can be written as a* $\mathcal{Z}$*-linear combination of the* q*-series* $\hat{g}$*.*

The proof given here will be that of [Bac4]. The idea will be to write $G(\mathbf{k};\tau)$ in a different way using convolutions of certain functions. In doing so, we will also get most of the way towards our second goal, which we will achieve in Section 3.3: To produce a "natural" stuffle-regularization for multiple Eisenstein series.

3.1 Multitangent functions and other ingredients

In preparation for the proof of Theorem 3.7, we will need to define several families of functions.

Definition 3.8. Let $\mathbf{k} = (k_1, \ldots, k_r)$ be any index. For $N > 0$, we define the **truncated Hurwitz multiple zeta functions** $\zeta_N(\mathbf{k};x), \overline{\zeta}_N(\mathbf{k};x) \in \mathcal{H}(\mathbb{H})$ by[1]

$$\begin{aligned}
\zeta_N(\mathbf{k};x) &= \sum_{N>n_1>\cdots>n_r>0} \frac{1}{(x+n_1)^{k_1}\cdots(x+n_r)^{k_r}},\\
\overline{\zeta}_N(\mathbf{k};x) &= \sum_{0>n_1>\cdots>n_r>-N} \frac{1}{(x+n_1)^{k_1}\cdots(x+n_r)^{k_r}} = (-1)^{k_1+\cdots+k_r}\zeta_N(k_r,\ldots,k_1;-x),
\end{aligned}$$

and $\zeta_N(\varnothing;x) = \overline{\zeta}_N(\varnothing;x) = 1$. When $\mathbf{k}$ is admissible, we define the **Hurwitz multiple zeta function** $\zeta \in \mathcal{H}(\mathbb{H})$ by

$$\zeta(\mathbf{k};x) = \lim_{N\to\infty}\zeta_N(\mathbf{k};x) = \sum_{n_1>\cdots>n_r>0}\frac{1}{(x+n_1)^{k_1}\cdots(x+n_r)^{k_r}},$$

and similarly when the reverse of $\mathbf{k}$ is admissible, we define $\overline{\zeta}(\mathbf{k};x) = \lim_{N\to\infty}\overline{\zeta}_N(\mathbf{k};x)$.

The proof that the series converge absolutely for (reverses of) admissible indices is analogous to that for the multiple zeta function, and since the convergence is locally uniform in x, we do indeed get holomorphic functions. We extend $\zeta_N, \overline{\zeta}_N$ to linear maps $\mathfrak{H}^1 \to \mathbb{H}$ in the usual way. If we define $\overline{\mathfrak{H}^0}$ to be the subspace of $\mathfrak{H}^1$ generated by words not ending with z_1, including the empty word (note that this is a subalgebra under the stuffle-product), then ζ and $\overline{\zeta}$ are well-defined on $\mathfrak{H}^0$ and $\overline{\mathfrak{H}^0}$ respectively. Analogously to multiple zeta values, an application of Lemma 2.12 shows that $\zeta_N, \overline{\zeta}_N, \zeta$ and $\overline{\zeta}$ are homomorphisms with respect to the stuffle-product.

The next family of functions to define are multitangent functions:

[1] In fact, ζ_N and $\overline{\zeta}_N$ are holomorphic functions on $\mathbb{C}\setminus(-\mathbb{N})$ and $\mathbb{C}\setminus\mathbb{N}$ respectively, as are ζ and $\overline{\zeta}$. Similarly, the multitangent functions defined below are holomorphic functions on $\mathbb{C}\setminus\mathbb{Z}$. However, to keep things consistent, we will restrict all our functions to the upper half-plane.

Definition 3.9. Let $\mathbf{k} = (k_1, \dots, k_r)$ be any index. For $N > 0$, we define the **truncated multitangent function** $\Psi_N(\mathbf{k}; x) \in \mathcal{H}(\mathbb{H})$ by

$$\Psi_N(\mathbf{k}; x) = \sum_{N > n_1 > \cdots > n_r > -N} \frac{1}{(x+n_1)^{k_1} \cdots (x+n_r)^{k_r}},$$

and $\Psi(\varnothing, x) = 1$. When $k_1 \geq 2$ and $k_r \geq 2$, we define the **multitangent function** $\Psi(\mathbf{k}; x) \in \mathcal{H}(\mathbb{H})$ by

$$\Psi(\mathbf{k}; x) = \lim_{N \to \infty} \Psi_N(x) = \sum_{n_1 > \cdots > n_r} \frac{1}{(x+n_1)^{k_1} \cdots (x+n_r)^{k_r}},$$

and we set $\Psi(\varnothing; x) = 1$.

The multitangent functions were studied extensively by [Bou]. Note that if we extend Ψ_N to $\mathfrak{H}^1$ in the usual way, we have

$$\Psi_N(-; x) = \zeta_N(-; x) \star C(-; x) \star \overline{\zeta}_N(-; x), \tag{15}$$

where $C(-; x) : \mathfrak{H}^1 \to \mathcal{H}(\mathbb{H})$ is defined on words by

$$C(w; x) = \begin{cases} 0 & \ell(w) \geq 2, \\ \frac{1}{x^k} & w = z_k, \\ 1 & w = 1. \end{cases}$$

In particular, this makes it clear that $\Psi(k_1, \dots, k_r; x)$ converges when just $k_1 \geq 2$ and $k_r \geq 2$, so Ψ can naturally be defined on $\mathfrak{H}^0 \cap \overline{\mathfrak{H}^0}$, and we get

$$\Psi(-; x) = \zeta(-; x) \star C \star \overline{\zeta}(-; x). \tag{16}$$

Using Lemma 2.10 and Lemma 2.9, this also shows that Ψ_N and Ψ are homomorphisms with respect to the stuffle product (one can also see this directly from Lemma 2.12).

The depth 1 case of multitangents, $\Psi(k; x)$ for $k \geq 2$, are called **monotangents**. Note that by the Lipschitz formula, we have

$$\Psi(k; x) = \frac{(-2\pi i)^k}{(k-1)!} \sum_{n > 0} n^{k-1} q^n, \qquad (k \geq 2, q = e^{2\pi i x}). \tag{17}$$

Comparing this with the definition of $\hat{g}$, we immediately get

Lemma 3.10. *For $k_1, \dots, k_r \geq 2$,*

$$\hat{g}(k_1, \dots, k_r; \tau) = \sum_{m_1 > \cdots > m_r > 0} \Psi(k_1; m_1\tau) \cdots \Psi(k_r; m_r\tau).$$

$\square$

An important property of multitangents proven by [Bou] (and here stated in the equivalent form stated in [Bac4]), is that they can be written in terms of monotangents:

Theorem 3.11 ([Bou, Theorem 3])**.** *For any nonempty index $\mathbf{k} = (k_1, \dots, k_r)$ with $k_1, k_r \geq 2$ of weight $k = k_1 + \cdots + k_r$,*

$$\Psi(\mathbf{k}; x) = \sum_{\substack{1 \leq j < r \\ l_1 + \cdots + l_r = k}} (-1)^{l_1 + \cdots + l_{j-1} + k_j + k} \left(\prod_{\substack{1 \leq i \leq r \\ i \neq j}} \binom{l_i - 1}{k_i - 1} \right) \zeta(l_1, \dots, l_{j-1}) \Psi(l_j; x) \zeta(l_r, l_{r-1}, \dots, l_{j+1}).$$

Moreover, the terms with $\Psi(1; x)$ vanish.

