ON THE STRUCTURE OF CONTINUOUS FUNCTIONS
OF SEVERAL VARIABLES

BY

DAVID A. SPRECHER[1]

1. Introduction and statement of the results. This paper deals with the representations of arbitrary real continuous functions, of any number of variables, as finite sums of real continuous functions of one variable. The first such representation theorem was proved by A. N. Kolmogorov [5] in 1957. Our Theorem 1 is a refinement of Kolmogorov's result, and contains the latter as a special case; it, in turn, is a special case of Theorem 2 below.

Let $\mathscr{E}^n$ denote the cartesian product

$$\mathscr{E}^n = \prod_{1 \leq p \leq n} \mathscr{E}_p$$

of closed unit intervals, the interval $\mathscr{E}_p$ being laid on the pth coordinate axis of a rectangular coordinate system in n-dimensional Euclidean space, $\mathscr{R}^n$; designate points in $\mathscr{R}^n$ by $\mathbf{x}$:

$$\mathbf{x} = (x_1, \cdots, x_n);$$

$\mathscr{R}$ will denote the real line and $\mathscr{E}$ the unit interval thereof. The image of a set $A \subset \mathscr{R}^n$ under a mapping ϕ is designated by ϕA:

$$\phi A = \{ \phi(\mathbf{x}) \mid \mathbf{x} \in A \}.$$

THEOREM 1. *For each integer $N \geq 2$, there exists a real, monotonic increasing function, $\psi(x) \in \mathrm{Lip}[\ln 2/\ln(2N + 2)]$, $\psi\mathscr{E} = \mathscr{E}$, dependent on N, and having the following property[2]:*

For each preassigned number $\delta > 0$ there is a rational number ϵ, $0 < \epsilon \leq \delta$, such that for $2 \leq n \leq N$ every real continuous function of n variables, $f(\mathbf{x})$, defined on $\mathscr{E}^n$, has a representation as

$$(1.1) \qquad f(\mathbf{x}) = \sum_{0 \leq q \leq 2n} \chi \left[\sum_{1 \leq p \leq n} \lambda^p \psi(x_p + \epsilon q) + q \right],$$

[2] $\psi(x)$ belongs to class $\mathrm{Lip}[\alpha]$ if there are constants c and α, $0 < \alpha \leq 1$, for which $|\psi(x) - \psi(y)| \leq |x - y|^\alpha$ for all points, x and y, in the domain of ψ.

where the function χ is real and continuous, and λ a constant independent of f.

With an obvious modification in requirement (4.7) and the proof of Lemma 2 below, one can prove Theorem 1 for representations of the form

$$(1.2) \qquad f(\mathbf{x}) = \sum_{0 \leq q \leq 2n} \chi \left[\sum_{1 \leq p \leq n} \lambda^{pq} \psi(x_p + \epsilon q) \right],$$

λ again being a constant independent of f. We mention in passing that the powers of the single number λ can be replaced with any set of n real numbers λ_p, $1 \leq p \leq n$, which do not satisfy any homogeneous linear equation with integral coefficients: this is readily verified by analysing the proof of Lemma 2 below.

Kolmogorov [5] proved the existence of representations of the form

$$(1.3) \qquad f(\mathbf{x}) = \sum_{0 \leq q \leq 2n} \chi_q \left[\sum_{1 \leq p \leq n} \psi_{pq}(x_p) \right];$$

his $2n^2 + n$ functions ψ_{pq} are here replaced by the single function ψ plus constants; the χ_q are replaced by a single χ, a simplification that follows, as was first noted by G. G. Lorentz [6], from refining one of Kolmogorov's own constructions; the intervals in [5, Lemma 2(3)] are modified so as to be disjoint for variable q.

The specific constructions in our paper are much simpler than the corresponding ones suggested in [5], because the construction of ψ is independent of the parameter q, which appears in the remaining constructions only as an additive constant; they do, however, share their topological properties. In particular, Lemmas 1-2 above imply the corresponding Lemmas 1-2 in [5], but not conversely. Kolmogorov's result depends on three lemmas, stated in [5] without proof. Elegant proofs of these have been given by J. Kim in his master's thesis (University of Maryland, 1960, unpublished); subsequent proofs of Kolmogorov's theorem were given by Arnol'd [2], Lorentz [6] and Tihomirov [9].

For the case $n = 3$, Theorem 1 contains as a special case a result of Arnol'd [1], who obtained a corresponding representation with nine summands and considerably more complicated arguments (see also [5]).

Theorem 1 depends on the lemmas stated below. Since our results are valid for any preassigned integer $N \geq n$, we shall assume that $N = n$. It will be evident from the construction of the function ψ that we can do so without any loss of generality. Throughout this paper, Γ will designate the set of natural numbers; p will be an index with domain $1 \leq p \leq n$. The diameter of a set $A \subset \mathscr{R}^n$ is the number

$$\delta(A) = \sup_{\mathbf{x}, \mathbf{y} \in A} |\mathbf{x} - \mathbf{y}|,$$

$|\mathbf{x} - \mathbf{y}|$ denoting the Euclidean distance between the points in question.

Let γ be a constant, $\gamma > 1$. For each $k \in \Gamma$, S_k^0 will denote a certain finite family of pairwise disjoint, nonempty, closed sets in $\mathcal{R}^n$, with diameters not exceeding γ^{-k-1}; the members of S_k^0 will be designated by $S_k^0(i)$, where i is an index with a certain finite domain Λ_k, for each k, Λ_k depending on the particular choice of sets involved.

Let $\mathbf{v}$ be a fixed vector, $\mathbf{v} = (v_1, \cdots, v_n)$; let q and k be fixed. For each $i \in \Lambda_k$, designate by $S_k^q(i)$ the set $S_k^0(i)$ displaced along $\mathbf{v}$ by a distance q:

$$(1.4) \qquad S_k^q(i) = \{\mathbf{x} + q \cdot \mathbf{v} | \ \mathbf{x} \in S_k^0(i), i \in \Lambda_k\};$$

denote the family of sets (1.4) by S_k^q.

