THE SHANNON FIELD OF NON-NEGATIVE INFORMATION FUNCTIONS

E. GSELMANN AND GY. MAKSA

Received December 9, 2008; revised December 27, 2008

ABSTRACT. Motivating the known result that there are non-negative information functions different from the Shannon information function, we investigate how large the set is on which every non-negative information function coincides with the Shannon's one. The structure of this set is also discussed.

1. INTRODUCTION

The characterizations of the Shannon entropy

$$H_n(p_1, \dots, p_n) = -\sum_{k=1}^{n} p_k \log_2 p_k, \quad (p_1, \dots, p_n \in [0,1], \sum_{k=1}^{n} p_k = 1, 0 \log_2 0 = 0)$$

based upon its recursive and symmetric (or just semisymmetric) properties lead to the so-called fundamental equation of information

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

where $f : [0,1] \to \mathbb{R}$ (the reals) and (1.1) holds for all $x, y \in [0,1[$, $x + y \le 1$. The solutions f of (1.1) satisfying $f(0) = f(1)$ and $f\left(\frac{1}{2}\right) = 1$ are the information functions. The basic reference concerning the characterization problems of information measures (like the Shannon entropy) and the fundamental equation of information is the monograph of Aczél and Daróczy [2]. In this book several results on (1.1) are collected which state that $f = S$ on $[0,1]$ if f satisfies one of the following conditions: (i) continuous, (ii) continuous at 0, (iii) increasing on $[0, 1/2]$, (iv) Lebesgue integrable, (v) measurable, (iv) non-negative and bounded from above, where

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

is the Shannon information function. Furthermore, the general form of the information functions is determined there, as well, by proving that f is an information function if, and only if,

$$f(x) = \varphi(x) + \varphi(1-x) \quad (x \in [0,1]) \tag{1.2}$$

with some function $\varphi : [0, +\infty[\to \mathbb{R}$ satisfying the functional equation

$$\varphi(xy) = x\varphi(y) + y\varphi(x) \quad (x, y \in [0, +\infty[) \tag{1.3}$$

and $\varphi\left(\frac{1}{2}\right) = \frac{1}{2}$. Obviously, if $\varphi(x) = -x \log_2 x$, $x \in [0, +\infty[$ then $\varphi\left(\frac{1}{2}\right) = \frac{1}{2}$, φ satisfies (1.3), and (1.2) implies that $f = S$. However, as it was pointed out in [1], f does not determine φ unambiguously by (1.2). Indeed, if $d : \mathbb{R} \to \mathbb{R}$ is a real derivation, that is, d satisfies both functional equations

$$d(x+y) = d(x) + d(y) \quad \text{and} \quad d(xy) = xd(y) + yd(x)$$

for all $x, y \in \mathbb{R}$ and d is not identically zero (such a function does exist, see e.g. Kuczma [6], pp. 352, Theorem 2.) then (1.2) implies that $f = S$ also with the choice $\varphi(x) = -x \log_2 x + d(x)$, $x \in [0, +\infty[$. This is the main difficulty in deriving the particular solutions of (1.1) from the general one.

2000 *Mathematics Subject Classification.* 39B50, 94A17.

Key words and phrases. Shannon field, Shannon kernel, non-negative information function.
This research has been supported by the Hungarian Scientific Research Fund (OTKA) Grants NK 68040 and K 62316.

The non-negativity property of an information function is very natural also from information theoretical point of view, since $f(x)$ is the measure of information belonging to the probability distribution $\{x, 1-x\}$.

Throughout the paper IF_+ will denote the set of all non-negative information functions. In [4], Daróczy and Kátai proved that $f(r) = S(r)$ for all $f \in \mathrm{IF}_+$ and $r \in [0,1] \cap \mathbb{Q}$ ($\mathbb{Q}$ denotes the set of the rational numbers), and they showed that $f = S$ on $[0,1]$ under the additional supposition that f is bounded from above. In [5], Daróczy and Maksa proved that $f(x) \geq S(x)$ for all IF_+ and for all $x \in [0,1]$ (the same was proved independently also in Lawrence–Mess–Zorzitto [8]) but there exists $f_0 \in \mathrm{IF}_+$ different from S, namely

$$f_0(x) = \begin{cases} S(x) + \frac{d(x)^2}{x(1-x)}, & \text{if} \quad x \in]0,1[\\ 0, & \text{if} \quad x \in \{0,1\} \end{cases} \tag{1.4}$$

defines such a function, where d is a not identically zero real derivation.

