Canad. Math. Bull. Vol **22** (4), 1979

BRANCHING MEASURES OF INFORMATION ON STRINGS

BY
BRUCE R. EBANKS[1]

In classical information theory, the amount of information provided by an experiment is measured by a function of the probability distribution of the outcomes of the experiment. In this paper, information measures are functions of sequences of elements of a monoid $(S, \circ)$ with identity e. It is assumed that the measures $\{\mu_n : S^n \to \mathbb{R}\}$ of information are *branching*. For several classes of monoids, it is found that μ_n has a representation in the form

$$\mu_n(s_1, s_2, \ldots, s_n) = \sum_{i=1}^{n} f_{n,i}(s_i) + \sum_{i=1}^{n-1} \psi_n(s_i, s_{i+1} \circ \cdots \circ s_n),$$

where ψ_n is anti-symmetric and bi-additive.

1. **Introduction.** The purpose of this paper is to consider the problem of measuring the amount of information provided by a string (sequence) of symbols from some universal set S. It is assumed that these symbols can be combined under some binary associative operation $\circ$ which makes $(S, \circ)$ a *monoid* (i.e., a semigroup with unity) with unity e. For example, we can think of the output of a keyboard which occasionally fails to space properly, causing a symbol to be superposed on another. The set of symbols and their finite combinations under the operation of superposition then constitutes a monoid in which the identity is the blank space.

The (real-valued) measure μ of information is in reality a set $\{\mu_n : S^n \to \mathbb{R}\}$ of measures of the information contents of strings of various lengths. We arbitrarily fix $n \geq 3$ and seek the form of $\mu_{n'}$, the information measure of strings of length n. The essential property is that of branching, which is defined as follows.

We assume that the measure under μ_n of a string $(s_1, s_2, \ldots, s_n) \in S^n$ differs from the measure of a distorted string, in which s_i and s_{i+1} are "mixed", by an amount which depends only on s_i, s_{i+1}, and their position in the string.

Received by the editors January 6, 1978 and, in revised form, September 12, 1978.

[1] This work constitutes part of a Ph.D. thesis written at the University of Waterloo, Ontario, Canada under the supervision of C. T. Ng and J. Aczél.

Precisely, μ_n is *branching* if there exist maps $\Delta_{n,i}: S^2 \to \mathbb{R}$ such that

(1.1) $$\mu_n(s_1, s_2, \ldots, s_n) = \mu_n(s_1, s_2, \ldots, s_{i-1}, s_i \circ s_{i+1}, e, s_{i+2}, \ldots, s_n) + \Delta_{n,i}(s_i, s_{i+1}),$$

for all $i = 1, 2, \ldots, n-1$ and all $(s_1, s_2, \ldots, s_n) \in S^n$. This is analogous to the branching property in classical information theory, described, for example in Aczél and Daróczy [11].

The following examples illustrate a few settings in which the model introduced may be applied.

EXAMPLE 1. The example mentioned in the first paragraph.

EXAMPLE 2. Let S be an algebra of subsets of some universal set U, and let $\circ$ denote the operation of set intersection (or union). A string of symbols (subsets) from S might represent outcomes of a sequence of experiments on U, and the combination of two symbols under $\circ$ would correspond to the mixing of outcomes of a pair of experiments.

EXAMPLE 3. Let S be a range of intensities (of sound, light, pressure, etc.) which is closed under the operation $\circ$ of composition of intensities. For instance, one might be interested in measuring some aspect of the sound of a symphony, based on the collection of intensities of the various instruments. The branching property would govern the comparison of different instrumentations.

We shall consider the branching property over several classes of semigroups which include, among others, the special ones described in the preceding examples.

2. **Functional equations; existence of branching measures.** One can use (1.1) several times to get

$$\begin{aligned}
&\mu_n(s_1, \ldots, s_{i-2}, s_{i-1} \circ s_i \circ s_{i+1}, e, e, s_{i+2}, \ldots, s_n) + \Delta_{n,i-1}(s_{i-1}, s_i \circ s_{i+1}) + \Delta_{n,i}(s_i, s_{i+1}) \\
&\quad = \mu_n(s_1, \ldots, s_{i-1}, s_i \circ s_{i+1}, e, s_{i+2}, \ldots, s_n) + \Delta_{n,i}(s_i, s_{i+1}) \\
&\quad = \mu_n(s_1, \ldots, s_n) \\
&\quad = \mu_n(s_1, \ldots, s_{i-2}, s_{i-1} \circ s_i, e, s_{i+1}, \ldots, s_n) + \Delta_{n,i-1}(s_{i-1}, s_i) \\
&\quad = \mu_n(s_1, \ldots, s_{i-2}, s_{i-1} \circ s_i, s_{i+1}, e, s_{i+2}, \ldots, s_n) + \Delta_{n,i}(e, s_{i+1}) + \Delta_{n,i-1}(s_{i-1}, s_i) \\
&\quad = \mu_n(s_1, \ldots, s_{i-2}, s_{i-1} \circ s_i \circ s_{i+1}, e, e, s_{i+2}, \ldots, s_n) \\
&\qquad + \Delta_{n,i-1}(s_{i-1} \circ s_i, s_{i+1}) + \Delta_{n,i}(e, s_{i+1}) + \Delta_{n,i-1}(s_{i-1}, s_i)
\end{aligned}$$

Comparing the two extremes of this line of equations, we have

(2.1) $$\Delta_{n,i-1}(x, y \circ z) + \Delta_{n,i}(y, z) = \Delta_{n,i-1}(x \circ y, z) + \Delta_{n,i}(e, z) + \Delta_{n,i-1}(x, y),$$

for all $(x, y, z) \in S^3$, $i = 2, 3, \ldots, n-1$. With $x = e$, (2.1) yields

$$\Delta_{n,i}(y, z) = -\Delta_{n,i-1}(e, y \circ z) + \Delta_{n,i-1}(y, z) + \Delta_{n,i}(e, z) + \Delta_{n,i-1}(e, y)$$

https://doi.org/10.4153/CMB-1979-057-3 Published online by Cambridge University Press

for $i=2,3,\ldots,n-1$, and, in particular,

$$\Delta_{n,n-2}(y,z)=\Delta_{n,n-1}(y,z)+\Delta_{n,n-2}(e,y\circ z)-\Delta_{n,n-1}(e,z)-\Delta_{n,n-2}(e,y).$$

These equations enable us to rewrite (2.1) as

$$\Delta_{n,i}(x,y\circ z)-\Delta_{n,i}(e,y\circ z)+\Delta_{n,i}(y,z)+\Delta_{n,i}(e,y)=\Delta_{n,i}(x\circ y,z)+\Delta_{n,i}(x,y), \tag{2.2}$$

for all $(x,y,z)\in S^3$ and $i=1,2,\ldots,n-1$. Now defining maps $F_{n,i}:S^2\to\mathbb{R}$ by

$$F_{n,i}(x,y):=\Delta_{n,i}(x,y)-\Delta_{n,i}(e,y),\qquad \forall x,y\in S, \tag{2.3}$$

we find that each $F_{n,i}$ $(i=1,2,\ldots,n-1)$ satisfies

$$F_{n,i}(x,y)+F_{n,i}(x\circ y,z)=F_{n,i}(x,y\circ z)+F_{n,i}(y,z) \tag{2.4}$$

for all $x,y,z\in S$.

