On functional equations and measures of information I

By PL. KANNAPPAN and C. T. NG (Waterloo, Ontario)

§ 1. Introduction

In analysing the additivity property of Shannon's entropy one comes across the functional equation

$$\sum_{i=1}^{n}\sum_{j=1}^{m} f(p_i q_j) = \sum_{i=1}^{n} f(p_i) + \sum_{j=1}^{m} f(q_j).$$

Recently Z. DARÓCZY and A. JÁRAI [1] solved this equation with $n=m=2$ for (Lebesgue) measurable functions f.

The purpose of the present paper is to obtain the measurable solutions f_i: $]0, 1[\to R$ of the functional equation

(1.1)

$$f_1(pq)+f_2\big(p(1-q)\big)+f_3\big((1-p)q\big)+f_4\big((1-p)(1-q)\big)=f_5(p)+f_6(q) \quad (p, q\in]0, 1[)$$

which is a generalization of the equation considered in [1].

§ 2. Some preliminary results

Let X be a set and $(Y, \mathscr{B}, m)$ be a finite measure space. Let $\mathscr{D}$ be the family of all subsets $D\subseteq X\times Y$ satisfying the property

$$\inf_{x\in X} m(D_x) > 0, \quad \text{where} \quad D_x := \{y\in Y \,|\, (x, y)\in D\}\in\mathscr{B}.$$

Let $\mathscr{F}(D)=\{g\colon D\to Y\}$ be the family of all functions g such that for all $\varepsilon>0$, there exists $\delta>0$ so that for each $E\in\mathscr{B}$ with $m(E)<\delta$ and each $x\in X$, the set $g_x^{-1}(E):=$ $:=\{y\in Y \,|\, g(x, y)\in E\}$ is contained in some set B (depending upon x and E) $\in\mathscr{B}$ with $m(B)<\varepsilon$. We present the following Lemma without a proof.

Lemma 1. *Let* $D\in\mathscr{D}$ *and* $g_i\in\mathscr{F}(D)$ *for* $i=1, 2, \ldots, n$. *Then there exists* $\delta>0$ *so that for each* $E\in\mathscr{B}$ *with* $m(E)<\delta$,

$$\bigcup_{i=1}^{n} (g_i)_x^{-1}(E) \subsetneqq D_x \quad \textit{for every} \quad x\in X,$$

that is, for each $x\in X$ *there exists* $y_x\in D_x$ *such that*

$$g_i(x, y_x)\notin E \quad \textit{for all} \quad i = 1, 2, \dots, n.$$

Using Lemma 1 we can prove the following Theorem.

Theorem 2. *Let* $D\in\mathscr{D}$, $g_i\in\mathscr{F}(D)$, $h_i\colon Y\to R$ *and* $h\colon X\to R$ *be functions satisfying the functional equation*

$$h(x) = \sum_{i=1}^{n} h_i\bigl(g_i(x, y)\bigr) \quad (x, y)\in D.$$

If the functions h_i *are measurable, then* h *is bounded on* X.

The proofs of the above results are analogous to those used in [1].

We now proceed to show that if f_i's are (Lebesgue) measurable functions satisfying the functional equation (1.1), then they are locally bounded. We shall make use of the following lemma ([2], Lemma 3, p. 210) in Natanson.

Lemma 3. *Let* $h\colon [\alpha, \beta]\to R$ *be strictly monotone and suppose* $E\subset[\alpha, \beta]$ *is such that* $|h'(x)|\geqq b$ *holds on* E *for some* $b\geqq 0$, *then* $m^*h(E)\geqq bm^*E$, *where* m^* *is the (Lebesgue) outer measure.*

We can rewrite (1.1) as

$$\begin{aligned} f_1(x) = -f_2\bigl(g_1(x, y)\bigr)-f_3\bigl(g_2(x, y)\bigr)-f_4\bigl(g_3(x, y)\bigr)+\\ +f_5\bigl(g_4(x, y)\bigr)+f_6\bigl(g_5(x, y)\bigr), \quad (x, y)\in T \end{aligned} \tag{2.1}$$

where $g_1(x, y)=\dfrac{x}{y}-x$, $g_2(x, y)=y-x$, $g_3(x, y)=1+x-y-\dfrac{x}{y}$, $g_4(x, y)=\dfrac{x}{y}$, $g_5(x, y)=y$ and $T=\{(x, y)\,|\,0<x<y<1\}$.

Let $X=[\alpha, \beta]\subset]0, 1[$ $(\alpha<\beta)$ be an arbitrary closed interval. We shall show that f_1 is bounded on X.

Let $Y=]0, 1[$ and $\mathscr{B}$ be the class of all Lebesgue measurable subsets of Y with Lebesgue measure m. We can choose a sufficiently small $\eta>0$ such that the set

$$D = \{(x, y)\in T\,|\,x\in X,\ (y+\eta)^2 \leqq x\}$$

is in $\mathscr{D}$.

