ON REPRESENTABILITY OF CONVEX FUNCTIONS AS MAXIMA OF LINEAR FAMILIES

Y. Yomdin

Max-Planck-Institut für Mathematik
Gottfried-Claren-Str. 26
5300 Bonn 3
Federal Republic of Germany

Ben Gurion University
of the Negev
Beer-Sheva 84105
Israel

SFB 40/MPI 84 - 40

Abstract. Some necessary conditions are given for representability of a given convex function g in a form $g(x)=\max_t \ell_t(x)$, where ℓ_t is a family of linear functions of x, C^k-smoothly defending on an m-dimensional parameter t.

Introduction:

Functions represented as $f(x)=\max_t h(x,t)$, where $h(x,t)$ is a differentiable function of variables x and t, appear in various questions of analysis and optimization. Recently some new results have been obtained ([1], [2], [6], [7], [8]), showing that the reach techniques from smooth analysis can be applied in study of maximum functions. In particular, it turns out, that the high order differentiability of h with respect to both variables x and t is strongly reflected in the intrinstic structure of the maximum function f ([7], [8]). The situation here is somewhat dual: from one side, if we assume, that the uniform bounds for the derivatives of h of some fixed order k are explicitly known, we obtain effective bounds for the geometry of critical and "almost critical" points and values of f ([8]). These restrictions can be used (and verified) in effective computations with finite accuracy.

From the other side, in order to obtain conditions of representability of f as maximum of C^k-smooth family, we must exclude the explicit bounds of derivatives (although because of a compactness of the domain and the parameter space, which we always assume, some bounds exist). The corresponding "passing to the limit" in geometric restrictions above leads to rather delicate invariants, such as the fractional dimensions of critical values ([7]).

The gap between these two kinds of restrictions is big and in general the present understanding of the structure of maximum functions is far from being complete. In particular, the following basic question is answered only in a few very special situations ([1], [2], [7]): what are the necessary and sufficient condituons for representability of a given function $f(x)$ as a maximum of a C^k-smooth family $h(x,t)$ with respect to a parameter t, belonging to an m-dimensional compact domain?

An important special case of this problem concerns the representability of a given convex function g in a form $g(x) = \max_t \ell_t(x)$, where ℓ_t is a family of linear functions of x , C^k-smoothly defending on a "compact" m-dimensional parameter t. The general question can be, in some sense, reduced to this special one (see proposition 2.2 below).

As far as "polyhedral" convex functions with a countable number of faces are concerned, the necessary conditions of representability, very close to the sufficient ones, were obtained in [7] (theorem 4.1, theorem 4.4). But the method, used in [7] does not work in general case.

The aim of the present paper is to give necessary conditions of a representability of general (not only "polyhedral") convex functions as maxima of smooth families of linear ones.

We give two different type of conditions: the first one in terms of points, where the Hessian of our function (which exists almost everywhere, by the Alexandrov-Buseman-Feller theorem) fails to have the maximal rank.

The second condition concerns the critical values of (nonconvex) functions, which can be obtained by substracting from our initial convex function any sufficiently smooth one.

These two kinds of conditions are shown to be independent; both allow to find many explicit examples of "nice" convex functions, not representable as a maximum of a "too smooth" family of linear ones.

For "polyhedral" convex function the first of our representability conditions coincides with the one of theorem 4.1 [7]. Presumably, it is also "almost sufficient" in this case (see section 4 below). In contrast, for general convex functions, there are strong indications that some additional "invariants of representability" exist. We discuss this question in section 4 below.

The author would like to thank the Max-Planck-Institut für Mathematik, where this paper was written, for the kind hospitality.

2. Definitions and results:

To simplify notations, we assume below, that the maximum functions considered are defined on the unit closed cube $I^n = [0,1]^n \subset R^n$. C^k-differentiability of the function on a closed set always means below, that this function can be extended to a C^k-smooth function on some open neighbourhood of the set.

<u>Definition 2.1.</u> For $n,m,k = 1,2,\ldots,$ $S^k_{n,m}$ denotes the set of functions f on I^n, representable as

$$f(x) = \max_{t\in T^m} h(x,t) \quad , \quad x \in I^n ,$$

where T^m is a compact smooth n-dimensional manifold and $h : I^n \times T^m \to \mathbb{R}$ is a k times continuously differentiable function.

