Mathematische Zeitschrift 230, 259-316 (1999)

A geometrical approach to monotone functions in $\mathbb{R}^n$

Giovanni Alberti[1], **Luigi Ambrosio**[2]

[1] Dipartimento di Matematica Applicata "U. Dini", via Bonanno 25/B, I-56126 Pisa, Italy
[2] Dipartimento di Matematica "F. Casorati", via Abbiategrasso 215, I-27100 Pavia, Italy

Received October 9, 1996; in final form April 21, 1997

Abstract. This paper is concerned with the fine properties of monotone functions on $\mathbb{R}^n$. We study the continuity and differentiability properties of these functions, the approximability properties, the structure of the distributional derivatives and of the weak Jacobians. Moreover, we exhibit an example of a monotone function u which is the gradient of a $C^{1,\alpha}$ convex function and whose weak Jacobian Ju is supported on a purely unrectifiable set.

Introduction

This paper is devoted to a systematic analysis of the properties of monotone functions defined on a finite dimensional space $\mathbb{R}^n$. The initial motivation of our work is the following: in the last years, several questions concerning the "right" definitions of Jacobian determinant and of graph area for a nonsmooth function (e.g., belonging to a Sobolev space) have been debated (see for instance [Ba], [GMS1], [GMS2], [Mu1], [Mu2], [Muc]). We will see that for maximal monotone functions there is a natural and completely satisfactory way to introduce these concepts, getting continuity properties and strong approximability by smooth functions. We notice that the monotonicity property allows discontinuities and the existence of "vertical parts" in the graph; however, maximal monotone functions are in some sense very smooth, because for them several weak objects are univocally defined and there is a good description of the singularities.

Our paper has also been conceived as a review on the fine properties of monotone functions, such as continuity, differentiability, structure of the distributional derivative. Our review collects many results scattered in the

Mathematics Subject Classification (1991): 26B05, 26B25, 49Q15, 58A25, 47H05

literature (see for instance [A1], [Ale], [AG], [AK], [Be], [Br], [Mi], [Mig]) and is, as far as possible, self-contained. We will assume only well-known results on Lipschitz functions, namely Rademacher's and Kirszbraun's theorems, and the area formula.

Our starting point is the one to one correspondence, observed and used by Minty in [Mi], between graphs of monotone functions and graphs of 1-Lipschitz functions. Specifically, the Cayley transformation

$$(x,y) \mapsto \frac{1}{\sqrt{2}}(x+y,-x+y) \qquad (x,y) \in \mathbb{R}^n \times \mathbb{R}^n$$

(a clockwise rotation of $\pi/4$ if $n=1$) transforms the graph

$$\Gamma u = \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : \ y \in u(x)\}$$

of a monotone function u in the graph of a 1-Lipschitz function. We will see that this correspondence is extremely useful to unify and simplify several proofs existing in the literature: for instance, Rademacher's differentiability theorem for Lipschitz functions leads to an analogous property for monotone functions and Kirszbraun's extension theorem for Lipschitz functions leads to a characterization of maximal monotone functions as those monotone functions u such that $u+I$ is surjective. We also notice that the theory of cartesian currents developed in [GMS1] is strongly founded in the identification between a function and its (generalized) graph.

Now we briefly describe in a more detailed way the content of our paper; bibliographical and historical comments will be given after the statements of main theorems.

In the first section we fix our notations and we prove in Proposition 1.1 the above mentioned correspondence between graphs of monotone functions and graphs of 1-Lipschitz functions. After the statement of some classical properties of monotone functions (Proposition 1.2 and in Corollaries 1.3, 1.4, 1.5) we define a natural topology on the space $\mathscr{M}on$ of maximal monotone functions which is related to the Kuratowski convergence of the associated graphs. In Proposition 1.7 we prove that any sequence (u_h) of maximal monotone functions either admits a convergent subsequence or "goes to infinity" (i.e., the intersection of the graph of u_h with any bounded set in $\mathbb{R}^n \times \mathbb{R}^n$ is empty for sufficiently large h).

In the second section (Theorem 2.2) we prove that the singular sets

$$\Sigma^k(u) = \big\{x \in \mathbb{R}^n : \ \dim u(x) \geq k\big\} \qquad \text{for } k=1,\dots,n,$$

are countably $\mathscr{H}^{n-k}$-rectifiable, i.e., $\mathscr{H}^{n-k}$-almost all of $\Sigma^k(u)$ can be covered by a sequence of C^1 surfaces of dimension $(n-k)$. We notice that u is continuous outside $\Sigma^1(u)$.

In the third section we prove a uniform estimate on the area of the graph Γu of a monotone function u (see Proposition 3.1)

$$\mathscr{H}^n\big(\Gamma u \cap B(x,r)\big) \leq 2^{n/2}\omega_n r^n \ , \tag{1}$$

and we infer, using the area formula and Rademacher's theorem, the following differentiability property (see Theorem 3.2): for almost every $\bar{x} \in \mathsf{Dm}\, u \backslash \Sigma^1(u)$

$$\lim_{\substack{x\to\bar{x} \\ y\in u(x)}} \frac{y - u(\bar{x}) - \nabla u(\bar{x})\cdot(x-\bar{x})}{|x-\bar{x}|} = 0 \ , \tag{2}$$

here $\mathsf{Dm}\, u$ is the set of all points where u has non-empty value, and then $\mathsf{Dm}\, u \setminus \Sigma^1(u)$ is the set of all points $\bar{x}$ such that the set $u(\bar{x})$ consist of one point, which is then identified with $u(\bar{x})$.

In section 4 we associate to any maximal monotone function u a n-current Tu in $\mathbb{R}^n \times \mathbb{R}^n$ without boundary associated to the integration on the graph Γu. Indeed, by Minty's correspondence with graphs of 1-Lipschitz functions, Γu can be oriented in such a way that Stokes's theorem holds

$$\int_{\Gamma u} d\omega = 0 \qquad \forall\, \omega \in \mathscr{D}^{n-1}(\mathbb{R}^n \times \mathbb{R}^n) \ .$$

The mapping $u \mapsto Tu$ enjoys natural continuity properties which are a direct consequence (after performing the Cayley transformation) of the continuity properties of $f \mapsto Tf$ in the class of functions satisfying an equi-Lipschitz condition.

In section 5 we prove that the representation of the graph of u as a n-current implies, among other things, that any mononotone function is of class BV_{loc} in the interior of $\mathsf{Dm}\, u$ (Proposition 5.1). Moreover, the monotonicity property is characterized in $\mathsf{Dm}\, u$ by the positivity of the distributional derivative Du (Theorem 5.3). Consequently, we analyze the structure of Du, splitting it into three parts: the absolutely continuous part $D_a u$, the jump part $D_j u$, and the Cantor part $D_c u$ (paragraph 5.6). We present in Theorem 5.10 a very simple proof of the rank-one property of the measures $D_j u$, $D_c u$, based on the area estimate (1) and on Reshetnyak's continuity theorem (this proofs seems to work for monotone functions only and the proof of the rank-one property for a general BV function is much harder, cf. [A1]).

For any maximal monotone function u we can also define the weak Jacobian Ju, as the positive measure given by

$$Ju(B) := \mathscr{L}_n\big(u(B)\big)$$

for any bounded Borel set B contained in the interior of $\mathsf{Dm}\, u$ (cf. Definition 5.12). It turns out that the map $u \mapsto Ju$ is continuous with respect to the natural topologies, and that Ju coincides with the measure $\det(\nabla u)\cdot\mathscr{L}_n$ when u is a C^1 function. Furthermore,

$$\int \phi(x)\,dJu(x) = Tu\big(\phi(x)\,dy_1\wedge\ldots\wedge dy_n\big)$$

for any smooth function ϕ with support contained in the interior of $\mathsf{Dm}\,u$.

In section 6 we prove that any maximal monotone function u may be approximated by a canonical sequence of $1/\varepsilon$-Lipschitz maximal monotone functions u_ε (the so-called Yosida approximation of u) so that $u_\varepsilon\to u$, $Tu_\varepsilon\to Tu$ and $|Tu_\varepsilon|\to|Tu|$ (see Theorem 6.2 for precise statements). The function u_ε is given by $(\varepsilon I+u^{-1})^{-1}$ and its graph Γu_ε is given by

$$\Gamma u_\varepsilon = \big\{(x+\varepsilon y, y):\ (x,y)\in\Gamma u\big\}\ .$$

Hence, the graph of u_ε is, for small ε, a slight deformation of the graph of u. The convergence of Tu_ε to Tu implies, in particular, the convergence of the distributional derivatives Du_ε to Du. We also notice that the convergence $|Tu_\varepsilon|\to|Tu|$ (often called convergence in area) is a quite strong property and implies, via Reshetnyak's theorem, that $|Du_\varepsilon|\to|Du|$, that is, the convergence in variation of the derivatives.

If we choose u_ε to be the standard mollified functions, then the convergence in variation of the derivatives still holds, but we don't know whether u_ε converge in area to u or not (see Remark 6.4). The difficulty comes from the nonlinearities involved in the computations of the determinants of all minors of ∇u_ε.

In section 7 we confine our attention to a special class of monotone functions, the subgradients ∂f of convex and lower semicontinuous functions $f:\mathbb{R}^n\to]-\infty,+\infty]$. Using the differentiability property (2) we quickly recover the Aleksandrov's theorem on the almost everywhere second order differentiability of a convex function (more precisely, we prove the existence almost everywhere of a second order Taylor expansion, see Theorem 7.10). Moreover, in Proposition 7.13 we show that setting $u:=\partial f$, the approximating functions u_ε of section 6 are exactly the gradients of the inf-convolutions

$$f_\varepsilon(x) := \min_{x'\in\mathbb{R}^n}\left\{f(x')+\frac{1}{2\varepsilon}|x'-x|^2\right\}\ .$$

Section 8 is completely dedicated to the construction of an example of maximal monotone function u whose weak Jacobian has a nontrivial structure. Indeed, we find a Hölder continuous monotone function $u:\mathbb{R}^2\to\mathbb{R}^2$ such that the weak Jacobian Ju is supported on a Lebesgue negligible, purely unrectifiable set A and such that $\det(\nabla u)=0$ a.e. in $\mathbb{R}^2$. Moreover, u is the gradient of a convex function of class $C^1\cap W^{2,p}$ for every $p<2$.

We will obtain u by a method related to the construction of Hutchinson's self-similar fractals. The requirement of being the gradient of convex function is a source of difficulties (see the discussion in paragraph 8.8) and forces us to choose a countable family of similitudes Ψ_i taking the unit ball $B(0,1)$ to balls $B(x_i,r_i)$ which cover almost all $B(0,1)$.

It turns out that the Hausdorff dimension of A is related to the so-called packing exponent of this covering, and is never lower than 1 (see paragraph 8.10 and Remark 8.11). In paragraph 8.12 we discuss some conjectures about the "dimension" of the measure Ju. In particular, we believe that the parameters of our construction can be arranged in such a way that the dimension of Ju is any prescribed number in $]0,2[$.

Finally, in paragraph 8.15 we discuss the possibility of closure theorems for special classes of monotone functions (e.g. monotone functions on the plane whose distributional derivative and weak Jacobian can be written as a sum of integer dimensional measures) in analogy with the theory of SBV functions. We modify the construction of the function u (see Definition 8.13 and Proposition 8.14) to exhibit a sequence of monotone functions which shows that no analogous of the closure theorem for SBV functions holds for monotone functions.

Acknowledgement. The first author gratefully acknowledges the hospitality and the support of Institute for Mathematics and its Applications of the University of Minnesota, where part of this paper was written.

1. Preliminary definitions and basic results

We first recall some basic notation. We denote by $\mathscr{L}_n$ the *Lebesgue measure* in $\mathbb{R}^n$; $|B|:=\mathscr{L}_n(B)$ for every Borel set B in $\mathbb{R}^n$ and ω_n is the Lebesgue measure of the unit ball; in addition, $\mathscr{H}^n$ is the *n-dimensional Hausdorff measure* (on any metric space). By *measure* we always mean a measure on Borel sets. The *restriction* of a measure μ to a Borel set E is denoted by $\mu\llcorner E$, i.e., $\mu\llcorner E(B):=\mu(E\cap B)$ for any Borel set B; we also set

$$f\mu(B) := \int_B f(x)\,d\mu(x)$$

provided that the integral at the right hand side is defined. We say that μ *charges* E when $|\mu|(E)>0$, and μ *is supported on* E when the restriction of μ to the complement of E is zero.

For every function f, we denote by Df the *distributional derivative* of f, and by $\nabla f(x)$ the *gradient* at the point x (both in the classical sense and in the approximate sense).

An $n \times n$ matrix A is termed *positive* if it is positively semi-definite, that is, if $\langle Av, v\rangle \geq 0$ for every $v \in \mathbb{R}^n$, and it is termed symmetric if $A = A^t$. We say that a matrix-valued function f is positive (resp. symmetric) if $f(x)$ is always a positive (resp. symmetric) matrix. A matrix-valued distribution Λ is positive (resp. symmetric) if $\langle \Lambda, \phi\rangle$ is a positive (resp. symmetric) matrix for every positive real valued test function ϕ.

For every set B in $\mathbb{R}^n$, $\mathsf{Int}\, B$ and $\mathsf{Conv}\, B$ denote respectively the *interior* and the *closed convex hull* of B.

In this paper we deal with set-valued maps $u : \mathbb{R}^n \to \mathbb{R}^n$ (that is, maps which take every point $x \in \mathbb{R}^n$ in some set $u(x) \subset \mathbb{R}^n$). We call these maps *multifunctions* (on $\mathbb{R}^n$), or simply functions when no ambiguities may arise. We denote by I both the identity function on $\mathbb{R}^n$ and the $n \times n$ identity matrix.

We say that u is *univalued* on some set B if $u(x)$ consists of at most one point for every $x \in B$, and that u is k-Lipschitz if $|y_1 - y_2| \leq k|x_1 - x_2|$ for every $x_i \in \mathbb{R}^n$, $y_i \in u(x_i)$, $i = 1, 2$ (clearly every Lipschitz multifunction is univalued).

Let be given multifunctions u, v, real numbers $\lambda, \mu \in \mathbb{R}$ and a set $B \subset \mathbb{R}^n$. For all $x \in \mathbb{R}^n$ we set

$$\begin{aligned}
&\textit{domain of } u,\ \mathsf{Dm}\, u := \{x :\ u(x) \neq \emptyset\}\ , \\
&\textit{image of } u,\ \mathsf{Im}\, u := \{y :\ \exists\, x,\ y \in u(x)\}\ , \\
&\textit{graph of } u,\ \Gamma u := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n :\ y \in u(x)\}\ . \\
&\textit{inverse of } u,\ [u^{-1}](x) := \{y :\ x \in u(y)\}\ , \\
&\qquad [\lambda u + \mu v](x) := \{\lambda y + \mu y' :\ y \in u(x),\ y' \in v(x)\}\ , \\
&\qquad u(B) := \{y :\ \exists x \in B,\, y \in u(x)\}\ .
\end{aligned}$$

We write $u \supset v$ when the graph of u includes the graph of v, i.e., when $u(x) \supset v(x)$ for all x. We say that u is *monotone* if

$$\langle y_1 - y_2, x_1 - x_2\rangle \geq 0 \qquad \forall x_i \in \mathbb{R}^n,\ y_i \in u(x_i),\, i = 1, 2\ . \tag{1.1}$$

A monotone function u is called *maximal* when it is maximal with respect to inclusion in the class of monotone functions, i.e., if the following implication holds:

$$v \supset u,\ v \text{ monotone } \Rightarrow\ v = u\ .$$

Clearly for any monotone function u there exists a maximal monotone function $\bar{u}$ which includes u. We remark that u is a (maximal) monotone function if and only if u^{-1} is a (maximal) monotone function. It is also immediate that the class of monotone functions is a cone.

The definition of (maximal) monotone function can be readily extended to multifunctions on Hilbert spaces. Monotonicity can be defined also for multifunctions from a Banach space X into its dual X^*, but there exists also a notion of *accretive* function for multifunctions from X into X; these two notions agree on Hilbert spaces (see for instance [Br], section II.1).

Monotone functions in Hilbert spaces play an essential rôle in evolution equations and in many other fields of functional analysis, but in this paper we are concerned essentially with the finite dimensional case. For further details and references about monotone functions we refer to [Br], chapter II.

Following [Mi], we prove that the graph of a maximal monotone function u is a Lipschitz submanifold without boundary of $\mathbb{R}^n \times \mathbb{R}^n$. Moreover, there exists a one-to-one correspondence between the class of maximal monotone functions on $\mathbb{R}^n$ and the class of 1-Lipschitz functions from $\mathbb{R}^n$ to $\mathbb{R}^n$.

Let $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ be the Cayley transformation, i.e., the linear isometry defined by

$$\Phi : (x, y) \longmapsto \frac{1}{\sqrt{2}}(x + y, -x + y) \qquad \forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n\ . \tag{1.2}$$

Then we have the following result (first proved in [Mi], Lemma 3 and Theorem 4):

Proposition 1.1. *Let u be a maximal monotone function. Then $(u + I)^{-1}$ is defined on the whole $\mathbb{R}^n$ and $\Phi(\Gamma u)$ is the graph of the 1-Lipschitz function $Fu : \mathbb{R}^n \to \mathbb{R}^n$ given by*

$$Fu : z \mapsto \left[z - \sqrt{2}(u + I)^{-1}(\sqrt{2}z)\right] \quad \forall z \in \mathbb{R}^n\ . \tag{1.3}$$

Conversely, for any 1-Lipschitz function $\phi : \mathbb{R}^n \to \mathbb{R}^n$ the set $\Phi^{-1}(\Gamma\phi)$ is the graph of a maximal monotone function on $\mathbb{R}^n$.

Proof. Let u be a monotone function, and let Fu be the multifunction whose graph is $\Phi(\Gamma u)$. We claim that Fu satisfies (1.3) and that Fu is 1-Lipschitz in its domain (and in particular it is univalued).

Let $z \in \mathbb{R}^n$. By (1.2), we have that v belongs to $Fu(z)$ if and only if there exist $x \in \mathbb{R}^n$ and $y \in u(x)$ such that $z = \frac{1}{\sqrt{2}}(x + y)$ and $v = \frac{1}{\sqrt{2}}(-x + y)$, or, equivalently, $x = \frac{1}{\sqrt{2}}(z - v)$ and $y = \frac{1}{\sqrt{2}}(z + v)$. This means that $v \in Fu(z)$ if and only if

$$\frac{z + v}{\sqrt{2}} \in u\Big(\frac{z - v}{\sqrt{2}}\Big) \tag{1.4}$$

and then $v \in z - \sqrt{2}(u + I)^{-1}(\sqrt{2}z)$. Let $z' \in \mathbb{R}^n$ and $v' \in Fu(z')$. By applying (1.4) both to (z, v) and (z', v'), and taking the monotonicity of u into account, we infer

$$\big\langle (v + z) - (v' + z'), (z - v) - (z' - v')\big\rangle \geq 0$$

which yields $|v - v'|^2 \le |z - z'|^2$. This shows that Fu is a 1-Lipschitz function.

The same argument shows that Φ^{-1} maps graphs of 1-Lipschitz functions in graphs of monotone functions. Assuming now that u is maximal monotone, we claim that the domain of Fu is $\mathbb{R}^n$. Indeed, were this not true we could use Kirzsbraun's theorem (see for instance [Fe1], 2.10.43) to extend Fu to a 1-Lipschitz function L defined on all $\mathbb{R}^n$ and then $\Phi^{-1}(\Gamma L)$ provides a monotone extension of u, which contradicts the maximality of u. □

Proposition 1.2. *Let u be a monotone function. Then*

(1) *if u is maximal, Γu is closed, and $u(x)$ is a convex, closed (possibly empty) set for every $x \in \mathbb{R}^n$;*
(2) *u is maximal if and only if $(u+I)$ is onto, i.e., if and only if the domain of $(u+I)^{-1}$ is $\mathbb{R}^n$;*
(3) *$(u+I)$ and $(u+I)^{-1}$ are monotone functions and $(u+I)^{-1}$ is 1-Lipschitz;*
(4) *for any set $X \subset \mathsf{Dm}\, u$, $\bar{x}$ in the interior of $\mathsf{Conv}\, X$ and $\bar{y} \in u(\bar{x})$ we have*

$$|\bar{y}| \le \frac{C}{\mathrm{dist}\big(\bar{x}, \mathbb{R}^n \setminus \mathsf{Conv}\, X\big)} \tag{1.5}$$

where $C := \Big[\sup_{x \in X} \inf_{y \in u(x)} |y|\Big] \cdot \mathrm{diam}(X)$.

Proof. The proof of statement (1) is trivial, while (2) follows from Proposition 1.1.

About (3), it is easy to see that $(u+I)$ is monotone, and thus $(u+I)^{-1}$ is monotone too. For any choice of $x_i \in \mathbb{R}^n$ and $y_i \in u(x_i)$ $(i = 1, 2)$ we have

$$\begin{aligned} |x_1 - x_2|^2 &\le \big\langle (y_1 - y_2) + (x_1 - x_2), x_1 - x_2 \big\rangle \\ &\le \big|(y_1 + x_1) - (y_2 + x_2)\big|\, |x_1 - x_2| \ , \end{aligned} \tag{1.6}$$

so that $|x_1 - x_2| \le |(y_1 + x_1) - (y_2 + x_2)|$. In particular, when $y_1 + x_1 = y_2 + x_2$ then $x_1 = x_2$, and this means that $(u+I)^{-1}$ is univalued. Thereafter $(u+I)^{-1}(y_i + x_i) = x_i$, and (1.6) yields the Lipschitz property of $(u+I)^{-1}$.

