Can. J. Math. Vol. 44 (3), 1992 pp. 463–482

DECOMPOSITION OF MULTIVARIATE FUNCTIONS

J. M. BORWEIN AND A. S. LEWIS

ABSTRACT. Given a bivariate function defined on some subset of the Cartesian product of two sets, it is natural to ask when that function can be decomposed as the sum of two univariate functions. In particular, is a pointwise limit of such functions itself decomposable? At first glance this might seem obviously true but, as we show, the possibilities are quite subtle. We consider the question of existence and uniqueness of such decompositions for this case and for many generalizations to multivariate functions and to cases where the sets and functions have topological or measure theoretic structure.

1. **Introduction.** Consider the following simply stated question. Suppose S_1 and S_2 are sets, with $E \subset S_1 \times S_2$, and $h: E \longrightarrow \mathbb{R}$. Suppose further that $(f_1^n)_{n=1}^{\infty}$ and $(f_2^n)_{n=1}^{\infty}$ are two sequences of functions, $f_i^n: S_i \longrightarrow \mathbb{R}$, for $i = 1,2$ and each n, with the property that

$$(1) \qquad \lim_{n\to\infty} \{f_1^n(s_1) + f_2^n(s_2)\} = h(s_1, s_2), \text{ for all } (s_1, s_2) \in E.$$

In other words, on the set E, the bivariate function h is the pointwise limit of sums of univariate functions. Is it true that h is itself the sum of univariate functions: do there exist functions $f_1: S_1 \longrightarrow \mathbb{R}$, $f_2: S_2 \longrightarrow \mathbb{R}$ with

$$(2) \qquad f_1(s_1) + f_2(s_2) = h(s_1, s_2), \text{ for all } (s_1, s_2) \in E?$$

It seems clear that the answer must be affirmative, and we shall prove this. Notice that if E is a finite set the result is easy: the set of functions h which have a decomposition of the form (2) clearly forms a subspace of the finite-dimensional vector space $\mathbb{R}^E$, so is closed, and the result now follows. One might ask the same question, but with the range of all the functions involved changed from $\mathbb{R}$ to $\mathbb{R}_+ = \{0 \le x \in \mathbb{R}\}$. Again we shall show that the answer is affirmative. (Again the proof is straightforward if E is finite).

However, to demonstrate the difficulties which can arise, consider the following example. Define

$$E := \{(m,m), (m+1,m) \mid m \in \mathbb{N}\} \subset \mathbb{N}^2,$$

and a function $h: E \longrightarrow \mathbb{R}$ by $h(m,m) := 0$ and $h(m+1,m) := -1$, for all $m \in \mathbb{N}$. It is easy to see that if $f_i: \mathbb{N} \longrightarrow \mathbb{R}$, for $i = 1,2$, have property (2) then $f_1(m) = k - m$, and $f_2(m) = m - k$, for all $m \in \mathbb{N}$, where k is some constant. Thus it is not possible to find $f_1: \mathbb{N} \longrightarrow \mathbb{R}_+$ and $f_2: \mathbb{N} \longrightarrow \mathbb{R}$ with property (2).

Research partially supported by the Natural Sciences and Engineering Research Council of Canada
Received by the editors May 31, 1991.
AMS subject classification: Primary 54C05, 05C38; Secondary 08A45, 28A20.
Key words and phrases: Separable function, conservative, potential difference, equational compactness.

However, suppose we define $f_1^n\colon \mathbb{N} \longrightarrow \mathbb{R}_+$, and $f_2^n\colon \mathbb{N} \longrightarrow \mathbb{R}$, by

$$f_1^n(m) := (n-m)^+,$$
$$f_2^n(m) := (m-n),$$

for each $m, n \in \mathbb{N}$ (where for $x \in \mathbb{R}$, $x^+ = \max\{x, 0\}$). Then, for all $m \in \mathbb{N}$

$$\begin{aligned} f_1^n(m) + f_2^n(m) &= (m-n)^+ \\ &\longrightarrow 0 \text{ as } n \longrightarrow \infty, \text{ and} \\ f_1^n(m+1) + f_2^n(m) &= (m+1-n)^+ - 1 \\ &\longrightarrow -1 \text{ as } n \longrightarrow \infty. \end{aligned}$$

To summarize, we have property (1), and yet there do not exist functions $f_1\colon \mathbb{N} \longrightarrow \mathbb{R}_+$ and $f_2\colon \mathbb{N} \longrightarrow \mathbb{R}$ with property (2).

Such questions as the above have many natural extensions to multivariate problems. For example, suppose $E \subset \prod_{i=1}^4 S_i$, $h\colon E \longrightarrow \mathbb{R}$, $f_i^n\colon S_i \times S_{i+1} \longrightarrow \mathbb{R}$ for $i = 1, 2, 3$, and each n, and

$$\lim_{n\to\infty} \sum_{i=1}^{3} f_i^n(s_i, s_{i+1}) = h(s_1, s_2, s_3, s_4), \text{ for all } (s_1, s_2, s_3, s_4) \in E.$$

Do there exist $f_i\colon S_i \times S_{i+1} \longrightarrow \mathbb{R}$ with

$$\sum_{i=1}^{3} f_i(s_i, s_{i+1}) = h(s_1, s_2, s_3, s_4), \text{ for all } (s_1, s_2, s_3, s_4) \in E?$$

This paper will be concerned with characterizing those functions h which have a decomposition of the form (2), and generalizations of this question. When the underlying sets S_i have additional topological or measure-theoretic structure we can ask about the existence and uniqueness of decompositions of the form (2) with additional requirements on the functions f_i, such as continuity or measurability. The present investigation was largely motivated by questions of this type, which arose in the context of an abstract optimization problem (see [2]). Indeed, we were trying to treat continuous 'DAD problems' (*cf.* [12] and [13]) by direct optimization methods rather than using the information-theoretic techniques of [3] applied in [12] and [13]. (For a recent survey of the optimization approach to DAD problems in the finite-dimensional case, see [16] and [17]). However, the relevant result in [3] (Corollary 3.1) is inadequate because it fails to address these questions.

As a typical example, suppose (S_1, ds_1) and (S_2, ds_2) are measure spaces, with $E \subset S_1 \times S_2$ measurable. Consider the set of all those functions $h \in L_1(E, ds_1 ds_2)$ which decompose, in the sense that there exist measurable $f_1 = S_1 \longrightarrow \mathbb{R}$ and $f_2 = S_2 \longrightarrow \mathbb{R}$ with

$$f_1(s_1) + f_2(s_2) = h(s_1, s_2) \text{ a.e. on } E.$$

Clearly the set of all such functions h forms a subspace of $L_1(E, ds_1 ds_2)$. Is this subspace closed? We shall answer this question positively, by elementary methods, in this paper (under reasonable conditions on E). This result allows us to give conditions ensuring the existence of a solution to the following continuous DAD problem: given two measure spaces S and T, a non-negative measurable 'kernel' $k: S\times T \longrightarrow \mathbb{R}$, and non-negative measurable 'marginals' $\alpha: S \longrightarrow \mathbb{R}$ and $\beta: T \longrightarrow \mathbb{R}$, find non-negative measurable 'weights' $f: S \longrightarrow \mathbb{R}$ and $g: T \longrightarrow \mathbb{R}$ so that

$$\int_T f(s)k(s,t)g(t)\,dt = \alpha(s), \text{ a.e. on } S, \text{ and}$$
$$\int_S f(s)k(s,t)g(t)\,ds = \beta(t), \text{ a.e. on } T\,.$$

See [2] for details.

As suggested by the above application, our interest is in the case where the underlying sets are infinite. For finite sets the results characterizing decomposability are straightforward and well-known: they are related to cyclic products on matrices (see for example [5] for a survey). Similar ideas have been used to study the approximation of bivariate by univariate functions. See [6], [7] and [14] for further references.

