ON POLY(ANA)LOGS I

PHILIPPE ELBAZ-VINCENT AND HERBERT GANGL

Abstract. We investigate a connection between the differential of polylogarithms (as considered by Cathelineau) and a finite variant of them. This allows to answer a question raised by Kontsevich concerning the construction of functional equations for the finite analogs, using in part the p-adic version of polylogarithms and recent work of Besser. Kontsevich's original unpublished note is supplied (with his kind permission) in an "Appendix" at the end of the paper.

Contents

Introduction and Motivation

In an unpublished note [22] (included as an Appendix) Kontsevich defined the "$1\frac{1}{2}$-logarithm", associated to a prime p, as the truncated power series of $-\log(1-x)$ (for which we propose the "truncated" letter £, pronounced "sterling") as a function from $\mathbb{Z}/p$ to $\mathbb{Z}/p$:

$$£_1(x) = £_1^{(p)}(x) = \sum_{k=1}^{p-1} \frac{x^k}{k} \qquad (\text{mod } p).$$

Date: June 26, 2000.
The first author was partly supported by a Marie Curie fellowship of the EU.
The second author is supported by a Habilitationsstipendium der Deutschen Forschungsgemeinschaft.

For reasons which become apparent below we refer to it as the *finite 1-logarithm*. Kontsevich observed that it satisfies a functional equation which is known in the literature as the fundamental equation of information theory (see [1]), and provided a cohomological interpretation of the equation.

Cathelineau [8] was led to the same equation by considering an "infinitesimal" version of a one-valued cousin of the dilogarithm function which is defined over $\mathbb{C}$. He had encountered the fundamental equation of information theory already in [6] where, motivated by questions arising from Hilbert's third problem, he deduced an infinitesimal version of the famous Bloch–Suslin complex (which calculates certain algebraic K-groups of a field). Furthermore, he provided a homological interpretation of the equation. Cathelineau extended his results to infinitesimal versions of higher polylogarithms, and in particular—by mimicking Goncharov's setup [19] which generalizes the Bloch–Suslin complex—deduced an infinitesimal analogue of Goncharov's complexes. In the process, he produced the generic functional equation for the infinitesimal trilogarithm which contains 22 terms in 3 variables.

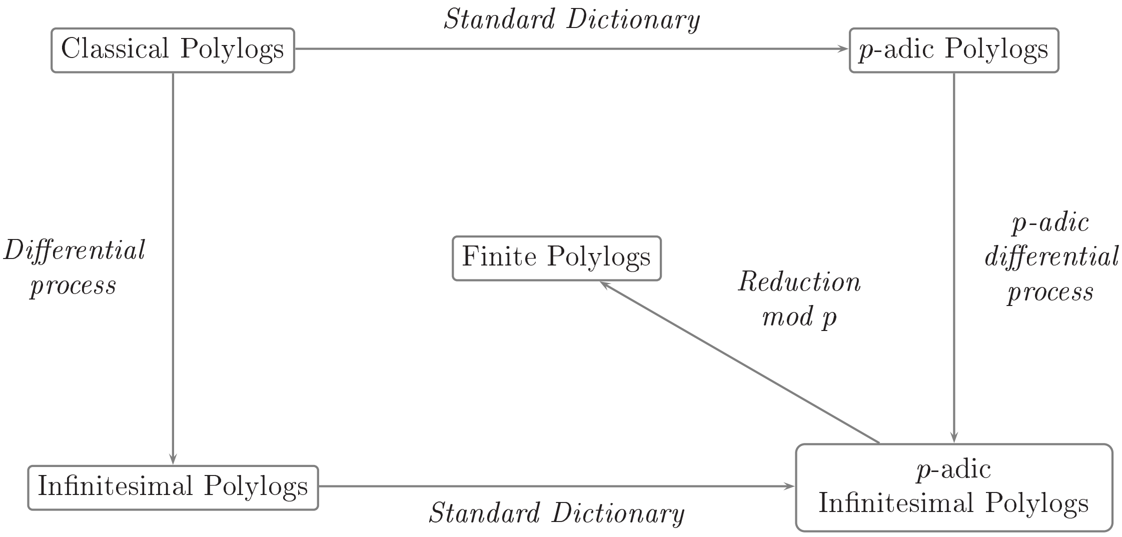
Kontsevich had asked explicitly in [22] for functional equations similar to the fundamental equation of information theory for the next case, i.e. for the case of the finite dilogarithm $\pounds_2(x) = \sum_{k=1}^{p-1} x^k/k^2$. Guided by the analogy between finite 1-logarithm and the infinitesimal dilogarithm, it was found that Cathelineau's equation for the infinitesimal trilogarithm is also satisfied by $\pounds_2$ and provides an answer to Kontsevich's question. Furthermore, $\pounds_2$ is characterized by the latter equation (actually, it is already characterized by certain specializations).

In fact we get a stronger statement: *each of the functional equations for the infinitesimal n-logarithm in this paper—and this includes the distribution formulas for any n—has been proved for the finite $(n-1)$-logarithm (whose definition should be clear by the above).*

What is more, there is a whole machinery to obtain this type of functional equations: on the one hand, Cathelineau had given a tangential procedure for elements in $\mathbb{Z}[F]$ (for certain fields F) which is compatible with the passage from functional equations for the dilogarithm to equations for the infinitesimal dilogarithm. It turns out (see §5) that the same is true for higher polylogarithms, and we will show how we can get a functional equation for an infinitesimal n-logarithm by "taking the derivative" of a functional equation for the classical n-logarithm relatively to an absolute derivation over F. On the other hand, since p-adic polylogarithms in the sense of Coleman [10] satisfy the same functional equations as the classical ones by work of Wojtkowiak [34] (for a more precise statement cf. §6), one arrives via Cathelineau's tangential procedure (proved by him in characteristic 0) at its p-adic equivalent and one could hope that there is a version of p-adic polylogarithms whose appropriate differential reduces to the finite polylogarithms. This hope (vaguely anticipated in [14]) has been made precise by Kontsevich (private communication) and was subsequently proved (in a slightly modified form) by Besser [2]. Combining the above, we obtain a recipe for deducing functional equations for $\pounds_{n-1}$ from functional equations for the n-logarithm, and thus we get analogues of distribution relations for each n and further "non-trivial" ones at least up to $n = 7$ (cf. [37], [17]).

The properties stated motivate the terminology of "poly(ana)logs" for the different analogues of polylogs. To help the reader to understand the interdependencies between the notions already discussed, we give the following picture, which can serve as a guideline for the paper:

The conceptual relationship between the different Poly(ana)logs



The present paper investigates the basic properties of the infinitesimal version of polylogarithms, including the p-adic ones, and their relationship with the finite polylogarithms and also with the classical polylogarithms via the "derivation map" (section 5). In particular, the answer to Kontsevich's question can be found in section 4 (Theorem 4.12), together with a proof of the unicity of $\pounds_2$ (Theorem 4.23). The sequel paper [15] exhibits interrelationships among the polylogarithmic groups and also among their infinitesimal versions, introduces finite versions of the so-called "multiple polylogarithms" (cf. e.g. [21]) and in particular some multiplicative structure related to them: it turns out that the proofs of the identities for the finite field case are far from trivial, and especially the most conceptual one found for Cathelineau's 22-term equation involves an identity expressing $\pounds_1(a)\pounds_1(b)$ in terms of $\pounds_2$ only. The special case of $a = b$ in the latter product is an identity found by Mirimanoff which is crucial for proving his criteria for Fermat's last theorem—the finite polylogarithms have appeared in the literature prominently in the guise of "Mirimanoff polynomials" (cf. Ribenboim's 13 Lectures [27]). Others of Mirimanoff's identities can be reinterpreted in terms of functional equations of finite polylogarithms (actually, "multiple polylogarithms") which might nurture the hope that further knowledge concerning the latter could provide more obstacles for a solution of FLT to exist (but this may well turn out to be a too pollyanna[1] attitude)...

The organisation of the present work is as follows:
Part I is dedicated to the introduction of classical and infinitesimal polylogarithms (in characteristic 0) and their associated functional equations and groups. In particular we re-introduce several notions of Cathelineau [6, 8] and give complementary properties.
Part II introduces the finite polylogs, the functional equations that they satisfy and give

[1] **Pollyanna.** The name of the heroine of stories written by Eleanor Hodgman Porter (1868-1920), American children's author, used with allusion to her skill at the 'glad game' of finding cause for happiness in the most disastrous situations; **one who is unduly optimistic or achieves happiness through self-delusion**. [Oxford English Dictionary 2]

their characterizations (section 4). We also introduce in the section 5 the construction of the "derivation map" and show that functional equations for classical polylogs give rise to functional equations for infinitesimal polylogs. The last section of this part (section 6) introduces the p-adic methods, and shows (Corollary 6.12), via Besser's result, that functional equations for infinitesimal p-adic polylogs produce functional equations for finite polylogs (under mild assumptions).
Finally, the main proofs of Part II are given in Part III.
The paper ends with a reproduction of the note of Kontsevich [22], originally written for a private booklet dedicated to Friedrich Hirzebruch on the occasion of his "Emeritierung" (retirement). We are grateful to him for letting us include it as an appendix.

Acknowledgements : We would like to express our sincere gratitude for financial support to DFG, EU (Marie Curie fellowship program), and to the following institutions for their hospitality: the Laboratoire Dieudonné UMR CNRS 6601 of the University of Nice–Sophia Antipolis where the whole project started out (and where the second author made several visits working on this project), the Institut für Experimentelle Mathematik (Essen), the FB6 Mathematik of the Universität-GH Essen, the I.H.É.S. and the MPI für Mathematik Bonn. We want to give our "Herzliches Dankeschön" to G. Faltings and G. Harder for their invitation to the MPI Conference on Polylogarithms at Schloss Ringberg. The first author wants to thank the Mathematical Institute of the University of Lausanne (and especially D. Arlettaz) where he was working on the subject. The second author wants to thank the Laboratoire G.T.A. UMR CNRS 5030 of the University Montpellier II for its hospitality while he was visiting. During the long period of gestation of this work we had the opportunity to discuss it with several people. It is a pleasure to thank them here: we are grateful to H. Esnault, G. Frey, A. Goncharov, G. Mersmann, J. Nekovař and Z. Wojtkowiak (and also to D. Grayson for pointing out the origin of Pollyanna). We particularly want to thank A. Besser for sending us a preliminary version of his note and for helpful remarks, P. Colmez for his patience explaining to us several p-adic features, C. Soulé for his constant interest and questions in our work, as well as D. Zagier for enlightening and stimulating comments. Last but not least, we want to express our warmest thanks to the (inadvertent) initiators of the story: J.-L. Cathelineau and M. Kontsevich, for their sustained encouragement and propulsive discussions, and without whom this work would not exist.

Part I. Preliminary Background

1. Definitions of polylogarithms and their analogues (in characteristic 0)

In the following we will recall some standard, and some less standard, facts about polylogarithms and their functional equations. The main references will be Zagier [36] and Goncharov [20] (for the classical case) as well as Cathelineau [8] (for the infinitesimal case).

1.1. Classical and one-valued Polylogarithms.

Let $n \geqslant 1$, and $\mathcal{D}_n : \mathbb{C} \to \mathbb{R}(n-1)$ be the Bloch/Wigner/Ramakrishnan/Zagier/Wojtkowiak function [36, 20, 8, 33], or *modified nth polylogarithm*, defined by

$$\mathcal{D}_n(z) = \Re_n \left(\sum_{k=0}^{n-1} \frac{2^k B_k}{k!} \log^k |z| Li_{n-k}(z) \right),$$

where $\Re_n$ denotes Re or iIm, and $\mathbb{R}(n) = \mathbb{R}$ or $i\mathbb{R}$, depending on whether n is even or odd. The B_k are the Bernoulli numbers ($B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$, $B_3 = 0, \ldots$), and Li_m

denotes the classical m-logarithm

$$Li_m(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^m}\,, \qquad |z| < 1,$$

which can be analytically continued to the cut plane $\mathbb{C} - [1, \infty)$ [36]. For example, we have,

$$\begin{aligned} \mathcal{D}_1(z) &= -\log|1-z|, \\ \mathcal{D}_2(z) &= i\,\mathrm{Im}\Big(Li_2(z) + \log(1-z)\log|z|\Big), \\ \mathcal{D}_3(z) &= \mathrm{Re}\Big(Li_3(z) - \log|z| Li_2(z) - \frac{1}{3}\log^2|z|\log(1-z)\Big). \end{aligned}$$

***Remark* 1.1.** 1. The virtue of these modifications of classical polylogarithms lies in the fact that they are one-valued functions on the whole complex plane (at the points 0 and 1 they are defined by continuity)—as opposed to the multi-valued classical polylogarithm functions—and that they satisfy "clean" functional equations (i.e. without lower order terms such as products of polylogarithms of lower degrees).

2. Instead of the above $\mathcal{D}_n$ there is also the closely related real-valued function P_n (originally introduced by Zagier [36]) widely used, and also denoted $\mathcal{L}_n$, e.g. [19]. It differs from $\mathcal{D}_n$ only by a possible factor of i.

3. *Polylogarithms of a real variable.* In a similar manner one can define real valued functions as given by Zagier [36] (eq. (31), p.412), cf. also Lewin [25] (eq.(16), p.7), which could be called *Rogers polylogarithms* in view of Rogers's investigations in the case $n = 2$ [28]:

$$L_n(x) = \sum_{j=0}^{n-1} \frac{(-\log|x|)^j}{j!} Li_{n-j}(x) \;+\; \frac{(-\log|x|)^{n-1}}{n!}\log|1-x|\,, \quad |x| \leqslant 1,$$

and for $|x| > 1$ via the inversion relation

$$L_n\Big(\frac{1}{x}\Big) = (-1)^{n-1} L_n(x)\,.$$

1.2. **Infinitesimal polylogarithms.** We mainly follow the presentation in Cathelineau [8]. Differentiating the functions $\mathcal{D}_n$ gives (see [8], p. 1328)

$$\frac{\partial}{\partial z}\mathcal{D}_n(z) = -\sum_{k=1}^{n-1} \frac{2^{k-1}B_k}{k!}\frac{\log^{k-1}|z|}{z}\mathcal{D}_{n-k}(z) + \frac{2^{n-2}B_{n-1}}{(n-1)!}\frac{\log^{n-1}|z|}{1-z},$$

and

$$\frac{\partial}{\partial \bar z}\mathcal{D}_n(z) = (-1)^{n-1}\overline{\frac{\partial}{\partial z}\mathcal{D}_n(z)}.$$

Finally we can deduce the expression for $d\mathcal{D}_n(z)$

$$\begin{aligned} d\mathcal{D}_n(z) &= \frac{\partial}{\partial z}\mathcal{D}_n(z)dz + \frac{\partial}{\partial \bar z}\mathcal{D}_n(z)d\bar z \\ &= -\sum_{k=1}^{n-1}\left(\frac{2^k B_k}{k!}\log^{k-1}|z|\,\mathcal{D}_{n-k}(z)\Upsilon_k(z)\right) - \frac{2^{n-1}B_{n-1}}{(n-1)!}\log^{n-1}|z|\,\Upsilon_{n-1}(1-z). \end{aligned}$$

If k is even : $\Upsilon_k(z) = d\log|z|$, and if k is odd : $\Upsilon_k(z) = di\arg(z)$. The main examples are

$$d\mathcal{D}_1(z) = -d\log|1-z|,$$
$$d\mathcal{D}_2(z) = -\log|1-z|\, di\arg(z) + \log|z|\, di\arg(1-z),$$
$$d\mathcal{D}_3(z) = \mathcal{D}_2(z) di\arg(z) + \frac{1}{3}\log|z| \big(\log|1-z|\, d\log|z| - \log|z|\, d\log|1-z|\big).$$

***Remark* 1.2.** Goncharov [19](Prop. 1.18) had deduced a slightly different, but equivalent, formula earlier (the terms which seem a priori different—he wrote $d\log|z|$ instead of $d\arg(z)$ — turn out to be multiplied by a Bernoulli number B_k which is zero since k is odd).

2. Groups related to polylogarithms

In the following, F will denote a field, and we abbreviate $F^{\bullet\bullet} = F - \{0,1\}$. We can think of it as a doubly punctured affine line over F.

2.1. The scissors congruence group.

We define the scissors congruence group $\mathfrak{p}(F)$ as the quotient of $\mathbb{Z}[F^{\bullet\bullet}]$ by the subgroup generated by the elements

$$[a] - [b] + \left[\frac{b}{a}\right] - \left[\frac{1-b}{1-a}\right] + \left[\frac{1-b^{-1}}{1-a^{-1}}\right],$$

whenever such an expression makes sense. The relation is the famous five term equation for the dilogarithm (first stated by Abel, cf. [23]). This group, which has a geometric origin (see for instance [11]), captures the algebraic properties of the dilogarithm, more precisely one has

Proposition 2.1. *If $F \subset \mathbb{C}$, then the dilogarithm $\mathcal{D}_2$ is defined on $\mathfrak{p}(F)$.*

Suslin's definition of the Bloch group of a field is given by the following exact sequence (see [30]),

$$0 \to B(F) \to \mathfrak{p}(F) \xrightarrow{\lambda} (F^\times \otimes_{\mathbb{Z}} F^\times)_s \to K_2^M(F) \to 0, \tag{2.1}$$

where $K_2^M(F)$ is the Milnor K_2 of the field F (see [29], chapter 4), $(F^\times \otimes_{\mathbb{Z}} F^\times)_s$ is the quotient of $F^\times \otimes_{\mathbb{Z}} F^\times$ by the subgroup generated by the elements of the kind $x\otimes y + y \otimes x$. The map λ is then defined by $\lambda([a]) = a \otimes (1-a)$ and the *Bloch group* of F is defined as the kernel of this map.

***Remark* 2.2.**

1. In [11], Dupont and Sah have studied in detail the scissors congruence group and also its connection to the dilogarithm.
2. If F is an infinite field, the precise relationship between $K_3(F)$ and $B(F)$ is described by Suslin in [30], and rationally we have $K_3(F)_{\mathbb{Q}} \cong B(F)_{\mathbb{Q}}$ which gives a description of $K_3(F)_{\mathbb{Q}}$ in terms of generators and relations.
3. Weibel [32] has computed the group $B(F)$ if F is a finite field and has shown that it has the same relationship to K_3 as in the case of infinite fields.

