

From polygons and symbols to polylogarithmic functions

Claude Duhr,^{a,b} Herbert Gangl^c and John R. Rhodes^c

^a*Institute for Particle Physics Phenomenology, University of Durham,
Durham, DH1 3LE, U.K.*

^b*Institut für theoretische Physik, ETH Zürich,
Wolfgang-Paulistr. 27, CH-8093, Switzerland*

^c*Department of Mathematical Sciences, University of Durham,
Durham, DH1 3LE, U.K.*

E-mail: duhrc@itp.phys.ethz.ch, herbert.gangl@durham.ac.uk,
j.r.rhodes@durham.ac.uk

ABSTRACT: We present a review of the symbol map, a mathematical tool introduced by Goncharov and used by him and collaborators in the context of $\mathcal{N} = 4$ SYM for simplifying expressions among multiple polylogarithms, and we recall its main properties. A recipe is given for how to obtain the symbol of a multiple polylogarithm in terms of the combinatorial properties of an associated rooted decorated polygon, and it is indicated how that recipe relates to a similar explicit formula for it previously given by Goncharov. We also outline a systematic approach to constructing a function corresponding to a given symbol, and illustrate it in the particular case of harmonic polylogarithms up to weight four. Furthermore, part of the ambiguity of this process is highlighted by exhibiting a family of non-trivial elements in the kernel of the symbol map for arbitrary weight.

KEYWORDS: Scattering Amplitudes, Supersymmetric gauge theory

ARXIV EPRINT: [1110.0458](https://arxiv.org/abs/1110.0458)

Contents

1	Introduction	1
2	Short review of multiple polylogarithms	3
3	Symbols and polygons	6
3.1	An example in a nutshell	6
3.2	Rules of symbol calculus	10
3.3	Relationship to the symbol of ref. [63].	14
4	A simple example	19
5	Integrating symbols: an algorithmic approach	22
5.1	Choosing the types of functions	23
5.2	Finding the arguments	24
5.3	Integrating the symbol (1)	27
5.4	A set of projectors	28
5.5	Integrating the symbol (2)	30
5.6	Elements in the kernel of the symbol map	32
6	Application: a spanning set for harmonic polylogarithms	34
6.1	Example	37
7	Conclusion	41
A	Review on shuffle algebras	41
B	Selected examples of symbols	43
B.1	The symbol of a generic multiple polylogarithm of weight one	45
B.2	The symbol of a generic multiple polylogarithm of weight two	45
B.3	The symbol of a generic multiple polylogarithm of weight three	46
B.4	The symbol of a generic multiple polylogarithm of weight four	46
C	Proof of Proposition 4	49
D	Some considerations on the implementation of the algorithm	57
E	Analytic continuation of the spanning set of functions for harmonic polylogarithms	58
E.1	Analytic representation inside the unit disc	58
E.2	Analytic representation outside the unit disc: inversion relations	59

F	Inversion formulas for the spanning set	60
F.1	Weight three	60
F.2	Weight four	62
G	Expression of HPL's in terms of the spanning set	67
G.1	Results for weight two	67
G.2	Results for weight three	67
G.3	Results for weight four	68

1 Introduction

Polylogarithms and their multivariable generalizations [1, 2] play an equally important role in modern mathematics and in physics. In mathematics they occur for instance in connection with algebraic K -theory and mixed Tate motives, e.g. [3–8], with Hilbert’s third problem (on scissors congruences), e.g. [9–11], as volume functions for hyperbolic spaces, e.g. [11–16, 18], and are also related to characteristic classes, e.g. [17], special values of L -functions in algebraic number theory, e.g. [5, 7, 18], algebraic cycles, e.g. [19–21] or, in the form of iterated integrals, in algebraic topology, e.g. [22–24]. In physics, the computation of higher order corrections to physical observables requires the analytical evaluation of Feynman integrals that can generally be expressed in terms of (special classes of) multiple polylogarithms, e.g. [25–59]. While in all of these applications it would be desirable to have a minimal spanning set — “basis functions” in physics parlance — for the polylogarithmic expressions involved in a given problem, it is well known that these latter functions satisfy various intricate functional equations among themselves, making the question of how to find a minimal spanning set very hard to answer in general. As a consequence, seemingly complicated results, say for a Feynman integral, may admit a much shorter analytic representation, the simplicity of the answer being hidden due to the existence of an abundance of functional equations among these functions. There is thus a strong interest for a better understanding of the functional equations among multiple polylogarithms, both from a formal mathematical standpoint and in view of practical applications in physics.

A way to approach functional equations among (multiple) polylogarithms is provided by the so-called *symbol map* [62, 63], a linear map that associates to each multiple polylogarithm of weight n an element in the n -fold tensor power of some vector space of one-forms. The virtue of the symbol map is that it captures to a good extent the main combinatorial and analytical properties of certain transcendental functions, and in particular it is expected that all functional equations among multiple polylogarithms are in the kernel of the symbol map. Loosely speaking, this means that a necessary condition for two expressions written in terms of multiple polylogarithms to be equal modulo functional equations is that they have the same symbol, a condition that is usually much easier to check than proving equality at the level of the functions. The inverse problem (sometimes called *integration of*

a symbol) of finding a function whose symbol matches a given tensor satisfying a certain integrability condition is much harder and we know of no general algorithm to construct such a function.

While special cases of the symbol map have been profitably used by mathematicians for over two decades (for example in connection with functional equations see e.g. refs. [5, 7, 60, 61]), it has only very recently been introduced into physics in the context of the $\mathcal{N} = 4$ Super Yang-Mills (SYM) theory in ref. [63], where it was applied to greatly simplify the analytic expression for the two-loop six-point remainder function obtained in ref. [64, 65]. In the wake of that work, the symbol map has seen various applications, mostly in the context of $\mathcal{N} = 4$ SYM. In particular, by now the symbols of all two-loop remainder functions are completely known [66], while at three loops the symbols of the remainder functions for the hexagon in general kinematics [67] and for the octagon in special kinematics [68] are known up to some free parameters that could not be fixed from general considerations. However, only in the latter octagon case an integrated form of the symbol is also known. Other approaches, aiming at the determination of the symbol of loop amplitudes by exploiting the operator product expansion in the collinear limit [69, 70] or the relationship between Feynman integrals and the volumes of polyhedra in non-euclidean spaces [71, 72], have also been considered. Furthermore, the symbol map was recently used to obtain compact analytic expressions for certain one-loop hexagon integrals in $D = 6$ dimensions [73–76]. More phenomenological applications, as e.g. in ref. [77], have also been considered.

The aim of this paper is twofold: While the symbol map has already been extensively used in the $\mathcal{N} = 4$ SYM community in physics, it seems still rather little known in other areas of physics in which the computation of Feynman integrals plays an important role. On the one hand, we therefore present a concise review on this topic, putting special emphasis on how to apply the symbol map to obtain simpler or shorter analytic results for functions arising from certain Feynman integrals. On the other hand, we believe that our work goes beyond the existing literature on the subject in various aspects. While so far the symbol of a transcendental function was defined recursively by considering iterated differentials, we introduce a simple diagrammatic rule that allows to directly read off the symbol from all possible “triangulations” of a certain decorated polygon associated to a multiple polylogarithm [21]. Furthermore, we also address the problem of how to integrate a symbol to a function by presenting an effective approach to construct a candidate spanning set of functions in terms of which the symbol might be integrated.

The structure of the paper is as follows: In section 2 we give a short review of multiple polylogarithms and of their properties. In section 3 we review the main properties of the symbol map and we show how to obtain the symbol of a multiple polylogarithm as the weighted sum of all possible maximal dissections of a certain decorated polygon associated to the polylogarithm. Moreover, we indicate how to relate it to a formula originally given by Goncharov in ref. [8]. In section 4 we give a short example of how to integrate a symbol following the approach introduced in refs. [63, 78], before generalizing this procedure to higher weights in section 5. In order to highlight remaining difficulties and ambiguities when trying to integrate to a function, we also give a family of non-trivial elements in the kernel of the symbol map. We illustrate these concepts in section 6 where we apply them

to derive a spanning set up to weight four for a special class of multiple polylogarithms, the so-called *harmonic* polylogarithms [25]. The five (rather long) appendices contain a summary of the mathematical notions used throughout the paper, as well as some technical details and proofs left out in the main text. We also include an appendix with a collection of symbols for multiple polylogarithms up to weight four.

Remark. The authors wish the reader to be aware that this paper contains the work of both physicists and mathematicians. As a consequence, it should be noted that the paper has been written to try to accommodate the language of both communities. We tried to find a compromise in the level of details given, and so while some arguments may go deeper than felt necessary by some, the text may be too sketchy at times for others.

2 Short review of multiple polylogarithms

Definition. Multiple polylogarithms can be defined recursively, for $n \geq 0$, via the iterated integral [1, 2]

$$G(a_1, \dots, a_n; x) = \int_0^x \frac{dt}{t - a_1} G(a_2, \dots, a_n; t), \quad (2.1)$$

with $G(x) = G(; x) = 1$, an exception being when $x = 0$ in which case we put $G(0) = 0$ (clearly any expression $\int_0^0 \dots$ should be zero), and with $a_i \in \mathbb{C}$ are chosen constants and x is a complex variable. In the following, we will also consider $G(a_1, \dots, a_n; x)$ to be functions of a_1, \dots, a_n . In the special case where all the a_i 's are zero, we define, using the obvious vector notation $\vec{a}_n = \underbrace{(a, \dots, a)}_n$, $a \in \mathbb{C}$,

$$G(\vec{0}_n; x) = \frac{1}{n!} \log^n x, \quad (2.2)$$

consistent with the case $n = 0$ above. Note that, while in the Mathematics literature these functions appear already in the early 20th century in the works of Poincaré and of Lappo-Danilevsky [79] as “hyperlogarithms”, as well as in the 1960’s in Chen’s work on iterated integrals (e.g., [22]),¹ in the physics literature these functions are often called *Goncharov polylogarithms*, due to the wealth of structure that the latter has established for them over the last 20 years. Throughout this paper, we follow the physics convention for the definition of the iterated integrals, which differs slightly from the mathematical one; e.g., in ref. [2], the function corresponding to $G(a_1, \dots, a_n; x)$ would be denoted $I(0; a_n, \dots, a_1; x)$, i.e., with the reverse order of the a_i but keeping the same variable x .

We will refer to the vector $\vec{a} = (a_1, \dots, a_n)$ as the *vector of singularities* attached to the multiple polylogarithm and the number of elements n , counted with multiplicities, in that vector is called the *weight* of the multiple polylogarithm.

Properties. We collect here a number of useful and well-known properties (cf. e.g. ref. [2, 8]). Iterated integrals form a *shuffle algebra* [81] (see appendix A for a short review of shuffle algebras), which allows one to express the product of two multiple polylogarithms of

¹In a sense, they already made an appearance in Kummer’s pioneering work [80] in 1840.

weight n_1 and n_2 as a linear combination with integer coefficients of multiple polylogarithms of weight $n_1 + n_2$, via

$$G(a_1, \dots, a_{n_1}; x) G(a_{n_1+1}, \dots, a_{n_1+n_2}; x) = \sum_{\sigma \in \Sigma(n_1, n_2)} G(a_{\sigma(1)}, \dots, a_{\sigma(n_1+n_2)}; x), \quad (2.3)$$

where $\Sigma(n_1, n_2)$ denotes the set of all shuffles of $n_1 + n_2$ elements, i.e., the subset of the symmetric group $S_{n_1+n_2}$ defined by (cf. ref. [22], eq. (1.5.6))

$$\Sigma(n_1, n_2) = \{ \sigma \in S_{n_1+n_2} \mid \sigma^{-1}(1) < \dots < \sigma^{-1}(n_1) \text{ and } \sigma^{-1}(n_1+1) < \dots < \sigma^{-1}(n_1+n_2) \}. \quad (2.4)$$

The algebraic properties of multiple polylogarithms imply that not all the $G(\vec{a}; x)$ for fixed x are independent, but that there are (polynomial) *relations* among them. In particular, we can reduce them, modulo products of lower weight functions, to functions whose rightmost index of all the vectors of singularities is non-zero (apart from objects of the form $G(\vec{0}_n; x)$), e.g.,

$$\begin{aligned} G(a, 0, 0; x) &= G(0, 0; x) G(a; x) - G(0, 0, a; x) - G(0, a, 0; x) \\ &= G(0, 0; x) G(a; x) - G(0, 0, a; x) - (G(0, a; x) G(0; x) - 2G(0, 0, a; x)) \quad (2.5) \\ &= G(0, 0; x) G(a; x) + G(0, 0, a; x) - G(0, a; x) G(0; x), \end{aligned}$$

where the middle summand is of the desired form (and the remaining summands are products of lower weight functions).

If the (rightmost) index a_n of \vec{a} is non-zero, then the function $G(\vec{a}; x)$ is *invariant under a rescaling* of all its arguments, i.e., for any $k \in \mathbb{C}^*$ we have

$$G(k \vec{a}; k x) = G(\vec{a}; x) \quad (a_n \neq 0). \quad (2.6)$$

Furthermore, multiple polylogarithms satisfy the *Hölder convolution* [82], i.e., whenever $a_1 \neq 1$ and $a_n \neq 0$, we have, $\forall p \in \mathbb{C}^*$,

$$G(a_1, \dots, a_n; 1) = \sum_{k=0}^n (-1)^k G\left(1 - a_k, \dots, 1 - a_1; 1 - \frac{1}{p}\right) G\left(a_{k+1}, \dots, a_n; \frac{1}{p}\right). \quad (2.7)$$

Below in section 5 we will be particularly interested in the limiting case $p \rightarrow \infty$ of this identity,

$$G(a_1, \dots, a_n; 1) = (-1)^n G(1 - a_n, \dots, 1 - a_1; 1). \quad (2.8)$$

Whenever they converge, multiple polylogarithms can equally well be represented [1] as *multiple nested sums* (e.g., for $|x_i| < 1$)

$$\text{Li}_{m_1, \dots, m_k}(x_1, \dots, x_k) = \sum_{0 < n_1 < n_2 < \dots < n_k} \frac{x_1^{n_1} x_2^{n_2} \dots x_k^{n_k}}{n_1^{m_1} n_2^{m_2} \dots n_k^{m_k}} = \sum_{n_k=1}^{\infty} \frac{x_k^{n_k}}{n_k^{m_k}} \sum_{n_{k-1}=1}^{n_k-1} \dots \sum_{n_1=1}^{n_2-1} \frac{x_1^{n_1}}{n_1^{m_1}}. \quad (2.9)$$

Note that we are using Goncharov's original summation convention [1]; other authors define $\text{Li}_{m_1, \dots, m_k}(x_1, \dots, x_k)$ using the reverse summation convention instead, i.e. $n_1 > \dots > n_k$.

The G and Li functions define in fact the same class of functions and are related by

$$\text{Li}_{m_1, \dots, m_k}(x_1, \dots, x_k) = (-1)^k G_{m_k, \dots, m_1} \left(\frac{1}{x_k}, \dots, \frac{1}{x_1 \dots x_k} \right), \quad (2.10)$$

(note the reverse order of the indices for G) where we used the notation

$$G_{m_1, \dots, m_k}(t_1, \dots, t_k) = G(\underbrace{0, \dots, 0}_{m_1-1}, t_1, \dots, \underbrace{0, \dots, 0}_{m_k-1}, t_k; 1). \quad (2.11)$$

It is possible to find *closed expressions* for (very few) special classes of multiple polylogarithms, for arbitrary weight, in terms of classical polylogarithm functions, e.g., for $a \neq 0$ we have

$$\begin{aligned} G(\vec{0}_n; x) &= \frac{1}{n!} \log^n x, & G(\vec{a}_n; x) &= \frac{1}{n!} \log^n \left(1 - \frac{x}{a} \right), \\ G(\vec{0}_{n-1}, a; x) &= -\text{Li}_n \left(\frac{x}{a} \right), & G(\vec{0}_n, \vec{a}_p; x) &= (-1)^p S_{n,p} \left(\frac{x}{a} \right), \end{aligned} \quad (2.12)$$

where $S_{n,p}$ denotes the Nielsen polylogarithm [83]. Moreover, up to weight three, multiple polylogarithms are well-known to be expressible in terms of ordinary logarithms, dilogarithms and trilogarithms (cf. ref. [84], section 8.4.3, implicitly, as well as refs. [11, 15]). In particular, if a and b are non-zero and different, we find

$$G(a, b; x) = \text{Li}_2 \left(\frac{b-x}{b-a} \right) - \text{Li}_2 \left(\frac{b}{b-a} \right) + \log \left(1 - \frac{x}{b} \right) \log \left(\frac{x-a}{b-a} \right). \quad (2.13)$$

Aim. The aim of this paper is to present an algorithmic approach how to deal with — in fact rather to circumvent — the complicated functional equations that relate multiple polylogarithms, and how to find, given a choice of certain singularities a_i , a (possibly minimal) spanning set for functions in which to express multiple polylogarithms with singularities only in these a_i , provided such a set exists. The approach we present is rather generic and can be applied to any expression involving multiple polylogarithms. This is made possible by using results closely related to work of Goncharov, Spradlin, Vergu and Volovich [63], which in turn was inspired by the theory of (mixed Tate) motives,² and in particular by using a certain tensor calculus associated to iterated integrals, which is called “symbol calculus” in the following (the name “symbol” originating from [63] and from ref. [62]), and which we will review in the next section.

An important remark is that the construction of a symbol seems to be a rather special case of a very general construction by Chen [22], where it appears as the image of an iterated integral as a 0-cocycle in the so-called “bar construction” attached to, say, X equal to the projective line minus a number of points (more generally, the construction has been investigated for a hyperplane configuration [92], section 3), and it lands in the n -fold

²Let us point out that this is far from being the first exhibit of a direct connection between mixed Tate motives and mathematical physics, as such a relationship has been explored, e.g., by Kreimer in work with Bloch and Esnault [85, 86], such a connection was clearly apparent from letter correspondence between Broadhurst and Deligne [87] resulting e.g. in ref. [88], work of Belkale-Brosnan [89] or more recently by Brown [90] and others. One should also mention work of Connes and Marcolli [91] in this direction.

tensor product of the vector space of 1-forms on the underlying space X . Moreover, Chen characterised the image as the formal words in these 1-forms satisfying a natural integrability condition. A polygon attached to an iterated integral enjoys the useful property that it gives a very concise way of explicitly producing integrable words of that kind. We also note that Goncharov has generalized the construction of a symbol to a considerably larger setting, for any so-called framed mixed Hodge-Tate structures (ref. [62]).

As an application, we restrict ourselves in section 6 to a specific subclass of multiple polylogarithms that are of particular importance in applications in high-energy physics. These so-called *harmonic* polylogarithms (HPL's) $H(\vec{a}; x)$ were first singled out and thoroughly studied in ref. [25]. HPL's correspond to a special case of the iterated integral defined in eq. (2.1) where $a_i \in \{-1, 0, 1\}$. More precisely, they are defined via

$$H(\vec{a}; x) = (-1)^k G(\vec{a}; x), \quad a_i \in \{-1, 0, 1\}, \quad (2.14)$$

where k is the number of elements in \vec{a} equal to $(+1)$ (this slightly puzzling sign arises from earlier ad-hoc conventions when authors used the 1-form $\frac{dt}{1-t}$ rather than $\frac{dt}{t-1}$). Many one-loop and two-loop Feynman integrals can be expressed in terms of HPL's up to weight four and generalizations thereof [26, 27]. As harmonic polylogarithms are just a special case of the multiple polylogarithms introduced at the beginning of this section, all HPL's through weight three can be expressed through classical polylogarithms. By contrast, similar to the general case of multiple polylogarithms, it is expected that HPL's of weight ≥ 4 are no longer expressible in terms of classical ones alone. In section 6 we illustrate our technique by constructing a spanning set of harmonic polylogarithms in weight 4 for which we indeed needed three non-classical functions.

3 Symbols and polygons

The differential structure of multiple polylogarithms can be captured very well combinatorially using a certain kind of decorated polygons with some additional structure, as developed in ref. [21], where they were called *R-deco polygons*. We note that there are related notions that had occurred previously in Goncharov's work, e.g. in refs. [8, 93]. There is an algebraic object attached to such a polygon, and hence to the corresponding multiple polylogarithm. This object, which has been dubbed a *symbol* in ref. [63], is an element in a tensor power of a certain vector space and contains a lot of information about the original function.

3.1 An example in a nutshell

In this subsection we give a quick idea of how, following ref. [21], one can associate to a multiple polylogarithm — or rather to an associated rooted decorated polygon — its symbol (we show in section 3.2 that this definition is equivalent to the definition given in ref. [63]). In the following subsection we then give a more detailed account of the construction.

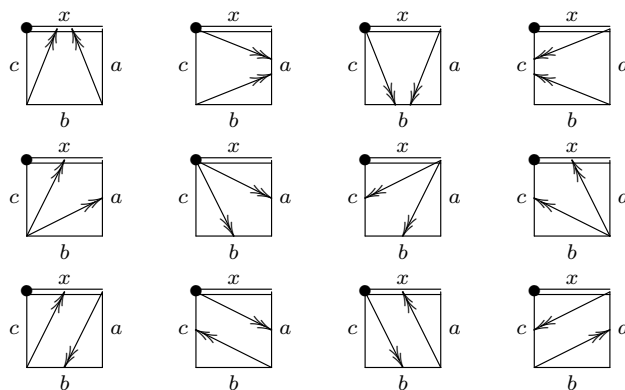
A multiple polylogarithm of weight n gives rise to a certain $(n + 1)$ -gon. As a “fore-shadowing” example, we first give the 4-gon $P = P(c, b, a, x)$ attached to some weight 3 multiple polylogarithm $G(a, b, c; x) = -Li_{1,1,1}(b/c, a/b, x/a) = I(0; c, b, a; x)$:

$$P(c, b, a, x) = \begin{array}{c} \bullet \\ \overline{x} \\ \square \\ \begin{array}{cc} c & a \\ & b \end{array} \end{array} \longleftrightarrow G(a, b, c; x)$$

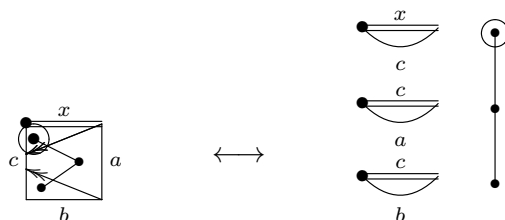
which comes equipped with distinct decorations (in this order) c, b, a and x , the latter decoration x being for the distinguished *root side* (drawn by a double line in the picture), and also carries information on the orientation of the polygon in the form of a fat vertex (which should be thought of as the “first” vertex, while the root side — adjacent to the first vertex — is the “last” side).

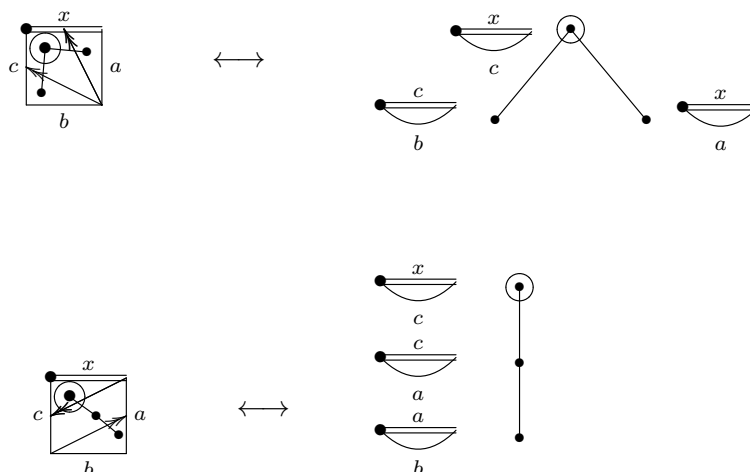
In a *first step*, one lists all possible ways to draw the maximal number of non-intersecting arrows (an arrow is a directed line from a vertex of P to a non-adjacent side), which for an $(n + 1)$ -gon amounts to $n - 1$ such arrows, and one formally adds the resulting objects (the framing polygon is identical, while each copy is equipped with a different maximal set of arrows).

In our example $n = 3$, such a maximal set contains $n - 1 = 2$ arrows, and there are precisely 12 different such sets, given by



In a *second step*, to each such maximal set A of arrows in P , we associate a rooted tree (as the tree dual to the polygon dissection defined by the arrows) whose decorations are (decorated and rooted) 2-gons. As an example, to the 4-gons in the last column above we attach





Any linear order on the vertices of such a rooted tree which is *compatible* in the sense discussed below in section 3.2 with the partial order on it (only the middle tree above is non-linear hence allows more than one such linear order) now gives a term in the *symbol* $\mathcal{S}(P)$ attached to P . In practice, this means that every branching in a tree contributes to the symbol by the shuffle of (the vertices that appear on each of) its branches (see below for a more detailed description).

Third step: each of the 2-gons B in one of the linear orders on the vertices now is mapped via a suitable map μ to a rational function in the original decorations of the polygons (in the example a natural target space would be the (multiplicative group of the) function field $\mathbb{Q}(a, b, c, x)$ of rational functions in the variables a, b, c, x). More precisely, if $B = \begin{matrix} y \\ \bullet \text{---} \text{---} \text{---} \bullet \\ x \end{matrix}$ for $x \neq y, y \neq 0$, where y denotes the root decoration, then we map B to $\mu(B) = 1 - y/x$, provided $x \neq 0$, and to $\mu(B) = y$ otherwise. The case of a 2-gon with identical decorations $y = x$ needs special attention. In this situation we naively obtain $\mu(B) = 0$, which morally corresponds to the divergent integral $G(x; y) = G(x; x)$. This problem can be circumvented by introducing a small ‘regulator’³ ε such that x and y do no longer coincide, and we drop in the symbol all terms of the form $\dots \otimes \varepsilon \otimes \dots$ before taking the limit $\varepsilon \rightarrow 0$. It should however be noted that the limit depends on the regularization, e.g. if we choose $y = (1 - \varepsilon)x$, we obtain $\mu(B) = 1 - (1 - \varepsilon) = \varepsilon$ and

$$\dots \otimes \mu(B) \otimes \dots = \dots \otimes \varepsilon \otimes \dots \rightarrow 0, \tag{3.1}$$

while the choice $y = x - \varepsilon$ leads to $\mu(B) = 1 - (x - \varepsilon)/x = \varepsilon/x$

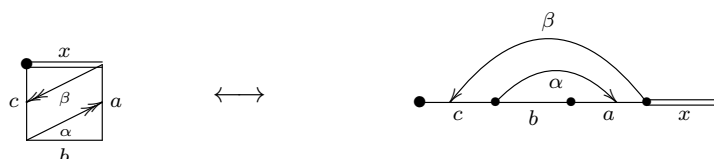
$$\dots \otimes \mu(B) \otimes \dots = \dots \otimes (\varepsilon/x) \otimes \dots \rightarrow -\dots \otimes x \otimes \dots \tag{3.2}$$

In practice we observe however that, at least for the two different regularizations presented above, the final result for the symbol of a finite quantity is independent of the way we

³For a detailed discussion on how to regularize multiple polylogarithms we refer to ref. [2].

regularized the 2-gons with identical decorations. For this reason we will from now on always define $\mu(B) = 1$, which corresponds to regularization prescription of eq. (3.1).

Last step: fixing the signs. We need to invoke a sign for the individual elementary tensors, and this sign is determined by using the number of *backward arrows* in a dissection. In order to see this quickly, it is convenient to “break up” the polygon at its *first vertex* (in the pictures it is typically indicated by a bullet). Then we “roll out” the sequence of sides and arrange it as a line from left to right, starting with the first vertex and ending with the root side; dissecting arrows inside the polygon will be stretched out (in a way that they still do not intersect). We give it for the third example above:



Now a *backward arrow* is one which, in the rolled-out version of the polygon, has its end point to the left of its starting point (i.e. points from right to left, like β above), while a *forward arrow* has it to its right (i.e. points from left to right, like α above).

Here is a more formal definition: there is a natural linear order on the sides e_1, \dots, e_n of an n -gon as above, starting with the non-root side e_1 incident with the first vertex and ending with the root side e_n (in the example above it is the linear order given by the sides e_1, \dots, e_4 decorated by c, b, a and x , and the vertices $v_1 = e_4 \cap e_1$ (the first vertex), $v_2 = e_1 \cap e_2$, $v_3 = e_2 \cap e_3$, $v_4 = e_3 \cap e_4$). This induces a linear order on the vertices v_j which arise as the intersection of e_j and e_{j-1} (indices taken modulo n), where the first vertex is the smallest element in that order. Then a (non-trivial) arrow can be encoded by a pair (v_j, e_k) with $k \notin \{j-1, j\}$, and it is *backward* if and only if $k < j-1$. With these notions, the *sign* attached to a polygon P with a maximal arrow set A is given by $(-1)^{\#\{\text{backward arrows of } A\}}$. In the three four-gons discussed above in more detail we get two backward arrows for the first maximal dissection of the square and one backward arrow for the remaining two dissections.

Putting all of the above ingredients together and writing τ_A for the tree dual to the maximal set of arrows A , and (τ_A, \prec) for its partial order, the final formula for the *symbol* $\mathcal{S}(P)$ of an n -gon P is

$$\mathcal{S}(P) = \sum_{\substack{\text{max sets } A \\ \text{of arrows in } P}} (-1)^{\#\{\text{backward arrows of } A\}} \sum_{\substack{\text{linear orders } \lambda \\ \text{compatible with } (\tau_A, \prec)}} \mu\left(\overset{a_1}{\curvearrowright}_{b_1}\right) \otimes \dots \otimes \mu\left(\overset{a_{n-1}}{\curvearrowright}_{b_{n-1}}\right). \tag{3.3}$$

As an example, the first and third of the three maximal sets of arrows above give

$$+\mu\left(\overset{x}{\curvearrowright}_c\right) \otimes \mu\left(\overset{c}{\curvearrowright}_a\right) \otimes \mu\left(\overset{c}{\curvearrowright}_b\right) = \left(1 - \frac{x}{c}\right) \otimes \left(1 - \frac{c}{a}\right) \otimes \left(1 - \frac{c}{b}\right)$$

and

$$-\mu\left(\begin{array}{c} \bullet \\ \xrightarrow{x} \\ \xleftarrow{c} \\ \bullet \end{array}\right) \otimes \mu\left(\begin{array}{c} \bullet \\ \xrightarrow{c} \\ \xleftarrow{a} \\ \bullet \end{array}\right) \otimes \mu\left(\begin{array}{c} \bullet \\ \xrightarrow{a} \\ \xleftarrow{b} \\ \bullet \end{array}\right) = -\left(1 - \frac{x}{c}\right) \otimes \left(1 - \frac{c}{a}\right) \otimes \left(1 - \frac{a}{b}\right),$$

respectively, while the middle term (corresponding to a *non-linear* dual tree, i.e. a dual tree with branchings) contributes via the shuffle product of the two branches

$$\begin{aligned} & -\mu\left(\begin{array}{c} \bullet \\ \xrightarrow{x} \\ \xleftarrow{c} \\ \bullet \end{array}\right) \otimes \left(\mu\left(\begin{array}{c} \bullet \\ \xrightarrow{c} \\ \xleftarrow{b} \\ \bullet \end{array}\right) \text{III} \mu\left(\begin{array}{c} \bullet \\ \xrightarrow{x} \\ \xleftarrow{a} \\ \bullet \end{array}\right)\right) \\ &= -\mu\left(\begin{array}{c} \bullet \\ \xrightarrow{x} \\ \xleftarrow{c} \\ \bullet \end{array}\right) \otimes \mu\left(\begin{array}{c} \bullet \\ \xrightarrow{c} \\ \xleftarrow{b} \\ \bullet \end{array}\right) \otimes \mu\left(\begin{array}{c} \bullet \\ \xrightarrow{x} \\ \xleftarrow{a} \\ \bullet \end{array}\right) - \mu\left(\begin{array}{c} \bullet \\ \xrightarrow{x} \\ \xleftarrow{c} \\ \bullet \end{array}\right) \otimes \mu\left(\begin{array}{c} \bullet \\ \xrightarrow{x} \\ \xleftarrow{a} \\ \bullet \end{array}\right) \otimes \mu\left(\begin{array}{c} \bullet \\ \xrightarrow{c} \\ \xleftarrow{b} \\ \bullet \end{array}\right) \\ &= -\left(1 - \frac{x}{c}\right) \otimes \left(1 - \frac{c}{b}\right) \otimes \left(1 - \frac{x}{a}\right) - \left(1 - \frac{x}{c}\right) \otimes \left(1 - \frac{x}{a}\right) \otimes \left(1 - \frac{c}{b}\right), \end{aligned}$$

where we introduced the symbol for the shuffle product

$$a \text{III} b = a \otimes b + b \otimes a. \tag{3.4}$$

Motivation and justification of this assignment has been given to an extent in ref. [21], where it forms part of an expression arising from the well-known bar construction in algebraic topology applied to a differential graded algebra on the polygons above (which in turn is motivated by certain algebraic cycles originally studied by Bloch [94] and Bloch-Kriz [19]).

The first (and rather similar) description of symbols, in terms of trivalent trees, has been given, under the name \otimes^m -invariant, by Goncharov (ref. [8], section 4.4). In section 3.3 below we discuss the relationship between the two descriptions.

To summarize: an important part of the differential structure of a weight n multiple polylogarithm is captured by a certain decorated $(n + 1)$ -gon. More precisely, if the arguments of the multiple polylogarithm are expressed in terms of variables/constants x_1, \dots, x_m for some m , the polygon is an $(n + 1)$ -gon with decoration by simple expressions in x_1, \dots, x_m ; now to this (rooted decorated oriented) polygon there is attached in a natural way an expression (its “symbol”) in $V^{\otimes n}$ where V is a finite rank submodule (it might be convenient for the reader to think of V as a finite dimensional vector space) of the space $\mathbb{Q}(x_1, \dots, x_m)^\times$ (of infinite rank), i.e., the invertible (= non-zero) rational functions in the variables x_1, \dots, x_m .

