Determination of All Semisymmetric Reçursive Information Measures of Multiplicative Type on n Positive Discrete Probability Distributions

J. Aczél and C. T. Ng
Centre for Information Theory
University of Waterloo
Waterloo, Ont. N2L 3G1, Canada

Dedicated to Alexander M. Ostrowski on his 90th birthday

Submitted by Chandler Davis

ABSTRACT

Information measures Δ_m (entropies, divergences, inaccuracies, information improvements, etc.), depending upon n probability distributions which we unite into a vector distribution, are recursive of type μ if

$$\Delta_m(p_1, p_2, p_3, \dots, p_m) = \Delta_{m-1}(p_1 + p_2, p_3, \dots, p_m) + \mu(p_1 + p_2)\Delta_2\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right).$$

If also a similar equation holds with three instead of two distinguished vectors, then μ has to be multiplicative, except if all Δ_m are identically 0. The information measure is semisymmetric if $\Delta_3(p_1, p_2, p_3) = \Delta_3(p_1, p_3, p_2)$. We determine all semisymmetric (in particular, symmetric) recursive information measures of multiplicative type, allowing first only positive probabilities. Previously the cases $n \leqslant 3$ have been examined mainly for $\mu(t) = \mu(\tau_1, \tau_2, \dots, \tau_n) = \tau_1^{\alpha_1}\tau_2^{\alpha_2}\cdots\tau_n^{\alpha_n}$, and some probabilities were allowed to be 0. This has made the proofs easier. But permitting certain probabilities to be 0 would exclude most information measures important for applications, so the description of appropriate domains became complicated. However, we show how the measures which we determine here can be extended to the "old" domains and to more general ones.

1. HISTORY, MOTIVATION, TERMINOLOGY

Entropies, deviations (directed divergences, inaccuracies, etc.) and information improvements (generalized directed divergences) depend on one, two,

52 Vanderbilt Ave., New York, NY 10017 0024-3795/83/$3.00

or three (for our purposes finite, discrete) probability distributions, respectively [6]. Some efforts have also been made recently to deal with information measures depending upon more than three probability distributions ([14], [35], etc.). We will deal here with information measures depending upon an arbitrary number n of finite probability distributions, each containing, say, m probabilities, the same m for each distribution (since they are approximations, estimates of each other, or in other ways connected). We will unite the corresponding probabilities in the n distributions into n-dimensional vectors, which we denote by Latin letters. So all lower case Latin letters, even with subscripts, will stand for vectors, except i, j, k, m, n etc. denoting natural numbers. Sets of vectors will be denoted by Latin capitals. Scalars (real numbers) will be denoted by Greek characters.

The purpose of this paper is to determine *all* such information measures satisfying certain natural conditions.

The class of information measures which were most thoroughly investigated are the symmetric recursive ones. (It seems that Faddeev [11] was the first to apply recursivity for characterization of an information measure; cf. [6]). *Recursivity* means $[p_j=(\pi_{j1},\pi_{j2},\dots,\pi_{jn})]$

$$\Delta_m(p_1,p_2,p_3,\dots,p_m)$$
$$=\Delta_{m-1}(p_1+p_2,p_3,\dots,p_m)+(\pi_{11}+\pi_{21})\Delta_2\left(\frac{p_1}{p_1+p_2},\frac{p_2}{p_1+p_2}\right)$$
$$(m=3,4,\dots). \tag{1}$$

(The operations on vectors—here additions and divisions, and later also subtractions and multiplications—are done *componentwise*.) If the probabilities belong to events, then (1) describes how the measure of information changes if an event is split into two. Recursivity is also connected to the Huffman coding procedure. It is even more natural to suppose that the Δ_m's are symmetric. For our purposes, it will be sufficient most of the time to suppose that they are *semisymmetric*, that is,

$$\Delta_3(p_1,p_2,p_3)=\Delta_3(p_1,p_3,p_2). \tag{2}$$

If also

$$\Delta_3(p_1,p_2,p_3)=\Delta_3(p_2,p_1,p_3), \tag{3}$$

then the information measure is *symmetric*. Examples of symmetric recursive information measures are the Shannon entropy

$$-\sum_{k=1}^{m} \pi_k \log \pi_k, \tag{4}$$

the Kullback directed divergence

$$\sum_{k=1}^{m} \pi_k \log \frac{\pi_k}{\sigma_k}, \tag{5}$$

and the Theil information improvement

$$\sum_{k=1}^{m} \pi_k \log \frac{\rho_k}{\sigma_k}. \tag{6}$$

We did not specify till now the *domains* of the Δ_m and of the equations (1), (2), (3). As n complete probability distributions, $(p_1, p_2, \ldots, p_m)$ should satisfy

$$p_1 + p_2 + \cdots + p_m = 1, \qquad p_j \geqslant 0 \quad (j = 1, 2, \ldots, m),$$

the additions (as before) and inequalities being meant componentwise, and $0, 1$ standing for $(0, 0, \ldots, 0)$ and $(1, 1, \ldots, 1)$, respectively. However, the examples (4) and, even more, (5) and (6) show that some caution has to be exercised with 0 probabilities, because these expressions are not defined if some π_k or ρ_k or σ_k are 0. The definition

$$0 \log 0 := 0$$

takes care of this problem for (4), and also for (5) if $\pi_k = 0$ was the problem. But how about $\sigma_k = 0$ in (5), and $\rho_k = 0$ or $\sigma_k = 0$ or both in (6) (in particular if $\pi_k \neq 0$ for the same k)? In fact, if such values were permitted in the suppositions, these important information measures would be excluded from our characterizations.

One way of getting around this difficulty is to require that, whenever σ_k or ρ_k or both are zero for a given k, then π_k should be zero too for that k [6, 13–17, 19, 20, 33, 36]. This leads to pretty complicated domains (cf. [7]). We get simpler domains (but potentially more difficult proofs) if we just exclude zero probabilities, except for the first distribution (π_k). In the case

$n=2$, this program has been carried out completely in [2, 4, 7, 8], both for (1) [and (2)] and for the measures of degree a which we will mention below. In the present paper, we completely eliminate zero probabilities (supposing even $\pi_k \neq 0$) and determine on this more difficult domain *all semisymmetric recursive information measures depending upon n probability distributions* (not just for $n \leqslant 3$), along with the generalizations to be outlined in the next paragraph. This has been done before only for $n=1$ (cf. [6, 9, 28]). We will show also how to derive from our results those on the previously handled domains.

As mentioned, we deal also with a rather wide range of generalizations: In [12] (cf. [9, 6]) a generalized recursivity has been introduced for $n=1$, first called [9] of type a and later [6] of degree a (because "type" sounded too general). Subsequently (see [6, 18, 21, 23, 32, 34] and the survey [36]) this has been investigated also for $n=2$ and $n=3$. An information measure $\{\Delta_m\}$ is *recursive of degree* $a=(\alpha_1,\alpha_2,\ldots,\alpha_n)$ if

$$\Delta_m(p_1,p_2,p_3,\ldots,p_m)=\Delta_{m-1}(p_1+p_2,p_3,\ldots,p_m)+(p_1+p_2)^a\Delta_2\left(\frac{p_1}{p_1+p_2},\frac{p_2}{p_1+p_2}\right), \tag{7}$$

where

$$x^a=(\xi_1,\xi_2,\ldots,\xi_n)^{(\alpha_1,\alpha_2,\ldots,\alpha_n)}:=\xi_1^{\alpha_1}\xi_2^{\alpha_2}\cdots\xi_n^{\alpha_n}.$$

As a further generalization, we now consider

$$\Delta_m(p_1,p_2,p_3,\ldots,p_m)=\Delta_{m-1}(p_1+p_2,p_3,\ldots,p_m)+\mu(p_1+p_2)\Delta_2\left(\frac{p_1}{p_1+p_2},\frac{p_2}{p_1+p_2}\right)\quad(m=3,4,\ldots). \tag{8}$$

We call the $\{\Delta_m\}$ satisfying this relation *recursive of type* μ, thus approaching the original name. As (7) shows, the case where μ is *multiplicative*, that is,

$$\mu(pq)=\mu(p)\mu(q), \tag{9}$$

is of particular importance. We call (8) with such μ *recursivity of multiplicative type*. Symmetric recursive entropies of multiplicative type have been examined in [22, 24, 31, 35] for $n=1,2,3$ and for the domains containing some zero probabilities. In this paper, we will first (Sections 2 and 3) determine all *semisymmetric recursive entropies of multiplicative type for all* *dimensions* n on the open domain described by

$$\sum_{j=1}^{m} p_j = 1, \qquad p_j > 0 \quad (j = 1,2,\ldots,m), \tag{10}$$

and then (Section 4) also on the older domains allowing some zero probabilities, which we have described above.

