ON FUNCTIONS REPRESENTABLE AS A SUPREMUM OF A FAMILY OF SMOOTH FUNCTIONS, II

83 31

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 83-31

ON FUNCTIONS REPRESENTABLE AS A SUPREMUM OF A FAMILY OF SMOOTH FUNCTIONS, II

Y. Yomdin*

ABSTRACT. The classes S_m^q of functions $f(x)$, representable as $\sup_t h(x,t)$, where t is an m - dimensional parameter and h is a C^q - smooth function of x and t, are studied. Considering the "massiveness" of the sets S_m^q in appropriate functional spaces, we show that these classes really differ for different q and m.

Studying geometric invariants of maximum functions, related to the critical values of smooth mappings involved, we give explicit examples of "nice" functions, nonrepresentable as the maximum of "too smooth" families.

1. Introduction. Functions, represented as a supremum of families of a certain type, arise naturally in many questions of analysis and optimization (see e.g.[1],[5],[6],[13] and many others). In [7],[9],[12] the class $H(D)$ has been

* 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 .

considered of functions f on the domain $D \subset R^n$, representable as $f(x) = \sup_{h \in Q} h(x)$, $x \in D$, where Q is a bounded in a C^2 - norm family of twice differentiabl functions on D. The functions $f \in H(D)$ have many nice properties, both geometric and analytic. The important point is also that the consideration of families of C^k - smooth functions, bounded in C^k - norm, for $k > 2$, does not restrict the class H. In fact, it is shown in [7],[9] that $H(D)$ can be described as the class of all f, representable as $f(x) = \sup_{p \in P} p(x)$, where P is a bounded subset in the space of all the quadratic polynomials on R^n.

Another important class of maximum functions appears when we assume that the family Q is smoothly parametrized.

Let $S^q_m(D)$ denote the set of functions f, representable as $f(x) = \max_{t \in T^m} h(x,t)$, where T^m is a compact m - dimensional smooth manifold and $h : D \times T^m \to R$ is a q - times continuously differentiable function, $q \geq 2$. Clearly, $S^q_m(D) \subset H(D)$.

A very precise information on the local structure of "generic" functions in $S^q_m(D)$ has been obtained by methods of Singularities theory (see e.g.[1],[5],[12],[13]).

However, the following important question seems to be untouched: do the classes S^q_m really depend on q and m This question is especially interesting in view of the independence of the class H above of the smoothness of functions involved.

In the present parer we answer this question, showing in many cases noncoincidence of S_m^q for different q and m, although our results are not strong enough to separate these classes completely.

Two different approaches are used: first we study the "massiveness" of the sets S_m^q in appropriate functional spaces, in a way similar to that used in [4], [10] for the problem of representability and approximations by means of compositions. This method allows to separate classes S_m^q rather accurately, but does not give explicit examples.

Another approach is based on the study of geometric invariants of maximum functions, related in one or another way to the structure of the sets of critical values of the smooth families, defining these functions. On this way we obtain various explicit examples of functions in $S_m^q \setminus S_{m'}^{q'}$.

To give the flavor of these examples, we state here the following

Theorem 4.4. Let f(x) be a convex piecewise-linear function of the single variable $x \in [0,1]$, with the countable number of "edges". Let $\delta_1 < \delta_2 < \dots < \delta_i < \dots$. be the slopes of these edges, and let $\alpha_i = \delta_{i+1} - \delta_i$.

Then f can be represented in a form

$$f(x) = \max_{t \in [0,1]} a(t)x + b(t) \quad , \quad x \in [0,1],$$

with a(t) and b(t) k times continuously differentiable functions on [0,1], if and only if $\sum_{i=1}^{\infty} \alpha_i^{1/k} < \infty$.

There are many open questions, concerning the structure of maximum functions, some of which we discuss in the last section.

The author would like to thank Y.Kanai for useful discussions and the Max-Planck-Institut für Mathematik, where this paper was written, for its kind hospitality.

2. ε - entropy of sets of maximum functions.

