# Chapter 5. — Dimension of the space of linear superpositions

Objekttyp: **Chapter**

Zeitschrift: **L'Enseignement Mathématique**

Band (Jahr): **23 (1977)**

Heft 1-2: **L'ENSEIGNEMENT MATHÉMATIQUE**

PDF erstellt am: **26.05.2024**

## Chapter 5. — Dimension of the space of linear superpositions

In this chapter we present a calculation of the functional dimension of the space of functions representable by means of linear superpositions and prove that a representation of analytic functions by means superpositions of smooth functions can not be stable.

### § 1. (ε, δ)-*entropy and the "dimension" of function spaces*

Let $G_n$ be a closed region of $n$-dimensional euclidean space, and $C(G_n)$ the space of all functions continuous in $G_n$. Two functions $f_1(x), f_2(x) \in C(G_n)$ are called $(\varepsilon, \delta)$-distinguishable if there exists an $n$-dimensional closed sphere $S_\delta \subset G_n$ of radius $\delta$ such that

$$\min_{x \in S_\delta} |f_1(x) - f_2(x)| \geqslant \varepsilon .$$

Let $F \subset C(G_n)$ be a set of continuous functions. A subset $K \subset F$ is called $(\varepsilon, \delta)$-distinguishable if any two of its elements are $(\varepsilon, \delta)$-distinguishable. We denote by $N_{\varepsilon,\delta}(F)$ the maximum number of elements in an $(\varepsilon, \delta)$-distinguishable subset of $F$.

*Definition* 5.1.1. The number $H_{\varepsilon,\delta}(F) = \log_2 N_{\varepsilon,\delta}(F)$, by analogy with the definition of ε-entropy, is called the $(\varepsilon, \delta)$-entropy of $F$.

Let $f_0 \in F$. We denote by $F_{\lambda\varepsilon}(f_0)$ the set of functions $f \in F$ such that $|f(x) - f_0(x)| \leqslant \lambda\varepsilon$. It follows immediately from the definition that the expression $\overline{\lim_{\delta\to 0}}\ \overline{\lim_{\varepsilon\to 0}} - \dfrac{\log_2 H_{\varepsilon,\delta}(F_{\lambda\varepsilon}(f_0))}{\log_2\delta}$ as a function of $\lambda$ does not decrease as $\lambda \to \infty$.

*Definition* 5.1.2. The number

$$r(F, f_0) = \lim_{\lambda\to\infty} \overline{\lim_{\delta\to 0}}\ \overline{\lim_{\varepsilon\to 0}} - \frac{\log_2 H_{\varepsilon,\delta}(F_{\lambda\varepsilon}(f_0))}{\log_2\delta}$$

is called the functional "dimension" of $F$ at $f_0$. The number $r(F) = \sup\,(F, f_0)$ is called the functional "dimension" of $F$.

The functional "dimension" $r(F)$ of a set of functions $F \subset C(G_n)$ has the following properties.

5.1.1. Let $\Phi \subset F$ be a set of functions. Then $r(\Phi) \leqslant r(F)$. Moreover, if $\Phi$ is everywhere dense in $F$ in the uniform metric, then $r(\Phi) = r(F)$.

*Proof.* The first part of the assertion follows immediately from the definition. For a proof of the second part it is sufficient to show that $r(\Phi, \varphi_0) \geqslant r(F, \varphi_0)$ for any element $\varphi_0 \in \Phi$. Suppose that the functions $f_1, \ldots, f_N$ from a $(2\varepsilon, \delta)$-distinguishable subset of $F_{\lambda\varepsilon}(\varphi_0)$. Since $\Phi$ is everywhere dense in $F$, there exist functions $\varphi_1, \ldots, \varphi_N \in \Phi$ such that $\max\limits_{x \in G_n} |f_i(x) - \varphi_i(x)| \leqslant \min\left(\frac{\varepsilon}{2}, \lambda\varepsilon\right) (i = 1, 2, \ldots, N)$. These functions form an $(\varepsilon,\delta)$-distinguishable subset of $F_{2\lambda\varepsilon}(\varphi_0)$. Consequently $N_{\varepsilon,\delta}(\Phi_{2\lambda\varepsilon}(\varphi_0)) \geqslant N_{2\varepsilon,\delta}(F_{\lambda\varepsilon}(\varphi_0))$. Hence $r(\Phi, \varphi_0) \geqslant r(F, \varphi_0)$.

5.1.2. For any set $F \subset C(G_n)$ we have $r(F) \leqslant n$.

*Proof.* Suppose that $f_0 \in F$ and $f_1, f_2, \ldots, f_p$ is a maximal set (with respect to $p$) of pairwise $(\varepsilon, \delta)$-distinguishable functions of $F_{\lambda\varepsilon}(f_0)$. Let $\sigma_1, \sigma_2, \ldots, \sigma_q$ be a maximal set (with respect to $q$) of spheres of radius $\delta/3$ in $G_n$, such that no two of them have common interior points. Then any pair of functions $f_i(x)$ and $f_j(x)$ of the given set satisfies on at least one of the spheres $\sigma_l$ the inequality $\min\limits_{x \in \sigma_l} |f_i(x) - f_j(x)| \geqslant \varepsilon$. For the functions $f_i(x)$ and $f_j(x)$ satisfy on some sphere $S_\delta \subset G_n$ the inequality $\min\limits_{x \in S_\delta} |f_i(x) - f_j(x)| \geqslant \varepsilon$. Since $q$ is maximal, it follows that one of the spheres $\sigma_l \subset S_\delta$. Consequently on this sphere the inequality we need is satisfied. We denote by $a_l$ the centre of the sphere $\sigma_l$ $(l = 1, 2, \ldots, q)$. Every set of functions $f_{i_1}, f_{i_2}, \ldots, f_{i_r}$ each pair of which has values differing by not less than $\varepsilon$ at one and the same point consists of a number $r \leqslant 2\lambda + 1$ of functions. (All functions are taken from the set indicated above.) Since every pair of functions $f_i(x)$ and $f_j(x)$ has values differing by not less than $\varepsilon$ at one of the points $a_l$ at least, we have $p \leqslant 2\lambda + 1$. But since the spheres $\{\sigma_i\}$ do not intersect, $q \leqslant C/\delta^n$, where $C$ is a constant depending only on $n$. Consequently,

$$r(F, f_0) \leqslant \lim_{\lambda \to \infty} \lim_{\delta \to 0} \lim_{\varepsilon \to 0} - \frac{\log_2 \log_2 (2\lambda + 1)^{\frac{C}{\delta^n}}}{\log_2 \delta} = n.$$

5.1.3. If $F$ is everywhere dense (in the uniform metric) in the space $C(G_n)$, then $r(F) = n$. In particular $r(C(G_n)) = n$.

