Representation of associative functions

By CHO-HSIN LING (Chicago, Ill.)

1. Introduction

In the course of their work on statistical metric spaces [17, 19, 20], B. SCHWEIZER and A. SKLAR were led to consider a class of 2-place real functions which play an important role in connection with generalized triangle inequalities for such spaces. Following K. MENGER [15], they called these functions *triangular norms* (briefly, t-norms).

In attacking the problem of representation of t-norms, SCHWEIZER and SKLAR found that the most crucial property of such norms is their associativity. Thus the representation problem for t-norms leads naturally to the consideration of representation for arbitrary associative functions, which in turn is one aspect of a larger problem: the representation of multi-place functions in general by composition ("superposition") of "simpler" functions, fixed functions, and functions of fewer places. We will briefly survey the larger problem before outlining the specific questions considered in this paper. [1])

In the following discussions, F will denote a 2-place real function to be represented, S a 2-place real, associative function, and $\sum$ the 2-place sum function. (I. e., $\sum(a, b) = a+b$).

Representation of Two-place Functions in general

W. SIERPIŃSKI was one of the first to study the representation of multiplace functions by superpositions of functions of fewer places [21, 22]. One of his representation theorems can be formulated as follows:

(1. 1) THEOREM. *There exist two* 1*-place functions h and k, such that for every* 2*-place function F, there exists a* 1*-place function g such that*

$$F(x, y) = g\big(h(x)+k(y)\big). \tag{SI}$$

For *symmetric* 2-place functions, it is reasonable to suppose, but not immediate from (SI) that we can take $h=k$, and so write

$$F(x, y) = g\big(h(x)+h(y)\big).$$

---

[1]) This work was supported in part by the National Science Foundation under Grants G 21110 and GP 2541.

The supposition is in fact true. YAO and LING [25], using a Hamel basis argument, have proved the following:

(1. 2) THEOREM. *There exists a* 1*-place function* f*, such that for every symmetric* 2*-place function* F*, there exists a* 1*-place function* g *such that*

$$\text{(YL)} \qquad F(x, y) = g\big(\sum(f(x), f(y))\big) = g\big(f(x)+f(y)\big).$$

Representation of Continuous Functions

HILBERT, in his 13[th] problem (1900) conjectured that not all continuous (Hilbert actually says "analytic") 3-place functions are superpositions of [continuous] 2-place functions. Recently, V. I. ARNOLD [4], using results of his own and A. N. KOLMOGOROV, disproved Hilbert's conjecture. KOLMOGOROV [14], going further, obtained the remarkable result that every continuous n-place function can be represented by a superposition of $\sum$'s and continuous 1-place functions. In particular, for $n=2$, his theorem can be formulated as follows:

(1. 3) THEOREM. *There exist ten continuous strictly increasing* 1*-place functions,* $f_1, f_2, \ldots, f_{10}$ *from the interval* $[0, 1]$ *to itself, such that for every continuous* 2*-place function* F *on the unit square, there exist five continuous* 1*-place functions,* $g_1, g_2, \ldots, g_5$ *such that*

$$\text{(K)} \qquad F(x, y) = \sum_{m=1}^{5} g_m\big(\sum(f_m(x), f_{m+5}(y))\big) = \sum_{m=1}^{5} g_m\big(f_m(x)+f_{m+5}(y)\big).$$

(The restriction to the unit interval and the unit square is a matter of convenience and is not essential to the validity of the result.)

It should be remarked that the outher summation cannot be eliminated in (K). In fact ARNOLD [5] has proved not only that there exist continuous 2-place functions that are not representable in the form $g\big(h(x)+k(y)\big)$ $(g, h, k$ continuous) [2]) but that the set of functions that can be so represented is nowhere dense in the space of all continuous 2-place functions.

Representation of Continuous and Associative Functions

When F is associative as well as continuous, then the representation (K) can be dramatically simplified. Such a representation was first obtained (*sui generis*, of course, not as a special case of (K)) by ABEL in 1828 [1], under the additional assumptions of symmetry, strict monotonicity and differentiability. His representation takes the form:

$$\text{(A1)} \qquad S(x, y) = f^{-1}\big(\sum(f(x), f(y))\big) = f^{-1}\big(f(x)+f(y)\big),$$

[2]) Among the non-representable functions is the associative function Min (see Section 6).

where f is continuous and strictly monotone, hence invertible, and f^{-1} is the inverse of f.

L. E. J. Brouwer [6] and É. Cartan [7] in working with continuous groups, in effect achieved the same representation without the assumption of differentiability.

Recent work has been done by J. Aczél [2, 3], D. Tamari [23], and W. M. Faucett [11].

In 1948, Aczél proved the following:

(1. 4) Theorem. *Let A be an open or half-open (but not closed) real interval and $S: A\times A \to A$ be a 2-place function. Suppose that S is continuous and strictly increasing in each of its places. Suppose further that S is associative, i. e., satisfies the functional equation*

$$S(S(x, y), z) = S(x, S(y, z)), \tag{A2}$$

for all x, y, z in A. Then there exists a function f, defined, continuous and strictly monotone in A such that S is representable in the form (A1).

(1. 5) *Remark*. If S satisfies the hypotheses of Theorem (1. 4), then (A1) shows that S is symmetric. Thus, every continuous, strictly increasing, associative function is symmetric, i. e., commutative.

Representation of Non-strict Associative Functions

The requirement of strict monotonicity in Aczél's theorem is rather severe, and it would be desirable to weaken it whenever possible. The author showed in her doctoral dissertation that this is indeed feasible in many cases: the result is the main theorem or this paper (Theorem (3. 3)) and its dual (Theorem (3. 5)).

It was later found that these theorems can be derived from previous results of P. S. Mostert and A. L. Shields [16], these results resting in turn on the work of A. D. Wallace [24] and W. M. Faucett [11], all in the general area of topological semigroups. Since our original proof used only the tools of elementary analysis, it has seemed desirable to present both proofs of the main theorem. Accordingly, the original elementary proof appears in Section 4, and the second proof in Section 5. In Section 5, we also show that the arguments can be reversed to yield the relevant results of Mostert and Shields as consequences of the main theorem.

In Section 6, we pursue the connection between Aczél's theorem and the main theorem by showing that every associative function satisfying the hypotheses of the main theorem is obtainable as a limit of "Aczélian" associative functions. Section 7 is devoted to various *non*-representability results, which indicate that the results of the main theorem are in a sense "best-possible".

I wish to take this opportunity to express my gratitude to Professor A. Sklar for his initially suggesting the problems investigated in this paper and his subsequent patient guidance and encouragement in its writing. I wish also to extend my thanks to Professors J. Aczél and A. D. Wallace for their many generous comments and valuable criticisms, and to Professor B. Schweizer for his many helpful suggestions.

2. Conventions and preliminaries

We list conventions that will be adhered to throughout this paper.

We work throughout with the extended real number system (the ordinary finite real numbers together with $\pm\infty$). Thus, a "real number" may be finite or infinite.

The capital letters I, J, K, R will be used for closed intervals. In particular, we will always denote the closed unit interval $[0, 1]$ by I and the closed interval $[0, \infty]$ by R. If J is a closed interval, then J^* will denote the corresponding half-open interval with the right endpoint removed, and J° the corresponding open interval with both endpoints removed. Thus, R^* denotes the set of non-negative finite real numbers.

The small letters, i, m, n will be used for indices ranging over sets of positive integers.

The capitalized letters F, S, T will be used for 2-place functions. Furthermore, the use of the letters S and T will always imply associtivity (cf. equation (A2)) of the functions. $S(x, y)$ will be written as $x \cdot y$ if there is no danger of possible confusion. Also, x^n will always mean $S(x^{n-1}, x)$, and the mapping: $x \to x^n$ will be denoted by s_n. Other 1-place functions will be denoted by the small letters f, g, h, k, j, r_n. In particular, the letter j will always be used for the identity function on any domain in question. The domain and range of a function f are denoted by Dom (f) and Ran (f), respectively. Composition of 1-place functions will be denoted by juxtaposition. The words "increasing", "decreasing" and "monotone" will be used in the sense of "strictly increasing", etc.

