SISSA

SCUOLA INTERNAZIONALE
SUPERIORE di STUDI AVANZATI
International School
for Advanced Studies

REGULARITY RESULTS
FOR HAMILTON-JACOBI EQUATIONS

Ph.D. Thesis

Supervisor
Prof. Stefano Bianchini

Candidate
Daniela Tonon

ACADEMIC YEAR 2010/2011

Il presente lavoro costituisce la tesi presentata da Daniela Tonon, sotto la direzione di ricerca del prof. Stefano Bianchini, al fine di ottenere l'attestato di ricerca post-universitaria di *Doctor Philosophiae in Analisi Matematica* presso la SISSA di Trieste. Ai sensi dell'art. 18, comma 3, dello Statuto della SISSA pubblicato sulla G.U. n. 62 del 15.03.2001, il predetto attestato è equipollente al titolo di *Dottore di Ricerca in Matematica*.

Trieste, anno accademico 2010/2011.

Contents

Introduction

The most part of this thesis is devoted to the study of the regularity of viscosity solutions of Hamilton-Jacobi equations under different hypotheses on the Hamiltonian.

In the last part of this thesis we present a decomposition theorem for BV functions, which extends the Jordan decomposition property.

0.1 Hamilton-Jacobi equations

In this thesis we consider the Hamilton-Jacobi equation

$$\partial_t u + H(t, x, D_x u) = 0 \qquad \text{in } \Omega \subset [0, T] \times \mathbb{R}^n, \tag{1}$$

where $H : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ is called Hamiltonian and Ω is an open domain in $\mathbb{R}^{n+1}$.

It is well known that, even under strong regularity assumptions on the Hamiltonian H and on the initial datum for the Cauchy problem related to (1), classical solutions exist only for a finite interval of time: indeed solutions of the Cauchy problem develop discontinuities of the gradient. The reason why Hamilton-Jacobi equations don't have in general smooth solutions for all times can be explained by the method of characteristics, see Evans [26]. However if we try to overcome the problem considering solutions which satisfy the equation only almost everywhere uniqueness is lost. The introduction of viscosity solutions, see Crandall and Lions [23], Crandall, Evans and Lions [22], and Lions [32], solves the problem of existence, uniqueness and stability even in the case of a continuous initial datum.

In general viscosity solutions are only locally Lipschitz continuous. The structure of the non differentiability set of viscosity solutions has been studied by Fleming [28], Cannarsa and Soner [21] and others. As a major assumption these authors restrict the problem to the case where the Hamiltonian $H(t, x, p)$ is convex with respect to p and smooth in all variables. Under this restriction the viscosity solution of (1) can be represented as the value function of a classical problem in Calculus of Variation, see Fleming and Rishel [29]. Indeed, adding suitable assumptions, the viscosity solution is precisely the value function

$$u(t, x) := \min \left\{ u_0(\xi(0)) + \int_0^t L(s, \xi(s), \dot{\xi}(s)) ds \;\middle|\; \xi(t) = x,\ \xi \text{ Lipschitz continuous in } [0, t] \right\},$$

where L, the Lagrangian, is the Legendre transform of H with respect to the last variable and u_0 is the initial datum at time $t = 0$. Moreover, when the Hamiltonian is strictly convex, the viscosity solution is locally semiconcave in both the variables, see Cannarsa and Sinestrari [20].

In particular, for every $K \subset\subset \Omega$, there is a constant $C > 0$ such that the function $(t, x) \mapsto u(t, x) - C(t^2 + |x|^2)$ is concave on K. The semiconcavity of u ensures that $Du = (\partial_t u, D_x u)$ is BV_{loc}, in particular u is twice differentiable almost everywhere and its distributional Hessian is a matric of measures with locally bounded variation. Intuitively one can figure out viscosity solutions as functions which are Lipschitz and whose gradient is piecewise smooth, undergoing jump discontinuities along a family of surfaces of codimension one (in space and time).

Deeper results on fine regularity properties of viscosity solutions have been proved. Using geometric measure theory and the classical method of characteristics, Fleming proved, in [28], that viscosity solutions inherit the regularity of the initial datum in the complement of the closed set $\Lambda \cup \Gamma$, Λ being the non differentiability set of u and Γ the set of conjugate points for the variational problem associated to (1). A significant result, which confirms the intuitive picture given above, was obtained by Cannarsa, Mennucci and Sinestrari in [19]. Requiring the strict convexity of H in the last variable, they proved the SBV regularity of the gradient Du, when u is the viscosity solution of the Cauchy problem of (1) with a regular initial datum $u(0, x) = u_0(x)$ belonging to $W^{1,\infty}(\mathbb{R}^n) \cap C^{R+1}(\mathbb{R}^n)$, with $R \geq 1$. Furthermore they give a sharper estimate on the set of regular conjugate points $\Gamma \setminus \Lambda$, which implies that this set has Hausdorff dimension at most $n - 1$ if the initial datum is C^∞. In particular they proved that the closure of the non differentiability set, which is equal to $\Lambda \cup \Gamma$, is $\mathcal{H}^n$-rectifiable. However the techniques used to obtain these results do not apply when the initial datum is less regular.

The question on the existence of such an SBV-regularizing effect for the gradient of a viscosity solution of the Hamilton-Jacobi equation was addressed by several other authors. For a recent survey on the topic see De Lellis [24]. The motivation for studying this kind of regularity arises from problems in Control Theory, in image segmentation and measure-theoretic questions.

A result, which confirms this SBV-regularizing effect, has been proved by Ambrosio and De Lellis, see [4], for the entropy solution U of the scalar conservation law

$$\partial_t U + D_x(H(U)) = 0 \qquad \text{in } \Omega \subset [0, T] \times \mathbb{R}. \tag{2}$$

Theorem 0.1 (Ambrosio and De Lellis)**.** *Let $U \in L^\infty(\Omega)$ be an entropy solution of (2) with $H \in C^2(\mathbb{R})$ and locally uniformly convex. Then, there exists $S \subset [0, T]$ at most countable such that $\forall t \in [0, T] \setminus S$ the following holds:*

$$U(t, \cdot) \in SBV_{loc}(\Omega_t) \quad \textit{with } \Omega_t := \{x \in \mathbb{R} |\ (t, x) \in \Omega\}.$$

In particular $U \in SBV_{loc}(\Omega)$.

Thanks to the equivalence between the entropy solution U and the gradient of a viscosity solution of a related Hamilton-Jacobi equation, which holds in the one-dimensional case, the same result applies to $D_x u$ when u is a solution of the Hamilton-Jacobi equation

$$\partial_t u + H(D_x u) = 0 \qquad \text{in } \Omega \subset [0, T] \times \mathbb{R}.$$

This equivalence is in general not true in the multi-dimensional case.

A recent generalization of the SBV-regularizing effect to the multi-dimensional case has been proved by Bianchini, De Lellis and Robyr in [9].

Theorem 0.2 (Bianchini, De Lellis and Robyr). *Let u be a viscosity solution of*

$$\partial_t u + H(D_x u) = 0, \tag{3}$$

in $\Omega \subset [0,T] \times \mathbb{R}^n$, assume H belongs to $C^2(\mathbb{R}^n)$ and

$$c_H^{-1} Id_n(p) \leq H_{pp}(p) \leq c_H Id_n(p)$$

for some $c_H > 0$. Then the set of times

$$S := \{t | \ D_x u(t,\cdot) \notin SBV_{loc}(\Omega_t)\}$$

is at most countable. In particular $D_x u \in [SBV_{loc}(\Omega)]^n, \partial_t u \in SBV_{loc}(\Omega)$.

Due to this result, when the Hamiltonian is uniformly convex, the singular part of the matrix of Radon measures $D_x^2 u$ is concentrated on a $\mathcal{H}^{n-1}$-rectifiable set, the measure theoretic jump set J of $D_x u$. This prevents the Hessian of u from having a Cantor part. Analogous considerations hold for $D_x \partial_t u$ and $\partial_t^2 u$.

The paper of Ambrosio and De Lellis and the one of Bianchini, De Lellis and Robyr prove the SBV-regularizing effect using a strategy whose idea originates from a conjecture pointed out by Bressan, during a conversation on the problem with De Lellis. If $D_x u(\bar{t},\cdot)$ is not SBV for a certain time $\bar{t}$, then at future times $\bar{t}+\varepsilon$ the Cantor part of $D_x^2 u(\bar{t},\cdot)$ gets transformed into jump singularities. Roughly speaking this allows to conclude that $D_x u(t,\cdot)$ is SBV out of a countable number of t's and that $D_x u$ is SBV as a function of two variables. Motivated by this conjecture, the general idea of the proof consists in constructing a monotone bounded functional $F(t)$, whose jumps are related to the presence of a Cantor part in $|D_x^2 u(t,\cdot)|$. Since the boundedness and the monotonicity of this functional imply that it can have only a countable number of jumps, the Cantor part of $|D_x^2 u(t,\cdot)|$ can be different from zero only for a countable number of t's. A key role is played by the map

$$X_{t,0}(x) := x - tH_p(D_x^+ u(t,x)),$$

where $D_x^+ u(t,x)$ is the superdifferential of the semiconcave function $u(t,\cdot)$, and by its restriction $\chi_{t,0}(\cdot)$ to the set U_t where $D_x^+ u(t,x)$ is single-valued. Indeed, the functional $F(t)$ measures exactly the area of a set transported along characteristics from time t to time 0

$$F(t) := \mathcal{H}^n(\chi_{t,0}(U_t)).$$

Thus the properties of characteristics are important. Since H depends only on p, characteristics are straight lines and are called optimal rays. Their no-crossing property ensures the injectivity of the map $\chi_{t,0}$. Other useful ingredients for the proof are two estimates on the measure of the area of a set transported along optimal rays. The first one ensures that, for any Borel set $A \subset \Omega_t$,

$$\mathcal{H}^n(X_{t,0}(A)) \geq c_1 \mathcal{H}^n(A) - c_2 t \int_A d(\Delta u(t,x)),$$

where c_1, c_2 are positive constants and $\Delta u(t,x)$ is the Laplacian of $u(t,\cdot)$, which is a Radon measure. The second ensures that, for any Borel set $A \subset \Omega_t$ and $0 < \delta < t$,

$$\mathcal{H}^n(X_{t,\delta}(A)) \geq \left(\frac{t-\delta}{t}\right)^n \mathcal{H}^n(X_{t,0}(A)).$$

The no-crossing property and the two estimates are enough to prove that, when the Cantor part of $|D_x^2 u(t,\cdot)|(\Omega_t)$ is positive, then there exists a set $A \subset U_t$ of null $\mathcal{H}^n$-measure over which the Cantor part is concentrated and such that

$$\chi_{t,0}(A) \cap \chi_{t+\delta,0}(U_{t+\delta}) = \emptyset$$

for all $\delta > 0$. Therefore the presence of a Cantor part in $|D_x^2 u(t,\cdot)|$ corresponds to a jump for $F(t)$. The geometrical theory of monotone functions, see for example Alberti and Ambrosio [1], plays a crucial role in handling the details for the multi-dimensional case.

Theorem 0.2 can be seen also as a kind of generalization of the result of Cannarsa, Mennucci and Sinestrari in [19]. Indeed, in the case of $H = H(p)$ uniformly convex, Theorem 0.2 contains part of that result since it proves SBV regularity reducing the regularity of the initial datum to bounded Lipschitz functions. Nothing is said, however, about the closure of the non differentiability set.

The result of Bianchini, De Lellis and Robyr suggested to us different directions of research.

i) A first question is about the preservation of the SBV-regularizing effect in the case of a Hamiltonian dependending on space and time, $H(t,x,p)$, which is uniformly convex in the last variable and a bounded Lipschitz initial datum. As already seen the result of Cannarsa, Mennucci and Sinestrari [19] applies to this kind of Hamiltonians when they are strictly convex in the last variable (a slightly weaker requirement) and in the case of a regular initial datum (a stronger requirement). Therefore a positive answer could be seen as a kind of generalization of this result. Moreover it will extend the result of Bianchini, De Lellis and Robyr [9] to the case of a general Hamiltonian depending also on time and space's variables.

ii) Theorem 0.2 implies that the Jacobian $J(t,\cdot) := \mathcal{H}^n(D_x^+ u(t,\cdot))$ is a measure which has only an absolute continuous part with respect to $\mathcal{H}^n$ and a part which is concentrated on a $\mathcal{H}^{n-1}$-rectifiable set and is absolute continuous with respect to $\mathcal{H}^{n-1}$. One can wonder if this measure has only integer parts, i.e. parts which are concentrated on a $\mathcal{H}^k$-rectifiable set and are absolute continuous with respect to $\mathcal{H}^k$ for $k \in \{0,1,\dots,n\}$.

iii) A third question is about the preservation of the SBV-regularizing effect in the case of a Hamiltonian which is only convex. In this case the property of semiconcavity of the solution is lost and its Hessian is no more a measure. However one can try to prove a kind of SBV regularity for the Radon measure $\mathrm{div} H_p(D_x u(t,\cdot))$. A positive answer will generalize Theorem 0.2 because for semiconcave functions the Cantor part of $D_x^2 u(t,\cdot)$ is controlled by the Cantor part of the spatial Laplacian $\Delta u(t,\cdot)$.

iv) One can look for applications of Theorem 0.2. In the one-dimensional case an easy application follows for Convection Theory and systems of sticky particles. One can wonder if that theorem applies also in the multi-dimensional case.

Let us consider in more details all these cases.

i) In Chapter 2 we consider the general Hamilton-Jacobi equation (1), introduced at the beginning of the section, and we require the following assumptions on H:

(H1) $H \in C^3([0,T]\times\mathbb{R}^n\times\mathbb{R}^n)$ with bounded second derivatives and there exist positive constants a, b, c such that

i) $H(t,x,p) \geq -c$,

ii) $H(t,x,0) \leq c$,

iii) $|H_{px}(t,x,p)| \leq a + b|p|$,

(H2) there exists $c_H > 0$ such that

$$c_H^{-1} Id_n(p) \leq H_{pp}(t,x,p) \leq c_H Id_n(p)$$

for any $t \in \mathbb{R}$, $x \in \mathbb{R}^n$.

As already seen these assumptions are necessary to relate our equation to a well defined problem in Calculus of Variations. The idea is to reproduce the strategy seen for the case of a Hamiltonian $H = H(p)$. The main difference from that case is that, due to the dependence of the Hamiltonian on (t,x), characteristics are curves in C^2 but in general they are not straight lines. This is practically the main difficulty to overcome since the strategy used in [9] heavily takes advantage on the simple form of characteristics, while in this case it is not so easy to find an explicit form for them. However, considering sufficiently small intervals of time, one can approximate characteristics with straight lines. Generalized backward characteristics ξ are solutions, together with their dual arc p, of the system

$$\begin{cases} \dot{\xi}(s) = H_p(s,\xi(s),p(s)) \\ \dot{p}(s) = -H_x(s,\xi(s),p(s)) \end{cases} \tag{4}$$

with final conditions

$$\begin{cases} \xi(t) = x \\ p(t) = p, \end{cases}$$

where $p \in D_x^+ u(t,x)$. We will show that one can establish a one to one correspondence between generalized backward characteristics and maximizers of the backward solution

$$u_{t,0}^-(\tau,y) := \max\left\{ u(t,\xi(t)) - \int_\tau^t L(s,\xi(s),\dot{\xi}(s))ds \;\middle|\; \xi(\tau) = y,\ \xi \in [C^2([\tau,t])]^n \right\}.$$

Thus the map $X_{t,\tau}(x)$, now defined as

$$X_{t,\tau}(x) := \{\xi(\tau) |\ \xi(\cdot) \text{ is a solution of (4), with } \xi(t) = x,\ p(t) = p \in D_x^+ u(t,x)\},$$

over the small interval of time $[\tau, \tau+\varepsilon]$, is injective. Handling with care the difference between characteristics and straight lines in a small interval of time, it is possible to recover the no-crossing property and the two area estimates up to an error of the order of ε. Therefore the usual strategy can be easily adapted.

Thus we are able to prove the following theorem.

Theorem 0.3. *Let u be a viscosity solution of (1), assume* H1, H2*. Then the set of times*

$$S := \{t \mid D_x u(t,\cdot) \notin [SBV_{loc}(\Omega_t)]^n\}$$

is at most countable. In particular $D_x u \in [SBV_{loc}(\Omega)]^n, \partial_t u \in SBV_{loc}(\Omega)$.

The results presented in Chapter 2 are contained in Bianchini and Tonon [13].

ii) In Chapter 3 we present one of the consequences of Theorem 0.2. In the case of a uniformly convex Hamiltonian, the Jacobian $J(t,\cdot) := \mathcal{H}^n(D_x^+ u(t,\cdot))$, defined on Ω_t, has a particular structure. Out of a countable number of t's, the measure $J(t,\cdot)$ has only an absolute continuous part with respect to $\mathcal{H}^n$ and a part which is concentrated on a $\mathcal{H}^{n-1}$-rectifiable set and is absolute continuous with respect to $\mathcal{H}^{n-1}$. This is a direct consequence of the fact that $|D_x^2 u(t,\cdot)|$ can have Cantor part for a countable number of t's only. This fact suggests that the Jacobian measure has a structure which admits only integer parts. That is, out of a countable number of t's, $J(t,\cdot)$ can have only parts which are concentrated on a $\mathcal{H}^k$-rectifiable set and are absolute continuous with respect to $\mathcal{H}^k$ for $k \in \{0,1,\ldots,n\}$. However, a counterexample in $\mathbb{R}^2$ shows that this is not true in general. It is possible to find a viscosity solution whose Jacobian has a positive part between $\mathcal{H}^1$ and $\mathcal{H}^0$.

iii) In Chapter 4 we consider the Hamilton-Jacobi equation (3) with a convex Hamiltonian H. When H is smooth and only convex, the Lagrangian L is strictly convex but no more regular. Therefore $u(t,\cdot)$ is no more semiconcave and $D_x u(t,\cdot)$ looses its BV regularity. The only regularity which is true in general for the viscosity solution u is local Lipschitzianity. Thus there is no hope to prove the SBV-regularizing effect for $D_x u(t,\cdot)$ apart from some particular cases. However, a kind of SBV regularity can be proven for the vector field $d(t,x) := H_p(D_x u(t,x))$. This vector field is defined only on the set of points (t,x) where $u(t,x)$ is differentiable in x but can be extended to the all Ω using the optimal rays of the forward solution. Once the vector field is extended, its divergence $\mathrm{div} d(t,\cdot)$ is shown to be a locally finite Radon measure. It is therefore reasonable to see if this measure admits a Cantor part for all t. When the vector field $d(t,\cdot)$ is BV and suitable hypotheses on the Lagrangian L are made, the measure $\mathrm{div} d(t,\cdot)$ has Cantor part only for a countable number of t's in $[0,T]$.

The strategy to obtain our result was suggested by an extension of Theorem 0.1, done by Robyr in [35], for the scalar conservation law

$$\partial_t U(t,x) + D_x(H(t,x,U(t,x))) + g(t,x,U(t,x)) = 0 \quad \text{in } \Omega \subset \mathbb{R}^+ \times \mathbb{R}. \tag{5}$$

Theorem 0.4 (Robyr)**.** *Let $H \in C^2(\mathbb{R}^+ \times \mathbb{R} \times \mathbb{R})$ be a flux function, such that*

$$\{p_i \in \mathbb{R} \mid H_{pp}(t,x,p_i) = 0\}$$

is at most countable for any fixed (t,x). Let $g \in C^1(\mathbb{R}^+ \times \mathbb{R} \times \mathbb{R})$ be a source term and let $U \in BV(\Omega)$ be an entropy solution of the balance law (5).

Then there exists a set $S \subset \mathbb{R}^+$ at most countable such that $\forall t \in \mathbb{R}^+ \setminus S$ the following holds:

$$U(t,\cdot) \in SBV_{loc}(\Omega_t).$$

In particular $U(t,x)$ belongs to $SBV_{loc}(\Omega)$.

First we prove that, in the one-dimensional case, the BV regularity of $d(t,x) = H_p(D_x u(t,x))$ follows automatically in the case of a convex smooth H and there is no need to add hypotheses to prove its SBV regularity out of a countable number of t's. This fact however does not necessarily implies that the same apply to $D_x u(t,x) = U(t,x)$, its BV regularity remains not true in general even in the one-dimensional case.

Moreover the proof of Theorem 0.4 suggested to us an idea to cope with the fact that H is only convex in the multi-dimensional case. As already said, when H is smooth and convex L is not C^2 in general. However, since we are looking at the Cantor part of $\mathrm{div} d(t,\cdot)$, we can reduce to the set $U := \{(t,x)|\ u(t,x)$ is differentiable in $x\}$. Indeed the set $\Omega_t \setminus U_t$ has null $\mathcal{H}^n$-measure for every t and $(\mathrm{div} d(t,\cdot))_c(\Omega_t \setminus U_t) = 0$. Moreover we can consider separately the set of points $(t,x) \in U$ where $L(d(t,x))$ is C^2 and the set of points where $L(d(t,x))$ is not twice differentiable. In the first set we reduce locally to the uniformly convex case. Thus we can apply Theorem 0.2 to obtain the SBV-regularity. In the second, we need to add some hypotheses to handle the problem. Since we are able to prove the regularity of $\mathrm{div} d(t,\cdot)$ in the one-dimensional case the idea is to reduce step by step to dimension one. A way to do this is to require that the vector field $d(t,\cdot)$ is BV and that the Lagrangian is such that the set of points where its Hessian is not defined is contained in a finite number of hyperplanes. We can then study our problem restricted to these hyperplanes reducing the dimension to $n-1$. Repeating the procedure, we need to ask the following hypotheses.

Let H be $C^2(\mathbb{R}^n)$, convex and such that $\lim_{|p|\to\infty} \frac{H(p)}{|p|} = +\infty$.

(HYP(0)) Suppose the vector field $d(t,\cdot)$ belongs to $[BV(\Omega_t)]^n$ for every $t \in [0,T]$.

Define V_{π_n} as

$$V_{\pi_n} := \{v \in \mathbb{R}^n |\ L(\cdot) \text{ is not twice differentiable in } v\},$$

and

$$\Sigma_{\pi_n} := \{(t,x) \in U |\ d(t,x) \in V_{\pi_n}\} \quad \text{and} \quad \Sigma^c_{\pi_n} := U \setminus \Sigma_{\pi_n}.$$

(HYP(n)) We suppose V_{π_n} to be contained in a finite union of hyperplanes Π_{π_n}.

For $j = n, \ldots, 3$ for every $(j-1)$-dimensional plane π_{j-1} in Π_{π_j}, let $L_{\pi_{j-1}} : \mathbb{R}^{j-1} \to \mathbb{R}$ be the $(j-1)$-dimensional restriction of L to π_{j-1} and

$$V_{\pi_{j-1}} := \{v \in \mathbb{R}^{j-1} |\ L_{\pi_{j-1}}(\cdot) \text{ is not twice differentiable in } v\}.$$

Define

$$\Sigma_{\pi_{j-1}} := \{(t,x) \in \Sigma_{\pi_j} |\ d(t,x) \in V_{\pi_j}\} \quad \text{and} \quad \Sigma^c_{\pi_{j-1}} := \Sigma_{\pi_j} \setminus \Sigma_{\pi_{j-1}}.$$

(HYP(j-1)) We suppose $V_{\pi_{j-1}}$ is contained in a finite union of $(j-2)$-dimensional planes $\Pi_{\pi_{j-1}}$, for every $\pi_{j-1} \in \Pi_{\pi_j}$.

Theorem 0.5. *Under the above assumptions (HYP(0)),(HYP(n)),...,(HYP(2)), the Radon measure* $\mathrm{div} d(t,\cdot)$ *has Cantor part on* Ω_t *only for a countable number of* t*'s in* $[0,T]$.

The question on the SBV regularity of $d(t,\cdot)$ without any additional hypothesis to the convexity of H is still open.

The results presented in Chapter 4 are contained in Bianchini and Tonon [12].

iv) In Chapter 5 we present an application of Theorem 0.1. In the one-dimensional case the Generalized Hydrostatic Boussinnesq equations of Convection Theory and sticky particles systems can be both described at a discrete level by a finite collection of particles that get stuck together right after they collide. At a continuous level, instead, they can be related to a scalar conservation law of type (2 with non decreasing initial datum and bounded Lipschitz flux function. When $U_0(x) := x$ is chosen as initial datum we can reduce to an equivalent Hamilton-Jacobi equation where the flux function is $\frac{1}{2}|x|^2$ and the initial datum is H. Therefore the result of Ambrosio and De Lellis can be used to prove the SBV regularity of the entropy solution of the scalar conservation law with non decreasing initial datum $U_0(x) := x$ and bounded Lipschitz flux function.

Considering the multi-dimensional case it is reasonable to ask whether Hamilton-Jacobi equations are again related to the multidimensional version of the Generalized Hydrostatic Boussinnesq equations of Convection Theory and sticky particles systems. Theorem 0.2 could be applied even to this case if the answer was affirmative. However, in the multi-dimensional case, Hamilton-Jacobi equations are no more able to describe Convection Theory and sticky particles systems. Indeed, viscosity solutions of the multi-dimensional Hamilton-Jacobi equations can behave in a way which is not allowed in Convection Theory and by sticky particles systems. For viscosity solutions, particles that collide eventually separate while in Convection Theory and in sticky particles systems particles get stuck together after a collision. We present a counterexample which shows precisely this behavior. A first counterexample was found by Vasseur in his PhD Thesis [38] but it was never published.

The results presented in Chapter 5 are contained in Tonon [37].

0.2 Decomposition for BV functions

One of the necessary and sufficient properties, which characterizes real valued BV functions of one variable, is the well-known Jordan decomposition: it states that a function $f : \mathbb{R} \to \mathbb{R}$ is of bounded variation if and only if it can be written as the difference of two monotone increasing functions.

In Chapter 6, we present a generalization of this property to real valued BV functions of many variables.

To this aim, we define a new concept of monotonicity for a real valued function of many variables. Many different definitions of monotone function already exist in literature.

One can in fact preserve the monotonicity of the product $\langle f(x) - f(y), x - y\rangle \geq 0$, defining that $f : \mathbb{R}^n \to \mathbb{R}^n$ is monotone if

$$\langle f(x) - f(y), x - y\rangle \geq 0,$$

where $\langle \cdot, \cdot \rangle$ is the scalar product in $\mathbb{R}^n$.

Another possibility is to preserve the maximum principle: the supremum (infimum) of f in every set is assumed at the boundary. Taken $\Omega \subset \mathbb{R}^n$, a Lebesgue monotone function is defined as a continuous function $f : \Omega \to \mathbb{R}$, which satisfies the maximum and minimum principles in every subdomain. Manfredi, in [33], and Hajlasz and Malý, in [31], give a weaker formulation. Here, a weakly monotone function is defined as a function $f : \Omega \to \mathbb{R}$ in the Sobolev space $W^{1,p}(\Omega)$, which satisfies the weak maximum and the weak minimum principles in every subdomain. A natural generalization is given in the case f is in the Sobolev space $W^{1,p}_{loc}(\Omega)$.

In our case we choose to define monotone a function whose sub-level and super-level sets are indecomposable and of finite perimeter for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}$.

This notion of monotonicity is no more a sufficient condition for a function to be of bounded variation, as it was in the Jordan decomposition. However, it allows a decomposition of BV functions in a countable sum of monotone functions.

Indeed, in the case of BV functions, sub-level and super-level sets are of finite perimeter for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}$. Moreover, sets with finite perimeter can be decomposed in the countable union of indecomposable sets, up to $\mathcal{H}^n$-negligible sets, see [3]. Therefore, playing with the indecomposable components of sub-level and super-level sets, one can decompose a BV function in the sum of monotone BV functions. In general more than one of such decompositions is possible.

The strategy above is the idea which lies in of the proof of the following theorem.

Theorem 0.6. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a $BV(\mathbb{R}^n)$ function. Then there exists a finite or countable family of monotone $BV(\mathbb{R}^n)$ functions $\{f_i\}_{i\in I}$, such that*

$$f = \sum_{i\in I} f_i \quad \textit{and} \quad |Df| = \sum_{i\in I} |Df_i|.$$

This result extends a theorem of Alberti, Bianchini and Crippa presented in Section 6.1 which proves a decomposition property for real valued Lipschitz functions of many variables.

Theorem 0.7 (Alberti, Bianchini and Crippa)**.** *Let f be a function in $\mathrm{Lip_c}(\mathbb{R}^n)$ with compact support. Then there exists a countable family $\{f_i\}_{i\in\mathbb{N}}$ of functions in $\mathrm{Lip_c}(\mathbb{R}^n)$ such that $f = \sum_i f_i$ and each f_i is monotone. Moreover there is a pairwise disjoint partition $\{\Omega_i\}_{i\in\mathbb{N}}$ of Borel sets of $\mathbb{R}^n$ such that ∇f_i is concentrated on Ω_i.*

In the case of a Lipschitz function monotonicity is given by the connectedness of level sets. Moreover the decomposition preserves the mutual singularity of every ∇f_i. This is no more true in the BV case. It can be found an example where the monotone functions given by the decomposition have distributional derivatives which in general are not mutually singular.

Theorem 0.6 is in a way optimal. We show with a counterexample that there is no hope for a further generalization of this decomposition to vector valued BV functions, apart from the case of a function $f : \mathbb{R} \to \mathbb{R}^m$ where the analysis is straightforward. Indeed, a Lipschitz function from $\mathbb{R}^2$ to $\mathbb{R}^2$ can be decomposed in a sum of monotone functions only if some of its level sets are of positive $\mathcal{H}^1$-measure. This is an additional property, which is clearly not shared by all the Lipschitz functions.

The results presented in Chapter 6 are contained in Bianchini and Tonon [11].

0.3 Notations

$\mathcal{H}^n$	n-dimensional Hausdorff measure
$\mathbb{R}^+$	set of all non negative real number
$[L^1(\mathbb{R}^n)]^m$	Lebesgue space of functions from $\mathbb{R}^n$ to $\mathbb{R}^m$
$L^1_{loc}(\mathbb{R}^n)$	space of functions from $\mathbb{R}^n$ to $\mathbb{R}$ which are locally $L^1(\mathbb{R}^n)$
$[\mathrm{Lip}_c(\mathbb{R}^n)]^m$	space of c-Lipschitz functions from $\mathbb{R}^n$ to $\mathbb{R}^m$
$[BV(\mathbb{R}^n)]^m$	space of bounded variation functions from $\mathbb{R}^n$ to $\mathbb{R}^m$
∇f	gradient of the Lipschitz function f
Df	distributional derivative of the BV function f
$\|Df\|$	total variation of the function f
$D_x u(t,x)$	spatial distributional derivative of the locally Lipschitz function $u(t,x)$
$\partial_t u(t,x)$	time distributional derivative of the locally Lipschitz function $u(t,x)$
$\mathrm{div} d(t,\cdot)$	spatial divergence of the vector field $d(t,\cdot)$
$H_p(t,x,p)$	derivative with respect to p
$H_{pp}(t,x,p)$	second derivative with respect to p
$H_{px}(t,x,p)$	derivative with respect to p and with respect to x
$\langle \cdot,\cdot \rangle$	scalar product in $\mathbb{R}^n$
$O(t)$	big O notation
E_t	$\{x \in \mathbb{R}^n \mid (t,x) \in E\}$ for $E \subset \mathbb{R}^+ \times \mathbb{R}^n$
E_x	$\{y \in \mathbb{R}^n \mid (x,y) \in E\}$ for $E \subset \mathbb{R}^{n+k}$, $x \in \mathbb{R}^k$
$P(E)$	perimeter of the set E
$\|x\|$	norm of the vector $x \in \mathbb{R}^n$
$\\|f\\|_V$	norm of a function in the space V
$\mathring{E}^M$	essential interior of the set E
$\overline{E}$	closure of the set E
χ_E	characteristic function of the set E
(mod $\mathcal{H}^n$)	up to $\mathcal{H}^N$-negligible sets
δ_x	Dirac measure
∂E	topological boundary of a set E
$d(\cdot,\cdot)$	Euclidean distance
$d_H(A,B)$	Hausdorff distance between the sets A and B

Chapter 1

Preliminaries

We list here some preliminary results which will be necessary in the following chapters. References with more detailed descriptions of the arguments treated can be found therein.

1.1 Rectifiable sets

We briefly introduce the concept of rectifiable sets, for a more comprehensive reference see [5].

Definition 1.1. An $\mathcal{H}^k$-measurable set $E \subset \mathbb{R}^n$ is *$\mathcal{H}^k$-rectifiable* if there exist countably many Lipschitz functions $f_i : \mathbb{R}^k \to \mathbb{R}^n$ such that

$$\mathcal{H}^k\left(E \setminus \bigcup_{i=0}^{\infty} f_i(\mathbb{R}^k)\right) = 0$$

and $\mathcal{H}^k(E) < \infty$.

1.2 Hausdorff distance and Hausdorff convergence of sets

Let $d(\cdot, \cdot)$ be the Euclidean distance in $\mathbb{R}^n$.

Definition 1.2. Given a set $A \subset \mathbb{R}^n$ and a point $x \in \mathbb{R}^n$ their distance is defined as

$$d(x, A) := \inf_{y \in A} d(x, y).$$

Definition 1.3. Given two sets $A, B \subset \mathbb{R}^n$ then their *Hausdorff distance* is defined as

$$d_H(A, B) := \max\left\{\sup_{x \in A} \inf_{y \in B} d(x, y), \sup_{y \in B} \inf_{x \in A} d(x, y)\right\},$$

or equivalently

$$d_H(A, B) = \inf\{\varepsilon > 0 |\ A \subset B_\varepsilon \text{ and } B \subset A_\varepsilon\},$$

where

$$E_\varepsilon := \cup_{x \in E} \{z \in \mathbb{R}^n |\ d(x, z) < \varepsilon\}$$

for any set $E \subset \mathbb{R}^n$.

Definition 1.4. A sequence of sets $E^\varepsilon \subset \mathbb{R}^n$ *converges in the Hausdorff sense* to $E \subset \mathbb{R}^n$ if

$$\lim_{\varepsilon \to 0} d_H(E^\varepsilon, E) = 0.$$

1.3 Generalized differentials

We recall the definition of generalized differential, see Cannarsa and Sinestrari [20] and Cannarsa and Soner [21].

In this section Ω will be an open subset of $\mathbb{R}^n$.

Definition 1.5. Let $u : \Omega \to \mathbb{R}$, for any $x \in \Omega$ the sets

$$D^- u(x) = \left\{ p \in \mathbb{R}^n | \ \liminf_{y \to x} \frac{u(y) - u(x) - \langle p, y - x \rangle}{|y - x|} \geq 0 \right\},$$

$$D^+ u(x) = \left\{ p \in \mathbb{R}^n | \ \limsup_{y \to x} \frac{u(y) - u(x) - \langle p, y - x \rangle}{|y - x|} \leq 0 \right\},$$

are called, respectively, the *subdifferential* and *superdifferential* of u at x.

$D^+ u$ and $D^- u$ are closed and convex sets, sometimes they can be empty.

Definition 1.6. Let $u : \Omega \to \mathbb{R}$ be locally Lipschitz. A vector $p \in \mathbb{R}^n$ is called a *reachable gradient* of u at $x \in \Omega$ if there exists a sequence $\{x_k\} \subset \Omega \setminus \{x\}$ such that u is differentiable at x_k for each $k \in \mathbb{N}$, and

$$\lim_{k \to \infty} x_k = x, \qquad \lim_{k \to \infty} Du(x_k) = p.$$

The set of all reachable gradients of u at x is denoted by $D^* u(x)$.