The following result provides a model for the problems which arise. For simplicity suppose $E \subset [0,1]^2$ is a convex open set in $\mathbb{R}^2$, and define two open intervals,

$$E_1 := \{ s_1 \mid \text{ there exists } s_2 \text{ with } (s_1, s_2) \in E\}, \text{ and}$$
$$E_2 := \{ s_2 \mid \text{ there exists } s_1 \text{ with } (s_1, s_2) \in E\}\,.$$

Suppose $h: E \longrightarrow \mathbb{R}$ is C^2.

THEOREM 1.1. *There exist C^1 functions $f_i: E_i \longrightarrow \mathbb{R}$, $i = 1, 2$, with*

$$f_1(s_1) + f_2(s_2) = h(s_1, s_2), \text{ for all } (s_1, s_2) \in E, \tag{3}$$

if and only if, for every closed, piecewise smooth, oriented curve γ in E we have

$$\int_\gamma \frac{\partial h}{\partial s_1} ds_1 = 0. \tag{4}$$

Since E is simply-connected this is equivalent to $\partial^2 h / \partial s_1 \partial s_2$ vanishing identically on E.

PROOF. Notice that, since we always have

$$\int_\gamma \Big(\frac{\partial h}{\partial s_1} ds_1 + \frac{\partial h}{\partial s_2} ds_2 \Big) = 0,$$

equation (4) is equivalent to

$$\int_\gamma \frac{\partial h}{\partial s_2} ds_2 = 0,$$

by [4, 9.4].

Suppose equation (4) holds. Pick $(s_1^0, s_2^0) \in E$. Now for any $t_1 \in E_1$, define

$$f_1(t_1) := \int_{\gamma_1} \frac{\partial h}{\partial s_1} ds_1, \tag{5}$$

where t_2 is any point with $(t_1, t_2) \in E$ and γ_1 is any piecewise smooth oriented curve in E from (s_1^0, s_2^0) to (t_1, t_2). Suppose $(t_1, t_2') \in E$ and γ_2 is a curve from (s_1^0, s_2^0) to (t_1, t_2'). Then

$$\begin{aligned}\int_{\gamma_1} (\partial h/\partial s_1)\, ds_1 &= \int_{\gamma_2} (\partial h/\partial s_1)\, ds_1 + \int_{[(t_1,t_2),(t_1,t_2')]} (\partial h/\partial s_1)\, ds_1 \\ &= \int_{\gamma_2} (\partial h/\partial s_1) ds_1,\end{aligned}$$

by (4), so $f_1 \colon E_1 \longrightarrow \mathbb{R}$ is well-defined by (5), and clearly C^1. Similarly (by the first remark), we can define a C^1 function $f_2 \colon E_2 \longrightarrow \mathbb{R}$ by

$$f_2(t_2) := \int_{\gamma_1} \frac{\partial h}{\partial s_2} ds_2 + h(s_1^0, s_2^0).$$

But now we have

$$\begin{aligned}f_1(t_1) + f_2(t_2) &= \int_{\gamma_1} \Big((\partial h/\partial s_1)\, ds_1 + (\partial h/\partial s_2)\, ds_2 \Big) + h(s_1^0, s_2^0) \\ &= h(t_1, t_2) - h(s_1^0, s_2^0) + h(s_1^0, s_2^0) \\ &= h(t_1, t_2),\end{aligned}$$

(since we are integrating an exact differential form, see [4, 9.4]), so (2) holds. The converse is obvious.

The last statement follows from the fact that (4) holds if and only if

$$(\partial h/\partial s_1)\, ds_1 + 0\, ds_2$$

is a closed differential form, which is equivalent to $\partial^2 h/\partial s_1 \partial s_2$ vanishing identically [4, 9.4.9]. ∎

The characterization that we shall develop for functions h satisfying (2) will be exactly the discrete analogue of Condition (4).

2. **Potential differences.** In this section we shall be concerned with graphs and with potential differences defined on the edges of a given graph. We will not restrict attention to graphs with a finite number of vertices, and we will allow potentials on the vertices to take their value in an arbitrary group.

Throughout this section, let (V, E) be an (undirected) graph with vertices V (not necessarily finite) and edges E (consisting of unordered pairs from V), and let G be a group, with identity element e. Recall that a *circuit* is a path of vertices $v_0 v_1 v_2 \cdots v_k v_{k+1}$ with $v_i v_{i+1} \in E$, $i = 1, \ldots, k$ and $v_{k+1} = v_0$.

DEFINITION 2.1. A function $h \colon V \times V \longrightarrow G$ is a *(left) potential difference* if there exists a function $p \colon V \longrightarrow G$ with $h(v_1, v_2) = p(v_1)^{-1} p(v_2)$ for all $v_1 v_2 \in E$.
Notice that if h is a potential difference and $v_1 v_2 \in E$, then

$$h(v_2, v_1) = h(v_1, v_2)^{-1}.$$

The following result is well-known in various forms. It is essentially Kirchoff's potential law (see for example [1] when $G = (\mathbb{R}, +)$). Notice though we do not require V finite.

THEOREM 2.2. *The function $h\colon V \times V \longrightarrow G$ is a potential difference if and only if for every circuit $v_0 v_1 \cdots v_k v_0$,*

$$h(v_0, v_1)\, h(v_1, v_2) \cdots h(v_k, v_0) = e \tag{6}$$

In this case $h(v_1, v_2) = p(v_1)^{-1} p(v_2)$ for all $v_1 v_2 \in E$, where the function $p\colon V \longrightarrow G$ is unique on each connected component, up to a constant multiple.

PROOF. One direction is immediate. On the other hand, suppose (6) holds. Clearly we can treat each connected component of the graph separately, so without loss of generality suppose (V, E) is connected. Fix an arbitrary vertex v_0.

We define $p\colon V \longrightarrow G$ in the following way. Set $p(v_0) := e$. Suppose $u \in V$. Since (V, E) is connected, there exists a path $v_0 v_1 \cdots v_k u$. Then define

$$p(u) := h(v_0, v_1)\, h(v_1, v_2) \cdots h(v_k, u). \tag{7}$$

To see that p is well-defined, suppose $v_0 v_1' \cdots v_j' u$ is another path, so

$$v_0 v_1 \cdots v_k u v_j' \cdots v_1' v_0$$

is a circuit and thus

$$h(v_0, v_1)\, h(v_1, v_2) \cdots h(v_k, u)\, h(u, v_j') \cdots h(v_1', v_0) = e, \tag{8}$$

Notice that whenever $u_1 u_2 \in E$, $u_1 u_2 u_1$ is a circuit, so (6) implies

$$h(u_1, u_2) = h(u_2, u_1)^{-1}.$$

From (8),

$$\begin{aligned} p(u) &= h(v_0, v_1)\, h(v_1, v_2) \cdots h(v_k, u) \\ &= h(v_1', v_0)^{-1} h(v_2', v_1')^{-1} \cdots h(u, v_j')^{-1} \\ &= h(v_0, v_1')\, h(v_1', v_2') \cdots h(v_j', u), \end{aligned}$$

as required. Furthermore, whenever $u_1 u_2 \in E$, by the definition of p, $p(u_2) = p(u_1) h(u_1, u_2)$, or $h(u_1, u_2) = p(u_1)^{-1} p(u_2)$, so h is a potential difference.

Suppose $h(v_1, v_2) = q(v_1)^{-1} q(v_2)$ whenever $v_1 v_2 \in E$, for some other function $q\colon V \longrightarrow G$. Then if $v_0 v_1 \cdots v_k u$ is a path, we must have

$$\begin{aligned} q(u) &= q(v_k) h(v_k, u) \\ &= q(v_{k-1}) h(v_{k-1}, v_k) h(v_k, u) \\ &\;\vdots \\ &= q(v_0) h(v_0, v_1) \cdots h(v_k, u) \\ &= q(v_0) p(u), \end{aligned}$$

and this must hold for every $u \in V$. The last assertion follows. ∎

The next result shows that if G is a topological group then the set of potential differences is closed under pointwise convergence.

COROLLARY 2.3. *Suppose G is a topological group and* $(h_\alpha)_{\alpha\in\Theta}$ *is a net of potential differences,* $h_\alpha\colon V\times V\longrightarrow G$, *for* $\alpha\in\Theta$. *Suppose that* $h\colon V\times V\longrightarrow G$ *and* $h_\alpha\longrightarrow h$ *pointwise: in other words,*

$$\lim_\alpha h_\alpha(v_1,v_2)=h(v_1,v_2),\ \textit{for all } v_1v_2\in E.$$

Then h is a potential difference.

