Finite polylogarithms, their multiple analogues and the Shannon entropy

Philippe Elbaz-Vincent, Herbert Gangl

Finite polylogarithms, their multiple analogues and the Shannon entropy

Philippe Elbaz-Vincent[1]* and Herbert Gangl[2]

[1] Institut Fourier, CNRS-Université Grenoble Alpes,
BP 74, F-38402 Saint Martin d'Hères, France
Philippe.Elbaz-Vincent@ujf-grenoble.fr
[2] Department of Mathematical Sciences, South Road,
University of Durham, United Kingdom
herbert.gangl@durham.ac.uk

Abstract. We show that the entropy function—and hence the finite 1-logarithm—behaves a lot like certain derivations. We recall its cohomological interpretation as a 2-cocycle and also deduce $2n$-cocycles for any n. Finally, we give some identities for finite multiple polylogarithms together with number theoretic applications.

1 Information theory, Entropy and Polylogarithms

It is well known that the notion of entropy occurs in many sciences. In thermodynamics, it means a measure of the quantity of disorder, or more accurately, the tendancy of a system to go toward a disordered state. In information theory, the entropy measures (in terms of real positive numbers) the quantity of information of a certain property [17],[20]. From a practical viewpoint, entropies play also a key role in the study of random bit generators (deterministic or not) [8], in particular due to the Maurer test [16]. A general definition of entropy has been given by Rényi [18]: 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 [20] can be recovered from the one of Rényi when $\alpha \to 1$

$$H_1(S) = \lim_{\alpha \to 1} H_\alpha(S) = -\sum_{i=1}^{n} p_i \log(p_i) .$$

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):

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

* Partially supported by the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01) and by the LabEx AMIES

Those different entropies are related by the following inequalities

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

The Shannon entropy can be characterised in the framework of information theory, assuming that the propagation of information follows a Markovian model [17],[20]. If H is the Shannon entropy, it fulfills the equation, often called the *Fundamental Equation of Information Theory* (FEITH),

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

In [2](section 5.4, pp.66-69), it is shown 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). There are several papers (e.g., [1],[7]) on the equation FEITH and the understanding of its structural properties, with the motivation to weaken either the probabilistic hypothesis or the analytical ones. The following generalisation of the equation FEITH has also been considered [14], for β positive and x and y in some admissible range,

$$H(x) + (1-x)^\beta H\left(\frac{y}{1-x}\right) - H(y) - (1-y)^\beta H\left(\frac{x}{1-y}\right) = 0\,. \qquad (1)$$

It turns out that FEITH can be derived, in a precise formal sense [9], from the 5-term equation of the classical (or p-adic) dilogarithm. Cathelineau [5] 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. Kontsevich [13] 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 (or its generalisation for $\beta = 1$ or p). In [9] 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. It was also shown that functional equations for finite polylogarithms often hold even as polynomial identities over finite fields. In particular, we have shown that the polynomial version of $\pounds_1$ fulfills (1) with $\beta = p$. Another approach, due to Bloch and Esnault [3], gives a more geometric version in terms of algebraic cycles, and further structural properties have been investigated by Cathelineau [6].

In this paper we propose some new formal characterisations of the entropy from an algebraic viewpoint, using formal derivations and a relation to cohomology (section 2), and we give complementary relations involving multiple analogues of the finite polylogarithms with a few applications to number theory (section

3). The details for this last work will be given in a subsequent paper. In the remainder of the paper, rings are assumed to be commutative. We will denote by $\mathbb{F}_q$ the finite field with q elements. If q is prime, we set $\mathbb{F}_q = \mathbb{Z}/q\mathbb{Z}$.

Acknowledgements: We would like to express our sincere gratitude to the reviewers for their valuable comments who have helped improve this paper.

2 Algebraic interpretation of the entropy function

2.1 Formal entropy as formal derivations

Definition 1. *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.