Let $D \subset R^n$ be a bounded closed domain. For $q = p + \alpha$, $p \geq 1$ - an integer, $0 < \alpha \leq 1$, we denote by $C^q(D)$ the space of p times continuously differentiable functions g on D, whose derivatives of order p satisfy the Hölder condition

$$(*) \qquad \|d^p g(x) - d^p g(y)\| \leq L\|x-y\|^\alpha ,$$

with some constant L.

$$\text{Let} \quad M_i(g) = \max_{y \in D} \|d^i g(y)\| \quad , \quad i = 0, \dots , p,$$

$$M_q(g) = \inf L \quad \text{in} \quad (*) .$$

(We consider all the Euclidean spaces R^s and the spaces of their linear and multilinear mappings with the usual Euclidean norms).

For $c > 0$ denote by $C^q(D,c)$ the set of all g in $C^q(D)$ with $M_i(g) \leq c$, $i = 0, \dots , p, q$.

Let F be a relatively compact set in a metric space X. For some $\varepsilon > 0$ a set $F^* \subset X$ is called an ε - net of F if for any $z \in F$ there exists $z^* \in F^*$ with

$d(z,z^*) \le \varepsilon$, where d denotes the distance in X.

Denote by $N_\varepsilon(F)$ the number of elements in a minimal ε - net of F. The number $H_\varepsilon(F) = \log_2 N_\varepsilon(F)$ is called the ε - entropy of the set F. It is convenient also to define the number $fd(F)$ as $\overline{\lim\limits_{\varepsilon \to 0}} \log_2 H_\varepsilon(F)/\log_2(1/\varepsilon)$.

The notion of entropy arises in a natural way in connection with various problems of analysis (see e.g.[4],[10], [14],[16]).

In this section the metric space X is the space $C(D)$ of continuous functions on D with the uniform norm. Now we turn back to maximum functions. Without loss of generality we can assume that the parameter manifold T^m is the unit m - dimensional cube I^m . For $c > 0$ denote by $S_m^q(D,c)$ the set of functions f on D, representable as $f(x) = \max_t h(x,t)$, with $h \in C^q(D \times I^m, c)$. Clearly, for any $c > 0$, $S_m^q(D,c)$ is a compact subset in $C(D)$.

Theorem 2.1. For any $c > 0$,

$$\frac{n+m}{q + 2m/n} \le fd(S_m^q(D,c)) \le \frac{n+m}{q} .$$

Proof. Consider the mapping $\mu : C(D \times I^m) \to C(D)$, defined by $\mu(h) = f$, $f(x) = \max\limits_{t \in I^m} h(x,t)$.

Clearly, μ does not increase the distance. From the definition of ε - entropy we obtain that for any relatively compact set $F \subset C(D \times I^m)$, $H_\varepsilon(\mu(F)) \le H_\varepsilon(F)$.

Since $S_m^q(D,c) = \mu(C^q(D,c))$ and since according to theorem 2.2.1 [10], $fd(C^q(D,c)) = \frac{n+m}{q}$, we obtain the right-hand side inequality.

To obtain the lower bound for $fd(S_m^q(D,c))$ it is sufficient to find in this set a suitable number of functions, any two of which differ at least by 2ε in C - norm.

Let us fix some $\delta > 0$ and let $\delta' = \frac{1}{2}\delta^{1+m/n}$.

Let $Z_\delta \subset R^n$ denote the net of points of the form $z = (k_1\delta, k_2\delta, \ldots, k_n\delta)$, $k_i \in Z$.

Consider also the points $z'_\alpha \in R^n$ of the form $(k_1\delta', \ldots, k_n\delta')$, $1 \le k_i \le [\delta/\delta']$, indexed in some fixed way, $1 \le \alpha \le [\delta/\delta']^n$. Finitely, let $z''_\beta \in I^m$ be the points of the form $(k_1\delta, \ldots, k_m\delta)$, $0 \le k_i \le [1/\delta]$, indexed in some fixed way, $1 \le \beta \le [1/\delta]^m$.

