Representation for Measures of Information with the Branching Property*

C. T. Ng

Faculty of Mathematics, University of Waterloo, Ontario, Canada

The representation for measures of information which are symmetric, expansible, and have the branching property in the form of a sum is provided. This class of measures includes, in particular, Shannon’s entropy, entropies of degree $\beta$, Kullback’s directed divergence, and Kerridge’s inaccuracy. Rényi’s entropy and information gain of order $\beta$ are, however, excluded from this class. The proof is based on an algebraic theorem concerning the representation of a two-place function by the superposition of a one-place function.

1. INTRODUCTION

Let $P = (p_1, p_2, \ldots, p_n)$ and $Q = (q_1, q_2, \ldots, q_n)$ with $p_i, q_i \geq 0$ and $\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i = 1$ be complete finite discrete probability distributions of arbitrary length. A mapping $I$, also called a measure of information, of the set of all such pairs $(P; Q)$ into the reals $\mathbb{R}$ will be considered under the following hypotheses.

(a) Symmetry. For all $P, Q$ and permutations $\alpha$ on $\{1, 2, \ldots, n\}$ we have

$$I(p_1, p_2, \ldots, p_n; q_1, q_2, \ldots, q_n) = I(p_{\alpha(1)}, p_{\alpha(2)}, \ldots, p_{\alpha(n)}; q_{\alpha(1)}, q_{\alpha(2)}, \ldots, q_{\alpha(n)}).$$

(b) Expansibility. For all $P, Q$

$$I(p_1, p_2, \ldots, p_n; q_1, q_2, \ldots, q_n) = I(p_1, p_2, \ldots, p_n, 0; q_1, q_2, \ldots, q_n, 0).$$

(c) Branching. For all $P, Q$ the difference between $I(p_1, p_2, \ldots, p_n; q_1, q_2, \ldots, q_n)$ and $I(p_1 + p_2, p_3, \ldots, p_n; q_1 + q_2, q_3, \ldots, q_n)$ depends only on $p_1, p_2, q_1, q_2,$ and $n (\geq 2)$. Thus, there exist functions $A_n$ such that

$$I(p_1, p_2, \ldots, p_n; q_1, q_2, \ldots, q_n) = I(p_1 + p_2, p_3, \ldots, p_n; q_1 + q_2, q_3, \ldots, q_n) + A_n(p_1, p_2; q_1, q_2).$$

* This work is supported by Canadian NRC under Grant No. A8212.
Measures of information such as Kullback's directed divergence

\[(P; Q) \mapsto \sum_{i=1}^{n} p_i \log_2 (p_i q_i^{-1})\]

and Kerridge's inaccuracy \((P; Q) \mapsto -\sum_{i=1}^{n} p_i \log q_i\) are examples of such mappings. The class of mappings \(I\) satisfying the above hypotheses is quite large. We shall give in Section 3 the proof of our main result:

**Theorem 1.1.** If a mapping \(I\) of all pairs \((P, Q)\) of complete finite discrete probability distributions into the reals \(\mathbb{R}\) satisfies (a) symmetry, (b) expansibility,

and (c) branching, then there exists a mapping \(f: [0, 1] \times [0, 1] \to \mathbb{R}\) with \(f(0; 0) = 0\) such that \(I\) is represented by

\[I(p_1, p_2, \ldots, p_n; q_1, q_2, \ldots, q_n) = I(1; 1) - f(1; 1) + \sum_{i=1}^{n} f(p_i; q_i)\]

for all \((P; Q)\). Conversely, if \(I\) can be so represented by an \(f\) with \(f(0; 0) = 0\), then \(I\) fulfills (a), (b), and (c).

**Remark 1.1.** We have just given the required terminology and our representation theorem for measures of information defined for pairs \((P, Q)\). We can extend the notion of being symmetric, expandible, and having the branching property to mappings defined for \(m\)-tuples of complete finite discrete probability distributions in a straightforward manner. The representation theorem can be given for a general \(m = 1, 2, \ldots\), without altering our argument used in Section 3 for the proof of Theorem 1.1. The algebraic theorem we give in Section 2 is for a general \(m\). We pick \(m = 2\) for our display so as to make the notations in writing simpler and for no reason other than this.

