Acta Academiae Paedagogicae Agriensis, Sectio Mathematicae **30** *(2003) 31–36*

A NOTE ON NON-NEGATIVE INFORMATION FUNCTIONS

Béla Brindza and Gyula Maksa (Debrecen, Hungary)

Dedicated to the memory of Professor Péter Kiss

Abstract. The purpose of the present paper is to make a first step to prove the conjecture, namely, that not every non-negative information function coincides with the Shannon's one on the algebraic elements of the closed unit interval.

1. Introduction

The characterization of the Shannon entropy, based upon its recursive and symmetric properties is strongly connected with the so-called fundamental equation of information, which is

$$f(x) + (1-x)f\left(\frac{y}{1-x}\right) = f(y) + (1-y)f\left(\frac{x}{1-y}\right) \tag{1.1}$$

where $f\colon [0,1] \to I\!R$ and (1.1) holds for all $x, y \in [0,1[$, $x + y \leq 1$.

The solutions of (1.1) satisfying $f(0) = f(1)$ and $f(\frac{1}{2}) = 1$ are the information functions. The basic monography Aczél and Daróczy [1] contains several results on these functions, like, if f is non-negative and bounded, then $f = S$, where

$$S(x) = -x\log_2 x - (1-x)\log_2(1-x), \qquad x \in [0,1],$$

($0\log_2 0$ is defined by 0). (See also Daróczy–Kátai [2]). A related result is

Theorem 1. (Daróczy–Maksa [3]). *If f is a non-negative information function, then*

$$f(x) \geq S(x), \qquad x \in [0,1] \tag{1.2}$$

moreover, there exists a non-negative information function different from S.

This research has been supported by the Hungarian Research Fund (OTKA) Grant T-030082 and by the Higher Educational Research and Development Fund (FKFP) Grant 0215/2001.

The proof of the second part of this theorem is based upon the existence of a non-identically zero real derivation $d\colon I\!R \to I\!R$ which is additive, that is

$$d(x+y) = d(x) + d(y) \qquad (x, y \in I\!R)$$

and satisfies the equation

$$d(xy) = xd(y) + yd(x), \qquad (x, y \in I\!R)$$

and different from 0 at some point. (See for example Kuczma [4]).

A computation shows that the function

$$f(x) = \begin{cases} S(x) + \dfrac{d(x)^2}{x(1-x)} & \text{if } x \in \,]0,1[\\ 0 & \text{if } x \in \{0,1\} \end{cases} \tag{1.3}$$

is a non-negative information function and different from S if d is a real derivation different from 0. (See Daróczy–Maksa [3]).

After this result some other natural questions arose, namely, the characterization of the non-negative information functions and (or at least) their Shannon kernel $\{x \in [0,1]\colon f(x) = S(x)\}$ where f is a fixed non-negative information function. (See Lawrence–Mess–Zorzitto [6], Maksa [7] and Lawrence [5].)

It is known that the real derivations are vanishing over the field of algebraic numbers (se Kuczma [4]), hence

$$f(\alpha) = S(\alpha) \tag{1.4}$$

if f is given by (1.3). It is noted that (1.4) holds for all non-negative information functions f and for all rational $\alpha \in [0,1]$. (See Daróczy–Kátai [2].)

Our conjecture is that there are non-negative information functions that are different from the Shannon's one at some algebraic element of $[0,1]$. In the next section we prove a partial result in this direction.

2. Results

The base of our investigations is the following theorem.

Theorem 2. *A function $f\colon [0,1] \to I\!R$ is a non-negative information function, if and only if, there exists an additive function $a\colon I\!R \to I\!R$ such that $a(1) = 1$,*

$$-xa(\log_2 x) - (1-x)a(\log_2(1-x)) \geq 0 \qquad \text{if} \quad x \in \,]0,1[\,, \tag{2.1}$$

and

$$(2.2)\qquad f(x) = \begin{cases} -xa(\log_2 x) - (1-x)a(\log_2(1-x)) & \textit{if } x \in \,]0,1[\\ 0 & \textit{if } x \in \{0,1\}. \end{cases}$$

