CRITICAL VALUES AND REPRESENTATION OF FUNCTIONS BY MEANS OF COMPOSITIONS

83 32

Y. Yomdin

Sonderforschungsbereich 40
Theoretische Mathematik
Beringstr. 4
D-5300 Bonn 1

Max-Planck-Institut
für Mathematik
Gottfried-Claren-Straße 26
D-5300 Bonn 3

MPI/SFB 83-32

CRITICAL VALUES AND REPRESENTATION OF FUNCTIONS BY MEANS OF COMPOSITIONS

Y. Yomdin*

ABSTRACT. The structure of the set of critical values of a composition of differentiable mappings is studied. On this base some explicit examples of differentiable functions, not representable by compositions of certain types, are given.

The question of a representability of functions of a given class by means of compositions of functions, belonging to some other given classes, has been studied in many publications (see e.g. [1],[2],[3], [4]). The known results on a nonrepresentability are mostly based on considerations of a "massiveness" of corresponding subsets in a suitable functional space. Therefore, showing the existence of non-representable functions, they do not give explicit examples.

In this paper we consider some special question of the above type, which roughly can be formulated as follows:

* Max-Planck-Institut für Mathematik, Gottfried-Claren-Strasse 26, D-5300 Bonn 3, Federal Republic of Germany and
Department of Mathematics, Ben-Gurion University of the Negev, Beer-Sheva 84120, Israel.

what functions of a given smoothness and a given number of variables can be represented as a function of a lower smoothness and of a smaller number of new variables, each of which, in turn, is a function of a higher smoothness of the initial variables?

We find an upper bound for the entropy dimension of the set of critical values of the function, represented in such a form. Using this bound we give the necessary condition for the representability, and some examples of non-representable functions.

Let $f : U \to R^m$ be a continuously differentiable mapping of an open domain $U \subset R^n$. For $\gamma \geq 0$ define the set of "γ - near-critical" points of f, $\Sigma(f,\gamma)$, by

$$\Sigma(f,\gamma) = \{ x \in U \,/\, \|df(x)\| \leq \gamma \},$$ and let

$\Delta(f,\gamma) = f(\Sigma(f,\gamma))$ be the corresponding set of "γ - near-critical" values of f. The set of critical points $\Sigma(f,0)$ we denote by $\Sigma(f)$ and the set of critical values $\Delta(f,0)$ by $\Delta(f)$.

For a compact domain $D \subset R^n$ let $C^k(D,m)$ denote the space of mappings $f : D \to R^m$, which can be extended to a k times continuously differentiable mapping of some open neighborhood of D. For $f \in C^k(D,m)$ let

$$M_i(f) = \max_{y \in D} \|d^i f(y)\| , \quad i = 0, \ldots, k.$$

(The Euclidean spaces R^q and the spaces of their linear and multilinear mappings are considered with the usual Euclidean norms.)

We recall the definition of the entropy dimension: for a bounded subset $A \subset R^m$ and $\varepsilon > 0$ let $M(\varepsilon, A)$ be the minimal number of balls of radius ε, covering A. Then the entropy dimension $\dim_e A$ is defined by

$$\dim_e A = \inf\{\beta \geq 0 \ / \ \exists K, \forall \varepsilon > 0, \ M(\varepsilon, A) \leq K(\tfrac{1}{\varepsilon})^\beta\}.$$

(See e.g. [3], [6] for some properties of $M(\varepsilon, A)$ and of the entropy dimension).

Now let $B^\alpha \subset R^{n_\alpha}$ be closed balls of radii ρ_α, respectively, $\alpha = 1, \ldots, s+1$, and let $f : B^1 \to B^{s+1}$ be given as a composition $f = f^s \circ f^{s-1} \circ \ldots \circ f^1$, where $f^\alpha : B^\alpha \to B^{\alpha+1}$, $f^\alpha \in C^{k_\alpha}(B^\alpha, n_{\alpha+1})$, $k_\alpha \geq 2$, $\alpha = 1, \ldots, s$.

Below we assume that the dimensions n_α do not increase: $n_1 \geq n_2 \geq \ldots \geq n_s \geq n_{s+1}$. (The problem of a factorization through lower-dimensional spaces has essentially different nature). Without loss of generality we can also assume that $k_1 > k_2 > \ldots > k_s$. Indeed, one can easily prove that if $k_2 \geq k_1$ and $n_1 \geq n_2 \geq n_3$, then all the mappings $f : B^1 \to B^3$, belonging to $C^{k_1}(B^1, n_3)$, and only these mappings, are representable as $f = f^2 \circ f^1$, $f^\alpha : B^\alpha \to B^{\alpha+1}$, $f^\alpha \in C^{k_\alpha}(B^\alpha, n_{\alpha+1})$, $\alpha = 1, 2$.