*Proof.* By 5.1.1 and 5.1.2 it is sufficient to show that $r(C(G_n)) \geqslant n$. We denote by $C_\varepsilon(G_n)$ the set of all $f(x) \in C(G_n)$ for which $\max\limits_{x \in G_n} |f(x)| \leqslant \varepsilon$. Let $\theta > 0$ be a constant such that for any $\delta > 0$ we can find $H = [\theta/\delta^n]$ closed and pairwise non-intersecting spheres $\sigma_1, \sigma_2, \ldots, \sigma_H$ of radius $\delta$ in $G_n$. For any system of numbers $\{\alpha_i\}$ ($\alpha_i = \pm 1$, $i = 1, 2, \ldots, H$) we construct a function $f_{\{\alpha_i\}}(x) \in C_\varepsilon(G_n)$ such that $f_{\{\alpha_i\}}(x) = a_i\varepsilon$ for $x \in \sigma_i$ ($i = 1, 2, \ldots, H$). These functions are obviously pairwise $(\varepsilon,\delta)$-distinguishable. The number of functions $f_{\{\alpha_i\}}(x)$ for all possible sets $\{\alpha_i\}$ is equal to $2^H$. Consequently $H_{\varepsilon,\delta}(C_\varepsilon(G_n)) \geqslant H = [\theta/\delta^n]$. Hence $r(C(G)) \geqslant n$.

COROLLARY 5.1.1. *The space of all polynomials in* $n$ *variables has functional "dimension"* $n$.

In the same way, the following properties are easily proved.

5.1.4. Let $G_n^1$ and $G_n^2$ be two non-intersecting closed regions in $n$- dimensional space, and $F(G_n^1 \cup G_n^2)$ a space of functions, defined and continuous on $G_n^1 \cup G_n^2$. Denote by $F(G_n^1)$ the space of all functions $\varphi(x)$, defined on the set $G_n^1$, for which there exists a function $\Phi(x) \in F(G_n^1 \cup G_n^2)$ such that $\varphi(x) \equiv \Phi(x)$ for $x \in G_n^1$. The space $F(G_n^2)$ is defined similarly. Then

$$r(F(G_n^1 \cup G_n^2)) = \max \{ r(F(G_n^1)) ;\ r(F(G_n^2)) \}.$$

5.1.5. If $F$ is a linear space, then $r(F) = r(F, f_0)$ for any function $f_0 \in F$. If $F$ is a finite-dimensional linear space, then $r(F) = 0$.

5.1.6. Let $F$ be a linear metric space with metric $\rho(\varphi, \psi)$ between a pair of functions $\varphi, \psi \in F$. We denote by $F(\rho_0)$ the set of all those functions $\varphi \in F$ for which $\rho(\varphi, 0) \leqslant \rho_0$. Then $r(F) = r(F(\rho_0))$.

COROLLARY 5.1.2. *The set of all polynomials in* $n$ *variables whose partial derivatives of order* $p$, *for any* $p = 1, 2, \ldots,$ *are bounded by a constant* $0 < K_p < \infty$ *has functional "dimension"* $n$.

5.1.7. Let $F$ be a complete linear metric space and $F = \bigcup\limits_{i=1}^{\infty} F_i$, where $\{F_i\}$ are sets of continuous functions. Then $r(F) = \max\limits_i r(F_i)$.

We now write down the main result on the functional "dimension" of a set of linear superpositions.

5.1.8. Let $q_i = q_i(x_1, x_2, \ldots, x_n)$ be continuously differentiable functions of $n$ variables, and $p_i = p_i(x_1, x_2, \ldots, x_n)$ continuous functions of $n$ variables ($i = 1, 2, \ldots, N$). We denote by $F(G_n, \{p_i\}, \{q_i\})$ the set of superpositions of the form $\sum_{i=1}^{N} p_i(x_1, x_2, ..., x_n) f_i(q_i(x_1, x_2, ..., x_n))$, where $(x_1, x_2, ..., x_n) \in G_n$, and $\{f_i(t)\}$ are arbitrary continuous functions of one variable. Then in any region $D_n$ there exists a closed subregion $G_n \subset D_n$ such that

$$r(F(G_n, \{p_i\}, \{q_i\})) \leqslant 1.$$

For ease of presentation we limit the proof to the case $n = 2$ (§ 3). It is interesting to compare the result 5.1.8 with the following proposition.

5.1.9. Let $\alpha_i(x_1, x_2, ..., x_n) = \sum_{j=1}^{n} \alpha_{ij}(x_j) \quad (i = 1, 2, ..., 2n+1)$

be the continuous functions involved in Kolmogorov's formula (I). We denote by $\psi(G_n, \alpha_i)$ the space of all functions of the form $\psi(\alpha_i(x_1, x_2, ..., x_n))$, where $\psi(t)$ is an arbitrary continuous function of one variable and $(x_1, x_2, ..., x_n) \in G_n$. Then for any $i$ and every region $G_n$, $r(\psi(G_n, \alpha_i)) = n$ (see 5.1.7).

Let $p_i(x_1, x_2, ..., x_n)$ be fixed continuous functions of $n$ variables, $q_{1,i}(x_1, x_2, ..., x_n)$, $q_{2,i}(x_1, x_2, ..., x_n)$, ..., $q_{k,i}(x_1, x_2, ..., x_n)$ fixed continuously differentiable functions of $n$ variables, and $f_i(t_1, t_2, ..., t_k)$ arbitrary continuous functions of $k$ variables, $k < n$ $(i = 1, 2, ..., N)$. One would expect that the set of superpositions of the form (V) (see Chapter I) has functional "dimension" not greater than $k$. However, in this direction, only the following partial result has so far been proved.

5.1.10. Denote by $F(\lambda, G_n, \{p_i\}, \{q_{1,i}\}, ..., \{q_{k,i}\})$ the set of all those continuous functions $\varphi(x_1, x_2, ..., x_n)$ for which there exist continuous functions $\{f_i(t_1, t_2, ..., t_k)\}$ such that in $G_n$.

$$\varphi(x_1, x_2, ..., x_n) = \sum_{i=1}^{N} p_i(x_1, x_2, ..., x_n) f_i(q_{1,i}(x_1, x_2, ..., x_n), ..., q_{k,i}(x_1, x_2, ..., x_n))$$

and

$$\max_i \sup_{(t_1, t_2, ..., t_k)} |f_i(t_1, t_2, ..., t_k)| \leqslant \lambda \sup_{(x_1, x_2, ..., x_n) \in G_n} |\varphi(x_1, x_2, ..., x_n)|$$

Then, for any $\lambda < \infty$, in any region $D_n$ there exists a closed subregion $G_n \subset D_n$ such that

$$r(F(\lambda, G_n, \{p_i\}, \{q_{1,i}\}, ..., \{q_{k,i}\}), 0) \leqslant k.$$

From the last result and Banach's open mapping theorem there follows

COROLLARY 5.1.3. *For any continuous functions* $p_i$ *and continuously differentiable functions* $q_{1,i}, q_{2,i}, \ldots, q_{k,i}$, $k < n$ $(i = 1, 2, \ldots, N)$ *and every region* $G_n$ *there exists a continuous function that is not equal in* $G_n$ *to any superposition of the form* (V).

