Entropy functions and functional equations

ESZTER GSELMANN\textsuperscript{1,*}

\textsuperscript{1} Institute of Mathematics, University of Debrecen, 4010 Debrecen, P. O. Box 12, Hungary

Received March 16, 2010; accepted December 3, 2010

Abstract. The purpose of this note is to give a general solution of two functional equations connected to the Shannon entropy and also to the Tsallis entropy. As a result of this, we present a regular solution of these equations as well. Furthermore, we point out that the regularity assumptions used in previous works can be weakened substantially.

AMS subject classifications: Primary 39B22; Secondary 94A17

Key words: functional equation, entropy, Shannon entropy, Tsallis entropy

1. Introduction and preliminaries

Since the celebrated paper of Claude E. Shannon (see [10]) appeared, the information theory has become an extensive branch of mathematics. Furthermore, it is known that information measures can be characterized via functional equations. Concerning this, the reader can consult the two basic monographs Aczél–Daróczy [1] and Ebanks–Sahoo–Sander [4].

Although the characterization problem of information measures nearly comes to the end, from time to time one can meet new functional equations from this area. A possible explanation for this is that the Shannon entropy and also the entropy of degree alpha (or Tsallis entropy) has been re-discovered by physicists and engineers, see Daróczy [3] and Tsallis [12].

The aim of this note is to give a general solution of two functional equations connected to the notion of the Shannon entropy and also that of the Tsallis entropy. More precisely, in the second section we will firstly solve the equation

\begin{equation}
f(xy) + f((1-x)y) - f(y) = (f(x) + f(1-x))y^q,
\end{equation}

which is supposed to hold for the unknown function $f : [0, 1] \rightarrow \mathbb{R}$ for all $x \in [0, 1]$ and $y \in [0, 1]$, where $q \in \mathbb{R}$ is a fixed parameter.

For the unknown function $f : [0, 1] \rightarrow \mathbb{R}$ the equation

\begin{equation}
f(xy) = \left(\frac{x^\alpha + x^\beta}{2}\right)f(y) + \left(\frac{y^\alpha + y^\beta}{2}\right)f(x)
\end{equation}

will also be solved which is assumed to hold for all $x, y \in [0, 1]$, where $\alpha, \beta \in \mathbb{R}$ are fixed parameters.

*Corresponding author. Email address: gselmann@math.unideb.hu (E. Gselmann)
These two functional equations were solved in Sharma–Taneja [11] and also in Furuichi [5] under the assumptions that the unknown function is nonnegative and differentiable and they called the solutions of these equations entropy functions – to which alludes the title of the present work. First, we give a general solution of these equations and then we point out that the regularity suppositions (that is, nonnegativity and differentiability) can essentially be weakened to get the same result as that of [5, 11].

In what follows some preliminary definitions and results will be listed, mainly from the theory of functional equation. These results can also be found in Kuczma [8].

**Definition 1.** Let $I \subset \mathbb{R}$ and $A = \{(x, y)|x, y, x + y \in I\}$. A function $a : I \to \mathbb{R}$ is called additive on $A$ if

$$a(x + y) = a(x) + a(y)$$

holds for all pairs $(x, y) \in A$.

Consider the set $I = \{(x, y)|x, y, xy \in I\}$. We say that $\mu : I \to \mathbb{R}$ is multiplicative on $I$ if the functional equation

$$\mu(xy) = \mu(x)\mu(y)$$

is fulfilled for all $(x, y) \in I$.

A function $\ell : I \to \mathbb{R}$ is called logarithmic on $I$ if it satisfies the functional equation

$$\ell(xy) = \ell(x) + \ell(y)$$

for all $(x, y) \in I$.

Henceforth, for all $n \geq 2$ we define the set $D_n$ by

$$D_n = \left\{(x_1, \ldots, x_n) \in \mathbb{R}^n|x_1, \ldots, x_n, \sum_{i=1}^{n} x_i \in [0, 1]\right\}.$$

As we wrote above, we will also determine the regular solutions of equations (1) and (2). To do this, the following regularity theorems will be applied.