Applying analogous arguments as for derivations (see for instance [15], chap. 9), we have

Proposition 1. *The set of unitary derivations over R, $Der^u(R)$, is an R-module, which has $Der_{\mathbb{Z}}(R)$ as a submodule. If D and D' are two unitary derivations, then the composition $D \circ D'$ and the Lie bracket $[D, D'] = D \circ D' - D' \circ D$ are unitary derivations.*
Let D be a unitary derivation over R.

1. *For all $x \in R$ and all $n \in \mathbb{N}$ we have $D(x^n) = nx^{n-1}D(x)$. Furthermore, if $x \in R^\times$ the rule is also true for $n \in \mathbb{Z}$.*
2. *For all $n \in \mathbb{N}$, $D((n+1)1_R) = \frac{n(n+1)}{2}D(-1)$, and $2D(-1) = 0$.*
3. *If R has no 2-torsion or if $2(-1) = 0$ in R, then for all $x \in R$ and all $n \in \mathbb{Z}$, we have $D(nx) = nD(x)$.*
4. *Suppose that R has no 2-torsion, or that $2(-1) = 0$ in R, and let $m \in \mathbb{Z}$ with $m \in R^\times$, then $D(\frac{1}{m}) = 0$. If moreover $\mathbb{Q} \subset R$, then $D(\mathbb{Q}) = 0$.*

Proof. 1. Works as the classical proof for derivations.
2. First, using the standard fact that $0 = 0 \cdot 0$, we deduce that $D(0) = 0$, and then $D(1) = 0$. Then we can see that $2D(-1) = 0$ and that $D(n+1) - D(n) = nD(-1)$. Thus an induction argument proves the formula.
3. If R has no 2-torsion, or if $2(-1) = 0$ in R, then $D(-1) = 0$, and using the previous result together with the fact that $D(-n) = -D(1+n)$, we deduce $D(n) = 0$ for all $n \in \mathbb{Z}$. Then the desired formula follows.
4. Direct consequence of the previous rules.

□

Remark 1. We can get nicer statements by working in $Der^u(R)/\langle D(-1)\rangle$, where $\langle D(-1)\rangle$ denotes the submodule of $Der^u(R)$ spanned by $D(-1)$.

Corollary 1. *Suppose that $nR = 0$, for a given $n \in \mathbb{N} - \{0\}$. Then if D is a unitary derivation over R and if $\lambda_n : R \longrightarrow R$ is defined by $\lambda_n(x) = x^n$, we then have $D \circ \lambda_n = 0$. In particular if p is a prime number, $\nu \in \mathbb{N} - \{0\}$ and $q = p^\nu$, then $D(\mathbb{F}_q) = 0$.*

Recall the following definition from [12].

Definition 2. *Let R be a commutative ring and k be a natural number. We say that R is k-fold stable if for any family of k unimodular vectors $(a_i, b_i)_{1 \leqslant i \leqslant k} \in R^2$ (i.e. $a_i R + b_i R = R$), there exists $t \in R$, such that $a_i + t b_i \in R^\times$ for all i.*

Proposition 2. "Unitary Derivations are almost Derivations"
Let R be a 2-fold stable ring, and suppose that R is of characteristic 2 (i.e. $2R = 0$) or that R has no 2-torsion. Then $Der_{\mathbb{Z}}(R) = Der^u(R)$.

Proof. According to Proposition 1, we have to show that any unitary derivation is additive. Let $D \in Der^u(R)$ and let $x, y \in R$. Suppose first that x is invertible. Then $x+y = x(1+\frac{y}{x})$, and by Leibniz's rule, we have $D(x+y) = xD(1+\frac{y}{x})+(1+\frac{y}{x})D(x)$. Using the additivity on partitions of unity, $D(1+\frac{y}{x}) = -D(-\frac{y}{x})$ and also $D(-\frac{y}{x}) = -D(\frac{y}{x})$. Hence we deduce $D(x+y) = D(x)+D(y)$. Now suppose that x is not invertible. Then applying the 2-fold stability to the unimodular vectors $(0,1)$, $(x,1)$, we deduce the existence of $t \in R^\times$ such that $x + t$ is invertible. Setting $x' = x+t$ and $y' = y-t$, we have $D(x+y) = D(x'+y')$. Then we can apply the previous arguments to x', y', and deduce that $D(x+y) = D(x+t)+D(y-t)$. Now we again apply the same arguments to x, t, and y, $-t$. Using the rules of Proposition 1, we conclude that $D(x+y) = D(x)+D(y)$, and the claim follows. □

Example 1. As any semilocal ring R such that any of its residue fields has at least 3 elements is 2-fold stable [12], we then deduce that $Der_{\mathbb{Z}}(R) = Der^u(R)$.