Note that if $l_1 = 1$ or $l_r = 1$, then the binomial coefficient $\binom{l_1-1}{k_2-1}$ or $\binom{l_r-1}{k_r-1}$ is zero, since $k_1 \geq 2$ and $k_r \geq 2$, so the formula does not involve multiple zeta-values with non-admissible indices. The fact that it does not involve $\Psi(1;x)$ is less obvious; it turns out to be a consequence of the antipodal relation with respect to the shuffle product, see [Bac4, Prop. 3.3] (alternatively, it can be proven analytically, which we do in a more general form in Theorem 3.18). We will not actually need the exact formula for reduction into monotangents; all that matters to us it that is of the form

$$\Psi(\mathbf{k};x) = \sum_{j=2}^{\mathrm{wt}(\mathbf{k})} b_{\mathbf{k}}(j)\Psi(j;x), \tag{18}$$

for certain coefficients $b_{\mathbf{k}}(j) \in \mathcal{Z}_{\mathrm{wt}(\mathbf{k})-j}$, as long as the first and last entry in $\mathbf{k}$ are both greater than or equal to 2. A particular consequence of this, and of (17), is that all multitangents have q-series with MZV-coefficients. Interestingly, we also see that the constant terms vanish, except of course in the case of $\Psi(\varnothing;x)$. We shall later define a stuffle-regularized version of multitangents which also have q-series with MZV-coefficients, and where the constant terms vanish in most cases (see Definition 3.15 and Theorem 3.18).

We need to define just one more family of functions before we can get on with the proof of Theorem 3.7.

Definition 3.12. For an index $\mathbf{k} = (k_1, \ldots, k_r)$ and $M, N > 0$, we define $\tilde{g}_{M,N}(\mathbf{k};\tau) \in \mathcal{H}(\mathbb{H})$ by

$$\tilde{g}_{M,N}(\mathbf{k}) = \sum_{\substack{\lambda_1 \succ \cdots \succ \lambda_r \succ 0 \\ \lambda_j \in (\mathbb{Z}_M\tau + \mathbb{Z}_N)\cap\mathbb{H}}} \frac{1}{\lambda_1^{k_1}\cdots\lambda_r^{k_r}},$$

and $g_{M,N}(\varnothing;\tau) = 1$. If $k_1, \ldots, k_r \geq 2$, we define $\tilde{g}(\mathbf{k};\tau) \in \mathcal{H}(\mathbb{H})$ by

$$\tilde{g}(\mathbf{k};\tau) = \lim_{M\to\infty}\lim_{N\to\infty} \tilde{g}_{M,N}(\mathbf{k};\tau).$$

We extend $\tilde{g}_{M,N}$ (resp. $\tilde{g}$) to linear maps $\mathfrak{H}^1 \to \mathcal{H}(\mathbb{H})$ (resp. $\mathfrak{H}^2 \to \mathcal{H}(\mathbb{H})$) in the usual way. Note that the definition of $\tilde{g}$ resembles that of G, the only difference is that in the former we only sum over lattice points strictly in the upper half-plane, while in the latter we also include the lattice points on the positive real axis.

3.2 q-series expansion of MES

With all the ingredients in place, we are now ready to prove Theorem 3.7.

Lemma 3.13. *For $M, N > 0$,*

$$G_{M,N} = \tilde{g}_{M,N} \star \zeta_N,$$

where ζ_N is the truncated multizeta-map $\zeta_N : \mathfrak{H}^1 \to \mathbb{Q}$, seen as a map $\zeta_N : \mathfrak{H}^1 \to \mathcal{H}(\mathbb{H})$ by including $\mathbb{Q}$ in $\mathcal{H}(\mathbb{H})$ as constant functions. Thus,

$$G = \tilde{g} \star \zeta,$$

where ζ is the multizeta-map $\zeta : \mathfrak{H}^0 \to \mathcal{Z}$, seen as a map $\zeta : \mathfrak{H}^2 \to \mathcal{H}(\mathbb{H})$.

Proof. Let $w = z_{k_1}\cdots z_{k_r}$ be a word in $\mathfrak{H}^1$. We can write $G_{M,N}(w;\tau)$ as a sum over $m_1, \ldots, m_r$ and $n_1, \ldots, n_r$ as

$$G_{M,N}(w;\tau) = \sum_{\substack{M > m_1 \geq \cdots \geq m_r \geq 0 \\ n_1,\ldots,n_r \in \mathbb{Z}_N \\ m_1\tau+n_1 \succ \cdots \succ m_r\tau+n_r \succ 0}} \frac{1}{(m_1\tau+n_1)^{k_1}\cdots(m_r\tau+n_r)^{k_r}},$$

Split this sum into $r+1$ cases according to how many m_j are zero:

$$G_{M,N}(w;\tau) = \left[\sum_{\substack{M>m_1\geq\cdots\geq m_r>0\\ n_1,\dots,n_r\in\mathbb{Z}_N\\ m_1\tau+n_1\succ\cdots\succ m_r\tau+n_r\succ 0}} + \sum_{\substack{M>m_1\geq\cdots\geq m_{r-1}>m_r=0\\ n_1,\dots,n_r\in\mathbb{Z}_N\\ m_1\tau+n_1\succ\cdots\succ m_r\tau+n_r\succ 0}} \right.$$
$$\left. + \ldots + \sum_{\substack{m_1=\cdots=m_r=0\\ n_1,\dots,n_r\in\mathbb{Z}_N\\ m_1\tau+n_1\succ\cdots\succ m_r\tau+n_r\succ 0}} \right] \frac{1}{(m_1\tau+n_1)^{k_1}\cdots(m_r\tau+n_r)^{k_r}}$$

The sum with $M > m_1 \geq \cdots \geq m_r > 0$ is nothing but $\tilde{g}_{M,N}(w;\tau)$, as requiring all m_j to be strictly possitive is equivalent to requiring that all $m_j\tau+n_j$ lie strictly in the upper half-plane. In the sum with $m_1 = \cdots = m_r = 0$, the requirement that $m_1\tau+n_1 \succ \cdots \succ m_r\tau+n_r \succ 0$ is equivalent to $n_1 > \ldots n_r > 0$, so we have

$$\sum_{\substack{m_1=\cdots=m_r=0\\ n_1,\dots,n_r\in\mathbb{Z}_N\\ m_1\tau+n_1\succ\cdots\succ m_r\tau+n_r\succ 0}} \frac{1}{(m_1\tau+n_1)^{k_1}\cdots(m_r\tau+n_r)^{k_r}} = \sum_{N>n_1>\cdots>n_r>0} \frac{1}{n_1^{k_1}\cdots n_r^{k_r}} = \zeta_N(w).$$