**Lemma 1.** Let $a : [0, 1] \to \mathbb{R}$ be an additive function on the set $D_2$ and assume that

(i) $a$ is bounded above or below on the subset of $[0, 1]$ that has a positive Lebesgue measure;

(ii) or $a$ is Lebesgue measurable.

Then there exists $c \in \mathbb{R}$ such that

$$a(x) = cx$$

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

Lemma 2. Let $\ell : [0,1] \to \mathbb{R}$ be a logarithmic function on the set $\tilde{L} = \{(x,y)|x,y,xy\in[0,1]\}$ and assume that

(i) $\ell$ is bounded above or below on the subset of $[0,1]$ that has a positive Lebesgue measure;

(ii) or $\ell$ is Lebesgue measurable.

Then there exists $c \in \mathbb{R}$ such that

$$\ell(x) = \ln(x)$$

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

We also mention that in case $a : [0,1] \to \mathbb{R}$ is an additive function on the set $D_2$, then it can be uniquely extended to the function $\tilde{a} : \mathbb{R} \to \mathbb{R}$ which is additive on $\mathbb{R}$ (cf. Kuczma [8]). For the sake of simplicity, we will always bear in mind this fact, and the extension of the function in question will always be denoted by the same character.

The notion of derivations will also be utilized in the next section, see Kuczma [8].

Definition 2. An additive function $a : \mathbb{R} \to \mathbb{R}$ is termed to be a real derivation, if

it also fulfills the equation

$$d(xy) = xd(y) + yd(x)$$

for all $x, y \in \mathbb{R}$.

From this definition it immediately follows that every real derivation vanishes at the rationals. Additionally, something more is true. Namely, every real derivation is identically zero on the set of algebraic numbers (over the rationals). Furthermore, if a real derivation is Lebesgue measurable or bounded above or below on the set that has a positive Lebesgue measure, then it is identically zero. Therefore, it can be seen that the non–trivial real derivations can be very irregular. Although it is surprising, there exists non identically zero real derivation, see Theorem 14.2.2 in Kuczma [8].

The following lemma was proved in [6].

Lemma 3. Suppose that the function $\varphi : [0,\infty[ \to \mathbb{R}$ is such that

$$\varphi(xy) = x\varphi(y) + y\varphi(x) \quad (x,y \in [0,1])$$

and the function $g : [0,1] \to \mathbb{R}$ defined by

$$g(x) = \varphi(x) + \varphi(1-x) \quad (x \in [0,1])$$

is Lebesgue measurable or it is bounded (above and below) on a subset of $[0,1]$ that has a positive Lebesgue measure. Then there exist $c \in \mathbb{R}$ and a real derivation $d : \mathbb{R} \to \mathbb{R}$ such that

$$\varphi(x) = cx \ln(x) + d(x)$$

is fulfilled for any $x \in [0,1]$.
In the proof of our main theorem concerning equation (1) we will apply a result concerning the so-called cocycle equation, see Jessen–Karpf–Thorup [7]. However, in the proof this equation (i.e., the cocycle equation) will not be satisfied on the whole domain but only on a restricted one. Therefore, we will apply a result of Ng [9], in which the author solves the cocycle equation on a restricted domain, see also Aczél–Ng [2] and Ebanks–Sahoo–Sander [4].

**Theorem 1** (See [9]). Let \( \mu : ]0, 1[ \rightarrow \mathbb{R} \) be a given multiplicative function and \( G : D_2 \rightarrow \mathbb{R} \) a function. Then a general solution of the system of functional equations

\[
G(x, y) = G(y, x); \quad ((x, y) \in D_2)
\]

\[
G(x, y) + G(x + y, z) = G(y, z) + G(x, y + z); \quad ((x, y, z) \in D_3)
\]

and

\[
G(tx, ty) = \mu(t)G(x, y) \quad (t \in ]0, 1[, (x, y) \in D_2)
\]

is given by in case \( \mu(x) = x \),

\[
G(x, y) = \varphi(x) + \varphi(y) - \varphi(x + y), \quad ((x, y) \in D_2)
\]

where \( \varphi : ]0, +\infty[ \rightarrow \mathbb{R} \) is such that

\[
\varphi(xy) = y\varphi(x) + x\varphi(y), \quad (x, y \in ]0, +\infty[)
\]

otherwise there exists \( c \in \mathbb{R} \) such that

\[
G(x, y) = c[\mu(x) + \mu(y) - \mu(x + y)]
\]

is fulfilled for all \((x, y) \in D_2\).

