JOURNAL OF APPROXIMATION THEORY 25, 360–368 (1979)

Superpositions of Continuous Functions

Y. STERNFELD

Institute for Advanced Studies, Hebrew University, Jerusalem, Israel, and Department of Mathematics, University of Texas, Austin, Texas 78712

Communicated by G. G. Lorentz

Received October 3, 1977

In this note we present relatively short and simple proofs for some theorems concerning superpositions of functions. We prove in particular the well known theorem of Kolmogorov [8], and its generalization due to Ostrand [10].

Our main observation is that by combining a general duality argument of functional analysis with the ideas introduced by Lorentz [9] and Hedberg [5], we obtain a better understanding of the nature of these theorems, and can avoid some of the difficulties which arose in former proofs.

We use the notation of [3]. $C(X)$ is the Banach space of real valued continuous functions on the compact metric space X, with the norm $\|f\| = \sup_{x\in X} |f(x)|$. We identify the dual $C(X)^*$ of $C(X)$ with the space of real regular Borel measures on X with the total variation as norm. μ^+ (resp. μ^-) denotes the positive (resp. negative) part of the real measure μ, and $|\mu| = \mu^+ + \mu^-$. Clearly

$$\|\mu\| = \||\mu|\| = \mu^+(X) + \mu^-(X). \tag{1}$$

If φ is a continuous function which maps X onto some (compact metric) space Y, and $\mu \in C(X)^*$, then $\mu \circ \varphi$ is the element of $C(Y)^*$ defined by

$$\mu \circ \varphi(V) = \mu(\varphi^{-1}(V)), \qquad V \subset Y. \tag{2}$$

We denote the interval [0, 1] by I, the n dimensional cube by I^n, and the circle by T. dim X is the covering dimension of X.

DEFINITION 1. Let X be a compact metric space. Let F be a family of continuous functions on X. We say that F *uniformly separates the Borel measures on* X if there exists a constant λ, $0 < \lambda \leqslant 1$, such that for each $\mu \in C(X)^*$, $\|\mu \circ \varphi\| \geqslant \lambda \|\mu\|$ for some $\varphi \in F$.

Let us say a word about the intuitive meaning of this concept: if F uniformly separates the Borel measures on X, and H_1, H_2 are disjoint closed subsets of X, then for some $\varphi \in F$ the intersection $\varphi[H_1] \cap \varphi[H_2]$ is "not too large,"

0021-9045/79/040360–09$02.00/0

where "not too large" depends on λ, and on a measure μ in $C(X)^*$ such that H_1, H_2 are the supports of $\mu^{\perp}$, μ^- respectively. In particular a family of functions which uniformly separates Borel measures, separates points. (Given $x_1 \neq x_2$ in X, apply the definition to $\mu = \delta_{x_1} - \delta_{x_2}$.) The converse is false: let $X = I^2$, let $F = \{\varphi_1, \varphi_2\}$ where $\varphi_1(x, y) = x$ and $\varphi_2(x, y) = y$. Clearly F separates points, but for $\mu = \delta_{(0,0)} + \delta_{(1,1)} - \delta_{(0,1)} - \delta_{(1,0)}$ we have $\| \mu \| = 4$, and $\| \mu \circ \varphi_i \| = 0$, $i = 1, 2$. i.e. F does not uniformly separate Borel measures. See [12] where this concept as well as related topics are studied.

The connection between uniform separation and superpositions is given in the following.

THEOREM 1. *Let $F = \{\varphi_i\}_{i=1}^k$ be a finite family of continuous functions on a compact metric space X, with $\varphi_i[X] = Y_i$, $1 \leqslant i \leqslant k$. The family F uniformly separates the Borel measures on X if and only if each $f \in C(X)$ can be represented as*

$$f(x) = \sum_{i=1}^{k} g_i(\varphi_i(x)) \tag{3}$$

with $g_i \in C(Y_i)$, $1 \leqslant i \leqslant k$.

