JOURNAL OF APPROXIMATION THEORY 63, 156–169 (1990)

Approximation from the Topological Viewpoint

YAKI STERNFELD

Department of Mathematics and Computer Science,
University of Haifa, Haifa 31999, Israel

Communicated by Allan Pinkus

Received June 2, 1989

Let X be a metric space. A family H of continuous functions of several variables of X with values in X is said to be generating if, whenever $A \subset C(K, X)$ separates points and H operates on A, then A is dense in $C(K, X)$. (For example, the family $H = \{x + y, xy, \text{ constants}\}$ in $C(R^2, R)$ is generating (for R) by the Stone–Weierstrass theorem.) We identify metric spaces which admit generating families (not all do), and among those, we search for spaces X that admit generating families in $C(X^2, X)$—such as R. (This may be considered a topological version of Hilbert's 13th problem.) Once we know this, we try to identify some (small) generating families in $C(X^2, X)$. (This is done in particular when $X = R$.) As a fringe benefit we obtain a "topological" proof of the Stone–Weierstrass theorem.

0. INTRODUCTION

We take the classical Stone–Weierstrass theorem as a starting point for our discussion. It says that if A is a subalgebra of a real $C(K)$ space which separates the points of K and contains the constants then A is dense in $C(K)$. The following reformulation of this theorem is more suitable for our purpose.

STONE–WEIERSTRASS THEOREM. *Let A be a subset of a $C(K)$ space which separates the points of K. If the functions $x + y$, xy and the constant functions operate on A then A is dense in $C(K)$.*

The following definition is included in the above theorem.

DEFINITION. Let X be a topological space, let $A \subset C(K, X)$, and let $h \in C(X^T, X)$. h operates on A if for $\{f_t\}_{t \in T} \subset A$, $h(\{f_t\}_{t \in T}) \in A$.

Thus, the general scheme of the theorem is the following. For some metric space X (namely $X = R =$ the real line) there exists a subfamily H of $\bigcup_{n \geqslant 1} C(X^n, X)$ (namely $H = \{x + y, xy, \text{ constants}\}$) with the following

0021-9045/90 $3.00

property: if H operates on some $A \subset C(K, X)$ which separates points, then A is dense in $C(K, X)$. Moreover, H is actually a subfamily of $C(X^2, X)$.

Remark. A constant function in $C(X, X)$ operates on $A \subset C(K, X)$ if and only if that constant is an element of A. We prefer to consider the constants as operators on A rather than as elements of A.

In this work we consider three problems:

1. For which metric spaces X do there exist subfamilies H of $\bigcup_{n \geq 1} C(X^n, X)$ as above?

2. If, for some metric space X, such families H exist, when can they be chosen in $C(X^2, X)$? (We see later that no such H can be a subset of $C(X, X)$.)

3. Assuming that $C(X^2, X)$ itself is such a family, what special property must a proper subfamily H of $C(X^2, X)$ have in order to satisfy the above; and, in particular, which subfamilies H of $C(R^2, R)$ do the job?

The problems are studied in the following sections, but first, we present some definitions and notation and mention some of the results.

All topological spaces in this work are assumed to be Hausdorff spaces. K always stands for a compact space. If (X, d) is a metric space and Y is a Hausdorff space then $C(Y, X)$ is the space of continuous X-valued functions on Y with the topology of uniform convergence on compact subsets of Y. $A \subset C(Y, X)$ separates points if, for $y_1 \neq y_2$ in Y, $f(y_1) \neq f(y_2)$ for some $f \in A$. $H \subset C(X^T, X)$ operates on A if each $h \in H$ operates on A.

DEFINITION. (i) A subfamily H of $\bigcup_{n \geq 1} C(X^n, X)$ is said to be a generating family if the following holds: for every compact K and every $A \subset C(K, X)$ which separates points, if H operates on A then A is dense in $C(K, X)$.

(ii) A metric space X is in Class 1 if $\bigcup_{n \geq 1} C(X^n, X)$ is a generating family.

(iii) A metric space X is in Class 2 if $C(X^2, X)$ is a generating family.

Now we can state our problems more precisely:

Problem 1. Identify the metric spaces in Class 1.

Problem 2. Among the members of Class 1, find those in Class 2.

Problem 3. For X in Class 2, find the generating families $H \subset C(X^2, X)$.

In Section 1 we show that absolute retracts (AR) and zero-dimensional metric spaces are in Class 1. Note that by this fact the Stone–Weierstrass theorem follows easily from the Weierstrass theorem. Indeed, since by the

Tietze theorem, R is an AR, it is in Class 1. If $A \subset C(K) = C(K, R)$ is a closed algebra which contains the constants, then each polynomial in $C(R^n, R) = C(R^n)$ operates on A and, by the Weierstrass theorem, $C(R^n)$ operates on A. It follows that if A separates points, then $A = C(K)$. (This proof applies the Weierstrass theorem in every finite dimension and thus can hardly be considered "simpler" than the classical proof which uses only the approximation of $|x|$ by polynomials on compact intervals. Still it provides us with a different viewpoint and indicates that the Stone–Weierstrass theorem is built out of two main ingredients: analytic (namely the Weierstrass theorem) and topological (the fact that R is in Class 1).)

The n-dimensional spheres S_n $(n \geqslant 1)$ are simple examples of spaces not in Class 1. (Let $A \subset C(S_n, S_n)$ consist of the null-homotopic elements. A is closed, separates points, and $C((S_n)^k, S_n)$ operates on A for every $k \geqslant 1$.) In Section 1 we also present examples of spaces in Class 1 which are neither ARs nor zero-dimensional.