LEMMA 1. *The family S_k^0 and the vector $\mathbf{v}$ can be so determined, that the union of families S_k^q, $0 \leqq q \leqq n + m$, covers $\mathcal{R}^n$ $m + 1$ times for each k. Furthermore, the members of each family can be so labeled that, for each $i = i_0$ in Λ_k, k being held fixed, there is a nonempty intersection of sets $S_k^q(i)$:*

$$(1.5) \qquad S_k(\mathbf{i}) = \bigcap_{0 \leqq q \leqq 2n} S_k^q(i_q) \neq \emptyset,$$

where $i_1, \cdots, i_{2n}$ are uniquely determined numbers in Λ_k, and $\mathbf{i} = (i_0, \cdots, i_{2n})$.

It is the special structure of the families S_k^q that permits us to prove Theorem 1 with a single function ψ.

LEMMA 2. *Let a number $\delta > 0$ be given; select a natural number $k = k_0$ so large, that $\epsilon = (\gamma - 1)^{-1}\gamma^{-k_0} \leqq \delta$. There exists a real, monotonic increasing function, $\psi(x)$, dependent on n, such that $\psi \mathcal{S} = \mathcal{S}$, and a real constant, λ, such that the functions*

$$(1.6) \qquad g^q(\mathbf{x}) = \sum_{1 \leqq p \leqq n} \lambda^p \psi(x_p + \epsilon q) + q,$$

$0 \leqq q \leqq 2n$, have the property

$$(1.7) \qquad g^q S_k^q(i) \cap g^r S_k^r(i') = \emptyset$$

for each k, unless $r = q$ and $i' = i$. Moreover, the function ψ can be constructed to belong to class $\mathrm{Lip}[\ln 2/\ln(2n + 2)]$.

A function f separates the distinct points $\mathbf{x}$ and $\mathbf{y}$ if their corresponding images under f are distinct. For an arbitrary continuous function, $f(\mathbf{x})$, imagine a representation such as

$$(1.8) \qquad f(\mathbf{x}) = \sum_{0 \leqq q \leqq m} \chi[h^q(\mathbf{x})],$$

the functions χ and h^q being continuous, and m arbitrary. Because the right side of (1.8) is invariant under all permutations of the functions h^q, it follows that it separates the points $\mathbf{x} \neq \mathbf{y}$ only if for no permutation $(\alpha_0, \cdots, \alpha_m)$ of the tuple $(0, 1, \cdots, m)$ is the set $(h^0(\mathbf{x}), \cdots, h^m(\mathbf{x}))$ a permuta-

tion of the set $(h^{\alpha 0}(\mathbf{y}), \cdots, h^{\alpha m}(\mathbf{y}))$. Clearly, this requirement is necessary for all representations of continuous functions in the form (1.8). We prove elsewhere [11] that it is not sufficient. We prove, namely, that there are continuous functions h^q which separate all points of $\mathscr{R}^n$, but for which not every continuous function $f(\mathbf{x})$, $\mathbf{x} \in E^n$, can be represented in the form (1.8).

The functions g^q of Lemma 2 separate the sets S_{ki}^q for each k: such functions have the property stated in

COROLLARY 1. *Any set of $2n + 1$ functions satisfying condition (1.7) separate all points of $\mathscr{S}^n$.*

THEOREM 2. *Let $\{h^q(\mathbf{x})\}$ be any set of $n + m$ continuous functions with domain E^n which, in addition to property (1.7), separates all points of $\mathscr{S}^n$. Then every continuous function, $f(\mathbf{x})$, defined on $\mathscr{S}^n$, can be represented in the form (1.8) with $n + m$ superpositions of a continuous function χ, where $m \geqq 1$.*

It is requirement (1.7) for variable i that restricts the smoothness of ψ; one can readily show that a function satisfying that condition cannot belong to class Lip[1]. Therefore, in order to determine the best possible ψ for which Theorem 1 can be proved, one has to know whether or not condition (1.7) is necessary; we do not know this to be the case.

The author wishes to express his gratitude to A. Douglis for his invaluable suggestions and criticism; his insight brought into focus many of the results presented here.

2. **Concomitant remarks.** The existence of continuous functions of several variables that are not finite superpositions of continuous functions of a lesser number of variables was conjectured by Hilbert in the thirteenth problem of his famous Paris lecture of 1900 [3]. Studying the roots of polynomial equations, he concluded that there are proper functions of more than one variable. Hilbert conjectured, namely, that the roots of the equation

$$x^7 + ax^3 + bx^2 + cx + 1 = 0$$

are functions of its coefficients that cannot be represented as finite superpositions of continuous functions of only two arguments. This conjecture was refuted in 1957 by Arnol'd [1] and Kolmogorov [5]; Kolmogorov, in fact, proved the general representation theorem quoted in the introduction.

In connection with his thirteenth problem, Hilbert also raised questions concerning the classification of continuous functions of n variables according to their representability in terms of continuous functions of m variables, $m \leqq n$, of a given class (say, analytic, algebraic, etc.) [4]. Owing to Theorem 1 and Theorem 2, such a classification fails when we wish to classify

continuous functions of n variables according to their representability in terms of continuous functions of m variables, for then $m = 1$. In this case, however, one could use as classification index the least number of summands in (1.8) for fixed functions h^q.

We mention in passing that every continuous function of n variables, $f(\mathbf{x})$, can be represented as

$$(2.1) \qquad\qquad f(\mathbf{x}) = g\left[\sum_{1 \le p \le n} \psi_p(x_p)\right],$$

where the functions ψ_p are monotonic increasing and independent of f, if the function g is permitted to be discontinuous [8].