**2. Main results**

In this section we will find the general solutions of equations (1) and (2). After this, the regular solutions of these equations will be presented. Furthermore, it will be pointed out that the regularity assumptions of [5] and [11] can be substantially weakened. Moreover, in some cases these suppositions can even be omitted.

**Theorem 2.** Let \( q \in \mathbb{R} \) be arbitrarily fixed, then the function \( f : ]0, 1[ \rightarrow \mathbb{R} \) fulfills equation

\[
f(xy) + f((1 - x)y) - f(y) = (f(x) + f(1 - x)) y^q
\]

for all \( x \in ]0, 1[ \) and \( y \in ]0, 1[ \), in case \( q \neq 1 \), if and only if, there exist \( c \in \mathbb{R} \) and an additive function \( a : \mathbb{R} \rightarrow \mathbb{R} \) such that

\[
f(x) = \begin{cases} a(x) + cx^q, & \text{if } x \in ]0, 1[ \\ 0, & \text{if } x = 1 \end{cases}
\]

Furthermore, in case \( q = 1 \), if and only if there exists an additive function \( a : \mathbb{R} \rightarrow \mathbb{R} \) and a function \( \varphi : ]0, +\infty[ \rightarrow \mathbb{R} \) such that

\[
\varphi(xy) = x\varphi(y) + y\varphi(x) \quad (x, y \in ]0, +\infty[)
\]
and
\[
f(x) = \begin{cases} 
  a(x) + \varphi(x), & \text{if } x \in ]0, 1[ \\
  0, & \text{if } x = 1 
\end{cases}
\]
is fulfilled.

**Proof.** Assume that the function \( f : ]0, 1[ \to \mathbb{R} \) fulfills equation (10). With the substitution \( y = 1 \) we immediately get that \( f(1) = 0 \). Therefore it is enough to restrict ourselves to the interval \( ]0, 1[ \). Let \((u, v) \in D_2\) and in equation (10) let us replace \( x \) by \( \frac{u}{u + v} \) and \( y \) by \( (u + v) \), respectively. In this case we obtain that
\[
f(u) + f(v) - f(u + v) = \left[ f \left( \frac{u}{u + v} \right) + f \left( \frac{v}{u + v} \right) \right] (u + v)^q 
\]
holds for all \((u, v) \in D_2\).

Define the functions \( \mathcal{C}_f \) and \( \mathcal{R}_f \) on the set \( D_2 \) by
\[
\mathcal{C}_f(u, v) = f(u) + f(v) - f(u + v) \quad ((u, v) \in D_2)
\]
and
\[
\mathcal{R}_f(u, v) = \left[ f \left( \frac{u}{u + v} \right) + f \left( \frac{v}{u + v} \right) \right] (u + v)^q. \quad ((u, v) \in D_2)
\]
With these notations equation (11) yields
\[
\mathcal{C}_f(u, v) = \mathcal{R}_f(u, v). \quad ((u, v) \in D_2)
\]
Let us observe that the function \( \mathcal{R}_f \) is \( q \)-homogeneous. Indeed, for all \( t \in ]0, 1[ \) and \((u, v) \in D_2\)
\[
\mathcal{R}_f(tu, tv) = \left[ f \left( \frac{tu}{tu + tv} \right) + f \left( \frac{tv}{tu + tv} \right) \right] (tu + tv)^q \\
= t^q \left[ f \left( \frac{u}{u + v} \right) + f \left( \frac{v}{u + v} \right) \right] (u + v)^q = t^q \mathcal{R}_f(u, v).
\]
This implies that the function \( \mathcal{C}_f \) is also a \( q \)-homogeneous function. Furthermore, the function \( \mathcal{C}_f \) is symmetric and it also fulfills the cocycle equation. All in all, this means that the function \( \mathcal{C}_f \) satisfies equations (6), (7) and (8) with the multiplicative function \( \mu(t) = t^q \). Thus by Theorem 1, in case \( q \neq 1 \) there exists \( c \in \mathbb{R} \) such that
\[
\mathcal{C}_f(x, y) = c [x^q + y^q - (x + y)^q], \quad ((x, y) \in D_2)
\]
and in case \( q = 1 \) there exists a function \( \varphi : ]0, +\infty[ \to \mathbb{R} \) such that
\[
\varphi(xy) = x\varphi(y) + y\varphi(x) \quad (x, y \in ]0, +\infty[)
\]
and
\[
\mathcal{C}_f(x, y) = \varphi(x) + \varphi(y) - \varphi(x + y) \quad ((x, y) \in D_2)
\]
is satisfied.
Firstly, we deal with the case $q \neq 1$. Define the function $\tilde{f} : [0, 1] \to \mathbb{R}$ by