In Section 2 we make some observations with respect to Problem 2 which can be regarded as a generalized topological version of Hilbert's 13th problem. The only useful information we have on Class 2 is that it contains all Banach spaces and some zero-dimensional spaces. Problem 3, which is considered in Section 3, is of an analytic nature. We show, among other things, that for every non-linear polynomial $p(x)$ in $C(R)$, the family $H = \{f(x, y) = p(x) - y, \text{ constants}\}$ is a generating family (for R), while the family $H_0 = \{-x^2 + y, \text{ constants}\}$ is not a generating family. Thus, $-H_0 = \{x^2 - y, \text{ constants}\}$ is generating, while H_0 is not. We also characterize all $g(x) \in C^1(R)$ so that $H = \{g(x) - y, \text{ constants}\}$ is generating.

1. Problem 1

To study Class 1, it is convenient to introduce another class, Class 1*, which, as we see later, is contained in Class 1 and whose members can be identified more easily.

Definition. A metric space X is in Class 1* if the following holds: For every compact K and every subset A of $C(K, X)$ which separates points, if for every power T, $C(X^T, X)$ operates on A, then $A = C(K, X)$.

Note that in this definition we assume more (namely that $C(X^T, X)$ operates on A for every T and not just for finite T) and require more (that $A = C(K, X)$) than in the definition of Class 1.

Definitions. Let T be a set of indices, and let $\{X_t, t \in T\}$ and X be metric spaces.

(i) Let T' be a subset of T, and let $f \in C(\prod_{t \in T} X_t, X)$. f is said to depend only on the variables of T' if $f(x) = g \circ P(x)$, where $P: \prod_{t \in T} X_t \to \prod_{t \in T'} X_t$ is the canonical projection and $g \in C(\prod_{t \in T'} X_t, X)$.

(ii) For a positive integer n, $C_n(\prod_{t \in T} X_t, X) = \{f \in C(\prod_{t \in T} X_t, X):$ f depends on some subset T' of T with cardinality $n\}$.

(iii) $C_F(\prod_{t \in T} X_t, X) = \bigcup_{n \geqslant 1} C_n(\prod_{t \in T} X_t, X) = \{f \in C(\prod_{t \in T} X_t, X):$ f depends on some finite $T' \subset T\}$.

LEMMA 1. *Let T, X_t, $t \in T$, and X be as above. Then $C_F(\prod_{t \in T} X_t, X)$ is dense in $C(\prod_{t \in T} X_t, X)$.*

Proof. Let d be the metric function on X, let $\varepsilon > 0$, let $f \in C(\prod_{t \in T} X_t, X)$, and let $K \subset \prod_{t \in T} X_t$ be compact. We must show that there is some $h \in C_F(\prod_{t \in T} X_t, X)$ so that $d(f(w), g(w)) < \varepsilon$ for every w in K. Let $k \in K$. Let $B_k = B(f(k), \varepsilon/2) = \{x \in X: d(x, f(k)) < \varepsilon/2\}$. $f^{-1}(B_k)$ is an open subset of $\prod_{t \in T} X_t$ which contains k. By the definition of the product topology there exists an open subset V_k of $\prod_{t \in T} X_t$ of the form $V_k = \prod_{t \in T_k} J_{t,k} \times \prod_{t \in T \setminus T_k} X_t$, where T_k is a finite subset of T and $J_{t,k}$ is an open subset of X_t, $t \in T_k$, so that $k \in V_k \subset f^{-1}(B_k)$. By the compactness of K, there exists a finite subset K' of K so that $\{V_k\}_{k \in K'}$ covers K. Set $T' = \bigcup\{T_k: k \in K'\}$. Then T' is finite. For $t \in T \setminus T'$ let y_t be some fixed (but otherwise arbitrary) point in X_t. Let $g: \prod_{t \in T'} X_t \to X$ be defined by $g(\prod_{t \in T'} x_t) = f(\prod_{t \in T'} x_t \times \prod_{t \in T \setminus T'} y_t)$. Let $h \in C_F(\prod_{t \in T} X_t, X)$ be defined by $h(w) = g(P_w)$, where $P: \prod_{t \in T} X_t \to \prod_{t \in T'} X_t$ is the canonical projection. We claim that h does the job. Indeed, let $w \in K$. Thus $w \in V_k$ for some $k \in K'$. Then both $f(w)$ and $h(w)$ are in B_k. For $f(w)$ this is trivial since $V_k \subset f^{-1}(B_k)$. Let $w' \in \prod_{t \in T} X_t$ be defined by $Pw' = Pw$, and for $t \in T \setminus T'$, $P_t w' = y_t$ (where $P_t: \prod_{t \in T} X_t \to X_t$ is the canonical projection). Then w' is also in V_k (since w' differs from w only in the coordinates of $T \setminus T'$) while V_k depends only on the coordinates of $T_k \subset T'$. Also, $h(w) = g(Pw) = g(Pw') = f(w')$ by the definition of g. Thus $h(w) \in f^{-1}(V_k) \subset B_k$. It follows that $d(f(w), h(w)) \leqslant d(f(w), f(k)) + d(f(k), h(w)) < \varepsilon$. ∎

LEMMA 2. *Let $h \in C(X^T, X)$ and let $A \subset C(K, X)$. If h operates on A then it operates on $\bar{A}$.*