For $n,m,k = 1,2,\ldots,$ $Q^k_{n,m}$ denotes the set of convex functions g on I^n, representable as

$$g(x) = \max_{t\in T^m} \ell_t(x) \quad , \quad x \in I^n ,$$

where $\ell_t(x) = a_1(t)x_1+\ldots+a_n(t)x_n + b(t)$ for $x= (x_1,\ldots,x_n) \in R^n$, and $a_1(t),\ldots, a_n(t), b(t)$ are k times continuously differentiable functions on a compact smooth m-dimensional manifold T^m.

One can easily show, that we can always assume the manifold T^m in definition 2.1 to be a closed unit m-cube $I^m \subset R^m$. Notice also, that the formula above defines functions from $Q^k_{n,m}$ on all the R^n and not only on I^n.

The following proposition relates the classes S and Q:

<u>Proposition 2.2.</u> Let $f \in S^k_{n,m}$, $k \geq 2$. Then there is $K \geq 0$ such

that the function $g = f + K\sigma$ is convex and belongs to $Q^{k-1}_{n,n+m}$, where $\sigma(x) = \|x\|^2 = x_1^2 + \dots + x_n^2$.

Conversely, if $g \in Q^k_{n,m}$, then for any k times continuously differentiable function u on I^n, $g + u \in S^k_{n,m}$.

<u>Proof.</u> Let $f \in S^k_{n,m}$. Fix some representation $f(x) = \max_{t \in T^m} h(x,t)$. Since h is k times continuously differentiable on $I^n \times T^m$, $k \geq 2$, all the second partial derivatives of h are bounded. Put

$$K = \sup_{x,t} \left\| \left(\frac{\partial^2 h(x,t)}{\partial x_i \, \partial x_j} \right) \right\| \quad i,j = 1,\dots,n \quad ,$$

where $\| \ \|$ denotes the Euclidean norm of the $(n \times n)$-matrix $\left(\frac{\partial^2 h}{\partial x_i \, \partial x_j} \right)$.

Consider the function $H(x,t) = h(x,t) + K\sigma(x)$. By the choice of K, for any x,t the matrix $\left(\frac{\partial^2 H}{\partial x_i \partial x_j} \right)$ is positively defined. Hence for any $t \in T^m$, the function $H_t = H(\cdot,t)$ is a convex function on I^n. Therefore,

$$g(x) = f(x) + K\sigma(x) = \max_t \, (h(x,t) + K\sigma(x)) = \max_t H_t(x)$$

is convex, as the maximum of a family of convex functions H_t.

Moreover, each H_t can be represented as $H_t = \max_{t' \in I^n} \ell_{t,t'}(x)$, where $\ell_{t,t'}$ is the supporting linear function of H_t at the point $t' \in I^n$. Clearly, $\ell_{t,t'}$ depends $k-1$ times differentiably on $(t,t') \in T^m \times I^n$. Thus $g(x) = \max_{(t,t') \in T^m \times I^n} \ell_{t,t'}(x) \in Q^{k-1}_{n,n+m}$

Conversely, if $g \in Q^k_{n,m}$, i.e. $g = \max_{t \in T^m} \ell_t(x)$, then

$$g(x) + u(x) = \max_{t\in T^m}(\ell_t(x) + u(x)) \in S^k_{n,m} .$$

Proposition is proved.

To give our conditions of representability we need the notion of the entropy dimension:

Definition 2.3. Let $A \subset R^q$ be a bounded subset. For any $\varepsilon > 0$ denote by $M(\varepsilon,A)$ the minimal number of balls of radius ε, covering A. The entropy dimension of A, $\dim_e A$, is defined as

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

We do not discuss here in detail the properties of the entropy dimension (see e.g. [3],[5],[9]). We mention only that for a "nice" sets A, e.g. for compact submanifolds in R^q, $\dim_e$ coincides with the Hausdorff dimension $\dim_h$ and with the usual topological dimension. In general, $\dim_e A \geq \dim_h K$ and the difference between these two dimensions becomes essential for "scarce" sets. E.g. $\dim_e \{1,\frac{1}{2^\alpha},\ldots,\frac{1}{p^\alpha},\ldots\} = \frac{1}{\alpha+1}$, while the Hausdorff dimension of any countable set is zero.

