Boundedness on a set of positive measure and the fundamental equation of information

By G. T. DIDERRICH*) (Milwaukee, Wis.)

1. Introduction. An information function f is a real valued function defined on $[0, 1]$ which satisfies the following functional equation and boundary conditions

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

$$\text{whenever} \quad x, y \in [0, 1[\quad \text{and} \quad x+y \leqq 1$$

$$\text{with } f(0)=f(1) \quad \text{and} \quad f\left(\frac{1}{2}\right)=1.$$

These conditions imply that $f(x)=f(1-x)$, $x \in [0, 1]$ and $f(0)=0$. The above functional equation is called the fundamental equation of information because the problem of characterizing Shannon's measure of entropy can be reduced to a study of it. (For details and terminology see ACZÉL and DARÓCZY [1]). The Shannon measure of entropy in a 2-event space

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

(with the conventions $\log = \log_2$ and $0 \log 0 = 0$) is the most important information function.

In [2] we have proved the following result: If f is an information function and f is bounded on an arbitrarily small nonvanishing interval contained in $]0, 1[$, then $f=S$. The following question seems quite natural: Can one replace "bounded on an arbitrary small nonvanishing interval" by "bounded on a set of positive measure" in the above statement (measurable is meant in the Lebesgue sense with μ denoting the measure). In this connection there is a result in [2].

Corollary 1.1. *Suppose that, for every $\varepsilon>0$, there is a positive constant K and a measurable set E_K on which the information function f is bounded by K and $\mu(E_K) \geqq 1-\varepsilon$, then $f=S$.*

In this paper we give a positive answer to the above question.

We prove the following:

---

*) Research Conducted Under a National Research Council Grant at the University of Waterloo, Waterloo, Ontario, Canada in the Pure Mathematics Department.

Theorem. *Let f be an information function. If f is bounded on a set of positive measure, then $f=S$.*

2. Proof of the Theorem. The proof is carried out by imitating the procedure in [1] (Section 3.4) and introducing a weaker definition of the multiplicative group G_f. The idea is to spread out the boundedness to larger and larger portions of $]0, 1[$ where intervals are replaced by measurable subsets together with the concept of a "right-sided density point". Some *definitions* follow.

Let E be a Borel subset on the real line. A point x is said to be a two-sided density of E, if

$$\lim_{\delta\to 0+}\frac{\mu(E\cap]-\delta+x, x+\delta[)}{2\delta}=1.$$

Almost all points of E are density points (see SAKS [7] Chapter 4 and RUDIN [6] Chapter 8). In an analogous way we say x is a right-sided density point, if

$$\lim_{\delta\to 0+}\frac{\mu(E\cap[x, x+\delta[)}{\delta}=1$$

with a similar definition for the concept of left-sided density point. We now proceed to the lemmas.

Lemma 2.1. (See JOU and BAKER [3].) *Let $E_1, E_2, \ldots, E_n$ and I be Borel subsets of the real line with $0<\mu(I)<\infty$. Then*

$$1-\frac{\mu(E_1\cap\ldots\cap E_n\cap I)}{\mu(I)}\leqq\sum_{i=1}^{n}\left[1-\frac{\mu(E_i\cap I)}{\mu(I)}\right].$$

From this lemma it is clear that a two-sided density point is also a right (left)-sided density point.

Lemma 2.2. *Let E_1, E_2 be Borel subsets of the real line. Suppose that x is a two-sided (right, left) density point of E_1 and E_2. Then x is a two-sided (right, left) density point of $E_1\cap E_2$.*

PROOF. We consider the two-sided case; the other cases follow in a similar fashion. From the definition

$$\lim_{\delta\to 0+}\frac{\mu(E_i\cap I_\delta)}{2\delta}=1\quad (i=1, 2)$$

where $I_\delta=]-\delta+x, x+\delta[$. The preceeding lemma gives

$$0\leqq 1-\frac{\mu(E_1\cap E_2\cap I_\delta)}{2\delta}\leqq 1-\frac{\mu(E_1\cap I_\delta)}{2\delta}+1-\frac{\mu(E_2\cap I_\delta)}{2\delta}$$

and hence the present lemma.

Lemma 2.3. *Let T be a continuously differentiable transformation from an open interval $V=]a, b[$ to an open interval $W=]c, d[$. Assume that $|T'|\neq 0$ on V. Let E be a set of positive measure in V and suppose that $x\in E$ is a two-sided density point of E. Then Tx is a two-sided density point of TE. In particular, if $T'>0$ $(T'<0)$*

and $x \in E$ *is a right-sided density point of* E, *then* Tx *is a right (left)-sided density point of* TE.

