Contents lists available at ScienceDirect
J. Math. Anal. Appl.
www.elsevier.com/locate/jmaa

The stability of the entropy of degree alpha ☆

Gyula Maksa *

Institute of Mathematics, University of Debrecen, H-4010 Debrecen, Pf. 12, Hungary

ARTICLE INFO

Article history:
Received 30 October 2007

Submitted by P.G.L. Leach

Keywords:
Stability
Entropy of degree α
Fundamental equation of information
Functional equation

ABSTRACT

In this paper, we first prove that the generalized fundamental equation of information depending on a positive real parameter α, is stable in the sense of Hyers and Ulam provided that $\alpha \neq 1$, then we apply this result to prove the stability of a system of functional equations that characterizes the entropy of degree alpha or Havrda–Charvát entropy which has recently often been called the Tsallis entropy.

1. Introduction

The basic problem in the stability theory of functional equations is weather an "approximate solution" of a functional equation or of a system of functional equations "can be approximated" by a solution of this equation or of this system of equations. This question was originally raised in a talk by Ulam in 1940 (see also [15]) concerning the Cauchy equation and was answered in the affirmative by Hyers [7] who proved that the Cauchy equation is stable. This terminology can, of course, be also applied to other functional equations (see e.g. the survey papers by Forti [4] and Ger [5]). In this paper, we first consider a functional equation that arises in a natural way from the characterization problem of the entropy of degree alpha or Havrda–Charvát entropy which is a well-known information measure.

In what follows we denote the set of real numbers and the set of positive integers by $\mathbb{R}$ and $\mathbb{N}$, respectively. For fixed $0 < \alpha \in \mathbb{R}$ and $2 \leqslant n \in \mathbb{N}$ define the set

$$\Gamma_n = \left\{ (p_1, \ldots, p_n) \in \mathbb{R}^n \colon\ p_i \geqslant 0,\ i = 1, \ldots, n,\ \sum_{i=1}^{n} p_i = 1 \right\}$$

and the function H_n^α on Γ_n by

$$H_n^\alpha(p_1, \ldots, p_n) = \begin{cases} (2^{1-\alpha} - 1)^{-1}\left(\sum_{i=1}^n p_i^\alpha - 1\right) & \text{for } \alpha \neq 1, \\ -\sum_{i=1}^n p_i \log p_i & \text{for } \alpha = 1. \end{cases}$$

Here $\log = \log_2$ and the convention $0 \log 0 = 0$ is adapted. It is well known and easy to see that

$$\lim_{\alpha \to 1} H_n^\alpha = H_n^1 \quad \text{on } \Gamma_n \quad \text{for all } 2 \leqslant n \in \mathbb{N}.$$

☆ This research has been supported by the Hungarian Scientific Research Fund (OTKA) Grants NK 68040 and K 62316.
* Fax: 36 52 416 857.
E-mail address: maksa@math.klte.hu.

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2008.05.034

The sequence (H_n^α) is the entropy of degree α, and particularly (H_n^1) is the Shannon entropy. (H_n^1) was first introduced to the statistical thermodynamics by Boltzmann and Gipps, to the information theory by Shannon [12], while (H_n^α) (for $\alpha \neq 1$) was first investigated from cybernetic point of view by Havrda and Charvát [6], from information theoretical point of view by Daróczy [2], and was rediscovered by Tsallis [14] for the Physics community. The basic reference concerning the entropy of degree α is the book by Aczél and Daróczy [1]. In this note we borrow the following definitions from it. We say that the sequence (I_n) of functions $I_n : \Gamma_n \to \mathbb{R}$ $(n \geqslant 2)$ is *α-recursive* if

$$I_n(p_1, \dots, p_n) = I_{n-1}(p_1 + p_2, p_3, \dots, p_n) + (p_1 + p_2)^\alpha I_2\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right) \tag{1.1}$$

holds for all $2 < n \in \mathbb{N}$, $(p_1, \dots, p_n) \in \Gamma_n$ (here, and through the paper, the convention $(0+0)^\alpha I_2(\frac{0}{0+0}, \frac{0}{0+0}) = 0$ will be adapted);