In [7], Lawrence introduced the concept of the Shannon kernel for $f \in \mathrm{IF}_+$ as

$$K_f = \{x \in [0,1] | f(x) = S(x)\},$$

formulating the conjecture that a subset $K \subset [0,1]$ is the Shannon kernel of some $f \in \mathrm{IF}_+$ if, and only if, $K = L \cap [0,1]$, where $L \subset \mathbb{R}$ is a subfield of $\mathbb{R}$, algebraically closed in $\mathbb{R}$ (real closed), and discussed what progress had been made towards proving this conjecture.

In this paper we show that, for all $f \in \mathrm{IF}_+$, the Shannon kernel K_f has the form $[0,1] \cap L_f$ where L_f is a subfield of $\mathbb{R}$ containing the square roots of its non-negative elements. We should remark that, in [7] it was announced, without proof and references, that this statement had been proved by 'K. Davidson and G. Maksa (independently)'. However, to the best of our knowledge, the proof has not been published till now.

In the present paper we also prove that all elements of $K = \bigcap \{K_f | f \in \mathrm{IF}_+\}$ are algebraic over $\mathbb{Q}$ and K contains all the algebraic elements of $[0,1]$ of degree at most 3. Throughout the paper, by algebraic number, we mean a real number algebraic over $\mathbb{Q}$.

2. Preliminary results

One of our tools is the following result which is a consequence of Theorem 1. in [5] that describes the general form of the non-negative information functions, Theorem 2. in the same paper about the minimal property of the Shannon information function and the mentioned Daróczy–Kátai's result in [4]. See also Theorem 1. and its proof in [8].

Theorem 2.1. *$f \in \mathrm{IF}_+$ if, and only if, there exists a function $\varphi : [0, +\infty[\to \mathbb{R}$ satisfying* (1.3) *such that*

$$\varphi(x+y) \leq \varphi(x) + \varphi(y), \quad (x, y \in [0, +\infty[) \tag{2.1}$$

$$\varphi(r) = 0, \quad (r \in [0, +\infty[\cap \mathbb{Q}) \tag{2.2}$$

and

$$f(x) = S(x) + \varphi(x) + \varphi(1-x). \quad (x \in [0,1]) \tag{2.3}$$

Proof. Due to Theorem 1. in [5], $f \in \mathrm{IF}_+$ if, and only if, there exists a function $\psi : [0, +\infty[\to \mathbb{R}$ with $\psi\left(\frac{1}{2}\right) = \frac{1}{2}$ such that

$$\psi(xy) = x\psi(y) + y\psi(x) \quad \text{and} \quad \psi(x+y) \leq \psi(x) + \psi(y) \quad (x, y \in [0, +\infty[) \tag{2.4}$$

and

$$f(x) = \psi(x) + \psi(1-x). \quad (x \in [0,1]) \tag{2.5}$$

First suppose that $f \in \mathrm{IF}_+$ and define the function φ on $[0, +\infty[$ by

$$\varphi(x) = x \log_2 x + \psi(x). \tag{2.6}$$

Then ψ satisfies (1.3) and (2.5) implies (2.3). Furthermore, by Theorem 2. in [5] and by [4], $f \geq S$ on $[0,1]$ and $f = S$ on $[0,1] \cap \mathbb{Q}$, respectively. Thus it follows from (2.5) and (2.6) that $\varphi(x) + \varphi(1-x) \geq 0$, $x \in [0,1]$ and $\varphi(r) + \varphi(1-r) = 0$, $r \in [0,1] \cap \mathbb{Q}$. On the other hand, (1.3) yields that

$$\varphi\left(\frac{x}{x+y}\right) + \varphi\left(1 - \frac{x}{x+y}\right) = \frac{1}{x+y}\left[\varphi(x) + \varphi(x) - \varphi(x+y)\right] \quad (x, y \in [0,+\infty[,\ x+y > 0)$$

whence (2.1) and the equation

$$\varphi(x+y) = \varphi(x) + \varphi(y) \quad (x, y \in [0,+\infty[\cap\mathbb{Q})$$

follow. Therefore $\varphi(r) = r\varphi(1) = 0$ for all $r \in [0,+\infty[\cap\mathbb{Q}$, i.e., (2.2) holds true, as well.