PROOF. We just need to check (6). If $v_0v_1\cdots v_kv_0$ is a circuit then

$$\begin{aligned} h(v_0,v_1)h(v_1,v_2)\cdots h(v_k,v_0) &= \big(\lim_\alpha h_\alpha(v_0,v_1)\big)\cdots\big(\lim_\alpha h_\alpha(v_k,v_0)\big)\\ &= \lim_\alpha\big(h_\alpha(v_0,v_1)\cdots h_\alpha(v_k,v_0)\big)\\ &= \lim_\alpha e\\ &= e, \end{aligned}$$

as required. ∎

We can restate this as follows.

COROLLARY 2.4. *Suppose G is a topological group and* $(p_\alpha)_{\alpha\in\Theta}$ *is a net of functions,* $p_\alpha\colon V\longrightarrow G$, *for* $\alpha\in\Theta$. *Suppose that* $h\colon V\times V\longrightarrow G$ *and*

$$\lim_\alpha\big(p_\alpha(v_1)^{-1}p_\alpha(v_2)\big)=h(v_1,v_2),\ \textit{for all } v_1v_2\in E.$$

Then there exists a function $p\colon V\longrightarrow G$ *(unique up to a constant multiple on each connected component) with*

$$p(v_1)^{-1}p(v_2)=h(v_1,v_2),\ \textit{for all } v_1v_2\in E.$$

Suppose in the above result we know that for each vertex v, $p_\alpha(v)\in G_v$ for all $\alpha\in\Theta$, where G_v is a closed subset of G. Can we choose the function p so that $p(v)\in G_v$ for all $v\in V$? A simple example shows that this need not be the case. Let the graph consist of just two vertices, $V=\{1,2\}$, and the single edge joining them, $E=\{12\}$. Let the group G be $(\mathbb{R},+)$ and suppose h is identically zero. Consider the sequence of functions $p_n\colon V\longrightarrow\mathbb{R}$ defined by (for $n\geq 2$) $p_n(1):=n$, and $p_n(2):=n+\frac{1}{n}$. Let

$$\begin{aligned} C_1 &:= \{n\in\mathbb{N}\mid n\geq 2\}\\ C_2 &:= \Big\{n+\frac{1}{n}\mid n\in\mathbb{N}, n\geq 2\Big\}, \end{aligned}$$

so certainly $p_n(i)\in C_i$ for each i and n, and C_1 and C_2 are closed in $\mathbb{R}$. Clearly

$$p_n(j)-p_n(i)\longrightarrow 0,\ \text{for all } i,j.$$

However, if $p(2)-p(1)=h(1,2)=0$, and $p(i)\in C_i$, for each i, then $0\in C_2-C_1$, which is a contradiction.

In the case where the graph (V, E) is *finite*, the group $G = (\mathbb{R}, +)$, and each set G_v is a closed interval in $\mathbb{R}$, there are well-known conditions characterizing those functions $h: V \times V \longrightarrow \mathbb{R}$ for which there exists a function $p: V \longrightarrow \mathbb{R}$ with $p(v) \in G_v$ for all $v \in V$, and

$$p(v_2) - p(v_1) = h(v_1, v_2) \text{ for all } v_1 v_2 \in E,$$

(see [1] for example). It is easy to see from these conditions that in this case the answer to the above question is affirmative.

However, even in this case, if we allow the graph to be infinite the result may fail. Examples analogous to that given in the introduction demonstrate this. What is lacking is a suitable compactness conditon.

COROLLARY 2.5. *Suppose G is a topological group, and $G_v \subset G$ is closed for each vertex $v \in V$. Suppose $(p_\alpha)_{\alpha\in\Theta}$ is a net of functions, $p_\alpha: V \longrightarrow G$, for $\alpha \in \Theta$, with the property that for each vertex $v \in V$, $p_\alpha(v) \in G_v$ for all $\alpha \in \Theta$. Suppose further that $h: V \times V \longrightarrow G$, with*

$$\lim_\alpha \big(p_\alpha(v_1)^{-1} p_\alpha(v_2)\big) = h(v_1, v_2), \textit{ for all } v_1 v_2 \in E. \tag{9}$$

Suppose finally that the following condition holds.

CONDITION A. *For each connected component V_γ, there exists a vertex $v_\gamma \in V_\gamma$ such that the net $\big(p_\alpha(v_\gamma)\big)_{\alpha\in\Theta}$ has a convergent subnet.*

Then there exists a function $p: V \longrightarrow G$ with

$$p(v_1)^{-1} p(v_2) = h(v_1, v_2), \textit{ for all } v_1 v_2 \in E,$$

and

$$p(v) \in G_v, \textit{ for all vertices } v \in V.$$

PROOF. We can treat each connected component independently, so without loss of generality suppose (V, E) is connected. Suppose $\big(p_\alpha(v_0)\big)_{\alpha\in\Theta}$ has a convergent subnet. Without loss of generality, $\lim_\alpha p_\alpha(v_0) = g \in G_{v_0}$. By Corollary 2.4 there exists a function $p: V \longrightarrow G$ with $p(v_1)^{-1} p(v_2) = h(v_1, v_2)$, for all $v_1 v_2 \in E$, and we can choose p so that $p(v_0) = g$.

We now claim $\lim_\alpha p_\alpha(v) = p(v)$ for all vertices $v \in V$. Suppose there is a path of length k linking v_0 to v. The proof will be induction on k. Clearly the result is true for $k = 0$. Suppose it holds for k, and $v_0 v_1 \cdots v_k v$ is a path. By hypothesis, $\lim_\alpha p_\alpha(v_k) = p(v_k)$. Furthermore, by assumption,

$$\begin{aligned}\lim_\alpha \big(p_\alpha(v_k)^{-1} p_\alpha(v)\big) &= h(v_k, v) \\ &= p(v_k)^{-1} p(v),\end{aligned}$$

so $\lim_\alpha p_\alpha(v) = p(v)$ as required. Thus the hypothesis holds for $k + 1$, and hence for all k.

Finally, since $p_\alpha(v) \in G_v$ for all $\alpha \in \Theta$, and G_v is closed, $p(v) \in G_v$ as required. ∎

EXAMPLES.

(1) If, for each connected component V_γ, there exists a vertex $v_\gamma \in V_\gamma$ with G_{v_γ} compact then Condition A holds. In particular, if G is compact, Condition A holds, as for example on $G = \left(\{ e^{2\pi i\theta} \mid 0 < \theta \leq 1\}, \cdot\right)$.

(2) Suppose $G = (\mathbb{R}, +)$ and in each connected component there exists an edge $v_1 v_2$ with $G_{v_1} = \mathbb{R}_+$, $G_{v_2} = -\mathbb{R}_+$. Then

$$p_\alpha(v_2) - p_\alpha(v_1) \longrightarrow h(v_1, v_2),$$

so for all α sufficiently large

$$p_\alpha(v_2) - p_\alpha(v_1) \in [h(v_1, v_2) - 1, h(v_1, v_2) + 1],$$

and thus

$$p_\alpha(v_2) \in [h(v_1, v_2) - 1, 0],$$

which is compact. Therefore $\left(p_\alpha(v_2)\right)$ has a convergent subnet, so Condition A holds.

(3) Much the same argument works with the multiplicative group

$$G = \left(\{0 < x \in \mathbb{R}\}, \cdot\right),$$

assuming in each connected component there exists an edge $v_1 v_2$ with $G_{v_1} = \{x \geq 1\}$ and $G_{v_2} = \{x \leq 1\}$.

3. **Decomposable functions.** We now return to our central question. We wish to characterize those bivariate (or, more generally, multivariate) functions which can be decomposed as products or sums of univariate functions. We will call such functions 'decomposable'.