1.4 Decomposition of a Radon measure

Given an $[L^\infty(\mathbb{R}^n)]^n$ vector field $d(x)$ such that $\mathrm{div} d(x) =: \mu(x)$ is a Radon measure on $\mathbb{R}^n$, we can decompose μ into three mutually singular measures:

$$\mu = \mu_a + \mu_c + \mu_j.$$

μ_a is the absolutely continuous part with respect to the Lebesgue measure. μ_j is the singular part of the measure which is concentrated on a $\mathcal{H}^{n-1}$-rectifiable set. μ_c, the Cantor part, is the remaining part.

1.5 BV and SBV functions

A detailed description of the spaces BV and SBV can be found in Ambrosio, Fusco and Pallara [5], Chapters 3 and 4.

Definition 1.7. A function $u : \mathbb{R}^n \to \mathbb{R}$, which belongs to $L^1(\mathbb{R}^n)$, is said to be of bounded variation if its distributional derivative is representable as an $\mathbb{R}^n$-valued measure Du with finite total variation, i.e.

$$\int_{\mathbb{R}^n} u \mathrm{div} \phi dx = -\sum_{i=1}^{n} \int_{\mathbb{R}^n} \phi_i dD_i u \quad \forall \phi \in [C_c^1(R^n)]^n.$$

The total variation $|Du|$ of a BV function is defined as the total variation of the vector measure Du. A function $u : \mathbb{R}^n \to \mathbb{R}^k$ is said to be of bounded variation if every components $u_j : \mathbb{R}^n \to \mathbb{R}$ is of bounded variation for $j = 1, \dots, k$.

Given $u \in BV(\mathbb{R}^n)$, it is possible to decompose the distributional derivative of u into three mutually singular measures:

$$Du = D_a u + D_c u + D_j u.$$

$D_a u$ is the absolutely continuous part with respect to the Lebesgue measure. $D_J u$ is the part of the measure which is concentrated on the rectifiable $(n-1)$-dimensional set J, where the function u has "jump" discontinuities, thus for this reason it is called jump part. $D_c u$, the Cantor part, is the singular part which satisfies $D_c u(E) = 0$ for every Borel set E with $\mathcal{H}^{n-1}(E) < \infty$. If this part vanishes, i.e. $D_c u = 0$, we say that $u \in SBV(\mathbb{R}^n)$. When $u \in [BV(\mathbb{R}^n)]^k$ the distributional derivative Du is a matrix of Radon measure and the decomposition can be applied to every component of the matrix.

We recall here some properties of BV functions which will be useful later on.

Definition 1.8. Let u in $[L^1_{loc}(\mathbb{R}^n)]^k$, we say that u has an *approximate limit* at $x \in \mathbb{R}^n$ if there exists $z \in \mathbb{R}^k$ such that

$$\lim_{\rho \to 0} \fint_{B_\rho(x)} |u(y) - z| dy = 0.$$

The set S_u of points where this property does not hold is called the *approximate discontinuity set*. For any $x \in \mathbb{R}^n \setminus S_u$ the vector z is called approximate limit of u at x and is denoted by $\tilde{u}(x)$.

Proposition 1.9. *Let u and v belong to $[BV(\mathbb{R}^n)]^k$. Let*

$$L := \{x \in \mathbb{R}^n \setminus (S_u \cup S_v) | \ \tilde{u}(x) = \tilde{v}(x)\}.$$

Then Du and Dv are equal when restricted to L.

Proof. See Remark 3.93 in [5]. □

Proposition 1.10. *Let u belongs to $[BV(\mathbb{R}^n)]^k$. Then $D_c u$ vanishes on sets which are σ-finite with respect to $\mathcal{H}^{n-1}$ and on sets of the form $\tilde{u}^{-1}(E)$ with $E \subset \mathbb{R}^k$ and $\mathcal{H}^1(E) = 0$.*

Proof. See Proposition 3.92 in [5]. □

Proposition 1.11. *Let u belongs to $[BV(\mathbb{R}^n)]^k$. For $j = 1, \dots, n-1$ define the $(n-j)$-dimensional restriction $u_{x_1,\dots,x_j}(\cdot) : \mathbb{R}^{n-j} \to \mathbb{R}^k$ as $u_{x_1,\dots,x_j}(\hat{x}) = u(x_1, \dots, x_j, \hat{x})$ for fixed $x_1, \dots, x_j \in \mathbb{R}^j$. Then $u_{x_1,\dots,x_j}(\cdot)$ is $[BV(\mathbb{R}^{n-j})]^k$ for $\mathcal{H}^j$-a.e. $x_1, \dots, x_j$ in $\mathbb{R}^j$.*

Proof. This is a well known result. The proof in the case $j = n-1$ can be found in [5] Section 3.11, in the other cases is similar. □

1.6 Semiconcave functions

For a complete introduction to the theory of semiconcave functions we refer to Cannarsa and Sinestrari [20], Chapter 2 and 3 and Lions [32]. For our purpose we define semiconcave functions with a linear modulus of semiconcavity. In general this class is considered only as a particular subspace of the class of semiconcave functions with general semiconcavity modulus. The proofs of the following statements can be found in the mentioned references.

In this section Ω will be an open subset of $\mathbb{R}^n$.

Definition 1.12. We say that a function $u : \Omega \to \mathbb{R}$ is *semiconcave* and we denote with $SC(\Omega)$ the space of functions with such a property, if for a $C > 0$ and for any $x, z \in \Omega$ such that the segment $[x - z, x + z]$ is contained in Ω

$$u(x + z) + u(x - z) - 2u(x) \leq C|z|^2.$$

Proposition 1.13. *Let $u : \Omega \to \mathbb{R}$ belongs to $SC(\Omega)$ with semiconcavity constant $C \geq 0$. Then the function*

$$\tilde{u} : x \mapsto u(x) - \frac{C}{2}|x|^2$$

is concave, i.e. for any x, y in Ω such that the whole segment $[x, y]$ is contained in Ω, $\lambda \in [0, 1]$

$$\tilde{u}(\lambda x + (1 - \lambda)y) \geq \lambda\tilde{u}(x) + (1 - \lambda)\tilde{u}(y).$$

Theorem 1.14. *Let $u : \Omega \to \mathbb{R}$ belongs to $SC(\Omega)$. Then the following properties hold.*

i) (Alexandroff's Theorem) u is twice differentiable $\mathcal{H}^n$-a.e.; that is, for $\mathcal{H}^n$-a.e. $x_0 \in \Omega$, there exist a vector $p \in \mathbb{R}^n$ and a symmetric matrix M such that

$$\lim_{x \to x_0} \frac{u(x) - u(x_0) - \langle p, x - x_0\rangle + \langle M(x - x_0), x - x_0\rangle}{|x - x_0|^2} = 0.$$

ii) The gradient of u, defined $\mathcal{H}^n$-a.e. in Ω, belongs to the class $BV_{loc}(\Omega, \mathbb{R}^n)$.

iii) Let $x \in \Omega$ then

$$D^+u(x) = \mathrm{co}D^*u(x),$$

where $\mathrm{co}A := \min\{B \mid B \supset A, B \text{ convex}\}$ is the convex hull of A. Thus D^+u is non empty at each point. Moreover D^+u is upper semicontinuous.

iv) The function $T(x) := -D^+\tilde{u}(x)$ is a maximal monotone function, i.e.

$$\langle y_1 - y_2, x_1 - x_2\rangle \geq 0 \qquad \forall x_i \in \Omega \; y_i \in T(x_i) \; i = 1, 2;$$

and it is maximal in following sense

$$V \supset T, \; V \text{ monotone} \implies V = T.$$

As stated in the above theorem at point ii), when u is semiconcave Du is a BV map, hence the distributional Hessian D^2u is a symmetric matrix of Radon measures and can be split into the three mutually singular parts $D^2_a u, D^2_j u, D^2_c u$. Moreover the following proposition holds.

Proposition 1.15. *Let u be a semiconcave function. If D denotes the set of points where D^+u is not single-valued, then $|D^2_c u|(D) = 0$.*

Proof. Indeed, the set of points where D^+u is not single-valued, i.e. the set of singular points, is a $\mathcal{H}^{n-1}$-rectifiable set. ☐

Definition 1.16. We say that a function $v : \Omega \to \mathbb{R}$ is *semiconvex* if $u := -v$ is semiconcave.

1.7 Viscosity solutions

A concept of generalized solutions to the equations

$$\partial_t u + H(t, x, D_x u) = 0 \qquad \text{in } \Omega \subset [0, T] \times \mathbb{R}^n, \tag{1.1}$$

and

$$H(x, Du) = 0 \qquad \text{in } \ \Omega \subset \mathbb{R}^n, \tag{1.2}$$

was found to be necessary since classical solutions break down and solutions which satisfy (1.1) almost everywhere are not unique. Crandall and Lions introduced in [23] the notion of viscosity solution to solve both these problems, see also Crandall, Evans and Lions [22]. A viscosity solution needs not be differentiable anywhere, the only regularity required in the definition is uniform continuity. This concept ensures existence, stability and uniqueness of solutions for a wider class of equations.

Definition 1.17. A bounded uniformly continuous function $u : \Omega \to \mathbb{R}$ is called a *viscosity solution* of (1.1) (resp. (1.2)) provided that

i) u is a viscosity subsolution of (1.1) (resp. (1.2)): for each $v \in C^\infty(\Omega)$ such that $u - v$ has a maximum at $(t_0, x_0) \in \Omega$ (resp. $x_0 \in \Omega$),

$$\partial_t v(t_0, x_0) + H(t_0, x_0, D_x v(t_0, x_0)) \leq 0 \qquad (\text{resp. } H(x_0, Dv(x_0)) \leq 0);$$

ii) u is a viscosity supersolution of (1.1) (resp. (1.2)): for each $v \in C^\infty(\Omega)$ such that $u - v$ has a minimum at $(t_0, x_0) \in \Omega$ (resp. $x_0 \in \Omega$),

$$\partial_t v(t_0, x_0) + H(t_0, x_0, D_x v(t_0, x_0)) \geq 0 \qquad (\text{resp. } H(x_0, Dv(x_0)) \geq 0).$$

1.8 Properties of viscosity solutions when $H = H(t, x, p)$

We will consider here only viscosity solutions of equation (1.1), similar results apply also to viscosity solutions of the Hamilton-Jacobi equation (1.2).

Let us introduce a locality property.

Proposition 1.18. *Let u be a viscosity solution of (1.1) in Ω, when the Hamiltonian H is convex in the last variable. Then u is locally Lipschitz. Moreover for any $(t_0, x_0) \in \Omega$, there exists a neighborhood $\mathcal{U}$ of (t_0, x_0), a positive number δ and a Lipschitz function v_0 on $\mathbb{R}^n$ such that*

(Loc) u coincides on $\mathcal{U}$ with the viscosity solution of

$$\begin{cases} \partial_t v + H(t, x, D_x v) = 0 & \text{in } [t_0 - \delta, \infty) \times \mathbb{R}^n \\ v(t_0 - \delta, x) = v_0(x). \end{cases}$$

Proof. The proof of Proposition 3.5, given in [9], still applies in our case. We only loose the property that minimizers of the representation formula for the viscosity solution (see later on 1.3) are straight lines which was unnecessary for the argument. Even the uniform convexity of H in the last variable was not necessary in the proof. □

Motivated by the above proposition, it is enough to consider the Cauchy problem

$$\begin{cases} \partial_t u + H(t, x, D_x u) = 0 & \text{in } \Omega \subset [0, T] \times \mathbb{R}^n, \\ u(0, x) = u_0(x) & \text{for all } x \in \Omega_0, \end{cases}$$

where $u_0(x)$ is a bounded Lipschitz function on $\Omega_0 := \{x \in \mathbb{R}^n |\ (0, x) \in \Omega\}$.

The proofs of the following statements can be found in Cannarsa and Sinestrari [20], Chapter 6. See also Fleming [28], Fleming and Rishel [29], Fleming and Soner [30] and Lions [32].

The convexity of the Hamiltonian in the p-variable relates Hamilton-Jacobi equations to a variational problem.

Let us require the following assumptions on H.

(H1) $H \in C^3([0, T] \times \mathbb{R}^n \times \mathbb{R}^n)$ with bounded second derivatives and there exist positive constants a, b, c such that

i) $H(t, x, p) \geq -c$,

ii) $H(t, x, 0) \leq c$,

iii) $|H_{px}(t, x, p)| \leq a + b|p|$,

(H2) H is uniformly convex in the last variable, to be more precise there exists $c_H > 0$ such that

$$c_H^{-1} Id_n(p) \leq H_{pp}(t, x, p) \leq c_H Id_n(p)$$

for any $t \in \mathbb{R}$, $x \in \mathbb{R}^n$.

Let L be the Lagrangian of our system, i.e. the Legendre transform of the Hamiltonian H with respect to the last variable, for any t, x fixed

$$L(t, x, v) = \sup_p \{\langle v, p \rangle - H(t, x, p)\}.$$

The Legendre transform inherits the properties of H, in particular L is $C^3([0, T] \times \mathbb{R}^n \times \mathbb{R}^n)$ and uniformly convex in the last variable.

In addition to the uniform convexity and C^3 regularity of L, the hypotheses on H, (H1) and (H2), ensure the existence of positive constants a, b, c such that

i) $L(t, x, v) \geq -c$,

ii) $L_x(t, x, 0) \leq c$,

iii) $|L_{vx}(t, x, v)| \leq a + b|v|$.

Define the value function $u(t, x)$ associated the bounded Lipschitz function $u_0(x)$, for $(t, x) \in \Omega$

$$u(t,x) := \min \left\{ u_0(\xi(0)) + \int_0^t L(s, \xi(s), \dot{\xi}(s))ds \;\middle|\; \xi(t) = x,\ \xi \in [C^2([0,t])]^n \right\}. \tag{1.3}$$

Less regularity can be asked to ξ, but it is unnecessary since any minimizing curve exists and is smooth, due to the regularity of L, see [20].

Theorem 1.19. *Taken a minimizing curve ξ in (1.3), for the point (t, x), such that $\xi(s) \in \Omega_s$ for all $s \in [0, t]$, the following holds.*

i) The map $s \mapsto L_v(s, \xi(s), \dot{\xi}(s))$ is absolutely continuous.

ii) ξ is a classical solution to the Euler-Lagrange equation

$$\frac{d}{ds} L_v(s, \xi(s), \dot{\xi}(s)) = L_x(s, \xi(s), \dot{\xi}(s)),$$

and to the Du Bois-Reymond equation

$$\frac{d}{ds}[L(s, \xi(s), \dot{\xi}(s)) - \langle \dot{\xi}(s), L_v(s, \xi(s), \dot{\xi}(s)) \rangle] = L_t(s, \xi(s), \dot{\xi}(s)),$$

for all $s \in [0, t]$, where $L_t(s, \xi(s), \dot{\xi}(s)), L_x(s, \xi(s), \dot{\xi}(s)), L_v(s, \xi(s), \dot{\xi}(s))$ are the derivatives of L with respect to t, x, v respectively.

iii) For any $r > 0$ there exists $K(r) > 0$ such that, if $(t, x) \in [0, r] \times B_r(0)$, then

$$\sup_{s \in [0,t]} |\dot{\xi}(s)| \leq K(r).$$

iv) There exists a dual arc or co-state

$$p(s) := L_v(s, \xi(s), \dot{\xi}(s)) \qquad s \in [0, t], \tag{1.4}$$

such that ξ, p solve the following system

$$\begin{cases} \dot{\xi}(s) = H_p(s, \xi(s), p(s)) \\ \dot{p}(s) = -H_x(s, \xi(s), p(s)). \end{cases}$$

v) $(s, \xi(s))$ is regular, i.e. for any $0 < s < t$, ξ is the unique minimizer for $u(s, \xi(s))$, and $u(s, \cdot)$ is differentiable at $\xi(s)$.

vi) Let p be the dual arc associated to ξ as in (1.4) then we have

$$p(t) \in D_x^+ u(t,x),$$

$$p(s) = D_x u(s, \xi(s)), \qquad s \in (0,t).$$

Theorem 1.20. *The value function u defined in (1.3) is a viscosity solution of (1.1) with bounded Lipschitz initial datum*

$$u(0,x) = u_0(x).$$

Definition 1.21. A point $(t,x) \in \Omega_t$ is called *regular* if there exists a unique minimizer for $u(t,x)$. All the other points are called *irregular.*

A point $(t,x) \in \Omega_t$ is called *conjugate* if $z \in \mathbb{R}^n$ exists such that $\xi(t,z) = x$, $\xi(\cdot, z)$ is a minimizer for $u(t,x)$ and

$$\det \xi_z(t,z) = 0.$$

We present below some properties of the unique viscosity solution to the Hamilton-Jacobi equation (1.1), which follow from the representation formula we have just seen. These properties are taken from [20].

Theorem 1.22 (Dynamic Programming Principle)**.** *Fix (t,x), then for all $t' \in [0,t]$*

$$u(t,x) := \min \left\{ u(t', \xi(t')) + \int_{t'}^{t} L(s, \xi(s), \dot{\xi}(s)) ds \;\middle|\; \xi(t) = x,\ \xi \in [C^2([t',t])]^n \right\}. \tag{1.5}$$

Moreover if ξ is a minimizer in (1.3) it is a minimizer also for (1.5) for any $t' \in [0,t]$.

Theorem 1.23. *Suppose* (H1), (H2) *hold and u_0 belongs to $C_b(\mathbb{R}^n)$. Then for any t in $(0,T]$, $u(t,\cdot)$ is locally semiconcave with semiconcavity constant $C(t) = \frac{C}{t}$. Thus for any fixed $\tau > 0$ there exists a constant $C = C(\tau)$ such that $u(t,\cdot)$ is semiconcave with constant less than C for any $t \geq \tau$.*

Moreover u is also locally semiconcave in both the variables (t,x) in $(0,T] \times \mathbb{R}^n$.

1.8.1 Minimizers and Generalized Backward Characteristics

We introduce the definition of generalized backward characteristics.

Definition 1.24. Given $x \in \Omega_t$ for t fixed in $[0,T]$, we call *generalized backward characteristic*, associated to u starting from x, the curve $s \mapsto (s, \xi(s))$, where $\xi(\cdot)$ and its dual arc $p(\cdot)$ solve the system

$$\begin{cases} \dot{\xi}(s) = H_p(s, \xi(s), p(s)) \\ \dot{p}(s) = -H_x(s, \xi(s), p(s)) \end{cases} \tag{1.6}$$

with final conditions

$$\begin{cases} \xi(t) = x \\ p(t) = p, \end{cases} \tag{1.7}$$

where $p \in D_x^+ u(t,x)$.

If $D_x^+ u(t,x)$ is single-valued then we call ξ a *classical backward characteristic.*

We state here some properties of minimizers which strictly relate them with classical and generalized characteristics, see [20].

Theorem 1.25. *For any $(t, x) \in \Omega$ the map that associates with any $(p_t, p_x) \in D^*u(t, x)$ the curve ξ obtained by solving the system (1.6) with the final conditions*

$$\begin{cases} \xi(t) = x \\ p(t) = p_x \end{cases}$$

provides a one-to-one correspondence between $D^*u(t, x)$ and the set of minimizers of $u(t, x)$.

Thus we can state the following theorem which follows from Theorem 1.19-(iv), Theorem 1.25 and Definition 1.24.

Theorem 1.26. *Let (t, x) in Ω be given, and let ξ be a C^2 curve such that $\xi(s) \in \Omega_s$ for all $0 \le s \le t$.*

Then ξ is a minimizer if and only if ξ and its dual arc p are solutions of the system (1.6) for any $s \in [0, t]$ with final conditions (1.7), where $(-H(t, x, p), p)$ belongs to $D^*u(t, x)$.

A minimizer ξ is a generalized backward characteristic. In particular ξ is a classical backward characteristic if and only if ξ is the unique minimizer for $u(t, x)$. The set of minimizers for $u(t, x)$ is a proper *subset of the set of generalized backward characteristics emanated from (t, x).*

Remark 1.27. Note that, the solutions ξ of the system (1.6) are in general curves and not straight lines, as in the case $H = H(p)$.

Remark 1.28. *No-crossing property of minimizers.* Fix a time t and consider a minimizing curve ξ such that $\xi(t) = x \in \Omega_t$. For $0 < s < t$ the curve ξ is the unique minimizer for $u(s, \xi(s))$, this ensures that any other minimizer cannot intersect ξ for any $0 < s < t$ (otherwise uniqueness would be lost, see point (v) of Theorem 1.19). As a consequence generalized backward characteristics which are also minimizers, i.e. solution of (1.6), (1.7), where $(-H(t, x, p), p)$ belongs to $D^*u(t, x)$, cannot intersect except than in 0 or t. Nothing can be said at this level for generalized backward characteristics solution to (1.6) with

$$\xi(t) = x \qquad p(t) = p \in D_x^+ u(t, x) \setminus D_x^* u(t, x),$$

which are not minimizers. In general they can cross.

1.8.2 Backward solutions

The introduction of a *backward solution*, as in Barron, Cannarsa, Jensen and Sinestrari [7], will allow us to see in Section 2.1 that, at least for a small interval of time, all the generalized backward characteristics share the no-crossing property.

Fix t in $(0, T]$ and define for $0 \le \tau < t$, $y \in \Omega_\tau$ the function

$$u_{t,0}^-(\tau, y) := \max \left\{ u(t, \xi(t)) - \int_\tau^t L(s, \xi(s), \dot{\xi}(s)) ds \;\middle|\; \xi(\tau) = y,\ \xi \in [C^2([\tau, t])]^n \right\}. \tag{1.8}$$

Note that the function $v(\tau, y) := u_{t,0}^-(t - \tau, y)$ is a viscosity solution of

$$\partial_\tau v - H(t - \tau, y, D_y v) = 0 \qquad \text{in } \Omega \subset [0, T] \times \mathbb{R}^n$$

with initial datum $v(0, y) = u_{t,0}^-(t, y) = u(t, y)$, for this reason $u_{t,0}^-$ is called backward solution.

Proposition 1.29. *In general*

$$u^-_{t,0}(\tau, y) \leq u(\tau, y)$$

and the equality holds if and only if the maximizer ξ in (1.8), defined for $\tau \leq s \leq t$, is part of a minimizing curve for $u(t, \xi(t))$.

Proof. Let ξ be a C^2-curve which is a maximizer for $u^-_{t,0}(\tau, y)$, i.e.

$$u^-_{t,0}(\tau, y) = u(t, \xi(t)) - \int_\tau^t L(s, \xi(s), \dot{\xi}(s))ds.$$

Thanks to the Dynamic Programming Principle,

$$u(t, \xi(t)) \leq u(\tau, y) + \int_\tau^t L(s, \xi(s), \dot{\xi}(s))ds.$$

Hence,

$$u^-_{t,0}(\tau, y) \leq u(\tau, y)$$

and the equality holds if and only if ξ is also a minimizer for $u(t, \xi(t))$, thus $D^+_x u(s, \xi(s))$ is single-valued for any $\tau \leq s < t$. ☐

Note that a curve ξ which is a minimizer for $u(t, x)$ is also a maximizer for $u^-_{t,0}(\tau, \xi(\tau)) = u(\tau, \xi(\tau))$ for any $0 \leq \tau < t$.

With suitable modifications Theorems 1.19, 1.20, 1.22 and 1.23 still hold for $u^-_{t,0}(\tau, y)$ and its maximizers, in particular $u^-_{t,0}$ is semiconvex (rather than semiconcave) with constant $\frac{C}{t-\tau}$.

Without adding any other assumption, the no-crossing property holds also for maximizers.

1.9 Properties of viscosity solutions when $H = H(p)$

We present here the case in which $H = H(p)$ is smooth and convex. Some of the results are just a particular case of the results obtained in the case $H = H(t, x, p)$.

As already noticed in the case of $H = (t, x, p)$ it is enough to consider the Cauchy problem

$$\begin{cases} \partial_t u + H(D_x u) = 0 & \text{in } \Omega \subset [0, T] \times \mathbb{R}^n, \\ u(0, x) = u_0(x) & \text{for all } x \in \Omega_0, \end{cases} \tag{1.9}$$

where $u_0(x)$ is a bounded Lipschitz function on Ω_0.

The proofs of the following statements can be found in Evans [26], Section 3.3 and Chapter 10. See also Cannarsa and Sinestrari [20], Fleming [28], Fleming and Rishel [29], Fleming and Soner [30] and Lions [32].

When $H = H(p)$ the Lagrangian of our system can be obtained as

$$L(v) = \sup_p \{\langle v, p \rangle - H(p)\}.$$

In the case of a smooth convex Hamiltonian the corresponding Lagrangian is strictly convex but non smooth in general.

Theorem 1.30. *The unique viscosity solution of the Cauchy problem (1.9) is the Lipschitz continuous function $u(t,x)$ defined for $(t,x) \in \Omega$ as*

$$u(t,x) = \min_{y \in \Omega_0} \left\{ u(0,y) + tL\left(\frac{x-y}{t}\right) \right\}. \tag{1.10}$$

Theorem 1.31. *Let $u(t,x)$ be a viscosity solution of the Cauchy problem (1.9).*

i) The minimum point y for $(t,x) \in \Omega$ in (1.10) is unique if and only if $u(t,x)$ is differentiable in x. Moreover in this case $y = x - tH_p(D_x u(t,x))$.

ii) (Dynamic Programming Principle) Fix $(t,x) \in \Omega$, then for all $t' \in [0,t]$

$$u(t,x) = \min_{z \in \Omega_{t'}} \left\{ u(t',z) + (t-t')L\left(\frac{x-z}{t-t'}\right) \right\}.$$

iii) Let $0 < s < t$, let $(t,x) \in \Omega$ and y a minimum point in (1.10). Let $z = \frac{s}{t}x + \left(1 - \frac{s}{t}\right)y$. Then y is the unique minimum point for

$$u(s,z) = \min_{w \in \Omega_0} \left\{ u(0,y) + sL\left(\frac{z-w}{s}\right) \right\}.$$

Remark 1.32. Note that in this case characteristics are straight lines, for this reason will be called *rays*.

Definition 1.33. Let $y \in \Omega_0$ be a minimizer for $u(t,x)$. We call *optimal ray* the segment $[x,y]$ defined in $[0,t]$.

Proposition 1.34. *Let $[x,y]$ and $[x',y']$ be two optimal rays in $[0,t]$, for $x, x' \in \Omega_t$ $y, y' \in \Omega_0$ then they cannot intersect except than at time 0 or t.*

Proof. It follows from Theorem 1.31-(iii). □

Proposition 1.35. *Let u_0 be a semiconcave function. Then the unique viscosity solution $u(t,x)$ of (1.9) is semiconcave in x, for all $t \in [0,T]$.*

Proof. See Lemma 3 in Section 3.3 [26]. □

Theorem 1.36. *Suppose H is locally uniformly convex. Then for any t in $(0,T]$, $u(t,\cdot)$ is locally semiconcave with semiconcavity constant $C(t) = \frac{C}{t}$. Thus for any fixed $\tau > 0$ there exists a constant $C = C(\tau)$ such that $u(t,\cdot)$ is semiconcave with constant less than C for any $t \geq \tau$.*

Proof. See Lemma 4 in Section 3.3 [26]. □

Moreover u is also locally semiconcave in both the variables (t,x) in $(0,T] \times \mathbb{R}^n$.

1.9.1 Duality solutions

We consider a fixed interval of time $[0,1]$, and we define duality solutions in this time interval.

Definition 1.37. Setting $u^+(1,z) := u(1,z)$, we define *duality solutions* for $s \in [0,1]$ and $z \in \Omega_s$, the *backward solution*

$$u^-(s,z) := \max_{x\in\Omega_1} \left\{ u^+(1,x) - (1-s)L\left(\frac{x-z}{1-s}\right) \right\}, \tag{1.11}$$

and the *forward solution*

$$u^+(s,z) := \min_{y\in\Omega_0} \left\{ u^-(0,y) + sL\left(\frac{z-y}{s}\right) \right\}. \tag{1.12}$$

Remark 1.38. Note that the function $v(\tau,y) := u^-(1-\tau,y)$ is a viscosity solution of

$$\begin{cases} \partial_\tau v - H(D_y v) = 0 & \text{in } \Omega \subset [0,1]\times\mathbb{R}^n \\ v(0,y) = u(1,y) & \text{for all } y \in \Omega_1. \end{cases}$$

Moreover the forward solution is the viscosity solution of

$$\begin{cases} \partial_t u + H(D_x u) = 0 & \text{in } \Omega \subset [0,1]\times\mathbb{R}^n, \\ u(0,x) = u^-(0,x) & \text{for all } x \in \Omega_0. \end{cases}$$

Thanks to the previous remark Theorems 1.31, 1.36 and Propositions 1.34, 1.35 hold for v and the forward solution u^+.

Proposition 1.39. *From the definitions above, u^+ and u^- satisfy the following properties for $x \in \Omega_1$, $y \in \Omega_0$ and $z \in \Omega_s$ for $s \in (0,1)$*

$$u^-(1,x) = u^+(1,x) = u(1,x), \quad u^+(0,y) = u^-(0,y) \le u(0,y), \quad u^-(s,z) \le u^+(s,z) \le u(s,z).$$

Proof. The first two equalities are a consequence of the fact that u^0 and u^1, defined as follows, are $L(x-y)$ conjugate functions. First, for $x \in \Omega_1$, set

$$u^1(x) := \min_{y\in\Omega_0}\{u(0,y) + L(x-y)\},$$

i.e. $u^1(x) = u(1,x)$.
Then, for $y \in \Omega_0$, set

$$u^0(y) := \max_{x\in\Omega_1}\{u^1(x) - L(x-y)\},$$

i.e. $u^0(y) = u^-(0,y)$.

From these definitions it follows $u^0(y) \le u(0,y)$ and

$$u^1(x) = \min_{y\in\Omega_0}\{u^0(y) + L(x-y)\}.$$

Indeed, let $\tilde{x} \in \Omega_1$ a maximizer for $u^0(y)$ then

$$u^0(y) = u^1(\tilde{x}) - L(\tilde{x}-y) \le u(0,y) + L(\tilde{x}-y) - L(\tilde{x}-y) = u(0,y).$$

Nevertheless, from $u^0(y) \le u(0,y)$, it follows

$$\min_{y\in\Omega_0}\{u^0(y) + L(x-y)\} \le \min_{y\in\Omega_0}\{u(0,y) + L(x-y)\} = u^1(x).$$

On the other hand, let $\tilde{y}$ be a minimizer for $\min_{y\in\Omega_0}\{u^0(y) + L(x-y)\}$, then we have

$$\begin{aligned}\min_{y\in\Omega_0}\{u^0(y) + L(x-y)\} &= u^0(\tilde{y}) + L(x-\tilde{y})\\ &\ge u^1(x) - L(x-\tilde{y}) + L(x-\tilde{y})\\ &= u^1(x).\end{aligned}$$

Note that the definition of $u^-(s,z)$ and $u^+(s,z)$ implies that $u^-(1,x) = u^1(x)$ and $u^+(0,y) = u^0(y)$.

The last inequality follows, for s in $(0,1)$, by

$$\begin{aligned}u^-(s,z) =& \max_{x\in\Omega_1}\left\{u^1(x) - (1-s)L\left(\frac{x-z}{1-s}\right)\right\}\\ =& \max_{x\in\Omega_1}\left\{\min_{y\in\Omega_0}\left\{u^0(y) + L(x-y) - (1-s)L\left(\frac{x-z}{1-s}\right)\right\}\right\}\\ \le& \min_{y\in\Omega_0}\left\{u^0(y) + sL\left(\frac{z-y}{s}\right)\right\} = u^+(s,z),\end{aligned}$$

where the inequality is given by the convexity of L

$$L(x-y) \le sL\left(\frac{z-y}{s}\right) + (1-s)L\left(\frac{x-z}{1-s}\right).$$

Note that, from the strict convexity of L, the equality holds if and only if $\frac{x-z}{1-s} = \frac{z-y}{s}$, i.e. $z = sx + (1-s)y$ that is z belongs to the segment joining the maximizer x to the minimizer y.

Furthermore, due to the fact that $u^-(0,y) \le u(0,y)$, we have $u^+(s,z) \le u(s,z)$. $\square$

Proposition 1.40. *Suppose H is a smooth uniformly convex Hamiltonian. Then a $C^{1,1}$-estimate holds in the regions where $u^-(s,z) = u^+(s,z)$, for $s \in (0,1)$.*

Proof. Fix s in $(0,1)$ and z such that $u^-(s,z) = u^+(s,z)$, then as observed in the previous proof there is a unique segment, connecting the unique minimizer $y(z)$ in (1.12) to the unique maximizer $x(z)$ in (1.11) and passing through z. Hence $z = (1-s)y(z) + sx(z)$. Moreover both $u^+(s,\cdot)$ and $u^-(s,\cdot)$ are differentiable in z since the minimizer and the maximizer are unique. Note that neither $u^-(s,z) = u^+(s,z)$ implies necessarily that $u^-(s,z) = u^+(s,z) = u(s,z)$, nor $u^+(s,z) = u(s,z)$ implies that $u^-(s,z) = u^+(s,z) = u(s,z)$. However, if for a $\tilde{z}$ $u^-(s,\tilde{z}) = u(s,\tilde{z})$ then $u^-(s,\tilde{z}) = u^+(s,\tilde{z}) = u(s,\tilde{z})$.

From the definition of u^+ and u^- for $z' \in \Omega_t$

$$u^1(x(z)) - (1-s)L\left(\frac{x(z)-z'}{1-s}\right) \le u^-(s,z') \le u^+(s,z') \le u^0(y(z)) + sL\left(\frac{z'-y(z)}{s}\right).$$

Since $z = (1-s)y(z) + sx(z)$ and

$$u^-(s,z) = u^1(x(z)) - (1-s)L\left(\frac{x(z)-z}{1-s}\right) = u^0(y(z)) + sL\left(\frac{z-y(z)}{s}\right) = u^+(s,z),$$

we obtain

$$-(1-s)\left(L\left(x(z)-y(z)-\frac{z'-z}{1-s}\right) - L\big(x(z)-y(z)\big)\right) \le u^-(s,z') - u^-(s,z)$$
$$\le u^+(s,z') - u^+(s,z) \le s\left(L\left(x(z)-y(z)+\frac{z'-z}{s}\right) - L\big(x(z)-y(z)\big)\right).$$

In particular, recalling the fact that both $u^+(s,\cdot)$ and $u^-(s,\cdot)$ are differentiable in z, and that L is C^1

$$D_x u^-(s,z) = D_x u^+(s,z) = L_v\big(x(z)-y(z)\big).$$

Moreover, thanks to the fact that we are considering the region where $u^+(s,z) = u^-(s,z)$, they are both semiconvex and semiconcave in this region, thus we can recover a Lipschitz estimate for $D_x u^+$ and $D_x u^-$.

$$-\frac{C}{1-s}|z|^2 \le u^-(s,x+z)+u^-(s,x-z)-u^-(s,x) = u^+(s,x+z)+u^+(s,x-z)-u^+(s,x) \le \frac{C}{s}|z|^2.$$

Hence we have proved that in the region where $u^- = u^+$ the dual solutions are $C^{1,1}$. □

Remark 1.41. In the proof of the above proposition we used the semiconcavity of $u^+(s,\cdot)$ and the semiconvexity of $u^-(s,\cdot)$ thus the hypothesis of uniform convexity of the Hamiltonian is necessary.

The definition of backward and forward solutions can be easily generalized for every time interval $[\tau,t] \subset [0,T]$. Propositions 1.39 and 1.40 hold even in this case.