Proof. We consider the $T'>0$ case in the last sentence of the lemma. The other cases follow in a similar way. Let x be a right-sided density point of E, so

$$\lim_{\delta\to 0+} \frac{\mu(E\cap I_\delta)}{\delta} = 1$$

where $I_\delta=[x, x+\delta[$. Clearly T is a $1-1$ open mapping of V into W, hence $T(E\cap I_\delta)=TE\cap TI_\delta$. Furthermore TI_δ is an interval in W of the form $[Tx, Tx+\delta_1[$ where $\delta_1=\mu(TI_\delta)$. We have by [5] (Theorem 8.26) that

$$\mu(TE\cap TI_\delta) = \int_{E\cap I_\delta} |T'(y)|\, d\mu(y). \tag{1}$$

It suffices to prove that

$$\lim_{\delta_1\to 0} \frac{\mu(TE\cap TI_\delta)}{\mu(TI_\delta)} = 1$$

to reach the desired conclusion. From (1) we obtain

$$\inf_{y\in E\cap I_\delta} |T'(y)|\,\mu(E\cap I_\delta) \leqq \mu(TE\cap TI_\delta) \leqq \sup_{y\in E\cap I_\delta} |T'(y)|\,\mu(E\cap I_\delta).$$

Therefore

$$\frac{l(E\cap I_\delta)}{\dfrac{\mu(TI_\delta)}{\mu(I_\delta)}}\frac{\mu(E\cap I_\delta)}{\mu(I_\delta)} \leqq \frac{\mu(TE\cap TI_\delta)}{\mu(TI_\delta)} \leqq \frac{L(E\cap I_\delta)}{\dfrac{\mu(TI_\delta)}{\mu(I_\delta)}}\frac{\mu(E\cap I_\delta)}{\mu(I_\delta)} \tag{2}$$

where $l(E\cap I_\delta)=\lim_{y\in E\cap I_\delta} |T'(y)|$ and $L(E\cap I_\delta)=\sup_{y\in E\cap I_\delta} |T'(y)|$. But it is clear, using the continuity of T' and the definitions, that the left and right extremities of (2) go to 1 as $\delta_1\to 0$. This completes the proof.

Definition. Let f be a real-valued function defined on $[0, 1]$. The positive real number λ belongs to G_f, if there exist positive numbers k_λ, δ_λ and a Borel set I_λ contained in $[0, 1]$ with $0\in I_\lambda$ as a right-sided density point such that

$$\left|f(y)-\lambda f\left(\frac{I}{\lambda}\right)\right| < k_\lambda$$

for every $y\in I_\lambda\cap[0, \delta[$ $(0<\delta\leqq\delta_\lambda)$.

Roughly speaking, λ belongs to G_f if $\left|f(y)-\lambda f\left(\frac{y}{\lambda}\right)\right|$ becomes eventually bounded near $[0, \delta[$ on a set whose measure becomes closer and closer to δ as $\delta\to 0$.

Lemma 2.4. G_f *is a multiplicative subgroup of* $]0, \infty[$, *the multiplicative group of positive reals.*

Proof.

(a) Clearly $1\in G_f$.

(b) Suppose that $\lambda\in G_f$, we want to show that $\frac{1}{\lambda}\in G_f$. The definition yields positive numbers k_1, δ_1 and a Borel set I_1 with $0\in I_1$ as a right-sided density point such that

$$\left|f(y)-\lambda f\left(\frac{y}{\lambda}\right)\right| < k_1 \tag{3}$$

for each $y\in I_1\cap[0, \delta[$ $(0<\delta\leqq\delta_1)$. Now put $I_0=\dfrac{I_1}{\lambda}$. Lemma 2.3 implies that $0\in I_0$ is a right-sided density point. Choose $0<\delta_0\leqq\dfrac{\delta_1}{\lambda}$. Hence, $x\in I_0\cap[0, \delta_0[$ implies that $x=\dfrac{y}{\lambda}$ where $y\in I_1\cap[0, \delta_1[$. Thus, dividing (3) by λ, we obtain

$$\left|f(x)-\frac{1}{\lambda}f(x\lambda)\right| = \left|f\left(\frac{y}{\lambda}\right)-\frac{1}{\lambda}f(y)\right| < \frac{k_1}{\lambda} \tag{4}$$

for each $x\in I_0\cap[0, \delta[$ $(0<\delta\leqq\delta_0)$. Hence we proved $\dfrac{1}{\lambda}\in G_f$.