Eventually we prove (4). We may assume with no loss in generality that $\bar{x} = 0$. We take $\bar{y} \in u(0)$, and a positive number $d < \mathrm{dist}\big(0, \mathbb{R}^n \setminus \mathsf{Conv}\, X\big)$. We set $z := d\bar{y}/|\bar{y}|$. Then z belongs to the interior of the convex hull of X, and we can find finitely many points $x_i \in X$ and positive real numbers α_i such that

$$\sum_i \alpha_i = 1 \ , \quad \sum_i \alpha_i x_i = z \ .$$

Moreover we take m so that

$$m > \Big[\sup_{x \in X} \inf_{y \in u(x)} |y|\Big] \ .$$

Since $X \subset \mathsf{Dm}\, u$ and $x_i \in X$, we can find $y_i \in u(x_i)$ so that $m \ge |y_i|$ for all i. By the monotonicity of u we infer $\langle \bar{y}, x_i \rangle \le \langle y_i, x_i \rangle$, and then

$$d|\bar{y}| = \langle \bar{y}, z \rangle = \sum_i \alpha_i \langle \bar{y}, x_i \rangle \le \sum_i \alpha_i \langle y_i, x_i \rangle \le m \cdot \sup_i |x_i| \le m \cdot \mathrm{diam}\, X \ .$$

Inequality (1.5) follows by taking the supremum over all admissible d and the infimum over all admissible m. □

Corollary 1.3. *Let u be a maximal monotone function. Then*

(1) *u is upper semicontinuous, i.e., if $x_h \to x$, $y_h \to y$ and $y_h \in u(x_h)$ then $y \in u(x)$;*
(2) *the domain of u contains the interior of its convex hull, that is,*

$$\mathsf{Int}\,\mathsf{Conv}\,(\mathsf{Dm}\, u) \subset \mathsf{Dm}\, u \subset \mathsf{Conv}\,(\mathsf{Dm}\, u) \ ; \tag{1.7}$$

(3) *for any B relatively compact in the interior of $\mathsf{Dm}\, u$, the image $u(B)$ is bounded;*
(4) *if $u(x)$ consists of exactly one point y, then x belongs to the interior of $\mathsf{Dm}\, u$, and u is continuous at x, i.e., $x_h \to x$ and $y_h \in u(x_h)$ yield $y_h \to y$.*

Proof. Property (1) is a restatement of the closure of the graph of u (Proposition 1.2(1)).

Take $\bar{x}$ in the interior of the convex hull of $\mathsf{Dm}\, u$. We claim that $\bar{x} \in \mathsf{Dm}\, u$; this would prove statement (2).

By the choice of $\bar{x}$ we may find finitely many points $x_i \in \mathsf{Dm}\, u$, $i \in I$, so that $\bar{x}$ belongs to the interior of the convex hull of $X := \{x_i : i \in I\}$. For every $i \in I$ we choose $y_i \in u(x_i)$, and for every $\varepsilon > 0$ we set $u_\varepsilon := (\varepsilon I + u^{-1})^{-1}$ (cf. Definition 6.1).

Since u^{-1} is maximal, then $\varepsilon I + u^{-1}$ is surjective (cf. Proposition 1.2(2)), and then the domain of u_ε is $\mathbb{R}^n$. Therefore we can choose $y_\varepsilon \in u_\varepsilon(\bar{x})$ for every $\varepsilon > 0$. Every cluster point y of the sequence (y_ε) belongs to $u(\bar{x})$: indeed $y_\varepsilon \in u_\varepsilon(\bar{x})$ implies $y_\varepsilon \in u(\bar{x} - \varepsilon y_\varepsilon)$, and passing to the limit as $\varepsilon \to 0$ we obtain $y \in u(\bar{x})$ (apply statement (1)). Hence the claim is proved if we show that the sequence (y_ε) is definitively bounded, which is obtained by a suitable application of estimate (1.5). For any $\varepsilon > 0$ we set $X_\varepsilon := \{x_i + \varepsilon y_i : i \in I\}$; since $y_i \in u_\varepsilon(x_i + \varepsilon y_i)$ for every i, then

$$\sup_{x \in X_\varepsilon} \inf_{y \in u_\varepsilon(x)} |y| \le \sup_{i \in I} |y_i| < +\infty \ .$$

Since $\bar{x}$ belongs to the interior of $\mathsf{Conv}\, X$, then there exists d such that $\mathrm{dist}(\bar{x}, \mathbb{R}^n \setminus \mathsf{Conv}\, X) > d > 0$, and then $\mathrm{dist}(\bar{x}, \mathbb{R}^n \setminus \mathsf{Conv}\, X_\varepsilon) > d$ definitively in ε. Hence we may apply inequality (1.5) with y_ε instead of $\bar{y}$ and X_ε instead of X to prove that (y_ε) is definitively bounded.

About statement (3), we recall that for every set B relatively compact in the interior of $\mathsf{Dm}\, u$, there exists a finite set $X \subset \mathsf{Dm}\, u$ such that $B \subset \mathsf{Conv}\, X$. Hence we may apply Proposition 1.2(4) (notice that the right hand side of inequality (1.5) is finite whenever X is a finite set).

About statement (4), we first prove that if $u(x)$ is a singleton, then x belongs to the interior of $\mathsf{Dm}\, u$. Take indeed x in the boundary of $\mathsf{Dm}\, u$. Then x belongs to the boundary of $\mathsf{Conv}\,(\mathsf{Dm}\, u)$, because the interior of $\mathsf{Conv}\,(\mathsf{Dm}\, u)$ is included in the interior of $\mathsf{Dm}\, u$ (statement (2)). Therefore Hahn-Banach theorem yields a non-trivial $e \in \mathbb{R}^n$ such that $\langle e, x\rangle \geq \langle e, x'\rangle$ for all $x' \in \mathsf{Dm}\, u$; if we take $y \in u(x)$, $x' \in \mathsf{Dm}\, u$ and $y' \in u(x')$ we obtain

$$\langle (y+e) - y', x - x'\rangle = \langle y - y', x - x'\rangle + \langle e, x - x'\rangle \geq 0 \ .$$

This means that $y + e$ belongs to $u(x)$ because u is maximal, and then $u(x)$ is not a singleton. The rest of statement (4) follows immediately by (1) and (3). □

Remark. About statement (2) of Corollary 1.3, we remark that the domain of u may be not convex. Indeed let u be the function on $\mathbb{R}^2$ defined by $u(\pm 1, 0) := (\pm 1, 0)$, $u(0,t) = (0, \log t)$ for every $t > 0$. It may be easily verified that u is monotone, and if we take any maximal extension $\bar{u}$ of u, then $\mathsf{Dm}\, \bar{u}$ contains the points $(\pm 1, 0)$, but the point $(0,0)$ never belongs to $\mathsf{Dm}\, \bar{u}$: assume by contradiction that $(y_1, y_2) \in \bar{u}(0,0)$, then the monotonicity condition yields

$$(\log t - y_2)t \geq 0 \qquad \text{for every } t > 0,$$

that is, $y_2 \leq \log t$ for any $t > 0$, which is impossible.

Corollary 1.4. *Let u be a monotone function which is upper semicontinuous and whose values are closed and convex. Then u agrees in the interior of its domain with every maximal monotone function which includes u. In particular every continuous univalued monotone function with domain $\mathbb{R}^n$ is maximal.*

Proof. Let u be a monotone function, let $\bar{u}$ be a monotone function which includes u, and let x be a point in the interior of $\mathsf{Dm}\, u$ such that u is upper semicontinuous at x, and $u(x)$ is closed and convex. We claim that $u(x) = \bar{u}(x)$.

Take r such that $B(x,r) \subset \mathsf{Dm}\, u$. Take $y \in \bar{u}(x)$, and assume by contradiction that $y \notin u(x)$. Since $u(x)$ is convex and closed, Hahn-Banach theorem yields a unitary vector $e \in \mathbb{R}^n$ and $\varepsilon > 0$ such that $\langle y' - y, e\rangle \geq \varepsilon$ for all $y' \in u(x)$. Therefore, if we take $y_t \in u(x + te)$ for every t with $0 < t < r$,

$$|y_t - y'| \geq \langle y_t - y', e\rangle = \frac{1}{t}\langle y_t - y, te\rangle + \langle y - y', e\rangle \geq \varepsilon \qquad \forall\, y' \in u(x)\, .$$

This means that the points y_t do not converge to $u(x)$ when $t \to 0$, which contradicts the assumption that u is upper semicontinuous at x. □

Corollary 1.5. *Let be given maximal monotone functions u and v and an open convex set A so that $u(x) \cap v(x) \neq \emptyset$ for every x in a dense subset of A. Then $u(x) = v(x)$ for every x in A.*

Proof. Let w the monotone function which takes every $x \in \mathbb{R}^n$ in the set $u(x) \cap v(x)$; thus $w(x)$ is always closed and convex (cf. Proposition 1.2(1)), and w is upper semicontinuous (cf. Corollary 1.3(1)).

Let $D := \mathsf{Dm}\, w$. By assumption we have that $\mathsf{Dm}\, u$ and $\mathsf{Dm}\, v$ contain D, and since D is dense in A we deduce that $\mathsf{Dm}\, u$ and $\mathsf{Dm}\, v$ contain A. Hence both u and v are locally bounded on A, therefore the same holds for w, and recalling that w is upper semicontinuous we deduce that $w(x)$ is non-empty for every $x \in A$, that is, $D \supset A$. Now we can apply Corollary 1.4 to obtain that $w(x) = u(x) = v(x)$ for every $x \in A$. □

Eventually we want to endow the class of all maximal monotone functions with a suitable topology. It turns out that the "natural" topology is related with the *Hausdorff distance* between closed sets. For more general results and details on the Hausdorff distance, Kuratowski convergence and related topics we refer to [CV], chapter II.

Let (E,d) be a metric space. We say that a sequence of closed sets $C_h \subset E$ converges to C *in the sense of Kuratowski* when the following two conditions are satisfied:

$$x_h \in C_h \Rightarrow \text{ every cluster point of } (x_h) \text{ belongs to } C, \tag{1.8}$$

$$\forall\, x \in C,\ \exists\, x_h \in C_h \text{ such that } x_h \to x. \tag{1.9}$$

We remark that C is uniquely determined by (1.8) and (1.9). When $C = \emptyset$, (1.9) is always verified and (1.8) is equivalent to say that any compact set $K \subset E$ intersects finitely many C_h only.

For every couple of closed sets C, $C' \subset E$ we define the *Hausdorff distance*

$$\delta(C, C') := \sup_{x \in C} \mathrm{dist}(x, C') \vee \sup_{x' \in C'} \mathrm{dist}(x', C)\ . \tag{1.10}$$

When E is compact the Hausdorff distance δ induces the Kuratowski convergence, and the class of all closed subsets of E is a compact metric space. Thereafter, taking into account that the empty set is an isolated point, we have that any sequence of non-empty closed sets admits a subsequence which converges to a non-empty closed set (see [CV], section II.4).

Definition 1.6. *Let X be the class of all closed subsets of $\mathbb{R}^n \times \mathbb{R}^n$, let $E = (\mathbb{R}^n \times \mathbb{R}^n) \cup \{\infty\}$ be the one-point compactification of $\mathbb{R}^n \times \mathbb{R}^n$ (endowed with a suitable metric) and let δ be the Hausdorff distance on the class of all closed subsets of E. Let i be the map which associates to any $C \in X$ the set $C \cup \{\infty\} \subset E$, and set*

$$d(C, C') := \delta\big(i(C), i(C')\big) \qquad \forall C, C' \in X \ . \tag{1.11}$$

It may be easily checked that d is a distance inducing the Kuratowski convergence in X, and then X is compact. In the following we identify maximal monotone functions with their graphs, and denote by $\mathscr{M}on$ the class of all maximal monotone multifunctions on $\mathbb{R}^n$. By Proposition 1.2(1) the graph of a maximal monotone function is closed subset of $\mathbb{R}^n \times \mathbb{R}^n$. Hence $\mathscr{M}on$ is a subset of X and we endow it with the distance d defined in (1.11). Then the following proposition holds:

Proposition 1.7. *$\mathscr{M}on \cup \{\emptyset\}$ is a closed subset of X. Since X is compact, this means that for every sequence $(u_h) \subset \mathscr{M}on$ either we may find a subsequence which converges to some $u \in \mathscr{M}on$, or (u_h) converges to ∞, which means that for every compact set $K \subset \mathbb{R}^n$*

$$\inf\big\{|y| : \ y \in u_h(x),\, x \in K\big\} \to +\infty \qquad \text{as } h \to +\infty.$$

Proof. It is enough to prove that for every sequence $(C_h) \subset \mathscr{M}on$ which converges to $C \neq \emptyset$ there holds $C \in \mathscr{M}on$. Let u_h be the maximal monotone multifunctions whose graphs are C_h, and let u be the multifunction whose graph is C, i.e.,

$$u(x) := \big\{y \in \mathbb{R}^n : \ (x, y) \in C\big\} \quad \forall x \in \mathbb{R}^n.$$

We first prove that u is monotone. Let be given $x_i \in \mathsf{Dm}\,(u)$ and $y_i \in u(x_i)$, $i = 1, 2$; by (1.9) we can find sequences $(x_{i,h}, y_{i,h}) \in C_h$ converging to (x_i, y_i). Hence

$$\langle y_1 - y_2, x_1 - x_2\rangle = \lim_{h\to\infty} \langle y_{1,h} - y_{2,h}, x_{1,h} - x_{2,h}\rangle \ \geq 0 \ ,$$

by the monotonicity of u_h.

In order to show that u is maximal, according to Proposition 1.2(2) it is enough to prove that $\mathsf{Dm}\,(u + I)^{-1} = \mathbb{R}^n$. By Proposition 1.2(2) and (3) the functions $(u_h + I)^{-1}$ are 1-Lipschitz functions defined on $\mathbb{R}^n$, and we claim that they converge to $(u+I)^{-1}$ uniformly on compact sets, which would imply $\mathsf{Dm}\,(u + I)^{-1} = \mathbb{R}^n$ (hence the maximality of u).

Notice that our claim is proved if we show that there always exists a subsequence which converges to some 1-Lipschitz function v uniformly on compact sets, because this would imply $v = (u + I)^{-1}$. Now, since the functions $(u_h + I)^{-1}$ are equi-continuous, it is enough to find a bounded sequence $z_h \in \mathbb{R}^n$ such that $(u_h + I)^{-1}(z_h)$ is bounded and then apply Ascoli-Arzelà theorem. Since we assumed C non-empty, we can take $(x, y) \in C$ and by (1.9) we may find $(x_h, y_h) \in C_h$ which converges to (x, y). Eventually we set $z_h := x_h + y_h$; the points z_h converge to $x + y$, and $(u + I)^{-1}(z_h) = x_h$ converge to x, and then both sequences are bounded. □

Remark. The mapping $u \mapsto Fu$ considered in Proposition 1.1 is an homeomorphism of $\mathscr{M}on$ in the space of all 1-Lipschitz functions from $\mathbb{R}^n$ into $\mathbb{R}^n$, endowed with the topology of uniform convergence on compact sets.

2. Rectifiability of the singular sets

In this section we are concerned with the structure of *singular sets* of a monotone function, that is, sets of points where the function is not univalued.

Definition 2.1. *Let u be a maximal monotone function and $k = 1, \dots, n$ an integer. We define*

$$\Sigma^k(u) := \big\{x \in \mathbb{R}^n : \ \dim u(x) \geq k\big\} \ . \tag{2.1}$$

where $\dim u(x)$ is the dimension of the set $u(x)$.

(By Proposition 1.2(1) $u(x)$ is a convex set, and thus its dimension is the dimension of the affine space spanned by it). We note that

$$\Sigma^n(u) \subset \Sigma^{n-1}(u) \subset \ldots \subset \Sigma^1(u) \ .$$

The following theorem provides an upper bound on the Hausdorff dimension of $\Sigma^k(u)$.

Theorem 2.2. *The Hausdorff dimension of the set $\Sigma^k(u)$ is at most $(n-k)$. More precisely $\Sigma^k(u)$ is countably $\mathscr{H}^{n-k}$-rectifiable, and this means that we can find countably many C^1 submanifolds $\Gamma_i \subset \mathbb{R}^n$ of dimension $(n-k)$ which cover $\mathscr{H}^{n-k}$-almost all of $\Sigma^k(u)$, i.e.,*

$$\mathscr{H}^{n-k}\Big(\Sigma^k(u) \setminus \bigcup \Gamma_i\Big) = 0 \ . \tag{2.2}$$

In particular $\Sigma^n(u)$ is at most countable.

Proof. It is clear that $\Sigma^k(u) = \Sigma^k(u+I)$. Let $\mathbf{G}(n, n-k)$ be the Grassmann manifold of (unoriented) $(n-k)$-planes in $\mathbb{R}^n$, let S be a countable dense subset of $\mathbb{R}^n$ and F a countable dense subset of $\mathbf{G}(n, n-k)$. We claim that

$$\Sigma^k(u+I) \subset (u+I)^{-1}\Big(\bigcup_{\substack{y\in S\\ P\in F}} y+P\Big) = \bigcup_{\substack{y\in S\\ P\in F}} (u+I)^{-1}(y+P) \ .$$

Indeed, if $x \in \Sigma^k(u+I)$ then $(u+I)(x)$ contains a closed convex set of dimension greater than k, so that we can find $y \in S$ and $P \in F$ such that $(y+P) \cap (u+I)(x) \neq \emptyset$. In particular, $x \in (u+I)^{-1}(y+P)$.

Since the class of countably $\mathscr{H}^{n-k}$-rectifiable sets is stable under countable union, it is enough to show that the set $(u+I)^{-1}(y+P)$ is countably $\mathscr{H}^{n-k}$-rectifiable for any $y \in \mathbb{R}^n$ and any $P \in \mathbf{G}(n, n-k)$. Since $(u+I)^{-1}$ is a Lipschitz function (cf. Proposition 1.2(3)), this follows by the fact that the image of an affine $(n-k)$-plane through a Lipschitz function is countably $\mathscr{H}^{n-k}$-rectifiable (cf. [Si], section 11). □

Remark 2.3. As a particular case of this theorem we have proved that for every (maximal) monotone function u the set of all points $x \in \mathsf{Dm}\, u$ such that $u(x)$ contains more than one element (that is, $\Sigma^1(u)$) has dimension $(n-1)$ and is σ-finite with respect to $\mathscr{H}^{n-1}$.

The proof we have given allows for a slight refinement of the statement of Theorem 2.2: for every k, the set Σ^k can be covered by countably many Lipschitz images of $(n-k)$-planes. In fact, Σ^k may be covered by countably many sets which are graphs of Lipschitz functions from $\mathbb{R}^{n-k}$ into $\mathbb{R}^k$. This strong form of Theorem 2.2 was first proved by Zajíček in [Z1].

Our proof can be extended to the case of monotone functions defined on a separable Hilbert space, while the technics used in [Z1] apply also to monotone functions on Banach spaces with separable dual. Another proof (of the strong form of Theorem 2.2) can be obtained by modifying the argument given in [AAC] for the rectifiability of the singular sets of subdifferential of semi-convex functions.

Theorem 2.2 is not optimal: Zajíček conjectured that the singular set $\Sigma^k(u)$ can be covered by countably many (c-c)-surfaces of codimension k, namely, graphs of functions $g : \mathbb{R}^{n-k} \to \mathbb{R}^k$ such that each component of g is the difference of two convex functions. In particular this would imply that Σ^k can be covered (up to an $\mathscr{H}^{n-k}$-negligible set) by countably many surfaces of codimension k and class C^2. This conjecture is proved in the case $n=2$, $k=1$ (see [Ve]) and when u agrees with the subdifferential of a convex function (see paragraph 7.9). As far as we know, the general case is still open.

3. Differentiability properties

Using Proposition 1.1 we give in this section a simple proof of the fact that a monotone function is almost everywhere differentiable. Moreover, we estimate the Hausdorff n-dimensional measure of the graphs of (maximal) monotone functions.

Proposition 3.1. *For any monotone function u and for any ball $B \subset \mathbb{R}^n \times \mathbb{R}^n$ of radius r we have*

$$\mathscr{H}^n(\Gamma u \cap B) \le 2^{n/2}\omega_n r^n \ , \tag{3.1}$$

and for every Borel set $A \subset \mathbb{R}^n$

$$\mathscr{H}^n\big(\Gamma u \cap (A \times \mathbb{R}^n)\big) \le 2^{n/2}\omega_n\big[\operatorname{diam} A + \operatorname{osc}(u, A)\big]^n \ , \tag{3.2}$$

where $\operatorname{osc}(u, A)$ stands for the supremum of $|y_1 - y_2|$ over all $y_1, y_2 \in u(A)$.

Proof. Let Fu be the function in Proposition 1.1. Estimate (3.1) follows by the fact that the Cayley transformation Φ in (1.2) is an isometry and by the fact that the Lipschitz constant of $z \mapsto (z, Fu(z))$ does not exceed $\sqrt{2}$. Inequality (3.2) follows from (3.1) because $\Gamma u \cap (A \times \mathbb{R}^n)$ is included in a ball of $\mathbb{R}^n \times \mathbb{R}^n$ of radius $r = \big[\operatorname{diam} A + \operatorname{osc}(u, A)\big]$. □

Now, using essentially Rademacher's differentiability theorem for Lipschitz functions and the area formula we can prove the following differentiability property of monotone functions.

Theorem 3.2. *Let u be a maximal monotone function and let D be the set of points x such that $u(x)$ is a singleton (that is, $u(x)$ consists of exactly one point which we still denote by $u(x)$). Then, u is differentiable at almost every $\bar{x} \in D$, that is, there exists an $n \times n$ matrix $\nabla u(\bar{x})$ such that*

$$\lim_{\substack{x\to\bar{x}\\ y\in u(x)}} \frac{y - u(\bar{x}) - \nabla u(\bar{x}) \cdot (x - \bar{x})}{|x - \bar{x}|} = 0 \ . \tag{3.3}$$

Moreover, the determinants of all minors of ∇u are integrable on every bounded set B such that $u(B)$ is bounded.

Remark 3.3. We already proved in Corollary 1.3(4) that every x where u is a singleton belongs to the interior of $\mathsf{Dm}\, u$, and that u is continuous at x, and in Theorem 2.2 we proved that $u(x)$ is a singleton for all $x \in \mathsf{Dm}\, u$ except the $(n-1)$ dimensional set $\Sigma^1(u)$.

To our knowledge, this result was originally proved in [Mig]. Since the gradient of a convex function is a monotone function (see section 7), the above theorem can be viewed as an extension of a classical differentiability theorem by Aleksandrov (see [Ale]) to monotone functions. We will see in Theorem 7.10 that Aleksandrov's theorem is an easy consequence of this result.

Proof of Theorem 3.2.

Since the graph of u is an n-dimensional Lipschitz manifold without boundary in $\mathbb{R}^n \times \mathbb{R}^n$ (Proposition 3.1), it admits an n-dimensional tangent space $\mathrm{Tan}(v)$ for all points $v \in \Gamma u$ outside an exceptional set N which is $\mathscr{H}^n$-negligible (this is an immediate and well-known corollary of Rademacher's differentiability theorem for Lipschitz functions).