4. The original definition of Bloch [4](Lecture 6, p.59) is given by the following exact sequence
$$0 \to \mathcal{B}(F) \to \mathcal{A}(F) \xrightarrow{\lambda} F^\times \otimes_{\mathbb{Z}} F^\times \to K_2(F) \to 0,$$
where $\mathcal{A}(F)$ is just the group $\mathbb{Z}[F^{\bullet\bullet}]$. Notice that he also generalized the definition to rings in order to prove some rigidity property [4](pp.62-68). Moreover, he obtained a map between $\mathcal{B}(F)$ and $K_3^{ind}(F)/Tor_1^{\mathbb{Z}}(F^\times, F^\times)$ for any algebraically closed field F [4](pp.71-72). (Here, $K_3^{ind}(F)$ denotes the quotient of $K_3(F)$ by the image of $K_3^M(F)$ in $K_3(F)$.) Later, Suslin [30] showed that we have an analogous map with $B(F)$ and that, modulo 2-torsion, this map is an isomorphism.

5. In fact the exact sequence (2.1) holds also for "rings with many units", such as semilocal rings with infinite residue fields (this is a consequence of results in [12]).

2.2. **Polylogarithmic groups and Goncharov complexes.** Zagier has generalized in [36] (section 8) the construction of the Bloch group to higher n and defined *higher Bloch groups*, on which the corresponding polylogarithm functions $\mathcal{D}_n$ are defined. They are constructed by an inductive procedure which has been made more conceptual by Goncharov whose framework we adopt here. Let $\mathbb{P}^1(F)$ be the projective line over F. The construction of an intermediate group $\mathcal{B}_n(F)$, descriptively called *polylogarithmic group* in [7], proceeds by induction on $n \geqslant 2$. We first need to construct certain subgroups $\mathcal{A}_n(F)$ and $\mathcal{R}_n(F)$ of $\mathbb{Z}[\mathbb{P}^1(F)]$. Suppose that $\mathcal{R}_n(F)$ is defined, then we set
$$\mathcal{B}_n(F) = \mathbb{Z}[\mathbb{P}^1(F)]/\mathcal{R}_n(F).$$

Define the morphisms
$$\begin{aligned} \delta_2 = \delta_{2,F} : \qquad & \mathbb{Z}[\mathbb{P}^1(F)] \to \frac{\bigwedge_{\mathbb{Z}}^2 F^\times}{(2 - torsion)}, \\ & [x] \mapsto \begin{cases} 0 & \text{if } x = 0, 1, \infty, \\ (1-x) \wedge x & \text{otherwise,} \end{cases} \end{aligned}$$
and for $n \geqslant 3$
$$\begin{aligned} \delta_n = \delta_{n,F} : \qquad & \mathbb{Z}[\mathbb{P}^1(F)] \to (\mathcal{B}_{n-1}(F) \otimes F^\times), \\ & [x] \mapsto \begin{cases} 0 & \text{if } x = 0, 1, \infty, \\ \{x\}_{n-1} \otimes x & \text{otherwise,} \end{cases} \end{aligned}$$
where $\{x\}_n$ denotes the class of x in $\mathcal{B}_n(F)$.
Although it is not used in the inductive definition, let us define $\mathcal{R}_1(F)$ to be the group generated by $[\infty]$ and $[x + y - xy] - [x] - [y]$, where $x, y \in F \backslash \{1\}$. Then $\mathcal{B}_1(F) \cong F^\times$.
For $n \geqslant 2$, we define $\mathcal{A}_n(F)$ as the kernel of δ_n and $\mathcal{R}_n(F)$ as the subgroup of $\mathbb{Z}[\mathbb{P}^1(F)]$ spanned by $[0]$, $[\infty]$ and the elements $\sum n_i([f_i(0)] - [f_i(1)])$, where the f_i are rational fractions in the indeterminate T, such that $\sum n_i[f_i] \in \mathcal{A}_n(F(T))$. Goncharov proved the following basic

Lemma 2.3. *For all $n \geqslant 2$, the group $\mathcal{R}_n(F)$ is contained in the kernel of δ_n.*

Proof. See [19](Lemma 1.16, p.221) and also [8](Proposition 1, p.1330). □

We then have a (cochain) complex, due to Goncharov [19, 20], with the group $\mathcal{B}_n(F)$ put in degree 1,
$$\mathcal{B}_n(F) \xrightarrow{\delta} \mathcal{B}_{n-1}(F) \otimes F^\times \xrightarrow{\delta} \mathcal{B}_{n-2}(F) \otimes \bigwedge^2 F^\times \xrightarrow{\delta} \cdots \xrightarrow{\delta} \mathcal{B}_2(F) \otimes \bigwedge^{n-2} F^\times \xrightarrow{\delta} \frac{\bigwedge^n F^\times}{(2-torsion)},$$

with

$$\delta(\{x\}_{n-i} \otimes y_1 \wedge \ldots \wedge y_i) = \{x\}_{n-i-1} \otimes x \wedge y_1 \wedge \ldots \wedge y_i\,, \qquad i = 0, \ldots, n-3,$$

and

$$\delta(\{x\}_2 \otimes y_1 \wedge \ldots \wedge y_{n-2}) = (1-x) \wedge x \wedge y_1 \wedge \ldots \wedge y_{n-2}.$$

Zagier's *higher Bloch groups* [36] arise in this context as the first cohomology group of the above complex, namely

$$B_n(F) = \frac{\mathcal{A}_n(F)}{\mathcal{R}_n(F)}\,.$$

Note the typographical difference: one has $B_n(F) \subset \mathcal{B}_n(F)$. (There are in the literature several similar definitions of the "set of relations" $\mathcal{R}_n(F)$, denoted also $\mathcal{C}_n(F)$ in [36].)

***Remark* 2.4.** 1. According to Zagier's main conjecture, the groups $B_n(F)$ in the case of a number field F are presumably rationally isomorphic to $K_{2n-1}(F)_{\mathbb{Q}}$. Using his complex, Goncharov was able to formulate a corresponding conjecture for *any* field and involving the γ-filtration of the K-theory of F.

2. One of the major achievements concerning the above complexes was Goncharov's proof [19] of Zagier's conjecture for $n = 3$ in the course of which he has given an explicit set of relations for (some version of) $\mathcal{R}_3(F)$ which enabled him to relate $B_3(F)$ to (some graded piece of) the algebraic K-group $K_5(F)$. It is not known, however, whether his relations generate *all* functional equations for the 3-logarithm.

2.2.1. ***Functions on the polylogarithmic groups***. The following proposition relates functional equations for polylogarithms and relations in $\mathcal{B}_n(F)$. (It is essentially the content of [36], Prop.3, in the form given in [19].)

Criterion 2.5. *The function $\mathcal{D}_n$ vanishes on $\mathcal{R}_n(F)$, assuming that $F \subset \mathbb{C}$.*

Let us end this section with a characterization of functions which actually can be defined on the corresponding $\mathcal{B}_n(F)$. For $n \leqslant 3$ one knows from work of Bloch and Goncharov, respectively, a characterization of the measurable functions which are defined on $\mathcal{B}_n(\mathbb{C})$:

Proposition 2.6. (Characterization of $\mathcal{D}_1$, $\mathcal{D}_2$ and $\mathcal{D}_3$)

1. *The function $\mathcal{D}_1(z) = -\log|1-z|$ is (up to a constant factor) the only measurable function defined on $\mathcal{B}_1(\mathbb{C})$.*

2. *The function $\mathcal{D}_2$ is (up to a constant factor) the only measurable function : $\mathbb{C} \to \mathbb{R}$ which vanishes on $\mathcal{R}_2(\mathbb{C})$ and thus defines a morphism on $\mathcal{B}_2(\mathbb{C})$.*

3. *The space of measurable functions : $\mathbb{C} \to \mathbb{R}$ which vanish on $\mathcal{R}_3(\mathbb{C})$ and thus define a morphism on $\mathcal{B}_3(\mathbb{C})$, is two-dimensional, spanned by $\mathcal{D}_3$ and $z \mapsto \log|z|\,\mathcal{D}_2(z)$.*

Proof. 1. is classical, 2. has been proved by Bloch [4], and 3. was given by Goncharov [19]. □

2.3. **The infinitesimal polylogarithmic groups.** Cathelineau [8] has given analogues of the Goncharov complexes for infinitesimal polylogarithms whose cohomology is expected to be computed by some graded piece of Hochschild homology (the latter can be viewed in a sense as arising from applying a certain tangent functor to algebraic K-theory).

One defines the group $\beta_2(F)$, for F any infinite field, as follows

$$\beta_2(F) = \frac{F[F^{\bullet\bullet}]}{r_2(F)},$$

where $r_2(F)$ is the kernel of the map

$$F[F^{\bullet\bullet}] \longrightarrow F^+ \otimes F^\times, \quad [a] \mapsto a \otimes a + (1-a) \otimes (1-a).$$

If $\mathcal{D}_2$ denotes the Bloch-Wigner dilogarithm function, as defined in (1.1), and if $F \subset \mathbb{C}$, then $\widetilde{d\mathcal{D}_2}$, a somewhat modified differential defined below, is zero on $r_2(F)$.
For $n \geqslant 3$, one defines inductively

$$\beta_n(F) = \frac{F[F^{\bullet\bullet}]}{r_n(F)},$$

where $r_n(F)$ is the kernel of the map

$$\partial_n = \partial_{n,F} : F[F^{\bullet\bullet}] \to (\beta_{n-1}(F) \otimes F^\times) \oplus (\mathcal{B}_{n-1}(F) \otimes F),$$
$$[a] \mapsto \langle a \rangle_{n-1} \otimes a + \{a\}_{n-1} \otimes (1-a),$$

and where $\langle a \rangle_k$ and $\{a\}_k$ denotes the class of $[a]$ in $\beta_k(F)$ and $\mathcal{B}_k(F)$, respectively.

The F-vector spaces $\beta_n(F)$ can be viewed as infinitesimal analogues of the groups $\mathcal{B}_n(F)$. The previous definition still makes sense in the case of a finite field F, but it would give $\beta_2(F) = 0$. But there is also a presentation of $\beta_2(F)$ in terms of generators and relations given in [6](section 1, pp.52-53). As we are mainly interested by the structural properties of infinitesimal polylogarithms, we introduce the following group

Definition 2.7. *Let F be an arbitrary field. The group $\mathfrak{b}_2(F)$ is defined as the F-vector space generated by symbols $\langle a \rangle$, $a \in F^{\bullet\bullet}$, subject to the relation*

$$\langle a \rangle - \langle b \rangle + a\Big\langle \frac{b}{a} \Big\rangle + (1-a)\Big\langle \frac{1-b}{1-a} \Big\rangle = 0,$$

for $a \neq b$.

We should notice that we always have a natural map $\mathfrak{b}_2(F) \to \beta_2(F)$. In characteristic 0, using [8](section 4.2, pp. 1336-1337), we have

Proposition 2.8. *If F is a field of characteristic 0 then the groups $\mathfrak{b}_2(F)$ and $\beta_2(F)$ are isomorphic.*

***Remark* 2.9.**

1. Definition 2.7 makes sense for any field.
2. It is not obvious that for a finite field $\mathbb{F}$ of characteristic p we have $\mathfrak{b}_2(\mathbb{F}) \neq 0$. It will be proven later that this is actually the case. As a counterpoint, if F is a finite field or, more generally, a perfect field of characteristic $p \neq 2$, we then have $\beta_2(F) = 0$ (see [6](Théorème 1, p.57)).

3. The infinitesimal analogue (in the sense of Cathelineau) of the above higher Bloch group $B_n(F)$ would be $\ker \partial_n / r_n(F)$ which turns out to be 0 for $n = 2, 3$ if F is any field of characteristic 0. In fact we can show that the analogue of the Bloch group $B_2(F)$ is given by the second Harrison homology group [13], proving that it is zero for any smooth $\mathbb{Q}$-algebra. The results and problems described in [13, 9], illustrate the (presumably) close connection between infinitesimal Bloch groups and smoothness properties.

***Observation* 2.10.** (Possible extension of generators in characteristic 0)

1. If we allow the symbols $\langle 1\rangle_n$ and $\langle 0\rangle_n$ in $\beta_n(F)$ then, using the distribution relation (3.10) below, we necessarily have $\langle 1\rangle_n = \langle 0\rangle_n = 0$ if $n = 2, 3$.
2. We have $\langle -1\rangle_{2k+1} = 0$ by the inversion relation.

2.3.1. ***Functions on infinitesimal polylogarithmic groups.*** The following proposition from [8] relates, for $F = \mathbb{C}$, functional equations for the infinitesimal polylogarithms and relations in the corresponding groups.

Proposition 2.11. [8] *For $n \geqslant 2$, the morphism of $\mathbb{R}$-vector spaces*

$$\begin{aligned} \widehat{d\mathcal{D}_n} : \quad \mathbb{C}[\mathbb{C}^{\bullet\bullet}] &\longrightarrow \mathbb{R}(n-1) \\ b[a] &\mapsto d\mathcal{D}_n(a)(a(1-a)b), \end{aligned}$$

is zero on $r_n(\mathbb{C})$, hence we get a morphism

$$\widetilde{d\mathcal{D}_n} : \beta_n(\mathbb{C}) \longrightarrow \mathbb{R}(n-1).$$

***Remark* 2.12.** The definition is to be understood as follows: consider $\mathbb{C}$ as a 2-dimensional $\mathbb{R}$-vector space with basis $(1, i)$ and with multiplication induced by the one in $\mathbb{C}$. Then $\mathcal{D}_n$ is seen as a map from $\mathbb{R}^2 \to \mathbb{R}$, $d\mathcal{D}_n(a)$ is given by the Jacobian matrix in a (i.e. a row matrix of length 2). Identifying $a(1-a)b$ as a column vector relative to the basis $(1, i)$, the expression $d\mathcal{D}_n(a)(a(1-a)b)$ is just the evaluation of the linear map $d\mathcal{D}_n(a)$ in $a(1-a)b$ (i.e. the product of a row matrix of length 2 by a column vector of same size).

Proposition 2.13. (Characterization of $d\mathcal{D}_2$)
The function $d\mathcal{D}_2$, restricted to $\mathbb{R}$, is (up to a constant factor) the only continuous function $G : \mathbb{R}^{\bullet\bullet} \to \mathbb{R}$ which satisfies the equation

$$a(1-a)\, G(a) - b(1-b)\, G(b) + \frac{b(a-b)}{a}\, G\left(\frac{b}{a}\right) + \frac{(1-b)(a-b)}{1-a}\, G\left(\frac{1-b}{1-a}\right) = 0\,.$$

whenever the terms are defined.

Proof. Define $H(a) = a(1-a)G(a)$, $a \in \mathbb{R}^{\bullet\bullet}$, and $H(0) = H(1) = 0$, then the above functional equation is reduced to the equation from 2.7, for which it is well-known (cf. e.g. [22]) that there is only the differentiable function $H(x) = -x \log|x| - (1-x)\log|1-x|$ (up to a multiplicative constant) which satisfies the latter equation. □

***Remark* 2.14.** Aczel and Dhombres [1](section 5.4, pp.66-69) have shown that if g is a real function locally integrable on $]0,1[$ and if, moreover, g fulfills the *Fundamental Equation of Information Theory*, namely

$$g(x) + (1-x)g\left(\frac{y}{1-x}\right) - g(y) - (1-y)g\left(\frac{x}{1-y}\right) = 0,$$

then there exists $c \in \mathbb{R}$ such that $g = cH$, where $H :]0,1[\to \mathbb{R}$, is the function $H(x) = -x\log(x) - (1-x)\log(1-x)$. For more detail on this topic see [16].

3. Functional equations

Definition 3.1. *A* functional equation of the n-logarithm *resp.* infinitesimal n-logarithm over the field F *is an element in* $\mathcal{R}_n(F)$ *resp. in* $r_n(F)$ *(cf. s.2.2).*

Let $F = K(t_1, \dots, t_r)$ *and* K' *be an extension of* K*. We will say that* $t_1 = z_1, \dots, t_r = z_r$*, with* $z_i \in K'$*, is an admissible* K'*-specialisation for a functional equation* $\xi(t_1, \dots, t_r) \in \mathcal{R}_n(F)$ *(resp.* $r_n(F)$*), if* $\xi(z_1, \dots, z_r)$ *is well defined as an element of* $\ker(\delta_{n,K'})$ *(resp.* $\ker(\partial_{n,K'})$*).*

***Remark* 3.2.** The restriction in the definition of a functional equation for the n-logarithm to *rational* arguments (in the definition of $\mathcal{R}_n(F)$), as opposed to algebraic arguments, is probably not a serious one, since the corresponding polylogarithmic groups are expected to be rationally isomorphic (cf. e.g. [19], pp.225, Conjecture 1.20). The above definition has the advantage of being more directly accessible to calculations.

3.1. Functional equations for classical polylogarithms.

We first list the equations which are true for general n: the inversion and distribution relations.

Proposition 3.3. (Functional equations for $\mathcal{D}_n$, any n)

1. *The inversion formula:*
$$\left\{\frac{1}{a}\right\}_n = (-1)^{n-1}\{a\}_n .$$
2. *The distribution formula*
$$\{a^m\}_n = m^{n-1} \sum_{\zeta^m=1} \{\zeta a\}_n$$
holds in $\mathcal{B}_n(\mathbb{C})$ *for* $m \in \mathbb{Z}$ *and reduces to the inversion relation for* $m = -1$*.*

***Remark* 3.4.** There is another symmetry coming from the complex conjugation:
$$\mathcal{D}_n(\overline{z}) = (-1)^{n-1}\mathcal{D}_n(z) .$$
Note that this does not come from a functional equation in the above sense, since the corresponding relation $\{\overline{z}\}_n + (-1)^n\{z\}_n$ is not zero in $\mathcal{B}_n(\mathbb{C})$.

