Kodai Mathematical Journal **15** (1992), 354–374

Regularity classes for operations in convexity theory

Christer O. Kiselman

Contents:

1. Introduction
2. Local comparison of functions
3. Imposing conditions on the gradient
4. Marginal functions
5. The convex envelope
6. Regularity of sets
7. Dual regularity classes
References

Resumo: *Glatecaj klasoj por operacioj en la teorio pri konvekseco*
Ni enkondukas klasojn de glateco kiuj estas adaptitaj al la plej gravaj operacioj en la teorio pri konvekseco. Plej ofte ili estas inter la klasoj C^1 kaj C^2.

Abstract: We introduce regularity classes which are adapted to the most important operations in convexity theory. They are typically between C^1 and C^2.

1. Introduction

The convex hull of a smoothly bounded set in $\mathbf{R}^n$ does not necessarily have a boundary of class C^2. This elementary observation is the origin of the present paper. We ask what regularity such a convex hull must have, and we construct regularity classes which are adapted to the operation $A \mapsto \operatorname{cvx} A$ of taking the convex hull of a set in $\mathbf{R}^n$, as well as to other operations which are of interest in convexity theory: that of forming the vector sum $A+B$ of two sets A and B and that of projecting a convex set into a space of lower dimension. In all these cases, we reduce the question of regularity to that of a marginal function $f(x) = \inf_y F(x,y)$.

The simple example with A as the union of two disjoint Euclidean balls shows that the convex hull $\operatorname{cvx} A$ need not have C^2 boundary, but it is easy to see that the boundary in this case is of Hölder class $C^{1,1}$ (i.e., the boundary is described by a function whose derivative is Lipschitz continuous). Our regularity classes are generalizations of this Hölder class.

To describe the simplest case of our results, let A be a compact set in $\mathbf{R}^n$. If the boundary of A is of class $C^{1,1}$, then A is a union of Euclidean balls with radii bounded from below; if A is convex, the converse holds. The property of being such a union of balls is easily seen to be stable under the operation $A \mapsto \operatorname{cvx} A$. Thus we see that ∂A of class $C^{1,1}$ implies $\partial \operatorname{cvx} A$ in the same class.

Given two compact convex sets A and B in $\mathbf{R}^n$, $n \geqslant 2$, with boundary of class C^k, $k = 1, 2, 3, \ldots$, we may ask whether the boundary of their vector sum $A+B$

This work was partially supported by the Swedish Natural Science Research Council.

is also of class C^k. The answer is in the affirmative if and only if $k = 1$ or $n = 2$ and $k = 2, 3, 4$. The result for $k = 1$, n arbitrary, is an easy exercise if you recall that a convex set has a C^1 boundary if and only if it possesses a unique tangent plane at every boundary point. For $n = 2$, the answer is in the affirmative for $k = 2$ (an easy exercise), for $k = 3$ (also an exercise), and for $k = 4$ (not at all easy; see Boman [1990a]). For $k = 5$ the answer is in the negative: there exist two compact convex sets in $\mathbf{R}^2$ with C^∞ boundary such that $\partial(A + B)$ is not of class $C^{4+\varepsilon}$ for any $\varepsilon > 0$ (Boman [1990a]). However, these sets possess infinitely flat points, so one may ask what the regularity is in case ∂A and ∂B are smooth and do not have infinitely flat points. The answer is $C^{20/3}$. More generally, if finitely many convex sets A_j are given in $\mathbf{R}^2$ with C^∞ boundaries without infinitely flat points, then $\partial(A_1 + \cdots + A_k)$ can be described in a local coordinate system either by a straight line or by a function $h(x) = x^{p+1}g(x, x^{2/m})$, where g is a C^∞ function of two variables satisfying $g(0, 0) > 0$, and where (p, m) is one of the pairs

$$(1,1), (3,1), (5,3), (7,15), (9,105), (11,315), (13,3465), (15,45\,045), (17,45\,045), \ldots$$

(Kiselman [1987, Theorem 5.4]). Such a function h is of Hölder class $C^{p+1+2/m}$ at least. In the case $(p, m) = (5, 3)$, which is the first class not contained in C^∞, the Hölder class $C^{p+1+2/m} = C^{20/3}$ is optimal.

In higher dimensions the regularity drops considerably: there exist two compact convex sets in $\mathbf{R}^3$ with boundary of class C^∞ such that $\partial(A + B)$ is not of class C^2 (Kiselman [1986, Theorem 3.4]). However, again these sets have infinitely flat points: is it true that $\partial(A + B)$ is of class C^2 if ∂A and ∂B are real analytic? The answer is no in $\mathbf{R}^n$ when $n \geqslant 4$ (Boman [1990b]) and unknown in $\mathbf{R}^3$. Therefore, looking for regularity results which are true in all dimensions, we must accept weaker regularity than C^2.

What about projections $\mathbf{R}^{n+m} \to \mathbf{R}^n$? If A is a compact convex set in $\mathbf{R}^3$, then its projection in $\mathbf{R}^2$ is of Hölder class $C^{2+\varepsilon}$ for some $\varepsilon > 0$ if ∂A is real analytic, but not necessarily of class C^2 if ∂A is of class C^∞ (Kiselman [1986, Theorems 2.2 and 3.4]). Boman's result about $\partial(A + B)$ in $\mathbf{R}^4$ not being of class C^2 shows that there is a compact convex set in $\mathbf{R}^7$ with real-analytic boundary whose projection in $\mathbf{R}^4$ has boundary not of class C^2.

Again, the conclusion is that to find regularity results which are true in all dimensions, we must look below C^2. The purpose of this paper is to study regularity classes of convex sets and functions which appear naturally and are independent of the dimension. They are typically between C^1 and C^2, in contrast to the classes that appear in low dimensions.

Of earlier work concerning dimension-independent regularity, let us mention that of Krantz & Parks [1991], where the authors prove that the boundary of the sum of finitely many compact convex sets is of Hölder class $C^{1,\varepsilon}$ if one of the sets has this regularity, and that of Griewank & Rabier [1990], who prove that the convex envelope of a function in $C^{1,\varepsilon}$ is in the same class, under a kind of coerciveness hypothesis. The present paper arose out of an attempt to simplify the approach of Krantz & Parks, whereas the work by Griewank & Rabier was unknown to the author until the paper was more or less completed. However, it unifies, generalizes, and perhaps even simplifies the earlier results, so maybe publication is nevertheless justified.

All our classes appear in two versions: with and without local uniformity in the constants. Under the Fenchel transformation a curious difference between these appear. In the last section we make a short study of this kind of duality. There is some relation here with the study by Azé & Penot [ms].

I am grateful to Jan Boman for discussions and helpful comments on earlier versions of this paper. I would also like to thank the referee for an unusually careful reading of the manuscript, thus eliminating several inaccuracies.

2. Local comparison of functions

Let h be a function on $\mathbf{R}^n$ with values in $[0,+\infty]$. We want to consider functions f such that $x \mapsto f(a+x)$ can be estimated from above by an affine function plus $h(x)$ near an arbitrary point a. We also want the comparisons to be stable under a change of scale in the independent as well as in the dependent variable; in other words, we compare $f(a+x)$ not only with $f(a)+L(x)+h(x)$ but with $f(a)+L(x)+Ch(cx)$ for all choices of positive constants c and C. We are thus led to the following definition.

Definition 2.1. *Let a function $h\colon \mathbf{R}^n \to [0,+\infty]$ be given. We shall say that a function $f\colon \mathbf{R}^n \to [-\infty,+\infty]$ is* ***locally controlled by*** *h* ***from above*** *if for every point $a \in \mathbf{R}^n$ such that $f(a) \in \mathbf{R}$ there exist constants $c = c_a$ and $C = C_a$ and a linear function L such that*

$$f(a+x) \leqslant f(a) + L(x) + Ch(cx) \qquad \text{for all } x \in \mathbf{R}^n. \tag{2.1}$$

We shall write $B^h = B^h(\mathbf{R}^n)$ for the set of all f which are locally controlled by h from above.

The sum in the right-hand side of (2.1) is a well-defined element of $]-\infty,+\infty]$, for at most one term, $Ch(cx)$, is infinite. We impose no condition if $f(a) = \pm\infty$; in particular functions with only infinite values belong to every class B^h.

Although Definition 2.1 has a sense for an arbitrary function h, we shall often have reason to impose some condition on it. A weak and reasonable such condition is

(2.2)

There is a constant M such that $0 \leqslant h(tx) \leqslant Mh(x)$ when $0 \leqslant t \leqslant 1$, $x \in \mathbf{R}^n$.