## § 2. (ε, δ)-*entropy of the set of linear superpositions*

We denote by $S(\delta, z)$ the disc of radius $\delta$ with centre at $z$. Let $p(z) = p(x, y)$ and $q(z) = q(x, y)$ be functions defined in a closed region $G$ of the $x, y$-plane and having the properties:

a) $p(x, y), \dfrac{\partial q(x, y)}{\partial x}, \dfrac{\partial q(x, y)}{\partial y}$ are continuous in $G$ and have modulus of continuity $\omega(\delta)$,

b) the inequalities $0 < \gamma \leqslant |\operatorname{grad}[q(r)]| \leqslant \dfrac{1}{\gamma}$ and $|p(z)| \leqslant \dfrac{1}{\gamma}$, where $\gamma$ is some constant, are satisfied everywhere in $G$.

LEMMA 5.2.1. *Let* $S(\delta, z) \subset G$ *and let* $\mu_q(t)$ *be the function equal to* $2\sqrt{\delta^2 - (t - q(z))^2 |\operatorname{grad}[q(z)]|^{-2}}$ *on*

$$q(z) - \delta |\operatorname{grad}[q(z)]| \leqslant t \leqslant q(z) + \delta |\operatorname{grad}[q(z)]|$$

*and equal to zero elsewhere. Then*

$$\int_{-\infty}^{\infty} |\mu_q(t) - h_1(e(q, t) \cap S(\delta, z))| \, dt \leqslant c_1(\gamma)\, \omega(\delta)\, \delta^2,$$

*where* $c_1(\gamma)$ *is a constant depending only on* $\gamma$.

*Proof.* Let $[a, b] \subset e(q, t) \cap S(\delta, z)$ be the segment of the level curve $e(q, t)$, endpoints $a$ and $b$, lying on the boundary of $S(\delta, z)$; $[z, a]$ and $[z, b]$ the vectors with origin at $z$ and endpoints at $a$ and $b$, respectively;

$$\alpha_1 = \gamma(\overrightarrow{[z, a]}, \operatorname{grad}[q(z)]), \quad \alpha_2 = \gamma(\overrightarrow{[z, b]}, \operatorname{grad}[q(z)]).$$

We have

$$|t - q(z)| = |q(a) - q(z)| = \left| \int_{s \in [z, a]} \frac{\partial q}{\partial s} \, ds \right|$$
$$= \delta \cos \alpha_1 |\operatorname{grad}[q(z)]| \, (1 + 0(1)\, \omega(\delta))$$

Hence

$$\delta \sin \alpha_1 = \sqrt{\delta^2 - (t - q(z) + 0(\gamma)\,\delta\omega(\delta))^2 \,|\operatorname{grad}\ [q(z)]\,|^{-2}}$$

and similarly

$$\delta \sin \alpha_2 = \sqrt{\delta^2 - (t - q(z) + 0(\gamma)\,\delta\omega(\delta))^2 \,|\operatorname{grad}\ [q(z)]\,|^{-2}}$$

By b) the size of the angle swept out by the tangent vector to the level curve $e(q, t)$ on moving along $[a, b]$ does not exceed $C_2(\gamma)\,\omega(\delta)$. Therefore

$$h_1([a, b]) = \delta(\sin\alpha_1 + \sin\alpha_2)(1 + 0(\gamma)\,\omega(\delta))$$
$$= 2\sqrt{\delta^2 - (t - q(z) + 0(\gamma)\,\delta\omega(\delta))^2 \,|\operatorname{grad}\ [q(z)]\,|^{-2}} + 0(\gamma)\,\delta\omega(\delta).$$

If $\alpha_1 \geqslant C_3(\gamma)\,\omega(\delta)$ ($C_3$ is a sufficiently large constant), then $[a, b] = e(q, t) \cap S(\delta, z)$. Consequently, for

$$|t - q(z)| \leqslant \theta = \delta \cos[C_3\,\omega(\delta)]\,|\operatorname{grad}[q(z)]| \times (1 + 0(1)\,\omega(\delta))$$

we have $h_1(e(q, t) \cap S(\delta, z)) = h_1([a, b])$. Since for every $t$ (by b))

$$h_1(e(q, t) \cap S(\delta, z)) \leqslant C_4(\gamma)\,\delta\,(1 + \omega(\delta)),$$

we have

$$\int_{-\infty}^{\infty} |h_1(e(q, t) \cap S(\delta, z)) - \mu_q(t)|\,dt =$$
$$= \int_{q(z)-\Theta}^{q(z)+\Theta} |h_1(e(q, t) \cap S(\delta, z)) - \mu_q(t)|\,dt + 0(\gamma)\,\delta^2\omega(\delta).$$

We now estimate

$$\int_{q(z)-\Theta}^{q(z)+\Theta} |h_1(e(q, t) \cap S(\delta, z)) - \mu_q(t)|\,dt =$$
$$= \int_{q(z)-\Theta}^{q(z)+\Theta} |h_1([a, b]) - \mu_q(t)|\,dt \leqslant$$
$$\leqslant 2 \int_{q(z)-\Theta}^{q(z)+\Theta} \Big(\sqrt{\delta^2 - (t - q(z) + 0(\gamma)\,\delta\omega(\delta))^2 \,|\operatorname{grad}\ [q(z)]\,|^{-2}}$$
$$- \sqrt{\delta^2 - (t - q(z))^2 \,|\operatorname{grad}\ [q(z)]\,|^{-2}}\Big)\,dt + 0(\gamma)\,\delta^2\omega(\delta)$$
$$= 0(\gamma)\,\delta^2\omega(\delta) \int_{-1}^{1} \frac{d\tau}{\sqrt{1 - \tau^2}} + 0(\gamma)\,\delta^2\omega(\delta) = 0(\gamma)\,\delta^2\omega(\delta).$$

Here we have the mean value theorem. This proves the lemma.

LEMMA 5.2.2. *Let* $p(z), q(z)$ *satisfy conditions a) and b);* $S(\delta, z) \subset G$; *let* $f(t)$ *be an arbitrary continuous function, uniformly bounded in modulus by the constant* $m$. *Then*

$$\iint_{(u,v) \in S(\delta, z)} p(u,v) f\big(q(u,v)\big)\, du dv$$

$$= p(z) \left| \operatorname{grad} [q(z)] \right|^{-1} \int_{-\infty}^{\infty} f(t)\, \mu_q(t)\, dt + \lambda(z)\, m \delta^2 \omega(\delta),$$

*where* $|\lambda(z)| \leqslant C_5(\gamma)$.