3-semi-symmetric if

$$I_3(p_1, p_2, p_3) = I_3(p_1, p_3, p_2) \quad \text{on } \Gamma_3; \tag{1.2}$$

and *normalized* if

$$I_2\left(\frac{1}{2}, \frac{1}{2}\right) = 1. \tag{1.3}$$

We extend these definitions by saying that (I_n) is *2-semi-symmetric* if

$$I_2(1, 0) = I_2(0, 1). \tag{1.4}$$

The sequence (H_n^α) satisfies the properties above and, as it is proved in [2] and [1], is characterized by (1.1)–(1.4) for $0 < \alpha \neq 1$. The idea for the characterization of (H_n^α) in [2] and [1] is that the (1.1) α-recursivity and (1.2) 3-semi-symmetry imply that

$$g(x) + (1-x)^\alpha g\left(\frac{y}{1-x}\right) = g(y) + (1-y)^\alpha g\left(\frac{x}{1-y}\right) \tag{1.5}$$

holds on $D = \{(x, y) \in \mathbb{R}^2\colon\ x, y \in [0, 1[,\ x + y \leqslant 1\}$ for the function $g : [0, 1] \to \mathbb{R}$ defined by

$$g(x) = I_2(1 - x, x).$$

In case $\alpha = 1$, Eq. (1.5) is called the fundamental equation of information (see [1]). In [2] Daróczy proved that the only solution $g : [0, 1] \to \mathbb{R}$ of (1.5) with $0 < \alpha \neq 1$ satisfying $g(0) = g(1)$, $g(\frac{1}{2}) = 1$ (equivalently with (1.4) 2-semi-symmetry and (1.3) normalization) is

$$g(x) = \left(2^{1-\alpha} - 1\right)^{-1}\left[x^\alpha + (1-x)^\alpha - 1\right] \quad \left(x \in [0, 1]\right).$$

Thus the initial element of the α-recursive sequence (I_n) (and therefore also the sequence (I_n) itself) is uniquely determined by the properties of 3-semi-symmetry, 2-semi-symmetry, and normalization. The purpose of this paper is to prove the stability of the system of equations of α-recursivity and 3-semi-symmetry in case $0 < \alpha \neq 1$. In case $\alpha = 1$ only a partial result is known (see Morando [10]).

2. Main results

We first prove a theorem which generalizes the result of [2] and implies the stability of Eq. (1.5).

Theorem 2.1. *Let $\alpha, \varepsilon \in \mathbb{R}$, $0 < \alpha \neq 1$, $0 \leqslant \varepsilon$, and $f : [0, 1] \to \mathbb{R}$. Suppose that*

$$\left| f(x) + (1-x)^\alpha f\left(\frac{y}{1-x}\right) - f(y) - (1-y)^\alpha f\left(\frac{x}{1-y}\right)\right| \leqslant \varepsilon \tag{2.1}$$

holds for all $(x, y) \in D$. Then there exist $a, b \in \mathbb{R}$ such that

$$\left| f(x) - \left(ax^\alpha + b(1-x)^\alpha - b\right)\right| \leqslant 7\left|2^{1-\alpha} - 1\right|^{-1}\varepsilon \quad \left(x \in [0, 1]\right). \tag{2.2}$$

Proof. In the proof we adapt some ideas from [2]. Define the function F on $]0, 1[\times [0, 1]$ by

$$F(p, q) = f(1 - p) + p^\alpha f(q) - f(pq) - (1 - pq)^\alpha f\left(\frac{1-p}{1-pq}\right).$$

Then inequality (2.1), with the substitutions

$$x = 1 - p, \qquad y = pq, \quad (p, q) \in]0, 1[\times [0, 1],$$

Please cite this article in press as: Gy. Maksa, The stability of the entropy of degree alpha, J. Math. Anal. Appl. (2008), doi:10.1016/j.jmaa.2008.05.034

implies that

$$\left|F(p,q)\right| \leqslant \varepsilon \quad \left((p,q) \in]0,1[\times [0,1]\right). \tag{2.3}$$

On the other hand, for all $p, q \in]0,1[$, we have

$$\begin{aligned}&\left[q^\alpha + (1-q)^\alpha - 1\right]\left[f(p) - f(1)p^\alpha\right] - \left[p^\alpha + (1-p)^\alpha - 1\right]\left[f(q) - f(1)q^\alpha\right] \\ &\quad = F(q,p) - F(p,q) - F(q,1) + F(p,1) + (1-pq)^\alpha \left[F\left(\frac{1-p}{1-pq}q, 1\right) + F\left(\frac{1-p}{1-pq}, 1\right) - F\left(\frac{1-p}{1-pq}, q\right)\right].\end{aligned}$$