A stronger condition is:

$$0 \leqslant h(tx) \leqslant th(x) \text{ for } 0 \leqslant t \leqslant 1,\ x \in \mathbf{R}^n. \tag{2.3}$$

Sometimes we will need

$$0 \leqslant h(x) = o(|x|) \text{ as } x \to 0, \tag{2.4}$$

which is weaker than

$$h \text{ is convex, } h(0) = 0, \text{ and } h'(0) = 0. \tag{2.5}$$

A typical case, important in geometry, is when $h(x) = |x|^2$ for $|x| \leqslant 1$, $h(x) = +\infty$ when $|x| > 1$.

If h is identically zero, then a function with values in $[-\infty, +\infty]$ and with at least one finite value belongs to B^h if and only if it is concave, does not take the value $+\infty$, and admits a supergradient at every point where it has a finite value.

Functions in B^h need not be differentiable, even in the case $h(x) = |x|^2$, for if $f, g \in B^h$, then also $\min(f, g) \in B^h$. In particular all real-valued concave functions are in B^h. Thus in general the linear function L is not unique. However, if f and h are differentiable and $h(0) = 0$, then L is uniquely determined and must be $L(x) = f'(a) \cdot x$, so that (2.1) can be written

$$f(a + x) \leqslant f(a) + f'(a) \cdot x + Ch(cx). \tag{2.6}$$

If h is differentiable with $h(0) = 0$, then real-valued convex functions in B^h are differentiable, and for them (2.6) can be improved to

$$f(a) + f'(a) \cdot x \leqslant f(a + x) \leqslant f(a) + f'(a) \cdot x + Ch(cx). \tag{2.7}$$

Here the first inequality just expresses the fact that the graph of f lies above its tangent plane. More generally, if $h(x) = o(|x|)$ as $x \to 0$ and $f, -f \in B^h$, then f is differentiable and

$$f(a) + f'(a) \cdot x - C_1 h(c_1 x) \leqslant f(a + x) \leqslant f(a) + f'(a) \cdot x + C_2 h(c_2 x). \tag{2.8}$$

If (2.2) holds we can get $c_1 = c_2$ and $C_1 = C_2$ in (2.8).

It is sometimes of interest to have some kind of uniformity in the constants c and C in Definition 2.1. The following definition seems to work well in many cases.

Definition 2.2. *We shall say that f is* ***locally uniformly controlled by** h **from above*** *if for every $a_0 \in \mathbf{R}^n$ such that $f(a_0) \in \mathbf{R}$ there are constants c and C such that (2.1) holds for all a near a_0 with $f(a)$ finite. We shall write $G^h = G^h(\mathbf{R}^n)$ for the set of all f such that f is locally uniformly controlled by h from above.*

A weaker form of uniformity will be introduced in Section 4.

Example. It is easy to see that B^h and G^h are different in general. Let for example $h(x) = x^2$ for $x \in \mathbf{R}$, $|x| \leqslant 1$, and $h(x) = +\infty$ elsewhere. We can then just take $f(x) = -\sqrt{|x|}$, or $f(x) = -|x| + |x|^{3/2}$. But it does not help to require that f be in C^1: take $f(x) = x^3 \sin(1/x)$ for $x \neq 0$, $f(0) = 0$. Then $f \in B^h \setminus G^h$. We can take $c_0 = C_0 = 1$, but when $a = 1/(2\pi m - \pi/2)$, $c_a^2 C_a \to +\infty$ as $m \to +\infty$.

It is also easy to construct examples of real-valued convex functions which are in B^h but not in G^h:

Proposition 2.3. *Assume that $h\colon \mathbf{R}^n \to [0, +\infty]$ satisfies (2.5) and that $h \in B^h$. Assume also that $h(x) > 0$ for $x \neq 0$. Then there is a convex function in $B^h \setminus G^h$.*

Proof. Let us take a sequence $a_j \in \mathbf{R}^n$ with $a_j \neq 0$, $a_j \to 0$ and such that $h(a_j)$ is finite. Define f as the convex envelope of

$$x \mapsto \min\big[2h(x), \min_j \big[h(x) + h(j(x - a_j))\big]\big].$$

Using Theorem 5.3 below we conclude that $f \in B^h$ (here we need that $h \in B^h$). Now in a neighborhood of a_j we must have $f(x) = h(x) + h(j(x - a_j))$, at least for infinitely many j's. (Here we need $h(x) = o(x)$.) If f did belong to G^h we would get, for x close to a_j,

$$h(x) + h(j(x - a_j)) = f(x) \leqslant f(a_j) + L(x - a_j) + Ch(c(x - a_j)),$$

which implies $h(jx) \leqslant Ch(cx)$ for small x. Now the inequality $h(jx) \leqslant Ch(cx)$ can hold near the origin only if $c \geqslant j$ or $Cc \geqslant j$. In fact, if $c < j$, then the inequality implies

$$h(jx) \leqslant Ch(cx) \leqslant C(1 - c/j)h(0) + Cch(jx)/j = Cch(jx)/j,$$

which is possible only if $1 \leqslant Cc/j$, since $h(x) > 0$ for $x \neq 0$. Thus we cannot choose both c_a and C_a bounded: we must have $c_{a_j} \geqslant j$ or $c_{a_j} C_{a_j} \geqslant j$ for infinitely many j's.

3. Imposing conditions on the gradient

It turns out that for convex functions, the kind of one-sided regularity expressed by (2.1) imposes a regularity on the function's gradient. To prove this is the purpose of the present section.

Definition 3.1. *Let Ω be an open set in $\mathbf{R}^n$, and let $h\colon \mathbf{R}^n \to [0, +\infty]$ be $+\infty$ outside the unit ball. We define the class $B^{1,h}(\Omega)$ as the class of all continuously differentiable functions $f\colon \Omega \to \mathbf{R}$ such that for every point $a \in \Omega$ there are constants $c = c_a$ and $C = C_a$ such that*

$$|f'(a + x) - f'(a)||x| \leqslant Ch(cx) \qquad \text{for all } x \in \mathbf{R}^n. \tag{3.1}$$

(The constant c shall be so large that if $a + x \notin \Omega$, then $|cx| > 1$, making the condition void.) The class $G^{1,h}(\Omega)$ is defined in the same way with the extra requirement that the constants c_a and C_a shall be chosen as locally bounded functions of $a \in \Omega$. Finally we define $B^{1,h}_{\mathrm{dot}}(\Omega)$ and $G^{1,h}_{\mathrm{dot}}(\Omega)$ by replacing (3.1) with the weaker inequality

$$|(f'(a + x) - f'(a)) \cdot x| \leqslant Ch(cx) \qquad \text{for all } x \in \mathbf{R}^n. \tag{3.2}$$

The Hölder classes $C^{1,\varepsilon}$ with $0 < \varepsilon \leqslant 1$ are special cases of the regularity classes $G^{1,h}$, viz. with $h(x) = |x|^{1+\varepsilon}$ for $|x| \leqslant 1$, $h(x) = +\infty$ for $|x| > 1$.

That $B^{1,h}_{\mathrm{dot}}$ differs from $B^{1,h}$ follows from the example $h(x) = x_1^2 + |x_2|^{1+\varepsilon}$, where $0 < \varepsilon < 1$. Here $|(h'(a + x) - h'(a)) \cdot x| \leqslant 2h(x)$, so that $h \in B^{1,h}_{\mathrm{dot}}(\mathbf{R}^2)$, but the inequality $|h'(x)||x| \leqslant Ch(x)$ is impossible (try $x_2 = |x_1|^\alpha$ with $1 < \alpha < 1/\varepsilon$), proving that h cannot belong to $B^{1,h}$.

Theorem 3.2. *Let $h\colon \mathbf{R}^n \to [0, +\infty]$ satisfy (2.3), and assume that $h(x) = +\infty$ for $|x| > 1$. Then $B^{1,h}(\Omega) \subset B^{1,h}_{\mathrm{dot}}(\Omega) \subset B^h$ and $G^{1,h}(\Omega) \subset G^{1,h}_{\mathrm{dot}}(\Omega) \subset G^h$ for every open set Ω. (We define the functions as being $+\infty$ outside Ω.)*

In particular the hypothesis on h is satisfied if h is convex, $h \geqslant 0$ and $h(0) = 0$, for then $h(tx) \leqslant (1 - t)h(0) + th(x) = th(x)$.

Proof. If $f \in C^1(\Omega)$ we can write

$$f(a + x) - f(a) - f'(a) \cdot x = \int_0^1 (f'(a + tx) - f'(a)) \cdot x \, dt$$

provided $a \in \Omega$ and $|x|$ is so small that the whole segment $[a, a+x]$ is contained in Ω. If now $f \in B^{1,h}_{\rm dot}(\Omega)$ we can apply (3.2) at the point a and with x replaced by tx, $0 < t \leqslant 1$:

$$|(f'(a+tx) - f'(a)) \cdot x| \leqslant \frac{Ch(ctx)}{t} \leqslant Ch(cx),$$

where the last inequality follows from (2.3). Thus

$$|f(a+x) - f(a) - f'(a) \cdot x| \leqslant \int_0^1 |(f'(a+tx) - f'(a)) \cdot x| \, dt \leqslant \int_0^1 Ch(cx) \, dt = Ch(cx),$$

which shows that (2.1) holds.