Extensive work along the lines suggested by Hilbert in [4] was done by A. G. Vituškin; we wish to single out one significant result [10]:

We consider those functions of n variables having all partial derivatives of orders $\le p$, and whose derivatives of order p belong to class $\mathrm{Lip}[\alpha]$: we associate with such functions the index $\mu = n/(p + \alpha)$, where we assume that $p + \alpha \ge 1$. Vituškin's theorem states that not all continuous functions with index μ can be represented with superpositions of functions with index μ_0, if $\mu < \mu_0$.

It is not yet known how to determine the class, say, of analytic functions of several variables that can be represented as finite superpositions of analytic functions of one variable; similar questions arise for other classes of functions. In particular, it would be of interest to determine whether or not a Lipschitz continuity of f in (1.1) implies a Lipschitz continuity of χ; in general, we do not expect f and χ to belong to the same exponent α. In this connection, also the following question arises: In what way is the smoothness of χ and ψ in (1.1) dependent on the number of superpositions?

The number of summands in representations of the form (1.1) can be reduced when one is willing to sacrifice some of the properties of our results (see, for example, (2.1) above). The function $f(x, y) = xy$ can be written as

$$f(x, y) = \chi_1(x + y) + \chi_2(x + (- y)),$$

where $\chi_1 = \xi^2/4$ and $\chi_2 = - \xi^2/4$: here, however, the arguments involved, $x + y$ and $x + (- y)$, depend on f. Arnol'd [1] obtained, for $n = 3$, representations with only three superpositions of two variables each, but there the functions vary over a space that is considerably more complicated than an interval; they take on values on the cartesian product on an interval and a tree.

3. **Proof of Theorem 2.** Of the families S_k^q we need in this section only their covering property and (1.5); the desired function will be obtained as a limit, $\chi = \lim_{r \to \infty} \chi_r$, of a sequence of polygonal arcs, $\{\chi_r\}$, to be deter-

mined by induction on r, simultaneously with a subsequence, $\{k_r\}$, of Γ. Except for a slight modification, the construction that follows is that of Kolmogorov, but the estimates in the convergence proof are new.

For each k, denote by S_k the union

$$(3.1) \qquad S_k = \bigcup_{i_q \in \Lambda_k} S_k(\mathbf{i})$$

(see (1.5)). Clearly, if the (fixed) function $\chi_{r-1}(\xi)$ is continuous for $\xi \in \mathscr{R}$, then the function

$$(3.2) \qquad f_{r-1}(\mathbf{x}) = \sum_{0 \le q \le n+m} \chi_{r-1}[h^q(\mathbf{x})]$$

is continuous for $\mathbf{x} \in \mathscr{E}^n$. Defining $\chi_0(\xi) \equiv 0$, so that $f_0(\mathbf{x}) \equiv 0$, we set with $S_{k_0} = \mathscr{E}^n$,

$$(3.3) \qquad \begin{aligned} \mu_r &= \sup_{S_{k_r} \cap \mathscr{E}^n} |f(\mathbf{x}) - f_r(\mathbf{x}), \\ \nu_r &= \sup_{\mathscr{E}^n} |f(\mathbf{x}) - f_r(\mathbf{x})|; \end{aligned}$$

clearly, $\mu_r \le \nu_r$, and, in particular, $\mu_0 = \nu_0$.

The numbers μ_{r-1} and ν_{r-1} are determined by $f_{r-1}(\mathbf{x})$ and k_{r-1}; since the inequality $\delta(S_k^q(i)) \le \gamma^{-k-1}$ is uniform in i, we can select k_r so large, that

$$(3.4) \qquad \begin{aligned} \rho_{r-1} &= \sup_{\substack{\bigcup \\ 0 \le q \le n+m} S_{k_r}^q} \left\{ \sup_{\mathbf{x},\mathbf{y} \in S_{k_r}^q(i)} |[f(\mathbf{x}) - f_{r-1}(\mathbf{x})] - [f(\mathbf{y}) - f_{r-1}(\mathbf{y})]| \right\} \\ &\le \frac{1}{n+m+1} \mu_{r-1}. \end{aligned}$$

Let $k_0 = 1$; by definition, $\chi_0 \equiv 0$; assuming that the continuous function χ_{r-1} and number k_{r-1} are already fixed; we select k_r such that (3.4) holds. In the sets $S_{k_r}(\mathbf{i})$ we fix arbitrary points, $x_{k_r,\mathbf{i}}$, for each $i_q \in \Lambda_{k_r}$; we assume tacitly that all these points belong to $\mathscr{E}^n$.

On every interval $h^q S_{k_r}^q(\mathbf{i})$ define the function $\chi_r(\xi)$ by the equation

$$(3.5) \qquad \chi_r(\xi) = \chi_{r-1}(\xi) + \frac{1}{n+m+1} [f(x_{k_r,\mathbf{i}}) - f_{r-1}(x_{k_r,\mathbf{i}})],$$

this being possible because the intervals under consideration are all mutually exclusive. For the specified domain of ξ, $\chi_r(\xi)$ satisfies the inequality

$$(3.6) \qquad |\chi_r(\xi) - \chi_{r-1}(\xi)| \le \frac{1}{n+m+1} \mu_{r-1}.$$

On the complement of the intervals $h^q S_{k_r}^q(\mathbf{i})$ complete the definition of $\chi_r(\xi)$ arbitrarily, but so that it remains subject to inequality (3.6) and continu-

ous, a process justified by the Tietze-Urysohn extension theorem.

We shall now estimate the difference, $f - f_r$, at arbitrary points in S_{k_r} and E^n, respectively: these estimates are possible because the functions $h^q(x)$ together separate all points of $\mathscr{E}^n$. The first estimate will enable us to show that the sequence $\chi_r(\xi)$ converges uniformly to a continuous limit function; the estimate over E^n will yield Theorem 2.