Proof. Let Y denote the disjoint union of the Y_i's, $1 \leqslant i \leqslant k$. Consider the bounded linear operator S: $C(Y) \to C(X)$ defined by

$$Sg(x) = \sum_{i=1}^{k} g(\varphi_i(x)), \qquad g \in C(Y), \qquad x \in X. \tag{4}$$

A routine check shows that the adjoint S^* of S acts according to the formula

$$S^*\mu = \sum_{i=1}^{k} \mu \circ \varphi_i, \qquad \mu \in C(X)^*, \tag{5}$$

and that

$$\| S^*\mu \| = \sum_{i=1}^{k} \| \mu \circ \varphi_i \|. \tag{6}$$

Each $f \in C(X)$ admits a representation of the form (3) if and only if S maps $C(Y)$ *onto* $C(X)$. This occurs if and only if S^* is an *isomorphism into*, i.e., there exists a constant $\alpha > 0$ such that $\| S^*\mu \| \geqslant \alpha \| \mu \|$ for all $\mu \in C(X)^*$ (see [3]).

By 6 this is equivalent to F being uniformly separating Borel measures on X. ∎

A very simple illustration of an application of Theorem 1 is

THEOREM 2. *There exists three real valued analytic functions* $\{\varphi_i\}_{i=1}^3$ *on the circle* T, *such that each* $f \in C(T)$ *can be represented as* $f(t) = \sum_{i=1}^3 g_i(\varphi_i(t))$ *with* $g_i \in C(I)$.

(*See Kahane* [6] *for a similar result. The number three in Theorem* 2 *cannot be reduced as proved in* [11].)

Proof. We realize T as the interval I with its endpoints identified. Set

$$I_1 = (0, \tfrac{1}{3}), \qquad I_2 = (\tfrac{1}{3}, \tfrac{2}{3}), \qquad I_3 = (\tfrac{2}{3}, 1) \quad \text{(Open intervals)} \tag{7}$$

and

$$J_i = T \backslash I_i, \qquad i = 1, 2, 3. \tag{8}$$

Let φ_i, $i = 1, 2, 3$ be any three elements of $C(T)$ such that φ_i / J_i is one to one. (φ_i / J_i is the restriction of φ_i to J_i; the same notation will be used later for measures.) We claim that $F = \{\varphi_i\}_{i=1}^3$ uniformly separates the Borel measures on T, with $\lambda = \frac{1}{3}$.

Indeed, let $\mu \in C(T)^*$ be of norm one. Then $|\mu|$ is a probability measure, and it is easily seen that

$$\sum_{i=1}^{3} |\mu|(J_i) = \sum_{i=1}^{3} \int 1_{J_i} \, d|\mu| = \int \left(\sum_{i=1}^{3} 1_{J_i} \right) d|\mu| \geqslant 2 \tag{9}$$

since $\sum_{1=i}^3 1_{J_i}(t) \geqslant 2$ for all $t \in T$. (1_{J_i} is the indicator function of J_i.) It follows that $|\mu|(J_{i_0}) \geqslant \frac{2}{3}$ for some i_0, $1 \leqslant i_0 \leqslant 3$. Thus, $\| \mu \circ (\varphi_{i_0} / J_{i_0}) \| \geqslant \frac{2}{3}$ since φ_{i_0} is one to one on J_{i_0}.

Clearly $|\mu|(I_{i_0}) \leqslant \frac{1}{3}$, hence, the mass of μ which is outside J_{i_0}, can reduce the norm of $\mu \circ (\varphi_{i_0} / J_{i_0})$ by at most $\frac{1}{3}$, i.e.,

$$\| \mu \circ \varphi_{i_0} \| \geqslant \| \mu \circ (\varphi_{i_0} / J_{i_0}) \| - |\mu|(I_{i_0}) \geqslant \tfrac{2}{3} - \tfrac{1}{3} = \tfrac{1}{3}. \tag{10}$$

Thus F uniformly separates the Borel measures on T with $\lambda = \frac{1}{3}$, and the theorem follows from Theorem 1. ▌

The proofs of the theorems of Kolmogorov and Ostrand require more machinery. We start with some more definitions.

DEFINITION 2. (a) A family U of subsets of a metric space X is said to be *discrete* if its elements have mutually disjoint closures.