Since we can control the constants c and C, this also proves that $G^{1,h}_{\rm dot} \subset G^h$. Theorem 3.2 is proved.

On the other hand, we cannot of course infer any regularity of f' from an inequality like (2.1) or (2.6). However, if we add convexity the situation is quite different; more generally we can assume that $-f$ belongs to some class G^h. Let us define $F(X)$ as the set of all functions $f\colon \mathbf{R}^n \to [-\infty, +\infty]$ which have real values in $X \subset \mathbf{R}^n$ and take the value $+\infty$ in $\mathbf{R}^n \setminus X$.

Theorem 3.3. *Let h and k be two functions satisfying (2.2) and (2.4) and taking the value $+\infty$ outside the unit ball. Then for every open set $\Omega \subset \mathbf{R}^n$ we have*

$$F(\Omega) \cap B^h \cap (-G^k) \subset B^{1,H}_{\rm dot}(\Omega) \qquad \text{and} \qquad F(\Omega) \cap G^h \cap (-G^k) \subset G^{1,H}_{\rm dot}(\Omega),$$

where $H(x) = h(x) + k(x) + k(-x)$. If h and k are radial (i. e., functions of $|x|$), then

$$F(\Omega) \cap B^h \cap (-G^k) \subset B^{1,h+k}(\Omega) \qquad \text{and} \qquad F(\Omega) \cap G^h \cap (-G^k) \subset G^{1,h+k}(\Omega).$$

It is not enough to assume that $\pm f \in B^h$ to conclude that $f \in B^{1,h}_{\rm dot}$; some uniformity is needed:

Example. Let $h(x) = x^2$ when $x \in \mathbf{R}$, $|x| \leqslant 1$, and $h(x) = +\infty$ outside this interval. Define $f(x) = x^2 \sin |x|^\alpha$, $x \neq 0$, $f(0) = 0$, with a number α satisfying $-1 < \alpha < 0$. Then $\pm f \in B^h$, and f' exists and is continuous. But $f \notin B^{1,h}(\mathbf{R}) = B^{1,h}_{\rm dot}(\mathbf{R})$.

Proof of Theorem 3.3. If $f \in B^h$, $-f \in B^k$, where $h(x), k(x) = o(|x|)$, then $f'(a)$ exists at every point a where f has a finite value; see (2.8). We want to prove (3.1) or (3.2) with h replaced by H, i. e., that

$$|f'(a+x) - f'(a)||x| \leqslant CH(cx) \tag{3.3}$$

or the weaker

$$|(f'(a+x) - f'(a)) \cdot x| \leqslant CH(cx) \tag{3.4}$$

for $a, a+x \in \Omega$. The hypotheses imply that for every $a \in \Omega$, there are constants c_a, C_a, d_a, D_a such that

$$\begin{aligned} f(a+x+y) - f(a) &\leqslant f'(a) \cdot (x+y) + C_a h(c_a(x+y)), \\ -f(a+x) + f(a) &\leqslant -f'(a) \cdot x + D_a k(d_a x), \\ -f(a+x+y) + f(a+x) &\leqslant -f'(a+x) \cdot y + D_{a+x} k(d_{a+x} y). \end{aligned}$$

Moreover d_{a+x} and D_{a+x} can be chosen to be bounded above for all $a+x$ near a, say for $x \in \omega$, ω a neighborhood of the origin. Adding these three inequalities we get

$$(3.5)\qquad (f'(a+x)-f'(a))\cdot y \leqslant C_a h(c_a(x+y)) + D_a k(d_a x) + D_{a+x} k(d_{a+x} y).$$

Now choose $y = \pm x$. This gives

$$\begin{aligned} (f'(a+x)-f'(a))\cdot x &\leqslant C_a h(2c_a x) + D_a k(d_a x) + D_{a+x} k(d_{a+x} x), \\ -(f'(a+x)-f'(a))\cdot x &\leqslant C_a h(0) + D_a k(d_a x) + D_{a+x} k(-d_{a+x} x), \end{aligned}$$

which we can simplify to

$$|(f'(a+x)-f'(a))\cdot x| \leqslant C'_a(h(c'_a x) + k(c'_a x) + k(-c'_a x)) = C'_a H(c'_a x),$$

thanks to the assumption (2.2), putting

$$c'_a = \max(2c_a, \sup_{x\in\omega} d_{a+x}) \quad \text{ and } \quad C'_a = M\max(C_a, D_a + \sup_{x\in\omega} D_{a+x}).$$

This is (3.4) with new constants. Note that we use c_a and C_a only at the point a, which is why we need not assume them to be locally bounded.

In the radial case we argue instead as follows. Choose $y = \lambda z = \lambda(f'(a+x) - f'(a))$ in (3.5). This gives

$$(3.6)\qquad \lambda|f'(a+x)-f'(a)|^2 \leqslant C_a h(c_a(x+\lambda z)) + D_a k(d_a x) + D_{a+x} k(d_{a+x}\lambda z).$$

If $z = f'(a+x) - f'(a) = 0$, (3.3) certainly holds. If not, we may choose $\lambda = |x|/|z|$, and we note that h is to be evaluated at the point

$$c_a(x+\lambda z) = c_a x + c_a \frac{|x|}{|z|} z,$$

whose distance to the origin is at most $2c_a|x|$. Thus (3.6) yields, since h and k are radial,

$$\begin{aligned} |f'(a+x)-f'(a)||x| = \lambda|z|^2 &\leqslant C_a h(2c_a x) + D_a k(d_a x) + D_{a+x} k(d_{a+x} x) \\ &\leqslant 2C'_a(h(c'_a x) + k(c'_a x)); \end{aligned}$$

this proves (3.3) (with new constants).

Since the argument keeps track of the constants c_a and C_a, it works also to prove that $f \in G^{1,H}_{\text{dot}}(\Omega)$, resp. $f \in G^{1,h+k}(\Omega)$, if we strengthen the hypothesis $f \in B^h$ to $f \in G^h$.

Corollary 3.4. *Let $f\colon \mathbf{R}^n \to]-\infty, +\infty]$ be convex, finite in a convex open set Ω and equal to plus infinity outside. Let h satisfy (2.3) and (2.4) and be $+\infty$ outside the unit ball. Then f belongs to $B^h(\mathbf{R}^n)$ if and only if f belongs to $B^{1,h}_{\text{dot}}(\Omega)$. Similarly, f belongs to $G^h(\mathbf{R}^n)$ if and only if f belongs to $G^{1,h}_{\text{dot}}(\Omega)$. If we assume in addition that h is radial, then $f \in B^h$ iff $f \in B^{1,h}(\Omega)$ iff $f \in B^{1,h}_{\text{dot}}(\Omega)$, and similarly for the G classes.*

Proof. Theorem 3.2 says that $B^{1,h}_{\text{dot}}(\Omega) \subset B^h$ and $G^{1,h}_{\text{dot}}(\Omega) \subset G^h$. For the other inclusions we just have to take $k = 0$ in Theorem 3.3 and note that $-f$ is in $G^k = G^0$ if f is convex, real-valued in Ω and $+\infty$ outside.

Every convex function of class C^1 belongs to some $G^{1,h}$ locally:

Proposition 3.5. *Let $f \in C^1(\Omega)$ be convex in a convex open set Ω. Then for every relatively compact convex open subset ω of Ω there exists a radial convex function H with $H \geqslant 0$, $H(0) = 0$ and $H'(0) = 0$ such that $f|_\omega \in G^{1,H}(\omega)$.*

Proof. The convexity gives

$$\begin{aligned} 0 \leqslant f(a+x) - f(a) - f'(a)\cdot x &\leqslant f'(a+x)\cdot x - f'(a)\cdot x \\ &\leqslant |f'(a+x) - f'(a)||x| \leqslant |x|g(|x|) \end{aligned}$$

for $a, a+x \in \omega$. Here g is the modulus of continuity of f' in ω. We now define

$$h(x) = \sup_a \bigl[f(a+x) - f(a) - f'(a)\cdot x;\ a, a+x \in \omega \bigr], \qquad x \in \mathbf{R}^n.$$

As a supremum of convex functions, h is convex, and $f|_\omega \in G^h$ by construction. Moreover, $h(x) \leqslant |x|g(|x|) = o(|x|)$, so Corollary 3.4 yields $f|_\omega \in G^{1,H}(\omega)$ if we define H as the smallest radial majorant of h.