Owing to (3.2) we may write

$$(3.7) \qquad f(\mathbf{x}) - f_r(\mathbf{x}) = f(\mathbf{x}) - f_{r-1}(\mathbf{x}) - \sum_{0 \leq q \leq n+m} [\chi_r(h^q) - \chi_{r-1}(h^q)],$$

where $h^q = h^q(\mathbf{x})$. For arbitrary points, $\mathbf{x}$, in S_{k_r}, each summand, $\chi_r - \chi_{r-1}$, can be represented by (3.5) for some vector i; with the help of (3.4), we find that the right side of (3.7) does not exceed $\mu_{r-1}/(n + m + 1)$ in absolute value so that, in particular, $\mu_r \leq \mu_{r-1}/(n + m + 1)$. Hence,

$$(3.8) \qquad \mu_r \leq (n + m + 1)^{-r}\mu_0 = (n + m + 1)^{-r}\nu_0;$$

applied to (3.6), the last inequality shows that

$$(3.9) \qquad |\chi_r(\xi) - \chi_{r-1}(\xi)| < (n + m + 1)^{-r}\nu_0,$$

this inequality being valid for all points $\xi \in \mathscr{R}$. It follows that the sequence χ_r converges uniformly to a continuous limit function, χ, as $r \to \infty$.

To justify representation (1.8), we shall show that $\nu_r \to \infty$ as $r \to \infty$. Consider equation (3.7) for an arbitrary point, $\mathbf{x} \in \mathscr{E}^n$. Because of the properties of the families S_k^q, there are at least $m + 1$ values, q', of q, for which $\mathbf{x} \in S_{k,r}^q(\mathbf{i})$ for some $m + 1$ vectors i. For these values of q we represent the difference, $\chi_r - \chi_{r-1}$, with the help of (3.5); in particular, we have:

$$(3.10) \qquad \chi_r(\xi) - \chi_{r-1}(\xi) = \frac{1}{n + m + 1} [f(\mathbf{x}) - f_{r-1}(\mathbf{x})] - \alpha,$$

where, owing to (3.5) and (3.9),

$$(3.11) \; \alpha \leq (n + m + 1)^{-1}\rho_{r-1} \leq (n + m + 1)^{-2}\mu_{r-1} \leq (n + m + 1)^{-r-1}\nu_0.$$

Substituting (3.10) into (3.7), we obtain

$$f(\mathbf{x}) - f_r(\mathbf{x}) = \frac{1}{n + m + 1} [f(\mathbf{x}) - f_{r-1}(\mathbf{x})] - \alpha + \sum_{q \neq q'} [\chi_r(h^q) - \chi_{r-1}(h^q)].$$

The n summands are estimated with the help of (3.9) which, together with (3.3) and (3.11) shows that

$$|f(\mathbf{x}) - f_r(\mathbf{x})| \leq n(n + m + 1)^{-1}\nu_{r-1}$$
$$+ [m(n + m + 1)^{-1} + n](n + m + 1)^{-r}\nu_0.$$

Because this inequality is valid at each point of $\mathscr{E}^n$, it holds also for ν_r,

the supremum of the difference $|f - f_r|$; an iteration leads to the inequality

$$\nu_r \leqq \left\{ \frac{n^r}{(n+m+1)^r} + \left[\frac{n}{(n+m+1)^r} + \frac{m}{(n+m+1)^{r-1}} \right] \frac{n^r - 1}{n-1} \right\} \nu_0 = o(1)$$

as $r \to \infty$. This completes the proof of Theorem 2.

4. Construction of the function $\psi(x)$. We shall construct in this section the monotonic increasing function $\psi(x)$, for $x \in \mathcal{R}$, belonging to class $\mathrm{Lip}[\ln 2/\ln(2n+2)]$. This function will be determined by a correspondence between certain point-sets, generated from suitably selected systems of nested intervals. Throughout this section, i will be an index with domain $\Gamma' = \{0, 1, 2, \cdots\}$.

Let γ denote a fixed integer, $\gamma \geqq 2n+2$; for each natural number k define the numbers

(4.1)
$$e_k(i) = i\gamma^{-k},$$
$$\delta_k = \frac{\gamma-2}{\gamma-1}\gamma^{-k},$$

with corresponding intervals

(4.2) $$E_k(i) = [e_k, e_k + \delta_k],$$

where $e_k = e_k(i)$. As i varies, with k held fixed, these intervals are separated by gaps of width $(\gamma-1)^{-1}\gamma^{-k}$; for increasing k they have intersection properties as follows:

(4.3) $$E_k(i) \cap E_{k+1}(i') \neq \emptyset$$

if and only if

(4.4) $$i' = \gamma i + t \qquad (0 \leqq t \leqq \gamma - 2).$$

Furthermore, for each i,

(4.5)
$$e_k(i) = e_{k+1}(\gamma i),$$
$$e_k(i) + \delta_k = e_{k+1}(\gamma i + \gamma - 2) + \delta_{k+1}$$

(see Figure 1); It follows that the $E_k(i)$ are nested or disjoint:

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

whenever (4.3) is valid.

As i varies over Γ', and as t admits values between 0 and $\gamma - 1$, the right side of (4.4) defines a single valued mapping of Γ' onto itself. Thus, each element of Γ' is completely determined by the pair (i, t), in a one-to-one manner, and conversely. Only for $t = \gamma - 1$ is the intersection (4.3) empty (see Figure 1).
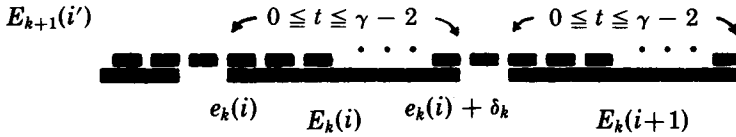
FIGURE 1

We construct by induction a sequence of natural numbers, $\{\beta_k\}$, as follows: Let $\lambda > 0$ be a fixed real number which is not a root of any polynomial equation of degree less than n with integral coefficients. For each k, consider the domain

$$\mathscr{A}_k = \{ -\gamma^{\beta_k}, \cdots, -1, 0, 1, \cdots, \gamma^{\beta_k} \};$$

let $h_1, \cdots, h_n$ designate arbitrary values in $\mathscr{A}_k$; let $\mathbf{H}_k$ denote the set of all nonzero vectors, $(h_1, \cdots, h_n)$, with components in $\mathscr{A}_k$.