3.1.1. ***The case*** $n = 2$. The following functional equations are well-known for the dilogarithm: apart from the distribution relations above it satisfies a 2-term relation relating the arguments x and $1 - x$, while the most important relation (which actually characterizes $\mathcal{D}_2$) is the five term relation which allows a formulation as a 3-cocycle relation.

Proposition 3.5. (Functional equations for $\mathcal{D}_2$)

1. *A two term relation.*
$$\{x\}_2 = -\{1-x\}_2 . \tag{3.1}$$
2. *The five term relation. We give two different formulations:*
   (a) *(as a cocycle relation in five variables): denote* $\mathrm{cr}(a, b, c, d) = \frac{a-c}{a-d}\frac{b-d}{b-c}$ *. Then*
$$\sum_{i=1}^{5} (-1)^i \left\{\mathrm{cr}(x_1, \dots, \hat{x}_i, \dots, x_5)\right\}_2 = 0 . \tag{3.2}$$

(b) *in two variables (using the arguments as in Suslin's definition of the Bloch group; this equation is a specialization of (a), putting* $(x_1,\ldots,x_5)=(\infty,0,1,a,b)$*):*

$$\{a\}_2-\{b\}_2+\left\{\frac{b}{a}\right\}_2-\left\{\frac{1-b}{1-a}\right\}_2+\left\{\frac{1-b^{-1}}{1-a^{-1}}\right\}_2=0\,. \tag{3.3}$$

3.1.2. ***The case*** $n=3$. For the trilogarithm one has, in addition to the inversion and distribution relations, an equation with $3(+1)$ terms (in one variable), the well-known Kummer-Spence equation with $9(+1)$ terms (in two variables) and, most important, Goncharov's equation with $22(+1)$ terms (in three variables, the "$+1$" referring to some constant term).

Proposition 3.6. (Functional equations for $\mathcal{D}_3$)

1. *There is a 3-term relation*

$$\{1-x\}_3+\{x\}_3+\left\{1-\frac{1}{x}\right\}_3=\{1\}_3\,. \tag{3.4}$$

2. *The Kummer-Spence equation:*

$$\begin{aligned}\left\{\frac{a\,(1-b)}{b\,(1-a)}\right\}_3+\left\{\frac{(1-a)\,a}{b\,(1-b)}\right\}_3+\left\{\frac{a\,b}{(1-b)\,(1-a)}\right\}_3&\\ -2\left\{\frac{1-a}{1-b}\right\}_3-2\left\{\frac{b}{b-1}\right\}_3-2\left\{\frac{a}{a-1}\right\}_3&\\ -2\left\{\frac{b}{a}\right\}_3-2\left\{\frac{a}{1-b}\right\}_3-2\left\{\frac{1-a}{b}\right\}_3+2\,\{1\}_3&=0\,.\end{aligned} \tag{3.5}$$

An equivalent version is given by

$$\begin{aligned}\left\{\frac{x(1-y)^2}{y(1-x)^2}\right\}_3+\{xy\}_3+\left\{\frac{x}{y}\right\}_3&\\ -2\left\{\frac{y(1-x)}{y-1}\right\}_3-2\left\{\frac{1-x}{1-y}\right\}_3-2\left\{\frac{y(1-x)}{x(1-y)}\right\}_3&\\ -2\left\{\frac{x-1}{x(1-y)}\right\}_3-2\,\{x\}_3-2\,\{y\}_3+2\,\{1\}_3&=0\,.\end{aligned} \tag{3.6}$$

3. *Goncharov's equation: Set*

$$f(a,b,c)=\{a\}_3+\left\{\frac{b(1-a)}{b-1}\right\}_3+\left\{\frac{a(1-b)}{a-1}\right\}_3+\left\{\frac{1-a}{1-abc}\right\}_3+\left\{\frac{cb(1-a)}{1-abc}\right\}_3 \tag{3.7}$$

$$-\{ab\}_3-\left\{-\frac{a(1-c)(1-c)}{(1-a)(1-abc)}\right\}_3\,. \tag{3.8}$$

Then

$$f(a,b,c)+f(b,c,a)+f(c,a,b)+\{abc\}_3=3\{1\}_3\,.$$

3.1.3. ***The case*** $n>3$. For general n, there are only the inversion relation and the distribution relations known (they are the so-called trivial ones), while the existence of non-trivial equations has only been established up to $n\leqslant 7$ (cf. [17]).

3.2. **Functional equations for infinitesimal polylogarithms.** Most of the functional equations for $d\mathcal{D}_n$ stated in this section can be viewed as analogues of equations for the corresponding $\mathcal{D}_n$. The main example which cannot be interpreted in this way (so far) is Cathelineau's equation for $d\mathcal{D}_3$.
We first list the equations which are true for general n: the analogues of the inversion and distribution relations.

Proposition 3.7. (Functional equations for $d\mathcal{D}_n$, any n)

1. *The inversion formula*

$$a\Big\langle\frac{1}{a}\Big\rangle_n = (-1)^{n-1}\langle a\rangle_n\,. \tag{3.9}$$

2. *The distribution formula*

$$\langle a^m\rangle_n = m^{n-2}\sum_{\zeta^m=1}\frac{1-a^m}{1-\zeta a}\,\langle\zeta a\rangle_n \tag{3.10}$$

holds in $\beta_n(\mathbb{C})$ for $m\in\mathbb{Z}$ and reduces to the inversion relation for $m=-1$. When $m=2$, we call this equation the ***duplication formula****.*

3.2.1. ***The case*** $n=2$. The following functional equations are true for the infinitesimal dilogarithm:

Proposition 3.8. (Functional equations for $d\mathcal{D}_2$)

1. *The* ***2-term relation****.*

$$\langle x\rangle_2 = \langle 1-x\rangle_2\,. \tag{3.11}$$

2. *A six term relation. Let $s\in F$. Then*

$$(1-y)\langle\frac{x-s}{1-y}\rangle_2 + y\langle\frac{s}{y}\rangle_2 + \langle y\rangle_2 \tag{3.12}$$

is symmetric in x and y. Specifically, we have for $s=0$ the ***fundamental equation of information theory***

$$(1-y)\langle\frac{x}{1-y}\rangle_2 - \langle x\rangle_2 = (1-x)\langle\frac{y}{1-x}\rangle_2 - \langle y\rangle_2 \tag{3.13}$$

which is equivalent to Cathelineau's version

$$\langle a\rangle_2 - \langle b\rangle_2 + a\Big\langle\frac{b}{a}\Big\rangle_2 + (1-a)\Big\langle\frac{1-b}{1-a}\Big\rangle_2 = 0\,. \tag{3.14}$$

3. *A family of five term relations is given by taking linear combinations of the following two equations in five variables: denote $\mathrm{cr}(a,b,c,d)=\frac{a-c}{a-d}\frac{b-d}{b-c}$ and $\mathrm{denom}(a,b,c,d)=(a-d)(b-c)$. Then one has*

$$\sum_{i=1}^{5}(-1)^i\,\mathrm{denom}(x_1,\dots,\widehat{x_i},\dots,x_5)\,\big\langle\mathrm{cr}(x_1,\dots,\widehat{x_i},\dots,x_5)\big\rangle_2 = 0\,, \tag{3.15}$$

and

$$\sum_{i=1}^{5}(-1)^i\,x_i\,\mathrm{denom}(x_1,\dots,\widehat{x_i},\dots,x_5)\,\big\langle\mathrm{cr}(x_1,\dots,\widehat{x_i},\dots,x_5)\big\rangle_2 = 0\,. \tag{3.16}$$

4. *The same family of five term relations can be stated with less parameters in the arguments:*

(3.17)
$$(b+t)\langle a\rangle_2-(a+t)\langle b\rangle_2+(1+t)a\Big\langle\frac{b}{a}\Big\rangle_2+t(1-a)\Big\langle\frac{1-b}{1-a}\Big\rangle_2+b(1-a)\Big\langle\frac{a(1-b)}{b(1-a)}\Big\rangle_2.$$

Proof. It is a straightforward matter to check that the above elements lie in the kernel of ∂_2. Nevertheless, we give some interrelationships between the various equations.

1. The symmetry of equation (3.12) is equivalent to (3.14):
   We have to write the following relation
   $$(1-y)\langle\frac{x-s}{1-y}\rangle_2+y\langle\frac{s}{y}\rangle_2+\langle y\rangle_2=(1-x)\langle\frac{y-s}{1-x}\rangle_2+x\langle\frac{s}{x}\rangle_2+\langle x\rangle_2$$
   as a sum of 4-term relations.
   On the left hand side of the equation we add the 4-term relation in the following form
   $$-y\left\langle\frac{s}{y}\right\rangle_2-\langle y\rangle_2+\langle s\rangle_2+(1-s)\left\langle\frac{1-y}{1-s}\right\rangle_2=0\,,$$
   and we do the same on the right hand side with y replaced by x. This leaves us with another form of the 4-term relation
   $$(1-y)\left\langle\frac{x-s}{1-y}\right\rangle_2+(1-s)\left\langle\frac{1-y}{1-s}\right\rangle_2=(1-x)\left\langle\frac{y-s}{1-x}\right\rangle_2+(1-s)\left\langle\frac{1-x}{1-s}\right\rangle_2$$
   (to see this we should replace, in (3.14), x by $\frac{x-s}{1-s}$ and y by $\frac{y-s}{1-s}$ and use (3.11)), thereby proving the first claim.
   The equivalence of (3.14) and (3.13) is easily shown using the inversion and the 2-term relation.
2. The second family of five term relations is almost direct to deduce: the combination given is the sum of t times the 4-term relation (3.14) and its following equivalent formulation

(3.18)
$$b\langle a\rangle_2-a\langle b\rangle_2+a\Big\langle\frac{b}{a}\Big\rangle_2+b(1-a)\Big\langle\frac{a(1-b)}{b(1-a)}\Big\rangle_2.$$

   (replace in (3.14) a and b by their inverses, respectively, then multiply the result by $-ab$ and finally use the inversion relation on three of the ensuing terms).

From this, we get a very simple proof of the five term relations in cocycle form, i.e. (3.15) and (3.16): in each of the two versions (3.14) and (3.18) of the 4-term relation we put $a=\mathrm{cr}(x_1,x_2,x_3,x_4)$ and $b=\mathrm{cr}(x_1,x_2,x_3,x_5)$. Introducing for the moment the notation $(ijkl):=\mathrm{cr}(x_i,x_j,x_k,x_l)$, we can rewrite the two equations in a concise way:
$$\begin{aligned}&\big\langle(1234)\big\rangle-\big\langle(1235)\big\rangle+(1234)\big\langle(1245)\big\rangle+(1324)\big\langle(1345)\big\rangle\,,\\&(1235)\big\langle(1234)\big\rangle-(1234)\big\langle(1235)\big\rangle+(1234)\big\langle(1245)\big\rangle+(1235)(1324)\big\langle(2345)\big\rangle\,.\end{aligned}$$
Given $\lambda\in\mathbb{Z}$, there is a linear combination of the two equations such that the coefficient of $\Big\langle\mathrm{cr}(x_1,x_3,x_4,x_5)\Big\rangle$ (which only occurs in the first equation) and of $\Big\langle\mathrm{cr}(x_2,x_3,x_4,x_5)\Big\rangle$ (only occurring in the second equation) is $-x_2^\lambda(x_1-x_5)(x_3-x_4)$ and $x_1^\lambda(x_2-x_5)(x_3-x_4)$, respectively. If, for $\lambda=0$ and $\lambda=1$, we compute the coefficients of the other three arguments, we obtain exactly the expressions given in the claim.

For example, let us compute the coefficient of the first argument in the case $\lambda = 1$: the first equation is multiplied by

$$-x_2(x_1 - x_5)(x_3 - x_4)\frac{(x_1 - x_4)(x_3 - x_2)}{(x_1 - x_2)(x_3 - x_4)}\,,$$

the second by

$$x_1(x_2 - x_5)(x_3 - x_4)\frac{(x_1 - x_5)(x_2 - x_3)}{(x_1 - x_3)(x_2 - x_5)}\,\frac{(x_1 - x_4)(x_3 - x_2)}{(x_1 - x_2)(x_3 - x_4)}\,,$$

so the coefficient becomes

$$-x_2(x_1 - x_5)\frac{(x_1 - x_4)}{(x_1 - x_2)}(x_3 - x_2) + x_1(x_2 - x_5)\frac{(x_1 - x_4)}{(x_1 - x_2)}(x_3 - x_2)\,,$$

which is equal to $x_5(x_1 - x_4)(x_2 - x_3)$.

□

***Remark* 3.9.** The generalized version of the fundamental equation of information theory, namely (3.12), is equivalent to the one given by both Kontsevich and Cathelineau (referring to Aczél-Dhombres), as was shown in the proof (part 1.) above. At first glance, it is somewhat surprising that we do not get anything new although we can achieve to insert a third (non-homogenizing) parameter—but there are related phenomena known for the five term relation. In particular, we do not gain new information for information theory.

3.2.2. ***The case*** $n = 3$. For the infinitesimal trilogarithm one has an equation with three terms (in one variable), a "derived version" of the Kummer-Spence equation with eight terms (in two variables) and, most important, Cathelineau's equation with 22 terms (in three variables).

The proposition below gives complementary information on $\beta_3(F)$.

Proposition 3.10. (Functional equations for $d\mathcal{D}_3$)

1. *There is a 3-term relation*

$$\langle 1 - x\rangle_3 - \langle x\rangle_3 + x\left\langle 1 - \frac{1}{x}\right\rangle_3 = 0\,. \tag{3.19}$$

2. *The Kummer-Spence analogue: the F-linear combination*

$$\begin{aligned}\frac{(1-b)\,b}{1-b-a}\left\langle\frac{(1-a)\,a}{b\,(1-b)}\right\rangle_3 &+ \frac{(1-b)\,(1-a)}{1-b-a}\left\langle\frac{a\,b}{(1-b)\,(1-a)}\right\rangle_3\\ &+ (1-b)\left\langle\frac{1-a}{1-b}\right\rangle_3 - (1-b)\left\langle\frac{b}{b-1}\right\rangle_3 - (1-a)\left\langle\frac{a}{a-1}\right\rangle_3\\ &- a\left\langle\frac{b}{a}\right\rangle_3 + \frac{(a-b-1)\,(1-b)}{1-b-a}\left\langle\frac{a}{1-b}\right\rangle_3 - \frac{(a-b+1)\,b}{1-b-a}\left\langle\frac{1-a}{b}\right\rangle_3\end{aligned} \tag{3.20}$$

vanishes in $\beta_3(F)$. An equivalent version, denoted $KS(x,y)$, is given by

$$\begin{aligned}(3.21)\quad \langle xy\rangle_3 + y\left\langle \frac{x}{y}\right\rangle_3 - (1-y)\left\langle \frac{y(1-x)}{y-1}\right\rangle_3 \\ + (1-y)\left\langle \frac{1-x}{1-y}\right\rangle_3 - x(1-y)\left\langle \frac{y(1-x)}{x(1-y)}\right\rangle_3 \\ + x(1-y)\left\langle \frac{x-1}{x(1-y)}\right\rangle_3 - (1+y)\langle x\rangle_3 - (1+x)\langle y\rangle_3\end{aligned}$$

Proof. The 3-term equation and (3.6) will follow directly from the equation in the next proposition. The equivalence of the two Kummer-Spence analogues becomes evident after applying the change of variables $x = \frac{a}{1-b}$, $y = \frac{1-a}{b}$, and multiplying the result by $\frac{b(1-b)}{1-a-b}$. □

We can also notice the following formal property, that we will give as

Lemma 3.11. *In $\beta_3(F)$, the inversion formula is a consequence of the 3-term equation.*

Proof. Add the 3-term equation to its variant where x is replaced by $1-x$. Four of the terms cancel and the remaining two give the inversion relation. □

Cathelineau has given a 22-term equation which completely describes the set of relations for the infinitesimal polylogarithmic group $\beta_3(F)$: In order to state it conveniently, we use his notation for a distinguished linear combination of seven terms

$$[[a,b]] = (b-a)\tau(a,b) + \frac{1-b}{1-a}\sigma(a) + \frac{1-a}{1-b}\sigma(b)\,, \tag{3.22}$$

where we have set

$$\tau(a,b) = \frac{[a]}{1-a} - \frac{[b]}{1-b} + \frac{a}{a-b}\left[\frac{b}{a}\right] - \frac{1-a}{b-a}\left[\frac{1-b}{1-a}\right] + \frac{b(1-a)}{b-a}\left[\frac{a(1-b)}{b(1-a)}\right],$$

(τ arises by taking the five term relation (3.3) and multiplying each $[z_i]$ with the coefficient $\frac{1}{1-z_i}$) and

$$\sigma(a) = a[a] + (1-a)[1-a]\,.$$

Then we can state the 22-term relation as follows

Definition 3.12. *We define the formal expression $J(a,b,c)$ in the indeterminates a,b,c as*

$$J(a,b,c) = [[a,c]] - [[b,c]] + a\left[\left[\frac{b}{a},c\right]\right] + (1-a)\left[\left[\frac{1-b}{1-a},c\right]\right].$$

***Remark* 3.13.** 1. Writing out all the terms, we obtain 22 different arguments:

$$\begin{aligned} J(a,b,c) = \; & c\,[a] - c\,[b] + (a-b+1)\,[c] \\ & + (1-c)\,[1-a] - (1-c)\,[1-b] + (b-a)\,[1-c] \\ & - a\left[\frac{c}{a}\right] + b\left[\frac{c}{b}\right] + ca\left[\frac{b}{a}\right] \\ & - (1-a)\left[\frac{1-c}{1-a}\right] + (1-b)\left[\frac{1-c}{1-b}\right] + c(1-a)\left[\frac{1-b}{1-a}\right] \\ & + c(1-a)\left[\frac{a(1-c)}{c(1-a)}\right] - c(1-b)\left[\frac{b(1-c)}{c(1-b)}\right] \\ & - b\left[\frac{ca}{b}\right] - (1-b)\left[\frac{c(1-a)}{1-b}\right] \\ & + (1-c)a\left[\frac{a-b}{a}\right] + (1-c)(1-a)\left[\frac{b-a}{1-a}\right] \\ & - (a-b)\left[\frac{(1-c)a}{a-b}\right] - (b-a)\left[\frac{(1-c)(1-a)}{b-a}\right] \\ & + c(a-b)\left[\frac{(1-c)b}{c(a-b)}\right] + c(b-a)\left[\frac{(1-c)(1-b)}{c(b-a)}\right]. \end{aligned}$$

2. When a, b, c are elements of an arbitrary field F, we will still use the notation $J(a,b,c)$ for the evaluation of J in the specified values.