(b) $\delta(U)$ is $\sup_{\mathscr{U} \in U}$ diameter $\mathscr{U}$.

(c) If φ is a function on X, we say that φ *separates* U if for each $\mathscr{U}_1, \mathscr{U}_2 \in U$, $\varphi[\mathscr{U}_1] \cap \varphi[\mathscr{U}_2] = \varnothing$.

(d) If $U_1, U_2, \ldots, U_k$ are k families of subsets of X we say that $\{U_i\}_{i=1}^k$ *covers* the set X n *times* ($n \leqslant k$) if each $x \in X$ is an element in some member of U_i for at least n values of i.

The following are trivial observations. (We do not distinguish between U_i and the union of its elements.)

PROPOSITION 1. *Let X be a set, and let $\{U_i\}_{i=1}^k$ be k families of subsets of X. The statements* (a), (b), (c), (d) *are equivalent and imply* (e).

(a) *$\{U_i\}_{i=1}^k$ covers X n times.*

(b) *Each $k - n + 1$ of the families $\{U_i\}_{i=1}^k$ cover X.*

(c) *$\sum_{i=1}^k 1_{U_i}(x) \geqslant n$ for all $x \in X$.*

(d) *$\sum_{i=1}^k \mu(U_i) \geqslant n$ for all probability measures μ on X.*

(e) *For each probability measure μ on X there exists some i_0, $0 \leqslant i_0 \leqslant k$, so that $\mu(U_{i_0}) \geqslant n/k$.*

LEMMA 1. *Let X be a compact metric space, and let $F = \{\varphi_i\}_{i=1}^k$ be a family of continuous functions on X. If for each $\epsilon > 0$, there exists k finite discrete families $U_1, U_2, \ldots, U_k$ of subsets of X so that*

(i) *$\{U_i\}_{i=1}^k$ covers X $\left[\frac{k}{2}\right] + 1$ times,*

(ii) *$\delta(U_i) < \epsilon$, $1 \leqslant i \leqslant k$,*

(iii) *φ_i separates U_i, $1 \leqslant i \leqslant k$.*

Then F uniformly separates the Borel measures on X with $\lambda = 1/k$.

Proof. We wish to show that for each $\mu \in C(X)^*$, $\| \mu \circ \varphi_i \| \geqslant (1/k) \| \mu \|$ for some $\varphi_i \in F$. The measures μ with μ^+ and μ^- having disjoint supports are norm dense in $C(X)^*$, (by regularity) and therefore we may consider such measures only.

So let $\mu \in C(X)^*$ be of norm one, and with $\operatorname{supp} \mu^+ \cap \operatorname{supp} \mu^- = \varnothing$. Let $\epsilon = d(\operatorname{supp} \mu^+, \operatorname{supp} \mu^-)$, and let $\{U_i\}_{i=1}^k$ be the families of sets corresponding to ϵ.

It follows that a member of U_i, $1 \leqslant i \leqslant k$ cannot intersect both $\operatorname{supp} \mu^+$ and $\operatorname{supp} \mu^+$.

By (i) and Proposition 1(e), there exists $1 \leqslant i_0 \leqslant k$ so that

$$| \mu | (U_{i_0}) \geqslant \frac{1}{k}\left(\left[\frac{k}{2}\right] + 1\right) \geqslant \frac{1}{k}\left(\frac{k}{2} + \frac{1}{2}\right) = \frac{1}{2} + \frac{1}{2k}. \qquad (11)$$

Now, since members of U_{i_0} intersect at most one of the sets supp μ^+ and supp μ^-, and since by (iii), φ_{i_0} separates U_{i_0}, it follows from (11) that

$$\| \mu \circ (\varphi_{i_0}/U_{i_0})\| \geqslant \frac{1}{2} + \frac{1}{2k}. \tag{12}$$

Clearly by (11)

$$| \mu | (X \backslash U_{i_0}) \leqslant \frac{1}{2} - \frac{1}{2k}. \tag{13}$$