$$\tilde{f}(x) = f(x) - cx^q, \quad (x \in [0, 1]),$$

then equation (12) yields that the function $\tilde{f}$ is additive on the set $D_2$. Regarding the function $f$ this shows that there exists an additive function $a : \mathbb{R} \to \mathbb{R}$ such that

$$f(x) = cx^q + a(x)$$

holds for any $x \in [0, 1]$. Since $f(1) = 0$, the function $f$ has the form which had to be proved.

If $q = 1$, then let us define the function $\tilde{f} : [0, 1] \to \mathbb{R}$ by

$$\tilde{f}(x) = f(x) - \varphi(x). \quad (x \in [0, 1])$$

Equation (13) yields that the function $\tilde{f}$ is an additive function on $D_2$. Concerning the function $f$, from this we get that there exists an additive function $a : \mathbb{R} \to \mathbb{R}$ such that

$$f(x) = \varphi(x) + a(x)$$

is fulfilled for any $x \in [0, 1]$, where the function $\varphi : [0, +\infty] \to \mathbb{R}$ satisfies

$$\varphi(xy) = x\varphi(y) + x\varphi(y). \quad (x, y \in [0, +\infty])$$

Since $f(1) = 0$ has to hold, the function $f$ is of the form

$$f(x) = \begin{cases} a(x) + \varphi(x), & \text{if } x \in [0, 1] \\ 0, & \text{if } x = 1 \end{cases}$$

that had to be proved. The converse direction is an easy computation.

The following corollary contains the regular solutions of equation (10).

**Corollary 1.** Let $q \in \mathbb{R}$ be arbitrary and suppose that the function $f : [0, 1] \to \mathbb{R}$ satisfies equation (10) for all $x \in [0, 1]$ and $y \in [0, 1]$. If $q \neq 1$, assume further that one of the following is true.

(i) $f$ is bounded above or below on a subset of $[0, 1]$ that has a positive Lebesgue measure;

(ii) $f$ is Lebesgue measurable.

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

$$f(x) = \begin{cases} c^* x + cx^q, & \text{if } x \in [0, 1] \\ 0, & \text{if } x = 1 \end{cases}$$

In case $q = 1$, suppose additionally that one of the statements below holds.

(i) $f$ is bounded (above and below) on a subset of $[0, 1]$ that has a positive Lebesgue measure;
(ii) $f$ is Lebesgue measurable.

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

$$f(x) = \begin{cases} cx \ln(x) + c^* x, & \text{if } x \in ]0, 1[ \\ 0, & \text{if } x = 1 \end{cases}$$

is fulfilled.

**Proof.** Firstly, we investigate the case $q \neq 1$. From Theorem 2 we obtain that

$$f(x) = a(x) + cx^q$$

holds for all $x \in ]0, 1[$ and $f(1) = 0$. If we rearrange this, it follows that

$$a(x) = f(x) - cx^q. \quad (x \in ]0, 1[)$$

By our assumptions $f$ is bounded above or below on a subset of $]0, 1[$ that has a positive Lebesgue measure, or $f$ is a Lebesgue measurable function.

This implies that the additive function $a$ fulfills condition (i) or (ii) of Lemma 1. From this, we obtain that there exists a constant $c^* \in \mathbb{R}$ such that $a(x) = c^* x$. This implies however that

$$f(x) = c^* x + cx^q$$

holds for all $x \in ]0, 1[$ and $f(1) = 0$.