DEFINITION 3.1. Suppose S_1, S_2 are sets, $E \subset S_1 \times S_2$, and G is a group. A function $h: E \longrightarrow G$ is *decomposable* if there exist functions $f_i: S_i \longrightarrow G$, $i = 1, 2$, with

$$h(s_1, s_2) = f_1(s_1) f_2(s_2), \quad \text{for all } (s_1, s_2) \in E. \tag{10}$$

Associated with the sets E, S_1, S_2 we can define a bipartite graph (V, E') with vertices $V := S_1 \cup S_2$, and edges $E' := \{s_1 s_2 \mid (s_1, s_2) \in E\}$. Corresponding to the function $h: E \longrightarrow G$ we can define a function

$$h': (S_1 \cup S_2) \times (S_1 \cup S_2) \longrightarrow G,$$

by

$$h'(u, v) = \begin{cases} h(u, v), & \text{if } u \in S_1,\ v \in S_2, \\ h(u, v)^{-1}, & \text{if } u \in S_2,\ v \in S_1, \\ e, & \text{otherwise,} \end{cases}$$

https://doi.org/10.4153/CJM-1992-030-9 Published online by Cambridge University Press

(where as before e is the identity). Now h' is a potential difference exactly when there exists a function $p: S_1 \cup S_2 \longrightarrow G$ with $h'(u,v) = p(u)^{-1}p(v)$ for all edges uv. Simply by making the identification

$$f_1(s_1) = p(s_1)^{-1}, \text{ for } s_1 \in S_1,$$
$$f_2(s_2) = p(s_2), \text{ for } s_2 \in S_2,$$

it is clear that h is decomposable if and only if h' is potential difference.

Suppose the connected components of the associated bipartite graph (V, E') are (V_γ, E'_γ), for $\gamma \in \Gamma$, so $V = \cup_{\gamma\in\Gamma} V_\gamma$ and $E' = \cup_{\gamma\in\Gamma} E'_\gamma$ are disjoint unions. For each $\gamma \in \Gamma$, $V_\gamma = S_1^\gamma \cup S_2^\gamma$, for some subsets $S_1^\gamma \subset S_1$, $S_2^\gamma \subset S_2$, and $\cup_{\gamma\in\gamma} S_1^\gamma = S_1$, $\cup_{\gamma\in\Gamma} S_2^\gamma = S_2$ are disjoint unions. If we write $E^\gamma := \{(s_1,s_2) \mid s_1s_2 \in E'_\gamma\}$, then $E^\gamma \subset S_1^\gamma \times S_2^\gamma$ for each $\gamma \in \Gamma$ and $E = \cup_{\gamma\in\Gamma} E^\gamma$ is a disjoint union.

We call $(S_1^\gamma, S_2^\gamma, E^\gamma)$ the *connected components* of (S_1, S_2, E). To define them directly, define a relation L on E by $(s_1,s_2)L(t_1,t_2)$ for $s,t \in E$ if either $s_1 = t_1$ or $s_2 = t_2$, and then define an equivalence relation $\sim$ on E by $s \sim t$ for $s,t \in E$ if there exists a sequence $s^1, s^2, \ldots, s^k$ in E with $s^1 = s$, $s^k = t$ and s^jLs^{j+1} for $j = 1, \ldots, k-1$. The relation $\sim$ partitions E into equivalence classes E^γ, for $\gamma \in \Gamma'$, and we then define

$$S_1^\gamma := \{s_1 \in S_1 \mid (s_1,s_2) \in E^\gamma \text{ for some } s_2 \in S_2\}, \tag{11}$$
$$S_2^\gamma := \{s_2 \in S_2 \mid (s_1,s_2) \in E^\gamma \text{ for some } s_1 \in S_1\}. \tag{12}$$

For each $s_1 \in S_1 \setminus \cup_{\gamma\in\Gamma'} S_1^\gamma$, we adjoin the triple $(\{s_1\}, \emptyset, \emptyset)$, and similary for each $s_2 \in S_2 \setminus \cup_{\gamma\in\Gamma'} S_2^\gamma$ we adjoin $(\emptyset, \{s_2\}, \emptyset)$. Together with the triples $(S_1^\gamma, S_2^\gamma, E^\gamma)$, $\gamma \in \Gamma'$, these make up the connected components of (S_1, S_2, E).

Corresponding to circuits in the associated bipartite graph are circuits in E. To describe them directly we use the following definition.

DEFINITION 3.2. A *circuit* in $E \subset S_1 \times S_2$ is a sequence $s^1s^2s^3 \cdots s^{2k}s^{2k+1}$ of points $s^j = (s_1^j, s_2^j) \in E$ with $s^{2k+1} = s^1$, and $s_1^j = s_1^{j+1}$, $s_2^j \neq s_2^{j+1}$, $s_1^{j+1} \neq s_1^{j+2}$ and $s_2^{j+1} = s_2^{j+2}$ for either all even j or all odd j in $\{1, \ldots, 2k\}$. (In other words, the component of s^j which changes as we follow the circuit alternates between the first and second.)

The following results are exact translations of the corresponding results in § 2. In each of them, $E \subset S_1 \times S_2$, G is a group, and $h: E \longrightarrow G$.

THEOREM 3.3. *The function h is decomposable if and only if for every circuit in E, $s^1, s^2 \cdots s^{2k}s^1$, we have*

$$h(s^1)h(s^2)^{-1}h(s^3)h(s^4)^{-1}\cdots h(s^{2k})^{-1} = e.$$

In this case the functions $f_i: S_i \longrightarrow G$, $i = 1,2$ in (10) are unique on each connected component of (S_1, S_2, E) up to a constant multiple. In other words, if f_1', f_2' also satisfy (10) then on each connected component $(S_1^\gamma, S_2^\gamma, E^\gamma)$, there exists $g_\gamma \in G$ with $f_1'(s_1) = f_1(s_1)g_\gamma$ for all $s_1 \in S_1^\gamma$, and $f_2'(s_2) = g_\gamma^{-1}f_2(s_2)$ for all $s_2 \in S_2^\gamma$.

PROOF. Theorem 2.2 ∎

https://doi.org/10.4153/CJM-1992-030-9 Published online by Cambridge University Press

COROLLARY 3.4. *Suppose G is a topological group and $(h_\alpha)_{\alpha\in\Theta}$ is a net of decomposable functions, $h_\alpha: E \longrightarrow G$, for $\alpha \in \Theta$. Suppose that $h: E \longrightarrow G$ and $h_\alpha \longrightarrow h$ pointwise: in other words,*

$$\lim_\alpha h_\alpha(p) = h(p), \textit{ for all } p \in E.$$

Then h is decomposable.

PROOF. Corollary 2.3. ■

It is worthwhile making the intuitive basis for the proof of this result clearer. For simplicity, consider the case where $G = (\mathbb{R}, +)$ and the nets are sequences, and suppose that (S_1, S_2, E) is connected. Thus we have, for all (s_1, s_2) in E, $f_1^n(s_1) + f_2^n(s_2) \longrightarrow h(s_1, s_2)$. The difficulty of course is that f_1^n and f_2^n may not converge themselves. However, it is not difficult to see that the only way this can happen is if $f_1^n = g_1^n + k_n$ and $f_2^n = g_2^n - k_n$ for some sequence of constants k_n, where the functions g_1^n and g_2^n do converge. To make this precise, pick an arbitrary $(\overline{s_1}, \overline{s_2})$ in E and define $k_n := f_1^n(\overline{s_1})$, $g_1^n := f_1^n - k_n$, and $g_2^n := f_2^n + k_n$. Then the proof above translates immediately into a demonstration that g_1^n and g_2^n converge pointwise to the desired decomposition of h.

COROLLARY 3.5. *Suppose $(f_i^\alpha)_{\alpha\in\Theta}$, $i = 1, 2$ are sets of functions, $f_i^\alpha: S_i \longrightarrow \mathbb{R}_+$, that $E \subset S_1 \times S_2$ and $h: E \longrightarrow \mathbb{R}_+$, and that*

$$f_1^\alpha(s_1) + f_2^\alpha(s_2) \longrightarrow h(s_1, s_2), \textit{ for all } (s_1, s_2) \in E.$$

Then there exist functions $f_i: S_i \longrightarrow \mathbb{R}_+$, $i = 1, 2$, unique up to a constant on each connected component (cf. Theorem 3.3), with

$$f_1(s_1) + f_2(s_2) = h(s_1, s_2), \textit{ for all } (s_1, s_2) \in E.$$

PROOF. Corollary 2.5 and the following examples. ■

This resolves the opening questions in the introduction.