2.2 Unitary Derivations and Symmetric Information Function of degree 1

For more details on this section related to information theory see [14].

Definition 3. *Let R be a commutative ring. 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"), then it is also an R-module (in fact the composition and the Lie bracket still hold in $Leib(R)$). The proof of the following proposition is a straightforward computation.

Proposition 3. *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, $Ker(H) = Der^u(R)$.*

Remark 2. The morphism H is not necessarily onto. If $R = \mathbb{F}_q$, a finite field, then $Leib(\mathbb{F}_q) = 0$, but $\mathcal{IF}_1(\mathbb{F}_q) \neq 0$.

2.3 Cohomological interpretation of formal entropy functions

The following results are classical in origin (see [4], pp.58–59, and also the references cited there, and also [13]). We try in this section to render the proofs (for the finite case) more transparent, and also emphasize the derivation aspect of the previous sections.

Theorem 1. *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.*

Proof. The fact that φ is a 2-cocycle is a straightforward consequence of the properties on H. In order to see this, we use the inversion relation, which in turn one can deduce from (FEITH), and the relation $H(x) = H(1-x)$. By setting $Y = \frac{x}{x+y+z}$ and $X = \frac{y}{x+y+z}$ (assuming some suitable admissibility conditions on x, y and z), and modulo some modifications using the other relations, the 2-cocycle condition is deduced from (FEITH). For the non-triviality, notice that φ is homogeneous and recall that as F is a field we can endow the cochains with a structure of F-vector space. Suppose that φ is a 2-coboundary. Then, there exists a map $Q : F \to F$, such that $\varphi(x, y) = Q(x+y) - Q(x) - Q(y)$. Notice that $Q(0) = 0$. As φ is homogeneous, we have $\varphi(\lambda x, \lambda y) = \lambda Q(x+y) - \lambda Q(x) - \lambda Q(y)$. Thus 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 $F^\times$. Indeed, assuming $F = \mathbb{Z}/p\mathbb{Z}$, if λ, μ are arbitrary elements of F, as $\mu\psi_\lambda(1) = \psi_\lambda(\mu)$, by a straightforward computation we deduce $\psi_{\lambda\mu}(1) = \psi_\lambda(\mu) + \lambda\psi_\mu(1)$. Thus we formally have $\psi_{\lambda^m}(1) = m\lambda^{m-1}\psi_\lambda(1)$. But $F^\times$ is generated by a primitive root, say ω. Let $p = card(F)$. Then $\omega^{p-1} = 1$. Moreover $0 = \psi_1(1) = (p-1)\omega^{p-2}\psi_\omega(1)$. Hence $\psi_\omega(1) = 0$ and then $Q(\lambda x) = \lambda Q(x)$ for all $\lambda, x \in F$. This implies that Q is an additive map and thus $\varphi = 0$, which contradicts the fact that it is a non-zero 2-cochain. □

Remark 3. We should notice that $H(\lambda) = \varphi(\lambda, 1-\lambda) = \psi_\lambda(1) + \psi_{(1-\lambda)}(1)$, which is very similar to the results of Maksa [14].