Let π be the projection of the product $\mathbb{R}^n \times \mathbb{R}^n$ on the first $\mathbb{R}^n$, and let S be the set of points $v \in \Gamma u \setminus N$ so that the restriction of π to $\mathrm{Tan}(v)$ is one-to-one (in other words, so that the projection of $\mathrm{Tan}(v)$ has dimension n). The proof of Theorem 3.2 will be achieved by proving the following two claims:

(i) *if $(x, y) \in S$, then $x \in D$ and u is differentiable at x;*

(ii) *$\pi(S)$ has full measure in* $\mathsf{Dm}\, u$.

Proof of claim (i).

With no loss in generality we may assume that $x = 0$, $y = 0$. We denote by M the tangent space $\mathrm{Tan}(0, 0)$.

Assume by contradiction that $u(0)$ is not a singleton. Thus there exists $y \in u(0)$ such that $y \neq 0$, and then $u(0)$ contains the interval $[0, y]$ by Proposition 1.2(1). Hence the tangent space M contains the vector $(0, y)$, which belongs to the kernel of π, and this contradicts the assumption that π is one-to-one on M.

We denote by p be the projection of $\mathbb{R}^n \times \mathbb{R}^n$ onto M and by $\bar\pi$ the restriction of π to M. Since $\bar\pi : M \to \mathbb{R}^n$ is one-to-one, there exists an n by n matrix A such that $x \mapsto (x, Ax)$ is the inverse of $\bar\pi$. Therefore for every $v = (x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ there holds

$$(0, y - Ax) = \big(I - \bar\pi^{-1}\pi\big)v = \big(I - \bar\pi^{-1}\pi\big)\,(v - p(v)) \;. \tag{3.4}$$

Take $x_h \in \mathbb{R}^n$, $x_h \to 0$, and $y_h \in u(x_h)$. Then $y_h \to 0$ because u is continuous at 0 (cf. Corollary 1.3(4)), and if we denote by v_h the points (x_h, y_h), (3.4) becomes

$$|y_h - Ax_h| \le \big\|I - \bar\pi^{-1}\pi\big\|\, |v_h - p(v_h)| \;. \tag{3.5}$$

Now, by the definition of tangent space, we have that $|v_h - p(v_h)| = o\big(|v_h|\big)$, and it may be verified that $|v_h| = O\big(|x_h|\big)$ (indeed, if this were not true we could find a non-zero vector v in M such that $\pi(v) = 0$). Hence (3.5) yields

$$|y_h - Ax_h| = o\big(|x_h|\big) \;,$$

and (3.3) follows with $\nabla u(0) := A$.

Proof of claim (ii).

Let N_1 be the set of all points in $\Gamma u \setminus N$ such that the projection π is not one-to-one on the tangent space. Then the set $\mathsf{Dm}\, u \setminus \pi(S)$ is given by $\pi(N \cup N_1)$. Since N is $\mathscr{H}^n$-negligible, so is $\pi(N)$ (recall that π is 1-Lipschitz). Moreover by area formula (see [Fe1], 3.2.22, or [Si], Chapt. 2) we get

$$\mathscr{H}^n(\pi(N_1)) \le \int_{N_1} J(v)\, d\mathscr{H}^n(v)$$

where J stands for the determinant of the restriction of π to the tangent space $\mathrm{Tan}(v)$ of Γu in v. Since by definition the restriction of π to $\mathrm{Tan}(v)$ is never one-to-one, then $J(v) = 0$ for every $v \in N_1$, and then $\pi(N_1)$ is $\mathscr{H}^n$-negligible too, proving the claim.

Eventually we remark that the integrability of the determinants of all minors $M_\alpha \nabla u$ of ∇u follows by the area formula and (3.2), taking into account that

$$\int_{B \cap \pi(S)} \Big(\sum_\alpha \big(\det(M_\alpha \nabla u)\big)^2\Big)^{1/2} d\mathscr{L}_n \le \mathscr{H}^n\big(\Gamma u \cap (B \times u(B))\big) \;. \tag{3.6}$$

Inequality (3.6) follows by the area formula and the fact that $\pi(S)$ can be covered by a sequence of sets D_h so that the restriction of u to each D_h is Lipschitz (see [Fe1], 3.1.8, or [GMS3]). □

4. The current associated to a graph

In this section we show that the graph of a maximal monotone function u may be endowed with the structure of an n-dimensional current Tu in $\mathbb{R}^n \times \mathbb{R}^n$, and that the maps which takes each u in the current Tu is continuous. Since Proposition 1.1 shows that the graph of a maximal monotone function agrees, up to an isometry of $\mathbb{R}^n \times \mathbb{R}^n$, with the graph of a 1-Lipschitz function, we begin with recalling the canonical way to endow the graph of a Lipschitz function with the structure of an n-current.

In the rest of this section k denotes a positive integer, π denotes the projection of $\mathbb{R}^n \times \mathbb{R}^k$ onto $\mathbb{R}^n$ and $f = (f^1, \dots, f^k) : \mathbb{R}^n \to \mathbb{R}^k$ is a Lipschitz function. An n-dimensional current in $\mathbb{R}^n \times \mathbb{R}^k$ is a continuous linear functional on the space $\mathscr{D}^n(\mathbb{R}^n \times \mathbb{R}^k)$ of all smooth n-forms in $\mathbb{R}^k$ with compact support (we refer to [Fe1] or to [Si], [Mo] for the general results and notations in current theory).

Definition 4.1. *Given an n-dimensional Lipschitz submanifold Γ of some euclidean space which is closed, oriented, and without boundary, we define* *a current T by integrating n-forms on Γ. More precisely, since Γ admits a tangent space $\mathrm{Tan}(\Gamma, y)$ for all $y \in \Gamma$ except an $\mathscr{H}^n$-negligible set N (this is an immediate corollary of Rademacher's differentiability theorem for Lipschitz functions), for all $y \in \Gamma \setminus N$ we can choose an orthonormal basis $\{e_1, \dots, e_n\}$ of $\mathrm{Tan}(\Gamma, y)$ which preserves the orientation, and we denote by $\sigma(y)$ the n-covector $e_1 \wedge \dots \wedge e_n$. Thus we set*

$$T(\omega) := \int_\Gamma \langle \omega, \sigma \rangle d\mathscr{H}^n \qquad \forall \omega \in \mathscr{D}^n(\mathbb{R}^n \times \mathbb{R}^k) \ . \tag{4.1}$$

The graph Γf of a Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}^k$ is an n-dimensional Lipschitz submanifold of $\mathbb{R}^n \times \mathbb{R}^k$, and we may orient it so that the orientation induced on $\mathbb{R}^n$ by the projection is the usual one (that is, for every point $x \in \Gamma f$ where f is differentiable we choose a basis $\{e_1, \dots, e_n\}$ of the tangent space $\mathrm{Tan}(\Gamma f, x)$ so that $\{\pi(e_1), \dots, \pi(e_n)\}$ is a positively oriented basis of $\mathbb{R}^n$). We denote by Tf the current associated to the graph of f.

The linear functional T in (4.1) is an n-current without boundary by Stokes's theorem; more precisely in Geometric Measure Theory T is called a multiplicity one integral current without boundary. For any Borel set $B \subset \mathbb{R}^n \times \mathbb{R}^k$ we have

$$|T|(B) = \mathscr{H}^n(\Gamma \cap B) \ , \tag{4.2}$$

where T is viewed as a vector measure on $\mathbb{R}^n \times \mathbb{R}^k$, and the positive measure $|T|$ is its total variation.

In order to prove convergence results, the following representation of Tf is often useful. Let's denote a point of $\mathbb{R}^n \times \mathbb{R}^k$ as $y = (y_1, \dots, y_{n+k})$, where $(y_1, \dots, y_n) \in \mathbb{R}^n$ and $(y_{n+1}, \dots, y_{n+k}) \in \mathbb{R}^k$, and let $I(n, n+k)$ be the set of all strictly increasing functions

$$\alpha : \{1, \dots, n\} \to \{1, \dots, n+k\} \ .$$

The set of all n-covectors $dy_{\alpha(1)} \wedge \dots \wedge dy_{\alpha(n)}$ with $\alpha \in I(n, n+k)$ is a basis of the space of all n-covectors in $\mathbb{R}^n \times \mathbb{R}^k$. Hence every n-form $\omega \in \mathscr{D}^n(\mathbb{R}^n \times \mathbb{R}^k)$ may be canonically decomposed as

$$\omega = \sum_{\alpha \in I(n,n+k)} \omega_\alpha(y) \, dy_{\alpha(1)} \wedge \dots \wedge dy_{\alpha(n)} \ . \tag{4.3}$$

For every $\alpha \in I(n, n+k)$ we denote by $J_\alpha f$ the Jacobian determinant of the map

$$x = (x_1, \dots, x_n) \longmapsto \big(x_{\alpha(1)}, \dots, x_{\alpha(k)}, f_{\alpha(k+1)-n}, \dots, f_{\alpha(n)-n} \big) \ ,$$

where k is the maximal h such that $\alpha(h) \le n$. Then the following representation formula holds:

Proposition 4.2. *Let $f : \mathbb{R}^n \to \mathbb{R}^k$ be a Lipschitz functions and let Tf be the n-current associated to the graph of f. Then*

$$Tf(\omega) = \sum_\alpha \int_{\mathbb{R}^n} \omega_\alpha(x, f(x)) \, J_\alpha f(x) \, d\mathscr{L}_n(x) \qquad \forall \, \omega \in \mathscr{D}^n(\mathbb{R}^n \times \mathbb{R}^k) \ . \tag{4.4}$$

Moreover, for any Borel set $B \subset \mathbb{R}^n \times \mathbb{R}^k$ there holds

$$|Tf|(B) = \mathscr{H}^n(B \cap \Gamma f) = \int_{\pi(B)} \Big(\sum_\alpha (J_\alpha f)^2 \Big)^{1/2} d\mathscr{L}_n \ . \tag{4.5}$$

Proof. Equality (4.4) is essentially a consequence of the area formula, which allows to carry the integration from Γf to the domain of f (see for instance [GMS1], [GMS3]). The second equality in (4.5) easily follows from (4.4). □

Remark. Notice that the integral representation of Tf given in (4.1) (or the one given in (4.4) as well) provides an extension of Tf to compactly supported n-forms with bounded Borel coefficients.

The continuity properties of the mapping T are given in the following proposition.

Proposition 4.3. *Let $f_h : \mathbb{R}^n \to \mathbb{R}^k$ be equi-Lipschitz functions which uniformly converge on compact sets to some Lipschitz function f. Then Tf_h converge to Tf in the sense of currents, i.e.,*

$$Tf_h(\omega) \to Tf(\omega) \quad \forall \, \omega \in \mathscr{D}^n(\mathbb{R}^n \times \mathbb{R}^k) \tag{4.6}$$

(In fact this holds for any continuous n-form ω with compact support).

Proof. Taking formula (4.4) into account, this proposition follows from the continuity of determinants of minors of derivatives as maps from $W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}^n, \mathbb{R}^k)$, endowed with the weak* topology, into $L^\infty_{\mathrm{loc}}(\mathbb{R}^n)$, endowed with the weak* topology (see [Da], section 4.2.2, Theorem 2.6). □

Using the results above and Proposition 3.1 we can define the current Tu for every maximal monotone function u.

Definition 4.4. *Let Φ be the isometry of $\mathbb{R}^n \times \mathbb{R}^n$ given in (1.2), and u a maximal monotone function. By Proposition 1.1, $\Phi(\Gamma u)$ is the graph of the Lipschitz function $Fu : \mathbb{R}^n \to \mathbb{R}^n$ and is oriented as in Definition 4.1, and* *then* Γu *is a Lipschitz submanifold of* $\mathbb{R}^n \times \mathbb{R}^k$ *equipped with the orientation induced by* Φ *(that is, for* $\mathscr{H}^n$ *a.e.* $y \in \Gamma u$ *we choose an orthonormal basis* $\{e_1, \dots, e_n\}$ *of* $\mathrm{Tan}(\Gamma u, y)$*, so that* $\{\pi\Phi(e_1), \dots, \pi\Phi(e_n)\}$ *is a positively oriented basis of* $\mathbb{R}^n$*). We denote by* Tu *the current associated to the graph of* u*.*

It may be checked that this definition of current associated to the graph of a function is consistent with the one given in Definition 4.1: one has simply to verify that for every function u which is Lipschitz and monotone, the orientation of Γu given in Definition 4.1 agrees with the one given in before.

Remark 4.5. The currents Tu and $T(Fu)$ are connected by the following (immediate) formula:

$$Tu(\omega) = T(Fu)(\omega(\Phi)) \qquad \forall\, \omega \in \mathscr{D}^n(\mathbb{R}^n \times \mathbb{R}^n)\ , \tag{4.8}$$

and then $|Tu|(B) = |T(Fu)|\,(\Phi(B))$ for all Borel sets $B \subset \mathbb{R}^n \times \mathbb{R}^n$. In addition, we observe that T is as a local operator, i.e., for any pair of maximal monotone functions u, v which agree on some open set Ω we have $Tu(\omega) = Tv(\omega)$ for all smooth n-forms ω with support included in $\Omega \times \mathbb{R}^n$ (more generally, when u and v agree on some Borel set B, then $Tu(\omega) = Tv(\omega)$ for every compactly supported n-form ω with bounded Borel coefficients which are 0 out of $B \times \mathbb{R}^n$).

The operator T maps the space $\mathscr{M}\!on$ of all maximal monotone function into the space $\mathscr{D}_n(\mathbb{R}^n \times \mathbb{R}^n)$ of all n-currents on $\mathbb{R}^n \times \mathbb{R}^n$ continuously. More precisely, the following statement holds:

Theorem 4.6. *If* (u_h) *is a sequence of maximal monotone functions converging to* u *in the sense specified in Definition 1.6, then* Tu_h *converges to* Tu *in the sense of currents, i.e.,*

$$\lim_{h\to\infty} Tu_h(\omega) = Tu(\omega) \quad \forall\, \omega \in \mathscr{D}^n(\mathbb{R}^n \times \mathbb{R}^n)\ .$$

Proof. We know from the proof of Proposition 1.7 that the functions $(u_h+I)^{-1}$ converge in $\mathbb{R}^n$ to $(u+I)^{-1}$ uniformly on compact sets; then the maps Fu_h given in (1.3) converge to Fu uniformly on compact sets, and since their Lipschitz constants do not exceed 1, we may apply Proposition 4.3, which yields that the currents $T(Fu_h)$ converge to $T(Fu)$. Hence the currents Tu_h converge to Tu by formula (4.8). $\square$

5. Distributional derivatives

In this section we are concerned with the structure of the distributional derivative of a monotone function.

Let be given a monotone function u, and an open set Ω relatively compact in the interior of $\mathsf{Dm}\, u$. By Corollary 1.3(3) and Remark 2.3, u is bounded and almost everywhere univalued in Ω, and then we may consider it as an element of $L^\infty(\Omega, \mathbb{R}^n)$. The following theorem holds:

Proposition 5.1. *The function* u*, viewed as an element of* $L^\infty(\Omega, \mathbb{R}^n)$*, belongs to* $BV(\Omega, \mathbb{R}^n)$*. Moreover*

$$\int_\Omega |Du| \le C_n \big[\mathrm{diam}\,\Omega + \mathrm{osc}(u, \Omega)\big]^n\ , \tag{5.1}$$

where C_n *is a constant which depends on* n *only, and* $\mathrm{osc}(u, \Omega)$ *is defined as in Proposition 3.1.*

Proof. Take integers i and j such that $1 \le i, j \le n$. We want to prove that the distributional derivative $D_i u_j$ is (represented by) a Borel measure on Ω.

We denote points of $\mathbb{R}^n \times \mathbb{R}^n$ as (x, y) with $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$. We denote by dx the n-covector $dx_1 \wedge \ldots \wedge dx_n$, and by $\widehat{dx_i}$ the $(n-1)$-covector $dx_1 \wedge \ldots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \ldots \wedge dx_n$. Take a function $\phi \in C^1_c(\Omega)$, and let ω be the form given by

$$\omega(x, y) := \phi(x)\, y_j\, \widehat{dx_i}\ . \tag{5.2}$$

Then ω is an $(n-1)$-form of class C^1 with compact support in $\mathbb{R}^n \times \mathbb{R}^n$. Since the graph of u is a Lipschitz manifold without boundary, we may apply Stokes's theorem to get

$$\int_{\Gamma u} d\omega = 0\ . \tag{5.3}$$

By (5.2) we obtain that

$$d\omega = (-1)^{i-1}(D_i\phi)\, y_j\, dx + \phi\, dy_j \wedge \widehat{dx_i}$$

and then (5.3) yields

$$\int_\Omega (D_i\phi)\, u_j\, d\mathscr{L}_n = \int_{\Gamma u} (D_i\phi)\, y_j\, dx = (-1)^i \int_{\Gamma u} \phi\, dy_j \wedge \widehat{dx_i}\ . \tag{5.4}$$

Thus, denoting by M the intersection of Γu with $\Omega \times \mathbb{R}^n$,

$$\Big| \int_\Omega (D_i\phi)\, u_j\, d\mathscr{L}_n \Big| \le \|\phi\|_\infty \mathscr{H}^n(M)\ . \tag{5.5}$$

Now, since u is bounded on Ω, by inequality (3.2) we obtain that the $\mathscr{H}^n(M)$ is finite, and then (5.5) implies that the partial derivative $D_i u_j$ is a bounded measure on Ω with total variation

$$\|D_i u_j\| \le \mathscr{H}^n(M) \le 2^{n/2}\omega_n\big[\operatorname{diam}\Omega + \operatorname{osc}(u,\Omega)\big]^n ,$$

and (5.1) immediately follows for a suitable constant C_n. □

Remark 5.2. A careful examination of this proof shows that inequality (5.1) can be improved; in fact there holds

$$\int_\Omega |Du| \le C_n\big(\operatorname{diam}\Omega\big)^{n-1}\operatorname{osc}(u,\Omega) ,$$

where C_n is a constant which depends on n only.

Proposition 5.1 is a particular case of the following more general result (which may be proved in the same way). Let $u : \Omega \to \mathbb{R}^m$ a bounded function, and let T a rectifiable n-current with multiplicity 1 in $\Omega \times \mathbb{R}^m$ and boundary with finite mass which is supported on a $\mathscr{H}^n$-rectifiable set S. For every $x \in \Omega$, let S_x be the set of all $y \in \mathbb{R}^m$ such that $(x, y) \in S$. If S_x consists of the point $u(x)$ only for a.e. $x \in \Omega$, then u belongs to $BV(\Omega, \mathbb{R}^m)$.

The following theorem characterizes monotone functions in term of distributional derivatives.

Theorem 5.3. *Let Ω be an open convex set in $\mathbb{R}^n$.*

(i) *If u is a (maximal) monotone function such that $\mathsf{Dm}\, u \supset \Omega$, then $u \in BV_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$, and Du is a positive (matrix-valued and locally bounded) measure.*

(ii) *Conversely, if $u \in L^1_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$ and Du is a positive (matrix-valued) distribution on Ω, then there exists a maximal monotone function v such that $\mathsf{Dm}\, v \supset \Omega$ and $v = u$ a.e. in Ω. Therefore Du is a locally bounded measure by statement* (i).

Remark 5.4. We remark that when Λ is a positive matrix-valued distribution, then the symmetric part of Λ, $\frac{1}{2}\big(\Lambda + \Lambda^{\mathrm{t}}\big)$, is a locally bounded measure, but nothing can be said about the skew-symmetric part of Λ. Statement (ii) of Theorem 5.3 shows that when Λ is a (distributional) derivative, then also the skew-symmetric part of Λ is a locally bounded measure (the proof we give actually works for derivatives of locally summable functions, but this hypothesis may be easily removed).

Remark 5.5. For every function $u \in L^1_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$, let $\mathcal{E}u$ denote the symmetric part of the derivative, that is, $\mathcal{E}u = \frac{1}{2}\big(Du + D^{\mathrm{t}}u\big)$. By Korn's inequalities, if $\mathcal{E}u$ belongs to L^p for some $p \in]1, +\infty[$, then Du belongs to L^p too.

Korn's inequality does not hold when $p = 1$: there exists a function u, smooth out of a point singularity, such that $\mathcal{E}u \in L^1$ and $Du \notin L^1$ (see for instance Example 7.7 in [ACD]). On the other hand we have just stated that this can never happen when u is monotone and bounded. Therefore the following question arises: is it possible to prove some Korn type inequality for monotone functions? in other words, we want to find a constant C (depending on the set Ω only) such that $\|Du\| \le C\|\mathcal{E}u\|$ for every monotone function u with $\mathsf{Dm}\, u \supset \Omega$. A simple application of Baire's theorem shows that such a constant C exists if (and only if) $\|Du\|$ is finite whenever $\|\mathcal{E}u\|$ is finite.

Proof of Theorem 5.3.

Statement (i) follows immediately from Proposition 5.1.

The converse is proved as follows. For every $\varepsilon > 0$, let Ω_ε the set of all $x \in \Omega$ such that the ball $B(x, \varepsilon)$ is relatively compact in Ω (then Ω_ε is convex and open), and take positive smooth mollifiers ρ_ε so that ρ_1 has support included in the ball $B(0, 1)$.

The mollified function $u_\varepsilon := u * \rho_\varepsilon$ is well-defined in Ω_ε, and $Du_\varepsilon = Du * \rho_\varepsilon$ is a positive matrix-valued smooth function on Ω_ε. Therefore u_ε is monotone: indeed, for every x, h such that x and $x + h$ belong to Ω_ε, there holds

$$\langle u_\varepsilon(x+h) - u_\varepsilon(x), h\rangle = \int_0^1 \big\langle Du_\varepsilon(x+th)\cdot h,\, h\big\rangle\, dt \ge 0 .$$

Now, for every $\varepsilon > 0$, let v_ε a maximal monotone function which includes u_ε. By Proposition 1.7, possibly passing to a subsequence we may assume that the functions v_ε converge to some maximal monotone function v (in the sense given in Definition 1.6). Moreover $u(x) \in v(x)$ for every x such that $u_\varepsilon(x) \to u(x)$ (cf. condition (1.5)), and since this happens for every Lebesgue point x of u, we deduce that $u(x) \in v(x)$ for a.e. $x \in \Omega$. □

Now we want to examine the structure of the distributional derivative of u. In order to do this, we recall some well-known facts about (vector-valued) BV functions.

5.6. *Decomposition of derivatives of BV functions*

Let Ω be an open subset of $\mathbb{R}^n$, k a positive integer, and v a function in $BV(\Omega, \mathbb{R}^k)$. Then the distributional derivative Dv is (represented by) a bounded Borel measure on Ω which takes values in $n \times k$ matrices.