4. **The multivariate case.** We would naturally expect such results as Corollary 3.4, which showed that pointwise limits of (bivariate) decomposable functions were decomposable, would extend to the multivariate case. In most of the circumstances in which we are interested (for example $G = (\mathbb{R}, +)$) this is indeed the case, as we shall see. Surprisingly however, these results can fail in the multivariate case even with $G = (\mathbb{Z}, +)$. We shall present an example of this. Beyond the two variable case such simple characterizations as Theorem 3.3 are no longer possible, and rather than attempting to generalize these, we will proceed directly to extensions of Corollary 3.4.

The proofs in this section, which could also be used in the bivariate case, are shorter but less constructive than those used in the bivariate case. They are more restrictive on the underlying group G, and provide no uniqueness information as in Corollary 3.5.

We will use the following ideas from universal algebra (see [8]). Suppose G is an abelian group, Φ and Λ are two index sets,

$$a: \Phi \times \Lambda \longrightarrow \{-1, 0, +1\}$$

has the property that $\{\varphi \in \Phi \mid a(\varphi, \lambda) \neq 0\}$ is finite for all $\lambda \in \Lambda$, and $b: \Lambda \longrightarrow G$. Consider a canonical set of linear equations, indexed by Λ, in variables $x_\varphi \in G$ for $\varphi \in \Phi$:

$$\sum_{\varphi \in \Phi} a(\varphi, \lambda) x_\varphi = b(\lambda), \text{ for } \lambda \in \Lambda. \tag{13}$$

It is easy to see in fact that systems of this form with arbitrary integer coefficients can be reduced to this special case by adding extra variables.

DEFINITION 4.1. (i) System (13) is *solvable* if there exist variables x_φ, for $\varphi \in \Phi$ satisfying the system.

(ii) System (13) is *finitely solvable* if every finite subsystem of (13) (*i.e.* indexed by $\lambda \in \Lambda_0$ where Λ_0 is a finite subset of Λ) is solvable.

(iii) G is *equationally compact* if any system of the form (13) which is finitely solvable, is solvable.

It is easy to see, by Tychonoff's theorem, that if G is a compact abelian group then it is equationally compact: corresponding to any finite subsystem is a collection of possible solutions which clearly form a closed subset of the compact set G^Φ, and which furthermore have the finite intersection property.

DEFINITION 4.2. An abelian group G is *divisible* if for any $x \in G$ and $n \in \mathbb{N}$, there exists $y \in G$ with $ny = x$.

Thus, for example, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ and $(\mathbb{C}, +)$ are divisible, but $(\mathbb{Z}, +)$ is not.

THEOREM 4.3. *Any divisible abelian group is equationally compact.*

PROOF. See for example [8, Appendix 6, by G.H. Wenzel]. ∎

Thus $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ and $(\mathbb{C}, +)$ are all equationally compact. The following example shows that $(\mathbb{Z}, +)$ is not equationally compact.

EXAMPLE. Consider the system of equations in $(\mathbb{Z}, +)$

$$x_n + 2x_{n+1} = 1, \quad n \in \mathbb{N}. \tag{14}$$

Clearly (14) is finitely solvable but not solvable.

The proof of the following decomposition result in the case in which we are particularly interested, where G is $(\mathbb{R}, +)$ (and equally when G is $(\mathbb{C}, +)$ or $(\mathbb{Q}, +)$), proceeds in two steps. The first step depends on the previous observation that for functions on finite sets, decomposable functions form a subspace of a finite-dimensional vector space, so are a closed set under pointwise convergence. We then use the equational compactness of $(\mathbb{R}, +)$ (or $(\mathbb{C}, +)$ or $(\mathbb{Q}, +)$) to extend to the infinite case. We shall use the notation that for $s \in \prod_{i=1}^n S_i$ and $I \subset \{1, 2, \ldots, n\}$, $s_I \in \prod_{i \in I} S_i$ is defined by $(s_I)_i = s_i$, for $i \in I$.

THEOREM 4.4. *Suppose*

$$V \subset \prod_{i=1}^{n} S_i,$$

G is a topological group, and $h\colon V \longrightarrow G$. Suppose $\emptyset \neq I_j \subset \{1,\ldots,n\}$, and $f_j^\alpha\colon \prod_{i\in I_j} S_i \longrightarrow G$ are nets of functions indexed by $\alpha \in \Theta$, for each $j = 1,\ldots,m$. Suppose that $h\colon \prod_{i=1}^n S_i \longrightarrow G$ has the property that

$$\lim_\alpha\big(f_1^\alpha(s_{I_1})f_2^\alpha(s_{I_2})\ldots f_m^\alpha(s_{I_m})\big) = h(s), \text{ for all } s \in V. \tag{15}$$

Suppose finally that one of the following assumptions holds

$$G \text{ is compact, or} \tag{16}$$

$$G = (\mathbb{Q},+),(\mathbb{R},+) \text{ or } (\mathbb{C},+). \tag{17}$$

Then there exist functions $f_j\colon \prod_{i\in I_j} S_i \longrightarrow G$, $j = 1,\ldots,m$, with,

$$f_1(s_{I_1})f_2(s_{I_2})\cdots f_m(s_{I_m}) = h(s), \text{ for all } s \in V. \tag{18}$$

PROOF. Assume (16) first. By Tychonoff's Theorem, for each $j = 1,\ldots,m$, the Cartesian product

$$F_j := G^{\prod_{i\in I_j} S_i}$$

is compact (in the product topology). Since the net $\{f_1^\alpha \mid \alpha \in \Theta\} \subset F_1$, there exists a convergent subnet $\{f_1^{\alpha_1} \mid \alpha_1 \in \Theta_1\}$ with $f_1 := \lim_{\alpha_1} f_1^{\alpha_1} \in F_1$. Selecting convergent subnets in turn for each net $\{f_j^\alpha\}$ we eventually arrive at convergent subnets $\{f_j^{\alpha_m} \mid \alpha_m \in \Theta_m\}$ with $f_j := \lim_{\alpha_m} f_j^{\alpha_m} \in F_j$, each $j = 1,\ldots,m$. The required conclusion, (18), now follows from (15).

Secondly, suppose (17) holds. By equational compactness, in order to prove that (18) is solvable for the functions $f_1,\ldots,f_m$, it suffices to show that (18) is finitely solvable. We can therefore without loss of generality assume V is finite, and therefore that the sets $S_1,\ldots,S_n$ are all finite.

We can now regard G^V as a finite-dimensional vector space over G (where G is either $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$). Consider the following condition on $p \in G^V$:

$$\text{There exist } f_j\colon \prod_{i\in I_j} S_i \longrightarrow G,\ j = 1,\ldots,m, \text{ with } \sum_{j=1}^{m} f_j(s_{I_j}) = p(s), \text{ for all } s \in V. \tag{19}$$

The set of $p \in G^V$ satisfying (19) is clearly a subspace of G^V, so is closed (under pointwise convergence). From (15), h is a pointwise limit of functions satisfying (19), so therefore $p := h$ satisfies (19). But this is exactly (18). ∎

EXAMPLE (MEKLER). The following counterexample, due to [11], shows that the result can fail with $G = (\mathbb{Z},+)$. Let $S_1 = S_2 = \mathbb{N} \times \{1,2,3\}$, and $S_3 = \mathbb{N}$. Define $V \subset S_1 \times S_2 \times S_3$ as the set of all points of the form

$$\begin{array}{ccc} \big((r,1),(r,1),2r-1\big) & \big((r,1),(r,2),2r\big) & \big((r,2),(r,1),2r\big) \\ \big((r,2),(r,3),2r-1\big) & \big((r,3),(r,2),2r+1\big) & \big((r,3),(r,3),2r+2\big) \end{array}$$

https://doi.org/10.4153/CJM-1992-030-9 Published online by Cambridge University Press

as r ranges over $\mathbb{N}$, and define $h\colon V \longrightarrow \mathbb{Z}$ by

$$h(s_1, s_2, s_3) = \begin{cases} 1, & \text{if } (s_1, s_2, s_3) = \big((r,3),(r,2),2r+1\big), \text{ some } r \in \mathbb{N}, \\ 0, & \text{otherwise.} \end{cases}$$