Since $\dfrac{\partial}{\partial y} g_i(x, y)$ $(i=1, 2, \dots, 5)$ are continuous and non-vanishing on the closure $\bar{D}$ of D in R^2, and since $\bar{D}$ is compact, there exists $b>0$ such that

$$\left|\frac{\partial}{\partial y}\, g_i(x, y)\right| \geqq b \quad \text{for all} \quad (x, y)\in\bar{D}.$$

Thus for each $x\in X$,

$$\left|\frac{\partial}{\partial y}\, g_i(x, y)\right| \geqq b \quad \text{for all} \quad y\in\overline{D_x},$$

and therefore the functions $g_i(x, \cdot)$ are strictly monotone on $\overline{D_x}$. So, by Lemma 3,

$$m^*\bigl(g_i(x, E)\bigr) \geqq bm^*E \quad \text{for all} \quad E\subset D_x.$$

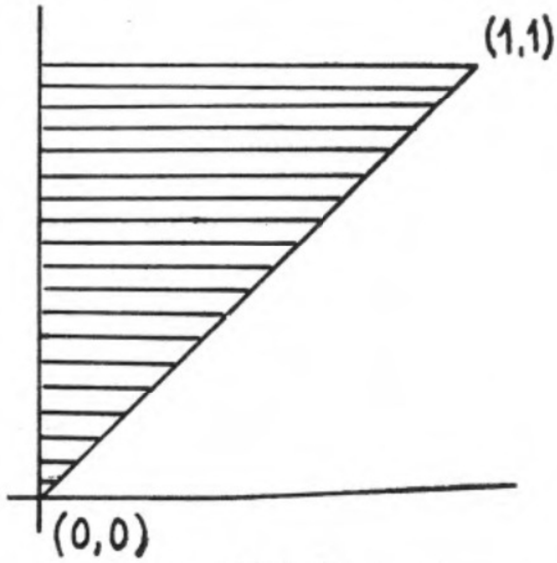

Figure-T

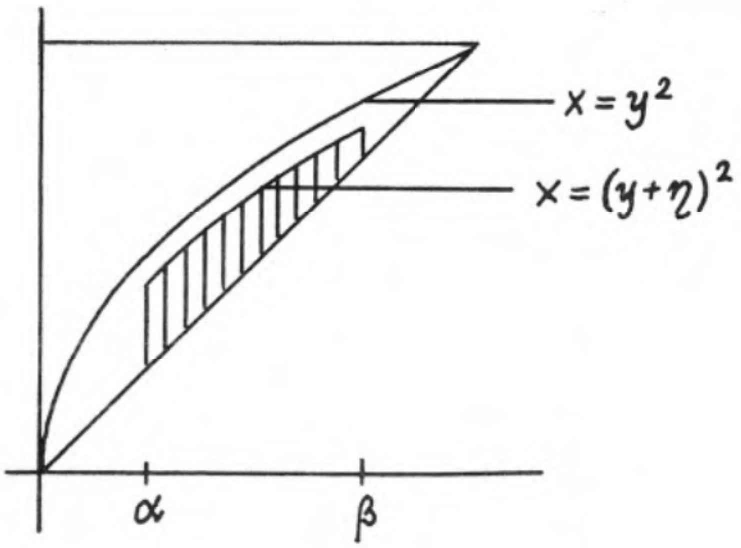

Figure-D

Now we claim that $g_i \in \mathscr{F}(D)$ and thereby the boundedness of f_1 on X follows from Theorem 2. In fact, let $\varepsilon > 0$ be given. Choose $\delta = b\varepsilon > 0$. Then for any $F \in \mathscr{B}$ with $m(F) < \delta$ and for each $x \in X$, we have

$$m^*\big((g_i)_x^{-1}(F)\big) \leqq \frac{1}{b}\, m^*\big(g_i(x,\, (g_i)_x^{-1}(F))\big) \leqq \frac{1}{b}\, m^*(F) < \varepsilon.$$

Hence $(g_i)_x^{-1}(F)$ is contained in some measureable set in Y of measure less than ε.

As the closed interval $[\alpha, \beta]$ is arbitrary, this proves the local boundedness of f_1 on $]0, 1[$. Since the f_i's $(i = 1, 2, 3, 4)$ play symmetric roles in (1.1) we have indeed established the following theorem.

Theorem 4. *If* f_i: $]0,1[\to R$ *are measurable functions satisfying* (1.1), *then they are all locally bounded.*

From Theorem 4, it follows that the measurable solutions f_i's of (1.1) are locally integrable and then by adapting similar techniques employed in ([1], Theorem 3) (which is now considered as standard in the theory of functional equations), it can be shown that the derivatives of all orders of the f_i's exist on $]0,1[$.