In summary, we have

LEMMA 2.1. *If a measure $\mu_n:S_n\to\mathbb{R}$ has the* (1.1) *branching property over a monoid $(S,\circ)$, then $\Delta_{n,i}$ can be represented in the form*

$$\Delta_{n,i}(x,y)=F_{n,i}(x,y)+\Delta_{n,i}(e,y),\quad \mathbb{R}\,\forall x,y\in S, \tag{2.5}$$

$i=1,2,\ldots,n-1$, *where $F_{n,i}:S^2\to\mathbb{R}$ satisfies* (2.4) *and $\Delta_{n,i}(e,.)$ is an arbitrary map of S into $\mathbb{R}$.*

It is easily verified that the following result provides a large class of solutions to equation (2.4) without using (2.3).

LEMMA 2.2. *Let $(S,\circ)$ be a semigroup, and let $F_{n,i}:S^2\to\mathbb{R}$ be defined by*

$$F_{n,i}(x,y):=f_{n,i}(x)+f_{n,i}(y)-f_{n,i}(x\circ y)+\psi_{n,i}(x,y), \tag{2.6}$$

for all $x,y\in S$, where $f_{n,i}:S\to\mathbb{R}$ and $\psi_{n,i}:S^2\to\mathbb{R}$ so that

$$\psi_{n,i}(x,y)=-\psi_{n,i}(y,x),\qquad \forall x,y\in S, \tag{2.7}$$

$$\psi_{n,i}(x\circ y,z)=\psi_{n,i}(x,z)+\psi_{n,i}(y,z),\qquad \forall x,y,z\in S. \tag{2.8}$$

Then $F_{n,i}$ satisfies (2.4).

Since our aim is to show that solutions of (2.4) under various alegebraic and/or topological conditions on S always have the form (2.6), let us now examine what representation of μ_n these solutions yield.

LEMMA 2.3. *Let μ_n be a* (1.1) *branching measure of information over a commutative monoid $(S,\circ)$. If the maps $\Delta_{n,i}$ $(i=1,2,\ldots,n-1)$ are of the form* (2.5) *with $F_{n,i}$ given by* (2.6) *for $\psi_{n,i}$ satisfying* (2.7) *and* (2.8), *then μ_n can be represented in the form*

$$\mu_n(s_1,s_2,\ldots,s_n)=k_{n,0}(s_1\circ s_2\circ\cdots\circ s_n)+\sum_{i=1}^{n}k_{n,i}(s_i)+\sum_{i=1}^{n-1}\psi_n(s_i,s_{i+1}\circ s_{i+2}\circ\cdots\circ s_n), \tag{2.9}$$

for all $(s_1, s_2, \ldots, s_n) \in S^n$, *with* $k_{n,i} : S \to \mathbb{R}$ *satisfying* $k_{n,i}(e) = 0$ *for* $n \geq i \geq 1$, *and* $\psi_{n,i} = \psi_n (\forall i)$.

Proof. (2.5) and (2.6) imply that for all $x, y \in S$,

$$\Delta_{n,i}(x, y) = f_{n,i}(x) + f_{n,i}(y) + \Delta_{n,i}(e, y) - f_{n,i}(x \circ y) + \psi_{n,i}(x, y) \tag{2.10}$$

If we put $x = e$ in (2.8), we find that

$$\psi_{n,i}(e, y) = 0, \qquad \forall y \in S. \tag{2.11}$$

Thus (2.10) with $x = e$ yields

$$f_{n,i}(e) = 0. \tag{2.12}$$

With these preliminaries established, we return to (2.10) and use it to write (2.1) as

$$\begin{aligned} f_{n,i-1}(y \circ z) + \Delta_{n,i-1}(e, y \circ z) + \psi_{n,i-1}(x, y \circ z) + f_{n,i}(y) + f_{n,i}(z) - f_{n,i}(y \circ z) + \psi_{n,i}(y, z) \\ = f_{n,i-1}(z) + \Delta_{n,i-1}(e, z) + \psi_{n,i-1}(x \circ y, z) + f_{n,i-1}(y) + \Delta_{n,i-1}(e, y) + \psi_{n,i-1}(x, y). \end{aligned}$$

By (2.8), this can be rewritten as

$$\begin{aligned} f_{n,i-1}(y \circ z) + \Delta_{n,i-1}(e, y \circ z) - f_{n,i}(y \circ z) + \psi_{n,i}(y, z) \\ = f_{n,i-1}(y) + \Delta_{n,i-1}(e, y) - f_{n,i}(y) + f_{n,i-1}(z) \\ + \Delta_{n,i-1}(e, z) - f_{n,i}(z) + \psi_{n,i-1}(y, z). \end{aligned} \tag{2.13}$$

Defining functions $h_{n,i} : S \to \mathbb{R}$ $(i = 1, 2, \ldots, n-2)$ by

$$h_{n,i}(x) := f_{n,i}(x) + \Delta_{n,i}(e, x) - f_{n,i+1}(x), \qquad \forall x \in S, \tag{2.14}$$

we obtain

$$h_{n,i-1}(y \circ z) + \psi_{n,i}(y, z) = h_{n,i-1}(y) + h_{n,i-1}(z) + \psi_{n,i-1}(y, z) \tag{2.15}$$

from (2.13). Equating symmetric and anti-symmetric parts (cf. (2.7) and the commutativity of $\circ$), we have

$$h_{n,i}(y \circ z) = h_{n,i}(y) + h_{n,i}(z), \qquad \forall y, z \in S, \tag{2.16}$$

$$\psi_{n,i}(y, z) = \psi_{n,i-1}(y, z), \qquad \forall y, z \in S. \tag{2.17}$$

If we define $\psi_n : S^2 \to \mathbb{R}$, $h_{n,n-1} : S \to \mathbb{R}$, and $f_{n,n} : S \to \mathbb{R}$ by

$$\psi_n(x, y) := \psi_{n,1}(x, y), \qquad \forall x, y \in S, \tag{2.18}$$

$$h_{n,n-1}(x) := 0, \qquad \forall x \in S, \tag{2.19}$$

$$f_{n,n}(x) := f_{n,n-1}(x) + \Delta_{n,n-1}(e, x), \qquad \forall x \in S, \tag{2.20}$$

then (2.14), (2.17), (2.18), (2.19), and (2.20) transform (2.10) into

$$\Delta_{n,i}(x, y) = f_{n,i}(x) + f_{n,i+1}(y) + h_{n,i}(y) - f_{n,i}(x \circ y) + \psi_n(x, y) \tag{2.21}$$

for all $x, y \in S$ and $i = 1, 2, \ldots, n-1$.