Following B. SCHWEIZER and A. SKLAR [17, 18, 19, 20], we introduce the notions of right-subinverse and left-neutralizer.

(2. 1) Definition. A right-subinverse of a function f is a function g such that

(RSI) $$\operatorname{Dom}(g) = \operatorname{Ran}(f),\ \operatorname{Ran}(g) \subseteqq \operatorname{Dom}(f), \text{ and } fg = j \text{ on } \operatorname{Ran}(f),$$

i. e., $f(g(x)) = x$ for all x in Ran (f).

Any function, whether it has an inverse or not, has at least one right-subinverse. (This assertion is in fact equivalent to the axiom of choice [20].) Moreover, if g is a right-subinverse of f, then g is itself invertible, and the inverse of g is the (generally proper) restriction of f to Ran (g).

(2. 2) Definition. A left-neutralizer of a real function f is a function g such that

(LN) $$gf \subseteqq j, \text{ i. e., } g(f(x)) = x \text{ for all } x \text{ in } \operatorname{Dom}(gf).$$

In the case of continuous monotonic functions, it is convenient to single out a particular left-neutralizer, as follows:

(2. 3) Definition. Let $f: J \to R$ be a continuous and increasing function from $[a, b]$ to $[0, \infty]$. The pseudo-inverse of f is the function $g: R \to J$ defined by

(PII) $$g(x) = \begin{cases} a, & \text{if } x \text{ is in } [0, f(a)], \\ f^{-1}(x), & \text{if } x \text{ is in } [f(a), f(b)], \\ b, & \text{if } x \text{ is in } [f(b), \infty]. \end{cases}$$

If f is continuous and decreasing, the pseudo-inverse g is defined by

$$\text{(PID)} \qquad g(x) = \begin{cases} b, & \text{if } x \text{ is in } [0, f(b)], \\ f^{-1}(x), & \text{if } x \text{ is in } [f(b), f(a)], \\ a, & \text{if } x \text{ is in } [f(a), \infty]. \end{cases}$$

It is easy to see that the pseudo-inverse g is continuous and weakly monotone. Furthermore, it is immediate that, if g is the pseudo-inverse of f, then f is a right-subinverse of g.

3. Representation of associative functions

For convenience we restate Aczél's theorem here as follows:

(3. 1) THEOREM. [1]. *Let A be an open or half-open (but not closed) interval and $S: A \times A \to A$ be an associative function satisfying the following conditions:*
(i) *S is continuous.*
(ii) *S is increasing in each place.*
Then there exists a continuous and monotone function $f: A \to R$, such that S is representable in the form

$$\text{(A3)} \qquad S(x, y) = g\big(f(x) + f(y)\big),$$

where g is the inverse of f.

(3. 2) Definition. A semigroup S satisfying the hypotheses of Aczél's theorem will be called Aczélian.

As stated in the introduction, our purpose is to weaken the isotonicity condition (ii). Now it is not possible to do this in general unless additional conditions are introduced (the t-norm Min shows this: see Section 6). We have tried to keep these new conditions "natural", in the sense that they actually hold in the most useful examples of Aczél's theorem.

There is another point to be observed: Aczél's theorem really involves three different cases: (1) A is open, i. e., of the form (a, b); (2) A is of the form $(a, b]$; (3) A is of the form $[a, b)$. It turns out that the generalisation of Aczél's theorem involves four different cases: (1), (2), (3) as in Aczèl's theorem, and (4): A is closed, which is expressedly excluded in Aczél's theorem. Moreover, it is most convenient to work with case (4); and since, as we shall see, the other cases can be reduced to this, it is case (4) alone that will be treated in detail.

Generalization of Aczél's Theorem

(3. 3) **Main Theorem.** *Let J be a closed interval $[a, b]$ of the extended real line and $S: J \times J \to J$ be an associative function satisfying the following conditions:*

(1) *S is continuous,*
(2) *S is nondecreasing in each place,*
(3) *The endpoint a is a left unit, i. e., $a \cdot x = x$ for all x in J,*
(4) *For all x in J°, $x^2 > x$.*

Then there exists a continuous and increasing function $f: J \to R$ such that S is representable in the form

$$S(x, y) = g(f(x) + f(y)), \tag{L}$$

where g is the pseudo-inverse of f.

(3. 4) *Remark*. The requirement that J be closed does not really restrict us in any way. For example, if we begin with $(a, b]$ or (a, b), we can replace postulate (3) by (3a):

(3a) For all y in $(a, b]$ or in (a, b),

$$\lim_{x \to a} S(x, y) = y, \quad \lim_{x \to a} S(x, x) = a.$$

Hence, upon adjoining the endpoint a to $(a, b]$, or both endpoints a and b to (a, b), by continuity, we immediately recover the hypotheses of Theorem (3. 3).

In later sections (Sections 4 and 5), we shall give two proofs of the foregoing theorem.

Dualization of the Main Theorem

For various applications, it is convenient to have a dual version of the main theorem at hand.

(3. 5) **Dual of the Main Theorem.** *Let J be a closed interval $[a, b]$ of the extended real line and $S: J \times J \to J$ be an associative function satisfying the following conditions:*

(1) *S is continuous,*
(2) *S is nondecreasing in each place,*
(3) *The endpoint b is a left unit, i. e., $b \cdot x = x$ for all x in J,*
(4) *For all x in J°, $x^2 < x$.*

Then there exists a continuous and decreasing function $f: J \to R$ such that S is representable in the form

$$S(x, y) = g(f(x) + f(y)), \tag{L*}$$

where g is the pseudo-inverse of f.

(3. 6) **Theorem.** *The main theorem implies its dual, and conversely.*

Proof. Let S satisfy all the hypotheses of the dual theorem. Let $h: J \to J$ be any order-reversing homeomorphism of J with itself. Then both h and its inverse k are continuous and decreasing.

We define a 2-place function $T: J \times J \to J$ as follows:

$$T(x, y) = kS(h(x), h(y)).$$

We show first that T satisfies the hypotheses of the main theorem. T is associative by invariance under isomorphism. T is continuous because it is the composite of continuous functions k, S, h.

T is nondecreasing because kS is nonincreasing in each place, being the composite of decreasing and nondecreasing functions k and S. Finally $kS(h(x), h(y))$ is nondecreasing because it is the composite of a nonincreasing function kS with a decreasing function h.

The endpoint a is a left unit because (i) $S(b, x)=x$ for all x, (ii) $S(h(a), h(y))=h(y)$ for all y, whence $kS(h(a), h(y))=y$ for all y, i. e., $T(a, y)=y$ for all y.

Finally let x be in J°, so that $h(x)$ is in J°. Then $S(h(x), h(x))<h(x)$. By the antitonicity of k, $kS(h(x), h(x))>x$, i. e., $T(x, x)>x$.

Now applying the main theorem to T, we obtain $T(x, y) = g(f(x)+f(y))$, where g is the pseudo-inverse of f.

Writing $u=h(x)$, and $v=h(y)$, whence $x=k(u)$ and $y=k(v)$, the representation of T by the generator f becomes a representation of S by the generator $fk: J\to R$ such that

$$S(u, v) = hg(fk(u)+fk(v)).$$

Since $f: J\to R$ is continuous and increasing, fk is continuous and decreasing. Also remember that $fk(a)=f(b)$ and $fk(b)=f(a)$. We have the following:

(i) If x is in $[0, f(a)]$, then $g(x)=a$. Therefore if x is in $[0, fk(b)]$, then $hg(x)=b$.

(ii) If x is in $[f(b), \infty]$, then $g(x)=b$. Therefore if x is in $[fk(a), \infty]$, then $hg(x)=a$.

(iii) If x is in $[f(a), f(b)]$, then $g(x)=f^{-1}(x)$. Therefore if x is in $[fk(b), fk(a)]$, then $hg(x)=k^{-1}f^{-1}(x)=(fk)^{-1}(x)$. It follows that hg is the pseudo-inverse of fk.

The converse is proved in the same manner, and this completes the proof of Theorem (3. 6).