Hence, for \(m = 1\) the representation for \(I\) is \(I(1) - f(1) + \sum_{i=1}^{n} f(p_i)\). The entropy of Shannon (1948) defined by \((p_1, p_2, \ldots, p_n) \mapsto -\sum_{i=1}^{n} p_i \log_2 p_i\) and entropies of degree \(\beta (\neq 1)\) defined by

\[(p_1, p_2, \ldots, p_n) \mapsto (2^{1-\beta} - 1)^{-1} \left(-1 + \sum_{i=1}^{n} p_i^\beta\right)\]

are examples of such measures.

Similarly, for \(m = 3\) the representation for \(I\) is

\[I(1; 1; 1) - f(1; 1; 1) + \sum_{i=1}^{n} f(p_i; q_i; r_i).\]
The measure generalized directed divergence defined by

\[(p_1, p_2, \ldots, p_n; q_1, q_2, \ldots, q_n; r_1, r_2, \ldots, r_n) \rightarrow \sum_{i=1}^{n} p_i \log_2(q_i/r_i^{-1})\]

is an example.

Remark 1.2. There have been various characterizations of the entropies we have mentioned. Aczél (1969) has summarized most of the relevant works. For entropies with the branching property on spaces without probability, we refer to Forte and Ng (1973).

2. Functional Equations and a Representation Theorem

We shall consider the Euclidean m-space \(\mathbb{R}^m\) as a linear space over the rationals \(\mathbb{Q}\) under its usual structures. We shall write \(J := [0, 1]^m\) for the unit rectangle in \(\mathbb{R}^m\).

**Theorem 2.1.** Let \(g: J \rightarrow \mathbb{R}\) be a mapping of the unit rectangle in \(\mathbb{R}^m\) into \(\mathbb{R}\) satisfying the boundary condition \(g(0) = 0\). Then the function \(G\) defined by

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

for all \(x, y \in J\) with \(x + y \in J\) satisfies the functional equation

\[G(x, y) = G(y, x),\]

\[G(x, 0) = 0,\]  

\[G(x, y) + G(x + y, z) = G(x, y + z) + G(y, z),\]

for all \(x, y, z \in J\) with \(x + y + z \in J\). Conversely, if a function \(G\) satisfies Eqs. (2.2), then there exists a function \(g: J \rightarrow \mathbb{R}\) with \(g(0) = 0\) representing \(G\) through Eq. (2.1).

**Proof.** The first part is straightforward. For the converse we construct \(g\) by transfinite induction through successive steps.

Step 1. The operation \(\oplus\) defined by

\[(x, u) \oplus (y, v) = (x + y, u + v - G(x, y))\]
for all \( x, y \in J \) with \( x + y \in J, u, v \in R \) has the following properties:

\[ \oplus 1. \ (x, u) \oplus (y, v) = (y, v) \oplus (x, u) \]

for all \( x, y \in J \) with \( x + y \in J \).

\[ \oplus 2. \ [(x, u) \oplus (y, v)] \oplus (z, w) = (x, u) \oplus [(y, v) \oplus (z, w)] \]

for all \( x, y, z \in J \) with \( x + y + z \in J \).

\[ \oplus 3. \ (x, u) \oplus (y, v) = (x, u) \oplus (y', v') \implies (y, v) = (y', v'). \]

\[ \oplus 4. \ (a) \ n(x, u) := (x, u) \oplus (x, u) \oplus \cdots \oplus (x, u) \ (n\text{-fold}) \]

is definable for each \( x \in J, n \in N^+ = \{1, 2, 3, \ldots\} \) such that \( nx \in J \). We also define \( 0(x, u) := (0, 0) \). With this definition we have the identities

\[
(m + n)(x, u) = m(x, u) \oplus n(x, u), \\
(mn)(x, u) = m(n(x, u)),
\]

\[
m[(x, u) \oplus (y, v)] = m(x, u) \oplus m(y, v),
\]

whenever both sides are defined.