*Proof.* Using a) and b) and Lemma 5.2.1 we have

$$\int_{S(\delta, z)} p(u,v) f\big(q(u,v)\big)\, du dv$$

$$= p(z) \iint_{(u,v) \in S(\delta, z)} f\big(q(u,v)\big)\, du dv + 0(1)\, m \delta^2 \omega(\delta)$$

$$= p(z) \int_{-\infty}^{\infty} \{ f(t) \int_{s \in e(q,t) \cap S(\delta, z)} \left| \operatorname{grad} [q(s)] \right|^{-2} ds \}\, dt + 0(1)\, m \delta^2 \omega(\delta)$$

$$= p(z) \left| \operatorname{grad} [q(z)] \right|^{-1} \int_{-\infty}^{\infty} \{ f(t) \int_{s \in e(q,t) \cap S(\delta, z)} ds \}\, dt + 0(\gamma)\, m \delta^2 \omega(\delta)$$

$$= p(z) \left| \operatorname{grad} [q(z)] \right|^{-2} \int_{-\infty}^{\infty} f(t)\, h_1\big(e(q,t) \cap S(\delta, z)\big)\, dt + 0(\gamma)\, m \delta^2 \omega(\delta)$$

$$= p(z) \left| \operatorname{grad} [q(z)] \right|^{-1} \int_{-\infty}^{\infty} f(t)\, \mu_q(t)\, dt + 0(\gamma)\, m \delta^2 \omega(\delta).$$

This proves the lemma.

LEMMA 5.2.3. *Suppose that a number* $\alpha > 0$ *and functions* $p(z)$, $q(z), f(t)$ *satisfying the conditions of Lemma 5.2.2. are given. If for every integer* $k$ *such that*

$$\min_{z \in G} q(z) \leqslant t_k = k\delta \frac{\alpha}{m} \leqslant \max_{z \in G} q(z)$$

*and any integer* $l$ *such that*

$$\min_{z \in G} \left| \operatorname{grad} [q(z)] \right| \leqslant t_l' = l \frac{\alpha}{m} \leqslant \max_{z \in G} \left| \operatorname{grad} [q(z)] \right|,$$

*the inequality*

$$\left| \int_{t_k - t_l' \delta}^{t_k + t_l' \delta} f(t) \sqrt{\delta^2 - \left( \frac{t - t_k}{t_l'} \right)^2}\, dt \right| \leqslant \alpha \delta^2$$

*is satisfied, then for every* disc $S(\delta, z) \subset G$

$$\Big| \iint_{(u,v) \in S(\delta,z)} p(u,v) f(q(u,v))\, du dv \Big| \leqslant c_6(\gamma)\left(\alpha\delta^2 + m\delta^2 \omega(\delta)\right).$$

*Proof.* Suppose that a disc $S(\delta, z) \subset G$ is given. By the condition of the lemma there are integers $k$ and $l$ such that $|q(z) - t_k| \leqslant \delta\alpha/m$ and $\big| |\operatorname{grad}[q(z)]| - t_l' \big| \leqslant \alpha/m$. From Lemma 5.2.2 we obtain

$$\Big| \int_{(u,v) \in S(\delta,z)} p(u,v) f(q(u,v))\, du dv \Big| \leqslant \frac{|p(z)|}{|\operatorname{grad}[q(z)]|} \Big| \int_{-\infty}^{\infty} f(t)\, \mu_q(t)\, dt \Big|$$

$$+ c_5(\gamma)\, m\delta^2 \omega(\delta) \leqslant \frac{2}{\gamma^2} \left| \int_{q(z) - \delta|\operatorname{grad}[q(z)]|}^{q(z) + \delta|\operatorname{grad}[q(z)]|} f(t) \sqrt{\delta^2 - \frac{(t - q(z))^2}{|\operatorname{grad}[q(z)]|^2}}\, dt \right.$$

$$\left. - \int_{t_k - t_l'\delta}^{t_k + t_l'\delta} f(t) \sqrt{\delta^2 - \left(\frac{t - t_k}{t_l'}\right)^2}\, dt \right| + \frac{2}{\gamma^2}\alpha\delta^2 + c_5(\gamma)\, m\delta^2\omega(\delta) \leqslant$$

(by the mean value theorem)

$$\leqslant \frac{2}{\gamma^2}\alpha\delta^2 + c_5(\gamma)\, m\delta^2 \omega(\delta) + \frac{2}{\gamma^2}\left( \int_{-1}^{1} \frac{\delta m d\tau}{\sqrt{1-\tau^2}} \right) \delta \frac{\alpha}{m}$$

$$+ \frac{2}{\gamma^2}\left( \int_{-1}^{1} \frac{\delta^2 m d\tau}{\sqrt{1-\tau^2}} \right) \frac{\alpha}{m} \leqslant c_6(\gamma)\left(\alpha\delta^2 + m\delta^2\omega(\delta)\right).$$

This proves the lemma.

We denote by $F_m = F_m(D; p_1, p_2, \ldots, p_N; q_1, q_2, \ldots, q_N)$ the set of superpositions of the form

$$f(x,y) = \sum_{i=1}^{N} p_i(x,y) f_i(q_i(x,y)), \text{ where } \{p_i(x,y)\}$$

and $\{q_i(x,y)\}$ are fixed functions, defined in the closed region $D$ of the $x, y$ plane and satisfying conditions a) and b) with a constant $\gamma$ not depending on $i$ and $\{f_i(t)\}$ are arbitrary continuous functions, defined on $\{[a_i, b_i]\} = \{[\min_{z \in D} q_i(z); \max_{z \in D} q_i(z)]\}$ and uniformly bounded in modulus by the constant $m$.