Let $\beta_1 = 1$; suppose β_k already determined for $k \geq 1$, and select β_{k+1} such that

$$(4.7) \qquad \gamma^{-\beta_{k+1}} < \gamma^{-\beta_k-1} \min_{\mathbf{H}_k} \left| \sum_{1 \leq p \leq n} h_p \lambda^p \right|:$$

this is clearly possible, since the right side does not vanish. It is convenient to let λ be an algebraic number of degree n; applying [7, Theorem 3] to the right side of (4.7), one can easily show that it exceeds $\gamma^{-n\beta_k-c}$, where c is a constant depending only on λ and n. Thus, (4.7) will be satisfied if we take $\beta_{k+1} \geq n\beta_k + c$; an iteration shows that

$$(4.8) \qquad \beta_{k+1} \geq \left(1 + \frac{c}{n-1} \right) n^k - \frac{c}{n-1}.$$

Now define the numbers

$$(4.9) \qquad h_k(j) = j\gamma^{-\beta_k}, \quad \epsilon_k = (\gamma - 2) \cdot \sum_{\nu \in \Gamma} \gamma^{-\beta_k+\nu}$$

for $j \in \Gamma'$, and corresponding intervals

$$(4.10) \qquad H_k(j) = [h_k, h_k + \epsilon_k],$$

where $h_k = h_k(j)$, laid on a coordinate axis perpendicular to $\mathscr{R}$. We observe that, for fixed k, consecutive intervals are separated by gaps of width $\gamma^{-\beta_k} - \epsilon_k$, and one verifies readily that

$$(4.11) \qquad H_k(j) \cap H_{k+1}(j') \neq \emptyset$$

if and only if

$$(4.12) \qquad j' = j\gamma^{\beta_{k+1}-\beta_k} + s \qquad (0 \leq s \leq \gamma - 2);$$

moreover,

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

whenever (4.11) holds. Thus, for increasing k, these intervals have inter-section properties similar to those of the intervals $E_k(i)$. For $\gamma - 1 \leq s \leq \gamma^{\beta_{k+1}-\beta_k} - 1$ the intersection (4.11) is empty.

For each k, we relate certain of the intervals $H_k(j)$ to the $E_k(i)$; the association is done inductively, as follows:

LEMMA (i). *To each value of the index i, k being fixed, we can associate a value, j_k, of j, different j_k to different i, such that*

(a) *For each k the association is monotonic increasing;*

(b) *The $H_k(j_k)$ will be subject to condition (4.13) for the values $j = j_k$ and $j' = j_{k+1}$, for each k;*

(c) *The widths of the gaps between the intervals $H_k(j_k)$ diminish to zero, uniformly in j_k, as $k \to \infty$.*

This interval correspondence will lead to a point-correspondence between the set of points that belong to infinitely many $E_k(i)$, and the analogous set, consisting of all points belonging to infinitely many $H_k(j_k)$. The two sets thereby related being dense (see below), this correspondence will enable us suitably to define ψ.

Proof. Set $j_1 = i$; for some value of $k \geq 1$, suppose j_k already determined as a function of i; define $j_{k+1} = j_{k+1}(i,t)$ as

(4.14) $j_{k+1} = \begin{cases} j_k \gamma^{\beta_{k+1}-\beta_k} + t & (0 \leq t \leq \gamma - 2), \\ \langle \frac{1}{2} \cdot [j_{k+1}(i, \gamma - 2) + j_{k+1}(i+1,0)] \rangle & (t = \gamma - 1), \end{cases}$

the brackets $\langle \, \rangle$ indicating the integral part of the enclosed quotient.

For fixed k, this correspondence is increasing, as demanded in (a). There-by, by induction, a value j_k is assigned to every pair (i,t), where $i \in \Gamma'$, and $0 \leq t \leq \gamma - 1$; in view of the one-to-one relation between these pairs and all numbers of Γ', it follows that the association (4.14) assigns a value, j_k, to each number, i', where $i' = \gamma i + t$, t admitting values between 0 and $\gamma - 1$ as described (see Figure 2). Also, owing to this correspondence, both (4.3) and (4.11) are valid if and only if $0 \leq t \leq \gamma - 2$; this proves the assertion made in (b). A simple calculation shows that the widest gap be-tween any two intervals $H_k(j_k)$ is bounded by 2^{-k-1}, thus establishing (c).
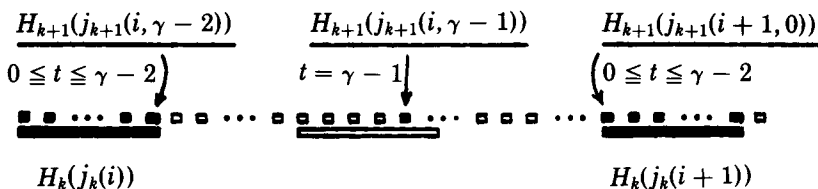

FIGURE 2

It now follows that for fixed k, but any integer $r > k$, $H_k(j_k) \supset H_r(j_r)$ whenever $E_k(i) \supset E_r(i')$, and conversely, where $j_k = j_k(i)$ and $j_r = j_r(i')$; this is easily verified from (4.14) and the fact that (4.5), together with the analogous property for the $H_k(j_k)$, are true for each k. Thus, to every nonempty infinite intersection of intervals $E_k(i)$ there corresponds a non-empty infinite intersection of intervals $H_k(j_k)$, and conversely.