In general, the sum with $M > m_1 \geq \cdots \geq m_\ell > m_{\ell+1} = \cdots = m_r = 0$ is

$$\begin{aligned}\sum_{\substack{M>m_1\geq\cdots\geq m_\ell>m_{\ell+1}=\cdots=m_r=0\\ n_1,\dots,n_r\in\mathbb{Z}_N\\ m_1\tau+n_1\succ\cdots\succ m_r\tau+n_r\succ 0}} \frac{1}{(m_1\tau+n_1)^{k_1}\cdots(m_r\tau+n_r)^{k_r}} &= \sum_{\substack{\lambda_1\succ\cdots\succ\lambda_\ell\succ 0\\ \lambda_1,\dots,\lambda_\ell\in(\mathbb{Z}_M\tau+\mathbb{Z}_N)\cap\mathbb{H}\\ N>n_{\ell+1}>\cdots>n_r>0}} \frac{1}{\lambda_1^{k_1}\cdots\lambda_\ell^{k_\ell} n_{\ell+1}^{k_{\ell+1}}\cdots n_r^{k_r}}\\ &= \tilde{g}_{M,N}(z_{k_1}\cdots z_{k_\ell};\tau)\zeta_N(z_{k_{\ell+1}}\cdots z_{k_r}),\end{aligned}$$

so we get

$$\begin{aligned}G_{M,N}(w;\tau) &= \tilde{g}_{M,N}(z_{k_1}\cdots z_{k_r};\tau) + \tilde{g}_{M,N}(z_{k_1}\cdots z_{k_{r-1}};\tau)\zeta_N(z_{k_r}) + \cdots + \zeta_N(z_{k_1}\cdots z_{k_r})\\ &= (\tilde{g}_{M,N}\star\zeta)(w).\end{aligned}$$

Thus $G_{M,N} = \tilde{g}_{M,N}\star\zeta_N$. Restricting to $\mathfrak{H}^2$ and taking limits gives $G = \tilde{g}\star\zeta$. □

Lemma 3.14. *For $M, N > 0$,*

$$\tilde{g}_{M,N}(-;\tau) = \mathop{\bigstar}_{m=1}^{M-1} \Psi_N(-;m\tau),$$

where the convolution should be taken in decreasing order, ie

$$\mathop{\bigstar}_{m=1}^{M-1} \Psi_N(-;m\tau) := \Psi_N(-;(M-1)\tau)\star\cdots\star\Psi_N(-;\tau).$$

Proof. For a word $w = z_{k_1}\cdots z_{k_r}$ in $\mathfrak{H}^0$, we can write $\tilde{g}_{M,N}(w;\tau)$ as a sum over m_j and n_j:

$$\tilde{g}_{M,N}(w;\tau) = \sum_{\substack{M>m_1\geq\cdots\geq m_r>0\\ n_1,\dots,n_r\in\mathbb{Z}_N\\ m_j\tau+n_j\succ m_{j+1}\tau+n_{j+1}\succ 0}} \frac{1}{(m_1\tau+n_1)^{k_1}\cdots(m_r\tau+n_r)^{k_r}}.$$

Again, we will split this up into several sums, this time into 2^{r-1} sums according to which inequalities $m_j \geq m_{j+1}$ are ">" and which are "=". Each of these sums will correspond to a way of splitting w up into nonempty subwords. Namely, if we have a string of equalities $m_{j-1} > m_j = \cdots = m_{j+\ell} > m_{j+\ell+1}$, then the requirement that $m_{j-1}\tau + n_{j-1} \succ \cdots \succ m_{j+\ell+1}\tau + n_{j+\ell+1}$ is equivalent to $n_j > \cdots > n_{j+\ell}$, so the sum will have a factor of

$$\sum_{\substack{M > m_j = \cdots = m_{j+\ell} > 0 \\ N > n_j > \cdots > n_{j+\ell} > -N}} \frac{1}{(m_j\tau + n_j)^{k_j} \cdots (m_{j+\ell}\tau + n_{j+\ell})^{k_{j+\ell}}} = \sum_{\substack{M > m > 0 \\ N > n_j > \cdots > n_{j+\ell} > -N}} \frac{1}{(m\tau + n_j)^{k_j} \cdots (m\tau + n_{j+\ell})^{k_{j+\ell}}}$$
$$= \sum_{M > m > 0} \Psi_N(z_{k_j} \cdots z_{k_{j+\ell}}; m\tau).$$

In general, if our string of inequalities is

$$M > m_1 = \cdots = m_{\ell_1} > m_{\ell_1+1} = \cdots = m_{\ell_1+\ell_2} > \cdots > m_{\ell_1+\cdots+\ell_{s-1}+1} = \cdots = m_{\ell_1+\cdots+\ell_s} > 0,$$

where $\ell_1 + \cdots + \ell_s = r$, then the corresponding sum over the n_j is

$$\sum_{M > \tilde{m}_1 > \cdots > \tilde{m}_s > 0} \sum_{\substack{N > n_1 > \cdots > n_{\ell_1} > -N \\ N > n_{\ell_1+1} > \cdots > n_{\ell_1+\ell_2} > -N \\ \vdots \\ N > n_{\ell_1+\cdots+\ell_{s-1}+1} > \cdots > n_{\ell_1+\cdots+\ell_s} > -N}} \prod_{i=1}^{s} \frac{1}{(\tilde{m}_i\tau + n_{\ell_1+\cdots+\ell_{i-1}+1})^{k_{\ell_1+\cdots+\ell_{i-1}+1}} \cdots (\tilde{m}_i\tau + n_{\ell_1+\cdots+\ell_i})^{k_{\ell_1+\cdots+\ell_i}}}$$
$$= \sum_{M > \tilde{m}_1 > \cdots > \tilde{m}_s > 0} \prod_{i=1}^{s} \psi_N(z_{k_{\ell_1+\cdots+\ell_{i-1}+1}} \cdots z_{k_{\ell_1+\cdots+\ell_i}}; \tilde{m}_i\tau)$$

where $\tilde{m}_j = m_{\ell_1+\cdots+\ell_j}$. This corresponds exactly to the terms in the convolution $\bigstar_{m=1}^{M-1} \Psi_N(-; m\tau)$ coming from splitting up w into s nonempty words of lengths $\ell_1, \ldots, \ell_s$, with the $\tilde{m}_j$ telling us at which indices to insert these nonempty words (i.e. $w = w_1 \cdots w_{M-1}$ with $\ell(w_{\tilde{m}_j}) = \ell_j$, and all other words being empty). Adding up all the 2^{r-1} sums of this type thus gives us exactly $\left(\bigstar_{m=1}^{M-1} \Psi_N(-; m\tau)\right)(w)$. □