Conversely, if φ is a function with the properties listed in Theorem 2.1. and f is defined by (2.3), then we have (2.4) for the function ψ defined by $\psi(x) = \varphi(x) - x\log_2 x$, $x \in [0,+\infty[$. Since $\varphi(1/2) = 1/2$ we obtain that $f \in \mathrm{IF}_+$. □

The function $-\varphi$, where $\varphi : [0,+\infty[\to \mathbb{R}$ has the properties (1.3), (2.1) and (2.2) is called near-derivation in [8]. In [9], we gave a representation of the near-derivations (See Theorems 1. and 2.). The following theorem is an easy consequence of this result and Theorem 2.1.

Theorem 2.2. *$f \in \mathrm{IF}_+$ if, and only if, there exists a function $A : \mathbb{R}^2 \to \mathbb{R}$ satisfying the conditions*

$$A(x,y) = A(y,x), \quad (x, y \in \mathbb{R}) \tag{2.7}$$

$$A(x+y,z) = A(x,z) + A(y,z), \quad (x, y, z \in \mathbb{R}) \tag{2.8}$$

$$A(x,x) \geq 0, \quad (x \in \mathbb{R}) \tag{2.9}$$

$$A(xy,z) + zA(x,y) = A(x,yz) + xA(y,z), \quad (x, y, z \in \mathbb{R}) \tag{2.10}$$

$$A\left(x, \frac{1}{x}\right) \leq 0 \quad (0 \neq x \in \mathbb{R}) \tag{2.11}$$

such that the series $\sum 2^{n-1}x^{1-2^{-n}}A\left(x^{2^{-n}}, x^{2^{-n}}\right)$ is convergent for all $x \geq 0$ (with the convention $0^0 = 0$) and

$$f(x) = S(x) + \sum_{n=1}^{\infty} 2^{n-1}\left[x^{1-2^{-n}}A\left(x^{2^{-n}}, x^{2^{-n}}\right) + (1-x)^{1-2^{-n}}A\left((1-x)^{2^{-n}}, (1-x)^{2^{-n}}\right)\right] \tag{2.12}$$

holds for all $x \in [0,1]$.

3. The main results

First we prove the following

Lemma 3.1. *Suppose that the function $A : \mathbb{R}^2 \to \mathbb{R}$ has the properties* (2.7)–(2.11). *Then the set $F = \{x \in \mathbb{R} | A(x,x) = 0\}$ is a subfield of $\mathbb{R}$ containing the square roots of its non-negative elements.*

Proof. Inequalities (2.9) and (2.11) imply that $1 \in F$. Since A is a symmetric, positive semi-definite bilinear form on the $\mathbb{Q}$-vector space $\mathbb{R} \times \mathbb{R}$, we have that $\mathbb{Q} \subset F$ and A satisfies the Cauchy-Schwarz inequality

$$|A(x,y)| \leq \sqrt{A(x,x)}\sqrt{A(y,y)}. \quad (x, y \in \mathbb{R}) \tag{3.1}$$