Denote by $\mathscr{E}$ the set of all points belonging to infinitely many of the intervals $E_k(i)$:

$$(4.15) \qquad \mathscr{E} = \left\{ x \mid x \in \bigcap_{\nu \in \Gamma} E_{k_\nu}(i_\nu) \right\},$$

$\{k_\nu\}$ and $\{i_\nu\}$ being arbitrary infinite sequences. Consider any point, x, of $\mathscr{E}$; this point is the intersection of infinitely many $E_{k_\nu}(i_\nu)$, as follows from the fact that the lengths of these intervals diminish to zero with in-increasing k. The corresponding intervals $H_{k_\nu}(j_{k_\nu})$ also intersect in a single point y. For $x \in \mathscr{E}$, we now define $\psi(x) = y$; the range of $\psi(x)$, for $x \in \mathscr{E}$, obviously is the set

$$\mathscr{D} = \left\{ y \mid y \in \bigcap_{\nu \in \Gamma} H_{k_\nu}(j_{k_\nu}) \right\}.$$

It is readily verified that if x and x' belong to $\mathscr{E}$, and $x < x'$, then $\psi(x) < \psi(x')$. It follows that the mapping $\psi \colon \mathscr{E} \to \mathscr{D}$ has a unique extension, again denoted by ψ, on the positive real line, $\mathscr{R}^+$. This extension is monotonic increasing and has the additional property

$$(4.16) \qquad \psi(x) \in H_k(j_k)$$

whenever $x \in E_k(i)$, j_k being determined according to Lemma (i).

This completes the construction of the function $\psi(x)$. It only remains to show that it belongs to class $\mathrm{Lip}[\ln 2 / \ln(2n + 2)]$: we shall merely show that this true for $x \in \mathscr{E}$. The argument below, however, easily extends to any interval. We need

LEMMA (ii). *For any two points, $x \neq x'$, in $\mathscr{E}$, there exists a value of k for which the two points are separated by at least one interval $E_k(i)$, or exactly one gap, but by not more than $\gamma - 1$ intervals.*

Proof. For $k = 1$, there are at most $\gamma - 1$ intervals $E_1(i)$ between any two points in $\mathscr{E}$, because $\mathscr{E}$ is covered by $\gamma + 1$ of them (except, of course, for the above-mentioned gaps). If two distinct points are never separated by exactly one gap, or at most one interval, then they are separated by N intervals $E_k(i)$ for some k. We recall that each $E_k(i)$ contains $\gamma - 1$ intervals $E_{k+1}(i')$, and $\gamma - 2$ gaps; each gap between the $E_k(i)$ contains one interval $E_{k+1}(i')$ and two gaps (see Figure 1); the result follows from the fact that the lengths of these diminish to zero with increasing k.

For given points, $x \neq x'$, let now k be such that there are N intervals $E_k(i)$ and M gaps between them; according to Lemma (ii), $N = 1$ if $M = 0$, and $M = 1$ if $N = 0$, $0 \leq N \leq \gamma - 1$ and $0 \leq M \leq \gamma$.

Clearly,

$$(4.17) \qquad |x - x'| \geq N\delta_k + M(\gamma - 1)^{-1}\gamma^{-k} > 2^{\gamma - k - 1};$$

at the same time, $\psi(x)$ and $\psi(x')$ must be contained in a span of at most $N + 2$ intervals $H_k(j_k)$ and $M + 2$ gaps. Thus,

$$|\psi(x) - \psi(x')| \leq (N + 2)\epsilon_k + (M + 2)\alpha_k,$$

where α_k denotes the maximum gap between the $H_k(j_k)$; we have noted in the proof of Lemma (i) that $\alpha_k \leq 2^{-k-1}$. Owing to (4.9) and the fact that $N \leq \gamma - 1$,

$$(N + 2)\epsilon_k < (\gamma - 1)(\gamma + 2)\gamma^{-\beta_k + 1} < 2^{-k-1}.$$

Consequently,

$$(4.18) \qquad |\psi(x) - \psi(x')| < 2^{-k-1} + (\gamma + 2)2^{-k-1} = (\gamma + 3)2^{-k-1}.$$

To obtain constants c and α for which

$$|\psi(x) - \psi(x')| \leq c|x - x'|^\alpha,$$

we compare (4.17) with (4.18) and set

$$(\gamma + 3)2^{-k-1} \leq c2^\alpha \gamma^{-(k+1)\alpha}.$$

For $c = 2^{-\alpha}(\gamma + 3)$, this inequality is valid if $\alpha = \ln 2 / \ln \gamma$. Taking $\gamma = 2n + 2$, as we may, we obtain the desired result. Clearly, better estimates for c are possible by improving the estimates in (4.17) and those leading to (4.18).

5. **Proof of Lemma 1.** For each k, we construct now $2n + 1$ finite families, S_k^q, of pairwise disjoint sets, as demanded in Lemma 1; the members of each family will, in fact, be cubes obtained as the cartesian products of the intervals below.

Let $\delta > 0$ be fixed; select $k = k_0$ such that $\epsilon = (\gamma - 1)^{-1}\gamma^{-k} \leq \delta$; for each q and k, define the intervals

$$(5.1) \qquad E_k^q(i) = [e_k(i) - \epsilon q, e_k(i) + \delta_k - \epsilon q],$$

obtained from the $E_k(i)$ through the indicated translation to the left. Obviously, $E_k^0(i) = E_k(i)$ for each k, and being only a translation of the $E_k(i)$, the intervals (5.1) have the same intersection properties for increasing k as the latter for each value of q.

For each q, let i_q denote the restriction of i to the domain

$$(5.2) \qquad \Lambda_k^q = \{(\gamma^k - 1)\epsilon q, \cdots, \gamma^k + (\gamma^k - 1)\epsilon q\};$$

for the remainder of this section we shall consider the intervals (5.1) only for the values $i = i_q$, and for $k \geqq k_0$.