To formulate our main condition of representability, we remind the notion of the second differential we need:

Definition 2.4. A quadratic polynimial $P(x)$ is called the second differential of a function f at $x_0 \in \mathring{I}^n$, if

$$|f(x) - P(x)| = o(\|x-x_0\|^2) .$$

If the second differential of f at x_0 exists, it is unique and is denoted by $d^2f_{x_0}(x)$.

By the Alexandrow-Buseman-Feller theorem, for any convex function g on I^n, the set G_g of points $z \in \mathring{I}^n$ where d^2g_z exists, has measure 1. For any $z \in G_g$ we define the first and the second order partial derivatives of g at z in the usual way through the coefficients of d^2g_z. In particular, we define the grad $g(z)$ as the vector $(\frac{\partial g}{\partial x_1},\dots,\frac{\partial g}{\partial x_n})_z$ and the Hessian matrix $H_g(z)$ as $H_g(z) = (\frac{\partial^2 g}{\partial x_i \partial x_j})_z$.

<u>Definition 2.5</u> Let for a convex function g on I^n, and for $q = 0,1,\dots,n$, $S_q(g) \subset G_g \subseteq \mathring{I}^n$ be the set of points $z \in G_g$, where the rank $H_g(z) \leq q$.

Let $\sigma_q(g) \subset R^n$ be the set

$$\{\text{grad } g(z), \quad z \in S_q(g)\}.$$

Theorem 2.6. For any $g \in Q^k_{n,m}$, $k \geq 2$, and for $q = 0,1,\dots,n$

$$\dim_e \sigma_q(g) \leq q' + \frac{m-q'}{k},$$

where $q' = \min(q,m)$.

<u>Proof.</u> Write g as $g(x) = \max_{t \in T^m} \ell_t(x)$,
where $\ell_t(x) = a_1(t)x_1+\dots+a_n(t)x_n + b(t)$, $a_1,\dots,a_n,b$ - k times continuously differentiable functions on T^m.

We consider an auxiliary mapping $\phi: T^m \to R^n$, defined by

$$\phi(t) = (a_1(t),\dots,a_n(t)) \in R^n.$$

Lemma 2.7. Let $z \in S_q(g)$, $q = 0,1,\ldots,n$, and let $t_z \in T^m$ be such that

$$g(z) = \max_{t \in T^m} \ell_t(z) = \ell_{t_z}(z).$$

Then t_z is a critical point of rank q of the mapping $\emptyset$ (I.e. the rank of the first differential $d\emptyset(t_z): R^m \to R^n$, or of the restriction of this differential to the boundary ∂T^m, if $t_z \in \partial T^m$, is at most q).

We give the proof of lemma 2.7 below, and now we complete the proof of theorem 2.6.

Clearly, for any $z \in S_q(g) \subset G_g$, grad $g(z) = \emptyset(t_z)$. Hence by lemma 2.7 we have an inclusion $\sigma_q(g) \subset \Delta^q(\emptyset)$, where $\Delta^q(\emptyset)$ is the set of critical values of rank q of $\emptyset$ (i.e. the values of $\emptyset$ at the points, where the rank $d\emptyset$ is at most q). But by theorem 1.3, [5], $\dim_e \Delta^q(\emptyset) \le q' + \frac{m-q'}{k}$, and hence $\dim_e \sigma_q(g) \le q' + \frac{m-q'}{k}$, $q' = \min(m,q)$.

Theorem is proved.

Proof of lemma 2.7.

Let $z \in S_q(g)$. After an appropriate affine coordinate transformation in R^n we can assume that $z = 0 \in R^n$.

By definition of $S_q(z)$ we have rank $H_g(z) \le q$. Hence we can find an n-q-dimensional subspace $V \subset R^n$, such that for any $v_1, v_2 \in V$, $v_1 H_g(z) v_2 = 0$.

Let $v_0,\ldots,v_{n-q}$ be the vertices of the regular simplex in V, centered at $0 \in V \in R^n$, $v_i = (v_i^1,\ldots,v_i^n)$. For $\rho > 0$ denote by

$z_i(\rho)$, $i = 0,\dots,n-q$, the points $z_i(\rho) = \rho v_i$.