The condition that h is decomposable is then that there exist functions u, v: $\mathbb{N} \times \{1,2,3\} \longrightarrow \mathbb{Z}$, and $w\colon \mathbb{N} \longrightarrow \mathbb{Z}$ such that for each $r \in \mathbb{N}$,

$$\begin{aligned} u(r,1) + v(r,1) + w(2r-1) &= 0, \\ u(r,1) + v(r,2) + w(2r) &= 0, \\ u(r,2) + v(r,1) + w(2r) &= 0, \\ u(r,2) + v(r,3) + w(2r-1) &= 0, \\ u(r,3) + v(r,2) + w(2r+1) &= 1, \\ u(r,3) + v(r,3) + w(2r+2) &= 0, \end{aligned} \tag{20}$$

or equivalently,

$$\begin{aligned} u(r,1) + v(r,2) + w(2r) &= 0, \\ u(r,2) + v(r,3) + w(2r-1) &= 0, \\ u(r,3) + v(r,3) + w(2r+2) &= 0, \\ v(r,1) - v(r,2) + w(2r-1) - w(2r) &= 0, \\ v(r,2) - v(r,3) + w(2r+1) - w(2r+2) &= 1, \\ 2\big(w(2r-1) - w(2r)\big) + \big(w(2r+1) - w(2r+2)\big) &= 1. \end{aligned} \tag{21}$$

Notice that having solved for w, the values of u and v may be found by back-substitution in (21). It follows, by making the substitution

$$x_r := w(2r-1) - w(2r), \text{ each } r \in \mathbb{N},$$

in the last equation in (21) and using the example after Theorem 4.3, that (21) (and therefore (20)) is finitely solvable for the variables u, v, w, but not solvable. Thus h is not decomposable.

However, since (20) is finitely solvable, we can find u^l, v^l, w^l such that putting $u := u^l, v := v^l, w := w^l$ solves (20) for all $r = 1, 2, \ldots, l$. It then follows that

$$\lim_{l \to \infty} \big(u^l(s_1) + v^l(s_2) + w^l(s_3)\big) = h(s_1, s_2, s_3), \text{for all } (s_1, s_2, s_3) \in V,$$

so h is a pointwise limit of decomposable functions, but is not decomposable.

https://doi.org/10.4153/CJM-1992-030-9 Published online by Cambridge University Press

5. **The continuous case.** In Section 3 we gave various conditions under which a bivariate function was decomposable, or in other words could be written as a product (or sum) of two univariate functions. In many circumstances the sets involved may have additional topological or measure-theoretic structure, in which case it is natural to impose the corresponding structure on the component functions. In the next two sections we will consider first the case of continuous functions, and then measurable functions. We will restrict attention to the bivariate case: the multivariate results will be analogous extensions.

Throughout this section we will suppose S_1, S_2 are topological spaces, and for a subset $E \subset S_1 \times S_2$ we write for the projections

$$
\begin{aligned}
S_1^E &:= \{ s_1 \in S_1 \mid (s_1, s_2) \in E \text{ for some } s_2 \in S_2 \}, \\
S_2^E &:= \{ s_2 \in S_2 \mid (s_1, s_2) \in E \text{ for some } s_1 \in S_1 \}.
\end{aligned} \tag{22}
$$

Clearly if E is open in $S_1 \times S_2$, S_1^E and S_2^E are open in S_1 and S_2 respectively. We shall also suppose that G is a topological group, with identity e.

The first result shows that if h is continuous on an open set E, and is decomposable, then it decomposes into a product of continuous functions.

PROPOSITION 5.1. *Suppose $E \subset S_1 \times S_2$ is open, $h: E \longrightarrow G$ is continuous, and $f_i: S_i^E \longrightarrow G$, $i = 1, 2$, with*

$$h(s_1, s_2) = f_1(s_1) f_2(s_2), \textit{ for all } (s_1, s_2) \in E.$$

Then f_1 and f_2 are continuous.

PROOF. Suppose $s_1^0 \in S_1^E$, $U \subset G$ open, and $f_1(s_1^0) \in U$. Thus for some $s_2^0 \in S_2^E$, $(s_1^0, s_2^0) \in E$, so $h(s_1^0, s_2^0) = f_1(s_1^0) f_2(s_2^0) \in U f_2(s_2^0)$. Since h is continuous, there exists an open neighbourhood V of s_1^0 in S_1, with

$$h(s_1, s_2^0) = f_1(s_1) f_2(s_2^0) \in U f_2(s_2^0),$$

or $f_1(s_1) \in U$, for all $s_1 \in V$ with $(s_1, s_2^0) \in E$. But since E is open there exists an open neighbourhood W of s_1^0 in S_1, with $(s_1, s_2^0) \in E$ for all $s_1 \in W$. Thus $V \cap W$ is an open neighbourhood of s_1^0 in S_1 with $f_1(s_1) \in U$ for all $s_1 \in V \cap W$, so f_1 is continuous. Similarly, f_2 is continuous. ∎

An easy example shows that this result may fail if E is not open. For example, suppose $S_1 := \{1, 2\}$ (with the discrete topology), $S_2 := [0, 1]$, $G = (\mathbb{R}, +)$ (with the usual topologies), and $E = \{(1, 0), (1, 1)\} \cup (\{2\} \times (0, 1])$, with $h(1, 1) = -1$ and $h \equiv 0$ elsewhere on E. Then defining $f_1(1) := 0, f_1(2) := 1$, and

$$f_2(s_2) = \begin{cases} 0, & \text{if } s_2 = 0, \\ -1, & \text{if } s_2 > 0, \end{cases}$$

we have $h(s_1, s_2) = f_1(s_1) + f_2(s_2)$ for all $(s_1, s_2) \in E$, although f_2 is not continuous. Furthermore, there is clearly no decomposition of h into continuous functions, as the given decomposition is unique up to a constant.

The other question of interest is that of uniqueness of decomposition. By Theorem 3.3, this amounts to checking that the graph associated with E is connected. In this topological setting we can be more concrete, using topological connectedness.

PROPOSITION 5.2. *Suppose* $f_i \colon S_i^E \longrightarrow G$, $i = 1,2$ *are continuous, with* $f_1(s_1)f_2(s_2) = e$, *for all* $(s_1, s_2) \in E$. *Suppose one of the following conditions holds:*

(i) S_1^E *is connected, and for all* $s_1 \in S_1^E$, *there exists* $s_2 \in S_2$ *and an open neighbourhood* U *of* s_1 *with* $U \times \{s_2\} \subset E$.

(ii) S_2^E *is connected, and for all* $s_2 \in S_2^E$, *there exists* $s_1 \in S_1$ *and an open neighbourhood* V *of* s_2 *with* $\{s_1\} \times V \subset E$.

(iii) E *is open and connected.*

Then for some contant $k \in G$, $f_1 \equiv k$ *on* S_1^E *and* $f_2 \equiv k^{-1}$ *on* S_2^E.

PROOF. (i) Suppose $s_1^0 \in S_1^E$. Pick $s_2 \in S_2$ and an open neighbourhood U of s_1^0 with $U \times \{s_2\} \subset E$. Then for all $s_1 \in U, f_1(s_1) = f_2(s_2)^{-1}$. Thus f_1 is locally constant on S_1^E, so by connectedness is constant, [18, p. 105]. The result follows.

(ii) Similar.