Corollary 2. *The map $F \to H^2(F, F)$, given by $\lambda \mapsto \lambda\varphi$, is an isomorphism and, up to a constant, $\pounds_1$ is unique.*

Using the (cup) product structure on the cohomology ring $H^*(F, F)$ (cf. [10], chap. 3), we can check the following property:

Corollary 3. *Let n be a positive integer. The map*

$$\varphi(x_1, \ldots, x_{2n}) = \prod_{i=1,\, i \text{ even}}^{2n-1} \varphi(x_i, x_{i+1})$$

induces a non-trivial cocycle in $H^{2n}(F, F)$, which corresponds to the cup product induced by φ. This cocycle corresponds to the product of n functions H, and is unique up to a constant.

3 Finite multiple polylogarithms

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 (e.g., [11]). 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

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

3.1 Expressing $\pounds_{1,1}$ via $\pounds_2$

Our arguably most interesting result, from which we will deduce a couple of consequences, is the following.

Theorem 2. *The finite* $(1,1)$*-logarithm* $\pounds_{1,1}(x,y)$ *can be expressed in terms of* $\pounds_2$*. More precisely, we have*

$$y\pounds_{1,1}(x,\frac{1}{y}) = \pounds_2\Big(-y^p\Big[\frac{x}{y}\Big] - (1-y)^p\Big[\frac{1-x}{1-y}\Big] + [1-x] + [1-y]\Big)\,. \qquad (2)$$

The proof of this result takes $(1-y)^p\pounds_2\big(\frac{1-x}{1-y}\big)$ and decomposes the (triangular) domain over which the summation variables run into an "open" part (a triangle) and three "boundary" parts (one diagonal, a vertical and a horizontal line) and identifies the former with the $\pounds_{1,1}$-expression and the latter with the three remaining terms in the equation. At a crucial step one uses the binomial identity

$$\sum_{r=0}^{N}\binom{N-r}{s}\binom{r}{t} = \binom{N+1}{s+t+1}\,.$$

3.2 Cathelineau's $\pounds_2$-identity

Combining the well-known shuffle identity $\pounds_{a,b}(x,y) + \pounds_{b,a}(y,x) + \pounds_{a+b}(xy) = \pounds_a(x)\pounds_b(y)$ for $a = b = 1$ with the above we find that the product $\pounds_1(x)\pounds_1(y)$ can indeed be expressed as a sum of $\pounds_2$-terms. In fact, the resulting expression is precisely Cathelineau's "double bracket" $[[x,y]]$ ([5], p.1344, Déf. 4). Now the sum obtained from the four terms of (FEITH) for one of the two arguments and while leaving the second argument fixed kills the products of $\pounds_1$-terms so we are left with only $\pounds_2$-terms, hence we have proved a functional equation—in fact, Cathelineau's 22-term equation in ([5], p.1346, (2)).

3.3 Further identities

We can prove an inversion formula for finite multiple polylogarithms

$$T_1^p\cdots T_\ell^p\,\pounds_{m_\ell,\ldots,m_1}\Big(\frac{1}{T_\ell},\ldots,\frac{1}{T_1}\Big) = (-1)^{m_1+\cdots+m_\ell}\,\pounds_{m_1,\ldots,m_\ell}(T_1,\ldots,T_\ell)\,,$$

and we can also build a four variable identity for $\pounds_{1,1}$.

Proposition 4. *Define* $[x,y]_s = \pounds_{1,1}(x,y) + \pounds_{1,1}(y,x)$ *and consider the following linear combination*

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

Then the following functional equation (purely in $\pounds_{1,1}$ *) holds:*

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

where

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

3.4 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. [19], VIII, (1.11))

$$\varphi_j(T) = \sum_{j=1}^{p-1} k^{j-1}T^k\,,$$

which are nothing else but finite polylogarithms:

$$\varphi_j(T) \equiv \pounds_{p-j}(T) \pmod p\,.$$

(Note that Mirimanoff's original polynomials correspond to $-\varphi_j(-T)$.)

Part of the groundwork for Mirimanoff's congruences was formed by the **crucial identity**

$$-\frac{1}{2}\big[\varphi_{p-1}(T)\big]^2 \equiv \varphi_{p-2}(T) + (T-1)^{2p}\varphi_{p-2}\Big(\frac{T}{T-1}\Big) \pmod p$$

([19], VIII, (1.29)) which is nothing but the special case product formula $x = y\,(=T)$ in our identity for $\pounds_1(x)\pounds_1(y)$ alluded to in 3.2.

