

ON THE STRUCTURE OF REPRESENTATIONS OF
CONTINUOUS FUNCTIONS OF SEVERAL
VARIABLES AS FINITE SUMS OF
CONTINUOUS FUNCTIONS OF
ONE VARIABLE

DAVID A. SPRECHER

1. **Introduction and summary.** Let \mathcal{C}_n denote the set of all real and continuous functions of n variables, defined on the n -dimensional unit cube, \mathcal{E}_n , in Euclidean space, \mathcal{R}_n : in particular, \mathcal{C} will mark the set of real and continuous functions defined on the real line, \mathcal{R} . Considered in this paper are the representations of arbitrary functions of \mathcal{C}_n ($n \geq 2$) as finite superpositions of functions of \mathcal{C} , using only addition. That is, we are interested in the representations of an arbitrary function, $f(x_1, \dots, x_n) \in \mathcal{C}_n$, as

$$(1) \quad f(x_1, \dots, x_n) = \sum_{1 \leq q \leq m} g_q \left[\sum_{1 \leq p \leq n} \psi_{pq}(x_p) \right],$$

where $g_q(\xi) \in \mathcal{C}$, $\psi_{pq}(\xi) \in \mathcal{C}$ and m is arbitrary.

It was proved by Kolmogorov in a remarkable paper [2] that such representations are possible with a fixed set of monotonic increasing, Hölder-continuous functions,¹ ψ_{pq} , $1 \leq p \leq n$, $1 \leq q \leq 2n+1$, the ψ_{pq} , namely, being independent of f . Modifying Kolmogorov's construction, we obtained in [5] the stronger version which states that all functions $f \in \mathcal{C}_n$ can be represented as

$$(2) \quad f(x_1, \dots, x_n) = \sum_{1 \leq q \leq 2n+1} g \left[\sum_{1 \leq p \leq n} \lambda^{pq} \psi(x_p + \epsilon q) \right],$$

where $g(\xi) \in \mathcal{C}$, λ and ϵ are constants, $\psi(\xi) \in \mathcal{C}$ is monotonic increasing and Hölder-continuous, λ , ϵ and ψ being independent of f (ϵ , in fact, can be taken to be any nonzero constant).

One naturally aims to obtain representations such as described with the smoothest possible fixed functions: it is regrettable that the function ψ in our construction is pathological, as described in

Presented to the Society, January 26, 1965 under the title *On the representations of continuous functions of several variables*; received by the editors March 15, 1965.

¹ A function, $\phi(x)$, is Hölder-continuous (with exponent α) if there are constants α and β , $0 < \alpha \leq 1$, for which $|\phi(x) - \phi(y)| \leq \beta|x - y|^\alpha$ for all points, x and y , in the domain of ϕ .

THEOREM 1. *The derivative of ψ is zero almost everywhere.*

An analysis of the constructions in [2] and the proof below reveal that Theorem 1 applies also to the functions ψ_{pq} .

That the dependence of ψ on the parameter q cannot be eliminated is established in

THEOREM 2. *Not every polynomial $u \in \mathcal{C}_n$, $n \geq 2$, can be represented in the form*

$$(3) \quad u(x_1, \dots, x_n) = \sum_{1 \leq q \leq m} g_q \left[\sum_{1 \leq p \leq n} \alpha_{pq} \psi_p(x_p) + \beta_q \right],$$

where $g_q(\xi) \in \mathcal{C}$, $\psi_p(\xi) \in \mathcal{C}$, α_{pq} and β_q are constants, and m is an arbitrary natural number.