Theorem 3.14. (Cathelineau, [8], Corollaire 1, p.1345) *Let F be a field of characteristic zero.*

1. *The image of $J(a,b,c)$ under the projection $F[F^{\bullet\bullet}] \to \beta_3(F)$ is zero.*
2. *Furthermore, $J(a,b,c)$, together with its specializations to $c = a,\ b,\ \frac{a}{b}$ or $\frac{1-a}{1-b}$, respectively, and the inversion relation generate the set of relations which define $\beta_3(F)$. Here we understand $\langle 1\rangle_3 = 0$.*

***Remark* 3.15.** 1. In the presentation of [8], Corollaire 1, one can replace his equation 1) coming from $[[a,b]] - [[b,a]]$ by the shorter inversion relation (3.9). (Proof: add his equation 1) to the same relation where a and b are replaced by $\frac{1}{a}$ and $\frac{1}{b}$ and where the result is multiplied by ab.)

2. The combinations $[[a,c]] + a[[1/a,c]]$ and $[[a,c]] - [[1-a,c]]$ give versions of the Kummer-Spence analogue. Since, e.g., $\langle a\rangle_2 - \langle 1-a\rangle_2 = 0$ results formally from the four term relation (at least up to 2-torsion), we get the Kummer-Spence analogue directly from $J(a,b,c)$.

3. By Observation 2.10, one can introduce elements $[a]$ for $a = 0, 1$ and set their image in $\beta_3(F)$ equal to zero. What is more, one can add a formal generator $[\infty]$ as well, for which we only require $0[\infty] = 0$. One can then formally deduce the 3-term equation (3.19) by specializing $a = 1$ in $\frac{1}{b-1}J(a,b,c)$ and one obtains the Kummer-Spence analogue (3.6)

by specializing $a = 0$ in $(1-x)(1-y)J(a,(1-x)^{-1},(1-y)^{-1})$, (these specializations are not allowed in Cathelineau's context, but will make sense in the "finite polylog" case below).

4. A different way to obtain the Kummer-Spence analogue is to symmetrize, i.e. to form
$$J(a,b,c) + J(b,a,c) + c\big(J(a,b,\frac{1}{c}) + J(b,a,\frac{1}{c})\big)\,,$$
and then to check that one obtains the difference of two Kummer-Spence analogues $KS(c,\frac{b}{a}) - KS(c,\frac{1-b}{1-a})$.

5. Alternatively, one can deduce the Kummer-Spence analogue or the 3-term relation (non-explicitly) from $J(a,b,c)$ by simply checking that the corresponding linear combinations lie in the kernel of ∂_3, and then use Cathelineau's theorem to deduce that each such combination must be a consequence of $J(a,b,c)$.

6. In the case of the classical trilogarithm, Goncharov has given a new functional equation in $22(+1)$ terms which *presumably* generates all functional equations for $\mathcal{D}_3$, i.e. the kernel of δ_3, but there are (infinitely many) functional equations (cf. [35], [18]) which are not known to be formal consequences of it. Cathelineau's result in the infinitesimal setting is stronger in the sense that it actually generates the kernel of ∂_3.

One of the major consequences of Theorem 3.14 is that it allows us to give a general definition for $\mathfrak{b}_3$.

Definition 3.16. *Let F be an arbitrary field. The group $\mathfrak{b}_3(F)$ is defined as the F-vector space generated by symbols $[a]$, $a \in F$, subject to the relations $J(a,b,c)$, together with its specializations to $c = a$, b, $\frac{a}{b}$ or $\frac{1-a}{1-b}$, respectively, the inversion relation and $[1] = [0] = 0$.*

If in $\beta_3(F)$ we introduce elements $[a]$ for $a = 0, 1$, we then have, in virtue of (2.10), a surjective map of F-vector spaces $\mathfrak{b}_3(F) \to \beta_3(F)$, which is an isomorphism in characteristic 0. As in the case $n = 2$, if F is a finite field of characteristic p, $\beta_3(F) = 0$ but it will be shown in part III that $\mathfrak{b}_3(F) \neq 0$. The groups $\mathfrak{b}_n(F)$, for $n = 2, 3$, measure how much the group $\beta_n(F)$ deviates from being generated by the main functional equations of infinitesimal polylogarithms.

3.2.3. ***The case*** $n > 3$. For general n, there are only the inversion relation and the distribution relations, as seen in (3.10), known. For each functional equation of the corresponding classical polylog, using the "derivation map" described in the section 5, there is associated a functional equation (actually many) for the infinitesimal polylogarithm. From what has been stated above for the classical case, this means that at least up to $n = 7$ there are non-trivial ones.

Part II. The Results

4. Finite versions of polylogarithms and their functional equations

In this section we will study what we can call *finite analogs* of the polylogarithms and also the groups $\mathfrak{b}_n(F)$ for $n = 2, 3$ in the case where F is a field of characteristic $p \neq 2$ (eventually finite). We will show that for $n = 2, 3$ the finite analogs of the polylogarithms define functions on $\mathfrak{b}_n(F)$, showing that surprisingly they behave like the infinitesimal polylogarithms. As for the previous cases, we will show, at least in low dimension, that these *finite polylogs* are uniquely characterized by their functional equations.

For the remainder of the paper, let us fix an odd prime p. We shall work over an arbitrary field $\mathbb{F}$ of characteristic p.

4.1. **Definition and first properties of finite polylogarithms.**

Definition 4.1. *For any field F of characteristic p, the* nth finite polylogarithm *or* finite n-logarithm *is given by the following polynomial in $\mathbb{F}[T]$:*

$$\pounds_n(T) = \sum_{k=1}^{p-1} \frac{T^k}{k^n}\,.$$

***Notation* 4.2.** For the remainder of this paper, we will denote $\widetilde{P}$ the function associated to the polynomial P.

***Remark* 4.3.** 1. "Extension by periodicity"
If $\mathbb{F}$ is of characteristic p, it has $\mathbb{F}_p$ as prime subfield, which is fixed by the Frobenius morphism $x \mapsto x^p$. As a result we have the $(p-1)-$periodicity $\pounds_{n+p-1} = \pounds_n$, and we need only consider $n < p$.
2. It is important to notice that the functions $\widetilde{\pounds_n}$ are not identically zero on $\mathbb{F}$.

The following differential equation relates the finite polylogarithms of different orders (just like in the classical case)

$$d\pounds_n(U) = \pounds_{n-1}(U)\, d\log(U),$$

where we denoted $\frac{dU}{U}$ by $d\log(U)$. Extending this formally, it is convenient to introduce the following notation:

Definition 4.4. *Let F be a field of characteristic p. Define the following* "Frobeniizing" map

$$\begin{aligned} \widehat{\pounds}_m \,:\, F[F^{\bullet\bullet}] &\to F\,,\\ c[f] &\mapsto c^p \pounds_m(f)\,. \end{aligned}$$

One observes that, for any c and f in F, the differential operator $\frac{\partial}{\partial x}$ acts linearly on the coefficient c of $\widehat{\pounds}_m\big(c[f]\big)$ and, as above, like $d\log$ on the generator $[f]$:

$$\frac{\partial}{\partial x}\widehat{\pounds}_m\big(c[f]\big) = \widehat{\pounds}_{m-1}\big(c[f]\big)\frac{\partial}{\partial x}\log(f)\,.$$

***Observation* 4.5.** 0. For $n = 0$ we have

$$\pounds_0(T) = \frac{T - T^p}{1-T},$$

and therefore

$$T\pounds_0(1-T) = -(1-T)\pounds_0(T). \tag{4.1}$$

1. For $n = 1$, by expanding $(1-T)^p$ and noticing that $\frac{1}{p}\binom{p}{k} = \frac{1}{k}\binom{p-1}{k-1} = \frac{(-1)^{k-1}}{k}$, we get a simple (and well-known) formula

$$\pounds_1(T) \equiv \frac{1 - T^p - (1-T)^p}{p} \pmod{p}\,.$$

Note that the term on the right hand side occurs in the polynomials which define the sum of the two Witt vectors $(1, 0, \dots, 0, \dots)$ and $(-T, 0, \dots, 0, \dots)$.

4.2. **Functional equations for finite polylogarithms.** *A priori*, there seems to be at least two natural candidates for functional equations for the finite n-logarithm: we could ask for linear combinations $\sum_i c_i[x_i]$ such that $\widetilde{\pounds_n}$ vanishes for all specializations of the parameters which "make sense" (i.e. no term "$\frac{0}{0}$" occurs); we will call those combinations weak functional equations. But this definition has the disadvantage that there are too many ambiguities involved (just think of a coefficient that is divisible by $x^p - x$). Instead, we will impose the stronger property that $\sum_i c_i \pounds_n(x_i)$ vanishes as a rational expression, and by multiplying with the common denominator, we can even assume it to vanish as a polynomial.

Definition 4.6. *A* functional equation in the strong sense *for the finite n-logarithm over a field F of characteristic p is a finite linear combination $\sum_i c_i \pounds_n(x_i) \in F(t)[F(t)]$ which vanishes identically as a polynomial.*
A functional equation in the weak sense *is a finite linear combination $\sum_i c_i \pounds_n(x_i) \in F(t)[F(t)]$ which vanishes for each specialization of parameters which makes sense.*

In the following we list a number of equations which are identical to the ones for the infinitesimal polylogarithms, apart from "Frobeniizing" the coefficients (i.e. raising them to the pth power). The proofs will be postponed to §7.

4.2.1. ***General functional equations for $\pounds_n$.***

Proposition 4.7. *Let $n \in \mathbb{Z}$ be arbitrary. We have the following identities*

1. *Inversion formula:* $\pounds_n(T) = (-1)^n T^p \, \pounds_n\left(\dfrac{1}{T}\right)$.
   It can be viewed as a special case $(m = -1)$ of the following
2. *Distribution formulae: assume F contains a primitive mth root of unity. Then*

$$\pounds_n(T^m) = m^{n-1} \sum_{\zeta^m = 1} \frac{1 - T^{pm}}{1 - \zeta^p T^p} \, \pounds_n(\zeta T) \, .$$

3. *Special values:* $\widetilde{\pounds_n}(1) = 0$ *if* $(p-1) \not| n$ *and* $= -1$ *else, while* $\widetilde{\pounds_{2n}}(-1) = 0$ *for any* n*. Let* $B_j =$ *be the* j*th Bernoulli number and set* $G_j = 2(1 - 2^j)B_j$*. Then for* $0 < m < (p-1)$ *we have that* $\widetilde{\pounds_{p-m}}(-1) = \dfrac{G_m}{m}$.

***Remark* 4.8.** Notice that the numbers G_m are integers by virtue of classical results (for instance it is a consequence of the Theorem of von Staudt-Clausen[31](Theorem 5.10, p.56)). These numbers are called the Genocchi numbers and we have $mG_{p-1+m} = (m-1)G_m \bmod p$ which is nothing else than the famous Kummer congruence for Bernoulli numbers.

Still mirroring the set-up in the infinitesimal case, we now state several functional equations specific to $n = 1, 2$.

4.2.2. ***Equations for $\pounds_1$.***

Proposition 4.9. 1. *The 2-term relation:* $\pounds_1(T) = \pounds_1(1 - T)$.

2. *The generalized fundamental equation of information theory: let* s , x *and* y *be indeterminates. The expression*

$$H(x,y,s) = (1-y)^p \pounds_1\left(\frac{x-s}{1-y}\right) + y^p \pounds_1\left(\frac{s}{y}\right) + \pounds_1(y)$$

in $\mathbb{F}[x,y,s]$ *is symmetric in* x *and* y *. Specifically, we have*

$$\pounds_1(a) - \pounds_1(b) + a^p\, \pounds_1\left(\frac{b}{a}\right) + (1-a)^p\, \pounds_1\left(\frac{1-b}{1-a}\right) = 0\,. \tag{4.2}$$

3. *The five term relations.*
Denote $\mathrm{cr}(a,b,c,d) = \frac{a-c}{a-d}\frac{b-d}{b-c}$ *and* $\mathrm{denom}(a,b,c,d) = (a-d)(b-c)$. *Then we have the polynomial identities in* $\mathbb{F}[x_1,\ldots,x_5]$

$$\sum_{i=1}^{5}(-1)^i \left(\mathrm{denom}(x_1,\ldots,\widehat{x_i},\ldots,x_5)\right)^p \pounds_1\big(\mathrm{cr}(x_1,\ldots,\widehat{x_i},\ldots,x_5)\big) = 0\,,$$

and

$$\sum_{i=1}^{5}(-1)^i\, x_i^p \left(\mathrm{denom}(x_1,\ldots,\widehat{x_i},\ldots,x_5)\right)^p \pounds_1\big(\mathrm{cr}(x_1,\ldots,\widehat{x_i},\ldots,x_5)\big) = 0\,.$$

Corollary 4.10. *The* $\mathbb{F}$*-vector space* $\mathfrak{b}_2(\mathbb{F})$*, as defined in (2.7), is of dimension at least 1. If, moreover,* $\mathbb{F}$ *is a perfect field, then* $\mathfrak{b}_2(\mathbb{F}) = \mathbb{F}$*.*

Proof. According to Proposition 4.9, the function $\widetilde{\pounds_1}$ is a well-defined function on $\mathfrak{b}_2(\mathbb{F})$, and as it is not identically zero on $\mathbb{F}$, the dimension of $\mathfrak{b}_2(\mathbb{F})$ is non-zero. By [6](Théorème 1, p.57), we know that $\beta_2(\mathbb{F}) = 0$. But as the relations in $\beta_2(\mathbb{F})$ are given by the 4-term equation (i.e. the Fundamental Equation of Information Theory) and the relation $\sum_{k=2}^{p-1}[k1_{\mathbb{F}}]$, (see section 1.1 and also Sah's Lemma in [6](pp.52-53)), and as we further know, again by Sah (see the remark on p.53 in op. cit.), that these two relations are independent, we can conclude that the kernel of the map $\mathfrak{b}_2(\mathbb{F}) \to \beta_2(\mathbb{F})$ is generated by the element $\sum_{k=2}^{p-1}[k1_{\mathbb{F}}]$. Evaluating $\widetilde{\pounds_1}$ on this element shows that it is non-zero, which ends the proof. □

4.2.3. ***Equations for*** $\pounds_2$. In this subsection we will give answers to the question raised by Kontsevich in [22]. Notice that we need to assume $p > 3$ throughout.

Proposition 4.11. *The 3-term relation and the Kummer-Spence analogue are functional equations for* $\pounds_2$*.*

Proof. This is a consequence of the following theorem, together with remark (3.15). □

Theorem 4.12. *The image of* $J(a,b,c)$ *under the map* $\widehat{\pounds}_2$ *is a polynomial which is identically zero in* $\mathbb{F}[a,b,c]$*.*

***Remark* 4.13.**
1. By §6, there is a better answer to Kontsevich's question, at least "quantitatively": each functional equation for $d\mathfrak{D}_3$ induces a functional equation (in the weak sense) for $\pounds_2$. This is true in particular for the 3-term equation and the Kummer-Spence analog.

2. One can find further equations (in the strong sense) for $\pounds_2$ and in general for $\pounds_n$ with $n \geqslant 3$ in [16].

By similar arguments as in the proof for $\pounds_1$, we get

Corollary 4.14. *The* $\mathbb{F}$*-vector space* $\mathfrak{b}_3(\mathbb{F})$ *is of dimension at least 1.*

4.3. **Characterization of finite polylogarithms.** We can characterize $\pounds_1$ and $\pounds_2$ by the functional equations they satisfy.