4. Marginal functions

In this section we study the marginal function $f(x) = \inf_y F(x,y)$ of a function F defined in $\mathbf{R}^n \times \mathbf{R}^m$. To conclude anything about regularity, it is necessary to assume that the infimum is attained for every x. Therefore it is natural to assume that the partial functions $F(x,\cdot)$ are lower semicontinuous. But this is not enough: we need some kind of coerciveness. We shall use the following terminology:

Definition 4.1. *A function $f\colon Y \to [-\infty,+\infty]$, Y being a topological space, is said to be* **coercive in Y** *if the sublevel sets $\{y \in Y;\ f(y) < s\}$ are relatively compact in Y for all real numbers s. We shall say that a function $F\colon X \times Y \to [-\infty,+\infty]$ is* **coercive in the y direction** *if all partial functions $F(x,\cdot)$ are coercive in Y. Finally we shall say that F is* **locally uniformly coercive in the y direction**, *if for every $x_0 \in X$, X being another topological space, the sublevel sets $\{y \in Y; F(x,y) < s\}$ are contained in a fixed compact subset of Y for all x in some neighborhood of x_0.*

Theorem 4.2. *Let $H\colon \mathbf{R}^n \times \mathbf{R}^m \to [0+\infty]$ and $F\colon \mathbf{R}^n \times \mathbf{R}^m \to [-\infty,+\infty]$ be given, let*

$$f(x) = \inf_{y \in \mathbf{R}^m} F(x,y), \qquad x \in \mathbf{R}^n, \tag{4.1}$$

be the marginal function of F, and define $h(x) = H(x,0)$ for $x \in \mathbf{R}^n$. Assume that the infimum in (4.1) is attained for every point $x \in \mathbf{R}^n$. (In particular this is true if all partial functions $F(x,\cdot)$ are lower semicontinuous and F is coercive in the y direction.) Then $F \in B^H(\mathbf{R}^n \times \mathbf{R}^m)$ implies $f \in B^h(\mathbf{R}^n)$.

In applications of this theorem, most often h is given, and we define H on $\mathbf{R}^n \times \mathbf{R}^m$ by

$$H(x,y) = \begin{cases} h(x), & y = 0, \\ +\infty, & y \neq 0. \end{cases} \tag{4.2}$$

For such H, the assumption $F \in B^H$ is a condition only on the partial functions $F(\cdot, y)$.

Proof. To every given $a \in \mathbf{R}^n$ with $f(a) \in \mathbf{R}$ there is a point $b \in \mathbf{R}^m$ such that $f(a) = F(a, b)$. Then by definition there is a linear function L and constants $c = c_{(a,b)}$, $C = C_{(a,b)}$ such that

$$F(a + x, b + y) \leqslant F(a, b) + L(x, y) + CH(cx, cy).$$

In particular, taking $y = 0$,

$$F(a + x, b) \leqslant F(a, b) + L(x, 0) + CH(cx, 0) = f(a) + L(x, 0) + Ch(cx).$$

Now by the definition of f, $f(a + x) \leqslant F(a + x, b)$. This proves that $f \in B^h(\mathbf{R}^n)$.

We would like to have a result analogous to Theorem 4.2 but with uniformity. A somewhat weaker uniformity than that of the classes G^h appears naturally. The definition is as follows.

Definition 4.3. *Let $h: \mathbf{R}^n \to [0, +\infty]$ be given. A function $f: \mathbf{R}^n \to [-\infty, +\infty]$ is said to belong to $E^h(\mathbf{R}^n)$ if for every point $a_0 \in \mathbf{R}^n$ such that $f(a_0) \in \mathbf{R}$ and every number A there exist a neighborhood U of a_0 and constants c and C such that for every $a \in U$ with $f(a) \leqslant A$ there exists a linear function L such that (2.1) holds.*

Thus we now allow the constants c and C to blow up in a neighborhood of a_0, but only if $f(a)$ tends to $+\infty$.

Obviously $G^h \subset E^h \subset B^h$. A function in E^h which is locally bounded above belongs to G^h, and therefore $E^h = G^h$ if h is bounded above in a neighborhood of the origin. For this reason E^h can often be replaced by G^h, and it was not needed in the previous sections. However, in general the classes E^h and G^h do differ as shown by the following two examples.

Example. Define $h(x) = 0$ for $x \in \mathbf{R}$ satisfying $-1 \leqslant x \leqslant 0$, and $h(x) = +\infty$ otherwise. Let $f(x) = 1/x$, $x > 0$; $f(x) = 0$, $x \leqslant 0$. Then $f \in E^h \setminus G^h$.

There are also convex functions in $E^h \setminus G^h$, at least in three variables:

Example. Define $h(x, y, z) = z^2$ when $(x, y) = (0, 0)$, $|z| \leqslant 1$; $h(x, y, z) = +\infty$ otherwise. There is a convex function g in $\mathbf{R}^3$ which is C^∞ except at the origin, satisfies $0 \leqslant g(x, y, z) \leqslant x^2 + y^2 + z^2$, and equals $x^2 + y^2 + z^2$ outside small neighborhoods U_j of the points $a_j = (1/j, 1/j^3, 0)$, let us say $g(x, y, z) = x^2 + y^2 + z^2$ when $|(x, y, z) - a_j| \geqslant 1/j^3$. We require that the second derivative $|\partial^2 g(a_j)/\partial z^2|$ tend to $+\infty$ as $j \to +\infty$. Then $g \in B^h \setminus G^h$: if c_j and C_j are constants that can serve in (2.1) at the point a_j, then the product $C_j c_j^2$ must blow up. Now define $f(x, y, z) = g(x, y, z) + x^2/y$ when $y > 0$, $f(0, 0, z) = z^2$, and $f(x, y, z) = +\infty$ otherwise. Then $f \in E^h \setminus G^h$. The constants blow up at the origin all the same, but now f tends to $+\infty$ at the special points; in fact, when $f \leqslant A$, then $y \geqslant x^2/A$, so only finitely many neighborhoods U_j are involved.

Theorem 4.4. *Let H, F and f be as in Theorem 4.2. Assume in addition that (2.2) holds for H and that F is lower semicontinuous in $\mathbf{R}^n \times \mathbf{R}^m$ and locally uniformly coercive in the y direction. Then $F \in E^H(\mathbf{R}^n \times \mathbf{R}^m)$ implies $f \in E^h(\mathbf{R}^n)$. If moreover f is locally bounded above where it is less than $+\infty$, then $f \in G^h(\mathbf{R}^n)$. (In* *particular this holds if h is bounded above in a neighborhood of the origin, or if F is locally bounded above where it is less than $+\infty$.)*

Proof. Let $a_0 \in \mathbf{R}^n$ and $A \in \mathbf{R}$ be given. In view of the semicontinuity and the locally uniform coerciveness, there exists a compact neighborhood U of a_0 such that

$$K = \{(a, b);\ a \in U \text{ and } F(a, b) \leqslant A\}$$

is compact in $\mathbf{R}^n \times \mathbf{R}^m$. The hypothesis $F \in E^H$ and condition (2.2) imply the existence of constants c and C such that for every $(a, b) \in K$ with $F(a, b) \in \mathbf{R}$ there is a linear function L such that, for all $(x, y) \in \mathbf{R}^n \times \mathbf{R}^m$,

$$F(a + x, b + y) \leqslant F(a, b) + L(x, y) + CH(cx, cy). \tag{4.3}$$

For every $a \in U$ such that $-\infty < f(a) \leqslant A$ there exists b such that $(a, b) \in K$ and $F(a, b) = f(a)$; hence (4.3) shows that

$$f(a + x) \leqslant F(a + x, b) \leqslant F(a, b) + L(x, 0) + CH(cx, 0) = f(a) + L(x, 0) + Ch(cx).$$

This proves that $f \in E^h$.

Example. One can construct $F \in C^\infty(\mathbf{R}^2)$ such that $F(x, y) = y^2$ when $|xy| \leqslant \frac{1}{2}$ or $|xy| \geqslant 2$, and $F(x, y) \geqslant -\sqrt{|x|}$ everywhere, with equality when $|xy| = 1$. Then the marginal function is $f(x) = -\sqrt{|x|}$. So if we put $H(x, y) = |(x, y)|^2$ for $|(x, y)| \leqslant 1$ and $+\infty$ elsewhere, we see that $F \in G^H$, $f \in B^h \setminus E^h$. (Here F is coercive in the y direction without being uniformly so.)

Example. Put

$$F(x, y) = \begin{cases} y^2 + 1/(y^2 - 1/x^2), & x > 0, \quad |y| > 1/x, \\ +\infty, & x > 0, \quad |y| \leqslant 1/x, \\ y^2, & x \leqslant 0. \end{cases}$$

Then the marginal function is $f(x) = 2 + 1/x^2$ for $x > 0$, $f(x) = 0$ for $x \leqslant 0$. Define $H(x, y) = 0$ if $y = 0$ and $-1 \leqslant x \leqslant 0$, $H(x, y) = +\infty$ elsewhere. Then $F \in G^H$, $f \in E^h \setminus G^h$. (Here f is not bounded above near the origin; all other hypotheses of Theorem 4.4 are satisfied.)