Since $\delta/\delta' = 2(1/\delta)^{m/n}$,

$[\delta/\delta']^n \sim 2^n(1/\delta)^m > [1/\delta]^m$ for δ sufficiently small, we can fix some one-to-one mapping ω from the set of indices β into the set of indices α.

Now consider in $R^n \times I^m$ the net of points y of the form $y = (z + z'_{\omega(\beta)}, z''_\beta)$, $z \in Z_\delta$, $1 \le \beta \le [1/\delta]^m$, and let y_ξ, $1 \le \xi \le K(\delta)$, denote those from the points y, whose projections x_ξ on R^n belong to D. Since D has a nonempty interior, $K(\delta) \ge K'(1/\delta)^{n+m}$, with some $K' > 0$, not depending on δ.

We shall use below only the following property of points

y_ξ : they form a net in $D \times I^m$ with the distance at least δ between any two points, while their projections x_ξ form a net in D with the distance at least δ' between any two points.

Let $\phi : R \to R$ be a C^∞ - smooth even function with the following properties:

i. $\phi(s) = 1 - s^2$ for $|s| \le \frac{1}{2}$

ii. $\phi(s) = 0$ for $|s| \ge 1$

iii. $\phi(s)$ is a decreasing function of $|s|$ for $\frac{1}{2} \le |s| \le 1$.

Let $\psi : R^{n+m} \to R$ be defined by $\psi(y) = \phi(\|y\|)$. Denote by M $2^q \cdot \max M_i(\psi)$, $i = 1, \ldots , p+1$.

Now for any subset $\kappa \subset \{1, \ldots , K(\delta)\}$ define the function $\psi_\kappa : D \times I^m \to R$ as follows: for $y = (x,t) \in D \times I^m$

$$\psi_\kappa(y) = \frac{c}{M} \delta^q \sum_{\xi \in \kappa} \psi(\tfrac{2}{\delta}(y-y_\xi)) .$$

Clearly, ψ_κ is a C^∞ function, and since the supports of $\psi(\frac{2}{\delta}(y-y_\xi))$ are disjoint for different ξ by the property ii of ϕ, we see that $\psi_\kappa \in C^q(D \times I^m, c)$.

Now consider the corresponding maximum function $f_\kappa = \mu(\psi_\kappa)$, which by definition belongs to $S^q_m(D,c)$.

Lemma 2.2. For $\kappa \ne \kappa'$, $\|f_\kappa - f_{\kappa'}\|_C \ge \frac{c}{M} \delta^{q + 2m/n}$.

Proof. Since the supports of $\psi(\frac{2}{\delta}(y-y_\xi))$ are disjoint, we have: $f_\kappa(x) = \frac{c}{M} \delta^q \max_{\xi \in \kappa} \max_{t \in I^m} (\frac{2}{\delta}(y-y_\xi)) =$

$\frac{c}{M} \delta^q \max_{\xi \in \kappa} \phi(\frac{2}{\delta} \|x-x_\xi\|)$, by construction of function ψ .

Now assume that $\kappa \setminus \kappa' \neq \phi$ and fix some $\eta \in \kappa \setminus \kappa'$. Then $f_\kappa(x_\eta) = \frac{c}{M}\delta^q \phi(0) = \frac{c}{M}\delta^q$, by the property i of ϕ. On the other hand, for any $\xi \neq \eta$,

$$\phi(\tfrac{2}{\delta}\|x_\eta - x_\xi\|) = 1 - (\tfrac{2}{\delta}\|x_\eta - x_\xi\|)^2 \leq 1 - 4(\tfrac{\delta'}{\delta})^2,$$

since $\|x_\eta - x_\xi\| \geq \delta'$ for $\xi \neq \eta$. Hence

$$f_{\kappa'}(x_\eta) \leq \frac{c}{M}\delta^q - \frac{c}{M}\delta^q\, 4(\tfrac{\delta'}{\delta})^2 = f_\kappa(x_\eta) - \frac{c}{M}\delta^{q+2m/n}$$