Proposition 4.15. *The space (over* $\mathbb{F}$*) of solutions of the "fundamental equation of information theory" is of dimension* 1 *generated by* $\pounds_1$*.*

Proof. Set $f(T) = \sum_{i=0}^{p-1} a_i T^i \in \mathbb{F}_q[T]$, and suppose that f verifies $f(0) = 0$ and the following identity in $\mathbb{F}_q[x,y]$

$$f(x) + (1-x)^p f\left(\frac{y}{1-x}\right) - f(y) - (1-y)^p f\left(\frac{x}{1-y}\right) = 0\,.$$

Differentiating the previous equation with respect to x gives,

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

with $df(T) = a_1 + \sum_{i=2}^{p-1} i a_i T^{i-1}$ and thus $df(0) = a_1$. Setting $x = 0$ in the previous identity gives

$$a_1 + y\, df(y) - \frac{1-y^p}{1-y}\, a_1 = 0.$$

But as $\frac{1-y^p}{1-y} = \sum_{i=0}^{p-1} y^i$, the previous equality implies $a_i = \frac{a_1}{i}$. In other words, since $f(0) = 0$ we have $f = a_1 \pounds_1$, which proves the claim. $\square$

In fact we have a stronger statement

Proposition 4.16. *The 2-term equation, the inversion and the duplication formulae characterize altogether* $\pounds_1$*.*

Proof. It is a consequence of the following lemma

Lemma 4.17. *Suppose that* a_k *is a sequence of integers with* $k = 1, \ldots, p-1$ *(p an odd prime fixed), which fulfills the following rules*

$$a_k = \begin{cases} -\frac{1}{2} \sum_{i=k+1}^{p-1} a_i \binom{i}{k}, & \text{if } k \text{ is odd,} \\ \frac{1}{2} a_{\frac{k}{2}} & \text{otherwise,} \end{cases}$$

and $a_{p-k} = -a_k$ *for all* $k = 1, \ldots, p-1$*. Then* $a_k = \frac{a_1}{k} \in \mathbb{F}_p$ *for all* $k = 1, \ldots, p-1$*.*

Proof of the lemma. The proof goes by descending induction starting from $p-1$. First we notice that by the third rule, we have $a_{p-1} = -a_1 = \frac{a_1}{p-1}$ modulo p. Suppose that $a_i = \frac{a_1}{i}$ modulo p for all $i > k$. Now compute a_k modulo p. Observe that we can assume $k \leqslant p-3$, since we can compute from the rules a_{p-1} and a_{p-2}. If k is odd then by the first rule we deduce directly a_k, but we still have to show that $a_k = \frac{a_1}{k}$ modulo p. This is done via the

Sub-Lemma 4.18. *If* k *is odd and* $a_i = \frac{a_1}{i}$ *modulo* p *for all* $i > k$*. Then* $a_k = \frac{a_1}{k}$ *modulo* p*.*

Proof of the sub-lemma. We have to show that, modulo p,

$$\frac{a_1}{k} = -\frac{1}{2} \sum_{i=k+1}^{p-1} \frac{a_1}{i} \binom{i}{k},$$

or equivalently, assuming $a_1 \neq 0$, that

$$-2 = \sum_{i=k+1}^{p-1} \frac{k}{i}\binom{i}{k}.$$

But

$$\frac{k}{i}\binom{i}{k} = \binom{i-1}{k-1}.$$

Using the usual rule $\binom{m}{n} = 0$ if $n > m$, we have

$$\sum_{i=k+1}^{p-1} \frac{k}{i}\binom{i}{k} = \sum_{i=0}^{p-2}\binom{i}{k-1} - 1.$$

But $\sum_{i=0}^{p-2}\binom{i}{k-1} = \binom{p-1}{k}$, and as, modulo p, we have $\binom{p-1}{k} = (-1)^k$, we finally get, using the fact that k is odd, the desired identity. □

Now return to the proof of the lemma and suppose that k is even. If $k = 2$ then the process ends, so we can suppose that $k > 3$. The idea is to show that we can compute directly a_{k-1} and to deduce a_k from the first rule (we will still need to show the desired property). As k is even, $k-1$ is odd and thus $p-k+1$ is even. Thus by the third rule we have $a_{k-1} = -a_{p-k+1}$ and by the second rule we have

$$a_{p-k+1} = \frac{1}{2}a_{\frac{p-k+1}{2}}.$$

But there exists $j \in \mathbb{N}$ such that $p = k+j$ with $3 \leqslant j < p$ (because $k \leqslant p-3$). Hence, applying once again the third rule gives

$$a_{\frac{p-k+1}{2}} = -a_{p-\frac{p-k+1}{2}}.$$

But

$$p - \frac{p-k+1}{2} = \frac{p+k-1}{2} = k + \frac{j-1}{2}.$$

And as $j \geqslant 3$, we have $\frac{j-1}{2} \geqslant 1$, which means, applying the induction, that $a_{p-\frac{p-k+1}{2}}$ is already known. We then get the value of a_{k-1} and by applying the first rule to it we deduce the value of a_k. Now to finish the proof we need to show that, in this case $a_k = \frac{a_1}{k}$ modulo p. Notice that we can also assume by the induction that $a_i = \frac{a_1}{i}$ modulo p for all $i > k$. First we show that in the previous process, we get $a_{k-1} = \frac{a_1}{k-1}$ modulo p. Indeed, by the induction we have

$$a_{p-\frac{p-k+1}{2}} = \frac{a_1}{p - \frac{p-k+1}{2}}.$$

Thus

$$a_{\frac{p-k+1}{2}} = -\frac{a_1}{p - \frac{p-k+1}{2}}.$$

And finally

$$\begin{aligned} a_{k-1} &= -a_{p-k+1}, \\ &= \frac{1}{2}\left(\frac{a_1}{p-\frac{p-k+1}{2}}\right), \\ &= \frac{a_1}{k-1}. \end{aligned}$$

To conclude we need to prove a variant of Sub-Lemma 4.18.

Sub-Lemma 4.19. *Suppose that k is even, $a_i = \frac{a_1}{i}$ modulo p for all $i > k$ and $a_{k-1} = \frac{a_1}{k-1}$ modulo p. Then $a_k = \frac{a_1}{k}$ modulo p.*

Proof of Sub-Lemma 4.19. We have the equality

$$\frac{a_1}{k-1} = -\frac{1}{2}a_k k - \frac{1}{2}\sum_{i=k+1}^{p-1}\frac{a_1}{i}\binom{i}{k-1}.$$

Using the same arguments than in the proof of Sub-Lemma 4.18, we get the following identities,

$$\begin{aligned} \sum_{i=k+1}^{p-1}\frac{k-1}{i}\binom{i}{k-1} &= \sum_{i=0}^{p-2}\binom{i}{k-2} - \binom{k-1}{k-2} - \binom{k-2}{k-2}, \\ &= (-1)^{k-1} - (k-1) - 1, \\ &= -1-k, \text{ as } k \text{ is even.} \end{aligned}$$

And we finally have

$$\frac{a_1}{k-1} = -\frac{1}{2}a_k k + \frac{(1+k)a_1}{2(k-1)},$$

from which we deduce $a_k = \frac{a_1}{k}$. □

Hence the proof of Lemma 4.17 is complete. □

Back to the proof of Proposition 4.16. Suppose that $P(T) = \sum_{i=0}^{p-1} a_i T^i \in \mathbb{F}_p[T]$ verifies the conditions of the proposition. Then applying the three equations to P gives $a_0 = 0$, and the other coefficients a_i fulfill the rules described in the Lemma 4.17. □

***Remark* 4.20.** "Cohomological characterization of $\pounds_1$"
Kontsevich showed that $\pounds_1$ gives a non-zero 2-cocycle in $H^2(\mathbb{Z}/p, \mathbb{Z}/p)$. Since the latter group is isomorphic to $\mathbb{Z}/p$, this characterizes $\pounds_1$ up to a scalar.

4.4. **Space of solutions for equations associated to $\pounds_2$.** As $J(a,b,c)$ is the main relation for $\mathfrak{b}_3$, we can expect that it characterizes $\pounds_2$. In fact, we can first give a family of polynomials (which form a space of dimension growing linearly with p) and then characterize $\pounds_2$ by imposing also the duplication relation (i.e. the distribution relation for $\pounds_2$ with $m=2$). Since these two equations are consequences of the Kummer-Spence analogue, and the latter in turn is a consequence of $J(a,b,c)$, we are done.

Proposition 4.21. *The dimension of the $\mathbb{F}_p$-space of solutions associated to the equation*

$$T^p P\left(1-\frac{1}{T}\right) - P(T) + P(1-T) = 0 \tag{4.3}$$

grows with p and is at least of dimension $\frac{p-1}{3}+1$. *The family of polynomials*

$$\tau_{i,p}(T) = T^i(1-T)^i(T^{p-3i}+(-1)^i),$$

with $i \in \mathbb{N}$ *such that the valuation of* $\tau_{i,p}$ *is* $\geqslant 0$ *(for instance if* $i \leqslant \lfloor \frac{p}{2} \rfloor$*), is a solution of (4.3). Moreover, for* $i = 0, \dots, \frac{p-1}{3}$*, this family is free.*

Proof. The fact that $\tau_{i,p}$ fulfills (4.3) is a direct computation. For $i = 1, \dots, \frac{p-1}{3}$, the family is free for degree reasons, since $deg(\tau_{i,p}) = p - i$. Furthermore $\tau_{0,p}$ does not belong to this family for valuation reasons. □

***Remark* 4.22.**

1. We already know, by Lemma 3.11, that the inversion formula is a consequence of the 3-term equation. But a straightforward computation shows that the polynomials $\tau_{i,p}$ fulfill the inversion formula for $\pounds_2$.
2. In fact the rank of the family $\tau_{i,p}$ is greater than $\frac{p-1}{3}$, but the proof is a little bit more involved. We can also notice that $\pounds_2$ is never expressible in terms of $\tau_{i,p}$ if i runs only through $0, \dots, \frac{p-1}{3} - 1$.

Thus the 3-term equation is insufficient for the characterization of $\pounds_2$. Nevertheless, we have the following main result

Theorem 4.23. *Let* P *be a polynomial of* $\mathbb{F}[T]$ *of degree less than or equal to* $p-1$*. Set* $h = TP'$*. Then if* P *fulfills the duplication relation and the 3-term equation, and if moreover* h *fulfills the 2-term equation then* P *is equal, up to a multiplicative constant, to* $\pounds_2$*.*

Proof. Let P be a polynomial of degree $\leqslant p-1$, and suppose that P fulfills the following two equations

$$T^p P\left(1 - \frac{1}{T}\right) - P(T) + P(1-T) = 0, \tag{4.4}$$

$$2(1+T^p)P(T) + 2(1-T^p)P(-T) - P(T^2) = 0. \tag{4.5}$$

Then observing that we can deduce the inversion formula as a consequence of the 3-term, and taking the derivative with respect to these equations shows that h fulfills the inversion formula and the duplication formula. As, by hypothesis, h fulfills also the 2-term equation, we conclude from Proposition 4.16 that h is $\pounds_1$ up to a constant, which implies that P is $\pounds_2$ up to a constant. □

***Remark* 4.24.** We actually expect a slightly stronger result to be true, inasmuch as already the duplication and 3-term relation characterize $\pounds_2$; this claim has been verified for all primes $3 < p < 200$.

As we can formally deduce the two equations in the proposition from the Kummer-Spence analogue and the Kummer-Spence analogue in turn from the Cathelineau equation $J(a,b,c)$ (because in this case the specialisation mentioned in (3.15) is allowed), we get

Corollary 4.25. *The space of solutions of the Kummer-Spence analogue and the space of solutions of the Cathelineau equation are both of dimension* 1 *generated by* $\pounds_2$*.*

Proof. We only need to show that if $P \in \mathbb{F}[T]$, assumed to be of degree less than or equal to $p-1$, setting $h = TP'$, h fulfills the 2-term equation. In order to do that let $KS(a,b)$ denote the formal Kummer-Spence analogue, then taking the derivative with respect to a, and rewriting the equation with h and finally specializing to $a = 0$, we can see that, modulo the inversion formula for h (which we can get directly by deriving the inversion formula for P), we have the identity $h(b) = h(1-b)$. $\square$

5. Deriving functional equations : construction of the derivation map

The main goal of this section is to prove that one can pass from functional equations for polylogarithms to functional equations for the corresponding infinitesimal polylogarithms. For this purpose we will construct a family of maps, parametrized by a given derivation, from $\mathcal{B}_n(F)$ to $\beta_n(F)$. The origin of such maps comes from the categorical setting which is behind the "tangential processing" involved in the construction of the infinitesimal polylogarithmic groups, which is to some extent discussed in [3, 6, 9], and will be treated in more detail in [15].

In subsection 5.1, we present the "derivation map" from polylogarithmic groups to infinitesimal polylogarithmic groups. In subsection 5.2, we prove, as an application, that the derivation of a functional equation for any polylogarithm gives rise to a functional equation for the corresponding infinitesimal polylogarithm, and we will show several examples.

5.1. **From classical polylogarithmic groups to infinitesimal polylogarithmic groups.**

For the construction of the polylogarithmic groups (see section 2 on page 6), we gave an initial procedure for $n = 2$ and an inductive procedure for higher n. The construction of the "derivation map" follows this principle.

5.1.1. ***The case*** $n = 2$.

Lemma 5.1. *Let F be a field and $D \in Der_{\mathbb{Z}}(F)$ be an absolute derivation. Consider the well-defined maps $f_D : \mathbb{Z}[F^{\bullet\bullet}] \to F[F^{\bullet\bullet}]$, $[a] \mapsto D(a)[a]$ and $g_D : \bigwedge^2(F^\times) \to F^\times \otimes_{\mathbb{Z}} F$, $x \wedge y \mapsto -x \otimes \frac{D(y)}{y} + y \otimes \frac{D(x)}{x}$. Then the following diagram*

$$\begin{array}{ccc} \mathbb{Z}[F^{\bullet\bullet}] & \xrightarrow{f_D} & F[F^{\bullet\bullet}] \\ {\scriptstyle\delta_2}\downarrow & & \downarrow{\scriptstyle\bar\partial_2} \\ \bigwedge^2(F^\times) & \xrightarrow[g_D]{} & F^\times \otimes_{\mathbb{Z}} F, \end{array}$$

is commutative, where $\bar\partial_2([a]) = \frac{1}{a} \otimes \frac{1}{1-a} + \frac{1}{1-a} \otimes \frac{1}{a}$.

Proof. First we observe that the map g_D is well defined. Indeed this is a consequence of the $d\log$ property of the map $y \mapsto \frac{D(y)}{y}$ defined on the units and of the fact that $g_D(x \otimes x) = 0$ which implies that $g_D(x \wedge x) = 0$. Then, the commutativity of the diagram is a direct check. $\square$

As a direct consequence we get a map from $\ker(\delta_2)$ to $\ker(\bar\partial_2)$. Similarly, we can obtain a map $\ker(\delta_2)$ to $\ker(\partial_2)$ by replacing f_D by $\tilde f_D : [a] \mapsto \frac{D(a)}{a(1-a)}[a]$ which induces a map $\tau_{2,D} : \mathcal{B}_2(F) \to \beta_2(F)$.

5.1.2. ***The case*** $n > 2$. Suppose we have defined the "derivation map" $\tau_{n-1,D} : \mathcal{B}_{n-1}(F) \to \beta_{n-1}(F)$ (with respect to a derivation D) for the level $n-1$. Then we can construct $\tau_{n,D}$ by induction as follows.

Proposition 5.2. *Let $D \in Der_{\mathbb{Z}}(F)$ be an absolute derivation for the field F. Then we have the following commutative diagram:*

$$\begin{array}{ccc} \mathbb{Z}[F^{\bullet\bullet}] & \xrightarrow{\tilde{f}_D} & F[F^{\bullet\bullet}] \\ {\scriptstyle \delta_n}\downarrow & & \downarrow{\scriptstyle \partial_n} \\ \mathcal{B}_{n-1}(F) \otimes F^{\times} & \xrightarrow[g_{n,D}]{} & \beta_{n-1}(F) \otimes_{\mathbb{Z}} F^{\times} \oplus \mathcal{B}_{n-1}(F) \otimes_{\mathbb{Z}} F \end{array}$$

where $\tilde{f}_D$ is defined on generators as $[a] \mapsto \frac{D(a)}{a(1-a)}[a]$, while $g_{n,D}$ is given by

$$g_{n,D} : \{x\}_{n-1} \otimes y \mapsto \tau_{n-1,D}\big(\{x\}_{n-1}\big) \otimes y + \{x\}_{n-1} \otimes \frac{D(y)}{y}$$

and ∂_n by

$$\partial_n([a]) = \langle a \rangle_{n-1} \otimes a \ + \ \{a\}_{n-1} \otimes (1-a)\,.$$

***Remark* 5.3.** We want to point out that despite their apparent simplicity, these crucial commutative diagrams do not show up at first sight.

This induces a map from $\ker(\delta_n)$ to $\ker(\partial_n)$ which in turn induces the desired "derivation map" $\tau_{n,D} : \mathcal{B}_n(F) \to \beta_n(F)$.

Definition 5.4. *Let F be a field and $D \in Der_{\mathbb{Z}}(F)$ be an absolute derivation for the field F. We will call the map $\tau_{n,D} : \mathcal{B}_n(F) \to \beta_n(F)$ the derivation map from $\mathcal{B}_n(F)$ to $\beta_n(F)$, with respect to D. If x is an element of $\mathcal{B}_n(F)$, the element $\tau_{n,D}(x) \in \beta_n(F)$ will be called the derivative of x with respect to D.*

As usual, if D is clear from the context we will omit it.

***Remark* 5.5.** We can notice that all the $\tau_{n,D}$, and also all the maps involved in the previous propositions, give rise to an F-linear map $\tau_n : Der_{\mathbb{Z}}(F) \to Hom_{\mathbb{Z}}(\mathcal{B}_n(F), \beta_n(F))$ for all $n \geqslant 2$.

5.2. **Explicit derivation of functional equations.** As a consequence of the previous setting we get

Corollary 5.6. *Each element in* $\ker \delta_n$ *induces (many) elements in* $\ker \partial_n$*.*

The crucial main consequence is the following result.

Corollary 5.7. *Let K be an arbitrary field and set $F = K(t_1, \ldots, t_r)$, with $(t_1, \ldots, t_r)$ a transcendence basis over K. Let $D \in Der_{\mathbb{Z}}(F)$. Then any functional equation of the n-logarithm over K induces, via the derivation map $\tau_{n,D}$, a functional equation of the infinitesimal n-logarithm over K.*

Proof. It is a direct consequence of the definition 3.1 and of the construction of $\tau_{n,D}$. ☐

***Remark* 5.8.** Notice that, in the above corollary, $Der_{\mathbb{Z}}(F) \neq 0$ since $Der_K(F) \neq 0$, at least if $r \geqslant 1$. In practice it could be interesting to have a differential basis, and thus we can assume that if K is of characteristic p then $(t_1, \ldots, t_r)$ is a p-basis over K.

It is a priori not clear that the procedure gives non-trivial equations, but the following examples show that it is actually the case:

***Example* 5.9.** The first example is taken from [8] and it retrieves the 4-term relation from the 5-term relation (3.3) by applying the above procedure with

$$D = a(1-a)\frac{\partial}{\partial a} + b(1-b)\frac{\partial}{\partial b}\,,$$

assuming that $F = K(a,b)$ with a, b indeterminates over the field K, and that $\frac{\partial}{\partial a}$ and $\frac{\partial}{\partial b}$ are the usual partial derivatives.

The following proposition gives a partial answer to Cathelineau's question concerning the relationship of his 22-term equation for $d\mathcal{D}_3$ and Goncharov's equation (3.7) for $\mathcal{D}_3$ (with the same number of terms). It is a consequence of the previous results but can also be verified directly.