Remind that $t_0 \in T^m$ is one of the values of the parameter t, for which ℓ_{t_0} is the support linear function of g at 0. Hence for any $i = 0,\dots,n-q$, and $\rho \geq 0$ we have

$$g(z_i(\rho)) \geq \ell_{t_0}(z_i(\rho)) \ ,$$

and by definition of the second differential,

$$(*) \quad g(z_i(\rho)) - \ell_{t_0}(z_i(\rho)) = o(\rho^2)$$

as $\rho \to 0$, since the points v_i belong to the null space of the Hessian $H_g(0)$.

Now we want to show that the derivatives of $\ell_i(z_i(\rho))$ with respect to t at t_0 are small. The following lemma can be proved by elementary computations:

<u>Lemma 2.8.</u> Let Ψ be a twice continuously differentiable function on the domain D in R^q, and let $\|d^2\Psi(y)\| \leq K$ for any $y \in D$. Let $M = \max\limits_{y \in D} \Psi(y)$.

Then for any $y \in D$, $\|d\Psi(y)\|^2 \leq 4K(M - \Psi(y))$, assuming that the distance of y to the boundary of D is at least $(\frac{M-\Psi(y)}{K})^{1/2}$.

Assume in additon that the point t_0 belongs to the interior of T^m . Considerations in the case $t_0 \in \partial T^m$ are completely similar.

Differentiating $\ell_t(x) = a_1(t)x_1+\dots+a_n(t)x_n + b(t)$ with respect to t we obtain:

$$d_t\ell_t(x) = da_1(t)x_1+\dots+da_n(t)x_n + db(t).$$

In particular, for $x = z = 0$ and for $t = t_0$ we obtain (since $\ell_t(0)$ attains its maximum at $t = t_0$)

$$d_t \ell_t(0)\,|_{t=t_0} = db(t_0) = 0.$$

At each point $z_i(\rho)$, $i = 0,\dots,n-q$, $\rho \geq 0$, we, therefore, have

$$d_t \ell_t(z_i(\rho))\,|_{t=t_0} = da_1(t_0)v_i^1\rho + \dots + da_n(t_0)v_i^n\rho \ .$$

Dividing these equations by ρ we get finally the following system of linear equations for $da_j(t_0)$:

$$(**)\quad v_i^1 da_1(t_0) + \dots + v_i^n da_n(t_0) = \frac{1}{\rho}\ \frac{d}{dt}\ \ell_t(z_i(\rho))\,|_{t=t_0} \ ,$$

$$i = 0,1,\dots,n-q \ .$$

Notice that the left hand side of this system does not depend on ρ . To estimate the right hand side, we apply lemma 2.8 to the function $\Psi(t) = \ell_t(z_i(\rho))$ on T^m at the point t_0 .

Since the family ℓ_t by assumptions is twice continuously differentiable on a compact manifold T^m , we can assume that $\|d^2\Psi\| \leq K$. $M = \max \Psi$ is in our case $g(z_i(\rho))$.

Now by inequality (*) above and since t_0 is an inner point of T^m, the conditions of lemma 2.8 are satisfied for ρ sufficiently small. We obtain:

$$\left\| \frac{d}{dt}\ \ell_t(z_i(\rho))\,|_{t=t_0} \right\| \leq 2(K|g(z_i(\rho)) - \ell_{t_0}(z_i(\rho))|)^{1/2} =$$

$$= o(\rho),\quad i = 0,\dots,n-q, \text{ by the inequality } (*).$$

Hence $\frac{1}{\rho} \frac{d}{dt} \ell_t(z_i(\rho))|_{t=t_0} \to 0$ as $\rho \to 0$, and passing to the limit in the system (**), we finally obtain the following linear system:

$$v_i^1 da_1(t_0) + \ldots + v_i^n da_n(t_0) = 0 \ , \ i = 0,\ldots,n-q.$$

Since $v_i = (v_i^1,\ldots,v_i^n)$ are the vertices of a nondegenerate simplex, the rank of this system is $n-q$, and hence the rank of $d\phi(t_0) = (da_1(t_0),\ldots da_n(t_0))$ is at most q.

Lemma 2.7 is proved.

The criterion of representability, given by theorem 2.6, concernes the behavior of the second order derivatives. Our next result gives the "first order" criterion.