Therefore $\|f_\kappa - f_{\kappa'}\|_C \geq \|f_\kappa(x_\eta) - f_{\kappa'}(x_\eta)\| \geq \frac{c}{M}\delta^{q+2m/n}$

Now for given $\varepsilon > 0$ let $\delta = (\frac{3M}{c}\varepsilon)^{1/q+2m/n}$. Then

$\frac{c}{M}\delta^{q+2m/n} = 3\varepsilon$ and by lemma 2.2 all the functions f_κ form the 3ε - separated net in $S_m^q(D,c)$. Since the number of elements in this net is equal to

$$2^{K(\delta)} \geq 2^{K'(1/\delta)^{n+m}} = 2^{K''(1/\varepsilon)^{\frac{n+m}{q+2m/n}}},$$ we have

$H_\varepsilon(S_m^q(D,c)) \geq K''(1/\varepsilon)^{\frac{n+m}{q+2m/n}}$. Theorem 2.1 is proved.

As an immediate consequence of theorem 2.1 we obtain the main result of this section:

Theorem 2.3. Let D be a compact domain in R^n. Then for any m, q and m', q', such that $\frac{n+m}{q+2m/n} > \frac{n+m'}{q'}$, the set of functions in $S_m^q(D) \subset C(D)$, not belonging to $S_{m'}^{q'}(D)$, is a set of second category. In particular, this set is everywhere dense in $S_m^q(D)$.

Proof. Clearly, $S_{m'}^{q'}(D) = \bigcup_{N=1}^{\infty} S_{m'}^{q'}(D,N)$. By theorem 2.1

for any ball Ω in $C(D)$, $fd(S_m^q(D) \cap \Omega) \geq \frac{n+m}{q+2m/n}$.

But by the same theorem, $fd(S_{m'}^{q'}(D,N)) \leq \frac{n+m'}{q'} < \frac{n+m}{q+2m/n}$.

Since $S_{m'}^{q'}(D,N) \cap S_m^q(D)$ is closed, it is therefore a nowhere dense subset in $S_m^q(D)$.

Now we give some corrolaries, showing what classes S_m^q can be separated by means of theorem 2.3.

Corollary 2.4. For given $D \subset R^n$ and m and for any q and $q' > q + 2m/n$, $S_m^{q'}(D) \subsetneq S_m^q(D)$. In particular, if $n > 2m$, then for any $q' > q$, $S_m^{q'}(D) \subsetneq S_m^q(D)$

Proof. For $q' > q + 2m/n$, $\frac{n+m}{q'} < \frac{n+m}{q+2m/n}$.

Corollary 2.5. For given $D \subset R^n$, q and m , and for any $m' < m - \frac{2n(n+m)}{2n+qm}$, $S_{m'}^q(D) \subsetneq S_m^q(D)$. In particular, for $q > \frac{2n^2-1}{m} + 2n$, $S_{m'}^q(D) \subsetneq S_m^q(D)$ for any $m' < m$, and for $q > 2n^2 + 2n - 1$, $S_{m'}^q(D) \subsetneq S_{m''}^q(D)$ for any $m' < m''$.

Proof. For $m' < m - \frac{2n(n+m)}{2n+qm}$, $\frac{n+m'}{q} < \frac{n+m}{q+2m/n}$.
For $q > \frac{2n^2-1}{m} + 2n$, $\frac{2n(n+m)}{2n+qm} < 1$. Finitely, for $q > 2n^2 + 2n - 1$, the last inequality is satisfied for any m .

We can use the result of theorem 2.1 also to compare $S^q_m(D)$ with $C^k(D)$. Indeed, by theorem 2.2.1 [10], $\mathrm{fd}\, C^k(D) = \frac{n}{k}$, and repeating the proof of theorem 2.3 we obtain the following:

Corollary 2.6. For $\frac{n+m}{q} < \frac{n}{k}$, the set of k times continuously differentiable functions on D, not belonging to $S^q_m(D)$, is a set of second category and, in particular, is everywhere dense in the uniform topology.