An advantage of considering (8) rather than (7) is the possibility of extension to more general fields than the reals. We don't elaborate on this, since we see no applications yet.

We can also give another explanation why it is natural to assume that μ in (8) is multiplicative (9): Let us suppose that, in addition to (8), the similar equation (cf. e.g. [6, p. 62], [35])

$$\begin{aligned}\Delta_m(p_1, p_2, p_3, p_4, \ldots, p_m) \\ &= \Delta_{m-2}(p_1 + p_2 + p_3, p_4, \ldots, p_m) \\ &\quad + \mu(p_1 + p_2 + p_3)\Delta_3\left(\frac{p_1}{p_1 + p_2 + p_3}, \frac{p_2}{p_1 + p_2 + p_3}, \frac{p_3}{p_1 + p_2 + p_3}\right)\end{aligned}$$

holds, at least for $m = 4$. This and (8) imply

$$\begin{aligned}\Delta_4[pq(1-r), pqr, p(1-q), 1-p] \\ &= \Delta_2(p, 1-p) + \mu(p)\Delta_3[q(1-r), qr, 1-q] \\ &= \Delta_2(p, 1-p) + \mu(p)\Delta_2(q, 1-q) + \mu(p)\mu(q)\Delta_2(1-r, r).\end{aligned}$$

On the other hand, two more applications of (8) give

$$\begin{aligned}\Delta_4[pq(1-r), pqr, p(1-q), 1-p] \\ &= \Delta_3[pq, p(1-q), 1-p] + \mu(pq)\Delta_2(1-r, r) \\ &= \Delta_2(p, 1-p) + \mu(p)\Delta_2(q, 1-q) + \mu(pq)\Delta_2(1-r, r).\end{aligned}$$

Thus comparison indeed yields $\mu(pq) = \mu(p)\mu(q)$, that is (9), if $\Delta_2(1-r, r) \not\equiv 0$. [If $\Delta_2(1-r, r) \equiv 0$, then, by (8), all Δ_m are 0 whether μ is multiplicative or not.]

In Section 2 the case where the multiplicative μ is *not additive* [$\mu(p+q) \not\equiv \mu(p) + \mu(q)$ $(p, q, p+q \in]0,1[^n)$] is settled with an argument analogous to one applied in linear algebra in order to establish the connection between bilinear and quadratic forms. In Section 3 we prove for the remaining case where μ is both additive and multiplicative that it is *linear* too. Therefore, and because of the multiplicativity, Equation (1) and the similar ones obtained by permuting the components are the only recursive equations left. They will then be solved completely.

A preliminary announcement of some of these results is contained in [3, 29].

2. MULTIPLICATIVE NONADDITIVE TYPES

We first derive a functional equation with a single unknown function (rather than the sequence $\{\Delta_m\}$) from (2) and (8) on (10). We introduce this function $\phi: I \to \mathbb{R}$ by the definition

$$\phi(q) := \Delta_2(1-q, q) \qquad (q \in I), \tag{11}$$

where

$$I :=]0,1[^n. \tag{12}$$

We use (8) for $m = 3$ and (2):

$$\Delta_2(p_1 + p_2, p_3) + \mu(p_1 + p_2)\Delta_2\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right)$$

$$= \Delta_3(p_1, p_2, p_3)$$

$$= \Delta_3(p_1, p_3, p_2) = \Delta_2(p_1 + p_3, p_2) + \mu(p_1 + p_3)\Delta_2\left(\frac{p_1}{p_1 + p_3}, \frac{p_3}{p_1 + p_3}\right)$$

and obtain, with (11) and with $x = p_3$, $y = p_2$,

$$\phi(x) + \mu(1-x)\phi\left(\frac{y}{1-x}\right) = \phi(y) + \mu(1-y)\phi\left(\frac{x}{1-y}\right) \tag{13}$$

for all $(x, y) \in D$, where

$$D := \{(x, y) | x, y, x + y \in I\} \tag{14}$$

[remember that $p_1 = 1 - x - y \in I$; cf. (12)]. We call functions $\phi: I \to \mathbb{R}$ which satisfy (13) on (14) *n-dimensional information functions of type* μ. As mentioned before, the *multiplicativity* (9) is also supposed for $p, q \in I$, so we are dealing with *n*-dimensional information functions *of multiplicative type*. The $\{\Delta_n\}$ can be determined from ϕ with the aid of (11) and (8).

In order to solve (13) in the case where μ is multiplicative but not additive, we prove the following.

LEMMA 1. *If ϕ satisfies (13) on (14) and μ satisfies (9) for all $p, q \in I$, then the function $\Psi : D \to \mathbb{R}$, defined by*

$$\Psi(x, y) = \phi(x) + \mu(1-x)\phi\left(\frac{y}{1-x}\right) - \phi(x+y) \qquad [(x, y) \in D], \quad (15)$$

satisfies

$$\Psi(x, y) = \Psi(y, x), \quad (16)$$

$$\Psi(x, y) + \Psi(x+y, z) = \Psi(x, y+z) + \Psi(y, z), \quad (17)$$

and

$$\Psi(tx, ty) = \mu(t)\Psi(x, y) \quad (18)$$

whenever $x, y, z, x+y+z, t \in I$.

Proof. In view of (15), Equation (16) is the same as (13).

We prove (17) and (18) by using (15) and the multiplicativity (9) repeatedly:

$$\begin{aligned}
&\Psi(x, y) + \Psi(x+y, z) \\
&\overset{(15)}{=} \phi(x) + \mu(1-x)\phi\left(\frac{y}{1-x}\right) - \phi(x+y) + \phi(x+y) \\
&\quad + \mu(1-x-y)\phi\left(\frac{z}{1-x-y}\right) - \phi(x+y+z) \\
&\overset{(9)}{=} \phi(x) + \mu(1-x)\left[\phi\left(\frac{y}{1-x}\right) + \mu\left(1 - \frac{y}{1-x}\right)\phi\left(\frac{z/(1-x)}{1-y/(1-x)}\right)\right] \\
&\quad - \phi(x+y+z) \\
&\overset{(15)}{=} \phi(x) + \mu(1-x)\left[\Psi\left(\frac{y}{1-x}, \frac{z}{1-x}\right) + \phi\left(\frac{y+z}{1-x}\right)\right] - \phi(x+y+z) \\
&\overset{(15)}{=} \Psi(x, y+z) + \mu(1-x)\Psi\left(\frac{y}{1-x}, \frac{z}{1-x}\right).
\end{aligned} \quad (19)$$

By the symmetry (16) of Ψ, the left-hand side is symmetric in x and y while the right-hand side is symmetric in y and z. Hence both sides are symmetric in x, y, z. This symmetry yields (17), while comparison of the right-hand sides of (17) and (19) gives

$$\Psi(y,z)=\mu(1-x)\Psi\left(\frac{y}{1-x},\frac{z}{1-x}\right),$$

which is (18) after renaming the variables. ■

It is easy to see that μ, being multiplicative (9), is identically zero on I [cf. (12)], if it is zero for one $q_0 \in I$. Indeed, then

$$\mu(pq_0)=0 \qquad \text{for all} \quad p\in I; \tag{20}$$

in particular,

$$\mu(x)=0 \qquad \text{if} \quad x \text{ is close enough to } 0. \tag{21}$$

For each $q \in I$, we will have q^k as close to 0 as we want to by choosing a large k. Then, by (21) and (9), $0=\mu(q^k)=\mu(q)^k$, so $\mu(q)=0$ on I. Since $\mu \equiv 0$ is both multiplicative and additive, we relegate this case to Section 3 and suppose here that μ is *nowhere zero* on I.

Now we can extend first μ to $P:=]0,\infty[^n$ and then Ψ to P^2 with their properties (9), (16), (17), and (18) intact. All $t \in P$ can be written as $t=p/q$ with $p, q \in I$. We define

$$\bar{\mu}(t)=\bar{\mu}\left(\frac{p}{q}\right):=\frac{\mu(p)}{\mu(q)} \qquad (p,q\in I) \tag{22}$$

[since $\mu(q)\neq 0$]. Because of (9), $\bar{\mu}$ is unambiguously defined and is an extension of μ satisfying (9) (cf. [5]):

$$\bar{\mu}(st)=\bar{\mu}(s)\bar{\mu}(t) \qquad (s,t\in P). \tag{23}$$

Again, $\bar{\mu}$ is nowhere 0 on P.