Secondly, we assume that $q = 1$. From Theorem 2, we obtain that $f(1) = 0$ and

$$f(x) = \varphi(x) + a(x), \quad (x \in ]0, 1[)$$

where $\varphi : ]0, +\infty[ \rightarrow \mathbb{R}$ fulfills equation

$$
\varphi(xy) = x \varphi(y) + y \varphi(x) \quad (x, y \in ]0, +\infty[)
$$

and $a : \mathbb{R} \rightarrow \mathbb{R}$ is an additive function. Define the function $g : ]0, 1[ \rightarrow \mathbb{R}$ by

$$g(x) = f(x) + f(1 - x) - a(1). \quad (x \in ]0, 1[)$$

In this case, the function $g$ is Lebesgue measurable or bounded on a subset of $]0, 1[$ that has a positive Lebesgue measure. Furthermore,

$$g(x) = f(x) + f(1 - x) - a(1) = \varphi(x) + a(x) + \varphi(1 - x) + a(1 - x) - a(1) = \varphi(x) + \varphi(1 - x),$$

where we used that $a$ is an additive function. All in all, this implies that the function $g$ satisfies the assumptions of Lemma 3. Thus there exist $c \in \mathbb{R}$ and a real derivation $d : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$
\varphi(x) = cx \ln(x) + d(x). \quad (x \in ]0, 1[)
$$

Concerning the function $f$ this yields that

$$f(x) = cx \ln(x) + d(x) + a(x), \quad (x \in ]0, 1[)$$
or if we rearrange this,

\[ f(x) - cx \ln(x) = d(x) + a(x). \quad (x \in ]0, 1[) \]

By our assumptions the function \( f \) is bounded on a subset of \( ]0, 1[ \) that has a positive Lebesgue measure or it is Lebesgue measurable. Furthermore, the function \( d(x) + a(x) \) is a sum of two additive functions, that is, this function is also additive, which is bounded on a subset of \( ]0, 1[ \) with a positive Lebesgue measure or it is Lebesgue measurable. In view of Lemma 1, there exists \( c^* \in \mathbb{R} \) such that \( d(x) + a(x) = c^*x \) holds for all \( x \in \mathbb{R} \). Thus

\[ f(x) = cx \ln(x) + c^*x \]

holds for all \( x \in ]0, 1[ \) and \( f(1) = 0 \).

\[ \square \]

\textbf{Remark 1.} Under the assumptions of the previous corollary, in case for the function

\[ \lim_{x \to 1^-} f(x) \]

exists and \( \lim_{x \to 1^-} f(x) = f(1) \), then in case \( q \neq 1 \),

\[ f(x) = c^*(x - x^q) \quad (]0, 1[) \]

is fulfilled with some \( c^* \in \mathbb{R} \). Furthermore, in case \( q = 1 \),

\[ f(x) = cx \ln(x) \quad (x \in ]0, 1[) \]

is satisfied with a certain \( c \in \mathbb{R} \).

At this point of the paper we turn to deal with equation (2). Before this, we present a more general equation. Thus the solutions of the above mentioned functional equation will be showed as a corollary of the following result. Additionally, the regular solutions of (2) will be dealt with, too.

The following lemma was proved by E. Vincze in 1962 for commutative groups. Although \( (]0, 1[, \cdot) \) is not a group, only a semigroup, we remark that the method used in Satz 5 in Vincze [13] is appropriate for commutative semigroups as well.

\textbf{Lemma 4.} Let \( g : ]0, 1[ \to \mathbb{R} \) be a given function, and \( f : ]0, 1[ \to \mathbb{R} \) be such that

\[ f(xy) = g(y)f(x) + g(x)f(y) \quad (14) \]

holds for all \( x, y \in ]0, 1[ \).