Proof. Let $\{f_t\}_{t \in T} \subset \bar{A}$ and $\varepsilon > 0$ be given. Define $f: K \to X^T$ by $(f(k))_t = f_t(k)$. Set $S = f(K) \subset X^T$. For $k \in K$, set $B_k = B(h(f(k)), \frac{1}{2}\varepsilon) \subset X$. Then $h^{-1}(B_k)$ is open in X^T and contains $s = f(k)$. Hence there exists a basic neighbourhood V_k of $f(k)$ in X^T so that $V_k \subset h^{-1}(B_k)$ is of the form

$$V_k = \prod_{t \in T_k} B(f_t(k), \delta_{t,k})) \times X^{T \setminus T_k}$$

with $T_k \subset T$ finite. Set $U_k = \prod_{t \in T_k} B(f_t(k), \frac{1}{2}\delta_{t,k}) \times X^{T \setminus T_k}$. Then

$$f(k) \in U_k \subset V_k \subset h^{-1}(B_k).$$

By the compactness of S there exists a finite subset K' of K so that $\{U_k\}_{k \in K'}$ covers S.

Set $\delta = \min\{\delta_{t,k} : k \in K', t \in T_k\}$. δ is positive. For each $t \in T$, let $g_t \in A$ so that for every $k \in K$, $d(g_t(k), f_t(k)) < \frac{1}{2}\delta$. Such g_t's exist since $f_t \in \bar{A}$, $t \in T$. Define $g: K \to X^T$ by $(g(k))_t = g_t(k)$. Let $w \in K$. Then $f(w) \in S$, so $f(w) \in U_k$ for some $k \in K'$. It follows then that $g(w)$ is in $V_k \subset h^{-1}(B_k)$. So, both $h(f(w))$ and $h(g(w))$ are in B_k, and it follows that $d(h(f(w)), h(g(w))) < \varepsilon$ for all $w \in K$. As h operates on A, $h \circ g \in A$, and thus, since ε is arbitrary, $h \circ f \in \bar{A}$ and we are done. ▌

LEMMA 3. *Let $D \subset C(X^T, X)$ and $A \subset C(K, X)$. If D operates on A then $\bar{D}$ operates on $\bar{A}$.*

Proof. Let $h \in \bar{D}$ and $\{f_t\}_{t \in T} \subset \bar{A}$ be given. Let $f: K \to X^T$ be defined by $(f(k))_t = f_t(k)$. Set $S = f(K) \subset X^T$. Let $\{h_n\}_{n \geqslant 1} \subset D$ be a sequence which converges to h uniformly on compact subsets of X^T, and in particular on S. Then by Lemma 2, $h_n \circ f \in \bar{A}$ and $h_n \circ f$ converges to $h \circ f$ uniformly on K. Hence $h \circ f$ is in $\bar{A}$ and the lemma follows. ▌

THEOREM 1. *Class 1 contains Class 1*.*

Proof. Let X be in Class 1*. Let $A \subset C(K, X)$ separate points, and assume that for all n, $C(X^n, X)$ operates on A. This clearly implies that for every power T, $C_F(X^T, X)$ operates on A. (The two statements are actually equivalent.) It follows from Lemma 3 that the closure of $C_F(X^T, X)$ in $C(X^T, X)$ operates on $\bar{A}$. So, by Lemma 1, $C(X^T, X)$ operates on $\bar{A}$, and since X is in Class 1*, $\bar{A} = C(K, X)$, i.e., X is in Class 1. ▌

The following theorem characterizes the elements of Class 1*.

THEOREM 2. *Let X be a metric space. Then X is in Class 1* if and only if the following holds:*

(*) *For every power T and for every compact subset S of X^T, every function $f \in C(S, X)$ is continuously extendable over X^T.*

COROLLARIES. (i) *Every absolute retract is in Class 1*.*

(ii) *Every zero-dimensional metric space is in Class 1*.*

(iii) *If X is in Class 1* and X contains a copy of [0, 1] (or, equivalently, if X contains any other absolute retract) then for every normal space Y and every compact subset S of Y, every f in $C(S, X)$ is extendable over Y.*

In particular if X is a compact element of Class 1* *which contains a copy of* [0, 1] *then X is an absolute retract.*

Proof of Corollaries. (i) is trivial and (ii) follows from [Kur, p. 333 (vi)]. For (iii), let X be in Class 1*. If S is a compact subset of a normal space Y, set $A=\{f\in C(S,X), f \text{ is extendable over } Y\}$. It is easy to check (see Proof of Theorem 2) that for every T, $C(X^T, X)$ operates on A. Also, if X contains [0, 1] then the [0, 1]-valued elements of $C(S, X)$ are in A and separate the points of S. It follows that $A=C(S,X)$. ∎

Proof of Theorem 2. Let X be in Class 1*. Let $K\subset X^T$ be compact. Set

$$A=\{f\in C(K,X): f \text{ is extendable over } X^T\}.$$

Then for every $t\in T$ the restriction $q_t=P_t/K$, where $P_t: X^T\to X_t$ is the canonical projection, is obviously in A, and since $\{q_t\}_{t\in T}$ separate the points of K, so does A. Also, for every set D of indices, $C(X^D, X)$ operates on A. Indeed, if $\{f_\delta\}_{\delta\in D}\subset A$, let $\{F_\delta\}_{\delta\in D}$ be their extensions to X^T. Then for every h in $C(X^D, X)$, $h\circ\{F_\delta\}_{\delta\in D}$ is an extension of $h\circ\{f_\delta\}_{\delta\in D}$, so $h\circ\{f_\delta\}_{\delta\in D}$ is in A. As X is in Class 1*, $A=C(X^T, X)$, so X satisfies (∗).

