# Finite polylogarithms, their multiple analogues and the Shannon entropy

## Geometric Sciences of Information 2015

### Session "Topological Forms and Information"

École Polytechnique (France), 28 October 2015

**Philippe Elbaz-Vincent**(Université Grenoble Alpes) & Herbert Gangl (Durham University)





# Content of this talk

- Information theory, Entropy and Polylogarithms (review of past works),
- Algebraic interpretation of the entropy function,
- Cohomological interpretation of formal entropy functions,
- Finite multiple polylogarithms, applications and open problems.

# Information theory, Entropy and Polylogarithms (1/4)

The Shannon entropy can be characterised in the framework of information theory, assuming that the propagation of information follows a Markovian model (Shannon, 1948).
If $H$ is the Shannon entropy, it fulfills the equation, often called the *Fundamental Equation of Information Theory* (FEITH)

$$H(x) + (1-x)H\left(\frac{y}{1-x}\right) - H(y) - (1-y)H\left(\frac{x}{1-y}\right) = 0\,. \tag{FEITH}$$

It is known (Aczel and Dhombres, 1989), that if $g$ is a real function locally integrable on $]0,1[$ and if, moreover, $g$ fulfills FEITH, then there exists $c \in \mathbb{R}$ such that $g = cH$ (we can also restrict the hypothesis to Lebesgue measurable).

# Information theory, Entropy and Polylogarithms (2/4)

It turns out that FEITH can be derived, in a precise formal sense (Elbaz-Vincent and Gangl, 2002), from the 5-term equation of the classical (or $p$-adic) dilogarithm.

Cathelineau (1996) found that an appropriate derivative of the Bloch–Wigner dilogarithm coincides with the classical entropy function, and that the five term relation satisfied by the former implies the four term relation of the latter.

More precisely, we define $Li_m(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^m}$, $\quad |z| < 1$, the $m$-logarithm. We set

$$\mathcal{D}_2(z) = i\,\mathrm{Im}\Big( Li_2(z) + \log(1-z)\log|z| \Big),$$

Then $\mathcal{D}_2$ satisfies the following 5-term equation

$$\mathcal{D}_2(a) - \mathcal{D}_2(b) + \mathcal{D}_2\left(\frac{b}{a}\right) - \mathcal{D}_2\left(\frac{1-b}{1-a}\right) + \mathcal{D}_2\left(\frac{1-b^{-1}}{1-a^{-1}}\right) = 0,$$

whenever such an expression makes sense. The relation is the famous *five term equation for the dilogarithm* (first stated by Abel).

# Information theory, Entropy and Polylogarithms (3/4)

It can be shown formarly (see Cathelineau, Elbaz-Vincent and Gangl) that FEITH is an infinitesimal version of this 5-term equation.

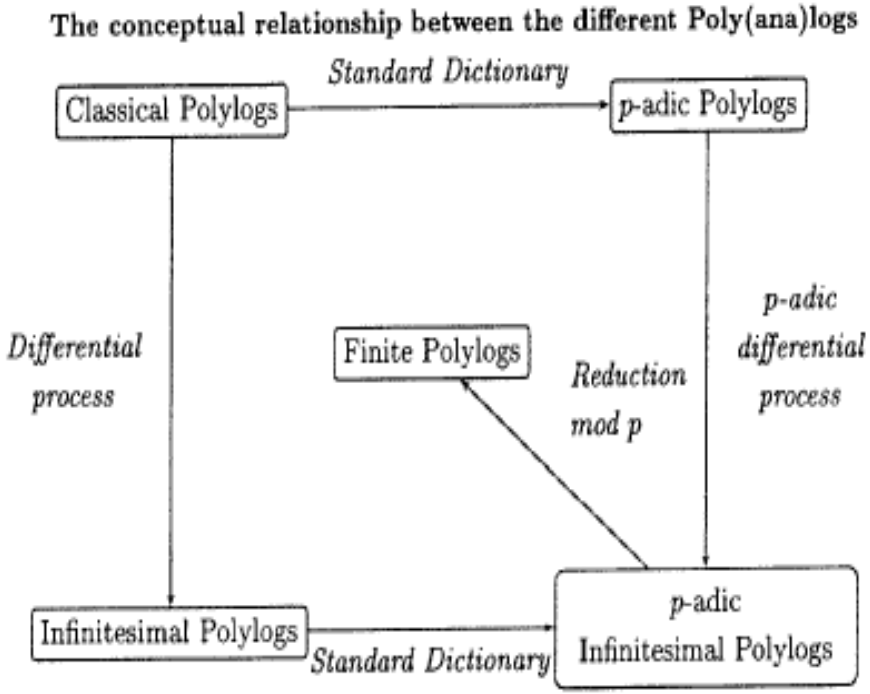
Kontsevich (1995) discovered that the truncated finite logarithm over a finite field $\mathbb{F}_p$, with $p$ prime, defined by