Hence, by the same reasoning as in the proof of Theorem 2, we get from (12) and (13) that

$$\begin{aligned} \| \mu \circ \varphi_{i_0} \| &\geqslant \| \mu \circ (\varphi_{i_0}/U_{i_0})\| - | \mu | (X \backslash U_{i_0}) \\ &\geqslant \frac{1}{2} + \frac{1}{2k} - \left(\frac{1}{2} - \frac{1}{2k}\right) = \frac{1}{k}. \end{aligned} \tag{14}$$

i.e., F uniformly separates the Borel measures on X with $\lambda = 1/k$. ∎

By Proposition 1, condition (i) of Lemma 1 is equivalent to the following: any $[(k+1)/2]$ of the families U_i cover X. Such a cover by $[(k+1)/2]$ families U_i is of order $[k/2]$ (i.e., at most $[k/2]+1$ of its elements intersect). It follows that the existence of families $\{U_i\}_{i=1}^k$ with (i) and (ii) of Lemma 1 for each $\epsilon > 0$ implies that dim $X \leqslant [k/2]$. Ostrand [10] proved the following converse assertion.

THEOREM 3. *Let X be an n-dimensional compact metric space, let $k \geqslant n+1$, and $\epsilon > 0$.*

There exist k discrete families $\{U_i\}_{i=1}^k$ of subsets of X which cover X $k-n$ times, so that $\delta(U_i) < \epsilon$, $1 \leqslant i \leqslant k$.

Our next lemma proves the existence of functions $\{\varphi_i\}_{i=1}^k$ as in the assumption of Lemma 1, if we are given a suitable sequence of nice coverings of X. We shall do this in a more general setting which will be used later on.

LEMMA 2. *Let X_j, $j = 1, 2, \ldots, L$ be compact metric spaces and let $X = X_1 \times X_2 \times \cdots \times X_L$.*

For each $1 \leqslant j \leqslant L$ let $\{U_m{}^j\}_{m=1}^\infty$ be a sequence of discrete families of subsets of X_j with $\delta(U_m{}^j) \to_{m\to\infty} 0$.

Let U_m, $m = 1, 2, \ldots$ be the family of subsets of X defined by

$$U_m = \{\mathscr{U}^1 \times \mathscr{U}^2 \times \cdots \times \mathscr{U}^L : \mathscr{U}^j \in U_m{}^j\}. \tag{15}$$

and let $\lambda_1, \lambda_2, ..., \lambda_L$ *be reals independent over the rationals. There exist functions* $\tau_j \in C(X_j)$, $1 \leqslant j \leqslant L$ *such that the function* $\varphi \in C(X)$ *defined by*

$$\varphi(x_1, x_2, ..., x_L) = \sum_{j=1}^{L} \lambda_j \tau_j(x_j)$$

separates U_m *for infinitely many m's.*

If $X_1 = X_2 = \cdots = X_L$ *and* $U_m^1 = U_m^2 = \cdots = U_m^L$ *then one can also take* $\tau_1 = \tau_2 = \cdots = \tau_L$.

Proof. Set $C = C(X_1) \times C(X_2) \times \cdots \times C(X_L)$ with the norm $\| (\tau_1, \tau_2, ..., \tau_L) \| \max_{1 \leqslant j \leqslant L} \| \tau_j \|$.

For each integer $\ell \geqslant 1$ let $A_\ell \subset C$ be defined by

$$A_\ell = \left\{ (\tau_1, \tau_2, ..., \tau_L) \colon \varphi(x_1, x_2, ..., x_L) = \sum_{j=1}^{L} \lambda_j \tau_j(x_j) \right. \\ \left. \text{separates } U_m \text{ for some } m \geqslant \ell \right\}. \tag{16}$$

We claim that A_ℓ is open and dense in C for all $\ell \geqslant 1$.

A_ℓ is open: Let $\tau = (\tau_1, \tau_2, ..., \tau_L) \in A_\ell$, i.e. $\varphi = \sum_{j=1}^{L} \lambda_j \tau_j(x_j)$ separates U_m for some $m \geqslant \ell$.