It follows from (2.3) that

$$\left|\left[q^\alpha + (1-q)^\alpha - 1\right]\left[f(p) - f(1)p^\alpha\right] - \left[p^\alpha + (1-p)^\alpha - 1\right]\left[f(q) - f(1)q^\alpha\right]\right| \leqslant 7\varepsilon.$$

Finally, with the substitution $q = \frac{1}{2}$ and with the notations

$$a = f(1) + \left(2^{1-\alpha} - 1\right)^{-1}\left(f\left(\frac{1}{2}\right) - f(1)2^{-\alpha}\right), \qquad b = a - f(1),$$

dividing both sides by $|2^{1-\alpha} - 1|$, and writing x in place of p we have that

$$\left|f(x) - \left(ax^\alpha + b(1-x)^\alpha - b\right)\right| \leqslant 7\left|2^{1-\alpha} - 1\right|^{-1}\varepsilon \quad \left(x \in]0,1[\right).$$

A direct calculation shows that this inequality holds also for $x = 0$ and $x = 1$. Indeed, for $x = 1$ the left-hand side is zero, while for $x = 0$ (with $y \to 0$), (2.1) implies that $|f(0)| \leqslant \varepsilon \leqslant 7|2^{1-\alpha} - 1|^{-1}\varepsilon$. □

We immediately get the following corollaries.

Corollary 2.2. *If $\varepsilon = 0$ then we have the general solution of* (1.5). (*See* [2].)

Corollary 2.3. *The functional equation* (1.5) *is stable and, because of the boundedness of the solutions of* (1.5), *it is superstable.* (*See e.g.* [4, *Definition* 3] *or* [11, *Définition* 6].)

Remark 2.4. The right-hand side of (2.2) tends to $+\infty$ whenever $\alpha \to 1$.

It is easy to see that the sequence of functions

$$(p_1, \ldots, p_n) \mapsto cH_n^\alpha(p_1, \ldots, p_n) + d\left(p_1^\alpha - 1\right), \quad (p_1, \ldots, p_n) \in \Gamma_n$$

is α-recursive and 3-semi-symmetric for all $c, d \in \mathbb{R}$. Thus the following theorem can also be considered as a stability theorem.

Theorem 2.5. *Let (I_n) be the sequence of functions $I_n : \Gamma_n \to \mathbb{R}$ $(n \geqslant 2)$ and suppose that there exist a sequence (ε_n) of non-negative real numbers and a real number $0 < \alpha \neq 1$ such that*

$$\left|I_n(p_1, \ldots, p_n) - I_{n-1}(p_1 + p_2, p_3, \ldots, p_n) - (p_1 + p_2)^\alpha I_2\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right)\right| \leqslant \varepsilon_{n-1} \tag{2.4}$$

holds for all $n \geqslant 3$ and $(p_1, \ldots, p_n) \in \Gamma_n$, and

$$\left|I_3(p_1, p_2, p_3) - I_3(p_1, p_3, p_2)\right| \leqslant \varepsilon_1 \quad \text{holds on } \Gamma_3. \tag{2.5}$$

Then there exist $c, d \in \mathbb{R}$ such that

$$\left|I_n(p_1, \ldots, p_n) - \left[cH_n^\alpha(p_1, \ldots, p_n) + d\left(p_1^\alpha - 1\right)\right]\right| \leqslant \sum_{k=2}^{n-1} \varepsilon_k + 7(n-1)(\varepsilon_1 + 2\varepsilon_2)\left|2^{1-\alpha} - 1\right|^{-1} \tag{2.6}$$

for all $n \geqslant 2$ and $(p_1, \ldots, p_n) \in \Gamma_n$. Here the convention $\sum_{k=2}^{1} \varepsilon_k = 0$ is adapted.