(3. 7) Definition. A semigroup S satisfying the hypotheses of the main theorem or its dual will be called Archimedean. (Cf. Lemma (4. 2).)

(3. 8) Definition. A function f solving equation (A3) or (L), $(L)^*$ is called an additive generator (or simply generator) of S. (Cf. SCHWEIZER and SKLAR [20].)

(3. 9) Definition. An Archimedean semigroup is called "properly Archimedean", if every additive generator is unbounded.

We could equivalently say: S is properly Archimedean if the function g in (L) or $(L)^*$ is strictly monotonic, therefore the inverse of f; otherwise S is *improperly* Archimedean.

Since any properly Archimedean S is easily seen to be Aczélian on J°, we will use the two terms: properly Archimedean, Aczélian, interchangeably.

(3. 10) **Converse of the Main Theorem.** *Let J be a closed interval of the real numbers, $f: J\to R$ be a continuous and increasing function, and $g: R\to J$ the pseudo-inverse of f. Then the 2-place function S defined by*

$$S(x, y) = g(f(x)+f(y)), \tag{CL}$$

is an Archimedean semigroup.

PROOF. By straightforward verfication, of the postulates (1)—(4) and associativity.

(3. 11) *Remark.* The converse to the main theorem (or rather the converse to the dual of main theorem) is essentially due to B. SCHWEIZER and A. SKLAR [20]. They phrased it in terms of triangular norms, but it can easily be translated into the language of Archimedean semigroups.

Application of Representation Theorems to T-norms

We illustrate the scope of these theorems with examples (and counterexamples) drawn from the class of t-norms and their duals.

(3. 12) Definition. A triangular norm (t-norm) is an associative function $T: I \times I \to I$ satisfying the following conditions:

(SS1) T is nondecreasing in each of its places,

(SS2) T is commutative,

(SS3) The endpoint 1 is a unit, i. e.,

$$1 \cdot x = x \cdot 1 = x \text{ for all } x \text{ in } I,$$

(SS4) $T(0, 0) = 0$.

(3. 13) Definition. A t-norm is strict if T satisfies the following additional conditions:

(SS5) T is continuous,

(SS6) T is increasing, in each of its places, on

$$(0, 1] \times (0, 1].$$

(3. 14) Definition. A t-norm T is called Archimedean if it is Archimedean as a semigroup (cf. Definition (3. 7)).

It is readily seen that the class of strict t-norms is a proper subclass of Archimedean t-norms.

(3. 15) *Examples.* Of particular importance are the t-norms T_w, T_0, Prod, and Min, defined respectively as follows:

$$T_w(x, y) = \begin{cases} x, & \text{if } y = 1, \\ y, & \text{if } x = 1, \\ 0, & \text{otherwise.} \end{cases}$$

$$T_0(x, y) = \max(x + y - 1, 0).$$

$$\text{Prod}\,(x, y) = xy \text{ (the ordinary product).}$$

$$\text{Min}\,(x, y) = \begin{cases} x, & \text{if } x \leqq y, \\ y, & \text{if } y \leqq x. \end{cases}$$

Of these t-norms, only Prod is strict, only Prod and T_0 are Archimedean, and only T_w is discontinuous.

The duals (conorms [20]) of these t-norms are also of importance. In particular we will need the dual S_0 of T_0, where

$$S_0(x, y) = \text{Min}\,(x + y, 1).$$

The following theorem applies Theorem (3. 5) to the class of Archimedean t-norms.

(3. 16) **Theorem.** *Let T be an Archimedean t-norm. Then there exists a continuous and decreasing function $f: I \to R$ such that T is representable in the form*

$$T(x, y) = g(f(x) + f(y)), \tag{LT)*}$$

where $g: R \to I$ is the pseudo-inverse of f, and $f(1) = 0$.

PROOF. T satisfies all the hypotheses of Theorem (3. 5). Therefore, there exists a continuous and decreasing function $f: I \to R$ solving equation (LT)*.

Let $f(0)=Z$ and $f(1)=W$. Then $Z > W$.

We want to show that $f(1)=W=0$.

Recall that on $[0, W]$ and only on $[0, W]$ is the value of g equal to 1. Now we have $1 = T(1, 1) = g(f(1)+f(1)) = g(2W)$, whence $2W$ is in $[0, W]$. This is only possible if $W=0$.

(3. 17) *Example.* Since the t-norm T_0 is Archimedean, it has a generator $f: I \to R$, e. g., $f(x) = 1-x$ for all x in I. The pseudoinverse $g: R \to I$ of f is given by:

$$g(x) = \begin{cases} 1-x, & \text{if } x \leqq 1, \\ 0, & \text{if } x \geqq 1. \end{cases}$$

The following identity is obvious:

$$T_0(x, y) = \max(x+y-1, 0) = g((1-x)+(1-y)).$$

Theorem (3. 16) has of course a dual which applies to conorms. In particular, since S_0 is Archimedean, it has an additive generator f_0. Both f_0 and its pseudoinverse g_0 are readily found (either by dualizing from T_0 or directly); accordingly we have the following example:

$$S_0(x, y) = g_0(f_0(x)+f_0(y)),$$

where

(3. 18) $$f_0(x)=x \text{ on } I, \text{ and}$$

(3. 19) $$g_0(x)=\mathrm{Min}(x, 1) \text{ on } R.$$

The t-norm Min

The t-norm Min provides an interesting counterexample to Theorem (3. 5). Min satisfies all the hypotheses except one (viz., $x^2<x$) of Theorem (3. 5). Since Theorem (3. 5) is a characterization theorem, it follows that Min has no continuous and decreasing generators. But we can go much further. In fact, ARNOLD and KYRILOV (see [5]) have shown that $\mathrm{Min}(x, y)$ is not representable in the form $g(h(x)+k(y))$ for any combination of continuous functions g, h, k. We can even drop the requirement of continuity for some of the functions involved, and then obtain the following results:

(3. 20) **Theorem.** *The t-norm Min has no continuous additive generator. More precisely, there exists no continuous function $f: I \to R$ such that Min is representable in the form*

(C) $$\mathrm{Min}(x, y) = g(f(x)+f(y)),$$

where g is some left-neutralizer (not necessarily continuous) of f.

PROOF. Since $gf=j_1$, f sends distinct elements into distinct elements. Let $f(0)=Z$ and $f(1)=W$.

Since $0=\mathrm{Min}(0, x)=g(f(0)+f(x))$ for all x in I, g will map the closed interval $[2Z, Z+W]$ (in case $Z<W$) or $[Z+W, 2Z]$ (in case $W<Z$) into the single number 0.

In case $Z<W$, then, for sufficiently small x in I, we have $2Z \leqq f(x)+f(x) < Z+W$. Thus, $g(f(x)+f(x))=0$. On the other hand, $g(f(x)+f(x)) = \mathrm{Min}(x, x) = x$. This is a contradiction.

In case $W<Z$, then, for all sufficiently small x, we have $Z+W < f(x)+f(x) \leqq 2Z$. Thus, $g(f(x)+f(x))=0$. On the other hand, $g(f(x)+f(x)) = \mathrm{Min}(x, x) = x$. This is a contradiction.

Hence, Min has no continuous generator.

(3. 21) **Theorem.** *The t-norm Min has no decreasing generator. More precisely, there exists no decreasing function $f: I \to R$ such that* Min *is representable in the form*

$$\text{(D)} \qquad \mathrm{Min}(x, y) = g(f(x)+f(y)),$$

where g is a nonincreasing left-neutralizer of f.

PROOF. Suppose that Min has a decreasing generator solving the equation (D).

Since f is decreasing, it has at most countably many points of discontinuity.

Let c in I° be a point of continuity of f. Choose an increasing sequence $\{c_n\}$ in I converging to c from below.

Although g is required to be only nonincreasing on Dom (g), g must be decreasing on Ran (f). Hence we have the decreasing sequence $\{f(c_n)\}$, which converges from above to $f(c)$ by the continuity of f at the point c.