Proof of Theorem 3.7. Let $w = z_{k_1} \cdots z_{k_r}$ be a nonempty word in $\mathfrak{H}^2$. From Lemma 3.14, we have

$$\tilde{g}_{M,N}(w; \tau) = \sum_{w_{M-1} \cdots w_1 = w} \prod_{m=1}^{M-1} \Psi_N(w_j; j\tau) = \sum_{s=1}^{r} \sum_{\substack{w_1 \cdots w_s = w \\ w_j \neq 1}} \sum_{M > m_1 > \cdots > m_s > 0} \prod_{j=1}^{s} \Psi_N(w_j; m_j\tau),$$

where the second equality is since $\Psi(1; j\tau)$ just contributes a factor of 1 to the product, so we can instead do the sum only over ways of splitting w into s nonempty words for $w_1, \ldots, w_s$, with $M > m_1 > \cdots > m_s > 0$ telling us at which indices to place these nonempty words. Taking limits, we get

$$\tilde{g}(w; \tau) = \sum_{s=1}^{r} \sum_{\substack{w_1 \cdots w_s = w \\ w_j \neq 1}} \sum_{m_1 > \cdots > m_s > 0} \prod_{j=1}^{s} \Psi(w_j; m_j\tau).$$

By Theorem 3.11, we can write each $\Psi(w_j; m_j\tau)$ as a $\mathcal{Z}$-linear combination of monotangents $\Psi(\ell; m_j\tau)$ with $\ell \geq 2$, so $\tilde{g}(w; \tau)$ is a $\mathcal{Z}$-linear combination of terms of the form

$$\sum_{m_1 > \cdots > m_s > 0} \prod_{j=1}^{s} \Psi(\ell_j; m_j\tau) = \hat{g}(\ell_1, \ldots, \ell_s; \tau), \qquad \ell_1, \ldots, \ell_s \geq 2,$$

where the second equality is by Lemma 3.10.

It follows from Lemma 3.13 that $G(w;\tau)$ is a $\mathcal{Z}$-linear combination of $\tilde{g}$-functions, and therefore, by the above, it is a $\mathcal{Z}$-linear combination of $\hat{g}$-functions, as claimed. In particular, $G(w;\tau)$ has a q-series with coefficients in $\mathcal{Z}[\pi i]$. As for the constant term, we observe in (17) that the monotangents have no constant term in their q-series, so neither does $\tilde{g}(w;\tau)$ for nonempty words w in $\mathfrak{H}^2$. Thus, the only contribution to the constant term of $G(w;\tau) = (\tilde{g} \star \zeta)(w)$ is

$$\tilde{g}(\varnothing;\tau)\zeta(w) = \zeta(w).$$

□

3.3 Stuffle-regularized MES

By Lemmas 3.13 and 3.14, the MES-map $G : \mathfrak{H}^2_* \to \mathcal{H}(\mathbb{H})$ is a convolution of the MZV-map ζ with the map $\tilde{g}$, which is in turn a limit of convolutions of multitangents, which are in turn convolutions of the Hurwitz zeta functions by (16). The Hurwitz zeta functions $\zeta, \overline{\zeta}$ are homomorphisms of $\mathfrak{H}^0_*$ and $\overline{\mathfrak{H}}^0_*$ respectively, so using[2] Lemma 2.16, they can be extended to homomorphisms $\zeta^*, \overline{\zeta}^* : \mathfrak{H}^1_* \to \mathcal{H}(\mathbb{H})$ by choosing values for $\zeta^*(1;x)$ and $\overline{\zeta}^*(1;x)$. Working from the bottom up, we will use this to regularize G by regularizing its convolution ingredients.

Definition 3.15. We define the following homomorphisms $\mathfrak{H}^1_* \to \mathcal{H}(\mathbb{H})$.

(i) For $N > 0$, define $\zeta_N * (-;x), \zeta^*(-;x) : \mathfrak{H}^1_* \to \mathcal{H}(\mathbb{H})$ by $\zeta^*_N(w;x) = \zeta_N(w;x)$ and $\zeta^*(w;x) = \zeta(w;x)$ for $w \in \mathfrak{H}^0$; and

$$\zeta_N(z_1;x) = \sum_{N>n>0} \left(\frac{1}{x+n} - \frac{1}{n} \right),$$

$$\zeta(z_1;x) = \sum_{n>0} \left(\frac{1}{x+n} - \frac{1}{n} \right).$$

(ii) For $N > 0$, define $\overline{\zeta}_N * (-;x), \overline{\zeta}^*(-;x) : \mathfrak{H}^1_* \to \mathcal{H}(\mathbb{H})$ by $\overline{\zeta}^*_N(w;x) = \overline{\zeta}_N(w;x)$ and $\overline{\zeta}^*(w;x) = \overline{\zeta}(w;x)$ for $w \in \overline{\mathfrak{H}}^0$; and

$$\overline{\zeta}_N(z_1;x) = \sum_{0>n>-N} \left(\frac{1}{x+n} - \frac{1}{n} \right),$$

$$\overline{\zeta}(z_1;x) = \sum_{0>n} \left(\frac{1}{x+n} - \frac{1}{n} \right).$$

(iii) Define

$$\Psi^*_N(-;x) = \zeta^*_N(-;x) \star C(-;x) \star \overline{\zeta}^*_N(-;x)$$

for $N > 0$, and

$$\Psi^*(-;x) = \zeta^*(-;x) \star C \star \overline{\zeta}(-;x).$$

[2]Lemma 2.16 remains valid if one replaces $\mathfrak{H}^0$ with $\overline{\mathfrak{H}}^0$. The proof is the same, except that one should use Theorem 2.14(ii) instead of Theorem 2.14(i)

(iv) For $M, N > 0$, define

$$\tilde{g}^*_{M,N}(-;\tau) = \mathop{\bigstar}_{m=1}^{M-1} \Psi^*_N(-;m\tau),$$

and for $M > 0$, define

$$\tilde{g}^*_M(-;\tau) = \mathop{\bigstar}_{m=1}^{M-1} \Psi^*(-;m\tau),$$

where, as in Lemma 3.14, the convolutions should be taken in decreasing order.

Note that all the truncated maps defined above converge to their untruncated versions (or, in the case of $\tilde{g}^*_{M,N}$, to a version that is only truncated by M) as $N \to \infty$. In the case of Ψ^*_N and $\tilde{g}^*_{M,N}$, this is clear. In the case of ζ^* and $\overline{\zeta}$, it follows from the uniqueness in Lemma 2.16, as both $\zeta^*(-;x)$ and $\lim_{N\to\infty} \zeta^*_N(-;x)$ are homomorphic extensions of $\zeta(-;x)$ that map z_1 to $\sum_{n>0} \left(\frac{1}{x+n} - \frac{1}{n}\right)$ (and similarly for $\overline{\zeta}$).

Our choice of regularization of the Hurwitz zeta functions, and thus of the multitangent functions, is the main one studied in [Bou]. While we could in principle have made any choice of $\zeta(z_1;x)$ and $\overline{\zeta}(z_1;x)$, this particular choice is natural in many ways, as many of the properties of multitangents are retained. Importantly, we keep the reduction into monotangents.