Notice that corrolaries 2.4, 2.5 and 2.6 do not give explicit examples of functions, not representable as a maximum of a suitably smooth family. Below we give such examples for any situation, covered by corollary 2.6.

3. Critical values of maximum functions.

Let $f : D \to R^s$ be a continuous mapping. The point $x \in \mathring{D}$ is called a critical point of f, if the first differential $df(x)$ exists and is equal to zero. Let $\Sigma(f)$ be the set of all the critical points of f and let $\Delta(f) = f(\Sigma(f)) \subset R^s$ be the set of critical values of f. (Actually, $\Sigma(f)$ is the set of critical points of rank zero of f, in usual terminology).

A well-known and widely used property of critical values of differentiable mappings is given by the Morse-Sard theorem (see [3],[8]): if the mapping is sufficiently smooth, the set of its critical values has the Lebesgue measure (or, more precisely, the Hausdorff measure of an appropriate dimension) zero.

In [14] the stronger property of critical values has been established: let A be a bounded subset in R^s. The entropy dimension of A, $\dim_e A$ is defined as

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

where $N_\varepsilon(A)$, as above, is the number of elements in a minimal ε - net of A in R^s.

For "nice" sets the entropy dimension coincides with the Hausdorff dimension $\dim_h$ and with the topological dimension. For any A $\dim_e A \geq \dim_h A$, and the important advantage of the entropy dimension, which we shall use below, is that it allows to distinguish countable sets, while the Hausdorff dimension of any countable set is zero.

The following result has been obtained in [14] :

Theorem 5.4 ([14]). Let $D \subset R^n$ be a compact domain and let $f : D \to R^s$ be a C^q - mapping. Then

$$\dim_e \Delta(f) \leq \frac{n}{q} .$$

(The Morse-Sard theorem, in its general form, proved in [3] , gives the same bound for the Hausdorff dimension of $\Delta(f)$).

It turns out that also the critical values of the function, representable as a maximum of a smooth family, cannot form "too big" set.

Theorem 3.1. Let $D \subset R^n$ be a compact domain. For any $f \in S_m^q(D)$,

$$\dim_e \Delta(f) \le \frac{n+m}{q} .$$

Proof. Let $f(x) = \max_{t \in I^m} h(x,t)$, $h \in C^q(D \times I^m)$.

Lemma 3.2. Let $x_0 \in D$ be a critical point of f and let $t_0 \in I^m$ be such that $f(x_0) = h(x_0,t_0)$. Then $(x_0,t_0) \in D \times I^m$ is a critical point of h.

Proof. Since $h(x_0,t)$ attains its maximum with respect to t at t_0, we have $d_t h(x_0,t_0) = 0$.

By the definition of critical points, f is differentiable at x_0 and $df(x_0) = 0$. But then from the expression of the generalized differential of maximum functions (see [2]) it follows immediately that $d_x h(x_0,t_0)$

Thus by lemma 3.2 any critical value $f(x_0)$ of f is a critical value $h(x_0,t_0)$ of h, i. e. $\Delta(f) \subset \Delta(h)$. But by theorem 5.4 [14], $\dim_e \Delta(h) \le \frac{n+m}{q}$.

Now to obtain examples of functions, nonrepresentable as maximum, we note that for any $\gamma < \frac{n}{k}$, by theorem 5.6 [14], the function $g \in C^k(D)$ can be built, with $\dim_e \Delta(g) \ge \gamma$. By theorem 3.1, g does not belong to any $S_m^q(D)$, with $\frac{n+m}{q} < \frac{n}{k}$, and therefore, we get explicit examples of nonrepresentable functions in any situation, covered by corollary 2.6.

In particular, let $g_n : I^n \to R$, $g_n \in C^{n-1}(I^n)$ be the Whitney function with $\Delta(g) = [0,1]$. (See [11]).

Since $\dim_e[0,1] = 1$, we have

Corollary 3.3. $g_n \notin S^q_m(I^n)$ for $q > n + m$.