Therefore $x \in F$ implies that $A(x,y) = A(y,x) = 0$ for all $y \in \mathbb{R}$. Thus, it follows from the identity

$$A(x-y, x-y) = A(x,x) - 2A(x,y) + A(y,y) \quad (x, y \in \mathbb{R})$$

that $x-y\in F$ if $x,y\in F$. Furthermore, if $x,y\in F$, $y\neq 0$ then substituting $\frac{x}{y^2}$ instead of x and $\frac{x}{y}$ instead of z, respectively into (2.10) we have that

$$A\left(\frac{x}{y},\frac{x}{y}\right)+\frac{x}{y}A\left(\frac{x}{y^2},y\right)=A\left(\frac{x}{y^2},x\right)+\frac{x}{y^2}A\left(y,\frac{x}{y}\right)$$

whence $A\left(\frac{x}{y},\frac{x}{y}\right)=0$, i.e., $\frac{x}{y}\in F$ follows. Thus we have proved that F is a field. Finally, let $0<x\in F$, $y=\frac{1}{\sqrt{x}}$ and $z=\sqrt{x}$. Then by (2.9), (2.10), (2.7) and (2.11),

$$0\leq A\left(\sqrt{x},\sqrt{x}\right)=A\left(x\frac{1}{\sqrt{x}},\sqrt{x}\right)+\sqrt{x}A\left(x,\frac{1}{\sqrt{x}}\right)=A(x,1)+xA\left(\frac{1}{\sqrt{x}},\sqrt{x}\right)=xA\left(\sqrt{x},\frac{1}{\sqrt{x}}\right)\leq 0,$$

hence $\sqrt{x}\in F$. □

One of our main results is the following

Theorem 3.2. *Let $f\in \mathrm{IF}_+$. Then there exists a unique subfield L_f of $\mathbb{R}$ containing the square roots of its non-negative elements such that $[0,1]\cap L_f$ is the Shannon kernel of f.*

Proof. Theorem 2.2. implies that there exists a function $A:\mathbb{R}^2\to\mathbb{R}$ satisfying (2.7)–(2.11) such that (2.12) holds for all $x\in[0,1]$. Let $L_f=\{x\in\mathbb{R}|A(x,x)=0\}$. Then, by Lemma 3.1., L_f is a subfield of $\mathbb{R}$ containing the square roots of its elements. Therefore, if x is an element of the Shannon kernel K_f of f then, by (2.12), $\sqrt{x}\in L_f$ thus $x\in L_f$. On the other hand, if $x\in[0,1]\cap L_f$ then $1-x\in[0,1]\cap L_f$, moreover $x^{2^{-n}},(1-x)^{2^{-n}}\in[0,1]\cap L_f$ for each positive integer n. Hence $x\in K_f$ follows from (2.12). The uniqueness in obvious. □

For $f\in\mathrm{IF}_+$ the field F_f will be called the Shannon field of f. So, Theorem 3.2. can be formulated as follows: The Shannon kernel of any non-negative information function is the closed unit interval of its Shannon field.

The Shannon field of the Shannon information function is, of course, $\mathbb{R}$ itself. However, the Shannon field of $f_0\in IF_+$ defined in (1.4) with a non identically zero derivation d, is a proper subfield of $\mathbb{R}$. The following theorem shows that some more is true.

Theorem 3.3. *Let L_f be the Shannon field of $f\in\mathrm{IF}_+$. Then all the elements of the field $L=\bigcap\{L_f|f\in\mathrm{IF}_+\}$ are algebraic.*

Proof. Suppose, in the contrary, that there is a transcendental element α in L. We may (and do) suppose that $\alpha\in]0,1[$. It follows from [6] (pp. 352, Theorem 1.) that there is a derivation d so that $d(\alpha)=1$. Define $f_0\in\mathrm{IF}_+$ by (1.4). Then $\alpha\notin L_{f_0}$ and so $\alpha\notin L$ which is a contradiction. □

In the remaining part of the paper we show that the intersection of all the Shannon fields in not too small since it contains the algebraic numbers of degree at most three. To do this, we need a consequence of a Hahn–Banach type theorem and the aid of the computer algebra package Maple V Release 9.