(b) \( n^{-1}(x, u) \) is definable for each \( x \in J, u \in R \), and \( n \in N^+ \) as the unique point such that

\[
n[n^{-1}(x, u)] = (x, u).
\]

(c) Furthermore,

\[
(mk)[(nk)^{-1}(x, u)] = m[n^{-1}(x, u)]
\]

for each \( x \in J, m \in N, n, k \in N^+ \) with \( mn^{-1}x \in J \). Thus, for each nonnegative rational \( r := mn^{-1} \) with \( m \in N \) and \( n \in N^+ \) and for each \( (x, u) \) with \( x \in J \) and \( rx \in J \), we may define

\[
r(x, u) := m[n^{-1}(x, u)].
\]

(d) \( (r_1 + r_2)(x, u) = r_1(x, u) \oplus r_2(x, u) \) for all \( x \in J \), nonnegative \( r_1, r_2 \in Q \) with \( r_1x + r_2x \in J \). \( (r_1r_2)(x, u) = r_1(r_2(x, u)) \) for all \( x \in J \), nonnegative rationals \( r_1, r_2 \) with \( r_2x, r_1r_2x \in J \).

(e) \( r[(x, u) \oplus (y, v)] = r(x, u) \oplus r(y, v) \) for all \( x, y \in J \), nonnegative \( r \in Q \) with \( x + y, rx + ry \in J \).

The proofs of \( \oplus 1-\oplus 3 \) and \( \oplus 4(a) \) are straightforward verifications using Eqs. (2.2) for \( G \).
For $\oplus 4(b)$ we have

$$n^{-1}(x, u) = \left[ n^{-1}x, n^{-1}u + n^{-1} \sum_{i=1}^{n-1} G(n^{-1}x, in^{-1}x) \right]$$

by direct computation. We observe that the additive groups of Euclidean spaces are divisible and can be regarded as $\mathbb{Q}$-linear spaces.

For $\oplus 4(c)$ we have

$$(mk)[(nk)^{-1}(x, u)] = m[n^{-1}(x, u)]$$

iff

$$m[k[(nk)^{-1}(x, u)]] = m[n^{-1}(x, u)]$$

iff

$$k[(nk)^{-1}(x, u)] = n^{-1}(x, u)$$

iff

$$n[k[(nk)^{-1}(x, u)]] = (x, u)$$

iff

$$(nk)[(nk)^{-1}(x, u)] = (x, u)$$

iff

$$(x, u) = (x, u).$$

For $\oplus 4(d)$ we let $r_1 = m_1n_1^{-1}$, $r_2 = m_2n_2^{-1}$, $x \in J$ with $m_i \in N$, $n_i \in N^+$, $r_1x + r_2x \in J$. Then

$$(r_1 + r_2)(x, u) = (m_1n_1^{-1} + m_2n_2^{-1})(x, u)$$

$$= [(m_1n_2 + m_2n_1)(n_1n_2)^{-1}](x, u)$$

$$= (m_1n_2 + m_2n_1)[(n_1n_2)^{-1}(x, u)]$$

$$= (m_1n_2)[(n_1n_2)^{-1}(x, u)] \oplus (m_2n_1)[(n_1n_2)^{-1}(x, u)]$$

$$= r_1(x, u) \oplus r_2(x, u).$$

Let $r_1 = m_1n_1^{-1}$, $r_2 = m_2n_2^{-1}$, $x \in J$ with $m_i \in N$, $n_i \in N^+$, $r_2x + r_1r_2x \in J$. Then