3.2 Rules of symbol calculus

Roughly, a symbol is a formal sum of elementary n -fold tensors $a_1 \otimes \dots \otimes a_n$, and one works in each tensor factor as with (a refined form of) $d \log$ terms. In other words, each factor a_i in a tensor product is tacitly understood as

$$d \log a_i \equiv \frac{da_i}{a_i}. \tag{3.5}$$

Furthermore, we use shuffle products and the following *rules* (essentially boiling down to multilinearity, but in an unusual form, as we pass from multiplicative to additive notation):

- **Distributivity.**

$$C \otimes (a \cdot b) \otimes D = C \otimes a \otimes D + C \otimes b \otimes D \tag{3.6}$$

and consequently

$$C \otimes a^n \otimes D = n(C \otimes a \otimes D), \quad n \in \mathbb{Z}, \tag{3.7}$$

where C and D denote fixed elementary tensors. Note that n here is a coefficient rather than part of the first tensor factor; in particular, putting $n = 0$ we see that $C \otimes 1 \otimes D = 0$.

- **Neglecting torsion.**

We will “work up to torsion”, which means that we will put

$$C \otimes \rho_n \otimes D = 0, \quad n \in \mathbb{Z}, \tag{3.8}$$

for ρ_n an n -th root of unity (by eq. (3.7), we have $n(C \otimes \rho_n \otimes D) = 0$, in any case).

- **Shuffle product.**

An important property of the symbol is that it preserves products: more precisely, it maps the product of two multiple polylogarithms to the *shuffle product* of their respective symbols, i.e.

$$\mathcal{S}(G(a_1, \dots, a_r; x)G(b_1, \dots, b_s; y)) = \mathcal{S}(G(a_1, \dots, a_r; x)) \text{III} \mathcal{S}(G(b_1, \dots, b_s; y)) \tag{3.9}$$

where **III** is the symbol used for the shuffle product of two tensors, defined on elementary tensors by

$$(a_1 \otimes \dots \otimes a_{n_1}) \text{III} (a_{n_1+1} \otimes \dots \otimes a_{n_1+n_2}) = \sum_{\sigma \in \Sigma(n_1, n_2)} a_{\sigma^{-1}(1)} \otimes \dots \otimes a_{\sigma^{-1}(n_1+n_2)}, \tag{3.10}$$

where $\Sigma(n_1, n_2)$ was defined in eq. (2.4). For more details on shuffle algebras, we refer to appendix A. We note that, on the left hand side of eq. (3.9), the shuffle permutations are applied to the arguments of the two functions (cf. e.g. eq. (2.3)), while on right hand side one shuffles the tensor factors instead, in a completely analogous fashion.

Note that eq. (3.9) is a rather non-trivial fact, as one can already see in the first non-obvious case:

$$\begin{aligned} \mathcal{S}(G(a; x)G(b; x)) &= \mathcal{S}(G(a, b; x) + G(b, a; x)) = \\ &= \left(\left(1 - \frac{x}{a}\right) \otimes \left(1 - \frac{x}{b}\right) - \left(1 - \frac{x}{a}\right) \otimes \left(1 - \frac{a}{b}\right) + \left(1 - \frac{x}{b}\right) \otimes \left(1 - \frac{b}{a}\right) \right) \\ &\quad + \left(\left(1 - \frac{x}{b}\right) \otimes \left(1 - \frac{x}{a}\right) - \left(1 - \frac{x}{b}\right) \otimes \left(1 - \frac{b}{a}\right) + \left(1 - \frac{x}{a}\right) \otimes \left(1 - \frac{a}{b}\right) \right), \end{aligned}$$

which agrees with $\mathcal{S}(G(a; x)) \text{III} \mathcal{S}(G(b; x)) = \left(1 - \frac{x}{a}\right) \otimes \left(1 - \frac{x}{b}\right) + \left(1 - \frac{x}{b}\right) \otimes \left(1 - \frac{x}{a}\right)$, due to cancellations of terms.

We will encounter expressions which involve both tensor and shuffle products — in order to avoid writing many parentheses, our convention is that a shuffle takes precedence over a tensor, i.e.

$$a \otimes b \text{ III } c \equiv a \otimes (b \text{ III } c). \tag{3.11}$$

Furthermore, we abbreviate elementary tensors with the same factors as follows:

$$a^{\otimes n} = \underbrace{a \otimes \cdots \otimes a}_{n \text{ times}}. \tag{3.12}$$

- **Refined $d \log$ terms.**

We emphasise here already, though, that we will not treat $d \log c$ for a rational constant c as zero (as opposed to ref. [63], but in line with the non-reduced symbol of ref. [62]) since we would lose a lot of important information this way. Instead we extend the above calculus to rational numbers in complete analogy with the above; so we have, e.g.,

$$C \otimes 2^m \cdot 3^n \cdot x^{-5} \otimes D = m(C \otimes 2 \otimes D) + n(C \otimes 3 \otimes D) - 5(C \otimes x \otimes D). \tag{3.13}$$

- **Root decoration 0 annihilates:**

since $G(\dots; 0) = 0$, we also need to put $\mathcal{S}(G(\dots; 0)) = 0$, and this indicates that we can (and will) ignore polygons whose root side is decorated by 0.

Linear orders of a tree. For a rooted tree T , which we view without a fixed embedding into the plane, hence e.g. we consider as equal the two trees



there is a natural partial order \prec on its vertices v_j ($j \in J$), given as follows: the root vertex $v_0 \prec v_j$ for any $j \in J$, and $v_j \prec v_k$ for $v_j \neq v_0$ if and only if there is a direct path from root to a leaf passing first through v_j and then through v_k .

A *linear order on the vertices* of T which is *compatible with the order* \prec is a sequence $(v_0, v_{j_1}, \dots, v_{j_r})$ of all the vertices v_j ($j \in J$) such that $v_{j_i} \prec v_{j_k}$ implies $j_i \leq j_k$. (This means that if two vertices are in a relation with respect to the partial order, then they should be related in any compatible linear order in the same way, while if they are not related in the partial order, there is no condition for how they should be related in that linear order.) In the example, there are precisely two linear orders which are compatible with the partial order, as the root vertex always comes first: (v_0, v_1, v_2) and (v_0, v_2, v_1) .

Definition of a symbol. Now we are ready to give a complete definition of the symbol attached to a (rooted decorated oriented) $(n + 1)$ -gon P with decoration sequence (t_1, \dots, t_n, x) , the last member x being the root decoration, and then extend it by linearity and shuffle product to any sum of (products of) polygons, hence also for multiple polylogarithms:

$$\mathcal{S}(P) = \sum_{\substack{\text{max sets } A \\ \text{of arrows in } P}} (-1)^{\#\{\text{backward arrows of } A\}} \sum_{\substack{\text{linear orders } \lambda \\ \text{compatible with } (\tau_A, \prec)} } \mu(\pi_1^{A,\lambda}) \otimes \cdots \otimes \mu(\pi_n^{A,\lambda}), \quad (3.14)$$

where the 2-gons $\pi_\nu^{A,\lambda}$ are determined by the maximal dissection A together with the linear order λ which is compatible with the partial order \prec on τ_A , the dual tree of the dissection A , in the manner given above in the second step of section 3.1 (i.e. for each 2-gon arising from the dissection of A there is a vertex of τ_A decorated by that 2-gon, and for any two 2-gons that are adjacent there is an edge in τ_A connecting the corresponding vertices).

Integrability condition. A very useful property of the rooted decorated polygons, found by the second author in collaboration with F. Brown and A. Levin, is that each polygon (or rather its symbol) satisfies a certain integrability condition. Indeed, an arbitrary sum of elementary tensors does not necessarily lie in the image of the symbol map. Instead, it was pointed out in ref. [92], making explicit in a special case the very general approach of Chen [22], that a necessary and sufficient condition for a symbol

$$S = \sum_{I=(i_1, \dots, i_m)} c_I \omega_{i_1} \otimes \cdots \otimes \omega_{i_m} \quad (c_I \in \mathbb{Q}), \quad (3.15)$$

to be integrable to a function is that

$$\sum_{I=(i_1, \dots, i_m)} c_I \omega_{i_1} \otimes \cdots \otimes (\omega_{i_j} \wedge \omega_{i_{j+1}}) \otimes \cdots \otimes \omega_{i_m} = 0 \quad \text{for all } 1 \leq j \leq m, \quad (3.16)$$

where $\omega_{i_j} \wedge \omega_{i_{j+1}}$ denotes the usual exterior product of two differential forms. We rewrite this for our purposes as

$$\sum_{I=(i_1, \dots, i_m)} c_I [d \log \omega_{i_j} \wedge d \log \omega_{i_{j+1}}] \omega_{i_1} \otimes \cdots \otimes \widehat{\omega}_{i_j} \otimes \widehat{\omega}_{i_{j+1}} \otimes \cdots \otimes \omega_{i_m} = 0, \quad (3.17)$$

where the hats indicate that we omit the corresponding factors in the tensor product. As an example, we indicate the statement for $G(a, b; x)$, whose symbol

$$\mathcal{S}(G(a, b; x)) = \left(1 - \frac{x}{b}\right) \otimes \left(1 - \frac{x}{a}\right) - \left(1 - \frac{x}{b}\right) \otimes \left(1 - \frac{b}{a}\right) + \left(1 - \frac{x}{a}\right) \otimes \left(1 - \frac{a}{b}\right) \quad (3.18)$$

satisfies

$$d \log \left(1 - \frac{x}{b}\right) \wedge d \log \left(1 - \frac{x}{a}\right) - d \log \left(1 - \frac{x}{b}\right) \wedge d \log \left(1 - \frac{b}{a}\right) + d \log \left(1 - \frac{x}{a}\right) \wedge d \log \left(1 - \frac{a}{b}\right) = 0, \quad (3.19)$$

where we recall that $d \log x = dx/x$. Indeed, writing

$$d \log \left(1 - \frac{\alpha}{\beta}\right) = \frac{d\alpha - \alpha d\beta}{\beta - \alpha} \Big|_{y=0}^\alpha \quad (3.20)$$

the left-hand side of eq. (3.19) becomes

$$\frac{dy - db}{y - b} \Big|_{y=0}^x \wedge \frac{dy - da}{y - a} \Big|_{y=0}^x - \frac{dy - db}{y - b} \Big|_{y=0}^x \wedge \frac{dy - da}{y - a} \Big|_{y=0}^b + \frac{dy - da}{y - a} \Big|_{y=0}^x \wedge \frac{dy - db}{y - b} \Big|_{y=0}^a \quad (3.21)$$

and we find, e.g., that the coefficient of $dx \wedge da$ is given by

$$\frac{1}{x - b} \cdot \frac{-1}{x - a} - \frac{1}{x - b} \cdot \frac{-1}{b - a} + \frac{1}{x - a} \cdot \frac{1}{a - b} = 0. \quad (3.22)$$

The coefficients of $dx \wedge db$ and $db \wedge da$ vanish in a similar way.

Note that the integrability condition also plays an important role in ref. [62].

3.3 Relationship to the symbol of ref. [63].

In ref. [63], the Goncharov, Spradlin, Vergu and Volovich use the differential equation for multiple polylogarithms recursively to arrive at the definition of a symbol. More precisely, if $F : \mathbb{C}^n \rightarrow \mathbb{C}$ denotes a complex valued function depending on n complex variables x_k , $1 \leq k \leq n$, the authors of ref. [63] define the “symbol of the transcendental function” F in the following recursive way: if the total differential of F can be expressed in the form

$$dF = \sum_i F_i d \log R_i, \quad (3.23)$$

where F_i and R_i are functions of the variables x_k , and R_i are moreover rational functions, then the symbol of F is defined recursively via

$$\mathcal{S}(F) = \sum_i \mathcal{S}(F_i) \otimes R_i. \quad (3.24)$$

In the case where F is a multiple polylogarithm, we can write down the differential of F in an explicit form. For example, in the special case where all the arguments of the multiple polylogarithm are generic (i.e., they are mutually different and do not take particular values), we obtain [2]

$$dG(a_{n-1}, \dots, a_1; a_n) = \sum_{i=1}^{n-1} G(a_{n-1}, \dots, \hat{a}_i, \dots, a_1; a_n) d \log \left(\frac{a_i - a_{i+1}}{a_i - a_{i-1}} \right). \quad (3.25)$$

The symbol of $G(a_1, \dots, a_{n-1}; a_n)$ is then defined in the form

$$\mathcal{S}(G(a_{n-1}, \dots, a_1; a_n)) = \sum_{i=1}^{n-1} \mathcal{S}(G(a_{n-1}, \dots, \hat{a}_i, \dots, a_1; a_n)) \otimes \left(\frac{a_i - a_{i+1}}{a_i - a_{i-1}} \right). \quad (3.26)$$

The symbol we just obtained looks seemingly different from the definition we gave in eq. (3.14), which consists in summing over all possible maximal sets of arrows of the polygon $P(a_1, \dots, a_{n-1}, a_n)$ associated to $G(a_{n-1}, \dots, a_1; a_n)$. In the following we show that the two definitions are equivalent up to a rearrangement of the terms in the sum, and hence give rise to the same tensor.

Let us consider the n -gon $P = P(a_1, \dots, a_n)$ (i.e. with sides decorated by a_i , the last one a_n decorating the root side). We will show that the symbol of P satisfies the recursion (3.26). For simplicity, we will concentrate here on the case of generic decorations. Let Λ_P be the set of all linear orders on the dual tree attached to any of the maximal sets of arrows of P . Then there is an obvious bijection between the terms in the double sum in eq. (3.14) and the elements in Λ_P . We can partition Λ_P by collecting all those linear orders into a subset which share the same last 2-gon that decorates the last vertex of this linear order. This partitions Λ_P into a priori $2n$ subsets, as those last vertices correspond precisely to the 2-gons that we can cut off from P .

Note that cutting off the last 2-gon in a linear order on a maximal dissection corresponds to contracting the associated edge in the dual tree. Note also that, clearly, the last vertex must be a *leaf* of the (rooted) dual tree, and hence each last 2-gon necessarily cuts off two successive sides of P .

Remark. For each n -gon P with $n > 2$, there are three ways to cut off a 2-gon, resulting in two different ones. None of these can become the last one in any linear order in Λ_P .

More explicitly, these are the two 2-gons $\bullet \begin{array}{c} \xrightarrow{a_n} \\ \text{---} \\ \xleftarrow{a_{n-1}} \end{array} \bullet$ and $\bullet \begin{array}{c} \xrightarrow{a_n} \\ \text{---} \\ \xrightarrow{a_1} \end{array} \bullet$. The former can arise

only by cutting off the root side, while the latter can arise both by cutting off the root side and by cutting off the first side and the first vertex. As a consequence, Λ_P partitions into only $2n - 3$ non-empty subsets (of same cardinality) of the above type.

In view of the above, it is clear that any such subset indexes exactly the linear orders $\Lambda_{\tilde{P}}$ on the (dual trees of the maximal dissections of the) subpolygon \tilde{P} of P which is obtained from P by cutting off a fixed 2-gon, followed by contracting the dissecting arrow to a point.

There are typically two ways of cutting off a 2-gon in which such a subpolygon \tilde{P} can occur: for $i = 2, \dots, n - 1$, cutting off the 2-gon $\bullet \begin{array}{c} \xrightarrow{a_{i\pm 1}} \\ \text{---} \\ \xrightarrow{a_i} \end{array} \bullet$ leaves a complementary

subpolygon \tilde{P}_i^\pm which is identified with \tilde{P}_i upon contraction of the dissecting arrow. Note that \tilde{P}_i^+ and \tilde{P}_i^- will give terms of opposite parity, as precisely one of them corresponds to a forward arrow. The only exception arises from cutting off $\bullet \begin{array}{c} \xrightarrow{a_2} \\ \text{---} \\ \xrightarrow{a_1} \end{array} \bullet$, which corresponds

to a forward arrow, for which only one complementary subpolygon \tilde{P}_1^+ can occur.

In summary, we get:


Claim. *There is a bijection of sets*

$$\Lambda_P \xleftrightarrow{1:1} \bigcup_{i=1}^{n-1} \left(\Lambda_{\tilde{P}_i^+} \times \left\{ \bullet \begin{array}{c} \xrightarrow{a_{i+1}} \\ \text{---} \\ \xrightarrow{a_i} \end{array} \bullet \right\} \right) \cup \bigcup_{i=2}^{n-1} \left(\Lambda_{\tilde{P}_i^-} \times \left\{ \bullet \begin{array}{c} \xrightarrow{a_{i-1}} \\ \text{---} \\ \xrightarrow{a_i} \end{array} \bullet \right\} \right). \quad (3.27)$$

Moreover, the sign of a maximal set of arrows of Λ_P agrees with the sign of the corresponding maximal set of arrows in $\Lambda_{\tilde{P}_i^+}$ and is opposite to the sign of the corresponding maximal set of arrows of $\Lambda_{\tilde{P}_i^-}$.

All we need to note here is that the above remark ensures that both $\Lambda_{\tilde{P}_n^\pm}$ and $\Lambda_{\tilde{P}_1^-}$ are empty, so are left out at the right hand side, and that cutting off a 2-gon of the form



corresponds to a forward arrow, hence contributes a sign +1 to the maximal dissection, while  corresponds to a backward arrow, hence contributes a sign -1.

Due to the bijection (3.27), we can rewrite eq. (3.14) by first summing over all subpolygons \tilde{P}_i^\pm , followed by a sum over all possible elements in $\Lambda_{\tilde{P}_i^\pm}$. The inner sum then evaluates to the symbol of the subpolygon \tilde{P}_i^\pm , and we are left with

$$\mathcal{S}(P) = \sum_{i=1}^{n-1} \mathcal{S}(\tilde{P}_i^+) \otimes \mu\left(\text{Diagram of 2-gon with vertices } a_{i+1} \text{ and } a_i \text{, arrow from } a_{i+1} \text{ to } a_i\right) - \sum_{i=2}^{n-1} \mathcal{S}(\tilde{P}_i^-) \otimes \mu\left(\text{Diagram of 2-gon with vertices } a_{i-1} \text{ and } a_i \text{, arrow from } a_{i-1} \text{ to } a_i\right), \quad (3.28)$$

where the relative minus sign arises because, as discussed above, $\Lambda_{\tilde{P}_i^\pm}$ contribute with opposite signs. After identification of \tilde{P}_i^+ and \tilde{P}_i^- , eq. (3.28) agrees with the recursion (3.26) modulo the additivity of the symbol. In order to finish the proof, we need to show that also the base steps of the recursions are the same. It is indeed easy to check by explicit computation that, e.g., the symbol of $G(a_2, a_1; a_3)$ obtained from the recursive definition (3.26) agrees with the symbol obtained from our polygon construction, eq. (3.14).

We note that in ref. [63] the d log of a constant is put to 0. Although this seems rather natural (and turns out to be sufficient in several cases), we advocate to use a refined version of this (which is what is typically used when working with symbols in a number field, in accordance with ref. [62]): for each element of a set of multiplicatively independent elements in a given number field one can choose a logarithm independently but then the logarithm of any product formed from those elements is determined. For example, we will see in section 6 that in the context of harmonic polylogarithms the only constants that need to be treated in this fashion are powers of 2, and hence it is sufficient to think of “2” as an irreducible element. The reason for considering this refined version is that it is very helpful for recognising functions from which a given symbol might originate. In particular, it has proved to be very useful, e.g., when “recognizing” HPL’s for keeping track of terms which come from a (shuffle) product of polylogarithms, see section 6.

While the symbol of a multiple polylogarithm obtained by considering the maximal set of arrows of the associated decorated polygon is equivalent to the symbol obtained from the recursive definition (3.26), we believe that both approaches have their virtues. While the latter might be easier to implement into a computer program in general, it is strictly speaking only valid in the case of generic arguments of the polylogarithms. Indeed, if the arguments are non-generic, we obtain divergences in the right-hand side of eq. (3.25), e.g. when $a_i = a_{i+1}$ for some i . It is then in principle necessary to resort to a careful regularization to deal with the degenerate cases [2]. The definition of the symbol based on the decorated polygons, being combinatorial in nature, avoids this problem by construction and allows to identify very easily the degenerate cases (they correspond, e.g., to arrows ending on a side whose decoration is zero), and to discard them from the start, avoiding in this way the need of regularization. Furthermore, as we will see in the next section,

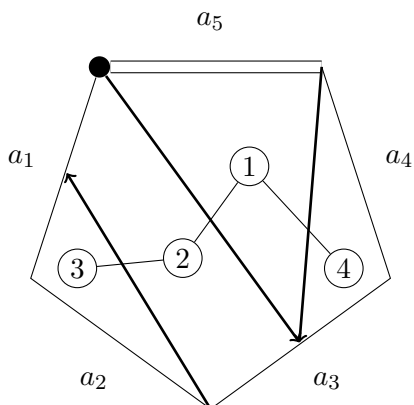


Figure 1. A (maximal) dissection and dual tree for a pentagon.

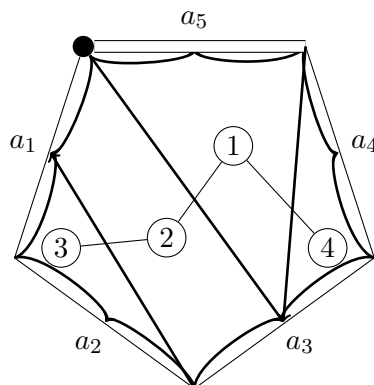


Figure 2. Figure 1, with all trivial arrows added.

the polygon approach has naturally built in the refined ‘ $d \log$ -prescription’, because the combinatorial nature of the construction does not make a distinction between constants as e.g. “2” for which one might be typically tempted to define $d \log 2 = 0$.

Relating polygon dissections to binary trees. We indicate how to directly relate the terms in the symbol definition via polygons in section 3.1 to the one via binary trees given in ref. [8], section 4.4, by working out a concrete example. For details, we refer to [97].

The map attaching to a maximal dissection a binary tree is not immediate and it is helpful to invoke associated ternary trees (i.e., trees where all the vertices are tetravalent) with level structure as a suitable index set. Note that this map is surjective but far from injective.

Let us concentrate on the terms in the symbol of a multiple polylogarithm arising from a given linear order on the dual tree of the maximal dissection shown in figure 1. Our goal is to map this maximal dissection for this given linear order to a binary tree. We start by adding all possible trivial arrows to the dissection, i.e., all possible arrows that end on adjacent sides (see figure 2).

Next, we ‘complete’ the dual tree of a maximal polygon dissection (including all the trivial arrows) in a natural way to a *ternary tree*, as shown in figure 3. The new edges correspond to (the parts cut-off by) the trivial arrows. We then remove the arrows (and trivial arrows) and extend the leaves of the dual tree to the sides of the polygon, thereby decomposing the polygon into $2n$ regions, half of which are naturally associated with one vertex each and are shaded in figure 4, while the remaining half are associated with one side each.

The root vertex of the ternary tree (marked with a square) is inherited from the root edge of the polygon (and is on the leaf closest to the first vertex of the polygon on the root edge). We ‘break’ the polygon at the root vertex of the ternary tree and ‘roll-out’ in a similar way to the last step in section 3.1. The edges of the polygon will then form a base line, which is not part of the ternary tree. Applying this to figure 4 we obtain figure 5. It is crucial to note that we induce a height ordering on the internal vertices of the level ternary tree inherited from the linear order on the dual tree.

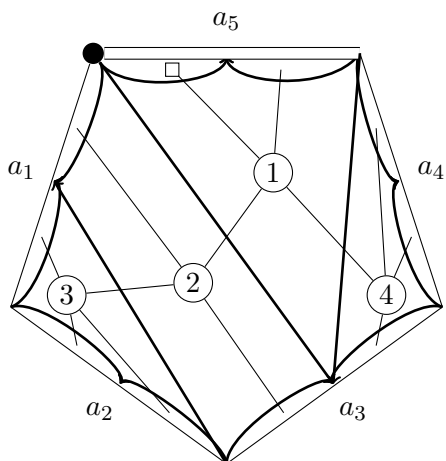


Figure 3. Figure 2, with its associated ternary tree (a square denoting its root side).

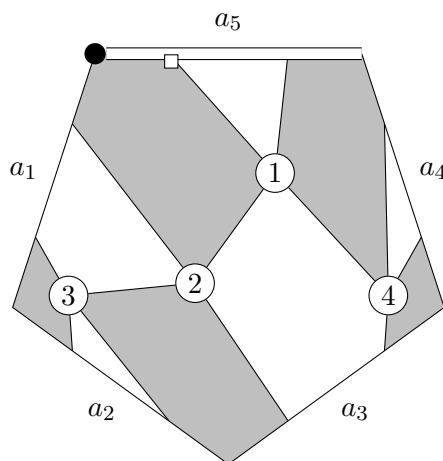


Figure 4. Figure 1, with the shaded regions induced by the ternary tree.

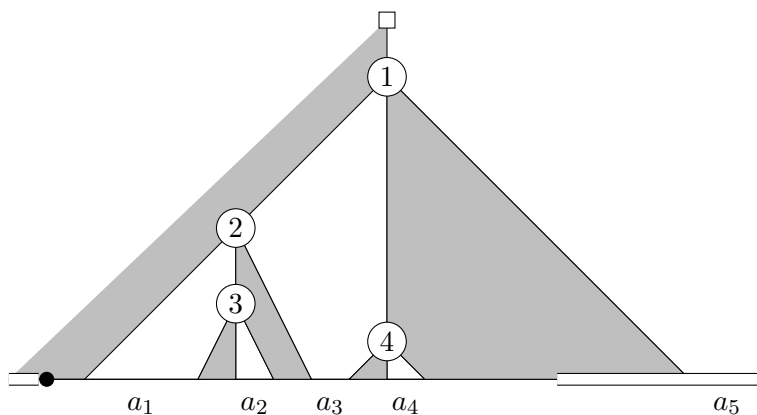


Figure 5. The ternary tree associated to the rolled-out version of the pentagon in figure 4.

The final step involves a horizontal “shrinking” of the shaded regions which, due to the choice of shading, will alternate along the base line. Within each shaded region we identify all points of the same height. The horizontal shrinking process will identify two outgoing edges for each internal vertex (and leave the other two in place), whence a binary tree.

For our example, we show a step which is part of the way through the horizontal contraction (Figure 6). This gives rise, by construction, to a binary tree (Figure 7).

We have thus shown how we can relate the maximal dissection with a given linear order on its dual tree to a binary level tree. It is easy to see that we obtain all the binary trees (with the same decoration as the polygon) in this way. Now a typical term $\frac{a_{i+1}-a_i}{a_{i-1}-a_i}$ in the setting of ref. [8], section 4.4, gives rise to two terms $\frac{a_{i+1}-a_i}{-a_i}$ and $\frac{a_{i-1}-a_i}{-a_i}$ (the latter giving rise to a sign) in the polygon setting, and in general an elementary tensor in weight w in the former setting gives rise to 2^k terms ($k \leq w$) in the latter one. Moreover, only the latter setting gives rise to signs, hence the binary tree description seems often preferable. A more detailed proof, establishing in fact a bijection between the terms arising from maximal dissections on the one hand and from the binary trees on the other, will be given in [97].

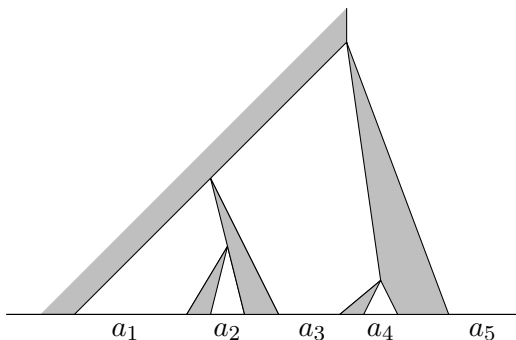


Figure 6. Intermediate stage for the horizontal contraction of Figure 5 to Figure 7.

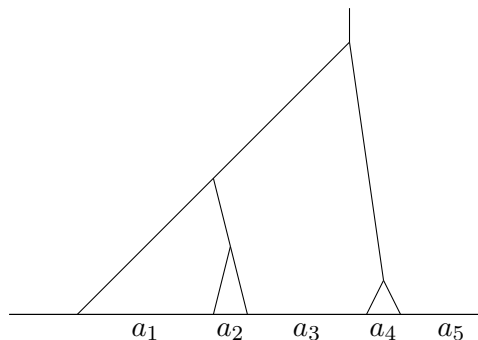


Figure 7. The ternary tree associated to the maximal dissection of figure 1.

Symbols for classical polylogarithms. The polygons attached to classical polylogarithms $Li_m(x) = -G(\underbrace{0, \dots, 0}_{m-1 \text{ terms}}, 1; x)$, are given by decorations 1 (for the first side) and 0 (for the remaining non-root sides) as well as x (for the root side). Their attached symbol consists of (the negative of) a *single* elementary tensor, in fact we have

$$\mathcal{S}(Li_m(x)) = - \left((1-x) \otimes \underbrace{x \otimes \dots \otimes x}_{m-1 \text{ factors}} \right), \tag{3.29}$$

where we have m factors (“weight” m) on the right hand side. (Note the parentheses which separate the coefficient, here -1 , from the actual tensor, to avoid a misinterpretation as $(x-1) \otimes x \otimes \dots \otimes x$.)

Such tensors have long been considered in connection with functional equations of polylogarithms — in fact, Zagier [5, 60] has given a criterion for such equations built on those tensors, which has been used (cf. ref. [61]) to find the first non-trivial equations for Li_6 and Li_7 (beyond weight 7 none are known), and the corresponding expressions for multiple polylogarithms are important already in Goncharov’s early work (e.g. [7]) where he generalises the underlying tensor algebra considerably.

4 A simple example

The symbol attached to $G(-1, 1; x)$. In this section we illustrate the fact that the symbol calculus provides a convenient tool to detect functional equations among (multiple) polylogarithms, on the example of $G(-1, 1; x)$ (which happens to coincide with the HPL $-H(-1, 1; x)$). Even though we could of course immediately apply eq. (2.13) to express $G(-1, 1; x)$ in terms of classical polylogarithms, we will derive a similar functional equation using the tensor calculus introduced in the previous section. The multiple polylogarithm $G(-1, 1; x)$ is associated to a trigon via

$$G(-1, 1; x) \iff P(1, -1, x) = \begin{matrix} & x & \\ \bullet & \triangle & \\ 1 & & -1 \end{matrix} . \tag{4.1}$$

The three maximal dissections of the trigon can easily be translated into the tensor associated to the polylogarithm,

$$\begin{array}{c}
 \begin{array}{c} \bullet \\ \text{---} x \\ \text{---} 1 \quad \text{---} -1 \\ \triangle \\ \text{---} 1 \quad \text{---} -1 \end{array} \\
 \downarrow \\
 \begin{array}{c} \bullet \\ \text{---} x \\ \text{---} 1 \quad \text{---} -1 \\ \triangle \\ \text{---} 1 \quad \text{---} -1 \end{array} + \begin{array}{c} \bullet \\ \text{---} x \\ \text{---} 1 \quad \text{---} -1 \\ \triangle \\ \text{---} 1 \quad \text{---} -1 \end{array} - \begin{array}{c} \bullet \\ \text{---} x \\ \text{---} 1 \quad \text{---} -1 \\ \triangle \\ \text{---} 1 \quad \text{---} -1 \end{array} \\
 \downarrow \\
 \mu\left(\begin{array}{c} \bullet \\ \text{---} x \\ \text{---} -1 \\ \text{---} 1 \end{array}\right) \otimes \mu\left(\begin{array}{c} \bullet \\ \text{---} -1 \\ \text{---} 1 \end{array}\right) + \mu\left(\begin{array}{c} \bullet \\ \text{---} x \\ \text{---} 1 \end{array}\right) \otimes \mu\left(\begin{array}{c} \bullet \\ \text{---} x \\ \text{---} -1 \end{array}\right) - \mu\left(\begin{array}{c} \bullet \\ \text{---} x \\ \text{---} 1 \end{array}\right) \otimes \mu\left(\begin{array}{c} \bullet \\ \text{---} 1 \\ \text{---} -1 \end{array}\right)
 \end{array}$$

The last line allows to read off the symbol of $G(-1, 1; x)$ as

$$\begin{aligned}
 \mathcal{S}(G(-1, 1; x)) &= \left(1 - \frac{x}{-1}\right) \otimes \left(1 - \frac{-1}{1}\right) + \left(1 - \frac{x}{1}\right) \otimes \left(1 - \frac{x}{-1}\right) - \left(1 - \frac{x}{1}\right) \otimes \left(1 - \frac{1}{-1}\right) \quad (4.2) \\
 &= (1+x) \otimes 2 + (1-x) \otimes (1+x) - (1-x) \otimes 2.
 \end{aligned}$$

Before turning to the question of how to attach a function to this symbol, let us briefly comment on how this symbol could have been obtained by using the recursive definition (3.26). Using eq. (3.25), we obtain

$$\begin{aligned}
 dG(-1, 1; x) &= G(-1; x) d \log \left(1 - (-1)\right) + G(1; x) d \log \left(\frac{(-1) - x}{(-1) - 1}\right) \quad (4.3) \\
 &= G(-1; x) d \log 2 + G(1; x) d \log(1+x) - G(1; x) d \log 2.
 \end{aligned}$$

The three terms in the last line of this equation are in one-to-one correspondence with the three terms in the symbol in eq. (4.2). Note, however, that we had to treat all the arguments of the (three-variable) function $G(\bullet, \bullet; \bullet)$ as generic, and to use the refined ‘ $d \log$ -prescription’, i.e. $d \log 2 \neq 0$. Putting to zero all the $d \log 2$ terms is equivalent to putting to zero all elementary tensors in the symbol where a factor inside a tensor product is constant [63]. As we will see below, we prefer to keep these terms as they provide us with valuable information about the function that should be associated to the symbol.

Attaching a function to a symbol. Since every multiple polylogarithm of weight two can be expressed as a combination of classical polylogarithms, we make the ansatz that $G(-1, 1; x)$ can be written, up to an additive constant, in the form

$$\sum_i c_i \text{Li}_2(f_i(x)) + \sum_{j,k} c_{jk} \log(g_j(x)) \log(h_k(x)), \quad (4.4)$$

such that the tensor associated to this expression corresponds to the tensor in eq. (4.2), where c_i and c_{jk} are rational numbers and f_i, g_j and $h_k \in \mathbb{Q}(x)$ are rational functions. We subdivide this problem into smaller ones by postulating that we can distinguish between

the three different contributions in eq. (4.2). By using a procedure suggested in ref. [63, 78] we can distinguish the first sum in eq. (4.4) from the second by projecting the respective symbols onto their symmetric or alternating parts: each term in the second sum will give zero contribution for the latter one, while each summand in the first sum will give a non-zero contribution. Indeed, the tensor associated to a product of logarithms is totally symmetric, and hence its contribution to the tensor vanishes if we project onto the antisymmetric component of the tensor in eq. (4.2).