First we prove the following

Lemma 3.4. *Let $n>1$ be a fixed integer, suppose that the function $A:\mathbb{R}^2\to\mathbb{R}$ has the properties (2.7)–(2.11), and $A(1,1)=0$. Furthermore, let $\alpha\in\mathbb{R}$ be an algebraic number of degree n and $F=\mathbb{Q}(\alpha)$ be the smallest subfield of $\mathbb{R}$ containing α. Then there exist functions $a_1,\dots,a_{n-1}:F\to\mathbb{R}$ such that a_k is additive, i.e.,*

$$a_k(x+y)=a_k(x)+a_k(y),\quad(x,y\in F)$$

(3.2) $$a_k(\alpha^j)=0$$

for all $k = 1, \dots , n-1$*;* $j = 0, \dots , k-1$*, and*

$$A(x, y) = \sum_{k=1}^{n-1} a_k(x)a_k(y). \quad (x, y \in F) \tag{3.3}$$

Proof. Due to the Cauchy–Schwarz inequality (3.1) the function $p : F \to \mathbb{R}$ defined by $p(x) = \sqrt{A(x, x)}$, $x \in F$ is sublinear in the sense of Definition 1.1. in Berz [3] (see also Remark 2.1. there), that is, $p(x + y) \le p(x) + p(y)$ and $p(mx) = mp(x)$ hold for all $x, y \in F$ and non-negative integers m. Therefore, by Lemma 1.3. in [3], there exists an additive function $a_1 : F \to \mathbb{R}$ such that

$$a_1(\alpha)^2 = A(\alpha, \alpha) \tag{3.4}$$

and $a_1(x) \le \sqrt{A(x, x)}$ for all $x \in F$. Since a_1 is odd and $x \mapsto \sqrt{A(x, x)}$, $x \in F$ is even function, this inequality implies that $a_1(x)^2 \le A(x, x)$, $x \in F$. Obviously, $a_1(1) = 0$.

This argument can be repeated also for the function $(x, y) \mapsto A(x, y) - a_1(x)a_1(y)$ $(x, y \in L)$ and for α^2 instead of α. We obtain that there exists an additive function $a_2 : F \to \mathbb{R}$ such that

$$a_1(\alpha^2)^2 + a_2(\alpha^2)^2 = A(\alpha^2, \alpha^2)$$

and

$$a_1(x)^2 + a_2(x)^2 \le A(x, x). \quad (x \in F)$$

Obviously, $a_1(1) = 0$ and, by (3.4), $a_2(\alpha) = 0$.

Repeating the argument again and again, finally we find additive functions $a_1, \dots , a_{n-1} : F \to \mathbb{R}$ such that

$$a_1(\alpha^k)^2 + \ldots + a_k(\alpha^k)^2 = A(\alpha^k, \alpha^k), \quad (k = 1, \dots , n-1) \tag{3.5}$$

$$a_1(x)^2 + \ldots + a_k(x)^2 \le A(x, x), \quad (x \in F)$$

and

$$a_k(\alpha^j) = 0. \quad (k = 1, \dots , n-1; j = 0, \dots , k-1) \tag{3.6}$$

Now we show that (3.3) holds. Indeed, define the function B on $F \times F$ by

$$B(x, y) = A(x, y) - \sum_{k=1}^{n-1} a_k(x)a_k(y).$$

Then B is a symmetric, positive semi-definite bilinear form on the n-dimensional $\mathbb{Q}$-vector space $F \times F$, furthermore it follows from (3.5), (3.6) and the Cauchy-Schwarz inequality for B that $B(\alpha^i, \alpha^j) = 0$ for all $i, j \in \{0, \dots , n-1\}$. Since $\{1, \alpha, \dots , \alpha^{n-1}\}$ is a base for F as a $\mathbb{Q}$-linear vector space we get that $B \equiv 0$ on $F \times F$, consequently (3.3) holds. □

Now we are ready to prove the following

Theorem 3.5. *The intersection of all the Shannon fields contains the algebraic numbers of degree at most three.*