Since $c<1$, we have $f(c)>f(1)\geqq 0$, whence

$$f(c)<f(c)+f(c).$$

Therefore for sufficiently large n, we have

$$f(c)<f(c_n)<f(c)+f(c).$$

Applying the nonincreasing left-neutralizer g to the above inequalities, we obtain

$$c>c_n \geqq g(f(c)+f(c)) = \mathrm{Min}(c, c) = c.$$

This contradiction shows the nonexistence of a decreasing generator f solving the representation equation (D).

4. First proof of the main theorem

To prove the main theorem, we first prove a sequence of lemmas which are consequences of postulates (1)—(4) together with the assumption of associativity.

(4. 1) **Lemma 1.** *For any x in J, the sequence $\{x^n\}$ is nondecreasing.*

PROOF. $x^{n+1} = x \cdot x^n \geqq a \cdot x^n = x^n$ by postulates (2) and (3).

(4. 2) **Lemma 2.** *For all x and y in J°, there exists an n such that $x^n > y$.* (I. e., the semigroup S is Archimedean.)

PROOF. Assume the lemma to be false. Then there exist two elements x and y in J° such that $x^n \leqq y$ for all n. Since the sequence $\{x^n\}$ is nondecreasing, it has a limit L. Furthermore, $x < L \leqq y$, consequently

$$L^2 = L \cdot L = (\lim x^n) \cdot (\lim x^n) = \lim x^{2n} = L.$$

But $L^2 = L$ contradicts postulate (4).

(4. 3). *Remark*. We could equivalently have postulated Lemma 2 in place of postulate (4). Because if the ordering is Archimedean, and there exists an x in J° such that $x^2 \leqq x$, then by mathematical induction $x^n \leqq x$ for all n. This contradicts Lemma 2 since we may take $y = x$ in that lemma.

(4. 4) **Lemma 3.** *For all x in J°, the limit of the sequence $\{x^n\}$ is the endpoint b.*

PROOF. If $\lim x^n = y \neq b$, then we have $x^n > y$ for some n by Lemma 2. Since $\{x^n\}$ is nondecreasing, we have a contradiction.

(4. 5) **Lemma 4.** *The endpoint b is an annihilator, i. e., $b \cdot x = x \cdot b = b$ for all x in J.*

PROOF. If x is in J°, then $b = \lim\limits_{n \to \infty} x^n$. Therefore $b \cdot x = (\lim x^n) \cdot x = \lim (x^{n+1}) = b$. Similarly $x \cdot b = b$, for x in J°. Consequently $b^2 \geqq x \cdot b = b$, whence $b^2 = b$.

If $x = a$, then $a \cdot b = b$ by postulate (3), and

$$b \cdot a = b \cdot (\lim_{x \to a} x) = \lim_{x \to a} (b \cdot x) = \lim (b) = b.$$

(4. 6) Definition. An element e of J is called idempotent if $e^2 = e$. Thus we see that the endpoints a and b are idempotents. But by postulate (4) J has no interior idempotents.

(4. 7) Definition. An element x of J is called a nilpotent if $x^n =$ the annihilator for some positive integer n.

(4. 8) **Lemma 5.** *The endpoint a is a unit, i. e., $a \cdot x = x \cdot a = x$ for all x in J.*

PROOF. It suffices to show that a is a right unit, i. e., $x \cdot a = x$ for all x in J.

Consider the continuous map $x \to x \cdot a$. It assumes the values $a (= a \cdot a)$ and $b (= b \cdot a)$. Therefore it assumes all the values from a to b, i. e., every x in J is of the form $y \cdot a$ for some y in J. Then

$$x \cdot a = (y \cdot a) \cdot a = y \cdot (a \cdot a) = y \cdot a = x.$$

(4. 9) **Lemma 6.** *If x and y are in J°, then $x \cdot y > y$ and $y \cdot x > y$.* (I. e., the semigroup S is positively ordered [9].)

PROOF. We have $x \cdot y \geqq a \cdot y = y$. If $x \cdot y = y$, then by mathematical induction, $y = x^n \cdot y$ for all n. But by Lemma 2, there is an n such that $x^n > y$. Hence $y < x^n = x^n \cdot a \leqq x^n \cdot y = y$.

This contradiction proves $x \cdot y > y$. The proof of $y \cdot x > y$ is similar.

(4. 10) **Lemma 7.** *If x and y are in J and $x < y$, then there exist z_1 and z_2 such that $y = x \cdot z_1$ and $y = z_2 \cdot x$.* (I. e., the semigroup S is naturally ordered [9].)

PROOF. Consider the continuous map $S(x, z)$ with a fixed x and variable z. It assumes the values $x \; (= x \cdot a)$, and $b \; (= x \cdot b)$. Therefore it assumes all the values from x to b, in particular y. Hence y is of the form $x \cdot z_1$. The proof of the other part is similar.

(4. 11) *Remark*. It follows immediately from Lemmas 2 and 7 that S is commutative (see [9], [12]). We do not, however, use this fact in the proof of the main theorem.

(4. 12) **Lemma 8.** *If $x < y$ and $x^n < b$, then $x^n < y^n$.*

PROOF. Assume that the lemma is false. Then there exist two elements x and y in J° and an integer n such that $x < y$ and $x^n < b$, but $x^n = y^n$.

By Lemma 7, there exists z in J° such that $y = x \cdot z$. Then we have

$$b > x^n = y^n = (y^{n-1}) \cdot y = (y^{n-1}) \cdot (x \cdot z) = ((y^{n-1}) \cdot x) \cdot z \geqq$$

$$\geqq ((x^{n-1}) \cdot x) \cdot z = x^n \cdot z > x^n.$$

This contradiction proves the lemma.

Thus we see that the continuous and nondecreasing function s_n (where $s_n(x) = x^n$ [cf. the list of conventions]) increases steadily from the value a to the value b, then remains constant at b.

(4. 13) **Lemma 9.** *For all x in J°, $x^n < b$ implies $x^n < x^{n+1}$.*

PROOF. This follows from Lemma 6 by taking y to be x^n. Thus we see that the sequence $\{x^n\}$ increases steadily from the value x until the value b is reached. From that term onward, the sequence has all terms equal to b.

(4. 14) **Lemma 10.** *Let x be in J and $R_n(x) = \{y | y \text{ in } J, y^n = x\}$. Then $R_n(x)$ has a least element.*

PROOF. Consider the continuous function $s_n : J \to J$. It assumes the values a $(= a^n)$ and b $(= b^n)$. Therefore it assumes all values from a to b, in particular the value x. Hence there exists y in J such that $y^n = x$.

Let L be the infimum of $R_n(x)$. We want to show that $L^n = x$. Choose a sequence $\{y_m\}$ in $R_n(x)$ converging to L.

Since $s_n(y_m) = x$, and $\lim_{m \to \infty} s_n(y_m) = s_n(L)$, we must also have $s_n(L) = x$. This proves the lemma.

(4. 15) Definition. We define a function $r_n : J \to J$ by:

$$r_n(x) = \min (R_n(x)).$$

(4. 16) **Lemma 11.** *The function r_n is a right-subinverse of the continuous function s_n, and r_n is continuous and increasing.*

PROOF. Dom $(r_n) = J =$ Ran (s_n). Ran (r_n) is contained in $J =$ Dom (s_n). Finally, $s_n(r_n(x)) = x$ for all x in J. Thus r_n is a right-subinverse of s_n.

By Lemma 8, s_n is continuous and increasing on the interval $[a, r_n(b)] = K$. Therefore s_n restricted to the domain K is an orderpreserving homeomorphism between K and J. By Definition (4. 15) $r_n : J \to K$ is the inverse of this homeomorphism. Hence r_n is also continuous and increasing.

(4. 17) **Lemma 12.** *For all x in J°, $r_n(x) > r_{n+1}(x)$ for all n.*

PROOF. Assume that the lemma is false. Then there exist an x in J° and an n such that

$$a < r_n(x) \leqq r_{n+1}(x) < b.$$

Since s_n is nondecreasing, we then have

$$a < s_n r_n(x) \leqq s_n r_{n+1}(x) < b.$$

Therefore $a < x \leqq s_n r_{n+1}(x) < s_{n+1} r_{n+1}(x) = x$ by Lemma 9, a contradiction. This proves the lemma.