(c) We now prove $\lambda_1, \lambda_2\in G_f$ imply $\lambda_1\lambda_2\in G_f$. Our hypothesis gives

$$\left|f(y)-\lambda_i f\left(\frac{y}{\lambda_i}\right)\right| < k_i \tag{5}$$

for $y\in I_i\cap[0, \delta[$ where $0<\delta\leqq\delta_i$ $(i=1, 2)$. Put $I_0=I_1\cap I_2\lambda_1$ and $\delta_i=\min\{\delta_1, \delta_2\lambda_1\}$. It follows from Lemma 2.2 and Lemma 2.3 that 0 is a right-sided density point of I_0. We claim that

$$\left|\lambda_1 f\left(\frac{x}{\lambda_1}\right)-\lambda_1\lambda_2 f\left(\frac{x}{\lambda_1\lambda_2}\right)\right| < \lambda_1 k_2 \tag{6}$$

for every $x\in I_0\cap[0, \delta[$ where $0<\delta\leqq\delta_0$. This follows by multiplying (5), in the $i=2$ case, by λ_1 and noting that $x\in I_0\subset I_2\lambda_1$ implies $\dfrac{x}{\lambda_1}\in I_2$. Also (5), in the $i=1$ case, yields

$$\left|f(x)-\lambda_1 f\left(\frac{x}{\lambda_1}\right)\right| < k_1 \tag{7}$$

for every $x\in I_0\cap[0, \delta[$ $(0<\delta\leqq\delta_0)$. Adding (6) and (7) results in

$$\left|f(x)-\lambda_1\lambda_2 f\left(\frac{x}{\lambda_1\lambda_2}\right)\right| < k_1+\lambda_1 k_2$$

for every $x\in I_0\cap[0, \delta[$ $(0<\delta\leqq\delta_0)$. This completes the proof of the lemma.

Lemma 2.5. *If an information function f is bounded on a set of positive measure E, then $G_f=]0, \infty[$.*

PROOF. Almost all points in E are two-sided density points, a fortori, right-sided density points. Fix a density point x in $E\cap]0, 1[$. Assume that $y\to 0+$ such that $x+y\leqq 1$. Then the fundamental equation gives

$$f(x)+(1-x)f\left(\frac{y}{1-x}\right)=f(y)+(1-y)f\left(\frac{x}{1-y}\right). \tag{8}$$

Now $\frac{x}{1-y}\in E$ if and only if $y\in I=(1-xE^{-1})$. An application of Lemma 2.3, with $T(z)=1-xz^{-1}$, shows that $0\in I$ is a right-sided density point of I. Equation (8) gives

$$\left|f(x)-(1-y)f\left(\frac{x}{1-y}\right)\right| = \left|f(y)-(1-x)f\left(\frac{y}{1-x}\right)\right| < 2k$$

for $y\in I\cap[0, \delta[$ and small enough $\delta>0$ where $k>0$ is the bound for f on E. The preceeding argument clearly proves that $1-x\in G_f$. Hence, $G_f\supset 1-E^*$ is the set of density points of $E\cap]0, 1[$.

The group $(G_f, \cdot)$ is isomorphic to a subgroup $(\mathbf{R}_f, +)$ of the additive group of reals $(\mathbf{R}, +)$ via the mapping $x\to\log x$. But $(\mathbf{R}_f, +)$ contains a set of positive measure, $F=\log(1-E^*)$. The linear Steinhaus theorem (see KEMPERMAN [4]. Section 2) implies that $F+F$ contains a non-degenerate interval. So $(\mathbf{R}_f, +)=(\mathbf{R}, +)$, hence $G_f=]0, \infty[$ which completes the proof of the lemma.

Lemma 2.6. *If $G_f=], \infty[$, then every point x in $]0, 1[$ is in a set E_x of positive measure with x as a right-sided density point on which f is bounded. Indeed, $E_x= =x(1-I_{1-x})^{-1}$ where I_{1-x} is the Borel set associated with $1-x$ in the definition of G_f.*

PROOF. Now $1-x\in G_f$ $(0<x<1)$. Therefore the definition of G_f yields positive numbers k_1, δ_1 and a Borel set I with $0\in I$ as a right-sided density point such that

$$\left|f(y)-(1-x)f\left(\frac{y}{1-x}\right)\right| < k_1 \tag{9}$$

for $y\in I\cap[0, \delta[$ $(0<\delta\leqq\delta_1)$. The fundamental equation (8) together with (9) then gives

$$\left|f(x)-(1-y)f\left(\frac{x}{1-y}\right)\right| < k_1 \tag{10}$$

for $y\in I\cap[0, \delta[$. But Lemma 2.3, using $T(z)=x(1-z)^{-1}$, shows that $E_x=x(1-I)^{-1}$ with (10) has the desired properties.