https://doi.org/10.4153/CMB-1979-057-3 Published online by Cambridge University Press

Now we are ready to evaluate the measure of a string. Indeed, by (1.1), (2.21), and (2.16), we have

$$\begin{aligned}
&\mu_n(s_1, s_2, \ldots, s_n)\\
&\quad = \mu_n(s_1, s_2, \ldots, s_{n-2}, s_{n-1}\circ s_n, e) + f_{n,n-1}(s_{n-1}) + f_{n,n}(s_n)\\
&\qquad + h_{n,n-1}(s_n) - f_{n,n-1}(s_{n-1}\circ s_n) + \psi_n(s_{n-1}, s_n)\\
&\quad = \cdots\\
&\quad = \mu_n(s_1\circ s_2\circ, \ldots \circ s_n, e, e, \ldots, e) + \sum_{i=1}^{n} f_{n,i}(s_i)\\
&\qquad + \sum_{i=2}^{n}\sum_{j=1}^{i-1} h_{n,j}(s_i) - f_{n,1}(s_1\circ s_2 \cdots \circ s_n)\\
&\qquad + \sum_{i=1}^{n-1} \psi_n(s_i, s_{i+1}\circ s_{i+2} \cdots \circ s_n),
\end{aligned}$$

for all $(s_1, s_2, \ldots, s_n) \in S^n$. Defining $k_{n,i} : S \to \mathbb{R}$ $(1 \leq i \leq n)$ by $k_{n,0}(x) := \mu_n(x, e, e, \ldots, e) - f_{n,1}(x)$, $k_{n,1}(x) := f_{n,1}(x)$,

$$k_{n,i}(x) := f_{n,i}(x) + \sum_{j=1}^{i-1} h_{n,j}(x),\ 1 < i \leq n, \tag{2.22}$$

for all $x \in S$, we have (2.9). Moreover, (2.16) with $y = z = e$ gives $h_{n,i}(e) = 0$. This, together with (2.22) and (2.12), yields $k_{n,i}(e) = 0$ for $1 \leq i \leq n$.

Finally, by (2.18), ψ_n satisfies (2.7) and (2.8). This completes the proof of Lemma 2.3.

Since we now know the form of μ_n when (2.4) has a solution of the form (2.6), we shall concentrate on solving (2.4) on various classes of semigroups. The subscripts of solutions $F_{n,i}$ of (2.4) will be suppressed where this leads to no ambiguity.

3. **Solution when $(S, \circ)$ is idempotent and commutative.**

THEOREM 3.1. *Let $(S, \circ)$ be a commutative idempotent semigroup, and let $F : S^2 \to \mathbb{R}$. Then F satisfies* (2.4) *only if there is a map $f : S \to \mathbb{R}$ such that*

$$F(x, y) = f(x) + f(y) - f(x\circ y), \qquad \forall x, y \in S. \tag{3.1}$$

[*N.B.* (3.1) *is just* (2.6) *with* $\Psi \equiv 0$.]

Proof. Putting $x = y$ in (2.4), we get

$$F(y, y) + F(y, z) = F(y, y\circ z) + F(y, z)$$

by virtue of the idempotence of $(S, \circ)$. Hence,

$$F(y, y\circ z) = F(y, y), \qquad \forall y, z \in S. \tag{3.2}$$

Using, in order, (2.4) with $z=x\circ y$, commutativity of $\circ$, idempotence of $\circ$, commutativity again, and finally (3.2), we obtain

$$\begin{aligned} F(x,y)+F(x\circ y,x\circ y) &= F(x,y\circ x\circ y)+F(y,x\circ y) \\ &= F(x,y\circ y\circ x)+F(y,y\circ x) \\ &= F(x,y\circ x)+F(y,y\circ x) \\ &= F(x,x\circ y)+F(y,y\circ x) \\ &= F(x,x)+F(y,y) \end{aligned}$$

for all $x, y\in S$. With $f:S\to\mathbb{R}$ defined by $f(x):=F(x,x)$, for all $x\in S$, this leads to (3.1), which was to be proved.

With respect to the problem of measuring information, we have the

THEOREM 3.2. *Let μ_n be a branching measure over an idempotent, abelian monoid $(S,\circ)$. Then μ_n is represented by*

$$\mu_n(s_1,s_2,\ldots,s_n)=k_{n,0}(s_1\circ s_2\circ\cdots\circ s_n)+\sum_{i=1}^{n}k_{n,i}(s_i), \tag{3.3}$$

for all $(s_1,s_2,\ldots,s_n)\in S^n$, with $k_{n,i}:S\to\mathbb{R}$ satisfying

$$k_{n,i}(e)=0, \qquad i=1,2,\ldots,n. \tag{3.4}$$

Proof. By Lemma 2.1, the $\Delta_{n,i}$'s are of the form (2.5) with $F_{n,i}$'s satisfying (2.4). Then Theorem 3.1 yields the representation (3.1) for each $F_{n,i}$. Now the hypotheses of Lemma 2.3 are fulfilled (where all $\Psi_{n,i}$ are identically zero), so μ_n is given by (2.9). Finally, $\Psi_{n,i}\equiv 0$ for $i=1,2,\ldots,n-1$ yields $\Psi_n\equiv 0$, reducing (2.9) to the form (3.3). (3.4) also follows from Lemma 2.3, establishing the theorem.

As one interpretation of Theorem 3.2, we can think of Example 1 with S the set generated by $\{0,1,2,\ldots,9\}$ under the operation $\circ$ of superposition, together with the blank space (which serves as the identity). If we define $k_{n,i}:S\to\mathbb{R}$ (with a given, fixed n) by

$$k_{n,i}(x):=\begin{cases} x\cdot 10^{n-i}, & \forall x\in\{0,1,2,\ldots,9\},\quad 0<i\leq n,\\ 0, & \text{if}\quad x\in S\setminus\{0,1,\ldots,9\},\quad\text{or}\quad i=0.\end{cases}$$

then $\mu_n(d_1,d_2,\ldots,d_n)=d_1\cdot 10^{n-1}+d_2\cdot 10^{n-2}+\cdots+d_n\cdot 10^0=$ the real number represented by juxtaposition $d_1d_2\cdots d_n$ of the digits $d_1,d_2,\ldots,d_n$. If some of the d_i's are blank spaces or superpositions of digits, rather than simply digits, then zeros appear in their places in the representation.