Notice that in all the constructions above it is sufficient to use the Hausdorff dimension of the sets of critical values. However, using the specific properties of the entropy dimension we can give examples of very simple functions, not representable as maximum:

Let $\psi_k : [0,1] \to R$ be defined as $\psi_k(x) = x^k \cos(\frac{1}{x})$. $\psi_k \in C^{[k/2]-1}([0,1])$ and the critical values of ψ_k form the sequence $-(\frac{1}{\pi})^k$, $(\frac{1}{2\pi})^k$, ..., $(-1)^i(\frac{1}{i\pi})^k$, ...

Hence $\dim_e \Delta(\psi_k) = \frac{1}{k+1}$ (see e.g. [15]).

Corollary 3.4. $\psi_k \notin S^q_m([0,1])$ for $q > (k+1)(m+1)$.

4. Maxima of smooth families of linear functions.

In this section we give simple examples of convex functions, nonrepresentable as a maximum of a sufficiently smooth family of linear functions. Once more we reduce the question of representability to the properties of critical values of some differentiable mappings. But here, in contrast to section 3, the arising sets of critical values are apriori at most countable. Thus the Morse-Sard theorem gives no information in this case and the use of the entropy dimension and the stronger theorem 5.4 [14] becomes essential. In fact, this theorem was found in attempt to give criteria of representability of convex functions as maximum of linear ones.

Let D be a convex compact domain in R^n. We consider a cone $Q(D)$ of convex functions f on D, which are extendable to convex functions on R^n and whose graph $\Gamma(f)$ over D is a polyhedron with possibly countable number of faces.

We study the representability of $f \in Q(D)$ as $f = \max_{t \in I^m}$

where for $x = (x_1, \dots, x_n) \in R^n$

$l_t(x) = a_1(t)x_1 + \dots + a_n(t)x_n + b(t)$, with $a_1, \dots, a_n, b \in C^q($

Denote the set of functions in $Q(D)$, representable in this form, by $Q^q_m(D)$.

Let L be a hyperplane in $R^n \times R$ with a nondegenerate projection on R^n. L is a graph of the linear function $l_L(x) = a_1(L)x_1 + \dots + a_n(L)x_n + b(L)$. Denote the point $(a_1(L), \dots, a_n(L)) \in R^n$ by $\delta(L)$.

For a given $f \in Q(D)$ let $\delta(f) \subset R^n$ be the set of $\delta(\gamma)$, where γ runs through all the faces of $\Gamma(f)$. $\delta(f)$ is at most countable bounded subset in R^n.

Theorem 4.1. For any $f \in Q^q_m(D)$, $\dim_e \delta(F) \leq \frac{m}{q}$.

Proof. Write f as $\max l_t$, $l_t(x) = a_1(t)x_1 + \dots + a_n(t)x_n + b(t)$

and for each face γ of $\Gamma(f)$ find some $t_\gamma \in I^m$, such that γ is a graph of l_{t_γ}.

Lemma 4.2. For any face γ of $\Gamma(f)$, t_γ is a critical point of each of functions $a_1, \dots, a_n, b$.

Proof. Let $x^0, x^1, \dots, x^n$ be the vertices of some nondegenerate simplex in the projection of γ on R^n. Since at each x^i, $f(x^i) = l_{t_\gamma}(x^i) = \max_t l_t(x^i)$, we have $d_t l_t(x^i)|_{t=t_\gamma} = 0$, or

$$d_t a_1(t_\gamma)x_1^i + \dots + d_t a_n(t_\gamma)x_n^i + d_t b(t_\gamma) = 0, \quad i = 0,\dots,n .$$

But since $x^0, \dots, x^n$ are the vertices of a nondegenerated simplex, this linear system has only zero solution.

Thus if we define the mapping $\Phi : I^m \to R^n$ by $\Phi(t) = (a_1(t), \dots, a_n(t))$, any point t_γ is a critical point of Φ and $\delta(\gamma) = \Phi(t_\gamma)$ is a critical value of Φ. Hence $\delta(f) \subset \Delta(\Phi)$, and since by theorem 5.4 [14] $\dim_e \Delta(\Phi) \leq \frac{m}{q}$, theorem 4.1 is proved.