Lemma 2.7. *If f is bounded on a Borel set of positive measure in $]0, 1[$, then f is bounded on Borel sets of measure arbitrarily close to 1.*

PROOF. We use a modified Heine—Borel type of argument. Let $F=[a, b]$ $(a<b)$ be a given closed interval in $]0, 1[$. Let $N\geqq 2$ be a given positive integer. Let T denote the set of real numbers t such that $a\leqq t\leqq b$ and there is a Borel set E contained in $[a, t]$ with $\mu(E)\geqq(t-a)\left(1-\frac{1}{N}\right)$ on which f is bounded. Let $c=\sup T$. By our hypothesis and Lemma 2.6, it is clear that $c>a$.

We now argue that $c \in T$. Lemma 2.6 shows that c is a left-sided density point of $E'_c = 1 - E_{1-c}$ on which f is bounded. Hence for t sufficiently close to c $(t < c)$ we can find a measurable set E_1 contained in $[t, c]$ on which f is bounded and

$$\mu(E_1) \geqq (c-t)\left(1-\frac{1}{N}\right). \tag{11}$$

But by the definition of sup and by our hypothesis, we can find a measurable set E_2 contained in $[a, t]$ on which f is bounded and

$$\mu(E_2) \geqq (t-a)\left(1-\frac{1}{N}\right). \tag{12}$$

Thus (11) and (12) yield the set $E = E_1 \cup E_2 \subset [a, c]$ on which f is bounded and

$$\mu(E) \geqq (c-a)\left(1-\frac{1}{N}\right)$$

which demonstrates that $c \in T$.

Let $E_1 \subset [a, c]$ be the set associated with c, so

$$\mu(E_1) \geqq (c-a)\left(1-\frac{1}{N}\right) \tag{13}$$

and f is bounded on E_1. Again by Lemma 2.6, we can find a Borel set E_2, disjoint from E_1, with c as a right-sided density point such that $E_2 \subset [c, c_1]$ for some $c_1 > c$ with

$$\mu(E_2) \geqq (c_1-c)\left(1-\frac{1}{N}\right) \tag{14}$$

on which f is bounded. Clearly f is bounded on $E = E_1 \cup E_2 \subset [a, c_1]$ and (13) and (14) yield

$$\mu(E) = \mu(E_1)+\mu(E_2) \geqq (c-a)\left(1-\frac{1}{N}\right)+(c_1-c)\left(1-\frac{1}{N}\right) \geqq (c_1-a)\left(1-\frac{1}{N}\right).$$

But this contradicts the definition of c and, as a consequence, c cannot be an interior point, so $c = b$. Hence we proved that every closed interval $[a, b]$ contains a measurable set E such that $\mu(E) \geqq (b-a)\left(1-\frac{1}{N}\right)$ on which f is bounded. To complete the proof of the lemma, take $[a, b] = \left[\frac{1}{k}, 1-\frac{1}{k}\right]$ with k sufficiently large and N sufficiently large.

The theorem now follows from Corollary 1.1 of [2]. The reader is referred to [5] for simplified proofs of the main result here and the result in [2].

References

[1] J. Aczél and Z. Daróczy, On measures of information and their characterizations, *Academic Press, New York*, 1975.
[2] G. T. Diderrich, Local boundedness and the Shannon entropy, *Information and Control*, **36** (1978), 292—308.
[3] S. S. Jou and J. Baker, On transformations of measurable sets in locally compact topological groups, *Glasnik Mat.* **9** (29) (1974), 35—41.
[4] J. H. B. Kemperman, A general functional equation, *Trans. Amer. Math. Soc.* **86** (1957), 28—56.
[5] Gy. Maksa, Bounded symmetric information functions, *C. R. Math. Rep. Acad. Sci. Canada* — **II** (1980) *No. 5*, 247—252.
[6] W. Rudin, Real and complex analysis, *McGraw-Hill Book Company, New York* 1974.
[7] S. Saks, Theory of the integral, *Hafner Publishing Company, New York* 1937.

JOHNSON CONTROLS, INC.
P.O. BOX 423
MILWAUKEE, WIS. 53201
U.S.A.

(Received July 6, 1982)