Preparatory steps: decomposing tensors into symmetric and antisymmetric parts. We recall that, for a vector space V (over \mathbb{C} , say, or more generally over a field of characteristic $\neq 2$) there is a direct sum decomposition into symmetric and antisymmetric parts $V \otimes V = (V \odot V) \oplus (V \wedge V)$ (other notations, as used e.g. in refs. [5] or [7], are $V \odot V = \text{Sym}^2(V)$, and $V \wedge V = \bigwedge^2 V$), and $V \odot V$ is generated by $a \odot b$ (for some $a, b \in V$), while $V \wedge V$ is generated by $a \wedge b$ where we introduce the following rather standard notations for symmetric and antisymmetric tensors,

$$\begin{aligned} a \odot b &\equiv a \otimes b + b \otimes a, \\ a \wedge b &\equiv a \otimes b - b \otimes a. \end{aligned} \tag{4.5}$$

Back to the example. With this notation, the decomposition of a generic elementary rank two tensor (i.e., $a \otimes b$ for some $a, b \in V$) into its symmetric and antisymmetric components can be expressed as

$$a \otimes b = \frac{1}{2}(a \odot b) + \frac{1}{2}(a \wedge b). \tag{4.6}$$

Concentrating solely on the antisymmetric component of eq. (4.2), and using the antisymmetry of the wedge product which, e.g., induces $2 \wedge 2 = 0$, we obtain

$$\begin{aligned} (1+x) \wedge 2 - (1-x) \wedge 2 + (1-x) \wedge (1+x) \\ = \frac{1-x}{2} \wedge \frac{1+x}{2} = \left(1 - \frac{1+x}{2}\right) \wedge \frac{1+x}{2}. \end{aligned} \tag{4.7}$$

As the tensor associated to a product of logarithms does not have an antisymmetric component, eq. (4.7) suggests that it is the antisymmetric part of the tensor associated to some sum of dilogarithms, and from eq. (3.29), it is easily identified as the antisymmetric part of $\mathcal{S}\left(-\text{Li}_2\left(\frac{1+x}{2}\right)\right)$.

Having identified the dilogarithm contributions to $G(-1, 1; x)$, we can proceed via a bootstrap procedure and subtract off this contribution, leaving only a totally symmetric tensor

$$\begin{aligned} \mathcal{S}\left(G(-1, 1; x) + \text{Li}_2\left(\frac{1+x}{2}\right)\right) &= 2 \odot (1+x) - \frac{1}{2}(2 \odot 2) \\ &= \mathcal{S}\left(\log 2 \log(1+x) - \frac{1}{2} \log^2 2\right). \end{aligned} \tag{4.8}$$

Fixing the constant. We have shown that the tensor associated to $G(-1, 1; x)$ equals the tensor associated to the combination $-\text{Li}_2\left(\frac{1+x}{2}\right) + \log 2 \log(1+x) - \frac{1}{2} \log^2 2$. It would be premature, however, to conclude that both expressions are equal, but they are equal only up to an additive constant independent of x . Indeed, specializing to $x = 0$, and using the fact that $G(-1, 1; x = 0) = 0$ and $\text{Li}_2(1/2) = \frac{\pi^2}{12} - \frac{1}{2} \log^2 2$, we see that

$$\left\{ G(-1, 1; x) - \left[-\text{Li}_2\left(\frac{1+x}{2}\right) + \log 2 \log(1+x) - \frac{1}{2} \log^2 2 \right] \right\}_{|x=0} = \frac{\pi^2}{12}. \quad (4.9)$$

Thus, we obtain

$$G(-1, 1; x) = -\text{Li}_2\left(\frac{1+x}{2}\right) + \log 2 \log(1+x) - \frac{1}{2} \log^2 2 + \frac{\pi^2}{12}. \quad (4.10)$$

Note that this additive constant is not detected by the symbol, because

$$\mathcal{S}(\pi^2) = -\mathcal{S}(\log^2(-1)) = 0. \quad (4.11)$$

5 Integrating symbols: an algorithmic approach

In the previous section we illustrated how the symbol calculus can be used to derive a functional equation among polylogarithms. While in that weight two example all the steps were easily carried out by hand, an algorithmic approach is desirable in more complicated cases. In this section, we present our approach that often allows to integrate a symbol of a (transcendental) function, i.e., to find a function F , written as a linear combination of (products of) multiple polylogarithms, whose symbol matches a given tensor S , which in the following we always assume to satisfy the integrability condition (3.17).

From the example of the previous section it should be clear that the main challenges to address when going to higher weight w are

1. choosing appropriate arguments of the desired function types (as a few examples of function types of weight four, we list Li_4 , $\text{Li}_{2,2}$ or $\text{Li}_2 \log \log$) such that their symbols span the vector space in which the tensor S lies;
2. finding the generalization of the decomposition of weight two tensors into symmetric and antisymmetric parts (indicated in the simple example of weight two in section 4) to higher weights. This problem was addressed up to weight four in refs. [63, 78].

Let us assume that we have a linear combination S (with rational coefficients) of elementary tensors where the factors in each elementary tensor are rational functions of some variables x_1, \dots, x_n . In the following we assume the tensor to be of “pure weight” w , i.e., each elementary tensor is assumed to have the same number w of factors. Without loss of generality we can then assume that S takes the form (all sums are assumed to be finite)

$$S = \sum_{i_1, \dots, i_w} c_{i_1, \dots, i_w} \left(R_{i_1}(x_1, \dots, x_n) \otimes \dots \otimes R_{i_w}(x_1, \dots, x_n) \right), \quad (5.1)$$

where $R_{i_\ell}(x_1, \dots, x_n)$ ($1 \leq i_\ell \leq m(S)$ for some $m(S)$ determined by the initial tensor S) are rational functions in the variables x_i and the c_{i_1, \dots, i_w} are rational numbers. Distributivity (cf. eqs. (3.6), (3.7)) then implies that, without loss of generality, S can be rewritten (with adapted constants $\tilde{c}_{j_1, \dots, j_w} \in \mathbb{Q}$) as

$$S = \sum_{j_1, \dots, j_w} \tilde{c}_{j_1, \dots, j_w} \pi_{j_1} \otimes \dots \otimes \pi_{j_w}, \tag{5.2}$$

where $\pi_j = \pi_j(x_1, \dots, x_n)$ ($1 \leq j \leq K$ for some K) are suitably chosen rational functions which are *multiplicatively independent* (i.e., there is no non-trivial relation of the form $\prod_{j=1}^K \pi_j^{r_j} = \pm 1$, for $r_j \in \mathbb{Z}$). In practice, we achieve this by simply factorizing the rational functions $R_i(x_1, \dots, x_n)$ in eq. (5.1) into irreducible polynomials over \mathbb{Q} , say (i.e., polynomials in $\mathbb{Q}[x_1, \dots, x_n]$ that cannot be written as the product of two non-constant polynomials in $\mathbb{Q}[x_1, \dots, x_n]$). Let us denote the set of these polynomials by $P_S = \{\pi_1, \dots, \pi_K\}$, which will constitute our building blocks in the following. The symbol can then be seen as an element of the tensor algebra over the \mathbb{Q} -vector space generated by the elements from the set P_S (more precisely, we should consider it as an element of the w th grading of the tensor algebra over the \mathbb{Z} -module $\langle \pm \prod_{j=1}^K \pi_j^{r_j} \mid r_j \in \mathbb{Z} \rangle$). Our goal is now to find a function, say F , that is a rational linear combination of (multiple) polylogarithms (and products thereof) whose arguments are rational functions in the x_i such that $\mathcal{S}(F) = S$. The method to achieve this proceeds in two steps: first we have to decide on the types of functions that should appear in F , and then we have to concoct suitable rational functions in the x_i as arguments of these functions such that for some linear combination of these functions the resulting expression fulfills the condition $\mathcal{S}(F) = S$. Note that this latter step is not algorithmic in general, as it may involve finding arguments for the functions that have singularities *outside* P_S .

5.1 Choosing the types of functions

Our first goal is to construct a set of function types (our ‘basic types’) out of which we can construct our candidate function F . This involves two steps, and we want both the functions and their arguments to be ‘as simple as possible’, but we need to take into account that all the possible function types one can write down for a given weight are related by an abundance of functional equations. The main criterion we will use in the following is that a function type that can be written as a product of lower weight function types is ‘simpler’ than a function type of *pure weight/transcendentality* (i.e., a function that cannot be written as a sum of terms, each of which being a product of lower weights). Furthermore, we are guided by the conjecture (which the second author learned many years ago from Goncharov) that a multiple polylogarithm $\text{Li}_{m_1, \dots, m_k}$ with $m_j = 1$ for some j can be expressed in terms of multiple polylogarithms where no index is equal to 1. This conjecture suggests to put a restriction on the function types of pure transcendentality that can appear for a given weight. Furthermore, the shuffle and stuffle relations provide us with further constraints. As an example, we can deduce from

$$\text{Li}_{m_1, m_2}(x, y) + \text{Li}_{m_2, m_1}(y, x) = \text{Li}_{m_1}(x) \text{Li}_{m_2}(y) - \text{Li}_{m_1+m_2}(xy). \tag{5.3}$$

Weight	Basic function types of pure weight
1	$\log x$
2	$\text{Li}_2(x)$
3	$\text{Li}_3(x)$
4	$\text{Li}_4(x), \text{Li}_{2,2}(x, y)$
5	$\text{Li}_5(x), \text{Li}_{2,3}(x, y)$
6	$\text{Li}_6(x), \text{Li}_{2,4}(x, y), \text{Li}_{3,3}(x, y), \text{Li}_{2,2,2}(x, y, z)$

Table 1. Indecomposable multiple polylogarithms of pure weight ≤ 6 .

that we can hence ignore Li_{m_2, m_1} with $m_2 < m_1$. For low weights, natural sets of (presumably independent) functions which are indecomposable, i.e. cannot be written in terms of products of lower order functions, are given in table 1.

5.2 Finding the arguments

Having at hand a suitable set to construct the basic function types from, we still need to find the *arguments* of these function types. In the context of particle physics it has proved helpful to use guidance from educated guesses, motivated by physical constraints (cf. refs. [67, 68]), to construct the symbol and/or the functions expressing the desired physical quantities. To see how one might proceed even without any such guidance, let us concentrate first on classical polylogarithms only.

We start by defining, for $P_S = \{\pi_j\}_j$ as above, the set

$$\overline{P}_S = P_S \cup P'_S, \tag{5.4}$$

where P'_S is the set of all irreducible polynomials (possibly including prime numbers as well) that appear in $\pi_i \pm \pi_j$ and $1 \pm \pi_i$, $\forall \pi_i, \pi_j \in P_S$. The reason for introducing this enlarged set of polynomials will be clarified below.⁴ Let us denote the elements of \overline{P}_S by $\overline{\pi}_i$. Since S is constructed out of the irreducible polynomials $\pi_i \in P_S \subset \overline{P}_S$, it is perhaps reasonable to hope that all arguments appearing in the polylogarithmic expressions needed for S can be written in the form

$$R_{n_1, \dots, n_k}^\pm(x_1, \dots, x_n) = \pm \overline{\pi}_1^{n_1}(x_1, \dots, x_n) \dots \overline{\pi}_k^{n_k}(x_1, \dots, x_n), \tag{5.5}$$

where n_1, \dots, n_k are integers. Let us denote the set of these functions $R_{n_1, \dots, n_k}^\pm(x_1, \dots, x_n)$ by \mathcal{R}_S , i.e. this is, up to sign, the multiplicative span of the $\overline{\pi}_j \in \overline{P}_S$. Note that in practice it is often enough to consider \mathcal{R}_S to be the span of only a *subset* of the polynomials in \overline{P} . Finally, note that the set \mathcal{R}_S carries a group structure, given by the multiplication of rational functions.

⁴In principle one could even consider more complicated definitions for \overline{P}_S by adding sums and differences of monomials in the π_i . In practice, we observe that the definition given above worked in all cases we have considered so far.

Choosing arguments: the case of classical polylogarithms. However, not all of the functions in \mathcal{R}_S are automatically good candidates for arguments of polylogarithms. Indeed, if for example such a function appears as an argument of a classical polylogarithm, then by eq. (3.29) for $R = R_{n_1, \dots, n_k}^\pm(x_1, \dots, x_n)$ we can write

$$\mathcal{S}(\text{Li}_n(R)) = -(1 - R) \otimes \underbrace{R \otimes \dots \otimes R}_{(n-1) \text{ times}}. \tag{5.6}$$

It is now easy to see that if we want this tensor, for given R , to be an element of the tensor algebra of the vector space generated by the set \overline{P}_S , then we need to impose the constraint

$$1 - R \in \mathcal{R}_S. \tag{5.7}$$

Let us introduce the notation

$$\mathcal{R}_S^{(1)} = \{R \in \mathcal{R}_S \mid 1 - R \in \mathcal{R}_S\} \subset \mathcal{R}_S. \tag{5.8}$$

It follows that, for $R \in \mathcal{R}_S^{(1)}$, the symbol of $\text{Li}_n(R)$ ($n \geq 1$) is a linear combination of tensors of the form $\overline{\pi}_{\ell_1} \otimes \dots \otimes \overline{\pi}_{\ell_n}$. Hence all the rational functions in the set $\mathcal{R}_S^{(1)}$ are good candidates for arguments of the classical polylogarithms that can appear in our function F . Note that $\mathcal{R}_S^{(1)}$ is no longer a group in general. It can, however, be given some more structure by considering the following action, on the group of non-zero rational functions, of the permutation group S_3 on three letters, defined for a rational function R and rational functions σ_i of one variable by

$$\begin{aligned} \sigma_1(R) &= R, & \sigma_2(R) &= 1 - R, & \sigma_3(R) &= 1/R, \\ \sigma_4(R) &= 1/(1 - R), & \sigma_5(R) &= 1 - 1/R, & \sigma_6(R) &= R/(R - 1). \end{aligned} \tag{5.9}$$

Note that the σ_j form a group (under composition of functions) isomorphic to S_3 . It is easy to check that $\mathcal{R}_S^{(1)}$ is closed under this action of S_3 . As S_3 is generated by the two elements σ_2 and σ_3 , it is enough to check that $\mathcal{R}_S^{(1)}$ is closed under these two maps. Closure under σ_2 is trivial by definition of $\mathcal{R}_S^{(1)}$. To see that it is also closed under σ_3 we have to check that $\forall R \in \mathcal{R}_S^{(1)}, 1 - \sigma_3(R) \in \mathcal{R}_S$. Indeed, we have

$$1 - \sigma_3(R) = 1 - 1/R = -(1 - R) R^{-1} \in \mathcal{R}_S, \tag{5.10}$$

because of the group structure of \mathcal{R}_S .

Before generalizing this construction to the case of polylogarithms of depth > 1 , let us briefly comment on why we needed to enlarge the set of irreducible polynomials out of which the set $\mathcal{R}_S^{(1)}$ is constructed from P_S to \overline{P}_S . In some cases it can happen that, if we are given a combination of polylogarithms whose arguments are rational functions drawn from the set $\mathcal{R}_S^{(1)}$, then there are cancellations in the symbol S of this function such that the irreducible polynomials that appear in S are only a subset of the irreducible polynomials in $\mathcal{R}_S^{(1)}$, i.e., $P_S \subset \overline{P}_S$. Let us illustrate this on an example, and let us consider the function

$$f(x) = -\frac{1}{3}\text{Li}_2(1 - x^3) + \text{Li}_2(1 - x). \tag{5.11}$$

The polynomials appearing as arguments of the polylogarithms are products of the irreducible polynomials x , $1 - x$ and $1 + x + x^2$. The symbol of f however only involves two out of the three polynomials,

$$\mathcal{S}[f(x)] = x \otimes (1 + x + x^2), \tag{5.12}$$

and so $P_S = \{x, 1+x+x^2\}$. Thus, writing $\pi_1(x) = \bar{\pi}_1(x) = x$ and $\pi_2(x) = \bar{\pi}_2(x) = 1+x+x^2$, we see that we have to include $\bar{\pi}_3(x) = 1 - \pi_1(x) = 1 - x$ into the definition of \bar{P}_S . Similar considerations lead to the definition of \bar{P}_S given at the beginning section. We stress, though, that it is in practice often sufficient to consider the case $\bar{P}_S = P_S$, and only to enlarge the set P_S if no solutions could be found in this way.

Choosing arguments: the case of polylogarithms of depth > 1 . So far we have only considered classical polylogarithms, but in general we should also be able to make a sensible ansatz for the arguments of *multiple* polylogarithms of depth greater than one. In the following, we find it more convenient to work with the functions G_{m_1, \dots, m_k} defined in eq. (2.10) rather than with the functions $\text{Li}_{m_1, \dots, m_k}$. As the two function types are related via eq. (2.10), one can easily convert from one representation to the other.

Let us consider a multiple polylogarithm of depth two, say $G_{2,2}$. We are hence looking for a pair of rational functions $(R_1, R_2) \in \mathcal{R}_S \times \mathcal{R}_S$ that are good candidates for the arguments of $G_{2,2}$. The symbol of $G_{2,2}(R_1, R_2)$ is given by

$$\begin{aligned} \mathcal{S}(G_{2,2}(R_1, R_2)) = & - \left(1 - \frac{1}{R_1}\right) \otimes R_1 \text{III} \left[\left(1 - \frac{R_1}{R_2}\right) \otimes \frac{R_1}{R_2} \right] \\ & - \left(1 - \frac{1}{R_2}\right) \otimes R_2 \text{III} \left[\left(1 - \frac{R_2}{R_1}\right) \otimes \frac{R_2}{R_1} \right] - \left(1 - \frac{1}{R_2}\right) \otimes \left(1 - \frac{R_2}{R_1}\right) \otimes R_1 \text{III} R_2 \\ & + 2 \left(1 - \frac{1}{R_2}\right) \otimes \left(1 - \frac{R_2}{R_1}\right) \otimes R_1 \otimes R_1 - 2 \left(1 - \frac{1}{R_2}\right) \otimes \left(1 - \frac{1}{R_1}\right) \otimes R_1 \otimes R_1 \\ & - \left(1 - \frac{1}{R_2}\right) \otimes R_2 \text{III} \left[\left(1 - \frac{1}{R_1}\right) \otimes R_1 \right], \end{aligned} \tag{5.13}$$

recalling our notation for the shuffle products (see eq. (3.11))

$$\begin{aligned} A \otimes B \otimes C \text{III} D &= A \otimes B \otimes C \otimes D + A \otimes B \otimes D \otimes C, \\ A \otimes B \text{III} (C \otimes D) &= A \otimes B \otimes C \otimes D + A \otimes C \otimes B \otimes D + A \otimes C \otimes D \otimes B. \end{aligned} \tag{5.14}$$

Using the same reasoning as for classical polylogarithms, we see that the candidate arguments for multiple polylogarithms of depth two are pairs of rational functions from the set

$$\mathcal{R}_S^{(2)} = \{(R_1, R_2) \in \mathcal{R}_S^{(1)} \times \mathcal{R}_S^{(1)} \mid R_1 - R_2 \in \mathcal{R}_S\}. \tag{5.15}$$

An important consequence is that no new rational functions are needed to construct the set $\mathcal{R}_S^{(2)}$, but all the information is already contained in $\mathcal{R}_S^{(1)}$. The new set $\mathcal{R}_S^{(2)}$ then consists of pairs of elements of $\mathcal{R}_S^{(1)}$, subject to the additional constraint that their difference must again be factorizable in terms of the same prime elements. Moreover, we saw that $\mathcal{R}_S^{(1)}$ is

endowed with a natural action of the group S_3 , defined in eq. (5.9). It is hence natural to ask for non-trivial symmetry groups that leave the set $\mathcal{R}_S^{(2)}$ invariant. First, it is easy to see that the defining condition for $\mathcal{R}_S^{(2)}$ is invariant (up to an overall sign) under swapping the two entries of any given pair. Second, the action of S_3 defined in eq. (5.9) induces a (simultaneous on both factors) action on $\mathcal{R}_S^{(1)} \times \mathcal{R}_S^{(1)}$, defined for $\sigma_i \in S_3$ by

$$(R_1, R_2) \xrightarrow{\sigma_i} (\sigma_i(R_1), \sigma_i(R_2)). \tag{5.16}$$

It is now easy to check that $\mathcal{R}_S^{(2)}$ is closed under this action. To see this, it is enough to check that $\sigma_i(R_1) - \sigma_i(R_2) \in \mathcal{R}_S$ for $i = 1, 2$ and whenever $(R_1, R_2) \in \mathcal{R}_S^{(2)}$. Indeed, we have

$$\begin{aligned} \sigma_2(R_1) - \sigma_2(R_2) &= (1 - R_1) - (1 - R_2) = -(R_1 - R_2) \in \mathcal{R}_S, \\ \sigma_3(R_1) - \sigma_3(R_2) &= 1/R_1 - 1/R_2 = -(R_1 - R_2)R_1^{-1}R_2^{-1} \in \mathcal{R}_S, \end{aligned} \tag{5.17}$$

where we used the fact that $R_1 - R_2 \in \mathcal{R}_S$ and that \mathcal{R}_S is a multiplicative group. Combining this S_3 symmetry with the invariance under an exchange of arguments, here R_1 and R_2 , we see that $\mathcal{R}_S^{(2)}$ is closed under the action of the group $S_3 \times S_2$, defined for $(\sigma, \rho) \in S_3 \times S_2$ by

$$(R_1, R_2) \xrightarrow{(\sigma, \rho)} (\sigma(R_{\rho(1)}), \sigma(R_{\rho(2)})), \tag{5.18}$$

i.e. the factor S_2 simply acts as a permutation of the entries.

The previous discussions for depths one and two readily generalize to higher depth. Our candidate arguments for the multiple polylogarithms of depth k are k -tuples of rational functions from the set

$$\mathcal{R}_S^{(k)} = \{(R_1, \dots, R_k) \in \mathcal{R}_S^{(1)} \times \dots \times \mathcal{R}_S^{(1)} \mid R_i - R_j \in \mathcal{R}_S, 1 \leq i < j \leq k\}, \tag{5.19}$$

and using exactly the same argument as in the depth two case, we see that $\mathcal{R}_S^{(k)}$ can be equipped with an action of the group $S_3 \times S_k$, acting on $(R_1, \dots, R_k) \in \mathcal{R}_S^{(k)}$ via

$$(R_1, \dots, R_k) \xrightarrow{(\sigma, \rho)} (\sigma(R_{\rho(1)}), \dots, \sigma(R_{\rho(k)})), \tag{5.20}$$

i.e. the factor S_k simply acts as a permutation of the entries.

5.3 Integrating the symbol (1)

We now turn to the problem of integrating a tensor S that satisfies the integrability condition (3.17). In practice, such tensors could come from computing a Feynman integral in terms of multiple polylogarithms, or from computing its symbol by other means [66–68, 70, 73, 76]. Our goal is to find a function F , more precisely a linear combination of (multiple) polylogarithms, such that $\mathcal{S}(F) = S$. The considerations of the previous section suggest writing down an ansatz for the set Φ of *functions* (i.e., function types, together with rational functions as arguments) to express F in. We assume the “folklore conjecture” that any functional equation of MPL’s involving different weights should actually split into

functional equations of pure weight, so we can assume S to be of pure weight. It is convenient to partition Φ into subsets $\Phi^{(w)}$ according to the weight w , say $\Phi^{(w)} = \{b_i^{(w)}\}_i$. Then we want to find rational coefficients $c_{i_1 \dots i_\ell}$ such that

$$\begin{aligned}
 S &= \sum_i c_i \mathcal{S} \left(b_i^{(w)} \right) + \sum_{\substack{i_1, i_2 \\ w_1 + w_2 = w}} c_{i_1 i_2} \mathcal{S} \left(b_{i_1}^{(w_1)} b_{i_2}^{(w_2)} \right) + \\
 &+ \sum_{\substack{i_1, i_2, i_3 \\ w_1 + w_2 + w_3 = w}} c_{i_1 i_2 i_3} \mathcal{S} \left(b_{i_1}^{(w_1)} b_{i_2}^{(w_2)} b_{i_3}^{(w_3)} \right) + \\
 &+ \dots + \sum_{i_1, \dots, i_w} c_{i_1 \dots i_w} \mathcal{S} \left(b_{i_1}^{(1)} \dots b_{i_w}^{(1)} \right)
 \end{aligned} \tag{5.21}$$

where we can assume non-increasing partitions of w , i.e. that $w_i \geq w_{i+1}$ for all i . In view of property (3.9), each term on the right hand side can be written as a shuffle of terms $\mathcal{S} \left(b_j^{(\ell)} \right)$. All the terms on both sides of this equation are elements of the grade w part of the tensor algebra over the vector space spanned by basis vectors labeled by the elements in \overline{P}_S . We know from linear algebra that a straightforward basis for the tensor space of weight w tensors over the vector space spanned by \overline{P}_S is given by $\{\overline{\pi}_{i_1} \otimes \dots \otimes \overline{\pi}_{i_w} | \overline{\pi}_{i_\ell} \in \overline{P}_S\}$. At this stage we have hence mapped the problem of finding a function F satisfying $\mathcal{S}(F) = S$ into a problem of linear algebra, more precisely the problem of finding the coefficients $c_{i_1 \dots i_\ell}$ such that eq. (5.21) is true. As we know a basis of the tensor algebra, we can just compute the coefficients by extracting and comparing the coefficient of each basis vector on either side of the equation and solve the ensuing linear system. Note that the function F obtained in this way is not unique. Indeed, the map \mathcal{S} is non-injective and hence it would be possible to find a different function F' such that $F - F'$ in the kernel of \mathcal{S} . This issue will be addressed in section 5.6.

5.4 A set of projectors

Even though we have solved the problem of integrating the symbol in principle and we have reduced it to a linear algebra problem, the linear system one has to solve can be quite large. It is therefore preferable to break it down into smaller problems, for example by introducing a suitable filtration on the target space⁵ which permits to successively solve the problem for the filtration pieces. Such a filtration would allow to separate the different contributions in eq. (5.21): to get the ball rolling, we would like to first solve for the coefficients of the functions $b_i^{(w)}$ without having to care about the product terms. This can be achieved by introducing a projector that sends to zero exactly all the product terms.

Definition 1. *Let V be a vector space. We define linear operators Π_w acting on elementary tensors of lengths $w \geq 1$ by $\Pi_1 = id$ and for $w \geq 2$ by*

$$\begin{aligned}
 &\Pi_w(a_1 \otimes \dots \otimes a_w) \\
 &= \frac{w-1}{w} [\Pi_{w-1}(a_1 \otimes \dots \otimes a_{w-1}) \otimes a_w - \Pi_{w-1}(a_2 \otimes \dots \otimes a_w) \otimes a_1] .
 \end{aligned} \tag{5.22}$$

⁵A filtration of a vector space V is a sequence $\{V_i\}_{1 \leq i \leq n}$ of subspaces of V such that $V_1 \subset V_2 \subset \dots \subset V_n = V$.

Similar sets of operators acting on shuffle algebras have already been defined in refs. [81, 95], differing from our definition only by an overall normalization (F. Brown, in a text in preparation on the representation theory of polylogarithms, uses essentially the same operators, referring to ref. [96], II section 2, Prop. 7a). More precisely, the operators in refs. [81, 95] are defined by $\rho_1 = \text{id}$, and for $w \geq 2$ by

$$\rho_w(a_1 \otimes \dots \otimes a_w) = \rho_{w-1}(a_1 \otimes \dots \otimes a_{w-1}) \otimes a_w - \rho_{w-1}(a_2 \otimes \dots \otimes a_w) \otimes a_1. \quad (5.23)$$

The exact correspondence between Π_w and ρ_w is simply, for all $w \geq 1$,

$$\Pi_w = \frac{1}{w} \rho_w. \quad (5.24)$$

It follows from refs. [81, 95] that the kernel of ρ_w corresponds exactly to the ideal⁶ generated by all shuffle products. Since Π_w and ρ_w only differ by an overall normalization, we immediately arrive at the following

Proposition 1. *The kernel of Π_w equals the ideal generated by all shuffle products, i.e., for every element ξ in a shuffle algebra, $\Pi_w(\xi) = 0$ if and only if ξ can be written as a linear combination of shuffle products.*

In other words, the projectors Π_w are by construction such that they annihilate precisely all shuffles. Conversely, if Π_w applied to some tensor ξ does not vanish, then it is not possible to express ξ as a linear combination of shuffle products.

The reason for the normalization factor $\frac{w-1}{w}$ in Definition 1 is that it makes Π_w idempotent, in other words, Π_w is a projector.

Proposition 2. *For any $w \in \mathbb{N}$, Π_w is idempotent, and hence a projector, i.e.*

$$\Pi_w^2 = \Pi_w. \quad (5.25)$$

Proof. In ref. [81], Lemma 1.2, and also in eq. (2) in ref. [95], it was shown that ρ_w satisfies the identity

$$w(a_1 \otimes \dots \otimes a_w) = \sum_{k=0}^{w-1} (a_1 \otimes \dots \otimes a_k) \Pi \rho_{w-k}(a_{k+1} \otimes \dots \otimes a_w). \quad (5.26)$$

We now act on both sides of this equation with ρ_w . Since ρ_w annihilates all shuffles, all the terms on the right-hand side vanish, except for $k = 0$. Hence we obtain

$$w \rho_w(a_1 \otimes \dots \otimes a_w) = \rho_w^2(a_1 \otimes \dots \otimes a_w), \quad (5.27)$$

and so $\Pi_w^2 = \Pi_w$, after dividing both sides by w^2 . □

⁶An ideal in a commutative algebra \mathcal{A} is an additive subgroup \mathcal{I} such that $\forall a \in \mathcal{A}$ and $\forall b \in \mathcal{I}$, one has $ab \in \mathcal{I}$.

5.5 Integrating the symbol (2)

The projectors defined in the previous section can be used to improve the chances of successful integration of the symbol described in section 5.3. As the kernel of Π_w corresponds to all possible linear combinations of shuffles, its effect on eq. (5.21) is to remove all product terms, i.e.,

$$\Pi_w S = \sum_i c_i \Pi_w \mathcal{S} \left(b_i^{(w)} \right), \tag{5.28}$$

and so we can solve for *fewer* coefficients c_i without having to worry about the product terms.

Having found the coefficients c_i , we are left with the determination of the coefficients of the product terms. In order to proceed via induction, we determine the behaviour of shuffles under tensor products of projectors (which are themselves projectors). We first study the instructive case of a projector $\Pi_\ell \otimes \Pi_{\ell'}$ which we apply to a shuffle

$$(a_1 \otimes \cdots \otimes a_k) \text{III} (b_1 \otimes \cdots \otimes b_{k'}), \tag{5.29}$$

where $k + k' = \ell + \ell' = w$ (still w denoting the weight), with $k \geq k' \geq 1$, $\ell \geq \ell' \geq 1$ and $k \neq \ell$. We can rewrite the shuffle in a form more suitable to applying $\Pi_\ell \otimes \Pi_{\ell'}$ by regrouping the sum (5.29) using deconcatenation (i.e. splitting into a left hand part and a right hand part) of the sequence $(1, \dots, k)$ into two blocks $(1, \dots, k_1)$ and $(k_1 + 1, \dots, k_1 + k_2)$, giving

$$\sum_{k_1=1}^{\min(k,\ell)} \left((a_1 \otimes \cdots \otimes a_{k_1}) \text{III} (b_1 \otimes \cdots \otimes b_{\ell-k_1}) \right) \otimes \left((a_{k_1+1} \otimes \cdots \otimes a_k) \text{III} (b_{\ell-k_1+1} \otimes \cdots \otimes b_{k'}) \right). \tag{5.30}$$

Applying $\Pi_\ell \otimes \Pi_{\ell'}$ to this sum will annihilate each of the $\min(k, \ell)$ summands individually, in fact it will annihilate the left hand factor except (possibly) for $k_1 = \ell$, since in the other cases there is a proper shuffle on the left which will be mapped to zero by Π_ℓ already, by virtue of Proposition 1. If the remaining case occurs, the corresponding right hand factor is mapped to zero by $\Pi_{\ell'}$ instead. For example, if $(k_1, k_2) = (3, 3)$ and $(\ell_1, \ell_2) = (4, 2)$, we rewrite

$$\begin{aligned} (a_1 \otimes a_2 \otimes a_3) \text{III} (b_1 \otimes b_2 \otimes b_3) &= (a_1 \text{III} (b_1 \otimes b_2 \otimes b_3)) \otimes (a_2 \otimes a_3) \\ &\quad + ((a_1 \otimes a_2) \text{III} (b_1 \otimes b_2)) \otimes (a_3 \text{III} b_3) \\ &\quad + ((a_1 \otimes a_2 \otimes a_3) \text{III} (b_1)) \otimes (b_2 \otimes b_3), \end{aligned} \tag{5.31}$$

where already the factor Π_4 of $\Pi_4 \otimes \Pi_2$ annihilates the left part of each of the terms.