Definition 1.42. Setting $u^+_{t,\tau}(t,z) := u(t,z)$, we define *duality solutions* for $s \in [\tau,t]$ and $z \in \Omega_s$, the *backward solution*

$$u^-_{t,\tau}(s,z) := \max_{x\in\Omega_t}\left\{u^+_{t,\tau}(t,x) - (t-s)L\left(\frac{x-z}{t-s}\right)\right\},$$

and the *forward solution*

$$u^+_{t,\tau}(s,z) := \min_{y\in\Omega_\tau}\left\{u^-_{t,\tau}(\tau,y) + (s-\tau)L\left(\frac{z-y}{s-\tau}\right)\right\}.$$

Chapter 2

SBV Regularity for Hamilton-Jacobi equations

In this chapter we study the regularity of the viscosity solution of the following Hamilton-Jacobi equation

$$\partial_t u + H(t, x, D_x u) = 0 \qquad \text{in } \Omega \subset [0, T] \times \mathbb{R}^n. \tag{2.1}$$

Under the following assumptions on H

(H1) $H \in C^3([0,T]\times\mathbb{R}^n\times\mathbb{R}^n)$ with bounded second derivatives and there exist positive constants a, b, c such that

  i) $H(t, x, p) \geq -c$,

  ii) $H(t, x, 0) \leq c$,

  iii) $|H_{px}(t, x, p)| \leq a + b|p|$,

(H2) there exists $c_H > 0$ such that

$$c_H^{-1} Id_n(p) \leq H_{pp}(t, x, p) \leq c_H Id_n(p)$$

for any t, x,

we will prove the following result.

Theorem 2.1. *Let u be a viscosity solution of (2.1), assume* (H1), (H2). *Then the set of times*

$$S := \{t \mid D_x u(t, \cdot) \notin [SBV_{loc}(\Omega_t)]^n\}$$

is at most countable. In particular $D_x u, \partial_t u \in [SBV_{loc}(\Omega)]^n$.

Moreover, under the hypotheses

(H1-bis) $H \in C^3(\mathbb{R}^n \times \mathbb{R}^n)$ with bounded second derivatives and there exist positive constants a, b, c such that

  i) $H(x, p) \geq -c$,

ii) $H(x,0) \le c$,

iii) $|H_{px}(x,p)| \le a + b|p|$,

(H2-bis) there exists $c_H > 0$ such that

$$c_H^{-1} Id_n(p) \le H_{pp}(x,p) \le c_H Id_n(p)$$

for any x,

as a consequence of the theorem above, the following corollary holds.

Corollary 2.2. *Under assumptions* (H1 − bis), (H2 − bis)*, the gradient of any viscosity solution u of*

$$H(x, Du) = 0 \qquad \text{in } \ \Omega \subset \mathbb{R}^n,$$

belongs to $[SBV_{loc}(\Omega)]^n$.

These results confirm and extend the SBV-regularizing effect of Hamilton-Jacobi equations which was already proved by Cannarsa, Mennucci and Sinestrari in [19] for strictly convex Hamiltonians and regular initial data and by Bianchini, De Lellis and Robyr in [9] for uniformly convex Hamiltonians depending only on $D_x u$.

The Lipschitzianity of the viscosity solution u in x allows us to conjecture that the SBV-regularizing effect is preserved even in the case of a Hamiltonian which depends also on the solution u,

$$H(x, u, Du) = 0 \qquad \text{in } \ \Omega \subset \mathbb{R}^n.$$

As we will see, the study of generalized characteristics, which are in general C^2-curve, and their approximations with straight lines in small intervals of time, allow us to prove the main result using a strategy which is standard in showing the SBV-regularizing effect.

The chapter is organized as follows. In Section 2.1 we prove the no-crossing property of generalized backward characteristics. Finally in Section 2.2 we prove all the necessary lemmas and Theorem 2.1.

2.1 No-crossing property for a small interval of time

We begin with the study of generalized backward characteristics, defined as in Definition 1.24. Let $t \in [0,T]$ be fixed, if we restrict to a $\tau > 0$ which is not too far from t, we can establish a one to one correspondence between generalized backward characteristics and maximizers of (1.8). Thus we are able to recover regularity and the no-crossing property for generalized backward characteristics. Moreover the backward solution $u^-_{t,0}(s,\cdot)$, defined in (1.8), belongs to $C^{1,1}(\Omega_s)$ for every $s \in (\tau, t)$, while in $[0,\tau]$ it can loose its regularity.

To prove the above fact let us first reduce to a simpler case which will be useful also later on during the proof of Theorem 2.1.

Lemma 2.3. *Consider the solutions to the system*

$$\begin{cases} \dot{\xi}(s) = H_p(s, \xi(s), p(s)) \\ \dot{p}(s) = -H_x(s, \xi(s), p(s)) \end{cases} \tag{2.2}$$

with final conditions

$$\begin{cases} \xi(t) = x \\ p(t) = p \end{cases} \tag{2.3}$$

where x is fixed in $\mathbb{R}^n$ and $p \in K$ a compact set in $\mathbb{R}^n$. For $t - \tau$ small enough there exists a one to one correspondence between p in K and $\xi(\tau)$ when $\xi(\cdot)$ is a solution of (2.2),(2.3).

Proof. Thanks to the Taylor expansion of the flow generated by (2.2), the solution to that system, with (2.3) as final conditions, is equal to

$$\xi(\tau) = x - (t-\tau)H_p(t,x,p) + O((t-\tau)^2),$$

and differentiating in p

$$\xi_p(\tau) = -(t-\tau)H_{pp}(t,x,p) + O((t-\tau)^2). \tag{2.4}$$

Note that ξ_p and p_p satisfy

$$\begin{cases} \dot{\xi}_p(s) = H_{px}(s,\xi(s),p(s))\xi_p(s) + H_{pp}(s,\xi(s),p(s))p_p(s) \\ \dot{p}_p(s) = -H_{xx}(s,\xi(s),p(s))\xi_p(s) - H_{xp}(s,\xi(s),p(s))p_p(s) \end{cases}$$

with final conditions

$$\begin{cases} \xi_p(t) = 0 \\ p_p(t) = Id_n(p). \end{cases}$$

Since they are smooth, equation (2.4) is precisely the Taylor expansion of $\xi_p(\tau)$.

Call $\omega := \frac{x-\xi(\tau)}{t-\tau}$. Last equation implies that ω_p is uniformly different from zero since

$$\omega_p = H_{pp}(t,x,p) + O(t-\tau).$$

Thus, restricting to $t-\tau$ small enough, we can locally invert this equation and obtain

$$p_\omega = L_{vv}\,(t,x,\omega) + O(t-\tau). \tag{2.5}$$

Moreover, from

$$\omega = H_p(t,x,p) + O(t-\tau),$$

integrating (2.5), we obtain

$$p = L_v\,(t,x,\omega) + O(t-\tau).$$

Thus we have reached a one to one correspondence between $\xi(\tau)$ and the value p of its dual curve at time t. $\square$

Integrating (2.4) in p between p_1 and p_2 we obtain

$$\frac{\xi_1(\tau) - \xi_2(\tau)}{\tau - t} = H_p(t,x,p_1) - H_p(t,x,p_2) + O(t-\tau)(p_1 - p_2)$$

where ξ_1 and ξ_2 are the generalized backward characteristics with initial data p_1 and p_2 respectively.

Proposition 2.4. *Consider a solution ξ to the system (2.2) with final conditions (2.3), let $y := \xi(\tau)$ and consider the straight line joining x to y*

$$\eta(s) = \frac{s-\tau}{t-\tau}x + \frac{t-s}{t-\tau}y. \tag{2.6}$$

Then we have the following estimates

$$\|\eta - \xi\|_{[C^0([\tau,t])]^n}, \|\eta_p - \xi_p\|_{[C^0([\tau,t])]^{n^2}}, \|\eta_{pp} - \xi_{pp}\|_{[C^0([\tau,t])]^{n^3}} \leq O((t-\tau)^2),$$

$$\|\dot\eta - \dot\xi\|_{[C^0([\tau,t])]^n}, \|\dot\eta_p - \dot\xi_p\|_{[C^0([\tau,t])]^{n^2}}, \|\dot\eta_{pp} - \dot\xi_{pp}\|_{[C^0([\tau,t])]^{n^3}} \leq O(t-\tau).$$

Proof. As we saw in the previous proposition

$$y = \xi(\tau) = x - (t-\tau)H_p(t,x,p) + O((t-\tau)^2),$$

and for $s \in [\tau, t]$

$$\xi(s) = x - (t-s)H_p(t,x,p) + O((t-s)^2).$$

Compute now the difference

$$\begin{aligned}
\sup_{s\in[\tau,t]} |\eta(s) - \xi(s)| &= \sup_{s\in[\tau,t]} \left| \frac{s-\tau}{t-\tau}x + \frac{t-s}{t-\tau}y - x + (t-s)H_p(t,x,p) + O((t-s)^2) \right| \\
&= \sup_{s\in[\tau,t]} \left| \frac{t-s}{t-\tau}\left(x - (t-\tau)H_p(t,x,p) + O((t-\tau)^2)\right) - \frac{t-s}{t-\tau}x \right. \\
&\qquad \left. + (t-s)H_p(t,x,p) + O((t-s)^2) \right| \\
&\leq O((t-\tau)^2).
\end{aligned}$$

Moreover from

$$y_p = \xi_p(\tau) = -(t-\tau)H_{pp}(t,x,p) + O((t-\tau)^2),$$

and from

$$\xi_p(s) = -(t-s)H_{pp}(t,x,p) + O((t-s)^2)$$

for $s \in [\tau, t]$, we obtain

$$\begin{aligned}
\sup_{s\in[\tau,t]} |\eta_p(s) - \xi_p(s)| &= \sup_{s\in[\tau,t]} \left| \frac{t-s}{t-\tau}y_p + (t-s)H_{pp}(t,x,p) + O((t-s)^2) \right| \\
&= \sup_{s\in[\tau,t]} \left| \frac{t-s}{t-\tau}\left(-(t-\tau)H_{pp}(t,x,p) + O((t-\tau)^2)\right) \right. \\
&\qquad \left. + (t-s)H_{pp}(t,x,p) + O((t-s)^2) \right| \\
&\leq O((t-\tau)^2).
\end{aligned}$$

In an analogous way, from

$$y_{pp} = \xi_{pp}(\tau) = -(t-\tau)H_{ppp}(t,x,p) + O((t-\tau)^2),$$

and from
$$\xi_{pp}(s) = -(t-s)H_{ppp}(t,x,p) + O((t-s)^2)$$
for $s \in [\tau, t]$, we obtain
$$\sup_{s\in[\tau,t]} |\eta_{pp}(s) - \xi_{pp}(s)| \le O((t-\tau)^2).$$

Observe now that
$$\dot\eta(s) = \frac{x-y}{t-\tau},$$
and
$$\dot\xi(s) = -H_p(t,x,p) + O(t-s),$$
hence
$$\begin{aligned}\sup_{s\in[\tau,t]} |\dot\eta(s) - \dot\xi(s)| &= \sup_{s\in[\tau,t]} \left| \frac{x-y}{t-\tau} - H_p(t,x,p) + O((t-s)) \right| \\ &= \sup_{s\in[\tau,t]} \Big| H_p(t,x,p) + O(t-\tau) - H_p(t,x,p) + O(t-s) \Big| \\ &\le O(t-\tau).\end{aligned}$$

In the same way we obtain
$$\sup_{s\in[\tau,t]} |\dot\eta_p(s) - \dot\xi_p(s)| \le O(t-\tau),$$
and
$$\sup_{s\in[\tau,t]} |\dot\eta_{pp}(s) - \dot\xi_{pp}(s)| \le O(t-\tau).$$

□

Now, fix $x \in \mathbb{R}^n$ and a compact set $K \subset \mathbb{R}^n$. Call $\xi(\tau, K)$ the subset of $\mathbb{R}^n$ defined as
$$\xi(\tau, K) := \{\xi(\tau) |\ \xi \text{ is a solution of (2.2) with final conditions (2.3)}\}.$$

For any y in $\xi(\tau, K)$ consider the function
$$\phi(\tau, y, t, x) := \min \left\{ \int_\tau^t L(s, \xi(s), \dot\xi(s))ds \;\middle|\; \xi \in [C^2([\tau,t])]^n,\ \xi(\tau) = y,\ \xi(t) = x, \right\},$$
and observe that for any $y \in \xi(\tau, K)$ there exists a unique ξ solution of (2.2) with final conditions (2.3) such that $y = \xi(\tau, p)$. Thus we can see y as $y = y(p)$ with a C^2 dependence of y from p.

Proposition 2.5. *It holds*
$$\left\| \phi(\tau, y(p), t, x) - (t-\tau) L\left(t, x, \frac{x - y(p)}{t-\tau}\right) \right\|_{C^2(K)} \le O((t-\tau)^2).$$

In particular for $t-\tau$ small enough $y \mapsto \phi(\tau, y, t, x)$ and $x \mapsto \phi(\tau, y, t, x)$ are convex with constant $\frac{\tilde C}{t-\tau}$.

Proof. Note that, from the definition, $y \mapsto \phi(\tau, y, t, x)$ and and $x \mapsto \phi(\tau, y, t, x)$ are automatically semiconvex.

Moreover, it is enough to consider the function $y \mapsto \phi(\tau, y, t, x)$ since there is a symmetry between $y \mapsto \phi(\tau, y, t, x)$ and $x \mapsto \phi(\tau, y, t, x)$. Thus the analysis of the two functions is similar.

From the definition, the function $y \mapsto \phi(\tau, y, t, x)$ has a unique minimum ξ which is the solution to system (2.2) with final conditions (2.3). Thus the C^2 dependence of y from p implies, for a small $t - \tau$, that $p \mapsto \phi(\tau, y(p), t, x)$ belongs to $C^2(K)$.

Let ξ be the unique minimizer for $\phi(\tau, y, t, x)$ and observe that $x = \eta(t)$ and $\frac{x-y}{t-\tau} = \dot{\eta}(t)$, where η is the straight line joining x to y as in (2.6).

$$
\begin{aligned}
&\sup_{p\in K} \left|\phi(\tau, y(p), t, x) - (t-\tau)L\left(t, x, \frac{x-y(p)}{t-\tau}\right)\right| = \\
&= \sup_{p\in K}\left|\int_\tau^t L(s,\xi(s),\dot{\xi}(s))ds - \int_\tau^t L(t,\eta(t),\dot{\eta}(t))ds\right| \\
&\le \sup_{p\in K}\Bigg\{\left|\int_\tau^t L(s,\xi(s),\dot{\xi}(s))ds - \int_\tau^t L(t,\xi(s),\dot{\xi}(s))ds\right| \\
&\qquad + \left|\int_\tau^t L(t,\xi(s),\dot{\xi}(s))ds - \int_\tau^t L(t,\eta(t),\dot{\xi}(s))ds\right| \\
&\qquad + \left|\int_\tau^t L(t,\eta(t),\dot{\xi}(s))ds - \int_\tau^t L(t,\eta(t),\dot{\eta}(t))ds\right|\Bigg\} \\
&\le \sup_{p\in K}\left\{C_1\int_\tau^t |s-t|ds + C_2\int_\tau^t|\xi(s)-\eta(t)|ds + C_3\int_\tau^t|\dot{\xi}(s)-\dot{\eta}(t)|ds\right\} \\
&\le \sup_{p\in K}\Bigg\{-\frac{C_1}{2}(t-\tau)^2 + C_2 H_p(t,x,p)\int_\tau^t|t-s|ds \\
&\qquad + C_2\int_\tau^t O((t-s)^2)ds + C_3\int_\tau^t O((t-s))ds\Bigg\} \\
&\le O((t-s)^2).
\end{aligned}
$$

Moreover for the first derivative

$$
\begin{aligned}
&\sup\nolimits_{p\in K}\left|\partial_p\left[\phi(\tau,y(p),t,x) - (t-\tau)L\left(t,x,\frac{x-y(p)}{t-\tau}\right)\right]\right| = \\
&= \quad \sup_{p\in K}\Bigg|\int_\tau^t L_x(s,\xi(s),\dot{\xi}(s))\xi_p(s)ds + \int_\tau^t L_v(s,\xi(s),\dot{\xi}(s))\dot{\xi}_p(s)ds \\
&\qquad -\int_\tau^t L_x(t,\eta(t),\dot{\eta}(t))\eta_p(t)ds - \int_\tau^t L_v(t,\eta(t),\dot{\eta}(t))\dot{\eta}_p(t)ds\Bigg|
\end{aligned}
$$

$$
\begin{aligned}
&\leq \sup_{p\in K}\left\{\left|\int_\tau^t L_x(s,\xi(s),\dot\xi(s))(-(t-s)H_p(t,x,p)+O(t-s)^2)ds\right|\right.\\
&\qquad\left.+\left|\int_\tau^t (L_v(s,\xi(s),\dot\xi(s))-L_v(t,\eta(t),\dot\eta(t)))(-H_{pp}(t,x,p)+O(t-\tau))ds\right|\right\}\\
&\leq \sup_{p\in K}\left\{C_1\int_\tau^t |(s-t)+O(t-s)|ds\right.\\
&\qquad\left.+\left|\int_\tau^t [L_v(s,\xi(s),\dot\xi(s))-L_v(t,\eta(t),\dot\eta(t))](C_2+O(t-\tau))ds\right|\right\}\\
&\leq O((t-s)^2).
\end{aligned}
$$

Analogously for the second derivative

$$
\sup_{p\in K}\left|\partial_{pp}\left[\phi(\tau,y(p),t,x)-(t-\tau)L\left(t,x,\frac{x-y(p)}{t-\tau}\right)\right]\right|\leq O((t-s)^2).
$$

The map $p\mapsto y(p)$ is $C^2(K)$, it has bounded derivative and the same holds true also for its inverse, due to Proposition 2.4. Thus it follows that $\phi(\tau,y,t,x)$ and $(t-\tau)L\left(t,x,\frac{x-y}{t-\tau}\right)$ are close in $C^2(\tilde K)$, where $\tilde K$ is the image of K through the map $p\mapsto y(p)$. Therefore $y\mapsto\phi(\tau,y,t,x)$ is convex with constant $\frac{\tilde C}{t-\tau}$, the same constant of $y\mapsto(t-\tau)L\left(t,x,\frac{x-y}{t-\tau}\right)$. □

Remark 2.6. All the estimates found strictly depend on the compact set K, however thanks to the finite speed of propagation of the minimizers ξ, see point (iii) of Theorem 1.19, the estimates can be made uniform for our $u^-_{t,0}$.

Let us now come back to our case.

Proposition 2.7. *For $0\leq\tau<t$ consider the backward solution defined in (1.8) for y in Ω_τ. Then for $t-\tau$ small enough the maximum is unique for all $y\in\Omega_\tau$.*

Proof. The backward solution can be written in this equivalent way

$$
u^-_{t,0}(\tau,y)=\max_{x\in\Omega_t}\{u(t,x)-\phi(\tau,y,t,x)\}. \tag{2.7}
$$

Recalling that $u(t,\cdot)$ is semiconcave with constant $\frac{C}{t}$ and that $-\phi(\tau,y,t,\cdot)$ is strictly concave with constant $\frac{\tilde C}{t-\tau}$, we can rewrite (2.7) as

$$
u^-_{t,0}(\tau,y)=\max_{x\in\Omega_t}\left\{u(t,x)-\frac{C}{t}|x|^2-\phi(\tau,y,t,x)+\frac{C}{t}|x|^2\right\}.
$$

Hence, since $u(t,x)-\frac{C}{t}|x|^2$ is concave and $-\phi(\tau,y,t,x)+\frac{C}{t}|x|^2$ remains strictly concave, the function $u^-_{t,0}(\tau,y)$ is the maximum of a strictly concave function, hence this maximum is unique. Thus there exists a unique $x\in\Omega_t$ such that

$$
u^-_{t,0}(\tau,y)=u(t,x)-\phi(\tau,y,t,x),
$$

i.e. there exists a unique curve $\xi \in [C^2([\tau,t])]^n$ such that $\xi(\tau)=y$, $\xi(t)=x$ and

$$u^-_{t,0}(\tau,y) = u(t,\xi(t)) - \int_\tau^t L(s,\xi(s),\dot{\xi}(s))ds.$$

□

Corollary 2.8. *For $t-\tau$ small enough and $s \in (\tau,t)$ the function $u^-_{t,0}(\tau,\cdot)$ is $C^{1,1}(\Omega_s)$.*

Proof. From the above proposition we know that $u^-_{t,0}(s,\cdot)$ is $C^1(\Omega_s)$ for every $s \in [\tau,t)$. Consider now the *forward solution* defined from $u^-_{t,0}(\tau,\cdot)$

$$u^+_{t,\tau}(s,x) := \min\left\{ u^-_{t,0}(\tau,\xi(\tau)) + \int_\tau^s L(l,\xi(l),\dot{\xi}(l))dl \;\middle|\; \xi(s)=x,\ \xi \in [C^2([\tau,s])]^n \right\}.$$

Due to the fact that $u^-_{t,0}(\tau,y)$ has a unique maximizer for every $y \in \Omega_\tau$ we have that $u^+_{t,\tau}(s,x) = u^-_{t,0}(s,x)$ for every $s \in [\tau,t]$ and $x \in \Omega_s$. Thus for $s \in (\tau,t)$, $u^-_{t,0}(s,\cdot)$ is both semiconvex and semiconcave, hence $C^{1,1}(\Omega_s)$. □

Remark 2.9. As a consequence of Proposition 2.7, for every $y \in \Omega_\tau$ there exists only one curve which is a maximizer for the function $\tilde{u}(\tau,y)$ and a generalized backward characteristic. Hence generalized backward characteristics which are also maximizers do not intersect even at time τ. It remains to prove the following.

Proposition 2.10. *Every generalized backward characteristic $\xi(s)$, i.e. a solution of (2.2) with final conditions (2.3) where $p \in D^+_x u(t,x)$, is a maximizer for $u^-_{t,0}(\tau,\xi(\tau))$ if $t-\tau$ is small enough.*

Proof. Let ξ be a generalized backward characteristic with $\xi(t)=x$, $p(t)=p \in D^+_x u(t,x)$ and $\xi(\tau)=y$. Then ξ is a minimizer for $\phi(\tau,t,y,x)$ and $p=p(t)=-D_y\phi(\tau,y,t,x)$.

Let $\tilde{\xi}$ be the unique maximizer for $u^-_{t,0}(\tau,y)$ and suppose by contradiction that $\tilde{\xi}$ differs from ξ, in particular $\tilde{\xi}(t)=\tilde{x} \neq x = \xi(t)$. Then by definition

$$u^-_{t,0}(\tau,y) = u(t,\tilde{x}) - \phi(\tau,y,t,\tilde{x}) > u(t,x) - \phi(\tau,y,t,x).$$

Thus, for the differentiability and the convexity of $\phi(\tau,y,t,\cdot)$

$$\begin{aligned} u(t,\tilde{x}) - u(t,x) &> \phi(\tau,y,t,\tilde{x}) - \phi(\tau,y,t,x) \\ &\geq \langle D_y\phi(\tau,y,t,x), \tilde{x}-x\rangle + \frac{\tilde{C}}{t-\tau}|\tilde{x}-x|^2. \end{aligned}$$

On the other hand for the semiconcavity of $u(t,\cdot)$

$$u(t,\tilde{x}) - u(t,x) < \langle p, \tilde{x}-x\rangle + \frac{C}{t}|\tilde{x}-x|^2.$$

Recalling that $p=-D_y\phi(\tau,y,t,x)$, for $t-\tau$ small enough we reach the absurd

$$\frac{C}{t} > \frac{\tilde{C}}{t-\tau}.$$

□

From the above proposition it follows

Corollary 2.11. *Generalized backward characteristics cannot intersect in $[\tau, t)$ if $t - \tau$ is small enough.*

2.2 Proof of Theorem 2.1

2.2.1 Preliminary remarks

Let u be a viscosity solution of (2.1). Applying Proposition 1.18 we can assume without loss of generality that u is a solution of the Cauchy Problem (2.1) with a bounded Lipschitz initial datum $u(0, x) = u_0(x)$ over a bounded domain $[0, \delta] \times \mathcal{U}$. Moreover assumptions (H1)-(H2) guarantee that the Hamiltonian is convex and has super-linear growth in the last variable.

We will prove the SBV regularity over the smaller interval of time $[\tau, \tau + \varepsilon]$ for a fixed $\tau > 0$, $\varepsilon > 0$ small enough and such that $[\tau, \tau + \varepsilon] \subset [0, \delta]$. As we have already seen, this is necessary to prevent intersections of generalized backward characteristics.

We consider a ball $B_R(0) \subset \mathbb{R}^n$ and a bounded convex set $\Omega \subset [\tau, \tau + \varepsilon] \times \mathbb{R}^n$ with the properties that

- $\{s\} \times B_R(0) \subset \Omega$ for every $s \in [\tau, \tau + \varepsilon]$;
- for any $(t, x) \in \Omega$ and for any C^2 curve ξ which minimizes $u(t, x)$ in (1.3), the entire curve $\xi(s)$ for $s \in [\tau, t]$ is contained in Ω.

Indeed, from the fact that $\|Du\|_\infty < \infty$, it is enough to choose

$$\Omega := \big\{(t, x) \in [\tau, \tau + \varepsilon] \times \mathbb{R}^n \big| \ |x| \leq R + C'(\tau + \varepsilon - t)\big\}$$

with C' sufficiently large and depending only on $\|Du\|_\infty$ and H.

The general idea of the proof is now standard, see [4], [9]. We construct a monotone bounded functional $F(t)$ defined on the interval $[\tau, \tau + \varepsilon]$. Then, we relate the presence of a Cantor part in the matrix $D_x^2 u(t, \cdot)$ for a certain t in $[\tau, \tau + \varepsilon]$ with a jump of the functional F in t. Since this functional can have only a countable number of jumps, the Cantor part of $D_x^2 u(t, \cdot)$ can be different from zero only for a countable number of t's.

Remark 2.12. Once we have formalized the above strategy and proved the SBV regularity for almost every t in $[\tau, \tau + \varepsilon]$ the conclusion that $D_x u$ belongs to $[SBV_{loc}(\Omega)]^n$ follows from the slicing theory of BV functions (see Theorem 3.108 of [5]). The local SBV regularity of $\partial_t u$ follows instead from the Volpert chain rule.

2.2.2 Construction of the functional F

Consider t belonging to $(\tau, \tau + \varepsilon]$ for a fixed $\tau > 0$ and $\varepsilon > 0$ small enough. For any $\tau \leq s < t$ we define the set-valued map

$$X_{t,s}(x) := \{\xi(s) | \ \xi(\cdot) \text{ is a solution of (2.2), with } \xi(t) = x, \ p(t) = p \in D_x^+ u(t, x)\}.$$

Moreover we will denote by $\chi_{t,s}$ the restriction of $X_{t,s}$ to the points where it is single-valued. According to Theorem 1.25, the domain of $\chi_{t,s}$, $dom(\chi_{t,s}) =: U_t$, consists of those points where $D_x^+ u(t,x)$ is single-valued, i.e. there exists a unique minimizer for $u(t,x)$. For that reason $\chi_{t,s}$ is clearly defined $\mathcal{H}^n$-a.e. in Ω_t. We will sometimes write $\chi_{t,s}(\Omega_t)$ meaning $\chi_{t,s}(U_t)$.

Remark 2.13. In the definition of $X_{t,s}$ we follow generalized backward characteristics starting at time $t > 0$ till time s. As we have already seen, if $t - s$ is small enough, generalized backward characteristics cannot intersect except than at time t. Thus if we choose $\varepsilon > 0$ small enough we have the injectivity of the set valued map $X_{t,\tau}$ over the interval of time $[\tau, \tau + \varepsilon]$.

Note that in the case $H = H(D_x u)$ the authors of [9] were able, in Proposition 5.2, to prove the injectivity of $X_{t,0}$, as a set-valued map, for every $t \in [0, \varepsilon]$ with ε small enough.

Therefore, equivalently to Proposition 5.2 in [9], we can state

Proposition 2.14. *Let t be fixed such that $\tau < t \leq \tau + \varepsilon$, for an $\varepsilon > 0$ small enough, which does not depend on t. Then taken any two solutions (ξ_1, p_1) and (ξ_2, p_2) of the system (2.2) with final condition*

$$\xi_i(t) = x_i \in \Omega_t \qquad p_i(t) \in D_x^+ u(t, x_i) \qquad i = 1, 2,$$

and $(\xi_1(t), p_1(t)) \neq (\xi_2(t), p_2(t))$ it follows that $\xi_1(\tau) \neq \xi_2(\tau)$. Hence, in particular, the map $x \mapsto X_{t,\tau}(x)$ is injective as a set-valued map.

Proof. It follows from Corollary 2.11. $\square$

For every $\tau < t \leq \tau + \varepsilon$, we can now define the functional

$$F(t) := \mathcal{H}^n(\chi_{t,\tau}(U_t)). \tag{2.8}$$

Lemma 2.15. *The functional F is non increasing,*

$$F(s) \geq F(t) \quad \text{for any } s, t \in (\tau, \tau + \varepsilon] \text{ with } s < t.$$

Proof. As in the proof of Lemma 4.1 in [9], the claim follows from the following consideration:

$$\chi_{t,\tau}(\Omega_t) \subset \chi_{s,\tau}(\Omega_s) \quad \text{ for every } \tau \leq s \leq t \leq \tau + \varepsilon.$$

Indeed, consider any $y \in \chi_{t,\tau}(\Omega_t)$. Then there exists a curve ξ in $[C^2([\tau, t])]^n$ and a point $x \in \Omega_t$ such that ξ is the unique minimizer in (1.3) with the following endpoints conditions $\xi(t) = x, \xi(\tau) = y$. Such a curve remains the unique minimizer also for $u(s, \xi(s))$ for any $\tau \leq s \leq t \leq \tau + \varepsilon$. Hence, setting $z = \xi(s)$, we have that the point y can be seen as $y = \chi_{s,\tau}(z)$ and $y \in \chi_{s,\tau}(\Omega_s)$. $\square$

2.2.3 Hille-Yosida transformation

Take a Borel set $A \subset \Omega_t$ for a fixed time $t \in (\tau, \tau + \varepsilon]$. In order to compute the measure $\mathcal{H}^n(X_{t,\tau}(A))$ we follow the evolution of the set along generalized backward characteristics till the time τ.

Let us recall how the characteristics and their dual arc evolve in time. They are solutions of the system (2.2), together with the final condition (2.3) where p belongs to $D_x^+ u(t,x)$.

We have to face the following problem: the function $D_x^+ u(t,\cdot)$ is a multi-valued function of bounded variation which is not Lipschitz in general. However it can be easily related to a maximal monotone function whose graph can be parametrized in a Lipschitz way as shown in Alberti and Ambrosio [1].

Let us consider the graph $(A, D_x^+ u(t,A))$ for a Borel set $A \subset \Omega_t$. Since $u(t,x)$ is semiconcave in x, $v(x) := -(u(t,x) - \frac{1}{2}C|x|^2)$ is a convex function. Note that the semiconcavity constant should depend on t, i.e. $C(t) = \frac{C}{t}$, however a uniform one can be taken due to the fact that t belongs to $(\tau, \tau+\varepsilon]$ where $\tau > 0$. Moreover, as seen in Theorem 1.14-(iv), the differential of v is a maximal monotone function. It can be proven, see for example [1], that the graph of a maximal monotone function is a Lipschitz submanifold without boundary. Adapting the same procedure to our case, we can parametrize the graph of the derivative of our semiconcave function with a 1-Lipschitz function.

Indeed, we pass from our graph $\{(x, D_x^+ u(t,x)) |\ x \in A\}$ to the graph of a maximal monotone function with the following transformation

$$\begin{cases} x = x \\ y = Cx - p, \end{cases}$$

where C is the semiconcavity constant of $u(t,\cdot)$. Then we apply an Hille-Yosida transformation to have a 1-Lipschitz parametrization of it.

$$\begin{cases} z = x + y \\ w = y. \end{cases}$$

Call $T(x) := D_x v(x)$ the maximal monotone function. Retracing the passages above, we can express w as a 1-Lipschitz single-valued function of z. Taking $z \in B := A + T(A)$

$$\begin{cases} z = z \\ w = (Id_n + (T)^{-1})^{-1}(z). \end{cases}$$

Thus, coming back to our original coordinates, we can describe our graph with the following Lipschitz parametrization

$$\begin{cases} x(z) = z - w(z) \\ p(z) = Cz - (C+1)w(z), \end{cases} \tag{2.9}$$

where $z \in B$, i.e. we have

$$\Gamma_A := \{(x, D_x^+ u(t,x)) |\ x \in A\} = \{(z - w(z), Cz - (C+1)w(z)) |\ z \in B\}.$$

Remark 2.16. As explained in [1] the 1-Lipschitz function $w(z)$ is exactly the derivative of the inf-convolution function of $v(x) = -(u(t,x) - \frac{1}{2}C|x|^2)$

$$f(z) = \min_{x \in \mathbb{R}^n} \left\{ v(x) + \frac{|x-z|^2}{2} \right\}.$$

Thus we have $w(z) = f_z(z)$ where f is a convex function.

When applying the flux backward in time, starting from our set Γ_A, characteristics $\xi(s,z)$ and $p(s,z)$ evolve according to

$$\begin{cases} \dot{\xi}(s,z) = H_p(s,\xi(s,z),p(s,z)) \\ \dot{p}(s,z) = -H_x(s,\xi(s,z),p(s,z)) \end{cases} \tag{2.10}$$

with final conditions

$$\begin{cases} \xi(t,z) = x(z) = z - w(z) \\ p(t,z) = p(z) = Cz - (C+1)w(z), \end{cases} \tag{2.11}$$

for z in B. Since the flux is described by smooth equations and thanks to the fact that the parametrization of our initial set is 1-Lipschitz, the solutions $\xi(s,z), p(s,z)$ are Lipschitz curves.

We can now rewrite $X_{t,\tau}$ in an equivalent way, for x in A

$$\begin{aligned} X_{t,\tau}(x) &= \{\xi(\tau) \mid \quad \xi(\cdot) \text{ is a solution of (2.2), with } \xi(t) = x,\ p(t) = p \in D_x^+ u(t,x)\} \\ &= \{\xi(\tau,z) \mid \xi(\cdot,z) \text{ is a solution of (2.10), with } \xi(t,z) = z - w(z), \\ &\qquad p(t,z) = Cz - (C+1)w(z), z \in x + T(x)\}. \end{aligned}$$

With an abuse of notation we will denote with $\xi(\tau,\cdot) : B \to \Omega_\tau$ the function $X_{t,\tau}(\cdot)$ when we are considering the Lipschitz parametrization; with this notation $X_{t,\tau}(A) = \xi(\tau, B)$. We can now apply the Area Formula to $\xi(\tau,\cdot)$

$$\int_{\xi(\tau,B)} \mathcal{H}^0((\xi(\tau,\cdot)^{-1}(w))dw = \int_B |\det(\xi_z(\tau,z))|dz. \tag{2.12}$$

Thanks to the injectivity of the map $X_{t,\tau}$ which is preserved when passing to the Lipschitz parametrization, the left term of (2.12) is precisely the measure of the set $\xi(\tau, B)$.

Hence, we have

$$\int_{\xi(\tau,B)} \mathcal{H}^0((\xi(\tau,\cdot)^{-1}(w))dw = \mathcal{H}^n(\xi(\tau,B)) = \mathcal{H}^n(X_{t,\tau}(A)).$$

To compute $\det(\xi_z(\tau,z))$ we differentiate in z the equations (2.10), (2.11) obtaining that ξ_z and p_z satisfy the system

$$\begin{cases} \dot{\xi}_z(s,z) = H_{px}(s,\xi(s,z),p(s,z))\xi_z(s,z) + H_{pp}(s,\xi(s,z),p(s,z))p_z(s,z) \\ \dot{p}_z(s,z) = -H_{xx}(s,\xi(s,z),p(s,z))\xi_z(s,z) - H_{xp}(s,\xi(s,z),p(s,z))p_z(s,z) \end{cases} \tag{2.13}$$

with the final conditions

$$\begin{cases} \xi_z(t,z) = Id_n(z) - w_z(z) \\ p_z(t,z) = CId_n(z) - (C+1)w_z(z), \end{cases} \tag{2.14}$$

for any $z \in B$.