Proof. The statement follows from Theorem 3.2. for algebraic numbers of degree at most two. Let $\alpha \in \mathbb{R}$ be algebraic of degree three. Then there exist unique rational numbers R_0, R_1, R_2 such that $0 \ne R_0$,

$$\alpha^3 + R_2\alpha^2 + R_1\alpha + R_0 = 0, \tag{3.7}$$

and $\{1, \alpha, \alpha^2\}$ forms a base for the field $L = \mathbb{Q}(\alpha)$ as a $\mathbb{Q}$-vector space. Since L is a field $0 \ne x \in L$ implies that $\frac{1}{x} \in L$. To apply the property

$$A\left(x, \frac{1}{x}\right) \le 0, \quad (0 \ne x \in L) \tag{3.8}$$

firstly we look for $\frac{1}{x}$ in the form

$$\frac{1}{x} = s_2\alpha^2 + s_1\alpha + s_0,$$

with certain $s_0, s_1, s_2 \in \mathbb{Q}$. However, before this, let us observe that

$$\alpha^3 = -R_2\alpha^2 - R_1\alpha - R_0$$

and

$$\alpha^4 = \left(R_2^2 - R_1\right)\alpha^2 + (R_2R_1 - R_0)\alpha + R_2R_0.$$

Let $0 \neq x \in L$, then $x = r_2\alpha^2 + r_1\alpha + r_0$, with certain $r_0, r_1, r_2 \in \mathbb{Q}$, therefore

$$\begin{aligned}(3.9) \quad 1 = x\frac{1}{x} &= (r_2\alpha^2 + r_1\alpha + r_0)(s_2\alpha^2 + s_1\alpha + s_0)\\ &= r_2s_2\alpha^4 + (r_2s_1 + r_1s_2)\alpha^3 + (r_2s_0 + r_1s_1 + r_0s_2)\alpha^2 + (r_1s_0 + r_0s_1)\alpha + r_0s_0\\ &= \left[r_2s_2(R_2^2 - R_1) - R_2(r_2s_1 + r_1s_2) + (r_2s_0 + r_1s_1 + r_0s_2)\right]\alpha^2\\ &\quad + \left[r_2s_2(R_2R_1 - R_0) - R_1(r_2s_1 + r_1s_2) + (r_1s_0 + r_0s_1)\right]\alpha\\ &\quad + \left[r_2s_2R_2R_0 - R_0(r_2s_1 + r_1s_2) + r_0s_0\right].\end{aligned}$$

Thus the following system of equations has to hold.

$$\left\{\begin{array}{rcl} r_2s_2(R_2^2 - R_1) - R_2(r_2s_1 + r_1s_2) + (r_2s_0 + r_1s_1 + r_0s_2) &=& 0\\ r_2s_2(R_2R_1 - R_0) - R_1(r_2s_1 + r_1s_2) + (r_1s_0 + r_0s_1) &=& 0\\ r_2s_2R_2R_0 - R_0(r_2s_1 + r_1s_2) + r_0s_0 &=& 1. \end{array}\right.$$

To solve this we use the computer algebra package Maple V Release 9. Therefore, first we give the system of equations that will be solved

```
> eq_1:=((R_2^2-R_1)*r_2-R_2*r_1+r_0)*s_2+(r_1-R_2*r_2)*s_1+r_2*s_0=0;
> eq_2:=((R_2*R_1-R_0)*r_2-R_1*r_1+r_0)*s_2+(r_0-R_1*r_2)*s_1+r_1*s_0=0;
> eq_3:=(R_2*R_0*r_2-R_0*r_1)*s_2-R_0*r_2*s_1+r_0*s_0=1;
```

Then we determine its solutions by the help of the command

```
> solve({eq_1, eq_2, eq_3}, {s_0, s_1, s_2});
```

This produces

$$\begin{aligned} s_0 = -\big(&-r_1R_0r_2 + R_2R_0{r_2}^2 - {r_0}^2 - r_0R_2r_2 - r_0r_2{R_2}^2\\ &+2r_0R_1r_2 + r_0R_2r_1 - {R_1}^2{r_2}^2 - {r_1}^2R_1 + r_0r_1 + r_1r_2R_2R_1\big)N^{-1}, \end{aligned}$$

$$s_1\left(-R_0r_2^2 + r_2^2R_2R_1 + r_0r_2 - r_2R_2^2r_1 + r_1^2R_2 - r_0r_1\right)N^{-1}$$

and

$$s_2\left(r_1^2 - R_2r_2r_1 - r_0r_2 + R_1\right)N^{-1},$$

where

$$\begin{aligned} N = &-2r_0^2R_1r_2 + r_0^2R_2r_2 + r_0^2r_2R_2^2 - r_0r_1r_2R_2R_1 - r_0R_0r_2^2 + r_0R_1^2r_2^2 + 3r_0r_2R_0r_1\\ &- 2r_0R_2R_0r_2^2 + r_0r_1^2R_1 + r_0^3 + r_1^2R_2R_0r_2 - r_1^3R_0 + R_0^2r_2^3 - r_0^2R_2r_1 - r_0^2r_1 - R_0r_2^2R_1r_1. \end{aligned}$$