Conversely, let X satisfy (∗). Let $A\subset C(K,X)$ be a subset that separates the points of K, and assume that for every T, $C(X^T, X)$ operates on A. Define $\psi: K\to X^A$ by $(\psi(k))_f=f(k)$, $f\in A$. As A separates points, ψ is one-to-one and, by compactness of K, ψ is a homeomorphism of K onto $S=\psi(K)\subset X^A$. Let $g\in C(K,X)$. Then $g\circ\psi^{-1}\in C(S,X)$. Since X satisfies (∗), $g\circ\psi^{-1}$ is extendable to a function $h\in C(X^A, X)$. As $C(X^A, X)$ operates on A, $h\circ\{f\}_{f\in A}=h\circ\psi\in A$. But for $k\in K$, $\psi(k)\in S$, so $h\circ\psi(k)=g\circ\psi^{-1}\circ\psi(k)=g(k)$. Thus $g\in A$, i.e., $A=C(K,X)$, and we are done. ∎

The following is an example of metric spaces in Class 1 but not in Class 1*.

EXAMPLE. Let $X_0=\{(t, \sin 1/t): 0<t\leqslant 1\}$ and let $X_0\subsetneq X\subset \bar{X}_0\subset R^2$ be any set which lies between X_0 and its closure in R^2. We claim that X is in Class 1 but not in Class 1*. The fact that X is not in Class 1* follows from Corollary (iii), since X contains a copy of [0, 1] in X_0, and if $a\in X_0$ and $b\in X\setminus X_0$, then the function $f: \{0,1\}\to X$ so that $f(0)=a$ and $f(1)=b$ is not extendable over [0, 1].

The fact that X is in Class 1 follows from the next more general lemma.

LEMMA 4. *Let* (X, d) *be a metric space. Let* $X_0\subset X$ *be an AR so that for every* $\varepsilon>0$ *there exists a continuous mapping* $r_\varepsilon: X\to X_0$ *such that* $d(x, r_\varepsilon x)<\varepsilon$ *for every* x *in* X. *Then X is in Class* 1.

Remarks. (i) The condition clearly implies that X_0 is dense in X.

(ii) Note that r_ε need not be a retraction.

(iii) In our example X_0 is homeomorphic to $(0, 1]$ and hence is an AR. Also, given $\varepsilon > 0$, r_ε can be constructed as follows: let n be so large that $\frac{1}{2}\pi(4n-3) > 1/\varepsilon$, and let r_ε be the retraction of $\bar{X}$ onto $X_n = \{(t, \sin 1/t)$: $2/\pi(4n-1) \leqslant t \leqslant 1\}$ obtained by moving a point $x \in \bar{X} \setminus X_n$ parallel to the x-axis until it meets X_n. One checks easily that $d(x, r_\varepsilon x) < \varepsilon$ for all x in $\bar{X}$.

Proof of Lemma 4. Let $A \subset C(K, X)$ separate points and assume that for all $n \geqslant 1$, $C(X^n, X)$ operates on A. We must show that A is dense in $C(K, X)$. Identify $C(K, X_0)$ with $\{f \in C(K, X), f(K) \subset X_0\}$ and set $A_0 = A \cap C(K, X_0)$. $C(K, X_0)$ is dense in $C(K, X)$ since, for f in $C(K, X)$, $r_\varepsilon \circ f \in C(K, X_0)$ and $d(f(k), r_\varepsilon f(k)) < \varepsilon$. We complete the proof by showing that A_0 is dense in $C(K, X_0)$. A_0 separates points since, if $f \in A$ distinguishes between two points in K, then so does $r_\varepsilon \circ f$ for sufficiently small $\varepsilon > 0$, and, as $r_\varepsilon \in C(X, X)$ which operates on A, $r_\varepsilon \circ f \in A \cap C(K, X_0) = A_0$. We claim that $C(X_0^n, X_0)$ operates on A_0. Indeed, let $h \in C(X_0^n, X_0)$ and let $\{f_i\}_{i=1}^n \subset A_0$. Set $Y = \bigcup_{i=1}^n f_i(K) \subset X_0$. Let $H \in C(X^n, X_0) \subset C(X^n, X)$ be an extension of h/Y^n: $Y^n \to X_0$. (H exists since X_0 is an AR and Y is compact.) Then $g = H(f_1, f_2, ..., f_n) \in A$ since $C(X^n, X)$ operates on A, and $g \in A_0$ as $g(K) \subset H(X^n) \subset X_0$. Also, for $k \in K$, $(f_1(k), ..., f_n(k)) \in Y^n$ and on Y^n, H agrees with h, so $g = h(f_1, ..., f_n) \in A_0$. Recall that by Theorem 1 and Corollary (i), X_0 is in Class 1. Hence A_0 is dense in $C(K, X_0)$ and we are done. ∎

2. Problem 2

Let X be a metric space and let n be a positive integer. In $C(X^n, X)$ consider the *composition* operation which assigns to every $(n+1)$-tuple h, $f_1, f_2, ..., f_n$ of elements of $C(X^n, X)$ another element $h(f_1, ..., f_n)$ in $C(X^n, X)$.