<u>Definition 2.9</u> Let f be a continuous function on I^n. The point $x \in I^n$ is called a critical point of f if the first differential $df(x)$ exists and is equal to zero. The values of f at all the critical points form the set $\Delta(f) \subset R$ of critical values of f.

Clearly, any convex function can have at most one critical value. Thus to get nontrivial restriction for a representability of a given convex function g, we first "destroy" its convexity by substracting some auxiliary smooth function u.

<u>Theorem 2.10</u> Let a convex function g on I^n belong to $Q^k_{n,m}$. Then for any k times continuously differentiable function u on I^n,

$$\dim_e \Delta(g-u) \leq \frac{n+m}{k} \quad .$$

Proof. By proposition 2.2, $g-u \in S^k_{n,m}$. Then by theorem 3.1 [7], $\dim_e \Delta(g-u) \leq \frac{n+m}{k}$.

Remark 1. The criterion of theorem 2.10 takes into account only the "smooth" critical points and values of $g-u$. Using the version of the Morse-Sard theorem for maximum functions, proved in [8], one can give similar bounds for the entropy dimension of "nonsmooth" critical values of $g-u$.

Remark 2. Both the results of theorem 2.6 and theorem 2.10 remain valid also in the case where the parameter manifold is not compact, with the only alteration: the entropy dimension should be replaced by the Hausdorff dimension.

3. Some examples

The criteria of representability of convex functions as maxima of linear families, given by theorem 2.6 and 2.10, respectively, are independent (see examples below). They also differ in their applicability. The first one is usually easier to check, but the construction of examples, based on this criterion, requires some accuracy. The second criterion allows to construct immediately a lot of examples of convex functions, not representable as a maximum of a "too smooth" linear family: it is enough to take any function on I^n with a "big" set of critical values, and to make it convex, adding $K\|x\|^2$ with sufficiently big K. However, the vertication of this criterion in concrete situations can be rather difficult.

The following propositions and examples make these remarks more precise.

Proposition 3.1 For any countable bounded set $\Sigma \subset R^n$ there exists a convex function g on I^n with $\sigma_0(g)$ containing Σ .

Moreover, if the cond $\Sigma \subset \partial$ conv Σ, i.e. if all the condensation points of Σ belong to the boundary of the convex hull of Σ, then there exists a convex function g on I^n, for which $\sigma_0(g) = \Sigma$, $\sigma_q(g) \setminus \sigma_0(g) = \emptyset$, $q = 1,\ldots,n$.

Proof. For a given bounded countable set $\Sigma \subset R^n$ one can easily construct by induction a piecewise linear convex function g on I^n , with a countable number of linearity domains, such that for any $v \in \Sigma$, grad $g(x) = v$ at any inverse point x of one of the linearity domains of g. Since $d^2g = 0$ at any such x, we obtain $\Sigma \subset \sigma_0(g)$.

If $\text{cond}\,\Sigma \subset \partial\,\text{conv}\Sigma$, then g can be constructed in such a way, that the linearity domains of g condensate only to the boundary of I^n.

Therefore, the inner points of I^n, where d^2g exists, are exactly the inner points of the linearity domains of g. By definition we obtain $\sigma_0(g) = \Sigma$, $\sigma_q(g) \setminus \sigma_0(g) = \emptyset$ for $q = 1,\ldots,n$.

<u>Remark.</u> In the last case, i.e. for $\text{cond}\,\Sigma \subset \partial\,\text{conv}\,\Sigma$, we can easily make our function g C^∞-smooth in the inferior of I^n.

<u>Corollary 3.2.</u> For any $q = 0,1,\ldots,n$, and for any countable bounded set $\Sigma \subset R^{n-q}$ with $\text{cond}\,\Sigma \subset \partial\,\text{conv}\,\Sigma$, there exist a convex function g on I^n, with

$$\sigma_p(g) = \emptyset \quad , \qquad p < q ,$$
$$\sigma_p(g) = \Sigma \times I^q \subset R^{n-q} \times R^q = R^n ,$$
$$\sigma_p(g) = \sigma_q(g) \quad , p \geq q .$$

<u>Proof.</u> First we find, using proposition 3.1, the convex function $\tilde{g}$ on I^{n-q}, with $\sigma_0(\tilde{g}) = \Sigma$. Now we define g on $I^n = I^{n-q} \times I^q$ as $g(x,y) = \tilde{g}(x) + \frac{1}{2}\|y\|^2$, $x \in I^{n-q}$, $y \in I^q$.