The extension of Ψ to P^2 is done by the definition

$$\overline{\Psi}(x,y)=\frac{1}{\mu(s)}\Psi(sx,sy), \qquad \text{where} \quad (sx,sy)\in D, \quad s\in I \tag{24}$$

(possible, since μ is nowhere 0 on I). Because of (18), $\overline{\Psi}$ is unambiguously defined and is an extension of Ψ. Indeed, if also $t \in I$ and $(tx, ty) \in D$, so much the more $(stx, sty) \in D$. By (18)

$$\mu(s)\Psi(tx, ty) = \Psi(stx, sty) = \mu(t)\Psi(sx, sy),$$

so we have

$$\frac{1}{\mu(s)}\Psi(sx, sy) = \frac{1}{\mu(t)}\Psi(tx, ty)$$

and (24) is unambiguous. If $(x, y) \in D$, then, by (18),

$$\frac{1}{\mu(s)}\Psi(sx, sy) = \frac{\mu(s)}{\mu(s)}\Psi(x, y) = \Psi(x, y),$$

so $\overline{\Psi}$, as defined in (24), is an extension of Ψ. We also see that $\overline{\Psi}$ and $\bar{\mu}$ satisfy

$$\overline{\Psi}(x, y) = \overline{\Psi}(y, x), \tag{25}$$

$$\overline{\Psi}(x, y) + \overline{\Psi}(x + y, z) = \overline{\Psi}(x, y + z) + \overline{\Psi}(y, z), \tag{26}$$

and

$$\overline{\Psi}(tx, ty) = \bar{\mu}(t)\overline{\Psi}(x, y) \tag{27}$$

for all $x, y, z, t \in P$. We prove for instance (27). Write $t = p/q$, where $p, q \in I$ and $(px, py) \in D$. Then, by (24),

$$\overline{\Psi}(x, y) = \frac{1}{\mu(p)}\Psi(px, py)$$

and

$$\overline{\Psi}(tx, ty) = \overline{\Psi}\left(\frac{px}{q}, \frac{py}{q}\right) = \frac{1}{\mu(q)}\Psi(px, py).$$

Because of (22) we have indeed (27).

We now proceed to determine $\overline{\Psi}$. First we note that, for $n = 1$ and $\bar{\mu}(\xi) = \xi^2$, the equations (25), (26), and (27) are satisfied by all symmetric bilinear forms. Just as symmetric bilinear forms can be represented by their diagonals, the quadratic forms, we may expect that $\overline{\Psi}$ can be represented through its diagonal and, ultimately, by $\bar{\mu}$. We follow up this idea by the following computation, first using (25) and (26) in order to get

$$\begin{aligned}
&\overline{\Psi}(px, py)+\overline{\Psi}(px+py, qx+qy)\\
&\quad=\overline{\Psi}(px, qx+(p+q)y)+\overline{\Psi}(py, qx+qy)\\
&\quad=[\overline{\Psi}(px, qx+(p+q)y)+\overline{\Psi}(qx,(p+q)y)]-\overline{\Psi}(qx,(p+q)y)\\
&\qquad+[\overline{\Psi}(qx, qy)+\overline{\Psi}(qx+qy, py)]-\overline{\Psi}(qx, qy)\\
&\quad=[\overline{\Psi}(px, qx)+\overline{\Psi}((p+q)x,(p+q)y)]-\overline{\Psi}(qx,(p+q)y)\\
&\qquad+[\overline{\Psi}(qx,(p+q)y)+\overline{\Psi}(qy, py)]-\overline{\Psi}(qx, qy)\\
&\quad=\overline{\Psi}(px, qx)+\overline{\Psi}((p+q)x,(p+q)y)+\overline{\Psi}(qy, py)-\overline{\Psi}(qx, qy),
\end{aligned}$$

or

$$\begin{aligned}
&\overline{\Psi}(px, py)+\overline{\Psi}(qx, qy)-\overline{\Psi}((p+q)x,(p+q)y)\\
&\qquad\qquad=\overline{\Psi}(px, qx)+\overline{\Psi}(py, qy)-\overline{\Psi}(p(x+y), q(x+y)).
\end{aligned}$$

With (27) we get

$$[\bar{\mu}(p)+\bar{\mu}(q)-\bar{\mu}(p+q)]\overline{\Psi}(x, y)=[\bar{\mu}(x)+\bar{\mu}(y)-\bar{\mu}(x+y)]\overline{\Psi}(p, q) \tag{28}$$

for all $x, y, p, q \in P$.

At this point we make use of the assumption that μ, and thus $\bar{\mu}$, is not additive, so that there exist $p_0, q_0 \in P$ for which $\bar{\mu}(p_0)+\bar{\mu}(q_0)-\bar{\mu}(p_0+q_0) \neq 0$. Put into (28) $p=p_0$ and $q=q_0$ in order to obtain, with $\alpha = \overline{\Psi}(p_0, q_0)/[\bar{\mu}(p_0)+\bar{\mu}(q_0)-\bar{\mu}(p_0+q_0)]$,

$$\overline{\Psi}(x, y)=\alpha\bar{\mu}(x)+\alpha\bar{\mu}(y)-\alpha\bar{\mu}(x+y) \qquad \text{for all} \quad x, y \in P. \tag{29}$$

In particular, Ψ is of the form (29) on D, that is,

$$\Psi(x, y)=\alpha\mu(x)+\alpha\mu(y)-\alpha\mu(x+y) \qquad \text{on } D. \tag{30}$$

Evidently, functions of the form (30) are indeed solutions to (16), (17), and

(18). With this general solution for Ψ, Equation (15) is reduced to

$$\psi(x)+\mu(1-x)\psi\left(\frac{y}{1-x}\right)-\psi(x+y)=0 \qquad \text{on } D, \tag{31}$$

where

$$\psi(x):=\phi(x)-\alpha\mu(x). \tag{32}$$

Putting $y=(1-x)z$ into (31), we get

$$\psi(x)+\mu(1-x)\psi(z)=\psi(x+z-xz) \tag{33}$$

for all $x, z\in I$. We distinguish two cases.

If $\mu(t)\not\equiv 1$, we use the symmetry of the right-hand side of (33) to get

$$\psi(x)+\mu(1-x)\psi(z)=\psi(z)+\mu(1-z)\psi(x).$$

By fixing $z=z_0$ with $\mu(1-z_0)\neq 1$ we get

$$\psi(x)=\beta\mu(1-x)-\beta,$$

where $\beta=\psi(z_0)/[\mu(1-z_0)-1]$ is a constant. This and (32) give

$$\phi(x)=\alpha\mu(x)+\beta\mu(1-x)-\beta \qquad (x\in I), \tag{34}$$

and then (8) and (11) yield

$$\Delta_m(p_1,p_2,\dots,p_m)=\beta\mu(p_1)+\alpha\sum_{j=2}^{m}\mu(p_j)-\beta \tag{35}$$

($\sum_{j=1}^{m}p_j=1$; $p_j>0$, $j=1,2,\dots,m$; $m=2,3,\dots$) with constant α,β.

If, on the other hand, $\mu(t)\equiv 1$ on I, then (33) is reduced to

$$\psi(x)+\psi(z)=\psi(x+z-xz),$$

or, with $\lambda(x):=\psi(1-x)$,

$$\lambda(xy)=\lambda(x)+\lambda(y) \qquad (x,y\in I). \tag{36}$$

Thus, from (32),

$$\phi(x)=\alpha+\lambda(1-x) \qquad (x\in I), \tag{37}$$

and, by (8) and (11),

$$\Delta_m(p_1,p_2,\dots,p_m)=(m-1)\alpha+\lambda(p_1)$$
$$\left(\sum_{j=1}^{m}p_j=1; \qquad p_j>0, \quad j=1,2,\dots,m; \quad m=2,3,\dots\right). \tag{38}$$

It is easy to verify that (34) always satisfies (13) and that (35) satisfies (2) and (8), while (37) satisfies (13), and (38) satisfies (2) and (8) *if* λ satisfies (36) and $\mu(t)\equiv 1$. We have proved the following.