2.2.4 Approximation and area estimates

If we choose $\varepsilon > 0$ small enough we can approximate our curves with straight lines for any t in $(\tau, \tau + \varepsilon]$, i.e. we can write

$$\xi(\tau, z) = \xi(t, z) - (t - \tau)\dot{\xi}(t, z) + O((t - \tau)^2).$$

Using this approximation and (2.13) we obtain

$$\begin{aligned}\det(\xi_z(\tau, z)) = \det &\Big(\xi_z(t, z) - (t - \tau)H_{px}(t, x(z), p(z))\xi_z(t, z) - (t - \tau)H_{pp}(t, x(z), p(z))p_z(t, z)\Big) \\ &+ O((t - \tau)^2).\end{aligned} \tag{2.15}$$

Since we are now considering nearly straight lines, instead of more general curves, we can expect that this approximation should allow us to adapt the techniques of [9] and recover the lemmas needed.

Before going on, let us give an explicit formula for the spatial Laplacian of our solution. Thanks to the semiconcavity of $u(t, \cdot)$ its spatial Laplacian is a measure. Moreover, using the 1-Lipschitz parametrization given by Hille-Yosida, the spatial Laplacian can be seen as the push-forward of a particular measure.

Lemma 2.17. *For any Borel set A, let $\{(x(z), p(z))|\ z \in A + T(A)\}$ be the 1-Lipschitz parametrization of the set $\{(x, D_x^+ u(t, x)|\ x \in A\}$ as seen above in (2.9). Then we have*

$$\Delta u(t, A) = x(z)_\sharp \left[\left(\sum_{i,k} \frac{\partial p_i(z)}{\partial z_k} [\mathrm{cof}\ x_z(z)]_{ik} \right) \mathcal{H}^n \right] (A).$$

Here cof A *is the cofactor matrix of the matrix A.*

This formula has been shown to the authors by C. De Lellis.

Proof. We can assume A open. Take any ϕ in $C_c^\infty(\mathbb{R}^n)$ and compute

$$\begin{aligned}\int_A \phi(x) d[D_x^2 u(t, x)]_{ij} &= -\int_A [D_x u(t, x)]_i \frac{\partial \phi(x)}{\partial x_j} dx \\ &= -\int_{A+T(A)} p_i(z) \frac{\partial \phi(x(z))}{\partial x_j} \det(x_z(z)) dz \\ &= -\int_{A+T(A)} p_i(z) \sum_k \left(\frac{\partial \phi(x(z))}{\partial z_k} \frac{\partial z_k(x(z))}{\partial x_j} \right) \det(x_z(z)) dz \\ &= -\int_{A+T(A)} p_i(z) \sum_k \left(\frac{\partial \phi(x(z))}{\partial z_k} [\mathrm{cof}\ x_z(z)]_{jk} \right) dz \\ &= \int_{A+T(A)} \phi(x(z)) \sum_k \left(\frac{\partial p_i(z)}{\partial z_k} [\mathrm{cof}\ x_z(z)]_{jk} \right) dz \\ &\quad + \int_{A+T(A)} \phi(x(z)) p_i(z) \sum_k \left(\frac{\partial}{\partial z_k} [\mathrm{cof}\ x_z(z)]_{jk} \right) dz.\end{aligned}$$

In the lines above we have used the 1-Lipschitz parametrization of the set $\{(x, D_x^+ u(t,x)) |\ x \in A\}$ and the fact that

$$\frac{\partial z_k(x(z))}{\partial x_j} = [x_z(z)]^{-1}_{kj} = \frac{1}{\det(x_z(z))}[\mathrm{cof}\ x_z(z)]_{jk}.$$

Now, repeating upside down the passages starting from the last term, one obtains that

$$\begin{aligned}
&\int_{A+T(A)} \phi(x(z))p_i(z) \sum_k \left(\frac{\partial}{\partial z_k}[\mathrm{cof}\ x_z(z)]_{jk} \right) dz \\
&= -\int_{A+T(A)} \sum_k \left(\frac{\partial}{\partial z_k} \Big(\phi(x(z))p_i(z) \Big) [\mathrm{cof}\ x_z(z)]_{jk} \right) dz \\
&= -\int_{A+T(A)} \sum_k \left(\frac{\partial}{\partial z_k} \Big(\phi(x(z))p_i(z) \Big) \frac{\partial z_k(x(z))}{\partial x_j} \det(x_z(z)) \right) dz \\
&= -\int_{A+T(A)} \frac{\partial}{\partial x_j} \Big(\phi(x(z))p_i(z) \Big) \det(x_z(z)) dz \\
&= -\int_A \frac{\partial}{\partial x_j} \Big(\phi(x)[D_x u(t,x)]_i \Big) dx
\end{aligned}$$

which is equal to zero due to the fact that ϕ has compact support. Hence

$$\sum_k \left(\frac{\partial}{\partial z_k}[\mathrm{cof}\ x_z(z)]_{jk} \right) = 0.$$

$\square$

We are now able to prove an analogous of Lemma 4.3 in [9].

Lemma 2.18. *For ε small enough (depending only on the bound M for $\|H_{px}\|$), let $t \in (\tau, \tau+\varepsilon]$ and $A \subset \Omega_t$ be a Borel set. Then*

$$\mathcal{H}^n(X_{t,\tau}(A)) \geq C_1 \mathcal{H}^n(A) - C_2(t-\tau) \int_A d\Delta u(t,\cdot) + O((t-\tau)^2),$$

where C_1, C_2 are positive constants (depending on C, c_H). $\Delta u(t,\cdot)$ is the spatial Laplacian of $u(t,\cdot)$.

Proof. Let us start from (2.15).

For $t-\tau$ small enough the matrix

$$Id_n(z) - (t-\tau)H_{px}(t, x(z), p(z))$$

is invertible. Indeed, since $\exists M > 0$ such that the norm $\|H_{px}(\cdot,\cdot,\cdot)\| < M$ it is sufficient to take $\varepsilon < \frac{1}{2nM}$. This condition ensures that

$$\det(Id_n(z) - (t-\tau)H_{px}(t, x(z), p(z))) > \frac{1}{2} > 0.$$

Thus this determinant can be put in evidence in (2.15)

$$
\begin{aligned}
|\det(\xi_z(\tau,z))| &= |\det(Id_n-(t-\tau)H_{px})||\det\big(\xi_z-(t-\tau)(Id_n-(t-\tau)H_{px})^{-1}H_{pp}p_z\big)|+O((t-\tau)^2)\\
&> \frac{1}{2}|\det(\xi_z-(t-\tau)H_{pp}p_z)|+O((t-\tau)^2).
\end{aligned}
$$

To lighten the computation above we have omitted the dependence of H_{px}, H_{pp} from $t, x(z), p(z)$ and of ξ_z, p_z from t, z. Moreover we used the fact that for $t-\tau$ small enough it is possible to expand the inverse

$$(Id_n-(t-\tau)H_{px})^{-1}=Id_n+(t-\tau)H_{px}+O((t-\tau)^2).$$

We are then left to expand the determinant in series

$$\det(\xi_z-(t-\tau)H_{pp}p_z)=\det(\xi_z)-(t-\tau)\mathrm{tr}\ \big([\mathrm{cof}\ \xi_z]^T H_{pp}p_z\big)+O((t-\tau)^2),$$

and use that $w=f_z$ as underlined in the Remark 2.16, so that, recalling (2.14),

$$\xi_z=Id_n-w_z=Id_n-f_{zz}, \qquad p_z=CId_n-(C+1)w_z=CId_n-(C+1)f_{zz}.$$

Call λ_i, for $i=1,...,n$, the eigenvalues of the positive semidefinite matrix f_{zz}. Hence we can compute

$$\det(\xi_z)=\prod_i(1-\lambda_i) \qquad [\mathrm{cof}\ \xi_z]_{ii}=\prod_{j\neq i}(1-\lambda_j).$$

The convexity of f and the 1-Lipschitzianity of f_z imply that all the eigenvalues are bounded from above and from below: $0\le\lambda_i\le 1$, for $i=1,\dots,n$. Thus, for every $i=1,\dots,n$, we have $0\le 1-\lambda_i\le 1$ and $-1\le C-(C+1)\lambda_i\le C$, in particular this last inequality suggests that we have to work a bit to bound our determinant, since $C-(C+1)\lambda_i$ has no definite sign.

$$
\begin{aligned}
&\frac{1}{2}\big(\det(\xi_z)-(t-\tau)\mathrm{tr}\ \big([\mathrm{cof}\ \xi_z]^T H_{pp}p_z\big)\big)+O((t-\tau)^2)=\\
&=\frac{1}{2}\left(\prod_i(1-\lambda_i)-(t-\tau)\mathrm{tr}\ \left(\mathrm{diag}\left[\prod_{j\neq i}(1-\lambda_j)\right]H_{pp}\mathrm{diag}[C-(C+1)\lambda_i]\right)\right)+O((t-\tau)^2)\\
&=\frac{1}{2}\left(\prod_i(1-\lambda_i)-(t-\tau)\sum_i\prod_{j\neq i}(1-\lambda_j)[H_{pp}]_{ii}(C-(C+1)\lambda_i)\right)+O((t-\tau)^2)\\
&=\frac{1}{2}\prod_i(1-\lambda_i)-(t-\tau)\frac{1}{2}\sum_i\prod_{j\neq i}(1-\lambda_j)[H_{pp}]_{ii}(C(1-\lambda_i)-\lambda_i)+O((t-\tau)^2)\\
&=\frac{1}{2}(1-(t-\tau)C\ \mathrm{tr}\ H_{pp})\prod_i(1-\lambda_i)+(t-\tau)\frac{1}{2}\sum_i\lambda_i[H_{pp}]_{ii}\prod_{j\neq i}(1-\lambda_j)+O((t-\tau)^2).
\end{aligned}
$$

Now that all the terms have positive sign for an ε small enough, we can use the uniform convexity of H in p and the bounds on λ_i to show that there exist constants C_1, C_2, all of them depending only on C, c_H, such that

$$\begin{aligned}
|\det(\xi_z(\tau,z))| &\geq C_1 \prod_i (1-\lambda_i) + (t-\tau) C_2 \sum_i \lambda_i \prod_{j\neq i}(1-\lambda_j) + O((t-\tau)^2) \\
&\geq C_1 \prod_i (1-\lambda_i) + (t-\tau) C_2 \sum_i \lambda_i \prod_{j\neq i}(1-\lambda_j) - n(t-\tau)C_2 C \prod_i (1-\lambda_j) \\
&\quad + O((t-\tau)^2) \\
&= C_1 \prod_i (1-\lambda_i) - (t-\tau) C_2 \sum_i (C(1-\lambda_i)-\lambda_i)) \prod_{j\neq i}(1-\lambda_j) + O((t-\tau)^2) \\
&= C_1 \prod_i (1-\lambda_i) - (t-\tau) C_2 \sum_i (C-(C+1)\lambda_i) \prod_{j\neq i}(1-\lambda_j) + O((t-\tau)^2).
\end{aligned}$$

Therefore if we compute the area formula (2.12) we obtain

$$\begin{aligned}
\int_B |\det(\xi_z(\tau,z))| dz \geq \int_B &\left[C_1 \prod_i (1-\lambda_i) - (t-\tau) C_2 \sum_i (C-(C+1)\lambda_i) \prod_{j\neq i}(1-\lambda_j) \right] dz \\
&+ O((t-\tau)^2).
\end{aligned}$$

Applying Lemma (2.17) and recalling that $1-\lambda_i$ are the eigenvalues of $\xi_z(t,z)$ we obtain the thesis.

$$\mathcal{H}^n(X_{t,\tau}(A)) \geq C_1 \mathcal{H}^n(A) - C_2(t-\tau)\int_A d\Delta u(t,\cdot) + O((t-\tau)^2)$$

where C_1, C_2 are constants depending only on C, c_H. □

In order to complete the proof of the main theorem we need to prove a Lemma which states the equivalent result of Lemma 5.1 in [9].

Lemma 2.19. *If $\varepsilon > 0$ is small enough, for any $t \in (\tau, \tau+\varepsilon]$, any $\delta \in [0, t-\tau]$ and any Borel set $A \subset \Omega_t$ we have*

$$\mathcal{H}^n(X_{t,\tau+\delta}(A)) \geq \left(\frac{1}{2}\right)^n \left(\frac{t-(\tau+\delta)}{t-\tau}\right)^n \mathcal{H}^n(X_{t,\tau}(A)).$$

Proof. Fix t in $(\tau, \tau+\varepsilon]$, and let A be a Borel set $A \subset \Omega_t$. Without loss of generality we can suppose A to be a compact set.

Consider an approximation of the vector field induced by our generalized backward characteristics by taking a dense sequence of points $\{x_i\}_{i=1}^\infty$ in A. Fix an integer $I > 0$, call $A_I := \{x_i |\ i = 1, \dots, I\}$ and define for any s such that $\tau \leq s < t$ and $y \in X_{t,s}(A)$

$$(u_I)^-_{t,=}(s,y) := \max\left\{ u(t,\xi(t)) - \int_s^t L(l,\xi(l),\dot\xi(l))dl \;\middle|\; \xi \text{ is a } C^2([s,t]) \text{ curve}, \xi(s)=y,\ \xi(t)\in A_I \right\}.$$

We assume in addition that the sequence $\{x_i\}_{i\in I}$ is big enough so that we can uniformly bound the speed of propagation of every maximizer ξ.

Remark 2.20. All the properties which we stated for maximizers of the backward solution and for the backward solution itself are preserved in each cone of propagation for the maximizers of this approximated backward solution (Euler equation, systems for maximizer and dual arc, no-crossing property, etc) and for $(u_I)^-_{t,0}$ (a.e. differentiability, dynamic programming principle, semiconvexity).

Through this approximation the set $E_s := X_{t,s}(A)$ is split into at most I open regions E^i_s, $i = 1, \dots, I$, defined by

$$E^i_s := \text{interior of } \{y \in X_{t,s}(A) | \; \exists \xi \text{ maximizer for } (u_I)^-_{t,0}(s,y) \text{ such that } \xi(t) = x_i\},$$

together with the set

$$J^I_s := \bigcup_{i \neq j} \left(\overline{E}^i_s \cap \overline{E}^j_s \right)$$

of negligible $\mathcal{H}^n$-measure. Indeed, even for $(u_I)^-_{t,0}(s,\cdot)$ the set of points with more than one maximum is the set of point of non differentiability and this set has $\mathcal{H}^n$-measure zero.

Call

$$X^I_{t,s}(x_i) := \{\xi(s) | \; \xi \text{ is a maximizer for } (u_I)^-_{t,0}(s,y) \text{ with } y \in \overline{E}^i_s\},$$

this is a multi-valued function defined on the set A_I.

The set $X^I_{t,s}(A_I)$ converges in the Hausdorff sense to the set $X_{t,s}(A)$ as I tends to infinity. Indeed, it follows from the strong convergence of the maximizers of $(u_I)^-_{t,0}$ to the maximizers of $u^-_{t,\delta}$ which is ensured by their bound on the derivative (Theorem 1.19-(iii)). Thus

$$\mathcal{H}^n(X_{t,s}(A)) \geq \limsup_{I \to \infty} \mathcal{H}^n(X^I_{t,s}(A_I)).$$

Let us decompose $\mathcal{H}^n(X^I_{t,s}(A_I))$ in the sum over $i \in I$ of $\mathcal{H}^n(X^I_{t,s}(x_i))$. Using the one to one correspondence of Lemma 2.3

$$\frac{\xi_p(\tau)}{\tau - t} = H_{pp}(t, x_i, p) + O(t - \tau)$$

and

$$\frac{\xi_p(\tau + \delta)}{\tau + \delta - t} = H_{pp}(t, x_i, p) + O(t - \tau).$$

Therefore

$$\left| \frac{\xi_p(\tau)}{\tau - t} - \frac{\xi_p(\tau + \delta)}{\tau + \delta - t} \right| \leq O(t - \tau),$$

and

$$\left| \left(\frac{t - (\tau + \delta)}{t - \tau} \right) \xi_p(\tau)(\xi_p(\tau + \delta))^{-1} - Id \right| \leq O(t - \tau).$$

Thus, passing to the determinant,

$$\det(\xi_p(\tau + \delta)) \geq \left(\frac{1}{2} \right)^n \left(\frac{t - (\tau + \delta)}{t - \tau} \right)^n \det(\xi_p(\tau)).$$

From which it follows

$$\mathcal{H}^n(X^I_{t,\tau+\delta}(x_i)) \geq \left(\frac{1}{2}\right)^n \left(\frac{t-(\tau+\delta)}{t-\tau}\right)^n \mathcal{H}^n(X^I_{t,\tau}(x_i)).$$

Summing up all the terms

$$\mathcal{H}^n(X^I_{t,\tau+\delta}(A_I)) \geq \left(\frac{1}{2}\right)^n \left(\frac{t-(\tau+\delta)}{t-\tau}\right)^n \mathcal{H}^n(X^I_{t,\tau}(A_I)).$$

Finally using the fact that $\mathcal{H}^n(X^I_{t,\tau}(A_I)) = \mathcal{H}^n(X_{t,\tau}(A))$ and the Hausdorff convergence we obtain

$$\begin{aligned}
\mathcal{H}^n(X_{t,\tau+\delta}(A)) &\geq \limsup_{I\to\infty} \mathcal{H}^n(X^I_{t,\tau+\delta}(A_I)) \\
&\geq \limsup_{I\to\infty} \left(\frac{1}{2}\right)^n \left(\frac{t-(\tau+\delta)}{t-\tau}\right)^n \mathcal{H}^n(X^I_{t,\tau}(A_I)) \\
&= \left(\frac{1}{2}\right)^n \left(\frac{t-(\tau+\delta)}{t-\tau}\right)^n \mathcal{H}^n(X_{t,\tau}(A)).
\end{aligned}$$

Hence the thesis is proved. □

2.2.5 Conclusion of the proof

The previous lemmas allow us to prove the following one.

We will denote the Cantor part of $D^2_x u(t,\cdot)$ with $D^2_c u(t,\cdot)$.

Lemma 2.21. *For ε small enough, for any t in $(\tau,\tau+\varepsilon]$ such that $|D^2_c u(t,\cdot)|(\Omega_t) > 0$ and δ in $(0,\tau+\varepsilon-t]$, there exists a Borel set $A \subset \Omega_t$ such that*

i) $\mathcal{H}^n(A) = 0$, $|D^2_c u(t,\cdot)|(A) > 0$ *and* $|D^2_c u(t,\cdot)|(\Omega_t \setminus A) = 0$;

ii) $X_{t,\tau}$ *is single-valued on* A*;*

iii) and

$$\chi_{t,\tau}(A) \cap \chi_{t+\delta,\tau}(\Omega_{t+\delta}) = \emptyset.$$

Proof. From Proposition 1.15 and the definition of Cantor part of a measure, there exists a Borel set A such that

- $D^+_x u(t,x)$ is single-valued for every $x \in A$,
- $\mathcal{H}^n(A) = 0$,
- $|D^2_c u(t,\cdot)|(\Omega_t \setminus A) = 0$ and $|D^2_c u(t,\cdot)|(A) > 0$.

By contradiction suppose there exists a compact set $K \subset A$ such that

$$|D_c^2 u(t,\cdot)|(K) > 0$$

and

$$X_{t,\tau}(K) = \chi_{t,\tau}(K) \subset \chi_{t+\delta,\tau}(\Omega_{t+\delta}).$$

Call $\omega := |D_c^2 u(t,\cdot)|(K)$.

Then there exists a Borel set $\tilde{K} \subset \Omega_{t+\delta}$ such that $\chi_{t,\tau}(K) = \chi_{t+\delta,\tau}(\tilde{K})$. Moreover, thanks to the fact that we are considering classical characteristics starting from $\tilde{K}$, we have

$$\chi_{t+\delta,t}(\tilde{K}) = K \quad \text{and} \quad \chi_{t+\delta,s}(\tilde{K}) = \chi_{t,s}(K) \;\; \forall s \in [\tau, t).$$

Using Lemma 2.19, for any $s \in [\tau, t)$,

$$\begin{aligned} \mathcal{H}^n(K) &= \mathcal{H}^n(X_{t+\delta,t}(\tilde{K})) \geq \left(\frac{1}{2}\right)^n \left(\frac{\delta}{t+\delta-s}\right)^n \mathcal{H}^n(X_{t+\delta,s}(\tilde{K})) \\ &= \left(\frac{1}{2}\right)^n \left(\frac{\delta}{t+\delta-s}\right)^n \mathcal{H}^n(X_{t,s}(K)). \end{aligned}$$

Hence

$$\mathcal{H}^n(K) \geq \left(\frac{1}{2}\right)^n \left(\frac{\delta}{t+\delta-s}\right)^n \mathcal{H}^n(X_{t,s}(K)). \tag{2.16}$$

Moreover if we choose s such that $t-s$ is small enough

$$\begin{aligned} \mathcal{H}^n(X_{t,s}(K)) &\geq C_1 \mathcal{H}^n(K) - C_2(t-s) \int_K d\Delta_s u(t,\cdot) + O((t-s)^2) \\ &\geq -C_2(t-s) \int_K d\Delta_c u(t,\cdot) + O((t-s)^2) \\ &\geq +C_2(t-s)\omega + O((t-s)^2) \\ &\geq \frac{C_2}{2}\omega^2, \end{aligned}$$

where we have used the fact that $\mathcal{H}^n(K) = 0$, that $\Delta_j u(t,K) \leq 0$, which is true due to semi-concavity, implies $\Delta_s u(t,K) \leq \Delta_c u(t,K)$, and $-\Delta_c u(t,K) \geq |D_c^2 u(t,\cdot)(K)| = \omega$. Thus

$$\mathcal{H}^n(X_{t,s}(K)) \geq \frac{C_2}{2}\omega^2. \tag{2.17}$$

Combining (2.16) with (2.17) we obtain

$$\mathcal{H}^n(K) \geq \left(\frac{1}{2}\right)^n \left(\frac{\delta}{t+\delta-s}\right)^n \frac{C_2}{2}\omega^2 > 0.$$

This is in contradiction with our hypothesis. □

We now have all the necessary Lemmas to prove the Theorem 2.1.

Proof. For $\varepsilon > 0$ sufficiently small such that Lemmas 2.15, 2.18, 2.19, and 2.21 hold, consider the functional F defined in (2.8) over the interval $[\tau, \tau + \varepsilon]$. F is bounded, and, from Lemma 2.15, F is a monotone function. Thus its points of discontinuity are at most countable.

We will prove that the presence of a Cantor part at a time t is related to a discontinuity of the functional F in t, hence there must be only a countable number of t's in $[\tau, \tau + \varepsilon]$ for which there is a Cantor part.

Suppose there exists a t in $(\tau, \tau + \varepsilon)$ such that

$$|D_c^2 u(t, \Omega_t)| > 0,$$

then for any $\delta > 0$ let A be the set of Lemma 2.21. Using Lemma 2.21-(iii) we get

$$F(t + \delta) \leq F(t) - \mathcal{H}^n(X_{t,\tau}(A)) \tag{2.18}$$

To compute $\mathcal{H}^n(X_{t,\tau}(A))$ call $\omega := |D_c^2 u(t, \cdot)|(A)$. As we saw in the previous lemma, if we choose $s \in [\tau, t)$ such that $t - s$ is small enough, we have

$$\mathcal{H}^n(X_{t,s}(A)) \geq \frac{C_2}{2}\omega^2.$$

Moreover for Lemma 2.19

$$\mathcal{H}^n(X_{t,\tau}(A)) \geq \left(\frac{1}{2}\right)^n \left(\frac{t-\tau}{t-s}\right)^n \mathcal{H}^n(X_{t,s}(A)).$$

Hence

$$\mathcal{H}^n(X_{t,\tau}(A)) \geq \left(\frac{1}{2}\right)^n \left(\frac{t-\tau}{t-s}\right)^n \frac{C_2}{2}\omega^2 \geq C\omega^2.$$

We can now use this estimate in (2.18) obtaining

$$F(t + \delta) \leq F(t) - C\omega^2.$$

Letting $\delta \to 0$

$$\limsup_{\delta \to 0} F(t + \delta) < F(t).$$

Therefore t is a point of discontinuity for F, as needed.

As already noticed this concludes the proof of our theorem, since F can have only a countable number of points of discontinuity. $\square$

Chapter 3

Jacobian's regularity

Let us consider a viscosity solution u of the Hamilton-Jacobi equation

$$\partial_t u + H(D_x u) = 0$$

with bounded Lipschitz initial datum $u_0(x)$ and uniformly convex Hamiltonian H.

Taken a Borel set $B \subset \Omega_t$ we call Jacobian the measure $J(t, \cdot)$ defined as

$$J(t, B) := \mathcal{H}^n(D_x^+ u(t, B)).$$

Since $D_x u(t, \cdot)$ is SBV out of a countable number of t's, the Jacobian J cannot have a positive part between $\mathcal{H}^n$ and $\mathcal{H}^{n-1}$, i.e., out of a countable number of t's, the measure $J(t, \cdot)$ has only an absolute continuous part with respect to $\mathcal{H}^n$ and a part which is concentrated on a $\mathcal{H}^{n-1}$-rectifiable set and is absolute continuous with respect to $\mathcal{H}^{n-1}$.

One can wonder if the Jacobian has only integer parts, that is, out of a countable number of t's, $J(t, \cdot)$ can have only parts which are concentrated on a $\mathcal{H}^k$-rectifiable set and are absolute continuous with respect to $\mathcal{H}^k$ for $k \in \{0, 1, \ldots, n\}$.

The following counterexample shows that this cannot be true.

Example 3.1. Let $V : \mathbb{R} \to [0, 1]$ be the Vitali function, and consider the concave function defined for $s \in [0, +\infty), y \in \mathbb{R}$, α, β positive constants

$$v(s, y) = -\alpha \int_0^y V(z)dz - \beta s.$$

Note that the y-derivative of this function is precisely the Vitali function, hence a function which sends the Cantor set $C \subset [0, 1]$, a set of positive $\mathcal{H}^{\frac{\log 2}{\log 3}}$-measure, in a set of positive $\mathcal{H}^1$-measure, precisely $\mathcal{H}^1(V(C)) = 1$. Moreover its derivative is a measure which gives a positive value to the Cantor set, $V'(C) = 1$.

We construct a viscosity solution of the two-dimensional Hamilton-Jacobi equation

$$\partial_t u + \frac{1}{2}|D_x u|^2 = 0, \tag{3.1}$$

whose behavior on vertical sections is exactly the behavior of the function v, i.e. for any fixed x_1 $u(\cdot, (x_1, \cdot)) = v(\cdot, \cdot)$.

Note that in this case the function

$$X_{t,0}(x) := x - tH_p(D_x^+ u(t,x))$$

defined on Ω_t is precisely

$$X_{t,0}(x) = x - tD_x^+ u(t,x).$$

Hence $J(t,x) = \mathcal{H}^n\left(\frac{x - X_{t,0}(x)}{t}\right)$.

Let us construct the initial datum for our viscosity solution as

$$u(0,(y_1,y_2)) := \max_{s\in[0,+\infty), y\in\mathbb{R}} \left\{ v(s,y) - \frac{y_1^2 + (y-y_2)^2}{2s} \right\}.$$

We are looking for an explicit form of this function so that it can be easily taken as the initial datum of our Hamilton-Jacobi equation.

To find a maximizer we take the derivatives with respect to s and y and equal them to 0, finding

$$-\beta + \frac{y_1^2 + (y-y_2)^2}{2s^2} = 0, \quad -\alpha V(y) - \frac{y - y_2}{s} = 0,$$

from which we deduce that (s,y) can be a maximizer for $u(0,(y_1,y_2))$ if these relations are invertible

$$y_2 = y + \alpha s V(y), \quad y_1 = \pm s\sqrt{2\beta - \alpha^2 V(y)^2}. \tag{3.2}$$

This is possible if $\beta \geq \frac{\alpha^2}{2}$. In this case the above system admits a solution, i.e. for every (y_1,y_2) one can find (s,y) that solve the system.

Let us verify that s,y are a true maximizer for our function

$$F(s,y) := v(s,y) - \frac{y_1^2 + (y-y_2)^2}{2s}.$$

The Jacobian matrix of $F(s,y)$ evaluated at s,y is

$$\begin{bmatrix} -\frac{y_1^2+(y-y_2)^2}{s^3} & \frac{y-y_2}{s^2} \\ \frac{y-y_2}{s^2} & -\alpha V'(y) - \frac{1}{s} \end{bmatrix} = \begin{bmatrix} -\frac{2\beta}{s} & -\frac{\alpha V(y)}{s} \\ -\frac{\alpha V(y)}{s} & -\alpha V'(s) - \frac{1}{s} \end{bmatrix}.$$

For $\beta \geq \frac{\alpha^2}{2}$, it has exactly two negative eigenvalues, thus (s,y) as in the system (3.2) is the unique global maximizer for $u(0,(y_1,y_2))$.

For simplicity we set $\beta = 1$, $\alpha = 1$.

With this parametrization the initial datum takes the form

$$u\Big(0, \Big(\pm s\sqrt{2 - V(y)^2}, y + sV(y)\Big)\Big) = -\int_0^y V(z)dz - 2s.$$

We can now recover the unique viscosity solution of the Hamilton-Jacobi equation with this initial datum through the Hopf-Lax formula

$$u(t,(x_1,x_2)) = \min_{(y_1,y_2)\in\mathbb{R}^2} \left\{ u(0,(y_1,y_2)) + \frac{(x_1-y_1)^2 + (x_2-y_2)^2}{2t} \right\}, \tag{3.3}$$

which, for $x_1 \geq 0$, is equivalent to

$$u(t,(x_1,x_2)) = \min_{s\in[0,+\infty), y\in\mathbb{R}} \Big\{ -\int_0^y V(z)dz - 2s + \frac{(x_1 - s\sqrt{2-V(y)^2})^2 + (x_2 - y + sV(y))^2}{2t} \Big\}.$$

To find the minimizer and obtain an explicit formula for our solution, take the derivatives in s and y and let them be equal to zero. What we obtain is

$$-2 + \frac{1}{t}\Big[(s\sqrt{2-V(y)^2} - x_1)\sqrt{2-V(y)^2} + (y + sV(y) - x_2)V(y) \Big] = 0$$

and

$$-V(y) + \frac{1}{t}\Big[-(s\sqrt{2-V(y)^2} - x_1)\frac{V(y)V'(y)}{\sqrt{2-V(y)^2}} + (y + sV(y) - x_2)(1 + sV'(y)) \Big] = 0,$$

where $V'(y)$ has to be intended as the distributional derivative of V which is null for $\mathcal{H}^1$-a.e. y in $\mathbb{R}$.

Therefore taken s and y, with $V'(y) = 0$, they are a minimizer for $u(t,(x_1,x_2))$ where

$$x_2 = y + (s-t)V(y), \qquad x_1 = \pm(s-t)\sqrt{2-V(y)^2},$$

and

$$u\Big(t, \Big(\pm(s-t)\sqrt{2-V(y)^2}, y + (s-t)V(y)\Big)\Big) = -\int_0^y V(z)dz - 2(s-t).$$

The derivative of such a solution is SBV and its Jacobian $J(t,\cdot)$ has no positive part in the interval $(1,2)$ out of a countable number of t's.

Let us see that $J(t,\cdot)$ has a positive part between $\mathcal{H}^1$ and $\mathcal{H}^0$ for every $t > 0$.

Taking $x_1 = 0$ and $x_2 = y$ in the Cantor set C, i.e. $(x_1,x_2) \in E := \{0\} \times C$, we compute the minimizers for $(t,(0,y))$ in (3.3). From the previous computation

$$u(t,(0,y)) = -\int_0^y V(z)dz.$$

Hence the two minimizers are $\Big(y_1 = \pm t\sqrt{2-V(y)^2}, y_2 = y + tV(y)\Big)$.

Thus for every $t > 0$ and every $y \in C$

$$D_x^+ u(t,(0,y)) = [-\sqrt{2-V(y)^2}, \sqrt{2-V(y)^2}] \times \{V(y)\}.$$

Therefore, we have found a set E which is of positive $\mathcal{H}^{\frac{log2}{log3}}$-measure such that $J(t,x) = 0$ for every $x \in E$ and every $t > 0$. Moreover, call

$$A := \Big\{[-\sqrt{2-V(y)^2}, \sqrt{2-V(y)^2}] \times \{V(y)\} \,|\, y \in C\Big\} \subset \mathbb{R}^2,$$

then $A = D_x^+ u((0,E)$ and $J(t,E) = \mathcal{H}^2(A) > 0$.

Hence the Jacobian has a positive part between $\mathcal{H}^1$ and $\mathcal{H}^0$.

Observe also that the set

$$[-t\sqrt{2-V(y)^2}, t\sqrt{2-V(y)^2}] \times \{y + tV(y)\}$$

does not contain the set

$$[-t'\sqrt{2-V(y)^2}, t'\sqrt{2-V(y)^2}] \times \{y + t'V(y)\}$$

for $t' > t$.

Chapter 4

SBV-like regularity for Hamilton-Jacobi equations: the convex case

In this chapter we consider the Hamilton-Jacobi equation

$$\partial_t u + H(D_x u) = 0 \quad \text{in } \Omega \subset [0,T] \times \mathbb{R}^n,$$

where H is a smooth convex Hamiltonian. A viscosity solution of such an equation is locally Lipschitz but in general it doesn't have any additional regularity. As already seen in the uniformly convex case instead the viscosity solution u is semiconcave, therefore $D_x u(t,\cdot)$ belongs to BV and $D_x^2 u(t,\cdot)$ is a matrix of Radon measures. Moreover, in [9], Bianchini, De Lellis and Robyr proved the theorem here presented as Theorem 0.2, which states that $D_x u(t,\cdot)$ belongs to $[SBV(\Omega_t)]^n$, $\Omega_t := \{x \in \mathbb{R}^n | \ (t,x) \in \Omega\}$, out of a countable number of t's in $[0,T]$. More precisely $D_x^2 u(t,\cdot)$ can have Cantor part only for a countable number of t's in $[0,T]$.

When H is just convex, $D_x u(t,\cdot)$ looses its BV regularity, an example can be found in Remark 3.7 in Bianchini [8]. However, in this chapter, we show that an SBV-like regularity result can be proven for the vector field

$$d(t,x) := H_p(D_x u(t,x)),$$

defined on the set U of points (t,x) where $u(t,x)$ is differentiable in x. Here H_p is the gradient of the Hamiltonian $H(p)$. Indeed the divergence $\mathrm{div} d(t,\cdot)$ is in general a locally finite Radon measure. When the vector field $d(t,\cdot)$ is BV and suitable hypotheses are made on the Lagrangian L, the Legendre transform of H, the measure $\mathrm{div} d(t,\cdot)$ has Cantor part only for a countable number of t's in $[0,T]$.

More precisely let H be $C^2(\mathbb{R}^n)$, convex and such that $\lim_{|p|\to\infty} \frac{H(p)}{|p|} = +\infty$.
(HYP(0)) Suppose the vector field $d(t,\cdot)$ belongs to $[BV(\Omega_t)]^n$ for every $t \in [0,T]$.

Define V_{π_n} as

$$V_{\pi_n} := \{v \in \mathbb{R}^n | \ L(\cdot) \text{ is not twice differentiable in } v\},$$

and

$$\Sigma_{\pi_n} := \{(t,x) \in U | \ d(t,x) \in V_{\pi_n}\} \quad \text{and} \quad \Sigma_{\pi_n}^c := U \setminus \Sigma_{\pi_n}.$$

(HYP(n)) We suppose V_{π_n} to be contained in a finite union of hyperplanes Π_{π_n}.