More generally, for a partition $\lambda = (\lambda_1, \dots, \lambda_r)$ of w , we call the λ -*shuffle* of a w -fold tensor $a_1 \otimes \cdots \otimes a_w$ the product

$$(a_1 \otimes \cdots \otimes a_{\lambda_1}) \text{III} (a_{\lambda_1+1} \otimes \cdots \otimes a_{\lambda_1+\lambda_2}) \text{III} \dots \text{III} (a_{\lambda_1+\dots+\lambda_{r-1}+1} \otimes \cdots \otimes a_w) \tag{5.32}$$

and given any partition $\lambda' = (\lambda'_1, \dots, \lambda'_s)$ with $s < r$, the projector $\Pi_{\lambda'}$ will vanish on a λ -shuffle, as we can again regroup its terms according to λ' and find for each so combined summand at least one projector factor $\Pi_{\lambda'_j}$ which vanishes on the corresponding part. This sketches a proof of

Proposition 3. *For two different non-increasing partitions λ and λ' of w , of length $\ell(\lambda)$ and $\ell(\lambda')$, respectively, we have that a λ -shuffle vanishes under $\Pi_{\lambda'}$ whenever $\ell(\lambda) \geq \ell(\lambda')$.*

This proposition suggests to define a sequence of subspaces, each of which contains the next one (hence forming a descending filtration). For this, we consider the standard lexicographic order \succ on non-increasing partitions of w given by (denoting $\lfloor x \rfloor$ the largest integer $\leq x$)

$$(w) \succ (w-1, 1) \succ (w-2, 2) \succ \dots \succ (w - \lfloor \frac{w}{2} \rfloor, \lfloor \frac{w}{2} \rfloor) \succ (w-2, 1, 1) \succ \dots \succ (2, \underbrace{1, \dots, 1}_{w-2 \text{ slots}}) \succ (\underbrace{1, \dots, 1}_w). \tag{5.33}$$

As an example, in weight five we have the following non-decreasing partitions, ordered as

$$(5) \succ (4, 1) \succ (3, 2) \succ (3, 1, 1) \succ (2, 2, 1) \succ (2, 1, 1, 1) \succ (1, 1, 1, 1, 1). \tag{5.34}$$

From these, we form the descending filtration alluded to above by taking the span of all shuffles of partition types greater than a given type λ , say, and denote it by

$$\mathcal{F}_\lambda = \langle \lambda' - \text{shuffles in } V \text{ for all } \lambda \succ \lambda', \lambda \neq \lambda' \rangle. \tag{5.35}$$

Hence $\mathcal{F}_{(1, \dots, 1)}$ is the zero space, $\mathcal{F}_{(2, 1, \dots, 1)}$ is generated by all $(1, \dots, 1)$ -shuffles (i.e. fully symmetric tensors), $\mathcal{F}_{(3, 1, \dots, 1)}$ is generated by all $(1, \dots, 1)$ -shuffles and $(2, 1, \dots, 1)$ -shuffles etc. The proposition now guarantees that $\mathcal{F}_\lambda \subset \ker \Pi_\lambda$. Similarly, we put, for convenience, the “shifted” sequence

$$\mathcal{F}_\lambda^\succ = \langle \lambda' - \text{shuffles in } V \text{ for all } \lambda \succeq \lambda' \rangle, \tag{5.36}$$

so that $\mathcal{F}_{\lambda^+}^\succ = \mathcal{F}_\lambda$ where λ^+ denotes the immediate successor of λ in the lexicographic order \succ above.

We proceed by induction on λ with respect to the order \succ , starting with the shuffle $\lambda = (w)$. By the analysis above, we can put $S_{(w)} \equiv \sum_i c_i \mathcal{S}(b_i^{(w)})$, provided we have been able to solve for the c_i . This gives us the basis step for the induction. Now assume we have found an (integrable) tensor S_λ “approximating” S in the sense that $S - S_\lambda \in \mathcal{F}_\lambda$. Then we try to construct a “better approximation” S_{λ^+} for the successor partition λ^+ of λ by finding integrable tensors $T_{\lambda^+} \in \mathcal{F}_{\lambda^+}^\succ = \mathcal{F}_\lambda$ such that $S_{\lambda^+} \equiv S_\lambda + T_{\lambda^+}$ satisfies $S - S_{\lambda^+} \in \mathcal{F}_{\lambda^+}$. If we are successful in finding such a T_{λ^+} , this finishes the induction step.

We expect to be able to find such a T_{λ^+} using (certain sums of products of) multiple polylogarithms, by selecting for each part λ_r^+ of λ^+ an “indecomposable function type” of weight λ_r^+ (see table 1) and taking their product. In other words, we assume that the tensor T_{λ^+} can be written, as a linear combination of λ^+ -shuffles, in the form

$$T_{\lambda^+} = \sum_{i_1, \dots, i_l} c_{i_1, \dots, i_l} \mathcal{S}(b_{i_1}^{(\lambda_1^+)}) \text{III} \dots \text{III} \mathcal{S}(b_{i_l}^{(\lambda_l^+)}) , \tag{5.37}$$

with $l = \ell(\lambda^+)$ and $\lambda^+ = (\lambda_1^+, \dots, \lambda_l^+)$. In the weight 5 example of table 1, the respective polylogarithmic functions we consider are

$$\begin{aligned}
 (5) &\leftrightarrow \text{Li}_5(R_1), \text{Li}_{2,3}(R_1, R_2), \\
 (4, 1) &\leftrightarrow \text{Li}_4(R_1) \log R_2, \text{Li}_{2,2}(R_1, R_2) \log R_3, \\
 (3, 2) &\leftrightarrow \text{Li}_3(R_1) \text{Li}_2(R_2), \\
 (3, 1, 1) &\leftrightarrow \text{Li}_3(R_1) \log R_2 \log R_3, \\
 (2, 2, 1) &\leftrightarrow \text{Li}_2(R_1) \text{Li}_2(R_2) \log R_3, \\
 (2, 1, 1, 1) &\leftrightarrow \text{Li}_2(R_1) \log R_2 \log R_3 \log R_4, \\
 (1, 1, 1, 1, 1) &\leftrightarrow \log R_1 \log R_2 \log R_3 \log R_4 \log R_5,
 \end{aligned}
 \tag{5.38}$$

where the R_i correspond to the rational functions as indicated in section 5.2.

Since the proposition implies $\Pi_\lambda \mathcal{F}_\lambda = 0$, we get $\Pi_\lambda(S - S_\lambda) = 0$. Therefore we are essentially working at each step in the quotient space $\mathcal{F}_\lambda^\times / \mathcal{F}_\lambda$ and have to solve for considerably fewer coefficients than if we worked in $\mathcal{F}_\lambda^\times$. If at any stage we cannot find a T_{λ^+} with the desired property, i.e. cannot solve for the corresponding coefficients using our limited spanning set of input functions $b_i^{(s)}$, the algorithm stops (we can of course try to rerun it with a larger set of input functions). Otherwise, it ends with producing an integrable tensor $S_{(1,\dots,1)}$ with $S - S_{(1,\dots,1)} = 0$, solving the main part of our integration problem.

5.6 Elements in the kernel of the symbol map

The algorithmic approach we described in the previous section often allows us, given a tensor S , to construct a function F such that $\mathcal{S}(F) = S$. Let us now assume that the tensor S was obtained in some way from an analytic expression F_0 (representing, say, a Feynman integral). It would be premature to conclude that the function F we constructed is equal to the original expression F_0 , because they are only equal up to terms that are in the kernel of \mathcal{S} .

In the following we describe a way that, at least in most of the cases we studied so far, allows to fix this remaining ambiguity by parametrizing $F - F_0$ in a suitable way. The parametrization we are proposing takes the form

$$F - F_0 = \sum_i \tilde{c}_i \tilde{k}_i + \sum_\ell \sum_{\substack{i_1, \dots, i_\ell \\ w_1 + \dots + w_\ell = w}} c_{i_1 \dots i_\ell} k_{i_1 \dots i_\ell} b_{i_1}^{(w_1)} \dots b_{i_\ell}^{(w_\ell)}, \tag{5.39}$$

where $b_{i_k}^{(w_k)}$ are defined in the previous section, and $c_{i_1 \dots i_\ell}$ and \tilde{c}_i are rational coefficients, and $k_{i_1 \dots i_\ell}$ and \tilde{k}_i are generators of the kernel of the symbol map. Below we give a (non-exhaustive) list of such generators, which cover a wide range of applications. The free coefficients can then be fixed by considering special values for the variables x_i , e.g., $x_i = 0$ or $x_i = 1$, yielding a linear system for the coefficients.

In order for the above procedure to work we need to know the generators of the kernel of \mathcal{S} . Even though this is a very difficult task to settle in general, we can compile a list of (presumably transcendental) numbers whose symbol should be defined as zero.

- $\mathcal{S}(\log(-1)) = 0$. An associated polygon would be $\bullet \underbrace{\hspace{1.5cm}}_{-1} \bullet$, and this gives the symbol -1 which is zero (modulo 2-torsion which we ignore). In particular, this shows that the symbol does not detect the multivaluedness of the logarithm (let alone of the polylogarithms), and in particular does not suggest how to fix the branch cuts of the polylogarithms.
- All multiple zeta values (MZV's) are in the kernel of \mathcal{S} . Indeed, MZV's can be defined as the values in $x_i = 1$ for all $i = 1, \dots, k$ of the multiple polylogarithms $\text{Li}_{m_1, \dots, m_k}(x_1, \dots, x_k)$. Then it is easy to see that the associated decorated polygon will have all decorations equal to 0 or 1, and hence the symbol vanishes.
- In addition, there are combinations of transcendental expressions that individually have a non-vanishing symbol,⁷ but there is a linear combination with zero symbol, e.g.,

$$\mathcal{S}\left(\text{Li}_4\left(\frac{1}{2}\right) + \frac{1}{24} \log^4 2\right) = -\left(1 - \frac{1}{2}\right) \otimes \frac{1}{2} \otimes \frac{1}{2} \otimes \frac{1}{2} + \frac{1}{24}(2 \odot 2 \odot 2 \odot 2) = 0. \quad (5.40)$$

The previous example is a special case of a more general result for so-called colored multiple zeta values, defined by the alternating sums

$$\zeta(m_1, \dots, m_k; s_1, \dots, s_k) = \sum_{0 < n_1 < n_2 < \dots < n_k} \frac{s_1^{n_1} s_2^{n_2} \dots s_k^{n_k}}{n_1^{m_1} n_2^{m_2} \dots n_k^{m_k}}, \quad (5.41)$$

with $m_i \in \mathbb{N}$ and $s_i \in \{\pm 1\}$. It is easy to check that $\zeta(1, 1, 1, 1; -1, -1, 1, 1) = -\text{Li}_4\left(\frac{1}{2}\right)$, and so eq. (5.40) can be written as

$$\mathcal{S}(\zeta(1, 1, 1, 1; -1, -1, 1, 1) - \frac{1}{24} \log^4 2) = 0. \quad (5.42)$$

More generically, we have

Proposition 4.

1. If at least one of the m_i is different from ± 1 and $(m_1, s_1) \neq (1, 1)$, then

$$\mathcal{S}(\zeta(m_1, \dots, m_k; s_1, \dots, s_k)) = 0. \quad (5.43)$$

2. $\forall s_i \in \{-1, 1\}, \forall m \geq 1$,

$$\mathcal{S}\left(\zeta(\underbrace{1, \dots, 1}_{m \text{ times}}; -1, s_2, \dots, s_m) - \frac{1}{m!} \log^m \frac{1}{2}\right) = 0. \quad (5.44)$$

The proof of this proposition will be given in appendix C.

⁷We recall that the “refined” symbol of a constant is not necessarily zero.

6 Application: a spanning set for harmonic polylogarithms

In this section we illustrate our approach by expressing all harmonic polylogarithms (HPL's) up to weight four in terms of a spanning set of functions. We start by determining the arguments our spanning set should depend upon. Indeed, in the case of harmonic polylogarithms we can classify all the arguments of the spanning set of functions under the mild assumption that all the arguments should be rational functions. Then, it is easy to see that the polygons associated to HPL's correspond to polygons where the root edge is decorated by the variable x , and all other sides are decorated only by 0 or ± 1 . This implies that the tensor associated to an HPL is an element of the tensor algebra of the vector space generated by the formal basis vectors $[x]$, $[1-x]$, $[1+x]$ and $[2]$. Indeed, the decorations of the polygon associated to an HPL are all ± 1 or 0, except for the root which is decorated by the variable x . It is then easy to see that any dissection of this polygon will only involve the following five bigons

$$\begin{aligned} \mu\left(\begin{array}{c} \text{---} x \text{---} \\ \text{---} \text{---} \\ \text{---} 0 \end{array}\right) &= x, & \mu\left(\begin{array}{c} \text{---} x \text{---} \\ \text{---} \text{---} \\ \text{---} 1 \end{array}\right) &= 1-x, \\ \mu\left(\begin{array}{c} \text{---} x \text{---} \\ \text{---} \text{---} \\ \text{---} -1 \end{array}\right) &= 1+x, & \mu\left(\begin{array}{c} \text{---} -1 \text{---} \\ \text{---} \text{---} \\ \text{---} 1 \end{array}\right) &= \mu\left(\begin{array}{c} \text{---} 1 \text{---} \\ \text{---} \text{---} \\ \text{---} -1 \end{array}\right) = 2. \end{aligned} \tag{6.1}$$

Hence, the sets P_{HPL} and \overline{P}_{HPL} of irreducible polynomials defined in section 5 are

$$P_{HPL} = \{2, x, 1-x, 1+x\} \text{ and } \overline{P}_{HPL} = P_{HPL} \cup \{2 \pm x, 3 \pm x, 1 \pm 2x\}. \tag{6.2}$$

The most general rational function we can construct out of the irreducible polynomials in \overline{P}_{HPL} then reads

$$\pm x^{\alpha_1} (1-x)^{\alpha_2} (1+x)^{\alpha_3} 2^{\alpha_4} (2-x)^{\alpha_5} (2+x)^{\alpha_6} (3-x)^{\alpha_7} (3+x)^{\alpha_8} (1-2x)^{\alpha_9} (1+2x)^{\alpha_{10}}, \tag{6.3}$$

with $\alpha_i \in \mathbb{Z}$. In the case of HPL's, however, it turns out (a posteriori) that we can restrict ourselves to the following set of rational functions

$$\mathcal{R}_{HPL} = \{\pm 2^\delta x^\alpha (1-x)^\beta (1+x)^\gamma \mid \alpha, \beta, \gamma, \delta \in \mathbb{Z}\}. \tag{6.4}$$

In the following we denote the elements of \mathcal{R}_{HPL} by

$$R_{\alpha\beta\gamma\delta}^\pm(x) = \pm 2^\delta x^\alpha (1-x)^\beta (1+x)^\gamma, \tag{6.5}$$

with $\alpha, \beta, \gamma, \delta \in \mathbb{Z}$. It follows then from the previous section that we should consider the subset of these rational functions (those contained in the set $\mathcal{R}_{HPL}^{(1)}$; see eq. (5.7) and eq. (5.8)) that satisfy the constraint

$$1 - R_{\alpha\beta\gamma\delta}^\pm(x) = R_{\alpha'\beta'\gamma'\delta'}^s(x), \tag{6.6}$$

for some integers α', β', γ' and δ' , and $s = \pm 1$. A little algebra shows that quadruples $(\alpha, \beta, \gamma, \delta)$ are confined to the values given in table 2. The first line corresponds to constant arguments, and will not be discussed any further. Note that we could also include

s	α	β	γ	δ	$R_{\alpha\beta\gamma\delta}^s(x)$	s	α	β	γ	δ	$R_{\alpha\beta\gamma\delta}^s(x)$
-	0	0	0	0	-1	+	0	0	0	-1	1/2
+	1	0	0	0	x	-	1	0	0	0	$-x$
+	0	1	0	0	$1-x$	+	0	0	-1	0	$1/(1+x)$
+	2	0	0	0	x^2	+	0	1	1	0	$1-x^2$
-	1	-1	0	0	$x/(x-1)$	+	1	0	-1	0	$x/(x+1)$
+	0	1	-1	0	$(1-x)/(1+x)$	-	0	1	-1	0	$(x-1)/(x+1)$
-	2	-2	-2	0	$x^2/(x^2-1)$	+	0	2	-2	0	$(1-x)^2/(1+x)^2$
+	0	1	0	-1	$(1-x)/2$	+	0	0	1	-1	$(1+x)/2$
-	1	-1	0	1	$2x/(x-1)$	+	1	0	-1	1	$2x/(x+1)$
+	2	0	-2	2	$4x/(1-x)^2$	-	2	-2	0	2	$-4x/(x+1)^2$

Table 2. Solutions to the constraint (6.6). Only half of the solutions are shown, all other solutions being related by $R_{(-\alpha)(-\beta)(-\gamma)(-\delta)}^s(x) = 1/R_{\alpha\beta\gamma\delta}^s(x)$.

the inverses of the arguments in table 2. Using the inversion formula for the classical polylogarithms, we can however always express Li_n functions with inverted arguments in terms of polylogarithms taken at the arguments in table 2,

$$\text{Li}_n\left(\frac{1}{x}\right) = (-1)^{n-1} \text{Li}_n(x) + \text{products of lower weight terms}, \quad n \geq 2. \quad (6.7)$$

Since furthermore the arguments in table 2 are less than 1 for $x \in [0, 1]$, we will in the following only consider these functions as arguments of the Li_n functions. Note however that even the functions in table 2 are not completely independent, because we can use the distribution formula for the Li_n function to relate three of them, e.g.,

$$\text{Li}_n(x^2) = 2^{n-1} (\text{Li}_n(x) + \text{Li}_n(-x)), \quad (6.8)$$

and three others using the same equation with x replaced by $\frac{1-x}{1+x}$. We finally arrive at the conclusion that, if we want to reduce all HPL's to some small (possibly minimal) set of (multiple) polylogarithms, the Li_n sector of that set contains classical polylogarithms with 16 different arguments. As we will see below, for lower weights we can find more relations among the spanning set of functions, reducing its size even further. In the following we will also have to deal with multiple polylogarithms of depth greater than one. From the previous section we know that the pairs of arguments (R_1, R_2) of these functions come from table 2, subject to the additional constraint $R_1 - R_2 \in \mathcal{R}_{HPL}$.

Finally, using the results from the previous section, as well as some elementary identities among classical polylogarithms, we find the following spanning set of indecomposable functions up to weight four,

- for weight one,

$$\mathcal{B}_1^{(1)}(x) = \log x, \quad \mathcal{B}_1^{(2)}(x) = \log(1-x), \quad \mathcal{B}_1^{(3)}(x) = \log(1+x), \quad \mathcal{B}_1^{(4)}(x) = \log 2, \quad (6.9)$$

- for weight two,

$$\mathcal{B}_2^{(1)}(x) = \text{Li}_2(x), \quad \mathcal{B}_2^{(2)}(x) = \text{Li}_2(-x), \quad \mathcal{B}_2^{(3)}(x) = \text{Li}_2\left(\frac{1-x}{2}\right), \quad (6.10)$$

- for weight three,

$$\begin{aligned} \mathcal{B}_3^{(1)}(x) &= \text{Li}_3(x), & \mathcal{B}_3^{(2)}(x) &= \text{Li}_3(-x), & \mathcal{B}_3^{(3)}(x) &= \text{Li}_3(1-x), \\ \mathcal{B}_3^{(4)}(x) &= \text{Li}_3\left(\frac{1}{1+x}\right), & \mathcal{B}_3^{(5)}(x) &= \text{Li}_3\left(\frac{1+x}{2}\right), & \mathcal{B}_3^{(6)}(x) &= \text{Li}_3\left(\frac{1-x}{2}\right), \\ \mathcal{B}_3^{(7)}(x) &= \text{Li}_3\left(\frac{1-x}{1+x}\right), & \mathcal{B}_3^{(8)}(x) &= \text{Li}_3\left(\frac{2x}{x-1}\right), \end{aligned} \quad (6.11)$$

- for weight four,

$$\begin{aligned} \mathcal{B}_4^{(1)}(x) &= \text{Li}_4(x), & \mathcal{B}_4^{(2)}(x) &= \text{Li}_4(-x), \\ \mathcal{B}_4^{(3)}(x) &= \text{Li}_4(1-x), & \mathcal{B}_4^{(4)}(x) &= \text{Li}_4\left(\frac{1}{1+x}\right), \\ \mathcal{B}_4^{(5)}(x) &= \text{Li}_4\left(\frac{x}{x-1}\right), & \mathcal{B}_4^{(6)}(x) &= \text{Li}_4\left(\frac{x}{x+1}\right), \\ \mathcal{B}_4^{(7)}(x) &= \text{Li}_4\left(\frac{1+x}{2}\right), & \mathcal{B}_4^{(8)}(x) &= \text{Li}_4\left(\frac{1-x}{2}\right), \\ \mathcal{B}_4^{(9)}(x) &= \text{Li}_4\left(\frac{1-x}{1+x}\right), & \mathcal{B}_4^{(10)}(x) &= \text{Li}_4\left(\frac{x-1}{x+1}\right), \\ \mathcal{B}_4^{(11)}(x) &= \text{Li}_4\left(\frac{2x}{x+1}\right), & \mathcal{B}_4^{(12)}(x) &= \text{Li}_4\left(\frac{2x}{x-1}\right), \\ \mathcal{B}_4^{(13)}(x) &= \text{Li}_4(1-x^2), & \mathcal{B}_4^{(14)}(x) &= \text{Li}_4\left(\frac{x^2}{x^2-1}\right), \\ \mathcal{B}_4^{(15)}(x) &= \text{Li}_4\left(\frac{4x}{(x+1)^2}\right). \end{aligned} \quad (6.12)$$

These functions are sufficient to express all harmonic polylogarithms up to weight three. Starting from weight four, we need to extend the set of functions by adjoining *multiple* polylogarithms. We find that it is enough to add the following three supplementary functions in order to express all harmonic polylogarithms up to weight four,

$$\mathcal{B}_4^{(16)}(x) = \text{Li}_{2,2}(-1, x), \quad \mathcal{B}_4^{(17)}(x) = \text{Li}_{2,2}\left(\frac{1}{2}, \frac{2x}{x+1}\right), \quad \mathcal{B}_4^{(18)}(x) = \text{Li}_{2,2}\left(\frac{1}{2}, \frac{2x}{x-1}\right). \quad (6.13)$$

Note that, if the vector of singularities only takes values in the set $\{0, 1\}$, we can restrict ourselves to the smaller spanning set,

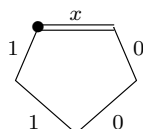
- for weight one: $\mathcal{B}_1^{(1)}(x), \mathcal{B}_1^{(2)}(x)$,
- for weight two: $\mathcal{B}_2^{(1)}(x)$,

- for weight three: $\mathcal{B}_3^{(1)}(x), \mathcal{B}_3^{(3)}(x),$
- for weight four: $\mathcal{B}_4^{(1)}(x), \mathcal{B}_4^{(3)}(x), \mathcal{B}_4^{(5)}(x).$

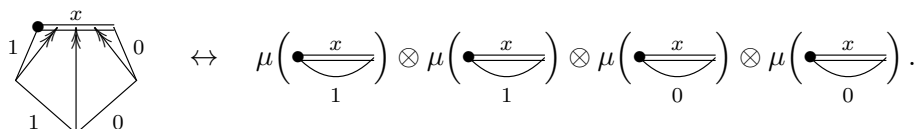
Our choice for the spanning set is of course not unique, and we might have chosen a different set of functions, related to the $\mathcal{B}_i^{(j)}$ functions via functional equations. Our choice was motivated by the fact that $\mathcal{B}_i^{(j)}(x)$ is manifestly real for $x \in [0, 1]$. Note that outside this interval, the branching structure of these functions can be more complicated. This issue is addressed in appendix E.

6.1 Example

Let us conclude this section by giving an example of how we can apply the procedure of section 5 to express a generic HPL of weight four in terms of the functions $\mathcal{B}_j^{(i)}$. We discuss in detail the example of $H(0, 0, 1, 1; x) = S_{2,2}(x)$, all other cases being similar. For the list of all results, we refer to appendix G. We start by deriving the tensor associated to $H(0, 0, 1, 1; x)$. The polygon associated to $H(0, 0, 1, 1; x) = G(0, 0, 1, 1; x)$ is



There is only one relevant maximal dissection of this pentagon (all other dissections give rise to twogons with decorations 0 and / or 1 which vanish by definition),



From this dissection we can immediately read of the symbol of $H(0, 0, 1, 1; x)$ as

$$\mathcal{S}(H(0, 0, 1, 1; x)) = (1 - x) \otimes (1 - x) \otimes x \otimes x. \tag{6.14}$$

Note that in general the symbol of a harmonic polylogarithm $H(a_1, \dots, a_w; x)$, with $a_i \in \{0, 1\}$ is simply given by $(-1)^k (a_w - x) \otimes \dots \otimes (a_1 - x)$, where k is the number of a_i 's equal to 1. Before turning to the question of how to express $H(0, 0, 1, 1; x)$ in terms of the spanning set, let us review how we could have obtained the symbol (6.14) using the recursive definition of the symbol (3.26). Note however, that in this case we cannot apply eq. (3.25) immediately, as the arguments of $G(0, 0, 1, 1; x) = H(0, 0, 1, 1; x)$ are not generic and we would hence obtain divergences in the right-hand side of eq. (3.25). We therefore need to use a regularized version of the differential equation (3.25) [2],

$$dH(0, 0, 1, 1; x) = H(0, 1, 1; x) d \log x, \tag{6.15}$$

and so

$$\mathcal{S}(H(0, 0, 1, 1; x)) = \mathcal{S}(H(0, 1, 1; x)) \otimes x. \tag{6.16}$$

We can iterate this procedure to compute the symbol of $H(0, 1, 1; x)$. The differential equation for $H(0, 1, 1; x)$ is

$$dH(0, 1, 1; x) = H(1, 1; x) d \log x, \tag{6.17}$$

and so we get

$$\mathcal{S}(H(0, 1, 1; x)) = \mathcal{S}(H(1, 1; x)) \otimes x. \tag{6.18}$$

The symbol of $H(1, 1; x) = \frac{1}{2} \log^2(1-x)$ is easy to obtain from eq. (3.9),

$$\mathcal{S}(H(1, 1; x)) = \frac{1}{2} \mathcal{S}(\log^2(1-x)) = \frac{1}{2} \mathcal{S}(\log(1-x)) \text{III} \mathcal{S}(\log(1-x)) = (1-x) \otimes (1-x). \tag{6.19}$$

Putting everything together, we immediately arrive at the symbol given in eq. (6.14).

Let us now turn to the actual problem we want to study, namely how to express $H(0, 0, 1, 1; x)$ in terms of the spanning set for HPL's defined in the previous section. The goal is to find rational numbers $c_{i_1 \dots i_\ell}^{(k)}$ such that the tensor associated to

$$\begin{aligned} & \sum_{i=1}^{18} c_i^{(1)} \mathcal{B}_4^{(i)}(x) + \sum_{i=1}^8 \sum_{j=1}^4 c_{ij}^{(2)} \mathcal{B}_3^{(i)}(x) \mathcal{B}_1^{(j)}(x) + \sum_{i,j=1}^3 c_{ij}^{(3)} \mathcal{B}_2^{(i)}(x) \mathcal{B}_2^{(j)}(x) \\ & + \sum_{i=1}^3 \sum_{j,k=1}^4 c_{ijk}^{(4)} \mathcal{B}_2^{(i)}(x) \mathcal{B}_1^{(j)}(x) \mathcal{B}_1^{(k)}(x) + \sum_{i,j,k,l=1}^4 c_{ijkl}^{(5)} \mathcal{B}_1^{(i)}(x) \mathcal{B}_1^{(j)}(x) \mathcal{B}_1^{(k)}(x) \mathcal{B}_1^{(l)}(x), \end{aligned} \tag{6.20}$$

equals the tensor given in eq. (6.14). The symbol of eq. (6.20) is easily obtained from the fact that \mathcal{S} is linear and maps products of polylogarithms to shuffles, i.e.,

$$\begin{aligned} & \sum_{i=1}^{18} c_i^{(1)} \mathcal{S}(\mathcal{B}_4^{(i)}(x)) + \sum_{i=1}^8 \sum_{j=1}^4 c_{ij}^{(2)} \mathcal{S}(\mathcal{B}_3^{(i)}(x)) \text{III} \mathcal{S}(\mathcal{B}_1^{(j)}(x)) \\ & + \sum_{i,j=1}^3 c_{ij}^{(3)} \mathcal{S}(\mathcal{B}_2^{(i)}(x)) \text{III} \mathcal{S}(\mathcal{B}_2^{(j)}(x)) \\ & + \sum_{i=1}^3 \sum_{j,k=1}^4 c_{ijk}^{(4)} \mathcal{S}(\mathcal{B}_2^{(i)}(x)) \text{III} \mathcal{S}(\mathcal{B}_1^{(j)}(x)) \text{III} \mathcal{S}(\mathcal{B}_1^{(k)}(x)) \\ & + \sum_{i,j,k,l=1}^4 c_{ijkl}^{(5)} \mathcal{S}(\mathcal{B}_1^{(i)}(x)) \text{III} \mathcal{S}(\mathcal{B}_1^{(j)}(x)) \text{III} \mathcal{S}(\mathcal{B}_1^{(k)}(x)) \text{III} \mathcal{S}(\mathcal{B}_1^{(l)}(x)). \end{aligned} \tag{6.21}$$

The symbol of each function $\mathcal{B}_j^{(i)}$ can be easily obtained, e.g.,

$$\begin{aligned}
 \mathcal{S}(\mathcal{B}_4^{(6)}(x)) &= \mathcal{S}\left(\text{Li}_4\left(\frac{x}{x+1}\right)\right) = -\left(1 - \frac{x}{x+1}\right) \otimes \left(\frac{x}{x+1}\right) \otimes \left(\frac{x}{x+1}\right) \otimes \left(\frac{x}{x+1}\right) \\
 &= -\left(\frac{1}{x+1}\right) \otimes \left(\frac{x}{x+1}\right) \otimes \left(\frac{x}{x+1}\right) \otimes \left(\frac{x}{x+1}\right) \\
 &= (x+1) \otimes x \otimes x \otimes x - (x+1) \otimes x \otimes x \otimes (x+1) - (x+1) \otimes x \otimes (x+1) \otimes x \\
 &\quad + (x+1) \otimes x \otimes (x+1) \otimes (x+1) - (x+1) \otimes (x+1) \otimes x \otimes x \\
 &\quad + (x+1) \otimes (x+1) \otimes x \otimes (x+1) + (x+1) \otimes (x+1) \otimes (x+1) \otimes x \\
 &\quad - (x+1) \otimes (x+1) \otimes (x+1) \otimes (x+1).
 \end{aligned} \tag{6.22}$$

The different shuffles in eq. (6.21) can be distinguished further by acting with the projectors defined in the previous section. In particular, we obtain

$$\Pi_4 \mathcal{S}(H(0, 0, 1, 1; x)) = \sum_{i=1}^{18} c_i^{(1)} \Pi_4 \mathcal{S}(\mathcal{B}_4^{(i)}(x)). \tag{6.23}$$

Equating the coefficients of the different elementary tensors on both sides of this equation, we obtain a linear system that allows us to solve for the coefficients $c_i^{(1)}$, $1 \leq i \leq 18$. The solution is easily obtained, and is given by

$$c_1^{(1)} = -c_3^{(1)} = c_5^{(1)} = 1, \tag{6.24}$$

and all other coefficients are zero. We now proceed recursively, and subtract the (integrable) tensor arising from the solution we have found. By construction, the symbol of this difference must vanish under the action of the projector Π_4 ,

$$\Pi_4 \mathcal{S}(H(0, 0, 1, 1; x) - L(x)) = 0, \tag{6.25}$$

where we defined

$$L(x) = \text{Li}_4(x) - \text{Li}_4(1-x) + \text{Li}_4\left(\frac{x}{x-1}\right). \tag{6.26}$$

We next turn to the determination of the coefficients $c_{ij}^{(2)}$, i.e., coefficients of terms of the form $\mathcal{B}_3^{(i)}(x)\mathcal{B}_1^{(j)}(x)$. We can isolate these terms by applying the projector $\Pi_3 \otimes \Pi_1$,

$$\begin{aligned}
 &(\Pi_3 \otimes \Pi_1) \mathcal{S}(H(0, 0, 1, 1; x) - L(x)) \\
 &= \sum_{i,j=1}^3 c_{ij}^{(2)} (\Pi_3 \otimes \Pi_1) [\mathcal{S}(\mathcal{B}_3^{(i)}(x)) \text{III} \mathcal{S}(\mathcal{B}_1^{(j)}(x))] \\
 &= \sum_{i,j=1}^3 c_{ij}^{(2)} [\Pi_3 \mathcal{S}(\mathcal{B}_3^{(i)}(x))] \otimes [\Pi_1 \mathcal{S}(\mathcal{B}_1^{(j)}(x))],
 \end{aligned} \tag{6.27}$$

and equating the coefficients of the elementary tensors on both sides of the equation we can solve for the coefficients $c_{ij}^{(2)}$. We find $c_{12}^{(2)} = -1$, and $c_{ij}^{(2)} = 0$ if $(i, j) \neq (1, 2)$. We

again subtract this contribution to find an expression that vanishes under the actions of both Π_4 and $\Pi_3 \otimes \Pi_1$. Next we act with the projector $\Pi_2 \otimes \Pi_2$ on this difference

$$(\Pi_2 \otimes \Pi_2)\mathcal{S}(H(0, 0, 1, 1; x) - L(x) + \text{Li}_3(x) \log(1 - x)) = 0, \quad (6.28)$$

and we immediately conclude that $c_{ij}^{(3)} = 0, \forall 1 \leq i, j \leq 3$. Similarly, by acting with the projector $\Pi_2 \otimes \Pi_1 \otimes \Pi_1$ we conclude that all the coefficients $c_{ijk}^{(4)}$ must vanish. The remaining terms are thus all associated to products of logarithms, which can immediately be read off from the tensor

$$\begin{aligned} & \mathcal{S}(H(0, 0, 1, 1; x) - L(x) + \text{Li}_3(x) \log(1 - x)) \\ &= \frac{1}{24} (1 - x) \odot (1 - x) \odot (1 - x) \odot (1 - x) - \frac{1}{6} x \odot (1 - x) \odot (1 - x) \odot (1 - x) \\ &= \mathcal{S} \left(\frac{1}{24} \log^4(1 - x) - \frac{1}{6} \log x \log^3(1 - x) \right). \end{aligned} \quad (6.29)$$

At this stage, we have found a combination of (a product of) functions in our spanning set that has the same symbol as $H(0, 0, 1, 1; x)$, and so the quantities are equal up to terms that are mapped to zero by \mathcal{S} . We make an ansatz assuming we have found sufficiently many elements in the kernel of \mathcal{S} , as

$$\begin{aligned} & a_1 \left(\text{Li}_4 \left(\frac{1}{2} \right) + \frac{1}{24} \log^4 2 \right) + a_2 \pi^4 + \zeta_3 \sum_{i=1}^4 b_i \mathcal{B}_1^{(i)}(x) \\ & + \pi^2 \left(\sum_{i=1}^3 c_i \mathcal{B}_2^{(i)}(x) + \sum_{i,j=1}^4 c_{ij} \mathcal{B}_1^{(i)}(x) \mathcal{B}_1^{(j)}(x) \right), \end{aligned} \quad (6.30)$$

where a_i, b_i, c_i and c_{ij} are rational numbers, and $\zeta_3 = \zeta(3)$ denotes the value in $s = 3$ of the Riemann ζ function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \quad (6.31)$$

The coefficients can now be fixed by looking at particular values of x . In particular, harmonic polylogarithms are known analytically up to weight four for $x = 0, x = \pm 1$ and $x = \pm \frac{1}{2}$. It turns out that in all cases these values are enough to fix all the free coefficients. In the case of $H(0, 0, 1, 1; x)$, we finally arrive at

$$\begin{aligned} H(0, 0, 1, 1; x) &= S_{2,2}(x) \\ &= -\text{Li}_4(1 - x) + \text{Li}_4(x) + \text{Li}_4 \left(\frac{x}{x - 1} \right) \\ &\quad - \text{Li}_3(x) \log(1 - x) + \frac{1}{24} \log^4(1 - x) - \frac{1}{6} \log x \log^3(1 - x) \\ &\quad + \zeta_3 \log(1 - x) + \frac{1}{12} \pi^2 \log^2(1 - x) + \frac{\pi^4}{90}. \end{aligned} \quad (6.32)$$

7 Conclusion

In this paper, we have provided a review of the symbol map, a linear map that associates to a multiple polylogarithm of weight n an n -fold tensor in a way that captures many of the combinatorial properties of polylogarithms, and also respects the functional equations they satisfy. While so far the symbol map was defined recursively via iterated differentials, we have given a diagrammatic rule where the symbol of a multiple polylogarithm is obtained directly via a weighted sum over the maximal dissections of the decorated polygon associated to the polylogarithm introduced in ref. [21].