According to [Am1], [Am2], Dv can be split into three parts: the *absolutely continuous part*, the *jump part*, the *Cantor part*. Indeed, let $D_a v$, $D_s v$ be respectively the absolutely continuous and the singular part of Dv with respect to Lebesgue measure, let Sv be the complement of Lebesgue points of v. Then, the jump part D_jv is defined as the restriction of D_sv to Sv, and the Cantor part D_cv is defined as the restriction of D_sv to $\Omega\setminus Sv$. In this way we have $Dv = D_av + D_cv + D_jv$.

The function v is almost everywhere approximately differentiable (see [CZ], or [EG], section 6.1), and the approximate gradient ∇v agrees almost everywhere with the density of D_av (and then of Dv) with respect to Lebesgue measure, that is, $D_av = \nabla v\cdot\mathscr{L}_n$.

The set Sv is countably $\mathscr{H}^{n-1}$-rectifiable, and for $\mathscr{H}^{n-1}$-almost every $x\in Sv$, v admits one sided traces $v^+(x)$, $v^-(x)$ with respect to a suitable direction $\nu(x)$, i.e., the functions v_ρ defined by $v_\rho(y) := v(x+\rho y)$ when $x+\rho y\in\Omega$, and 0 otherwise, converge in $L^1_{\rm loc}(\mathbb{R}^n,\mathbb{R}^k)$ as $\rho\searrow 0$ to the step function

$$v_0(y) := \begin{cases} v^+(x) & \text{if } \langle y,\nu(x)\rangle > 0,\\[2ex] v^-(x) & \text{if } \langle y,\nu(x)\rangle < 0.\end{cases} \tag{5.6}$$

Vol'pert proved in [Vo] that the jump part of the derivative D_jv can be written in term of the traces v^+, v^-, and of the direction ν (see also [Fe1], 4.5.9):

$$D_jv = (v^+ - v^-)\otimes\nu\cdot\mathscr{H}^{n-1}\llcorner Su\ . \tag{5.7}$$

Very little is presently known about the Cantor part of derivative. In [Am1] the second author proved that $|D_cv|(B) = 0$ for any Borel set B such that $\mathscr{H}^{n-1}(B) < +\infty$. Hence the $(n-1)$-dimensional part of the derivative is given by D_jv only, and the derivative does not charge any set which is $\mathscr{H}^{n-1}$-negligible.

Remark 5.7. When u is a maximal monotone function, then u belongs to $BV(\Omega,\mathbb{R}^n)$ for every open set Ω relatively compact in $\mathsf{Dm}\,u$ (Proposition 5.1). Moreover $Su = \Sigma^1(u)\cap\Omega$ and $\big[u^-(x),u^+(x)\big] = u(x)$ for $\mathscr{H}^{n-1}$-almost every $x\in Su$ (in fact, for every $x\in\Sigma^1(u)\setminus\Sigma^2(u)$).

The inclusion $Su\subset\Sigma^1(u)$ is immediate because u is continuous in the complement of $\Sigma^1(u)$ (cf. Definition 2.1 and Corollary 1.3(4)), while the opposite inclusion is slightly more delicate. Take indeed $\bar x\in(\Sigma^1(u)\setminus Su)\cap\Omega$, and let $\bar y$ be the approximate value of u in $\bar x$. We may assume $\bar x = 0$ and $\bar y = 0$. Since $u(0)$ is not a singleton, we may find $y'\in u(0)$ such that $y'\neq 0$. Let C be the cone of all x such that $\langle x,y'\rangle\geq|x|\,|y'|/2$. The monotonicity of u yields $\langle y,x\rangle\geq\langle y',x\rangle$ for every $x\in\mathbb{R}^n$ and $y\in u(x)$, and if in addition x belongs to C then

$$|y|\geq\frac{\langle y,x\rangle}{|x|}\geq\frac{\langle y',x\rangle}{|x|}\geq\frac{|y'|}{2}>0\ .$$

Since C has positive density in 0, this contradicts the assumption that u has approximate limit equal to 0 at 0.

5.8. *Rank-one property of derivatives of BV functions*

Identity (5.7) shows that the density of the jump part of the derivative D_jv with respect to its total variation $|D_jv|$ is a rank-one matrix for $|D_jv|$-almost every point. The first author proved in [A1] that the same property holds for the Cantor part too, that is, the density of the singular part of the derivative D_sv with respect to its total variation $|D_sv|$ is a rank-one matrix for $|D_sv|$-almost every point.

Using the rank-one property and a blow-up argument one can show that any vector BV function asymptotically depends only on one variable in the neighborhood of $|D_sv|$-almost every point. In the analysis of functionals defined in spaces of vector-valued BV functions this information is crucial (see [AD], [FM]).

The original proof of the rank-one property for a general BV function is very long and complicated. In Theorem 5.10 below we present a simple proof which works for monotone functions (and then for gradients of a convex functions as well, cf. section 7). A partial result in this direction was also obtained in [AG]. Our proof of the rank property is mainly based on the estimate (3.4) and on the following version of Reshetnyak's continuity theorem (see [Re1], or the appendix of [LM]):

Theorem 5.9. *Let $\Omega\subset\mathbb{R}^N$ be an open set and let (μ_h) be a sequence of vector measures in Ω with p components, weakly converging to μ in Ω, and assume that*

$$\lim_{h\to\infty}|\mu_h|(\Omega) = |\mu|(\Omega)\ .$$

Then

$$\lim_{h\to\infty}\int_\Omega g\Big(x,\frac{\mu_h}{|\mu_h|}(x)\Big)\,d|\mu_h|(x) = \int_\Omega g\Big(x,\frac{\mu}{|\mu|}(x)\Big)\,d|\mu|(x)$$

for any bounded continuous function $g:\Omega\times S^{p-1}\to\mathbb{R}$.

Here $\mu/|\mu|$ stands for the density of the measure μ with respect to its total variation $|\mu|$.

Theorem 5.10. *Let u be a monotone function on $\mathbb{R}^n$, and let Ω be an open set such that u belongs to $BV(\Omega,\mathbb{R}^n)$. Then the density of the singular part of the derivative D_su with respect to its total variation $|D_su|$ is a rank-one matrix for $|D_su|$-almost every $x\in\Omega$.*

Proof. We assume $\Omega = \mathbb{R}^n$, the proof of the general case being a straightforward generalization.

For all $\varepsilon > 0$ let ρ_ε be positive smooth mollifiers. Set $u_\varepsilon := u*\rho_\varepsilon$, and let f be the density of Du with respect to $|Du|$. For any $n\times n$ matrix A we set

$$M(A) := \Big(\sum_B|\det B|^2\Big)^{1/2} \tag{5.8}$$

where the sum is taken over all 2×2 minors B of the matrix A. We remark that a matrix A has rank 1 or 0 if and only if $M(A) = 0$.

The proof depends on the following two properties:

(a) for any ball B such that $|Du|(\partial B) = 0$ the total variations $|Du_\varepsilon|(B)$ converge to $|Du|(B)$;
(b) for any compact set $K \subset \mathbb{R}^n$ there exists a constant C_K such that

$$\int_K M(Du_\varepsilon)\, d\mathscr{L}_n \le C_K \qquad \forall \varepsilon \in (0,1)\ .$$

The first property is well-known. Indeed, if $B = B(x,r)$, and we take $\bar r$ so that $B(0,\bar r)$ includes the support of ρ_1, it may be easily verified that

$$\int_B |Du_\varepsilon| = \int_B |Du * \rho_\varepsilon| \le \int_{B(x,r+\varepsilon\bar r)} |Du|\ .$$

We deduce that

$$\limsup_{\varepsilon\to 0} \int_B |Du_\varepsilon| \le |Du|(\overline{B}) = |Du|(B)\ ,$$

and the liminf inequality follows by the lower semicontinuity of $v \mapsto |Dv|(B)$ in the $L^1_{\mathrm{loc}}(B)$ convergence.

The second property follows by the estimate (3.2) on the area of the graph of a monotone function by remarking that the functions u_ε are smooth, monotone and uniformly bounded on every compact set of $\mathbb{R}^n$.

Now, let us fix a ball B such that $|Du|(\partial B) = 0$ and a continuous function $g \in C_c(B)$. By (a) the measures Du_ε converge to Du in variation, as matrix-valued measures on the ball B. Since the function $(x, A) \mapsto g(x)\sqrt{M(A)}$ is uniformly continuous on $B \times \mathbb{R}^{n\times n}$, and 1-homogeneous in the second variable, by Theorem 5.9 we infer

$$\lim_{\varepsilon\to 0} \int_B g\sqrt{M(\nabla u_\varepsilon)}\, d\mathscr{L}_n = \int_B g\sqrt{M(f)}\, d|Du|\ . \tag{5.9}$$

Since (5.9) holds for every $g \in C_c(B)$, we obtain that

$$\sqrt{M(\nabla u_\varepsilon)} \cdot \mathscr{L}_n \to \sqrt{M(f)} \cdot |Du| \qquad \text{weakly*} \tag{5.10}$$

as positive Borel measures on B.

On the other hand, by (b) the functions $\sqrt{M(\nabla u_\varepsilon)}$ are uniformly bounded in $L^2(B)$. Therefore (5.10) implies that the measure $\sqrt{M(f)} \cdot |Du|$ belongs to $L^2(B)$, and then it has no singular part. This means that $\sqrt{M(f)}$ is almost everywhere 0 on B with respect to $|D_s u|$, and then $f(x)$ is a rank-one matrix for $|D_s u|$-almost every $x \in B$. Since the collection of all open balls B such that $|Du|(\partial B) = 0$ covers $\mathbb{R}^n$, the proof is achieved. □

In the last part of this section we briefly discuss "the right way" to extend the definition of Jacobian determinant to monotone functions. Taking into account section 4, the "natural" approach is to use the fact that the graph of u is a Lipschitz manifold without boundary (clearly this construction works as well for the determinant of every minor of the derivative of u).

Let Ω a bounded open convex set in $\mathbb{R}^n$; we denote by J the operator which takes every function $u \in C^1(\Omega, \mathbb{R}^n)$ in the Jacobian determinant $Ju := \det(\nabla u)$, considered as a real valued locally bounded measure in $\mathscr{M}_{\mathrm{loc}}(\Omega)$.

We denote by $\mathscr{Mon}(\Omega)$ the class of all functions on Ω which agree almost everywhere with some (maximal) monotone function whose domain includes Ω. Since every monotone function is locally bounded in the interior of its domain (cf. Corollary 1.3(3)), we may regard any function in $\mathscr{Mon}(\Omega)$ as an element of $L^\infty_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$.

Thus J is well-defined on the subset $C^1 \cap \mathscr{Mon}(\Omega)$, and the following theorem provides an extension of J to the whole of $\mathscr{Mon}(\Omega)$.

Theorem 5.11. *The operator $J : C^1 \cap \mathscr{Mon}(\Omega) \to \mathscr{M}_{\mathrm{loc}}(\Omega)$ admits a unique continuous extension, which we still denote by J, from $\mathscr{Mon}(\Omega)$, endowed with the L^1_{loc} topology, to $\mathscr{M}_{\mathrm{loc}}(\Omega)$, endowed with the weak* topology.*

For every $u \in \mathscr{Mon}(\Omega)$, Ju is a locally bounded positive measure, and if $\bar u$ is any maximal monotone function whose domain includes Ω and which agrees a.e. with u in Ω, then for every Borel set B relatively compact in Ω the following generalization of the area formula holds:

$$Ju(B) = \int_{\Gamma \bar u} 1_{B\times\mathbb{R}^n}\, dy_1 \wedge \ldots \wedge dy_n = |\bar u(B)|\ . \tag{5.11}$$

Eventually, J is a local operator, that is, for every couple of functions $u, v \in \mathscr{Mon}(\Omega)$ which agree a.e. on the open set A, the measures Ju and Jv agree on A.

Before proving this theorem, we remark that it provides a definition of Jacobian determinant for maximal monotone functions so that area formula holds (cf. (5.11)), and then the following definition is clearly justified:

Definition 5.12. *For every function u in $\mathscr{Mon}(\Omega)$ we define the weak Jacobian of u as the positive measure $Ju \in \mathscr{M}_{\mathrm{loc}}(\Omega)$, where J is the extended operator given in Theorem 5.11, that is*

$$Ju(B) := |\bar u(B)| \qquad \textit{for every Borel set } B \subset \Omega.$$

Proof of Theorem 5.11:

The uniqueness of the extension follows from the density of $C^1 \cap \mathscr{M}on(\Omega)$ in $\mathscr{M}on(\Omega)$, which can be easily proved by mollification (see also Theorem 6.2 and Remark 6.3).

For every $u \in \mathscr{M}on(\Omega)$, let $\bar{u}$ be a maximal monotone f unction whose domain includes Ω, and which agrees a.e. with u, and let μ_u be the set function given by

$$\mu_u(B) := \int_{\Gamma\bar{u}} 1_{B\times\mathbb{R}^n}\, dy_1 \wedge \ldots \wedge dy_n \quad \forall B \subset\subset \Omega\,. \tag{5.12}$$

The μ_u is a well-defined locally bounded measure (use inequality (3.2) recalling that u is bounded on any set relatively compact in Ω by Corollary 1.3(3)).

First of all we claim that given $u, v \in \mathscr{M}on(\Omega)$ which agree a.e. on a convex open set $A \subset \Omega$, then the measures μ_u and μ_v agree on A (this shows implicitly that μ_u does not depend on the choice of the maximal monotone function $\bar{u}$, and that $u \mapsto \mu_u$ is a local operator). More precisely, we claim that given maximal monotone functions $\bar{u}$ and $\bar{v}$ corresponding to u and v respectively, then $\bar{u}(x) = \bar{v}(x)$ for every $x \in A$. Indeed, since $u = v$ a.e. in A, then $\bar{u}(x) = \bar{v}(x)$ a.e. in A (recall that $\bar{u}$ and $\bar{v}$ are a.e. univalued), and then $\bar{u}(x) = \bar{v}(x)$ for every $x \in A$ by Corollary 1.5.

Now, since $\mu_u = Ju$ for every function u of class C^1, the existence of the continuous extension and the first identity in (5.11) will be proved once we have shown that

$$u \mapsto \int_\Omega g\, d\mu_u$$

is a continuous map on $\mathscr{M}on(\Omega)$ for every continuous test function g with compact support in Ω.

Now, let (u_h) a sequence converging to u in $\mathscr{M}on(\Omega)$, and let $\bar{u}_h$ the corresponding maximal monotone functions. Possibly passing to a subsequence we may assume that the functions u_h converge to u a.e. in Ω, and the functions $\bar{u}_h$ converge to some maximal monotone function v (in the sense given in Proposition 1.7), and then $u(x) \in v(x)$ for a.e. x. Therefore the domain of v includes Ω and $u(x) = v(x)$ a.e. in Ω.

Hence by (5.12) and Theorem 4.6 we obtain that

$$\int_\Omega g y d\mu_{u_h} = \int_{\Gamma\bar{u}_h} g(x)\, dy_1 \wedge \ldots \wedge dy_n = \langle T\bar{u}_h, \omega\rangle \to$$

$$\to \langle Tv, \omega\rangle = \int_{\Gamma v} g(x)\, dy_1 \wedge \ldots \wedge dy_n = \int_\Omega g\, d\mu_u$$

where $\omega(x,y) := g(x)\, dy_1 \wedge \ldots \wedge dy_n$ (actually ω does not have compact support in $\mathbb{R}^n \times \mathbb{R}^n$, and the above convergence can be derived from Theorem 4.6 by recalling that the functions $\bar{u}_n$ are uniformly bounded on the support of the test function g).

Eventually we prove the second equality in (5.11). We denote by π the function which takes each $(x,y) \in \Gamma\bar{u}$ into $y \in \mathbb{R}^n$, then π is a Lipschitz function from the n-dimensional oriented Lipschitz manifold without boundary $\Gamma\bar{u}$ into $\mathbb{R}^n$, and for every Borel set $B \subset \Omega$ there holds

$$\int_{\Gamma\bar{u}} 1_{B\times\mathbb{R}^n}\, dy_1 \wedge \ldots \wedge dy_n = \int_{\mathbb{R}^n} \deg(\pi, B\times\mathbb{R}^n, y)\, d\mathscr{L}_n(y)\,, \tag{5.13}$$

where $\deg(\pi, B\times\mathbb{R}^n, y)$ is the Brower degree of the function π relative to the open set $B \times \mathbb{R}^n$ and the point y (see for instance [Fe1], 4.1.26, or [GMS3]). By Theorem 2.2, $\Sigma^1(\bar{u}^{-1})$ is Lebesgue negligible, and then $\pi^{-1}(y)$ consists of one point for a.e. y. Hence $\deg(\pi, B\times\mathbb{R}^n, y) = 0$ for every $y \notin \bar{u}(B)$ and $\big|\deg(\pi, B\times\mathbb{R}^n, y)\big| = 1$ for a.e. $y \in \bar{u}(B)$, and taking into account our choice of the orientation of $\Gamma\bar{u}$ (see Definition 4.4), we obtain $\deg(\pi, B\times\mathbb{R}^n, y) = 1$ for a.e. $y \in \bar{u}(B)$. Thus $\deg(\pi, B\times\mathbb{R}^n, \cdot)$ agrees a.e. with the characteristic function of the set $\bar{u}(B)$. Hence (5.13) becomes

$$\int_{\Gamma\bar{u}} 1_{B\times\mathbb{R}^n}\, dy_1 \wedge \ldots \wedge dy_n = |\bar{u}(B)|\,. \qquad \square$$

Remark 5.13. Notice that the weak Jacobian Ju may not agree with the pointwise determinant $\det(\nabla u)$. Take indeed $u(x) := x/|x|$ on $\mathbb{R}^2$: by approximating u with smooth monotone functions and using the continuity of J, we obtain that Ju is the Dirac mass concentrated in 0. On the other hand a direct computation shows that $\det(\nabla u) = 0$ almost everywhere. Nevertheless the pointwise determinant is connected to the weak Jacobian by the following relation:

Proposition 5.14. *For every monotone function $u \in \mathscr{M}on(\Omega)$, the absolutely continuous part of the measure Ju with respect to Lebesgue measure is represented by the pointwise determinant* $\det(\nabla u)$ *(recall that u is almost everywhere differentiable by Theorem 3.2).*

This statement was originally proved by S. Müller for the distributional determinant $\operatorname{Det} Du$ of a Sobolev function (provided that $\operatorname{Det} Du$ is defined and is a measure), but the proof in that case is remarkably more difficult (see [Mu2], Theorem 1 and following remarks).

Proof. Since u is of class $BV_{\rm loc}(\Omega)$, then we may find an increasing sequence of Borel sets D_h which cover almost all of Ω and such that u agrees in D_h

with a function u_h of class C^1 (see for instance [EG], section 6.6). Hence for every h, $\nabla u = \nabla u_h$ a.e. in D_h, and $\Gamma u \cap (D_h \times \mathbb{R}^n) = \Gamma u_h \cap (D_h \times \mathbb{R}^n)$, and then for every Borel set $B \subset D_h$ there holds

$$\begin{aligned}\int_B \det(\nabla u)\, d\mathscr{L}_n &= \int_B \det(\nabla u_h)\, d\mathscr{L}_n \\ &= \int_{\Gamma u_h} 1_{B\times\mathbb{R}^n}\, dy_1 \wedge \ldots \wedge dy_n = \int_{\Gamma u} 1_{B\times\mathbb{R}^n}\, dy_1 \wedge \ldots \wedge dy_n = Ju(B)\end{aligned}$$

(the second equality holds because u_h is a function of class C^1, while the fourth one follows from the definition of Ju). Hence for every Borel set $B \subset D_h$ there holds $\int_B \det(\nabla u)\, d\mathscr{L}_n = Ju(B)$. Moreover this equality may be extended to every Borel set B included in the union of the sets D_h, and since this union cover almost all of Ω the proof is complete. □

Remark 5.15. We recall that $\det(\nabla u) \in L^1_{\mathrm{loc}}(\Omega)$ for every $u \in W^{1,n}_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$, and $u \mapsto \det(\nabla u)$ is continuous with respect to the corresponding strong topologies (see [Da], section 4.2.2); on the other hand $\mathrm{Det}\, Du$ is a well-defined distribution on Ω for every locally bounded $u \in W^{1,n-1}_{\mathrm{loc}}$, and is continuous with respect to the natural topologies (see for instance [Ba], [Mu1]). Clearly $\det(\nabla u) = \mathrm{Det}\, Du$ for every $u \in W^{1,n}_{\mathrm{loc}}$.

Using the continuity of the extended operator J and a density argument, we deduce that Ju agrees with the pointwise determinant $\det(\nabla u)$ for every monotone function u which belongs to $W^{1,n}_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$, and with the distributional determinant $\mathrm{Det}\, Du$ for every monotone function u which belongs to $W^{1,n-1}_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$.

6. Approximation by smooth functions

In this section we approximate a maximal monotone function u by Lipschitz maximal monotone functions defined in $\mathbb{R}^n$ in such a way that Tu_h converge to Tu and the measures $|Tu_h|$ converge to $|Tu|$. Simple examples show that, in general, the convergence of Tu_h to Tu does not imply the convergence of the variations $|Tu_h|$ to $|Tu|$.

Definition 6.1. *For every $\varepsilon > 0$ we set $\Psi_\varepsilon(x, y) := (x + \varepsilon y, y)$ for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, and for every maximal monotone function u we define u_ε as the multifunction whose graph is $\Psi_\varepsilon(\Gamma u)$, that is, $\Gamma u_\varepsilon = \big\{(x + \varepsilon y, y) : (x, y) \in \Gamma u\big\}$. Hence*

$$u_\varepsilon := (\varepsilon I + u^{-1})^{-1} . \tag{6.1}$$

Then we have the following result (cf. [Br], section 2.4, or [At], section 3.7.1):

Theorem 6.2. *Let u be a maximal monotone function, and let u_ε be given as in Definition 6.1. Then*

(1) *u_ε is a $1/\varepsilon$-Lipschitz maximal monotone function on $\mathbb{R}^n$ for every $\varepsilon > 0$;*
(2) *u_ε converges to u as $\varepsilon \to 0$ in the sense specified in Definition 1.6;*
(3) *Tu_ε converges to Tu as $\varepsilon \to 0$ in the sense of currents (cf. section 4);*
(4) *$|Tu_\varepsilon|$ converges to $|Tu|$ as $\varepsilon \to 0$, in the sense of measures, i.e.,*

$$\int g\, d|Tu_\varepsilon| \to \int g\, d|Tu| \qquad \forall\, g \in C_c(\mathbb{R}^n \times \mathbb{R}^n)\,;$$

(5) *for every convex open set $\Omega \subset \mathsf{Dm}\, u$, $Du_\varepsilon \to Du$ and $|Du_\varepsilon| \to |Du|$ in the sense of measures on Ω (cf. Theorem 5.3), i.e.,*

$$\begin{aligned}&\int g\, Du_\varepsilon\, d\mathscr{L}_n \to \int g\, d(Du) \\ &\int g\, |Du_\varepsilon|\, d\mathscr{L}_n \to \int g\, d|Du| \quad \forall\, g \in C_c(\Omega).\end{aligned}$$

Remark 6.3. Theorem 6.2 shows that we may approximate any maximal monotone function u with Lipschitz maximal monotone functions u_ε (with domain $\mathbb{R}^n$) so that properties (1)–(5) hold. Clearly we may also ask that the functions u_ε are smooth (regularize the approximating Lipschitz functions by convolution), and that they are diffeomorphisms (add to each smooth approximating function the term εI).