(4. 18) **Lemma 13.** *For all x in J°,*

$$\lim_{n\to\infty} r_n(x) = a.$$

PROOF. The sequence $\{r_n(x)\}$ is decreasing by Lemma 12. Let L be the limit of the sequence. Then $L < r_n(x)$ for all n. Hence $L^n \leqq x$ for all n, by the monotonicity of s_n. Hence $L = a$ by the Archimedean order of S.

(4. 19) **Lemma 14.** *For all i, m, n, we have*

$$s_m r_n = s_{im} r_{in}.$$

PROOF. Assume that the lemma is false. Then there exist i, m, n and an element x in J° such that

$$s_m r_n(x) \neq s_{im} r_{in}(x).$$

Writing $r_n(x) = y$ and $r_{in}(x) = z$, we have $y^n = x$ and $z^{in} = x$. Hence both y and z^i are in $R_n(x)$, and by Definition (4. 15),

$$y \leqq z^i, \quad \text{whence} \quad y^m \leqq z^{im}.$$

But by choice, $y^m \neq z^{im}$. Consequently $y^m < z^{im}$. This shows that $y < z^i$.

Now consider the continuous function s_i. It assumes the value $a\ (= a^i)$ and the value z^i, therefore it assumes all values from a to z^i. In particular it assumes the value y. Therefore there exists an element w in the open interval (a, z) such that $y = w^i$.

Then $w^{in} = y^n = x$. Hence w is in $R_{in}(x)$ and $w < z$. This contradicts the definition of z, whence the lemma.

Proof of the Main Theorem

With the help of the preceding lemmas, we are in a position to prove the main theorem.

Let Q denote the set of all positive rational numbers. First we construct a relation $g^* : Q \to J$ as follows:

We choose an element c in (a, b) and keep c fixed throughout the remainder of the discussion. We define $g^*(m/n) = s_m r_n(c)$ for any m/n in Q. By virtue of Lemma 14, the relation g^* is in fact a function.

Now we proceed to prove the following four propositions:
Proposition (1). *The function g^* is nondecreasing.*

PROOF. Let x and y be two positive rational numbers. Reduce them to a common denominator d, i. e.,

$$x = m/d \text{ and } y = n/d.$$

If $x < y$, then $m < n$. Hence

$$g^*(x) = g^*(m/d) = s_m r_d(c) \leqq s_n r_d(c) = g^*(n/d) = g^*(y).$$

Proposition (2). *For all x, y in* $\mathrm{Dom}(g^*)$, *if $x < y$ and $g^*(x) < b$, then $g^*(x) < < g^*(y)$.*

PROOF. Let x and y be in Q. Reduce them to a common denominator d, i. e.,

$$x = m/d \quad \text{and} \quad y = n/d.$$

If $x < y$, then $m < n$. If $g^*(x) < b$, then by Lemma 9,

$$g^*(x) = s_m r_d(c) < s_{m+1} r_d(c) \leqq s_n r_d(c) = g^*(y).$$

Proposition (3). *The function g^* satisfies the functional equation*

$$g^*(x) \cdot g^*(y) = g^*(x+y) \text{ for all } x \text{ in } \mathrm{Dom}(g^*).$$

PROOF. Let x, y be in Q. Reduce them to a common denominator d, i. e.,

$$x = m/d \text{ and } y = n/d.$$

Then $g^*(x) \cdot g^*(y) = s_m(r_d(c)) \cdot s_n(r_d(c)) = s_{m+n}(r_d(c)) = g^*((m+n)/d) = g^*(x+y)$, by the definition of g^*.
Proposition (4). *The function g^* is continuous on* $\mathrm{Dom}(g^*)$.

PROOF. Since the function g^* is nondecreasing, it has limits on both sides at any x in Q. Therefore, we need only to show, e. g., that

$$\lim_{n \to \infty} g^*(x + 1/n) = g^*(x),$$

and

$$\lim_{n \to \infty} g^*(x - 1/n) = g^*(x),$$

for all x in Q.
First observe that

$$\lim_{n \to \infty} g^*(1/n) = g^*(0) = a,$$

since

$$\lim_{n \to \infty} g^*(1/n) = \lim_{n \to \infty} r_n(c) = a,$$

by Lemma 13. Also, we have

$$g^*(x + 1/n) = g^*(x) \cdot g^*(1/n),$$

by Proposition (3). Therefore

$$\lim_{n \to \infty} g^*(x + 1/n) = g^*(x) \cdot \lim_{n \to \infty} g^*(1/n) = g^*(x) \cdot a = g^*(x).$$

Next for $x>0$ in Q, we have

$$\lim_{n\to\infty} g^*(x-1/n) = \big[\lim_{n\to\infty} g^*(x-1/n)\big]\cdot a =$$

$$= \big[\lim_{n\to\infty} g^*(x-1/n)\big]\cdot\big[\lim_{n\to\infty} g^*(1/n)\big] = \lim_{n\to\infty}[g^*(x-1/n)g^*(1/n)],$$

by virtue of the continuity of the semigroup operation. But $g^*(x-1/n)\cdot g^*(1/n) = g^*(x)$, by Proposition (3). Hence

$$\lim_{n\to\infty} g^*(x-1/n) = \lim_{n\to\infty} g^*(x) = g^*(x).$$

This shows that g^* is continuous on Dom (g^*).

We can now extend the function g^* by continuity to a unique function $g: R\to J$. All the Propositions (1)—(4) remain valid for the extended function g. In particular, according to Proposition (2), g starts from a and increases strictly until the value b is reached. Let B be the first point in R such that $g(B)=b$. Then by Propositions (2) and (4), the restriction of g to $[0, B]$ is a homeomorphism from $[0, B]$ onto J. Let $f: J\to[0, B]$ be the inverse homeomorphism. Then f is increasing and continuous.

By Proposition (3), we have for all x, y in R,

$$g(x)\cdot g(y) = g(x+y).$$

Writing $u=g(x)$ and $v=g(y)$, we see that, since x and y are both in $[0, B]$, $x=f(u)$ and $y=f(v)$. The above equation then reduces to

$$S(u, v) = u\cdot v = g\big(f(u)+f(v)\big).$$

By examining all relevant definitions, and straightforward computation, we see that g is the pseudo-inverse of f. Therefore, the representation of S by a continuous and strictly increasing additive generator f is established. This completes the first proof of the main theorem.

5. Second proof of the main theorem

We will now show the main theorem and its dual can be deduced from known results on topological semigroups [11, 16]. To this end we first restate a theorem of Faucett and one of Mostert and Shields.

(5. 1) **Theorem** of FAUCETT [11]: *Let J be a closed interval $[a, b]$ of the extended real line and $S: J\times J\to J$ an associative function continuous on J. If the left endpoint a is the identity element for the binary operation S and the right endpoint b the annihilator (or zero), and if no interior point of J is nilpotent, then the semigroup J under S is order-isomorphic to the semigroup $[0, \infty]$ under addition* (i. e., *there exists an increasing and bi-continuous function $f: J\to[0, \infty]$ such that $f\big(S(x, y)\big) = f(x)+f(y)$.*)

(5. 2) **Theorem** of MOSTERT and SHIELDS [16]: *Let J and S satisfy all the hypotheses of Faucett's theorem except that J has at least one interior nilpotent element. Then J under S is order-isomorphic to the semigroup $I=[0, 1]$ under the binary operation S_0* (i. e., *there exists an increasing and bi-continuous function $f: J\to I$ such that* $f\big(S(x, y)\big) = S_0\big(f(x), f(y)\big) = \mathrm{Min}\,\big(f(x)+f(y), 1\big)$.

(5. 3) *Deduction of the Main Theorem.* Let $S:J\times J \to J$ satisfy the postulates (1)—(4) of the main theorem. Then we have the following two cases.