Furthermore, we have addressed the problem of integrating a symbol, i.e., the problem of finding a function whose symbol matches a given tensor that satisfies the integrability condition (3.17). We have presented a systematic approach of how to find a candidate spanning set for such a function. Once this candidate spanning set has been constructed, and working under the assumptions that its elements suffice to express the integrated symbol in, we showed how a set of projectors can be defined which help to find a function whose symbol matches the initial tensor. While our approach falls short of a complete algorithmic proof and is surely not adequate in all possible scenarios, we nevertheless believe that it can be applied in many situations, as was for example demonstrated in ref. [74–76] where our method was successfully applied to obtain new compact analytic results for certain one-loop hexagon integrals in $D = 6$ dimensions. Finally, we have used our approach to derive a spanning set for harmonic polylogarithms up to weight 4, and this spanning set was recently used to obtain an efficient numerical implementation of harmonic polylogarithms up to weight four [77].

Acknowledgments

CD is grateful to B. Anastasiou, S. Buehler, P. Heslop, M. Spradlin, C. Vergu and A. Volovich for useful discussions. HG expresses his thanks to F. Brown and D. Kreimer for helpful comments. The authors thank E.W.N. Glover for helpful remarks on a preliminary version of the text. JR would like to thank EPSRC for their support. This work was supported by the Research Executive Agency (REA) of the European Union under the Grant Agreement number PITN-GA-2010-264564 (LHCPhenoNet).

A Review on shuffle algebras

As shuffle algebras are a recurrent theme when working with multiple polylogarithms and their symbols, we review in this appendix the most important notions. Before turning to the special case of shuffle algebras, we first review some basic notions about algebras in general.

Algebras over a field. An algebra \mathcal{A} over a field F ($F = \mathbb{R}$ or $F = \mathbb{C}$, say) is a vector space⁸ over F together with an associative and distributive multiplication $\mathcal{A} \otimes \mathcal{A} \rightarrow \mathcal{A}$. In

⁸More generally, we could consider \mathcal{A} to be a module over a ring R .

the case the multiplication has a unit element, the algebra is called unital. Furthermore, an algebra is said to be *graded* if \mathcal{A} can be written as a direct sum as a vector space,

$$\mathcal{A} = \bigoplus_{n \in I} \mathcal{A}_n, \tag{A.1}$$

and if $\forall a \in \mathcal{A}_m$ and $\forall b \in \mathcal{A}_n$, we have

$$a \cdot b \in \mathcal{A}_{m+n}. \tag{A.2}$$

One of the most prominent representatives of a graded algebra is the tensor algebra associated to an F -vector space V , defined by

$$\mathcal{T}(V) = \bigoplus_{n=0}^{\infty} \mathcal{T}_n(V), \tag{A.3}$$

where $\mathcal{T}_0(V) = F$ and $\mathcal{T}_1(V) = V$, and for $n \geq 2$ we define

$$\mathcal{T}_n(V) = \underbrace{V \otimes \dots \otimes V}_{n \text{ times}}. \tag{A.4}$$

The multiplication on $\mathcal{T}(V)$ is defined on elementary tensors $a_1 \otimes \dots \otimes a_n$ by

$$\begin{aligned} \mathcal{T}_m(V) \otimes \mathcal{T}_n(V) &\rightarrow \mathcal{T}_{m+n}(V) \\ (a_1 \otimes \dots \otimes a_m) \otimes (b_1 \otimes \dots \otimes b_n) &\mapsto a_1 \otimes \dots \otimes a_m \otimes b_1 \otimes \dots \otimes b_n, \end{aligned} \tag{A.5}$$

making the tensor algebra into a *graded* algebra in an obvious way. Furthermore, the tensor algebra is also unital, the unit being the unit $1 \in F$.

A *homomorphism* between two algebras \mathcal{A} and \mathcal{B} is a linear map ϕ that preserves the multiplication, i.e., a linear map ϕ such that $\forall a, b \in \mathcal{A}$, $\phi(a \cdot b) = \phi(a) \cdot \phi(b)$.

Ideals in algebras. An *ideal* in an algebra \mathcal{A} (or more generally in a ring) is an additive subgroup \mathcal{I} of \mathcal{A} such that

$$\forall a \in \mathcal{A}, \forall b \in \mathcal{I}, \quad a \cdot b \in \mathcal{I} \quad \text{and} \quad b \cdot a \in \mathcal{I}. \tag{A.6}$$

An easy example of an ideal is given by considering the ring of integer numbers \mathbb{Z} and its subset $n\mathbb{Z}$, $n \in \mathbb{Z}$, i.e., the set of all integer multiples of n . The set $n\mathbb{Z}$ is obviously an additive subgroup of \mathbb{Z} , and every time we multiply an element of $n\mathbb{Z}$ by an integer number, we obtain another multiple of n . Hence $n\mathbb{Z}$ is an ideal of \mathbb{Z} . Another example of an ideal is the kernel of an algebra homomorphism ϕ . Indeed, $\text{Ker } \phi$ is a sub-vector space of \mathcal{A} , and hence an additive subgroup. Furthermore, $\forall a \in \mathcal{A}$ and $\forall b \in \text{Ker } \phi$, we have

$$\phi(a \cdot b) = \phi(a) \cdot \phi(b) = \phi(a) \cdot 0 = 0, \tag{A.7}$$

and so $a \cdot b \in \text{Ker } \phi$, making $\text{Ker } \phi$ into an ideal in \mathcal{A} .

Shuffle algebras. After this rather general discussion on algebras, let us from now on focus exclusively on the example of shuffle algebras. As a starting point, let us consider a set \mathcal{L} , whose elements we will refer to as *letters*, and consider the set \mathcal{W} of all words constructed from elements in \mathcal{L} , i.e., the set of all possible concatenations of letters in \mathcal{L} , together with the *empty word* ε , consisting of no letters (more precisely, \mathcal{W} is the free monoid generated by the elements in \mathcal{L}). We can define a multiplication on \mathcal{W} given by concatenation of words, the empty word being the unit element.

Let us now consider the vector space \mathcal{A} over some field F given by all formal linear combinations of words in \mathcal{W} . \mathcal{A} is then in fact an algebra over F , the multiplication given simply by the concatenation of words. Furthermore, it is easy to see that \mathcal{A} is also graded by the length of the words (the concatenation of two words with length m and n gives a word of length $m + n$).

We can define another multiplication on \mathcal{A} , the so-called *shuffle product*, defined on words by

$$(a_1 \dots a_m) \mathbb{III} (a_{m+1} \dots a_{m+n}) = \sum_{\sigma \in \Sigma(n_1, n_2)} a_{\sigma^{-1}(1)} \dots a_{\sigma^{-1}(n_1+n_2)}, \quad (\text{A.8})$$

where $\Sigma(n_1, n_2)$ denotes the set of all shuffles of $n_1 + n_2$ elements, i.e., the subset of the symmetric group $S_{n_1+n_2}$ defined by

$$\Sigma(n_1, n_2) = \{ \sigma \in S_{n_1+n_2} \mid \sigma^{-1}(1) < \dots < \sigma^{-1}(n_1) \text{ and } \sigma^{-1}(n_1+1) < \dots < \sigma^{-1}(n_1+n_2) \}. \quad (\text{A.9})$$

The vector space \mathcal{A} together with the shuffle product is called a *shuffle algebra*. A shuffle algebra is again graded by the length of the words, and the empty word ε is the unit element of the shuffle algebra. The definition (A.8) of the shuffle product is equivalent to the following recursive definition, $\forall x, y \in \mathcal{L}, \forall u, v \in \mathcal{W}$,

$$\begin{aligned} \varepsilon \mathbb{III} u &= u \mathbb{III} \varepsilon = u, \\ (xu) \mathbb{III} (yv) &= x(u \mathbb{III} (yv)) + y((xu) \mathbb{III} v). \end{aligned} \quad (\text{A.10})$$

Note that the tensor algebra of a vector space V can be equipped with a shuffle product in a natural way: the set of letters \mathcal{L} is simply a basis of V , the set of words corresponds to the elementary tensors $a_1 \otimes \dots \otimes a_n$, the empty word is simply the scalar 1, and the concatenation of words corresponds to the multiplication (A.5) of two tensors.

In section 2 we have encountered another example of a shuffle algebra, the shuffle algebra of multiple polylogarithms, eq. (2.3). In that case letters are the elements a_i of the vector of singularities (a_1, \dots, a_n) , the latter being the words. Concatenation is simply defined by the concatenation of the vectors of singularities. The length of a word corresponds to the weight of the polylogarithm, i.e., the number of components of the vector of singularities. In other words, the shuffle algebra of polylogarithms is graded by the weight.

B Selected examples of symbols

In this appendix we compile a list of the symbols of the most commonly used (multiple) polylogarithms. First of all, the symbol of an ordinary logarithm is simply the argument

of the logarithm,

$$\mathcal{S}(\log x) = x. \tag{B.1}$$

From eq. (3.9) it follows then, for every non-negative integer n ,

$$\mathcal{S}\left(\frac{1}{n!} \log^n x\right) = \underbrace{x \otimes \dots \otimes x}_{n \text{ times}} \equiv x^{\otimes n}. \tag{B.2}$$

The symbol of the classical polylogarithms have a similarly simple form, i.e., for every non-negative integer n we obtain

$$\mathcal{S}(\text{Li}_n(x)) = -(1-x) \otimes x^{\otimes(n-1)} = -(1-x) \otimes \underbrace{x \otimes \dots \otimes x}_{(n-1) \text{ times}}, \tag{B.3}$$

where we used the notation of eq. (B.2). Note that for $n = 1$, the classical polylogarithm can be expressed as an ordinary logarithm, $\text{Li}_1(x) = -\log(1-x)$, which is consistent with the symbols given in eqs. (B.1) and (B.3),

$$\mathcal{S}(\text{Li}_1(x)) = -(1-x) = \mathcal{S}(-\log(1-x)). \tag{B.4}$$

Finally, the symbol of a Nielson polylogarithm reads,

$$\mathcal{S}(S_{n,p}(x)) = (-1)^p (1-x)^{\otimes p} \otimes x^{\otimes n} = (-1)^p \underbrace{(1-x) \otimes \dots \otimes (1-x)}_{p \text{ times}} \otimes \underbrace{x \otimes \dots \otimes x}_{n \text{ times}}. \tag{B.5}$$

Again we note that the Nielson polylogarithms contain the classical polylogarithms as a special case, $S_{n-1,1}(x) = \text{Li}_n(x)$, an identity which is easily verified at the level of the symbols (B.3) and (B.5).

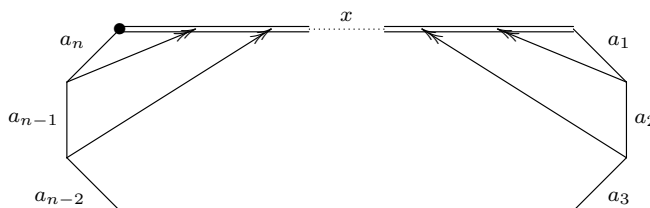
The previous examples of the classical and Nielson polylogarithms are both just special cases of harmonic polylogarithms where the components of the vector of singularities take values in $\{0, 1\}$,

$$\text{Li}_n(x) = H(\vec{0}_{n-1}, 1; x) \quad \text{and} \quad S_{n,p}(x) = H(\vec{0}_n, \vec{1}_p; x). \tag{B.6}$$

The symbol of a harmonic polylogarithm $H(a_1, \dots, a_n; x)$, with $a_i \in \{0, 1\}$, can be written in the compact form

$$\mathcal{S}(H(a_1, \dots, a_n; x)) = (-1)^k (a_n - x) \otimes \dots \otimes (a_1 - x), \tag{B.7}$$

where k is the number of components in the vector of singularities (a_1, \dots, a_n) equal to 1. Indeed, the polygon $P(a_n, \dots, a_1, x)$ associated to $G(a_1, \dots, a_n; x) = (-1)^k H(a_1, \dots, a_n; x)$ has the root side decorated by x , and all other sides decorated by 0 or 1. It is easy to see that the only relevant maximal dissection of such a polygon is



All other maximal dissection give rise to bigons decorated only by 0 and / or 1, which give a zero contribution to the symbol. The non-vanishing dissection produces a term in the symbol given by

$$\mu\left(\begin{array}{c} x \\ \curvearrowright \\ a_n \end{array}\right) \otimes \mu\left(\begin{array}{c} x \\ \curvearrowright \\ a_{n-1} \end{array}\right) \otimes \dots \otimes \mu\left(\begin{array}{c} x \\ \curvearrowright \\ a_2 \end{array}\right) \otimes \mu\left(\begin{array}{c} x \\ \curvearrowright \\ a_1 \end{array}\right), \quad (\text{B.8})$$

which is equal to the tensor in the right-hand side of eq. (B.7) (apart from the sign).

Generic harmonic polylogarithms where the components of the vector of singularities take values in $\{-1, 0, 1\}$ do not admit a compact expression for the symbol. In this case, the symbol is however easily obtained from the symbols of generic multiple polylogarithms, which are reviewed up to weight four in the next subsections.

B.1 The symbol of a generic multiple polylogarithm of weight one

A generic multiple polylogarithm of weight one can be written as $G(a; x)$, with $a, x \in \mathbb{C}$. We can associate a bigon to this function via

$$G(a; x) \leftrightarrow P(a, x) = \begin{array}{c} x \\ \curvearrowright \\ a \end{array}. \quad (\text{B.9})$$

The symbol of $G(a; x)$ is then

$$\mathcal{S}(G(a; x)) = \mu\left(\begin{array}{c} x \\ \curvearrowright \\ a \end{array}\right). \quad (\text{B.10})$$

B.2 The symbol of a generic multiple polylogarithm of weight two




To a generic multiple polylogarithm $G(a, b; x)$ of weight two we associate a trigon

$$G(a, b; x) \leftrightarrow P(b, a, x) = \begin{array}{c} x \\ \triangle \\ b \quad a \end{array}. \quad (\text{B.11})$$

The symbol of $G(a, b; x)$ is then obtained by looking at all the maximal dissections of the trigon, obtained by inserting a single arrow. In the following we give the three maximal dissections, together with the term in the symbol they correspond to. We use the shorthand

$$ab|cd \equiv \mu\left(\begin{array}{c} b \\ \curvearrowright \\ a \end{array}\right) \otimes \mu\left(\begin{array}{c} d \\ \curvearrowright \\ c \end{array}\right). \quad (\text{B.12})$$

The three maximal dissections of the trigon in eq. (B.11), together with the term in the symbol $\mathcal{S}(G(a, b; x))$ they correspond to, are

		
$+ax ba$	$+bx ax$	$-bx ab$

B.3 The symbol of a generic multiple polylogarithm of weight three

To a generic multiple polylogarithm $G(a, b, c; x)$ of weight three we associate a tetragon

$$G(a, b, c; x) \leftrightarrow P(c, b, a, x) = \begin{array}{c} \bullet \\ \text{---} x \text{---} \\ | \\ c \quad | \quad a \\ | \\ \text{---} b \text{---} \end{array} . \quad (\text{B.13})$$

The symbol of $G(a, b, c; x)$ is then obtained by looking at all the maximal dissections of the tetragon, obtained by inserting two non-intersecting arrows. In the following we give the twelve maximal dissections, together with the term in the symbol they correspond to. We use the shorthand

$$ab|cd|ef \equiv \mu\left(\begin{array}{c} b \\ \text{---} \\ \bullet \\ \text{---} \\ a \end{array}\right) \otimes \mu\left(\begin{array}{c} d \\ \text{---} \\ \bullet \\ \text{---} \\ c \end{array}\right) \otimes \mu\left(\begin{array}{c} f \\ \text{---} \\ \bullet \\ \text{---} \\ e \end{array}\right), \quad (\text{B.14})$$

as well as the notation for shuffles,

$$A|B \text{ III } C \equiv A|B|C + A|C|B. \quad (\text{B.15})$$

The twelve maximal dissections of the tetragon in eq. (B.13), together with the term in the symbol $\mathcal{S}(G(a, b, c; x))$ they correspond to, are

$+cx bx ax$	$+ax ca ba$	$-bx ab \text{ III } cb$	$+cx ac bc$
$+cx ax ba$	$+ax ba cb$	$+cx bc ab$	$-cx ax \text{ III } bc$
$-cx bx ab$	$-ax ca bc$	$bx ax \text{ III } cb$	$-cx ac ba$

B.4 The symbol of a generic multiple polylogarithm of weight four

To a generic multiple polylogarithm $G(a, b, c, d; x)$ of weight two we associate a pentagon

$$G(a, b, c, d; x) \leftrightarrow P(d, c, b, a, x) = \begin{array}{c} \bullet \\ \text{---} x \text{---} \\ | \\ d \quad | \quad a \\ | \quad \quad | \\ \text{---} c \quad \quad | \quad \quad b \end{array} . \quad (\text{B.16})$$

The symbol of $G(a, b, c, d; x)$ is then obtained by looking at all the maximal dissections of the pentagon, obtained by inserting three non-intersecting arrows. In the following we give

the 55 maximal dissections, together with the term in the symbol they correspond to. We use the shorthand

$$ab|cd|ef|gh \equiv \mu\left(\begin{array}{c} b \\ \longleftarrow \\ a \end{array}\right) \otimes \mu\left(\begin{array}{c} d \\ \longleftarrow \\ c \end{array}\right) \otimes \mu\left(\begin{array}{c} f \\ \longleftarrow \\ e \end{array}\right) \otimes \mu\left(\begin{array}{c} h \\ \longleftarrow \\ g \end{array}\right), \quad (\text{B.17})$$

as well as the notation for shuffles

$$\begin{aligned} A|B|C \text{ III } D &\equiv A|B|C|D + A|B|D|C, \\ A|B \text{ III } (C|D) &\equiv A|B|C|D + A|C|B|D + A|C|D|B, \\ A|B \text{ III } C \text{ III } D &\equiv \sum_{\sigma \in S_3} A|\sigma(B)|\sigma(C)|\sigma(D), \end{aligned} \quad (\text{B.18})$$

where in the last equation the sum runs over all permutations of the set $\{B, C, D\}$. The 55 maximal dissections of the pentagon in eq. (B.16), together with the term in the symbol $\mathcal{S}(G(a, b, c, d; x))$ they correspond to, are

$+dx cx bx ax$	$+ax da ca ba$	$-bx ab \text{ III } (db cb)$	$+cx dc \text{ III } (ac bc)$	$-dx ad bd cd$
$+cx dc \text{ III } (bx ax)$	$-dx ad ca ba$	$+bx ax \text{ III } (db cb)$	$-cx dc \text{ III } (ac ba)$	$+dx ad bd cb$
$+ax da bd cd$	$-dx cx bx ab$	$-ax da ca bc$	$+bx ab \text{ III } (db cd)$	$+dx cx ac bc$
$-cx dc \text{ III } (bx ab)$	$+dx ad ca bc$	$-bx ax \text{ III } (db cd)$	$-dx cx ac ba$	$-ax da bd cb$

$+bx ax \text{ III } (cb dc)$	$+dx cd ac ba$	$-dx ax \text{ III } (bd cb)$	$+cx dc \text{ III } (ax ba)$	$-dx ad ba cb$
$+ax ba db cb$	$+cx dc \text{ III } (bc ab)$	$-dx ad cd bc$	$-dx cd \text{ III } (bx ax)$	$+dx ax ca ba$
$-ax ba db cd$	$+dx cx bc ab$	$+ax da cd bc$	$+dx cd \text{ III } (bx ab)$	$-dx ax ca bc$
$-bx ab \text{ III } (cb dc)$	$-dx cd ac bc$	$+dx ax \text{ III } (bd cd)$	$+dx cx ax ba$	$+ax da ba cb$
$+ax ca dc \text{ III } ba$	$+dx bd ab \text{ III } cb$	$-cx dc \text{ III } bc \text{ III } ax$	$+dx ad cd \text{ III } ba$	$+dx bx ax \text{ III } cb$
$-ax ca dc \text{ III } bc$	$-dx bd cd \text{ III } ab$	$-dx cx ax \text{ III } bc$	$-ax da cd \text{ III } ba$	$-dx bx cb \text{ III } ab$
$+ax ba cb dc$	$-dx cd bc ab$	$+dx ax \text{ III } (cd bc)$	$-dx cd \text{ III } (ax ba)$	$+dx ax ba cb$

C Proof of Proposition 4

In this section we present the proof of Proposition 4. The proof uses the combinatorics of the decorated polygons introduced in section 3. In order to be able to map the symbol of colored multiple zeta values (CMZV's) to polygons, we first have to relate the CMZV's to multiple polylogarithms. From the series representations (2.9) and (5.41), it is easy to see that one has the relation (provided that the CMZV's converge)

$$\zeta(m_1, \dots, m_k; s_1, \dots, s_k) = (-1)^w G_{m_1, \dots, m_k}(\hat{s}_1, \dots, \hat{s}_k), \tag{C.1}$$

where we defined $w = m_1 + \dots + m_k$ and

$$\hat{s}_j = \prod_{i=1}^j s_i. \tag{C.2}$$

Hence, using the correspondence between multiple polylogarithms and decorated polygons, we can associate to the CMZV $\zeta(m_1, \dots, m_k; s_1, \dots, s_k)$ the polygon⁹

$$P(\underbrace{0, \dots, 0}_{m_k-1 \text{ times}}, \hat{s}_k, \dots, \underbrace{0, \dots, 0}_{m_1-1 \text{ times}}, \hat{s}_1, 1). \tag{C.3}$$

We start by introducing concepts needed to prove, and then prove Proposition 4. We will in fact proof results at the level of the polygons that are slightly more general than the results given in Proposition 4 but not imposing the restriction $(m_1, s_1) \neq (1, 1)$ (which corresponds to divergent CMZV's). The first proposition we give is a generalization of statement 2 in the proposition. Statement 1 in Proposition 4 is equivalent to Proposition C.6.

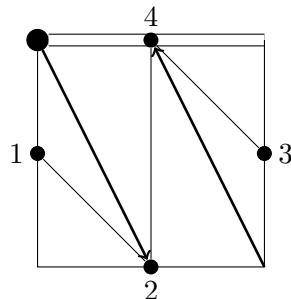
Proposition C.1. *The symbol of $P(\varepsilon_1, \dots, \varepsilon_n, 1)$ for some $\varepsilon_i = \pm 1$ is equal to $\lambda_{a,n}(2^{\otimes n})$ for*

$$\lambda_{a,n} = (-1)^a \binom{n-1}{a} \quad \text{and} \quad a = n - \max\{i \mid \varepsilon_i = -1\}. \tag{C.4}$$

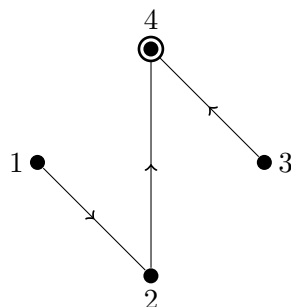
We start by noting that from the Hölder convolution (2.8) with $p = \infty$ it follows that $P(x_1, \dots, x_m, 1)$ has the same symbol as the polygon $P(1 - x_m, \dots, 1 - x_1, 1)$. So, without loss of generality, we consider the polygon $P(\underbrace{0, \dots, 0}_{t_0}, 2, \underbrace{0, \dots, 0}_{t_1}, 2, \dots, 2, \underbrace{0, \dots, 0}_{t_m}, 1)$ and will find its symbol, remembering to take into account the factor of $(-1)^n$. The move from sides labelled 1 and -1 to sides labelled 0 and 2 increases the number of dissections that have coefficient 0. The combinatorics of the dissections of polygons of this type is best captured by a certain type of planar trees, the so-called *Hook-arrow trees*, which is a change of view on the maximal dissections, and hence the symbol, of a polygon.

⁹There is of course the factor $(-1)^w$ of eq. (C.1) to be kept in mind.

Hook-arrow trees. Every full dissection of a polygon uniquely defines a certain spanning tree τ on the vertices which are the midpoints of the polygon sides, and vice-versa. These vertices, v_1, \dots, v_n , inherit the label of the side they sit on and they form the vertices of τ . We induce the edges of τ as all possible lines, between the v_i , that do not cross arrows from the dissection. Here is an example of a full dissection of a 4-gon with the spanning tree induced:



We also induce a root on τ as the vertex lying on the final side of the polygon. The edges are then oriented towards the root. For the above example of a dissected 4-gon the rooted spanning tree is:



The edges of the tree will not cross by construction; we define interlacing to reflect this for use in the definition of a hook-arrow tree.

Definition C.2. A tree with a linear order on its vertices w_j is said to be **interlaced** if there exists a choice of four vertices $w_1 < \dots < w_4$ such that both edges $\{w_1, w_3\}$ and $\{w_2, w_4\}$ are contained in the tree.

We now give a formal definition of a hook-arrow tree and illustrate the definition with an example.

Definition C.3. A **hook-arrow tree** is a rooted spanning tree on a set of vertices in a linear order, $v_1 < \dots < v_n$, which is not interlaced and has root v_n .

We can think of a hook-arrow tree as being a tree embedded in the plane on the vertices arranged in a circle.

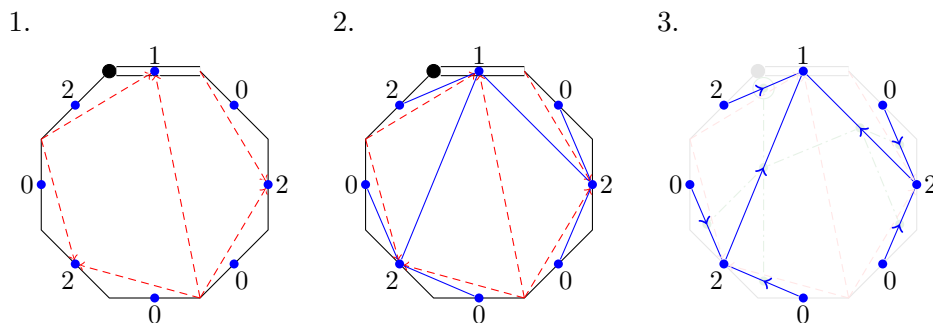
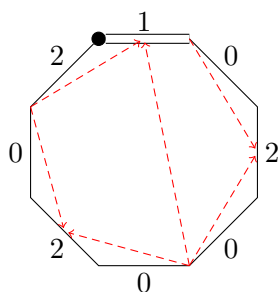


Figure 8. Construction of a hook arrow tree corresponding to a maximal dissection. Details are given in the text.

Example. For the polygon $P(2,0,2,0,0,2,0,1)$ (attached to the multiple polylogarithm $G(0, 2, 0, 0, 2, 0, 2; 1)$) we have the following possible maximal dissection



We note that, for clarity, in this example we give the arrows in the dissection of a polygon dashed shafts.

We now construct the hook-arrow tree for this dissection, following the method outlined in Definition C.3 (see also figure 8):

1. add vertices to the middle of each side of the polygon;
2. join all vertices that can be connected without crossing the shaft of an arrow;
3. remove the polygon and arrows and direct tree towards distinguished vertex representing final side of polygon.

We now reintroduce the dual tree view of a dissection from section 3.1 of ref. [21] as this is also beneficial in finding the symbol attached to a polygon. As with the dissection of a polygon using arrows, the dual tree can easily be seen in the hook-arrow tree view. In figure 9 we illustrate the relationship between maximal dissections, dual trees and hook arrow trees. For clarity we give the dual tree a dash-dotted line.

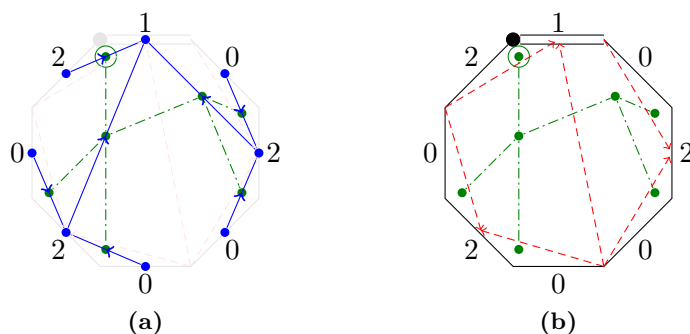
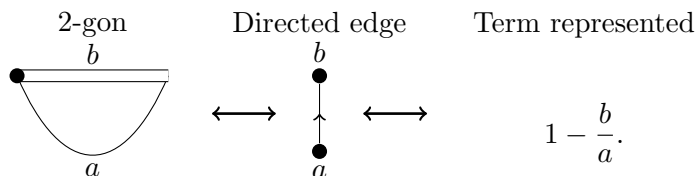


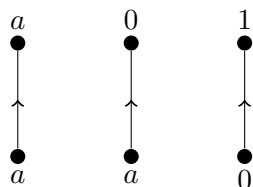
Figure 9. Hook-arrow tree and dual tree (a). Polygon dissection and dual tree (b).

We now elaborate on the interpretation of a hook-arrow tree.

The association between a 2-gon in a polygon dissection, an edge of a hook-arrow tree and a term of a tensor product in the symbol is



In the dissection of a polygon, we set the coefficient of certain terms in the symbol to 0 if the sides of 2-gons have certain combinations of labels, e.g., $\mu\left(\begin{matrix} a \\ 0 \end{matrix}\right) = 0$. We do the same in a hook-arrow tree when we have the corresponding edges



for any a .

To obtain the term in the symbol, the edges of the hook-arrow tree are chosen in the same order as the corresponding 2-gons in the dissection of a polygon. This can also be seen by viewing the dual tree (this correspondence is shown in the above example). Note that an explicit algorithm for this purely from the view of the hook-arrow tree can be written.

The sign of a maximal dissection is determined, in a similar way to the polygons by first determining the number, α , of ‘backward’ edges in the hook-arrow tree. These correspond to edges that, respecting their direction, go from one vertex in the order of the vertices defined by the polygon, to a previous vertex. The sign is then $(-1)^\alpha$.

Remark C.4. *As previously mentioned, it is important to note that for every full dissection of a polygon there is a unique corresponding hook-arrow tree and that the method*

of extracting the symbol from hook-arrow trees is simply a different view of the procedure for polygons. We include an overview of the construction of an hook-arrow tree because constructing all possible dissections for the functions required in Proposition C.1 is a lot easier to view from this perspective. The actual proof then takes place in the dual tree view.

Finally, before the proof of Proposition C.1 we require the following proposition for which we give a sketch proof involving generating functions.