Now consider, for instance, the set $\Delta_\alpha \subset R^n$, consisting of points of the form $(\frac{1}{k_1^\alpha}, \frac{1}{k_2^\alpha}, \dots, \frac{1}{k_n^\alpha})$, $k_i = 1,2,\dots$, $\alpha > 0$. We have $\dim_e \Delta_\alpha = \frac{n}{\alpha+1}$ (see e.g. [15]).

One can easily find functions $f_\alpha \in Q(I^n)$ with $\delta(f) = \Delta_\alpha$. We obtain the following:

Corollary 4.3. $f_\alpha \notin Q_m^q(I^n)$ for $\frac{m}{q} < \frac{n}{\alpha+1}$.

Using the metric invariant V_β, defined in [15], one can give more precise version of theorem 4.1. We shall consider only one special case, where the question of representability can be answered completely.

Let $f \in Q([0,1])$ be a function, for which $\delta(f)$ is a sequence $\delta_1 < \delta_2 < \dots < \delta_i < \dots$ and let $\alpha_i = \delta_{i+1} - \delta_i$

Theorem 4.4. Function f can be represented as $f(x) = \max_{t \in [0,1]} a(t)x + b(t)$, $x \in [0,1]$, with a and b k times continuously differentiable functions on $[0,1]$, if and only if $\sum_{i=1}^{\infty} \alpha_i^{\frac{1}{k}} < \infty$.

Proof. If f has a required representation, then, by lemma 4.2, $\delta(f) = \{\delta_1, \delta_2, \dots\} \subset \Delta(a)$. But then by theorem 4.1 [15], $\sum_{i=1}^{\infty} \alpha_i^{\frac{1}{k}} < \infty$.

Now assume that $\sum_{i=1}^{\infty} \alpha_i^{\frac{1}{k}} < \infty$. In proof of theorem 4.1 [15] it is shown that we can find the function $a : [0,1] \to R$ with the following properties:

i. $a \in C^{\infty}([0,1))$ and all the derivatives of a up to order k at $t \in [0,1)$ tend to zero as t tends to 1 (and, in particular, a is k times continuously differentiable on $[0,1]$.

ii. a increases on $[0,1]$.

iii. There is a sequence of points $t_1, t_2, \dots$ in $[0,1)$, converging to 1, such that $a(t_i) = \delta_i$

and all the derivatives of a at t_i vanish, $i = 1,2,\ldots$.

Using $a(t)$ and f, we now define $b(t)$ as follows: $b(t)$ is the constant term in the equation of the support line to the graph $\Gamma(f)$ with the slope $a(t)$.

Clearly, $f(x) = \max_t a(t)x + b(t)$. It remains to prove that b is k times continuously differentiable on $[0,1]$.

Let (x_i,y_i) be the coordinates of the vertex of $\Gamma(f)$, belonging to the edges of $\Gamma(f)$ with the slopes δ_i and δ_{i+1} , $i = 1, 2, \ldots$. Then for $\delta_i \le a(t) \le \delta_{i+1}$, i.e. for $t_i \le t \le t_{i+1}$, $b(t) = y_i - a(t)x_i$. Hence $b(t)$ is C^∞ - smooth on each segment $[t_i,t_{i+1}]$, and its derivatives coincide with the derivatives of $a(t)$ up to a coefficient $-x_i$. But by the condition iii, all the derivatives of a vanish at t_i. Hence $b \in C^\infty([0,1))$. Since by i , all the derivatives of $a(t)$ up to order k tend to zero as t tends to 1 , the same is true for $b(t)$ and hence b is a k times continuously differentiable function on $[0,1]$.

5. Some open questions.

Of course, the results above are far from being complete. However, they show that there is a rich vriety of interesting phenomena concerning the maxima of smooth families, as well as various connections with the deep properties of smooth mappings.