Theorem 3.16 ([Bou, Theorem 6])**.** *For any index* $\mathbf{k} = (k_1, \ldots, k_r)$*, there exist coefficients* $b_{\mathbf{k}}(j) \in \mathcal{Z}_{\mathrm{wt}(\mathbf{k})-j}$ *such that*

$$\Psi^*(\mathbf{k};x) = \delta_{\mathbf{k}} + \sum_{j=1}^{\mathrm{wt}(\mathbf{k})} b_{\mathbf{k}}(j)\Psi^*(j;x),$$

where

$$\delta_{\mathbf{k}} = \begin{cases} \frac{(i\pi)^r}{r!} & \text{if } \mathbf{k} = (\underbrace{1,\ldots,1}_{r}) \text{ and } r \text{ is even,} \\ 0 & \text{otherwise.} \end{cases}$$

The coefficients $b_{\mathbf{k}}(j)$ are in fact the same as in Theorem 3.11, except that ζ should be replaced with $\zeta^{ш}$. The difference from Theorem 3.11, then, is the occasional presence of a constant term $\delta_{\mathbf{k}}$, as well as the presence of $\Psi^*(1;x)$. Note that with out choice of regularization

$$\begin{aligned}
\Psi^*(1;x) &= \zeta(1;x) + \frac{1}{x} + \overline{\zeta}(1;x) \\
&= \frac{1}{x} + \lim_{N\to\infty} \left[\sum_{N>n>0} \left(\frac{1}{x+n} - \frac{1}{n} \right) + \sum_{0>n>-N} \left(\frac{1}{x+n} - \frac{1}{n} \right) \right] \\
&= \frac{1}{x} + \lim_{N\to\infty} \left[\sum_{N>n>0} \frac{1}{x+n} + \sum_{0>n>-N} \frac{1}{x+n} \right] \\
&= \lim_{N\to\infty} \sum_{N>n>-N} \frac{1}{x+n} \\
&= \frac{\pi}{\tan \pi x},
\end{aligned}$$

so by (12), we have

$$\Psi^*(1;x) = -\pi i - 2\pi i \sum_{r>0} q^r. \tag{19}$$

Since the other monotangents also have q-series by (17), it follows from the monotangent decomposition that all our regularized multitangents have q-series. We shall show that the q^0-term in fact vanishes in most cases, which will be crucial for the convergence of $\tilde{g}_M^*$ as $M \to \infty$.

Lemma 3.17. *As $x \to i\infty$ along the imaginary axis, we have the following asymptotics for any $\varepsilon > 0$:*

(i) If $\mathbf{k}$ is admissible, then $\zeta(\mathbf{k}; x) = o(x^{-\operatorname{wt}(\mathbf{k})+\operatorname{dep}(\mathbf{k})+\varepsilon})$. If the reverse of $\mathbf{k}$ is admissible, then $\overline{\zeta}(\mathbf{k}; x) = o(x^{-\operatorname{wt}(\mathbf{k})+\operatorname{dep}(\mathbf{k})+\varepsilon})$

(ii) For any index $\mathbf{k}$, $\zeta(\mathbf{k}; x) = o(x^{\varepsilon})$ and $\overline{\zeta}(\mathbf{k}; x) = o(x^{\varepsilon})$.

(iii) if $\mathbf{k} = (k_1, \ldots, k_r)$ with $k_j > 1$ for at least one j, then $\zeta^*(\mathbf{k}; x) = o(x^{-1+\varepsilon})$ and $\overline{\zeta}^*(\mathbf{k}; x) = o(x^{-1+\varepsilon})$

Proof. We only prove the statements about ζ, as the statements about $\overline{\zeta}$ are similar.

(i) For $\mathbf{k} = \varnothing$, this is clear, since $1 = o(x^{\varepsilon})$ as $x \to i\infty$. Suppose then that $\mathbf{k} = (k_1, \ldots, k_r)$ with $k_1 \geq 2$. Let $k = k_1 + \cdots + k_r$. Then for x on the imaginary axis, we have

$$\begin{aligned}
\left|x^{k-r-\varepsilon}\zeta(\mathbf{k}; x)\right| &= \left| x^{-\varepsilon/2} \sum_{n_1 > \cdots > n_r > 0} \frac{x^{k_1-1-\varepsilon/2} x^{k_2-1} \cdots x^{k_r-1}}{(x+n_1)^{k_1}(x+n_2)^{k_2} \cdots (x+n_r)^{k_r}} \right| \\
&\leq \left|x^{-\varepsilon/2}\right| \sum_{n_1 > \cdots > n_r > 0} \frac{|x|^{k_1-1-\varepsilon/2}|x|^{k_2-1} \cdots |x|^{k_r-1}}{|x+n_1|^{k_1}|x+n_2|^{k_2} \cdots |x+n_r|^{k_r}} \\
&\leq \left|x^{-\varepsilon/2}\right| \sum_{n_1 > \cdots > n_r > 0} \frac{|x+n_1|^{k_1-1-\varepsilon/2}|x+n_2|^{k_2-1} \cdots |x+n_r|^{k_r-1}}{|x+n_1|^{k_1}|x+n_2|^{k_2} \cdots |x+n_r|^{k_r}} \\
&= \left|x^{-\varepsilon/2}\right| \sum_{n_1 > \cdots > n_r > 0} \frac{1}{|x+n_1|^{1+\varepsilon/2}|x+n_2| \cdots |x+n_r|} \\
&\leq \left|x^{-\varepsilon/2}\right| \sum_{n_1 > \cdots > n_r > 0} \frac{1}{|n_1|^{1+\varepsilon/2}|n_2| \cdots |n_r|} \qquad \xrightarrow{x \to i\infty} 0,
\end{aligned}$$

since the sum on the last line converges to a finite value by (5), while of course $x^{-\varepsilon/2} \to 0$ as $x \to i\infty$ along the imaginary axis. Thus $\zeta(\mathbf{k}; x) = o(x^{-k+r+\varepsilon})$ as $x \to i\infty$ along the imaginary axis.

(ii) Note that by termwise differentiation, $\frac{d}{dx}\zeta^*(1; x) = -\zeta(2; x)$. Thus, by L'Hopital's rule and (i),

$$\lim_{\substack{x \to i\infty \\ \operatorname{Re}(x)=0}} \frac{\zeta^*(1; x)}{x^{\varepsilon}} = \lim_{\substack{x \to i\infty \\ \operatorname{Re}(x)=0}} \frac{-\zeta(2; x)}{\varepsilon x^{-1+\varepsilon}} = 0,$$

so $\zeta^*(1; x) = o(x^{\varepsilon})$ as $x \to i\infty$ along the imaginary axis. Since in general $\zeta^*(\mathbf{k}; x)$ is a polynomial in $\zeta^*(1; x) = o(x^{\varepsilon})$ whose coefficients are linear combinations of terms of the form $\zeta(\mathbf{k}'; x) = o(x^{\varepsilon})$ for admissible indices $\mathbf{k}'$, we see that $\zeta^*(\mathbf{k}; x) = o(x^{m\varepsilon})$ for some integer $m > 0$ as $x \to i\infty$ along the imaginary axis. As ε can be picked arbitrarily small, we get $\zeta^*(\mathbf{k}; x) = o(x^{\varepsilon})$ as $x \to i\infty$ along the imaginary axis.