Without loss of generality, we may assume that the functions ψ_p are monotonic increasing, for suppose that $\psi_p(x_p) = \psi_p(y_p)$ for some points, $x_p \neq y_p$: then the right side of (3) will yield the same number for the points (x_1, \dots, x_n) and (y_1, \dots, y_n) , whereas we can always find a polynomial $u \in \mathcal{C}_n$ for which $u(x_1, \dots, x_n) \neq u(y_1, \dots, y_n)$. In a given polynomial, $u(x_1, \dots, x_n)$, we substitute $x_p = \psi_p(t_p)$ and write

$$u(x_1, \dots, x_n) = u[\psi_1(t_1), \dots, \psi_n(t_n)] = h(t_1, \dots, t_n).$$

Applying formula (3) to $h(t_1, \dots, t_n)$ and then replacing $\psi_p(t_p)$ by x_p for each p we obtain the equivalent representation

$$(4) \quad u(x_1, \dots, x_n) = \sum_{1 \leq q \leq m} g_q \left[\sum_{1 \leq p \leq n} \alpha_{pq} x_p + \beta_q \right].$$

Thus, Theorem 2 is proved through

LEMMA 1. *Not every polynomial $u \in \mathcal{C}_n$ has a representation as (4).*

The constants α_{pq} can be chosen so as to have the arguments of the functions g_q separate all points of \mathcal{E}_n . This yields

COROLLARY 1. *It is necessary but not sufficient that the arguments of any representation of polynomials in the form (2) separate all points of \mathcal{E}_n .*

We first remark that the functions g_q can always be replaced by a single function g by suitably selecting the constants β_q , an observation for which the author is indebted to G. G. Lorentz.

A result of a similar nature to Theorem 2 is due to Pólya-Szegö

[4], who proved that the function $f(x, y, z) = xy + xz + yz$ cannot be constructed by composing only three functions of two variables each, if the functions used are infinitely differentiable and are defined on all of \mathcal{R}_3 ; Ostrovski [3] has shown that the function

$$\zeta(x, y) = \sum_{1 \leq \nu \leq \infty} x^\nu / y^\nu$$

cannot be represented (in any domain) with a finite number of analytic functions of only one variable, and algebraic functions involving any number of arguments.

The impossibility of (4) was first suggested by A. Douglis, who noted that every continuous function which admits such a representation will necessarily satisfy a partial differential equation with constant coefficients: indeed, we shall show that the right side of (4) is weakly annihilated by a certain m th order differential operator.

2. **Proof of Theorem 1.** To investigate the function ψ , it is necessary to recall the manner in which it was constructed. We do so briefly, omitting entirely the required proofs: these are readily supplied from [5]. For convenience, ψ is being constructed for the positive real line, \mathcal{R}_+ .

Γ will stand for the set of natural numbers; i, j and k will designate indices ranging over Γ . We define the following constants:

$$\begin{aligned} \gamma &\geq 2n + 2, \\ \beta_k &\geq (n + 1)^{k-1}, \\ (5) \quad \delta_k &= \frac{\gamma - 2}{\gamma - 1} \gamma^{-k}, \\ \epsilon_k &= (\gamma - 2) \sum_{\nu \in \Gamma} \gamma^{-\beta_{k+\nu}}, \end{aligned}$$

the last three constants being defined for each k , and corresponding closed intervals,

$$\begin{aligned} (6) \quad E_k(i) &= [i\gamma^{-k}, i\gamma^{-k} + \delta_k], \\ H_k(j) &= [j\gamma^{-\beta_k}, j\gamma^{-\beta_k} + \epsilon_k], \end{aligned}$$

laid, respectively on orthogonal coordinate axes in the real plane.

Let us first note the relevant properties of the $E_k(i)$:

For fixed k , the intervals are separated by gaps of width $(\gamma - 2)^{-1}\delta_k$; for increasing k , they are either nested or disjoint, as follows:

$$(7) \quad E_k(i) \supset E_{k+1}(i')$$

if, and only if, $i' = \gamma i + t$, $0 \leq t \leq \gamma - 2$; the intervals corresponding to values $i' = \gamma i + \gamma - 1$ are located in the gaps separating the $E_k(i)$ for each k . For subsequent calculations, it is important to observe that the initial points of the intervals $E_k(i)$ and $E_{k+1}(i')$ coincide when $i' = \gamma i$, whereas their terminal points are indistinguishable for all values $i' = \gamma i + \gamma - 2$.