Definition. Let H be a subset of $C(X^n, X)$. Then comp H denotes the smallest subset of $C(X^n, X)$ which contains H and is invariant under the composition operation. $\overline{\text{comp}}\, H$ is the closure of comp H. Clearly comp $H = \bigcup_{k \geqslant 0} H_k$, where $H_0 = H$, and H_{k+1} consists of compositions of elements of H_k. Hilbert's 13th problem contains the conjecture that comp $C_2(R^3, R) \neq C(R^3, R)$. (See Definition (ii).) This has been refuted by Arnold [A] and Kolmogorov [Kol]. Kolmogorov's result is surprisingly strong. It shows in particular that for each $n \geqslant 2$,

$$\text{comp}(C_1(R^n, R) \cup \{x + y\}) = C(R^n, R)$$

(where $x+y$ is the addition function in $C_2(R^n, R)$.) See [St1] for a survey of related results.

DEFINITION. Class 3 consists of the metric spaces X such that $\overline{\text{comp}}\, C_2(X^n, X) = C(X^n, X)$ for every $n \geqslant 3$.

From Kolmogorov's theorem it follows that R is in Class 3. However, as the polynomials in $C(R^n)$ are in comp $C_2(R^n, R)$, the Weierstrass theorem implies this as well. Malcev [M] proved that for every pair (n, k) of non-negative integers, $X = I^n \times (S_1)^k$ satisfies the following Kolmogorov type result (I is the interval $[0, 1]$ and S_1 the 1-sphere): There exists some h in $C(X^2, X)$ so that $\text{comp}(C_1(X^m, X) \cup \{h\}) = C(X^m, X)$, $m \geqslant 2$. Thus, $I^n \times (S_1)^k$ are in Class 3 for all n and k. From the results of [St–W] and [St2] it follows that every Banach space is in Class 3.

Gadzhiev proved in [G1] that comp $C_2((S_3)^3, S_3) = C((S_3)^3, S_3)$, where S_3 is the 3-sphere; however, in [G2] he showed that comp $C_3((S_3)^4, S_3)$ differs from $C((S_3)^4, S_3)$, and it seems that his argument shows that S_3 is not in Class 3.

Our interest in Class 3 results from the following theorem.

THEOREM 3. *Class* 2 *consists of the intersection of Classes* 1 *and* 3.

Proof. Let X be in Class 2. Then X is in Class 1 and $C(X^2, X)$ is a generating family. Set $A = \text{comp}\, C_2(X^n, X)$. Then $A \subset C(X^n, X)$ separates the points of X^n and obviously $C(X^2, X)$ operates on A. As $C(X^2, X)$ is generating, A must be dense, and thus X is in Class 3. Note that by definition the operation of a generating family implies density when $A \subset C(K, X)$ with K compact. We applied this here for $A \subset C(X^n, X)$, where X^n may fail to be compact. But this is still valid since the topology in $C(X^n, X)$ is that of uniform convergence on compacts. Conversely, assume that X is in both Classes 1 and 3. Let $A \subset C(K, X)$ separate points, and let $C(X^2, X)$ operate on A. Then, for each n, comp $C_2(X^n, X)$ operates on A. By Lemma 3, $\overline{\text{comp}}\, C_2(X^n, X)$ operates on $\bar{A}$. Thus $\bar{A}$ is dense in $C(K, X)$ and X is in Class 2. ∎

It seems as though Class 3 contains Class 1. By Theorem 3 this would imply that Classes 1 and 2 agree. We leave this as an open question.

Question 1. Does Class 3 contain Class 1?

Question 2. Does Class 3 contain Class 1*?

Question 3. Does Class 3 contain all AR's and all zero-dimensional spaces?

For zero-dimensional spaces we have the following partial result.

THEOREM 4. *Let X be a finite dimensional compact metric space. If X^2 is homeomorphic to a subset of X, then for all $n \geqslant 2$, comp $C_2(X^n, X) = C(X^n, X)$, and in particular X is in Class 3, and hence in Class 2.*

Proof. Let h be a homeomorphism of X^2 into X. Define $h_n: X^n \to X$ by $h_2 = h$, $h_{n+1}(x_1, x_2, ..., x_{n+1}) = h(h_n(x_1, ..., x_n), x_{n+1})$. h_n is then in comp $C_2(X^n, X)$ and is a homeomorphism. Note that if $\dim X \geqslant 1$, then $\dim X^2 \geqslant \dim X + 1$. (See [E, p. 98, 1.9.E(b)]). Hence $\dim X = 0$. Thus, by [Kur, p. 333, (vi)], $h_n^{-1}: h_n(X^n) \to X^n$ is extendable to a mapping $H_n: X \to X^n$. Let $f \in C(X^n, X)$. Define $\tau \in C(X, X)$ by $\tau(x) = f(H_n(x))$. Then for $w \in X^n$

$$\tau(h_n(w)) = f \circ H_n \circ h_n(w) = f(w).$$

As $\tau \in C_1(X^n, X)$ it follows that $f \in \text{comp}\, C_2(X^n, X)$. ∎

3. PROBLEM 3

In this section we consider generating families $H \subset C(X^2, X)$, where X is in Class 2. Since, aside from some zero-dimensional spaces, the only members of Class 2 we know of are Banach spaces, we concentrate on those, and in particular on the most interesting case, $X = R$. Still in the general frame we have the following characterization of generating families.