$$(r_1r_2)(x, u) = (m_1m_2)[(n_1n_2)^{-1}(x, u)]$$

$$= (m_1m_2)((n_1n_2)^{-1}((n_1n_2)^{-1}(x, u)]$$

$$= r_1(r_2(x, u)).$$

For $\oplus 4(e)$ we have

$$(mn^{-1})[(x, u) \oplus (y, v)] = (mn^{-1})(x, u) \oplus mn^{-1}(y, v)$$

iff

$$m(n^{-1}[(x, u) \oplus (y, v)]) = m(n^{-1}(x, u)) \oplus m(n^{-1}(y, v))$$

iff

$$m(n^{-1}[(x, u) \oplus (y, v)]) = m(n^{-1}(x, u)) \oplus n^{-1}(y, v))$$

iff

$$n^{-1}[(x, u) \oplus (y, v)] = n^{-1}(x, u) \oplus n^{-1}(y, v)$$

iff

$$(x, u) \oplus (y, v) = n[n^{-1}(x, u) \oplus n^{-1}(y, v)]$$

iff

$$(x, u) \oplus (y, v) = n[n^{-1}(x, u)] \oplus n[n^{-1}(y, v)]$$

iff

$$(x, u) \oplus (y, v) = (x, u) \oplus (y, v).$$
Step 2. The class \( \mathcal{A} = \{(S, g)\} \) of pairs \((S, g)\) satisfying

\begin{itemize}
  \item \(\emptyset \subseteq S \subseteq J\),
  \item \(s \in S\) implies \(n^{-1}s \in S\) for all \(n \in \mathbb{N}^+\),
  \item \(s_1, s_2 \in S\) and \(s_1 + s_2 \in J\) imply \(s_1 + s_2 \in S\),
  \item \(g : S \to \mathbb{R}\) satisfies \([s_1 + s_2, g(s_1 + s_2)] = [s_1, g(s_1)] \oplus [s_2, g(s_2)]\)
\end{itemize}

for all \(s_1, s_2 \in S\) with \(s_1 + s_2 \in S\) is nonempty and complete under the natural partial ordering \(\subseteq\) defined as

\[(S, g) \subseteq (T, h) \iff S \subseteq T \text{ and } h | S = g,\]

where \(h | S\) denotes the restriction of \(h\) to the set \(S\).

In fact, \((\emptyset, \text{zero map})\) is in \(\mathcal{A}\) and \(\mathcal{A}\) is nonempty. If \(\mathcal{B} = \{(S_i, g_i) | i \in I\}\) is a nonempty chain in \(\mathcal{A}\), then \((S, g), \text{ defined by } S = \bigcup\{S_i | i \in I\}\) and \(g(s) = g_i(s)\) if \(s \in S_i\), is again in \(\mathcal{A}\) and is an upper bound for \(\mathcal{B}\).

Step 3. By Zorn's lemma there exists \((S_0, g_0)\) which is a maximal element of \(\mathcal{A}\). We shall prove that \(S_0 = J\) and hence

\[[x + y, g_0(x + y)] = [x, g_0(x)] \oplus [y, g_0(y)]\]

for all \(x, y \in J\) with \(x + y \in J\). This is equivalent to the existence of \(g_0 : J \to \mathbb{R}\) such that Eq. (2.1) is satisfied by \(g_0\) and \(G\).

We suppose, if possible, that \(S_0 \subsetneq J\). We shall show that this leads to a contradiction to the maximality of \((S_0, g_0)\).

Case 1. Suppose there exists a point \(x_0 \in J \setminus S_0\) such that \(n^{-1}x_0 \notin S_0 - S_0\) for all \(n \in \mathbb{N}^+\).

In this case \(s_1 + m_1n_1^{-1}x_0 = s_2 + m_2n_2^{-1}x_0\) (where \(s_i \in S_0\), \(m_i \in \mathbb{N}\), \(n_i \in \mathbb{N}^+\), \(s_1 + m_1n_1^{-1}x_0 \in J\)) if \(s_1 = s_2\) and \(m_1n_1^{-1} = m_2n_2^{-1}\). For otherwise we may suppose \(s_1 > s_0\) so that

\[s_1 = s_2 + (m_2n_1 - m_1n_2)(n_1n_2)^{-1}x_0\]

with \(m_2n_1 - m_1n_2 > 0\). This implies

\[\begin{align*}
  (m_2n_1 - m_1n_2)^{-1}s_1 &= (m_2n_1 - m_1n_2)^{-1}s_2 + (n_1n_2)^{-1}x_0.
\end{align*}\]