The intervals $H_k(j)$, being evidently pairwise disjoint for fixed k , are, like the above intervals, nested or disjoint for increasing k :

$$(8) \quad H_k(j) \supset H_{k+1}(j')$$

if, and only if, $j' = j\gamma^{\beta k - \beta k + 1} + s$, $0 \leq s \leq \gamma - 2$, whereas for the excluded values of s , the $H_{k+1}(j')$ lie in the gaps separating the $H_k(j)$ for fixed k . Again, the initial and terminal points of the intervals $H_{k+1}(j')$ are related to those of the $H_k(j)$ for the values $s = 0$ and $s = \gamma - 1$, respectively.

To construct ψ , we aim to relate inductively on k certain of the intervals $H_k(j)$ to the $E_k(i)$ through a single valued function, $j_k = j_k(i)$, $j_k \in \Gamma$: the correspondence will not be detailed here beyond stating that the relation is subject to the following conditions:

- (i) For each k , the association is monotonic increasing.
- (ii) The $H_k(j_k)$ will be subject to condition (8) with $j = j_k$ and $j' = j_{k+1}$.
- (iii) The widths of the gaps between the $H_k(j_k)$ diminish to zero, uniformly in j_k , as $k \rightarrow \infty$.

ψI will mark the image of an interval I under the mapping ψ :

$$\psi I = \{\psi(\xi) : \xi \in I\}.$$

The structure of the function ψ in formula (2) is described in

THEOREM 3. *The biunique correspondence*

$$(9) \quad \psi E_k(i) = H_k(j_k)$$

defines a point-correspondence,

$$(10) \quad \begin{aligned} \psi(i\gamma^{-k}) &= j_k\gamma^{-\beta k}, \\ \psi(i\gamma^{-k} + \delta_k) &= j_k\gamma^{-\beta k} + \epsilon_k, \end{aligned}$$

which has a unique continuous extension, again denoted ψ , such that $\psi \mathbb{R}_+ = \mathbb{R}_+$. This extension is monotonic increasing, Hölder-continuous (with exponent $\ln 2 / \ln \gamma$), as proved in [5], and has, in addition, the following structure:

Designate by V the set of all points belonging to infinitely many of the intervals $E_k(i)$:

$$V = \left\{ \xi : \xi \in \bigcap_{\nu \in \Gamma} E_{k_\nu}(i_\nu) \right\},$$

$\{k_\nu\}$ and $\{i_\nu\}$ being arbitrary infinite sequences in Γ . Denote by $E'_k(i)$ the closed gaps separating the $E_k(i)$ when k is fixed:

$$E'_k(i) = [i\gamma^{-k} + \delta_k, (i + 1)\gamma^{-k});$$

let W stand for the set of those points of \mathcal{R}_+ which are excluded for some value of k from the interiors of all intervals $E_{k+\nu}(i)$ ($\nu \in \Gamma$): that is, we define the set

$$W = \left\{ \xi : \xi \in \bigcap_{\nu \in \Gamma} E'_{k+\nu}(i_\nu) \right\}.$$

Upon examination W is seen to be expressible as the countable union of perfect sets of measure zero, from which it follows that the set V has full measure. Expressing \mathcal{R}_+ as the union of mutually exclusive sets,

$$\mathcal{R}_+ = (V - W) \cup (W - V) \cup (V \cap W),$$

we can describe the differentiability of ψ as follows:

- (i) $\psi' = 0$ at all points of $V - W$ where the derivative exists.
- (ii) $\psi' = +\infty$ at those points of $W - V$ where ψ' exists.
- (iii) ψ is not differentiable on the set $V \cap W$, this set consisting of the points $i\gamma^{-k}$ and $i\gamma^{-k} + \delta_k$, for all admitted i and k .

The number

$$D\psi(\xi) = \lim_{k \rightarrow \infty} \frac{\psi(\xi + h_k) - \psi(\xi)}{h_k}$$

is a derived number of ψ at ξ if the limit exists for some null-sequence $\{h_k\}$: the notation $D\psi(\xi-)$ and $D\psi(\xi+)$ is self explanatory.