(iii) Let $k = k_1 + \cdots + k_r$. Note that since at least one k_j is greater than 1, we have $k > r$. Write $z_{k_1} \cdots z_{k_r} = \sum_{j=0}^{r} \alpha_j z_1^j$ with $\alpha_j \in \mathfrak{H}^0$. Each α_j is a weight-homogeneous linear combination of words $z_{\mathbf{k}'} \in \mathfrak{H}^0$ of weight $k - j > r - j \geq 0$. In particular, each $z_{\mathbf{k}'}$ is a nonempty word, so $-\operatorname{wt}(\mathbf{k}') + \operatorname{dep}(\mathbf{k}') \leq -1$. It then follows from (i) that

$$\zeta(\alpha_j; x) = o(x^{-1+\varepsilon})$$

as $x \to i\infty$ along the imaginary axis. From this and (ii), we get

$$\zeta^*(\mathbf{k}; x) = \sum_{j=0}^{r} \zeta(\alpha_j; x)\zeta^*(1; x)^j = o(x^{-1+(r+1)\varepsilon})$$

as $x \to i\infty$ along the imaginary axis.. Since ε can be picked arbitrarily small, we get $\zeta^*(\mathbf{k}; x) = o(x^{-1+x})$ as $x \to i\infty$ along the imaginary axis.

□

Theorem 3.18. *If $\mathbf{k} = (k_1, \ldots, k_r)$ is an index with $k_j > 1$ for at least one j, then the constant term in the q-series of $\Psi^*(\mathbf{k}; x)$ is zero.*

Proof. The constant term can be extracted from $\Psi^*(\mathbf{k}; x)$ by taking the limit $x \to i\infty$. It will suffice to take this limit along the imaginary axis. Letting $w = z_{k_1} \cdots z_{k_r}$, we have

$$\Psi^*(w; x) = \sum_{w_1 w_2 w_3 = w} \zeta^*(w_1; x) C(w_2; x) \overline{\zeta}^*(w_3; x).$$

In each term of the sum, all three factors are at worst $o(x^{\varepsilon})$ as $x \to i\infty$ along the imaginary axis, and at least one of the factors is $o(x^{-1+\varepsilon})$. Namely, whichever factor has a z_{k_j} with $k_j > 1$ in its word will be $o(x^{-1+\varepsilon})$: For ζ and $\overline{\zeta}$, this is by Lemma 3.17, and for C it is trivial. Thus, $\Psi^*(w; x) = o(x^{-1+3\varepsilon})$ as $x \to i\infty$ along the imaginary axis. Picking ε small enough that $-1 + 3\varepsilon \leq 0$, this shows that $\lim_{x \to i\infty} \Psi^*(w; x) = 0$. □

Note that the only contributions to the q^0-coefficient in the monotangent decomposition of $\Psi^*(\mathbf{k}; x)$ come from $\delta_{\mathbf{k}}$ and $\Psi^*(1; x)$. In particular, it follows that the $\Psi^*(1; x)$-term vanishes whenever $\mathbf{k}$ is not of the form $(1, \ldots, 1)$. This implies a relation among multiple zeta values, namely [Bac3, Prop. 3.3] can be seen to hold for any index not of the form $(1, \ldots, 1)$, not just when the first and last entry are greater than 1.

Returning to the matter of MES-regularization, Theorem 3.18 implies the following:

Proposition 3.19. *If $w \in \mathfrak{H}^0$, then $\tilde{g}^*_M(w; \tau)$ converges to a holomorphic function on the upper half-plane as $M \to \infty$.*

Proof. Let $w = z_{k_1} \cdots z_{k_r}$ with $k_1 \geq 2$. Then

$$\begin{aligned} \tilde{g}^*_M(w; \tau) &= \sum_{w_{M-1} \cdots w_1 = w} \prod_{j=1}^{M-1} \Psi^*(w_j; j\tau) \\ &= \sum_{M > m > 0} \sum_{\ell=1}^{r} \Psi^*(z_{k_1} \cdots z_{k_\ell}; m\tau) \sum_{w_{m-1} \cdots w_1 = z_{k_{\ell+1}} \cdots z_{k_r}} \prod_{j=1}^{m-1} \Psi^*(w_j; j\tau), \end{aligned} \tag{20}$$

where m and ℓ are the index and length of the first nonempty word. Since $k_1 \geq 2$, it follows from Theorem 3.18 that the q-series of $\Psi^*(z_{k_1} \cdots z_{k_\ell}; \tau)$ is of the form $\sum_{r \geq 1} a_r q^r$, so the q-series of $\Psi^*(z_{k_1} \cdots z_{k_\ell}; m\tau)$ is of the form $\sum_{r \geq 1} a_r q^{mr}$. Thus, when we let $M \to \infty$, we get a sum $\sum_{m > 0}$

where only the first r terms can contribute to the coefficient of q^r. This implies that the limit exists as a formal power series $\sum_{r\geq 0} c_r q^r$.

To see that it converges to a holomorphic function of the upper half-plane, we need to show that this formal power series has radius of convergence at least 1, since $\operatorname{Im}(\tau) > 0$ is equivalent to $|q| < 1$. To show this, it will suffice to show that the coefficients c_r grow at most polynomially in r. Note that, by (17) and (19), the q-series coefficients of monotangents are of polynomial growth, and so it follows from Theorem 3.16 that the q-series coefficients of all multitangents are of polynomial growth. Since only finitely many different multitangents appear in (20), we can pick the implied constants so as to work for all of them. Noting also that the number of nonzero terms in the inner sum of (20) is of polynomial growth in m, it follows that the coefficients c_r are of polynomial growth in r. $\square$

We can thus use Lemma 2.16 to extend $\tilde{g}$ to $\mathfrak{H}_1$ by simply picking a value for $\tilde{g}^*(1;\tau)$. The most general choice is to let $\tilde{g}^*(1;\tau)$ be a free variable.

Definition 3.20. We define the following homomorphisms:

(i) Define $\tilde{g}^* : \mathfrak{H}^1_* \to \mathcal{H}(\mathbb{H})[Y]$ by $\tilde{g}^*(w;\tau;Y) = \lim_{M\to\infty} \tilde{g}^*_M(w;\tau)$ for $w \in \mathfrak{H}^0$, and $\tilde{g}^*(z_1;\tau;Y) = Y$.

(ii) Define $G^* : \mathfrak{H}^1_* \to \mathcal{H}(\mathbb{H})[X,Y]$ by

$$G^*(-;\tau;X,Y) = \tilde{g}^*(-;\tau,Y) \star \zeta^*(-;X),$$

where $\tilde{g}^* : \mathfrak{H}^1_* \to \mathcal{H}(\mathbb{H})[Y] \hookrightarrow \mathcal{H}(\mathbb{H})[X,Y]$ is the map defined above, and $\zeta^* : \mathfrak{H}^1_* \to \mathcal{Z}[X] \hookrightarrow \mathcal{H}(\mathbb{H})[X,Y]$ is the stuffle-regularized multizeta-map of Definition 2.17.