Due to (3.8),

$$\begin{aligned} 0 \geq A\left(x, 1/x\right) A&\left(r_2\alpha^2 + r_1\alpha + r_0, s_2\alpha^2 + s_1\alpha + s_0\right)\\ &= r_2s_2\left[a_1(\alpha^2)^2 + a_2(\alpha^2)^2\right] + r_1s_1a_1(\alpha)^2 + (r_2s_1 + r_1s_2)a_1(\alpha)a_1(\alpha^2). \end{aligned}$$

Therefore,

$$
\begin{aligned}
(3.10)\quad 0 \geq A(x, \frac{1}{x})& \\
= \big(& r_2r_1^2B^2 + r_2r_1^2C^2 - R_2r_2^2r_1B^2 - R_2r_2^2r_1C^2 - r_0r_2^2B^2 - r_0r_2^2C^2 + R_1r_2^3B^2 + R_1r_2^3C^2 \\
& - r_1A^2R_0r_2^2 + r_1A^2r_2^2R_2R_1 + r_1A^2r_0r_2 - r_1^2A^2r_2R_2^2 + r_1^3A^2R_2 - r_1^2A^2r_0 - ABR_0r_2^3 \\
& + ABR_2r_2^3R_1 + ABr_0r_2^2 - ABR_2^2r_2^2r_1 - 2ABr_1r_0r_2 + ABr_1^3 + ABr_1R_1r_2^2\big)\, N^{-1}
\end{aligned}
$$

holds for all $r_0, r_1, r_2 \in \mathbb{Q}$, where we used the notations

$$A = a_1(\alpha), \quad B = a_1(\alpha^2) \quad \text{and} \quad C = a_2(\alpha^2).$$

Let us observe that the denominator of $A\,(x, 1/x)$ is a polynomial of r_0 of degree 3 (It can be checked by using the command `> degree(denom(A(x, 1/x)), r_0);`). Thus the denominator of $A\,(x, 1/x)$ has at least one root having odd multiplicity, say x_0. Taking the limit $r_0 \to x_0$ one can observe that the denominator of $A\,(x, 1/x)$ changes sign, nevertheless $A\,(x, 1/x) \leq 0$. This implies that the numerator of $A(x, 1/x)$ has to be zero, if we take the limit $r_0 \to x_0$, that is,

(3.11)

$$
\begin{aligned}
&\big(-r_2^2C^2 - r_1^2A^2 + r_1A^2r_2 - r_2^2B^2 - 2ABr_1r_2 + ABr_2^2\big)\, x_0 + r_2r_1^2B^2 + r_2r_1^2C^2 - R_2r_2^2r_1B^2 - R_2r_2^2r_1C^2 \\
&\quad - ABR_2^2r_2^2r_1 + r_1A^2r_2^2R_2R_1 + R_1r_2^3B^2 + R_1r_2^3C^2 - r_1A^2R_0r_2^2 - ABR_0r_2^3 - r_1^2A^2r_2R_2^2 \\
&\quad r_1^3A^2R_2 + ABR_2r_2^3R_1 + ABr_1R_1r_2^2 + ABr_1^3 = 0
\end{aligned}
$$

holds for all $r_1, r_2 \in \mathbb{Q}$. Since $\mathbb{Q}$ is everywhere dense in $\mathbb{R}$, the same holds for all $r_1, r_2 \in \mathbb{R}$.