Meanwhile, both \((m_2n_1 - m_1n_2)^{-1}s_1\) and \((m_2n_1 - m_1n_2)^{-1}s_2\) are in \(S_0\) by \(\mathcal{A}2\), and so \((n_1n_2)^{-1}x_0 \in S_0 - S_0\) is a contradiction to our assumption on \(x_0\).
Hence, on the set
\[ S_0 := \{ s + mn^{-1}x_0 \mid s \in S_0, m \in \mathbb{N}, n \in \mathbb{N}^+ \text{ with } s + mn^{-1}x_0 \in J \} \]
we may define a mapping \( g_0 \) by
\[
[s + mn^{-1}x_0, g_0(s + mn^{-1}x_0)] = [s, g_0(s)] \oplus mn^{-1}[x_0, g_0(x_0)],
\]
where \( g_0(x_0) \) is an arbitrary constant. The pair \((S_0, g_0)\) is in \( \mathcal{A} \), and \((\tilde{S}_0, \tilde{g}_0) \not\in (S_0, g_0)\) contradicts the maximality of \((S_0, g_0)\).

Hence, Case 1 cannot occur and so for every point \( x \in J \backslash S_0 \) there exists \( n \in \mathbb{N}^+ \) such that \( n^{-1}x \in S_0 \). On the other hand, this is also true for \( x \in S_0 \) as \( x = 0 + x \) while \( 0 \in S_0 \) as well. We carry on our discussion in Case 2.

**Case 2.** Suppose for each \( x \in J \) there exist \( s_1, s_2 \in S_0 \) and \( n \in \mathbb{N}^+ \) such that \( s_1 = s_2 + n^{-1}x \). In this case the mapping \( \tilde{g} : J \to \mathbb{R} \) defined by
\[
\tilde{g}(x) = \text{the unique point } a \in \mathbb{R} \text{ such that }
\[ [s_1, g_0(s_1)] = [s_2, g_0(s_2)] \oplus n^{-1}(x, a) \quad (2.3)\]
is well defined.

To prove the above assertion we suppose, for a given \( x \in J \), we can write
\[
s_1 = s_2 + n^{-1}x \quad (2.4)
\]
and
\[
t_1 = t_2 + m^{-1}x
\]
at the same time with \( s_i, t_i \in S_0 \). Let \( a \) and \( b \) be unique points such that
\[
[s_1, g_0(s_1)] = [s_2, g_0(s_2)] \oplus n^{-1}(x, a) \quad (2.5)
\]
and
\[
[t_1, g_0(t_1)] = [t_2, g_0(t_2)] \oplus m^{-1}(x, b). \quad (2.6)
\]
We must show that \( a = b \). For this purpose we consider the following identities from Eqs. (2.4):
\[
(3m)^{-1}s_1 = (3m)^{-1}s_2 + (3mn)^{-1}x
\]
and
\[
(3n)^{-1}t_1 = (3n)^{-1}t_2 + (3mn)^{-1}x,
\]
which yield \((3m)^{-1}s_2 + (3n)^{-1}t_1 = (3m)^{-1}s_1 + (3n)^{-1}t_2 \in J\) while each term is again in \(S_0\). Hence, by \(S4\) on \((S_0, g_0)\) we get

\[
[3^{-1}m^{-1}s_2, g_0(3^{-1}m^{-1}s_2)] \oplus [3^{-1}n^{-1}t_1, g_0(3^{-1}n^{-1}t_1)] = [3^{-1}m^{-1}s_1, g_0(3^{-1}m^{-1}s_1)] \oplus [3^{-1}n^{-1}t_2, g_0(3^{-1}n^{-1}t_2)].
\]  

(2.7)

We multiply (2.5) by \(3^{-1}m^{-1}\) and get

\[
3^{-1}m^{-1}[s_1, g_0(s_1)] = 3^{-1}m^{-1}[s_2, g_0(s_2)] \oplus 3^{-1}m^{-1}n^{-1}(x, a).
\]