$\epsilon = \inf_{\mathscr{U}, \mathscr{T} \in U_m} d(\varphi[\mathscr{U}], \varphi[\mathscr{T}])$ is positive since U_m is discrete. Let $\delta > 0$ be so small that $\| \tau - \tau' \|_C < \delta$ implies $\| \varphi - \varphi' \|_{C(X)} < \epsilon/2$ where $\varphi'(x_1, x_2, ..., x_L) = \sum_{j=1}^{L} \lambda_j \tau_j'(x_j) \in C(X)$. Then φ' separates U_m too, i.e., $\tau' \in A_\ell$.

A_ℓ is dense: Let $\psi = (\psi_1, \psi_2, ..., \psi_L) \in C$ and $\epsilon > 0$ be given.

We shall construct $\tau \in A_\ell$ with $\| \tau - \psi \| < \epsilon$.

Let $m \geqslant \ell$ be so big that the oscillation of ψ_j on elements of U_m^j is smaller than ϵ for all $1 \leqslant j \leqslant L$. Such an m exists since $\delta(U_m^j) \to_{m \to \infty} 0$.

Let τ_j be defined as follows: τ_j is constant on elements of U_m^j, these constants being *distinct* rationals so that $\| \tau_j / U_m^j - \psi_j / U_m^j \| < \epsilon$. This being possible by the above choice of m, we extend τ_j to the whole of X_j by Tietze's theorem so that $\| \tau_j - \psi_j \| < \epsilon$. Then clearly $\| \tau - \psi \| < \epsilon$ where $\tau = (\tau_1, \tau_2, ..., \tau_L)$. We claim that $\tau \in A_\ell$. Indeed, let $\mathscr{U} = \mathscr{U}^1 \times \mathscr{U}^2 \times \cdots \times \mathscr{U}^L \in U_m$, with $\mathscr{U}^j \in U_m^j$ and $\tau_j[\mathscr{U}^j] = r_j$—the rational value of τ_j on the element $\mathscr{U}^j$ of U_m^j.

If $\varphi(x_1, x_2, ..., x_L) = \sum_{j=1}^{L} \lambda_j \tau_j(x_j)$, then φ attains the constant value $\sum_{j=1}^{L} \lambda_j r_j$ on $\mathscr{U}$. But all the reals $\sum_{j=1}^{L} \lambda_j r_j$ are distinct, since the λ_j's are independent over the rationals, and the values of τ_j on members of U_m^j are distinct rationals. It follows that φ separates U_m, i.e., $\tau \in A_\ell$.

Let $A = \bigcap_{\ell=1}^{\infty} A_\ell$. By the Baire category theorem A is a dense G_δ in C, and each $\tau = (\tau_1, \tau_2, ..., \tau_L) \in A$ has the desired property.

If $X_1 = X_2 = \cdots = X_L$ and $U_m^1 = U_m^2 = \cdots = U_m^L$ the same arguments can be applied with the sets $A_\ell \subset C(X_1)$, $A_\ell = \{\tau \in C(X_1): \varphi(x_1, x_2, ..., x_L) = \sum_{j=1}^{L} \lambda_j \tau(x_j)$ separates U_m for some $m \geqslant \ell\}$. This proves Lemma 2. ▌

Remark. If $X_1 = X_2 = \cdots = X_L = I$, (i.e., $X = I^n$) and the elements of $U_m^1 = U_m^2 = \cdots = U_m^L$ are intervals then one can extend the τ_j's from U_m^j to I by letting them being linear on the intervals in the complement of U_m^j, provided the length of these complementing intervals tends to 0 together with the intervals in U_m^j. (This will be the case in our proof of Kolmogorov's theorem.)

Moreover: $C(X_j) = C(I)$ can be replaced in this case by $\mathrm{Lip}_\alpha(I)$, $0 < \alpha < 1$, i.e., the τ_j's can be chosen to be (nondecreasing) Lip α functions. (See [5].)