Proposition 5.10. 1. *The infinitesimal functional equation below, which is derived from the Goncharov functional equation for the trilogarithm is zero in* $\beta_3(F)$ *.*

2. *If* $F \supset \mathbb{Q}$*, the Goncharov equation is expressible in terms of an* F*-linear combination of* $J(a,b,c)$ *.*

We give an example of such a derived version in the case $F = K(a,b,c)$ with a, b, c indeterminates over the field K, applying the above procedure with

$$D = a(1-a)\frac{\partial}{\partial a} + b(1-b)\frac{\partial}{\partial b} + c(1-c)\frac{\partial}{\partial c}$$

to the equation stated in (3.7). Let us set

$$\begin{aligned}\varphi(a,b,c) =& [a] - \frac{(b-1)(a-1)}{ab-1}\left(\left[-\frac{b(a-1)}{b-1}\right] + \left[-\frac{a(b-1)}{a-1}\right]\right)\\ &+ \frac{(c^2b+cb^2-3cb+1)}{cb-1}\left[\frac{a-1}{abc-1}\right] - \frac{(abc-a-b-c+2)}{cb-1}\left[\frac{cb(a-1)}{abc-1}\right]\\ &- \frac{(a+b-2)}{ab-1}[ab] - \frac{(a^2bc-2abc+b+c-1)(a-1)}{(ac-1)(ab-1)}\left[-\frac{a(c-1)(b-1)}{(a-1)(abc-1)}\right].\end{aligned}$$

Then, modulo the inversion formula,

$$\varphi(a,b,c) + \varphi(b,c,a) + \varphi(c,a,b) - \frac{(a+b+c-3)}{abc-1}[abc]$$

is the differential of the Goncharov equation and vanishes in $\beta_3(F)$ by virtue of Corollary 5.6.

***Observation* 5.11.** We should notice that we have not yet proved that the infinitesimal Goncharov equation also holds in characteristic p and to know that this equation is expressible in terms of an F-linear combination of $J(a,b,c)$ is not enough to ensure this (unless we know that this linear combination is independant of F). It will be seen in the next section that it is the case, at least if we see $\pounds_2$ as a function from $\mathbb{Z}/p$ to $\mathbb{Z}/p$.

6. Reduction of functional equations mod p via the p-adic realm

In this section, we want to prove the following statements (which are made more precise below):

❶ Each functional equation for the classical n-logarithm $\mathcal{D}_n$ induces a functional equation for certain p-adic n-logarithm functions (those which satisfy Wojtkowiak's p-adic version of Zagier's criterion).

❷ Each functional equation for the classical n-logarithm induces a functional equation for the infinitesimal n-logarithm (via the derivation procedure given in the previous section). A similar statement holds for the p-adic case.

❸ Each functional equation for the infinitesimal n-logarithm $d\mathcal{D}_n$ induces a functional equation for the corresponding p-adic infinitesimal n-logarithm denoted DF_n (see Definition 6.6).

❹ Each "good $\mathbb{Q}_p$-specialization", as defined in (6.10) below, of a functional equation for the p-adic infinitesimal polylogarithm induces a functional equation (in the weak sense) for the finite $(n-1)$-logarithm.

Combining the four statements, we arrive at the somewhat more surprising statement:

Surprise: *Each functional equation for the classical n-logarithm induces a functional equation for the finite $(n-1)$-logarithm.*

Throughout this section, we denote by $Li_n(z)$ Coleman's p-adic n-logarithm [10]. Let us first look for the p-adic combinations which should play the same role as the modified polylogarithms $\mathcal{D}_n$.

***Remark* 6.1.** The inversion relation $\big($in its clean form $P_n(z) = (-1)^{n-1}P_n(1/z)\big)$ for a combination $P_n(z) = \sum_{k=0}^{n-1} a_k \log^k(z) Li_{n-k}(z)$ is equivalent to the following condition on the coefficients:

$$\sum_{k=0}^{n-1} \frac{a_k}{(n-k)!} = 0 \tag{6.1}$$

(cf. [34], Lemma 4.2). Since the inversion relation is in the kernel of ∂_n, we can restrict our investigations to combinations $P_n(z)$ satisfying those conditions.

While one needs to work harder in the "classical" case to find functions which satisfy cleanly their functional equations, it turns out that in the p-adic case the above condition is already good enough, and we can state the above claim ❶ more precisely as

Proposition 6.2. (Wojtkowiak, [34], Proposition 4.4)
Let $\xi \in \ker \delta_{n,\mathbb{Q}_p(t_1,\dots,t_r)}$. Then each admissible $\mathbb{C}_p$-specialization of ξ is mapped to a constant by the p-adic functions

$$P_n(z) = \sum_{k=0}^{n-1} a_k \log^k(z) Li_{n-k}(z)\,, \tag{6.2}$$

if the coefficients satisfy condition (6.1).

This motivates the following definition:

Definition 6.3. *A linear combination of p-adic polylogarithms of the form* (6.2) *whose coefficients satisfy* (6.1) *is called a* clean p-adic polylogarithm.

***Remark* 6.4.**
1. For $n=2$, there is, up to a multiplicative constant, only one clean p-adic 2-logarithm P_2 satisfying (6.1).
2. The original statement was actually somewhat stronger: $\mathbb{Q}_p(t_1,\dots,t_r)$ was replaced by $\mathbb{C}_p(t)$, where $\mathbb{C}_p$ denotes a completion of an algebraic closure of $\mathbb{Q}_p$.

The claim in ❷ follows immediately from the "derivation map" in §5.

Before we show a more precise version of ❸ by imitating Proposition 7 of [8], we state our intermediate goal: We are looking for a morphism $F[F]\to F_0$, where $F=\mathbb{C}_p(z)$ and $F_0=\mathbb{C}_p$. More precisely, we want to have a family of morphisms $(DP_n)_{n\geqslant 2}$ on $\beta_n(\mathbb{C}_p)$ expressed in terms of the differential operator $D=z(1-z)\frac{d}{dz}$ and some clean p-adic polylogarithms P_n. There are many candidates:

Proposition 6.5. *Let $(P_n)_{n\geqslant 2}$ be a family of clean p-adic polylogarithms such that for $n\geqslant 3$*

$$DP_n(z)=\lambda_n\,(1-z)P_{n-1}(z)\;+\;\mu_n\,\log(z)\,DP_{n-1}(z)\,, \tag{6.3}$$

for some $\lambda_n,\ \mu_n\in\mathbb{C}_p^\times$.
Then, for any n, DP_n defines a morphism on $\beta_n(\mathbb{C}_p)$.

Proof. P_n is defined on $\mathcal{B}_n(\mathbb{C}_p)$ by assumption. For $n=2$, we have seen that the function is essentially unique:

$$P_2(z)=-2\,Li_2(z)+\log(z)Li_1(z)\,,$$

and the resulting infinitesimal dilogarithm

$$DP_2(z)=(1-z)\log(1-z)+z\log(z)$$

vanishes on $r_2(\mathbb{C}_p)$ (due to Proposition 2.8, it is enough to check that it vanishes on the four term relation, which is straightforward).
Now suppose the claim is true for $n-1$. The maps $DP_{n-1}\otimes\log:\beta_{n-1}(\mathbb{C}_p)\otimes\mathbb{C}_p^\times\to\mathbb{C}_p$, $x\langle y\rangle_{n-1}\otimes z\mapsto xDP_{n-1}(y)\log(z)$ resp. $P_{n-1}\otimes Id:\mathcal{B}_{n-1}(\mathbb{C}_p)\otimes\mathbb{C}_p\to\mathbb{C}_p$, $\{y\}_{n-1}\otimes z\mapsto zP_{n-1}(y)$, are well-defined by the inductive assumption resp. by assumption (P_{n-1} is clean). Furthermore, an element $\xi\in r_n(\mathbb{C}_p)$ lies in the kernel of each of the "components" of ∂_n, say $\partial_n':\mathbb{C}_p[\mathbb{C}_p]\to\beta_{n-1}(\mathbb{C}_p)\otimes\mathbb{C}_p^\times$ and $\partial_n'':\mathbb{C}_p[\mathbb{C}_p]\to\mathcal{B}_{n-1}(\mathbb{C}_p)\otimes\mathbb{C}_p$, and therefore

$$\Big(\mu_n\,DP_{n-1}\otimes\log\;+\lambda_n\;P_{n-1}\otimes Id\Big)\big(\partial_n\xi\big)=\Big(\mu_n\,DP_{n-1}\otimes\log\circ\partial_n'+\lambda_n\;P_{n-1}\otimes Id\circ\partial_n''\Big)\xi=0\,,$$

which shows that the function defined by (6.3) can be linearly extended to a well-defined function on $\beta_n(\mathbb{C}_p)$. □

Definition 6.6. *Besser's p-adic n-logarithm is defined as*

$$F_n(z)=\sum_{k=0}^{n-1}a_{k,n}\log^k(z)Li_{n-k}(z) \tag{6.4}$$

with

$$a_{k,n}=\frac{(-1)^k}{k!}(k-n)\,.$$

We will call DF_n the distinguished infinitesimal p-adic n-logarithm.

Proposition 6.7 (Existence)**.** *There exist families of clean p-adic polylogarithms satisfying* (6.3) *for some λ_n, $\mu_n \in \mathbb{C}_p^\times$.*
In particular, Besser's family (6.4) *satisfies* (6.3) *with* $(\lambda_n, \mu_n) = (\frac{1}{n-1}, -\frac{1}{n-1})$*,* $n \geqslant 3$.
There are many other possibilities.

Proof. Again, the case $n = 2$ gives the unique choice for P_2 (up to multiplicative constant). Inductively, starting from P_{n-1} and DP_{n-1}, one can form an arbitrary linear combination of them using λ_n and μ_n which gives a candidate for DP_n, with coefficients $b_{k,n}$, say; a subsequent "integration" (putting $a_{0,n} = -n$ and successively $a_{k+1,n} = -n(b_{kn} - a_{kn})/(k+1)$, $k = 0, \ldots, n-2$) provides a candidate P_n whose coefficients a_{kn} have to satisfy the further condition (6.1)—this gives a linear restriction on the possible (λ_n, μ_n) at each step. We thus obtain inductively an extra degree of freedom at each level.
For example, normalizing $P_n(z)$ such that $a_0 - n$, we obtain successively

$$\lambda_3 - \mu_3 = 1\,, \qquad \lambda_4 - \mu_4 = \frac{1}{2 - \lambda_3}\ , \text{ etc.}$$

It remains to check that Besser's choice (6.4) does satisfy

$$(n-1)\,DF_n(z) = (1-z)\,F_{n-1}(z) - \log(z) DF_{n-1}(z) \tag{6.5}$$

which is straightforward. Also, the $a_{k,n}$ satisfy condition (6.1) since

$$-\sum_{k=0}^{n-1} \frac{(-1)^k}{k!(n-k)!}(n-k) = \frac{1}{(n-1)!}(1-1)^{n-1} = 0\,.$$

□

***Remark* 6.8.** 1. Writing $\Phi_n(z) = (n-1)! F_n(z)$ and noticing that $(1-z) = D\log(z)$, we can reformulate (6.5) more suggestively, using the ad-hoc convention $D^-(a \otimes b) := D(a)b - aD(b)$, as

$$D\Phi_n(z) = D^-\big(\log(z) \otimes \Phi_{n-1}(z)\big)\,.$$

2. We have just seen that, a priori, there are many choices for the P_n individually, but the condition that the morphisms at level n and $n-1$ be linked via the condition $\rho\, DP_n(z) = (1-z)P_{n-1}(z) - \log(z) DP_{n-1}(z)$ for some $\rho \in \mathbb{C}_p$ provides us with a unique function, up to a multiplicative factor, the condition (6.1) still being true for P_n. We have not found a "natural" justification for the condition (6.5), though. A normalization condition for the above P_n is then $a_{0,n} + a_{1,n} = -1$ which entails $\rho = n-1$. The resulting family coincides with Besser's functions (6.4)—his choice of coefficients was forced by two rather natural requirements: first, a certain p-adic power series expansion becomes independent of the "direction" in which to expand; second, one retrieves the finite $(n-1)$-logarithm by reducing DF_n mod p^n (or, more precisely, reducing $p^{1-n}DF_n$ mod p) on elements in $\mathbb{Z}_p^\times \cap (1-\mathbb{Z}_p)^\times \subset \mathbb{C}_p$ (for an improved statement of this and of the following theorem cf. [2]).

The F_n can be characterized by the following

Theorem 6.9. (Besser, [2], Theorem 1.1)
Let $X = \{z \in \mathbb{Z}_p\,, |z| = |1-z| = 1\}$*. For* $p > n+1$*, one has* $DF_n(\mathbb{Z}_p) \subset p^{n-1}\mathbb{Z}_p$*, and for* $z \in X$*:*

$$p^{1-n} DF_n(z) \equiv \pounds_{n-1}(z) \pmod{p}\,.$$

The choice of coefficients (in $\mathbb{Q}$) for F_n is unique for a clean p-adic polylogarithm which satisfies the above property for all $p > n+1$.

In order to formulate the subsequent statements conveniently, we introduce the following notion:

Definition 6.10. *A* good $\mathbb{Q}_p$-specialization for $\sum n_i[x_i] \in F[F]$, $F \subset \mathbb{Q}_p(t_1,\dots,t_r)$, *is a family of numbers $u_j \in \mathbb{Q}_p$, $j = 1,\dots,r$, such that the images of $n_i = n_i(t_1,\dots,t_r)$, $x_i = x_i(t_1,\dots,t_r)$ and $1-x_i = 1-x_i(t_1,\dots,t_r)$ under the specialization map $t_j \mapsto u_j$, $j = 1,\dots,r$, are in $\mathbb{Z}_p^\times$.*

The virtue of a good $\mathbb{Q}_p$-specialization lies in the fact that we can reduce it modulo $p\mathbb{Z}_p$. As we can notice, a good $\mathbb{Q}_p$-specialization is, in particular, an admissible $\mathbb{Q}_p$-specialization. Now, putting Proposition 6.5 and Theorem 6.9 together, we can make ❹ more precise:

Corollary 6.11. *Let $n \geqslant 2$, $p > n+1$, and $\eta \in \ker \partial_{n,\mathbb{Q}_p(t_1,\dots,t_r)}$. Then we have*

a) *For each admissible $\mathbb{C}_p$-specialization η^{spec} for η, $DF_n(\eta^{\mathrm{spec}}) = 0$.*
b) *For each good $\mathbb{Q}_p$-specialization η^{spec} for η, the reduction mod p gives*

$$\pounds_{n-1}(\eta^{\mathrm{spec}}) \equiv 0 \pmod{p}\,.$$

Proof. The infinitesimal polylogarithm DF_n vanishes on η by Proposition 6.5, and reducing mod p obviously conserves this vanishing property. Besser's result now says that the reduction of $p^{1-n}DF_n(\eta^{\mathrm{spec}})$ is equal to $\pounds_{n-1}\big(\eta^{\mathrm{spec}} \pmod{p}\big)$. □

Going even one step further, we can state a more precise version of the above "surprise":

Corollary 6.12. *Let $n \geqslant 2$, $p > n+1$, and $\xi \in \ker \delta_{n,\mathbb{Q}(t_1,\dots,t_r)}$. Then we have*

a) *For each admissible $\mathbb{C}$-specialization resp. $\mathbb{C}_p$-specialization ξ^{spec} for ξ, the quantities $\mathcal{D}_n(\xi^{\mathrm{spec}})$ resp. $F_n(\xi^{\mathrm{spec}})$ are constants.*
b) *For each absolute derivation $\Delta \in Der_{\mathbb{Z}}(\mathbb{Q}(t_1,\dots,t_r))$, ξ induces $\xi_\Delta \in \ker \partial_{n,\mathbb{Q}(t_1,\dots,t_r)}$, and therefore, for each admissible $\mathbb{C}$-specialization resp. $\mathbb{C}_p$-specialization,*

$$d\mathcal{D}_n(\xi_\Delta) = 0, \quad \textit{resp.} \quad DF_n(\xi_\Delta) = 0\,.$$

c) *For each good $\mathbb{Q}_p$-specialization $\xi_\Delta^{\mathrm{spec}}$ for ξ_Δ, the reduction mod p gives*

$$\pounds_{n-1}(\xi_\Delta^{\mathrm{spec}}) \equiv 0 \pmod{p}\,.$$

Proof. a) Follows from Zagier [36] and Wojtkowiak [34], respectively.
b) This follows via the "derivation map" (see §5).
c) $0 = p^{1-n}DF_n(\xi_\Delta) \equiv \pounds_{n-1}(\xi_\Delta^{\mathrm{spec}})$.

□

Alas, although being quite powerful, the above strategy does not give the full answer to our problem.

***Remark* 6.13.** 1. The virtues of the procedure described above lie in its generality: we do not need to (find and) prove functional equations for (p-adic) infinitesimal or finite polylogs, since they "drop out" using the machinery.

2. The drawbacks of the machinery lie in its lack of control:
   (a) We do not get the functional equations as polynomial identities but only "on points", i.e. in the form of (good) specializations.

(b) A more mundane reason for proving functional equations for $\pounds_n$ in the strong sense is the fact that all the ones which have occurred in our investigations are not only true for $\mathbb{F}_p$ but actually hold more generally for any field of characteristic p.

(c) (a minor point, given the range in which we mostly work) We need to assume that $p > n+1$.
This restriction is not (always) necessary for the polynomial identities to hold: there are examples of equations for $\pounds_3$ which are still true in characteristic 3.

In summary, there are still plenty of reasons which leave us with the task of finding proofs of functional equations for the finite polylogarithms. The final section will therefore be dedicated to this issue.