For $j = n, \ldots, 3$ for every $(j-1)$-dimensional plane π_{j-1} in Π_{π_j}, let $L_{\pi_{j-1}} : \mathbb{R}^{j-1} \to \mathbb{R}$ be the $(j-1)$-dimensional restriction of L to π_{j-1} and

$$V_{\pi_{j-1}} := \{v \in \mathbb{R}^{j-1} | \ L_{\pi_{j-1}}(\cdot) \text{ is not twice differentiable in } v\}.$$

Define

$$\Sigma_{\pi_{j-1}} := \{(t,x) \in \Sigma_{\pi_j} | \ d(t,x) \in V_{\pi_j}\} \quad \text{and} \quad \Sigma^c_{\pi_{j-1}} := \Sigma_{\pi_j} \setminus \Sigma_{\pi_{j-1}}.$$

(HYP(j-1)) We suppose $V_{\pi_{j-1}}$ is contained in a finite union of $(j-2)$-dimensional planes $\Pi_{\pi_{j-1}}$, for every $\pi_{j-1} \in \Pi_{\pi_j}$.

Theorem 4.1. *Under the above assumptions (HYP(0)),(HYP(n)),...,(HYP(2)), the Radon measure* $\mathrm{div} d(t,\cdot)$ *has Cantor part on* Ω_t *only for a countable number of* t*'s in* $[0,T]$.

This result can be seen as the multi-dimensional version of Theorem 0.4 proved by Robyr (see [35] for its proof). Furthermore, we prove that in the one-dimensional case the BV regularity of $d(t,x)$, which was an hypothesis in the theorem of Robyr, follows automatically in the case of a convex smooth Hamiltonian.

The question on the SBV regularity of $d(t,\cdot)$ without any additional hypothesis is still open.

The chapter is organized as follows. In Section 4.1 we extend the definition of the vector field d to the all Ω, we prove that $\mathrm{div} d(t,\cdot)$ is a locally finite Radon measure on Ω_t, for all $t \in [0,T]$. In Section 4.2 we present the general strategy used to prove that $\mathrm{div} d(t,\cdot)$ has a Cantor part only for a countable number of t's in $[0,T]$. In Section 4.3 we study the one-dimensional case and we prove that $\mathrm{div} d(t,\cdot)$ belongs to $SBV(\Omega_t)$, out of a countable number of t's in $[0,T]$, without any additional hypothesis. In Section 4.4 we study the multi-dimensional case and prove Theorem 4.1. We also state some easy corollaries.

4.1 Extension and preliminary properties of the vector field d

We consider a viscosity solution u of the Hamilton-Jacobi equation

$$\partial_t u + H(D_x u) = 0 \quad \text{in } \Omega \subset [0,T] \times \mathbb{R}^n,$$

where H is $C^2(\mathbb{R}^n)$ convex and

$$\lim_{|p| \to \infty} \frac{H(p)}{|p|} = +\infty.$$

As already noticed, thanks to the time invariance of the equation and to Proposition 1.18, it is enough to consider the unique viscosity solution of the following Cauchy problem

$$\begin{cases} \partial_t u + H(D_x u) = 0 & \text{in } \Omega \subset [0,T] \times \mathbb{R}^n, \\ u(0,x) = u_0(x) & \text{for all } x \in \Omega_0, \end{cases}$$

where $u_0(x)$ is a bounded Lipschitz function on Ω_0.

The vector field $d(t,x) = H_p(D_x u(t,x))$ is well defined where $u(t,x)$ is differentiable in x, i.e. $\mathcal{H}^n$-a.e. on Ω_t, for every $t \in [0,T]$.

Thanks to the Lipschitz regularity of $u(t,\cdot)$ and the fact that H is smooth, the vector field $d(t,\cdot)$ belongs to $[L^\infty(\Omega_t)]^n$.

Moreover d is constant along optimal rays. Indeed, thanks to Theorem 1.31-(iii), we have

$$d(t,x) = d(s, x-(t-s)d(t,x))$$

for all $0 \le s \le t$.

A natural extension of d to Ω is $\mathcal{D}(\cdot) : \Omega \to \mathbb{R}^n$

$$\mathcal{D}(t,x) := \left\{ \frac{x-y}{t} \mid \; y \text{ is a minimum for } u^+_{t,0}(t,z) \right\},$$

where $u^+_{t,0}$ is the forward solution as in Definition 1.42.

$\mathcal{D}(t,x)$ is a multi-valued function which coincides with $d(t,x)$ in the points (t,x) where $u(t,x)$ is differentiable in x. Indeed, where $u(t,\cdot)$ is differentiable, $u(t,x) = u^+_{t,0}(t,x)$ and they both admit as unique minimizer $y = x - tH_p(D_x u(t,x))$ in Ω_0.

Following the results of Bianchini and Gloyer in [10], we can prove that $\mathcal{D}(t,x)$ has closed graph and thanks to the fact that $\mathcal{D}(t,x)$ is closed

$$\mathcal{D}(t,x) \subset \mathcal{D}(t,x') + B(0,\varepsilon)$$

for $x, x' \in \Omega_t$. Moreover $\mathcal{D}(t,x)$ is a Borel measurable function and $\mathrm{div} d(t,\cdot)$ a locally finite Radon measure. We repeat the proof for the reader's convenience.

Theorem 4.2. *For every $t \in (0,T]$, the divergence $\mathrm{div} d(t,\cdot)$ is a locally finite Radon measure with negative singular part.*

Proof. Consider an approximation of our vector field done by taking a dense sequence of points $\{y_i\}_{i=1}^\infty$ in Ω_0. Fix an integer $I > 0$, call $\Omega_0^I := \{y_i | \; i = 1, \dots, I\}$ and define for any $x \in \Omega_t$

$$u^+_I(t,x) := \min_{i \in I} \left\{ u^-_{t,0}(0,y_i) + tL\left(\frac{x-y_i}{t}\right)\right\},$$

where $u^-_{t,0}$ is the backward solution as in Definition 1.42.

Through this approximation the set Ω_t is split into at most I open regions Ω_t^i, $i = 1, \dots, I$, defined by

$$\Omega_t^i := \text{interior of } \{x \in \Omega_t | \; \exists y_i \text{ minimizer for } u^+_I(t,x)\},$$

together with the set

$$J_t^I := \bigcup_{i \neq j} \left(\overline{\Omega}_t^i \cap \overline{\Omega}_t^j \right)$$

of negligible $\mathcal{H}^n$-measure. Indeed, even for $u^+_I(t,\cdot)$ the set of points with more than one minimum is the set of points of non differentiability of $u^+_I(t,\cdot)$ and this set has $\mathcal{H}^n$-measure zero. We define the vector field d^I on Ω so that on each open set Ω_t^i

$$d^I(t,x) := \frac{x-y_i}{t}.$$

Using explicitly the definition of d^I and the fact that $\mathcal{H}^n(J_t^I) = 0$,

$$\mathrm{div} d^I(t,x) \leq \frac{n}{t}.$$

Thanks to the pointwise convergence of d^I to d

$$\mathrm{div} d(t,\cdot) - \frac{n}{t}\mathcal{H}^n \leq 0,$$

i.e. $\mathrm{div} d(t,\cdot) - \frac{n}{t}\mathcal{H}^n$ is a negative definite distribution, hence it is a locally finite Radon measure. Thus $\mathrm{div} d(t,\cdot)$ is itself a locally finite Radon measure.

Moreover

$$\mathrm{div} d(t,\cdot) \leq \frac{n}{t}\mathcal{H}^n,$$

implies that the singular part of this measure can be only negative. □

From now on we will denote $\mu(t,\cdot) := \mathrm{div} d(t,\cdot)$.

Since we have proven that $\mu(t,\cdot)$ is a locally finite Radon measure, it makes sense to ask whether is possible or not that $\mu(t,\cdot)$ has Cantor part for all t in $[0,T]$. Note that if a Cantor part is different from zero then it must be negative for Theorem 4.2.

4.2 General strategy

In order to prove that $\mu(t,\cdot)$ has Cantor part only for a countable number of t's, the general idea is now standard, see [4], [9] and Chapter 2.

We present this strategy in the form useful in our case. We reduce to a smaller interval $[\tau,T]$, for a fixed $\tau > 0$, and we construct, on this interval, a monotone bounded functional $F(t)$. Then, we relate the presence of a Cantor part for the measure $\mu(t,\cdot)$, for a certain t in $[\tau,T]$, with a jump of the functional F in t. Since this functional is bounded monotone it can have only a countable number of jumps. Thus, the Cantor part of $\mu(t,\cdot)$ can be different from zero only for a countable number of t's.

To define F we consider the following maps: $X_{t,\tau}(x) : \Omega_t \to \Omega_\tau$

$$X_{t,\tau}(x) := x - (t-\tau)\mathcal{D}(t,x),$$

and its restriction to the set U_t of points where $\mathcal{D}(t,x)$ is single-valued, $\chi_{t,\tau}(x) : U_t \to U_\tau$

$$\chi_{t,\tau}(x) := x - (t-\tau)d(t,x).$$

We will sometimes write $\chi_{t,\tau}(\Omega_t)$ for $\chi_{t,\tau}(U_t)$.

We define the functional $F : (\tau,T] \to \mathbb{R}$

$$F(t) := \mathcal{H}^n(\chi_{t,\tau}(U_t)).$$

The functional F is bounded, and, due to the fact that optimal rays do not intersect except than at time t or 0, F is a monotone decreasing functional.

In order to apply the strategy above we need two estimates of the following type:

i) For any Borel set $A \subset U_t$ for t in $(\tau, T]$

$$\mathcal{H}^n(X_{t,\tau}(A)) \geq C_1 \mathcal{H}^n(A) - (t-\tau) C_2 \mu(t, A), \tag{4.1}$$

where C_1, C_2 are fixed positive constants.

ii) For any Borel set $A \subset \Omega_t$, for t in $(\tau, T]$ and for every $0 \leq \delta \leq t - \tau$

$$\mathcal{H}^n(X_{t,\tau+\delta}(A)) \geq \left(\frac{t - (\tau+\delta)}{t-\tau} \right)^m \mathcal{H}^n(X_{t,\tau}(A)), \tag{4.2}$$

where $m \in \mathbb{N}, m > 0$ is fixed.

Indeed with the estimates above we can prove the following lemma.

Lemma 4.3. *For any t in $(\tau, T]$ such that $\mu_c(t, \Omega_t) < 0$ and δ in $(0, T-t]$, there exists a Borel set $A \subset U_t$ such that*

i) $\mathcal{H}^n(A) = 0$, $\mu_c(t, A) < 0$ *and* $\mu_c(t, \Omega_t \setminus A) = 0$;

ii) $X_{t,\tau}$ *is single-valued on A;*

iii) and

$$\chi_{t,\tau}(A) \cap \chi_{t+\delta,\tau}(\Omega_{t+\delta}) = \emptyset.$$

Proof. The set of points where $d(t, \cdot)$ is not single-valued, which coincides with the set of points where $u(t, \cdot)$ is differentiable, is an $\mathcal{H}^{n-1}$-rectifiable set, due to the Lipschitz regularity of $u(t, \cdot)$. Hence, the Radon measure $\mu(t, \cdot)$ has null Cantor part on it. This and the definition of Cantor part of a measure imply the existence of a Borel set A such that

- $d(t, x)$ is single-valued for every $x \in A$,
- $\mathcal{H}^n(A) = 0$,
- $\mu_c(\Omega_t \setminus A) = 0$ and $\mu_c(A) < 0$.

By contradiction suppose there exists a compact set $K \subset A$ such that

$$\mu_c(t, K) < 0$$

and

$$X_{t,\tau}(K) = \chi_{t,\tau}(K) \subset \chi_{t+\delta,\tau}(\Omega_{t+\delta}).$$

Then there exists a Borel set $\tilde{K} \subset \Omega_{t+\delta}$ such that $\chi_{t,\tau}(K) = \chi_{t+\delta,\tau}(\tilde{K})$. Moreover, thanks to the fact that we are considering optimal rays starting from $\tilde{K}$, we have

$$\chi_{t+\delta,t}(\tilde{K}) = K \quad \text{and} \quad \chi_{t+\delta,\tau}(\tilde{K}) = \chi_{t,\tau}(K).$$

Using the estimate (4.2),

$$\mathcal{H}^n(K) = \mathcal{H}^n(X_{t+\delta,t}(\tilde{K})) \geq \left(\frac{\delta}{t+\delta-\tau} \right)^m \mathcal{H}^n(X_{t+\delta,\tau}(\tilde{K})) = \left(\frac{\delta}{t+\delta-\tau} \right)^m \mathcal{H}^n(X_{t,\tau}(K)).$$

Hence
$$\mathcal{H}^n(K) \geq \left(\frac{\delta}{t+\delta-\tau}\right)^m \mathcal{H}^n(X_{t,\tau}(K)).$$
Moreover applying estimate (4.1)
$$\mathcal{H}^n(K) \geq \left(\frac{\delta}{t+\delta-\tau}\right)^m \left(C_1\mathcal{H}^n(K) - (t-\tau)C_2\mu(t,A)\right).$$
Since $\mathcal{H}^n(K) = 0$ we obtain $\mu(t,A) \geq 0$ in contrast with the fact that $\mu_c(t,A) < 0$. □

The estimate (4.1) and Lemma 4.3 lead us to the expected conclusion.
Suppose there exists a t in (τ, T) such that
$$\mu_c(t,\Omega_t) < 0,$$
then, for any $\delta > 0$, let A be the set of Lemma 4.3. According to Lemma 4.3-(iii) we have
$$F(t+\delta) \leq F(t) - \mathcal{H}^n(X_{t,\tau}(A)).$$
Moreover, the estimate (4.1) gives
$$F(t+\delta) \leq F(t) + (t-\tau)C_2\mu_c(t,A).$$
Hence, letting $\delta \to 0$, we obtain
$$\limsup_{\delta\to 0} F(t+\delta) < F(t).$$
Therefore t is a point of discontinuity for F, as we wanted to prove.

4.3 One-dimensional case

We first consider the one-dimensional case. In this case we don't need any further assumption on d or L to prove the following theorem.

Theorem 4.4. *The vector field $d(t,\cdot)$ belongs to $SBV(\Omega_t)$, out of a countable number of $t \in [0,T]$.*

In the uniformly convex case, Theorem 4.4 is a corollary of Theorem 0.1 of the Introduction proved by Ambrosio and De Lellis in [4].

Proof. Since we are in the one-dimensional case, $\text{div} d(t,x) = \frac{\partial}{\partial x} d(t,x)$. Hence, Theorem 4.2 implies that $d(t,x)$ belongs to $BV(\Omega_t)$, for every $t \in (0,T]$.

Moreover, $\mathcal{D}(t,\cdot)$ is semimonotone. Indeed, since we are following optimal rays for $u^+_{t,0}$, they do not intersect except than at time 0 or t. Thus for $x_1, x_2 \in \Omega_t$, $x_1 < x_2$ and $d_1 \in \mathcal{D}(t,x_1), d_2 \in \mathcal{D}(t,x_2)$, it must hold
$$x_1 - td_1 \leq x_2 - td_2,$$

otherwise the rays cross each other at a time $s \in (0, t)$. Hence the function $\frac{1}{t}x - \mathcal{D}(t, x)$ is monotone increasing and $\mathcal{D}(t, x)$ is semimonotone with constant $C = \frac{1}{t}$.

Let us consider the map $X_{t,\tau}$ for any $t \in (\tau, T]$, $\tau > 0$ fixed. The fact that we are in the one-dimensional case implies that for t, x, such that $\mathcal{D}(t, x)$ is multi-valued,

$$\mathcal{D}(t, x) = [d_1, d_2],$$

where $d_1, d_2 \in \mathbb{R}$ are the speeds of the optimal rays for $u(t, x)$. Indeed, for every $\bar{d}, \tilde{d} \in [d_1, d_2]$, the ray $[x, x - t\bar{d}]$ cannot cross $[x, x - t\tilde{d}]$, since they are straight lines starting in the same point. So they fill the triangle delimited by $[x, x - td_1]$, $[x, x - td_2]$. Moreover, optimal rays starting in other points cannot cross $[x, x - td_1]$ and $[x, x - td_2]$, at intermediate time, since they are optimal. Thus they cannot cross any other ray $[x, x - td]$, where $d \in [d_1, d_2]$. For this reason these rays are optimal for $u^+_{t,0}(t, x)$. Thus optimal rays for the forward solution completely fill the set $\{\Omega_s \mid s \in [0, t]\}$.

Remark 4.5. This argument holds also in the multi-dimensional case but only for a set of points of non differentiability of zero-dimension. The argument is not true in general when the points of non differentiability lie on a surface of dimension greater than zero, since rays starting in two different points of this surface can intersect even at intermediate times.

The above consideration ensures that the map $X_{t,\tau}$ is injective for $\tau > 0$, however this map is multi-valued. To recover the Lipschitzianity we use the Hille-Yosida transformation as seen in [1] and Chapter 2.

For any Borel set $A \subset \Omega_t$, let $z \in B := A + T(A)$, $T(x) := (Cx - \mathcal{D}(t, x))$ and $w(z) := (Id_1 + (T)^{-1})^{-1}(z)$. Then the following 1-Lipschitz transformations

$$\begin{cases} x(z) = z - w(z) \\ p(z) = Cz - (C + 1)w(z), \end{cases} \tag{4.3}$$

transform our graph

$$\{(x, p) |\ x \in A,\ p \in \mathcal{D}(t, x)\}$$

into the equivalent graph of a maximal monotone function

$$\{(z - w(z), Cz - (C + 1)w(z)) |\ z \in B\}.$$

Recall that C is the semimonotonicity constant of $\mathcal{D}(t, \cdot)$.

Following optimal rays starting in A with speed in $\mathcal{D}(t, A)$, we can now pass from $X_{t,\tau}(x)$ to a Lipschitz map defined on B

$$\xi(\tau, z) := z - w(z) - (t - \tau)(Cz - (C + 1)w(z)).$$

Note that

$$\{(Cz - (C + 1)w(z)) |\ z \in x + T(x)\} = \mathcal{D}(t, x)$$

so that $X_{t,\tau}(x) = \{\xi(\tau, z) |\ z \in x + T(x)\}$ and $X_{t,\tau}(A) = \xi(\tau, B)$.

We can now apply the Area Formula to $\xi(\tau, \cdot)$

$$\int_{\xi(\tau,B)} \mathcal{H}^0(\xi(\tau, \cdot)^{-1}(w))dw = \int_B |\xi_z(\tau, z)|dz. \tag{4.4}$$

Thanks to the injectivity of the map $X_{t,\tau}$, which is preserved when passing to the Lipschitz parametrization, the left term of (4.4) is precisely the measure of the set $\xi(\tau, B)$. Hence, we have

$$\int_{\xi(\tau,B)} \mathcal{H}^0(\xi(\tau,\cdot)^{-1}(w))dw = \mathcal{H}^1(\xi(\tau,B)) = \mathcal{H}^1(X_{t,\tau}(A)).$$

Moreover, differentiating ξ we respect to z we denote

$$\xi_z(\tau, z) = \xi_z(t, z) - (t-\tau)\dot{\xi}_z(t, z),$$

where $\xi_z(t, z) := \frac{\partial}{\partial z}(z - w(z))$ and $\dot{\xi}_z(t, z) := \frac{\partial}{\partial z}(Cz - (C+1)w(z))$.

Thus we have

$$\mathcal{H}^1(X_{t,\tau}(A)) = \int_B |\xi_z(t,z) - (t-\tau)\dot{\xi}_z(t,z)|dz \geq \int_B \xi_z(t,z)dz - (t-\tau)\int_B \dot{\xi}_z(t,z)dz.$$

Observing that

$$\int_B \dot{\xi}_z(t,z)dz = \int_B \frac{\partial}{\partial z}(Cz - (C+1)w(z))dz = \mu(t, A),$$

we have proven the following estimate: given a Borel set $A \subset \Omega_t$ for t in $(\tau, T]$, we have

$$\mathcal{H}^1(X_{t,\tau}(A)) \geq \mathcal{H}^1(A) - (t-\tau)\mu(t, A). \tag{4.5}$$

Moreover, since for every $0 \leq \delta \leq t - \tau$

$$\xi_z(t,z) - (t-(\tau+\delta))\dot{\xi}_z(t,z) = \frac{\delta}{t-\tau}\xi_z(t,z) + \frac{t-(\tau+\delta)}{t-\tau}(\xi_z(t,z) - (t-\tau)\dot{\xi}_z(t,z)),$$

and $\xi_z(t, z) > 0$, we have

$$\xi_z(t,z) - (t-(\tau+\delta))\dot{\xi}_z(t,z) \geq \frac{t-(\tau+\delta)}{t-\tau}(\xi_z(t,z) - (t-\tau)\dot{\xi}_z(t,z)).$$

Thus, integrating the last equation over B, we obtain the following estimate: given a Borel set $A \subset \Omega_t$ for t in $(\tau, T]$, then for every $0 \leq \delta \leq t - \tau$ we have

$$\mathcal{H}^1(X_{t,\tau+\delta}(A)) \geq \frac{t-(\tau+\delta)}{t-\tau}\mathcal{H}^1(X_{t,\tau}(A)). \tag{4.6}$$

The estimates (4.5) and (4.6) are of type (4.1) and (4.2) respectively, thus they are enough to prove the SBV regularity of d, as seen in Subsection 4.2.

□

4.4 The multi-dimensional case

In [9] Bianchini, De Lellis and Robyr proved that the estimates (4.1) and (4.2) hold for the uniformly convex Hamiltonian $H_\epsilon(p) := H(p) + \frac{\epsilon}{2}|p|^2$ for every $\varepsilon > 0$ in a small interval of time and with constants strictly depending on ϵ. Thus, the two estimates cannot pass to the limit.

Nevertheless, we can prove that the divergence div$d(t,\cdot)$ has Cantor part only for a countable number of t's, adding some hypothesis on the regularity of d and on the structure of the the set of points where L is not twice differentiable.

As already noticed, the Lagrangian corresponding to a smooth convex Hamiltonian is strictly convex but non smooth in general. Particular conditions on the set of points where L is not twice differentiable will allow us to reduce iteratively our problem to a problem of lower dimension, down to the one-dimensional case, where, as we have seen, SBV regularity can be proven without additional assumptions.

Before going on with the proof we set some notations. We will denote with $(x_1, x_2, \ldots, x_n)$ the components of the vector $x \in \mathbb{R}^n$ and, to contract the notation, for a fixed $j = 1, \ldots, n-1$ we call $\hat{x} \in \mathbb{R}^{n-j}$ the vector defined so that

$$(x_1, \ldots, x_j, \hat{x}) = (x_1, x_2, \ldots, x_n).$$

Given a set $E \subset [0,T] \times \mathbb{R}^n$ we will denote with

$$E_t := \{x \in \mathbb{R}^n |\ (t,x) \in E\}$$

and for $j = 1, \ldots, n-1$

$$E_{x_1,\ldots,x_j} := \Big\{(t, x_{j+1}, \ldots, x_n) |\ (t, x_1, \ldots, x_j, x_{j+1}, \ldots x_n) \in E\Big\}.$$

As before we will sometimes denote with $\mu(t,\cdot)$ the Radon measure div$d(t,\cdot)$ defined on Ω_t.

(HYP(0)) Suppose that the vector field $d(t,\cdot)$ belongs to $[BV(\Omega_t)]^n$ for any $t \in [0,T]$.

The measure div$d$ can have Cantor part only on a subset of the points of differentiability in x of $u(t,x)$, i.e. the points where $\mathcal{D}(t,x)$ is single-valued. Thus we can reduce to the study of our measure on the set

$$U := \Omega \setminus \{(t,x) |\ \mathcal{D}(t,x) \text{ is multi-valued}\}.$$

Call V the set of points where L is not twice differentiable:

$$V := \{v \in \mathbb{R}^n |\ L(\cdot) \text{ is not twice differentiable in } v\}.$$

Then the set U can be split into two subsets:

$$\Sigma := \{(t,x) \in U |\ d(t,x) \in V\} \quad \text{and} \quad \Sigma^c := U \setminus \Sigma.$$

(HYP(n)) Suppose V is contained in a finite union of hyperplanes.

Claim 1.(n) The vector field $d(t,\cdot)$ belongs to $[SBV(\Sigma_t^c)]^n$ out of a countable number of t's in $[0,T]$.

Claim 2.(n) The Radon measure div$d(t,\cdot)$, restricted to Σ_t, can have Cantor part only for a countable number of t's in $[0,T]$.

The regularity of div$d$ will follow from the previous claims and the fact that $U = \Sigma \cup \Sigma^c$.

Proof of Claim 1.(n). For a fixed $(\bar{t},\bar{x}) \in \Sigma^c$, the Hessian of L exists and is continuous in $\bar{v} := d(\bar{t},\bar{x})$. Thus there exist $r > 0$ and a $(n+1)$-dimensional ball $B_r^{n+1}(\bar{t},\bar{x}) \subset \Omega \setminus \Sigma$ where L and H are uniformly convex.

We can also find an open cone $C_{n+1}(\bar{t},\bar{x}) \subset B_r^{n+1}(\bar{t},\bar{x})$, properly containing $(\bar{t},\bar{x})$, over which an Hamilton-Jacobi equation can be solved. Indeed, we take an n-dimensional ball as base,

$$B^n \subset (B_r^{n+1}(\bar{t},\bar{x}))_{\bar{t}-\sigma} \subset (\Omega \setminus \Sigma)_{\bar{t}-\sigma},$$

for a certain $0 < \sigma < r$, and we fix the height of length $l \in \mathbb{R}$, $0 < l < 2r$. The height must be chosen according to the speed of propagation of the solution and such that $\bar{t} < \bar{t} - \sigma + l$.

Consider now the viscosity solution $\bar{u}$ of the Cauchy problem

$$\begin{cases} \partial_t \bar{u} + H(D_x \bar{u}) = 0 & \text{in } C_{n+1}(\bar{t},\bar{x}), \\ \bar{u}(t-\sigma,x) = u(t-\sigma,x)\mathbb{1}_{B^n}(x), & \end{cases}$$

where $\mathbb{1}_E(x)$ is the indicator function of the set E. Note that $u(t,x) = \bar{u}(t,x)$ on $C_{n+1}(\bar{t},\bar{x})$.

Thanks to the uniform convexity of H over $C_{n+1}(\bar{t},\bar{x})$, the main theorem of [9] ensures that the vector field

$$\bar{d}(t,\cdot) := H_p(D_x \bar{u}(t,\cdot))$$

is SBV out of a countable number of t's in $[\bar{t}-\sigma, \bar{t}-\sigma+l]$.

The vector fields $d(t,\cdot)$ and $\bar{d}(t,\cdot)$ are both BV and coincide on $(C_{n+1}(\bar{t},\bar{x}))_t$, thus, for Proposition 1.9,

$$D_x d(t,\cdot) = D_x \bar{d}(t,x).$$

Therefore $d(t,\cdot)$ belongs to $SBV((C_{n+1}(\bar{t},\bar{x}))_t)$ out of a countable number of t's in $[\bar{t}-\sigma, \bar{t}-\sigma+l]$.

Finally, using the fact that $\mathbb{R}^n$ is a countable union of bounded sets, we can apply Besicovitch covering Theorem, see [5], to prove that the set Σ^c can be fully covered by a countable number of cones C_{n+1}^i, for $i \in \mathbb{N}$, with the property stated above. Thus $d(t,\cdot)$ belongs to $[SBV(\Sigma_t^c)]^n$ out of a countable number of t's in $[0,T]$. □

We consider now the behavior of div$d$ on the set Σ. In order to prove Claim 2.(n), in the n-dimensional case, $n > 2$, we need some other hypothesis on L and its restriction to the set of points where L is not twice differentiable. No additional hypotheses are needed in the case $n = 2$.

Proof of Claim 2.(n). 2**-dimensional case.** First, suppose V is a single straight line. Without loss of generality we can fix $V = \{v \in \mathbb{R}^2 | \; v_1 = 0\}$.

Call $L_V : \mathbb{R} \to \mathbb{R}$ the restriction of the Lagrangian L to V,

$$L_V(v_2) := L(0, v_2)$$

for any $v_2 \in \mathbb{R}$. Call $I \subset \mathbb{R}$ the set of every x_1 in $\mathbb{R}$ such that Σ_{x_1} is non empty. Note that if $(t,x_2) \in \Sigma_{x_1}$ then $(0, x_2 - t d_2(t,(x_1,x_2)))$ belongs to Σ_{x_1} because $d(t,(x_1,x_2)) = (0, d_2(t,(x_1,x_2)))$.

For every $x_1 \in I$, we consider the one-dimensional Hamilton-Jacobi equation for the function $u_{x_1}(t,x_2)$.

$$\begin{cases} \partial_t u_{x_1} + H_V(D_{x_2} u_{x_1}) = 0 & \text{in } \Sigma_{x_1}, \\ u_{x_1}(0,x_2) = u(0,(x_1,x_2)) & \forall x_2 \in (\Sigma_{x_1})_0, \end{cases}$$

where $H_V(p)$ is the Hamiltonian associated to $L_V(v)$.

The viscosity solution $u_{x_1}(t,x_2)$ is equal to $u(t,(x_1,x_2))$ for every $(t,(x_1,x_2)) \in \Sigma$. Indeed

$$\begin{aligned} u_{x_1}(t,x_2) = &\min_{y_2\in\mathbb{R}}\left\{u(0,x_1,y_2)+tL_V\left(\frac{x_2-y_2}{t}\right)\right\} \\ = &\min_{y_2\in\mathbb{R}}\left\{u(0,x_1,y_2)+tL\left(0,\frac{x_2-y_2}{t}\right)\right\} \\ = &\,u(t,(x_1,x_2)), \end{aligned}$$

where the last equality follows from the fact that, for (t,x) in Σ, the unique minimizer in the representation formula (1.10) is $y=(x_1-td_1(t,x),x_2-td_2(t,x))$ and $d(t,x)=(0,d_2(t,x))$.

Let us define as usual

$$d_{x_1}(t,x_2) := (H_V)_{p_2}(D_{x_2}u_{x_1}(t,x_2))$$

and

$$\mu_{x_1}(t,\cdot) := \frac{\partial}{\partial x_2}d_{x_1}(t,\cdot).$$

The vector field $d_{x_1}(t,\cdot)$ is one-dimensional. Hence, for Theorem 4.2, $d_{x_1}(t,\cdot)$ belongs to $BV((\Sigma_{x_1})_t)$ for any $x_1\in I$, for any $t\in[0,T]$.

On the set $\Sigma\subset U$, the matrix of Radon measures D_xd has no jump part. Moreover, since Σ_t is contained on the set $\{x|\ d_1(t,x)=0\}$ and $d(t,\cdot)$ is BV, Proposition 1.10 implies

$$\frac{\partial}{\partial x_1}d_1(t,\Sigma_t)=0 \quad \text{and} \quad \frac{\partial}{\partial x_2}d_1(t,\Sigma_t)=0.$$

Therefore

$$\mathrm{div} d(t,\cdot)=\frac{\partial}{\partial x_2}d_2(t,\cdot) \quad \text{on } \Sigma_t.$$

For every $(t,x)\in\Sigma$, $u_{x_1}(t,x_2)=u(t,(x_1,x_2))$ implies

$$d_2(t,x)=d_{x_1}(t,x_2).$$

The vector field $d_2(t,(x_1,\cdot))$ is a one-dimensional restriction of $d_2(t,\cdot)$ thus, for Proposition 1.11, belongs to $BV((\Sigma_{x_1})_t)$ for $\mathcal{H}^1$-a.e. $x_1\in I$. Since even $d_{x_1}(t,\cdot)$ is BV on $(\Sigma_{x_1})_t$, Proposition 1.9 implies

$$\frac{\partial}{\partial x_2}d_2(t,(x_1,\cdot))=\frac{\partial}{\partial x_2}d_{x_1}(t,\cdot)$$

for $\mathcal{H}^1$-a.e. $x_1\in I$. Therefore taken a Borel set $A\subset\Sigma_t$ and any $\phi\in C_c^\infty(\Sigma_t)$,

$$\int_A \phi(x)d\mu(t,x)=\int_I\int_{A_{x_1}}\phi(x)d\mu_{x_1}(t,x_2)dx_1.$$

Thanks to the convexity of L_V, we can apply Theorem 4.4 to $\mu_{x_1}(t,\cdot)$ and obtain the following estimates.

For any $\tau>0$, let A be a Borel set in Σ_t, for $t\in(\tau,T]$. Then for any $0\le\delta\le t-\tau$, and every section A_{x_1}, for $x_1\in I$, we have

$$\mathcal{H}^1(X^{x_1}_{t,\tau}(A_{x_1})) \geq \mathcal{H}^1(A_{x_1}) - (t-\tau)\mu_{x_1}(t, A_{x_1}),$$

$$\mathcal{H}^1(X^{x_1}_{t,\tau+\delta}(A_{x_1})) \geq \frac{t-(\tau+\delta)}{t-\tau}\mathcal{H}^1(X^{x_1}_{t,\tau}(A_{x_1})).$$

Here we denote with $X^{x_1}_{t,\tau}(x_2)$ the one-dimensional map defined on $(\Sigma_{x_1})_t$

$$X^{x_1}_{t,\tau}(x_2) := x_2 - (t-\tau)d_{x_1}(t, x_2).$$

The corresponding 2-dimensional map

$$X_{t,\tau}(x) := x - (t-\tau)d(t,x),$$

reduces to

$$X_{t,\tau}(x) = (x_1, X^{x_1}_{t,\tau}(x_2))$$

for every $x \in \Sigma_t$.

We can integrate the previous estimates with respect to $\mathcal{H}^1$ on $I \subset \mathbb{R}$ to recover estimates of type (4.1) and (4.2).
For any $\tau > 0$, given a Borel set $A \subset \Sigma_t$, for t in $(\tau, T]$, we have

$$\mathcal{H}^2(X_{t,\tau}(A)) \geq \mathcal{H}^2(A) - (t-\tau)\mu(t, A). \tag{4.7}$$

For any $\tau > 0$, given a Borel set $A \subset \Sigma_t$, for t in $[\tau, T]$ and $0 \leq \delta \leq t-\tau$ we have

$$\mathcal{H}^2(X_{t,\tau+\delta}(A)) \geq \frac{t-(\tau+\delta)}{t-\tau}\mathcal{H}^2(X_{t,\tau}(A)). \tag{4.8}$$

Thus the strategy seen in the Subsection 4.2 can be easily applied to prove that $\mu(t,\cdot)$, restricted to Σ_t, can have Cantor part only for a countable number of t's in $[0, T]$.

Remark 4.6. Note that in this case nothing can be said about the Cantor part of $\frac{\partial}{\partial x_1}d_2(t,\cdot)$. Thus we cannot say that $d(t,\cdot)$ belongs to $[SBV(\Omega_t)]^2$.

Consider now the case in which V consists of a finite number of straight lines. When we consider $\mu(\cdot,\cdot)$ restricted to the points of Σ such that $d(t,x)$ belongs only to a part of one of the straight lines, we can apply the considerations done in the case where V consists only of a single straight line. On the other hand, when we consider $\mu(\cdot,\cdot)$ restricted to the points of Σ such that $d(t,x)$ belongs to an intersection point (v_1, v_2) of two, or more, straight lines, the divergence div$d(t,\cdot)$ must be zero on every Borel subset of $\{x|\ d_1(t,x) = v_1, d_2(t,x) = v_2\}$, for Proposition 1.10. Thus the measure $\mu(t,\cdot)$, restricted to Σ_t, can have Cantor part only for a countable number of t's in $[0,T]$ even when V consists of a finite number of straight lines. The case in which V is contained in a finite number of straight lines is analogous.

***n*-dimensional case.** We prove the claim iterating a subdivision of Σ down to the dimension one.

Call $V_n := V$. At the step $n-j$, for $j = n, \dots, 3$, we first suppose that V_j consists of a single $(j-1)$-dimensional plane, without loss of generality we can fix

$$V_j = \{v \in \mathbb{R}^n|\ v_1 = 0, \dots, v_{n+1-j} = 0\}.$$

Call $L_{V_j} : \mathbb{R}^{j-1} \to \mathbb{R}$ the restriction of $L_{V_{j+1}}$ to V_j,

$$L_{V_j}(\hat{v}) := L_{V_{j+1}}(0, \hat{v}) = L(0, \dots, 0, \hat{v})$$

for any $\hat{v} \in \mathbb{R}^{j-1}$.