Remark 6.4. Notice that when u is a C^1 function with domain $\mathbb{R}^n$, then the standard regularization by convolution provides a *good* approximation of u with smooth functions u_ε (and good means in particular that $Tu_\varepsilon \to Tu$ and $|Tu_\varepsilon| \to |Tu|$).

When u is not a C^1 functions, this trick does not work, even when the domain of u is $\mathbb{R}^n$. In fact, we may still define the regularized functions $u_\varepsilon = u * \rho_\varepsilon$; the functions u_ε are still smooth maximal monotone functions and $Tu_\varepsilon \to Tu$, but we are able neither to prove that $|Tu_\varepsilon| \to |Tu|$ nor to find counterexamples showing that this convergence might not hold.

Proof of Theorem 6.2.

By (6.1) we obtain that $u_\varepsilon = \frac{1}{\varepsilon}\big(I + \frac{1}{\varepsilon}u^{-1}\big)^{-1}$. Since u^{-1} is maximal, u^{-1}/ε is maximal too, and then $\big(I + u^{-1}/\varepsilon\big)^{-1}$ is a 1-Lipschitz function defined on $\mathbb{R}^n$ (by statements (2) and (3) of Proposition 1.2). Hence u_ε is a $(1/\varepsilon)$-Lipschitz function on $\mathbb{R}^n$, and then it is maximal (Corollary 1.4), and statement (1) is proved.

To prove statement (2) we show that conditions (1.8) and (1.9) holds. Let be given sequences (ε_h) and $(x_h, y_h) \in \Gamma u_{\varepsilon_h}$ such that $\varepsilon_h \to 0$ and $(x_h, y_h) \to (x, y)$. Since $(x_h - \varepsilon_h y_h, y_h)$ belongs to Γu (by definition of u_ε) and Γu is closed, we get $(x, y) \in \Gamma u$, and then (1.8) is verified. Conversely, given $(x, y) \in \Gamma u$, the points $(x + \varepsilon y, y)$ belongs to Γu_ε and converge to (x, y) as $\varepsilon \to 0$, hence (1.9) holds.

Statement (3) follows by Theorem 4.6 and statement (2).

By formula (4.2) we have that for any Borel set $B \subset \mathbb{R}^n \times \mathbb{R}^n$,

$$|Tu_\varepsilon|(B) = \mathscr{H}^n(\Gamma u_\varepsilon \cap B) = \mathscr{H}^n(\Psi_\varepsilon(\Gamma u) \cap B) .$$

Therefore statement (4) follows by Lemma 6.5 below (recall that Γu has locally finite $\mathscr{H}^n$ measure). In particular

$$|Tu_\varepsilon|(A) \to \|Tu\|(A) \tag{6.2}$$

for any bounded open set A such that $|Tu|(\partial A) = 0$.

Finally we prove statement (5). By formula (5.4) we obtain the following representation of the partial derivative $D_i u_j$:

$$\langle \phi,\, D_i u_j \rangle = \int_{\Gamma u} (-1)^{i+1} \phi \, dy_j \wedge \widehat{dx_i} \qquad \forall\, g \in C_c(\Omega) ,$$

so that, denoting by $T^{ij}u$ the (i,j)-th component of Tu which appears at the right hand side,

$$|D_i u_j|(B) = |T^{ij} u|(B \times \mathbb{R}^n)$$

for any Borel set $B \subset \Omega$. Therefore $Du_\varepsilon \to Du$ weakly* in $\mathscr{M}_{\mathrm{loc}}(\Omega, \mathbb{R}^{n\times n})$ by statement (3), and $|Du_\varepsilon| \to |Du|$ weakly* in $\mathscr{M}_{\mathrm{loc}}(\Omega)$ by statement (4), (6.2) and Theorem 5.9. □

Lemma 6.5. *Let be given integers k, n with $k \geq n$, and a set $M \subset \mathbb{R}^k$ with locally finite $\mathscr{H}^n$ measure. Let $\Psi_h : \mathbb{R}^k \to \mathbb{R}^k$ be C^1 diffeomorphisms such that $\Psi_h(y) \to y$ and $D\Psi_h(y) \to I$ uniformly on compact sets. Then for every continuous function g with compact support in $\mathbb{R}^k$ there holds*

$$\int_{\Psi_h(M)} g \, d\mathscr{H}^n \to \int_M g \, d\mathscr{H}^n . \tag{6.3}$$

Proof. By the assumptions we deduce that the maps Ψ_h^{-1} converge to the identity map uniformly on compact sets, and since also $D\Psi_h$ converge to I uniformly on compact sets, for every $\varepsilon > 0$, $R > 0$, there exist $\bar{h}$ such that $|\Psi_h^{-1}(y) - y| \leq R$ and $|D\Psi_h(y) - I| \leq \varepsilon$ for every $h \geq \bar{h}$, $y \in B(0, 2R)$. Hence, for every $h \geq \bar{h}$, and every set $A \subset B(0, 2R)$,

$$(1 - \varepsilon) \operatorname{diam} A \leq \operatorname{diam} \Psi_h(A) \leq (1 + \varepsilon) \operatorname{diam} A .$$

By the definition of Hausdorff measures this yields, for every Borel set $A \subset B(0, 2R)$,

$$(1 - \varepsilon)^n \, \mathscr{H}^n(A) \leq \mathscr{H}^n(\Psi_h(A)) \leq (1 + \varepsilon)^n \, \mathscr{H}^n(A) . \tag{6.4}$$

Moreover for every set $C \subset B(0, R)$, $\Psi_h^{-1}(C) \subset B(0, 2R)$ and then (6.4) yields

$$(1-\varepsilon)^n \, \mathscr{H}^n\big(\Psi_h^{-1}(C) \cap M\big) \leq \mathscr{H}^n\big(C \cap \Psi_h(M)\big) \leq (1+\varepsilon)^n \, \mathscr{H}^n\big(\Psi_h^{-1}(C) \cap M\big) .$$

Hence for every positive Borel function g with support included in $B(0, R)$ we get

$$(1 - \varepsilon)^n \int_M g(\Psi_h) \, d\mathscr{H}^n \leq \int_{\Psi_h(M)} g \, d\mathscr{H}^n \leq (1 + \varepsilon)^n \int_M g(\Psi_h) \, d\mathscr{H}^n . \tag{6.5}$$

Taking into account that ε and R are arbitrarily taken, (6.5) shows that for every positive Borel function g with compact support

$$\lim_{h\to\infty} \int_{\Psi_h(M)} g \, d\mathscr{H}^n = \lim_{h\to\infty} \int_M g(\Psi_h) \, d\mathscr{H}^n . \tag{6.6}$$

If in addition g is continuous, by the dominated convergence theorem we get

$$\lim_{h\to\infty} \int_M g(\Psi_h) \, d\mathscr{H}^n = \int_M g \, d\mathscr{H}^n . \tag{6.7}$$

Equalities (6.6) and (6.7) yield (6.3). □

7. Convex functions

An important class of monotone functions is represented by the gradients (subdifferentials) of convex functions. In this section we examine some properties of convex functions and extend the results previously established for monotone functions to the gradients of convex functions.

Definition 7.1. *We denote by $\mathscr{C}onv$ the class of all lower semicontinuous convex functions $f : \mathbb{R}^n \to]-\infty, +\infty]$. For every $f \in \mathscr{C}onv$, we denote by*

$\mathsf{Dm}\, f$ *the set of points* $x \in \mathbb{R}^n$ *such that* $f(x) < +\infty$ *and we say that* $y \in \mathbb{R}^n$ *is a* subgradient *of* f *at* x *if*

$$x \in \mathsf{Dm}\, f\ , \quad f(x') \geq f(x) + \langle y, x' - x\rangle \quad \forall x' \in \mathbb{R}^n\ . \tag{7.1}$$

We denote by ∂f *the* subdifferential *of* f*, i.e., the multifunction which takes each* $x \in \mathbb{R}^n$ *in the set of all subgradients of* f *at* x.

Unless differently stated, throughout this section by convex function we mean a function in $\mathscr{C}onv$.

Remark 7.2. It follows immediately from definition that x is a minimum point of f if and only if $0 \in \partial f(x)$.

A convex function f is differentiable at the point x if and only $\partial f(x)$ is a singleton, and in this case $\partial f(x)$ consists exactly of the gradient of f at x, which we denote by $\nabla f(x)$.

Notice moreover that for every $x \in \mathsf{Dm}\, f$, y belongs to $\partial f(x)$ if and only if (7.1) holds for all x' in some neighborhood of x, and this yields that the subdifferential is a local operator on open sets, i.e., given convex functions f and g which agree on some open set Ω, then ∂f and ∂g agree on Ω.

Remark 7.3. By applying the Hahn-Banach theorem to the epigraph of f (which is a closed convex set in $\mathbb{R}^n \times \mathbb{R}$), we obtain that $\partial f(x)$ is never empty when x is in the interior of domain of f, and then

$$\mathsf{Dm}\, f \supset \mathsf{Dm}\,(\partial f) \supset \mathsf{Int}\,\mathsf{Dm}\, f\ .$$

Notice that both inclusions may be strict, and $\mathsf{Dm}\,\partial f$ may be not even convex (cf. Corollary 1.3(2) and the following remark).

Moreover ∂f is a maximal monotone function unless $f \equiv +\infty$ (cf. [Br], example 2.3.4). Indeed, the monotonicity of ∂f is an easy consequence of the convexity, and by Proposition 1.2, the maximality is equivalent to the subjectivity of $(\partial f + I)$. Let $y \in \mathbb{R}^n$ be given, and take $x \in \mathbb{R}^n$ the (unique) minimizer of the coercive lower semicontinuous function

$$x \mapsto f(x) + \frac{1}{2}|x|^2 - \langle y, x\rangle\ .$$

Then 0 is a subgradient of this function at x; hence $y \in \partial\big(f(x) + |x|^2/2\big) = (\partial f + I)(x)$, and the subjectivity of $(\partial f + I)$ is proved.

The subdifferentials of convex functions have been characterized as maximal *cyclically monotone* functions (cf. [Br], section II.7).

Remark 7.4. Since ∂f is monotone, then it is bounded on every open set relatively compact in the interior of its domain (Corollary 1.3(3)), and by Remark 7.3 $\mathsf{Int}\,\mathsf{Dm}\, f = \mathsf{Int}\,\mathsf{Dm}\,(\partial f)$. Using the non-smooth version of the mean value theorem (see [Cl], Theorem 2.3.7), we deduce that every convex function f is locally Lipschitz in the interior of its domain.

7.5. *A notion of convergence in* $\mathscr{C}onv$

We may associate to every convex function f its epigraph $Ef := \big\{(x,t) \in \mathbb{R}^n \times \mathbb{R} :\ t \geq f(x)\big\}$, and Ef is a closed convex set; accordingly we say that a sequence of convex functions (f_h) converges to f *in the sense of epigraphs* if the (Ef_h) converge to Ef in the sense of Kuratowski as $h \to +\infty$. This yields a compact metrizable topology on the class of all lower semicontinuous convex functions $\mathscr{C}onv$ (cf. section 1, Definition 1.6 and Proposition 1.7). Note that if $f_h \to f$ and K is relatively compact in the interior of $\mathsf{Dm}\, f$, then $K \subset \mathsf{Int}\,(\mathsf{Dm}\, f_h)$ for h large enough and f_h converge to f uniformly on K.

The convergence of epigraphs of lower semicontinuous (convex) functions in the sense of Kuratowski has been widely studied in the infinite dimensional case (namely when $\mathbb{R}^n$ is replaced by some functional space) and is equivalent to the notion of Γ-convergence for the corresponding functions (see for instance [DM], chapter 4, and [At]). This notion is particularly relevant from the variational viewpoint, since it carries over the convergence of minimizers: this means that if the convex functions f_n converge to f in the sense of epigraphs and x_n is a minimizer of f_n, then every cluster point of the sequence (x_n) is a minimizer of f.

Remark 7.6. Let $\Omega \subset \mathbb{R}^n$ be an open convex set, and let $\mathscr{C}onv(\Omega)$ be the space of all convex functions $f : \Omega \to \mathbb{R}$. We endow as usual $\mathscr{C}onv(\Omega)$ with the (metrizable) topology of uniform convergence on compact subsets of Ω. Thus a set of functions in $\mathscr{C}onv(\Omega)$ is relatively compact if and only if the functions are locally uniformly bounded.

We may extend a function $f \in \mathscr{C}onv(\Omega)$ to a convex function $\tilde{f} \in \mathscr{C}onv$ by setting

$$\tilde{f}(x) := \begin{cases} f(x) & \text{if } x \in \Omega, \\ \liminf\limits_{x' \to x} f(x') & \text{if } x \in \overline{\Omega} \setminus \Omega, \\ +\infty & \text{if } x \in \mathbb{R}^n \setminus \overline{\Omega}. \end{cases} \tag{7.2}$$

This extension is a continuous map from $\mathscr{C}onv(\Omega)$ (endowed with the topology of uniform convergence on compact sets) to $\mathscr{C}onv$ (endowed with the topology of convergence of epigraphs).

Theorem 7.7. *Let* (f_h) *be a sequence of convex functions which converges to* f *and assume that* $f \not\equiv +\infty$*. Then* (∂f_h) *converges to* ∂f *(cf. Definition 1.6).*

Proof. By Proposition 1.7, we may assume that ∂f_h converge to some u where

u is a monotone function. We claim that $u \supset \partial f$. Were this true, we would get $u = \partial f$ because ∂f is maximal. Take $(x,y) \in \partial f$; we want to show that y belongs to $u(x)$. We may assume with no loss in generality that $(x,y) = (0,0)$. Then 0 is the (unique) minimum point of $x \mapsto f(x) + |x|^2$. For each h, let x_h be the unique minimum point of $x \mapsto f_h(x) + |x|^2$; then $0 \in \partial f_h(x_h) + 2x_h$, i.e., $-2x_h \in \partial f_h(x_h)$, and moreover $x_h \to 0$ because minimum points converge to minimum points (cf. Remark 7.5). Hence $(x_h, -2x_h)$ belongs to $\Gamma(\partial f_h)$ and converge to $(0,0)$, and by (1.8) this means that $(0,0)$ belongs to Γu. □

Notice that the following converse holds: if ∂f_h converge to some non-empty multifunction, then there exist constants c_h such that $f_h + c_h$ converge to some convex function f, and hence $\partial f_h \to \partial f$.

Remark 7.8. Let T be the operator which associates to the graph of a maximal monotone function u the current Tu (cf. Definition 4.4). It follows immediately from Theorem 7.7 that the mapping $f \mapsto T(\partial f)$ is continuous, i.e., that for any sequence of convex functions f_h which converge to f in the sense of Remark 7.5 (or in Remark 7.6 as well), the currents $T(\partial f_h)$ converge to $T(\partial f)$ in the sense of n-currents of $\mathbb{R}^n \times \mathbb{R}^n$.

7.9. *Singular sets of convex functions*

As we did for monotone functions (cf. Definition 2.1), for every $f \in \mathscr{C}onv$ and every integer $k = 1, 2, \dots, n$ we denote by $\Sigma^k(f)$ the set of all points $x \in \mathsf{Dm}\, f$ such that ∂f has dimension greater than or equal to k. From Theorem 2.2 we obtain that $\Sigma^k(f)$ is a countably $\mathscr{H}^{n-k}$-rectifiable set. This kind of result was first proved by Anderson and Klee ([AK], see also [Be]).

This rectifiability property is not optimal: $\Sigma^k(f)$ is actually a countably $\mathscr{H}^{n-k}$-rectifiable set of class C^2, i.e., we can find countably many C^2 submanifolds of dimension $(n-k)$ in $\mathbb{R}^n$ which cover $\mathscr{H}^{n-k}$-almost all of $\Sigma^k(f)$. This was proved by the first author in [A2], but it is also an immediate corollary of Theorem 1 in [Z2]. In fact in both papers a slightly more precise result is proved: $\Sigma^k(f)$ can be covered by countably many graphs of functions $g : \mathbb{R}^{n-k} \to \mathbb{R}^k$ such that each component of g is the difference of two convex functions.

Notice that Theorem 2 in [Z2] generalizes this result to lower semicontinuous convex functions on a separable Banach space. In [AAC] the rectifiability of singular sets is proved for semi-convex functions.

We proved in Theorem 3.2 that a monotone function is differentiable at almost every point of the domain. If we apply this result to the subdifferential of a convex function f, we obtain that for almost every $\bar{x} \in \mathsf{Dm}\, f$ where f is differentiable (i.e., where ∂f is a singleton, cf. Remark 7.2) there exists a matrix $\nabla^2 f(\bar{x})$ such that

$$\lim_{\substack{x \to \bar{x} \\ y \in \partial f(x)}} \frac{y - \nabla f(\bar{x}) - \nabla^2 f(\bar{x}) \cdot (x - \bar{x})}{|x - \bar{x}|} = 0 \ . \tag{7.3}$$

Using (7.3) we can recover Aleksandrov's differentiability theorem for convex function (see [Ale]); simple proofs were also given in [CIL], Theorem A.2, and in [BCP]. Notice that Fréchet and Gateaux (first order) differentiability of convex functions on Banach spaces have been widely investigated (see [Asp], [PZ]; see also [Ph] for a survey on this topic),

Theorem 7.10. *Let x be a point where* (7.3) *holds. Then, f has a second order differential at x, i.e.,*

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - \langle \nabla f(x), h\rangle - \frac{1}{2}\langle \nabla^2 f(x)\, h,\, h\rangle}{|h|^2} = 0 \ . \tag{7.4}$$

Proof. It is not restrictive to assume $x = 0$, $f(x) = 0$ and $\nabla f(x) = 0$.

Let A be the matrix $\nabla^2 f(0)$. Let $\phi(h) := f(h) - \frac{1}{2}\langle Ah, h\rangle$. We have to prove that $\phi(h) = o(|h|^2)$.

Let be given $h \neq 0$. By the non-smooth version of the mean value theorem (see [Cl], Theorem 2.3.7) there exist y in the segment joining 0 to h, and $p \in \partial\phi(y)$, such that $\phi(h) - \phi(0) = \langle p, h\rangle$, i.e.,

$$\phi(h) = \langle q - Ay, h\rangle \tag{7.5}$$

for some vector $q \in \partial f(y)$. From (7.3) we infer

$$|q - Ay| = o(|y|) = o(|h|) \ ,$$

and in conjunction with (7.5) this implies $\phi(h) = o(|h|^2)$. □

As we did in Theorem 5.3 for monotone functions, we can characterize convex functions in term of distributional derivatives (this is a well-known result, see for instance [Du], [Re2]).

Proposition 7.11. *Let Ω a convex open set of $\mathbb{R}^n$. If $f : \Omega \to \mathbb{R}$ is convex, then Df is monotone, and $D^2 f$ is a positive and symmetric (matrix-valued and locally bounded) measure. Conversely if $f \in L^1_{\mathrm{loc}}(\Omega)$ and $D^2 f$ is a positive (matrix-valued) distribution on Ω, then f agrees almost everywhere on Ω with a convex function g such that $\Omega \subset \mathsf{Dm}\, g$.*

The proof of this proposition is the same as Theorem 5.3 and we omit it.

A standard way to approximate from below a l.s.c. convex function is the inf-convolution, also known as Moreau-Yosida approximation (see for instance [Br], section 2.7, or [At], section 2.7):

Definition 7.12. *Given $f \in \mathscr{C}onv$, for every $\varepsilon > 0$, $x \in \mathbb{R}^n$, we set*

$$f_\varepsilon(x) := \min_{x' \in \mathbb{R}^n} \left\{ f(x') + \frac{1}{2\varepsilon}|x' - x|^2 \right\} . \tag{7.6}$$

The function f_ε is well-defined and finite for every $x \in \mathbb{R}^n$ because $x' \mapsto f(x') + \frac{1}{2\varepsilon}|x' - x|^2$ is lower semicontinuous and coercive, moreover it may be easily verified that the functions f_ε converge to f in the sense of epigraphs (see paragraph 7.5).

This approximation is tightly connected to the approximation of monotone functions given in section 6, as shown in the following statement (cf. [Br], Proposition 2.11).

Proposition 7.13. *For any $\varepsilon > 0$, f_ε is a $C^{1,1}$ convex function and $\partial(f_\varepsilon) = (\partial f)_\varepsilon$, where $(\partial f)_\varepsilon$ is given in Definition 6.1.*

Proof. A simple computation shows that f_ε is convex. Hence f_ε is of class $C^{1,1}$ once proved that $\partial(f_\varepsilon) = (\partial f)_\varepsilon$, because $(\partial f)_\varepsilon$ is a Lipschitz function (Theorem 6.2(1)).