Computations of the second differential of g show immediately that g has the required properties.

Now for any α , $0 \le \alpha < p$, one can easily find bounded countable sets $\Sigma_\alpha \subset R^p$ with $\dim_e \Sigma_\alpha = \alpha$. Let e.g. $A_\beta \subset [0,1]$ be the set $\{1, \frac{1}{2^\beta}, \ldots, \frac{1}{r^\beta}, \ldots\}$. $\dim_e A_\beta = \frac{1}{1+\beta}$ (see e.g. [3]). Now taking $\Sigma_\alpha = A_{\beta'}^p \subset R^p$, with $\beta' = \frac{p}{\alpha} - 1$, we have $\dim_e \Sigma_\alpha = \frac{p}{1+\beta'} = \alpha$. Clearly, cond $\Sigma_\alpha \subset \partial$ conv Σ_α. One also has $\dim_e \Sigma_\alpha \times I^r = \alpha + r$.

Hence the proposition 3.1 and corolary 3.2 show, that the invariant $\dim_e \sigma_q(g)$, $q = 0,1,\ldots,n$, are independent in the following sense: for any $q = 0,1,\ldots,n$ and α, $q \le \alpha < n$, one can find a convex function $g_{q,\alpha}$ on I^n with $\dim_e \sigma_p(g_{q,\alpha}) = 0$, $p < q$, $\dim_e \sigma_p(g_{q,\alpha}) = \alpha$, $p \ge q$.

By theorem 2.6, $g_{q,\alpha} \notin Q^k_{n,m}$ for $q' + \frac{m-q'}{k} < \alpha$, or for $k > \frac{m-q'}{\alpha - q'}$, where $q' = \min(m,q)$.

Now we turn to applications of theorem 2.10.

<u>Proposition 3.3.</u> Let φ be a ℓ times continuously differentiable function on $I^n, \ell \ge 2$, and let $\dim_e \Delta(\varphi) = \alpha$. Then for any sufficientl big K , the function $g = \varphi + K \|x\|^2$ is convex and ℓ-smooth and $g \notin Q^k_{n,m}$ for any k and m, such that $\frac{n+m}{k} < \alpha$.

<u>Proof.</u> We have $\varphi = g - K\|x\|^2$, and by theorem 2.10, $\alpha = \dim_e \Delta(\varphi) \le \frac{n+m}{k}$, if $g \in Q^k_{n,m}$.

As an example consider the Whitney function $\varphi_n : I^n \to R$, where φ_n is C^{n-1}-smooth on I^n and $\Delta(\varphi_n) = [0,1]$. (See [4]).

Since $\dim_e[0,1] = 1$, we obtain the following:

Corollary 3.4. For any $n \geq 3$ and for any sufficiently big K, the convex function $g_n = \varphi_n + K\|x\|^2$ does not belong to $Q^k_{n,m}$, if $k > n+m$.

The sensibility of the entropy dimension to countable sets allows also here to give examples of very simple and "nice" convex functions, not representable as maxima of "too smooth" families.

Corollary 3.5. Let for $s = 4,5,\ldots$ Ψ_s be the convex function on $[0,1]$, defined by $\Psi_s(x) = x^s \cos(\frac{1}{x}) + s^2x^2$. Then $\Psi_s \notin Q^k_{1,m}$ for $k > (s+1)(m+1)$.

Proof. One can easily check that $\dim_e \Delta(x^s \cos(\frac{1}{x})) = \frac{1}{s+1}$.

Notice that Ψ_s is an $[s/2]-1$ -smooth function on $[0,1]$ and it is real-analytic on $(0,1]$.

Now we consider two examples, showing that the criteria of representability, given by theorem 2.6 and 2.10, are independent.

First of all, corollary 3.5 gives an example of the situation, where theorem 2.10 works, and theorem 2.6 gives not restrictions. Indeed, by definition, $\Psi''_s(x) > 0$ for any $x \in [0,1]$. Hence $\sigma_0(\Psi_s) = \emptyset$. But the bound for the entropy dimension $\dim_e \sigma_1(\Psi_s) \leq 1$ is apriori satisfied, since $\sigma_1(\Psi_s) \subset R$.