Let f be represented as a superposition of s mappings as above. We call $S = (n_1, k_1, n_2, k_2, \ldots, n_s, k_s)$ the diagram of the representation and say that f is representable with the diagram S.

For a given diagram S define $\sigma(S)$ as

$$\frac{n_1-n_2}{k_1-1} + \frac{n_2-n_3}{k_2-1} + \dots + \frac{n_{s-1}-n_s}{k_{s-1}-1} + \frac{n_s}{k_s-1} .$$

Theorem 1. If f is representable with the diagram S, then

$$\dim_e \Delta(f) \le \sigma(S) .$$

Proof. Fix some $\varepsilon > 0$, $\varepsilon \le 1$. Below we denote by K_j the constants, depending only on the set of data $Q = \{\rho_\alpha, n_\alpha, k_\alpha, M_i(f^\alpha), \alpha = 1, \dots, s+1, i = 0, \dots, k_\alpha\}$, but not on ε.

For each $\alpha = 1, \dots, s$ let $r_\alpha = \varepsilon^{\frac{1}{k_\alpha - 1}}$. Consider in each $B^\alpha \subset R^{n_\alpha}$ the points with the coordinates of the form $m \frac{r_\alpha}{\sqrt{n_\alpha}}$, $m \in Z$, and denote these points by x^α_β, $1 \le \beta \le d_\alpha$. The balls of the radius r_α centered at x^α_β, cover B^α.

Let P^α_β be the Taylor polynomial of order k_α of the mapping f^α at the point x^α_β.

For any s-tuple $(\beta_1, \dots, \beta_s)$, $1 \le \beta_\alpha \le d_\alpha$, denote by $P_{\beta_1, \dots, \beta_s}$ the polynomial mapping

$$P^s_{\beta_s} \circ P^{s-1}_{\beta_{s-1}} \circ \dots \circ P^1_{\beta_1} : R^{n_1} \to R^{n_{s+1}} .$$

We also denote by $D_{\beta_1, \dots, \beta_s}$ the set

$$\{ x \in B^1_{\beta_1} / \quad f^\alpha \circ f^{\alpha-1} \circ \dots \circ f^1(x) \in B^{\alpha+1}_{\beta_{\alpha+1}}, \alpha = 1, \dots, s-1 \}.$$

Lemma 2. For any s - tuple $(\beta_1,\ldots,\beta_s)$ and for any $x \in D_{\beta_1,\ldots,\beta_s}$,

i. $$\|f(x) - P_{\beta_1,\ldots,\beta_s}(x)\| \leq K_1\varepsilon ,$$

ii. $$\|df(x) - dP_{\beta_1,\ldots,\beta_s}(x)\| \leq K_1\varepsilon .$$

Proof. Induction on α. Denote $f^\alpha \cdot f^{\alpha-1} \cdot \ldots \cdot f^1$ by F^α, and $P^\alpha_{\beta_\alpha} \cdot \ldots \cdot P^1_{\beta_1}$ by Q^α and assume that i and ii are satisfied for $F^{\alpha-1}$, $Q^{\alpha-1}$ with the constant K_1^α.

By the choice of r_α and by the Taylor formula we have for any $y \in B^\alpha_{\beta_\alpha}$:

$$\|f^\alpha(y) - P^\alpha_{\beta_\alpha}(y)\| \leq K_2 r_\alpha^k = K_2\varepsilon^{\frac{k_\alpha}{k_\alpha - 1}} \leq K_2\varepsilon ,$$

$$\|df^\alpha(y) - dP^\alpha_{\beta_\alpha}(y)\| \leq K_3 r_\alpha^{k_\alpha - 1} = K_3\varepsilon .$$

Now denote $F^{\alpha-1}(x)$ by y_1, $Q^{\alpha-1}(x)$ by y_2. Since $x \in D_{\beta_1,\ldots,\beta_s}$, $y_1 \in B^\alpha_{\beta_\alpha}$. We have :

$$\|F^\alpha(x) - Q^\alpha(x)\| = \|f^\alpha(y_1) - P^\alpha_{\beta_\alpha}(y_2)\| \leq$$

$$\|f^\alpha(y_1) - P^\alpha_{\beta_\alpha}(y_1)\| + \|P^\alpha_{\beta_\alpha}(y_1) - P^\alpha_{\beta_\alpha}(y_2)\| \leq$$

$$K_2\varepsilon + M_1(P^\alpha_{\beta_\alpha}) \|y_1 - y_2\| \leq K_2\varepsilon + K_4 K_1^{\alpha-1}\varepsilon = K_5\varepsilon .$$