Theorem 2.1 gives bounds for the functional dimension fd of the sets of maximum functions. What is the precise value of $fd(S_m^q(D,c))$?

Notice also that the invariants, obtained in sections 2 and 3 , "mixe" the smoothness q and the dimension m of the parameter space. Could one find invariants of maximum functions, separating the influence of these factors?

There is a big similarity between the study of maximum functions above and the study of functions, representable by means of compositions of functions of some given classes. (see [10], [16]). In both cases the consideration of the ε - entropy allows to prove the existance of nonrepresentable functions, while the study of some invariants, related to critical values, gives explicit examples of such functions. Are there direct connections between these two problems?

The necessary condition for the representability of a polyhedral convex function as a maximum of a smooth family of linear functions, given by theorem 4.1, is very close to the sufficient one (see theorem 4.4). However, for an arbitrary (not polyhedral) convex function, the method used here breaks. Could one define invariants of a general convex function, responsible for the representability of this function as a maximum of a smooth family of linear functions?

The maximum functions of smooth families have nice differentiability properties, which can be formulated, in particular, in terms of their generalized derivatives (both in sense of distributions - see [12], and in sense of optimization theory - see [2],[6],[7], [9]). Could one give criteria of representability in these terms? In particular, could one find functional spaces, appropriate for treatment of maximum functions?

There is one particular question, concerning the differentiability properties of maximum functions. A continuous function $f : R^n \to R$ is said to have the k - th differential at x_0, if there exists a polynomial $P : R^n \to R$ of degree k, such that $\|f(x)-P(x)\| = o(\|x-x_0\|^k)$. Convex functions are known to have the second differential almost everywhere. Is it true that functions in $S_m^k(D)$ have the k-th differential almost everywhere in D ? Some variant of this question (for finite families) is considered in [17] .

REFERENCES

1. L.N.Bryzgalova, Singularities of a maximum of a function, depending on parameters, Functional Anal. Appl., 11(1977) N1, 49 - 50 .

2. F.H.Clarke, Generalized gradients and applications, Trans. Amer. Math. Soc., 205(1975), 247 - 262.

3. H.Federer, Geometric measure theory, Die Grundlehren der mathematischen Wissenschaften, Berlin, Heidelberg, New York, Springer, 1969.

4. A.N.Kolmogorov and V.M.Tihomirov, ε - entropy and ε - capacity of sets in functional spaces, Usp. Mat. Nauk 2 (86), 14 (1959), 3 - 86, Transl. Amer. Math. Soc. 72, (1961), 277 - 364.

5. V.I.Matov, Topological classification of germs of maximum and minimax functions of generic families, Usp. Mat. Nauk 37 , 4 (1982), 167 - 168.

6. B.N.Pschenichnyi, Necessary conditions for an extremum, Marsel Dekker, New York, 1971.

7. R.T.Rockafellar, Favorable classes of Lipschitz continuous functions in subgradient optimization, in Progress in nondifferentiable optimization (E.Nurminski, ed.), Pergamon Press, 1981.

8. A.Sard, The measure of the critical values of differentiable maps, Bull. Amer. Math. Soc. 48 (1942), 883 - 890.

9. A.Shapiro and Y.Yomdin, On functions representable as a difference of two convex functions, and necessary conditions in a constrained optimization, preprint, 1981.

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

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

12. Y.Yomdin, On functions representable as a supremum of a family of smooth functions, SIAM J. Math. Anal., 14, N2 (1983), 239 - 246.

13. Y.Yomdin, On the local structure of a generic central set, Compositio Math., 43 (1981), 225 - 238.

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

15. Y.Yomdin, β - spread of sets in metric spaces and critical values of smooth functions, to appear.

16. Y.Yomdin, Critical values and representation of functions by means of compositions, to appear.

17. Y.Yomdin, Some results on finite determinacy and stability not requiring the explicit use of smoothness, Proceedings of Symp. Pure Math., 40 (1983), 667 -674.