Part III. The Main Proofs

7. Proofs of functional equations over fields of characteristic p.

7.1. Straightforward demonstrations.

Proof. (of Proposition 4.7)

1. The inversion relation can be checked via a straightforward algebraic manipulation.
2. In order to prove the distribution relation, let us fix a primitive mth root of unity ζ. Dividing both sides by m^n and developing the fraction into a (finite) series leaves us to prove:

$$\begin{aligned}\sum_{k=1}^{p-1}\frac{T^{km}}{(km)^n} &= \frac{1}{m}\sum_{\zeta^m=1}\left(1+(\zeta T)^p+(\zeta T)^{2p}+\cdots+(\zeta T)^{(m-1)p}\right)\sum_{k=1}^{p-1}\frac{(\zeta T)^k}{k^n}\\ &= \frac{1}{m}\sum_{k=1}^{p-1}\frac{1}{k^n}\sum_{\zeta^m=1}\left((\zeta T)^k+(\zeta T)^{p+k}+(\zeta T)^{2p+k}+\cdots+(\zeta T)^{(m-1)p+k}\right)\\ &= \frac{1}{m}\sum_{k=1}^{p-1}\sum_{\zeta^m=1}\left(\frac{(\zeta T)^k}{k^n}+\frac{(\zeta T)^{p+k}}{(p+k)^n}+\frac{(\zeta T)^{2p+k}}{(2p+k)^n}+\cdots+\frac{(\zeta T)^{(m-1)p+k}}{\left((m-1)p+k\right)^n}\right)\\ &= \frac{1}{m}\sum_{\substack{r=1\\ p\nmid r}}^{pm-1}\Big(\sum_{\zeta^m=1}\zeta^r\Big)\frac{T^r}{r^n},\end{aligned}$$

and this is true due to the character relations

$$\sum_{\zeta^m=1}\zeta^r = \begin{cases} m, & \text{if } m|r\\ 0, & \text{otherwise.}\end{cases}$$

3. (Proof of the special values) $\widetilde{\pounds_n}(1)=0$ if $(p-1)\nmid n$ follows from the well-known fact that $\sum_{k=0}^{p-1}P(k)=0$ for any polynomial $P\in\mathbb{Z}/p\mathbb{Z}[x]$ of degree $\leqslant p-2$ (here we apply it to the monomials $x,\ldots,x^{p-2}$), the statement for $(p-1)|n$ being obvious.
The assertion for $\widetilde{\pounds_{2n}}(-1)=0$ is a direct consequence of the inversion relation.

To prove the last formula of Proposition 4.7 we only need to take $m = 2n$ (the odd values correspond to the above identities). For this, one can use the special case $a = 2$, in [27](Proposition (5B), p.108), $p-1 \not| 2n$:

$$(1 - 2^{2n})B_{2n} \equiv 2n\, 2^{2n-1} \sum_{1 \leqslant j < \frac{p}{2}} \frac{1}{j^{1-2n}} \pmod p \tag{7.1}$$

and the fact that $£_{1-2n}(-1)$ is equal to the sum in (7.1): rewriting

$$\widetilde{£_{p-2n}}(-1) = \widetilde{£_{1-2n}}(-1) = \sum_{j=1}^{(p-1)/2} (2j)^{2n-1} - \sum_{j=1}^{(p-1)/2} (2j-1)^{2n-1} = 2 \sum_{j=1}^{(p-1)/2} (2j)^{2n-1} - \sum_{j=1}^{p-1} j^{2n-1}\,,$$

one sees that the first sum is equal to 2^{2n} times the sum in (7.1), while the second one equals $-\widetilde{£_{1-2n}}(1)$ and therefore is zero (for $0 < n < \frac{p-1}{2}$) by the above special value.

□

7.2. **A recipe for proving functional equations.** Let R be a domain of characteristic p. In order to show that a polynomial $Q(T) \in R[T]$ is zero, we divide it into three parts:

$$Q(T) = Q(0) + Q_1(T) + Q_2(T^p)\,,$$

where $Q_1(T)$ involves only powers of T which are not divisible by p. Then we verify separately that $Q_2(T^p)$ and the constant $Q(0)$ vanish and that $\frac{d}{dT}Q_1(T)$ is zero as well. We can iterate this procedure in an obvious way.

Proof of Proposition 4.9. 1. We will apply the recipe above. We have

$$\begin{aligned}\frac{d}{dT}£_1(1-T) &= -\frac{1}{1-T}£_0(1-T),\\ &= \frac{1}{T}£_0(T) \quad \text{by (4.1)},\\ &= \frac{d}{dT}£_1(T),\end{aligned}$$

and as the degree of either polynomials is less than $p-1$, we conclude that $£_1(T) = £_1(1-T) + c$ where c is a constant.
This, in turn, implies that $2c = 0$ (specialize $T = 0$ and $T = 1$, respectively), and therefore we get as a by-product $£_1(1) = £_1(0) = 0$ (in characteristic $\neq 2$).

2. The following proof is a slight variation of the recipe, in that it uses two iterated derivatives.
Denote by ∂_x and ∂_y the derivatives to respect to x and y. We can check, using the differential equation for $£_1$ and the rational expression (..) for $£_0$, that

$$\partial_y \partial_x H(x, y, s) = \frac{1 - y^p - x^p + s^p}{(1 - y - x + s)^2},$$

which is an expression that is symmetric in x and y. Thus

$$\partial_y \partial_x (H(x, y, s) - H(y, x, s)) = 0.$$

But the maximum degree for each indeterminate in the polynomial $H(x,y,s)$ is never greater than $p-1$, and as a consequence the above identity implies that

$$H(x,y,s)-H(y,x,s)=R_0(s)+R_1(s)x+R_2(s)y,$$

where $R_0,R_1,R_2\in\mathbb{F}[s]$. But setting $x=y$ implies both $R_0=0$ and $R_1+R_2=0$, and the construction of R_1 and R_2 shows directly that they are both zero $\big($the coefficients of x and y in $H(x,y,s)$ are both equal to $\sum_{k=0}^{p-2}(-s)^k\big)$.

□

Proof of Proposition 4.11. 1. Set

$$E(T)=£_2(1-T)-£_2(T)+T^p£_2(1-\frac{1}{T}).$$

We want to prove that E is 0 in $\mathbb{F}[T]$. Computing $\frac{d}{dT}E$ we get

$$\frac{d}{dT}E(T)=-\frac{1}{1-T}£_1(1-T)-\frac{1}{T}£_1(T)+\frac{T^{p-1}}{T-1}£_1(1-\frac{1}{T}).$$

But by Proposition 4.9, $£_1(1-T)=£_1(T)$ and $£_1(1-\frac{1}{T})=£_1(\frac{1}{T})$. Moreover by the inversion formula (see Proposition 4.7) we have $£_1(\frac{1}{T})=-\frac{1}{T^p}£_1(T)$. Hence,

$$\begin{aligned}\frac{d}{dT}E(T)&=-\frac{1}{1-T}£_1(T)-\frac{1}{T}£_1(T)-\frac{1}{(T-1)T}£_1(T),\\&=0.\end{aligned}$$

As $E(0)=0$ and $\deg(E)\leqslant p$, we know that $E(T)=cT^p$ and therefore $T^pE(\frac{1}{T})=c$, but using the inversion relation one sees that $T^pE(\frac{1}{T})=E(T)$, which implies $c=0$.

□

***Remark* 7.1.** A different way to prove that $c=0$: For this we look directly at $E(T)$ and try to compute this coefficient which can only appear in the expression

$$T^p£_2(1-\frac{1}{T})=\sum_{i=1}^{p-1}\frac{T^{p-i}(T-1)^i}{i^2}.$$

But, for each i, the coefficient of T^p is $\frac{1}{i^2}$, and thus $c=£_2(1)=0$.

Proof of Theorem 4.12. The strategy of proof could be summarized as follows:

(i) Prove that $\partial_c\widehat{£_2}(J(a,b,c))=0$ in $\mathbb{F}[a,b,c]$.
(ii) Prove that $\widehat{£_2}(J(a,b,0))=0$ in $\mathbb{F}[a,b]$.
(iii) Prove that the coefficient of c^p in $\widehat{£_2}(J(a,b,c))$ is 0.

For the proof of this functional equation we will need several preliminary formulas. First we will use these two relations, in $\mathbb{F}[x,y]$, coming from the 4-term equation for $£_1$

$$(1-y)^p£_1\left(\frac{x}{1-y}\right)=£_1(x)+(1-x)^p£_1\left(\frac{y}{1-x}\right)-£_1(y),\tag{7.2}$$

$$£_1(y)-£_1(x)=(1-x)^p£_1\left(\frac{1-y}{1-x}\right)+x^p£_1\left(\frac{y}{x}\right).\tag{7.3}$$

We use implicitly the following formal derivation rules, where t is an indeterminate and λ a constant independent of t:

$$\frac{d}{dt}\pounds_2\big(\lambda(1-t)\big) = -\frac{1}{1-t}\,\pounds_1(\lambda(1-t)),$$
$$\frac{d}{dt}\pounds_2\left(\lambda\Big(1-\frac{1}{t}\Big)\right) = -\frac{1}{t(1-t)}\,\pounds_1\left(\lambda\Big(1-\frac{1}{t}\Big)\right).$$

We also point out that the following simple formula will be often used:

$$\frac{1}{t}+\frac{1}{1-t}=\frac{1}{t(1-t)}.$$

For the convenience of the reader we will give detailled computations in order to make checking almost straightforward.

Let's first split $\widehat{\pounds_2}(J(a,b,c))$ into six pieces to facilitate the identification of the cancellation in the forthcoming computations:

$$\begin{aligned}
A_1 &= c^p\pounds_2(a) - c^p\pounds_2(b) + (a-b+1)^p\pounds_2(c)\\
&\quad + (1-c)^p\pounds_2(1-a) - (1-c)^p\pounds_2(1-b) + (b-a)^p\pounds_2(1-c),\\
A_2 &= -a^p\pounds_2\left(\frac{c}{a}\right) + b^p\pounds_2\left(\frac{c}{b}\right) + c^pa^p\pounds_2\left(\frac{b}{a}\right)\\
&\quad - (1-a)^p\pounds_2\left(\frac{1-c}{1-a}\right) + (1-b)^p\pounds_2\left(\frac{1-c}{1-b}\right) + c^p(1-a)^p\pounds_2\left(\frac{1-b}{1-a}\right),\\
A_3 &= c^p(1-a)^p\pounds_2\left(\frac{a(1-c)}{c(1-a)}\right) - c^p(1-b)^p\pounds_2\left(\frac{b(1-c)}{c(1-b)}\right),\\
A_4 &= -b^p\pounds_2\left(\frac{ca}{b}\right) - (1-b)^p\pounds_2\left(\frac{c(1-a)}{1-b}\right),\\
A_5 &= -(a-b)^p\pounds_2\left(\frac{(1-c)a}{a-b}\right) - (b-a)^p\pounds_2\left(\frac{(1-c)(1-a)}{b-a}\right)\\
&\quad + c^p(a-b)^p\pounds_2\left(\frac{(1-c)b}{c(a-b)}\right) + c^p(b-a)^p\pounds_2\left(\frac{(1-c)(1-b)}{c(b-a)}\right),\\
A_6 &= (1-c)^pa^p\pounds_2\left(\frac{a-b}{a}\right) + (1-c)^p(1-a)^p\pounds_2\left(\frac{b-a}{1-a}\right).
\end{aligned}$$

Set $d=\frac{\partial}{\partial c}$.
First step: prove that $\sum dA_i = 0$.

It is immediate that $dA_6 = 0$. Using the rules described above, we get the following equalities:

$$\begin{aligned} dA_1 &= \frac{1}{c}\pounds_1(c) + \frac{(a-b)^p}{c(1-c)}\pounds_1(c), \\ dA_2 &= -\frac{a^p}{c}\pounds_1\left(\frac{c}{a}\right) + \frac{b^p}{c}\pounds_1\left(\frac{c}{b}\right) \\ &\quad + \frac{(1-a)^p}{1-c}\pounds_1\left(\frac{1-c}{1-a}\right) - \frac{(1-b)^p}{1-c}\pounds_1\left(\frac{1-c}{1-b}\right), \\ dA_3 &= -\frac{c^p(1-a)^p}{c(1-c)}\pounds_1\left(\frac{a(1-c)}{c(1-a)}\right) + \frac{c^p(1-b)^p}{c(1-c)}\pounds_1\left(\frac{b(1-c)}{c(1-b)}\right), \\ dA_4 &= -\frac{b^p}{c}\pounds_1\left(\frac{ca}{b}\right) - \frac{(1-b)^p}{c}\pounds_1\left(\frac{c(1-a)}{1-b}\right), \\ dA_5 &= \frac{(a-b)^p}{1-c}\pounds_1\left(\frac{(1-c)a}{a-b}\right) + \frac{(b-a)^p}{1-c}\pounds_1\left(\frac{(1-c)(1-a)}{b-a}\right) \\ &\quad - \frac{c^p(a-b)^p}{c(1-c)}\pounds_1\left(\frac{(1-c)b}{c(a-b)}\right) - \frac{c^p(b-a)^p}{c(1-c)}\pounds_1\left(\frac{(1-c)(1-b)}{c(b-a)}\right). \end{aligned}$$

Then, applying consecutively (7.2) to dA_5, with $x = 1-c$, $y = \frac{b}{a}$, with $x = 1-\frac{1}{c}$, $y = \frac{b}{a}$, and with $x = 1-\frac{1}{c}$, $y = \frac{1-a}{1-b}$, and to dA_3 with $x = 1-\frac{1}{c}$, $y = \frac{1}{a}$, and using (7.3) for simplification as well as the basic relations for $\pounds_1$, we get

$$\begin{aligned} dA_4 + dA_5 + dA_3 &= -\frac{\pounds_1(b)}{1-c} + \frac{\pounds_1(a)}{1-c} \\ &\quad + \frac{c^p}{c(1-c)}\left(\pounds_1\left(\frac{b}{c}\right) - \pounds_1\left(\frac{a}{c}\right)\right) - \pounds_1(c) + \frac{(b-a)^p}{c(1-c)}\pounds_1(c), \end{aligned}$$

then

$$dA_1 + dA_4 + dA_5 + dA_3 = -\frac{\pounds_1(b)}{1-c} + \frac{\pounds_1(a)}{1-c} + \frac{c^p}{c(1-c)}\left(\pounds_1\left(\frac{b}{c}\right) - \pounds_1\left(\frac{a}{c}\right)\right).$$

It remains to transform dA_2, but using (7.3), we have e.g.

$$(1-b)^p\pounds_1\left(\frac{1-c}{1-b}\right) = \pounds_1(c) - \pounds_1(b) - b^p\pounds_1\left(\frac{c}{b}\right),$$

then

$$dA_2 = \frac{\pounds_1(b)}{1-c} - \frac{\pounds_1(a)}{1-c} + \frac{b^p}{c(1-c)}\pounds_1\left(\frac{c}{b}\right) - \frac{a^p}{c(1-c)}\pounds_1\left(\frac{c}{a}\right).$$

Now by invoking the inversion formula we see that

$$\sum_{i=1}^{5} dA_i = 0.$$

Second step: *Prove that the relation is true for $c = 0$.*
Putting $c = 0$ in $\sum_{i=1}^{6} A_i$ gives

$$\pounds_2(1-a) - \pounds_2(1-b) - (1-a)^p \pounds_2\left(\frac{1}{1-a}\right) + (1-b)^p \pounds_2\left(\frac{1}{1-b}\right)$$
$$a^p \pounds_2\left(\frac{a-b}{a}\right) + (1-a)^p \pounds_2\left(\frac{b-a}{1-a}\right) - (a-b)^p \pounds_2\left(\frac{a}{a-b}\right) - (b-a)^p \pounds_2\left(\frac{1-a}{b-a}\right)$$

and applying the inversion formula for $\pounds_2$ we get 0.

Third step: *Prove that the coefficient of c^p is 0.*
Notice first that if λ is an expression independent of c, then the coefficient of c^p in the sum $\sum_{i=1}^{p-1} \frac{\lambda^i}{i^2} c^{p-i}(1-c)^i$ is $\pounds_2(-\lambda)$. Using this fact, we can see that the coefficient of c^p in the expression $\sum_{i=1}^{6} A_i$ is given by

$$\begin{aligned}
&\pounds_2(a) - \pounds_2(1-a) + (1-a)^p \pounds_2\left(\frac{-a}{1-a}\right) \\
&- \pounds_2(b) + \pounds_2(1-b) - (1-b)^p \pounds_2\left(\frac{-b}{1-b}\right) \\
&+ a^p \pounds_2\left(\frac{a}{b}\right) - a^p \pounds_2\left(\frac{a-b}{a}\right) + (a-b)^p \pounds_2\left(\frac{-b}{a-b}\right) \\
&+ (1-a)^p \pounds_2\left(\frac{1-b}{1-a}\right) - (1-a)^p \pounds_2\left(\frac{b-a}{1-a}\right) + (b-a)^p \pounds_2\left(-\frac{1-b}{b-a}\right).
\end{aligned}$$

But each of the previous lines are 0 by using the 3-term functional equation of $\pounds_2$ (see Proposition 4.12 1.) and this completes the proof of the 22-term functional equation for $\pounds_2$. □

Remark 7.2. We want to stress some more structural properties in the rather computational parts of the previous proof—thereby also giving an indication that there should exist a common proof for both the finite and the infinitesimal case:

(i) we first use the ("$d\log$-like") behaviour (cf. the comment after (Definition 4.4))
$$\frac{d}{dc}\widehat{\pounds}_m\Big(c^\alpha(1-c)^\beta\Big) = \Big(\frac{\alpha}{c} - \frac{\beta}{1-c}\Big)\widehat{\pounds}_{m-1}\Big(c^\alpha(1-c)^\beta\Big)$$
to group the terms of $\frac{d}{dc}\widehat{\pounds}_2\big(J(a,b,c)\big)$ with a coefficient $\frac{1}{c}$ (resp. $\frac{1}{1-c}$) together—these are exactly the terms with a factor c (resp. $1-c$). For instance, the terms with $\frac{1}{c}$ are as follows:
$$\begin{aligned}
\frac{1}{c}\widehat{\pounds}_1\Bigg(&(a-b+1)[c] \\
&- a\Big[\frac{c}{a}\Big] + b\Big[\frac{c}{b}\Big] - b\Big[\frac{ca}{b}\Big] - (1-b)\Big[\frac{c(1-a)}{1-b}\Big] \\
&- c(1-a)\Big[\frac{1-c^{-1}}{1-a^{-1}}\Big] + c(1-b)\Big[\frac{1-c^{-1}}{1-b^{-1}}\Big] \\
&- c(a-b)\Big[\frac{1-c^{-1}}{1-\frac{a}{b}}\Big] - c(b-a)\Big[\frac{1-c^{-1}}{1-\frac{1-a}{1-b}}\Big]\Bigg)
\end{aligned}$$

In order to verify that this expression vanishes, we rewrite it in a slightly more convenient fashion (in order to be able to apply the four term relation line by line), neglecting the factor $\frac{1}{c}$, we get:

$$\begin{aligned}\widehat{\pounds}_1\Bigg(& (a-b+1)[c] \\ & - a\left[\frac{c}{a}\right] - c(1-a)\left[\frac{1-c^{-1}}{1-a^{-1}}\right] \\ & + b\left[\frac{c}{b}\right] + c(1-b)\left[\frac{1-c^{-1}}{1-b^{-1}}\right] \\ & - b\left[\frac{c}{\left(\frac{b}{a}\right)}\right] - c(a-b)\left[\frac{1-c^{-1}}{1-\left(\frac{b}{a}\right)^{-1}}\right] \\ & - (1-b)\left[\frac{c}{\left(\frac{1-a}{1-b}\right)^{-1}}\right] - c(b-a)\left[\frac{1-c^{-1}}{1-\frac{1-a}{1-b}}\right]\Bigg)\end{aligned}$$

Applying the 4-term equation (3.18) "linewise" to the 2nd, 3rd, 4th and 5th line above with $x=a$, $x=b$, $x=\frac{b}{a}$ and $x=\frac{1-b}{1-a}$, respectively, this latter expression is seen to reduce to

$$\begin{aligned}& \widehat{\pounds}_1\Bigg((a-b+1)[c] \\ & \quad - a[c] + c[a] \ + b[c] - c[b] \\ & \quad - a\left(\frac{b}{a}[c] - c\left[\frac{b}{a}\right]\right) \ - \ (1-a)\left(\frac{1-b}{1-a}[c] - c\left[\frac{1-b}{1-a}\right]\right)\Bigg) \\ = \ & \widehat{\pounds}_1\left(c\left([a] - [b] + a\left[\frac{b}{a}\right] + (1-a)\left[\frac{1-b}{1-a}\right]\right)\right)\end{aligned}$$

which vanishes, again in view of the four term equation (and because the coefficients for $[c]$ add up to zero). The terms with $\frac{1}{1-c}$ can be treated in a completely analogous way.