In a similar way, for the first derivatives we have:

$$\|dF^\alpha(x) - dQ^\alpha(x)\| = \|df^\alpha(y_1) \cdot dF^{\alpha-1}(x) - dP^\alpha_{\beta_\alpha}(y_2) \cdot dQ^{\alpha-1}(x)\|$$

$$\|df^\alpha(y_1) - dP^\alpha_{\beta_\alpha}(y_1)\| \; \|dF^{\alpha-1}(x)\| + \|dP^\alpha_{\beta_\alpha}(y_1)\| \; \|dF^{\alpha-1}(x) - dQ^{\alpha-1}(x)\| +$$

$$+ \|dP^\alpha_{\beta_\alpha}(y_1) - dP^\alpha_{\beta_\alpha}(y_2)\| \; \|dQ^{\alpha-1}(x)\| \leq$$

$$K_3\varepsilon K_6 + K_7 K_1^{\alpha-1}\varepsilon + M_2(P^\alpha_{\beta_\alpha}) K_1^{\alpha-1}\varepsilon K_8 = K_9\varepsilon .$$

Denoting $\max(K_5, K_9)$ by K_1^{α}, we obtain the required inequalities, with $K_1 = K_1^s$.

For any s - tuple $(\beta_1,\ldots,\beta_s)$ let

$\Sigma_{\beta_1,\ldots,\beta_s} = \Sigma(f) \cap D_{\beta_1,\ldots,\beta_s} \subset B^1$ and let

$$\Delta_{\beta_1,\ldots,\beta_s} = f(\Sigma_{\beta_1,\ldots,\beta_s}) .$$

Lemma 3. For any $(\beta_1,\ldots,\beta_s)$, $\Delta_{\beta_1,\ldots,\beta_s}$ can be covered by K_{10} balls of radius ε.

Proof. By lemma 2, ii, $\Sigma_{\beta_1,\ldots,\beta_s} \subset \Sigma(P_{\beta_1,\ldots,\beta_s}, K_1\varepsilon)$. By i, $\Delta_{\beta_1,\ldots,\beta_s}$ is thus contained in a $K_1\varepsilon$ - neighborhood of $\Delta(P_{\beta_1,\ldots,\beta_s}, K_1\varepsilon)$. Now by corollary 2.14, [6], this last set, being the set of near critical values of the polynomial mapping $P_{\beta_1,\ldots,\beta_s}$ on the ball B^1 of radius ρ_1, can be covered by N balls of radius $\rho_1 K_1 \varepsilon$, where $N = N(n_1, n_{s+1}, k_1 k_2 \ldots k_s)$ depends only on the dimensions n_α and on the degrees of differentiability k_α. Hence $\Delta_{\beta_1,\ldots,\beta_s}$ can be covered by the same number of balls of radius $(\rho_1+1)K_1\varepsilon$, or by $N[\mathfrak{m}_{s+1}(2\rho_1+2)K_1]^{n_{s+1}} = K_{10}$ balls of radius ε.

Of course, all the sets $\Delta_{\beta_1,\ldots,\beta_s}$, $1 \le \beta_\alpha \le d_\alpha$, cover $\Delta(f)$. But in fact many of these sets are empty. So to prove theorem 1 it remains to estimate the number of nonempty $\Delta_{\beta_1,\ldots,\beta_s}$, which, in turn, does not exceed the number of nonempty $D_{\beta_1,\ldots,\beta_s}$.

Lemma 4. Let $\beta_1, \ldots, \beta_{\alpha-1}$ be fixed, $\alpha = 1,\ldots,s$. Then the number of the indeces β_α , for which $D_{\beta_1,\ldots,\beta_{\alpha-1},\beta_\alpha,\beta_{\alpha+1},\ldots,\beta_s}$ is not empty for some $\beta_{\alpha+1}$, $\ldots$, β_s , does not exceed $K_{11}(\frac{1}{\varepsilon})^{-\frac{n_\alpha}{k_{\alpha-1}-1}+\frac{n_\alpha}{k_\alpha-1}}$ (where $k_0 \overset{\text{def}}{=} \infty$).

Proof. By definition, $D_{\beta_1,\ldots,\beta_s} = \{x \in B^1 / F^{\alpha-1}(x) \in B^\alpha_{\beta_\alpha}$, $\alpha = 1, \ldots, s\}$. Hence $F^{\alpha-1}(D_{\beta_1,\ldots,\beta_s}) \subset B^\alpha_{\beta_\alpha} \cap f^{\alpha-1}(B^{\alpha-1}_{\beta_{\alpha-1}})$.