Case 1. (This is actually a special case of Aczél's theorem, but we include it here for the sake of completeness.) If $J=[a, b]$ has no interior nilpotent element, then by Lemmas 4 and 5, the semigroup J under S satisfies all the hypotheses of Faucett's theorem. Consequently there exists an increasing and bi-continuous function $f:J\to[0, \infty]$ such that $f(S(x, y)) = f(x)+f(y)$ for all x, y in J. Let $g:[0, \infty]\to J$ be the inverse function of f. Then

$$S(x, y) = g(f(x)+f(y)),$$

for all x and y in J.

Case 2. If $J=[a, b]$ has an interior nilpotent element, then the semigroup J under S satisfies all the hypotheses of the theorem of Mostert and Shields. Consequently there exists an increasing and bicontinuous function $f:J\to I$ such that $f(S(x, y))=S_0(f(x), f(y))=\mathrm{Min}(f(x)+f(y), 1)$ for all x and y in J. Let $g:[0, \infty]\to J$ be the pseudo-inverse of f, i. e.,

$$g(x) = \begin{cases} f^{-1}(x), & \text{if } x \text{ is in } [0, 1], \\ b, & \text{the greatest element of } J, \text{ if } x\geqq 1. \end{cases}$$

Then we have the following:

(i) If $f(x)+f(y) < 1$, then

$$fS(x, y) = \mathrm{Min}\,(f(x)+f(y), 1) = f(x)+f(y)$$

$$S(x, y) = f^{-1}(f(x)+f(y)),$$

and

$$g(f(x)+f(y)) = f^{-1}(f(x)+f(y)).$$

Therefore

$$S(x, y) = g(f(x)+f(y).$$

(ii) If $f(x)+f(y) \geqq 1$, then

$$fS(x, y) = \mathrm{Min}\,(f(x)+f(y), 1) = 1,$$

and

$$f(b)=1,$$

$$g(f(x)+f(y)) = b,$$

Therefore

$$S(x, y) = b = g(f(x)+f(y)).$$

Hence we have $S(x, y) = g(f(x)+f(y))$ for all x and y in J, and this completes the second proof of the main theorem.

Having deduced the main theorem from a theorem of Mostert and Shields, we shall now show that, conversely, their theorem is a consequence of the main theorem. To do this, we recall that if $S_0:I\times I \to I$ is the associative function given by

$$S_0(x, y) = \mathrm{Min}\,(x+y, 1),$$

then S_0 has an additive generator f_0, where f_0 and its pseudo-inverse g_0 are given, by $f_0(x)=x$ for all x in I and $g_0(x)=\mathrm{Min}\,(x, 1)$ for all x in $R=[0, \infty]$. (Cf. (3. 18) (3. 19).)

Let $S: J\times J \to J$ be any semigroup satisfying all the hypotheses of Mostert and Shields's theorem. Then S satisfies all the hypotheses of the main theorem and is therefore Archimedean. Since there exists at least one nilpotent element in the interior of the interval $J=[a, b]$, S must be improperly Archimedean. Hence S has a bounded additive generator.

Let $f: J\to[0, B]$ be an additive generator of S and $g: R\to J$ its pseudo-inverse. Let the function $h: J\to I$ be defined as follows:

$$h(x)=f(x)/B, \text{ where } B=f(b).$$

Hence the inverse $h^{-1}: I\to J$ is given by

$$h^{-1}(x)=f^{-1}(Bx).$$

Then, with the above definitions, we have the following:

(5. 4) **Lemma.** *$g_0(x)=h\big(g(Bx)\big)$ for all x in I.*

PROOF. By the definition of h, we have

$$x=h\big(f^{-1}(Bx)\big) \text{ for all } x \text{ in } I.$$

Hence

$$\begin{aligned} g_0(x) &= h(f^{-1}(Bg_0(x))) = h(f^{-1}(B\operatorname{Min}(x,1))) = \\ &= \begin{cases} hf^{-1}(Bx), & \text{if } x\leqq 1, \\ hf^{-1}(B)=h(b)=1, & \text{if } x\geqq 1, \end{cases} \\ &= \begin{cases} hg(Bx), & \text{if } x\leqq 1, \\ 1 & \text{if } x\geqq 1. \end{cases} \\ &= hg(Bx), \text{ for all } x \text{ in } R=[0,\infty]. \end{aligned}$$

Hence, the lemma is proved.

(5. 5) **Theorem.** *The main theorem implies Mostert and Shields' theorem* (5. 2).

PROOF. Since S and S_0 are Archimedean, we have, for all x, y in I, the following:

$$h\big(S(x,y)\big) = hg\big(f(x)+f(y)\big) = hg\big(B(f(x)/B+f(y)/B\big) = g_0\big(f(x)/B+f(y)/B\big) =$$

(by Lemma (5. 4))

$$= g_0\big(h(x)+h(y)\big) = \operatorname{Min}\big(h(x)+h(y), 1\big) = S_0\big(h(x), h(y)\big).$$

This shows that $h: J\to I$ is an order-preserving isomorphism.

Thus Mostert and Shields' theorem is obtainable as a consequence of the main theorem restricted to the improperly Archimedean semigroups. Similarly Faucett's theorem (a special case of Aczél's theorem) is equivalent to the main theorem restricted to properly Archimedean semigroups. The main theorem has the advantage of treating the two cases in a unified manner.

The history of such isomorphism theorems for ordered semigroups goes back at least as far as HÖLDER [13]. For comprehensive surveys of the field and additional references, see CLIFFORD [8, 9] and FUCHS [12].

We note that another theorem of MOSTERT and SHIELDS characterizes all the continuous semigroups on the unit interval that have 0 as an annihilator and 1 as a unit. Using this we can characterize the most general continuous t-norms.

(5. 6) Definition. Let A be a totally ordered set and $\{S_a\}_{a\in A}$ be a collection of disjoint semigroups indexed by A. Then the ordinal sum of $\{S_a\}$ is the set-theoretic union $\bigcup_{a\in A} S_a$ under the following binary operation:

$$x\cdot y = \begin{cases} x\cdot y, & \text{if } x \text{ and } y \text{ are in one and same } S_a \text{ for some } a\in A. \\ x, & \text{if } x\in S_a \text{ and } y\in S_b \text{ for some } a \text{ and } b \text{ in } A \text{ and } a<b. \\ y, & \text{if } x\in S_a \text{ and } y\in S_b \text{ for some } a \text{ and } b \text{ in } A \text{ and } a>b. \end{cases}$$

It is immediately verified that the ordinal sum is a semigroup under the above defined operation.

The theorem of Mostert and Shields ([16], p. 130, Theorem B) when applied to t-norms takes the following form:

Every continuous t-norm is either the t-norm Min, or is an ordinal sum of Archimedean semigroups and one-point semigroups. This theorem generalizes the theorem of Climescu [10] on the ordinal sum of two semigroups, which was applied to t-norms by Schweizer and Sklar in [20].

6. Archimedean semigroups as limits of Aczélian semigroups

The purpose of this section is to show that every Archimedean semigroup S is obtainable as a limit of Aczélian semigroups. The precise statement is as follows:

(6. 1) **Theorem.** *Let J be a closed interval $[a, b]$, and S an Archimedean semigroup on J. (Cf. Definition (3. 7)). Then there exists a sequence $\{S_n\}$ of semigroups on J, where each S_n is continuous on J and Aczélian on J° (cf. Definition (3. 2)), and such that for every (x, y) in $J\times J$,*

$$\lim_{n\to\infty} S_n(x, y) = S(x, y).$$

Proof. It is clear that we need only consider the case of semigroups satisfying the hypotheses of the main theorem (3. 3), since the dual case follows by dual arguments. In what follows, then, all notation will conform to that of the main theorem.

Furthermore, since the case where S is itself Aczélian is trivial, we consider only non-Aczélian S.

By the main theorem (3. 3), S has a continuous and increasing generator $f: J\to\to [0, B]$, such that S is representable in the form

$$\text{(L)} \qquad S(x, y) = g\big(f(x)+f(y)\big),$$

where g is the pseudo-inverse of f, and $B=f(b)$.

The function $g: R\to J$ is nondecreasing, is the inverse of f on $[0, B]$, and is constant on $[B, \infty]$.

The idea of the proof is to construct a sequence of increasing, continuous function g_n converging pointwise to the continuous function g.