$$\pounds_1(x) = \sum_{k=1}^{p-1} \frac{x^k}{k},$$

satisfies FEITH.

In our previous work, we showed how one can expand this relationship for "higher analogues" in order to produce and prove similar functional identities for finite polylogarithms from those for classical polylogarithms (using mod $p$ reduction of $p$-adic polylogarithms and their infinitesimal version). It was also shown that functional equations for finite polylogarithms often hold even as polynomial identities over finite fields.

# Information theory, Entropy and Polylogarithms (4/4)




☞ *Entropy and FEITH arise from the infinitesimal picture (for both archimedean and non-archimedean structure) and their finite analogs associated to the dilogarithm*. Does their exist higher analogue of the Shannon entropy associated to *m*-logarithms ? It could be connected to the higher degrees of the *information cohomology space* of Baudot and Bennequin (Entropy 2015).

# Algebraic interpretation of the entropy function (1/2)

Let $R$ be a (commutative) ring and let $D$ be a map from $R$ to $R$. We will say that $D$ is a *unitary derivation* over $R$ if the following axioms hold :

1. "Leibniz's rule" : for all $x, y \in R$, we have $D(xy) = xD(y) + yD(x)$.
2. "Additivity on partitions of unity" : for all $x \in R$, we have $D(x) + D(1-x) = 0$.

We will denote by $Der^u(R)$ the set of unitary derivations over $R$. We will say that a map $f : R \to R$ is an abstract symmetric information function of degree 1 if the two following conditions hold : for all $x, y \in R$ such that $x, y, 1-x, 1-y \in R^\times$, the functional equation FEITH holds and for all $x \in R$, we have $f(x) = f(1-x)$. Denote by $\mathcal{IF}_1(R)$ the set of abstract symmetric information functions of degree 1 over $R$. Then $\mathcal{IF}_1(R)$ is an $R$-module. Let $Leib(R)$ be the set of Leibniz functions over $R$ (i.e. which fulfill the "Leibniz rule").

# Algebraic interpretation of the entropy function (2/2)

**Proposition :** We have a morphism of $R$-modules $h : Leib(R) \to \mathcal{IF}_1(R)$, defined by $h(\varphi) = \varphi + \varphi \circ \tau$, with $\tau(x) = 1 - x$. Furthermore, $Ker(h) = Der^u(R)$.

☞ Hence, if $h$ is onto, abstract information function are naturally associated to formal derivations. Nevertheless, $h$ can be also 0. Indeed, if $R = \mathbb{F}_q$, is a finite field, then $Leib(\mathbb{F}_q) = 0$, but $\mathcal{IF}_1(\mathbb{F}_q) \neq 0$ (it is generated by $\pounds_1$).

# Cohomological interpretation of formal entropy functions

The following results are classical in origin (Cathelineau, 1988 and Kontsevich, 1995)

**Proposition :** Let $F$ be a finite prime field and $H : F \to F$ a function which fulfills the following conditions : $H(x) = H(1-x)$, the functional equation (FEITH) holds for $H$ and $H(0) = 0$. Then the function $\varphi : F \times F \to F$ defined by $\varphi(x,y) = (x+y)H(\frac{x}{x+y})$ if $x + y \neq 0$ and 0 otherwise, is a non-trivial 2-cocycle.

**sketch of proof :** Suppose that $\varphi$ is a 2-coboundary. Then, there exists a map $Q : F \to F$, such that $\varphi(x,y) = Q(x+y) - Q(x) - Q(y)$. The function $\psi_\lambda(x) = Q(\lambda x) - \lambda Q(x)$ is an additive morphism $F \to F$, hence entirely determined by $\psi_\lambda(1)$. The map $\psi_\lambda(1)$ fulfills the Leibniz chain rule on $F^\times$. We deduce from it that $\varphi = 0$ (which is not possible, so it is not a coboundary !)