Let us observe that (3.11) is a polynomial of r_1 of degree at most 3, and it can also be considered as a polynomial of r_2. Furthermore, (3.11) yields that this polynomial is zero for all $r_1, r_2 \in \mathbb{R}$. Thus its partial derivatives with respect to r_1 and r_2 have to be zero, as well. If we differentiate (3.11) with respect to r_1 three times, we get that

$$6AB + 6A^2R_2 = 0,$$

the third order partial derivative of (3.11) with respect to r_2 has to be also zero, that is,

$$6R_1B^2 + 6R_1C^2 - 6ABR_0 + 6ABR_2R_1 = 0,$$

finally, for instance, differentiate (3.11) two times with respect to r_1, then with respect to r_2, then we obtain that

$$2B^2 + 2C^2 - 2A^2R_2^2 = 0.$$

Therefore A, B and C have to satisfy the system of equations

$$
\begin{cases}
A^2R_2 + AB &= 0 \\
R_1B^2 + R_1C^2 + ABR_2R_1 - ABR_0 &= 0 \\
B^2 + C^2 - R_2^2A^2 &= 0.
\end{cases}
$$

After solving this system of equations we obtain that

$$A = 0, \quad B = 0 \quad \text{and} \quad C = 0,$$

in case $R_2 \neq 0$. If $R_2 = 0$ then the third equation of the system implies that $B = C = 0$ and $A = 0$ follows from (3.11) with $r_1 = r_2 = 1$, that is,

$$a_1(\alpha) = 0, \quad a_1(\alpha^2) = 0 \quad \text{and} \quad a_2(\alpha^2) = 0.$$

This means that the function $A : L \times L \to \mathbb{R}$ is identically zero on the base. The biadditivity of the function $A : L \times L \to \mathbb{R}$ implies however, that A is identically zero on $L \times L$. Therefore the field $L = \mathbb{Q}(\alpha)$ is contained in $\bigcap \{K_f | f \in \mathrm{IF}_+\}$ indeed. $\square$

Similarly to that of Lawrence [7], our conjecture is that this argument will work also for algebraic numbers of degree greater than three and, in spite of technical difficulties, it will be possible to prove that the intersection of all the Shannon fields is just the field of the algebraic numbers.

References

[1] J. Aczél, *Derivations and information functions (a tale of two surprises and a half)*, Contributions to probability, pp. 191–200, Academic Press, New York-London, 1981.
[2] J. Aczél, Z. Daróczy, *On measures of information and their characterization*, Academic Press, New York – San Francisco – London, 1975.
[3] E. Berz, *Sublinear functions on* $\mathbb{R}$, Aequationes Math. **12** (1975), no. 2/3, 200–206.
[4] Z. Daróczy, I. Kátai, *Additive zahlentheoretische Funktionen und das Mass der Information*, (German) Ann. Univ. Sci. Budapest. Eötvös Sect. Math. **13** (1970), 83–88.
[5] Z. Daróczy, Gy. Maksa, *Nonnegative information functions*, Analytic function methods in probability theory (Proc. Colloq. Methods of Complex Anal. in the Theory of Probab. and Statist., Kossuth L. Univ. Debrecen, Debrecen, 1977), pp. 67–78, Colloq. Math. Soc. János Bolyai, 21, North-Holland, Amsterdam-New York, 1979.
[6] M. Kuczma, *An Introduction to the Theory of Functional Equations and Inequalities. Cauchy's equation and Jensen's inequality*, Prace Naukowe Uniwersytetu Śląskiego w Katowicach (PWN), Warsaw, 1985.
[7] J. Lawrence, *The Shannon kernel of a nonnegative information function*, Aequationes Math. **23** (1981), no. 2-3, 233–235.
[8] J. Lawrence, G. Mess, F. Zorzitto, *Near derivations and information functions*, Proc. Amer. Math. Soc. **76** (1979), no. 1, 117–122.
[9] Gy. Maksa, *On near derivations*, Proc. Amer. Math. Soc. **81** (1981), no. 3, 406–408.

Institute of Mathematics, University of Debrecen, P. O. Box: 12., Debrecen, Hungary, H–4010
E-mail address: gselmann@math.klte.hu

Institute of Mathematics, University of Debrecen, P. O. Box: 12., Debrecen, Hungary, H–4010
E-mail address: maksa@math.klte.hu