Construction of opposite examples is more complicated. Denote by W the set in the plane x,y, defined by the inequalities $0 \le x \le 1$, $\sqrt{x} \le y \le 2\sqrt{x}$ (See fig. 1).

Let $\Delta = \{1 = \sigma_0 > \sigma_1 > \sigma_2 > \dots > 0\}$ be some sequence, converging to zero. Let ℓ_i be the horizontal line $\{y = \sigma_i\}$, and let $z_i = (\sigma_i^2, \sigma_i)$ be the intersection point of ℓ_i with the curve $y = \sqrt{x}$.

Let also ℓ_i' be the vertical line $\{x = \sigma_i^2\}$. We consider the function h_Δ on $[0,1]$, defined as follows:

1. If the lines ℓ_i and ℓ_{i+1}' intersect inside W, then $h_\Delta = \sigma_i$ on $(\sigma_{i+1}^2, \sigma_i^2]$.

2. If the lines ℓ_i and ℓ_{i+1}' intersect outside W, then $h_\Delta = \sigma_i$ on $(\frac{\sigma_i^2}{4}, \sigma_i^2]$, and $h_\Delta = 2\sqrt{x}$ on $(\sigma_{i+1}^2, \frac{\sigma_i^2}{4}]$. (See Fig. 1).

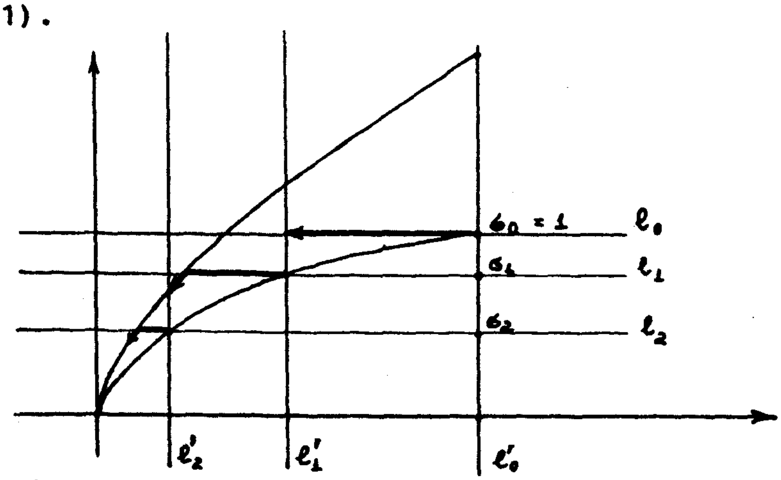

Figure 1

h_Δ is a monotone piecewise-continuous function. Let $g_\Delta(x) = \int_0^x h_\Delta(t)\,dt$, $x \in [0,1]$. Then g_Δ is a convex function, and by construction, $\sigma_0(g_\Delta) = \Delta$.

Hence theorem 2.6 gives the following restriction for representability of g_Δ :

$$g_\Delta \notin Q^k_{1,m} \quad \text{if} \quad k > \frac{m}{\dim_e \Delta}$$

Notice, that the sequence Δ can be chosen to have any entropy dimension between 0 and 1.

To show that the criterion of theorem 2.10 does not work for functions g_Δ , we need the following simple lemma:

<u>Lemma 3.6</u> Let u' be a continuously differentiable function on $[0,1]$. Then there is $\eta > 0$, such that the graph of u' does not intersect W for $0 \leq x \leq \eta$.

<u>Proof.</u> Assuming that the graph of u' intersects W for arbitrarily small x, we obtain immediately that $(u')'$ tends to infinity for some sequence of x_i , converging to 0.

Now fix some $k \geq 2$ and consider the function $g_\Delta - u$, where u is k times continuously differentiable on $[0,1]$. By lemma 3.6, the critical points of $g_\Delta - u$ belong to $[\eta,1]$ for some $\eta > 0$. But the interval $[\eta,1]$ can be subdivided into a finite number of subintervals, on which g_Δ is analytic, and hence $g_\Delta - u$ is k times smooth. By theorem 1.3 [5], the entropy dimension of the critical values of $g_\Delta - u$ on each of this intervals does not exceed $\frac{1}{k}$. Since the number of intervals is finite, we always have $\dim_e \Delta(g_\Delta - u) \leq \frac{1}{k}$. Thus the inequality of theorem 2.10,

$\dim_e \Delta(g_\Delta - u) \le \frac{1+m}{k}$, is always satisfied for any $k \ge 2$, $m = 1,2,\ldots,$ and a k-smooth u , and this theorem gives no restrictions for representability of g_Δ .