Since $(\partial f)_\varepsilon$ is maximal monotone, we need only to show that the inclusion $(\partial f)_\varepsilon(x) \subset \partial(f_\varepsilon)(x)$ holds for every x. Let x and $y \in (\partial f)_\varepsilon(x)$ be fixed. Since $(x - \varepsilon y, y)$ belongs to $\Gamma(\partial f)$, we have

$$f(x') \geq f(x - \varepsilon y) + \langle y, x' - x + \varepsilon y\rangle \quad \forall\, x' \in \mathbb{R}^n . \tag{7.7}$$

On the other hand, the Cauchy inequality yields $|\varepsilon\, y|^2 + |x'' - x'|^2 \geq 2\langle \varepsilon\, y, x'' - x'\rangle$, and then

$$\frac{1}{2\varepsilon}|x'' - x'|^2 \geq -\frac{\varepsilon}{2}|y|^2 + \langle y, x'' - x'\rangle \quad \forall\, x', x'' \in \mathbb{R}^n . \tag{7.8}$$

By summing (7.7) and (7.8) side by side, we obtain

$$f(x') + \frac{1}{2\varepsilon}|x' - x''|^2 \geq f(x - \varepsilon y) + \frac{\varepsilon}{2}|y|^2 + \langle y, x'' - x\rangle \quad \forall\, x',\, x'' \in \mathbb{R}^n . \tag{7.9}$$

Now, setting $x'' := x$ in (7.9) yields

$$f(x') + \frac{1}{2\varepsilon}|x' - x|^2 \geq f(x - \varepsilon y) + \frac{\varepsilon}{2}|y|^2 \quad \forall\, x' \in \mathbb{R}^n ,$$

so that, taking (7.6) and equality $\frac{\varepsilon}{2}|y|^2 = \frac{1}{2\varepsilon}|(x - \varepsilon y) - x|^2$ into account, we get

$$f_\varepsilon(x) = f(x - \varepsilon y) + \frac{\varepsilon}{2}|y|^2 . \tag{7.10}$$

If we replace (7.10) in the right hand side of (7.9) we get

$$f(x') + \frac{1}{2\varepsilon}|x' - x''|^2 \geq f_\varepsilon(x) + \langle y, x'' - x\rangle \quad \forall\, x',\, x'' \in \mathbb{R}^n ,$$

and taking the infimum over all x' we obtain

$$f_\varepsilon(x'') \geq f_\varepsilon(x) + \langle y, x'' - x\rangle \quad \forall\, x'' \in \mathbb{R}^n .$$

This proves that $y \in \partial f_\varepsilon(x)$. □

8. An example

In this section we give an example of monotone function whose weak Jacobian (see Definition 5.12 and following remarks) has a non-trivial structure. In fact we exhibit a function $u : \mathbb{R}^2 \to \mathbb{R}^2$ so that

(1) u is Hölder continuous and belongs to $W^{1,p}_{\mathrm{loc}}$ for every $p < 2$; moreover u is the gradient of a convex function of class C^1;
(2) the weak Jacobian Ju is supported on a purely unrectifiable and Lebesgue negligible set A, and $\det(\nabla u) = 0$ a.e. in $\mathbb{R}^2$.

Our example is given in dimension 2, but similar constructions can be generalized to higher dimensions. We obtain u as the limit of a sequence of functions u_k. The construction and the description of the approximating functions u_k is divided into several steps (from paragraph 8.2 to Lemma 8.5), and in Theorem 8.6 we show that these functions converge to a limit u, and describe the main features of u (see also paragraphs 8.8 – 8.12). We conclude this section by constructing another sequence of subdifferentials of convex functions which converges to u (see Definition 8.13 and Proposition 8.14) and then we discuss some consequences of this example (see paragraph 8.15 and Remark 8.16).

Other examples of Sobolev functions with Jacobian determinant supported on (fractal) sets of arbitrary dimension were constructed in [Mu3] and, with a slightly different viewpoint, in [Po].

8.1. *Some notation*
Throughout this section the capitol letter B always denotes an open ball in $\mathbb{R}^2$, and $B = B(x, r)$ is the open ball with center x and radius r. For every $a > 0$, aB denotes the rescaled ball with same center and radius multiplied by a, that is, $aB := B(x, ar)$.

A *similitude* is an affine map $\Psi : \mathbb{R}^2 \to \mathbb{R}^2$ such that for every couple of points $x, y \in \mathbb{R}^2$, there holds $\big|\Psi(x) - \Psi(y)\big| = a\,|x-y|$, where $a \in (0,+\infty)$ is the *scaling factor* of Ψ.

We say that a set is *purely unrectifiable* if its intersection with any Lipschitz curve is $\mathscr{H}^1$-negligible. We remark that by the coarea formula (see for instance [EG], section 5.5), the derivative of a BV function can never charge a purely unrectifiable set.

8.2. *Construction parameters*

Our construction basically depends on the following parameters: $\alpha > 1$ is a fixed real number, and $\big\{B_i : i \in I\big\}$ is a countable collection of pairwise disjoint open balls which cover almost all of the ball $B(0,1)$. We denote by ρ_i the radius of each ball B_i, and by ρ the maximum of all ρ_i (thus $\rho < 1$). The family $\big\{B_i : i \in I\big\}$ is called *basic cover*.

Definition 8.3. *For every $i \in I$, we denote by Ψ_i the similitude which takes the unit ball $B(0,1)$ into $\frac{1}{\alpha}B_i$, that is, the map given by*

$$\Psi_i(x) := x_i + \frac{\rho_i}{\alpha}(x - x_i) \quad \textit{for } x \in \mathbb{R}^2$$

(here x_i is the center of B_i). For every integer $k = 1, 2, \dots$ and every $\mathbf{i} = (i_1, \dots, i_k) \in I^k$ we set (cf. figure 1 below)

$$B^k_{\mathbf{i}} := \big(\Psi_{i_1} \circ \cdots \circ \Psi_{i_k}\big)\big(B(0,1)\big) \tag{8.1}$$

(in particular $B^1_i = \frac{1}{\alpha}B_i$ for every $i \in I$). Then we set $A_0 := B(0,1)$ and

$$A_k := \bigcup_{\mathbf{i} \in I^k} B^k_{\mathbf{i}}\ , \qquad A := \bigcap_{k=1}^{\infty} A_k\ . \tag{8.2}$$

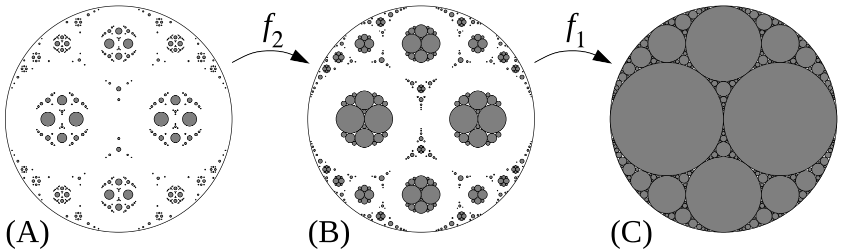

Fig. 1 (A): in grey, the set A_2, i.e., the union of the balls $\big\{B^2_{\mathbf{i}} : \mathbf{i} \in I^2\big\}$; **(B):** in grey, the set A_1, covered by the balls $\big\{\alpha B^2_{\mathbf{i}} : \mathbf{i} \in I^2\big\}$; **(C):** in grey, the set $A_0 = B(0,1)$, covered by the balls $\big\{B_i = \alpha B^1_i : i \in I\big\}$.

Let us fix $k \geq 1$ and $\mathbf{i} \in I^k$, and let $\Psi_{\mathbf{i}} := (\Psi_{i_1} \circ \cdots \circ \Psi_{i_k})$; then $B^k_{\mathbf{i}} = \Psi_{\mathbf{i}}(B(0,1))$ and $B^{k+1}_{(\mathbf{i},i)} = \Psi_{\mathbf{i}}\big(\frac{1}{\alpha}B_i\big)$ for every $i \in I$, and moreover $\Psi_{\mathbf{i}}$ is a similitude with scaling factor $\alpha^{-k}\rho_{i_1}\cdots\rho_{i_k}$, and then $B^k_{\mathbf{i}}$ is a ball with radius $\alpha^{-k}\rho_{i_1}\cdots\rho_{i_k}$. Since the balls $\big\{B_i : i \in I\big\}$ are pairwise disjoint and cover almost all of $B(0,1)$, then the balls $\big\{\alpha B^{k+1}_{(\mathbf{i},i)} : i \in I\big\}$ are pairwise disjoint and cover almost all of $B^k_{\mathbf{i}}$. We deduce that for every k the balls $\big\{\alpha B^{k+1}_{\mathbf{i}} : \mathbf{i} \in I^{k+1}\big\}$ are pairwise disjoint and cover almost all of A_k, and then we immediately derive the following result:

Proposition 8.4. *The following facts hold:*

(1) *the sets $\{A_k\}$ are a decreasing sequence of open sets with intersection A;*
(2) *for every $k > 0$, the balls $\big\{\alpha B^k_{\mathbf{i}} : \mathbf{i} \in I^k\big\}$ cover almost all of A_{k-1};*
(3) *each set A_k has measure $\pi\alpha^{-2k}$, and A has measure 0;*
(4) *each ball $B^k_{\mathbf{i}}$ has radius $\alpha^{-k}\rho_{i_1}\cdots\rho_{i_k} \leq (\rho/\alpha)^k$.*

Now, we briefly outline the construction of the function u. First we construct for every k a map $f_k : \mathbb{R}^2 \to \mathbb{R}^2$ which takes each ball $B^k_{\mathbf{i}}$ onto $\alpha B^k_{\mathbf{i}}$ for every $\mathbf{i} \in I^k$. Thus f_k takes A_k onto the union of the balls $\alpha B^k_{\mathbf{i}}$ for $\mathbf{i} \in I^k$, and by Proposition 8.4(2) this implies that $f_k(A_k)$ covers almost all of A_{k-1}.

Then we set $u_k := f_0 \circ \cdots \circ f_k$, and hence $u_k(A_k)$ covers almost all of A_0. Using the area formula we show that the Jacobian determinant of u_k is therefore supported on the set A_k and has integral always equal to π, while the measure of A_k decreases to 0 (cf. Proposition 8.4(3)).

Eventually we define u as the limit of the functions u_k as $k \to +\infty$, and we show that the weak Jacobian of u is a positive measure with mass π supported on the set A, which is negligible and purely unrectifiable. Moreover we show that u, like the functions u_k, is the gradient of a convex function and belongs to $W^{1,p}_{\mathrm{loc}}$ for every $p < 2$.

This construction is divided in several steps.

Step 1. For every ball $B = B(\bar{x}, r)$ we set

$$f_B(x) := \begin{cases} \bar{x} + \alpha(x - \bar{x}) & \text{if } |x - \bar{x}| \leq r, \\[2ex] \bar{x} + \alpha r \dfrac{x - \bar{x}}{|x - \bar{x}|} & \text{if } |x - \bar{x}| \geq r. \end{cases} \tag{8.3}$$

Then function f_B satisfies the following (immediate) properties:

(1) *f_B takes B onto αB linearly, takes $\alpha\overline{B} \setminus B$ onto $\partial(\alpha B)$, and agrees with the identity on the boundary of αB;*
(2) *f_B is α-Lipschitz on $\mathbb{R}^2$, and agrees in B with a similitude with scaling factor α;*

(3) $\nabla f_B(x)$ *is a positive symmetric matrix for a.e.* $x \in B$. *In fact,* f_B *is the gradient of the convex function* F *defined by*

$$F(x) := \begin{cases} \langle \bar{x}, x \rangle + \frac{\alpha}{2}|x-\bar{x}|^2 & \text{if } |x-\bar{x}| \le r, \\ \langle \bar{x}, x \rangle + \alpha r|x-\bar{x}| - \frac{\alpha r^2}{2} & \text{if } |x-\bar{x}| \ge r. \end{cases}$$

Step 2. For every integer $k = 1, 2, \dots$, we set

$$f_k(x) := \begin{cases} f_{(B_{\mathbf{i}}^k)}(x) & \text{if } x \in \alpha B_{\mathbf{i}}^k \text{ for some } \mathbf{i} \in I^k, \\ x & \text{otherwise.} \end{cases} \tag{8.4}$$

Then the functions f_k satisfy the following properties:

(1) $f_k(A_0) \subset \overline{A}_0$, $f_k(A_k)$ *cover almost all of* A_{k-1}, *and* $f_k(x) = x$ *for* $x \in \mathbb{R}^2 \setminus A_{k-1}$;
(2) f_k *is* α*-Lipschitz in* $\mathbb{R}^2$, *and everywhere differentiable in* A_k *and in* $\mathbb{R}^2 \setminus \overline{A}_{k-1}$;
(3) $\left|f_k(x) - f_k(y)\right| \le |x-y| + 4\alpha^{1-k}\rho^k$ *for every* $x, y \in \mathbb{R}^2$;
(4) $\nabla f_k(x) = \alpha I$ *for every* $x \in A_k$;
(5) $\nabla f_k(x)$ *is a positive symmetric matrix for a.e.* $x \in \mathbb{R}^2$.

Proof. Once recalled that the union of the balls $\{B_{\mathbf{i}}^k : \mathbf{i} \in I^k\}$ is A_k and that the balls $\{\alpha B_{\mathbf{i}}^k : \mathbf{i} \in I^k\}$ cover almost all of A_{k-1}, all statements but (3) easily follow from the properties of the functions f_B stated in Step 1.

The proof of (3) is slightly more delicate. Take indeed $x \in \mathbb{R}^2$. If x does not belong to the union of the balls $\alpha B_{\mathbf{i}}^k$, then $f_k(x) = x$, while if x belongs to $\alpha B_{\mathbf{i}}^k$ for some $\mathbf{i} \in I^k$, then $f_k(x)$ belongs to $\alpha B_{\mathbf{i}}^k$ too (cf. (8.4) and Step 1, statement (1)). In both cases there holds $\left|f_k(x) - x\right| \le 2\alpha r$, where r is the supremum of the radii of the balls $B_{\mathbf{i}}^k$, and since $r \le (\rho/\alpha)^k$ by Proposition 8.4(4),

$$\left|f_k(x) - x\right| \le 2\alpha^{1-k}\rho^k \qquad \forall\, x \in \mathbb{R}^2 . \tag{8.5}$$

Therefore for every $x, y \in \mathbb{R}^2$ we obtain

$$\left|f_k(x) - f_k(y)\right| \le |x-y| + \left|f_k(x) - x\right| + \left|f_k(y) - y\right| \le |x-y| + 4\alpha^{1-k}\rho^k . \ \square$$

Step 3. We define f_0 by setting $f_0(x) := x$ if $|x| \le 1$ and $f_0(x) := x/|x|$ if $|x| > 1$. Then for every integer $k = 0, 1, \dots$ we set

$$u_k := f_0 \circ \cdots \circ f_k . \tag{8.6}$$

The functions u_k enjoy the following properties:

(1) $u_k(\mathbb{R}^2) \subset \overline{A}_0$ *and* $u_k(A_k)$ *covers almost all of* A_0, $u_k(x) = u_h(x)$ *for every* $x \in \mathbb{R}^2 \setminus A_h$ *and every* $h < k$, *and* $u_k(x) = f_0(x) = x/|x|$ *for every* $x \in \mathbb{R}^2 \setminus A_0$;
(2) u_k *is* (α^k)*-Lipschitz;*
(3) *for every* $x, y \in \mathbb{R}^2$ *and every* h, k *there holds*

$$\left|u_k(x) - u_k(y)\right| \le \alpha^h|x-y| + \frac{4\rho^{h+1}}{1-\rho/\alpha} \ ; \tag{8.7}$$

(4) $\nabla u_k(x) = \alpha^k I$ *for every* $x \in A_k$ *and* $\nabla u_k(x) = \alpha^{k-1}\nabla f_k(x)$ *for almost every* $x \in A_{k-1}$;
(5) $\nabla u_k(x)$ *is a positive symmetric matrix for a.e.* $x \in \mathbb{R}^2$.

Proof. Statements (1) and (2) follow from the definition of u_k and statements (1) and (2) of Step 2 respectively.

Let us prove (3). Take $x, y \in \mathbb{R}^2$; if $k < h$ then (8.7) follows by statement (2). Assume then that $k \ge h \ge 0$: we claim that

$$\left|u_k(x) - u_k(y)\right| \le \alpha^h \Big(|x-y| + 4\rho \sum_{i=h}^{k-1} (\rho/\alpha)^i\Big) . \tag{8.8}$$

This inequality may be proved by induction on k: the case $k = h$ reduces to $\left|u_k(x) - u_k(y)\right| \le \alpha^k|x-y|$, which follows from statement (2), and the inductive step may be easily proved using statement (3) of Step 2 and the recursive formula $u_k = u_{k-1} \circ f_k$. Now inequality (8.7) follows from (8.8) by replacing the finite sum $\sum_h^{k-1} (\rho/\alpha)^i$ at the right hand side of (8.8) with $\sum_h^{\infty} (\rho/\alpha)^i$.

Since the sets A_k are a decreasing sequence of open sets (cf. Proposition 8.4(1)), from statement (4) of Step 2 we immediately deduce that $\nabla u_k(x) = \alpha^k I$ everywhere in A_k. The rest of statement (4) follows from the identity $u_k = u_{k-1} \circ f_k$.

Statement (5) is proved by induction on k. For $k = 0$, $u_0 = f_0$ and then statement (5) follows by explicit computation. Take now $k > 0$. From statement (4) we know that $\nabla u_k(x) = \alpha^{k-1}\nabla f_k(x)$ for almost every $x \in A_{k-1}$ and then statement (5) of Step 2 yields that $\nabla u_k(x)$ is a positive symmetric matrix a.e. in A_{k-1}. On the other hand u_k agrees with u_{k-1} a.e. in $\mathbb{R}^2 \setminus A_{k-1}$ (statement (1)), and then ∇u_k agrees a.e. in $\mathbb{R}^2 \setminus A_{k-1}$ with ∇u_{k-1}, which is a positive symmetric matrix by the inductive hypothesis. $\square$

Now we can prove the following result:

Lemma 8.5. *The functions* u_k *defined in Step 3 are Lipschitz and enjoy the following properties:*

(1) *$u_k(\mathbb{R}^2) \subset \overline{A}_0$ and $u_k(A_k)$ covers almost all of A_0, $u_k(x) = u_h(x)$ for every $x \in \mathbb{R}^2 \setminus A_h$ and every $h < k$, and $u_k(x) = f_0(x) = x/|x|$ for every $x \in \mathbb{R}^2 \setminus A_0$;*
(2) *for $k \geq 1$, the functions u_k satisfy the inequality*

$$\big|u_k(x) - u_k(y)\big| \leq \Big(1 + \frac{4}{1-\rho/\alpha}\Big)\, |x-y|^\gamma \qquad \text{whenever } |x-y| \leq 1, \tag{8.9}$$

with $\gamma := \log\rho / \log(\rho/\alpha)$;
(3) *the gradients ∇u_k are uniformly bounded in $L^p(\Omega)$ for every $p < 2$ and every bounded open set $\Omega \subset \mathbb{R}^2$;*
(4) *the functions u_k are gradients of convex functions.*

Proof. Statement (1) is statement (1) of Step 3, and statement (4) follows from statement (5) of Step 3 and Proposition 7.11. Let us proved statement (2). Fix $k \geq 0$ and $x, y \in \mathbb{R}^2$ such that $|x-y| \leq 1$, and let h be the largest integer below $\log|x-y| / \log(\rho/\alpha)$. Then $h \geq 0$ and $h > \big(\log|x-y| / \log(\rho/\alpha)\big) - 1$, and by statement (3) of Step 3 there holds

$$\big|u_k(x) - u_k(y)\big| \leq \alpha^h |x-y| + \frac{4\rho^{h+1}}{1-\rho/\alpha} \leq \Big(1 + \frac{4}{1-\rho/\alpha}\Big)\, |x-y|^\gamma \ .$$

Concerning statement (3), since $\nabla u_k(x) = \nabla\big(x/|x|\big)$ for every $x \in \mathbb{R}^2 \setminus \overline{A}_0$, and $\nabla\big(x/|x|\big)$ is bounded in $\mathbb{R}^2 \setminus \overline{A}_0$, then it is enough to prove that the gradients ∇u_k are uniformly bounded in $L^p(A_0)$. By Step 3, statement (2), $|\nabla u_k(x)| \leq \alpha^k$ a.e. in $\mathbb{R}^2$, and by Step 3, statement (1), $\nabla u_k(x) = \nabla u_{k-1}(x)$ a.e. in $\mathbb{R}^2 \setminus A_{k-1}$. Hence

$$\begin{aligned} \int_{A_0} \big|\nabla u_k(x)\big|^p \, d\mathscr{L}_2 &\leq \alpha^{kp} |A_{k-1}| + \int_{A_0 \setminus A_{k-1}} \big|\nabla u_{k-1}(x)\big|^p \, d\mathscr{L}_2 \\ &\leq \pi \alpha^{2+(p-2)k} + \int_{A_0} \big|\nabla u_{k-1}(x)\big|^p \, d\mathscr{L}_2 \ . \end{aligned}$$

Therefore by induction on k we obtain

$$\int_{A_0} |\nabla u_k|^p d\mathscr{L}_2 \leq \pi\, \alpha^p \sum_0^{k-1} \alpha^{(p-2)i} \leq \frac{\pi\, \alpha^2}{\alpha^{2-p} - 1} \ . \qquad \square$$

Lemma 8.5 implies the convergence of the functions u_k. More precisely, we have the following theorem:

Theorem 8.6. *The functions u_k converge uniformly on compact sets to a continuous function $u : \mathbb{R}^2 \to \mathbb{R}^2$. Moreover*

(1) *$u(\mathbb{R}^2) \subset \overline{A}_0$ and $u(A)$ covers almost all of A_0;*
(2) *$u \in C^{0,\gamma}$ with $\gamma := \log\rho / \log(\rho/\alpha)$;*
(3) *$u \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^2)$ for every $p < 2$;*
(4) *u is the gradient of a convex function of class C^1;*
(5) *the weak Jacobian Ju is a positive singular measure supported on the set A, and $\det(\nabla u) = 0$ a.e. in $\mathbb{R}^2$;*
(6) *the set A is purely unrectifiable.*

Remark 8.7. Since $\rho < 1$ and $\alpha > 1$, we have that $0 < \gamma < 1$. Moreover we can choose α and ρ so that γ is any prescribed real number between 0 and 1. We remark that the set A is a decreasing intersection of open sets. Therefore it is not immediate to prove that A is non-empty: indeed this follows from the fact that the measure Ju is non-trivial and supported on A.

Proof. By Lemma 8.5(1), the sequence $(u_k(x))$ is definitively constant for every $x \notin A$, and this set has measure 0 (Proposition 8.4(3)). Hence the functions u_k converge almost everywhere, and then we have the uniform convergence on compact sets because the functions u_k are equi-continuous (cf. (8.9)). Hence statements (1)–(4) follow immediately from (1)–(4) of Lemma 8.5, respectively.

Since u is the gradient of a convex function of class C^1, it is a maximal monotone function (cf. Remark 7.3), and then Ju is a positive Borel measure on $\mathbb{R}^2$ which satisfies $Ju(B) = \big|u(B)\big|$ for every Borel set $B \subset \mathbb{R}^2$ (cf. Theorem 5.11 and Definition 5.12). Taking statement (1) into account we get $Ju(\mathbb{R}^2) = Ju(A) = \pi$. Therefore Ju is supported on the negligible Borel set A. In particular this implies that $\det(\nabla u) = 0$ a.e. in A_0 (cf. Proposition 5.14).