The infimal convolution

$$f \mathbin{\square} g(x) = \inf_{y \in \mathbf{R}^n} \big[f(x - y) + g(y);\ f(x - y) < +\infty,\ g(y) < +\infty\big], \qquad x \in \mathbf{R}^n,$$

of two functions is a special kind of marginal function. We formulate the following particular case of Theorem 4.2 for them.

Theorem 4.5. *Let $f, g\colon \mathbf{R}^n \to \,]-\infty, +\infty]$ be two lower semicontinuous functions. Assume that one of them is coercive and the other is bounded below. If one of f, g is in $B^h(\mathbf{R}^n)$, then their infimal convolution $f \mathbin{\square} g \in B^h(\mathbf{R}^n)$. Similarly for E^h.*

Proof. The function $F(x, y) = f(x - y) + g(y)$ is lower semicontinuous and coercive in the y direction, so $\inf_y F(x, y)$ is attained. Define H by (4.2). Then F is in B^H

if $f \in B^h(\mathbf{R}^n)$. (If $g \in B^h$ we consider $f(y) + g(x-y)$ instead.) It follows from Theorem 4.2 that the marginal function of F is in B^h. Now this marginal function is exactly $f \mathbin{\square} g$. The result for E^h follows because under the hypothesis made, F is automatically locally uniformly coercive in the y direction.

Remark. Of course we actually need only that F is coercive in the y direction, or locally uniformly coercive in the y direction, respectively; the theorem describes a simple case when this is true.

Corollary 4.6. *Let $f, g\colon \mathbf{R}^n \to \,]-\infty, +\infty]$ be two lower semicontinuous convex functions which are coercive in $\mathbf{R}^n$ and assume that one of them is of class B^h for some radial function h satisfying (2.2) and (2.4) and being $+\infty$ outside the unit ball. Then $f \mathbin{\square} g \in B^{1,h}(\Omega)$, where Ω is the interior of the set where $f \mathbin{\square} g$ is finite. If in addition one of f, g is in E^h, then $f \mathbin{\square} g \in G^{1,h}(\Omega)$.*

Proof. We only need to remember that $f \mathbin{\square} g$ is convex and apply Theorem 3.3.

5. The convex envelope

Given a function $f\colon \mathbf{R}^n \to [-\infty, +\infty]$ we denote by $\operatorname{cvx} f$ the largest convex minorant of f. We can express $\operatorname{cvx} f$ as

(5.1)
$$\operatorname{cvx} f(x) = \inf \Big[\sum_1^p \lambda_j f(x_j);\ \lambda_j \geqslant 0,\ \sum_1^p \lambda_j = 1,\ \sum_1^p \lambda_j x_j = x,\ f(x_j) < +\infty \Big].$$

(We put $0 \cdot (-\infty) = 0$.) Here the infimum is taken over all $p \geqslant 1$, all numbers λ_j and all points $x_j \in \mathbf{R}^n$ satisfying the conditions indicated. Let us denote by $f_p(x)$ the infimum in (5.1) with p fixed. Then for every convex function $g \leqslant f$ we have

$$f = f_1 \geqslant f_2 \geqslant \cdots \geqslant f_p \geqslant f_{p+1} \geqslant \cdots \geqslant g.$$

It is also elementary to prove that for all t between 0 and 1 we have

$$f_{p+q}((1-t)x + ty) \leqslant (1-t) f_p(x) + t f_q(y).$$

This proves immediately that $f_\infty = \inf f_p = \lim f_p$ is equal to $\operatorname{cvx} f$. However, it is important here not to let p tend to infinity. And in fact the sequence (f_p) is stationary; more precisely $f_p = f_{n+1}$ for all $p \geqslant n+1$. In other words, the following lemma holds.

Lemma 5.1. *In (5.1) it is enough to take $p = n+1$.*

Proof. This follows from Carathéodory's theorem [1911: 200], but for completeness we shall include a proof. Assume first that $f > -\infty$ so that only finite values $f(x_j) \in \mathbf{R}$ appear in (5.1). If $p \geqslant n+2$ and p points $x_1, ..., x_p$ are given, there exist real numbers $\mu_1, ..., \mu_p$, not all zero, such that $\sum \mu_j = 0$ and $\sum \mu_j x_j = 0$. Hence in any linear combination $x = \sum \lambda_j x_j$ the numbers λ_j can be replaced by $\lambda_j(t) = \lambda_j + t\mu_j$ as long as the latter are nonnegative. We choose t as

$$t = \max_j \big[-\lambda_j/\mu_j;\ \mu_j > 0 \big] \leqslant 0 \quad \text{if} \quad \sum_1^p \mu_j f(x_j) \geqslant 0,$$

and

$$t = \min_j \big[-\lambda_j/\mu_j;\ \mu_j < 0 \big] \geqslant 0 \quad \text{if} \quad \sum_1^p \mu_j f(x_j) < 0.$$

In both cases $\lambda_j(t) \geqslant 0$ with equality for at least one index k. Then the number

$$\sum_1^p \lambda_j(t) f(x_j) = \sum_1^p \lambda_j f(x_j) + t \sum_1^p \mu_j f(x_j) \leqslant \sum_1^p \lambda_j f(x_j)$$

competes in the infimum (5.1), but now since $\lambda_k(t) = 0$, the point x_k is no longer needed. Thus we can eliminate one point at a time as long as $p \geqslant n+2$.

In the general case we apply this result to $f_c(x) = \max(f(x), c)$, $c \in \mathbf{R}$, and let $c \to -\infty$; then $f_c \to f$ and $\operatorname{cvx} f_c \to \operatorname{cvx} f$.

Simple examples like $1/(1+x^2)$ and $|x| + \sqrt{|x|}$ show that the infimum in (5.1) is not always attained, not even for coercive functions. However, if we add a strong coerciveness hypothesis this is true. Let us say that f ***grows faster than any linear function*** if $f - L$ is coercive in $\mathbf{R}^n$ for every linear function L. In particular this holds if f is $+\infty$ outside a bounded set.

Lemma 5.2. *Let $f\colon \mathbf{R}^n \to]-\infty, +\infty]$ be lower semicontinuous and assume that f grows faster than every linear function. Then the infimum in (5.1) is attained at every point x with $\operatorname{cvx} f(x) < +\infty$.*

The result in Lemma 5.2 does not hold if f assumes the value $-\infty$. We shall denote by $\operatorname{dom} f$ the set $\{x \in \mathbf{R}^n;\ f(x) < +\infty\}$ (the ***effective domain*** of f).

Proof. Suppose $\operatorname{cvx} f(x) < +\infty$ and take a hyperplane H in $\mathbf{R}^{n+1}$ which contains $(x, \operatorname{cvx} f(x))$ and has the graph of f on one side. If this hyperplane is vertical, say equal to $K \times \mathbf{R}$ where K is a hyperplane in $\mathbf{R}^n$, then $(\operatorname{cvx} f)|_K$ is equal to $\operatorname{cvx}(f|_K)$ and we can assume that the result is already proved by induction. If H is not vertical, it is equal to the graph of an affine function $C+L$, and we take sequences $(x_j^k)_{k\in\mathbf{N}}$ and $(\lambda_j^k)_{k\in\mathbf{N}}$, $j = 1, ..., n+1$, with $f(x_j^k) < +\infty$ such that $\sum \lambda_j^k f(x_j^k)$ tends to $\operatorname{cvx} f(x)$. By compactness the sequence $(\lambda_j^k)_k$ has an accumulation point λ_j. If $\lambda_j = 0$ we can say nothing about (x_j^k), but we simply discard it, since it does not contribute to the sum. If $\lambda_j > 0$ on the other hand, the sequence (x_j^k) must have an accumulation point, for the value of f at x_j^k must appoach that of $C + L$, and the coerciveness of $f - L$ keeps all the x_j^k in a bounded set. Let x_j be an accumulation point of $(x_j^k)_k$, the limit of a subsequence $(x_j^{k(i)})_i$. The semicontinuity yields $f(x_j) \leqslant \liminf f(x_j^{k(i)})$, which shows that $\sum \lambda_j f(x_j) \leqslant \operatorname{cvx} f(x)$, so that the infimum is attained.