Proposition C.5. *If $c, n \in \mathbb{Z}^{\geq 0}$ then*

$$\sum_{i=0}^n (-1)^i \binom{n-i+c}{n-i} \binom{n+c+1}{i} = (-1)^n$$

Proof. (Sketch)

Let $r = n-i$ and view the right hand side of the identity as coefficients of the generating function

$$\Phi(x) = \sum_{n=0}^{\infty} x^n \sum_{r=0}^n (-1)^{n-r} \binom{r+c}{r} \binom{n+c+1}{n-r}.$$

Firstly since for $r > n$ we have $\binom{n+c+1}{n-r} = 0$, we can change the summation of r to run over all positive integers. Then by re-ordering the summation signs for small x , and applying some basic properties of binomial coefficients we can arrive at

$$\Phi(x) = \sum_{r=0}^{\infty} \frac{(-1)^{r-c-1}}{x^{c+1}} \binom{r+c}{r} \sum_{n=0}^{\infty} \binom{n+c+1}{r+c+1} (-x)^{n+c+1}.$$

Then by using

$$\sum_m \binom{m}{k} x^m = \frac{x^k}{(1-x)^{k+1}} \quad \text{and then} \quad \sum_m \binom{m+k}{k} x^m = \frac{1}{(1-x)^{k+1}}$$

we can show that

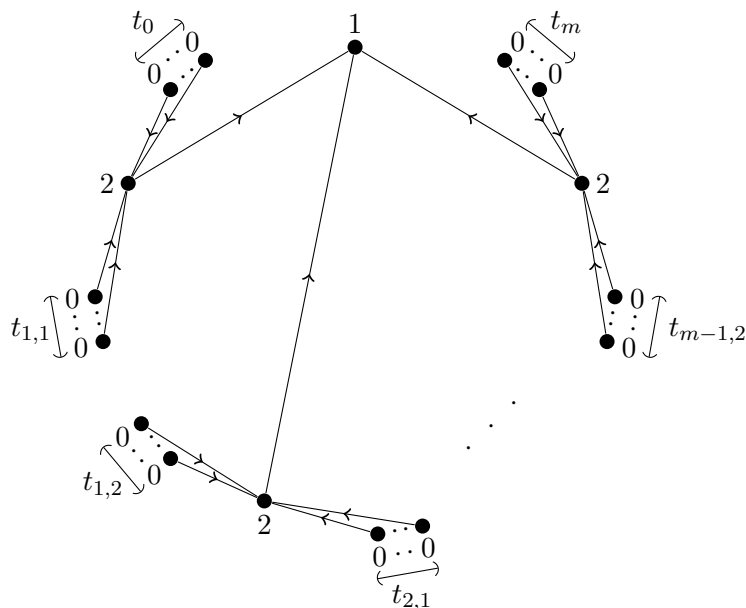
$$\Phi(x) = \frac{1}{1+x} = \sum_{n=0}^{\infty} (-x)^n$$

which proves the result. □

Proof. (of Proposition C.1) After applying the Hölder convolution with $p = \infty$, and without loss of generality, we attempt to find all hook-arrow trees relating to the polygon

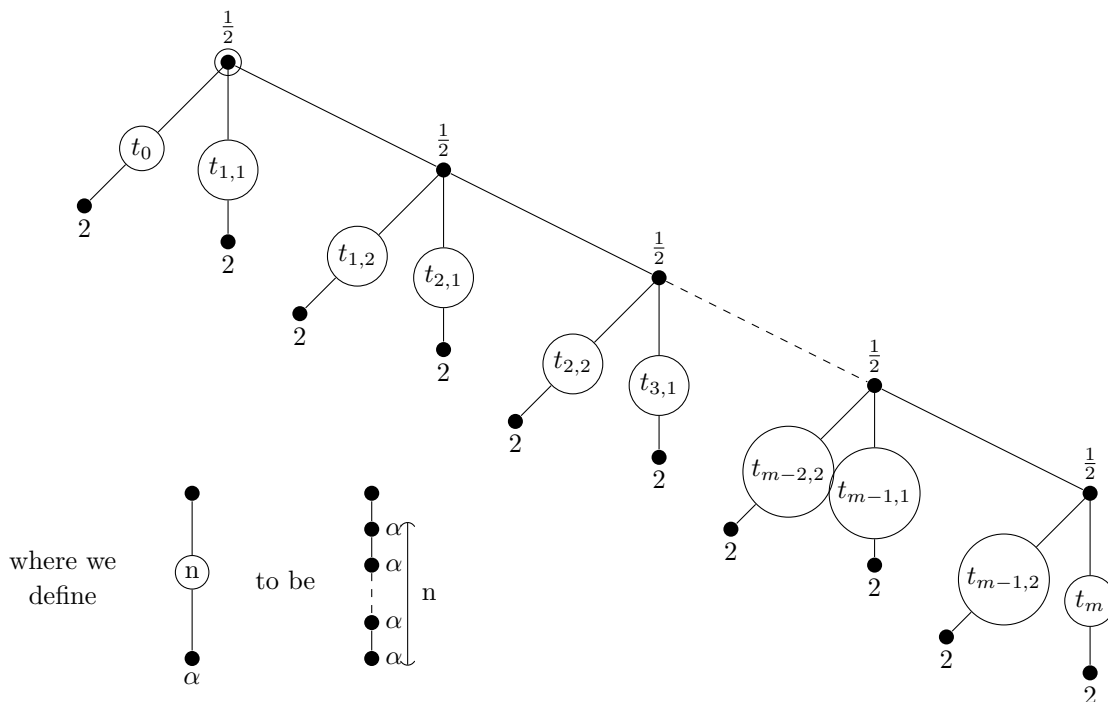
$$P(\underbrace{0, \dots, 0}_{t_0}, 2, \underbrace{0, \dots, 0}_{t_1}, 2, 0, \dots, 0, 2, \underbrace{0, \dots, 0}_{t_m}, 1)$$

which do not represent terms with coefficient 0 in the symbol. After some consideration we see that these hook-arrow trees must take the following form.

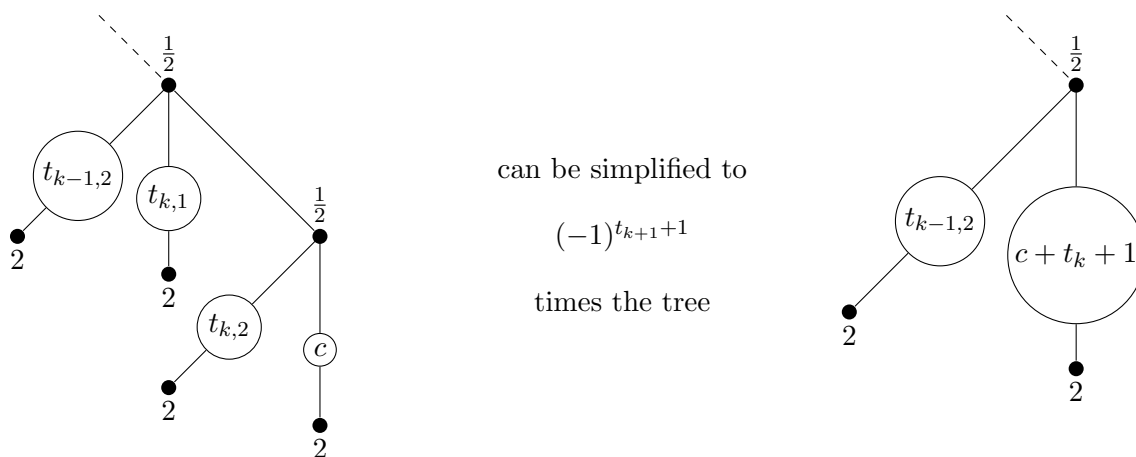


Each $t_{i,1}$ and $t_{i,2}$, for $i = 1, \dots, m - 1$ are chosen integers $0 \leq t_{i,1}, t_{i,2} \leq t_i$ such that $t_{i,1} + t_{i,2} = t_i$. The choice of the $t_{i,j}$ arises from the fact that we can choose where to partition each group of t_i vertices labelled 0, for $i = 1, \dots, m - 1$, and attach them to the vertices labelled 2, remembering that the vertices must not cross. In the case of the function $P(2, 0, 2, 0, 0, 2, 0, 1)$ from the example above, where $m = 3, t_0 = 0, t_1 = 1, t_2 = 2$ and $t_3 = 1$, we would have 6 possible valid dissections, arising from two choices of the $t_{i,j}$ in $t_{1,1} + t_{1,2} = 1$ and three choices from $t_{2,1} + t_{2,2} = 2$. We note that the example above explored the particular dissection where $t_{1,1} = 0, t_{1,2} = 1, t_{2,1} = 1$ and $t_{2,2} = 1$.

We will now show how it is possible to simplify this tree by effectively removing the edges joining vertices labelled 2 and 1 and replacing them with edges connecting vertices labelled 0 and 2. For this we return to the dual tree notation described above, and in section 3.1. The dual tree of the above hook-arrow tree above is



We claim that



We will now write the tensor product of the symbol of the left hand dual tree part in the above claim. Let us recall our convention (3.11) that shuffles takes precedence over a tensor

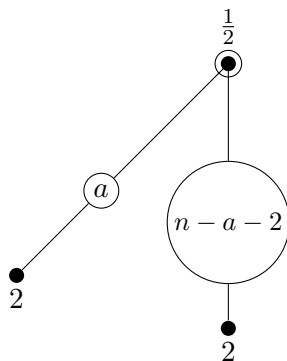
$$a^{\otimes b} = \underbrace{a \otimes \dots \otimes a}_b \quad \text{and that} \quad a^{\otimes b} \text{III} c = \underbrace{(a \otimes \dots \otimes a)}_b \text{III} c.$$

The left hand dual tree part in the claim will have the symbol

$$\begin{aligned}
 & \sum_{t_{k,1}=0}^{t_k} (-1)^{t_{k,1}} \left(\frac{1}{2} \otimes 2^{\otimes t_{k-1,2}} \text{III} 2^{\otimes t_{k,1}} \text{III} \left(\frac{1}{2} \otimes 2^{\otimes t_{k,2}} \text{III} 2^{\otimes c} \right) \right) \\
 &= - \sum_{t_{m-1,1}=0}^{t_{m-1}} (-1)^{t_{m-1,1}} \binom{t_{m-1}-t_{m-1,1}+t_m}{t_{m-1}-t_{m-1,1}} \binom{t_{m-1}+t_m+1}{t_{m-1,1}} \left(\frac{1}{2} \otimes 2^{\otimes t_{k-1,2}} \text{III} 2^{\otimes (t_{m-1}+t_m+1)} \right) \\
 &= (-1)^{t_{k+1}+1} \left(\frac{1}{2} \otimes 2^{\otimes t_{k-1,2}} \text{III} 2^{\otimes (c+t_k+1)} \right)
 \end{aligned}$$

which is exactly the symbol for the tree on the right side. Note that we used Proposition C.5 in the last line of the calculation.

By repeated application of this simplification, starting with $c = t_m$ and $k = m - 1$, we will arrive at a much simplified tree. By noting that $n - a - 1 = \sum_{i=1}^m (t_i + 1)$ and recalling that $t_0 = a$ we see that this tree is



times a factor of $(-1)^{(n-a-1)}$. This represents the symbol

$$\begin{aligned}
 & (-1)^{n-a-1} \left(\frac{1}{2} \otimes 2^{\otimes a} \text{III} 2^{\otimes (n-a-2)} \right) \\
 &= (-1)^{n-a} \binom{n-1}{a} 2^{\otimes n}
 \end{aligned}$$

Finally, by including the factor of $(-1)^n$ from the application of Hölder involution we find

$$\lambda_{a,n} = (-1)^a \binom{n-1}{a}.$$

□

Proposition C.6. *The polygon $P(\underbrace{0, \dots, 0}_{m_1-1}, \varepsilon_1, \dots, \underbrace{0, \dots, 0}_{m_k-1}, \varepsilon_k, 1)$ for $\varepsilon_i = \pm 1$ and at least one of the $m_i \neq 1$, has a symbol equal to 0.*

Proof. (Sketch) After applying Hölder involution (2.8) with $p = \infty$ we try to find possible hook-arrow trees which do not correspond to terms with coefficient 0 in the symbol. The vertices of the hook-arrow tree will be labelled corresponding to the sides of the polygon

$$P(\gamma_{1,1}, \dots, \gamma_{t_0,1}, 2, \gamma_{1,2}, \dots, \gamma_{t_1,2}, 2, \dots, 2, \gamma_{1,m}, \dots, \gamma_{t_m,m}, 1).$$

where all the $\gamma_{i,j}$ are equal to either 0 or 1. As in the proof of Proposition C.1, the vertices labelled 2 must connect directly to the final 1 and the vertices labelled 0 must connect to a 2. However, there is no way to connect the 1's to any other vertex without setting the coefficient of the term to 0. There is therefore no possible non-trivial dissection, and hence the symbol is zero. \square

D Some considerations on the implementation of the algorithm

The algorithmic approach presented in section 5 relies on the construction of the sets $\mathcal{R}^{(k)}$ defined in eq. (5.19). The algorithm then consists in selecting the elements of \mathcal{R} which satisfy certain factorization properties. Even though this is a mathematically well-defined prescription, implementing this algorithm into a computer program can be hampered by several issues,

1. The set \mathcal{R} is infinite, and so we cannot just proceed by ‘trial and error’ to select the elements that have the right factorization properties.
2. Factorization of polynomials is rather slow (at least on most computer algebra systems), leading to serious speed issues.

The first issue can be dealt with by decomposing \mathcal{R} as

$$\mathcal{R} = \bigcup_{n=0}^{\infty} \mathcal{R}_n, \tag{D.1}$$

where \mathcal{R}_n is defined as the subset of \mathcal{R} consisting of those elements $\pm \prod_{i=1}^k \pi_i^{n_i}$ such that $|n_1| + \dots + |n_k| = n$. In order to construct the set $\mathcal{R}^{(1)}$, which is the basis out of which $\mathcal{R}^{(k)}$ for arbitrary k is constructed, one can then limit oneself to truncating the tower of sets in eq. (D.1) to some finite value $N < \infty$. Indeed, in practical applications one does not expect rational functions for large values of the sum n of the exponents.¹⁰

The second item seems to be harder to solve, since it is related to the capabilities of the chosen computer algebra system. It is however possible to circumvent this problem by deriving from the factorization constraints a necessary condition that must be fulfilled and that can be checked in a fast way by a computer. Since in practice most of the elements of \mathcal{R} will fail this constraint, one can filter out these elements and discard them in an easier way. In the following we discuss the example of $\mathcal{R}^{(1)}$. The generalization to $\mathcal{R}^{(k)}$ is immediate.

¹⁰Empirically, we observe that the set $\mathcal{R}^{(1)}$ seems to be finite in general.

Let us consider a generic element R in \mathcal{R} . Without loss of generality, we can assume that we can write

$$R = s \frac{\bar{\pi}_1^{n_1} \dots \bar{\pi}_\ell^{n_\ell}}{\bar{\pi}_{\ell+1}^{n_{\ell+1}} \dots \bar{\pi}_k^{n_k}}, \tag{D.2}$$

with $s = \pm 1$ and $n_i \in \mathbb{N}$, and $\bar{\pi}_i \in \bar{P}$. Checking if $R \in \mathcal{R}^{(1)}$ is equivalent to checking whether $1 - R \in \mathcal{R}$, i.e., whether $1 - R$ can be written, up to a sign, as a ratio of elements from the set \bar{P} . Writing

$$1 - R = \frac{\bar{\pi}_{\ell+1}^{n_{\ell+1}} \dots \bar{\pi}_k^{n_k} - s \bar{\pi}_1^{n_1} \dots \bar{\pi}_\ell^{n_\ell}}{\bar{\pi}_{\ell+1}^{n_{\ell+1}} \dots \bar{\pi}_k^{n_k}}, \tag{D.3}$$

it is easy to see that this condition can only be fulfilled if the polynomial in the numerator can be factored into a product of elements in \bar{P} . Let us call Π this numerator. A necessary condition for $R \in \mathcal{R}^{(1)}$ is thus that Π can be divided by at least one element in \bar{P} . We can further simplify this condition by reducing it from a problem of division of polynomials to a problem of division of integers. Indeed, we can choose prime numbers $\{p_1, \dots, p_m\}$ such that $\bar{\pi}_i(p_1, \dots, p_m) \neq \bar{\pi}_j(p_1, \dots, p_m)$, for $i \neq j$. The necessary condition for $R \in \mathcal{R}^{(1)}$ then reduces to checking that the integer number $\Pi(p_1, \dots, p_m)$ can be divided by at least one of the numbers $\bar{\pi}_i(p_1, \dots, p_m)$, which is in general much quicker to test on a computer. Note however that this is a necessary, but not necessarily sufficient, condition for $R \in \mathcal{R}^{(1)}$.

E Analytic continuation of the spanning set of functions for harmonic polylogarithms

E.1 Analytic representation inside the unit disc

The analytic expressions for the spanning set $\{\mathcal{B}_j^{(i)}(x)\}$ introduced in section 6 are valid for $x \in [0, 1]$, but the functions might have different analytic representations in different regions of the complex plane. In particular, since the $\mathcal{B}_i^{(j)}$ functions are real when x is in the range $[0, 1]$, Schwarz's reflection principle implies that the elements of the spanning set must satisfy

$$\mathcal{B}_j^{(i)}(x^*) = \mathcal{B}_j^{(i)}(x)^*, \tag{E.1}$$

where x^* denotes the complex conjugate of the complex number x . We checked numerically that the analytic expressions for the elements of the spanning set are valid everywhere inside the unit disc, except for $\mathcal{B}_4^{(13)}(x)$, where the correct expression for $|x| < 1$ is

$$\mathcal{B}_4^{(13)}(x) = \begin{cases} \text{Li}_4(1-x^2), & \text{if } \text{Re}(x) > 0 \text{ or } (\text{Re}(x) = 0 \text{ and } \text{Im}(x) \geq 0), \\ \text{Li}_4(1-x^2) - \frac{i\pi}{3} \sigma(x) \log^3(1-x^2), & \text{otherwise,} \end{cases} \tag{E.2}$$

where $\sigma(x) = \text{sign}(\text{Im}(x))$. In order to understand this structure, let us look at the simpler case of weight two. Then we get

$$\text{Li}_2(1-x^2) = -\log(x^2) \log(1-x^2) - \text{Li}_2(x^2) + \frac{\pi^2}{6}. \tag{E.3}$$

The first term in this expression contains $\log(x^2) = 2\log x$, which is real for $x > 0$, but develops an imaginary part for $x < 0$, displaying the complicated branch cut structure of $\text{Li}_2(1 - x^2)$. A similar reasoning leads to eq. (E.2).

E.2 Analytic representation outside the unit disc: inversion relations

We now have analytic representations of the elements of the spanning set everywhere inside the unit disc, and so we can analytically continue them outside the unit disc via inversion relations, i.e., functional equations of the form

$$\mathcal{B}_j^{(i)}(x) = \sum_{k,l} c_{ijkl} \mathcal{B}_k^{(l)}\left(\frac{1}{x}\right) + \text{products of lower weight.} \quad (\text{E.4})$$

These functional equations can easily be obtained for the whole spanning set. Below we show the explicit inversion formulas for weight one and two. For the the complete list of inversion formulas for higher weights, we refer to appendix F. Letting $\sigma(x) = \text{sign}(\text{Im}(x))$, we get, for $|x| > 1$, x not real,

- for weight one:

$$\begin{aligned} \mathcal{B}_1^{(1)}(x) &= -\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right), \\ \mathcal{B}_1^{(2)}(x) &= -\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) + \mathcal{B}_1^{(2)}\left(\frac{1}{x}\right) - i\pi\sigma(x), \\ \mathcal{B}_1^{(3)}(x) &= \mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) - \mathcal{B}_1^{(1)}\left(\frac{1}{x}\right), \end{aligned} \quad (\text{E.5})$$

- for weight two:

$$\begin{aligned} \mathcal{B}_2^{(1)}(x) &= -i\pi\sigma(x)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) - \frac{1}{2}\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 - \mathcal{B}_2^{(1)}\left(\frac{1}{x}\right) + \frac{\pi^2}{3}, \\ \mathcal{B}_2^{(2)}(x) &= -\frac{1}{2}\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 - \mathcal{B}_2^{(2)}\left(\frac{1}{x}\right) - \frac{\pi^2}{6}, \\ \mathcal{B}_2^{(3)}(x) &= -\log 2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) - \frac{1}{2}\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 + \mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) + \mathcal{B}_2^{(1)}\left(\frac{1}{x}\right) \\ &\quad - \mathcal{B}_2^{(2)}\left(\frac{1}{x}\right) + \mathcal{B}_2^{(3)}\left(\frac{1}{x}\right) - \frac{\pi^2}{4}, \end{aligned} \quad (\text{E.6})$$

Note that the value of σ is ambiguous for real values of x . This ambiguity can be resolved by the ‘ $i\varepsilon$ ’ prescription commonly used in the physics literature. According to this prescription, we need to assign a small imaginary part to real value of x , i.e., if x is real, we need to perform the replacement $x \rightarrow x + i\varepsilon$, and this replacement fixes at the same time the value of σ . Note however that some care is needed when applying the inversion formulas to real

values of x because Schwarz' reflection principle implies

$$\begin{aligned}
 \mathcal{B}_j^{(i)}(x \pm i\varepsilon) &= \sum_{k,l} c_{ijkl} \mathcal{B}_k^{(l)}\left(\frac{1}{x \pm i\varepsilon}\right) + \text{products of lower weight} \\
 &= \sum_{k,l} c_{ijkl} \mathcal{B}_k^{(l)}\left(\frac{1}{x} \mp i\varepsilon\right) + \text{products of lower weight} \\
 &= \sum_{k,l} c_{ijkl} \mathcal{B}_k^{(l)}\left(\frac{1}{x} \pm i\varepsilon\right)^* + \text{products of lower weight.}
 \end{aligned}
 \tag{E.7}$$

If x lies on the unit circle, $|x| = 1$, there is an ambiguity whether to use the expression for $\mathcal{B}_j^{(i)}(x)$ valid inside or outside the unit circle. We checked numerically that the two values agree in all cases, except for $\mathcal{B}_5^{(14)}(x) = \text{Li}_4\left(\frac{4x}{(1+x)^2}\right)$. In this case, we find

$$\mathcal{B}_4^{(15)}(e^{i\varphi}) = \text{Li}_4\left(\frac{1}{\cos^2 \frac{\varphi}{2}}\right),
 \tag{E.8}$$

i.e., the argument of Li_4 is real and greater than 1 for every x on the unit circle (except for $x = -1$, where the result is divergent), and we have an ambiguity on the imaginary part of $\mathcal{B}_4^{(15)}(e^{i\varphi})$. This ambiguity can be lifted by requiring the function to be continuous in a neighborhood of the unit circle. To study this, let us consider a circle which is infinitesimally close to the unit circle, i.e., we choose $x = (1 - \varepsilon) e^{i\varphi}$, for some infinitesimal ε . We then find

$$\frac{4x}{(1+x)^2} = \frac{1}{\cos^2 \frac{\varphi}{2}} \left(1 + i\varepsilon \tan \frac{\varphi}{2} + \mathcal{O}(\varepsilon^2)\right),
 \tag{E.9}$$

i.e., we see that for $|x| = 1$, we have,

$$\mathcal{B}_4^{(15)}(x) = \text{Li}_4\left(\frac{4x}{(1+x)^2} + i\sigma\varepsilon\right),
 \tag{E.10}$$

with $\sigma(x) = \text{sign}(\text{Im}(x))$.

F Inversion formulas for the spanning set

In this appendix we present the inversion formulas for the spanning set for weight three and four, valid for $x \in \mathbb{C}^\times$.

F.1 Weight three

$$\begin{aligned}
 \mathcal{B}_3^{(1)}(x) &= \frac{1}{2} i\pi\sigma(x)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 + \frac{1}{6}\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^3 - \frac{1}{3}\pi^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) + \mathcal{B}_3^{(1)}\left(\frac{1}{x}\right), \\
 \mathcal{B}_3^{(2)}(x) &= \frac{1}{6}\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^3 + \frac{1}{6}\pi^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) + \mathcal{B}_3^{(2)}\left(\frac{1}{x}\right),
 \end{aligned}
 \tag{F.1}$$

$$\begin{aligned}
 \mathcal{B}_3^{(3)}(x) &= \frac{1}{6}\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^3 - \frac{1}{2}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 + \frac{1}{6}\pi^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) - \mathcal{B}_3^{(1)}\left(\frac{1}{x}\right) \\
 &\quad - \mathcal{B}_3^{(3)}\left(\frac{1}{x}\right) + \zeta_3, \\
 \mathcal{B}_3^{(4)}(x) &= \frac{1}{3}\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^3 - \frac{1}{2}\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 - \frac{1}{6}\pi^2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) - \mathcal{B}_3^{(2)}\left(\frac{1}{x}\right) \\
 &\quad - \mathcal{B}_3^{(4)}\left(\frac{1}{x}\right) + \zeta_3, \\
 \mathcal{B}_3^{(5)}(x) &= i\pi\sigma(x)\log 2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) - i\pi\sigma(x)\log 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) + \frac{1}{2}\log^2 2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) \\
 &\quad - \frac{1}{2}\log^2 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) + \frac{1}{2}\log 2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 - \log 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) \\
 &\quad - \frac{1}{2}\log 2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2 + \log 2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) + \frac{1}{2}i\pi\sigma(x)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 \\
 &\quad - i\pi\sigma(x)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) + \frac{1}{2}i\pi\sigma(x)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 + \frac{1}{6}\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^3 \\
 &\quad - \frac{1}{2}\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 - \frac{1}{2}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) \\
 &\quad + \mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) - \frac{1}{3}\pi^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) + \frac{1}{6}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^3 \\
 &\quad + \frac{1}{6}\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^3 - \frac{1}{2}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 + \frac{1}{6}\pi^2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right) + \frac{1}{6}\pi^2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) \\
 &\quad - \mathcal{B}_3^{(7)}\left(\frac{1}{x}\right) - \mathcal{B}_3^{(8)}\left(\frac{1}{x}\right) + \frac{1}{2}i\pi\sigma(x)\log^2 2 + \frac{1}{6}\log^3 2 - \frac{1}{3}\pi^2\log 2 + \zeta_3, \\
 \mathcal{B}_3^{(6)}(x) &= \frac{1}{2}\log^2 2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) - \frac{1}{2}\log^2 2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right) + \frac{1}{2}\log 2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 \\
 &\quad - \log 2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) + \frac{1}{2}\log 2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2 + \frac{1}{6}\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^3 \\
 &\quad - \frac{1}{2}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 + \frac{1}{2}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) + \frac{1}{6}\pi^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) \\
 &\quad - \frac{1}{6}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^3 - \frac{1}{6}\pi^2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right) + \mathcal{B}_3^{(8)}\left(\frac{1}{x}\right) + \frac{1}{6}\log^3 2 + \frac{1}{6}\pi^2\log 2, \\
 \mathcal{B}_3^{(7)}(x) &= -\frac{1}{2}\log^2 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) - \frac{1}{2}\log^2 2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right) + \log 2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) \\
 &\quad + \frac{1}{6}\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^3 - \frac{1}{2}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 + \frac{1}{6}\pi^2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) - \mathcal{B}_3^{(5)}\left(\frac{1}{x}\right) \\
 &\quad - \mathcal{B}_3^{(6)}\left(\frac{1}{x}\right) + \frac{1}{3}\log^3 2 - \frac{1}{6}\pi^2\log 2 + \zeta_3,
 \end{aligned}$$

$$\begin{aligned}
 \mathcal{B}_3^{(8)}(x) &= i\pi\sigma(x)\log 2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right) + \frac{1}{2}\log^2 2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right) - \frac{1}{2}\log 2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2 \\
 &\quad - \frac{1}{2}i\pi\sigma(x)\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2 + \frac{1}{6}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^3 - \frac{1}{3}\pi^2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right) + \mathcal{B}_3^{(6)}\left(\frac{1}{x}\right) \\
 &\quad - \frac{1}{2}i\pi\sigma(x)\log^2 2 - \frac{1}{6}\log^3 2 + \frac{1}{3}\pi^2\log 2.
 \end{aligned}$$

F.2 Weight four

$$\begin{aligned}
 \mathcal{B}_4^{(1)}(x) &= -\frac{1}{6}i\pi\sigma(x)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^3 - \frac{1}{24}\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^4 + \frac{1}{6}\pi^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 \\
 &\quad - \mathcal{B}_4^{(1)}\left(\frac{1}{x}\right) + \frac{\pi^4}{45}, \\
 \mathcal{B}_4^{(2)}(x) &= -\frac{1}{24}\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^4 - \frac{1}{12}\pi^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 - \mathcal{B}_4^{(2)}\left(\frac{1}{x}\right) - \frac{7\pi^4}{360}, \\
 \mathcal{B}_4^{(3)}(x) &= -\frac{1}{24}\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^4 + \frac{1}{6}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^3 - \frac{1}{4}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 \\
 &\quad - \frac{1}{12}\pi^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 + \frac{1}{6}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^3\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) + \frac{1}{6}\pi^2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) \\
 &\quad - \frac{1}{24}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^4 - \frac{1}{12}\pi^2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2 - \mathcal{B}_4^{(5)}\left(\frac{1}{x}\right) - \frac{7\pi^4}{360}, \\
 \mathcal{B}_4^{(4)}(x) &= \mathcal{B}_4^{(6)}\left(\frac{1}{x}\right), \\
 \mathcal{B}_4^{(5)}(x) &= \frac{1}{6}i\pi\sigma(x)\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^3 - \frac{1}{24}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^4 + \frac{1}{6}\pi^2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2 \\
 &\quad - \mathcal{B}_4^{(3)}\left(\frac{1}{x}\right) + \frac{\pi^4}{45}, \\
 \mathcal{B}_4^{(6)}(x) &= \mathcal{B}_4^{(4)}\left(\frac{1}{x}\right), \\
 \mathcal{B}_4^{(7)}(x) &= -\frac{1}{2}i\pi\sigma(x)\log^2 2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) + \frac{1}{2}i\pi\sigma(x)\log^2 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) \\
 &\quad - \frac{1}{2}i\pi\sigma(x)\log 2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 + i\pi\sigma(x)\log 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) \\
 &\quad - \frac{1}{2}i\pi\sigma(x)\log 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 - \frac{1}{6}\log^3 2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) + \frac{1}{6}\log^3 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) \\
 &\quad - \frac{1}{4}\log^2 2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 + \frac{1}{2}\log^2 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) - \frac{1}{4}\log^2 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 \\
 &\quad - \frac{1}{6}\log 2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^3 + \frac{1}{2}\log 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2
 \end{aligned} \tag{F.2}$$

$$\begin{aligned}
 & -\frac{1}{2} \log 2\mathcal{B}_1^{(3)} \left(\frac{1}{x}\right)^2 \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right) + \frac{1}{3} \pi^2 \log 2\mathcal{B}_1^{(1)} \left(\frac{1}{x}\right) + \frac{1}{6} \log 2\mathcal{B}_1^{(3)} \left(\frac{1}{x}\right)^3 \\
 & -\frac{1}{3} \pi^2 \log 2\mathcal{B}_1^{(3)} \left(\frac{1}{x}\right) - \frac{1}{6} i\pi\sigma(x) \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right)^3 + \frac{1}{2} i\pi\sigma(x) \mathcal{B}_1^{(3)} \left(\frac{1}{x}\right) \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right)^2 \\
 & -\frac{1}{2} i\pi\sigma(x) \mathcal{B}_1^{(3)} \left(\frac{1}{x}\right)^2 \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right) + \frac{1}{6} i\pi\sigma(x) \mathcal{B}_1^{(3)} \left(\frac{1}{x}\right)^3 - \frac{1}{24} \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right)^4 \\
 & + \frac{1}{6} \mathcal{B}_1^{(3)} \left(\frac{1}{x}\right) \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right)^3 - \frac{1}{4} \mathcal{B}_1^{(3)} \left(\frac{1}{x}\right)^2 \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right)^2 + \frac{1}{6} \pi^2 \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right)^2 \\
 & + \frac{1}{6} \mathcal{B}_1^{(3)} \left(\frac{1}{x}\right)^3 \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right) - \frac{1}{3} \pi^2 \mathcal{B}_1^{(3)} \left(\frac{1}{x}\right) \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right) - \frac{1}{24} \mathcal{B}_1^{(3)} \left(\frac{1}{x}\right)^4 \\
 & + \frac{1}{6} \pi^2 \mathcal{B}_1^{(3)} \left(\frac{1}{x}\right)^2 - \mathcal{B}_4^{(11)} \left(\frac{1}{x}\right) - \frac{1}{6} i\pi\sigma(x) \log^3 2 - \frac{1}{24} \log^4 2 + \frac{1}{6} \pi^2 \log^2 2 + \frac{\pi^4}{45}, \\
 \mathcal{B}_4^{(8)}(x) &= -\frac{1}{6} \log^3 2\mathcal{B}_1^{(1)} \left(\frac{1}{x}\right) + \frac{1}{6} \log^3 2\mathcal{B}_1^{(2)} \left(\frac{1}{x}\right) - \frac{1}{4} \log^2 2\mathcal{B}_1^{(1)} \left(\frac{1}{x}\right)^2 \\
 & + \frac{1}{2} \log^2 2\mathcal{B}_1^{(2)} \left(\frac{1}{x}\right) \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right) - \frac{1}{4} \log^2 2\mathcal{B}_1^{(2)} \left(\frac{1}{x}\right)^2 - \frac{1}{6} \log 2\mathcal{B}_1^{(1)} \left(\frac{1}{x}\right)^3 \\
 & + \frac{1}{2} \log 2\mathcal{B}_1^{(2)} \left(\frac{1}{x}\right) \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right)^2 - \frac{1}{2} \log 2\mathcal{B}_1^{(2)} \left(\frac{1}{x}\right)^2 \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right) \\
 & - \frac{1}{6} \pi^2 \log 2\mathcal{B}_1^{(1)} \left(\frac{1}{x}\right) + \frac{1}{6} \log 2\mathcal{B}_1^{(2)} \left(\frac{1}{x}\right)^3 + \frac{1}{6} \pi^2 \log 2\mathcal{B}_1^{(2)} \left(\frac{1}{x}\right) \\
 & - \frac{1}{24} \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right)^4 + \frac{1}{6} \mathcal{B}_1^{(2)} \left(\frac{1}{x}\right) \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right)^3 - \frac{1}{4} \mathcal{B}_1^{(2)} \left(\frac{1}{x}\right)^2 \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right)^2 \\
 & - \frac{1}{12} \pi^2 \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right)^2 + \frac{1}{6} \mathcal{B}_1^{(2)} \left(\frac{1}{x}\right)^3 \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right) + \frac{1}{6} \pi^2 \mathcal{B}_1^{(2)} \left(\frac{1}{x}\right) \mathcal{B}_1^{(1)} \left(\frac{1}{x}\right) \\
 & - \frac{1}{24} \mathcal{B}_1^{(2)} \left(\frac{1}{x}\right)^4 - \frac{1}{12} \pi^2 \mathcal{B}_1^{(2)} \left(\frac{1}{x}\right)^2 - \mathcal{B}_4^{(12)} \left(\frac{1}{x}\right) - \frac{1}{24} \log^4 2 \\
 & - \frac{1}{12} \pi^2 \log^2 2 - \frac{7\pi^4}{360}, \\
 \mathcal{B}_4^{(9)}(x) &= \mathcal{B}_4^{(10)} \left(\frac{1}{x}\right), \\
 \mathcal{B}_4^{(10)}(x) &= \mathcal{B}_4^{(9)} \left(\frac{1}{x}\right), \\
 \mathcal{B}_4^{(11)}(x) &= -\frac{1}{2} i\pi\sigma(x) \log^2 2\mathcal{B}_1^{(3)} \left(\frac{1}{x}\right) + \frac{1}{2} i\pi\sigma(x) \log 2\mathcal{B}_1^{(3)} \left(\frac{1}{x}\right)^2 + \frac{1}{6} \log^3 2\mathcal{B}_1^{(3)} \left(\frac{1}{x}\right) \\
 & - \frac{1}{4} \log^2 2\mathcal{B}_1^{(3)} \left(\frac{1}{x}\right)^2 + \frac{1}{6} \log 2\mathcal{B}_1^{(3)} \left(\frac{1}{x}\right)^3 - \frac{1}{3} \pi^2 \log 2\mathcal{B}_1^{(3)} \left(\frac{1}{x}\right) \\
 & - \frac{1}{6} i\pi\sigma(x) \mathcal{B}_1^{(3)} \left(\frac{1}{x}\right)^3 - \frac{1}{24} \mathcal{B}_1^{(3)} \left(\frac{1}{x}\right)^4 + \frac{1}{6} \pi^2 \mathcal{B}_1^{(3)} \left(\frac{1}{x}\right)^2 - \mathcal{B}_4^{(7)} \left(\frac{1}{x}\right) \\
 & + \frac{1}{6} i\pi\sigma(x) \log^3 2 - \frac{1}{24} \log^4 2 + \frac{1}{6} \pi^2 \log^2 2 + \frac{\pi^4}{45},
 \end{aligned}$$