§ 3. Main results

We provide the measurable solutions of (1.1) through a sequence of auxiliary results. The first two propositions follow from ([1], Lemma 7, Theorem 4 and Lemma 5).

Proposition 5. *A measurable function* F: $]0,1[\to R$ *satisfies the equation*

$$F(pq)+F(p(1-q))+F((1-p)q)+F((1-p)(1-q))=0$$

for all $p,q\in]0,1[$ *if, and only if,*

$$F(p)=4ap-a \quad \textit{for all} \quad p\in]0,1[,$$

where a *is an arbitrary constant.*

Proposition 6. *The measurable solutions* F: $]0,1[\to R$ *of*

$$F(pq)-F(p(1-q))-F((1-p)q)+F((1-p)(1-q))=0$$

for all $p,q\in]0,1[$ *are given by*

$$F(p)=ap^2-ap+b\log p+c, \quad p\in]0,1[$$

where a,b,c *are arbitrary constants.*

Remark. From now on all functions occurring in this section are from $]0,1[$ into R and all equations displayed are supposed to hold for all variables in $]0,1[$ (unless otherwise specified).

Proposition 7. *The measurable solutions of*

$$F(pq)+G(p(1-q))+G((1-p)q)+F((1-p)(1-q))=0 \tag{3.1}$$

are given by

$$\begin{cases} F(p)=bp^2+(2a-b)p-c\log p-d-a, \\ G(p)=-bp^2+(2a+b)p+c\log p+d, \end{cases} \tag{3.2}$$

where a,b,c,d *are arbitrary constants.*

Proof. Replacing p by $1-p$ in (3.1) and adding the resultant to (3.1), we obtain

$$(F+G)(pq)+(F+G)(p(1-q))+(F+G)((1-p)q)+(F+G)((1-p)(1-q))=0,$$

from which it follows by Proposition 5 that

$$F(p)+G(p)=4ap-a, \tag{3.3}$$

where a is a constant. With the use of (3.3) and (3.1) we get

$$H(pq)-H(p(1-q))-H((1-p)q)+H((1-p)(1-q))=0$$

where $H(p):=G(p)-2ap$. By Proposition 6, we obtain

$$H(p)=-bp^2+bp+c\log p+d$$

which in turn yields the asserted form of G in (3.2). This, together with (3.3) yields the sought for form of F in (3.2).

Proposition 8. *The solutions of*

$$F(pq)+G(p(1-q))-G((1-p)q)-F((1-p)(1-q))=0$$

are given by

$$F=a \quad \text{and} \quad G=b$$

where a, b are arbitrary constants.

Proof. From

$$F(pq)-F((1-p)(1-q))=G((1-p)q)-G(p(1-q))$$

and the symmetry of the left side and the antisymmetry of the right side in p and q we get

$$F(pq)-F((1-p)(1-q))=0=G((1-p)q)-G(p(1-q)).$$

Putting $u=pq$ in the equation of F we can rewrite it as

$$F(u)=F\left(1-p-\frac{u}{p}+u\right) \quad \text{for all} \quad p\in]u,1[$$

for each $u\in]0,1[$. This is equivalent to

$$F(u)=F(t) \quad \text{for all} \quad t\in]0,(1-\sqrt{u})^2[$$

for each $u\in]0,1[$. This constancy of F on the intervals $]0,(1-\sqrt{u})^2[$ for each $u\in]0,1[$ implies that F is constant on $]0,1[$. Similarly G is also constant.

Proposition 9. *The measurable solutions of*

$$f_1(pq)+f_2(p(1-q))+f_3((1-p)q)+f_4((1-p)(1-q))=0 \tag{3.4}$$

are given by

$$\begin{cases} f_i(p)=bp^2+(a-b)p-c\log p+d_i, & i=1,4,\\ f_i(p)=-bp^2+(a+b)p+c\log p+d_i, & i=2,3,\end{cases} \tag{3.5}$$

where a, b, c, d_i's are constants with $a+d_1+d_2+d_3+d_4=0$.