Proof. Let $(x, y) \in D$ and $n = 3$, $p_1 = 1 - x - y$, $p_2 = y$, $p_3 = x$ in (2.4). Then

$$\left|I_3(1-x-y, y, x) - I_2(1-x, x) - (1-x)^\alpha I_2\left(1 - \frac{y}{1-x}, \frac{y}{1-x}\right)\right| \leqslant \varepsilon_2 \tag{2.7}$$

and, by interchanging x and y in (2.7), we have that

$$\left|I_3(1-y-x, x, y) - I_2(1-y, y) - (1-y)^\alpha I_2\left(1 - \frac{x}{1-y}, \frac{x}{1-y}\right)\right| \leqslant \varepsilon_2. \tag{2.8}$$

Please cite this article in press as: Gy. Maksa, The stability of the entropy of degree alpha, J. Math. Anal. Appl. (2008), doi:10.1016/j.jmaa.2008.05.034

Therefore, by (2.7), (2.5), and (2.8), for the function $f:[0,1]\to\mathbb{R}$ defined by $f(x)=I_2(1-x,x)$, we obtain that

$$\begin{aligned}\left|f(x)+(1-x)^\alpha f\left(\frac{y}{1-x}\right)-f(y)-(1-y)^\alpha f\left(\frac{x}{1-y}\right)\right| &\leqslant \left|f(x)+(1-x)^\alpha f\left(\frac{y}{1-x}\right)-I_3(1-x-y,y,x)\right|\\ &\quad+\left|I_3(1-x-y,y,x)-I_3(1-x-y,x,y)\right|\\ &\quad+\left|I_3(1-y-x,x,y)-f(y)-(1-y)^\alpha f\left(\frac{x}{1-y}\right)\right|\\ &\leqslant 2\varepsilon_2+\varepsilon_1\end{aligned}$$

for all $(x,y)\in D$. Thus (2.1) holds with $\varepsilon=2\varepsilon_2+\varepsilon_1$. Therefore, by Theorem 2.1., we get (2.2) with $\varepsilon=2\varepsilon_2+\varepsilon_1$ and with some $a,b\in\mathbb{R}$. Let now $(p_1,p_2)\in\Gamma_2$. Then, with the notations $c=a(2^{1-\alpha}-1)$, $d=b-a$ and $x=1-p_1$, it follows from (2.2) that

$$\left|I_2(p_1,p_2)-cH_2^\alpha(p_1,p_2)-d\left(p_1^\alpha-1\right)\right|\leqslant 7(\varepsilon_1+2\varepsilon_2)\left|2^{1-\alpha}-1\right|^{-1}, \tag{2.9}$$

that is, (2.6) holds for $n=2$.

We continue the proof by induction on n. Suppose that (2.6) holds and, for the sake of brevity, introduce the notation

$$J_n(p_1,\dots,p_n)=cH_n^\alpha(p_1,\dots,p_n)+d\left(p_1^\alpha-1\right)$$

for all $2\leqslant n\in\mathbb{N}$, $(p_1,\dots,p_n)\in\Gamma_n$. Since (J_n) is α-recursive, for all $(p_1,\dots,p_{n+1})\in\Gamma_{n+1}$, we obtain that

$$\begin{aligned}&I_{n+1}(p_1,\dots,p_{n+1})-J_{n+1}(p_1,\dots,p_{n+1})\\ &\quad=I_{n+1}(p_1,\dots,p_{n+1})-J_n(p_1+p_2,p_3,\dots,p_{n+1})-(p_1+p_2)^\alpha J_2\left(\frac{p_1}{p_1+p_2},\frac{p_2}{p_1+p_2}\right)\\ &\quad=I_{n+1}(p_1,\dots,p_{n+1})-I_n(p_1+p_2,p_3,\dots,p_{n+1})-(p_1+p_2)^\alpha I_2\left(\frac{p_1}{p_1+p_2},\frac{p_2}{p_1+p_2}\right)\\ &\qquad+I_n(p_1+p_2,p_3,\dots,p_{n+1})-J_n(p_1+p_2,p_3,\dots,p_{n+1})\\ &\qquad+(p_1+p_2)^\alpha I_2\left(\frac{p_1}{p_1+p_2},\frac{p_2}{p_1+p_2}\right)-(p_1+p_2)^\alpha J_2\left(\frac{p_1}{p_1+p_2},\frac{p_2}{p_1+p_2}\right).\end{aligned}$$

Thus (2.4), the induction hypothesis, and (2.9) imply that

$$\begin{aligned}\left|I_{n+1}(p_1,\dots,p_{n+1})-J_{n+1}(p_1,\dots,p_{n+1})\right| &\leqslant \varepsilon_n+\sum_{k=2}^{n-1}\varepsilon_k+7(n-1)(\varepsilon_1+2\varepsilon_2)\left|2^{1-\alpha}-1\right|^{-1}\\ &\quad+7(\varepsilon_1+2\varepsilon_2)\left|2^{1-\alpha}-1\right|^{-1},\end{aligned}$$

that is, (2.6) holds also for $n+1$ instead of n. □

3. Open problems

Our arguments do not work if $\alpha=1$ or if we exclude zero probabilities and allow the non-positivity of α. Thus we have the following problems.