THEOREM 5.2.1. *There exist constants $A$ and $B$ such that if $\varepsilon > Am\omega(\delta)$ then for the $(\varepsilon, \delta)$-entropy of the set of functions $F_m$, $H_{\varepsilon,\delta}(F_m) \leqslant \frac{B}{\delta}\left(\frac{m}{\varepsilon}\right)^2$, where $A$ and $B$ depend only on $\gamma$, $N$ and $D$.*

*Proof.* We put

$$R\left(f(z), \delta\right) = \max_{S(\delta, z) \subset D} \left| \frac{1}{\pi\delta^2} \iint_{(u,v) \in S(\delta,z)} f(u,v)\, du dv \right| .$$

We denote by $\mathscr{H}_{\varepsilon,\delta}(F_m)$ the $\varepsilon$-entropy of the space $F_m$, taking as the distance between the functions $f_1(z), f_2(z) \in F_m$ the number $R\left(f_1(z) - f_2(z), \delta\right)$. The inequality $H_{2\varepsilon,\delta}(F_m) \leqslant \mathscr{H}_{\varepsilon,\delta}(F_m)$ holds owing to the fact that if two functions $f_1(z)$ and $f_2(z)$ are $(\varepsilon, \delta)$-distinguishable, then they are $\varepsilon$-distinguishable also in the sense of the metric $R\left(f_1(z) - f_2(z), \delta\right)$. We now estimate the value of $\mathscr{H}_{\varepsilon,\delta}(F_m)$. Let $k$ and $l$ be integers such that

$$\min_{z \in D} q_i(z) \leqslant t_k = k\delta \frac{\alpha}{m} \leqslant \max_{z \in D} q_i(z)$$

and

$$\min_{z \in D} \left| \operatorname{grad} \left[q_i(z)\right] \right| \leqslant t_l' = l \frac{\alpha}{m} \leqslant \max_{z \in D} \left| \operatorname{grad} \left[q_i(z)\right] \right| .$$

To compute the function

$$f_\delta(z) = \frac{1}{\pi\delta^2} \iint_{(u,v) \in S(\delta,z)} f(u,v)\, du dv ,$$

where $f(x, y) \in F_m$, $S(\delta, z) \subset D$ to within $\varepsilon$, it is sufficient by Lemma 5.2.3 to give the values of

$$v_i(t_k, t_l') = \frac{1}{\pi\delta^2} \int_{t_k - t_l'\delta}^{t_k + t_l'\delta} f_i(t) \sqrt{\delta^2 - \left(\frac{t - t_k}{t_l'}\right)^2}\, dt$$

to within $\alpha = \pi\varepsilon / \left(2 N C_B(\gamma)\right)$ and to assume that $\delta$ is small enough so that

$$\varepsilon > \frac{2NC_B(\gamma)\, m\omega(\delta)}{\pi} = A(\gamma, N)\, m\omega(\delta) .$$

Since $\left| v_i(t_k, t_l') \right| \leqslant C_1 m$, to write the numbers $v_i(t_k, t_l')$ $(i, k, l$ fixed) $\log_2 (C_1 m/\alpha)$ binary digits are sufficient. Since

$$\left| v_i(t_{k+1}, t_l') - v_i(t_k, t_l') \right| \leqslant c_8 \frac{1}{\delta^2} \left( \int_{-1}^{1} \frac{\delta m d\tau}{\sqrt{1-\tau^2}} \right) \delta \frac{\alpha}{m} = c_9(\gamma)\alpha$$

(here we again use the mean value theorem), to store the numbers $v_i(t_{k+1}, t_l') - v_i(t_k, t_l')$ to within $\alpha$, $\log_2 C_9$ binary digits are sufficient. Therefore to write the numbers $v_i(t_k, t_l')$ ($i$, $l$ fixed; $k$ any admissible number) $C_{10}(\gamma)\left[\log_2 \frac{m}{\alpha} + (b_i - a_i)\frac{m}{\delta\alpha}\right] = \mathscr{H}_{i,l}$ binary digits are sufficient. Consequently the total number of digits sufficient to store all the numbers $v_i(t_k, t_l')$ to within $\alpha$, that is, to store the functions $f_\delta(z)$ to within $\varepsilon$, is

$$\mathscr{H} = \sum_{i,l} \mathscr{H}_{i,l} \leqslant N c_{10}(\gamma)\left[\log_2 \frac{m}{\alpha} + (b_i - a_i)\frac{m}{\delta\alpha}\right]\frac{1}{\gamma}\frac{m}{\alpha} \leqslant \frac{B(\gamma, N, D)}{\delta}\left(\frac{m}{\varepsilon}\right)^2.$$

This proves the theorem.

## § 3. *Functional "dimension" of the space of linear superpositions*

Suppose that continuous functions $p_i(x, y)$ and continuously differentiable functions $q_i(x, y)$ $(i = 1, 2, \ldots, N)$ are fixed. Let $G$ be a closed region of the $x, y$ plane. We denote by $F = F(G, \{p_i\}, \{q_i\})$ the set of superpositions of the form $f(x, y) = \sum_{i=1}^{N} p_i(x, y) f_i(q_i(x, y))$, where $(x, y) \in G$ and $\{f_i(t)\}$ are arbitrary continuous functions of one variable. We are interested in the functional dimension of the set $F$.

THEOREM 5.3.1. *In every region $D$ of the $x, y$ plane there exists a closed subregion $G \subset D$ such that*

$$r(F(G, \{p_i\}, \{q_i\})) \leqslant 1.$$

*Proof.* By Theorem 4.5.1, in $D$ there exists a closed subregion $G^* \subset D$ such that the set of superpositions $F(G^*, \{p_i\}, \{q_i\})$ is closed (in the uniform metric) in $C(G^*)$, and the functions $\{q_i(x, y)\}$ satisfy the condition: for any $i$, either grad $[q_i(x, y)] \neq 0$ on $G^*$ or $q_i(x, y) \equiv$ const on $G^*$. We show that $r(F(G^*, \{p_i\}, \{q_i\})) \leqslant 1$. By Banach's open mapping theorem, there exists a constant $K$ such that for any superposition $\sum_{i=1}^{N} p_i(x, y) f_i(q_i(x, y)) = f(x, y) \in F(G^*, \{p_i\}, \{q_i\})$ there are continuous functions $\{f_i^*(t)\}$, defined on the sets $\{q_i(G^*)\}$ and satisfying the conditions

8) $$f(x, y) = \sum_{i=1}^{N} p_i(x, y) f_i^*(q_i(x, y)) \text{ for all } (x, y) \in G^*;$$

9) $$\max_i \max_{t \in q_i(G^*)} |f_i^*(t)| \geqslant K \max_{(x,y) \in G^*} |f(x, y)|.$$

Denote by $F_{\lambda\varepsilon} = F_{\lambda\varepsilon}(G^*, \{p_i\}, \{q_i\})$ the set of superpositions $f(x,y) \in F(G^*, \{p_i\}, \{q_i\})$ such that $\max_{(x,y) \in G^*} |f(x, y)| \leqslant \lambda\varepsilon$. By Theorem 5.2.1 and (8), (9), there exist constants $A$ and $B$ such that if $\omega(\delta) \leqslant (\lambda AK)^{-1}$ then $H_{\varepsilon,\delta}(F_{\lambda\varepsilon}) \leqslant B(\lambda K)^2/\delta$. Hence the functional dimension

$$r(F_i(G^*, \{p_i\}, \{q_i\})) \leqslant \lim_{\lambda\to\infty} \lim_{\delta\to 0} \lim_{\varepsilon\to 0} \frac{\log_2 \log_2 \frac{B(\lambda K)^2}{\delta}}{\log_2 \delta} = 1$$

This proves the theorem.