By construction, we have

Theorem 3.21. $G^*(w;\tau;X,Y) = G(w;\tau)$ *whenever* $w \in \mathfrak{H}^2$. $\square$

Even though we are formally free to choose any value for the variables X and Y, some values are more "natural" than others. For example, since we picked the function $\sum_{n>0}\left(\frac{1}{x+n} - \frac{1}{n}\right)$ as $\zeta^*(1;x)$ (Hurwitz multizeta function), setting $X = \sum_{n>0}\left(\frac{1}{0+n} - \frac{1}{n}\right) = 0$ seems most natural, as this preserves the property that evaluating a Hurwitz multiple zeta function at $x = 0$ gives the corresponding multiple zeta value $\zeta^*(1) = 0$. If one wants to retain this property while also keeping the freedom of choice of X, one would need to pick a more general choice of $\zeta^*(1;x)$, e.g. $\zeta^*(1;x;X) = X + \sum_{n>0}\left(\frac{1}{x+n} - \frac{1}{n}\right)$. However, this would change the behaviour of the multitangent functions, and might thus cause issues with the convergence of $\tilde{g}_M$ as $M \to \infty$. Still, it is possible that other choices of $\zeta^*(1;x)$ and $\overline{\zeta}^*(1;x)$ than the ones we made would preserve the periodicity of the multitangent functions and the polynomial growth of their Fourier coefficients, as well as the growth conditions of Lemma 3.17. In that case, the proofs of Theorem 3.18 and Proposition 3.19 would remain valid, and one would thus obtain a different version of G^*.

As for Y, note that Theorem 3.2 can be restated as

$$G(k;\tau) = \zeta(k) + \hat{g}(k;\tau)$$

for $k \geq 2$ (including odd values). On the other hand, we have

$$G^*(k;\tau;X,Y) = \tilde{g}^*(k;\tau;Y) + \zeta^*(k;X),$$

and in particular $G^*(1;\tau;X,Y) = X + Y$, which suggests that the choice $Y = \hat{g}(1;\tau)$ would be natural, as it extends Theorem 3.2 to also hold for $k = 1$.

Let us finish by calculating a few examples. We will express some stuffle-regularized MES in terms of the following "double-indexed" g-functions, which were introduced in [Bac1].

Definition 3.22. For integers $k_1, \ldots, k_r$ and $d_1, \ldots, d_r$ with $k_j \geq 1$ and $d_j \geq 0$, we define

$$g\left(\begin{smallmatrix}k_1,\ldots,k_r\\d_1,\ldots,d_r\end{smallmatrix};\tau\right) = \sum_{\substack{m_1>\cdots>m_r>0\\n_1,\ldots,n_r>0}} \frac{n_1^{k_1-1}m_1^{d_1}}{(k_1-1)!}\cdots\frac{n_r^{k_r-1}m_r^{d_r}}{(k_r-1)!}q^{m_1n_1+\cdots+m_rn_r},$$

$$\hat{g}\left(\begin{smallmatrix}k_1,\ldots,k_r\\d_1,\ldots,d_r\end{smallmatrix};\tau\right) = (-2\pi i)^{k_1+d_1+\cdots+k_r+d_r}g\left(\begin{smallmatrix}k_1,\ldots,k_r\\d_1,\ldots,d_r\end{smallmatrix};\tau\right).$$

Note that the "single-indexed" g-functions of Definition 3.12 are the special cases $d_1 = \cdots = d_r = 0$.

Example 3.23.

(i) Let us start by calculating $G^*(2,1;\tau;X,Y)$. We have

$$\begin{aligned}G^*(2,1;\tau;X,Y) &= \tilde{g}^*(2,1;\tau;Y) + \tilde{g}^*(2;\tau;Y)\zeta^*(1;X) + \zeta^*(2,1;X)\\ &= \lim_{M\to\infty}\tilde{g}^*_M(2,1;\tau) + \hat{g}(2;\tau)X + \zeta(2,1)\end{aligned} \tag{21}$$

To compute $\lim_{M\to\infty}\tilde{g}^*_M(2,1;\tau)$, we note that

$$\tilde{g}^*_M(2,1;\tau) = \sum_{M>m_1>m_2>0}\Psi^*(2;m_1\tau)\Psi^*(1;m_2\tau) + \sum_{M>m_1>0}\Psi^*(2,1;m_1\tau),$$

where the two sums are over ways of splitting the word z_2z_1 up into two nonempty words and one nonempty word respectively. It turns out that $\Psi^*(2,1;x)$ is identically zero, which can be seen by explicitly computing its reduction into monotangents ([Bou, Table 7] gives the monotangent reductions for divergent multitangent functions of weight $2,3,4,5$), so the second sum vanishes. For the first sum, we can use the q-series (17) and (19) to get in the limit $M \to \infty$

$$\begin{aligned}\lim_{M\to\infty}\tilde{g}^*_M(2,1;\tau) &= (-2\pi i)^3\sum_{m_1>m_2>0}\left(\sum_{n_1>0}n_1q^{m_1n_1}\right)\left(\frac{1}{2}+\sum_{n_2>0}q^{m_2n_2}\right)\\ &= (-2\pi i)^3\sum_{\substack{m_1>m_2>0\\n_1,n_2>0}}n_1q^{m_1n_1+m_2n_2} + \frac{(-2\pi i)^3}{2}\sum_{\substack{m_1>m_2>0\\n_1>0}}n_1q^{m_1n_1}\\ &= \hat{g}(2,1;\tau) + \frac{(-2\pi i)^3}{2}\sum_{\substack{m_1>0\\n_1>0}}(m_1-1)n_1q^{m_1n_1}\\ &= \hat{g}(2,1;\tau) + \frac{1}{2}\hat{g}\left(\begin{smallmatrix}2\\1\end{smallmatrix};\tau\right) + \pi i\hat{g}(2;\tau),\end{aligned} \tag{22}$$

Inserting this back into (21), we get

$$G^*(2,1;\tau;X,Y) = \hat{g}(2;\tau)X + \zeta(2,1) + \hat{g}(2,1;\tau) + \frac{1}{2}\hat{g}\left(\begin{smallmatrix}2\\1\end{smallmatrix};\tau\right) + \pi i\hat{g}(2;\tau).$$

Note the presence of a "double-indexed $\hat{g}$" and of the apparently lower-weight term $\pi i\hat{g}(2)$ (we can still think of this as having weight 3, if we think of πi as having weight 1). The presence of these term can be traced back to the fact that Lemma 3.10 fails when the index contains a 1, due to the presence of the constant term$-\pi i$ in the q-series of $\Psi^*(1;x)$.