(iii) This condition implies both (i) and (ii), since S_1^E and S_2^E are the continuous projections of E onto S_1 and S_2 respectively, so are connected [18, p. 97]. ∎

This result clearly resolves the uniqueness question since if

$$p_1(s_1)p_2(s_2) = q_1(s_1)q_2(s_2), \text{for all } (s_1, s_2) \in E,$$

then

$$\big(q_1(s_1)^{-1}p_1(s_1)\big)\big(p_2(s_2)q_2(s_2)^{-1}\big) = e,$$

so for some $k \in G$,

$$q_1(s_1)^{-1}p_1(s_1) = k, \text{for all } s_1 \in S_1^E,$$
$$p_2(s_2)q_2(s_2)^{-1} = k^{-1}, \text{for all } s_2 \in S_2^E,$$

or $q_1(s_1) = p_1(s_1)k^{-1}$ and $q_2(s_2) = kp_2(s_2)$.

EXAMPLE. Suppose $S_1 = S_2 = [0,1]$ with the usual topology, and $E \subset [0,1]^2$ is open with the main diagonal $\{(s,s) \mid s \in [0,1]\} \subset E$. Then it is easy to see (i) applies.

6. **The measurable case.** The previous section showed that if a bivariate function was continuous and decomposable then it decomposed into continuous univariate functions. We also showed uniqueness results. In this section we will develop the analogue results for the measurable case. Again we will restrict our attention to the bivariate case: extensions to multivariate results will be analogous.

We suppose (S_1, μ_1) and (S_2, μ_2) are measure spaces, and G is a topological group with identity e. We will take the Borel sets of a topological space to be the σ-algebra generated by the open sets (as in [15]) rather than the (possibly smaller) σ-algebra generated by the compact sets (as in [9]). In a locally compact, σ-compact Hausdorff space (such as $\mathbb{R}$) the two will be the same (see [9, p. 219]).

We shall suppose $E \subset S_1 \times S_2$ is $\mu_1 \times \mu_2$ - measurable and we will lose no generality by assuming $S_1^E = S_1$ and $S_2^E = S_2$ (defined as in (22)). Our first result is the analogue of Proposition 5.1. We show that, under a simple condition on E, if a measurable function on E is decomposable, then it decomposes into a product of measurable functions.

PROPOSITION 6.1. Suppose $E \subset S_1 \times S_2$ is a countable union of measurable rectangles, $h: S_1 \times S_2 \longrightarrow G$ is measurable $(\mu_1 \times \mu_2)$, and $f_1: S_1 \longrightarrow G, f_2: S_2 \longrightarrow G$, with

$$h(s_1, s_2) = f_1(s_1)f_2(s_2), \quad (\mu_1 \times \mu_2)\text{ -a.e. on } E.$$

Then f_i is μ_i-measurable, $i = 1, 2$.

PROOF. Suppose $E = \cup_{j=1}^{\infty}(S_1^j \times S_2^j)$, where each S_i^j is μ_i-measurable and non-null. Then for each j,

$$h(s_1, s_2) = f_1(s_1)f_2(s_2), \quad (\mu_1 \times \mu_2)\text{ -a.e. on } S_1^j \times S_2^j,$$

so by [9, p. 147], for μ_1-almost all $s_1 \in S_1^j$,

$$h(s_1, s_2) = f_1(s_1)f_2(s_2) \quad \mu_2\text{ -a.e. on } S_2^j.$$

In particular, there exists $s_1^j \in S_1^j$ with

$$f_2(s_2) = (f_1(s_1^j))^{-1}h(s_1^j, s_2), \quad \mu_2\text{ -a.e. on } S_2^j,$$

so $f_2|_{S_2^j}$ is μ_2-measurable, since h is measurable, and every section of a measurable function is measurable (*cf.* [9, p. 142]).

Now by assumption, $\cup_{j=1}^{\infty} S_2^j = S_2$. Suppose $U \subset G$ is open. Then

$$f_2^{-1}(U) = \cup_{j=1}^{\infty}(f_2|_{S_2^j})^{-1}(U)$$

which is measurable. Thus f_2 is measurable, and similarly, so is f_1. ∎

NOTE. At least when $G = \mathbb{R}$ it makes no difference whether we assume h is $(\mu_1 \times \mu_2)$-measurable or measurable with respect to the completion $(\overline{\mu_1 \times \mu_2})$, since in the latter case there exists a $(\mu_1 \times \mu_2)$-measurable function h' with $h = h'$, $(\overline{\mu_1 \times \mu_2})$-a.e. [15, p. 145].

Certainly not every measurable subset of $S_1 \times S_2$ will be a countable union of measurable rectangles. However, the following result gives a simple condition.

LEMMA 6.2. *Suppose S_1, S_2 are separable metric spaces, with Borel measures μ_1 and μ_2 respectively. If $E \subset S_1 \times S_2$ is open then it is a countable union of measurable rectangles.*

PROOF. For each $e = (e_1, e_2)$ in E choose open neighbourhoods S_1^e of e_1 and S_2^e of e_2 with $S_1^e \times S_2^e \subset E$. These measurable rectangles form an open cover of E, and since E is Lindelöf there is a countable subcover. ∎

We next prove a uniqueness result which is analogous to Proposition 5.2. We first need an analogue of the result that a locally constant function on a connected space is constant.

LEMMA 6.3. *Suppose S is a connected, locally compact, σ-compact Hausdorff space, and μ is a Borel measure on S with full support, in other words every nonempty open set in S has strictly positive measure. Suppose $f: S \longrightarrow G$ is measurable and locally a.e. constant: given any $s \in S$ there exists a neighbourhood of s on which f is constant a.e. Then f is constant a.e. on S.*

PROOF. Define a function $F: S \longrightarrow G$ as follows. Given $t \in S$, there exists an open neighbourhood U_t of t and a constant $g_t \in G$ with $f(s) = g_t$, a.e. on U_t: then define $F(t) := g_t$. The function F is well-defined, since if U' is another neighbourhood of t and $f(s) = g'$, a.e. on U', then since μ has full support, $\mu(U_t \cap U') > 0$, so $g_t = g'$.

Now F is locally constant on S. To see this, suppose $s_0 \in S$. By assumption, there exists an open neighbourhood U_0 of s_0 and $g_0 \in G$ with $f(s) = g_0$, a.e. on U_0, and then $F(s_0) = g_0$. Suppose $s_1 \in U_0$. Then again by assumption, there exists a neighbourhood U_1 of s_1 and $g_1 \in G$ with $f(s) = g_1$, a.e. on U_1, and then $F(s_1) = g_1$. Since μ has full support, $\mu(U_0 \cap U_1) > 0$, so $g_0 = g_1$. Thus $F(s) = F(s_0)$ for all $s \in U_0$, as required.

It follows immediately that F is continuous, and therefore constant, since S is connected. Suppose $F(s) = k \in G$ for all $s \in S$. Thus $f(s) = k$, a.e. on U_t for all $t \in S$. Now $\cup_{t\in S} U_t$ is an open cover of S, and since S is σ-compact it is Lindelöf [10, p. 172] so there exists a countable subcover $\cup_{j=1}^{\infty} U_{t_j}$. But $f(s) = k$, a.e. on each U_{t_j}, so $f(s) = k$, a.e. on S. ∎

THEOREM 6.4. *Suppose S_1 and S_2 are locally compact Hausdorff spaces with Borel measures μ_1 and μ_2 respectively, each with full support. Suppose E is a measurable subset of $S_1 \times S_2$, $f_1: S_1 \longrightarrow G$ and $f_2: S_2 \longrightarrow G$ are measurable, and*

$$f_1(s_1)f_2(s_2) = e, \quad (\mu_1 \times \mu_2)\text{-a.e. on } E.$$

Suppose the following conditions hold:

(1) S_1 is σ-compact and connected.

(2) For each $s_1 \in S_1$ there exists $s_2 \in S_2$ with (s_1, s_2) in the interior of E.

(3) $\mu_1\{s_1 \in S_1 \mid (s_1, s_2) \in E\} > 0$, μ_2-a.e. on S_2.