For this purpose, let $\{b_n\}$ be an increasing sequence of numbers with limit b. We can always take $b_1> a$. Let h_n be the restriction of g to the interval $[0, f(b_n)]$, k_n a continuous increasing function from $[f(b_n), \infty]$ to $[b_n, b]$, and g_n the union of

h_n and k_n. Then g_n is a continuous increasing function from R to J which coincides with g on $[0, f(b_n)]$; accordingly, g_n has a continuous, increasing inverse which we denote by f_n.

We have so constructed the g_n that they converge pointwise to the function g. For example, if b is finite, we can take the following as a formula for g_n:

$$g_n(x) = \begin{cases} g(x), & \text{if} \quad 0 \leqq x \leqq f(b_n), \\ b-(b-b_n)\exp\left(\dfrac{f(b_n)-x}{b}\right), & \text{if} \quad f(b_n) \leqq x \leqq \infty. \end{cases}$$

As $b \to \infty$, the last expression becomes $\big(x-f(b_n)\big)+b_n$.

Next, we construct semigroups S_n as follows:

$$S_n(x, y) = g_n\big(f_n(x)+f_n(y)\big),$$

for all x, y in J.

The converse of Aczél's theorem shows that the S_n are Aczélian, i. e., $S_n(x, y)$ is increasing in each of its places for all x, y in J°.

It remains to show that S_n converges pointwise to S. There are two cases to consider:

Case 1. If $S(x, y)$ is an $[a, b)$, i. e., $S(x, y) < b$. Then we have $S(x, y) = g\big(f(x)+f(y)\big) < b$, whence $f(x)+f(y) < f(b)$. But $\lim_{n\to\infty} b_n = b$, and $\lim_{n\to\infty} f(b_n) = f(b)$. Hence, for all sufficiently large n, we have

$$f(x)+f(y) < f(b_n).$$

It follows that, by the definitions of g_n and f_n,

$$g_n(x) = g(x), \quad f_n(x)+f_n(y) = f(x)+f(y).$$

Consequently, we have, for all sufficiently large n,

$$S_n(x, y) = g_n\big(f_n(x)+f_n(y)\big) = g\big(f(x)+f(y)\big) = S(x, y).$$

Case 2. If $S(x, y) = b$, then we have

$$S(x, y) = g\big(f(x)+f(y)\big) = b.$$

Therefore

$$f(x)+f(y) \geqq f(b).$$

Consider the following three sub-cases.

(i) If $f_n(x) = f(x)$ and $f_n(y) = f(y)$,

then

$$f_n(x)+f_n(y) = f(x)+f(y) \geqq f(b) > f(b_n) = f_n(b_n).$$

Hence, by the monotonicity of g_n, we have

$$S_n(x, y) = g_n\big(f_n(x)+f_n(y)\big) > g_n f_n(b_n) = b_n.$$

(ii) If $f_n(x) \neq f(x)$,

then by the definition of g_n and f_n, we have

$$f_n(x) > f_n(b_n).$$

Therefore, by monotonicity of g_n, we have

$$S_n(x, y) = g_n(f_n(x) + f(y)) > g_n f_n(b_n) = b_n.$$

(iii) If $f_n(y) \neq f(y)$, we have $S_n(x, y) < b_n$

by arguments similar to those of case (ii). Hence for all three cases, we have $S_n(x, y) > > b_n$. But $\lim_{n\to\infty} b_n = b$, so we must have

$$\lim_{n\to\infty} S_n(x, y) = b = S(x, y).$$

Consequently, $\lim_{n\to\infty} S_n(x, y) = S(x, y)$ for all (x, y) in $J \times J$.

The following theorem is a special case of the dual version of Theorem (6. 1).

(6. 2) **Theorem.** *Every Archimedean t-norm is the limit of a pointwise convergent sequence of strict t-norms.*

7. Nonexistence of monotone generators for discontinuous semigroups

The existence of continuous and monotone generators for any Aczélian semigroup S is deduced from the following two conditions:

(i) S is continuous,

(ii) S is increasing in each of its places.

In Section 3 we have obtained a proper generalization of Aczél's theorem by relaxing condition (ii). The purpose of this section is to investigate the situation arising from relaxing condition (i) while keeping condition (ii). In other words, we consider the question of the existance of monotone generators for a discontinuous semigroup S which satisfies condition (ii).

The following nonexistence theorem answers this question in the negative.

(7. 1) **Theorem.** *Let J^* be the half-open interval $[a, b)$ and $S: J^* \times J^* \to J^*$ be an associative function satisfying the following conditions:*

(S1) *S is discontinuous,*

(S2) *S is commutative,*

(S3) *S is increasing in each of its places,*

(S4) *The endpoint a is a unit, i. e., $a \cdot x = x \cdot a = x$ for all x in J^*.*

Then there exists no monotone function $f: J^* \to [0, \infty)$ such that S is representable in the form

(S) $$S(x, y) = g(f(x) + f(y)),$$

where g is the inverse of f.

Naturally Theorem (7. 1) can be dualized as follows:

(7. 2) **Theorem.** *Let* $(a, b]$ *be half-open interval and* $S:(a, b]\times(a, b]\rightarrow(a, b]$ *be an associative function satisfying the following conditions:*

(1S) *S is discontinuous,*
(2S) *S is commutative,*
(3S) *S is increasing in each of its places,*
(4S) *The endpoint b is a unit, i. e., $x\cdot b=b\cdot x=x$ for all x in $(a, b]$.*
Then there exists no monotone function $f:(a, b]\rightarrow(-\infty, 0]$ such that S is representable in the form

$$S(x, y) = g\big(f(x)+f(y)\big), \tag{S*}$$

where g is the inverse function of f.

Both Theorem (7. 1) and its dual are consequences of the following theorem on groups:

The Group Case

(7. 3) **Theorem.** *Let J° be an open interval (a, b) and $S:J^\circ\times J^\circ\rightarrow J^\circ$ be an associative function satisfying the following conditions:*

(G1) *S is discontinuous,*
(G2) *S is commutative,*
(G3) *S is increasing in each of its places,*
(G4) *There is a unit e in J°, i. e., $e\cdot x=x\cdot e=x$ for all x in J°,*
(G5) *For each x in J°, there is an inverse y, i. e., y is in J° and $x\cdot y=y\cdot x=e$.*

Then there exists no monotone function $f:J^\circ\rightarrow(-\infty, \infty)$ such that S is representable in the form

$$S(x, y) = g\big(f(x)+f(y)\big), \tag{G}$$

where g is the inverse function of f.

We remark that by applying f to both sides of the equation (G), we have

$$f(xy) = f(x)+f(y),$$

for all x and y in J°. Hence if f exists, it is an isomorphism of the group J° relative to S with a subgroup of the real numbers relative to addition.

In the proof of Theorem (7. 3) we need the following lemma:

(7. 4) **Lemma.** *If a subgroup G of the real numbers (relative to addition) has a least positive element, then G is countable.*

Proof. Let p be the least positive element of G. Suppose that G is not countable. Then there exists a positive element x in G such that $x\neq np$ for all n. But then x lies between two consecutive multiples of p, say

$$(n-1)p<x<np.$$

Hence $0<np-x<p$, and $(np-x)$ is a positive element of G smaller than p. This contradiction proves the lemma.

PROOF *of Theorem* (7. 3). Suppose that there exists a monotone generator f such that S is representable in the form (G). Then f is an isomorphism of the group J° with a subgroup G of the real numbers relative to addition. Since J° is uncountable, and f is a $1-1$ correspondence, G is also uncountable. Therefore, G has no least positive element by the preceding lemma. This means that for all $\eta>0$, there exists an x in G such that $0<x<\eta$.

On the orher hand, f, being already monotone, cannot be continuous. Otherwise g, the inverse of f, would be also both continuous and monotone. Then by the representation (G), S would be continuous, which contradicts (G1).