THEOREM 5. *Let X be in Class 2. A subset H of $C(X^2, X)$ is generating if and only if $\overline{\text{comp}}\{H \cup \{x\} \cup \{y\}\} = C(X^2, X)$, where x is the projection $h(x, y) = x$ and y is the projection $g(x, y) = y$.*

Proof. Let H operate on a subset A of $C(K, X)$ which separate points. Then clearly comp$\{H \cup \{x\} \cup \{y\}\}$ also operates on A, and by Lemma 3, $\overline{\text{comp}}\{H \cup \{x\} \cup \{y\}\}$ operates on $\bar{A}$. Thus, if $\overline{\text{comp}}\{H \cup \{x\} \cup \{y\}\} = C(X^2, X)$, then $\bar{A}$ is dense in $C(K, X)$ since X is in Class 2. *Conversely*, $A = \text{comp}\{H \cup \{x\} \cup \{y\}\} \subset C(X^2, X)$ separates points, and H operates on A. Thus, if H is generating, then A must be dense, and we are done. ∎

DEFINITIONS. Let $H \subset C(X^2, X)$.

1. $E(H) = \text{comp}\{H \cup \{x\} \cup \{y\}\}$.

2. Let $\tau: X \to Y$ be a homeomorphism of X onto Y. For $h \in C(X^2, X)$, let $h^* = \tau h \tau^{-1} \in C(Y^2, Y)$ be defined by

$$h^*(y_1, y_2) = \tau h \tau^{-1}(y_1, y_2) = \tau(h(\tau^{-1} y_1, \tau^{-1} y_2))$$

and for $H \subset C(X^2, X)$,

$$H^* = \tau H \tau^{-1} = \{\tau h \tau^{-1} : h \in H\} \subset C(Y^2, Y).$$

PROPOSITION 1. *Let X be in Class 2 and let $\tau: X \to Y$ be a homeomorphism. Then $H \subset C(X^2, X)$ is generating if and only if $H^* \subset C(Y^2, Y)$ is generating.*

Proof. One checks easily that $h^* \circ (f^*, g^*) = (h \circ (f, g))^*$. Hence $(E(H))^* = E(H^*)$. Thus, H is generating iff $E(H)$ is dense in $C(X^2, X)$, iff $(E(H))^*$ is dense in $C(Y^2, Y)$, and iff H^* is generating. ∎

From now on we restrict the discussion to real Banach spaces. Let X be a Banach space. Let $x+y$, $x-y$, tx denote the functions $f(x, y) = x+y$, $g(x, y) = x-y$, and $h(x) = tx$ $(t \in R)$, respectively. Let $C^1(X^n, X)$ denote the continuously differentiable elements of $C(X^n, X)$. Let also

$$S.W.(X) = \{h \in C(X, X) : \{x-y, tx, t \in R, \text{constants}, h\}$$
$$\text{is a generating family}\}$$

and

$$D(X) = \{h \in C(X, X) : \{x-y, \text{constants}, h\} \text{ is generating}\}.$$

Note that $D(X) \subset S.W.(X)$ and that the Stone–Weierstrasse theorem says that $h(x) = x^2$ is in $S.W.(R)$.

The families $S.W.(X)$ and $D(X)$ were studied in [St–W] and [St2], respectively. In particular it is proved in [St–W] that for every Banach space X, $S.W.(X)$ is dense in X; that $S.W.(R)$ consists of the non-affine functions of $C(R)$; and the elements of $S.W.(R^n)$ are identified explicitly. Similar results are obtained in [St2] for $D(X) \cap C^1(X, X)$.

In particular it follows from those results that $f \in D(R) \cap C^1(R)$ if and only if f is non-affine, and no proper closed subgroup G of R is a generalized period of f. (G is a generalized period of f if for $x, y \in R$, $x-y \in G$ implies that $f(x) - f(y) \in G$. If $G = aZ$, $a > 0$, this is equivalent to the identity $f(x+a) = f(x) + f(a) - f(0)$ for all $x \in R$. See [St2].) Thus every non-linear polynomial and every exponential $e^{p(x)}$, p a polynomial, are in $D(R)$.

THEOREM 6. *Let X be a Banach space. Let $f \in C(X, X)$, and let $h(x, y) = f(x) - y \in C(X^2, X)$. Then $H = \{h(x, y), \text{constants}\}$ is generating if and only if $f \in D(X)$.*

Proof. Clearly,

$$\text{comp}\{f(x)-y, \text{constants}, x, y\}$$
$$\subset \text{comp}\{x-y, f(x), \text{constants}, x, y\}.$$

If H is generating then, by Theorem 5, comp$\{f(x)-y$, constants, $x, y\}$ is dense in $C(X^2, X)$, and thus comp$\{x-y, f(x)$, constants, $x, y\}$ is also dense there. By Theorem 5 again, $f \in D(X)$. Let $f \in D(X)$. We must show that $E = \text{comp}\{f(x)-y$, constants, $x, y\}$ is dense in $C(X^2, X)$. We use the terminology and notation of [St2].

(1) *If $f \in D(X)$ then the closed additive subgroup G of X generated by the range of f is the whole of X.*

Proof. G is a generalized period of f; indeed, for x, y in X, $f(x)-f(y) \in G$, and in particular for x, y in G. Thus, if $G = \{0\}$ then f is constant, and since it is not a constant, G must be X.