Using \(S4\) essentially we get

\[
[3^{-1}m^{-1}s_1, g_0(3^{-1}m^{-1}s_1)] = [3^{-1}m^{-1}s_2, g_0(3^{-1}m^{-1}s_2)] \oplus 3^{-1}m^{-1}n^{-1}(x, a). \quad (2.8)
\]

Similarly, from (2.6) we get

\[
[3^{-1}n^{-1}t_1, g_0(3^{-1}n^{-1}t_1)] = [3^{-1}n^{-1}t_2, g_0(3^{-1}n^{-1}t_2)] \oplus 3^{-1}n^{-1}m^{-1}(x, b). \quad (2.9)
\]

Since \(3^{-1}m^{-1}s_1 + 3^{-1}n^{-1}t_2 + 3^{-1}m^{-1}n^{-1}x \in J\), we can cross add (2.8) and (2.9) to get

\[
[3^{-1}m^{-1}s_1, g_0(3^{-1}m^{-1}s_1)] \oplus [3^{-1}n^{-1}t_2, g_0(3^{-1}n^{-1}t_2)] \oplus 3^{-1}m^{-1}n^{-1}(x, b) = [3^{-1}m^{-1}s_2, g_0(3^{-1}m^{-1}s_2)] \oplus [3^{-1}n^{-1}t_1, g_0(3^{-1}n^{-1}t_1)] \oplus 3^{-1}m^{-1}n^{-1}(x, a).
\]

(2.10)

Compare (2.7) to (2.10); we can cancel common terms by \(\oplus 3\) and get

\[
3^{-1}m^{-1}n^{-1}(x, a) = 3^{-1}m^{-1}n^{-1}(x, b).
\]

Hence, \(a = b\) as we claimed.

To show that \((J, \tilde{g})\) is in \(\mathcal{A}\) we must prove the identity \([x + y, \tilde{g}(x + y)] = [x, \tilde{g}(x)] \oplus [y, \tilde{g}(y)]\) on \(J\). For this purpose suppose that

\[
s_1 = s_2 + n^{-1}x
\]

and

\[
t_1 = t_2 + m^{-1}y,
\]

where \(s_i, t_i \in S_0, \ m, n \in N^+.\) From these equations we get

\[
2^{-1}m^{-1}s_1 = 2^{-1}m^{-1}s_2 + 2^{-1}m^{-1}n^{-1}x, \\
2^{-1}n^{-1}t_1 = 2^{-1}n^{-1}t_2 + 2^{-1}m^{-1}n^{-1}y,
\]

C. T. NG
and
\[ 2^{-1}m^{-1}s_1 + 2^{-1}n^{-1}t_1 = (2^{-1}m^{-1}s_2 + 2^{-1}n^{-1}t_2) + 2^{-1}m^{-1}n^{-1}(x + y). \]

From the definition of \( \tilde{g} \), we get
\[
[2^{-1}m^{-1}s_1, g_0(2^{-1}m^{-1}s_1) ] = [2^{-1}m^{-1}s_2, g_0(2^{-1}m^{-1}s_2) ] \oplus 2^{-1}m^{-1}n^{-1}[x, \tilde{g}(x)],
\]
(2.11)

\[
[2^{-1}n^{-1}t_1, g_0(2^{-1}n^{-1}t_1) ] = [2^{-1}n^{-1}t_2, g_0(2^{-1}n^{-1}t_2) ] \oplus 2^{-1}m^{-1}n^{-1}[y, \tilde{g}(y)],
\]
(2.12)

and
\[
[2^{-1}m^{-1}s_1 + 2^{-1}n^{-1}t_1, g_0(2^{-1}m^{-1}s_1 + 2^{-1}n^{-1}t_1) ]
= [2^{-1}m^{-1}s_2 + 2^{-1}n^{-1}t_2, g_0(2^{-1}m^{-1}s_2 + 2^{-1}n^{-1}t_2) ]
\oplus 2^{-1}m^{-1}n^{-1}[x + y, \tilde{g}(x + y)],
\]
(2.13)