From Theorem 5.3.1 and the properties of functional dimension (§ 1) we have the following result, which is a stronger form of Theorem 4.6.1.

COROLLARY 5.3.1. *For any continuous functions* $\{p_i(x, y)\}$ *and continuously differentiable functions* $\{q_i(x, y)\}$ *and every region* $D$ *the set of linear superpositions* $F(D, \{p_i\}, \{q_i\})$ *is nowhere dense in any space of functions that has in every region* $G \subset D$ *functional "dimension" greater than* 1.

*Remark* 5.3.1. All the results about linear superpositions of the form $\sum_{i=1}^{N} p_i(x, y) f_i(q_i(x, y))$ remain valid if we assume that $\{f_i(t)\}$ are arbitrary bounded measurable functions.

## § 4. *Variation of superpositions of smooth functions*

Let $G_n$ be a closed region of the space of the variables $x_1, x_2, \ldots, x_n$ $(n \geqslant 2)$. A function $F(x) = F(x_1, x_2, \ldots, x_n)$ is called a superposition of order $s$ generated by the functions of $k$ $(k > 1)$ variables

$$f_{\beta_1,\beta_2\ldots,\beta_\alpha}(t_1, t_2, \ldots, t_k) \quad (\alpha = 0, 1, 2, \ldots, s;\ \beta_i = 1, 2, \ldots, k)$$

if it is defined in $G$ by relations

$$\left\{\begin{array}{ll} F = f(q_1, q_2, \ldots, q_k)\,, & \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ q_{\beta_1,\beta_2,\ldots,\beta_\alpha} = f_{\beta_1,\ldots,\beta_\alpha}(q_{\beta_1,\ldots,\beta_\alpha,1} q_{\beta_1,\ldots,\beta_\alpha,2}, \ldots, q_{\beta_1,\ldots,\beta_\alpha,k}), & \text{(VI)} \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ q_{\beta_1,\beta_2,\ldots,\beta_{s+1}} = x_{\gamma(\beta_1,\beta_2,\ldots,\beta_{s+1})}\,, & \end{array}\right.$$

where $\gamma(\beta_1, \beta_2, \ldots, \beta_{s+1})$ is a function of the indices $\beta_1, \beta_2, \ldots, \beta_{s+1}$ and takes one of the values $1, 2, \ldots, n$. As before, we assume that the functions $\{\varphi_{\beta_1,\beta_2,\ldots,\beta_\alpha}(t_1, t_2, \ldots, t_k)\}$ are defined for all values of the arguments.

A superposition of any order, generated by functions of one variable, is again a function of one variable. Therefore in this case ($k=1$) we consider superpositions of functions of one variable and the operation of addition, that is, superpositions definable in the following way.

A function $F(x) = F(x_1, x_2, \ldots, x_n)$ ($n > 1$) is called a superposition of order $s$ of the functions $f_{\beta_1,\ldots,\beta_\alpha}(t)$ ($\alpha = 0, 1, 2, \ldots, s$; $\beta_i = 1, 2$) if the following relations are satisfied:

$$\left.\begin{array}{l} F = f(q_1 + q_2)\,, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ q_{\beta_1,\beta_2,\ldots,\beta_\alpha} = f_{\beta_1,\beta_2,\ldots,\beta_\alpha}(q_{\beta_1,\beta_2,\ldots,\beta_\alpha,1} + q_{\beta_1,\beta_2,\ldots,\beta_\alpha,2}) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ q_{\beta_1,\beta_2,\ldots,\beta_{s+1}} = x_{\gamma(\beta_1,\beta_2,\ldots,\beta_{s+1})}\,, \end{array}\right\} \text{(VII)}$$

where $\gamma(\beta_1, \beta_2, \ldots, \beta_{s+1})$ takes one of the values $1, 2, \ldots, n$.

Note that we can represent as superpositions of the form (VII), for example, all rational functions of $x_1, x_2, \ldots, x_n$ since we can write any arithmetic operation by such superpositions, for example, $u \cdot v = e^{\ln u + \ln v} = f(f_1(u) + f_2(v))$.

Let $F(x_1, x_2, \ldots, x_n)$ be a superposition of order $s$ of the continuously differentiable functions $\{f_{\beta_1,\beta_2,\ldots,\beta_\alpha}(t_1, t_2, \ldots, t_k)\}$ and $\tilde{F}(x_1, x_2, \ldots, x_n)$ the superposition of the same form of the continuously differentiable functions $\{\tilde{f}_{\beta_1,\beta_2,\ldots,\beta_\alpha}(t_1, t_2, \ldots, t_k)\}$. We put

$$\varphi_{\beta_1,\beta_2,\ldots,\beta_\alpha} = \tilde{f}_{\beta_1,\ldots,\beta_\alpha} - f_{\beta_1,\ldots,\beta_\alpha} \quad (\alpha = 0, 1, 2, \ldots, s;\ \beta_i = 1, 2, \ldots, k)$$

$$\mu = \max_{\alpha,\beta_1,\ldots,\beta_\alpha} \sum_{i=1}^{k} \sup_t \left| \frac{\partial f_{\beta_1,\ldots,\beta_\alpha}(t_1, \ldots, t_k)}{\partial t_i} \right|,$$

$$\varepsilon = \max_{\alpha,\beta_1,\ldots,\beta_\alpha} \sup_t \left| \varphi_{\beta_1,\ldots,\beta_\alpha}(t_1, t_2, \ldots, t_k) \right|$$

LEMMA 5.4.1. *The inequality*

$$\sup_{x \in G} \left| \tilde{F}(x_1, x_2, \ldots, x_n) - F(x_1, x_2, \ldots, x_n) \right| \leqslant A(\mu, s)\, \varepsilon .$$

*holds, where the constant* $A(\mu, s)$ *depends only on* $\mu$ *and* $s$.

*Proof.* We proceed by induction on $s$. For definiteness suppose that $k < 1$. Having verified the statement of the lemma for $s = 1$ and having made an appropriate inductive assumption for superpositions of order $s - 1$, we have

$$\begin{aligned} &\sup_{x \in G} \left| \tilde{F}(x_1, x_2, \ldots, x_n) - F(x_1, x_2, \ldots, x_n) \right| \\ &\quad \leqslant \left| f(\tilde{q}_1, \ldots, \tilde{q}_k) - f(q_1, \ldots, q_k) \right| + \left| \varphi(\tilde{q}_1, \tilde{q}_2, \ldots, \tilde{q}_k) \right| \\ &\quad \leqslant \mu \max_{\beta_1} \sup_{x \in G} \left| \tilde{q}_{\beta_1} - q_{\beta_1} \right| + \varepsilon \leqslant \mu \cdot A(\mu, s-1)\, \varepsilon + \varepsilon = A(\mu, s)\, \varepsilon . \end{aligned}$$

(the last by the indictive assumption). This proves the lemma.