(HYP(j-1)) We require that the restriction L_{V_j} is twice $(j-1)$-differentiable out of the set V_{j-1},

$$V_{j-1} := \{\hat{v} \in \mathbb{R}^{j-1} |\ L_{V_j}(\cdot) \text{ is not twice differentiable in } \hat{v}\},$$

and V_{j-1} is contained in a finite number of $(j-2)$-dimensional planes.

Then we can subdivide Σ_j into two set:

$$\Sigma_{j-1} := \{(t,x) \in \Sigma_j |\ d(t,x) \in V_{j-1}\} \quad \text{and} \quad \Sigma^c_{j-1} := \Sigma_j \setminus \Sigma_{j-1}.$$

Thus, at every step, we have to prove the following claims.

Claim 1.(j-1) The Radon measure div$d(t,\cdot)$, restricted to $(\Sigma^c_{j-1})_t$, can have Cantor part only for a countable number of t's in $[0,T]$.

Claim 2.(j-1) The Radon measure div$d(t,\cdot)$, restricted to $(\Sigma_{j-1})_t$, can have Cantor part only for a countable number of t's in $[0,T]$.

Proof of Claim 1.(j-1) . We will prove it for $j = n$, in the other cases the proof is similar.

For a fixed $(\bar{t}, \bar{x}) \in \Sigma^c_{n-1}$, the Hessian of L_V exists and is continuous in $\hat{v} := (d_2(\bar{t},\bar{x}), \dots, d_n(\bar{t},\bar{x})) \in \mathbb{R}^{n-1}$. Thus there exist $r > 0$ and a $(n+1)$-dimensional ball $B^{n+1}_r(\bar{t},\bar{x}) \subset \Omega \setminus \Sigma_{n-1}$ where L_V and H_V are uniformly convex.

We can also find, as we did in the proof of Claim 1.(n), an open cone $C_{n+1}(\bar{t},\bar{x}) \subset B^{n+1}_r(\bar{t},\bar{x})$ of height $[\bar{t}-\sigma, \bar{t}-\sigma+l]$, for a certain $0 < \sigma < r$, $\bar{t} < \bar{t}-\sigma+l$ and base B^n, which contains properly $(\bar{t},\bar{x})$. On every section $(C_{n+1}(\bar{t},\bar{x}))_{x_1}$, for every $x_1 \in I := \{z \in \mathbb{R} |\ (C_{n+1}(\bar{t},\bar{x}))_z \neq \emptyset\}$, we can consider the viscosity solution $\bar{u}_{x_1}$ of the $(n-1)$-dimensional Hamilton-Jacobi equation

$$\begin{cases} \partial_t \bar{u}_{x_1} + H_V(D_{\hat{x}} \bar{u}_{x_1}) = 0 & \text{in } (C_{n+1}(\bar{t},\bar{x}))_{x_1}, \\ \bar{u}_{x_1}(\bar{t}-\sigma, \hat{x}) = u(\bar{t}-\sigma, x)\mathbb{1}_{B^n}(x). \end{cases}$$

As usual we define

$$\bar{d}_{x_1}(t,\hat{x}) := (H_V)_{\hat{p}}(D_{\hat{x}} \bar{u}_{x_1}(t,\hat{x})).$$

and

$$\bar{\mu}_{x_1}(t,\cdot) := \mathrm{div}_{n-1} \bar{d}_{x_1}(t,\cdot).$$

The vector field $\bar{d}_{x_1}(t,\cdot)$ belongs to $[BV(((C_{n+1}(\bar{t},\bar{x}))_{x_1})_t)]^{n-1}$ for any $x_1 \in I$, for any $t \in [\bar{t}-\sigma, \bar{t}-\sigma+l]$. Indeed in every $(C_{n+1}(\bar{t},\bar{x}))_{x_1}$ H_V is uniformly convex.

Since we have a uniform convexity constant for H_V, which holds on every $(C_{n+1}(\bar{t},\bar{x}))_{x_1}$, for $x_1 \in I$, we can arrange l small enough, eventually subdividing the cone, so that the following two estimates hold with uniform constants $C_1, C_2 > 0$, which do not depend on x_1.

Let $\bar{t}-\sigma < \tau < \bar{t}-\sigma+l$, let A be a Borel set in $(C_{n+1}(\bar{t},\bar{x}))_t$, for t in $[\tau, \bar{t}-\sigma+l]$. Then, for any $0 \leq \delta \leq t-\tau$ and every set A_{x_1}, for $x_1 \in I$, we have

$$\mathcal{H}^{n-1}(\bar{X}^{x_1}_{t,\tau}(A_{x_1})) \geq C_1 \mathcal{H}^{n-1}(A_{x_1}) - (t-\tau) C_2 \bar{\mu}_{x_1}(t, A_{x_1}),$$

$$\mathcal{H}^{n-1}(\bar{X}^{x_1}_{t,\tau+\delta}(A_{x_1})) \geq \left(\frac{t-(\tau+\delta)}{t-\tau}\right)^{n-1} \mathcal{H}^{n-1}(\bar{X}^{x_1}_{t,\tau}(A_{x_1})).$$

Here the $(n-1)$-dimensional map $\bar{X}^{x_1}_{t,\tau}(\hat{x})$ is defined

$$\bar{X}^{x_1}_{t,\tau}(\hat{x}) := \hat{x} - (t-\tau)\bar{d}_{x_1}(t,\hat{x}).$$

Consider now the vector field d.

On the set $C_{n+1}(\bar{t},\bar{x}) \subset U$, the matrix of Radon measures $D_x d$ has no jump part. Moreover, since $(C_{n+1}(\bar{t},\bar{x}))_t$ is contained on the set $\{x|\ d_1(t,x)=0\}$ and $d(t,\cdot)$ is BV, Proposition 1.10 implies

$$\frac{\partial}{\partial x_j} d_1(t,(C_{n+1}(\bar{t},\bar{x}))_t) = 0 \quad \text{ for } j=1,\dots,n.$$

Therefore

$$\mathrm{div} d(t,\cdot) = \mathrm{div}_{n-1}\hat{d}(t,\cdot) \quad \text{on } (C_{n+1}(\bar{t},\bar{x}))_t,$$

$\hat{d}(t,x) := (d_2(t,x),\dots,d_n(t,x))$.

For every $(t,x) \in C_{n+1}(\bar{t},\bar{x})$, $u_{x_1}(t,x_2) = u(t,(x_1,x_2))$ implies

$$\hat{d}(t,x) = d_{x_1}(t,\hat{x}).$$

The vector field $\hat{d}(t,x_1,\cdot)$, being a $(n-1)$-dimensional section of the BV vector field $d(t,\cdot)$, belongs, for Proposition 1.11, to $[BV((\Sigma_{n-1})_{x_1})]^{n-1}$ for $\mathcal{H}^1$-a.e. x_1 such that $(\Sigma_{n-1})_{x_1}$ is non empty.

Since even $\bar{d}_{x_1}(t,\cdot)$ is BV on $(C_{n+1}(\bar{t},\bar{x}))_t$, Proposition 1.9 implies

$$\mathrm{div}_{n-1}\hat{d}(t,(x_1,\cdot)) = \mathrm{div}_{n-1}\bar{d}_{x_1}(t,\cdot)$$

for almost every x_1 such that $(\Sigma_{n-1})_{x_1}$ is non empty. Therefore taken a Borel set $A \subset (C_{n+1}(\bar{t},\bar{x}))_t$ and any $\phi \in C^\infty_c((C_{n+1}(\bar{t},\bar{x}))_t)$,

$$\int_A \phi(x)d\mu(t,x) = \int_I \int_{A_{x_1}} \phi(x)d\bar{\mu}_{x_1}(t,\hat{x})dx_1.$$

Moreover, for every $x \in (C_{n+1}(\bar{t},\bar{x}))_t$

$$X_{t,\tau}(x) = x - (t-\tau)d(t,x) = (x_1, \bar{X}^{x_1}_{t,\tau}(\hat{x})).$$

The uniformity on every A_{x_1} allow us to integrate with respect to $\mathcal{H}^1$, over the set I, to obtain the following estimates.

Let $\bar{t}-\sigma < \tau < t$, let A be a Borel set in $(C_{n+1}(\bar{t},\bar{x}))_t$, for t in $[\bar{t}-\sigma, \bar{t}-\sigma+l]$. Then for any $0 \leq \delta \leq t-\tau$, it holds

$$\mathcal{H}^n(X_{t,\tau}(A)) \geq C_1\mathcal{H}^n(A) - (t-\tau)C_2\mu(t,A),$$

$$\mathcal{H}^n(X_{t,\tau+\delta}(A)) \geq \left(\frac{t-(\tau+\delta)}{t-\tau}\right)^{n-1} \mathcal{H}^n(X_{t,\tau}(A)).$$

Therefore, repeating the standard procedure seen in Subsection 4.2, we can prove that $\mu(t,\cdot) := \mathrm{div} d(t,\cdot)$ has Cantor part only for a countable number of t's in $[\bar{t}-\sigma, \bar{t}-\sigma+l]$.

Finally, using again Besicovitch Theorem, the set Σ^c_{n-1} can be fully covered by a countable number of cones C^i_{n+1} for $i \in \mathbb{N}$ with the property stated above. Thus the Radon measure $\mathrm{div} d(t,\cdot)$ can have Cantor part on $(\Sigma^c_{n-1})_t$ only for a countable number of t's in $[0,T]$. □

We iterate the procedure subdividing Σ_{j-1} in Σ_{j-2} and Σ^c_{j-2}. Hence to prove Claim 2.(j-1) is enough to prove Claim 2.(2), i.e. for $j=3$.

Claim 2.(2) The Radon measure $\mathrm{div} d(t,\cdot)$, restricted to $(\Sigma_2)_t$, can have Cantor part only for a countable number of t's in $[0,T]$.

Proof. The proof is equal to the one done in the 2-dimensional case. We rewrite it with the notation which applies in this case.

First, suppose V_2 is a single straight line. Without loss of generality we can fix

$$V_2 = \{v \in \mathbb{R}^n |\ v_1 = 0, \dots, v_{n-1} = 0\}.$$

Recall that V_2 is a straight line in $V_3 = \{v \in \mathbb{R}^n |\ v_1 = 0, \dots, v_{n-2} = 0\}$.

Call $L_{V_2} : \mathbb{R} \to \mathbb{R}$ the restriction of the Lagrangian L_{V_3} to V_2,

$$L_{V_2}(v_n) := L_{V_3}(0, v_n) = L(0, \dots, 0, v_n)$$

for any $v_n \in \mathbb{R}$. For $i = 1 \dots, n-1$, call $I_i \subset \mathbb{R}$ the set of every x_i in $\mathbb{R}$ such that $(\Sigma_2)_{x_i}$ is non empty and $I := I_1 \times \cdots \times I_{n-1} \subset \mathbb{R}^{n-1}$.

For every $(x_1, \dots, x_{n-1}) \in I$, we consider the one-dimensional Hamilton-Jacobi equation for the function $u_{x_1,\dots,x_{n-1}}(t, x_n)$.

$$\begin{cases} \partial_t u_{x_1,\dots,x_{n-1}} + H_{V_2}(D_{x_n} u_{x_1,\dots,x_{n-1}}) = 0 & \text{in } (\Sigma_2)_{x_1,\dots,x_{n-1}}, \\ u_{x_1,\dots,x_{n-1}}(0, x_n) = u(0, (x_1, \dots, x_n)) & \forall x_n \in ((\Sigma_2)_{x_1,\dots,x_{n-1}})_0, \end{cases}$$

where $H_{V_2}(p_n)$ is the Hamiltonian associated to $L_{V_2}(v_n)$.

The viscosity solution $u_{x_1,\dots,x_{n-1}}(t, x_n)$ is equal to $u(t, (x_1, \dots, x_n))$ for $(t, (x_1, \dots, x_n)) \in \Sigma_2$. Indeed

$$u_{x_1,\dots,x_{n-1}}(t, x_n) = \min_{y_n \in \mathbb{R}} \left\{ u(0, (x_1, \dots, x_{n-1}, y_n)) + t L_{V_2}\left(\frac{x_n - y_n}{t}\right) \right\} = u(t, (x_1, \dots, x_n)),$$

where the last equality follows from the fact that, for (t,x) in Σ_2, the unique minimizer in (1.10) is $y = (x_1 - t d_1(t,x), \dots, x_n - t d_n(t,x))$ and $d(t,x) = (0, \dots, 0, d_n(t,x))$ on Σ_2.

Let us define as usual

$$d_{x_1,\dots,x_{n-1}}(t, x_n) := (H_{V_2})_{p_n}(D_{x_n} u_{x_1,\dots,x_{n-1}}(t, x_n)),$$

and

$$\mu_{x_1,\dots,x_{n-1}}(t,\cdot) := \frac{\partial}{\partial x_n} d_{x_1,\dots,x_{n-1}}(t,\cdot).$$

The vector field $d_{x_1,\dots,x_{n-1}}(t,\cdot)$ is one-dimensional. Hence, for Theorem 4.2, $d_{x_1,\dots,x_{n-1}}(t,\cdot)$ belongs to $BV(((\Sigma_2)_{x_1,\dots,x_{n-1}})_t)$ for any $(x_1, \dots, x_{n-1}) \in I$, for any $t \in [0,T]$.

On the set $\Sigma_2 \subset U$, the matrix of Radon measures $D_x d$ has no jump part. Moreover, since $(\Sigma_2)_t$ is contained on the set $\{x|\ d_1(t,x) = 0, \ldots, d_{n-1}(t,x) = 0\}$ and $d(t,\cdot)$ is BV, Proposition 1.10 implies

$$\frac{\partial}{\partial x_l} d_i(t, (\Sigma_2)_t) = 0 \quad \text{for } i = 1, \ldots, n-1 \text{ and } l = 1, \ldots, n.$$

Therefore

$$\mathrm{div} d(t,\cdot) = \frac{\partial}{\partial x_n} d_n(t,\cdot) \quad \text{on } (\Sigma_2)_t.$$

For every $(t,x) \in \Sigma_2$, $u_{x_1,\ldots,x_{n-1}}(t,x_n) = u(t,(x_1,\ldots,x_n))$ implies

$$d_n(t,x) = d_{x_1,\ldots,x_{n-1}}(t,x_n).$$

The vector field $d_n(t,(x_1,\ldots,x_{n-1},\cdot))$ is a one-dimensional restriction of $d_n(t,\cdot)$ thus, for Proposition 1.11, belongs to $BV(((\Sigma_2)_{x_1,\ldots,x_{n-1}})_t)$ for almost every $(x_1,\ldots,x_{n-1}) \in I$. Since even $d_{x_1,\ldots,x_{n-1}}(t,\cdot)$ is BV on $((\Sigma_2)_{x_1,\ldots,x_{n-1}})_t$, Proposition 1.9 implies

$$\frac{\partial}{\partial x_n} d_n(t,(x_1,\ldots,x_{n-1},\cdot)) = \frac{\partial}{\partial x_n} d_{x_1,\ldots,x_{n-1}}(t,\cdot)$$

for $\mathcal{H}^{n-1}$-a.e. $(x_1,\ldots,x_{n-1}) \in I$. Therefore taken a Borel set $A \subset (\Sigma_2)_t$ and any $\phi \in C_c^\infty((\Sigma_2)_t)$,

$$\int_A \phi(x) d\mu(t,x) = \int_I \int_{A_{x_1,\ldots,x_{n-1}}} \phi(x) d\mu_{x_1,\ldots,x_{n-1}}(t,x_n) d(x_1,\ldots,x_{n-1}).$$

Thanks to the convexity of L_{V_2}, we can apply Theorem 4.4 to $\mu_{x_1,\ldots,x_{n-1}}(t,\cdot)$ and obtain the following estimates.

For any $\tau > 0$, let A be a Borel set in $(\Sigma_2)_t$, $t \in (\tau, T]$. Then for any $0 \le \delta \le t - \tau$ and every section $A_{x_1,\ldots,x_{n-1}}$, for $(x_1,\ldots,x_{n-1}) \in I$, we have

$$\mathcal{H}^1(X_{t,\tau}^{x_1,\ldots,x_{n-1}}(A_{x_1,\ldots,x_{n-1}})) \ge \mathcal{H}^1(A_{x_1,\ldots,x_{n-1}}) - (t-\tau)\mu_{x_1,\ldots,x_{n-1}}(t, A_{x_1,\ldots,x_{n-1}}),$$

$$\mathcal{H}^1(X_{t,\tau+\delta}^{x_1,\ldots,x_{n-1}}(A_{x_1,\ldots,x_{n-1}})) \ge \frac{t-(\tau+\delta)}{t-\tau} \mathcal{H}^1(X_{t,\tau}^{x_1,\ldots,x_{n-1}}(A_{x_1,\ldots,x_{n-1}})).$$

Here we denote with $X_{t,\tau}^{x_1,\ldots,x_{n-1}}(x_n)$ the one-dimensional map defined on $((\Sigma_2)_{x_1,\ldots,x_{n-1}})_t$

$$X_{t,\tau}^{x_1,\ldots,x_{n-1}}(x_n) := x_n - (t-\tau) d_{x_1,\ldots,x_{n-1}}(t,x_n).$$

The corresponding n-dimensional map defined on $(\Sigma_2)_t$

$$X_{t,\tau}(x) := x - (t-\tau) d(t,x),$$

reduces to

$$X_{t,\tau}(x) = (x_1,\ldots,x_{n-1}, X_{t,\tau}^{x_1,\ldots,x_{n-1}}(x_n))$$

for every $x \in (\Sigma_2)_t$.

We can integrate the previous estimates with respect to $\mathcal{H}^{n-1}$ over I to recover estimates of type (4.1) and (4.2). For any $\tau > 0$, given a Borel set $A \subset (\Sigma_2)_t$, for t in $[\tau, T]$, we have

$$\mathcal{H}^n(X_{t,\tau}(A)) \geq \mathcal{H}^n(A) - (t-\tau)\mu(t, A). \tag{4.9}$$

For any $\tau > 0$, given a Borel set $A \subset \Sigma_t$, for t in $[\tau, T]$ and $0 \leq \delta \leq t - \tau$ we have

$$\mathcal{H}^n(X_{t,\tau+\delta}(A)) \geq \frac{t-(\tau+\delta)}{t-\tau}\mathcal{H}^n(X_{t,\tau}(A)). \tag{4.10}$$

Thus the strategy seen in the Subsection 4.2 can be easily applied to prove that $\mu(t,\cdot)$, restricted to $(\Sigma_2)_t$, can have Cantor part only for a countable number of t's in $[0,T]$.

Consider now the case in which V_2 consists of a finite number of straight lines. When we consider $\mu(\cdot,\cdot)$ restricted to the points of Σ_2 such that $d(t,x)$ belongs only to a part of one of the straight lines, we can apply the considerations done in the case where V_2 consists only of a single straight line. On the other hand, when we consider $\mu(\cdot,\cdot)$ restricted to the points of Σ_2 such that $d(t,x)$ belongs to an intersection point of two, or more, straight lines, the divergence $\mathrm{div} d(t,\cdot) = \mu(t,\cdot)$ must be zero on every Borel set, as seen in the 2-dimensional case. The case in which V_2 is contained in a finite number of straight lines is analogous.

Thus the measure $\mu(t,\cdot)$ can have Cantor part only for a countable number of t's in $[0,T]$ even when V_2 consists of a finite number of straight lines. $\square$

Once Claim 2.(2) is proved, we can iteratively prove all the others Claims 2.(j-1) for $j = 4, \cdots, n$ just by repeating the same considerations for the general case in which V_j consists of a finite union of $(j-1)$-dimensional planes. This case can be treated as usual distinguishing the two cases. When we consider $\mu(\cdot,\cdot)$ restricted to the points of Σ_j such that $d(t,x)$ belongs only to a part of one of the $(j-1)$-dimensional planes, we can apply the considerations done in the case where V_j consists only of a single $(j-1)$-dimensional plane. On the other hand, when we consider $\mu(\cdot,\cdot)$ restricted to the points of Σ_j such that $d(t,x)$ belongs to a $(j-2)$-dimensional plane intersection of two, or more, $(j-1)$-dimensional planes, we can reduce the problem to the $(j-2)$-dimensional case. Indeed in this case we can apply again the iterative proof. The case in which V_j is contained in a finite number of $(j-1)$-dimensional planes is analogous.

The considerations above done for $j = n+1$ concludes even the proof of Claim 2.(n). $\square$

Let us recall all the necessary assumptions.

Suppose H is $C^2(\mathbb{R}^n)$ convex and

$$\lim_{|p|\to\infty} \frac{H(p)}{|p|} = +\infty.$$

(HYP(0)) The vector field $d(t,\cdot)$ belongs to $[BV(\Omega_t)]^n$ for every $t \in [0,T]$.

Define V_{π_n} as

$$V_{\pi_n} := \{v \in \mathbb{R}^n |\ L(\cdot) \text{ is not twice differentiable in } v\},$$

and

$$\Sigma_{\pi_n} := \{(t,x) \in U |\ d(t,x) \in V_{\pi_n}\} \quad \text{and} \quad \Sigma^c_{\pi_n} := U \setminus \Sigma_{\pi_n}.$$

(HYP(n)) We suppose V_{π_n} to be contained in a finite union of hyperplanes Π_{π_n}.

For $j = n, \dots, 3$ for any $(j-1)$-dimensional plane π_{j-1} in Π_{π_j}, let $L_{\pi_{j-1}} : \mathbb{R}^{j-1} \to \mathbb{R}$ be the $(j-1)$-dimensional restriction of L to π_{j-1} and

$$V_{\pi_{j-1}} := \{v \in \mathbb{R}^{j-1} | \; L_{\pi_{j-1}}(\cdot) \text{ is not twice differentiable in } v\}.$$

Define

$$\Sigma_{\pi_{j-1}} := \{(t,x) \in \Sigma_{\pi_j} | \; d(t,x) \in V_{\pi_j}\} \quad \text{and} \quad \Sigma^c_{\pi_{j-1}} := \Sigma_{\pi_j} \setminus \Sigma_{\pi_{j-1}}.$$

(HYP(j-1)) We suppose $V_{\pi_{j-1}}$ is contained in a finite union of $(j-2)$-dimensional planes $\Pi_{\pi_{j-1}}$, for every $\pi_{j-1} \in \Pi_{\pi_j}$.

Remark 4.7. There is no need to ask any assumption on the one-dimensional restriction of L to a straight line in any of the V_{π_2} for a plane π_2, since in the one-dimensional case the SBV regularity is proven without any further assumptions on L.

Theorem 4.8. *With the above assumptions (HYP(0)),(HYP(n)),...,(HYP(2)), the Radon measure* $\mathrm{div}d(t,\cdot)$ *has Cantor part on* Ω_t *only for a countable number of* t*'s in* $[0,T]$.

The following corollaries are easily obtained from Theorem 4.8.

Corollary 4.9. *Let* $D_x u(t,\cdot)$ *belongs to* $[BV(\Omega_t)]^n$ *for every* $t \in [0,T]$ *and let* L *satisfy the assumptions (HYP(n)),..., (HYP(2)), then the Radon measure* $\mathrm{div}d(t,\cdot)$ *has Cantor part on* Ω_t *only for a countable number of* t*'s in* $[0,T]$.

Proof. If $D_x u(t,\cdot)$ belongs to $[BV(\Omega_t)]^n$ for every $t \in [0,T]$, then $d(t,\cdot) = H_p(D_x u(t,\cdot))$ belongs to $[BV(\Omega_t)]^n$ for every $t \in [0,T]$. □

Corollary 4.10. *Let* $u(0,\cdot)$ *be semiconcave and let* L *satisfy (HYP(n)),..., (HYP(2)), then the Radon measure* $\mathrm{div}d(t,\cdot)$ *has Cantor part on* Ω_t *only for a countable number of* t*'s in* $[0,T]$.

Proof. It follows from Proposition 1.35. □

Chapter 5

Some applications

In this chapter we present some simple applications of Ambrosio and De Lellis's SBV regularity theorem 0.1, for entropy solutions of the one-dimensional scalar conservation laws

$$\partial_t U + D_x(H(U)) = 0 \qquad \text{in } \Omega := \mathbb{R}^+ \times (a,b). \tag{5.1}$$

That theorem can be easily extended to one-dimensional Hamilton-Jacobi equations. Indeed, the potential, given by

$$\begin{cases} \partial_t u = -H(U) \\ D_x u = U, \end{cases}$$

is a viscosity solution of the Hamilton-Jacobi equation

$$\partial_t u + H(D_x u) = 0 \qquad \text{in } \Omega \tag{5.2}$$

if and only if U is an entropy solution to (5.1). Therefore, Theorem 0.1 applies also to the distributional derivative of a viscosity solution of Hamilton-Jacobi equation (5.2) when H is $C^2(\Omega)$ and locally uniformly convex.

In Sections 5.1 and 5.2 we describe Generalized Hydrostatic Boussinesq (GHB) equations and the model of sticky particles, then, in Section 5.3, we show how Theorem 0.1 of Ambrosio and De Lellis (whose proof can be found in [4]) applies to them in the one-dimensional case. In the last section we present a counterexample which prevent us from using the same approach for the multi-dimensional case. A similar counterexample was shown by Vasseur in [38], but it was never published.

The results presented here can be found in Tonon [37].

5.1 Generalized Hydrostatic Boussinesq equations

Generalized Hydrostatic Boussinesq (GHB) equations can be seen as the most degenerate version of Generalized Navier-Stokes Boussinesq (GNSB) equations, where both the inertia terms and the dissipative operator are neglected. These equations rule the dynamic of a fluid under fast convection. In terms of the temperature of the fluid they take the form

$$y = x + \nabla p, \qquad \nabla \cdot v = 0,$$

$$\partial_t y + (v \cdot \nabla) y = G(x), \tag{5.3}$$

here, being $D \subset \mathbb{R}^n$ a smooth bounded domain where the fluid is placed, the function $y(t,x) : \mathbb{R}^+ \times D \to \mathbb{R}^n$ is the generalized temperature field of the fluid, $v(t,x) : \mathbb{R}^+ \times D \to \mathbb{R}^n$ its velocity, $p(t,x) : \mathbb{R}^+ \times D \to \mathbb{R}$ the pressure, $G(x) : \mathbb{R}^n \to \mathbb{R}^n$ the generalized heat source term, an $[L^\infty(\mathbb{R}^n)]^n$ function. Equation (5.3) can be seen as a generalization of the hydrostatic balance in Convection Theory.

The fact that G depends only on the position of the fluid allows us to apply Theorem 0.1. However, this is a very particular assumption since the heat source can depend also on time and temperature $G = G(t,x,y)$.

The complete description of this system can be found in [17], Chapter 3. Passing to Lagrangian coordinates, Brenier proved there that a generalized solution can be constructed.

Since we need some concepts of Optimal Transport Theory let us recall some preliminary definitions and results.

First, we introduce rearrangements and measure preserving maps. Given two $[L^2(D)]^n$ maps Y and Z, we say that they are rearrangement of each other if they define the same image measure, i.e. for all continuous f on $\mathbb{R}^n$, such that $|f(y)| \leq 1 + |y|^2$,

$$\int_D f(Y(a))da = \int_D f(Z(a))da.$$

We say that Y in $[L^2(D)]^n$ is a measure preserving map, when it is a rearrangement of the identity map, i.e. for all continuous f on $\mathbb{R}^n$, such that $|f(y)| \leq 1 + |y|^2$,

$$\int_D f(a)da = \int_D f(Y(a))da.$$

Next we define the class of maps with convex potential. We say that an $[L^2(D)]^n$ map Y belongs to the class C of maps with a convex potential, if there is a lower semi-continuous convex function $p : \mathbb{R}^n \to (-\infty, +\infty]$ such that, for $\mathcal{H}^n$-a.e. point x in D, the gradient $\nabla p(x)$ coincides with Y.

Then, looking for a rearrangement with convex potential, we have the following Brenier's Theorem which can be found in [14]:

Theorem 5.1 (Brenier). *Let Y be a non degenerate $[L^2(D)]^n$ map. Then there is a unique polar factorization*

$$Y = Y^R \circ X,$$

where Y^R belongs to C and X is a Lebesgue measure preserving map of D.

In this decomposition, Y^R is the unique rearrangement of Y in C and X is the unique measure preserving map of D that minimizes

$$\int_D |X(a) - Y(a)|^2 da.$$

In addition, X can be written:

$$X(a) = (\nabla \phi)(Y(a)), \quad \mathcal{H}^n\text{-a.e. } a \in D,$$

where ϕ is a convex Lipschitz function defined on $\mathbb{R}^n$.

Coming back to our problem and passing to Lagrangian coordinates the system (5.1,5.3) looks like

$$Y(t,a) = X(t,a) + \nabla p(t, X(t,a)), \tag{5.4}$$

$$\partial_t Y(t,a) = G(X(t,a)), \tag{5.5}$$

where, for all t, $D \ni a \mapsto X(t,a)$ is a measure preserving map as a consequence of the fact that v is a smooth divergence-free vector field. $X(t,a)$ denotes the position of a fluid particle a at a time t, therefore its velocity and its temperature are

$$\partial_t X(t,a) = v(t, X(t,a)), \qquad Y(t,a) = y(t, X(t,a)).$$

Note that, due to the above equations, if particles reach the same position they will have the same velocity, the same temperature and will no more separate. In fact GHB equations are very closed to systems of sticky particles as we can see in the next section.

Assuming a priori that the map $x \mapsto x + \nabla p(t,x)$ has convex potential, we deduce from (5.4) that $x \mapsto x + \nabla p(t,x)$ is the unique convex rearrangement $Y^R(t,\cdot)$ of $Y(t,\cdot)$, due to the fact that $Y(t,a) = Y^R(t,\cdot) \circ X(t,a)$. Moreover (5.5) implies that for all C^1 function f compactly supported on $\mathbb{R}^n$, $Y^R(t,\cdot)$ satisfies

$$\frac{d}{dt}\int_D f(Y^R(t,a))da = \int_D (\nabla f)(Y^R(t,a)) \cdot G(a)da. \tag{5.6}$$

With these considerations Brenier naturally introduced a more general concept of solution to GHB system. We say that Y^{CR} in $C^0([0,T],[L^2(D)]^n)$ is the convex rearrangement (CR) solution to the GHB equations (5.4, 5.5), if :

- $Y^{CR}(t,\cdot)$ belongs to the set C of all maps with convex potential, for all $t \in [0,T]$,
- for all compactly supported C^1 function f on $\mathbb{R}^n$, $Y^{CR}(t,\cdot)$ satisfies (5.6).

In [17], he proved the following existence theorem.

Theorem 5.2 (Brenier)**.** *For each initial condition Y^0 in $[L^2(D)]^n$, there is at least one CR-solution $Y^{CR}(t,a)$ such that $Y^{CR}(0,\cdot) = (Y^0)^R(\cdot)$.*

This solution can be obtained as the limit in $C^0([0,T],[L^2(D)]^n)$ as $h \to 0$, of a time discrete approximation $Y^h(t,a)$ defined, first at discrete times $t = nh$, by:

$$Y^h(nh+h,a) = [Y^h(nh,a) + hG(a)]^R, \quad n = 0,1,2,\dots$$

(where, as seen before, $(\cdot)^R$ is the convex rearrangement operator) and then linearly interpolated in t.

The time discrete approximation, given by the theorem above, tells us that starting from an initial temperature datum Y^0, the CR-solution evolves linearly as $(Y^0)^R(a) + tG(a)$ as far as this function remains with convex potential. When this is no more the case, it is rearranged in order to preserve the membership to the space of maps with convex potential.

Note that CR-solutions are $\mathcal{H}^n$-a.e. equal to functions with convex potential, i.e. for all t there exists a convex function $\psi_t : D \to \mathbb{R}$ such that

$$Y^{CR}(t,a) = D\psi_t(a),$$

for $\mathcal{H}^n$-a.e. a in D. Taking now the Legendre transform of this convex function

$$u(t,x) = \sup_{a\in D} \{x \cdot a - \psi_t(a)\},$$

we obtain a function $u(t,\cdot)$ which is again convex and its distributional derivative $D_x u(t,\cdot)$ is the generalized inverse of $Y^{CR}(t,\cdot)$. We are interested in the regularity of $D_x u(t,\cdot)$. What we can say so far is that it is a function of bounded variation.

5.1.1 One-dimensional case

In the one-dimensional case it is possible to look at $D_x u(t,\cdot)$ as a solution of a scalar conservation law.

First we can observe that, taking $D = [0,1]$, the convex rearrangement is the monotone non decreasing rearrangement defined by

$$Y^R(s) = \inf\{t \in \mathbb{R} \mid \mu_Y(t) > s\},$$

for s in [0,1], where

$$\mu_Y(t) = \mathcal{H}^1(\{Y < t\}),$$

is the distribution function. For a detailed description of monotone non decreasing rearrangement we refer to [34], Chapter 1.

One of the properties of monotone non decreasing rearrangement is that it is non expansive in $L^2([0,1])$, i.e.

$$\int_D |Y^R(a) - Z^R(a)|^2 da \leq \int_D |Y(a) - Z(a)|^2 da.$$

This property guarantees the uniqueness of the solution of (5.6).

Moreover, as explained in [16], the limit of the time discrete approximation, defined in Theorem 5.2, satisfies the sub-differential inclusion:

$$G(x) \in \partial_t Y + \partial\Psi[Y],$$

where $\Psi[Y] = 0$ if Y is a non decreasing function of x in D, and $\Psi[Y] = +\infty$ otherwise.

The generalized inverse of the solution, in the one-dimensional case, can be found using the Heaviside function. Looking at its behavior, Brenier proved in [15], the following theorem. In the proof he used a Transport Collapse method, which involves the same time discrete approximation scheme seen in Theorem 5.2.

Theorem 5.3 (Brenier). *Let $Y^{CR}(t,a)$ be the CR-solution found in Theorem 5.2, then the generalized inverse*

$$U(t,y) = \int_0^1 \mathbb{H}(y - Y^{CR}(t,a))da,$$

where $\mathbb{H}$ *is the Heaviside function, is an entropy solution of the scalar conservation law*

$$\partial_t U + D_x(H(U)) = 0,$$

where H *is the primitive of* G, $H_p(p) = G(p)$.

We are interested in the regularity of an entropy solution of a scalar conservation law with non decreasing initial conditions and Lipschitz flux function H. Applying what we have already said, since we are in the one-dimensional case, the entropy solution above can be seen as the derivative of the unique viscosity solution of the following Hamilton-Jacobi equation

$$\partial_t u + H(D_x u) = 0,$$

with a convex initial datum and Lipschitz Hamiltonian.

5.2 Sticky particles

At a discrete level pressureless gases with sticky particles can be modeled by a finite collection of particles that get stuck together right after they collide with conservation of mass and momentum. On the other hand at a continuous level the model is governed by the following one-dimensional system of conservation laws in $(0, +\infty) \times \mathbb{R}$

$$\partial_t \rho + D_x(\rho v) = 0,$$

$$\partial_t(\rho v) + D_x(\rho v^2) = 0,$$

where $\rho(t, x)$ is the density field, while $v(t, x)$ is the velocity one. This set of equations can be seen as the limit, when pressure goes to zero, of the usual Euler equations. In [18], Brenier and Grenier showed that the continuous model can be fully described, in an alternative way, by scalar conservation laws, with non decreasing initial conditions, general flux functions and the usual Kruzhkov entropy condition.

In particular they proved that if (ρ, v) is a solution corresponding to sticky particles, then there exist $H \in \mathrm{Lip}(\mathbb{R})$ and U entropy solution of

$$\partial_t U + D_x(H(U)) = 0,$$

where $U(t, x) = D_x u(t, x)$ is such that $\rho(t, x) = D_x^2 u(t, x)$ is a cumulative distribution function associated to the probability measure ρ and $H_p(p) = v(0, p)$.