THEOREM 1. *Suppose that μ is multiplicative* (9) *on I^2 but not additive. Then the general solution of* (13) *on* (14) *is given by* (34) *for $\mu(t)\not\equiv 1$ and by* (37) *with* (36) *for $\mu(t)\equiv 1$. Further, the general semisymmetric recursive entropies of type μ are given by* (35) *for $\mu(t)\not\equiv 1$ and by* (38) *for $\mu(t)\equiv 1$, where λ is an arbitrary solution of* (36) *and α, β are real constants. If also* (3) *is supposed, then the Δ_m's are fully symmetric and $\beta=\alpha$ in* (35), *$\lambda=0$ in* (38).

We notice that no regularity conditions were needed. If ϕ [or Δ_2; cf. (11)] is even weakly regular, say bounded on a set of positive measure, then, in (37) and (38),

$$\lambda(x)=\sum_{k=1}^{n}\gamma_k\log\xi_k \qquad [\xi_k\in]0,1[,\ k=1,2,\dots,n; \quad x=(\xi_1,\xi_2,\dots,\xi_n)],$$

$\gamma_1,\gamma_2,\dots,\gamma_n$ being arbitrary constants (cf. [30]).

3. MULTIPLICATIVE ADDITIVE TYPES

We now go over to the remaining case where $\mu\colon I\to\mathbb{R}$ is both multiplicative

$$\mu(pq)=\mu(p)\mu(q) \qquad (p,q\in I) \tag{9}$$

and additive

$$\mu(s+t)=\mu(s)+\mu(t) \qquad [(s,t)\in D]. \tag{39}$$

We introduce the function $\nu:]0,1[\to \mathbb{R}$ by

$$\nu(\tau):=\mu(\tau,\tau,\ldots,\tau) \qquad (\tau\in]0,1[).$$

It immediately follows from (9) and (39) that ν is both additive and multiplicative. So we have either

$$\nu(\tau)\equiv\tau \tag{40}$$

or

$$\nu(\tau)\equiv 0 \tag{41}$$

on $]0,1[$. (This is well known for real-valued functions additive and multiplicative for all reals, but can be just as easily proved on our restricted domains: $\nu(\tau^2)=\nu(\tau)^2\geqslant 0$, so ν is bounded from below, and therefore all such solutions of $\nu(\sigma+\tau)=\nu(\sigma)+\nu(\tau)$, even on the open triangle $\{(\sigma,\tau)|\sigma,\tau,\sigma+\tau\varepsilon]0,1[\}$, are of the form $\nu(\tau)=\gamma\tau$, which satisfies $\nu(\sigma\tau)=\nu(\sigma)\nu(\tau)$ if, and only if, γ is either 0 or 1.)

In both cases

$$\begin{aligned}\mu(\tau x)&=\mu(\tau\xi_1,\tau\xi_2,\ldots,\tau\xi_n)=\mu(\tau,\tau,\ldots,\tau)\mu(\xi_1,\xi_2,\ldots,\xi_n)\\&=\nu(\tau)\mu(\xi_1,\xi_2,\ldots,\xi_n).\end{aligned} \tag{42}$$

In the case (41) this gives with $y=\tau x$, that μ is identically 0:

$$\mu(y)=\mu(\tau x)=0. \tag{43}$$

In the case (40), on the other hand, (42) establishes

$$\mu(\tau x)=\tau\mu(x)$$

so, together with (39), we have that μ is linear (more exactly, μ can be extended to a linear map $\mathbb{R}^n\to\mathbb{R}$). So

$$\mu(x)=\mu(\xi_1,\xi_2,\ldots,\xi_n)=\sum_{k=1}^{n}\alpha_k\xi_k.$$

But this satisfies (9), that is,

$$\sum_{k=1}^{n} \alpha_k \xi_k \eta_k = \sum_{i=1}^{n} \alpha_i \xi_i \sum_{k=1}^{n} \alpha_k \eta_k,$$

if, and only if,

$$\alpha_i \alpha_k = 0 \quad \text{for } i \neq k \quad \text{and} \quad \alpha_k^2 = \alpha_k.$$

So $\alpha_k = 0$ or $\alpha_k = 1$, but if an $\alpha_k = 1$, then all other $\alpha_i = 0$ $(i \neq k)$. Thus, in the case (40), μ is a *projection*:

$$\mu(x) = \mu(\xi_1, \xi_2, \dots, \xi_n) = \xi_k \qquad \text{for some fixed} \quad k = \{1, 2, \dots, n\}. \quad (44)$$

We have proved the following.

PROPOSITION. *A function* μ: $I \to \mathbf{R}$ *is additive and multiplicative, i.e. satisfies* (9) *and* (39) [*cf.* (12), (14)], *if, and only if*, μ *is either identically zero or a projection* (44) *of* I *to* $]0,1[$.

So the only cases not settled yet are, in addition to

$$\Delta_m(p_1, p_2, p_3, \dots, p_m) = \Delta_{m-1}(p_1 + p_2, p_3, \dots, p_m),$$

$$\left(\sum_{j=1}^{m} p_j = 1; \qquad p_j > 0, \quad j = 1, 2, \dots, m; \qquad m = 3, 4, \dots \right), \quad (45)$$

the equation (1) and the similar ones with $\pi_{1k} + \pi_{2k}$ $(k = 2, \dots, n)$ in place of $\pi_{11} + \pi_{21}$. For the latter, we may restrict ourselves to (1) [and (2)] without restricting generality.

The argument at the beginning of Section 2 translates these into

$$\phi(x) = \phi(y) \quad (46)$$

or

$$\phi(x) + (1 - \xi_1)\phi\left(\frac{y}{1-x}\right) = \phi(y) + (1 - \eta_1)\phi\left(\frac{x}{1-y}\right), \quad (47)$$

respectively, on D for ϕ: $I \to \mathbf{R}$, as defined by (11), (12), and (14).

Equation (46) simply means that ϕ is constant

$$\phi(x)=\delta \qquad (x\in I), \tag{48}$$

and, in view of (11) and (45),

$$\Delta_m(p_1,p_2,\ldots,p_m)=\delta \quad \text{(constant)}$$
$$\left(\sum_{j=1}^{m} p_j=1; \qquad p_j>0, \quad j=1,2,\ldots,m; \qquad m=2,3,\ldots\right). \tag{49}$$

As to (47), for convenience we change our notation slightly in this section. We write ξ,η for ξ_1,η_1 and $u=(\xi_2,\ldots,\xi_n)$, $v=(\eta_2,\ldots,\eta_n)$ (so, in most of this section, Latin characters stand for $(n-1)$-component vectors).

We write also

$$\phi(\xi,u)=\phi(\xi,\xi_2,\ldots,\xi_n)=\phi(x) \qquad (x\in I), \tag{50}$$

so that (47) goes over into

$$\phi(\xi,u)+(1-\xi)\phi\left(\frac{\eta}{1-\xi},\frac{v}{1-u}\right)=\phi(\eta,v)+(1-\eta)\phi\left(\frac{\xi}{1-\eta},\frac{u}{1-v}\right) \tag{51}$$

for

$$\xi,\eta,\xi+\eta\in\,]0,1[, \tag{52}$$

$$u,v,u+v\in\,]0,1[^{n-1}. \tag{53}$$

We fix u and v temporarily so that, with the notation

$$\begin{aligned}\phi_1(\tau)&=\phi(\tau,u), &\qquad \phi_3(\tau)&=\phi(\tau,v),\\ \phi_2(\tau)&=\phi\left(\tau,\frac{v}{1-u}\right), &\qquad \phi_4(\tau)&=\phi\left(\tau,\frac{u}{1-v}\right),\end{aligned} \tag{54}$$

Equation (51) becomes

$$\phi_1(\xi)+(1-\xi)\phi_2\left(\frac{\eta}{1-\xi}\right)=\phi_3(\eta)+(1-\eta)\phi_4\left(\frac{\xi}{1-\eta}\right)$$

for (52). It has been proved in [27] that this implies, among others,

$$\phi_1(\tau)=\tau\tilde{\lambda}(\tau)+(1-\tau)\tilde{\lambda}(1-\tau)+\alpha_1\tau+\beta_1,$$

$$\phi_3(\tau)=\tau\tilde{\lambda}(\tau)+(1-\tau)\tilde{\lambda}(1-\tau)+\alpha_3\tau+\beta_3,$$

where $\tilde{\lambda}:]0,1[\to\mathbb{R}$ is a solution of

$$\tilde{\lambda}(\sigma\tau)=\tilde{\lambda}(\sigma)+\tilde{\lambda}(\tau)\qquad(\sigma,\tau\in]0,1[). \tag{55}$$