respectively. We can add (2.11) and (2.12) and use \( \otimes 4 \) on \((S_0, g_0)\) and get
\[
[2^{-1}m^{-1}s_1 + 2^{-1}n^{-1}t_1, g_0(2^{-1}m^{-1}s_1 + 2^{-1}n^{-1}t_1) ]
= [2^{-1}m^{-1}s_2 + 2^{-1}n^{-1}t_2, g_0(2^{-1}m^{-1}s_2 + 2^{-1}n^{-1}t_2) ]
\oplus 2^{-1}m^{-1}n^{-1}[x + y, \tilde{g}(x + y)].
\]

This equation, when compared to (2.13) and using \( \otimes 3 \), leads to
\[
2^{-1}m^{-1}n^{-1}[x + y, \tilde{g}(x + y)] = 2^{-1}m^{-1}n^{-1}[x, \tilde{g}(x)] \oplus 2^{-1}m^{-1}n^{-1}[y, \tilde{g}(y)],
\]
and hence
\[
[x + y, \tilde{g}(x + y)] = [x, \tilde{g}(x)] \oplus [y, \tilde{g}(y)]
\]
as desired.

The fact that \((f, \tilde{g}) \not\equiv (S_0, g_0)\) in Case 2 is again a contradiction to the maximality of \((S_0, g_0)\). Thus Case 2 cannot occur.

This completes the proof that \( S_0 = j \), and \( G \) is represented by \( g_0 \) through (2.1).

**Remark 2.1.** Theorem 2.1 remains true when we drop the boundary condition \( g(0) = 0 \) and the identity \( G(x, 0) = 0 \) for all \( x \in j \) at the same time. For a function \( G \) satisfying \( G(x, y) = G(y, x) \) and \( G(x, y) + G(x + y, z) = G(x, y + z) + G(y, z) \), a translation of \( G \) to \( \tilde{G} = G - G(0, 0) \) will satisfy all three identities in (2.2).

The interval \( j \subset \mathbb{R}^m \) can be replaced by some general ones and Theorem 2.1
remains true without altering the present proof. For example, we can replace $J$ by the positive cone of $\mathbb{R}^m$. We can also generalize $\mathbb{R}^m$ to more general $\mathbb{Q}$-linear spaces (or divisible Abelian groups) while taking an appropriate subset in place of $J$.

**Remark 2.2.** Theorem 2.1 is known when the domain of the functions $g$ and $G$ is the full $\mathbb{R}^m$ and $\mathbb{R}^m \times \mathbb{R}^m$, respectively. It is also known when the domain of $g$ and $G$ is the positive cone $C_+^m$ and $C_+^m \times C_+^m$, respectively. The restriction of the domains under consideration in Theorem 2.1 makes the proof difficult. We do not prove it by extensions of the functions and their equations. Our proof is constructive and yields the extendability of the equations as a consequence. Related literature can be found in the work of Jessen, Karft, and Thorup (1968).

Questions such as whether a continuous $G$ can be represented by a continuous $g$ can be answered by combining the present result and the work of Kemperman (1957).

3. **Proof of Theorem 1.1**

The converse part of the theorem is trivial.

Let $I$ be symmetric, expansible, and for each $n$ there exists $A_n$ such that

$$I(p_1, p_2, \ldots, p_n; q_1, q_2, \ldots, q_n) = I(p_1 + p_2, p_3, \ldots, p_n; q_1 + q_2, q_3, \ldots, q_n) + A_n(p_1, p_2; q_1, q_2). \quad (3.1)$$

Since $I$ is expansible, $A_n$ is independent of $n$ and we can write $A$ in place of $A_n$. We rewrite (3.1) as

$$I(p_1, p_2, \ldots, p_n; q_1, q_2, \ldots, q_n) = I(p_1 + p_2, p_3, \ldots, p_n; q_1 + q_2, q_3, \ldots, q_n) + A(p_1, p_2; q_1, q_2). \quad (3.2)$$