Let us prove that A is purely unrectifiable. This means that for every Lipschitz curve C in $\mathbb{R}^2$ there holds $\mathscr{H}^1(A \cap C) = 0$. In fact it is enough to prove this statement when C is a curve which agrees (up to an isometry) with the graph of a Lipschitz function $g : \mathbb{R} \to \mathbb{R}$, and thus the proof relies on the following fact (which we shall prove below): if B is a ball in $\mathbb{R}^2$ and αB is the corresponding rescaled ball, then

$$\mathscr{H}^1(B \cap C) \leq \delta \mathscr{H}^1(\alpha B \cap C) \ , \tag{8.10}$$

where δ is strictly lower than 1 and depends on α and the Lipschitz constant of g only. Assume for the moment that (8.10) holds. For every $k > 0$ and every $\mathbf{i} \in I^k$, the balls $\big\{\alpha B^{k+1}_{(\mathbf{i},i)} : i \in I\big\}$ are pairwise disjoint and included in $B^k_{\mathbf{i}}$, and then (8.10) yields

$$\sum_{i \in I} \mathscr{H}^1\big(B^{k+1}_{(\mathbf{i},i)} \cap C\big) \leq \delta \sum_{i \in I} \mathscr{H}^1\big(\alpha B^{k+1}_{(\mathbf{i},i)} \cap C\big) \leq \delta \mathscr{H}^1\big(B^k_{\mathbf{i}} \cap C\big) \ .$$

Taking the sum over all $\mathbf{i} \in I^k$ we obtain

$$\mathscr{H}^1(A_{k+1} \cap C) \leq \delta \mathscr{H}^1(A_k \cap C) \ ,$$

and by iterating this inequality,

$$\mathscr{H}^1(A \cap C) \leq \mathscr{H}^1(A_{k+1} \cap C) \leq \delta^k \mathscr{H}^1(A_1 \cap C) \ ,$$

and then $\mathscr{H}^1(A \cap C) = 0$ by passing to the limit as $k \to +\infty$.

Finally we prove (8.10). Assume that $\alpha B \cap C$ is not empty. Let c be the Lipschitz constant of g, let r be the radius of the ball B, and let E the set of all $t \in \mathbb{R}$ such that $(t, g(t)) \in B$. The set E is included in an interval of length $2r$, and then

$$\mathscr{H}^1(B \cap C) = \int_E \sqrt{1 + (\dot{g})^2} \leq 2r \sqrt{1 + c^2} \ . \tag{8.11}$$

Since the curve C touches both the ball B and the boundary of the rescaled ball αB, the intersection of C and $\alpha B \setminus B$ has length $(\alpha - 1)r$ at least, and by (8.11) we get

$$\mathscr{H}^1\big((\alpha B \setminus B) \cap C\big) \geq (\alpha - 1)\, r \geq \frac{\alpha - 1}{2\sqrt{1 + c^2}} \mathscr{H}^1(B \cap C) \ ,$$

and then

$$\begin{aligned} \mathscr{H}^1(\alpha B \cap C) &= \mathscr{H}^1\big((\alpha B \setminus B) \cap C\big) + \mathscr{H}^1(B \cap C) \\ &\geq \Big(1 + \frac{\alpha - 1}{2\sqrt{1 + c^2}}\Big) \mathscr{H}^1(B \cap C) \ . \end{aligned} \tag{8.12}$$

Inequality (8.10) follows from (8.12) by letting $\delta := \Big(1 + \frac{\alpha - 1}{2\sqrt{1+c^2}}\Big)^{-1}$. $\square$

Now we want to discuss some details about the function u. In particular we address the following questions:

(a) Do we need such a complicate construction?
(b) What can be said about the dimension of the set A?
(c) What can be said about the dimension of the measure Ju?

8.8. *Why such a complicate construction?*
The construction of the function u is rather complicate, and it is not immediately clear whether it can be simplified or not. For example, one of the key point of this construction is the choice of the basic cover (cf. paragraph 8.2): obviously we had better take a finite cover rather than a countable one, but

this seems impossible. Indeed the balls $\{B_i : i \in I\}$ must cover almost all of the unit ball, otherwise the set A would support no part of the Jacobian determinant Ju, and it is clear that finitely many pairwise disjoint balls can never cover almost all of a ball.

On the other hand, we may consider the possibility of replacing the ball with any other plane figure C which can be covered, up to a negligible subset, by finitely many pairwise disjoint rescaled copies of C (e.g., rectangles, or triangles). But in this case the construction of the equivalent of the function f_B in Step 1 would fail: C has to be a polygon, and it may be proved that this being the case there is no continuous monotone function f which takes C into C, agrees with the identity on ∂C, and takes a smaller copy of C onto C.

Nevertheless if we drop the monotonicity assumption on the functions u_k and u, we are allowed to replace the basic cover in the construction with a finite one (e.g., by taking four squares which cover the unit square), and in this case u is an example of function in $W^{1,p}_{\mathrm{loc}} \cap C^{0,\gamma}$ whose weak Jacobian determinant is supported on A, and A is a compact self-similar fractal in the sense of Hutchinson [Hu]; moreover we can choose the construction parameter α so that the dimension of A is any fixed number between 0 and 2 (cf. paragraph 8.10). Similar examples were described in [Mu3].

If we drop the assumption that u is a Sobolev function, and we simply look for a convex function whose subdifferential u has weak Jacobian Ju supported on a purely unrectifiable set, then we can consider the following simpler construction (cf. [Am3]).

8.9. *A simple construction*
Take a finite non-atomic positive measure μ on $\mathbb{R}$ which is supported on a Lebesgue negligible compact set C. Let v be a continuous function on $\mathbb{R}$ whose distributional derivative is μ, and let g be a function whose derivative is v, e.g.

$$v(x) := \mu\big(]-\infty, x]\big) \ , \quad g(x) := \int_{-\infty}^{x} (x - y)\, d\mu(y) \qquad \forall x \in \mathbb{R} \, .$$

Then v is a continuous increasing function and g is a convex function of class C^1.

Now we set $f(x,y) := g(x) + g(y)$ and $u(x,y) := \big(v(x), v(y)\big)$. Then f is a convex function of class C^1 on the plane, u is the gradient of f and $D_1 u_1 = \mu \otimes \mathscr{L}_1$, $D_2 u_2 = \mathscr{L}_1 \otimes \mu$, $D_1 u_2 = D_2 u_1 = 0$. Moreover we claim that $Ju = \mu \otimes \mu$ (and then it is supported on $C \times C$). Take indeed Borel sets $B, B' \subset \mathbb{R}$. Thus $u(B \times B') = v(B) \times v(B')$, and then

$$Ju(B \times B') = \big|u(B \times B')\big| = |v(B)| \cdot |v(B')| = \mu(B) \cdot \mu(B') \ ,$$

and the claim is proved.

We remark that the set $C \times C$ has negligible projection on both axes, and therefore is purely unrectifiable. Moreover if C is a self-similar fractal in the sense of Hutchinson (e.g. a Cantor type set), and μ is a self-similar measure on C, then the same holds for $C \times C$ and Ju, and we may choose C and μ so that the dimension of Ju is any prescribed real number between 0 and 2 (cf. paragraphs 8.10 and 8.12).

8.10. *The dimension of the set A*

We have proved in Theorem 8.6 that A is a purely unrectifiable set. Now we want to compute its Hausdorff dimension. We begin by recalling that A satisfies the identity $A = T(A)$ where $T : \mathcal{P}(\mathbb{R}^2) \to \mathcal{P}(\mathbb{R}^2)$ is defined by

$$T(C) := \bigcup_{i \in I} \Psi_i(C) \qquad \text{for all } C \subset \mathbb{R}^2. \tag{8.13}$$

According to Hutchinson [Hu], if $\{\Psi_i : i \in I\}$ is a finite family of similitudes with scaling factor $r_i < 1$ and there exists a bounded open set D such that the sets $\{\Psi_i(D) : i \in I\}$ are pairwise disjoint and included in D, then the map T admits only one compact fixed point $C \subset \overline{D}$. The set C is a *self-similar fractal* with Hausdorff dimension d, where d is the (unique) solution of the equation

$$\sum_i r_i^d = 1 \ . \tag{8.14}$$

(In fact there holds a stronger result: $0 < \mathscr{H}^d(C) < +\infty$, and also the box-counting dimension of C is equal to d. For a good exposition of Hutchinson's results see also [Fa1], section 8.3, or [Fa2], chapter 9).

Our case is slightly different: the set A is not closed, and $\{\Psi_i : i \in I\}$ is a *countable* family of similitudes with scaling factor ρ_i/α. Nevertheless similar conclusions holds for A. Indeed let Φ be defined by

$$\Phi(t) := \sum_i (\rho_i/\alpha)^t \qquad \text{for every } t \geq 0. \tag{8.15}$$

Then Φ is convex, lower semicontinuous, decreasing, and $\Phi(0) = +\infty$, $\Phi(2) = \sum(\rho_i/\alpha)^2 = 1/\alpha^2 < 1$ (since the balls B_i cover almost all of the unit ball, then $\sum \rho_i^2 = 1$). We claim that A has dimension

$$d := \inf\{t : \Phi(t) \leq 1\} \ . \tag{8.16}$$

Proof. Take t such that $\Phi(t) \leq 1$ and fix $\delta > 0$. Now choose a positive integer k so that $2(\rho/\alpha)^k \leq \delta$. Then the balls $\{B^k_{\mathbf{i}} : \mathbf{i} \in I^k\}$ cover A (cf. (8.2)) and have diameters equal to $2\alpha^{-k}\rho_{i_1}\cdots\rho_{i_k} \leq 2(\rho/\alpha)^k \leq \delta$ (cf. Proposition 8.4(4)). Hence

$$\begin{aligned}\mathscr{H}^t_\delta(A) \leq \frac{\omega_t}{2^t}\sum_{\mathbf{i}\in I^k}\big(\mathrm{diam} B^k_{\mathbf{i}}\big)^t &= \omega_t \sum_{\mathbf{i}\in I^k}\big(\rho_{i_1}\cdots\rho_{i_k}\alpha^{-k}\big)^t \\ &= \omega_t\Big(\sum_{i\in I}(\rho_i/\alpha)^t\Big)^k = \omega_t\big(\Phi(t)\big)^k \leq \omega_t\end{aligned}$$

(here $\mathscr{H}^t_\delta$ is the t-dimensional Hausdorff pre-measure, and ω_t is a renormalization factor which agrees with the measure of the unit ball in $\mathbb{R}^t$ when t is integer).

Since δ is arbitrarily taken, we obtain that $\mathscr{H}^t(A) \leq \omega_t$ and then $\dim(A) \leq t$, and since this holds for every t such that $\Phi(t) \leq 1$, by the definition of d we deduce that $\dim(A) \leq d$.

It remains to prove the opposite inequality. Take $s < d$. Then $\Phi(s) > 1$, and so there exists a finite subset J of I such that

$$\sum_{i\in J}(\rho_i/\alpha)^s \geq 1 \ . \tag{8.17}$$

Let $T_J : \mathcal{P}(\mathbb{R}^2) \to \mathcal{P}(\mathbb{R}^2)$ be the map associated to the finite family of similitudes $\{\Psi_i : i \in J\}$ as in (8.13). Notice that each Ψ_i has scaling factor $r_i := \rho_i/\alpha$, and if we set $D := B(0,1)$, then the balls $\{\Psi_i(D) : i \in J\}$ are pairwise distant and included in D. Hence T_J admits a closed fixed point $C \subset \overline{D}$, and the dimension of C is the unique solution of the equation

$$\sum_{i\in J}(\rho_i/\alpha)^t = 1 \ , \tag{8.18}$$

so that $\dim(C) \geq s$ by (8.17). Moreover C is given by (see [Hu])

$$C = \bigcap_{k=1}^{\infty}\Big(\bigcup_{\mathbf{i}\in J^k}\big(\Psi_{i_1}\circ\cdots\circ\Psi_{i_k}\big)(\overline{D})\Big) = \bigcap_{k=1}^{\infty}\Big(\bigcup_{\mathbf{i}\in J^k}\overline{B^k_{\mathbf{i}}}\Big) \ ,$$

and since $\overline{B^k_{\mathbf{i}}} \subset \alpha B^k_{\mathbf{i}} \subset A_k$ for every k and every $\mathbf{i} \in I^k$, we obtain that $C \subset A$, and then $\dim(A) \geq \dim(C) \geq s$. Since s is any real number smaller than d, we have proved that $\dim(A) \geq d$. □

Remark 8.11. In (8.16) we give an explicit formulation for the Hausdorff dimension d of the set A, and d depends both on the choice of the parameter α and of the basic cover $\{B_i : i \in I\}$ (cf. the definition of the function Φ in (8.15)), but it is not clear which values d can actually assume.

In particular we remark that the basic cover $\{B_i\}$ may be taken so that $\sum_i \rho_i^t = +\infty$ for every $t < 2$, hence $\Phi(t) = +\infty$ for every $t < 2$ (regardless of the value of α), and then $\dim(A) = 2$ even if A is negligible (cf. (8.15) and (8.16)).

It turns out that the dimension of A is tightly connected with the so-called exponent of the packing $\{\overline{B}_i : i \in I\}$. Let us recall some notation: a *(spherical) packing* of a closed domain Ω in the plane is any family $\{B_h : h \in \mathbb{N}\}$ of closed balls included in Ω with pairwise disjoint interiors which cover almost all of $\overline{\Omega}$; a packing is called *osculatory* if the ball B_{h+1} is chosen in order to maximize the radius among all closed balls contained in the closure of $\Omega \setminus \big(B_1 \cup \cdots \cup B_h\big)$; the *residual set* of the packing is the set $E = \Omega \setminus \cup B_h$; the *exponent* of the packing is the infimum e of all real numbers t such that the radii ρ_h of the balls B_h are t-summable, that is, $e := \inf\big\{t \geq 0 : \sum_h \rho_h^t < +\infty\big\}$. When Ω is a closed ball, it is known that $e \geq \dim(E) \geq 1.03$ for every packing, and $1.315 \geq e = \dim(E) \geq 1.300$ for every osculatory packing.

For a complete survey of the results about plane spherical packings, and for detailed references as well, we refer to the remarks at the end of section 8.4 in Falconer [Fa1]. For recent results about general packings in arbitrary dimension see also [Gr].

Now let e be the exponent of the basic cover $\{B_i : i \in I\}$, and fix any real number t such that $e < t < 2$. Then $\sum \rho_i^t < +\infty$, and we may choose the parameter α so that $\Phi(t) = \alpha^{-t} \sum \rho_i^t = 1$. Hence $\dim(A) = t$. On the other hand $\Phi(t) = +\infty$ for every $t < e$ (regardless of the value of α), and then $\dim(A) \geq e$. Taking the previous estimates on e into account, we deduce that for every choice of the basic cover $\{B_i : i \in I\}$ and of the parameter α, $\dim(A) \geq 1.03$, and if the basic cover is an osculatory packing, then we may choose α so that $\dim(A)$ is any number between 1.315 and 2.

Eventually we give a simple example of packing with exponent strictly lower than 2 which was suggested to us by E. De Giorgi. Let Ω be any bounded open set in $\mathbb{R}^2$ with smooth boundary, and take a, b so that $0 < b < 1$ and $0 < a < 1/2$. Now let $C_0 := \mathbb{R}^2 \setminus \Omega$, and let $\mathcal{F}_1$ be the family of all closed balls $B(x, ab)$ whose center x belongs to the lattice $b\mathbb{Z}^2$ and whose distance from C_0 is not lower than b. Now let C_1 be the union of C_0 and all balls in $\mathcal{F}_1$, and let $\mathcal{F}_2$ be the family of all closed balls $B(x, ab^2)$ whose center x belongs to the lattice $b^2\mathbb{Z}^2$ and whose distance from C_1 is not lower than b^2. Define $\mathcal{F}_h$ in the same way for $h = 3, 4, \ldots$ and let $\mathcal{F}$ be the union of all families $\mathcal{F}_h$. It is not difficult to prove that $\mathcal{F}$ is a packing of Ω provided that, say, $b \leq (1 - 2a)(\sqrt{2} + 2 - 2a)^{-1}$, and that the exponent of $\mathcal{F}$ is lower than $2 - \log(1 - a^2/8) \big/ \log b$ (both inequalities are far from optimal).

8.12. *The dimension of the measure Ju*

The pointwise dimension of a finite positive measure μ is defined at every

point x in the support of μ as the following limit (if it exists)

$$\lim_{r \to 0} \frac{\log \mu\big(B(x, r)\big)}{\log r} , \tag{8.19}$$

and we say that μ has dimension d if it has pointwise dimension equal to d for μ-almost every point (the number d is unique, but it does not necessarily exist).

Let μ be a measure with dimension d. It follows immediately from the previous definition that the t-dimensional density of μ is μ-a.e. equal to $+\infty$ if $t > d$, and μ-a.e. equal to 0 if $t < d$. Therefore by Proposition 4.9 in [Fa2], we obtain that

(i) for every Borel set B, $\dim(B) < d$ implies $\mu(B) = 0$;
(ii) there exists a Borel set E with $\dim(E) = d$ which supports μ.

(We remark that (i) and (ii) do not imply that μ may be represented as $f \cdot \mathscr{H}^d$ for some $f \in L^1(\mathscr{H}^d)$; in order to obtain such representation we would need that the d-dimensional upper density of μ is finite and strictly positive μ-a.e.).

In general a measure μ has no specific dimension. Yet the dimension is well-defined, and more or less explicitly computable, for a large class of "self-similar" measures. As in paragraph 8.10, let be given finitely many similitudes $\{\Psi_i : i \in I\}$ with scaling factor r_i, let D be a bounded open set such that the sets $\big\{\Psi_i(D) : i \in I\big\}$ are pairwise distant and included in D, and let C be the compact set such that $T(C) = C$, where $T : \mathcal{P}(\mathbb{R}^2) \to \mathcal{P}(\mathbb{R}^2)$ is defined as in (8.13). Fix now positive numbers $\{p_i : i \in I\}$ so that $\sum p_i = 1$. Then there exists a unique probability measure μ supported on C such that $\mu(B) = p_i\, \mu\big(\Psi_i(B)\big)$ for every Borel set $B \subset C$ (which is therefore referred to as the *self-similar probability measure* associated with the families $\{\Psi_i : i \in I\}$ and $\{p_i : i \in I\}$). This measure is determined by the condition

$$\mu\big(\Psi_{i_1} \circ \cdots \circ \Psi_{i_k}(\overline{D})\big) = p_{i_1} \cdots p_{i_k} \qquad \text{for every } k = 1, 2, \ldots \text{ and } (i_1, \ldots, i_k) \in I^k, \tag{8.20}$$

and it may be proved that μ has dimension equal to

$$\dim(\mu) = \frac{\sum p_i \log p_i}{\sum p_i \log r_i} . \tag{8.21}$$

In particular we obtain that μ and C have the same dimension if and only if the ratio $\log r_i \big/ \log p_i$ does not depend on i, otherwise $\dim(\mu) < \dim(C)$. Moreover, we conjecture that the measure μ can be represented as $f \cdot \mathscr{H}^d$ for some d and some $f \in L^1(\mathscr{H}^d)$ if and only if $\dim(\mu) = \dim(C)$ (this fact is proved when I consists of two elements).

The equality (8.21) seems to be a well-known result, but somehow hidden in the literature about multifractal decomposition: for example it may be found in [CM], putting together Remark 2.12 and Theorem 3.3 (in the notation of that paper ρ plays the rôle of μ, t_i of r_i, and $\alpha(1) = f(\alpha(1))$ turns out to be equal to the right hand side of (8.21)).

Now we want to show that the measure Ju may be regarded as an example of self-similar measure in the above sense. We already remarked that Ju is supported on A, and that A is the fixed point of the map T associated in (8.13) with the countable family of similitudes $\{\Psi_i : i \in I\}$ given in Definition 8.3. Moreover each Ψ_i has scaling factor ρ_i/α, and it may be verified that for every $k = 1, 2, \dots$ and $\mathbf{i} = (i_1, \dots, i_k) \in I^k$, the map u (in fact, the map u_h for every $h \geq k$) takes the ball $B^k_{\mathbf{i}}$ into a ball of radius $\rho_{i_1} \cdots \rho_{i_k}$, and then Ju satisfies the following analogous of (8.20):

$$Ju(B^k_{\mathbf{i}}) = \big|u(B^k_{\mathbf{i}})\big| = \pi\big(\rho_{i_1} \cdots \rho_{i_k}\big)^2 \qquad \text{for every } k \text{ and } \mathbf{i} \in I^k. \tag{8.22}$$

Thus in this case ρ_i/α plays the rôle of r_i and ρ_i^2 plays the rôle of p_i, and therefore we *conjecture* that, at least when $\sum \rho_i^2 \log \rho_i$ is finite, the dimension of Ju is given by the following analogous of (8.21):

$$\dim(Ju) = \frac{\sum \rho_i^2 \log(\rho_i^2)}{\sum \rho_i^2 \log(\rho_i/\alpha)} = \frac{2\sum \rho_i^2 \log \rho_i}{-\log\alpha + \sum \rho_i^2 \log \rho_i} \,. \tag{8.23}$$

Notice that when α ranges in $(1, +\infty)$ the right hand side of (8.23) takes all values between 0 and 2. For technical reasons we have carried out the proof of (8.21) only when I is a finite index set, and so identity (8.23) is only a conjecture.

We conclude this section by showing that the function u given in Theorem 8.6 is also the limit of a sequence of monotone functions $(\bar u_k)$ whose derivatives and Jacobians have a particular structure (see Definition 8.13 and Proposition 8.14), and then we use this example to discuss the possibility of an equivalent of the SBV theory for monotone functions, which involves also the weak Jacobian (paragraph 8.15). A certain knowledge of the SBV theory is requested.

Throughout the rest of this section we assume $\alpha \leq 1/\rho$.