4. **Solution when $(S,\circ)$ is a thread or has a zero.**

THEOREM 4.1. *Let $(S,\circ)$ be a semigroup with a zero 0, and let $F:S^2\to\mathbb{R}$ be a solution of (2.4). Then there exists a map $f:S\to\mathbb{R}$ for which F has the form (3.1).*

Proof. Putting $z=0$ in (2.4), we get

$$F(x, y)+F(x\circ y, 0)=F(x, 0)+F(y, 0)$$

for all $x, y \in S$. With $f(t):=F(t, 0)$ for all $t\in S$, this is (3.1), and we are finished.

A result related to Theorem 4.1 concerns semigroups known as threads. Let $(S, \circ)$ be a totally ordered semigroup which is connected and has continuous multiplication with respect to the order topology. If, in addition, S has a greatest and least element with respect to the order, and if these endpoints are idempotent, then $(S, \circ)$ is called a *thread*. As an example, think of the real interval $[0, 1]$ with the usual multiplication. (There are many other non-isomorphic examples of threads.) For threads in general, we have

THEOREM 4.2. *Let* $(S, \circ)$ *be a thread, and let* $F: S^2 \to \mathbb{R}$ *satisfy* (2.4). *Then there is a map* $f: S\to\mathbb{R}$ *representing* F *through* (3.1).

Proof. Since S is connected and has endpoints, S is compact. Numakura [7] and Wallace [10] have shown, independently, that S contains a kernel K (i.e., a minimal closed ideal), and that K is connected. Hence K is a subinterval $[\alpha, \beta]\subseteq S$. K is a single point if and only if $\alpha=\beta=$ a zero for $(S, \circ)$. If this is the case, we apply Theorem 4.1 to complete the proof.

Now suppose $\alpha<\beta$. Faucett [2] has shown that multiplication in K is either left-trivial, i.e.,

$$\lambda\circ\mu=\mu, \qquad \lambda, \mu\in K, \tag{4.1}$$

or right-trivial ($\lambda\circ\mu=\lambda$, for all $\lambda, \mu\in K$). Without loss of generality, let us take (4.1) to be the case. With $x, y, z\in K$ in (2.4), using (4.1), we get

$$F(x, y)=F(x, z), \qquad x, y, z\in K. \tag{4.2}$$

With $x, z\in K$ and $y\in S$ in (2.4), taking (4.2) into account, we have

$$F(x, y)+F(x\circ y, w)=F(x, w)+F(y, z),$$

where w is an arbitrary element of K. Thus, $F(y, z)$ does not depend on $z\in K$. Define $f: S\to\mathbb{R}$ by

$$f(t):=F(t, w), \qquad \forall t\in S, \qquad w\in K. \tag{4.3}$$

Finally, restrict z to K while $x, y\in S$ in (2.4) to get, by (4.3),

$$F(x, y)+f(x\circ y)=f(x)+f(y), \qquad \forall x, y\in S,$$

since $y\circ z\in K$.

5. **Solution on groups and threads in the wider sense.** The solution of (2.4) (with one additional assumption) when $(S, \circ)$ is an ordered commutative group was obtained by Jessen, Karpf, and Thorup [4] and in quite a different context. In fact, Jessen [3] used their result in a simplified proof of Sydler's theorem on polyhedra. We quote part of their results here.

THEOREM 5.1. *Let $(A, \oplus)$ be an ordered commutative group, let $(X, +)$ be a divisible commutative group, and define $A_+ := \{a \in A \mid a > 0\}$. Then the class of functions $F: A_+^2 \to X$ (or $F: A^2 \to X$) determined by means of a function $f: A_+ \to X$ (or $f: A \to X$) through*

$$F(a, b) = f(a) + f(b) - f(a \oplus b) \tag{5.1}$$

is identical with the class of functions $F: A_+^2 \to X$ (or $F: A^2 \to X$) satisfying the equations

$$F(a, b) + F(a \oplus b, c) = F(a, b \oplus c) + F(b, c), \tag{5.2}$$

$$F(a, b) = F(b, a). \tag{5.3}$$

We shall use the following strengthening of their result, in which the (5.3) symmetry is discarded.

COROLLARY 5.2. *Under the hypotheses of Theorem 5.1, suppose a function $F: A_+^2 \to X$ ($F: A^2 \to X$) satisfies equation (5.2). Then (and only then, by Lemma 2.2) F is determined through*

$$F(a, b) = f(a) + f(b) - f(a \oplus b) + \Psi(a, b), \tag{5.4}$$

by means of a function $f: A_+ \to X$ ($f: A \to X$) and a function $\Psi: A_+^2 \to X$ ($\Psi: A^2 \to X$) satisfying the equations

$$\Psi(a, b) = -\Psi(b, a), \qquad \forall a, b \in A_+(A), \tag{5.5}$$

$$\Psi(a \oplus b, c) = \Psi(a, c) + \Psi(b, c), \qquad \forall a, b, c \in A_+(A). \tag{5.6}$$

Proof. Consider the decomposition $F = F_s + \Psi$ of F into its canonical symmetric and anti-symmetric parts,

$$F_s(a, b) = \tfrac{1}{2}[F(a, b) + F(b, a)], \Psi(a, b) = \tfrac{1}{2}[F(a, b) - F(b, a)]. \tag{5.7}$$

Now F_s satisfies (5.2), (5.3), hence it has the representation (5.1) for some map $f: A_+ \to X$ ($f: A \to X$). This, in turn, gives the form (5.4) for F, where Ψ satisfies (5.5) because of (5.7). It only remains to show that Ψ satisfies (5.6), which is accomplished by use of (5.2) three times, as follows.

$$\begin{aligned}
2\Psi(x \oplus y, z) &= F(x \oplus y, z) - F(z, x \oplus y) \\
&= [F(x, y \oplus z) + F(y, z) - F(x, y)] - [F(z \oplus x, y) + F(z, x) - F(x, y)] \\
&= F(x, z \oplus y) - F(x \oplus z, y) + F(y, z) - F(z, x) \\
&= F(x \oplus z, y) + F(x, z) - F(z, y) - F(x \oplus z, y) + F(y, z) - F(z, x) \\
&= [F(x, z) - F(z, x)] + [F(y, z) - F(z, y)] \\
&= 2\Psi(x, z) + 2\Psi(y, z), \qquad \forall x, y, z \in A_+(A).
\end{aligned}$$

We shall use Corollary 5.2 to solve (2.4) on semigroups which satisfy all the properties of a thread except the requirements concerning extremal elements.

https://doi.org/10.4153/CMB-1979-057-3 Published online by Cambridge University Press

Such semigroups are called *threads in the wider sense*, or, simply, *w-threads*. In other words, a w-thread is a connected, ordered (not necessarily compact) topological semigroup. The following lemma is an extension of works of Aczél [1] and Tamari [9].