Furthermore $f = S$ holds, if and only if, there is a real derivation $d\colon I\!R \to I\!R$ such that

$$(2.3)\qquad a(x) = x + 2^x d(2^{-x}) \qquad \textit{if} \quad x \in I\!R.$$

Proof. The first part of the theorem is an easy consequence of Theorem 1 of Daróczy–Maksa [3]. To prove the second part, first suppose that the non-negative information function f coincides with S on $[0,1]$. Therefore, by the definition of S and by (2.2), we get that

$$(2.4)\qquad -xa(\log_2 x) - (1-x)a(\log_2(1-x)) = -x\log_2 x - (1-x)\log_2(1-x)$$

holds for all $x \in \,]0,1[$ where a is an additive function that exists by the first part of the theorem. Define the function $\varphi\colon]0,+\infty[\, \to I\!R$ by

$$(2.5)\qquad \varphi(x) = -xa(\log_2 x) + x\log_2 x.$$

An easy calculation shows that

$$(2.6)\qquad \varphi(xy) = x\varphi(y) + y\varphi(x) \qquad \text{if} \quad x>0,\ y>0$$

and, because of (2.4),

$$\varphi(x) + \varphi(1-x) = 0 \qquad \text{if} \quad 0<x<1.$$

This implies that

$$\varphi\left(\frac{x}{x+y}\right) + \varphi\left(\frac{y}{x+y}\right) = 0$$

for all $x>0$, $y>0$ whence, applying (2.6), we have that

$$\begin{aligned} 0 &= x\varphi\left(\frac{1}{x+y}\right) + \frac{1}{x+y}\varphi(x) + y\varphi\left(\frac{1}{x+y}\right) + \frac{1}{x+y}\varphi(y) \\ &= (x+y)\varphi\left(\frac{1}{x+y}\right) + \frac{1}{x+y}(\varphi(x)+\varphi(y)) \\ &= \varphi(1) - \frac{1}{x+y}(\varphi(x+y) - \varphi(x) - \varphi(y)). \end{aligned}$$

Since $\varphi(1) = 0$, we dotain that

$$\varphi(x+y) = \varphi(x) + \varphi(y) \quad \text{if} \quad x > 0,\ y > 0. \tag{2.7}$$

If $x \in I\!R$ define the function $d\colon I\!R \to I\!R$ by

$$d(x) = \varphi(u) - \varphi(v)$$

where $u > 0$, $v > 0$ and $x = u - v$. Equation (2.7) garantees that the definition of d is correct, d is additive, and moreover, by (2.6) and (2.7), d is a real derivation that is an extension of φ to $I\!R$. Thus, by (2.5),

$$d(x) = -xa(\log_2 x) + x \log_2 x \qquad \text{if} \quad x > 0$$

whence we obtain (2.3) replecing x by 2^{-x}.

Finally, if d is an arbitrary real derivation then the function a defined by (2.3) is additive, $a(1) = 1$ and the function f given in (2.2) coincides with S on $[0, 1]$.

Since every real derivation vanishes at all algebraic points (see, for example Kuczma [4]), in order to prove our conjecture, by (2.3), we have to construct an additive function a for which $a(1) = 1$, $a(\log_2 \beta) \neq \log_2 \beta$ for some positive algebraic number β and (2.1) holds for all $x \in \,]0,1[$.

Instead of this we can proof the following weaker result only.

Theorem 3. *Let $Q(\alpha)$ be a real algebraic extension of Q of degree $n > 1$. If $Q[\alpha]$ (the ring of algebraic integers in $Q(\alpha)$) is a unique factorization domain then there exists an additive $a\colon I\!R \to I\!R$ with $a(1) = 1$ satisfying*

$$-xa(\log_2 x) - (1-x)a(\log_2(1-x)) \geq S(x) \qquad \text{if} \quad x \in \,]0,1[\, \cap Q[\alpha] \tag{2.8}$$

and

$$a(\log_2 \beta) \neq \log_2 \beta \tag{2.9}$$

for some positive algebraic number β.

Proof. Let U be the unitgroup of $Q[\alpha]$ generating by a set of fundamental units $\{\varepsilon_1, \dots, \varepsilon_{n-1}\}$ and $P = \{\pi_1, \dots, \pi_s, \dots\}$ be the set of primes in $Q[\alpha]$. Since the group of the roots of unity is $\{-1, 1\}$, only, we may assume that

$$0 < \varepsilon_i, \quad i = 1, \dots, n-1; \qquad 0 < \pi_j, \quad j = 1, 2, \dots .$$

and every non-zero element x of $Q[\alpha]$ can uniquely be written in the form

$$x = \pm \left(\prod_{i=1}^{n-1} \varepsilon_i^{k_i} \right) \left(\prod_{j=1}^{\infty} \pi_j^{\ell_j} \right) \tag{2.10}$$

where the exponents are (rational) integers and $\ell_j \geq 0$, $j = 1, 2, \ldots$. The set P is multiplicatively independent, hence the set $\{\log_2 \pi : \pi \in P\}$ is linearly independent (over $\mathbb{Q}$). Therefore there is a Hamel basis $\mathcal{H} \subset \mathbb{R}$ for which $1 \in \mathcal{H}$ and $\log_2 \pi \in \mathcal{H}$ if $\pi \in P$.