Further, let $\omega(\delta)$ be the common modulus of continuity of all the functions $\left\{ \dfrac{\partial f_{\beta_1, \ldots, \beta_\alpha}(t_1, \ldots, t_k)}{\partial t_i} \right\}$ and, in addition, put

$$\varepsilon' = \max_{\alpha, \beta_1, \ldots, \beta_\alpha} \sum_{i=1}^{k} \sup_t \left| \frac{\partial \varphi_{\beta_1, \ldots, \beta_\alpha}(t_1, \ldots, t_k)}{\partial t_i} \right|$$

LEMMA 5.4.2. *We have (for case* $k > 1$*)*

$$\begin{aligned} \tilde{F}(x_1, \ldots, x_n) - F(x_1, \ldots, x_n) &= \sum_{\alpha, \beta_1, \ldots, \beta_\alpha} p_{\beta_1, \ldots, \beta_\alpha}(x_1, x_2, \ldots, x_n) \\ &\times \varphi_{\beta_1, \ldots \beta_\alpha}\left(q_{\beta_1, \ldots, \beta_\alpha, 1}(x_1, \ldots, x_n), \ldots, q_{\beta_1, \ldots, \beta_\alpha, k}(x_1, \ldots, x_n)\right) \\ &+ R(x_1, x_2, \ldots, x_n) , \end{aligned}$$

*where*

$$\left| R(x_1, x_2, \ldots, x_n) \right| \leqslant B(\mu, s, k) \left[ \varepsilon' + \omega\left(A(\mu, s)\, \varepsilon\right) \right] \varepsilon ,$$

$$p_{\beta_1, \ldots, \beta_\alpha}(x_1, x_2, \ldots, x_n) = \prod_{i=0}^{\alpha-1} \frac{\partial f_{\beta_1, \ldots, \beta_i}}{\partial q_{\beta_1, \ldots, \beta_{i+1}}}$$

(for $\alpha = 0$ $p(x_1, x_2, \ldots, x_n) \equiv 1$),

$B(\mu, s, k)$ *is a constant depending only on* $\mu, s, k$. *For* $k = 1$ *the corresponding equation is slightly different (see Chapter I, (III)) :*

$$\tilde{F}(x_1, \dots, x_n) - F(x_1, \dots, x_n)$$
$$= \sum_{\alpha, \beta_1, \dots, \beta_\alpha} p_{\beta_1,\dots,\beta_\alpha}(x_1, x_2, \dots, x_n)\, \varphi_{\beta_1,\dots,\beta_\alpha}(q_{\beta_1,\dots,\beta_\alpha,1}(x_1, \dots, x_n)$$
$$+ q_{\beta_1,\dots,\beta_\alpha,2}(x_1, \dots, x_n)) + R(x_1, \dots, x_n)\,.$$

*Proof.* As in the preceding lemma we proceed by induction on $s$. Again for definiteness we limit ourselves to the case $k > 1$. For $s = 1$ the assertion of the lemma is easily verified. We assume that it is true for superpositions of order $s - 1$. By Lemma 5.4.1, for superpositions of order $s$ we have

$$\tilde{F}(x_1, \dots, x_n) - F(x_1, \dots, x_n) = f(\tilde{q}_1, \tilde{q}_2, \dots, \tilde{q}_k) - f(q_1, q_2, \dots, q_k)$$
$$+ \varphi(\tilde{q}_1, \tilde{q}_2, \dots, \tilde{q}_k) = \varphi(q_1, q_2, \dots, q_k) + \sum_{\beta_1=1}^{k} \frac{\partial f}{\partial q_{\beta_1}} (\tilde{q}_{\beta_1} - q_{\beta_1})$$
$$+ A(\mu, s)\, \varepsilon' \cdot \varepsilon + k \cdot A(\mu, s)\, \omega\big(A(\mu, s)\, \varepsilon\big)\, \varepsilon\,.$$

Since $\tilde{q}_{\beta_1}$ and $q_{\beta_1}$ $(\beta_1 = 1, 2, \dots, k)$ are superpositions of order $s - 1$, by the inductive hypothesis we have

$$\tilde{q}_{\beta_1} - q_{\beta_1} = \sum_{\substack{\alpha > 0 \\ \beta_2, \beta_3, \dots, \beta_\alpha}} \hat{p}_{\beta_1,\dots,\beta_\alpha}(x_1, x_2, \dots, x_n)$$
$$\times \varphi_{\beta_1,\dots\beta_\alpha}\big(q_{\beta_1,\dots,\beta_\alpha,1}(x_1, x_2, \dots, x_n), \dots, q_{\beta_1,\dots,\beta_\alpha,k}(x_1, x_2, \dots, x_n)\big)$$
$$+ \acute{R}(x_1, x_2, \dots, x_n)\,,$$

where

$$\big|\hat{R}(x_1, x_2, \dots, x_n)\big| \leqslant B(\mu, s-1, k)\big[\varepsilon' + \omega\big(A(\mu, s-1)\, \varepsilon\big)\big]\, \varepsilon\,,$$
$$\hat{p}_{\beta_1,\dots,\beta_\alpha}(x_1, \dots, x_n) = \prod_{i=1}^{\alpha-1} \frac{\partial f_{\beta_1,\beta_2,\dots,\beta_i}}{\partial q_{\beta_1,\dots,\beta_{i+1}}}$$

(for $\alpha = 1$, $\hat{p}_{\beta_1}(x_1, \dots, x_n) \equiv 1$).

When we now substative the expressions for the differences $\tilde{q}_{\beta_1} - q_{\beta_1}$ in the formula for $\tilde{F} - F$ above, we obtain the required representation of the difference of two superpositions $\tilde{F} - F$. This proves the lemma.

## § 5. *Instability of the representation of functions as superpositions of smooth functions*

Let $A$ be a set of functions of $n$ variables and $B$ a set of functions of $k$ variables ($k<n$). Suppose that a function $F(x_1, \ldots, x_n) \in A$ is in a region $G_n$ of the space $x_1, x_2, \ldots, x_n$ an $s$-fold superposition, generated by a system of functions $\{f_{\beta_1,\ldots,\beta_\alpha}(t_1, \ldots, t_k)\}$ of $B$.

