Polylogarithms, their multiple analogues and the Shannon entropy

TGSI 2017

Session “Information and Topology”

CIRM (France), 29 August 2017

Philippe Elbaz-Vincent (Univ. Grenoble Alpes) & Herbert Gangl (Durham University)

INSTITUT FOURIER   UNIVERSITÉ Grenoble Alpes   Durham University   ANR   PERSYVAL-Lab

Content of this talk

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

Information theory, Entropy and Polylogarithms (1/8)

A general definition of entropy for information theory has been given by Rényi (1960) : Let $S = \{s_1, \ldots, s_n\}$ be a set of discrete events for which the probabilities are given by $p_i = P(s = s_i)$ for $i = 1, \ldots, n$. The Rényi entropy S is then defined for $\alpha > 0$ and $\alpha \neq 1$ as

$$H_\alpha(S) = \frac{1}{1-\alpha} \log\left(\sum_{i=1}^{n} p_i^\alpha\right) .$$

The Shannon entropy (1948) can be recovered from the one of Rényi when $\alpha \to 1$. We also often use the minimal entropy which is related to the probability of the most predictable event (while the Shannon entropy gives an averaged measure) :

$$H_{\min}(S) = \lim_{\alpha\to\infty} H_\alpha(S) = -\log(\max_{i=1,\ldots,n}(p_i)) .$$

Those different entropies are related by the following inequalities

$$H_{\min}(S) \leqslant \cdots \leqslant H_2(S) \leqslant H_1(S) \leqslant \log(\mathrm{card}(S)) = \lim_{\alpha\to 0} H_\alpha(S) .$$

Information theory, Entropy and Polylogarithms (2/8)

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 (3/8)

Then enter the polylogarithms...
We define $Li_m(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^m}$, $\qquad |z| < 1$, 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 (4/8)

It turns out that FEITH can be derived from the 5-term equation. 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.
This construction has been extended to higher functional equations of polylogarithms (Elbaz-Vincent and Gangl, 2002).

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

$$\Downarrow (\partial)$$

$$d\mathcal{D}_2(x) + (1-x)d\mathcal{D}_2\left(\frac{y}{1-x}\right) - d\mathcal{D}_2(y) - (1-y)d\mathcal{D}_2\left(\frac{x}{1-y}\right) = 0\,.$$

Information theory, Entropy and Polylogarithms (5/8)

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

$$£_1(x) = \sum_{k=1}^{p-1} \frac{x^k}{k},$$

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 (6/8)

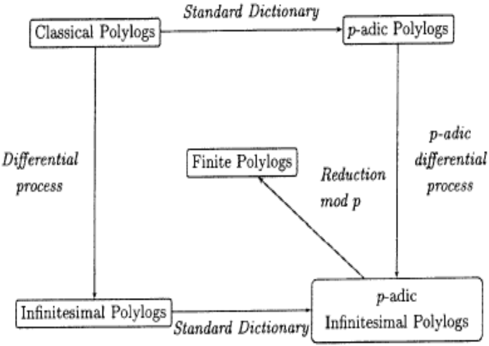
Using the recent works of Kaneko, Zagier, Sakugawa and Seki, we can give a global version of the finite polylogarithms. Consider the $\mathbb{Q}$-algebra $\mathcal{A}$ defined as

$$\mathcal{A} = \left(\prod_p \mathbb{F}_p\right) / \left(\bigoplus_p \mathbb{F}_p\right) .$$

Then we can define $\pounds_{\mathcal{A},1}(x) = (\pounds_1(x))_p$ (x can be either an indeterminate or a "$\mathbb{Z}$-value“, in the former we have to consider $\mathbb{F}_p[x]$ instead of $\mathbb{F}_p$). Using such framework, we can extend the FEITH to a new setting (which also involve multiple zeta values and periods).

Information theory, Entropy and Polylogarithms (7/8)

The conceptual relationship between the different Poly(ana)logs



☞ *Entropy and FEITH arise from the infinitesimal picture (for both archimedean and non-archimedean structure) and their finite analogs associated to the dilogarithm.*

Information theory, Entropy and Polylogarithms (8/8)

The previous picture does rise several ”mystical“ questions :

- Do their exist higher analogues of the Shannon entropy associated to *m*-logarithms ?
- What is their interpretations in terms of information theory ?

☞ It could be connected to the higher degrees of the *information cohomology space* of Baudot and Bennequin (Entropy 2015). Notice that the polylogarithms are already of cohomological nature...

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 φ 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 $\pounds_{1,1}(x,y)$ can be expressed in terms of $\pounds_2$. More precisely, we have

$$y\pounds_{1,1}(x,\frac{1}{y}) = \quad \pounds_2\Big(-y^p\Big[\frac{x}{y}\Big]-(1-y)^p\Big[\frac{1-x}{1-y}\Big]+[1-x]+[1-y]\Big).$$

Define $[x,y]_s=\pounds_{1,1}(x,y)+\pounds_{1,1}(y,x)$ 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 $\pounds_{1,1}$) 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

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

☞ **New mystical question :** what is the interpretation in term of information theory (even geometrical or topological) 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

$$\pounds_1(t) = 0\,, \quad \pounds_j(t)\pounds_{p-j}(t) = 0 \qquad (j = 2, \ldots, \frac{p-1}{2})\,.$$

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