$$\begin{aligned}
 \mathcal{B}_4^{(12)}(x) &= \frac{1}{2}i\pi\sigma(x)\log^2 2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right) - \frac{1}{2}i\pi\sigma(x)\log 2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2 + \frac{1}{6}\log^3 2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right) \\
 &\quad - \frac{1}{4}\log^2 2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2 + \frac{1}{6}\log 2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^3 - \frac{1}{3}\pi^2\log 2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right) \\
 &\quad + \frac{1}{6}i\pi\sigma(x)\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^3 - \frac{1}{24}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^4 + \frac{1}{6}\pi^2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2 - \mathcal{B}_4^{(8)}\left(\frac{1}{x}\right) \\
 &\quad - \frac{1}{6}i\pi\sigma(x)\log^3 2 - \frac{1}{24}\log^4 2 + \frac{1}{6}\pi^2\log^2 2 + \frac{\pi^4}{45}, \\
 \mathcal{B}_4^{(13)}(x) &= -\frac{2}{3}\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^4 + \frac{4}{3}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^3 + \frac{4}{3}\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^3 \\
 &\quad - \mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 - \mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 \\
 &\quad - 2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 - \frac{1}{3}\pi^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 \\
 &\quad + \frac{1}{3}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^3\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) + \frac{1}{3}\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^3\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) \\
 &\quad + \mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) + \frac{1}{3}\pi^2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) \\
 &\quad + \mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) + \frac{1}{3}\pi^2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) \\
 &\quad - \frac{1}{24}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^4 - \frac{1}{24}\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^4 - \frac{1}{6}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^3 \\
 &\quad - \frac{1}{12}\pi^2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2 - \frac{1}{4}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 - \frac{1}{12}\pi^2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 \\
 &\quad - \frac{1}{6}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^3\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) - \frac{1}{6}\pi^2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) \\
 &\quad - \mathcal{B}_4^{(14)}\left(\frac{1}{x}\right) - \frac{7\pi^4}{360}, \\
 \mathcal{B}_4^{(14)}(x) &= \frac{1}{6}i\pi\sigma(x)\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^3 + \frac{1}{2}i\pi\sigma(x)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2 \\
 &\quad + \frac{1}{2}i\pi\sigma(x)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right) + \frac{1}{6}i\pi\sigma(x)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^3 \\
 &\quad - \frac{1}{24}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^4 - \frac{1}{6}\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^3 - \frac{1}{4}\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2 \\
 &\quad + \frac{1}{6}\pi^2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2 - \frac{1}{6}\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^3\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right) + \frac{1}{3}\pi^2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right) \\
 &\quad - \frac{1}{24}\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^4 + \frac{1}{6}\pi^2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 - \mathcal{B}_4^{(13)}\left(\frac{1}{x}\right) + \frac{\pi^4}{45},
 \end{aligned}$$

$$\begin{aligned}
 \mathcal{B}_4^{(15)}(x) &= 4i\pi\sigma(x)\log^2 2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) - 8i\pi\sigma(x)\log^2 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) \\
 &\quad + 2i\pi\sigma(x)\log 2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 - 8i\pi\sigma(x)\log 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) \\
 &\quad + 8i\pi\sigma(x)\log 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 + \frac{1}{3}i\pi\sigma(x)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^3 \\
 &\quad - 2i\pi\sigma(x)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 + 4i\pi\sigma(x)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) \\
 &\quad - \frac{8}{3}i\pi\sigma(x)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^3 + \mathcal{B}_4^{(15)}\left(\frac{1}{x}\right) + \frac{8}{3}i\pi\sigma(x)\log^3 2, \\
 \mathcal{B}_4^{(16)}(x) &= \frac{1}{12}i\pi^3\sigma(x)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) - \frac{3}{2}\zeta_3\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) + \frac{1}{24}\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^4 \\
 &\quad + \frac{1}{2}\mathcal{B}_2^{(1)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 + \frac{1}{12}\pi^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)^2 - 2\mathcal{B}_3^{(1)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right) \\
 &\quad + \frac{1}{6}\pi^2\mathcal{B}_2^{(1)}\left(\frac{1}{x}\right) + 3\mathcal{B}_4^{(1)}\left(\frac{1}{x}\right) + \mathcal{B}_4^{(2)}\left(\frac{1}{x}\right) + \mathcal{B}_4^{(16)}\left(\frac{1}{x}\right) - \frac{37\pi^4}{720}, \\
 \mathcal{B}_4^{(17)}(x) &= -\frac{19}{24}\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^4 + \frac{7}{6}\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^3 + \frac{5}{2}\log 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^3 \\
 &\quad - \frac{9}{4}\log^2 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 - \mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 \\
 &\quad - \frac{1}{2}\mathcal{B}_2^{(1)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 + \frac{1}{2}\mathcal{B}_2^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 \\
 &\quad - \mathcal{B}_2^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 - \mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)\log 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 \\
 &\quad - \mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\log 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 + \frac{19}{24}\pi^2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)^2 - \frac{1}{3}\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^3\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) \\
 &\quad + \frac{5}{6}\log^3 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) + \mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) \\
 &\quad + \frac{3}{2}\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)\log^2 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) - \frac{1}{4}\pi^2\mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) \\
 &\quad - \frac{1}{6}\pi^2\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) + \mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)\mathcal{B}_2^{(3)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) \\
 &\quad + 2\mathcal{B}_3^{(1)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) - 2\mathcal{B}_3^{(5)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) + 2\mathcal{B}_3^{(7)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) \\
 &\quad + 2\mathcal{B}_3^{(8)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) + \mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)^2\log 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) \\
 &\quad - \mathcal{B}_1^{(1)}\left(\frac{1}{x}\right)\mathcal{B}_1^{(2)}\left(\frac{1}{x}\right)\log 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) - \frac{2}{3}\pi^2\log 2\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) \\
 &\quad + \frac{1}{2}i\pi\log^2 2\sigma(x)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) - \frac{1}{12}i\pi^3\sigma(x)\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right) + 3\zeta_3\mathcal{B}_1^{(3)}\left(\frac{1}{x}\right)
 \end{aligned}$$

$$\begin{aligned}
 & + \frac{1}{8} \log^4 2 - \frac{1}{3} \mathcal{B}_1^{(1)} \left(\frac{1}{x} \right) \log^3 2 + \frac{1}{2} \mathcal{B}_2^{(2)} \left(\frac{1}{x} \right)^2 + \frac{1}{12} \pi^2 \log^2 2 \\
 & + \frac{1}{6} \pi^2 \mathcal{B}_2^{(3)} \left(\frac{1}{x} \right) + 2 \mathcal{B}_1^{(1)} \left(\frac{1}{x} \right) \mathcal{B}_3^{(5)} \left(\frac{1}{x} \right) - \frac{3}{2} \mathcal{B}_4^{(1)} \left(\frac{1}{x} \right) - \frac{1}{2} \mathcal{B}_4^{(2)} \left(\frac{1}{x} \right) \\
 & - 2 \mathcal{B}_4^{(4)} \left(\frac{1}{x} \right) - 2 \mathcal{B}_4^{(6)} \left(\frac{1}{x} \right) - 3 \mathcal{B}_4^{(7)} \left(\frac{1}{x} \right) + 3 \mathcal{B}_4^{(11)} \left(\frac{1}{x} \right) - \frac{3}{4} \mathcal{B}_4^{(15)} \left(\frac{1}{x} \right) \\
 & - \mathcal{B}_4^{(16)} \left(\frac{1}{x} \right) - \mathcal{B}_4^{(17)} \left(\frac{1}{x} \right) + \frac{1}{6} \pi^2 \mathcal{B}_1^{(1)} \left(\frac{1}{x} \right) \log 2 - \frac{1}{6} \pi^2 \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right) \log 2 \\
 & - \frac{1}{2} i \pi \log^3 2 \sigma(x) + \frac{1}{12} i \pi^3 \log 2 \sigma(x) - \frac{7}{4} \mathcal{B}_1^{(1)} \left(\frac{1}{x} \right) \zeta_3 + 6 \text{Li}_4 \left(\frac{1}{2} \right) - \frac{\pi^4}{160}, \\
 \mathcal{B}_4^{(18)}(x) = & \frac{35}{96} \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right)^4 - \frac{5}{12} \mathcal{B}_1^{(1)} \left(\frac{1}{x} \right) \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right)^3 - \frac{1}{24} \mathcal{B}_1^{(3)} \left(\frac{1}{x} \right) \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right)^3 \\
 & - \frac{1}{2} \log 2 \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right)^3 - \frac{1}{6} i \pi \sigma(x) \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right)^3 - \frac{1}{16} \mathcal{B}_1^{(3)} \left(\frac{1}{x} \right)^2 \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right)^2 \\
 & + \frac{1}{4} \log^2 2 \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right)^2 + \frac{1}{4} \mathcal{B}_1^{(1)} \left(\frac{1}{x} \right) \mathcal{B}_1^{(3)} \left(\frac{1}{x} \right) \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right)^2 \\
 & + \frac{1}{2} \mathcal{B}_2^{(1)} \left(\frac{1}{x} \right) \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right)^2 - \frac{1}{2} \mathcal{B}_2^{(2)} \left(\frac{1}{x} \right) \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right)^2 + \mathcal{B}_2^{(3)} \left(\frac{1}{x} \right) \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right)^2 \\
 & + \frac{1}{48} \pi^2 \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right)^2 - \frac{1}{24} \mathcal{B}_1^{(3)} \left(\frac{1}{x} \right)^3 \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right) - \frac{1}{6} \log^3 2 \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right) \\
 & + \frac{1}{4} \mathcal{B}_1^{(1)} \left(\frac{1}{x} \right) \mathcal{B}_1^{(3)} \left(\frac{1}{x} \right)^2 \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right) + \frac{1}{2} \mathcal{B}_1^{(1)} \left(\frac{1}{x} \right) \log^2 2 \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right) \\
 & - \frac{1}{12} \pi^2 \mathcal{B}_1^{(1)} \left(\frac{1}{x} \right) \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right) - \frac{1}{24} \pi^2 \mathcal{B}_1^{(3)} \left(\frac{1}{x} \right) \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right) \\
 & - \mathcal{B}_1^{(1)} \left(\frac{1}{x} \right) \mathcal{B}_2^{(3)} \left(\frac{1}{x} \right) \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right) - 2 \mathcal{B}_3^{(6)} \left(\frac{1}{x} \right) \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right) \\
 & - 2 \mathcal{B}_3^{(8)} \left(\frac{1}{x} \right) \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right) - i \pi \mathcal{B}_2^{(3)} \left(\frac{1}{x} \right) \sigma(x) \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right) \\
 & + \frac{7}{2} \zeta_3 \mathcal{B}_1^{(2)} \left(\frac{1}{x} \right) + \frac{55}{96} \mathcal{B}_1^{(3)} \left(\frac{1}{x} \right)^4 + \frac{1}{8} \log^4 2 - \frac{7}{12} \mathcal{B}_1^{(1)} \left(\frac{1}{x} \right) \mathcal{B}_1^{(3)} \left(\frac{1}{x} \right)^3 \\
 & - \frac{1}{3} \mathcal{B}_1^{(1)} \left(\frac{1}{x} \right) \log^3 2 - \mathcal{B}_1^{(3)} \left(\frac{1}{x} \right) \log^3 2 - \frac{29}{48} \pi^2 \mathcal{B}_1^{(3)} \left(\frac{1}{x} \right)^2 + \frac{1}{2} \mathcal{B}_2^{(1)} \left(\frac{1}{x} \right)^2 \\
 & + \frac{3}{2} \mathcal{B}_1^{(3)} \left(\frac{1}{x} \right)^2 \log^2 2 - \frac{1}{12} \pi^2 \log^2 2 - \mathcal{B}_2^{(1)} \left(\frac{1}{x} \right) \mathcal{B}_2^{(2)} \left(\frac{1}{x} \right) - \frac{1}{6} \pi^2 \mathcal{B}_2^{(3)} \left(\frac{1}{x} \right) \\
 & - 2 \mathcal{B}_1^{(3)} \left(\frac{1}{x} \right) \mathcal{B}_3^{(1)} \left(\frac{1}{x} \right) + 2 \mathcal{B}_1^{(1)} \left(\frac{1}{x} \right) \mathcal{B}_3^{(6)} \left(\frac{1}{x} \right) + \frac{3}{2} \mathcal{B}_4^{(1)} \left(\frac{1}{x} \right) + \frac{1}{2} \mathcal{B}_4^{(2)} \left(\frac{1}{x} \right) \\
 & - 2 \mathcal{B}_4^{(3)} \left(\frac{1}{x} \right) + 4 \mathcal{B}_4^{(4)} \left(\frac{1}{x} \right) + 2 \mathcal{B}_4^{(5)} \left(\frac{1}{x} \right) + 4 \mathcal{B}_4^{(6)} \left(\frac{1}{x} \right) + 6 \mathcal{B}_4^{(7)} \left(\frac{1}{x} \right) \\
 & + 3 \mathcal{B}_4^{(8)} \left(\frac{1}{x} \right) - 6 \mathcal{B}_4^{(11)} \left(\frac{1}{x} \right) - 3 \mathcal{B}_4^{(12)} \left(\frac{1}{x} \right) + \frac{1}{4} \mathcal{B}_4^{(13)} \left(\frac{1}{x} \right) - \frac{1}{4} \mathcal{B}_4^{(14)} \left(\frac{1}{x} \right)
 \end{aligned}$$

$$\begin{aligned}
& + \frac{3}{4} \mathcal{B}_4^{(15)} \left(\frac{1}{x} \right) + \mathcal{B}_4^{(16)} \left(\frac{1}{x} \right) - \mathcal{B}_4^{(18)} \left(\frac{1}{x} \right) - \mathcal{B}_1^{(3)} \left(\frac{1}{x} \right)^3 \log 2 \\
& + \frac{1}{6} \pi^2 \mathcal{B}_1^{(1)} \left(\frac{1}{x} \right) \log 2 + \frac{1}{2} \pi^2 \mathcal{B}_1^{(3)} \left(\frac{1}{x} \right) \log 2 + \frac{1}{6} i\pi \log^3 2 \sigma(x) + 2i\pi \mathcal{B}_3^{(6)} \left(\frac{1}{x} \right) \sigma(x) \\
& + \frac{1}{12} i\pi^3 \log 2 \sigma(x) - \frac{7}{4} i\pi \zeta_3 \sigma(x) - \frac{7}{4} \mathcal{B}_1^{(1)} \left(\frac{1}{x} \right) \zeta_3 - \frac{3}{2} \mathcal{B}_1^{(3)} \left(\frac{1}{x} \right) \zeta_3 \\
& - 6\text{Li}_4 \left(\frac{1}{2} \right) - \frac{37\pi^4}{1440}.
\end{aligned}$$

G Expression of HPL's in terms of the spanning set

In this appendix we present the results for expressing all HPL's up to weight four in terms of the spanning set $\{\mathcal{B}_i^{(j)}\}$. We restrict ourselves to giving the expression for a minimal set of HPL's out of which all other cases can be obtained via shuffle relations.¹¹

G.1 Results for weight two

$$\begin{aligned}
H(-1, 1; x) &= \log 2 \log(1-x) - \log(1-x) \log(1+x) - \frac{1}{2} \log^2 2 - \text{Li}_2 \left(\frac{1-x}{2} \right) + \frac{\pi^2}{12}, \\
H(0, -1; x) &= -\text{Li}_2(-x), \\
H(0, 1; x) &= \text{Li}_2(x).
\end{aligned}$$

G.2 Results for weight three

$$\begin{aligned}
H(-1, 1, -1; x) &= -\text{Li}_2 \left(\frac{1-x}{2} \right) \log(1+x) - \frac{3}{2} \log^2 2 \log(1+x) \\
& + \log 2 \log^2(1+x) - \log(1-x) \log^2(1+x) + \log 2 \log(1-x) \log(1+x) \\
& + \frac{1}{4} \pi^2 \log(1+x) + \frac{1}{3} \log^3 2 - \frac{1}{6} \pi^2 \log 2 - 2\text{Li}_3 \left(\frac{1+x}{2} \right) + \frac{7\zeta_3}{4}, \\
H(-1, 1, 1; x) &= \text{Li}_2 \left(\frac{1-x}{2} \right) \log(1-x) - \frac{1}{2} \log 2 \log^2(1-x) + \frac{1}{2} \log^2(1-x) \log(1+x) \\
& + \frac{1}{6} \log^3 2 - \frac{1}{12} \pi^2 \log 2 - \text{Li}_3 \left(\frac{1-x}{2} \right) + \frac{7\zeta_3}{8}, \\
H(0, -1, -1; x) &= -\text{Li}_2(-x) \log(1+x) + \frac{1}{6} \log^3(1+x) - \frac{1}{2} \log x \log^2(1+x) - \frac{\pi^2}{6} \log(1+x) \\
& - \text{Li}_3 \left(\frac{1}{1+x} \right) + \zeta_3, \\
H(0, -1, 1; x) &= \text{Li}_2(-x) \log(1-x) - \frac{1}{6} \log^3(1-x) - \frac{1}{2} \log^2 2 \log(1-x) + \frac{1}{2} \log 2 \log^2(1-x) \\
& + \frac{1}{2} \log^2(1-x) \log x - \frac{\pi^2}{12} \log(1-x) + \frac{1}{6} \log^3 2 - \frac{\pi^2}{12} \log 2 - \text{Li}_3 \left(\frac{1-x}{2} \right) \\
& + \text{Li}_3(1-x) - \text{Li}_3(-x) + \text{Li}_3(x) + \text{Li}_3 \left(\frac{2x}{x-1} \right) - \frac{1}{8} \zeta_3,
\end{aligned}$$

¹¹A set of text files containing the expressions for all HPL's up to weight four (for $x \in [0, 1]$) in Mathematica is included in the arXiv distribution.

$$H(0, 0, -1; x) = -\text{Li}_3(-x),$$

$$H(0, 0, 1; x) = \text{Li}_3(x),$$

$$\begin{aligned} H(0, 1, -1; x) &= \text{Li}_2(x) \log(1+x) + \frac{1}{6} \log^3(1-x) - \frac{1}{2} \log^2 2 \log(1+x) - \frac{1}{2} \log 2 \log^2(1-x) \\ &\quad - \frac{1}{2} \log(1-x) \log^2(1+x) - \frac{1}{2} \log^2(1-x) \log x + \log 2 \log(1-x) \log(1+x) \\ &\quad + \frac{\pi^2}{6} \log(1-x) + \log(1-x) \log x \log(1+x) + \frac{\pi^2}{12} \log(1+x) + \frac{1}{6} \log^3 2 \\ &\quad - \frac{\pi^2}{12} \log 2 + \text{Li}_3(-x) - \text{Li}_3(x) - \text{Li}_3\left(\frac{2x}{x-1}\right) + \text{Li}_3\left(\frac{1}{1+x}\right) \\ &\quad - \text{Li}_3\left(\frac{1-x}{1+x}\right) - \text{Li}_3\left(\frac{1+x}{2}\right) + \frac{7}{8} \zeta_3, \end{aligned}$$

$$H(0, 1, 1; x) = -\text{Li}_2(x) \log(1-x) - \frac{1}{2} \log x \log^2(1-x) + \frac{\pi^2}{6} \log(1-x) - \text{Li}_3(1-x) + \zeta_3.$$

G.3 Results for weight four

$$\begin{aligned} H(-1, 1, -1, -1; x) &= -\frac{1}{2} \text{Li}_2\left(\frac{1-x}{2}\right) \log^2(1+x) - 2\text{Li}_3\left(\frac{1+x}{2}\right) \log(1+x) \\ &\quad - \frac{7}{8} \zeta_3 \log(1+x) - \frac{1}{6} \log^3 2 \log(1+x) + \frac{1}{2} \log 2 \log^3(1+x) \\ &\quad - \frac{1}{2} \log(1-x) \log^3(1+x) - \frac{1}{2} \log^2 2 \log^2(1+x) \\ &\quad + \frac{1}{2} \log 2 \log(1-x) \log^2(1+x) + \frac{\pi^2}{12} \log^2(1+x) + \frac{\pi^2}{12} \log 2 \log(1+x) \\ &\quad + 3\text{Li}_4\left(\frac{1+x}{2}\right) - 3\text{Li}_4\left(\frac{1}{2}\right), \end{aligned}$$

$$\begin{aligned} H(-1, 1, -1, 1; x) &= \frac{1}{2} \text{Li}_2\left(\frac{1-x}{2}\right) \log^2 2 - \text{Li}_2\left(\frac{1-x}{2}\right) \log 2 \log(1-x) \\ &\quad + \text{Li}_2\left(\frac{1-x}{2}\right) \log(1-x) \log(1+x) + 2\text{Li}_3\left(\frac{1+x}{2}\right) \log(1-x) \\ &\quad - 2\zeta_3 \log(1-x) + \frac{1}{4} \zeta_3 \log(1+x) + \frac{1}{12} \log^4(1+x) - \frac{7}{6} \log^3 2 \log(1-x) \\ &\quad + \frac{1}{3} \log^3 2 \log(1+x) - \frac{1}{3} \log(1-x) \log^3(1+x) + \log^2 2 \log^2(1-x) \\ &\quad - \frac{1}{2} \log^2 2 \log^2(1+x) + \frac{3}{2} \log^2 2 \log(1-x) \log(1+x) \\ &\quad - 2 \log 2 \log^2(1-x) \log(1+x) + \log^2(1-x) \log^2(1+x) + \frac{\pi^2}{6} \log^2(1+x) \\ &\quad + \frac{5}{12} \pi^2 \log 2 \log(1-x) - \frac{\pi^2}{6} \log 2 \log(1+x) - \frac{5}{12} \pi^2 \log(1-x) \log(1+x) \\ &\quad + \frac{1}{8} \log^4 2 - \frac{\pi^2}{24} \log^2 2 + \frac{1}{2} \text{Li}_2\left(\frac{1-x}{2}\right)^2 - \frac{\pi^2}{12} \text{Li}_2\left(\frac{1-x}{2}\right) + 2\text{Li}_4\left(\frac{1-x}{2}\right) \\ &\quad + 2\text{Li}_4\left(\frac{x-1}{x+1}\right) - 2\text{Li}_4\left(\frac{1+x}{2}\right) + \frac{11}{480} \pi^4, \end{aligned}$$

$$\begin{aligned}
 H(0, -1, -1, -1; x) &= -\frac{1}{2}\text{Li}_2(-x)\log^2(1+x) - \text{Li}_3\left(\frac{1}{1+x}\right)\log(1+x) + \frac{1}{8}\log^4(1+x) \\
 &\quad - \frac{1}{3}\log x \log^3(1+x) - \frac{\pi^2}{12}\log^2(1+x) - \text{Li}_4\left(\frac{1}{1+x}\right) + \frac{\pi^4}{90}, \\
 H(0, -1, -1, 1; x) &= \text{Li}_{2,2}(-1, x) + \text{Li}_{2,2}\left(\frac{1}{2}, \frac{2x}{x+1}\right) + \frac{1}{2}\text{Li}_2\left(\frac{1-x}{2}\right)\log^2(1+x) \\
 &\quad - \frac{1}{2}\text{Li}_2(-x)\log^2(1+x) + \frac{1}{2}\text{Li}_2(x)\log^2(1+x) + \frac{1}{2}\text{Li}_2(-x)\log^2 2 \\
 &\quad + \text{Li}_2(-x)\log(1-x)\log(1+x) - 2\text{Li}_3(x)\log(1+x) \\
 &\quad - 2\text{Li}_3\left(\frac{2x}{x-1}\right)\log(1+x) - 2\text{Li}_3\left(\frac{1-x}{1+x}\right)\log(1+x) \\
 &\quad - \text{Li}_2(-x)\log 2\log(1-x) + \text{Li}_3\left(\frac{1}{1+x}\right)\log(1-x) - \frac{5}{4}\zeta_3\log(1+x) \\
 &\quad - \frac{7}{8}\zeta_3\log(1-x) + \frac{5}{8}\log^4(1+x) - \frac{3}{2}\log 2\log^3(1+x) \\
 &\quad - \frac{2}{3}\log(1-x)\log^3(1+x) - \log x \log^3(1+x) - \frac{1}{2}\log^3 2\log(1+x) \\
 &\quad + \frac{1}{3}\log^3(1-x)\log(1+x) + \log^2 2\log^2(1+x) + \frac{3}{2}\log 2\log(1-x)\log^2(1+x) \\
 &\quad + 2\log(1-x)\log x \log^2(1+x) - \frac{5}{12}\pi^2\log^2(1+x) \\
 &\quad - \log 2\log^2(1-x)\log(1+x) - \log^2(1-x)\log x \log(1+x) \\
 &\quad + \frac{\pi^2}{4}\log 2\log(1+x) + \frac{5}{12}\pi^2\log(1-x)\log(1+x) - \frac{1}{4}\text{Li}_4(1-x^2) \\
 &\quad - \frac{1}{2}\text{Li}_2(-x)^2 + \text{Li}_2\left(\frac{1-x}{2}\right)\text{Li}_2(-x) - \frac{\pi^2}{12}\text{Li}_2(-x) + \text{Li}_4(1-x) \\
 &\quad + \text{Li}_4(-x) + \text{Li}_4(x) + \frac{1}{2}\text{Li}_4\left(\frac{4x}{(x+1)^2}\right) - \frac{1}{2}\text{Li}_4\left(\frac{1-x}{1+x}\right) + \frac{1}{2}\text{Li}_4\left(\frac{x-1}{x+1}\right) \\
 &\quad + 2\text{Li}_4\left(\frac{x}{x+1}\right) - 2\text{Li}_4\left(\frac{2x}{x+1}\right) + 3\text{Li}_4\left(\frac{1+x}{2}\right) - 3\text{Li}_4\left(\frac{1}{2}\right) + \frac{\pi^4}{480}, \\
 H(0, -1, 1, -1; x) &= -\text{Li}_{2,2}(-1, x) - \text{Li}_{2,2}\left(\frac{1}{2}, \frac{2x}{x+1}\right) - \frac{1}{2}\text{Li}_2\left(\frac{1-x}{2}\right)\log^2(1+x) \\
 &\quad + \frac{1}{2}\text{Li}_2(-x)\log^2(1+x) - \frac{1}{2}\text{Li}_2(x)\log^2(1+x) - \frac{1}{2}\text{Li}_2(-x)\log^2 2 \\
 &\quad - \text{Li}_3\left(\frac{1-x}{2}\right)\log(1+x) + \text{Li}_3(1-x)\log(1+x) - \text{Li}_3(-x)\log(1+x) \\
 &\quad + 3\text{Li}_3(x)\log(1+x) + 3\text{Li}_3\left(\frac{2x}{x-1}\right)\log(1+x) \\
 &\quad + 2\text{Li}_3\left(\frac{1-x}{1+x}\right)\log(1+x) + \text{Li}_2(-x)\log 2\log(1-x) + \frac{17}{8}\zeta_3\log(1+x) \\
 &\quad - \frac{19}{24}\log^4(1+x) + 2\log 2\log^3(1+x) + \frac{1}{2}\log(1-x)\log^3(1+x) \\
 &\quad + \frac{7}{6}\log x \log^3(1+x) + \frac{7}{6}\log^3 2\log(1+x) - \frac{1}{2}\log^3(1-x)\log(1+x) \\
 &\quad - \frac{7}{4}\log^2 2\log^2(1+x) - \frac{3}{2}\log 2\log(1-x)\log^2(1+x) \\
 &\quad - 2\log(1-x)\log x \log^2(1+x) + \frac{17}{24}\pi^2\log^2(1+x) \\
 &\quad + \frac{3}{2}\log 2\log^2(1-x)\log(1+x) - \frac{1}{2}\log^2 2\log(1-x)\log(1+x)
 \end{aligned}$$

$$\begin{aligned}
& + \frac{3}{2} \log^2(1-x) \log x \log(1+x) - \frac{7}{12} \pi^2 \log 2 \log(1+x) \\
& - \frac{5}{12} \pi^2 \log(1-x) \log(1+x) + \frac{1}{2} \text{Li}_2(-x)^2 - \text{Li}_2\left(\frac{1-x}{2}\right) \text{Li}_2(-x) \\
& + \frac{\pi^2}{12} \text{Li}_2(-x) - \frac{1}{2} \text{Li}_4(-x) - \frac{3}{2} \text{Li}_4(x) - \frac{3}{4} \text{Li}_4\left(\frac{4x}{(x+1)^2}\right) \\
& - \text{Li}_4\left(\frac{1}{1+x}\right) - 2\text{Li}_4\left(\frac{x}{x+1}\right) + 3\text{Li}_4\left(\frac{2x}{x+1}\right) - 6\text{Li}_4\left(\frac{1+x}{2}\right) \\
& + 6\text{Li}_4\left(\frac{1}{2}\right) + \frac{\pi^4}{90},
\end{aligned}$$

$$\begin{aligned}
H(0, -1, 1, 1; x) = & -\frac{1}{2} \text{Li}_2(-x) \log^2(1-x) + \text{Li}_3\left(\frac{1-x}{2}\right) \log(1-x) - \text{Li}_3(1-x) \log(1-x) \\
& + \text{Li}_3(-x) \log(1-x) - \text{Li}_3(x) \log(1-x) - \text{Li}_3\left(\frac{2x}{x-1}\right) \log(1-x) \\
& + \frac{7}{4} \zeta_3 \log(1-x) + \frac{19}{96} \log^4(1-x) + \frac{23}{96} \log^4(1+x) - \frac{1}{3} \log 2 \log^3(1-x) \\
& - \frac{7}{12} \log x \log^3(1-x) - \frac{1}{24} \log(1+x) \log^3(1-x) - \frac{1}{24} \log^3(1+x) \log(1-x) \\
& - \frac{1}{3} \log 2 \log^3(1+x) - \frac{1}{4} \log x \log^3(1+x) - \frac{1}{3} \log^3 2 \log(1+x) \\
& + \frac{1}{4} \log^2 2 \log^2(1-x) - \frac{1}{16} \log^2(1+x) \log^2(1-x) + \frac{1}{4} \log x \log(1+x) \log^2(1-x) \\
& + \frac{3}{16} \pi^2 \log^2(1-x) + \frac{1}{4} \log x \log^2(1+x) \log(1-x) + \frac{1}{2} \log^2 2 \log^2(1+x) \\
& - \frac{13}{48} \pi^2 \log^2(1+x) - \frac{\pi^2}{24} \log(1+x) \log(1-x) + \frac{\pi^2}{6} \log 2 \log(1+x) \\
& + \frac{1}{4} \text{Li}_4(1-x^2) - \frac{1}{4} \text{Li}_4\left(\frac{x^2}{x^2-1}\right) - \text{Li}_4\left(\frac{1-x}{2}\right) - \text{Li}_4(1-x) \\
& - \frac{1}{2} \text{Li}_4(-x) + \frac{1}{2} \text{Li}_4(x) + 2\text{Li}_4\left(\frac{x}{x-1}\right) - \text{Li}_4\left(\frac{2x}{x-1}\right) + \frac{1}{4} \text{Li}_4\left(\frac{4x}{(x+1)^2}\right) \\
& + 2\text{Li}_4\left(\frac{1}{1+x}\right) + 2\text{Li}_4\left(\frac{x}{x+1}\right) - 2\text{Li}_4\left(\frac{2x}{x+1}\right) + 2\text{Li}_4\left(\frac{1+x}{2}\right) \\
& - \text{Li}_4\left(\frac{1}{2}\right) - \frac{\pi^4}{72},
\end{aligned}$$

$$\begin{aligned}
H(0, 0, -1, -1; x) = & -\text{Li}_3(-x) \log(1+x) + \zeta_3 \log(1+x) + \frac{1}{12} \log^4(1+x) - \frac{1}{6} \log x \log^3(1+x) \\
& - \frac{\pi^2}{12} \log^2(1+x) + \text{Li}_4(-x) + \text{Li}_4\left(\frac{1}{1+x}\right) + \text{Li}_4\left(\frac{x}{x+1}\right) - \frac{\pi^4}{90},
\end{aligned}$$

$$\begin{aligned}
H(0, 0, -1, 1; x) = & \text{Li}_3(-x) \log(1-x) + \frac{3}{4} \zeta_3 \log(1-x) + \frac{1}{32} \log^4(1-x) + \frac{23}{96} \log^4(1+x) \\
& - \frac{1}{12} \log x \log^3(1-x) - \frac{1}{24} \log(1+x) \log^3(1-x) - \frac{1}{24} \log^3(1+x) \log(1-x) \\
& - \frac{1}{3} \log 2 \log^3(1+x) - \frac{1}{4} \log x \log^3(1+x) - \frac{1}{3} \log^3 2 \log(1+x) \\
& - \frac{1}{16} \log^2(1+x) \log^2(1-x) + \frac{1}{4} \log x \log(1+x) \log^2(1-x) + \frac{\pi^2}{16} \log^2(1-x) \\
& + \frac{1}{4} \log x \log^2(1+x) \log(1-x) + \frac{1}{2} \log^2 2 \log^2(1+x) - \frac{13}{48} \pi^2 \log^2(1+x) \\
& - \frac{\pi^2}{24} \log(1+x) \log(1-x) + \frac{\pi^2}{6} \log 2 \log(1+x) + \frac{1}{4} \text{Li}_4(1-x^2)
\end{aligned}$$