Comparing this with (54) and letting u, v vary again, we get

$$\phi(\tau,u)=\tau\tilde{\lambda}(\tau)+(1-\tau)\tilde{\lambda}(1-\tau)+\alpha(u)\tau+\beta(u)$$
$$\left(\tau\in]0,1[,\quad u\in]0,1[^{n-1}\right). \tag{56}$$

If we put (56) into (51), we get, in view of (55),

$$\alpha(u)\xi+\beta(u)+\alpha\left(\frac{v}{1-u}\right)\eta+\beta\left(\frac{v}{1-u}\right)(1-\xi)$$
$$=\alpha(v)\eta+\beta(v)+\alpha\left(\frac{u}{1-v}\right)\xi+\beta\left(\frac{u}{1-v}\right)(1-\eta).$$

Comparison of the coefficients of ξ and of the terms independent of ξ and η yields

$$\alpha(u)=\alpha\left(\frac{u}{1-v}\right)+\beta\left(\frac{v}{1-u}\right), \tag{57}$$

$$\beta(u)+\beta\left(\frac{v}{1-u}\right)=\beta(v)+\beta\left(\frac{u}{1-v}\right) \tag{58}$$

for (53). We solve these equations here by a method similar to [8]. Another proof can be found in [3]. Equation (58) is just (13) for β with $\mu(t)\equiv 1$ and $n-1$ instead of n. By Theorem 1, (37),

$$\beta(u)=\bar{\lambda}(1-u)+\varepsilon\qquad\left(u\in]0,1[^{n-1}\right), \tag{59}$$

where ε is constant and $\bar{\lambda}:]0,1[^{n-1}\to\mathbb{R}$ satisfies

$$\bar{\lambda}(tu)=\bar{\lambda}(t)+\bar{\lambda}(u)\qquad\left(t,u\in]0,1[^{n-1}\right); \tag{60}$$

consequently also

$$\bar{\lambda}\left(\frac{u}{t}\right)=\bar{\lambda}(u)-\bar{\lambda}(t)\qquad\left(t,u,\frac{u}{t}\in\left]0,1\right[^{n-1}\right). \tag{61}$$

We take this into account when substituting (59) into (57) and get

$$\begin{aligned}\alpha(u)&=\alpha\left(\frac{u}{1-v}\right)+\bar{\lambda}(1-u-v)-\bar{\lambda}(1-u)+\varepsilon\\ &=\alpha\left(\frac{u}{1-v}\right)+\bar{\lambda}(1-u-v)-\bar{\lambda}(1-v)+\bar{\lambda}(1-v)\\ &\quad-\bar{\lambda}(u)+\bar{\lambda}(u)-\bar{\lambda}(1-u)+\varepsilon\end{aligned}$$

for (53). Introducing $t=u/(1-v)$ and using (61) again, this goes over into

$$\alpha(u)+\bar{\lambda}(1-u)-\bar{\lambda}(u)=\alpha(t)+\bar{\lambda}(1-t)-\bar{\lambda}(t)+\varepsilon,$$

which shows that both sides have to be constant (say γ) and that $\gamma=\gamma+\varepsilon$, so $\varepsilon=0$. Thus we have, for all $u\in\left]0,1\right[^{n-1}$,

$$\alpha(u)=\lambda^{-}(u)-\lambda^{-}(1-u)+\gamma \tag{62}$$

and [cf. (59)]

$$\beta(u)=\bar{\lambda}(1-u)\qquad\left(u\in\left]0,1\right[^{n-1}\right). \tag{63}$$

Since (62) and (63) with (60) always satisfy (57) and (58), we have proved the following.

LEMMA 2. *The general solution of the system* (57), (58) *for* $u,v,u+v\in\left]0,1\right[^{n-1}$ *is given by* (62) *and* (63). *Here* γ *is an arbitrary real constant and* $\bar{\lambda}:\left]0,1\right[^{n-1}\to\mathbb{R}$ *is an arbitrary solution of* (60).

In view of (56), (62), and (63), we have

$$\phi(\xi,u)=\xi\left[\tilde{\lambda}(\xi)+\bar{\lambda}(u)\right]+(1-\xi)\left[\tilde{\lambda}(1-\xi)+\bar{\lambda}(1-u)\right]+\gamma\xi. \tag{64}$$

If we write

$$\lambda(x)=\lambda(\xi,u):=\tilde{\lambda}(\xi)+\bar{\lambda}(u),$$

this $\lambda :]0,1[^n \to \mathbb{R}$ satisfies (36) and we can write (64) as

$$\phi(x) = \xi\lambda(x) + (1-\xi)\lambda(1-x) + \gamma\xi. \tag{65}$$

It is easy to check that (65), with an arbitrary constant γ and an arbitrary solution λ of (36), always satisfies (47) on D.

Remembering that in our case and in the other cases of nontrivial $[\mu(t) \not\equiv 0]$ additive and multiplicative μ, we have (44), so (65) and the similar formulas, with other components distinguished instead of the first, can be written as

$$\phi(x) = \mu(x)[\lambda(x) + \gamma] + \mu(1-x)\lambda(1-x) \qquad (x \in I). \tag{66}$$

Making use of (8), (9), (39), and (11), we get

$$\Delta_m(p_1, p_2, \ldots, p_m) = \mu(p_1)\lambda(p_1) + \sum_{j=2}^{m} \mu(p_j)[\lambda(p_j) + \gamma]$$

$$\left(\sum_{j=1}^{m} p_j = 1; \qquad p_j > 0; \quad j = 1, 2, \ldots, m; \qquad m = 2, 3, \ldots \right). \tag{67}$$

Of course, (66) satisfies (13), and (67) satisfies (2) and (8) always when μ is both additive and multiplicative. We have proved the following.

Theorem 2. *If μ is both additive* (39) *and multiplicative* (9) *on I^2, then the general solution of* (13) *on* (14) *is given by* (48) *if $\mu(t) \equiv 0$ and by* (66) *if $\mu(t) \not\equiv 0$, and the general semisymmetric recursive entropies of type μ are in this case given by* (49) *if $\mu(t) \equiv 0$ and by* (67) *if $\mu(t) \not\equiv 0$, where γ, δ are arbitrary real constants and λ is an arbitrary solution of* (36). *If also* (3) *is supposed (symmetry), then $\gamma = 0$ in* (66) *and* (67).

If, as after Theorem 1, we wish to determine the weakly regular, say *measurable* solutions *or* those *bounded on a set of positive measure*, we can (cf. [6,10,26]), change λ so that it has the same regularity properties. So, under these conditions, we have in (66) and (67) [6,10,26,30]

$$\lambda(x) = \sum_{k=1}^{n} \gamma_k \log \xi_k \qquad [\xi_k \in]0,1[,\ k = 1, 2, \ldots, n; \quad x = (\xi_1, \xi_2, \ldots, \xi_n)], \tag{68}$$

again with arbitrary real constants $\gamma_1, \gamma_2, \ldots, \gamma_n$.

4. EXTENSIONS TO BOUNDARY POINTS

We are now showing how our solutions can be extended to the "old" domains mentioned in Section 1. Since that has been the case most completely researched before, we start with (44), in particular (1) and (47). We will deal simultaneously with the cases where just the first [or, in the general case (44), the kth] component of the vectors is 0 (or 1), or several, say n_1, components, including the first [the kth] are 0 (or 1). Therefore, we divide in this section every (n-component) vector into an n_1-component part with subscript 1 and an $(n - n_1)$-component vector with subscript 2. So we write (47) [cf. (51)] as

$$\phi(x_1, x_2) + (1-\xi)\phi\left(\frac{y_1}{1-x_1}, \frac{y_2}{1-x_2}\right) = \phi(y_1, y_2) + (1-\eta)\phi\left(\frac{x_1}{1-y_1}, \frac{x_2}{1-y_2}\right). \tag{69}$$