Then for some constant $k \in G$,

$$f_1(s_1) = k, \quad \mu_1\text{-a.e. on } S_1,$$

and

$$f_2(s_2) = k^{-1}, \quad \mu_2\text{-a.e. on } S_2.$$

PROOF. Suppose $t_1 \in S_1$. By Condition (2), there exists $t_2 \in S_2$ with (t_1, t_2) in the interior of E, so for some open neighbourhoods U_1 and U_2 of t_1 and t_2 respectively, $U_1 \times U_2 \subset E$. Now

$$f_1(s_1)f_2(s_2) = e, \quad (\mu_1 \times \mu_2)\text{-a.e. on } U_1 \times U_2.$$

Therefore by [9, p. 147], for almost all $s_2 \in U_2$,

$$f_1(s_1)f_2(s_2) = e, \quad \mu_1 \text{-a.e. on } U_1.$$

Since μ_2 has full support, so $\mu_2(U_2) > 0$, there exists $s_2^0 \in U_2$ with

$$f_1(s_1) = (f_2(s_2^0))^{-1}, \quad \mu_1 \text{-a.e. on } U_1.$$

Thus f_1 is locally a.e. constant, so constant a.e. on S_1, by Condition 1 and Lemma 6.3. Thus for some $k \in G, f_1(s_1) = k$, μ_1 -a.e. on S_1.

Finally, by assumption,

$$(\mu_1 \times \mu_2)\{\,(s_1, s_2) \in E \mid f_1(s_1)f_2(s_2) \neq e\} = 0,$$

so by [9, p. 147], for μ_2-almost all $s_2 \in S_2$,

$$\mu_1\{\,s_1 \in S_1 \mid (s_1, s_2) \in E, f_1(s_1)f_2(s_2) \neq e\} = 0,$$

and thus by Condition 3,

$$\mu_1\{\,s_1 \in S_1 \mid (s_1, s_2) \in E, f_1(s_1)f_2(s_2) = e\} > 0.$$

But $f_1(s_1) = k$, μ_1 -a.e. on S_1, so certainly there exists $s_1^0 \in S_1$ with (s_1^0, s_2) in E, $f_1(s_1^0)f_2(s_2) = e$, and $f_1(s_1^0) = k$, so $f_2(s_2) = k^{-1}$. Since this holds μ_2 -a.e. on S_2, the result follows. ∎

Just as in Section 5, this resolves the uniqueness question, since if

$$p_1(s_1)p_2(s_2) = q_1(s_1)q_2(s_2), \text{ a.e. on } E,$$

then

$$\big(q_1(s_1)^{-1}p_1(s_1)\big)(p_2(s_2)q_2(s_2)^{-1}) = e, \text{ a.e. on } E,$$

so for some $k \in G$,

$$q_1(s_1)^{-1}p_1(s_1) = k, \text{ a.e. on } S_1,$$
$$p_2(s_2)q_2(s_2)^{-1} = k^{-1}, \text{ a.e. on } S_2,$$

or $q_1 = p_1k^{-1}$, a.e. and $q_2 = kp_2$, a.e.

Clearly we could interchange S_1 and S_2.

EXAMPLES. The conditions in the above result will hold in each of the following cases.

(i) $S_1 = S_2 = I$, a closed interval in $\mathbb{R}$ with Lebesgue measure, E an open set in $I \times I$ containing the main diagonal $\{\,(s, s) \mid s \in I\}$.

(ii) Suppose S_1 is σ-compact and connected, and

$$E = \{\,(s_1, s_2) \mid p(s_1, s_2) > 0\},$$

where $p: S_1 \times S_2 \longrightarrow \mathbb{R}$ is lower semicontinuous (so E is open). Suppose there exists surjective $q: S_1 \longrightarrow S_2$ with $p\big(s_1, q(s_1)\big) > 0$ for all $s_1 \in S_1$. Then the conditions hold.

(iii) $S_1 = S_2 = [0,1]$ and $p(s_1, s_2) = |s_1 - s_2|$. Then

$$q(s_1) := \begin{cases} s_1 + \frac{1}{2}, & s_1 \in [0, \frac{1}{2}], \\ s_1 - \frac{1}{2}, & s_1 \in (\frac{1}{2}, 1], \end{cases}$$

works in (ii).

We conclude with a discussion of a problem from the introduction. Suppose S_1 and S_2 are compact metric spaces, with Borel measures μ_1 and μ_2 respectively, each with full support, and suppose $E \subset S_1 \times S_2$ is open. Consider the subspace $\mathcal{D}$ of $L_1(E, \mu_1 \times \mu_2)$ consisting of these integrable functions $h: E \longrightarrow \mathbb{R}$ which can be written

$$h(s_1, s_2) = f_1(s_1) + f_2(s_2), \text{ a.e. on } E,$$

where $f_1: S_1 \longrightarrow \mathbb{R}$ and $f_2: S_2 \longrightarrow \mathbb{R}$ are measurable. We shall show that $\mathcal{D}$ is closed.

Suppose therefore that $h_n \in \mathcal{D}, n = 1, 2, \ldots$, that $h \in L_1(E, \mu_1 \times \mu_2)$, and $\|h_n - h\|_1 \longrightarrow 0$. Thus for some measurable functions $f_1^n = S_1 \longrightarrow \mathbb{R}$ and $f_2^n = S_2 \longrightarrow \mathbb{R}$,

$$h^n(s_1, s_2) = f_1^n(s_1) + f_2^n(s_2), \text{ a.e. on } E.$$

By [15, 3.12], for some subsequence (n'), $h^{n'} \longrightarrow h$ pointwise a.e. on E. It now follows by Corollary 2.4 that there exist functions $f_1: S_1 \longrightarrow \mathbb{R}$ and $f_2: S_2 \longrightarrow \mathbb{R}$ with

$$h(s_1, s_2) = f_1(s_1) + f_2(s_2), \text{ a.e. on } E.$$

Since S_1 and S_2 are compact metric spaces they are separable, so Lemma 6.2 shows E is a countable union of measurable rectangles. Now Proposition 6.1 applies to show f_1 and f_2 are measurable, so $h \in \mathcal{D}$ as required.

REFERENCES

1. C. Berge, *Graphs and hypergraphs*, North-Holland, Amsterdam, 1975.

2. J. M. Borwein, A. S. Lewis and R. Nussbaum, *Entropy and DAD problems*, in preparation.

3. I. Csiszár, *I-Divergence geometry of probability distributions and minimization problems*, Annals of Probability **3**(1975), 146–158.

4. A. Devinatz, *Advanced calculus*, Holt, Rinehart and Winston, New York, 1968.

5. G. M. Engel and H. Schneider, *Cyclic and diagonal products on a matrix*, Linear Algebra and its Applications **7**(1973), 301–335.

6. M. v. Golitschek, *Optimal cycles in doubly weighted graphs and approximation of bivariate functions by univariate ones*, Numerische Mathematik **39**(1982), 65–84.

7. M. v. Golitschek and H. Schneider, *Applications of shortest path algorithms to matrix scalings*, Numerische Mathematik **44**(1984), 111-120.

8. G. Grätzer, *Universal algebra*, Springer-Verlag, New York, 1979.

9. P. R. Halmos, *Measure theory*, Van Nostrand, New York, 1950.

10. J. L. Kelley, *General topology*, Van Nostrand, New York, 1955.

11. A. Mekler, private communication, (1990).

12. R. D. Nussbaum, *Hilbert's projective metric and iterated nonlinear maps*, A.M.S., Providence, R.I., 1988.

13. ———, *Iterated nonlinear maps and Hilbert's projective metric II*, A.M.S., Providence, R.I., 1989.

14. U. R. Rothblum and H. Schneider, *Characterizations of optimal scalings of matrices*, Mathematical Programming **19**(1980), 121–130.

15. W. Rudin, *Real and complex analysis*, McGraw-Hill, New York, 1966.

16. M. H. Schneider, *Matrix scaling, entropy minimization and conjugate duality I, existence conditions*, Linear Algebra and its Applications **114**(1989), 785–813.

17. M. H. Schneider, *Matrix scaling, entropy minimization and conjugate duality II, the dual problem*, Mathematical Programming B **48**(1990), 103–124.

18. W. A. Sutherland, *Introduction to metric and topological spaces*, Oxford University Press, Oxford, 1975.

Department of Combinatorics & Optimization
University of Waterloo
Waterloo, Ontario, N2L 3G1