The *Mirimanoff congruences* ([19], VIII, (1B)) 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 $\pounds_n(T)$: denoting the Bernoulli numbers by B_n, we have

$$\pounds_{p-j-1,1+j}(1,T) \equiv \frac{1}{j+1}\sum_{n=0}^{j}\binom{j+1}{n}B_n\pounds_n(T) \qquad j = 1,\ldots,p-2\,. \qquad (3)$$

Also, *Wieferich's criterion* states that if the first case of FLT for the prime p is false then p^2 divides $2^p - 1$ (only two such primes are known for which that latter holds: $p = 1093$ and $p = 3511$). This criterion can be rephrased in terms of finite polylogarithms as saying $\pounds_1(-1) = 0$ for such primes.

References

1. Aczél, J., Entropies Old and New (and Both New and Old) and Their Characterizations, CP707, Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 23rd International Workshop, edited by G. Erickson and Y. Zhai, American Institute of Physics, 2004.
2. Aczél, J. and J. Dhombres, Functional equations in several variables, Encyclopedia of Math. and its Applications, Vol 31, Cambridge Univ. Press 1989.
3. S. Bloch and H. Esnault, An additive version of higher Chow groups, Ann. Sci. École Norm. Sup. (4) 36 (2003), p. 463-477.
4. Cathelineau, J.-L., Sur l'homologie de SL_2 à coefficients dans l'action adjointe, Math. Scand. 63 (1988), 51-86.
5. Cathelineau, J.-L., Remarques sur les différentielles des polylogarithmes uniformes, Ann. Inst. Fourier, Grenoble, 46, 5 (1996), 1327–1347.
6. Cathelineau, J.-L., The tangent complex to the Bloch-Suslin complex, Bull. Soc. math. France, 135 (4), 2007, 565-597.
7. Csiszár, Imre., Axiomatic Characterizations of Information Measures, Entropy 2008, 10, 261-273; DOI: 10.3390/e10030261.
8. De Julis, Guenaëlle, Analyse d'accumulateurs d'entropie pour les générateurs aléatoires cryptographiques, PhD thesis, Université Grenoble Alpes, December 2014. `https://tel.archives-ouvertes.fr/tel-01102765v1`.
9. Elbaz-Vincent, Ph., Gangl, H., On poly(ana)logs I, Compos. Math. 130 (2002), no. 2, 161–210.
10. Evens, L.; The Cohomology of Groups, Oxford Math. Monographs, Clarendon Press, 1991.
11. Goncharov, A.B., Galois symmetries of fundamental groupoids and noncommutative geometry. Duke Math. J., 128, no. 2 (2005), 209–284.
12. van der Kallen, W., The K_2 of rings with many units, Ann. Scient. Ec. Norm. Sup. 10, 473-515 (1977).
13. Kontsevich, M., The $1\frac{1}{2}$-logarithm, Appendix to [9].
14. Maksa, Gy., The general solution of a functional equation related to the mixed theory of information, Aeq. Math. 22 (1981), 90-96.
15. Matsumura, H., Commutative ring theory, Cambridge studies in advanced Math. 8. (1986).
16. Maurer, U.M., A universal statistical test for random bit generators. Journal of Cryptology, 5(2), p.89-105, 1992.
17. Ollivier, Yan, Aspects de l'entropie en mathématiques et en physique, Technical report, 2002. `http://www.yann-ollivier.org/entropie/entropie.pdf`.
18. Rényi, A.; On measures of entropy and information, in Proc. 4th Berkeley Symposium on Mathematical Statistics and Probability. vol. 1, 1960, p. 547-561.
19. Ribenboim, P., 13 Lectures on Fermat's Last Theorem. New York: Springer Verlag. 1979.
20. Shannon, C., A Mathematical Theory of Communication, The Bell System Technical Journal, Vol. 27, pp. 379-423, 623-656, July, October, 1948.