Considering boundary probabilities in (1) (in the first components or in several, including the first) means permitting, in addition to D [as defined in (14)], also the points with

$$x_1 = y_1 = 0 = (0,\ldots,0) \tag{70}$$

or

$$x_1 = 0, \quad y_1 \in]0,1[^{n_1} \quad \text{or} \quad y_1 = 0, \quad x_1 \in]0,1[^{n_1} \tag{71}$$

or

$$x_1 + y_1 = 1 = (1,\ldots,1), \qquad x_1, y_1 \in]0,1[^{n_1} \tag{72}$$

in the domain of (69), while we keep

$$x_2, y_2, x_2 + y_2 \in]0,1[^{n-n_1}. \tag{73}$$

We first note that, for $x_1 \in]0,1[^{n_1}$, $x_2, y_2 \in]0,1[^{n-n_1}$, $x_2 \leqslant y_2$,

$$\lambda(x_1, x_2) - \lambda(x_1, y_2)$$

depends only upon x_2/y_2. Indeed, by (36),

$$\lambda(x_1u_1, x_2u_2) = \lambda(x_1, x_2) + \lambda(u_1, u_2) \tag{74}$$

for $x_1, u_1 \in]0,1[^{n_1}$, $x_2, u_2 \in]0,1[^{n-n_1}$. So, if also $s_2, t_2 \in]0,1[^{n-n_1}$, $s_2 \leqslant t_2$, and

$$\frac{x_2}{y_2} = \frac{s_2}{t_2}, \quad \text{that is,} \quad t_2x_2 = s_2y_2,$$

and $s_1 \in]0,1[^{n_1}$, then

$$\lambda(s_1, s_2) + \lambda(x_1, y_2) = \lambda(s_1x_1, s_2y_2) = \lambda(s_1x_1, t_2x_2) = \lambda(s_1, t_2) + \lambda(x_1, x_2),$$

that is,

$$\lambda(s_1, s_2) - \lambda(s_1, t_2) = \lambda(x_1, x_2) - \lambda(x_1, y_2),$$

as asserted. Thus we can extend the definition of $\lambda(u_1, u_2)$ to $u_1 = 1$ by

$$\lambda\left(1, \frac{x_2}{y_2}\right) := \lambda(x_1, x_2) - \lambda(x_1, y_2)$$
$$\left(x_1 \in]0,1[^{n_1} \text{ arbitrary}, \qquad x_2, y_2, \frac{x_2}{y_2} \in]0,1[^{n-n_1}\right), \tag{75}$$

and the extended λ will still satisfy (74) and, as a consequence, also

$$\lambda\left(\frac{t_1}{u_1}, \frac{t_2}{u_2}\right) = \lambda(t_1, t_2) - \lambda(u_1, u_2) \tag{76}$$

[compare (75)]. The extension (75) of λ is similar to the extension (22) of μ and can also be carried out to $]0,\infty[^n$ if desired.

Now we substitute into (69) first $y_1 = 0$ (including $\eta = 0$) and $x_1 \in]0,1[^{n_1}$ [cf. (71)], and get, by (65),

$$\xi\lambda(x_1, x_2) + (1-\xi)\lambda(1-x_1, 1-x_2) + \gamma\xi + (1-\xi)\phi\left(0, \frac{y_2}{1-x_2}\right)$$
$$= \phi(0, y_2) + \xi\lambda\left(x_1, \frac{x_2}{1-y_2}\right) + (1-\xi)\lambda\left(1-x_1, \frac{1-x_2-y_2}{1-y_2}\right) + \gamma\xi$$

or, in view of (75),

$$\xi\lambda(1,1-y_2)+(1-\xi)\phi\left(0,\frac{y_2}{1-x_2}\right)$$
$$=\phi(0,y_2)+(1-\xi)\lambda\left(1,\frac{1-x_2-y_2}{(1-x_2)(1-y_2)}\right),$$

and, with (75) and $t_2=y_2/(1-x_2)$,

$$\phi(0,y_2)-\lambda(1,1-y_2)=(1-\xi)\phi(0,t_2)-(1-\xi)\lambda(1,1-t_2). \tag{77}$$

But this implies

$$\phi(0,y_2)-\lambda(1,1-y_2)=\beta \quad \text{(constant)},$$

and, substituting $\phi(0,y_2)=\lambda(1,1-y_2)+\beta$ into (77), we see that $\beta=0$ and

$$\phi(0,y_2)=\lambda(1,1-y_2) \qquad \left(y_2\in\,]0,1[^{n-n_1}\right). \tag{78}$$

Now, that is exactly what we expect as the extension of (65) to $x_1=0$ (consequently $\xi=0$), if we agree upon

$$0\cdot\lambda(0,x_2):=0, \tag{79}$$

which is customary (like $0\cdot\log 0:=0$). The function given by (78) indeed satisfies (69) for (71) and also for (70), in view of (74).

We substitute now (72) into (69) and get, by use of (65),

$$\xi\lambda(x_1,x_2)+(1-\xi)\lambda(1-x_1,1-x_2)+\gamma\xi+(1-\xi)\phi\left(1,\frac{y_2}{1-x_2}\right)$$
$$=(1-\xi)\lambda(1-x_1,y_2)+\xi\lambda(x_1,1-y_2)+\gamma(1-\xi)+\xi\phi\left(1,\frac{x_2}{1-y_2}\right).$$

We transform this by using (75) again:

$$(1-\xi)\left[\phi\left(1,\frac{y_2}{1-x_2}\right)-\lambda\left(1,\frac{y_2}{1-x_2}\right)-\gamma\right]$$
$$=\xi\left[\phi\left(1,\frac{x_2}{1-y_2}\right)-\lambda\left(1,\frac{x_2}{1-y_2}\right)-\gamma\right] \tag{80}$$

or, with $s_2 = y_2/(1-x_2)$, $t_2 = x_2/(1-y_2)$,

$$(1-\xi)[\phi(1,s_2)-\lambda(1,s_2)-\gamma] = \xi[\phi(1,t_2)-\lambda(1,t_2)-\gamma].$$

This implies that $\phi(1,t_2)-\lambda(1,t_2)-\gamma = 0$, i.e.

$$\phi(1,t_2) = \lambda(1,t_2)+\gamma \qquad (t_2 \in]0,1[^{n-n_1}).$$

This is again what we expect as the extension of (65) to $x_1 = 1$, with the convention (79).

Going over to the remaining trivial case $\mu(t) \equiv 0$ [cf. (43)] of additive and multiplicative μ, it is clear that, by extending (45) or (46) to some boundary points of their domains, the validity of (49) and (48), respectively, will also be extended to the values 0 (or 1) of the respective components. Thus, we have proved the following (cf. [8] for a special case).

THEOREM 3. *If, in addition to D [see (14)], also the points satisfying (70) or (71) or (72), with (73), are included into the domain of (13) and μ is both additive and multiplicative, then the general solutions continue to be presented by $\phi(x) = \delta$ if $\mu(t) \equiv 0$, and by*

$$\begin{aligned}\phi(x) = \phi(x_1,x_2) &= \xi_k\lambda(x_1,x_2)+(1-\xi_k)\lambda(1-x_1,1-x_2)+\gamma\xi_k \\ &= \xi_k[\lambda(x)+\gamma]+(1-\xi_k)\lambda(1-x) \end{aligned} \tag{81}$$

[cf. (65)] if $\mu(t) = \tau_k$ [see (44)], where now $x = (x_1,x_2)$, $x_1 \in]0,1[^{n_1} \cup \{0,1\}$, $x_2 \in]0,1[^{n-n_1}$ with the convention (79). Here γ, δ are arbitrary constants. In (81), x_1 contains the kth component ξ_k of x, and λ is an arbitrary solution of (74) with (75).

The same conventions extend the respective solutions $\Delta_m(p_1,p_2,\ldots,p_m) = \delta$ and

$$\Delta_m(p_1,p_2,\ldots,p_m) = \pi_{1k}\lambda(p_1)+\sum_{j=2}^{m}\pi_{jk}[\lambda(p_j)+\gamma]$$

of (2) and (8) (μ additive and multiplicative) to the case where some probabilities may be 0, with the same restrictions.

We consider now (13) with $\mu(t) \equiv 1$, that is,

$$\phi(x_1,x_2)+\phi\left(\frac{y_1}{1-x_1},\frac{y_2}{1-x_2}\right) = \phi(y_1,y_2)+\phi\left(\frac{x_1}{1-y_1},\frac{x_2}{1-y_2}\right), \tag{82}$$

allowing again, in addition to D, points satisfying (70), (71), or (72) with (73). By substituting, again, $y_1 = 0$, $x_1 \in]0,1[^{n_1}$ [cf. (71)] into (82), we get, in view of (37),

$$\alpha + \lambda(1-x_1, 1-x_2) + \phi\left(0, \frac{y_2}{1-x_2}\right) = \phi(0, y_2) + \alpha + \lambda\left(1-x_1, \frac{1-x_2-y_2}{1-y_2}\right)$$

or, with (75),

$$\phi\left(0, \frac{y_2}{1-x_2}\right) - \phi(0, y_2) = \lambda\left(1, \frac{1-x_2-y_2}{(1-x_2)(1-y_2)}\right)$$
$$= \lambda\left(1, 1-\frac{y_2}{1-x_2}\right) - \lambda(1, 1-y_2).$$

Introducing $t_2 = y_2/(1-x_2)$ again, this goes over into

$$\phi(0, y_2) - \lambda(1, 1-y_2) = \phi(0, t_2) - \lambda(1, 1-t_2) = \beta \quad \text{(constant)}.$$

So

$$\phi(0, y_2) = \beta + \lambda(1, 1-y_2) \qquad \text{for all} \quad y_2 \in]0,1[^{n-n_1}. \tag{83}$$

This satisfies (82) for (71) and also for (70).