The proof uses a scheme in which a finite number of particles are described by weight, position and velocity, under the assumption that the speed of a particle is constant as long as it meets no new particles and it changes only when shocks occur. Only a finite number of shocks can occur because particles remain stuck together after a collision. Moreover particles having the same position at a time t move together at the same speed and their total momentum is the sum of their initial momentum. This scheme is strongly reminiscent of Dafermos's polygonal approximation methods for scalar conservation laws, where each particle corresponds to a jump of an entropy solution of a scalar conservation law with a piecewise linear continuous flux function. Thus, it is reasonable to expect, as it is, that the continuous limit of the sticky particles dynamics is properly described by a scalar conservation law.

The fact that the distribution function U is a non decreasing entropy solution of that scalar conservation law strictly relates sticky particle system to Convection Theory. Indeed if we take the generalized inverse of U, which is precisely the monotone rearrangement of the measure ρ, it turns out that it is exactly the limit of the time discrete approximation seen in Theorem 5.2.

As we did for GHB equations we can relate the non decreasing entropy solutions to the viscosity solution of an Hamilton-Jacobi equation with convex initial datum.

5.3 Convex solutions of Hamilton-Jacobi equations in the multi-dimensional case

Let us now consider the following Hamilton-Jacobi equation

$$\partial_t u + H(D_x u) = 0,$$

with initial datum $u(0,x) = \frac{1}{2}|x|^2$, and Lipschitz Hamiltonian H. Thus we are in a particular case of the ones considered above. As proved in [6], by Bardi and Evans, the unique viscosity solution to such an equation has the form

$$u(t,x) = \sup_y \inf_z \left\{ \frac{1}{2}|z|^2 + y \cdot (x-z) - tH(y) \right\}.$$

This representation formula is true even in the multi-dimensional case and an analogous one works as well for general initial datum but convex Hamiltonian. Moreover it is equivalent to

$$u(t,x) = \sup_y \left\{ x \cdot y - \frac{1}{2}|y|^2 - tH(y) \right\}. \tag{5.7}$$

Here the sup becomes a maximum under suitable hypotheses on H.

Note that equation (5.7) is equivalent to saying that u is the Legendre transform of $\frac{1}{2}|y|^2 + tH(y)$. On the other hand since u is, in the GHB equation case, the Legendre transform of $\psi_t(a)$, we have the following geometric representation for the CR-solution, for $\mathcal{H}^1$-a.e. a,

$$Y^{CR}(t,a) = \nabla convex(\psi_0(a) + tH(a)),$$

where $convex(f) = \max\{g \leq f |\ g \text{ convex}\}$.

Define

$$v(t,x) := -\frac{1}{t}\left(u(t,x) - \frac{1}{2}|x|^2\right),$$

then

$$v(t,x) = \min_y \left\{ H(y) + \frac{|x-y|^2}{2t} \right\}$$

is the unique viscosity solution of

$$\partial_t v + \frac{|D_x v|^2}{2} = 0$$

with Lipschitz initial datum $v_0(x) = H(x)$.

Since the Hamiltonian $\frac{|x|^2}{2}$ is uniformly convex we can use directly Theorem 0.1 of Ambrosio and De Lellis, to prove that $D_x v(t, \cdot)$ belongs to SBV for a.e. t and the same is true also for $D_x u(t, \cdot)$.

Remark 5.4. From what we have seen, in the one-dimensional case, SBV regularity holds for the generalized inverse of a solution of the GHB equation with the identity as initial datum and for the cumulative distribution function associated to the density of the pressureless gas.

5.4 Multi-dimensional case

We wonder if Hamilton-Jacobi equations are a good model for GHB systems or sticky particles models even in the multi-dimensional case. Are they able to describe the behavior of our solution? If this was the case we could automatically state SBV regularity applying Theorem 0.2. Unfortunately the answer to our question is negative. In the following subsection we show a counterexample in which a multi-dimensional solution of an Hamilton-Jacobi equation has a behavior which is not allowed for GHB systems or sticky particles models, i.e. Theorem 0.2 does not suit our problem. However, this does not mean that SBV regularity cannot be proved in some other way.

5.4.1 A counterexample

A first counterexample was found by Vasseur in [38] but it was never published. With that counterexample Vasseur showed a discrepancy between the density distribution $\tilde{\rho}(t, x) = \det(D_x^2 u(t, x))$, associated to the solution u of the Hamilton-Jacobi equation $\partial_t u + H(D_x u) = 0$ with initial datum $u(0, x) = \frac{1}{2}$, and the density distribution $\rho(t, x)$, generated from the identity $\rho_0(x) = 1$ in a sticky particles process with speed $v = H_p(p)$. Indeed, he proved the existence of a time t at which the two density distribution differ.

The following counterexample shows the same discrepancy, underlining in addiction the cause of it. Hamilton-Jacobi equations allow separations of particles after collisions.

Consider the viscosity solution of the Hamilton-Jacobi equation

$$v_t + \frac{1}{2}|D_x v|^2 = 0,$$

in $\mathbb{R}^2$, with initial datum

$$v(0, x) = \begin{cases} -\frac{|x|^2}{2} & \text{for } x \in \overline{B(0,1)} \\ f(x) & \text{for } x \in \overline{B(0,2)} \setminus B(0,1) \\ -|x_1| & \text{for } x \in \mathbb{R}^2 \setminus B(0,2), \end{cases}$$

where $f(x)$ joins smoothly $-\frac{|x|^2}{2}$ to $-|x_1|$ and satisfies $f(x) > -\frac{|x|^2}{2}$ in $B(0,2) \setminus \overline{B(0,1)}$, being $B(x, r)$ the open ball with center in x and radius $r > 0$.

Note that following upside down the passages seen in Section 5.3 we can recover from v a convex viscosity solution of the equation

$$\partial_t u + H(D_x u) = 0, \qquad u(0,x) = \frac{|x|^2}{2}$$

where

$$u(t,x) = -tv(t,x) + \frac{|x|^2}{2}$$

and $H(x) = v(0,x)$ is a smooth function. We are thus considering a viscosity solution of Hamilton-Jacobi with convex initial datum. If Hamilton-Jacobi were the good model for GHB and sticky particles systems, passing to the Legendre transform of our viscosity solution we should recover the CR-solution limit of the time-discrete approximation scheme.

Using the Hopf-Lax formula for convex Hamiltonians we recover the viscosity solution for any time t

$$v(t,x) = \min_y \left\{ v(0,y) + \frac{(x_1 - y_1)^2 + (x_2 - y_2)^2}{2t} \right\}.$$

Let us compute the value of v for $t = 1$ in the origin:

$$v(1,(0,0)) = \min_y \left\{ v(0,y) + \frac{y_1^2 + y_2^2}{2} \right\}. \tag{5.8}$$

Observe that for any y in $\overline{B(0,1)}$ we have

$$v(0,y) + \frac{y_1^2 + y_2^2}{2} = 0.$$

For any y in $B(0,2) \setminus \overline{B(0,1)}$

$$v(0,y) + \frac{y_1^2 + y_2^2}{2} > 0.$$

For any y in $\mathbb{R}^2 \setminus B(0,2)$ we have

$$v(0,y) + \frac{y_1^2 + y_2^2}{2} = -|y_1| + \frac{y_1^2 + y_2^2}{2} \geq -|y_1| + \frac{y_1^2}{2},$$

hence we can restrict the minimum in the region $\mathbb{R}^2 \setminus B(0,2)$ to points with $y_2 = 0$, moreover for that points we have

$$-|y_1| + \frac{y_1^2}{2} \geq 0,$$

and the equality occurs only for $(-2,0)$ and $(2,0)$.

Thus the minimum in (5.8) is obtained if and only if y belongs to the set $B(0,1) \cup \{(-2,0),(2,0)\}$. The origin is therefore a point of non differentiability for $v(1,\cdot)$ with the convex hull of the set of all minima as superdifferential. This means that all the points in the set $B(0,1) \cup \{(-2,0),(2,0)\}$, which is of positive $\mathcal{H}^2$-measure, are transported by the flux along straight line trajectories which collide at time $t = 1$ in the position $(0,0)$.

However, for any $\delta > 0$, we have to compute

$$v(1+\delta,(0,0)) = \min_{y} \left\{ v(0,y) + \frac{y_1^2+y_2^2}{2(1+\delta)} \right\}. \tag{5.9}$$

For any y in $\overline{B(0,1)}$ we have

$$v(0,y) + \frac{y_1^2+y_2^2}{2(1+\delta)} = -\delta \frac{y_1^2+y_2^2}{2(1+\delta)} \geq -\frac{\delta}{2(1+\delta)}.$$

For any y in $B(0,2) \setminus \overline{B(0,1)}$

$$v(0,y) + \frac{y_1^2+y_2^2}{2(1+\delta)} > -\delta \frac{y_1^2+y_2^2}{2(1+\delta)} > -\frac{4\delta}{2(1+\delta)}.$$

Here we note that $-\frac{4\delta}{2(1+\delta)} < -\frac{\delta}{2(1+\delta)}$.

For any y in $\mathbb{R}^2 \setminus B(0,2)$ we have

$$v(0,y) + \frac{y_1^2+y_2^2}{2(1+\delta)} = -|y_1| + \frac{y_1^2+y_2^2}{2(1+\delta)} \geq -|y_1| + \frac{y_1^2}{2(1+\delta)},$$

hence we can restrict the minimum to the points in $\mathbb{R}^2 \setminus B(0,2)$ with $y_2 = 0$. Moreover, for that points we have that the minimum value is reached for $|y_1| = 2$ if $1+\delta < 2$, for $|y_1| = 1+\delta$ otherwise.

In the first case

$$-|y_1| + \frac{y_1^2}{2(1+\delta)} = -\frac{4\delta}{2(1+\delta)}.$$

In the second one

$$-|y_1| + \frac{y_1^2}{2(1+\delta)} = -\frac{(1+\delta)^2}{2(1+\delta)} < -\frac{4\delta}{2(1+\delta)}.$$

Thus (0,0) is a point of non differentiability even for $t = 1+\delta$ for any $\delta > 0$. Moreover its superdifferential is the set $[(-2,0),(2,0)]$ for $0 < \delta < 1$, or the set $[(-(1+\delta),0),((1+\delta),0)]$ for $\delta > 1$. In any case it is a set of positive $\mathcal{H}^1$-measure. This set has non empty intersection with the superdifferential of v in the origin at time $t = 1$ but does not contain the whole of it. Recall that, the superdifferential of v in the origin contains, at time $t = 1$, the set $\text{convex}(B(0,1) \cup \{(-2,0),(2,0)\})$ which is a set of positive $\mathcal{H}^2$-measure.

Points, being in $B(0,1) \cup \{(-2,0),(2,0)\}$ at time $t = 0$, collide at time $t = 1$ and separate at time $t = 1+\delta$ for any $\delta > 0$.

We have thus shown an example of a viscosity solution in which a point of non differentiability of zero codimension evolves in a point of non differentiability of codimension one.

As we have already said, coming back to $u(t,x) = -tv(t,x) + \frac{|x|^2}{2}$ and passing to the Legendre transform of our viscosity solution, we should obtain the CR-solution of the GSB equation. However for this function a flat part of dimension two would evolve in a flat part of dimension one, in contrast with propagation of flat parts. Particles stuck together could have different velocities but this is not the case for GHB and the sticky particles model.

Hence GHB and the sticky particles model cannot be truly described by Hamilton-Jacobi equation in the multi-dimensional case.

Chapter 6

Decomposition of BV functions

The aim of this chapter is to give a generalization of Jordan decomposition property to real valued BV functions of many variables.

The starting point is a recent result presented to us by Alberti, Bianchini and Crippa, which shows that a real Lipschitz function of many variables with compact support can be decomposed in sum of monotone functions. Precisely they give the following definition of monotone function

Definition 6.1. A function $f : \mathbb{R}^n \to \mathbb{R}$, which belongs to $\mathrm{Lip}(\mathbb{R}^n)$, is said to be *monotone* if the level sets $\{f = t\} := \{x \in \mathbb{R}^n | \ f(x) = t\}$ are connected for every $t \in \mathbb{R}$.

and state the following theorem.

Theorem 6.2 (Alberti, Bianchini and Crippa)**.** *Let f be a function in $\mathrm{Lip_c}(\mathbb{R}^n)$ with compact support. Then there exists a countable family $\{f_i\}_{i\in\mathbb{N}}$ of functions in $\mathrm{Lip_c}(\mathbb{R}^n)$ such that $f = \sum_i f_i$ and each f_i is monotone. Moreover there is a pairwise disjoint partition $\{\Omega_i\}_{i\in\mathbb{N}}$ of Borel sets of $\mathbb{R}^n$ such that ∇f_i is concentrated on Ω_i.*

In the case of BV functions, which are defined $\mathcal{H}^n$-a.e., an appropriate generalization of the concept of monotone function has to involve super-level sets, sub-level sets and the concept of indecomposable set, as given in [3].

Definition 6.3. A set $E \subseteq \mathbb{R}^n$ with finite perimeter is said to be *decomposable* if there exists a partition (A, B) of E such that $P(E) = P(A) + P(B)$ and both $\mathcal{H}^n(A)$ and $\mathcal{H}^n(B)$ are strictly positive. A set E is said to be *indecomposable* if it is not decomposable.

Definition 6.4. A function $f : \mathbb{R}^n \to \mathbb{R}$, which belongs to $L^1_{loc}(\mathbb{R}^n)$, is said to be *monotone* if the super-level sets $\{f > t\} := \{x \in \mathbb{R}^n | \ f(x) > t\}$ and the sub-level sets $\{f < t\} := \{x \in \mathbb{R}^n | \ f(x) < t\}$ are of finite perimeter and indecomposable for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}$.

As proved in Section 6.3, in the case of Lipschitz functions, Definition 6.1 and Definition 6.4 are equivalent.

When comparing the case of functions of one variables with the case of functions of many variables differences and analogies arise.

On the one hand, it can be found an L^1 monotone function, which is not of bounded variation, that is a counterexample to the fact that monotonicity is a sufficient condition for being of bounded variation (Example 6.14).

On the other hand, it can be stated that a BV function is decomposable in a countable sum of monotone functions, similarly to the case of BV functions of one real variable.

The main result of this chapter is the following.

Theorem 6.5. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a $BV(\mathbb{R}^n)$ function. Then there exists a finite or countable family of monotone $BV(\mathbb{R}^n)$ functions $\{f_i\}_{i\in I}$, such that*

$$f = \sum_{i\in I} f_i \quad \text{and} \quad |Df| = \sum_{i\in I} |Df_i|.$$

This decomposition is in general not unique, see Remark 6.12, and it can generate monotone BV functions without mutually singular distributional derivatives, see Example 6.13. Thus we loose the property true in the Lipschitz case.

The main tool for proving this theorem is a decomposition theorem for sets of finite perimeter, presented here in the form given in [3].

Theorem 6.6 (Ambrosio, Caselles, Masnou and Morel)**.** *Let E be a set with finite perimeter in $\mathbb{R}^n$. Then there exists a unique finite or countable family of pairwise disjoint indecomposable sets $\{E_i\}_{i\in I}$ such that*

$$\mathcal{H}^n(E_i) > 0 \quad \text{and} \quad P(E) = \sum_{i\in I} P(E_i).$$

Moreover, denoting with

$$\mathring{E}^M := \left\{ x \in \mathbb{R}^n | \lim_{r\to 0^+} \frac{|E \cap B(x,r)|}{|B(x,r)|} = 1 \right\}$$

the essential interior of the set E, it holds

$$\mathcal{H}^{n-1}\left(\mathring{E}^M \setminus \bigcup_{i\in I} \mathring{E}_i^M \right) = 0$$

and the E_i's are maximal indecomposable sets, i.e. any indecomposable set $F \subseteq E$ is contained, up to $\mathcal{H}^n$-negligible sets, in some set E_i.

The chapter is organized as follows.

In Section 6.1 we prove the decomposition theorem for Lipschitz functions.

In Section 6.2 we generalize the decomposition theorem to BV functions and show that this decomposition can generate monotone BV functions without mutually singular distributional derivatives.

In Section 6.3 we give two counterexamples: the first to the fact that a monotone function is always a BV function, the second to a further extension of the Theorem 6.5 to vector valued functions. We also give a proof of the fact that for Lipschitz functions Definition 6.1 and Definition 6.4 are equivalent.

6.1 The Decomposition Theorem for Lipschitz functions from $\mathbb{R}^n$ to $\mathbb{R}$

Before proving the decomposition theorem for Lipschitz functions we state some results on the structure of their level sets.

We first set some notations.

Let $f : \mathbb{R}^n \to \mathbb{R}$ belong to $\mathrm{Lip}_c(\mathbb{R}^n)$. For every $t \in \mathbb{R}$ we call $E_t := \{x|\, f(x) = t\}$, we denote with $\mathcal{C}_t$ the family of all connected components C of E_t such that $\mathcal{H}^{n-1}(C) > 0$ and we denote with E_t^c the union of all C in $\mathcal{C}_t$.

Theorem 6.7. *Let $f : \mathbb{R}^n \to \mathbb{R}$ belong to $\mathrm{Lip}_c(\mathbb{R}^n)$ and have compact support. Then for almost every $t \in \mathbb{R}$*

i) E_t is $\mathcal{H}^{n-1}$-rectifiable and $\mathcal{H}^{n-1}(E_t) < +\infty$;

ii) the map f is differentiable in x for $\mathcal{H}^{n-1}$-a.e. $x \in E_t$;

iii) the family $\mathcal{C}_t$ of open connected components of E_t is countable and $\mathcal{H}^{n-1}(E_t \setminus E_t^c) = 0$.

Proof. We refer to Theorem 2.5 in [2]. □

Lemma 6.8. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz function with compact support. Then the set E_t^c for any $t \in \mathbb{R}$ and the set $E^c := \cup_{t\in\mathbb{R}} E_t^c$ are a countable union of closed sets in $\mathbb{R}^n$; in particular they are Borel measurable.*

Proof. We refer to Lemma 6.1 in [2]. □

We show now the proof of Theorem 6.2 as presented to us by Alberti, Bianchini and Crippa.

Proof of Theorem 6.2. We divide the proof in several steps.

Step 1. Assume that $f \geq 0$ and that f has a strictly positive maximum. Take a countable dense sequence of $0 < t_i \in \mathbb{R}^+$, such that the conclusions of Theorem 6.7 hold. For $0 < t_i < \max f$, let G_i be the connected unbounded open component of $\{f < t_i\} := \{x \in \mathbb{R}^n|\, f(x) < t_i\}$, note that due to the fact that f has compact support there can be only one of such components. Let F_i be the compact set $\mathbb{R}^n \setminus G_i$, and decompose it into the connected compact components F_{ij} with positive $\mathcal{H}^{n-1}$-measure. At least one of such components exists thanks to the choice of the t_i's. It is clear that for $t_i \leq t_{i'}$

$$G_i \subseteq G_{i'}, \quad d_H(F_i, F_{i'}) \geq \frac{1}{c}|t_i - t_{i'}|, \tag{6.1}$$

the first holds because $G_i \subset \{f < t_{i'}\}$ and G_i is unbounded thus $G_i \subset G_{i'}$, the second follows from the Lipschitz estimate, when d_H is the Hausdorff distance.

Moreover $G_{ij} = \mathbb{R}^n \setminus F_{ij}$ is open and connected: in fact, is the complement of a closed set and it can be written as the union of G_i with the neighborhoods of each connected components $C \subset E_t \setminus F_{ij}$ not intersecting F_{ij}.

Step 2. Define the following partial order relation on the countable family F_{ij} as

$$F_{ij} \leq F_{i'j'} \quad \text{if} \quad t_i \leq t_{i'}, F_{ij} \supseteq F_{i'j'}.$$

Let F_{ij}, $i \in I$, $j \in J$ be a maximal countable ordered sequence. Note that we do not need the Axiom of Choice here because the family is countable.

From the definition of partial order the index j must be a function of i: $j = j(i)$, $i \in I$. Since the sequence must be maximal the sequence t_i is dense in a segment $[0, \bar{t}\,]$, by (6.1). Here $\bar{t} := \max f$.

Step 3. Define the function $\tilde{f}$ by

$$\tilde{f}(x) := \sup \Big\{ t_i \big|\ x \in F_{ij(i)}, i \in I \Big\}.$$

By the definition of $\tilde{f}$, $\tilde{f}(x) - \tilde{f}(y) \geq k$ implies that for all $\varepsilon > 0$ there are $i, i' \in I$ such that $y \notin F_{ij(i)}$, $x \in F_{i'j(i')}$ and $t_{i'} - t_i > k - \varepsilon$. Hence

$$|\tilde{f}(x) - \tilde{f}(y)| \leq |t_i - t_{i'}| \leq c|x - y|.$$

Thus $\tilde{f}$ is c-Lipschitz. Since the sets F_{ij} are uniformly bounded, $\tilde{f}$ has bounded support and $\max \tilde{f} = \bar{t} = \max f$.

Step 4. We observe that for each $0 \leq t_i \leq \bar{t}$ we have $\{\tilde{f} \geq t_i\} = F_{i,j(i)}$.

Moreover for $0 \leq t_i < t_{i'} \leq \bar{t}$ the set

$$\{t_i \leq \tilde{f} < t_{i'}\} = F_{ij(i)} \setminus F_{i'j(i')} = F_{ij(i)} \cap G_{i'j(i')}$$

is arc connected, because $G_{i'j(i')}$ is open connected, hence arc connected, and $d_H(G_{i'j(i')}, F_{ij(i)}) > 0$.

It follows that its closure $\{t_i \leq \tilde{f} \leq t_{i'}\}$ is compact connected, and the intersection as $t_i \nearrow t$, $t_i \searrow t$ is connected.

The case $\{0 \leq f < t_i\} = \cap_{t_i \searrow 0} \overline{G}_{ij(i)}$ can be treated similarly because of the ordering of $\overline{G}_{ij(i)}$ and the compactness connectedness of $\overline{G}_{ij(i)} \cap B(0, R)$, for $R \gg 1$.

Therefore $\tilde{f}$ is monotone.

Step 5. We now use the fact that for each i, j one has $\partial F_{ij} \subset E^c_{t_i}$ by construction. Let $\tilde{E}_t := \{\tilde{f} = h\}$ be a level set with empty interior: hence each $x \in \tilde{E}_t$ is the limit of a sequence of points in $\cup_i \partial F_{ij(i)}$, and by the continuity of f, $\tilde{f}$ we conclude that $f = \tilde{f}$ on $\tilde{E}_t$.

In particular if $\tilde{E}^c$ is defined as in Lemma 6.8 for the function $\tilde{f}$, then $\tilde{E}^c \cap E_t = \tilde{E}_t$, and by the Coarea Formula

$$\int_{\tilde{E}^c} \nabla f(x) d\mathcal{H}^n(x) = \int_{\mathbb{R}} \mathcal{H}^{n-1}(\tilde{E}^c \cap E_t) d\mathcal{H}^1(t) = \int_{\mathbb{R}} \mathcal{H}^{n-1}(\tilde{E}^c \cap \tilde{E}_t) d\mathcal{H}^1(t) = \int_{\tilde{E}^c} \nabla \tilde{f}(x) d\mathcal{H}^n(x).$$

We conclude thus that $\nabla f = \nabla \tilde{f}$ $\mathcal{H}^n$-a.e. on $\tilde{E}^c$.

Step 6. Since $\nabla \tilde{f} = 0$ $\mathcal{H}^n$-a.e. on $\mathbb{R}^n \setminus \tilde{E}^c$, we conclude that $f' = f - \tilde{f}$ is again c-Lipschitz, but its total variation is diminished by the total variation of $\tilde{f}$.

Since the total variation of f is bounded, there is at most a countable family of $f_i \neq 0$ such that $f = \sum_i f_i$, and if we denote with $E^c(i)$ the sets defined in Lemma 6.8 for the Lipschitz function f_i, then $E^c(i) \setminus E^c(j)$ is still a Borel set. Moreover, $\nabla f_j = 0$ $\mathcal{H}^n$-a.e. on $E^c(i)$ for $i \neq j$, so that $f_{j\sharp}\mathcal{H}^n \llcorner E^c(i)$ is singular w.r.t. $\mathcal{H}^1$.

The proof is complete by defining $\Omega_i := E^c(i) \setminus (\cup_{j<i} E^c(j))$.

□

6.2 The Decomposition Theorem for BV functions from $\mathbb{R}^n$ to $\mathbb{R}$

To generalize the Jordan decomposition property, let us concentrate on functions $f : \mathbb{R}^n \to \mathbb{R}$, which belong to $BV(\mathbb{R}^n)$. From now on $n > 1$.

Since we will consider functions of bounded variation, the Definition 6.4 of monotone function becomes the following:

Definition 6.9. A BV function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *monotone* if the super-level sets $\{f > t\} = \{x \in \mathbb{R}^n | \ f(x) > t\}$ and the sub-level sets $\{f < t\} = \{x \in \mathbb{R}^n | \ f(x) < t\}$ are indecomposable, for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}$.

Indeed, we recall that, for BV functions, super-level sets and sub-level sets are of finite perimeter for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}$.

We now prove the main theorem of this chapter.

Proof of Theorem 6.5. The proof will be given in several steps.

Before entering into details, let us consider the following simple case.
Let $f = \chi_E$ with $E \subseteq \mathbb{R}^n$ a decomposable set of finite perimeter such that $\mathbb{R}^n \setminus E$ is indecomposable. Thanks to Theorem 6.6, there exists a unique finite or countable family of pairwise disjoint indecomposable sets $\{E_i\}_{i\in I}$ such that

$$\mathcal{H}^n(E_i) > 0 \quad \text{and} \quad P(E) = \sum_{i \in I} P(E_i).$$

To see the properties of $\mathbb{R}^n \setminus E_i$ let us consider the following lemma.

Lemma 6.10. *Let E be a decomposable set of finite perimeter such that $\mathbb{R}^n \setminus E$ is indecomposable. Let $\{E_i\}_{i\in I}$ be the family of its indecomposable components given by Theorem 6.6. Then $\mathbb{R}^n \setminus E_i$ is indecomposable for every $i \in I$.*

Proof. Let $\hat{\imath} \in I$ be fixed. Without loss of generality we can relabel $\hat{\imath} = 1$.

By contradiction, suppose $\mathbb{R}^n \setminus E_1$ is decomposable and let $\{F_j\}_{j\in J}$ be the family of its indecomposable components given by Theorem 6.6.

It holds

$$\mathbb{R}^n \setminus E_1 = (\mathbb{R}^n \setminus E) \cup \bigcup_{i \in I, i \neq 1} E_i \pmod{\mathcal{H}^n},$$

where, we recall, $(\mathbb{R}^n \setminus E) \cup \{E_i\}_{i \in I, i \neq 1}$ is a family of indecomposable and pairwise disjoint sets.

From the maximal indecomposability of $\{F_j\}_{j\in J}$ and $\{E_i\}_{i\in I}$, it follows that

$$\exists!\ \hat{j}\in J \quad \text{s.t.} \quad \mathbb{R}^n\setminus E\subseteq F_{\hat{j}} \pmod{\mathcal{H}^n}$$

and

$$\forall j\in J,\ j\neq\hat{j},\ \exists!\ i\in I,\ i\neq 1, \quad \text{s.t.} \quad F_j=E_i \pmod{\mathcal{H}^n}.$$

We relabel $\hat{j}=1$.
Moreover, we can found two sub-families $\{E_{i_l}\}_{l\in L}$ and $\{E_{i_k}\}_{k\in K}$ of $\{E_i\}_{i\in I}$ such that

$$\{E_i\}_{i\in I}=\{E_{i_l}\}_{l\in L}\cup\{E_{i_k}\}_{k\in K},$$

and

$$F_1=(\mathbb{R}^n\setminus E)\cup\bigcup_{l\in L}E_{i_l} \pmod{\mathcal{H}^n},$$

$$\forall k\in K \quad \exists! j\neq 1\in J \quad \text{s.t.} \quad E_{i_k}=F_j \pmod{\mathcal{H}^n}.$$

Observe that

$$\mathbb{R}^n\setminus F_1=E_1\cup\bigcup_{k\in K}E_{i_k} \pmod{\mathcal{H}^n},$$

where $\{E_1,E_{i_k}\ k\in K\}$ is precisely the family of indecomposable sets given by Theorem 6.6. Therefore

$$P(\mathbb{R}^n\setminus F_1)=P(E_1)+\sum_{k\in K}P(E_{i_k}).$$

On the other hand

$$\begin{aligned}P(\mathbb{R}^n\setminus E_1)=&\sum_{j\in J}P(F_j)\\ =&P(F_1)+\sum_{k\in K}P(E_{i_k}),\end{aligned}$$

thus

$$P(E_1)=P(E_1)+2\sum_{k\in K}P(E_{i_k}).$$

This implies

$$\sum_{k\in K}P(E_{i_k})=\sum_{j\in J,j\neq 1}P(F_j)=0,$$

i.e. $\mathbb{R}^n\setminus E_1$ is equal to F_1, up to $\mathcal{H}^n$-negligible sets.
Therefore $\mathbb{R}^n\setminus E_1$ must be indecomposable. □

From this lemma, for every $i\in I$, E_i and $\mathbb{R}^n\setminus E_i$ are indecomposable. Therefore the functions χ_{E_i} are $BV(\mathbb{R}^n)$ and monotone, so that the decomposition of χ_E,

$$\chi_E=\sum_{i\in I}\chi_{E_i},$$

gives $|D\chi_E|=\sum_{i\in I}|D\chi_{E_i}|$ as required.

Step 0. We can assume without loss of generality that $f\geq 0$: in the general case one can decompose f^+ and f^- separately.

Step 1. The sets $E^t := \{f > t\}$ are of finite perimeter for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}^+$, thanks to the hypothesis that f is $BV(\mathbb{R}^n)$ and Coarea Formula. Therefore, Theorem 6.6 gives, for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}^+$, pairwise disjoint indecomposable sets $\{E_i^t\}_{i\in I_t}$ such that

$$\mathcal{H}^n\left(E^t \setminus \bigcup_{i\in I_t} E_i^t\right) = 0.$$

In particular, the property of maximal indecomposability yields a natural partial order relation between these sets: since $t_1 \geq t_2$ gives $E^{t_1} \subseteq E^{t_2}$, it follows that, for $\mathcal{H}^1$-a.e. $t_1 \geq t_2 \in \mathbb{R}^+$,

$$\forall i \in I_{t_1}\ \exists!\ i' \in I_{t_2} \quad \text{s.t.} \quad E_i^{t_1} \subseteq E_{i'}^{t_2} \pmod{\mathcal{H}^n}.$$

Taken a countable dense subset $\{t_j\}_{j\in J}$ of $\mathbb{R}^+$, such that, for all $j \in J$, the sets $E^j := E^{t_j}$ are of finite perimeter, the countable family $\{E_i^j\}_{j\in J, i\in I_{t_j}}$ can be equipped with the partial order relation

$$E_i^j \leq E_{i'}^{j'} \iff t_j \leq t_{j'},\ E_i^j \supseteq E_{i'}^{j'} \pmod{\mathcal{H}^n}.$$

Therefore there exists at least one maximal countable ordered sequence (here we do not need the Axiom of Choice).

Let $\{E_{i(j)}^j\}_{j\in J}$ one of these maximal countable ordered sequences.
Notice that, once one of these sequences is fixed, the index i is a function of j, by the uniqueness of the decomposition $\{E_i^j\}_{i\in I_{t_j}}$.
Step 2. Define

$$\tilde{f}(x) := \begin{cases} 0 & x \notin \bigcup_{j\in J} E_{i(j)}^j \\ \sup\{t_j|\ j \in J, x \in E_{i(j)}^j\} & \text{otherwise} \end{cases}$$

Clearly $0 \leq \tilde{f}(x) \leq f(x)$ for all $x \in \mathbb{R}^n$. Indeed, the set

$$\left\{t_j|\ j \in J, x \in E_{i(j)}^j\right\} \subseteq \left\{t_j|\ j \in J, x \in E^j\right\} \quad \forall x \in \mathbb{R}^n,$$

passing to the supremum one has $\tilde{f}(x) \leq f(x)$ for all $x \in \mathbb{R}^n$. Moreover $f \in \mathrm{L}^1_{loc}(\mathbb{R}^n)$ and $0 \leq \tilde{f} \leq f$ give $\tilde{f} \in \mathrm{L}^1_{loc}(\mathbb{R}^n)$.
Step 3. Fix $t \in \mathbb{R}^+$ such that E^t is a set of finite perimeter. Define $\tilde{E}^t := \{\tilde{f} > t\}$ and let $E_{i(t)}^t$ the indecomposable component of E^t which is contained in a set $E_{i(j)}^j$ of the maximal countable ordered sequence and contains another $E_{i(j')}^{j'}$, for certain $j, j' \in J$, up to $\mathcal{H}^n$-negligible sets. This is possible for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}^+$.

Due to the maximal indecomposability property, one has that

$$E_{i(j')}^{j'} \subseteq E_{i(t)}^t \subseteq E_{i(j)}^j \pmod{\mathcal{H}^n} \quad \forall t_{j'}, t_j,$$

where $t_{j'} > t > t_j$.
Notice that, for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}^+$, there exists only one of such an $E_{i(t)}^t$ among all the indecomposable sets $E_i^t, i \in I_t$.

We show that $\tilde{E}^t = E_{i(t)}^t \pmod{\mathcal{H}^n}$, for $\mathcal{H}^1$-a.e. t in $\mathbb{R}^+$, in two steps.

- First we show that $\tilde{E}^t \subseteq E^t_{i(t)}$ (mod $\mathcal{H}^n$) for $\mathcal{H}^1$-a.e t in $\mathbb{R}^+$.

  For $x \in \tilde{E}^t = \{\tilde{f} > t\}$, there exist $j_1 = j_1(x), j_2 = j_2(x)$ such that

  $$\tilde{f}(x) > t_{j_1} > t > t_{j_2} \quad \text{and} \quad x \in E^{j_1}_{i(j_1)} \cap E^{j_2}_{i(j_2)}.$$

  Since for all $t_{j_1} > t > t_{j_2}$ it holds

  $$E^{j_1}_{i(j_1)} \subseteq E^t_{i(t)} \subseteq E^{j_2}_{i(j_2)} \ (\text{mod } \mathcal{H}^n),$$

  it follows that for $\mathcal{H}^n$-a.e $x \in \tilde{E}^t$ $x \in E^t_{i(t)}$, hence

  $$\tilde{E}^t \subseteq E^t_{i(t)} \ (\text{mod } \mathcal{H}^n).$$

- Next we show the other inclusion up to countably many values of t.

  Observe that set $E^t_{i(t)}$ is contained in $\tilde{E}^{t'}$ for all $t' < t$. In fact $x \in E^t_{i(t)}$ implies $f(x) > t > t_j > t'$ for some $j \in J$, hence $\tilde{f}(x) \geq t_j > t'$. Thus for every $t'_n \nearrow t$ one has $\bigcap_{t'_n<t} \tilde{E}^{t'_n} \supseteq E^t_{i(t)}$.

  Suppose $\mathcal{H}^n(E^t_{i(t)} \setminus \tilde{E}^t) > 0$: from $\tilde{E}^t \subseteq E^t_{i(t)}$ it follows

  $$0 < \mathcal{H}^n\left(\bigcap_{t'_n<t} \tilde{E}^{t'_n} \setminus \tilde{E}^t\right) = \mathcal{H}^n\left(\{\tilde{f} \geq t\} \setminus \tilde{E}^t\right)$$

  and this implies $\mathcal{H}^n(\{\tilde{f} = t\}) > 0$. This last condition can be satisfied only for a countable number of $t \in \mathbb{R}^+$.