(2) *Let $n_1, n_2, ..., n_k$ be integers. If $\sum_{i=1}^k n_i$ is even then for every $t_1, t_2, ..., t_k$ in X, $\sum_{i=1}^k n_i f(t_i) + y \in E$, and if $\sum_{i=1}^k n_i$ is odd then for every $t_1, ..., t_k$ in X, $\sum_{i=1}^k n_i f(t_i) - y \in E$.*

Proof. By induction on $\sum_{i=1}^k |n_i| = m$. If $m = 1$, then $n_1 = k = 1$ and, as $f(x) - y \in E$, we apply the substitution $x = t$ (recall that E contains the constants) and obtain $f(t) - y \in E$. Assume (2) for $\sum_{i=1}^k |n_i| < m$, and let $\sum_{i=1}^k |n_i| = m$. Let m be even (the other case is similar). Then $m-1$ is odd. By subtracting 1 from one of the n_i's, n_1, say, we obtain that $-((n_1-1)f(t_1) + \sum_{i=2}^k n_i f(t_i)) - y \in E$. As $f(x) - y \in E$ we may substitute the above for y in $f(x) - y$ and conclude that $f(x) - (-((n_1-1)f(t_1) + \sum_{i=2}^k n_i f(t_i)) - y) \in E$. Set $x = t_1$ and obtain $\sum_{i=1}^k n_i f(t_i) + y \in E$.

(3) *For every α in G, $\alpha \pm y \in \bar{E}$, and by (1), $t - y \in \bar{E}$ for all $t \in X$, and in particular, $-y \in \bar{E}$.*

Proof. $\{\sum_{i=1}^k n_i f(t_i), n_i \in Z, t_i \in X\}$ is dense in G.

Set $B = \{v \in C(X, X) \colon v(x) - y \in \bar{E}\}$.

(4) *$B \subset \bar{E}$ since $v \in B$ implies $v(x) - y \in \bar{E}$ and set $y = 0$.*

(5) *B contains the constants (by (3)) and in particular $0 \in B$.*

(6) *B is a group.*

Proof. Let $u, v \in B$. Then $v(x) - y$, $u(x) - y$ are in $\bar{E}$. Hence (set

$y := v(x) - y$, $u(x) - (v(x) - y) = u(x) - v(x) + y \in \bar{E}$. As $-y \in \bar{E}$, $u(x) - v(x) - y \in \bar{E}$, i.e., $u(x) - v(x) \in B$.

(7) *f operates on B.*

Proof. Let $v \in B$. Then $v \in \bar{E}$. Hence the composition $h \circ (v, y) = f \circ v(x) - y \in \bar{E}$, so $f \circ v \in B$.

(8) *$f \in B$ (since $f(x) - y \in E$).*

(9) *Let $v \in B$ and $a \in X$; then $v(-x)$ and $v(x+a)$ are in B.*

Proof. $g(x, y) = v(x) - y \in \bar{E}$, and by (3), $-x$ and y are in $\bar{E}$. Hence $g(-x, y) = v(-x) - y \in \bar{E}$ so $v(-x) \in B$. By (3), $a \pm y \in \bar{E}$ so by substitution, $y := x$, $a \pm x \in \bar{E}$. In $g(x, y) = v(x) - y$ replace x by $x + a$ and obtain $v(x+a) - y \in \bar{E}$. So $v(x+a) \in B$.

Set $L = \bigcap_{v \in B} v^{-1}(0) = \{a \in X : v(a) = v(0) \text{ for all } v \in B\}$.

(10) *L is a closed subgroup of X.*

Proof. Clearly $0 \in L$ and, by (9), $a \in L$ implies that $-a \in L$. Let $a, b \in L$. We show that $v(0) = v(a+b)$, i.e., $a + b \in L$. Let $v \in B$. By (9), $u(x) = v(x+b) \in B$. Hence, as $b \in L$, $u(0) = v(b) = v(0)$, and since $a \in L$, $u(0) = u(a) = v(a+b)$. So $v(0) = v(a+b)$.

(11) *Let $v \in B$, $a \in L$, and $e \in X$. Then $v(e+a) = v(e) + v(a) - v(0)$.*

Proof. Set $u(x) = v(x+e) - v(x)$. $u \in B$ by (6) and (9). Thus, as $a \in L$, $u(0) = u(a)$, i.e., $v(e) - v(0) = v(a+e) - v(a)$.

(12) *Let $g \in \bar{E}$ be a function of x only so that $g(0) = 0$. Then $g(L) \subset L$.*

Proof. Let $w \in L$. Set $z = g(w)$. We must show that $z \in L$, i.e., $v(z) = v(0)$ for all $v \in B$. Let $v \in B$. Then $u(x) = v(g(x)) \in B$. So $u(w) = u(0)$, i.e., $v(g(w)) = v(g(0))$, and as $g(0) = 0$, $v(z) = v(0)$.

(13) *L is a generalized period for every element of B.*

Proof. We must show that for $x, y \in X$, $x - y \in L$ implies that $g(x) - g(y) \in L$ for $g \in B$. Note first that we may assume $g(0) = 0$, since if $g \in B$ then $\tilde{g}(x) = g(x) - g(0) \in B$ and $\tilde{g}(x) - \tilde{g}(y) = g(x) - g(y)$. So, let $x - y = w \in L$ and $g \in B$ with $g(0) = 0$ be given. By (11),

$$g(x) = g(y+w) = g(y) + g(w) - g(0) = g(y) + g(w).$$

Hence $g(x) - g(y) = g(w) \in g(L) \subset L$ by (12).