For $i \in \Lambda_k^q$, the $E_k^q(i)$ have the covering property stated in

LEMMA (iii). *For fixed k, each point of $\mathscr{S}$ is contained in at least $2n$ intervals $E_k^q(i)$, $0 \leqq q \leqq 2n$.*

Proof. Let k be fixed. We observe that the points 0 and 1 are contained, respectively, in the intervals corresponding to the values $i_q = (\gamma^k - 1)\epsilon q$ and $i_q = \gamma^k + (\gamma^k - 1)\epsilon q$. Hence, it suffices to show that no two gaps from those separating the intervals $E_k^q(i)$ intersect as q and i_q vary in their respective domains.

The shift ϵq is an integral multiple of the gap-width between the $E_k^q(i)$: so is the length of the intervals. Consequently, two gaps will either coincide or be mutually exclusive (except, possibly, for their end points). They will overlap if and only if $i_q\gamma^{-k} - \epsilon q = i_q'\gamma^{-k} - \epsilon q'$; unless $q = q'$ and $i = i'$ equality is impossible, owing to the inequality $q - q' < \gamma - 1$.
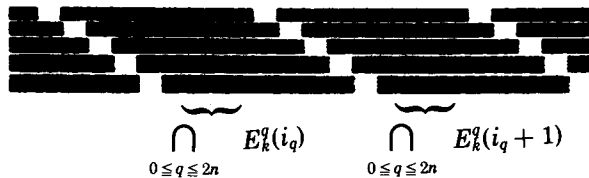

$$\bigcap_{0 \leqq q \leqq 2n} E_k^q(i_q) \qquad \bigcap_{0 \leqq q \leqq 2n} E_k^q(i_q + 1)$$

FIGURE 3

LEMMA (iv). *Let k be fixed. For each set of admitted values, $i_0, \cdots, i_{2n}$, given by*

$$(5.3) \qquad\qquad i_q = i_0 + (1 + \epsilon(\gamma^k - 1))q,$$

the intersection of corresponding intervals is not empty:

$$(5.4) \qquad\qquad \bigcap_{0 \leqq q \leqq 2n} E_k^q(i_q) \neq \emptyset$$

for all values $i_0 \in \Lambda_k^0$ and i_q given by (5.3).

Proof. According to the foregoing, $E_k^{q+1}(i_{q+1})$ will contain a gap separating the $E_k^q(i_q)$ if and only if the initial point of the former coincides with the terminal point of the latter (see Figure 3). Equating these points, we find that

$$i_{q+1} = i_q + 1 + \epsilon(\gamma^k - 1),$$

relation (5.4) being obtained through an iteration. To verify (5.4), we have only to show that $e_k(i_{2n}) + \delta_k - n\epsilon - e_k(i_0) > 0$; we omit this simple calculation.

For each q and k, let $i_{1q}, \cdots, i_{nq}$ be arbitrary values in Λ_k^q; let the corresponding intervals, $E_k^q(i_{pq})$, be laid on the pth coordinate axis of a rec-

tangular coordinate system in $\mathscr{R}^n$. The cartesian products

$$(5.5) \qquad S_k^q(\mathbf{i}_q) = \prod_{1 \leq p \leq n} E_k^q(i_{pq}),$$

in which $\mathbf{i}_q$ denotes the vector $\mathbf{i}_q = (i_{1q}, \cdots, i_{nq})$, define cubes in $\mathscr{R}^n$ with diameters $\sqrt{n} \cdot \delta_k$, which, for fixed q and k, are separated by gaps of width $(\gamma - 1)^{-1}\gamma^{-k}$; except for these, the cubes cover $\mathscr{E}^n$.

For each q and k, S_k^q will denote the set of all cubes defined by (5.5):

$$(5.6) \qquad S_k^q = \left\{ S_k^q(\mathbf{i}_q) \mid i_{pq} \in \Lambda_k^q \right\}.$$

We are now ready to complete the proof of Lemma 1. To establish the asserted covering property, we shall show that every point of $\mathscr{E}^n$ is contained in at least $m + 1$ cubes (5.5) as q varies in the domain $0 \leq q \leq n + m$, k being held fixed.

According to the properties of the intervals $E_k^q(i_{pq})$ already established, each point, $x_p \in \mathscr{E}_p$, appears for at least $n + m$ of them. This being true for each p, it follows that there are at most n values of q for which any point, $x \in \mathscr{E}^n$, is excluded from the cubes $S_k^q(\mathbf{i}_q)$; this point belongs, therefore, to at least $m + 1$ of them (see Figure 4). That condition (1.5) is satisfied follows from Lemma (iv); since the intersections (5.4) are nonempty, neither are their cartesian products.

Clearly, $\mathscr{E}^n$ can be expressed as the union of point sets generated from infinite intersections of cubes $S_k^q(\mathbf{i}_q)$ for each q.

6. Proof of Lemma 2 and Corollary 1. Throughout this section, q will vary in the domain $0 \leq q \leq 2n$. We first show that the functions $g^q(\mathbf{x})$, as defined by (1.6), satisfy Corollary 1 if they meet condition (1.7).

Let $(\alpha_0, \cdots, \alpha_{2n})$ and $(\beta_0, \cdots, \beta_{2n})$ denote any permutations of the tuple $(0, 1, \cdots, 2n)$; let k be fixed. According to Lemma 1, each point of $\mathscr{E}^n$ is contained in at least $n + 1$ cubes $S_k^q(\mathbf{i}_q)$. Let the points $\mathbf{x} \neq \mathbf{x}'$ belong to cubes $S_k^{\alpha q}(\mathbf{i}_q)$ and $S_k^{\beta q}(\mathbf{i}_q')$, respectively, for $0 \leq q \leq n$; on account of (1.7), the numbers $g^{\alpha q}(\mathbf{x})$ and $g^{\beta q}(\mathbf{x}')$ are all pairwise distinct. Thus, regardless of the values of the remaining n pairs of images corresponding to $n + 1 \leq q \leq 2n$, Corollary 1 is necessarily satisfied.