The function ψ being monotonic, its derived numbers are uniquely determined almost everywhere, when $D\psi = \psi'$, as follows from Lebesgue's well-known theorem. Therefore, the assertion made in (i) will be proved once we demonstrate that ψ has a vanishing derived number at all points of the set $V - W$ (we shall show, in fact, that this is true throughout the set V); the claim in (ii) is established by showing that each point of $W - V$ has a derived number $D\psi = +\infty$.

Let ξ be an arbitrary point in \mathcal{R}_+ : if $\xi \in V$, then this point is the infinite intersection of a nested sequence of intervals $E_{k_\nu}(i_\nu)$ for certain infinite sequences of natural numbers, $\{k_\nu\}$ and $\{i_\nu\}$. From these we can extract infinite subsequences, $\{k_{\nu_\mu}\}$ and $\{i_{\nu_\mu}\}$, for which

at least one of the following inequalities is valid for all values of μ :

$$(11) \quad \xi - i_{\nu\mu}\gamma^{-k\nu\mu} \geq \frac{1}{2}\delta_{k\nu\mu}$$

or

$$(12) \quad i_{\nu\mu}\gamma^{-k\nu\mu} + \delta_{k\nu\mu} - \xi \geq \frac{1}{2}\delta_{k\nu\mu}.$$

In the first case, we designate the difference by α_μ and find that

$$(13) \quad 0 \leq D\psi(\xi-) = \lim_{\mu \rightarrow \infty} \frac{\psi(\xi) - \psi(\xi - \alpha_\mu)}{\alpha_\mu} \leq \lim_{\mu \rightarrow \infty} 2\delta_{k\nu\mu}^{-1} \epsilon_{k\nu n} = 0.$$

Similarly, we show that $D\psi(\xi+) = 0$ in case (12) is valid.

The claim made in (ii) is proved in a like fashion, except for an estimate which has to be deduced from the constructions with which ψ was defined. If $\xi \in W$, then ξ can be expressed as the infinite intersection of appropriate gaps, $E'_{k+\nu}(i_\nu)$, k being fixed: according to (6), each of these gaps has width $(\gamma - 2)^{-1}\delta_{k+\nu}$. On the other hand, one readily verifies from [5, §4] that the corresponding gaps between the $H_k(j_k)$ are bounded from below by $\gamma^{-\beta k}$: one then shows that the gaps between the intervals $H_{k+\nu}(j_{k+\nu})$ are of width not less than

$$(14) \quad b = \frac{1}{2}c\gamma^{-\beta k}2^{-\nu},$$

c being a fixed positive constant. As in the previous case, we can find an infinite sequence $\{\nu_\mu\}$ for which either

$$(15) \quad \psi(\xi) - \psi((i_{\nu_\mu} + 1)\gamma^{-k-\nu_\mu}) \geq \frac{1}{2}b$$

or

$$(16) \quad \psi(i_{\nu_\mu}\gamma^{-k-\nu_\mu} + \delta_{k+\nu_\mu}) - \psi(\xi) \geq \frac{1}{2}b$$

for all μ . Estimating now the appropriate difference quotient, we deduce that one of the relations, $D\psi(\xi-) = +\infty$ or $D\psi(\xi+) = +\infty$, must hold, and (ii) follows.

To prove (iii), we observe that for each point $i\gamma^{-k}$, i and k being held fixed, there are numbers $i_\nu\gamma^{-k-\nu}$ such that $i\gamma^{-k} = i_\nu\gamma^{-k-\nu}$ for all $\nu \in \Gamma$, and furthermore,

$$i\gamma^{-k} = \lim_{\nu \rightarrow \infty} (i_\nu\gamma^{-k-\nu} + \delta_{k+\nu}) = \lim_{\nu \rightarrow \infty} (i_\nu\gamma^{-k-\nu\mu} - (\gamma - 2)^{-1}\delta_{k+\nu}).$$

With $\xi = i\gamma^{-k}$ and $\xi = i\gamma^{-k} + \delta_k$, respectively, it follows from (11-16) that $D\psi(i\gamma^{-k}+) = 0$, $D\psi(i\gamma^{-k}-) = +\infty$, $D\psi((i\gamma^{-k} + \delta_k)+) = +\infty$, and $D\psi((i\gamma^{-k} + \delta_k)-) = 0$ for all i and k , thereby completing the proof.