Now $B^{\alpha-1}_{\beta_{\alpha-1}}$ is a ball of radius $r_{\alpha-1}$, and therefore $f^{\alpha-1}(B^{\alpha-1}_{\beta_{\alpha-1}})$ is contained in some ball B in R^{n_α} of radius

$M_1(f^{\alpha-1})r_{\alpha-1} \le K_{12}r_{\alpha-1}$. But clearly B has nonempty intersection with not more than $(2\sqrt{n_\alpha}K_{12}r_{\alpha-1}/r_\alpha)^{n_\alpha}$ = $K_{11}(1/\varepsilon)^{\frac{n_\alpha}{k_\alpha-1}-\frac{n_\alpha}{k_{\alpha-1}-1}}$ balls B^α_β .

Thus the number of nonempty $D_{\beta_1,\ldots,\beta_s}$ is bounded by

$$K^s_{11}(1/\varepsilon)^{\frac{n_1}{k_1-1}-\frac{n_2}{k_1-1}+\frac{n_2}{k_2-1}-\cdots-\frac{n_s}{k_{s-1}-1}+\frac{n_s}{k_s-1}} = K_{13}(1/\varepsilon)^{\sigma(S)} .$$

Since by lemma 3 each $\Delta_{\beta_1,\ldots,\beta_s}$ can be covered by K_{10} balls of radius ε , we have $M(\varepsilon,\Delta(f)) \le K_{10}K_{13}(\frac{1}{\varepsilon})^{\sigma(S)}$.

Theorem 1 is proved.

Theorem 1 can be applied to the representability question in the following way : clearly, any mapping f , representable with the diagram $S = (n_1,k_1, \dots ,n_s,k_s)$, $n_1 \geq n_2 \geq \dots \geq n_s$ $k_1 > k_2 > \dots > k_s$, is at least k_s - smooth. On the other hand, any k_1 - smooth mapping is representable with the diagram S by the remark above. Hence, the question is nontrivial for mappings $f : B^1 \to B^{s+1}$ of smoothness q, $k_1 > q > k_s$.

Now we use the sharp estimate of the entropy dimension of the set of critical values, obtained in [6]. By theorem 5.6 ,[6] , for any $f \in C^q(B^1,n_{s+1})$, $\dim_e \Delta(f) \leq \frac{n_1}{q}$, and for any $\eta < \frac{n_1}{q}$ there are mappings $f \in C^q(B^1,n_{s+1})$ with $\dim_e \Delta(f) > \eta$. Thus we have the following:

Corollary 5. Let S be a given diagram and let $k_1 > q > k_s$. If $\frac{n_1}{q} > \sigma(S)$, then there are mappings $f \in C^q(B^1,n_{s+1})$, not representable with the diagram S.

Notice that the consideration of the ε - entropy of correspondings sets in the space of continuous mappings of B^1 (see [3],[4]) also implies the result of corollary 5. But in this way no specific examples can be found, while the mappings with the "big" set of critical values can be built explicitely. Hence in any situation , covered by corollary 5, we can find explicit examples of nonrepresentable mappings. In particular, let $B \subset R^n$ be the closed unit ball and let $h_n \in C^{n-1}(B,1)$ be the Whitney function (see [5]) with $\Delta(f) = [0,1]$. Since $\dim_e [0,1] = 1$, we obtain

Corollary 6. h_n cannot be represented with the diagram $S = (n,k_1, n_2,k_2)$, if

$$\frac{n-n_2}{k_1-1} + \frac{n_2}{k_2-1} < 1 \quad .$$

In particular, $h_{10}(x_1,\ldots,x_{10})$ cannot be written as $h_{10}(x_1,\ldots,x_{10}) = \psi(y_1,\ldots,y_5)$, $y_i = \psi_i(x_1,\ldots,x_{10})$, $i = 1,\ldots,5$, with $\psi \in C^7$ and $\psi_i \in C^{32}$.

There is some similarity in the properties of functions representable by means of compositions of smooth functions and maximum functions of smooth families (compare [7]), although the last class contains nondifferentiable functions.

It is interesting, whether direct connections between these two classes can be found?

REFERENCES

1. V.I.Arnold, On functions of three variables, Dokl. Akad. Nauk SSSR, 114 (1957), 679 - 681 .

2. A.N.Kolmogorov, On 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.

3. A.N.Kolmogorov and V.M.Tihomirov, ε - entropy and ε - capacity of sets in function spaces, Uspehi Mat. Nauk 14, N 2 (1959), 3 - 86, Amer. Math. Soc. Transl. (2) 17 (1961), 277 - 364.

4. A.G.Vitushkin, On representation of functions by means of superpositions and related topics, Enseignement Math. 23 (1977), 255 - 320.

5. H.Whitney, A function not constant on a connected set of critical points, Duke Math. J. 1 (1935), 514 - 517.

6. Y.Yomdin, The geometry of critical and near-critical values of differentiable mappings, Math. Ann. 264 (1983), 495 - 515.

7. Y.Yomdin, On functions representable as a supremum of a family of smooth functions, II , to appear.