LEMMA 5.3. *A cancellative w-thread is (topologically) isomorphic with a subthread of R.*

Throughout this paper, isomorphisms are of the topological type, i.e., algebraic isomorphisms that are also homeomorphisms. Based on Corollary 5.2 and Lemma 5.3 is

THEOREM 5.4. *Let $(S, \circ)$ be a cancellative w-thread which either has an identity or is the result of removing the identity from a cancellative w-thread. If a map $F: S^2 \to R$ satisfies (2.4), then there exist a map $f: S \to \mathbb{R}$ and a map $\Psi: S^2 \to \mathbb{R}$, with Ψ satisfying (2.7) and (2.8), representing F through*

$$F(x, y) = f(x) + f(y) - f(x \circ y) + \Psi(x, y), \qquad \forall x, y \in S. \tag{2.6}$$

Proof. Let Φ be an isomorphism of $(S, \circ)$ onto a subthread $(T, +)$ of $(R, +)$, as provided by Lemma 5.3, and define $F^*: T^2 \to \mathbb{R}$ by

$$F^*(t, u) := F(\Phi^{-1}(t), \Phi^{-1}(u)), \qquad \forall t, u \in T. \tag{5.8}$$

Furthermore, because of our assumptions concerning the presence of an identity in $(S, \circ)$, T must be one of the real intervals $]-\infty, 0[$, $]-\infty, 0]$, $[0, \infty[$, $]0, \infty[$, or R. Hence, all of the hypotheses of Corollary 5.2 are satisfied, by taking $(A, \oplus)$ to be $(T, +)$, $(X, +)$ to be $(R, +)$, and F to be F^*. Therefore, F^* has the form

$$F^*(t, u) = f^*(t) + f^*(u) - f^*(t+u) + \Psi^*(t, u), \qquad \forall t, u \in T, \tag{5.9}$$

with Ψ^* a solution of (5.5), (5.6). Defining $f: S \to \mathbb{R}$, $\Psi: S^2 \to \mathbb{R}$ by

$$f(x) := f^*(\Phi(x)), \qquad \Psi(x, y) := \Psi^*(\Phi(x), \Phi(y)), \quad \forall x, y \in S,$$

we have (2.4) with Ψ satisfying (2.7) and (2.8), by (5.9), (5.8), (5.5), (5.6). This finishes the proof.

6. **Solution on near-threads.** We begin this section with some results from the structure theory of threads and w-threads, upon which we shall draw later in the section. The fundamental object is the *standard thread* (or I-semigroup), a thread in which the lower endpoint acts as a zero and the upper as an identity. Related to this is the *positive thread*, which is a w-thread with a zero as its least element, an identity, and no largest element.

Encompassing both of the above kinds of semigroups is the *globally idempotent thread*, which is a w-thread $(S, \circ)$ with a zero as least element and the property $S \circ S = S$. Indeed, Storey [8] has shown the following.

LEMMA 6.1. *A globally idempotent thread must be one of the following: a standard thread, a standard thread with identity removed, or a positive thread.*

Moreover, Mostert and Shields [5] have revealed the structure of positive threads through

LEMMA 6.2. *In a positive thread, there is a largest idempotent M less than the identity, $[0, M]$ is a standard thread, $\{s \mid M<s\}$ is isomorphic with the group of positive reals, and $s\circ t=t\circ s=s$ if $s\leq M\leq t$.*

Define the *unit thread* J_1 and the *nil thread* J_2 by

$$J_1:=([0,1],\cdot), \qquad J_2:=([\tfrac{1}{2},1],*),$$

where $\cdot$ denotes the usual multiplication of real numbers, and $*$ the operation defined by $s*t:=\max(\frac{1}{2}, s\cdot t)$. Mostert and Shields [6] proved

LEMMA 6.3. *The set Λ of idempotents of a standard thread $(S, \circ)$ is closed, and so its complement P in S is a disjoint union of open intervals. The closure of each of these intervals is isomorphic to either J_1 or J_2, and multiplication of elements x, y not in the same component of P is $x\circ y=\min(x, y)$.*

Let us call a semigroup obtained by removing the zero from a globally idempotent thread a *near-thread*. We now turn our attention (and Lemmas 6.1–6.3) to the solution of (2.4) on near-threads.

We need one more preliminary result.

LEMMA 6.4. *Let b be an idemportent of a w-thread, and let $\langle a, b], [v, c\rangle$ be two components (either closed or half-open) of this w-thread, so that $x\circ y=y\circ x=x$ if $x\in\langle a,b]$, $y\in[b,c\rangle$. Furthermore, suppose that $F:\langle a,c\rangle^2\to\mathbb{R}$ satisfies (2.4) and is partially determined by*

$$F(x,y)=\begin{cases} f_1(x)+f_1(y)-f_1(x\circ y), & \forall(x,y)\in[b,c\rangle^2 \\ f_2(x)+f_2(y)-f_2(x\circ y)+\Psi_2(x,y), & \forall(x,y)\in\langle a,b]^2, \end{cases} \tag{6.1}$$

for $f_1:[b,c\rangle\to\mathbb{R}$, $f_2:\langle a,b]\to\mathbb{R}$, and $\Psi_2:\langle a,b]^2\to\mathbb{R}$ satisfying (2.7), (2.8). Then there exist a map $\Psi:\langle a,c\rangle^2\to\mathbb{R}$, satisfying (2.7) and (2.8), and a map $f:\langle a,c\rangle\to\mathbb{R}$ representing F through (2.6) on $\langle a,c\rangle^2$.

Proof. Define new maps $f:\langle a,c\rangle\to\mathbb{R}$, $\Psi:\langle a,c\rangle^2\to\mathbb{R}$ by

$$f(x):=\begin{cases} f_1(x), & \forall x\in[b,c\rangle, \\ f_2(x), & \forall x\in\langle a,b], \end{cases} \tag{6.2}$$

$$\Psi(x,y):=\begin{cases} 0, & \forall(x,y)\in\langle a,c\rangle^2\setminus\langle a,b]^2, \\ \Psi_2(x,y), & \forall(x,y)\in\langle a,b]^2. \end{cases} \tag{6.3}$$

To verify that $f(b)$ is well defined, we note that $f_1(b)=F(b,b)=f_2(b)+\Psi_2(b,b)$. But $\psi_2(b,b)=0$, as a consequence of (2.7).