At last, we substitute (72), that is $x_1 + y_1 = 1$, into (82), using (37):

$$\alpha + \lambda(1-x_1, 1-x_2) + \phi\left(1, \frac{y_2}{1-x_2}\right) = \alpha + \lambda(x_1, 1-y_2) + \phi\left(1, \frac{x_2}{1-y_2}\right) \tag{84}$$

and specify $x_1 = \frac{1}{2} = (\frac{1}{2}, \dots, \frac{1}{2})$, $s_2 = x_2/(1-y_2)$, $t_2 = y_2/(1-x_2)$ in order to get

$$\phi(1, t_2) - \phi(1, s_2) = \lambda(\tfrac{1}{2}, 1-y_2) - \lambda(\tfrac{1}{2}, 1-x_2-y_2)$$
$$+ \lambda(\tfrac{1}{2}, 1-x_2-y_2) - \lambda(\tfrac{1}{2}, 1-x_2)$$
$$= \lambda(1, 1-t_2) - \lambda(1, 1-s_2).$$

So, for some constant γ,

$$\phi(1,t_2)=\gamma+\lambda(1,1-t_2) \qquad \text{for all} \quad t_2\in\,]0,1[^{n-n_1}. \tag{85}$$

Putting (85) back into (84) or just choosing $y_2=x_2$ in (84), we are in for a surprise:

$$\lambda(1-x_1,1-x_2)=\lambda(x_1,1-x_2). \tag{86}$$

Equation (74) implies (cf. [1, 25])

$$\lambda(x_1,x_2)=\lambda_1(x_1)+\lambda_2(x_2), \tag{87}$$

where

$$\lambda_i(x_iy_i)=\lambda_i(x_i)+\lambda_i(y_i) \qquad (i=1,2). \tag{88}$$

So (86) goes over into

$$\lambda_1(x_1)=\lambda_1(1-x_1),$$

that is,

$$\lambda_1\left(\frac{x_1}{1-x_1}\right)=0.$$

But, as we have mentioned before, every solution of (88), bounded on an interval (and $\{z_1|z_1=x_1/(1-x_1),\ x_1\in\,]0,\tfrac{1}{2}[^{n_1}\}$ is an interval) is of the form

$$\lambda_1(\xi_1,\ldots,\xi_{n_1})=\sum_{k=1}^{n_1}\gamma_k\log\xi_k,$$

and this can be 0 on an interval only if

$$\lambda_1(x_1)\equiv 0.$$

Therefore (87) becomes

$$\lambda(x_1,x_2)=\lambda_2(x_2)=\lambda(1,x_2), \tag{89}$$

and (37) is reduced to

$$\phi(x)=\alpha+\lambda_2(1-x_2).$$

Thus we have the remarkable fact that extending the domain D of (82) to include some boundary points [those which satisfy (72)] restricts even the solution on the original domain D.

In view of (37), (83), (85), and (89), we have proved the following (cf. [2] for the special case $n=2$, $n_1=1$).

THEOREM 4. *The general solution of* (82) *for D and for* (70), (71), (72) *with* (73) *is given by*

$$\phi(x_1,x_2)=\begin{cases}\alpha+\lambda_2(1-x_2) & \text{if } x_1\in\,]0,1[^{n_1},\\ \beta+\lambda_2(1-x_2) & \text{if } x_1=0,\\ \gamma+\lambda_2(1-x_2) & \text{if } x_1=1,\end{cases}$$

where α,β,γ are arbitrary real constants and $\lambda_2:]0,1[^{n-n_1}\to\mathbb{R}$ is an arbitrary solution of (88).

From this we can build up Δ_m $(m=2,3,\dots)$ from ϕ with the aid of (11) and (8) $[\mu(t)\equiv 1]$.

Finally, if μ is multiplicative but not additive and not identically 1, we write (13) as

$$\begin{aligned}\phi(x_1,x_2)&+\mu(1-x_1,1-x_2)\phi\left(\frac{y_1}{1-x_1},\frac{y_2}{1-x_2}\right)\\ &=\phi(y_1,y_2)+\mu(1-y_1,1-y_2)\phi\left(\frac{x_1}{1-y_1},\frac{x_2}{1-y_2}\right)\end{aligned}\tag{90}$$

and put into it (71) ($y_1=0$, $x_1\in\,]0,1[^{n_1}$) and (34):

$$\begin{aligned}\alpha\mu(x_1,x_2)&+\beta\mu(1-x_1,1-x_2)-\beta+\mu(1-x_1,1-x_2)\phi\left(0,\frac{y_2}{1-x_2}\right)\\ &=\phi(0,y_2)+\mu(1,1-y_2)\left[\alpha\mu\left(x_1,\frac{x_2}{1-y_2}\right)+\beta\mu\left(1-x_1,\frac{1-x_2-y_2}{1-y_2}\right)-\beta\right].\end{aligned}\tag{91}$$

Here and in what follows we suppose that μ has been extended with the aid of (22) to $]0,\infty[^n$ (and omit the bars). Then μ satisfies [compare to (74) and (76)]

$$\mu(t_1u_1, t_2u_2)=\mu(t_1,t_2)\mu(u_1,u_2) \quad \text{and} \quad \mu\left(\frac{t_1}{u_1},\frac{t_2}{u_2}\right)=\frac{\mu(t_1,t_2)}{\mu(u_1,u_2)}$$

$$\text{for all} \quad t_1,t_2\in\,]0,\infty[^{n_1}, \quad u_1,u_2\in\,]0,\infty[^{n-n_1} \tag{92}$$

(μ is nowhere 0, as shown in Section 2). So (91) can be written as

$$\mu(1-x_1,1-x_2)\left[\phi\left(0,\frac{y_2}{1-x_2}\right)+\beta-\beta\mu\left(1,\frac{1-x_2-y_2}{1-x_2}\right)\right]$$
$$=\phi(0,y_2)+\beta-\beta\mu(1,1-y_2).$$

If we divide this by $\mu(1,y_2)$ and write $t_2=y_2/(1-x_2)$ again, we get

$$\frac{\phi(0,t_2)+\beta-\beta\mu(1,1-t_2)}{\mu(1,t_2)}\mu(1-x_1,1)=\frac{\phi(0,y_2)+\beta-\beta\mu(1,1-y_2)}{\mu(1,y_2)}$$
$$=\alpha' \quad \text{(constant)}, \tag{93}$$

that is,

$$\phi(0,y_2)=\alpha'\mu(1,y_2)+\beta\mu(1,1-y_2)-\beta \qquad \left(y_2\in\,]0,1[^{n-n_1}\right). \tag{94}$$

and

$$\alpha'\mu(1-x_1,1)=\alpha'.$$

Hence either $\alpha'=0$ and

$$\phi(0,y_2)=\beta\mu(1,1-y_2)-\beta \qquad \left(y_2\in\,]0,1[^{n-n_1}\right), \tag{95}$$

or

$$\mu(t_1,1)=1,$$

that is [cf. (92)]

$$\mu(t_1,t_2)=\mu(1,t_2)=\mu_2(t_2), \qquad \text{where} \quad \mu_2(t_2u_2)=\mu_2(t_2)\mu_2(u_2). \tag{96}$$

The functions given by (95) with arbitrary constant β and, if (96) holds, also (94) with arbitrary constants α', β satisfy (91) and also

$$\phi(0,x_2)+\mu(1,1-x_2)\phi\left(0,\frac{y_2}{1-x_2}\right)=\phi(0,y_2)+\mu(1,1-y_2)\phi\left(0,\frac{x_2}{1-y_2}\right)$$

obtained from (90) by the substitution (70).