Theorem 5.3. *Let $f \in B^h(\mathbf{R}^n)$ be lower semicontinuous and satisfy $-\infty < f \leqslant +\infty$. Assume that f grows faster than every linear function L. (In particular this is true if f is equal to plus infinity outside a bounded set.) Let h be any function with values in $[0, +\infty]$. Then $f \in B^h$ implies $\operatorname{cvx} f \in B^h$. If h is radial and satisfies (2.2) and (2.4), then $\operatorname{cvx} f \in B^{1,h}(\Omega)$, where Ω denotes the interior of $\operatorname{cvx} \operatorname{dom} f$.*

Proof. For every point $a \in \mathbf{R}^n$ with $\operatorname{cvx} f(a) < +\infty$ there are points a_j and numbers $\lambda_j \geqslant 0$, $j = 1, ..., n+1$, such that $\sum \lambda_j = 1$ and $g(a) = \operatorname{cvx} f(a) = \sum \lambda_j f(a_j)$

(Lemmas 5.1 and 5.2). By definition of the convex envelope, we have for any x,

$$g(a+x) = \operatorname{cvx} f(a+x) \leqslant \lambda_k f(a_k + y) + \sum_{j \neq k} \lambda_j f(a_j),$$

if we can choose y so that

$$a + x = \lambda_k(a_k + y) + \sum_{j \neq k} \lambda_j a_j = a + \lambda_k y.$$

This is possible if $\lambda_k > 0$: take $y = x/\lambda_k$. Using that f is controlled from above by h at the point a_k we see that

(5.2)

$$g(a+x) \leqslant \lambda_k(f(a_k) + L(y) + Ch(cy)) + \sum_{j \neq k} \lambda_j f(a_j) = g(a) + L(x) + \lambda_k Ch(cx/\lambda_k).$$

This proves that $g = \operatorname{cvx} f \in B^h$. The statement about $B^{1,h}(\Omega)$ follows from Theorem 3.3.

Remark. It is also possible to prove Theorem 5.3 using Theorem 4.2 if h satisfies (2.2). We take $\lambda_j \in \mathbf{R}$, $x_j \in \mathbf{R}^n$, and define

$$y = (\lambda_1, ..., \lambda_{n+1}, x_1, ..., x_{n+1}) \in \mathbf{R}^{(n+1)^2},$$

$$Y = \{y \in \mathbf{R}^{(n+1)^2};\ \lambda_j \geqslant 0,\ \sum \lambda_j = 1,\ \sum \lambda_j x_j = 0\},$$

and

$$F(x,y) = \begin{cases} \sum \lambda_j f(x + x_j), & x \in \mathbf{R}^n,\ y \in Y, \\ +\infty, & x \in \mathbf{R}^n,\ y \notin Y. \end{cases}$$

(Here as usual $0 \cdot (+\infty) = 0$.) We want to apply Theorem 4.2 to F, its marginal function (4.1) being $\operatorname{cvx} f$. Define H by (4.2). Then $F \in B^H(\mathbf{R}^n \times \mathbf{R}^{(n+1)^2})$ (here we need (2.2)). Clearly F is lower semicontinuous in all variables. But it is not necessarily coercive in the y direction as can be seen from simple examples (for $n = 1$, $F(x, 1, 0, 0, x_2) = f(x)$ does not tend to plus infinity as $x_2 \to \infty$). However, Lemma 5.2 is a satisfactory substitute for this, since all we need to know is that the infimum in $\operatorname{cvx} f(x) = \inf_y F(x,y)$ is attained. Therefore Theorem 4.2 applies, and we can conclude that $\operatorname{cvx} f \in B^h$.

The uniformity of the constants is preserved under the operation $f \mapsto \operatorname{cvx} f$ as shown by the following result.

Theorem 5.4. *Let $h\colon \mathbf{R}^n \to [0, +\infty]$ satisfy (2.2). Assume that $f\colon \mathbf{R}^n \to \,]-\infty, +\infty]$ is lower semicontinuous and grows faster than any linear function. Then $f \in E^h$ implies $\operatorname{cvx} f \in E^h$.*

Proof. It seems we cannot use Theorem 4.4 due to lack of coerciveness. We can instead argue as follows. Consider again (5.2), but take k this time so that $\lambda_k \geqslant 1/(n+1)$. There is at least one such index. The function f is bounded below, and it is no restriction to assume that $f \geqslant 0$. We claim that $f(a_k) \leqslant (n+1)g(a)$. Indeed, since $f \geqslant 0$,

$$\frac{1}{n+1} f(a_k) \leqslant \lambda_k f(a_k) \leqslant \sum_1^{n+1} \lambda_j f(a_j) = g(a).$$

Thus, when $g(a)$ is bounded above, f is bounded above in at least one of the points a_k. Let now a vary in a neighborhood of a fixed point a_0 under the restriction $g(a) \leqslant A$. Then with our choice of k (depending on a), $f(a_k) \leqslant (n+1)A$, so that a_k varies in a bounded set and, if $f \in E^h$, the constants c and C in (5.2) remain bounded. We now use (2.2) and the fact that there is a lower bound for λ_k. Thus (5.2) shows that g is controlled from above by h in a uniform way for all points a, in other words that $g \in E^h$.

Corollary 5.5. *Let $h\colon \mathbf{R}^n \to [0, +\infty]$ satisfy (2.2). Let $f\colon \mathbf{R}^n \to\ ]-\infty, +\infty]$ be lower semicontinuous and grow faster than any linear function. Assume in addition that* $\operatorname{dom} f$ *is open. Then $f \in E^h$ implies* $\operatorname{cvx} f \in G^h$. *If h is radial and satisfies (2.2) and (2.4) and is $+\infty$ outside the unit ball, then* $\operatorname{cvx} f \in G^{1,h}(\Omega)$, *where* $\Omega = \operatorname{cvx} \operatorname{dom} f = \operatorname{dom} \operatorname{cvx} f$.

Proof. When $\operatorname{dom} f$ is open, then also $\operatorname{cvx} \operatorname{dom} f = \operatorname{dom} \operatorname{cvx} f$ is open, so $\operatorname{cvx} f$ is locally bounded above there. This proves the corollary. We note that if h is bounded above in a neighborhood of the origin, then $f \in B^h$ implies that $\operatorname{dom} f$ is open.

Example. If the boundary of Ω is not smooth, then its lack of smoothness can be inherited by the convex envelope of a smooth function. Let us take for example a convex domain

$$\Omega = \{x \in \mathbf{R}^n;\ |x_n| + g(x_1, ..., x_{n-1}) < 0\},$$

where g is a convex function which is negative in some subset of $\mathbf{R}^{n-1}$, and let us define $f(x) = -|x_n|$ in Ω, and as $+\infty$ outside. Then $\operatorname{cvx} f(x) = g(x_1, ..., x_{n-1})$ in the set where g is negative. If we choose g and h so that $g \notin B^h$ we therefore get $\operatorname{cvx} f \notin B^h$; similarly for G^h and E^h. But $f \in G^h$ for every $h \geqslant 0$. A slight modification gives $f \in C^\infty(\Omega)$. (In this case f is not lower semicontinuous in $\mathbf{R}^n$.)

Example. It may of course happen, even if the boundary of Ω is smooth, that the constants c_a and C_a degenerate as $a \to \partial\Omega$. For $\Omega = \mathbf{R}^2$, we let φ be a test function which is equal to 1 near the origin and satisfies $0 \leqslant \varphi \leqslant 1$. Then

$$f(x) = (1 - \varphi(x_1x_2))|x_1|^{3/2} + \varphi(x_1x_2)(1 + x_1^2), \qquad x \in \mathbf{R}^2,$$

is in $C^\infty(\mathbf{R}^2)$, so in particular $f \in G^{1,h}(\mathbf{R}^2)$, putting $h(x) = |x|^2$ for $|x| \leqslant 1$ and equal to $+\infty$ elsewhere. But $\operatorname{cvx} f(x) = |x_1|^{3/2}$ which is not in B^h. (The function is not coercive.) To get a similar example with Ω as the unit ball in $\mathbf{R}^n$, we can take $\varphi \in C^\infty(\mathbf{R})$, $0 \leqslant \varphi \leqslant 1$, with $\varphi(t) = 0$ for $t \leqslant 1$ and $\varphi(t) = 1$ for $t \geqslant 2$. Then

$$f(x) = \begin{cases} (1 - |x_1|^{3/2})\varphi\big((1 - |x|^2)/x_1^2\big) + |x_1|^{3/2}, & |x| < 1,\ x_1 \neq 0, \\ 1, & |x| < 1,\ x_1 = 0, \end{cases}$$

is C^∞ in the open unit ball, and as before $\operatorname{cvx} f(x) = |x_1|^{3/2}$. (Here the infimum in (5.1) is attained for every x with $x_1 \neq 0$, although the function is not lower semicontinuous as a function in $\mathbf{R}^n$, nor coercive in Ω.)

6. Regularity of sets

If f is a function of $n-1$ variables in a regularity class B^h, we would like to say that its epigraph

$$\operatorname{epi} f = \{x \in \mathbf{R}^n;\ f(x_1, ..., x_{n-1}) \leqslant x_n\}$$

is a set of class B^h. This notion should be invariant under change of coordinates. A convenient class of sets are those with C^1 boundary, and then the admissible coordinate systems are those which do not have the x_n-axis (the line $x_1 = x_2 = \cdots = x_{n-1} = 0$) in the tangent plane. Now it is easy to see that if h is radial and satisfies (2.2), the condition on the epigraph to belong to the class B^h is the same for all such coordinate systems. (It scarcely has a sense to define regularity classes of sets for nonradial h.)