$$\begin{aligned}
& -\frac{1}{4}\text{Li}_4\left(\frac{x^2}{x^2-1}\right) - \text{Li}_4(1-x) - \frac{3}{2}\text{Li}_4(-x) + \frac{1}{2}\text{Li}_4(x) + \text{Li}_4\left(\frac{x}{x-1}\right) \\
& + \frac{1}{4}\text{Li}_4\left(\frac{4x}{(x+1)^2}\right) + 2\text{Li}_4\left(\frac{1}{1+x}\right) + 2\text{Li}_4\left(\frac{x}{x+1}\right) - 2\text{Li}_4\left(\frac{2x}{x+1}\right) \\
& + 2\text{Li}_4\left(\frac{1+x}{2}\right) - 2\text{Li}_4\left(\frac{1}{2}\right) - \frac{\pi^4}{72},
\end{aligned}$$

$$H(0, 0, 0, -1; x) = -\text{Li}_4(-x),$$

$$H(0, 0, 0, 1; x) = \text{Li}_4(x),$$

$$\begin{aligned}
H(0, 0, 1, -1; x) &= \text{Li}_3(x) \log(1+x) + \frac{3}{4}\zeta_3 \log(1+x) - \frac{1}{6} \log^4(1+x) + \frac{1}{3} \log 2 \log^3(1+x) \\
& + \frac{1}{6} \log x \log^3(1+x) + \frac{1}{3} \log^3 2 \log(1+x) - \frac{1}{2} \log^2 2 \log^2(1+x) + \frac{\pi^2}{6} \log^2(1+x) \\
& - \frac{\pi^2}{6} \log 2 \log(1+x) + \frac{1}{2}\text{Li}_4(-x) - \frac{3}{2}\text{Li}_4(x) - \frac{1}{4}\text{Li}_4\left(\frac{4x}{(x+1)^2}\right) - \text{Li}_4\left(\frac{1}{1+x}\right) \\
& - \text{Li}_4\left(\frac{x}{x+1}\right) + 2\text{Li}_4\left(\frac{2x}{x+1}\right) - 2\text{Li}_4\left(\frac{1+x}{2}\right) + 2\text{Li}_4\left(\frac{1}{2}\right) + \frac{\pi^4}{90},
\end{aligned}$$

$$\begin{aligned}
H(0, 1, -1, -1; x) &= \frac{1}{2}\text{Li}_2(x) \log^2(1+x) + \text{Li}_3(-x) \log(1+x) - \text{Li}_3(x) \log(1+x) \\
& - \text{Li}_3\left(\frac{2x}{x-1}\right) \log(1+x) + \text{Li}_3\left(\frac{1}{1+x}\right) \log(1+x) - \text{Li}_3\left(\frac{1-x}{1+x}\right) \log(1+x) \\
& - \text{Li}_3\left(\frac{1+x}{2}\right) \log(1+x) - \frac{3}{4}\zeta_3 \log(1+x) + \frac{1}{6} \log^4(1+x) - \frac{1}{2} \log 2 \log^3(1+x) \\
& - \frac{1}{2} \log(1-x) \log^3(1+x) - \frac{1}{6} \log x \log^3(1+x) - \frac{1}{3} \log^3 2 \log(1+x) \\
& + \frac{1}{6} \log^3(1-x) \log(1+x) + \frac{1}{4} \log^2 2 \log^2(1+x) + \log 2 \log(1-x) \log^2(1+x) \\
& + \log(1-x) \log x \log^2(1+x) - \frac{\pi^2}{8} \log^2(1+x) - \frac{1}{2} \log 2 \log^2(1-x) \log(1+x) \\
& - \frac{1}{2} \log^2(1-x) \log x \log(1+x) + \frac{\pi^2}{6} \log 2 \log(1+x) + \frac{\pi^2}{6} \log(1-x) \log(1+x) \\
& - \frac{1}{2}\text{Li}_4(-x) + \frac{1}{2}\text{Li}_4(x) + \frac{1}{4}\text{Li}_4\left(\frac{4x}{(x+1)^2}\right) + \text{Li}_4\left(\frac{1}{1+x}\right) - \text{Li}_4\left(\frac{2x}{x+1}\right) \\
& + 3\text{Li}_4\left(\frac{1+x}{2}\right) - 3\text{Li}_4\left(\frac{1}{2}\right) - \frac{\pi^4}{90},
\end{aligned}$$

$$\begin{aligned}
H(0, 1, -1, 1; x) &= -\frac{47}{96} \log^4(1-x) + \log 2 \log^3(1-x) + \frac{11}{12} \log x \log^3(1-x) \\
& + \frac{1}{24} \log(1+x) \log^3(1-x) - \frac{1}{4} \log^2 2 \log^2(1-x) + \frac{9}{16} \log^2(1+x) \log^2(1-x) \\
& - \log 2 \log(1+x) \log^2(1-x) - \frac{5}{4} \log x \log(1+x) \log^2(1-x) \\
& - \frac{1}{2}\text{Li}_2\left(\frac{1-x}{2}\right) \log^2(1-x) + \frac{1}{2}\text{Li}_2(-x) \log^2(1-x) - \frac{1}{2}\text{Li}_2(x) \log^2(1-x) \\
& - \frac{17}{48} \pi^2 \log^2(1-x) - \frac{1}{6} \log^3 2 \log(1-x) + \frac{1}{24} \log^3(1+x) \log(1-x) \\
& - \frac{1}{4} \log x \log^2(1+x) \log(1-x) + \frac{\pi^2}{12} \log 2 \log(1-x) \\
& + \frac{1}{2} \log^2 2 \log(1+x) \log(1-x) - \frac{\pi^2}{24} \log(1+x) \log(1-x) \\
& + \log 2 \text{Li}_2(x) \log(1-x) - \log(1+x) \text{Li}_2(x) \log(1-x) - \text{Li}_3(-x) \log(1-x)
\end{aligned}$$

$$\begin{aligned}
& + \text{Li}_3(x) \log(1-x) + 3\text{Li}_3\left(\frac{2x}{x-1}\right) \log(1-x) - \text{Li}_3\left(\frac{1}{1+x}\right) \log(1-x) \\
& + \text{Li}_3\left(\frac{1-x}{1+x}\right) \log(1-x) + \text{Li}_3\left(\frac{1+x}{2}\right) \log(1-x) - \frac{29}{8}\zeta_3 \log(1-x) \\
& - \frac{55}{96} \log^4(1+x) + \log 2 \log^3(1+x) + \frac{7}{12} \log x \log^3(1+x) - \frac{3}{2} \log^2 2 \log^2(1+x) \\
& + \frac{29}{48} \pi^2 \log^2(1+x) - \frac{\text{Li}_2(x)^2}{2} - \text{Li}_{2,2}(-1, x) + \text{Li}_{2,2}\left(\frac{1}{2}, \frac{2x}{x-1}\right) \\
& + \log^3 2 \log(1+x) - \frac{\pi^2}{2} \log 2 \log(1+x) - \frac{1}{2} \log^2 2 \text{Li}_2(x) - \text{Li}_2\left(\frac{1-x}{2}\right) \text{Li}_2(x) \\
& + \text{Li}_2(-x) \text{Li}_2(x) + \frac{\pi^2}{12} \text{Li}_2(x) + 2 \log(1+x) \text{Li}_3(x) + 3\text{Li}_4(1-x) \\
& - \frac{1}{2} \text{Li}_4(-x) - \frac{3}{2} \text{Li}_4(x) - 2\text{Li}_4\left(\frac{x}{x-1}\right) + 3\text{Li}_4\left(\frac{2x}{x-1}\right) \\
& - \frac{3}{4} \text{Li}_4\left(\frac{4x}{(x+1)^2}\right) - 4\text{Li}_4\left(\frac{1}{1+x}\right) - 4\text{Li}_4\left(\frac{x}{x+1}\right) + 6\text{Li}_4\left(\frac{2x}{x+1}\right) \\
& - 6\text{Li}_4\left(\frac{1+x}{2}\right) - \frac{1}{4} \text{Li}_4(1-x^2) + \frac{1}{4} \text{Li}_4\left(\frac{x^2}{x^2-1}\right) + \frac{3}{2} \log(1+x) \zeta_3 \\
& + 6\text{Li}_4\left(\frac{1}{2}\right) + \frac{\pi^4}{72},
\end{aligned}$$

$$H(0, 1, 0, -1; x) = -\text{Li}_{2,2}(-1, x),$$

$$\begin{aligned}
H(0, 1, 0, 1; x) & = 2\text{Li}_3(x) \log(1-x) - 2\zeta_3 \log(1-x) - \frac{1}{12} \log^4(1-x) + \frac{1}{3} \log x \log^3(1-x) \\
& - \frac{\pi^2}{6} \log^2(1-x) + \frac{\text{Li}_2(x)^2}{2} + 2\text{Li}_4(1-x) - 2\text{Li}_4(x) - 2\text{Li}_4\left(\frac{x}{x-1}\right) - \frac{\pi^4}{45},
\end{aligned}$$

$$\begin{aligned}
H(0, 1, 1, -1; x) & = \text{Li}_{2,2}(-1, x) - \text{Li}_{2,2}\left(\frac{1}{2}, \frac{2x}{x-1}\right) + \frac{1}{2} \text{Li}_2\left(\frac{1-x}{2}\right) \log^2(1-x) \\
& - \frac{1}{2} \text{Li}_2(-x) \log^2(1-x) + \frac{1}{2} \text{Li}_2(x) \log^2(1-x) + \frac{1}{2} \text{Li}_2(x) \log^2 2 \\
& - \text{Li}_2(x) \log 2 \log(1-x) - 2\text{Li}_3\left(\frac{2x}{x-1}\right) \log(1-x) - \text{Li}_3(1-x) \log(1+x) \\
& - 2\text{Li}_3(x) \log(1+x) + \frac{7}{4} \zeta_3 \log(1-x) - \frac{5}{8} \zeta_3 \log(1+x) + \frac{7}{24} \log^4(1-x) \\
& + \frac{3}{8} \log^4(1+x) - \frac{2}{3} \log 2 \log^3(1-x) - \frac{1}{3} \log x \log^3(1-x) - \frac{1}{6} \log^3 2 \log(1-x) \\
& - \frac{2}{3} \log 2 \log^3(1+x) - \frac{1}{3} \log x \log^3(1+x) - \frac{2}{3} \log^3 2 \log(1+x) \\
& + \frac{1}{2} \log^2 2 \log^2(1-x) + \frac{\pi^2}{8} \log^2(1-x) + \log^2 2 \log^2(1+x) - \frac{3}{8} \pi^2 \log^2(1+x) \\
& + \frac{\pi^2}{12} \log 2 \log(1-x) + \frac{\pi^2}{3} \log 2 \log(1+x) + \frac{1}{4} \text{Li}_4(1-x^2) + \frac{1}{2} \text{Li}_2(x)^2 \\
& + \text{Li}_2\left(\frac{1-x}{2}\right) \text{Li}_2(x) - \text{Li}_2(-x) \text{Li}_2(x) - \frac{\pi^2}{12} \text{Li}_2(x) + \text{Li}_4\left(\frac{1-x}{2}\right) \\
& - 2\text{Li}_4(1-x) + \text{Li}_4(-x) + \text{Li}_4(x) - 2\text{Li}_4\left(\frac{2x}{x-1}\right) + \frac{1}{2} \text{Li}_4\left(\frac{4x}{(x+1)^2}\right) \\
& + 3\text{Li}_4\left(\frac{1}{1+x}\right) - \frac{1}{2} \text{Li}_4\left(\frac{1-x}{1+x}\right) + \frac{1}{2} \text{Li}_4\left(\frac{x-1}{x+1}\right) + 2\text{Li}_4\left(\frac{x}{x+1}\right) \\
& - 4\text{Li}_4\left(\frac{2x}{x+1}\right) + 4\text{Li}_4\left(\frac{1+x}{2}\right) - 5\text{Li}_4\left(\frac{1}{2}\right) - \frac{\pi^4}{288},
\end{aligned}$$

$$\begin{aligned}
H(0, 1, 1, 1; x) &= \frac{1}{2} \text{Li}_2(x) \log^2(1-x) + \text{Li}_3(1-x) \log(1-x) + \frac{1}{3} \log x \log^3(1-x) \\
&\quad - \frac{\pi^2}{12} \log^2(1-x) - \text{Li}_4(1-x) + \frac{\pi^4}{90},
\end{aligned}$$

References

- [1] A.B. Goncharov, *Multiple polylogarithms, cyclotomy and modular complexes*, *Math. Res. Lett.* **5** (1998) 497 [[arXiv:1105.2076](#)].
- [2] A.B. Goncharov, *Multiple polylogarithms and mixed Tate motives*, [math/0103059](#).
- [3] S. Bloch, *Higher regulators, algebraic K-theory and zeta functions of elliptic curves*, Volume 11 of CRM Monograph Series, AMS, Providence, RI, U.S.A. (2000).
- [4] A. Beilinson, *Polylogarithms and cyclotomic elements*, MIT preprint (1989).
- [5] D. Zagier, *Polylogarithms, Dedekind zeta functions and the algebraic K-theory of fields*, in G. van der Geer, F. Oort, J. Steenbrink eds., *Arithmetic Algebraic Geometry*, *Prog. Math.* **89** Birkhäuser (1991), pg. 391–430.
- [6] A. Beilinson and P. Deligne, *Interprétation motivique de la conjecture de Zagier reliant polylogarithmes et régulateurs*, *Proc. Symp. Pure Math.* **55** (1994) 97.
- [7] A.B. Goncharov, *Geometry of configurations, polylogarithms and motivic cohomology*, *Adv. Math.* **144** (1995) 197.
- [8] A.B. Goncharov, *Galois symmetries of fundamental groupoids and noncommutative geometry*, *Duke Math. J.* **128** (2005) 209 [[math/0208144](#)].
- [9] J.L. Dupont and C.H. Sah, *Scissors Congruences II*, *J. Pure Appl. Algebra* **25** (1982) 159.
- [10] J.L. Cathelineau, *Quelques aspects du troisième problème de Hilbert*, *Gaz. Math.* **52** (1992) 45.
- [11] A.B. Goncharov, *Volumes of hyperbolic manifolds and mixed Tate motives*, *J. Amer. Math. Soc.* **12** (1999) 569 [[alg-geom/9601021](#)].
- [12] J. Böhm, *Inhaltsmessung in konstanter Krümmung*, *Arch. Math.* **11** (1960) 298.
- [13] W.D. Neumann and D. Zagier, *Volumes of hyperbolic three-manifolds*, *Topology* **24** (1985) 307.
- [14] D. Zagier, *Hyperbolic manifolds and special values of Dedekind zeta-functions*, *Invent. Math.* **83** (1986) 285.
- [15] R. Kellerhals, *Volumes in hyperbolic 5-space*, *Geom. Funct. Anal.* **5** (1995) 640.
- [16] W.D. Neumann and J. Yang, *Bloch invariants of hyperbolic 3-manifolds*, *Duke Math. J.* **96** (1999) 29. [math/9712224](#).
- [17] A.M. Gabriellov, I.M. Gelfand and M.V. Losik, *Combinatorial computation of characteristic classes*, *Funct. Anal. Appl.* **9** (1975) 5.
- [18] D. Zagier, *The dilogarithm function in Geometry and Number Theory*, in P.E. Cartier, B. Julia and P. Moussa, P. Vanhove eds., *Frontiers in number theory, physics, and geometry II*, Springer (2007), pg. 3–65
- [19] S. Bloch and I. Kriz, *Mixed Tate Motives*, *Ann. Math.* **140** (1994) 557.
- [20] H. Gangl and S. Mueller-Stach, *Polylogarithmic identities in cubical higher Chow groups*, in *Algebraic K-Theory. Proceedings of Symposia in Pure Mathematics* **67**, AMS, Providence, RI, U.S.A (1999) 25.

- [21] H. Gangl, A.B. Goncharov and A. Levin, *Multiple polylogarithms, polygons, trees and algebraic cycles*, in proceedings of *Summer Institute in Algebraic Geometry*, Seattle, U.S.A. (2005) [*Proc. Symp. Pure Math.* **80** (2009) 547] [[math/0508066](#)].
- [22] K.T. Chen, *Iterated path integrals*, *Bull. Amer. Math. Soc.* **83** (1977) 831.
- [23] R. Hain, *Classical polylogarithms*, in U. Jannsen, S. Kleiman and J.-P. Serre eds., *Motives, Proc. Symp. Pure Math.* **55** (1994) 3 [[alg-geom/9202022](#)].
- [24] Z. Wojtkowiak, *Mixed Hodge structures and iterated integrals I*, in F. Bogomolov and L. Katzarkov eds. *Motives, Polylogarithms and Hodge Theory (Part I: Motives and Polylogarithms)*, *Int. Press Lect. Ser.* **3** (2002) 121.
- [25] E. Remiddi and J. Vermaseren, *Harmonic polylogarithms*, *Int. J. Mod. Phys. A* **15** (2000) 725 [[hep-ph/9905237](#)] [[INSPIRE](#)].
- [26] T. Gehrmann and E. Remiddi, *Two loop master integrals for $\gamma^* \rightarrow 3$ jets: The Planar topologies*, *Nucl. Phys. B* **601** (2001) 248 [[hep-ph/0008287](#)] [[INSPIRE](#)].
- [27] J. Ablinger, J. Blumlein and C. Schneider, *Harmonic Sums and Polylogarithms Generated by Cyclotomic Polynomials*, *J. Math. Phys.* **52** (2011) 102301 [[arXiv:1105.6063](#)] [[INSPIRE](#)].
- [28] J. Vermaseren, A. Vogt and S. Moch, *The Third-order QCD corrections to deep-inelastic scattering by photon exchange*, *Nucl. Phys. B* **724** (2005) 3 [[hep-ph/0504242](#)] [[INSPIRE](#)].
- [29] S. Moch, J. Vermaseren and A. Vogt, *The Longitudinal structure function at the third order*, *Phys. Lett. B* **606** (2005) 123 [[hep-ph/0411112](#)] [[INSPIRE](#)].
- [30] A. Vogt, S. Moch and J. Vermaseren, *The Three-loop splitting functions in QCD: The Singlet case*, *Nucl. Phys. B* **691** (2004) 129 [[hep-ph/0404111](#)] [[INSPIRE](#)].
- [31] S. Moch, J. Vermaseren and A. Vogt, *The Three loop splitting functions in QCD: The Nonsinglet case*, *Nucl. Phys. B* **688** (2004) 101 [[hep-ph/0403192](#)] [[INSPIRE](#)].
- [32] R. Bonciani, P. Mastrolia and E. Remiddi, *Master integrals for the two loop QCD virtual corrections to the forward backward asymmetry*, *Nucl. Phys. B* **690** (2004) 138 [[hep-ph/0311145](#)] [[INSPIRE](#)].
- [33] W. Bernreuther, R. Bonciani, T. Gehrmann, R. Heinesch, T. Leineweber, P. Mastrolia and E. Remiddi, *Two-loop QCD corrections to the heavy quark form-factors: The Vector contributions*, *Nucl. Phys. B* **706** (2005) 245 [[hep-ph/0406046](#)] [[INSPIRE](#)].
- [34] W. Bernreuther, R. Bonciani, T. Gehrmann, R. Heinesch, T. Leineweber, P. Mastrolia and E. Remiddi, *Two-loop QCD corrections to the heavy quark form-factors: Axial vector contributions*, *Nucl. Phys. B* **712** (2005) 229 [[hep-ph/0412259](#)] [[INSPIRE](#)].
- [35] W. Bernreuther, R. Bonciani, T. Gehrmann, R. Heinesch, T. Leineweber, E. Remiddi, *Two-loop QCD corrections to the heavy quark form-factors: Anomaly contributions*, *Nucl. Phys. B* **723** (2005) 91 [[hep-ph/0504190](#)] [[INSPIRE](#)].
- [36] P. Mastrolia and E. Remiddi, *Two loop form-factors in QED*, *Nucl. Phys. B* **664** (2003) 341 [[hep-ph/0302162](#)] [[INSPIRE](#)].
- [37] R. Bonciani, A. Ferroglia, P. Mastrolia, E. Remiddi and J. van der Bij, *Two-loop $N(F) = 1$ QED Bhabha scattering: Soft emission and numerical evaluation of the differential cross-section*, *Nucl. Phys. B* **716** (2005) 280 [[hep-ph/0411321](#)] [[INSPIRE](#)].
- [38] R. Bonciani, A. Ferroglia, P. Mastrolia, E. Remiddi and J. van der Bij, *Two-loop $N(F) = 1$ QED Bhabha scattering differential cross section*, *Nucl. Phys. B* **701** (2004) 121 [[hep-ph/0405275](#)] [[INSPIRE](#)].

- [39] M. Czakon, J. Gluza and T. Riemann, *Master integrals for massive two-loop bhabha scattering in QED*, *Phys. Rev. D* **71** (2005) 073009 [[hep-ph/0412164](#)] [[INSPIRE](#)].
- [40] Z. Bern, M. Czakon, L.J. Dixon, D.A. Kosower and V.A. Smirnov, *The Four-Loop Planar Amplitude and Cusp Anomalous Dimension in Maximally Supersymmetric Yang-Mills Theory*, *Phys. Rev. D* **75** (2007) 085010 [[hep-th/0610248](#)] [[INSPIRE](#)].
- [41] G. Heinrich and V.A. Smirnov, *Analytical evaluation of dimensionally regularized massive on-shell double boxes*, *Phys. Lett. B* **598** (2004) 55 [[hep-ph/0406053](#)] [[INSPIRE](#)].
- [42] V.A. Smirnov, *Analytical result for dimensionally regularized massive on-shell planar double box*, *Phys. Lett. B* **524** (2002) 129 [[hep-ph/0111160](#)] [[INSPIRE](#)].
- [43] L. Bork, D. Kazakov and G. Vartanov, *On form factors in $N = 4$ SYM*, *JHEP* **02** (2011) 063 [[arXiv:1011.2440](#)] [[INSPIRE](#)].
- [44] J.M. Henn, S.G. Naculich, H.J. Schnitzer and M. Spradlin, *More loops and legs in Higgs-regulated $N = 4$ SYM amplitudes*, *JHEP* **08** (2010) 002 [[arXiv:1004.5381](#)] [[INSPIRE](#)].
- [45] U. Aglietti, R. Bonciani, G. Degrassi and A. Vicini, *Analytic Results for Virtual QCD Corrections to Higgs Production and Decay*, *JHEP* **01** (2007) 021 [[hep-ph/0611266](#)] [[INSPIRE](#)].
- [46] U. Aglietti, R. Bonciani, G. Degrassi and A. Vicini, *Master integrals for the two-loop light fermion contributions to $gg \rightarrow H$ and $H \rightarrow \gamma\gamma$* , *Phys. Lett. B* **600** (2004) 57 [[hep-ph/0407162](#)] [[INSPIRE](#)].
- [47] U. Aglietti, R. Bonciani, G. Degrassi and A. Vicini, *Two loop light fermion contribution to Higgs production and decays*, *Phys. Lett. B* **595** (2004) 432 [[hep-ph/0404071](#)] [[INSPIRE](#)].
- [48] T. Gehrmann and E. Remiddi, *Two loop master integrals for $\gamma^* \rightarrow 3$ jets: The Nonplanar topologies*, *Nucl. Phys. B* **601** (2001) 287 [[hep-ph/0101124](#)] [[INSPIRE](#)].
- [49] C. Anastasiou, S. Beerli, S. Bucherer, A. Daleo and Z. Kunszt, *Two-loop amplitudes and master integrals for the production of a Higgs boson via a massive quark and a scalar-quark loop*, *JHEP* **01** (2007) 082 [[hep-ph/0611236](#)] [[INSPIRE](#)].
- [50] S. Moch, P. Uwer and S. Weinzierl, *Two loop amplitudes with nested sums: Fermionic contributions to $e^+ e^- \rightarrow q \bar{q} g$* , *Phys. Rev. D* **66** (2002) 114001 [[hep-ph/0207043](#)] [[INSPIRE](#)].
- [51] S. Moch, P. Uwer and S. Weinzierl, *Two loop amplitudes for $e^+ e^- \rightarrow q \bar{q} g$: The $n(f)$ contribution*, *Acta Phys. Polon. B* **33** (2002) 2921 [[hep-ph/0207167](#)] [[INSPIRE](#)].
- [52] U. Aglietti, V. Del Duca, C. Duhr, G. Somogyi and Z. Trócsányi, *Analytic integration of real-virtual counterterms in NNLO jet cross sections. I.*, *JHEP* **09** (2008) 107 [[arXiv:0807.0514](#)] [[INSPIRE](#)].
- [53] V. Del Duca, C. Duhr, E. Nigel Glover and V.A. Smirnov, *The One-loop pentagon to higher orders in epsilon*, *JHEP* **01** (2010) 042 [[arXiv:0905.0097](#)] [[INSPIRE](#)].
- [54] T. Gehrmann and E. Remiddi, *Numerical evaluation of harmonic polylogarithms*, *Comput. Phys. Commun.* **141** (2001) 296 [[hep-ph/0107173](#)] [[INSPIRE](#)].
- [55] D. Maître, *HPL, a mathematica implementation of the harmonic polylogarithms*, *Comput. Phys. Commun.* **174** (2006) 222 [[hep-ph/0507152](#)] [[INSPIRE](#)].
- [56] D. Maître, *Extension of HPL to complex arguments*, *Comput. Phys. Commun.* **183** (2012) 846 [[hep-ph/0703052](#)] [[INSPIRE](#)].
- [57] J. Vollinga and S. Weinzierl, *Numerical evaluation of multiple polylogarithms*, *Comput. Phys. Commun.* **167** (2005) 177 [[hep-ph/0410259](#)] [[INSPIRE](#)].

- [58] A.I. Davydychev and M.Y. Kalmykov, *New results for the ϵ -expansion of certain one, two and three loop Feynman diagrams*, *Nucl. Phys. B* **605** (2001) 266 [[hep-th/0012189](#)] [[INSPIRE](#)].
- [59] A.I. Davydychev and M.Y. Kalmykov, *Massive Feynman diagrams and inverse binomial sums*, *Nucl. Phys. B* **699** (2004) 3 [[hep-th/0303162](#)] [[INSPIRE](#)].
- [60] D. Zagier, *Special Values and Functional Equations of Polylogarithms*, in L. Lewin ed., *Structural Properties of Polylogarithms*, AMS, Providence, RI, U.S.A. (1991), appendix A.
- [61] H. Gangl, *Functional equations of higher logarithms*, *Selecta Math.* **9** (2003) 361 [[math/0207222](#)].
- [62] A.B. Goncharov, *A simple construction of Grassmannian polylogarithms*, [arXiv:0908.2238](#).
- [63] A.B. Goncharov, M. Spradlin, C. Vergu and A. Volovich, *Classical Polylogarithms for Amplitudes and Wilson Loops*, *Phys. Rev. Lett.* **105** (2010) 151605 [[arXiv:1006.5703](#)] [[INSPIRE](#)].
- [64] V. Del Duca, C. Duhr and V.A. Smirnov, *An Analytic Result for the Two-Loop Hexagon Wilson Loop in $N = 4$ SYM*, *JHEP* **03** (2010) 099 [[arXiv:0911.5332](#)] [[INSPIRE](#)].
- [65] V. Del Duca, C. Duhr and V.A. Smirnov, *The Two-Loop Hexagon Wilson Loop in $N = 4$ SYM*, *JHEP* **05** (2010) 084 [[arXiv:1003.1702](#)] [[INSPIRE](#)].
- [66] S. Caron-Huot, *Superconformal symmetry and two-loop amplitudes in planar $N = 4$ super Yang-Mills*, *JHEP* **12** (2011) 066 [[arXiv:1105.5606](#)] [[INSPIRE](#)].
- [67] L.J. Dixon, J.M. Drummond and J.M. Henn, *Bootstrapping the three-loop hexagon*, *JHEP* **11** (2011) 023 [[arXiv:1108.4461](#)] [[INSPIRE](#)].
- [68] P. Heslop and V.V. Khoze, *Wilson Loops @ 3-Loops in Special Kinematics*, *JHEP* **11** (2011) 152 [[arXiv:1109.0058](#)] [[INSPIRE](#)].
- [69] L.F. Alday, D. Gaiotto, J. Maldacena, A. Sever and P. Vieira, *An Operator Product Expansion for Polygonal null Wilson Loops*, *JHEP* **04** (2011) 088 [[arXiv:1006.2788](#)] [[INSPIRE](#)].
- [70] D. Gaiotto, J. Maldacena, A. Sever and P. Vieira, *Pulling the straps of polygons*, *JHEP* **12** (2011) 011 [[arXiv:1102.0062](#)] [[INSPIRE](#)].
- [71] A.I. Davydychev and R. Delbourgo, *A Geometrical angle on Feynman integrals*, *J. Math. Phys.* **39** (1998) 4299 [[hep-th/9709216](#)] [[INSPIRE](#)].
- [72] M. Spradlin and A. Volovich, *Symbols of One-Loop Integrals From Mixed Tate Motives*, *JHEP* **11** (2011) 084 [[arXiv:1105.2024](#)] [[INSPIRE](#)].
- [73] L.J. Dixon, J.M. Drummond and J.M. Henn, *The one-loop six-dimensional hexagon integral and its relation to MHV amplitudes in $N = 4$ SYM*, *JHEP* **06** (2011) 100 [[arXiv:1104.2787](#)] [[INSPIRE](#)].
- [74] V. Del Duca, C. Duhr and V.A. Smirnov, *The massless hexagon integral in $D = 6$ dimensions*, *Phys. Lett. B* **703** (2011) 363 [[arXiv:1104.2781](#)] [[INSPIRE](#)].
- [75] V. Del Duca, C. Duhr and V.A. Smirnov, *The One-Loop One-Mass Hexagon Integral in $D = 6$ Dimensions*, *JHEP* **07** (2011) 064 [[arXiv:1105.1333](#)] [[INSPIRE](#)].
- [76] V. Del Duca, L.J. Dixon, J.M. Drummond, C. Duhr, J.M. Henn and V. A. Smirnov, *The one-loop six-dimensional hexagon integral with three massive corners*, *Phys. Rev. D* **84** (2011) 045017 [[arXiv:1105.2011](#)] [[INSPIRE](#)].

- [77] S. Buehler and C. Duhr, *CHAPLIN - Complex Harmonic Polylogarithms in Fortran*, [arXiv:1106.5739](#) [INSPIRE].
- [78] L.F. Alday, *Some analytic results for two-loop scattering amplitudes*, *JHEP* **07** (2011) 080 [[arXiv:1009.1110](#)] [INSPIRE].
- [79] J.A. Lappo-Danilevskij, *Mémoires sur la théorie des systèmes des équations différentielles linéaires. Vol. II, Travaux Inst. Physico-Math. Stekloff* **7** (1935) 5.
- [80] E.E. Kummer, *Über die Transcendenten, welche aus wiederholten Integrationen rationaler Formeln entstehen*, *J. Reine Angew. Math.* **21** (1840) 74.
- [81] R. Ree, *Lie elements and an algebra associated with shuffles*, *The Annals of Mathematics* **68** (1958) 210.
- [82] J.M. Borwein, D.M. Bradley, D.J. Broadhurst and P. Lisonek, *Special values of multiple polylogarithms*, *Trans. Am. Math. Soc.* **353** (2001) 907 [[math/9910045](#)] [INSPIRE].
- [83] N. Nielsen, *Der Eulersche Dilogarithmus und seine Verallgemeinerungen*, *Nova Acta Leopoldina* **90** (1909) 123.
- [84] L. Lewin, *Polylogarithms and associated functions*, North-Holland, New York (1981).
- [85] S. Bloch, H. Esnault and D. Kreimer, *On Motives associated to graph polynomials*, *Commun. Math. Phys.* **267** (2006) 181 [[math/0510011](#)] [INSPIRE].
- [86] S. Bloch and D. Kreimer, *Mixed Hodge Structures and Renormalization in Physics*, *Commun. Num. Theor. Phys.* **2** (2008) 637 [[arXiv:0804.4399](#)] [INSPIRE].
- [87] D. Broadhurst, P. Deligne, email correspondence (1997).
- [88] D.J. Broadhurst, *Massive three - loop Feynman diagrams reducible to SC^* primitives of algebras of the sixth root of unity*, *Eur. Phys. J. C* **8** (1999) 311 [[hep-th/9803091](#)] [INSPIRE].
- [89] P. Belkale and P. Brosnan, *Matroids, motives and conjecture of Kontsevich*, *Duke Math. J.* **116** (2003) 147 [[math/0012198](#)].
- [90] F. Brown, *The Massless higher-loop two-point function*, *Commun. Math. Phys.* **287** (2009) 925 [[arXiv:0804.1660](#)] [INSPIRE].
- [91] A. Connes and M. Marcolli, *Noncommutative geometry, quantum fields and motives*, AMS, Colloquium Publications (2008).
- [92] F.C.S. Brown, *Multiple zeta values and periods of moduli spaces $M_{0,n}$* , *Ann. Sci. Ecole. Norm. S.* **42** (2009) 371 [[math/0606419](#)].
- [93] A.B. Goncharov, *The dihedral Lie algebras and Galois symmetries of $\pi_1^{(1)}(\mathbf{P}^1 - (\{0, \infty\} \cup \mu_N))$* , *Duke Math. J.* **110** (2001) 397 [[math/0009121](#)].
- [94] S. Bloch, *Algebraic cycles and the Lie algebra of mixed Tate motives*, *J. Amer. Math. Soc.* **4** (1991) 771.
- [95] G. Griffing, *Dual Lie Elements and a Derivation for the Cofree Coassociative Coalgebra*, *P. Am. Math. Soc.* **123** (1995) 3269
- [96] N. Bourbaki, *Groupes et algèbres de Lie*, Hermann, Paris (1972), chapters 2 and 3.
- [97] J.R. Rhodes, *On the kernel of the symbol map for multiple polylogarithms*, Ph.D. Thesis, University of Durham (2012).