KOLMOGOROV'S THEOREM. *Let $n \geqslant 2$. There exist functions ψ_i, $i = 1, 2, ..., 2n + 1$ in* $C(I)$ and reals $\lambda_1, \lambda_2, ..., \lambda_n$ such that each $f \in C(I^n)$ can be represented as

$$f(x_1, x_2, ..., x_n) = \sum_{i=1}^{2n+1} g_i\left(\sum_{j=1}^{n} \lambda_j \psi_i(x_j)\right), \qquad g_i \in C(R).$$

Proof. Set $k = 2n + 1$. For each m, consider a partition of I into m intervals of length $1/m$ each, indexed from 1 to m by the natural order (i.e., the first is $[0, 1/m]$ and the last $[(m-1)/m, 1]$). Let $V_{m,i}$, $1 \leqslant i \leqslant k$ be the family of intervals generated by removing from I those intervals of the above partition with index congruent to i mod k. (All intervals in $V_{m,i}$, except the two extreme ones, are of length $(k-1)/m$, and for each m, $\{V_{m,i}\}_{1=1}^{k}$ covers I $k - 1$ times.)

Set $U_{m,i} = \{I_1 \times I_2 \times \cdots \times I_n: I_j \in V_{m,i}\}$, $1 \leqslant i \leqslant k$. It is easy to check that $\{U_{m,i}\}_{i=1}^{k}$ covers I^n, $k - n = [k/2] + 1$ times.

Let $\{\lambda_j\}_{1=i}^{n}$ independent over the rationals (e.g., $\lambda_j = e^{j-1}$). By Lemma 2 there exists functions ψ_i in $C(I)$, $1 \leqslant i \leqslant k$ and a subsequence $\{m_r\}_{r=1}^{\infty}$ of the integers such that $\varphi_i(x_1, x_2, ..., x_n) = \sum_{j=1}^{n} \lambda_j \psi_i(x_j) \in C(I^n)$ separates $U_{m_r,i}$ for $r = 1, 2, ...$. By Lemma 1, $\{\varphi_i\}_{i=1}^{k}$ uniformly separates the Borel measures on I^n, and the theorem follows from Theorem 1. ▌

OSTRAND'S THEOREM. *Let $X = X_1 \times X_2 \times \cdots \times X_L$ where X_j is a compact metric space of dimension n_j, $1 \leqslant j \leqslant L$. Let $n = \sum_{j=1}^{L} n_j$.*

I. *There exists functions $\{\psi_{i,j}\}_{i=1}^{2n+1}$, $1 \leqslant j \leqslant L$ in $C(X_j)$ such that each $f \in C(X)$ can be represented as*

$$f(x_1, x_2, ..., x_L) = \sum_{i=1}^{2n+1} g_i\left(\sum_{j=1}^{L} \psi_{i,j}(x_j)\right) \qquad \text{with} \qquad g_i \in C(R).$$

II. *If $X_1 = X_2 = \cdots = X_L$, then one can take $\psi_{i,j} = \lambda_j \psi_i$ where $\{\lambda_j\}_{j=1}^{L}$ are reals independent over the rationals.*

Proof. Set $k = 2n + 1$. For each $m \geqslant 1$ let $U^j_{m,i}$ be a discrete family of subsets of X_j, $1 \leqslant i \leqslant k$, so that

(a) $\{U^j_{m,i}\}_{i=1}^k$ covers X_j, $k - n_j$ times, for $1 \leqslant j \leqslant L$.

(b) $\delta(U^j_{m,i}) \to_{m\to\infty} 0$ for $1 \leqslant j \leqslant L$, and $1 \leqslant i \leqslant k$.

Such families exist by Theorem 3. Set

$$U_{m,i} = \{\mathscr{U}_1 \times \mathscr{U}_2 \times \cdots \times \mathscr{U}_L : \mathscr{U}_j \in U^j_{m,i}\}, \quad 1 \leqslant i \leqslant k, \quad m = 1, 2, \ldots .$$

From (a) and (b) it follows that

(i) $\{U_{m,i}\}_{i=1}^k$ cover X $k - n = [k/2] + 1$ times for each $m = 1, 2, \ldots$,

(ii) $\delta(U_{m,i}) \to_{m\to\infty} 0$ for $1 \leqslant i \leqslant k$.