3. Proof of Theorem 2. As noted, all we need prove is Lemma 1.

This is proved with the help of the theory of generalized functions (distributions), whereby every continuous function is differentiable. The simplest approach for our purpose is that of Bochner [1]. We recall that in this version of the theory, $\partial f/\partial x_p$ designates the equivalence class of weakly convergent sequences with representative $\{\Delta_p f\}$, where

$$\Delta_p f = \frac{1}{\epsilon} [f(x_1, \dots, x_p + \epsilon, \dots, x_n) - f(x_1, \dots, x_n)];$$

we mention in passing that every ordinary derivative is also a weak derivative (we use the same notation for both).

With the notation

$$h^q = \sum_{1 \leq p \leq n} \alpha_{pq} x_p + \beta_q$$

we consider the partial derivatives $\partial g(h^q)/\partial x_p$ for fixed p and q . We consider the operators

$$\alpha_{2q} \frac{\partial}{\partial x_1} - \alpha_{1q} \frac{\partial}{\partial x_2},$$

where, for each q ,

$$\left(\alpha_{2q} \frac{\partial}{\partial x_1} - \alpha_{1q} \frac{\partial}{\partial x_2} \right) g(h^q) = \left\{ \alpha_{2q} \frac{1}{\epsilon_1} [g(h^q + \alpha_{1q}\epsilon_1) - g(h^q)] - \alpha_{1q} \frac{1}{\epsilon_2} [g(h^q + \alpha_{2q}\epsilon_2) - g(h^q)] \right\}.$$

If we set $\epsilon_1 = \epsilon'/\alpha_{1q}$ in the first difference quotient, and $\epsilon_2 = \epsilon'/\alpha_{2q}$ in the second (as we may without destroying the equivalence class), then we find that each of these difference quotients equals

$$\alpha_{1q}\alpha_{2q} \frac{1}{\epsilon'} [g(h^q + \epsilon') - g(h^q)],$$

from which it follows that

$$\left(\alpha_{2q} \frac{\partial}{\partial x_1} - \alpha_{1q} \frac{\partial}{\partial x_2} \right) g(h^q) = 0.$$

This being valid for each q , we have

$$(17) \quad \prod_{1 \leq q \leq m} \left(\alpha_{2q} \frac{\partial}{\partial x_1} - \alpha_{1q} \frac{\partial}{\partial x_2} \right) \sum_{1 \leq q \leq m} g(h^q) = 0.$$

On the other hand, let $v(x_2, \dots, x_n)$ be any polynomial (or any continuous function for that matter); the function

$$u(x_1, \dots, x_n) = x_1^{m+1} + v(x_2, \dots, x_n)$$

is not annihilated by the differential operator in (15), thereby proving that at least this function cannot be represented in the manner of Theorem 3.

BIBLIOGRAPHY

1. S. Bochner, *Lectures on Fourier integrals*, Princeton Univ. Press, Princeton, N. J., 1959; Chapter VI.
2. A. N. Kolmogorov, *On the representation of continuous functions of several variables by superposition of continuous functions of one variable and addition*, Dokl. Akad. Nauk. SSSR 114 (1957), 953–956.
3. A. Ostrovski, *Über Dirichletschen Reihen und algebraische Differentialgleichungen*, Math. Z. 8 (1920), 241–298.
4. G. Pólya and G. Szegő, *Aufgaben und Lehrsätze aus der Analysis*, Problems 119–119(a), 2nd ed., Springer-Verlag, Berlin, 1954, pp. 61–62, 220–223.
5. D. A. Sprecher, *On the structure of continuous functions of several variables*, Trans. Amer. Math. Soc. 115 (1964), 340–355.

SYRACUSE UNIVERSITY