It is clear, by (6.1), (6.2), (6.3), that F is represented by (2.6) on $\langle a, b]^2 \cup [b, c\rangle^2$. We proceed to show that F also has this form on $\langle a, b]\times[b, c\rangle$. (The proof is similar for $(x, y)\in[b, c\rangle\times\langle a, b]$.) Indeed, with $x\in\langle a, b]$, $y=b$, $z\in [b, c\rangle$, (2.4) gives

$$F(x, z)=F(b, z), \qquad \forall x\in\langle a, b], \qquad z\in[b, c\rangle. \tag{6.4}$$

Thus, since $x\circ z=x$ when $x\in\langle a, b]$, $z\in[b, c\rangle$, we obtain

$$\begin{aligned} F(x, z)=F(b, z)&=f_1(b)+f_1(z)-f_1(b\circ z)\\ &=f_1(z)\\ &=f(z)\\ &=f(x)+f(z)-f(x\circ z), \end{aligned}$$

by (6.4), (6.1), and (6.2). This establishes (2.6) for $(x, z)\in\langle a, b]\times[b, c\rangle$, because (6.3) shows that

$$\Psi(x, z)=0 \qquad x\in\langle a, b], \qquad z\in[b, c\rangle.$$

Finally, it is easy to check that ψ satisfies (2.7) and (2.8). One uses the fact that $x, y\in\langle a, b]$ (resp. $[b, c\rangle$) implies $x\circ y\in\langle a, b]$ (resp. $[b, c\rangle$), which is implicit in (6.1) since f_1 is defined only on $[b, c\rangle$ and f_2 on $\langle a, b]$.

We are now ready to give the solution of (2.4) on an arbitrary near-thread. We have

THEOREM 6.5. *Let $(S, \circ)$ be a near-thread, and let $F: S^2\to\mathbb{R}$ be a solution of* (2.4). *Then F has the form* (2.6) *for some map $f: S\to R$ and a map $\Psi: S^2\to\mathbb{R}$ satisfying* (2.7), (2.8). *Furthermore, if the set $\Lambda:=\{s\in S \mid s\circ s=s\}$ is not empty, then*

$$\Psi(x, y)=0, \qquad \forall(x, y)\in S^2\setminus]0, \lambda]^2, \tag{6.5}$$

where $\lambda:=\operatorname{Inf}\Lambda$.

Proof. By definition of a near-thread, we can adjoin a zero 0 to S and extend the definition of $\circ$ in the obvious way, making $(S\cup\{0\}, \circ)$ a globally idempotent thread. Lemma 6.1 implies that $S\cup\{0\}$ is either a standard thread (possibly with identity removed) or a positive thread. If it is a standard thread without identity, we consider $S\cup\{0\}\cup\{e\}$, a standard thread. At this point, we divide the proof into two cases.

CASE 1. On the one hand, suppose $S\cup\{0\}$ $(\cup\{e\})$ is a standard thread, and suppose Λ, as defined in the statement of the theorem, is empty. Then the set P in Lemma 6.3 is S itself, hence the closure $\bar{S}$ is isomorphic to either J_1 or J_2. But S is a semigroup, which is not the case for J_2 with the (zero) point $\frac{1}{2}$ removed. Therefore, $\bar{S}$ is isomorphic to J_1. Now S is a cancellative w-thread (possibly with identity removed), and Theorem 5.4 gives the representation (2.6) for F.

If $\Lambda \neq \emptyset$, let λ be as defined in the statement of the theorem. If $\lambda > 0$, then we divide $S \cup \{0\}$ into two parts, $S_\lambda := \{s \in S \mid \lambda \leq s\}$ and $(S \cup \{0\}) \setminus S_\lambda$. By Lemma 6.3, S_λ is a semigroup with zero (λ), and the closure of $S \setminus S_\lambda$ is isomorphic with either J_1 or J_2. For the reason given in the preceding paragraph, the closure of $S \setminus S_\lambda$ must be isomorphic with J_1. Thus F has the form (3.1) on S_λ, by Theorem 4.1, while F has the representation (2.6) on $S \setminus S_\lambda$, by Theorem 5.4. Applying Lemma 6.4 with $b = \lambda$, we have the desired representation (2.6) for F on all of S, with Ψ satisfying (2.7) and (2.8). In addition to this, (6.3) shows that (6.5) holds.

Still with $S \cup \{0\}$ ($\cup \{e\}$) a standard thread, $\Lambda \neq \emptyset$, suppose $\lambda = 0$, and let (λ_i) be a sequence of idempotents in S decreasing to 0. For any n, $\{s \in S \mid \lambda_n \leq s\}$ is a semigroup with zero (λ_n), so Theorem 3.1 gives

$$F(x, y) = f_n^*(x) + f_n^*(y) - f_n^*(x \circ y), \qquad \forall x, y \geq \lambda_n, \tag{6.6}$$

for some map $f_n^* : \{s \mid \lambda_n \leq s\} \to \mathbb{R}$. Furthermore, $[\lambda_{n+1}, \lambda_n]$ is a semigroup with zero (λ_{n+1}), hence

$$F(x, y) = f_{n+1}(x) + f_{n+1}(y) - f_{n+1}(x \circ y), \qquad \forall x, y \in [\lambda_{n+1}, \lambda_n],$$

for some map $f_{n+1} : [\lambda_{n+1}, \lambda_n] \to \mathbb{R}$. Thus, by Lemma 6.4, there is an extension $f_{n+1}^* : \{s \mid \lambda_{n+1} \leq s\} \to \mathbb{R}$, identical with f_n^* on $\{s \mid \lambda_n \leq s\}$ and identical with f_{n+1} on $[\lambda_{n+1}, \lambda_n]$, such that F is given by

$$F(x, y) = f_{n+1}^*(x) + f_{n+1}^*(y) - f_{n+1}^*(x \circ y), \qquad \forall x, y \geq \lambda_{n+1}.$$

Now the set of pairs (S_n, f_n^*) forms a bounded chain, where $S_n = \{s \in S \mid \lambda_n \leq s\}$ and $f_n^* : S_n \to \mathbb{R}$ represents F through (6.6). By Zorn's Lemma, this chain has a maximal element (S^*, f); furthermore, $S^* = S$ since for any $s \in S$, there exists an N such that $s > \lambda_N$. This establishes (2.6) on all of S.

CASE 2. On the other hand, suppose $S \cup \{0\}$ is a positive thread. If the $M \in S$ defined by Lemma 6.2 is 0, then that lemma states that S is isomorphic with the group of positive reals. Now, by Corollary 5.2 or Theorem 5.4, we have the representation (2.6).

Alternatively, and finally, suppose $M > 0$. Then $M \in \Lambda$, so $\Lambda \neq \emptyset$. Also, $[0, M]$ is a standard thread, by Lemma 6.2. Let $\lambda = \operatorname{Inf} \Lambda$ again, and proceed exactly as in the case $\Lambda \neq \emptyset$ above, when $S \cup \{0\}$ ($\cup \{e\}$) was a standard thread. We again get (2.6), and Theorem 6.5 is proved.