Now a monotone function f has only jump discontinuities, if any. So our f must have a jump at some point c. Let H be the height of the jump at c. Then we have

$$\lim_{y\to c-} f(x) = L \leqq f(c) \leqq M = \lim_{y\to c+} f(y),$$

and

$$M-L = H.$$

There are two cases to consider:

Case 1. $L=f(c)$ or $M=f(c)$. Then G and the open interval (L, M) are disjoint. On the other hand, G has no least positive element, so G has an element x such that $0<x<H$. Then some integral multiple of x, $\pm nx$ must fall inside (L, M), which is a contradiction.

Case 2. $L<f(c)<M$. Let $\eta = \min\big(f(c)-L, M-f(c)\big)$. Then there exists no element of G inside $\big(L, f(c)\big)$. But again G has an element x such that $0<x<\eta$. So some multiple of x, $\pm nx$ must fall inside $\big(L, f(c)\big)$, which is a contradiction.

Hence, the nonexistence of a monotone generator f for S is proved.

Reduction of the Semigroup Case to the Group Case

We need only show that Theorem (7. 3) (group case) implies Theorem (7. 1).

(7. 5) **Theorem.** *Theorem* (7. 1) *follows from Theorem* (7. 3).

PROOF. Suppose that the semigroup S on J^*, which satisfies all the hypotheses of Theorem (7. 1), has a monotone generator $f: J^*\to[0, \infty)$ such that S is representable in the form

$$\text{(S)} \qquad S(x, y) = g\big(f(x)+f(y)\big),$$

where g is the inverse of f.

Let $q(x)$ denote the point symmetric to x with respect to the point a. We shall extend the function $f: J^*\to[0, \infty)$ to the function $h:\big(q(b), b\big) \to (-\infty, \infty)$ by symmetry. This means that we define h by

$$h(x) = \begin{cases} f(x), & \text{if } x \text{ is in } J^*, \\ -f(q(x)), & \text{if } x \text{ is in } q(J^*). \end{cases}$$

There is no conflict of the definition of h at the point a. This is because $f(a)=0$, which is a consequence of representation (S) and the fact that a is the unit.

Observe that the extension h is also monotone.

If k is the inverse function of h, it can be shown by a simple calculation that

$$k(x)=\begin{cases} g(x) & \text{if} \quad x \text{ is in } f(J^*), \\ qg(-x), & \text{if} \quad x \text{ is in } -f(J^*). \end{cases}$$

Now construct the following 2-place function T:

$$T(x, y) = k\big(h(x)+h(y)\big),$$

for all x and y in $\big(q(b), b\big)$.

It is obvious that T defines a group on the open interval $\big(q(b), b\big)$ and T extends S on J^*.

A straightforward verification shows that T satisfies all the hypotheses of Theorem (7. 3).

By the definition of T, T has a monotone generator h, which is impossible by Theorem (7. 3).

This contradiction proves the nonexistence of monotone generators f for S, and so concludes the proof of Theorem (7. 1). Theorem (7. 2) then follows dually.

The t-norm T_w

The t-norm T_w (see p. 14) serves as a noteworthy counterexample to the nonexistence theorem (7. 2). T_w (when restricted to the subdomain $(0, 1]\times(0, 1]$) satisfies all the hypotheses except one (isotonicity in each place) of Theorem (7. 2). Yet T_w has in fact a decreasing additive generator $f:I\to R$ such that T_w is representable in the form

$$T_w(x, y) = g\big(f(x)+f(y)\big), \tag{W}$$

where g is a nonincreasing left-neutralizer of f.

For example, we can take $f:I\to R$ to be the following function:

$$f(x)=\begin{cases} 2-x, & \text{if} \quad x<1, \\ 0, & \text{if} \quad x=1. \end{cases}$$

And the left-neutralizer $g: R\to I$ to be the following:

$$g(x)=\begin{cases} 1, & \text{if} \quad 0\leqq x\leqq 1, \\ 2-x, & \text{if} \quad 1\leqq x\leqq 2, \\ 0, & \text{if} \quad x\geqq 2. \end{cases}$$

That f and g satisfy the representation equation (W) is a matter of simple calculation.

It is noteworthy that f is decreasing but discontinuous; however, g is nonincreasing and continuous.

Bibliography

[1] N. H. Abel, Untersuchungen der Funktionen zweier unabhängig veränderlichen Größen x and y wie $f(x, y)$, welche die Eigenschaft haben, daß $f(z, f(x, y))$ eine symmetrische Funktion von x, y und z ist, *J. Reine Angew. Math.* **1** (1826), 11–15. (Oeuvres completes de N. H. Abel, v. 1, Christiania (1881), 61–65.)

[2] J. Aczél, Sur les opérations définis pour nombres réels, *Bull. Soc. Math. France* **76** (1949), 59–64.

[3] J. Aczél, Vorlesungen über Funktionalgleichungen und ihre Anwendungen, *Basel, Stuttgart* und *Berlin*, 1961.

[4] V. I. Arnold, On functions of three variables, *Amer. Math. Soc. Transl.* **28** (1963), 51–54.

[5] V. I. Arnold, Concerning the representability of functions of two variables in the form: $X[\Phi(x)+\Psi(y)]$. Uspehi Mat. Nauk **12**, Issue 2 (74), (1957), 119–121.

[6] L. E. J. Rrouwer, Die Theorie der endlichen kontinuierlichen Gruppen unabhängig von den Axiomen von Lie, *Math. Ann.* **67** (1909), 127–136.

[7] É. Cartan, La théorie des groupes finis et continus et l'Analyse Situs, (*Mem. Sci. Math.*, 42) *Paris*, 1930.

[8] A. H. Clifford, Connected ordered topological semigroups with idempotent endpoints I, *Trans. Amer. Math. Soc.* **88** (1958), 88–98.

[9] A. H. Clifford, Totally ordered commutative semigroups, *Bull. Amer. Math. Soc.* **64** (1958), 305–316.

[10] Al. C. Climescu, Sur l'équation fonctionelle de l'associativité. *Bull. École Polytechn. Jassy* **1** (1946), 1–16.

[11] W. M. Faucett, Compact semigroups irreducibly connected between two idempotents, *Proc. Amer. Math. Soc.* **6** (1955), 741–747.

[12] L. Fuchs, Partially Ordered Algebraic Systems, *New York*, 1963.

[13] O. Hölder, Die Axiome der Quantität and die Lehre vom Mass, *Ber. Verh. Sächs. Ges. Wiss. Leipzig. Math.–Phys. Cl.* **53** (1901), 1–64.

[14] A. N. Kolmogorov, On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition, *Amer. Math. Soc. Transl.* **28** (1963), 55–59.

[15] K. Menger, Statistical metrics, *Proc. Nat. Acad. Sci. U.S.A.* **28** (1942), 535–537.

[16] P. S. Mostert–A. L. Shields, On the structure of semigroups on a compact manifold with boundary, *Ann. of Math.* **65** (1957), 117–143.

[17] B. Schweizer–A. Sklar, Statistical metric spaces, *Pacific J. Math.* **10** (1960), 313–334.

[18] B. Schweizer–A. Sklar, The algebra of functions, *Math. Ann.* **139** (1960), 366–382.

[19] B. Schweizer–A. Sklar, Associative functions and statistical triangle inequalities, *Publ. Math. Debrecen* **8** (1961), 169–186.

[20] B. Schweizer–A. Sklar, Associative functions and abstract semigroups, *Publ. Math. Debrecen* **10** (1963), 69–81.

[21] W. Sierpiński, Remarques sur les fonctions de plusieurs variables réelles, *Prace mat.-fiz.* **41** (1934), 171–175.

[22] W. Sierpiński, Algèbre des ensembles, *Monografie Mat.*, 23, *Warszawa–Wroclaw*, 1951.

[23] D. Tamari, Caracterisation des semigroups à un paramétre, *C. R. Acad. Sci. Paris* **228** (1949), 1092–1094.

[24] A. D. Wallace, The structure of topological semigroups, *Bull. Amer. Math. Soc.* **61** (1955), 95–112.

[25] J. Z. Yao–C. H. Ling, On the existence of additive generators for a t-norm, *Notices Amer. Math. Soc.* **11** (1964), 64T–6, p. 127.

(Received September 1, 1964.)