4. Some open questions.

We discuss in this section some additional properties of convex functions, representable as maxima of smooth linear families, which seem to be important in attempts to give necessary and sufficient conditions of such a representation.

Consider first of all the case of piecewise-linear convex functions g with a countable number of "faces". In this case it seems to be possible to give the necessary and sufficient conditions of representability in terms of the geometry of the set $\sigma_0(g)$ only. But the geometric invariant, more precise than the entropy dimension, should be considered.

Definition 4.1 (cf. [9]). Let for $\beta > 0$ and for $x_1, \ldots, x_p \in R^q$, $\rho_\beta(x_1, \ldots, x_p)$ be the [minimum of the] sum of β-th degrees of lengths of edges in all the polygonal lines, connecting $x_1, \ldots, x_p$. For a bounded subset $A \subset R^q$ define the β-spred of A,

$$V_\beta(A) \ , \text{ as } \ V_\beta(A) = \sup_{p, x_1, \ldots, x_p \in A} \rho_\beta(x_1, \ldots, x_p).$$

Some properties of the β-spread and the applications of this invariant are given in [9].

Conjecture 1. The piecewise-linear convex function g [with a countable number of fac…] on I^n belongs to $Q^k_{n,m}$ if and only if $V_{\frac{m}{k}}(\sigma_0(g)) < \infty$.

For convex functions of one variable and for $m = 1$ this conjecture is true ([7], theorem 4.4). For $n = 1$ and arbitrary m the condition $V_{\frac{m}{k}}(\sigma_0(g)) < \infty$ (which is stronger, than the condition $\dim_e \sigma_0(g) \le \frac{m}{k}$ of theorem 2.6) is necessary for $g \in Q^k_{1,m}$.

In contrast, for general convex function g the geometry of sets $\sigma_q(g)$, together with the structure of critical values of $g - u$, seems to be insufficient for complete characterizing the representability of g as maximum of linear family.

One additional property we discuss here, concerns the existence of high-order derivatives almost everywhere.

Definition 4.2. A continuous function f on R^n is said to have the k-th differential at $x_0 \in R^n$, if there exists a polynomial $P: R^n \to R$ of degree k, such that $|f(x) - P(x)| = o(\|x-x_0\|^k)$ (Compare definition 2.4 above).

By the Alexandrov-Buseman-Feller theorem any convex function has the second differential almost everywhere.

Conjeture 2. Any convex function $g \in Q^k_{n,m}$ has the k-th differential almost everywhere.

This conjecture is true in some special cases, say for functions in $Q^k_{1,1}$.

The last type of invariant, we mention here, concerns the geometry of nonsmoothness points, of convex functions considered. In [6] one such an invariant was considered, which was defined, roughly, as the integral along the "edges" of the absolute value of the jump of the gradient of our function. This ivariant was shown to be effectively bounded for functions, representable as the maximum of twice differentiable families.

Conjecturally, for functions from $Q^k_{n,m}$ we can integrate along the edges some power α of the jump of the gradient, where $\alpha < 1$ depends on n,m,k, and still the integral converges.

Once more, in some special cases, e.g. for piecewise-linear convex functions of one variable, this conjecture is true.

REFERENCES

1. R.T. Rockafeller, Favorable classes of Lipschitz continuous functions in subgradient optimigation in "Progress in nondifferentiable optimization", E. Nurminsky, ed., Pergamon Press, 1981.

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

3. C. Tricot, Douse définitions de la densité logarithmique, C. R. Acad. Sc. Paris, t. 293, série I, 1981, 549-552.

4. M. Whitney, A function non constant on a connected set of critical points, Duke Math. J. 1 (1935), 514-517.

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

6. Y. Yomdin, On functions, representable as a supremum of a family of smooth functions, SIAM J. Math. Anal. 14, 2, 1983, 239-246.

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

8. Y. Yomdin, Maxima of smooth families III: Morse-Sard theorem, to appear.

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