7. **Summary, consequences, and remarks.** To summarize the results thus far, let X denote the class of all commutative monoids belonging to one of the following classes: idempotent commutative monoids, threads, monoids with zero, ordered commutative groups, cancellative w-threads, and near-threads. By Lemma 2.3 and the results of sections 3–6, we have proven

https://doi.org/10.4153/CMB-1979-057-3 Published online by Cambridge University Press

THEOREM 7.1. *Let* $\{\mu_n\}$ *be a* (1.1) *branching measure of information on a monoid* $(S, \circ)$ *from class X. Then* μ_n $(n=3,4,\ldots)$ *has a representation in the form* (2.9) *for all* $(s_1, s_2, \ldots, s_n) \in S^n$, *with maps* $k_{n,i}: S \to \mathbb{R}$ *satisfying* $k_{n,i}(e)=0$ $(1 \le i \le n)$ *and with a map* $\Psi_n: S^2 \to \mathbb{R}$ *satisfying* (2.7), (2.8).

Actually, all of our results hold in a slightly more general setting, as indicated in the next remark.

REMARK 7.2. We can interpret (1.1) for $\mu_n: S^n \to G$ and $\Delta_{n,i}: S^2 \to G$ and for any binary operation + on a set G under which $(G, +)$ becomes a divisible abelian group. Under this interpretation, Theorem 7.1 and all previous results still stand (with appropriate changes of $\mathbb{R}$ to G). This is explicit in Theorem 5.1 and implicit in the methods of proof elsewhere.

Two other properties for measures of information are *expansibility* and *symmetry* (cf. [11] for classical measures), defined respectively by

$$\mu_{n+1}(s_1, s_2, \ldots, s_n, e) = \mu_n(s_1, s_2, \ldots, s_n), \qquad n=1,2,\ldots, \tag{7.1}$$

$$\mu_n(s_1, s_2, \ldots, s_n) = \mu_n(s_{\pi(1)}, s_{\pi(2)}, \ldots, s_{\pi(n)}), \tag{7.2}$$

for all $(s_1, s_2, \ldots, s_n) \in S^n$ and permutations π on $(1, 2, \ldots, n)$.

COROLLARY 7.3. *Let* $\{\mu_n\}$ *be an information measure of the form* (2.9) *with maps* $k_{n,i}: S \to \mathbb{R}$ *satisfying*

$$k_{n,i}(e)=0, \qquad 1 \le i \le n, \tag{7.3}$$

and maps $\psi_n: S^2 \to \mathbb{R}$ *satisfying* (2.7), (2.8). *If* $\{\mu_n\}$ *is also* (7.1) *expansible, then* μ_n *has a representation in the form*

$$\mu_n(s_1, s_2, \ldots, s_n) = f_0(s_1 \circ s_2 \circ \cdots \circ s_n) + \sum_{i=1}^{n} f_i(s_i) + \sum_{i=1}^{n-1} \psi(s_i, s_{i+1} \circ s_{i+2} \circ \cdots \circ s_n), \tag{7.4}$$

where f_i $(i \ge 1)$ *satisfies* $f_i(e)=0$, *and* ψ *satisfies* (2.7) *and* (2.8).

Proof. If $n \ge 3$, then (2.9) enables us to write (7.1) as

$$\begin{aligned} k_{n,0}(s_1 \circ s_2 \circ \cdots \circ s_n) + \sum_{i=1}^{n} k_{n,i}(s_i) + \sum_{i=1}^{n-1} \psi_n(s_i, s_{i+1} \circ s_{i+2} \circ \cdots \circ s_n) \\ = k_{n+1,0}(s_1 \circ s_2 \circ \cdots \circ s_n) + \sum_{i=1}^{n} k_{n+1,i}(s_i) \\ + \sum_{i=1}^{n-1} \psi_{n+1}(s_i, s_{i+1} \circ s_{i+2} \circ \cdots \circ s_n), \end{aligned}$$

where we have used (7.3) and

$$\psi_n(x, e) = 0, \qquad \forall x \in S, \tag{7.5}$$

https://doi.org/10.4153/CMB-1979-057-3 Published online by Cambridge University Press

which follows from (2.11) by (2.18) and (2.7). Hence $(k_{n+1,0}, k_{n+1,1}, \ldots, k_{n+1,n}; \psi_{n+1})$ can be replaced by $(k_{n,0}, k_{n,1}, \ldots, k_{n,n}; \psi_n)$. Now, defining

$$f_i(x) := \begin{cases} k_{3,i}(x), & \forall x \in S, \quad 0 \le i \le 3, \\ k_{i,i}(x), & \forall x \in S, \quad i \ge 3, \end{cases} \tag{7.6}$$

$$\psi(x, y) := \psi_3(x, y), \qquad \forall x, y \in S, \tag{7.7}$$

we obtain (7.4) for all $n \ge 3$.

Similarly, (7.1) allows us to extend (7.4) to μ_2 and μ_1 through (7.6) and (7.7), which completes the proof.

COROLLARY 7.4. *Let $\{\mu_n\}$ be an information measure of the form* (2.9) *with $k_{n,i}$ satisfying* (7.3) *and ψ_n satisfying* (2.7) *and* (2.8). *If $\{\mu_n\}$ is also* (7.2) *symmetric, then μ_n has a representation in the form*

$$\mu_n(s_1, s_2, \ldots, s_n) = k_{n,0}(s_1 \circ s_2 \circ \cdots \circ s_n) + \sum_{i=1}^{n} f_n(s_i) \tag{7.8}$$

for $f_n : S \to \mathbb{R}$ satisfying $f_n(e) = 0$.

Proof. Using (2.9) in (7.2) with the permutation which interchanges i with $i+1$ $(1 \le i \le n-1)$, we get

$$k_{n,i}(s_i) + k_{n,i+1}(s_{i+1}) + \psi_n(s_i, s_{i+1}) = k_{n,i}(s_{i+1}) + k_{n,i+1}(s_i) + \psi_n(s_{i+1}, s_i), \tag{7.9}$$

where we have used the fact that ψ_n satisfies (2.7) and (2.8). Setting $s_{i+1} = e$ in (7.9) and using (7.3), (7.5), and (2.11) (with (2.18)), we obtain

$$k_{n,i}(s_i) = k_{n,i+1}(s_i), \qquad \forall s_i \in S. \tag{7.10}$$

Then (7.9) and (7.10) imply

$$\psi_n(s_i, s_{i+1}) = \psi_n(s_{i+1}, s_i), \qquad \forall s_i, s_{i+1} \in S,$$

which in turn yields

$$\psi_n(s, t) = 0, \qquad \forall s, t \in S, \tag{7.11}$$

by the (2.7) anti-symmetry. With $f_n : S \to \mathbb{R}$ defined by

$$f_n(x) := k_{n,1}(x), \qquad \forall x \in S,$$

(7.10) and (7.11) reduce (2.9) to (7.8), which was to be shown. Moreover, $f_n(e) = 0$ follows from (7.3), and the proof is finished.