We say that this superposition is $(A, B)$-stable in $G_n$ if every function $\tilde{F}(x_1, \ldots, x_n) \in A$ can be represented in $G_n$ as the $s$-fold superposition of the same form of functions $\{\tilde{f}_{\beta_1,\ldots,\beta_\alpha}(t_1, t_2, \ldots, t_k)\}$ of $B$ such that

$$\max_{\alpha;\,\beta_1,\ldots,\beta_\alpha} \sup_t \left|\tilde{f}_{\beta_1,\ldots,\beta_\alpha}(t_1, \ldots, t_k) - f_{\beta_1,\ldots,\beta_\alpha}(t_1, \ldots, t_k)\right|$$
$$\leqslant \lambda \sup_{x \in G_n} \left|\tilde{F}(x_1, \ldots, x_n) - F(x_1, \ldots, x_n)\right|,$$

where $\lambda$ is a constant not depending either on $\tilde{F}$ or on the $\{\tilde{f}_{\beta_1,\ldots,\beta_\alpha}\}$.

We denote by $C^{(1)}_{\omega(\delta)}$ the space of all continuously differentiable functions of $k$ variables whose partial derivatives have modulus of continuity $\omega(\delta)$ ($\omega(\delta) \to 0$ as $\delta \to 0$).

THEOREM 5.5.1. *Suppose that each function $F(x_1, \ldots, x_n) \in A$ is in some region $D_n$ of the space $x_1, \ldots, x_n$ a superposition of order $s$ of functions of $k$ variables $\{f_{\beta_1,\ldots,\beta_\alpha}(t_1, \ldots, t_k)\}$ belonging to $C^{(1)}_{\omega(\delta)}$ ($k<n$). If for any subregion $G_n \subset D_n$ the functional "dimension" of $A$ at $F(x_1, \ldots, x_n) \in A$ is greater than $k$, then the function $F(x_1, \ldots, x_n)$ cannot be an $(A, C^{(1)}_{\omega(\delta)})$-stable superposition in any such region $G \subset D_n$.*

*Proof.* Assume the contrary, that is, in a region $G_n \subset D_n$ the function $F(x_1, \ldots, x_n) \in A$ is an $(A, C^{(1)}_{\omega(\delta)})$-stable $s$-fold superposition of functions $\{f_{\beta_1,\ldots,\beta_\alpha}(t_1, \ldots, t_k)\}$ of $C^{(1)}_{\omega(\delta)}$. Then any function $\tilde{F}(x_1, \ldots, x_n) \in A$ can be represented as the superposition of the same form of functions $\{\tilde{f}_{\beta_1,\ldots,\beta_\alpha}(t_1, \ldots, t_k)\}$ of $C^{(1)}_{\omega(\delta)}$ such that

$$\max_{\alpha;\,\beta_1,\ldots,\beta_\alpha} \sup_t \left|\varphi_{\beta_1,\ldots,\beta_\alpha}(t_1, \ldots, t_k)\right| \leqslant \lambda \sup_{x \in G_n} \left|\tilde{F} - F\right|,$$

where $\varphi_{\beta_1,\ldots,\beta_\alpha} = \tilde{f}_{\beta_1,\ldots,\beta_\alpha} - f_{\beta_1,\ldots,\beta_\alpha}$. By Lemma 5.4.2 we have (for definiteness, $k > 1$)

$$\tilde{F} - F = \sum_{\alpha;\,\beta_1,\ldots,\beta_\alpha} p_{\beta_1,\ldots,\beta_\alpha}(x_1, \ldots, x_n)$$

$$\times\ \varphi_{\beta_1,\ldots,\beta_\alpha}(q_{\beta_1,\ldots,\beta_\alpha,1}(x_1, \ldots, x_n), \ldots, q_{\beta_1,\ldots,\beta_\alpha,k}(x_1, \ldots, x_n)) + R(x_1, \ldots, x_n),$$

where $| R(x_1, \ldots, x_n) | \leqslant \gamma(\varepsilon)\,\varepsilon$, $\gamma(\varepsilon) \to 0$ as $\varepsilon \to 0$, and

$$\varepsilon = \max_{\alpha;\,\beta_1,\ldots,\beta_\alpha} \sup_t | \varphi_{\beta_1,\ldots,\beta_\alpha}(t_1, \ldots, t_k) |$$

$$\leqslant \lambda \sup_{x \in G_n} | \tilde{F}(x_1, \ldots, x_n) - F(x_1, \ldots, x_n) | .$$

That $\gamma(\varepsilon) \to 0$ as $\varepsilon \to 0$ follows from the fact that as $\varepsilon \to 0$ the quantity

$$\varepsilon' = \max_{\alpha;\,\beta_1,\ldots,\beta_\alpha} \sum_{i=1}^{k} \sup \left| \frac{\partial \varphi_{\beta_1,\ldots,\beta_\alpha}(t_1, \ldots, t_k)}{\partial t_i} \right| \to 0,$$

provided only that the modulus of continuity of the partial derivatives of the functions $\{ \varphi_{\beta_1,\ldots,\beta_\alpha}(t_1, \ldots, t_k) \}$ is fixed. By 5.1.10 it follows that $r(A, F) \leqslant k$ in some subregion $G_n \subset D_n$. So we have obtained a contradiction to the assumption that $r(A, F) > k$ in any subregion $G_n \subset D_n$ and this proves the theorem.

## REFERENCES


[1] Hilbert, D. Mathematische Probleme. *Nachr. Akad. Wiss. Göttingen* (1900), 253-297; *Gesammelte Abhandlungen*, Bd. 3 (1935), 290-329.

[2] Ostrowski, A. Über Dirichletsche Reihen und algebraische Differentialgleichungen. *Math. Z. 8* (1920), 241-298.

[3] Hilbert, D. Über die Gleichung neunten Grades. *Math. Ann. 97* (1927), 243-250; *Gesammelte Abhandlungen*, Bd. 2 (1933), 393-400.

[4] Vitushkin, A. G. On Hilbert's thirteenth problem. *Dokl. Akad. Nauk SSSR 95* (1954), 701-704.

[5] Bieberbach, L. Bemerkung zum dreizehnten Hilbertschen Problem. *J. Reine Angew. Math. 165* (1931), 89-92.

[6] —— Einfluss von Hilberts Pariser Vortrag über „Mathematische Probleme". *Naturwissenschaften 51* (1930), 1101-1111.

[7] Kolmogorov, A. N. On the representation of continuous functions of several variables by superpositions of continuous functions of fewer variables. *Dokl. Akad. Nauk SSSR 108* (1956), 179-182. *Amer. Math. Soc. Transl.* (2) 17 (1961), 369-373.

[8] Arnol'd, V. I. On functions of three variables. *Dokl. Akad. Nauk SSSR 114* (1957), 679-681.

[9] Kolmogorov, A. N. On the representation of continuous functions of several 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), 55-59.