Definition 8.13. *For every ball $B = B(\bar x, r)$ and every $\sigma < 1$, we define the multifunction $\bar f_{(\sigma,B)}$ on $\mathbb{R}^2$ by setting*

$$\bar f_{(\sigma,B)}(x) := \begin{cases} \bar x & \text{if } |x-\bar x| < \sigma r, \\ \bar x + \Big[0, r\frac{x-\bar x}{|x-\bar x|}\Big] & \text{if } |x-\bar x| = \sigma r, \\ \bar x + r\frac{x-\bar x}{|x-\bar x|} & \text{if } |x-\bar x| > \sigma r. \end{cases} \tag{8.24}$$

For every $k = 1, 2 \dots$ and every $\mathbf{i} \in I^k$, we set $r_{\mathbf{i}} := \alpha^{-k}\rho_{i_1} \cdots \rho_{i_k}$, and $\sigma_{\mathbf{i}} := \alpha^{2k} r_{\mathbf{i}}/2$. Thus $r_{\mathbf{i}}$ is the radius of the ball $B^k_{\mathbf{i}}$ (cf. proposition 8.4(4)), and $\sigma_{\mathbf{i}} \leq 1/2$ (because we assumed $\alpha \leq 1/\rho$). Hence the following definitions are well-posed:

$$\bar f_k(x) := \begin{cases} \bar f_{(\sigma_{\mathbf{i}}, B^k_{\mathbf{i}})}(x) & \text{if } x \in B^k_{\mathbf{i}} \text{ for some } \mathbf{i} \in I^k, \\ x & \text{otherwise,} \end{cases} \tag{8.25}$$

and

$$\bar u_k := u_k \circ \bar f_k \,, \qquad D_k := \bigcup_{\mathbf{i} \in I^k} \partial\big(\sigma_{\mathbf{i}} B^k_{\mathbf{i}}\big) \,. \tag{8.26}$$

Notice that $\bar f_k$ and $\bar u_k$ are multifunctions.

The function $\bar f_{(\sigma,B)}$ takes the rescaled ball σB into the center of B, takes $\partial(\sigma B)$ onto $\overline{B}$, and $B \setminus \sigma\overline{B}$ onto the boundary of B, and $\bar f_{(\sigma,B)}(x) = x$ for every $x \in \partial B$. Moreover $\bar f_{(\sigma,B)}$ is the subdifferential of the convex function $\langle \bar x, x\rangle + r\,\max\big\{\sigma; |x-\bar x|\big\}$, it is smooth in every point out of $\partial(\sigma B)$, and $\Sigma^1 f_{(\sigma,B)} = \partial(\sigma B)$ while $\Sigma^2 f_{(\sigma,B)}$ is empty (cf. Definition 2.1).

Proposition 8.14. *The functions $\bar u_k$ given in (8.26) enjoy the following properties:*

1. *the sequence $(\bar u_k)$ converges to u uniformly on compact sets;*
2. *$\bar u_k$ is the subdifferential of a convex function, $\Sigma^1 \bar u_k = D_k$ and $\Sigma^2 \bar u_k$ is empty;*
3. *the distributional derivative of $\bar u_k$ has no Cantor part (cf. paragraph 5.6), and since the singular set of $\bar u_k$ is D_k, this means that*
$$D\bar u_k = \nabla \bar u_k \cdot \mathscr{L}_n + (\bar u_h^+ - \bar u_k^-) \otimes \nu \cdot \mathscr{H}^1 \llcorner D_k \;; \tag{8.27}$$
moreover $|\nabla \bar u_k| \leq |\nabla u_k|$ a.e., and $|\bar u_h^+ - \bar u_k^-| = \alpha^k r_{\mathbf{i}} \leq \rho^k$ in D_k;
4. *the weak Jacobian of u_k is given by $Ju_k = \mathscr{H}^1 \llcorner D_k$, and $\mathscr{H}^1(D_k) = \pi$ for every k.*
5. *the gradients $|\nabla \bar u_k|$ are uniformly bounded in $L^p(\Omega)$ for every bounded open set Ω and every $p < 2$,*

Proof. Take $x \in \mathbb{R}^2$. If x does not belong to A_k then $\bar f_k(x) = x$, and if x belongs to $B^k_{\mathbf{i}}$ for some $\mathbf{i} \in I^k$ then $\bar f_k(x) \subset B^k_{\mathbf{i}}$. Hence for every $z \in \bar f_k(x)$ there holds $|z - x| \leq 2(\rho/\alpha)^k$ (because the diameter of $B^k_{\mathbf{i}}$ is lower than $2(\rho/\alpha)^k$), and then $\big|u_k(z) - u_k(x)\big| \leq 2\rho^k$ because u_k is (α^k)-Lipschitz (cf. Step 3, statement (2)). This means that for every $x \in \mathbb{R}^2$ and every $y \in \bar u_k(x)$ we have $\big|u_k(x) - y\big| \leq 2\rho^k$, and since the functions u_k converge to u uniformly on compact sets, the same holds for the functions $\bar u_k$. Statement (1) is proved.

Let k be fixed. In order to prove statement (2) we consider the functions $\bar f_k^\varepsilon$ defined as in (8.25) by replacing $\bar f_{(\sigma,B)}$ with

$$\bar f^\varepsilon_{(\sigma,B)} := \begin{cases} \bar x & \text{if } |x-\bar x| < (1-\varepsilon)\sigma r, \\ \bar x + \frac{r}{\varepsilon}\Big(1 - \frac{(1-\varepsilon)\sigma}{|x-\bar x|}\Big)(x-\bar x) & \text{if } (1-\varepsilon)\sigma r \le |x-\bar x| \le \sigma r, \\ \bar x + r\frac{x-\bar x}{|x-\bar x|} & \text{if } |x-\bar x| > \sigma r. \end{cases}$$

It may be checked that $\bar f_k^\varepsilon$ is a Lipschitz function whose gradient is a symmetric matrix for almost every point, and reasoning as in the proof of statement (5) of Step 3, we obtain that also $u_k \circ \bar f_k^\varepsilon$ is a Lipschitz function whose gradient is a symmetric matrix for almost every point, and then $u_k \circ \bar f_k^\varepsilon$ is the gradient of a convex function g_ε of class $C^{1,1}$.

Moreover when ε tends to 0 the functions $u_k \circ \bar f_k^\varepsilon$ converge to $u_k \circ \bar f_k = \bar u_k$ with respect to the Kuratowski convergence of graphs (cf. Definition 1.6), and then $\bar u_k$ is the subdifferential of a convex function (more precisely, the subdifferential of the limit of the functions g_ε, cf. Theorem 7.7), and the rest of statement (2) follows by recalling that u_k is Lipschitz, $\Sigma^1 \bar f_k = D_k$, and $\Sigma^2 \bar f_k$ is empty.

To prove statement (3) we take an increasing sequence of *finite* sets J_{kh} whose union is I^k, and consider the functions $\bar f_{kh}$ defined as in (8.25) with I^k replaced by J_{kh}. Then $\bar f_{kh}$ and $u_k \circ \bar f_{kh}$ are of class $W^{1,\infty}$ in the complement of the set D_{kh} given by the union of the circles $\partial(\sigma_{\mathbf i} B^k_{\mathbf i})$ over all $\mathbf i \in J_{kh}$ (notice that D_{kh} is closed because is a *finite* union of closed circles).

Then the distributional derivative of $u_k \circ \bar f_{kh}$ has no singular part in $\mathbb R^2 \setminus D_{kh}$, and since D_{kh} is rectifiable, then it has no Cantor part at all (cf. paragraph 5.6). Moreover it may be verified that when h tends to infinity, the functions $u_k \circ \bar f_{kh}$ converge to $u_k \circ \bar f_k = \bar u_k$ in the BV norm on every bounded open set, and then $D\bar u_k$ has no Cantor part. The rest of statement (3) follows from statement (4) in Step 3, paragraph 5.6 and formulas (8.24) and (8.25).

From equality (5.13) we obtain that $J\bar u_k(B) = |\bar u_k(B)|$ for every Borel set B, and moreover, for every $\mathbf i \in I^k$, $\bar f_k$ takes the circle $\partial(\sigma_{\mathbf i} B^k_{\mathbf i})$ onto $B^k_{\mathbf i}$, and u_k takes $B^k_{\mathbf i}$ onto a ball of radius $\alpha^k r_{\mathbf i}$ (recall that by statement (4) in Step 3, u_k agrees on $B^k_{\mathbf i}$ with a similitude with scaling factor α^k). Hence

$$J\bar u_k(\partial(\sigma_{\mathbf i} B^k_{\mathbf i})) = \big|\bar u_k(\partial(\sigma_{\mathbf i} B^k_{\mathbf i}))\big| = \big|u_k(B^k_{\mathbf i})\big| = \pi(\alpha^k r_{\mathbf i})^2 = 2\pi\sigma_{\mathbf i} r_{\mathbf i}$$

Moreover it is clear from the construction that the measure $J\bar u_k$ is uniformly distributed on the circle $\partial(\sigma_{\mathbf i} B^k_{\mathbf i})$, which has radius $\sigma_{\mathbf i} r_{\mathbf i}$, and then $J\bar u_k \llcorner \partial(\sigma_{\mathbf i} B^k_{\mathbf i}) = \mathscr H^1 \llcorner \partial(\sigma_{\mathbf i} B^k_{\mathbf i})$. Hence

$$J\bar u_k \llcorner D_k = \mathscr H^1 \llcorner D_k \ . \tag{8.28}$$

Eventually we remark that $J\bar u_k(D_k) = |\bar u_k(D_k)| = |u_k(A_k)| = \pi$ and $J\bar u_k(\mathbb R^2) = |\bar u_k(\mathbb R^2)| = |A_0| = \pi$ (cf. Lemma 8.5(1)). Hence $J\bar u_k$ is supported on the set D_k, and then (8.28) implies statement (4). Statement (5) follows from Lemma 8.5(3) and statement (3). □

8.15. *About special classes of monotone functions*

The class BV of all functions with no Cantor part in the derivative (see paragraph 5.6), the so-called SBV space, has been widely used in the recent years to study variational problems involving volume and surface energies (see for instance [Am1], [Am2] and references therein). The basic theorem of the theory of SBV functions is essentially a closure result: if we denote by $\mathcal G(p,C)$ the class of all SBV functions on a bounded open set Ω satisfying the condition

$$\int_\Omega |\nabla u|^p \, d\mathscr L_n + \mathscr H^{n-1}(Su) \le C < +\infty \ , \tag{8.29}$$

then the closure of $\mathcal G(p,C)$ with respect to the weak convergence in $BV(\Omega)$ is contained in SBV provided that $p > 1$ (a more general version of this theorem was given by the second author in [Am1], see also [AM] for a simple proof).

Roughly speaking, the point of this closure theorem is the following: given a sequence (u_h) in SBV which converges weakly in $BV(\Omega)$, condition (8.29) forces the absolutely continuous part of derivatives $D_a u_h$ and the jump parts $D_j u_h$ to converge respectively to an n-dimensional and a $(n-1)$-dimensional measure, which turn out to be the absolutely continuous part and the jump part of the derivative of the limit function, which has therefore no Cantor part.

At least for $n = 2$ we would like to define the analogous of the SBV class for monotone functions as a class of functions in $\mathscr{M}on(\Omega)$ whose distributional derivatives and weak Jacobians belong to a restricted class of measures. This would be a first step towards theorems for more general classes of functions.

There are different possible definitions: for example we can introduce the class $\mathscr{M}on_1(\Omega)$ of all monotone functions in $SBV(\Omega)$ whose weak Jacobian is a sum of a 2-dimensional and a 1-dimensional part, that is, Ju may be represented as (cf. Proposition 5.14)

$$Ju = \det(\nabla u)\cdot \mathscr L_2 + \Phi^1_u \cdot \mathscr H^1 \ , \tag{8.30}$$

where Φ^1_u is an $\mathscr H^1$-summable function (we may assume in addition that Φ^1_u is supported on a countably $\mathscr H^1$-rectifiable set). We can also define the class $\mathscr{M}on_2(\Omega)$ of all monotone functions in $SBV(\Omega)$ whose weak Jacobian is a sum of h-dimensional parts with $h = 0,1,2$, that is, Ju may be represented as

$$Ju = \det(\nabla u)\cdot \mathscr L_2 + \Phi^1_u \cdot \mathscr H^1 + \Phi^0_u \cdot \mathscr H^0 \ , \tag{8.31}$$

where Φ_u^1 and Φ_u^0 are respectively $\mathscr{H}^1$ and $\mathscr{H}^0$-summable.

Now the sequence $(\bar{u}_k)$ given in Definition 8.13 shows that for both classes there holds no analogous of the closure theorem for SBV functions. Indeed the sequence $(\bar{u}_k)$ is included in $\mathcal{G}(p, C)$ for every $p < 2$ for suitable C (statements (3) and (5) of Proposition 8.14), and the weak Jacobians can be written as $J\bar{u}_k = \mathscr{H}^1 \llcorner D_k$ where the sets D_k are $\mathscr{H}^1$-rectifiable and the measure $\mathscr{H}^1(D_k)$ are uniformly bounded (statement (4) of Proposition 8.14), while the weak Jacobian of the limit u can be written neither in the form (8.30) nor in the form (8.31) (cf. statement (6) of Theorem 8.6; see also paragraphs 8.10 and 8.12).

Remark 8.16. The conclusions of the previous paragraph hold even if we restrict our attention to subdifferentials of convex functions (cf. Proposition 8.14(2)).

We remark here that given a convex function f on the plane, the distributional derivative and the weak Jacobian of the subdifferential are tightly connected with the weak definitions of the second fundamental form and the Gaussian curvature of the graph of f, viewed as a 2-dimensional convex surface in the space (see [Fe2] for a precise definition of curvature measures), and paragraph 8.15 is therefore related to the existence of closure theorems for special classes of convex surfaces (e.g., surfaces whose curvature measures can be written as a sum of integer dimensional measures) considered in [AO], section 5 (see also [Os]).

References

[A1] G. Alberti: Rank-one property for derivatives of functions with bounded variation, Proc. Royal Soc. Edinburgh **123-A** (1993), 239-274.

[A2] G. Alberti: On the structure of singular sets of convex functions, Calc. Var. **2** (1994), 17-27.

[AAC] G. Alberti, L. Ambrosio, P. Cannarsa: On the singularities of convex functions, Manuscripta Math. **76** (1992), 421-435.

[ACD] L. Ambrosio, A. Coscia, G. Dal Maso: Fine properties of functions in $BD(\Omega)$, Arch. Rat. Mech. Anal. **139** (1997), 201-238.

[AD] L. Ambrosio, G. Dal Maso: On the relaxation in $BV(\Omega;\mathbf{R}^m)$ of quasi-convex integrals, J. Funct. Anal. **109** (1992), 76-97.

[AG] P. Aviles, Y. Giga: Singularities and rank one properties of hessian measures, Duke Math. J. **58** (1989), 441-467.

[AK] R.D. Anderson, V.L. Klee: Convex functions and upper semicontinuous functions, Duke Math. J. **19** (1952), 349-357.

[Ale] A.D. Aleksandrov: Almost everywhere existence of the second differential of a convex function and some properties of convex surfaces connected with it, Leningrad Univ. Ann. (Math. ser.) **6** (1939), 3-35 (in Russian).

[Am1] L. Ambrosio: A compactness theorem for a special class of functions of bounded variation, Boll. Un. Mat. Ital. **3-B** (1989), 857-881.

[Am2] L. Ambrosio: Existence theory for a new class of variational problems, Arch. Rat. Mech. Anal. **111** (1990), 291-322.

[Am3] L. Ambrosio: Su alcune proprietà delle funzioni convesse, Rend. Mat. Acc. Lincei (ser. IX) **3** (1992), 193-202.

[AM] G. Alberti, C. Mantegazza: A note on the theory of SBV functions, Boll. Un. Mat. It. (7), **11-B** (1997), 375-382.

[AO] G. Anzellotti, E. Ossanna: Singular sets of convex bodies and surfaces with generalized curvatures, Manuscripta Math. **86** (1995), 417-433.

[Asp] E. Asplund: Frechét differentiability of convex functions, Acta Math. **121** (1968), 31-48.

[At] H. Attouch: Variational Convergence for Functions and Operators, Research Notes in Math., Pitman, London 1984.

[Ba] J.M. Ball: Convexity conditions and existence theorems in nonlinear elasticity, Arch. Rat. Mech. Anal. **63** (1977), 337-403.

[BCP] G. Bianchi, A. Colesanti, C. Pucci: On the second differentiability of convex surfaces, Geom. Dedicata **60** (1996), 39-48.

[Be] A.S. Besicovitch: On singular points of convex surfaces, in Proceedings of Symposia in Pure Mathematics **VII**, pp. 21-23, Am. Math. Soc., Providence 1963.

[Br] H. Brezis: Opérateurs Maximaux et Semi-groupes de Contractions dans les Espaces de Hilbert, North Holland, Amsterdam 1973.

[CIL] M.G. Crandall, H. Ishii, P.L. Lions: User's guide to viscosity solutions of second order partial differential equations, Bull. of the Amer. Math. Soc. **27** (1992), 1-67.

[Cl] F.H. Clarke: Optimization and Nonsmooth Analysis, John Wiley & Sons, New York 1983.

[CM] R. Cawley, R.D. Mauldin: Multifractal decomposition of Moran fractals, Adv. in Math. **92** (1992), 196-236.

[CV] C. Castaing, M. Valadier: Convex Analysis and Measurable Multifunction, Lect. Notes in Math. **580**, Springer-Verlag, Berlin 1977.

[CZ] A.P. Calderón, A. Zygmund: On the differentiability of functions which are of bounded variation in Tonelli's sense, Rev. Union Mat. Argentina **20** (1960), 102-121.

[Da] B. Dacorogna: Direct Methods in the Calculus of Variations, Applied Math. Sciences **78**, Springer-Verlag, New York 1989.

[DM] G. Dal Maso: An Introduction to Γ-convergence, Progress in Nonlinear Diff. Eq. and Appl. **8**, Birkhäuser, Boston 1993.

[Du] R.M. Dudley: On second derivatives of convex functions, Math. Scand. **41** (1977), 159-174.

[EG] L.C. Evans, R.F. Gariepy: Lecture Notes on Measure Theory and Fine Properties of Functions, Studies in Advanced Math., CRC Press, Boca Raton 1992.

[Fa1] K.J. Falconer: The Geometry of Fractal Sets, Cambridge University Press, Cambridge 1985.

[Fa2] K.J. Falconer: Fractal Geometry. Mathematical Foundations and applications, John Wiley & Sons, Chichester 1990.

[Fe1] H. Federer: Geometric Measure Theory, Springer-Verlag, Berlin 1969.

[Fe2] H. Federer: Curvature measures, Trans. Am. Math. Soc. **93** (1959), 418-491.

[FM] I. Fonseca, S. Müller: Relaxation of quasiconvex functionals in $BV(\Omega,\mathbf{R}^p$ for integrands *f(x,u,Du)*, Arch. Rat. Mech. Anal. **123** (1993), 1-49.

[GMS1] M. Giaquinta, G. Modica, J. Souček: Cartesian currents, weak diffeomorphism, and existence theorems in nonlinear elasticity, Arch. Rat. Mech. Anal.

106 (1989), 97-159. Erratum and addendum, Arch. Rat. Mech. Anal. **109** (1990), 385-392.

[GMS2] M. Giaquinta, G. Modica, J. Souček: Graphs of finite mass which cannot be approximated in area by smooth graphs, Manuscripta Math. 78 (1993), 259-271.

[GMS3] M. Giaquinta, G. Modica, J. Souček: Area and the area formula, Rend Semin. Mat Fis. Milano **62** (1992), 53-87.

[Gr] P.M. Gruber: How well can space be packed with smooth bodies? Measure theoretic results, J. London Math. Soc. **52** (1995), 1-14.

[Hu] J.E. Hutchinson: Fractals and self-similarity, Indiana Univ. Math. J. **30** (1981), 713-747.

[LM] S. Luckhaus, L. Modica: The Gibbs-Thompson relation within the gradient theory of phase transitions, Arch. Rat. Mech. Anal. **107** (1989), 71-83.

[Mi] G. Minty: Monotone nonlinear operators on a Hilbert space, Duke Math. J. **29** (1962), 341-346.

[Mig] F. Mignot: Contrôle optimal dans les inéquations variationelles elliptiques, J. Funct. Anal. **22** (1976), 130-185.

[Mor] F. Morgan: Geometric Measure Theory. A beginner's guide. Academic Press, Boston 1988.

[Mu1] S. Müller: Weak continuity of determinants and nonlinear elasticity, C.R. Acad. Sci. Paris, **307-I** (1988), 501-506.

[Mu2] S. Müller: Det=det. A remark on the distributional determinant, C.R. Acad. Sci. Paris, **311-I** (1990), 13-17.

[Mu3] S. Müller: On the singular support of the distributional determinant, Ann. I.H.P. Analyse non linéaire **10** (1993), 657-696.

[Muc] D. Mucci: Approximation in area of graphs with isolated singularities, Manuscripta Math. **88** (1995), 135-146.

[Os] E. Ossanna: Superfici Generalizzate Dotate di Curvature Deboli, PhD thesis, University of Trento.

[Ph] R.R. Phelps: Convex Functions, Monotone Operators and Differentiability, Lect. Notes in Math. **1364**, Springer-Verlag, New York 1989.

[Po] S.P. Ponomarev: Property N of homeomorphisms of the class $W^{1,p}$, *Siberian Math. J.* **28** (1987), 291-298.

[PZ] D. Preiss, L. Zajíček: Fréchet differentiation of convex functions in a Banach space with separable dual, Proc. Amer. Math. Soc. **91** (1984), 202-204.

[Re1] Yu.G. Reshetnyak: Weak convergence of completely additive vector functions on a set, Siberian Math. J. **9** (1968), 1039-1045.

[Re2] Yu.G. Reshetnyak: Generalized derivatives and differentiability almost everywhere, Math. USSR Sb. **4** (1968), 293-302.

[Ru] W. Rudin: Functional Analysis, Mc Graw-Hill, New York 1973.

[Si] L.M. Simon: Lectures on Geometric Measure Theory. Proc. of the Centre for Mathematical Analysis **3**, Australian National Univ., Canberra 1983.

[Ve] L. Veselý: On the multiplicity points of monotone operators on a separable Banach spaces I and II, Comment. Math. Univ. Carolinæ**27** (1986), 551-570, and **28** (1987), 295-299.

[Vo] A.I. Vol'pert: Spaces BV and quasi-linear equations, Math. USSR Sb. **17** (1967), 225-267.

[Z1] L. Zajíček: On the points of multiplicity of monotone operators, Comment. Mat. Univ. Carolinæ**19** (1978), 179-189.

[Z2] L. Zajíček: On the differentiability of convex functions in finite and infinite dimensional spaces, Czechoslovak Math. J. **29 (104)** (1979), 340-348.