☞ We deduce that $\pounds_1$ is unique (up to a constant). In the real or complex we use other type of cohomological arguments (see also the relationship with Baudot and Bennequin, 2015).

# Finite multiple polylogarithms (1/3)

While *classical* polylogarithms play an important role in the theory of mixed Tate motives over a field, it turns out that it is often preferable to also consider the larger class of *multiple* polylogarithms (cf. Goncharov's work). In a similar way it is useful to investigate their finite analogues. We are mainly concerned with finite double polylogarithms which are given as functions $\mathbb{Z}/p \times \mathbb{Z}/p \to \mathbb{Z}/p$ by

$$\pounds_{a,b}(x,y) = \sum_{0<m<n<p} \frac{x^m}{m^a}\frac{y^n}{n^b}\,.$$

# Finite multiple polylogarithms (2/3)

The finite $(1,1)$-logarithm $\mathfrak{L}_{1,1}(x,y)$ can be expressed in terms of $\mathfrak{L}_2$. More precisely, we have

$$y\mathfrak{L}_{1,1}(x,\frac{1}{y}) = \quad \mathfrak{L}_2\Big( - y^p\Big[\frac{x}{y}\Big] - (1-y)^p\Big[\frac{1-x}{1-y}\Big] + [1-x] + [1-y]\Big)\,.$$

Define $[x,y]_s = \mathfrak{L}_{1,1}(x,y) + \mathfrak{L}_{1,1}(y,x)$ and consider the following linear combination

$$\begin{aligned} K(x,y) =& [x,y]_s + x^p\Big[\frac{1}{x},y\Big]_s - (1-y)^p\Big[1-x,\frac{y}{y-1}\Big]_s + (1-y)^p\Big[1-x,\frac{1}{1-y}\Big]_s \\ & - x^p(1-y)^p\Big[1-\frac{1}{x},\frac{y}{y-1}\Big]_s + x^p(1-y)^p\Big[1-\frac{1}{x},\frac{1}{1-y}\Big]_s\,. \end{aligned}$$

Then the following functional equation (purely in $\mathfrak{L}_{1,1}$) holds :

$$I(x,y;z,w) - I(x,z;y,w) = 0\,,$$

where

$$I(x,y;z,w) = (1+z)\,(1+w)\,K(x,y) + (1+x)\,(1+y)\,K(z,w)\,.$$

# Finite multiple polylogarithms (3/3)

We have the following equation for which *no analogue is known for the classical case*

Let $n > 0$ be divisible by 3, and put $\omega = n/3 - 1$. Then

$$\sum_{j=0}^{\omega} \frac{\binom{\omega}{j}}{\binom{2\omega}{j}} \pounds_{n-(j+1),j+1}\Big([a,b]-[\frac{1}{a},a\,b]-a^p b^p[b,\frac{1}{a\,b}]+b^p[a\,b,\frac{1}{b}]\Big) = 0.$$

**Questions :** what is the interpretation in term of information theory for the multiple polylogs ?

# Finite polylogarithms and Fermat's last theorem

Several classical criteria used by Kummer, Mirimanoff and Wieferich to prove certain cases of Fermat's Last Theorem can be rephrased in terms of functional equations and evaluations of finite (multiple) polylogarithms. For example, Mirimanoff was led to the study of (nowadays called) *Mirimanoff polynomials* (cf. Ribenboim book on FLT) $\varphi_j(T) = \sum_{j=1}^{p-1} k^{j-1} T^k$, which are nothing else but finite polylogarithms...

The *Mirimanoff congruences* (op.cit) can be reformulated as follows : for any solution $(x, y, z)$ of $x^p + y^p + z^p = 0$ in pairwise prime integers not divisible by $p$ (i.e. a *Fermat triple*) and for $t = -\frac{x}{y}$ we have

$$\mathfrak{L}_1(t) = 0\,, \quad \mathfrak{L}_j(t)\mathfrak{L}_{p-j}(t) = 0 \qquad (j = 2, \ldots, \frac{p-1}{2})\,.$$

One can prove these congruences using an identity expressing $\mathfrak{L}_{p-j-1,j+1}(1, T)$ in terms of $\mathfrak{L}_n(T)$.