To prove (1.7), we begin with

LEMMA (v). *Let* $0 < \lambda < 1/n$, *then*

$$\sup_{\mathscr{E}^n} g^q(\mathbf{x}) < \inf_{\mathscr{E}^n} g^{q+1}(\mathbf{x})$$

for all admitted values of q.

Proof. According to (4.16), $\psi(x) \in H_k(j_k)$ whenever $x \in E_k(i)$; from (4.14) we derive the fact that $j_k(0) = 0$ and $j_k(\gamma^k) = \gamma^\beta$ for all k. Applied

to (4.11), this shows that $\psi(0) = 0$ and $\psi(1) = 1$; because ψ is monotonic, we have $\psi\mathscr{S} = \mathscr{S}$; it is easily verified that $\sup_{\mathscr{S}} \psi(x + \epsilon q) < 3/2$; (6.1) is now obtained through a simple estimate.

It follows that (1.7) is valid at least for $q \neq r$; hence, it remains only to show that this relation is also valid when $q = r$. We point out that in our present notation, i in (1.7) is replaced by i_q.

Let us consider again the intervals $E_k^q(i)$ for $i \in \Gamma'$; each $E_k^q(i)$ is related to $E_k(i)$, in a one-to-one manner, through the equation

$$(6.1) \qquad\qquad x^q = x - \epsilon q,$$

$x^q \in E_k^q(i)$, $x \in E_k(i)$. It follows at once that

$$(6.2) \qquad\qquad \psi(x + \epsilon q) \in H_k(j_k)$$

for $x \in E_k^q(i)$. For each p, q and k, define the intervals $H_k^{pq}(j_k)$ by

$$(6.3) \qquad H_k^{pq}(j_k) = [h_k(j_k) + q/n\lambda^p, h_k(j_k) + \epsilon_k + q/n\lambda^p]$$

(see 4.10): Obviously,

$$(6.4) \qquad\qquad \psi(x + \epsilon q) + q/n\lambda^p \in H_k^{pq}(j_k)$$

whenever $x \in E_k^q(i)$. It is also clear that if we replace x by x_p in (6.4) and multiply each of these functions by λ^p, then a summation over p will yield the functions $g^q(\mathbf{x})$, defined for all $x_p \geq 0$.

For convenience of notation, we replace now j_k by j. For each q, let j_q denote arbitrary admitted values of j corresponding to i_q, where i_q is restricted to the domain Λ_k^q (see 5.4); let $\mathbf{j}_q = (j_{1q}, \cdots, j_{nq})$, where the indices j_{pq} are admitted values of j_q. Set

$$h_k^q(\mathbf{j}_q) = \sum_{1 \leq p \leq n} \lambda^p h_k(j_{pq}) + q/n\lambda^p,$$
$$(6.5)$$
$$\epsilon_k^q = \epsilon_k \sum_{1 \leq p \leq n} \lambda^p,$$

and consider the intervals

$$(6.6) \qquad\qquad H_k^q(\mathbf{j}_q) = [h_k^q(\mathbf{j}_q), h_k^q(\mathbf{j}_q) + \epsilon_k^q].$$

The vectors $\mathbf{j}_q$ being uniquely determined by the vectors $\mathbf{i}_q$, and their relation being one-to-one for each q, it follows that, for fixed q and k, to each cube, $S_k^q(\mathbf{i}_q)$, there corresponds one, and only one, interval $H_k^q(\mathbf{j}_q)$ under the mapping g^q; evidently,

$$(6.7) \qquad\qquad H_k^q(\mathbf{j}_q) = g^q S_k^q(\mathbf{i}_q).$$

We are now ready to prove (1.7) for the case $r = q$. The end points of

the intervals (6.6) do not coincide, because λ satisfies no polynomial equation of degree less than n with integral coefficients. For fixed k, consider then such an inequality as

$$h_k^q(\mathbf{j}_q) > h_k^q(\mathbf{j}_q');$$

our aim is to prove that

$$h_k^q(\mathbf{j}_q) > h_k^q(\mathbf{j}_q') + \epsilon_k^q.$$

But, according to (4.7),

$$h_k^q(\mathbf{j}_q) - h_k^q(\mathbf{j}_q') = \gamma^{-\beta_k} \cdot \sum_{1 \leq p \leq n} \lambda^p (j_{pq} - j_{pq}') > \gamma^{-\beta_{k+1}+1}$$

$$> \gamma^{-\beta_{k+1}+1} \sum_{1 \leq p \leq n} \lambda^p > \epsilon_k^q,$$

the last inequality owing to (4.9). This completes the proof of Theorem 1.

BIBLIOGRAPHY

1. V. I. Arnol'd, *On functions of three variables*, Dokl. Akad. Nauk SSSR **114** (1957), 953-956.
2. _____, *The representation of functions of several variables*, Mat. Prosvešč **3** (1958), 41-61.
3. D. Hilbert, *Mathematical problems*, Bull. Amer. Math. Soc. **8** (1902), 461-462.
4. _____, *Über die Gleichung neunten Grades*, Math. Ann. **97** (1927), 243-250.
5. 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), 369-373.
6. G. G. Lorentz, *Metric entropy, widths, and superpositions of functions*, Amer. Math. Monthly **69** (1962), 469-485.
7. T. Schneider, *Einführung in die transzendenten Zahlen*, Springer, Berlin, 1957.
8. D. A. Sprecher, *A representation theorem for continuous functions of several variables*, Proc. Amer. Math. Soc. **16** (1965), 200-203.
9. V. M. Tihomirov, *The works of A. N. Kolmogorov on ϵ-entropy of function classes and superpositions of functions*, Uspehi Mat. Nauk **18** (1963), no. 5 (113), 55-92.
10. A. G. Vituškin, *On Hilbert's thirteenth problem*, Dokl. Akad. Nauk SSSR **95** (1954), 701-704.
11. D. A. Sprecher, *On the structure of representations of continuous functions of several variables as finite sums of continuous functions of one variable*, (to appear).

SYRACUSE UNIVERSITY,
SYRACUSE, NEW YORK