Proposition 6.1. *Let $h\colon \mathbf{R}^{n-1} \to [0, +\infty]$ be radial, bounded in some neighborhood of the origin, and equal to $+\infty$ outside some neighborhood of the origin. Assume that h satisfies (2.2). Let f be a Lipschitz function with $f(0) = 0$. Consider the epigraph $\operatorname{epi} f$ in $\mathbf{R}^n$ and make a linear mapping T of $\mathbf{R}^n$ onto itself such that, for some neighborhood ω of the origin, $T(\operatorname{epi} f \cap \omega)$ agrees with the epigraph of a Lipschitz function g in a neighborhood of the origin. Then $f \in B^h$ if and only if $g \in B^h$. Similarly for G^h.*

Proof. This is really obvious. The assumptions on h are natural since we are interested in local regularity properties.

Definition 6.2. *Let $h\colon \mathbf{R}^{n-1} \to [0, +\infty]$ be radial, equal to $+\infty$ outside some neighborhood of the origin and satisfy (2.2) and (2.4). Let A be a set with C^1 boundary in $\mathbf{R}^n$. Then we shall say that A is of class B^h or G^h if, in an admissible coordinate system, the function whose epigraph agrees with A near an arbitrarily given point of ∂A is of the same class as a function of $n-1$ variables.*

Theorem 6.3. *Let A be a compact set in $\mathbf{R}^n \times \mathbf{R}^m$ of regularity class B^H for some function H in $\mathbf{R}^{n+m-1}$ satisfying the properties of Definition 6.2. Assume that its projection $\pi(A)$ in $\mathbf{R}^n$ has a C^1 boundary (in particular this is true if A is convex). Then $\pi(A)$ is of class B^h, where $h(x_1, ..., x_{n-1}) = H(x_1, ..., x_{n-1}, 0, 0, ..., 0)$. Similarly for G^h.*

Proof. This follows from Theorems 4.2 and 4.4.

Theorem 6.4. *Let A and B be two compact convex sets one of which is of class B^h for a function h satisfying the requirements of Definition 6.2. Then the boundary of $A + B$ is of class B^h. Similarly for G^h.*

Proof. This is immediate from Theorem 4.5.

Krantz & Parks [1991: Theorem 3] prove this result in the case of Hölder classes $C^{1,\varepsilon}$, $0 < \varepsilon \leqslant 1$.

Theorem 6.5. *Let A be a compact set in $\mathbf{R}^n$ of regularity class B^h for some function h as in Definition 6.2. Then its convex hull $\operatorname{cvx} A$ is of class B^h. Similary for G^h. In particular, if ∂A is of class C^1 this is true also of $\partial \operatorname{cvx} A$; and similarly for the Hölder classes $C^{1,\varepsilon}$, $0 < \varepsilon \leqslant 1$.*

Proof. This is immediate from Theorem 5.3, Corollary 5.5, and Proposition 3.5.

Griewank & Rabier [1990: Theorem 4.2] prove this result in the case of Hölder classes $C^{1,\varepsilon}$, $0 < \varepsilon \leqslant 1$.

Note that Corollary 3.4 enables us to translate properties in terms of B^h, G^h into properties in terms of $B^{1,h}$ and $G^{1,h}$, under some mild conditions on h.

7. Dual regularity classes

We shall now define classes of functions which are analogous to the B^h by using instead estimates from below.

Definition 7.1. *Given a function $h\colon \mathbf{R}^n \to [0,+\infty]$ we shall say that $f\colon \mathbf{R}^n \to [-\infty,+\infty]$ is **locally controlled by** h **from below** if for every point $a \in \mathbf{R}^n$ such that $f(a) \in \mathbf{R}$ there are constants $c = c_a > 0$ and $C = C_a > 0$ and a linear function L such that*

$$f(a+x) \geqslant f(a) + L(x) + Ch(cx) \qquad \text{for all } x \in \mathbf{R}^n. \tag{7.1}$$

*We shall say that f is **locally uniformly controlled by** h **from below** if for every $a_0 \in \mathbf{R}^n$ such that $f(a_0) \in \mathbf{R}$ there are constants $c > 0$ and $C > 0$ such that (7.1) holds for all a near a_0 with $f(a) \in \mathbf{R}$. We shall write $B_h = B_h(\mathbf{R}^n)$ for the set of all f which are locally controlled by h from below, and $G_h = G_h(\mathbf{R}^n)$ for the set of all f such that f is locally uniformly controlled by h from below.*

Proposition 7.2. *Every strictly convex function is locally in some class G_h for a function h such that $h(0) = 0$ and $h(x) > 0$ for $x \neq 0$.*

Proof. Let Γ_a denote the set of subgradients at a and define for a compact set K contained in the interior of $\operatorname{dom} f$,

$$h(x) = \inf_{a,\xi} \big[f(a+x) - f(a) - \xi \cdot x;\ a \in K \text{ and } \xi \in \Gamma_a \big], \qquad x \in \mathbf{R}^n.$$

Then $f|_K \in G_h$ by construction. In view of the strict convexity, $h(x) > 0$ for $x \neq 0$.

The classes B_h and G_h are dual to the classes defined by control from above. This duality is expressed by means of the Fenchel transformation. Given any function $f\colon \mathbf{R}^n \to [-\infty,+\infty]$ we define its ***Fenchel transform*** (or ***Legendre transform*** or ***conjugate function***) as

$$\widetilde{f}(\xi) = \sup_{x \in \mathbf{R}^n} \big[\xi \cdot x - f(x) \big], \qquad \xi \in \mathbf{R}^n.$$

We shall see that essentially $f \in G^h$ if and only if $\widetilde{f} \in G_{\tilde{h}}$. As an example consider the functions that we use to define the Hölder classes:

$$h(x) = \begin{cases} \dfrac{1}{1+\varepsilon} |x|^{1+\varepsilon}, & |x| \leqslant 1, \\ +\infty, & |x| > 1, \end{cases}$$

for $0 < \varepsilon \leqslant 1$. Their transforms are

$$\widetilde{h}(\xi) = \begin{cases} \dfrac{1}{1+1/\varepsilon} |\xi|^{1+1/\varepsilon}, & |\xi| \leqslant 1, \\ |\xi| - \dfrac{1}{1+\varepsilon}, & |\xi| \geqslant 1. \end{cases}$$

More generally, writing $f \approx g$ if there exists a constant C such that $C^{-1}g \leqslant f \leqslant Cg$ in a neighborhood of the origin, we see that $h(x) \approx \sum |x_j|^{p_j}$ with $1 < p_j < +\infty$, $h(x) = +\infty$ outside a bounded set, implies $\widetilde{h}(\xi) \approx \sum |\xi_j|^{q_j}$, where $q_j = p_j/(p_j - 1)$.

Theorem 7.3. *Let $f\colon \mathbf{R}^n \to [-\infty, +\infty]$ be a lower semicontinuous function in $\mathbf{R}^n$ which grows faster than any linear function. Suppose that $h(x) = o(|x|)$ as $x \to 0$. If $f \in B^h$, then $\widetilde{f} \in B_{\tilde{h}}$. Also $f \in E^h$ implies $\widetilde{f} \in G_{\tilde{h}}$.*

Proof. If f takes the value $-\infty$, then $\widetilde{f}$ is identically $+\infty$, and if f is identically $+\infty$, then $\widetilde{f}$ is $-\infty$ identically; in these two cases the result is true, so suppose that this is not the case. Then $\widetilde{f}$ is finite everywhere. For any given $\alpha \in \mathbf{R}^n$, there is a point a such that $\widetilde{f}(\alpha) = \alpha \cdot a - f(a)$. The regularity means that

$$f(x) \leqslant f(a) + L(x - a) + Ch(c(x - a)) \tag{7.2}$$

for all x. The linear function L must be $L(x) = \alpha \cdot x$. In fact, we have

$$\alpha \cdot (a + x) - \widetilde{f}(\alpha) \leqslant f(a + x) \leqslant f(a) + L(x) + Ch(cx),$$

which yields $\alpha \cdot x - L(x) \leqslant Ch(cx)$, and this is possible only if $\alpha \cdot x = L(x)$. We now calculate the Fenchel transform of the right-hand side of (7.2):

$$\begin{aligned}
&\sup_x \left[\xi \cdot x - f(a) - \alpha \cdot (x - a) - Ch(c(x - a))\right] \\
&\qquad = -f(a) + \xi \cdot a + C \sup_y \left[((\xi - \alpha)/cC) \cdot y - h(y)\right] \\
&\qquad = -f(a) + \xi \cdot a + C\widetilde{h}((\xi - \alpha)/cC) \\
&\qquad = \widetilde{f}(\alpha) + (\xi - \alpha) \cdot a + C\widetilde{h}((\xi - \alpha)/cC).
\end{aligned} \tag{7.3}$$

So

$$\widetilde{f}(\xi) \geqslant \widetilde{f}(\alpha) + (\xi - \alpha) \cdot a + C\widetilde{h}((\xi - \alpha)/cC)$$

for all ξ. This means precisely that $\widetilde{f} \in B_{\tilde{h}}$. The statement for the class $G_{\tilde{h}}$ follows because the constants c and C are bounded when α varies in a compact set.