By Lemma 2 there exists a subsequence $\{m_r\}_{r=1}^\infty$ of the integers, and functions $\{\psi_{i,j}\}_{i=1}^k$ in $C(X_j)$ so that $\varphi_i(x_1, x_2, \ldots, x_L) = \sum_{j=1}^L \psi_{i,j}(x_j) \in C(X)$ separates $U_{m_r,i}$ for all $r = 1, 2, \ldots$ and $1 \leqslant i \leqslant k$. (We apply Lemma 2 for $i = 1$ first to get $\{\tau_{1,j}\}_{j=1}^L$ and set $\psi_{1,j} = \lambda_j \tau_{1,j}$. $\varphi_1(x_1, x_2, \ldots, x_L) = \sum_{j=1}^L \psi_{1,j}(x_j)$ separates $U_{m,1}$ for infinitely many m's, and we can apply Lemma 2 again with $i = 2$ on this subsequence to get $\{\psi_{2,j}\}_{j=1}^L$ and so on.)

By Lemma 1 $\{\varphi_i\}_{i=1}^k$ separates the Borel measures on X, and the theorem follows from Theorem 1. For II just apply the second part of Lemma 2. ∎

Remarks. The number $2n + 1$ in both Kolmogorov's and Ostrand's theorems cannot be reduced, at least not for $n = 2, 3, 4$ (see [11] and [12]).

As remarked after the proof of Lemma 2, the functions ψ_i in Kolmogorov's theorem can be chosen in $\text{Lip}_\alpha(I)$, $\alpha < 1$. Fridman [4] proved that the ψ_i's can even be Lip 1 functions. (See also Kahane [14] for a short proof.) However, the ψ_i's cannot be chosen to be continuously differentiable, as proved by Vituskin and Henkin [13], and Kaufman [7].

Demko [1] recently extended Kolmogorov's theorem to bounded continuous functions on R^n, while Doss [2] proved that addition can be replaced by multiplication in this theorem.

References

1. S. Demko, A superposition theorem for bounded continuous functions, *Proc. Amer. Math. Soc.*, in press.
2. R. Doss, Representation of continuous functions of several variables, *Amer. J. Math.* (2) **98** (1976), 375–383.
3. N. Dunford and J. T. Schwartz, "Linear Operators, I. General Theory," Interscience, New York, 1957.
4. B. L. Fridman, Improvement in the smoothness of function in Kolmogorov superposition theorem, *Dokl. Akad. Nauk SSSR* **177** (1967), 1019–1022; MR 38, No. 663; *Soviet Math. Dokl.* **8** (1967), 1550–1553.

5. T. Hedberg, The Kolmogorov superposition theorem, Appendix II, *in* "Topics Approximation Theory," by H. S. Shapiro, Springer–Verlag, New York/Berlin, 1971.
6. J. P. Kahane, "Séries de Fourier absolument convergentes," Springer–Verlag, New York/Berlin, 1970.
7. R. Kaufman, Linear superpositions of smooth functions, *Proc. Amer. Math. Soc.* **46** (1974), 360–362.
8. A. N. Kolmogorov, On the representation of continuous functions of many variables by superpositions of continuous functions of one variable and addition, *Dokl. Akad. Nauk SSSR* **114** (1957), 953–956; *Amer. Math. Soc. Transl.* (2) **28** (1963), 55–61.
9. G. G. Lorentz, "Approximation of Functions," H. H. Rinehart & Winston, New York, 1966.
10. P. A. Ostrand, Dimension of metric spaces and Hilbert's problem 13, *Bull. Amer. Math. Soc.* **71** (1965), 619–622.
11. Y. Sternfeld, Dimension theory and superpositions of continuous functions, *Israel J. Math.* **20** (1975), 300–320.
12. Y. Sternfeld, Uniformly separating families of functions, *Israel J. Math.* **29**(1) (1978), 61–91.
13. A. G. Vituškin and G. M. Henkin, Linear superpositions of functions, *Russian Math. Surveys* (1) **22** (1967), 77–126.
14. J. P. Kahane, Sur le théorem de Kolmogorov, *J. Approximation Theory* **13** (1975), 229–234.