PROOF. Replace p by $1-p$ and q by $1-q$ in (3.4), then add and subtract the resultant to (3.4) respectively to get

$$(f_1+f_4)(pq)+(f_2+f_3)(p(1-q))+(f_2+f_3)((1-p)q)+ +(f_1+f_4)((1-p)(1-q))=0, \tag{3.6}$$

$$(f_1-f_4)(pq)+(f_2-f_3)(p(1-q))-(f_2-f_3)((1-p)q)- -(f_1-f_4)((1-p)(1-q))=0. \tag{3.7}$$

By applying Propositions 7 and 8 to (3.6) and (3.7) respectively we get

$$(f_1+f_4)(p)=2bp^2+(2a-2b)p-2c\log p+e_1,$$
$$(f_2+f_3)(p)=-2bp^2+(2a+2b)p+2c\log p+e_2$$

with $a+e_1+e_2=0$, and

$$(f_1-f_4)(p)=e_3,$$
$$(f_2-f_3)(p)=e_4.$$

This proves (3.5) with $d_1=(e_1+e_3)/2$, $d_4=(e_1-e_3)/2$, $d_2=(e_2+e_4)/2$ and $d_3==(e_2-e_4)/2$.

Theorem 10. *The measurable solutions of the functional equation* (1.1) *are given by*

$$f_i(p)=4ap^3+(b-9a)p^2+c_ip+cp\log p+e_i\log p+d_i,\quad i=1,4$$
$$f_i(p)=4ap^3-(b+9a)p^2+c_ip+cp\log p+e_i\log p+d_i,\quad i=2,3$$
$$f_5(p)=-6ap^2+(6a-2b+c_2-c_4)p+c[p\log p+(1-p)\log(1-p)]+ +(e_1+e_2)\log p+(e_3+e_4)\log(1-p)+d_5,$$
$$f_6(p)=-6ap^2+(6a-2b+c_3-c_4)p+c[p\log p+(1-p)\log(1-p)]+ +(e_1+e_3)\log p+(e_2+e_4)\log(1-p)+d_6$$

where $a, b, c, c_i, d_i, e_i,$ *are constants with*

$$4b+c_1-c_2-c_3+c_4=0 \quad and \quad b-5a+c_4+d_1+d_2+d_3+d_4=d_5+d_6.$$

PROOF. From section 2 we may assume that the f_i's have derivatives of all orders. Differentiate (1.1) twice, first with respect to p and then with respect to q, to obtain

$$F_1(pq)-F_2(p(1-q))-F_3((1-p)q)+F_4((1-p)(1-q))=0,$$

where

$$F_i(p):=pf_i''(p)+f_i'(p),\quad (i=1,2,3,4). \tag{3.8}$$

By Proposition 9 the F_i's are given by

$$F_i(p)=Bp^2+(A-B)p-C\log p+D_i,\quad i=1,4,$$
$$F_i(p)=Bp^2-(A+B)p-C\log p-D_i,\quad i=2,3, \tag{3.9}$$

where A, B, C, D_i are constants with $A+D_1+D_2+D_3+D_4=0$. Using (3.8) and (3.9) we get

$$f_i(p) = \frac{B}{9}p^3 + \frac{A-B}{4}p^2 - C(p\log p - 2p) +$$

$$+D_i p + e_i \log p + d_i, \quad i = 1, 4,$$

$$f_i(p) = \frac{B}{9}p^3 - \frac{A+B}{4}p^2 - C(p\log p - 2p) -$$

$$-D_i p + e_i \log p + d_i, \quad i = 2, 3,$$

where e_i, d_i are constants. This establishes the asserted form of f_i's for $i=1, 2, 3, 4$, which in turn, from (1.1), yields the form of f_5 and f_6. This proves the theorem.

Corollary 11. *The measurable solutions of the functional equation* (1.1) *with* $f_1=f_2=f_3=f_4=f$ *are given by*

$$f(p) = 4ap^3 - 9ap^2 + c_0 p + cp\log p + e_0 \log p + d,$$

$$f_i(p) = -6ap^2 + 6ap + c[p\log p + (1-p)\log(1-p)] +$$

$$+2e_0 \log p(1-p) + d_i, \quad i = 5, 6,$$

where a, c_0, c, e_0, d, d_i *are constants with* $-5a+c_0+4d_0=d_5+d_6$.

Corollary 12. *The measurable solutions of the functional equation*

$$f(pq)+f(p(1-q))+f((1-p)q)+f((1-p)(1-p)) =$$

$$= f(p)+f(1-p)+f(q)+f(1-q)$$

for all $p, q \in]0, 1[$ *are given by*

$$f(p) = 4ap^3 - 9ap^2 + 5ap + cp\log p + d$$

where a, c, d *are constants*. (This is Theorem 5 in [1]).

References

[1] Z. Daróczy and A. Járai, On the Measurable Solution of a Functional Equation of the Information Theory, *Acta Math. Acad. Sci. Hungaricae* **34** (1979), 105—116.

[2] I. P. Natanson, Theory of Functions of a Real Variable, Volume 1, *Frederick Ungar Publishing Co., New York*, 1964.

FACULTY OF MATHEMATICS
UNIVERSITY OF WATERLOO
WATERLOO, ONTARIO
CANADA N2L 3G1

(Received February 8, 1978; in revised form September 7, 1984.)