Let $\pi_1 \in P$ be fixed. We may assume that $\pi_1 \neq 2$. Define the function a_0 on $\mathcal{H}$ by $a_0(\log_2 \pi_1) = \log_2 \frac{\pi_1}{2}$, $a_0(h) = h$ if $h \in \mathcal{H}$, $h \neq \log_2 \pi_1$, and let a be the additive extension of a_0 to $\mathbb{R}$. It is obvious that $a(1) = 1$ and (2.9) is satisfied by $\beta = \pi_1$. To prove (2.8) first suppose that the exponent of π_1 is positive in the decomposition (2.10) of $x \in \,]0,1[\, \cap \mathbb{Q}[\alpha]$. Then the exponent of π_1 in the decomposition of $(1-x)$ is zero. Of course, the same is true also for $(1-x)$ instead of x. Therefore

$$a(\log_2(1-x)) = \log_2(1-x) \tag{2.11}$$

or

$$a(\log_2 x) = \log_2 x \tag{2.12}$$

holds for all $x \in \,]0,1[\, \cap \mathbb{Q}[\alpha]$. Supposing (2.11) we have that

$$\begin{aligned}
&-xa(\log_2 x) - (1-x)a(\log_2(1-x)) \\
&= -xa\left(\log_2 \frac{x}{\pi_1^{\ell_1}} + \log_2 \pi_1^{\ell_1}\right) - (1-x)\log_2(1-x) \\
&= -xa\left(\log_2 \frac{x}{\pi_1^{\ell_1}}\right) - xa(\log_2 \pi_1^{\ell_1}) - (1-x)\log_2(1-x) \\
&= -x\log_2 \frac{x}{\pi_1^{\ell_1}} - x\ell_1 a(\log_2 \pi_1) - (1-x)\log_2(1-x) \\
&= -x\log_2 x - (1-x)\log_2(1-x) + x\ell_1\left[\log_2 \pi_1 - a(\log_2 \pi_1)\right] \\
&= -x\log_2 x - (1-x)\log_2(1-x) + x\ell_1\left[\log_2 \pi_1 - \log_2 \frac{\pi_1}{2}\right] \\
&> -x\log_2 x - (1-x)\log_2(1-x) = S(x).
\end{aligned}$$

Thus (2.8) holds. In case (2.12) the proof is similar. Finally, if the exponent of π_1 is zero in the decompositions of both x and $(1-x)$ then, of course, the equality is valid in (2.8).

Remark. According to the classical approximation result of Dirichlet the set $D = \{\, x \in]0,1[\cap \mathbb{Q}[\alpha] : \ell_1 > 0 \text{ in } (2.10)\}$ is dense in $[0,1]$. Thus the strict inequality holds on the dense set D in (2.8).

References

[1] Aczél, J. and Daróczy, Z., *On measures of information and their characterizations,* Academic Press, New York, 1975.

[2] Daróczy, Z. and Kátai, I., Additive zahlentheoretische Funktionen und das Mass der Information, *Ann. Univ. Sci. Budapest, Eötvös Sect. Math.*, **13** (1970), 83–88.

[3] Daróczy, Z. and Maksa, Gy., *Nonnegative information functions.* In: Proc. Colloqu. Methods of Complex Anal. in the Theory of Probab. and Statist., Debrecen, 1977, Colloquia Mathematica Societatis János Bolyai Vol. 21, North-Holland, Amsterdam, 1979, 67–68.

[4] Kuczma, M., *An Introduction to the Theory of Functional Equations and Inequalities*, Państwowe Wydawnictwo Naukowe, Warszawa–Kraków–Katowice, 1985.

[5] Lawrence, J., The Shannon kernel of non-negative information function, *Aequationes Math.*, **23** (1981), 233–235.

[6] Lawrence, J., Mess, G. and Zorzitto, F., Near-derivations and information functions, *Proc. Amer. Math. Soc.*, **76** (1979), 117–122.

[7] Maksa, Gy., On near-derivations, *Proc. Amer. Math. Soc.*, **81**, (1981), 406–408.

Béla Brindza
University of Debrecen
Institute of Mathematics
4010 Debrecen P.O. Box 12.
Hungary
e-mail: brindza@math.klte.hu

Gyula Maksa
University of Debrecen
Institute of Mathematics
4010 Debrecen P.O. Box 12.
Hungary
e-mail: maksa@math.klte.hu