To analyze the structure of $A$, we consider $J = [0, 1]^{2} \subseteq \mathbb{R}^2$ and the function $G$ defined by

$$G(x, y) = A(p_1, p_2; q_1, q_2), \text{ where } x = (p_1, q_1) \text{ and } y = (p_2, q_2). \quad (3.3)$$

From the symmetry of $I$ we get $I(p_1, p_2, \ldots, p_n; q_1, q_2, \ldots, q_n) = I(p_2, p_1, \ldots, p_n; q_2, q_1, \ldots, q_n)$, and so it follows from (3.2) that $A(p_1, p_2; q_1, q_2) = A(p_2, p_1; q_2, q_1)$. This gives the symmetry of $G$:

$$G(x, y) = G(y, x) \quad \text{for all } x, y \in J \text{ with } x + y \in J. \quad (3.4)$$
The expansibility and symmetry of $I$, together with (3.2) while choosing $p_2 = q_2 = 0$, allows us to get $A(p_1, 0; q_1, 0) = 0$. This gives the identity

$$G(x, 0) = 0 \quad \text{for all } x \in J. \quad (3.5)$$

We write $x = (p_1, q_1)$, $y = (p_2, q_2)$, and $z = (p_3, q_3)$ and, using (3.2) twice, we get

$$I(p_1, p_2, \ldots, p_n; q_1, q_2, \ldots, q_n)$$

$$= I(p_1 + p_2, p_3, \ldots, p_n; q_1 + q_2, q_3, \ldots, q_n) + G(x, y)$$

$$= I(p_1 + p_2 + p_3, p_4, \ldots, p_n; q_1 + q_2 + q_3, q_4, \ldots, q_n)$$

$$+ G(x + y, z) + G(x, y).$$

The symmetry of $I$ in $x$, $y$ and $z$ gives

$$G(x + y, z) + G(x, y) = G(y + z, x) + G(y, z),$$

and by the symmetry of $G$ we can rewrite it as

$$G(x, y) + G(x + y, z) = G(x, y + z) + G(y, z)$$

for all $x, y, z \in J$ with $x + y + z \in J. \quad (3.6)$

The mapping $G$ thus verifies the hypotheses of Theorem 2.1. Hence, there exists a mapping $g: J \rightarrow \mathbb{R}$ with $g(0) = 0$ representing $G$ through

$$G(x, y) = g(x) + g(y) - g(x + y). \quad (3.7)$$

We are ready to represent $I$ by $g$. In fact, we use (3.2) successively, with $x_i := (p_i, q_i)$ and $1 = (1, 1)$:

$$I(p_1, p_2, \ldots, p_n; q_1, q_2, \ldots, q_n)$$

$$= I(p_1, p_2, \ldots, p_n, 0; q_1, q_2, \ldots, q_n, 0)$$

$$= I(p_1 + p_2, p_3, \ldots, p_n, 0; q_1 + q_2, q_3, \ldots, q_n, 0) + G(x_1, x_2)$$

$$= I(p_1 + p_2 + p_3, p_4, \ldots, p_n, 0; q_1 + q_2 + q_3, q_4, \ldots, q_n, 0)$$

$$+ G(x_1 + x_2, x_3) + G(x_1, x_2)$$

$$= \cdots$$

$$= I(1; 1) + G(1, 0) + \sum_{j=1}^{n-1} G\left(\sum_{i=1}^{j} x_i, x_{j+1}\right)$$

$$= I(1; 1) - g(1) + \sum_{i=1}^{n} g(x_i).$$
With the notation $f(p_i; q_i)$ for $g(x_i)$, the above identity is our asserted representation.

**Remark 3.1.** It is sufficient to assume the symmetry of $I$ for $P, Q$ of length 4, as can be seen from our proof for Theorem 3.1.

**References**


Forte, B. and Ng, C. T. (unpublished data).


Rathie, P. N. and Kannappan, P. L. (1972), A directed-divergence function of type $\beta$, Information and Control 20, 38–45.