(ii) The constant term in c of the polynomial $\widehat{\pounds_2}(J(a,b,c))$, i.e. the polynomial $\widehat{\pounds_2}(J(a,b,0))$, is zero—this corresponds in the infinitesimal case to the degenerate case where we also put $c=0$ but where we need to give sense to expressions like $a[\frac{b}{a}]$ for $a=0$, the consistant choice being that it should be zero.

(iii) Instead of considering the coefficient of c^p in the polynomial $\widehat{\pounds_2}(J(a,b,c))$ we can equivalently check that the constant coefficient in $c^p\widehat{\pounds_2}(J(a,b,\frac{1}{c}))$ is zero. In the infinitesimal case we can perform the same check using $c\widehat{\pounds_2}(J(a,b,\frac{1}{c}))$ (so we can use the analogy again).

References

[1] **Aczél, J. and J. Dhombres** *Functional equations in several variables*, Encyclopedia of Math. and its Applications, Vol 31, Cambridge Univ. Press 1989.

[2] **Besser, A.** *Finite and p-adic polylogarithms*, preprint 2000.

[3] **Bloch, S.** *On the tangent space to Quillen K-theory*, Proc. Conf., Battelle Memorial Inst., Seattle, Wash., 1972, Springer Lect. Notes in Math. n^o 341, (1973), pp. 205–210.

[4] **Bloch, S.** *Higher regulators, algebraic K-theory, and zeta-functions of elliptic curves*, Lecture Notes, U.C. Irvine, 1977.

[5] **Bloch, S.** *Applications of the dilogarithm function in algebraic K-theory and algebraic geometry*, Proc. of the Int. Symp. on Alg. Geom. (Kyoto Univ., Kyoto, 1977), pp. 103–114, Kinokuniya Book Store, Tokyo, 1978.

[6] **Cathelineau, J.-L.** *Sur l'homologie de SL_2 à coefficients dans l'action adjointe*, Math. Scand. 63 (1988), 51-86.

[7] **Cathelineau, J.-L.** *Homologie du groupe linéaire et polylogarithmes (d'après Goncharov et d'autres)*, Séminaire Bourbaki, 772 (1992-93), Astérisque 216 (1993) 311-341.

[8] **Cathelineau, J.-L.** *Remarques sur les différentielles des polylogarithmes uniformes*, Ann. Inst. Fourier, Grenoble, **46**, 5(1996), 1327-1347.

[9] **Cathelineau, J.-L.** *Infinitesimal polylogarithms, multiplicative presentations of Kähler differentials and Goncharov complexes*, talk at the Workshop on Polylogarithms, Essen, May 1-4. (see http://www.exp-math.uni-essen.de/~herbert/polyloquy.html)

[10] **Coleman, R.** *Dilogarithms, regulators and p-adic L-functions*, Invent. Math. 69 (1982), no. 2, 171–208.

[11] **Dupont, J. L.; Sah, C.-H.** *Scissors congruences. II.*, J. Pure Appl. Algebra 25, (1982) 159-195.

[12] **Elbaz-Vincent, Ph.** *The indecomposable K_3 of rings and homology of SL_2*, Jo. Pure. Appl. Alg. 132,(1998) 27-71.

[13] **Elbaz-Vincent, Ph.** *Homology of Linear Groups with Coefficients in the Adjoint Action and K-theory*, K-Theory **16**, (1999), 35-50.

[14] **Elbaz-Vincent, Ph.** *Arithmetic and infinitesimal variations on a polylogarithm theme* (report on a work in progress with H. Gangl), talk at the Workshop on Polylogarithms, Essen, May 1-4. (see http://www.exp-math.uni-essen.de/~herbert/polyloquy.html)

[15] **Elbaz-Vincent, Ph.; Gangl, H.** *On Poly(ana)logs II*, (in preparation).

[16] **Elbaz-Vincent, Ph.; Gangl, H.** *Arithmetical and Infinitesimal Variations on a Polylogarithmic Theme* (in preparation)

[17] **Gangl, H.** *Funktionalgleichungen von Polylogarithmen*, (German) Dissertation, Bonn, 1995. Bonner Mathematische Schriften **278**. Universität Bonn.

[18] **Gangl, H.** *Functional equations for higher logarithms*, in preparation.

[19] **Goncharov, A.B.** *Geometry of configurations, polylogarithms and motivic cohomology*, Adv. Math. 114 (1995) 197-318.

[20] **Goncharov, A.B.** *Polylogarithms and motivic Galois groups*, Proc. of the Seattle conf. on motives, Seattle July 1991, AMS Proc. Symp. in Pure Math. 55 (1994) 2, 43-96.

[21] **Goncharov, A.B.** *Multiple polylogarithms, cyclotomy and modular complexes*, Math. Res. Lett. 5 (1998), no. 4, 497–516.

[22] **Kontsevich, M.** *The $1\frac{1}{2}$-logarithm* (unpublished note 1995).

[23] **Lewin, L.** *Polylogarithms and associated functions* North-Holland Publishing Co., New York-Amsterdam, 1981.

[24] **Lewin, L. (Ed)** *Structural Properties of Polylogarithms*, Mathematical Surveys and Monographs, vol 37, AMS, 1991.

[25] **Lewin, L.** *The order-independence of the polylogarithmic ladder structure—implications for a new category of functional equations*, Aequationes Math. 30 (1986), no. 1, 1–20.

[26] **Oesterlé, J.** *Polylogarithmes*, Sém. Bourbaki, 762 (1992-93), Astérisque 216 (1993) 49-67.

[27] **Ribenboim, P.** *13 Lectures on Fermat's last theorem*, New York - Heidelberg - Berlin: Springer-Verlag. (1979).

[28] **Rogers, L.J.** *On function sum theorems connected with the series $\sum_1^\infty x^n/n^2$*, Proc London Math. Soc. **4** (1907), 169–189.

[29] **Rosenberg, J.** *Algebraic K-Theory and Its Applications*, GTM 147, Springer-Verlag, 1994.

[30] **Suslin, A.A.** *K_3 of a field, and the Bloch group. (Russian)* Galois theory, rings, algebraic groups and their applications (Russian). Trudy Mat. Inst. Steklov. 183 (1990), 180–199, 229.
[31] **Washington, L.C.** *Introduction to cyclotomic fields*, GTM 83, Springer-Verlag 1997 (2nd edition).
[32] **Weibel, C.** *Private e-mail correspondence*, October 1999.
[33] **Wojtkowiak, Z.** *A construction of analogs of the Bloch-Wigner function*, Math. Scand. **65**, 1989, p.140-142.
[34] **Wojtkowiak, Z.** *A note on functional equations of the p-adic polylogarithms*, Bull. Soc. math. France **119**, 1991, p.343-370.
[35] **Wojtkowiak, Z.** *The basic structure of polylogarithmic functional equations*, Chapter 10 in [24], 205–231.
[36] **Zagier, D.** *Polylogarithms, Dedekind zeta functions and the algebraic K-theory of fields*, Proc. Texel Conf. on Arithm. Alg. Geometry 1989, Birkhäuser, Boston (1991) 391-430.
[37] **Zagier, D.** *Special values and functional equations of polylogarithms*, Appendix A in [24], 377–400.

Laboratoire G.T.A., UMR CNRS 5030, case courrier 51, Université Montpellier II, 34095 Montpellier Cedex 5, France.

E-mail address : pev@math.univ-monpt2.fr

MPI Bonn, Vivatsgasse 7, D-53111 Bonn, Deutschland.

E-mail address : herbert@mpim-bonn.mpg.de

URL : www.exp-math.uni-essen.de/~herbert

The $1\frac{1}{2}$-logarithm

(APPENDIX TO "ON POLY(ANA)LOGS I" BY PH. ELBAZ-VINCENT AND H. GANGL)

MAXIM KONTSEVICH

This appendix to the paper of Elbaz-Vincent and Gangl is included for historical purpose. It reproduces the text of [22], *initially written for the private booklet "Friedrich Hirzebruchs Emeritierung".*

Let $p > 2$ be a prime. Define a map from $\mathbb{Z}/p\mathbb{Z}$ to itself by the formula

$$H_p(x) = \sum_{k=1}^{p-1} \frac{x^k}{k} = x + \frac{x^2}{2} + \cdots + \frac{x^{p-1}}{p-1} \pmod p\,.$$

This function appears in explicit formulas for abelian extensions of cyclotomic fields. It looks like a truncated version of $\log(\frac{1}{1-x})$. Of course, it could not be a logarithm because there is no nonzero homomorphism from $(\mathbb{Z}/p\mathbb{Z})^\times \simeq \mathbb{Z}/(p-1)\mathbb{Z}$ to $\mathbb{Z}/p\mathbb{Z}$. I claim that H_p is analogous to another well-known function of a real variable. I will derive the analogy by writing several functional equations for H_p. These equations will be independent of p and I will suppress the index p from the notations.

(A): $H(1-x) = H(x)\,.$

Proof. we can compute explicitly the coefficients of the polynomial $H(1-x)$. First of all, its zeroth coefficient is $H(1) = 1 + \frac{1}{2} + \dots \frac{1}{p-1} = 1 + 2 + \cdots + (p-1) = \frac{p(p-1)}{2} = 0 \pmod p$. For l between 1 and $p-1$ the l-th coefficient of $H(1-x)$ is equal to

$$\sum_{k=1}^{p-1} \frac{1}{k}\,(-1)^l\,\frac{k(k-1)\dots(k-l+1)}{l!} = \frac{(-1)^l}{l!}\sum_{k=1}^{p-1}(k-1)\dots(k-l+1) =$$

$$= -\frac{(-1)^l}{l!}\,(0-1)(0-2)\dots(0-l+1) = \frac{(l-1)!}{l!} = \frac{1}{l}\,.$$

We use here the standard fact that

$$\sum_{k=0}^{p-1} P(k) = 0$$

for any polynomial $P \in \mathbb{Z}/p\mathbb{Z}[x]$ of degree at most $p-2$. □

A simple generalization of the previous argument shows that

(B): $H(x+y) = H(y) + (1-y)\,H(\frac{x}{1-y}) + y\,H(-\frac{x}{y})$ for $y \neq 0, 1\,.$

Also there is a very elementary identity

(C): $x\,H(\frac{1}{x}) = -H(x)$ for $x \neq 0$.

Claim: there is only one (up to a scalar factor) nonzero continuous solution of equations (A), (B), (C) in maps from $\mathbb{R}$ to itself. It is

$$H_\infty(x) = -\left(x \ \log|x| + (1-x) \ \log|1-x|\right) .$$

Also the function H_p is the unique (up to scalar factor) solution in maps from $\mathbb{Z}/p\mathbb{Z}$ to itself.

Cohomological interpretation of functional equations. Let F be a field and suppose that $H : F \to F$ satisfies (A) and (B). The equation (C) will be irrelevant. We associate with H a homogeneous function $\phi : F \times F \to F$ of degree 1:

$$\phi(x,y) := \begin{cases} (x+y)\, H\left(\frac{x}{x+y}\right) & \text{if} \quad x+y \neq 0\,, \\ 0 \quad \text{if} \quad x+y = 0\,. \end{cases}$$

Equation (A) implies that $\phi(x,y) = \phi(y,x)$. Equation (B) is equivalent to the identity

$$\phi(x,y) - \phi(x,y+z) + \phi(x+y,z) - \phi(y,z) = 0\,.$$

Thus, ϕ is a 2-cocycle of the abelian group F (the additive group of the field) with coefficients in itself as a trivial module. Because this cocycle is invariant under the usual action of the multiplicative group $F^\times$ (acting both on the group and on the coefficients), we get a 2-cocycle of the group of affine transformations of the line over F

$$\mathrm{Aff}(1,F) = \{t \mapsto at + b | a \in K^\times, b \in K\}$$

with coefficients in the non-trivial 1-dimensional representation given by the first coefficient. This 2-cocycle defines an extension of $\mathrm{Aff}(1,F)$ by F. The resulting group G can be identified as a set with $F \times F \times F^\times$.

Now we consider the case of $F = \mathbb{R}$ and assume that H is a measurable map. There are no non-trivial measurable cohomology classes in $H^2(\mathbb{R},\mathbb{R})$, hence ϕ should be a coboundary. This means that there exists a function $\psi : \mathbb{R} \to \mathbb{R}$ such that

$$\phi(x,y) = \psi(x) + \psi(y) - \psi(x+y)\,.$$

The homogeneity of ϕ implies that for any $\lambda \neq 0$ the function $\psi_\lambda(x) := \psi(\lambda x) - \lambda\psi(x)$ is additive in x. If we deal with measurable maps only then ψ_λ is a linear function. From this one can easily deduce that

$$\psi(x)/x = a \ \log|x| + b$$

for some $a, b \in \mathbb{R}$. Thus we get the solution of functional equations for $F = \mathbb{R}$.

Now we turn to the case $F = \mathbb{Z}/p\mathbb{Z}$. If the cohomology class in $H^2(F,F) \simeq \mathbb{Z}/p\mathbb{Z}$ corresponding to H is zero then by arguments parallel to the previous one obtains a homomorphism from $(\mathbb{Z}/p\mathbb{Z})^\times$ to $\mathbb{Z}/p\mathbb{Z}$. This homomorphism (a “logarithm”) vanishes inevitably, thus giving the uniqueness of H up to a scalar factor.

The group G in the case of $F = \mathbb{R}$ is a 3-dimensional solvable Lie group. The Lie algebra of G is defined over $\mathbb{Z}$ and it has a base x, y, z in which the commutation relations are

$$[x,y] = y, \ \ [x,z] = y+z, \ \ [y,z] = 0\,.$$

This Lie algebra cannot be the Lie algebra of any algebraic group over $\mathbb{Z}$ or over $\mathbb{Q}$. Nevertheless, we have defined groups of points over $\mathbb{R}$ and over $\mathbb{Z}/p\mathbb{Z}$ for all odd primes p.

Entropy. The function H is the entropy of a random variable taking two values. More generally, if ξ takes a finite number of values with probabilities $p_1, \ldots, p_k$, $\sum p_i = 1$ then the entropy of ξ is defined as

$$H(\xi) := -\sum_{i=1}^{k} p_i \ \log(p_i)\,.$$

We will consider the entropy also as a function of the collection of probabilities of elementary events, $H(\xi) = H(p_*)$. The main property of entropy is that if one random variable (say, ξ) is a function of another random variable (say, η) then the entropy of η can be computed as follows. Let us denote probabilities of all possible values of η by $p_{1,1}, p_{1,2}, \ldots, p_{1,l_1}; p_{2,1}, \cdots : \ldots p_{k,l_k}$ in such a way that $p_{1,1} + p_{1,2} + \cdots + p_{1,l_1} = p_1$ etc. Then we have k conditional distributions of probabilities $p_{i,*}/p_i$ for each $i \leq k$. The main identity of entropies is

$$H(p_{*,*}) = H(p_*) + \sum_{i=1}^{k} p_i H(\frac{p_{i,*}}{p_i})\,, \quad H(\eta) = H(\xi) + H_\xi(\eta)\,.$$

The last term in the formula above is the average value of the entropies of η with given values of ξ and it is called the relative entropy.

Using the main identity one can reduce by induction the calculation of the entropy of any random variable to the case of a two-valued variable, i.e. our function $H(x)$. One can check easily that the entropy of random variables computed using $H(x)$ is well-defined iff functional equations (A) and (B) are satisfied.

Conclusion: If we have a random variable ξ which takes finitely many values with all probabilities in $\mathbb{Q}$ then we can define not only the transcendental number $H(\xi)$ but also its "residues modulo p" for almost all primes p !

I propose calling the functions H_p "$1\frac{1}{2}$-logarithms," because their functional equation contains 4 terms, which is between 3 (the logarithm) and 5 (the dilogarithm giving an element in $H^3(Sl(2,\mathbb{C}),\mathbb{R})$).

The natural question is to find functional equations for the map $x \mapsto \sum_{k=1}^{p-1} x^k/k^2$ from $\mathbb{Z}/p\mathbb{Z}$ to itself. I don't know how to do it.

Maxim Kontsevich
Institut des Hautes Études Scientifiques, 35, Route de Chartres, 91440, Bures-sur-Yvette, France.
E-mail address : maxim@ihes.fr