If \( g \) is a multiplicative function, then

\[ f(x) = g(x)\ell(x), \quad (x \in ]0, 1[) \]

where \( \ell : ]0, 1[ \to \mathbb{R} \) is a logarithmic function, and in case \( g \) is not a multiplicative function, that is, there exist \( t_1, t_2 \in ]0, 1[ \) such that \( g(t_1t_2) \neq g(t_1)g(t_2) \), then

\[ f(x) = \frac{[g(t_1x) - g(t_1)g(x)] f(t_1)}{g(t_1t_2) - g(t_1)g(t_2)} \]

holds for all \( x \in ]0, 1[ \).
Corollary 2. Let $\alpha, \beta \in \mathbb{R}$ be arbitrary, $f : [0, 1] \to \mathbb{R}$ a function for which
\[ f(xy) = \left( \frac{x^\alpha + x^\beta}{2} \right) f(y) + \left( \frac{y^\alpha + y^\beta}{2} \right) f(x) \] (15)
holds for any $x, y \in [0, 1]$. Then we have the following two possibilities.

(i) if $\alpha \neq \beta$, then there exists $c \in \mathbb{R}$ such that
\[ f(x) = c \left( x^\alpha - x^\beta \right), \quad (x \in [0, 1]) \]

(ii) if $\beta = \alpha$, then there exists a logarithmic function $\ell : [0, 1] \to \mathbb{R}$ such that
\[ f(x) = cx^\alpha \ell(x) \]
holds for all $x \in [0, 1]$.

Proof. Firstly, let us suppose that $\alpha \neq \beta$. In this case the function $g : [0, 1] \to \mathbb{R}$ defined by $g(x) = \frac{1}{2} (x^\alpha + x^\beta)$ is not a multiplicative function. Furthermore, with this notation, equation (14) follows from (15). Thus, by Lemma 4,
\[ f(x) = \frac{[g(t_1 x) - g(t_1)g(x)] f(t_1)}{g(t_1 t_2) - g(t_1)g(t_2)} \]
holds for all $x \in [0, 1]$, where $t_1, t_2 \in [0, 1]$ are arbitrarily fixed. After using the form of the function $g$, we get that
\[ f(x) = \frac{f(t_1)}{t_2^\alpha - t_1^\alpha} \left( x^\alpha - x^\beta \right), \quad (x \in [0, 1]) \]
that is, there exists a constant $c \in \mathbb{R}$ such that
\[ f(x) = c \left( x^\alpha - x^\beta \right) \]
is satisfied for all $x \in [0, 1]$.

Secondly, assume that $\alpha = \beta$. In this case the function $g : [0, 1] \to \mathbb{R}$ defined by $g(x) = x^\alpha$ is a multiplicative function. Additionally, let us observe that with these notations (15) becomes equation (14). Therefore, due to Lemma 4, there exist $c \in \mathbb{R}$ and a logarithmic function $\ell : [0, 1] \to \mathbb{R}$ such that
\[ f(x) = cx^\alpha \ell(x) \]
is fulfilled for all $x \in [0, 1]$.

Finally, a facile computation shows the correctness of the converse direction. \(\square\)

From this corollary we can effortlessly get the regular solutions of equation (15). Let us observe that in case $\alpha \neq \beta$, the solutions of this equation are regular already. Therefore, in this case the regularity assumptions (that is, the nonnegativity and the differentiability) are superfluous in the papers [5] and [11]. Furthermore, if $\alpha = \beta$, then the above mentioned regularity suppositions can be significantly weakened. Namely, making use of Lemma 2 the following statement can be proved.
Corollary 3. Let \( \alpha \in \mathbb{R} \) be arbitrarily fixed and assume that the function \( f : ]0,1] \rightarrow \mathbb{R} \) fulfills equation
\[
 f(xy) = y^\alpha f(x) + x^\alpha f(y) \tag{16}
\]
for all \( x, y \in ]0,1] \). Suppose further that one of the following statements is true.

(i) \( f \) is bounded above or below on a subset of \( ]0,1] \) that has a positive Lebesgue measure;

(ii) \( f \) is Lebesgue measurable.

Then there exists \( c \in \mathbb{R} \) such that
\[
 f(x) = cx^\alpha \ln(x). \quad (x \in \mathbb{R})
\]

Acknowledgement

The author is indebted to the anonymous referee. His/her suggestions improved the manuscript significantly.

This research has been supported by the Hungarian Scientific Research Fund (OTKA) Grant NK 814 02 and by the TÁMOP 4.2.1./B-09/1/KONV-2010-0007 project implemented through the New Hungary Development Plan co-financed by the European Social Fund and the European Regional Development Fund.

References