(14) *B separates the points of X.*

Proof. $f \in B$ and L is a generalized period of f. As $f \in D(X)$, L must be trivial, i.e., $L = \{0\}$. Let $a, b \in X$ and assume that $v(a) = v(b)$ for all $v \in B$. As $v(x-a) \in B$ whenever $v \in B$, $v(0) = v(b-a)$ for all $v \in B$. Hence $b - a \in L$. So $b - a = 0$, i.e., $a = b$.

Proof of the Theorem (Concluded). B is a closed subgroup of $C(X, X)$ which separates points and contains the constants, and f operates on B. As $f \in D(X)$, $B = C(X, X)$. Hence the identity function x is in B, and the function $x - y$ is in $\bar{E}$. Hence $\{x - y, f,$ constants$\}$ operates on $\bar{E}$ and as $f \in D(X)$, $\bar{E} = C(X^2, X)$. ▮

It follows from Theorem 6 that for every non-linear polynomial $p(x)$, $H_p = \{p(x) - y,$ constants$\}$ is generating in $C(R^2)$. It turns out that $p(x) + y$ is very different.

PROPOSITION 2. *Let $p(x)$ be a polynomial whose leading coefficient is an integer. Then $\{x + y, p(x),$ constants$\}$ is not generating.*

COROLLARY. *$\{p(x) + y,$ constants$\}$ is not generating either.*

Proof. Let p be of degree $k \geqslant 2$ and let ε be the sign of the leading coefficient of p (which is assumed to be an integer). We distinguish among the following three cases.

(1) $\varepsilon = 1$, (2) $\varepsilon = -1$ and k is even, (3) $\varepsilon = -1$ ad k is odd.

In case (1), let $A \subset C(R)$ consist of the constants and all the polynomials of positive degree with leading coefficient in the set N of positive integers. From the classical theory of Chebyshev (see, e.g., [L, p. 32]) it follows that A is a closed subset of $C(R)$. (The restriction of A to every interval $J \subset R$ of length $\geqslant 4$ is closed in $C(J)$.) It is clear that A separates points and that $H = \{x + y, p(x),$ constants$\}$ operates on A.

As $A \neq C(R)$, H is not generating. (Note that the same proof shows that $A \cup \{x + y\}$ is not generating.) In case (2), note that $H = \{p(x), x + y,$ constants$\}$ is generating if and only if $H^* = \{-p(-x), x + y,$ constants$\}$ is generating. (Apply Proposition 1 with $\tau(x) = -x$, $X = Y = R$). As case (1) applies to $-p(-x)$, H is not generating.

In case (3) we need a slightly more complicated test set A. Let $A = \bigcup_{l \geqslant -1} A_l$, where $A_{-1} =$ the constants.

$A_l = \{$polynomials of degree k^l with leading coefficients in $(-1)^l N\}$. As in case (1), A is closed in $C(R)$ and $H = \{p(x), x + y,$ constants$\}$ operates on A. ▮

EXAMPLE. $H = \{x^2 - y, \text{ constants}\}$ is generating, by Theorem 6, while $-H = \{-x^2 + y, \text{ constants}\}$ is not generating, by Proposition 2.

REFERENCES

[A] V. I. ARNOLD, Representation of continuous functions of three variables by the superpositions of continuous functions of two variables. *Math. USSR-Sb.* **48**, No. 90 (1959), 3–74 (*Amer. Math. Soc. Transl. Ser. 2* **28** (1963), 61–149).

[E] R. ENGELKING, "Dimension Theory," North-Holland, Amsterdam, 1978.

[G1] F. A. GADZHIEV, The Kolmogorov–Slupcki theorem for a three dimensional sphere, *Usperkhi Mat. Nauk* **39**, No. 5 (1984), 51–54 (*Russian Math. Surveys* **39**, No. 5 (1984), 57–61).

[G2] F. A. GADZHIEV. On function of four variables on the three sphere that are not representable as composition of functions of fewer variables, *Soviet Math. Dokl.* 34, No. 1 (1987) 150–153.

[Kol] A. N. KOLMOGOROV, On the representation of continuous functions of many variables by superpositions of continuous functions of one variable and addition, *Dokl. Acad. Nauk SSSR* **114** (1957), 953–956 (*Amer. Math. Soc. Transl. Ser. 2* **28** (1963), 55–61).

[Kur] K. KURATOWSKI, "Topology II," Academic Press, New York, 1968.

[L] G. G. LORENTZ, Approximation of functions, Holt, Rinehart & Winston, New York, 1966.

[M] A. A. MAL'CEV, A topological version of Slupecki's Theorem for certain compacta, *Dokl. Akad. Nauk. SSSR* **188**, No. 1 (1964) (*Soviet Math. Dokl.* **10**, No. 5 (1964), 1071–1074).

[St1] Y. STERNFELD, Hilbert's 13th problem and dimension, *in* "Geometrical Aspects of Functional Analysis, 1987–88," pp. 1–49, Lecture Notes in Mathematics, No. 1376, Springer-Verlag, New York/Berlin.

[St2] Y. STERNFELD, Dense subgroups of $C(K)$ (Stone–Weierstrass type theorems for groups), *Constr. Approx.* to appear.

[St–W] Y. STERNFELD AND Y. WEIT, An approximation theorem for vector valued functions, *in* "Geometrical Aspects of Functional Analysis, 1987–88," pp. 126–138, Lecture Notes in Mathematics, No. 1376, Springer-Verlag, New York/Berlin.