Problem 3.1. What about the stability of (1.5) on D for $\alpha=1$?

Problem 3.2. In [13] Székelyhidi asked the following. Let $f:]0,1[\to\mathbb{R}$ be a function so that the function Af which is defined on the open triangle $D^\circ=\{(x,y)\in\mathbb{R}^2:\ x,y,\ x+y\in\,]0,1[\}$ by

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

is bounded. Is it true that $f=g+b$ on $]0,1[$ where $g:]0,1[\to\mathbb{R}$ satisfies $Ag=0$ on D° and $b:]0,1[\to\mathbb{R}$ is bounded? This problem is open also for all $\alpha\in\mathbb{R}$, if Af is replaced by $A_\alpha f$ where

$$A_\alpha f(x,y)=f(x)+(1-x)^\alpha f\left(\frac{y}{1-x}\right)-f(y)-(1-y)^\alpha f\left(\frac{x}{1-y}\right).$$

We mention that the solutions $g:]0,1[\to\mathbb{R}$ of the equation $A_\alpha g=0$ are known (Maksa [8], Maksa and Ng [9]).

Problem 3.3. In the literature (see e.g. Ebanks, Sahoo and Sander [3]) higher dimensional information measures of multiplicative type are also considered. The problem of the stability of the connected functional equations and the system of equations is partly still open.

References

[1] J. Aczél, Z. Daróczy, On Measures of Information and Their Characterizations, Academic Press, New York, San Francisco, 1975.
[2] Z. Daróczy, Generalized information functions, Inform. and Control 16 (1970) 36–51.
[3] B. Ebanks, P. Sahoo, W. Sander, Characterizations of Information Measures, Word Scientific Publishing Co. Inc., River Edge, NJ, 1998.
[4] G.L. Forti, Hyers–Ulam stability of functional equations in several variables, Aequationes Math. 50 (1–2) (1995) 143–190.
[5] R. Ger, A survey of recent results on stability of functional equations, in: Proceedings of the 4th International Conference on Functional Equations and Inequalities, Pedagogical University in Cracow (1994), pp. 5–36.
[6] J. Havrda, F. Charvát, Quantification method of classification processes. Concept of structural α-entropy, Kybernetika 3 (1967) 30–35.
[7] D.H. Hyers, On the stability of the linear functional equations, Proc. Nat. Acad. Sci. USA 27 (1941) 222–224.
[8] Gy. Maksa, Solution on the open triangle of the generalized fundamental equation of information with four unknown functions, Util. Math. 21 (1982) 267–282.
[9] Gy. Maksa, C.T. Ng, The fundamental equation of information on open domain, Publ. Math. Debrecen 33 (1–2) (1986) 9–11.
[10] A. Morando, A stability result concerning Shannon entropy, Aequationes Math. 62 (2001) 286–296.
[11] Z. Moszner, Sur les définitions différentes de la stabilitédes équation fonctionelles, Aequationes Math. 68 (2004) 260–274.
[12] C.E. Shannon, A mathematical theory of communication, Bell System Tech. J. 27 (1948) 379–423, and 623–656.
[13] L. Székelyhidi, 38. Problem, Aequationes Math. 41 (1991) 302.
[14] C. Tsallis, Possible generalization of Boltzmann–Gibbs statistics, J. Stat Phys. 52 (1–2) (1988) 479–487.
[15] S.M. Ulam, A Collection of Mathematical Problems, Interscience Publ., New York, 1961;
S.M. Ulam, Problems in Modern Mathematics, Wiley, New York, 1964.

Please cite this article in press as: Gy. Maksa, The stability of the entropy of degree alpha, J. Math. Anal. Appl. (2008), doi:10.1016/j.jmaa.2008.05.034