We conclude by putting (72) and (34) into (90):

$$\alpha\mu(x_1,x_2)+\beta\mu(1-x_1,1-x_2)-\beta+\mu(1-x_1,1-x_2)\phi\left(1,\frac{y_2}{1-x_2}\right)$$

$$=\alpha\mu(1-x_1,y_2)+\beta\mu(x_1,1-y_2)-\beta+\mu(x_1,1-y_2)\phi\left(1,\frac{x_2}{1-y_2}\right). \quad (97)$$

We substitute here $x_1=\frac{1}{2}$, $s_2=x_2/(1-y_2)$, $t_2=y_2/(1-x_2)$, divide (97) by $\mu(\frac{1}{2},1-x-y)$, and apply (92) in order to get

$$\frac{\phi(1,s_2)+\beta-\alpha\mu(1,s_2)}{\mu(1,1-s_2)}=\frac{\phi(1,t_2)+\beta-\alpha\mu(1,t_2)}{\mu(1,1-t_2)}=\beta' \quad \text{(constant)}.$$

Putting

$$\phi(1,t_2)=\alpha\mu(1,t_2)+\beta'\mu(1,1-t_2)-\beta \quad (98)$$

into (97), we get, with (92) and with $t_1=x_1/(1-x_1)$,

$$\beta'=\beta'\mu(t_1,1).$$

So, either $\beta'=0$ and

$$\phi(1,t_2)=\alpha\mu(1,t_2)-\beta \quad (99)$$

or

$$\mu(t_1,1)=1$$

and (96) holds. The functions given by (99) (with arbitrary constant α, β and, if (96) holds, also (98) with constant α, β, β' indeed satisfy (97). So we have proved the following (cf. [2, 4] for $n = 2$, $n_1 = 1$, $\mu(\tau_1, \tau_2) = \tau_1^{\alpha_1}\tau_2^{\alpha_2}$).

THEOREM 5. *The general solutions of* (90) (*with* $\mu \not\equiv 1$ *multiplicative but not additive*), *for the domain covering* D *and* (70), (71), (72) *along with* (73), *are given by* (34), (95), (99), *and, if* (96) *holds, also by*

$$\phi(x_1, x_2) = \begin{cases} \alpha\mu_2(x_2) + \beta\mu_2(1 - x_2) - \beta & \text{for} \quad x_1 \in]0,1[^{n_1}, \\ \alpha'\mu_2(x_2) + \beta\mu_2(1 - x_2) - \beta & \text{for} \quad x_1 = 0, \\ \alpha\mu_2(x_2) + \beta'\mu_2(1 - x_2) - \beta & \text{for} \quad x_1 = 1, \end{cases}$$

where $\alpha, \alpha', \beta, \beta'$ *are arbitrary constants and* $\mu_2(x_2) = \mu(1, x_2)$.

Again, it is easy to build up Δ_m ($m = 2,3,\ldots$) from ϕ with the aid of (11) and (8). In particular, since (95) and (99) can be considered as extensions of (34) with $\mu(0, t_2) = 0$, in this case Δ_m is given by (35) also if some components of a vector p_j are 0. Values of ϕ and Δ_m at other boundary points can also be easily determined by extending the domains of (13) and (8), respectively.

REFERENCES

1 J. Aczél, On a system of functional equations determining price and productivity indices, *Utilitas Math.* 7:345–362 (1975).

2 J. Aczél, Information functions on degree $(0, \beta)$, *Utilitas Math.* 18:15–26 (1980).

3 J. Aczél, Information functions on open domain. III *C.R. Math. Rep. Acad. Sci. Canada* 2:281–285 (1980).

4 J. Aczél, Notes on generalized information functions, *Aequationes Math.* 22:97–107 (1981).

5 J. Aczél, J. A. Baker, D. Ž. Djoković, Pl. Kannappan, and F. Radó, Extensions of certain homomorphisms of semigroups to homomorphisms of groups, *Aequationes Math.* 6:263–271 (1971).

6 J. Aczél and Z. Daróczy, *On Measures of Information and Their Characterizations*, Academic, New York, San Francisco, 1975.

7 J. Aczél and Pl. Kannappan, General two-place information functions, *Resultate Math.* 5:99–106 (1982).

8 J. Aczél and C. T. Ng, On general information functions, *Utilitas Math.* 19:157–170 (1981).

9 Z. Daróczy, Generalized information functions, *Inform. and Control* 16:36–51 (1971).

10 G. T. Diderrich, Boundedness on a set of positive measure and the fundamental equation of information, *Publ. Math. Debrecen*, to appear.

11 D. K. Faddeev, On the concept of entropy of a finite probabilistic scheme (Russian), *Uspehi Mat. Nauk* (*N.S.*) 11, No. 1 (67):227–231 (1956).
12 J. Havrda and F. Charvát, Quantification method of classification processes. Concept of structural a-entropy, *Kybernetika* (*Prague*) 3:30–35 (1967).
13 A. Kaminski and Pl. Kannappan, A note on some theorem in information theory, *Bull. Acad. Polon. Sci. Sér. Sci. Math.* 25:925–928 (1977).
14 A. Kaminski and Pl. Kannappan, Distributional solutions in information theory. II. *Ann. Polon. Math.* 38:289–294 (1980).
15 Pl. Kannappan, On two functional equations from information theory, *J. Indian Math. Soc.* (*N.S.*), to appear.
16 Pl. Kannappan and C.T. Ng, Measurable solutions of functional equations related to information theory, *Proc. Amer. Math. Soc.* 38:303–310 (1973).
17 Pl. Kannappan and C.T. Ng, A functional equation and its application to information theory, *Ann. Polon. Math.* 30:105–112 (1974).
18 Pl. Kannappan and P.N. Rathie, An application of a functional equation to information theory, *Ann. Polon. Math.* 26:95–101 (1972).
19 Pl. Kannappan and P. N. Rathie, An axiomatic characterization of generalized directed-divergence, *Kybernetika* (*Prague*) 9:330–337 (1973).
20 Pl. Kannappan and P. N. Rathie, On a characterization of directed divergence, *Inform. and Control* 22:163–171 (1973).
21 Pl. Kannappan and P. N. Rathie, On a generalized directed-divergence function, *Czechoslovak Math. J.* 24 (99):5–14 (1974).
22 Pl. Kannappan and P. N. Rathie, On the solution of a functional equation connected with inaccuracies, *J. Indian Math. Soc.* 39:131–147 (1975).
23 Pl. Kannappan and P. N. Rathie, On generalized information function, *Tôhoku Math. J.* 27:207–212 (1975).
24 Pl. Kannappan and P. N. Rathie, On a generalized directed divergence and related measures, in *Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes and of the Eighth European Meeting of Statisticians* (*Tech. Univ. Prague 1974*), Vol. B, Academia, Prague, 1978, pp. 255–265.
25 M. Kuczma, Note on additive functions of several variables, *Uniw. Śląski w Katowicach Prace Nauk.–Prace Mat.* 2:49–51 (1972).
26 Gy. Maksa, Bounded symmetric information functions, *C.R. Math. Rep. Acad. Sci. Canada* 2:247–252 (1980).
27 Gy. Maksa, Solution on the open triangle of the generalized fundamental equation of information with four unknown functions, *Utilitas Math.*, to appear.
28 Gy. Maksa and C. T. Ng, The fundamental equation of information on open domain, *Publ. Math. Debrecen*, to appear.
29 C. T. Ng, Information functions on open domains. II, *C.R. Math. Rep. Acad. Sci. Math. Canada* 2:155–158 (1980).
30 A. Ostrowski, Mathematische Miszellen. XIV. Über die Funktionalgleichung der Exponentialfunktion und verwandte Funktionalgleichungen, *Jahresber. Deutsch. Math.-Verein.* 38:54–62 (1929).
31 P. N. Rathie and Pl. Kannappan, On a functional equation connected with Shannon's entropy, *Funkcial. Ekvac.* 14:153–159 (1971).

32 P. N. Rathie and Pl. Kannappan, A directed-divergence function of type β, *Inform. and Control* 20:38–45 (1972).

33 P. N. Rathie and Pl. Kannappan, On a new characterization of directed divergence in information theory, in *Transaction of the Sixth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes* (*Prague 1971*), Academia, Prague, 1973, pp. 733–745.

34 P. N. Rathie and Pl. Kannappan, An inaccuracy function of type β, *Ann. Inst. Statist. Math.* 25:205–214 (1973).

35 I. J. Taneja, On information-improvement and its generalization, *Metrika* 23:249–257 (1976).

36 I. J. Taneja, Some contributions to information theory—I. (A survey: On measures of information), *J. Combin. Inform. System Sci.* 4:253–274 (1979).

Received 30 August 1981