Theorem 7.4. *Let $f\colon \mathbf{R}^n \to [-\infty, +\infty]$ be a lower semicontinuous function which grows faster than any linear function. Assume that f is differentiable at every point where it is finite. Let $h\colon \mathbf{R}^n \to [0, +\infty]$. If $f \in B_h$, then $\widetilde{f} \in B^{\tilde{h}}$.*

Proof. As before, if f is $+\infty$ identically or if f takes the value $-\infty$ the theorem holds, so we can suppose that $\widetilde{f}$ is finite everywhere. Moreover we know that to every given $\alpha \in \mathbf{R}^n$ there is a point a such that $\widetilde{f}(\alpha) = \alpha \cdot a - f(a)$. The hypothesis $f \in B_h$ means, for every a such that $f(a)$ is finite, that

$$f(x) \geqslant f(a) + L(x - a) + Ch(c(x - a)). \tag{7.4}$$

Here the only choice for L is $L(x) = f'(a) \cdot x$, since we now assume that $f'(a)$ exists. The Fenchel transform of the right-hand side of (7.4) at α is again given by (7.3), so it follows that

$$\widetilde{f}(\xi) \leqslant \widetilde{f}(\alpha) + (\xi - \alpha) \cdot a + C\widetilde{h}((\xi - \alpha)/cC)$$

for all ξ. Thus $\widetilde{f} \in B^{\tilde{h}}$.

The hypothesis that f' exists is of course not satisfactory, but it is necessary as the following example shows.

Example. Let $f(x) = |x| + \frac{1}{4}x^4$, $x \in \mathbf{R}$. Then $f \in B_h \setminus G_h$ if we take $h(x) = x^2/2$ for $|x| \leqslant 1$ and equal to $|x| - 1/2$ outside this interval. But $\widetilde{f}(\xi) = \frac{3}{4}((|\xi| - 1)^+)^{4/3}$ which is not in $B^{\tilde{h}}$.

As this example suggests, things improve if we strengthen the hypothesis $f \in B_h$ to $f \in G_h$:

Theorem 7.5. *Let $h\colon \mathbf{R}^n \to [0, +\infty]$ be given, and let $f\colon \mathbf{R}^n \to \left]-\infty, +\infty\right]$ be a lower semicontinuous convex function which grows faster than any linear function. Assume that* $\operatorname{dom} f$ *is open and nonempty. Then if $f \in G_h$ it follows that $\widetilde{f} \in G^{\tilde{h}}$.*

Proof. We know that $\widetilde{f}$ is finite everywhere and that there exists, to every given α, a point a such that $\widetilde{f}(\alpha) = \alpha \cdot a - f(a)$. Let b be any point such that f is differentiable at b; this happens for a dense set of points in the effective domain of f. At such a point we have, as in the proof of Theorem 7.4,

$$\widetilde{f}(\beta + \eta) \leqslant \widetilde{f}(\beta) + \eta \cdot b + \widetilde{h}(\eta), \qquad \eta \in \mathbf{R}^n, \tag{7.5}$$

where $\beta = f'(b)$. Here we have assumed that $c = C = 1$, which is legitimate since $f \in G_h$ and all points under consideration remain in a compact set. Let Γ_a denote the set of subgradients at a point a; thus $\Gamma_b = \{\beta\}$. Let further Θ_a be the subset of Γ_a formed by all limits of $f'(b_k)$ where b_k is such that f is differentiable at b_k and $b_k \to a$. Both Γ_a and Θ_a are compact, and $\Gamma_a = \operatorname{cvx} \Theta_a$. Given any $\beta \in \Theta_a$ we have

$$\widetilde{f}(\beta_k + \eta) \leqslant \widetilde{f}(\beta_k) + \eta \cdot b_k + \widetilde{h}(\eta), \qquad \eta \in \mathbf{R}^n, \tag{7.6}$$

for some $\beta_k = f'(b_k) \to \beta$, $b_k \to a$. We can pass to the limit in (7.6). Indeed, all terms except perhaps $\widetilde{h}$ are continuous in $\mathbf{R}^n$. Thus $\widetilde{f}(\beta + \eta) \leqslant \widetilde{f}(\beta) + \eta \cdot a + \widetilde{h}(\eta)$, or equivalently

$$\widetilde{f}(\xi) \leqslant \widetilde{f}(\beta) + (\xi - \beta) \cdot a + \widetilde{h}(\xi - \beta), \qquad \beta \in \Theta_a, \quad \xi \in \mathbf{R}^n. \tag{7.7}$$

We shall now extend the validity of (7.7) from Θ_a to all of Γ_a:

$$\widetilde{f}(\xi) \leqslant \widetilde{f}(\alpha) + (\xi - \alpha) \cdot a + \widetilde{h}(\xi - \alpha), \qquad \alpha \in \Gamma_a, \quad \xi \in \mathbf{R}^n. \tag{7.8}$$

We shall of course use that $\widetilde{f}$ is affine in Γ_a, specifically that $\widetilde{f}(\beta) + (\xi - \beta) \cdot a = \widetilde{f}(\alpha) + (\xi - \alpha) \cdot a$ for all $\xi \in \mathbf{R}^n$ when $\alpha, \beta \in \Gamma_a$. Thus (7.7) implies, if we fix α and let β vary in Θ_a,

$$\widetilde{f}(\xi) \leqslant \widetilde{f}(\alpha) + (\xi - \alpha) \cdot a + \inf_{\beta \in \Theta_a} \widetilde{h}(\xi - \beta), \qquad \alpha \in \Gamma_a, \quad \xi \in \mathbf{R}^n.$$

Since $\widetilde{f}$ is convex, this inequality can be replaced by the formally stronger

$$\widetilde{f}(\xi) \leqslant \widetilde{f}(\alpha) + (\xi - \alpha) \cdot a + \varphi_a(\xi), \qquad \alpha \in \Gamma_a, \quad \xi \in \mathbf{R}^n,$$

where φ_a denotes the convex envelope of the function $\xi \mapsto \inf_{\beta\in\Theta_a} \widetilde{h}(\xi-\beta)$. To get (7.8) it suffices to prove that $\varphi_a(\xi) \leqslant \widetilde{h}(\xi-\alpha)$. This is easy. In fact, if $\alpha = \sum \lambda_j\beta_j$ with $\beta_j \in \Theta_a$, $\lambda_j \geqslant 0$, $\sum \lambda_j = 1$, then $\xi = \sum \lambda_j \xi_j$, where $\xi_j = \beta_j + \xi - \alpha$, so by definition

$$\varphi_a(\xi) \leqslant \sum \lambda_j \varphi_a(\xi_j) \leqslant \sum \lambda_j \widetilde{h}(\xi_j - \beta_j) = \sum \lambda_j \widetilde{h}(\xi-\alpha) = \widetilde{h}(\xi-\alpha).$$

This proves the theorem. It is the limiting process of passing from (7.6) to (7.7) which saves us, and which cannot be done for $f \in B_h$: then the constants may blow up.

These theorems say that the transform of a function in a regularity class G^h must be strictly convex in a certain sense, and conversely.

Let $\widetilde{\widetilde{h}} = h$ satisfy the hypotheses of Theorems 7.3 and 7.5. Then $f \in G^h$ implies $\widetilde{\widetilde{f}} \in G^h$. This gives a new proof of Corollary 5.5 under the stated conditions.

References

Azé, Dominique, & Penot, Jean-Paul

ms — Uniformly convex functions and uniformly smooth convex functions. Manuscript, 29 pages.

Boman, Jan

1990a — The sum of two plane convex C^∞ sets is not always C^5. *Math. Scand.* **66**, 216–224.

1990b — Smoothness of sums of convex sets with real analytic boundaries. *Math. Scand.* **66**, 225–230.

Carathéodory, C.

1911 — Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Funktionen. *Rend. Circ. mat. Palermo* **32**, 193–217.

Griewank, A., & Rabier, P. J.

1990 — On the smoothness of convex envelopes. *Trans. Amer. Math. Soc.* **322**, 691–709.

Kiselman, Christer O.

1986 — How smooth is the shadow of a smooth convex body? *J. London Math. Soc. (2)* **33**, 101–109.

1987 — Smoothness of vector sums of plane convex sets. *Math. Scand.* **60**, 239–252.

Krantz, Steven G., & Parks, Harold R.

1991 — On the vector sum of two convex sets in space. *Canad. J. Math.* **43**, 347–355.