(ii) Let us now calculate $G^*(1,2;\tau;X,Y)$. Note that $z_1z_2 = z_1 * z_2 - z_2z_1 - z_3$, so

$$\begin{aligned}\tilde{g}^*(1,2;\tau;Y) &= \tilde{g}^*(2;\tau;Y)\tilde{g}^*(1;\tau;Y) - \tilde{g}^*(2,1;\tau;Y) - \tilde{g}^*(3;\tau;Y) \\ &= \hat{g}(2;\tau)Y - \hat{g}(2,1;\tau) - \frac{1}{2}\hat{g}\left(\begin{smallmatrix}2\\1\end{smallmatrix};\tau\right) - \pi i\hat{g}(2;\tau) - \hat{g}(3;\tau)\end{aligned}$$

and

$$\begin{aligned}\zeta^*(1,2;X) &= \zeta(2)X - \zeta(2,1) - \zeta(3) \\ &= \zeta(2)X - 2\zeta(3),\end{aligned}$$

where we used $\zeta(2,1) = \zeta(3)$. We then get

$$\begin{aligned}G^*(1,2;\tau;X,Y) &= \tilde{g}^*(1,2;\tau;Y) + \tilde{g}^*(1;\tau;Y)\zeta^*(2;X) + \zeta^*(1,2;X) \\ &= \hat{g}(2;\tau)Y - \hat{g}(2,1;\tau) - \frac{1}{2}\hat{g}\left(\begin{smallmatrix}2\\1\end{smallmatrix};\tau\right) - \pi i\hat{g}(2;\tau) - \hat{g}(3;\tau) + Y\zeta(2) + \zeta(2)X - 2\zeta(3)\end{aligned}$$

(iii) Let us finally calculate $G^*(2,1,2;\tau;X,Y)$. First note that

$$\begin{aligned}\tilde{g}^*_M(2,1,2;\tau) = &\sum_{M>m_1>m_2>m_3>0} \Psi^*(2;m_1\tau)\Psi^*(1;m_2\tau)\Psi^*(2;m_3\tau) \\ &+ \sum_{M>m_1>m_2>0} \big(\Psi^*(2;m_1\tau)\Psi^*(1,2;m_2\tau) + \Psi^*(2,1;m_1\tau)\Psi^*(2;m_1\tau)\big) \\ &+ \sum_{M>m_1>0} \Psi^*(2,1,2;m_1\tau),\end{aligned}$$

where the three sums correspond to ways of splitting the word $z_2z_1z_2$ into one, two and three nonempty subwords respectively. Consulting [Bou, tables 1 and 7], we see that $\Psi^*(2,1;x)$, $\Psi^*(1,2;x)$ and $\Psi^*(2,1,2;x)$ are all identically zero, so only the first sum does not vanish. Similarly to (22), we then get

$$\begin{aligned}\lim_{M\to\infty} \tilde{g}^*_M(2,1,2;\tau) &= (-2\pi i)^5 \sum_{m_1>m_2>m_3>0} \left(\sum_{n_1>0} n_1 q^{m_1n_1}\right)\left(\sum_{n_2>0}\frac{1}{2} + \sum_{n_2>0} q^{m_2n_2}\right)\left(\sum_{n_3>0} n_3 q^{m_3n_3}\right) \\ &= (-2\pi i)^5 \sum_{\substack{m_1>m_2>m_3>0\\ n_1,n_2,n_3>0}} n_1n_3 q^{m_1n_1+m_2n_2+m_3n_3} + \frac{(-2\pi i)^5}{2} \sum_{\substack{m_1>m_2>m_3>0\\ n_1,n_3>0}} n_1n_3 q^{m_1n_1+m_3n_3} \\ &= \hat{g}(2,1,2;\tau) + \frac{(-2\pi i)^5}{2} \sum_{\substack{m_1>m_3>0\\ n_1,n_3>0}} (m_1-m_3-1)n_1n_3 q^{m_1n_1+m_3n_3} \\ &= \hat{g}(2,1,2;\tau) + \frac{1}{2}\hat{g}\left(\begin{smallmatrix}2,2\\1,0\end{smallmatrix};\tau\right) - \frac{1}{2}\hat{g}\left(\begin{smallmatrix}2,2\\0,1\end{smallmatrix};\tau\right) + \pi i\hat{g}(2,2).\end{aligned}$$

Thus,

$$\begin{aligned}G^*(2,1,2;\tau;X,Y) &= \tilde{g}^*(2,1,2;\tau;Y) + \tilde{g}^*(2,1;\tau;Y)\zeta^*(2,X) + \tilde{g}^*(2;\tau;Y)\zeta^*(1,2;X) + \zeta^*(2,1,2;X) \\ &= \hat{g}(2,1,2;\tau) + \frac{1}{2}\hat{g}\left(\begin{smallmatrix}2,2\\1,0\end{smallmatrix};\tau\right) - \frac{1}{2}\hat{g}\left(\begin{smallmatrix}2,2\\0,1\end{smallmatrix};\tau\right) + \pi i\hat{g}(2,2) \\ &\quad + (\hat{g}(2,1;\tau) + \frac{1}{2}\hat{g}\left(\begin{smallmatrix}2\\1\end{smallmatrix};\tau\right) + \pi i\hat{g}(2;\tau))\zeta(2) \\ &\quad + \hat{g}(2;\tau)(\zeta(2)X - 2\zeta(3)) \\ &\quad + \zeta(2,1,2),\end{aligned}$$

cf. [Bac1, Example 6.15(i)] upon making the"natural" choice $X = 0$, and replacing $\zeta(3)$ with $\zeta(2,1)$.

References

[Bac1] Henrik Bachmann. The algebra of bi-brackets and regularized multiple eisenstein series. *Journal of Number Theory*, 200:260–294, 2019.

[Bac2] Henrik Bachmann. Lectures on multiple zeta values and modular forms, nagoya university, 2020.

[Bac3] Henrik Bachmann. Multiple eisenstein series and q-analogues of multiple zeta values. *Periods in Quantum Field Theory and Arithmetic, Springer Proceedings in Mathematics & Statistics*, 2020.

[Bac4] Henrik Bachmann. Stuffle regularized multiple eisenstein series revisited. *RIMS Kôkyûroku No.2238*, pages 73–86, 2022.

[BGHZ] Jan Hendrik Bruinier, Gerard Geer, Günter Harder, and Don Zagier. *The 1-2-3 of Modular Forms*. Springer Berlin, Heidelberg, 2008.

[BK] Henrik Bachmann and Ulf Kühn. The algebra of generating functions for multiple divisor sums and applications to multiple zeta values. *The Ramanujan Journal*, 40(3):605–648, Aug 2016.

[Bou] Olivier Bouillot. The algebra of multitangent functions. *Journal of Algebra*, 410:148–238, 2014.

[GKZ] Herbert Gangl, Masanobu Kaneko, and Don Zagier. Double zeta values and modular forms. *Automorphic forms and zeta functions*, pages 71–106, 2006.

[HI] Michael E. Hoffman and Kentaro Ihara. Quasi-shuffle products revisited. *Journal of Algebra*, 481:293–326, 2017.

[Hof] Michael E. Hoffman. Quasi-shuffle products. *Journal of Algebraic Combinatorics*, 11:49–68, 1999.

[IKZ] Kentaro Ihara, Masanobu Kaneko, and Don Zagier. Derivation and double shuffle relations for multiple zeta values. *Compositio Mathematica*, 142(2):307–338, 2006.