Combining the two preceding results, we have

COROLLARY 7.5. *Let $\{\mu_n\}$ satisfy all hypotheses of Corollary* 7.3 *and the* (7.2) *symmetry. Then μ_n can be represented in the form*

$$\mu_n(s_1, s_2, \ldots, s_n) = f_0(s_1 \circ s_2 \circ \cdots \circ s_n) + \sum_{i=1}^{n} f(s_i)$$

for all $(s_1, s_2, \ldots, s_n) \in S^n$ $(n = 1, 2, \ldots)$, by two maps $f_0, f : S \to \mathbb{R}$ with $f(e) = 0$.

https://doi.org/10.4153/CMB-1979-057-3 Published online by Cambridge University Press

8. **A generalization of (2.4).** We show that the equation

$$L(a, b)+G(a\circ b, c)=H(a, b\circ c)+K(b, c), \qquad \forall a, b, c\in A, \tag{8.1}$$

where A is any set with an identity e under a binary operation $\circ$, and L, G, H, K map A^2 into an abelian group X (whose operation we denote $+$), can be reduced to (2.4).

Setting $b=e$ in (8.1), we get

$$H(a, c)=G(a, c)+f(a)-g(c), \tag{8.2}$$

where $f(a):=L(a, e)$ and $g(c):=K(e, c)$. With $a=e$, (8.1) gives

$$K(b, c)=G(b, c)+h(b)-G(e, b\circ c)-f(e)+g(b\circ c), \tag{8.3}$$

by defining $h(b):=L(e, b)$ and using (8.2). Putting $c=e$ in (8.1) and using (8.2) and (8.3), we obtain

$$L(a, b)=G(a, b)+f(a)+G(b, e)+h(b)-G(e, b)-f(e)-G(a\circ b, e) \tag{8.4}$$

after simplification.

Now, substituting (8.2), (8.3), and (8.4) back into (8.1) and simplifying, we have

$$\begin{aligned} G(a, b)+G(b, e)-G(e, b)-G(a\circ b, e)+G(a\circ b, c)& \\ =G(a, b\circ c)+G(b, c)-G(e, b\circ c).& \end{aligned} \tag{8.5}$$

Subtracting $G(a, e)+G(e, c)$ from both sides of (8.5), and defining $F: A^2\to X$ by

$$F(a, b):=G(a, b)-G(a, e)-G(e, b), \qquad \forall a, b\in A \tag{8.6}$$

we find that F satisfies (2.4). So we can use this to express G, H, K, L in terms of an arbitrary solution of (2.4) and some arbitrary one-place functions. Indeed, defining $k, \ell: A\to X$ by

$$k(a):=G(a, e), \qquad \ell(a):=G(e, a) \qquad \forall a\in A, \tag{8.7}$$

we can rewrite (8.6) as

$$G(a, b)=F(a, b)+k(a)+\ell(b), \qquad \forall a, b\in A. \tag{8.8}$$

Furthermore, define $m, n, p: A\to X$ by

$$\begin{aligned} m(a)&:=G(a, e)+f(a),\\ n(a)&:=G(a, e)+h(a)-f(e),\\ p(a)&:=G(e, a)-g(a), \qquad \forall a\in A. \end{aligned} \tag{8.9}$$

By (8.9), (8.6), and (8.2), H has the form

$$H(a, c)=F(a, c)+m(a)+p(c), \qquad \forall a, c\in A. \tag{8.10}$$

https://doi.org/10.4153/CMB-1979-057-3 Published online by Cambridge University Press

And by (8.9), (8.7), (8.6), (8.3), and (8.4), K and L can be represented in the respective forms

$$(8.11) \qquad K(b, c) = F(b, c) + n(b) + \ell(c) - p(b \circ c), \qquad \forall b, c \in A,$$

$$(8.12) \qquad L(a, b) = F(a, b) + m(a) + n(b) - k(a \circ b), \qquad \forall a, b \in A.$$

Conversely, substituting (8.8), (8.10), (8.11), (8.12) for G, H, K, L in (8.1), and taking into account the fact that F is a solution of (2.4), we immediately verify that these functions constitute a solution of (8.1). Thus we have proved

THEOREM 8.1. *Let* $(A, \circ)$ *be a set with a binary operation and an identity,* $(X, +)$ *an abelian group, and* G, H, K, L *maps from* $A \times A$ *into* X. *Then* G, H, K, L *satisfy* (8.1) *if, and only if, there exist functions* k, ℓ, m, n, $p : A \to X$ *and a map* $F : A^2 \to X$ *satisfying* (2.4), *such that* G, H, K, L *are given by* (8.8), (8.10), (8.11), (8.12), *respectively.*

REFERENCES

1. J. Aczél, *Sur les opérations définies pour nombres réels*, Bull. Soc. Math. Franc. **76** (1948), 59–64.
2. W. M. Faucett, *Topological semigroups and continua with cut points*, Proc. Amer. Math. Soc. **6** (1955), 748–756.
3. B. Jessen, *The algebra of polyhedra and the Dehn-Sydler Theorem*, Math. Scand. **22** (1968), 241–256.
4. B. Jessen, J. Karpf, and A. Thorup, *Some functional equations in groups and rings*, Math. Scand. **22** (1968), 257–265.
5. P. S. Mostert and A. L. Shields, *On a class of semigroups on* E_n, Proc. Amer. Math. Soc. **7** (1956), 729–734.
6. P. S. Mostert and A. L. Shields, *On the structure of semigroups on a compact manifold with boundary*, Ann. Math. **65** (1957), 117–143.
7. K. Numakura, *On bicompact semigroups*, Math. J. Okayama Univ. **1** (1952), 99–108.
8. C. R. Storey, *The structure of threads*, Pacific J. Math. **10** (1960), 1429–1445.
9. D. Tamari, *Caractérisation des semi-groupes à un paramètre*, C. R. Acad. Sci. Paris **228** (1949), 1092–1094.
10. A. D. Wallace, *A note on mobs*, Anais de Academia Brasileira de Ciencias **24** (1952), 329–334.
11. J. Aczél and Z. Daróczy, *On Measures of Information and Their Characterizations* (1975), Academic Press, New York–London.

DEPARTMENT OF MATHEMATICS
TEXAS TECH. UNIVERSITY
LUBBOCK, TEXAS 79409

https://doi.org/10.4153/CMB-1979-057-3 Published online by Cambridge University Press