Therefore the set of t's such that $E^t_{i(t)}$ does not coincide with $\tilde{E}^t$ has zero n-dimensional Hausdorff measure, i.e. for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}^+$ the sets $\tilde{E}^t$ coincide with $E^t_{i(t)}$ up to $\mathcal{H}^n$-negligible sets. Since the property of being indecomposable is invariant up to $\mathcal{H}^n$-negligible sets, they are indecomposable.

In the following we will denote with $\tilde{t}_k,\ k \in K$, the countable family of values such that

$$H_k := \{\tilde{f} = \tilde{t}_k\}, \quad \mathcal{H}^n(H_k) > 0.$$

Step 4. The function $\tilde{f}$ is $BV(\mathbb{R}^n)$ and has indecomposable super-level sets.

The indecomposability of the super-level sets of $\tilde{f}$ was proved in the previous step.

Using Coarea Formula, see for example Theorem 2.93 of [5], we get

$$\begin{aligned}
|D\tilde{f}| &= \int_{-\infty}^{+\infty} P(\{\tilde{f} > t\})dt \\
&= \int_{-\infty}^{+\infty} P(E^t_{i(t)})dt \\
&\leq \int_{-\infty}^{+\infty} P(E^t)dt \\
&= |Df| < +\infty.
\end{aligned}$$

Thus the function $\tilde{f}$ is $BV(\mathbb{R}^n)$.

Step 5. Define the function $\hat{f} := f - \tilde{f}$. Clearly $\hat{f}$ is $BV(\mathbb{R}^n)$. The aim of the following steps is to show that its total variation satisfies

$$|D\hat{f}| = |Df| - |D\tilde{f}|.$$

Denote with E_1^t the super-level sets used to generate the function $\tilde{f}$: this can be done setting $i(t) = 1$ for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}^+$.

It has been proved that, for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}^+$, one has $\{\tilde{f} > t\} = E_1^t$, up to $\mathcal{H}^n$-negligible sets, therefore for such t's

$$\begin{aligned} P(\{f > t\}) &= \sum_{i \in I_t} P(E_i^t) \\ &= \sum_{i \in I_t,\ i > 1} P(E_i^t) + P(\{\tilde{f} > t\}). \end{aligned}$$

We would like to show that, for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}^+$, for every $i \in I_t,\ i > 1$, E_i^t is equal, up to $\mathcal{H}^n$-negligible sets, to one of the indecomposable components $\hat{E}_i^{\hat{t}}$ of $\{\hat{f} > \hat{t}\ \}$, where $\hat{t} = t - \tilde{t}_i$ for a certain $\tilde{t}_i$.
The index i in $\tilde{t}_i$ refers to the fact that its value varies with the indecomposable component E_i^t, $i \in I_t, i > 1$.

We prove it in the following three steps.

Step 6. Let t be such that the set E^t is of finite perimeter and $\{E_i^t\}_{i \in I_t}$ are its indecomposable components.

Let us prove that there exists a unique $k \in K$ such that the set E_i^t, $i \in I_t, i > 1$, is contained in H_k, up to $\mathcal{H}^n$-negligible sets.

The set E_i^t is indecomposable and $E_i^t \cap E_1^t = \emptyset$. Being $E_1^j \subseteq E_1^t$ for all $t_j \geq t$, up to $\mathcal{H}^n$-negligible sets, it follows

$$\mathcal{H}^n\left(E_i^t \cap E_1^j\right) = 0\ \forall t_j \geq t.$$

Therefore, from the definition of $\tilde{f}$, for $\mathcal{H}^n$-a.e. $x \in E_i^t$ one has $\tilde{f}(x) \leq t$.

Again from the indecomposability of E_i^t and from the fact that E_i^t is contained in $\{f > t_j\}$ for all $t_j \leq t$, it follows that there exists a unique $l \in I_{t_j}$ such that,

$$E_i^t \subseteq E_l^j \ (\mathrm{mod}\ \mathcal{H}^n) \quad \text{and} \quad \mathcal{H}^n\left(E_i^t \cap E_m^j\right) = 0\ \forall m \neq l,\ m \in I_{t_j},$$

for all $t_j \leq t$.
If there exists a j' such that $\mathcal{H}^n\left(E_i^t \cap E_1^{j'}\right) = 0$ then

$$\forall t_j,\ 0 \leq t_{j'} \leq t_j \leq t \quad \mathcal{H}^n\left(E_i^t \cap E_1^j\right) = 0,$$

on the other hand if there exists a j'' such that $E_i^t \subseteq E_1^{j''}$, up to $\mathcal{H}^n$-negligible sets, then

$$\forall t_j,\ 0 \leq t_j \leq t_{j''} \quad E_i^t \subseteq E_1^j\ (\mathrm{mod}\ \mathcal{H}^n).$$

Thus, being the definition

$$\tilde{f}(x) := \begin{cases} 0 & x \notin \bigcup_{j \in J} E_1^j \\ \sup\{t_j \mid j \in J, x \in E_1^j\} & \text{otherwise} \end{cases}$$

equivalent to

$$\tilde{f}(x) := \inf\{t_j \mid j \in J, x \notin E_1^j\},$$

it follows that, up to $\mathcal{H}^n$-negligible subsets of E_i^t, $\tilde{f}\mid_{E_i^t} = constant$, which belongs to $\{\tilde{t}_k\}_{k \in K}$.

In particular, we can order the sets $E_i^t, i \in I_t, i > 1$, as $E_{(k,i)}^t$ where

$$\{E_{(k,i)}^t \mid i \in B_k^t\} = \Big\{E_i^t \mid i \in I_t,\ i > 1,\ \ E_i^t \subseteq H_k \ (\mathrm{mod}\ \mathcal{H}^n)\Big\}.$$

Note that B_k^t could be empty for some $t \in \mathbb{R}^+,\ k \in K$.

Step 7. Let $\hat{t} > 0$ such that the set $\hat{E}^{\hat{t}}$ is of finite perimeter and $\{\hat{E}_i^{\hat{t}}\}_{i \in \hat{I}_{\hat{t}}}$ are its indecomposable components, for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}^+$.

Let us prove that there exists a unique $k \in K$, such that the set $\hat{E}_i^{\hat{t}}$ is contained in H_k, up to $\mathcal{H}^n$-negligible sets.

Define

$$\bar{t} := \sup\big\{0,\ t_j \mid j \in J,\ \hat{E}_i^{\hat{t}} \subseteq E_1^j \ (\mathrm{mod}\ \mathcal{H}^n)\big\}.$$

It follows that

$$f|_{\hat{E}_i^{\hat{t}}} = \hat{f}|_{\hat{E}_i^{\hat{t}}} + \tilde{f}|_{\hat{E}_i^{\hat{t}}} > \hat{t} + \bar{t} > \bar{t}.$$

For every t_j in the countable dense sequence such that $\bar{t} < t_j < \bar{t} + \hat{t}$ there exists a unique $\bar{i} \in I_{t_j}$ such that

$$\hat{E}_i^{\hat{t}} \subseteq E_{\bar{i}}^j \ (\mathrm{mod}\ \mathcal{H}^n).$$

Due to the indecomposability of $\hat{E}_i^{\hat{t}}$, and, for the definition of $\bar{t}$, the index $\bar{i}$ must be greater than 1.

Therefore $\tilde{f}|_{\hat{E}_i^{\hat{t}}} = \bar{t}$ and $\bar{t}$ belongs to $\{\tilde{t}_k\}_{k \in K}$.

In particular, we can order the sets $\hat{E}_i^{\hat{t}},\ i \in \hat{I}_{\hat{t}}$, as $\hat{E}_{(k,i)}^{\hat{t}}$ where

$$\{\hat{E}_{(k,i)}^{\hat{t}} \mid i \in \hat{B}_k^{\hat{t}}\} = \Big\{\hat{E}_i^{\hat{t}} \mid i \in \hat{I}_{\hat{t}},\ \hat{E}_i^{\hat{t}} \subseteq H_k \ (\mathrm{mod}\ \mathcal{H}^n)\Big\}.$$

Note that $\hat{B}_k^{\hat{t}}$ could be empty for some $\hat{t} \in \mathbb{R}^+,\ k \in K$.

Step 8. In this step we prove that, for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}^+$, $k \in K$ fixed,

$$\{E_{(k,i)}^t \mid i \in B_k^t\} = \{\hat{E}_{(k,i)}^{t-\tilde{t}_k} \mid i \in \hat{B}_k^{t-\tilde{t}_k}\}.$$

Indeed, fix $i \in B_k^t$

$$\hat{f}|_{E_{(k,i)}^t} = f|_{E_{(k,i)}^t} - \tilde{f}|_{E_{(k,i)}^t} > t - \tilde{t}_k.$$

Let us consider only the t's such that the set $\{\hat{f} > t - \tilde{t}_k\}$ is of finite perimeter.

For its indecomposability, $E^t_{(k,i)}$ must be contained, up to $\mathcal{H}^n$-negligible sets, in $\hat{E}^{t-\tilde{t}_k}_{(k,i')}$ for a unique $i' \in \hat{I}_{t-\tilde{t}_k}$.
Take then the set $\hat{E}^{t-\tilde{t}_k}_{(k,i')}$:

$$f|_{\hat{E}^{t-\tilde{t}_k}_{(k,i')}} = \tilde{f}|_{\hat{E}^{t-\tilde{t}_k}_{(k,i')}} + \hat{f}|_{\hat{E}^{t-\tilde{t}_k}_{(k,i')}} > \tilde{t}_k + t - \tilde{t}_k = t.$$

For its indecomposability, $\hat{E}^{t-\tilde{t}_k}_{(k,i')}$ must be contained, up to $\mathcal{H}^n$-negligible sets, in $E^t_{(k,i'')}$ for a unique $i'' \in I_t, i'' > 1$. Thus $i'' = i$ and $E^t_{(k,i)} = \hat{E}^{t-\tilde{t}_k}_{(k,i')}$, up to $\mathcal{H}^n$-negligible sets.
Hence

$$\{E^t_{(k,i)} |\ i \in B^t_k\} \subseteq \{\hat{E}^{t-\tilde{t}_k}_{(k,i)} |\ i \in \hat{B}^{t-\tilde{t}_k}_k\}.$$

The same argument, reversed, shows that, once $i' \in \hat{B}^{t-\tilde{t}_k}_k$ is fixed, $\hat{E}^{t-\tilde{t}_k}_{(k,i')} = E^t_{(k,i)}$, up to $\mathcal{H}^n$-negligible sets, for a certain $i \in B^t_k$. Hence

$$\{E^t_{(k,i)} |\ i \in B^t_k\} \supseteq \{\hat{E}^{t-\tilde{t}_k}_{(k,i)} |\ i \in \hat{B}^{t-\tilde{t}_k}_k\}.$$

In an equivalent way, we can also say that, for $\mathcal{H}^1$-a.e. $\hat{t} \in \mathbb{R}^+$, $k \in K$ fixed,

$$\{\hat{E}^{\hat{t}}_{(k,i)} |\ i \in \hat{B}^{\hat{t}}_k\} = \{E^{\hat{t}+\tilde{t}_k}_{(k,i)} |\ i \in B^{\hat{t}+\tilde{t}_k}_k\}.$$

In the following we relabel $\hat{E}^{\hat{t}}_{(k,i)}$ and $E^{\hat{t}+\tilde{t}_k}_{(k,i)}$ in order to have

$$\hat{E}^{\hat{t}}_{(k,i)} = E^{\hat{t}+\tilde{t}_k}_{(k,i)} \ (\mathrm{mod}\ \mathcal{H}^n).$$

Step 9. Coarea Formula gives

$$\begin{aligned} |Df| &= \int_{-\infty}^{+\infty} P(\{f > t\})dt \\ &= \int_{-\infty}^{+\infty} \sum_{i \in I_t, i>1} P(E^t_i)dt + \int_{-\infty}^{+\infty} P(\{\tilde{f} > t\})dt. \end{aligned}$$

The final steps consist in showing that

$$\int_{-\infty}^{+\infty} \sum_{i \in I_t, i>1} P(E^t_i)dt = |D\hat{f}|.$$

Step 10. The set $\{\tilde{t}_k |\ k \in K\}$ is the countable set of values such that $\mathcal{H}^n\left(\{\tilde{f} = \tilde{t}_k\}\right) > 0$ for all $k \in K$.

Step 6 shows that, for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}^+$ and for all $i \in I_t,\ i > 1$, there exists a unique $k \in K$ such that $\tilde{f}\ |_{E^t_i} = \tilde{t}_k$.

For every $k \in K$, let $\{E^t_{(k,i)} |\ i \in B^t_k\}$ be the set of indecomposable components of E^t such that $\tilde{f}\ |_{E^t_{(k,i)}} = \tilde{t}_k$, $i > 1$.

Observe that $\sum_{i\in B_k^t} P(E^t_{(k,i)})$ are measurable functions of t, for all $k\in K$: indeed we have

$$\begin{aligned}\left|D((f-\tilde{t}_k)\chi_{H_k})\right| &= \int_{\tilde{t}_k}^{+\infty} \sum_{\substack{i\in I_t, i>1\\ \{f>t\}_i\subseteq\{\tilde{f}=\tilde{t}_k\}}} P(\{f>t\}_i)dt\\ &= \int_{\tilde{t}_k}^{+\infty} \sum_{i\in B_k^t} P(\{f>t\}_i)dt \le |Df|(\mathbb{R}^n) < +\infty.\end{aligned}$$

Therefore the function $t\mapsto \sum_{i\in B_k^t} P(E^t_i)$ is integrable for all $k\in K$.

Using this notation, we can write

$$\begin{aligned}\int_{-\infty}^{+\infty}\sum_{i\in I_t, i>1} P(E^t_i)dt &= \int_{-\infty}^{+\infty}\sum_{k\in K}\sum_{i\in B_k^t} P(E^t_{(k,i)})dt\\ &= \sum_{k\in K}\int_{-\infty}^{+\infty}\sum_{i\in B_k^t} P(E^t_{(k,i)})dt\\ &= \sum_{k\in K}\int_{-\infty}^{+\infty}\sum_{i\in \hat{B}_k^{t-\tilde{t}_k}} P(\{\hat{f}>t-\tilde{t}_k\ \}_{(k,i)})dt\\ &= \sum_{k\in K}\int_{-\infty}^{+\infty}\sum_{i\in \hat{B}_k^{\hat{t}}} P(\{\hat{f}>\hat{t}\ \}_{(k,i)})d\hat{t}\\ &= \int_{-\infty}^{+\infty}\sum_{k\in K}\sum_{i\in \hat{B}_k^{\hat{t}}} P(\{\hat{f}>\hat{t}\ \}_{(k,i)})d\hat{t}.\end{aligned}$$

From Step 7 it holds

$$\begin{aligned}\hat{E}^{\hat{t}} &= \bigcup_i \{\hat{E}^{\hat{t}}_i \mid i\in \hat{I}_{\hat{t}}\}\\ &= \bigcup_i\bigcup_{k\in K}\{\hat{E}^{\hat{t}}_{(k,i)} \mid \tilde{f}\ |_{\hat{E}^{\hat{t}}_i} = \tilde{t}_k, i\in \hat{I}_{\hat{t}}\}\\ &= \bigcup_{k\in K}\bigcup_i \{\hat{E}^{\hat{t}}_{(k,i)} \mid i\in \hat{B}^{\hat{t}}_k\},\end{aligned}$$

we can write

$$\begin{aligned}\int_{-\infty}^{+\infty}\sum_{k\in K}\sum_{i\in \hat{B}_k^{\hat{t}}} P(\{\hat{f}>\hat{t}\ \}_{(k,i)})d\hat{t} &= \int_{-\infty}^{+\infty}\sum_{i\in \hat{I}_{\hat{t}}} P(\{\hat{f}>\hat{t}\}_i)d\hat{t}\\ &= \int_{-\infty}^{+\infty} P(\{\hat{f}>\hat{t}\})d\hat{t} = |D\hat{f}|.\end{aligned}$$

Step 11. Finally we have

$$\begin{aligned}|Df| &= \int_{-\infty}^{+\infty} P(\{f > t\})dt \\ &= \int_{-\infty}^{+\infty} P(\{\hat{f} > t\})dt + \int_{-\infty}^{+\infty} P(\{\tilde{f} > t\})dt \\ &= |D\hat{f}| + |D\tilde{f}|.\end{aligned}$$

Since f has bounded variation we can iterate this process at most a countable number of times generating the family of functions $\tilde{f}_l \in BV(\mathbb{R}^n)$, such that everyone of them has indecomposable super-level sets, for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}^+$.

Step 12. Let $\tilde{f} := \tilde{f}_l$ be one of the functions generated in the previous steps.

If $\{\tilde{f} < t\}$ is indecomposable for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}^+$, then $\tilde{f}$ is already monotone. Otherwise we must again decompose $\tilde{f}$. If we succeed in decomposing $\tilde{f}$ in a countable sum of monotone BV functions which preserves total variation we are done, since the decomposition of every function of a countable family in a countable family gives at the end a countable family as required.

In that case define $\tilde{F}^t := \{\tilde{f} < t\}$ and let $\{\tilde{F}^t_i\}_{i \in I_t}$ be the family of indecomposable sets given by Theorem 6.6 for $\mathcal{H}^1$-a.e. t in $\mathbb{R}^+$.
As for the super-level sets, we equip the family $\{\tilde{F}^j_i\}_{i \in I_{t_j}}$ with the natural partial order relation

$$\tilde{F}^j_i \leq \tilde{F}^{j'}_{i'} \iff t_j \geq t_{j'},\ \tilde{F}^j_i \supseteq \tilde{F}^{j'}_{i'} \pmod{\mathcal{H}^n}$$

and call $\{\tilde{F}^j_1\}_{j \in J}$ one of the maximal countable ordered sequences.

Define

$$\tilde{\tilde{f}}(x) := \inf\{t_j |\ j \in J, x \in \tilde{F}^j_1\}.$$

As in the previous case, one has that

- $\tilde{\tilde{f}}$ is $BV(\mathbb{R}^n)$,
- $\{\tilde{\tilde{f}} < t\} = \tilde{F}^t_1$ up to $\mathcal{H}^n$-negligible sets and for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}^+$,
- define $\hat{\hat{f}} := \tilde{f} - \tilde{\tilde{f}}$ then $\hat{\hat{f}}$ is $BV(\mathbb{R}^n)$ and

$$|D\tilde{f}| = |D\hat{\hat{f}}| + |D\tilde{\tilde{f}}|.$$

Recall that, for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}^+$, $\{\tilde{f} < t\}$ is decomposable and $\mathbb{R}^n \setminus \{\tilde{f} < t\}$ indecomposable. Since $\{\tilde{f} < t\} = \bigcup_{i \in I_t} \tilde{F}^t_i$ and $\{\tilde{\tilde{f}} < t\} = \tilde{F}^t_1$ up to $\mathcal{H}^n$-negligible sets, Lemma 6.10 implies that $\mathbb{R}^n \setminus \{\tilde{\tilde{f}} < t\}$ is indecomposable, hence the super-level set $\{\tilde{\tilde{f}} > t\}$ is indecomposable for $\mathcal{H}^1$-a.e. $t \in \mathbb{R}^+$. Therefore $\tilde{\tilde{f}}$ is monotone as required.

Since $\tilde{f}$ has bounded variation we can iterate this process at most a countable number of times generating the family of monotone functions $f_i \in BV(\mathbb{R}^n)$, which satisfies the theorem.

□

Remark 6.11. Notice that in Step 10 we have also proved that

$$\hat{f}|_{\bigcup_{k\in K} H_k} = \sum_{k\in K} f|_{H_k} - \tilde{t}_k.$$

We show now another proof of Theorem 6.6, which uses a variational argument.

Proof of Theorem 6.5. This proof is divided into 4 steps.

Step 1. The previous proof shows that if $\{f > t\}$ is indecomposable for $\mathcal{H}^1$-a.e. t in $\mathbb{R}$, then the decomposition

$$-f = \sum_{i\in I} f_i,$$

where $\{f_i > t\}$ is indecomposable for $\mathcal{H}^1$-a.e. t in $\mathbb{R}$, implies that also $\{-f_i > t\}$ is indecomposable. Hence, it is enough to show that there exists a decomposition $f = \sum_i f_i$ such that $\{f_i > t\}$ is indecomposable for $\mathcal{H}^1$-a.e. t in $\mathbb{R}$.

Step 2. Let $f \geq 0$ be a BV function and let E^t_i be an indecomposable component for the super-level set $E^t = \{f > t\}$. Consider the variational problem

$$\inf \left\{ \int |u(x)|dx, u \geq t\chi_{E^t_i}, |Du| + |D(f-u)| = |Df| \right\}.$$

Since $|Du| \leq |Df| < +\infty$, the above problem admits a minimum f_1, and this minimum satisfies $0 \leq f_1 \leq t$.

Step 3. Assume that for some $t > t_1 > 0$ the level set $E = \{f_1 > t_1\}$ is decomposable: let E_1 and $E_2 = E \setminus E_1$ be a decomposition such that $E^t_i \subseteq E_1$ and

$$P(E) = P(E_1) + P(E_2), \quad P(E_1), P(E_2) > 0.$$

Define the truncated function

$$\tilde{f}_1(x) = \begin{cases} f_1(x) & x \in \mathbb{R}^n \setminus E_2 \\ t_1 & x \in E_2. \end{cases}$$

Clearly $\tilde{f}_1 \geq t\chi_{E^t_i}$ and $\|\tilde{f}_1\|_{L^1} < \|f_1\|_{L^1}$. Moreover, since for $t_1 < t_2 < t$ one has $\{f_1 > t_2\} \subseteq \{f_1 > t_1\}$, it follows from the indecomposability that

$$P(\{f_1 > t_2\}) = P\big(\{f_1 > t_2\} \cap E_1\big) + P\big(\{f_1 > t_2\} \cap E_2\big)$$

so that

$$f_1 = \tilde{f}_1 + (f_1 - t_1)\chi_{E_2}, \quad |Df_1| = |D\tilde{f}_1| + |D((f_1 - t_1)\chi_{E_2})|.$$

Hence if $\{f_1 > t_1\}$ is not indecomposable for $0 < t_1 < t$, the the function f_1 can be decomposed as the sum of two positive functions $\tilde{f}_1$, $\hat{f}_1$ such that $\|\tilde{f}_1\|_{L^1}, \|\hat{f}_1\|_{L^1} > 0$ and

$$|Df_1| = |D\tilde{f}_1| + |D\hat{f}_1|.$$

Step 4. The subadditivity of the norm and the previous step implies that

$$\begin{aligned}|D\tilde{f}_1| + |D(f - \tilde{f}_1)| &= |D\tilde{f}_1| + |D\big(f - (\tilde{f}_1 + \hat{f}_1) + \hat{f}_1\big)| \\ &\leq |D\tilde{f}_1| + |D(f - f_1)| + |D\hat{f}_1| \\ &= |Df_1| + |D(f - f_1)|,\end{aligned}$$

and this with the fact that $\|\tilde{f}_1\|_{L^1} < \|f_1\|_{L^1}$ yields a contradiction to the minimality of f_1 . $\square$

Remark 6.12. In general the decomposition of f in BV monotone functions is not unique as the following example shows.

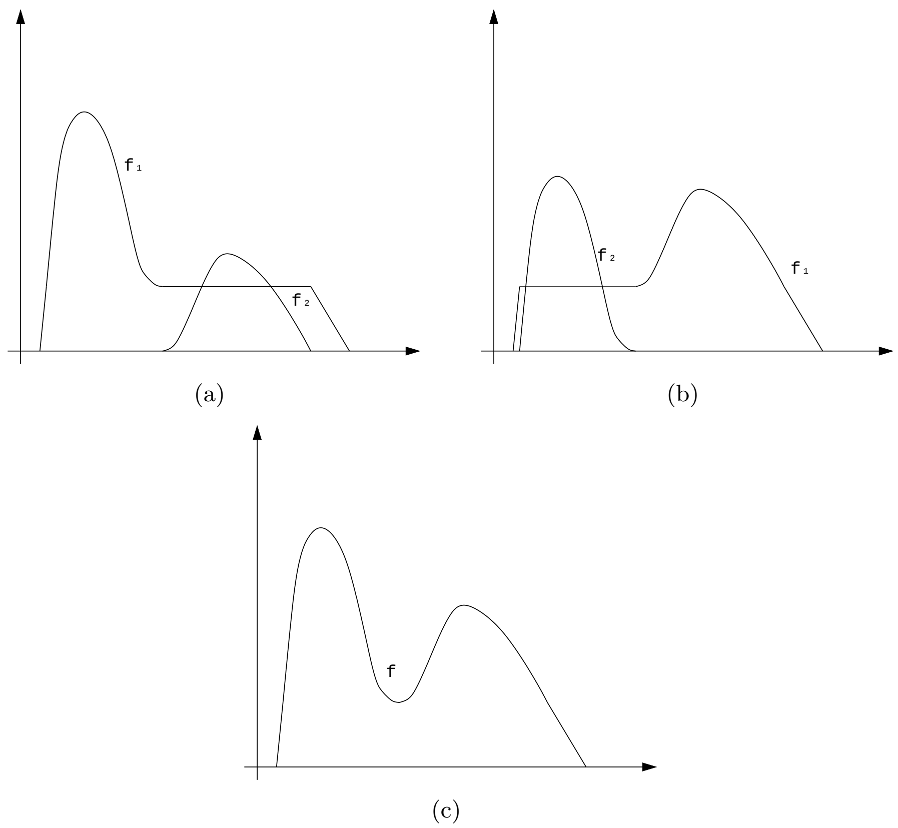

Figure 6.1:

The function f in Figure 6.1(c) can be decomposed either in the way shown in Figure 6.1(a) or in Figure 6.1(b).

In the simple case, where f is the characteristic function of a set of finite perimeter with an indecomposable complementary set, there exists a unique subdivision of f as a countable sum of BV monotone characteristic functions. Moreover in that case, due to the fact that the sets E_i are pairwise disjoint, $D\chi_{E_i}$ are mutually singular for all $i \in I$.

This property, which has been proved also for the decomposition of Lipschitz functions in Theorem 6.2, can be false in the general case. As shown in the example below, one can have monotone BV functions, whose distributional derivatives are concentrated on sets with non empty intersection.

Example 6.13. Let us consider a BV function f as in the Figure 6.2.

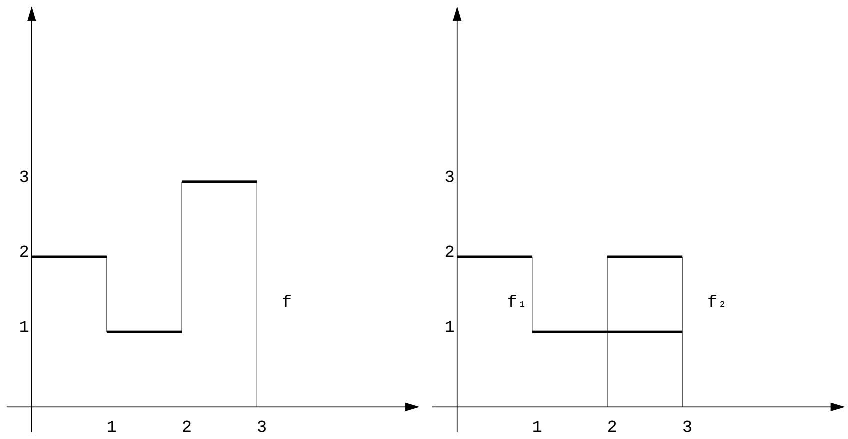

Figure 6.2:

In this case Theorem 6.5 gives two BV monotone functions f_1 and f_2 such that $f = f_1 + f_2$. Their distributional derivatives are

$$|Df_1| = 2\delta_0 - \delta_1 - \delta_3 \quad \text{and} \quad |Df_2| = 2\delta_2 - 2\delta_3,$$

where δ_x is the Dirac measure, $\delta_x(A) = 1$ if x belongs to the set A, $\delta_x(A) = 0$ otherwise. Clearly these distributional derivatives are not mutually singular, since both have an atom in $x = 3$.

One can easily show that for any other monotone decomposition it is impossible to find two disjoint sets on which the distributional derivatives are concentrated.

6.3 Counterexamples

As we have already said, the definition of monotone function could be given even for a function which is only $L^1_{loc}(\mathbb{R}^n)$. In that case one has to require that this function must have super-level sets with finite perimeter, which is true $\mathcal{H}^1$-a.e. $t \in \mathbb{R}$ for the super-level sets of a BV function.

The Jordan decomposition states that monotonicity is a sufficient condition for a function of one variable to be of bounded variation. However, we cannot say that every monotone function $f : \mathbb{R}^n \to \mathbb{R}$ defined as in Definition 6.4 is of bounded variation.

A counterexample is given below by a function, whose super-level sets are progressive configurations of the construction of a Koch snowflake.

Example 6.14. The Koch snowflake is a curve generated iteratively from a unitary triangle T adding each time, on each edge, a smaller centered triangle with edges one third of the previous edge, see Figure 6.3.
More precisely letting T_0 be the equilateral triangle T with unitary edge, and T_i the successive iterations of the curve, one has that at every stage

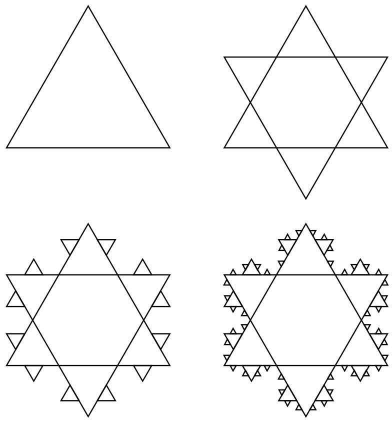

Figure 6.3: Progressive configurations of the construction of a Koch snowflake

- the number of edges is $N_k = 3 \cdot 4^k$,
- the length of the edges is $L_k = \left(\frac{1}{3}\right)^k$,
- the perimeter of the iterated curve is $P(T_k) = 3 \cdot \left(\frac{4}{3}\right)^k$,
- the area of the iterated curve is

$$\mathcal{H}^2(T_k) = \left[1 + \frac{1}{3}\sum_{j=1}^{k}\left(\frac{4}{9}\right)^j\right] \cdot \frac{\sqrt{3}}{2}.$$

Denote with B the ball

$$B = \{x \in \mathbb{R}^2 |\ \|x\| < R\},$$

which contains the unitary triangle T centered in the origin: hence $T_i \subseteq B$ for all $i \in \mathbb{N}$.

Let $E_k := B \setminus T_k$ for $k \in \mathbb{N}$ and define $f : B \to \mathbb{R}$ in this way

$$f(x) := \sum_k \left(\frac{3}{4}\right)^k \chi_{E_k}(x).$$

Clearly $0 \leq f < 4$, therefore f belongs to $L^1(B)$ and Coarea Formula can be used to obtain its variation.

Let us note which are the super-level sets and their perimeter:

- for $t < 0$ the set $\{f > t\} = B$ and $P(B, B) = 0$,
- for $t = 0$ the set $\{f > t\} = E_0$ and $P(E_0, B) = 3$,
- for $0 < t < 4$ the set $\{f > t\} = E_{\bar{k}}$ for the first $\bar{k}$ such that $\sum_{k=0}^{\bar{k}} \left(\frac{3}{4}\right)^k > t$ and $P(E_{\bar{k}}, B) = 3 \cdot \left(\frac{4}{3}\right)^{\bar{k}}$,

- for $t \geq 4$ the set $\{f > t\} = \emptyset$ and $P(\emptyset, B) = 0$.

Thus this function is monotone and computing its variation one has

$$\begin{aligned} |Df|(B) &= \int_{-\infty}^{+\infty} P(\{f > t\}, B) dt \\ &= \int_{0}^{4} P(\{f > t\}, B) dt \\ &= \sum_{k=0}^{+\infty} 3 \cdot \left(\frac{4}{3}\right)^k \cdot \left(\frac{3}{4}\right)^k = +\infty \end{aligned}$$

which implies that f does not belong to $BV(B)$.

In the case of Lipschitz functions Definition 6.1 and Definition 6.4 are equivalent.

Proposition 6.15. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz function, then f is monotone in the sense of Definition 6.1 if and only if f is monotone in the sense of Definition 6.4.*

Proof. ($\Rightarrow$) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz function which is monotone in the sense of Definition 6.1, then for all t in $\mathbb{R}$ the set $\{f = t\}$ is connected.

We claim that $\{f > t\}$ and $\{f < t\}$ are open connected sets. Indeed, let us concentrate on $\{f > t\}$, the other case is similar.
By contradiction suppose $\{f > t\}$ disconnected, then $\{f > t\}$ must have at least two connected components. For $t' > t$, such that $t' - t$ is sufficiently small, the set $\{f = t'\}$ is contained at least in two of the connected components of $\{f > t\}$. Thus we have a connected set $\{f = t'\}$ contained in two connected components of a disconnected set, absurd.

Since for $\mathcal{H}^1$-a.e. t in $\mathbb{R}$ the sets $\{f > t\}$ and $\{f < t\}$ are of finite perimeter Proposition 2 in [3] gives that the open and connected sets $\{f > t\}$ and $\{f < t\}$ are indecomposable for $\mathcal{H}^1$-a.e. t in $\mathbb{R}$.

Therefore f is monotone in the sense of Definition 6.4.

($\Leftarrow$) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz function which is not monotone in the sense of Definition 6.1, then there exists a t in $\mathbb{R}$ such that the set $\{f = t\}$ is disconnected.

For Theorem 6.1.23 in [25], every connected components of $\{f = t\}$ coincides with a quasi-connected component of $\{f = t\}$, because $\{f = t\}$ is compact.

This implies that there exists an open set G in $\mathbb{R}^n$ such that

$$\partial G \cap \{f = t\} = \emptyset, \quad G \cap \{f = t\} \neq \emptyset$$

and

$$(\mathbb{R}^n \setminus \overline{G}) \cap \{f = t\} \neq \emptyset.$$

From its continuity, f must be greater than t or lower than t over the all ∂G. Let us fix $f|_{\partial G} < t$.

The compactness of $\{f = t\}$ gives the existence of a $\delta > 0$ such that $f|_{\partial G} \leq t - \delta$. Thus, for all $\varepsilon \in (0, \delta)$,

$$\partial G \cap \{f > t - \varepsilon\} = \emptyset \quad \text{and} \quad \{f = t\} \subseteq \{f > t - \varepsilon\}.$$

Therefore

$$G \cap \{f > t - \varepsilon\} \neq \emptyset, \quad (\mathbb{R}^n \setminus \overline{G}) \cap \{f > t - \varepsilon\} \neq \emptyset.$$

In addiction, defining L the Lipschitz constant of f,

$$d(\{f \geq t - \varepsilon\}, \partial G) \geq \frac{\delta - \varepsilon}{L}.$$

It follows that the open set $\{f > t - \varepsilon\}$ can be decomposed into two open sets with positive distance, in particular it is decomposable.

In the case

$$f|_{\partial G} > t,$$

one can similarly show that, for all ε in $(0, \delta)$, the set $\{f < t - \varepsilon\}$ is decomposable. Therefore f is not monotone in the sense of Definition 6.4. $\square$

The Decomposition Theorem for real valued BV functions of $\mathbb{R}^n$ is in some sense optimal. Considering BV functions from $\mathbb{R}^2$ to $\mathbb{R}^2$ one can find counterexamples to this theorem, i.e. BV functions which cannot be decomposed in sum of BV monotone functions preserving total variation.

The crucial point is that we require to our decomposition, besides being the sum of BV monotone functions, to preserve the the total variation, i.e.

$$|Df| = \sum_{i \in I} |Df_i|.$$

Remark 6.16. For example, let us generalize as follows our definition of BV monotone function to functions with values in a space of a greater dimension.

Definition 6.17. A function $f : \mathbb{R}^n \to \mathbb{R}^m$, which belongs to $[BV(\mathbb{R}^n)]^m$, is said to be *monotone* if the super-level sets

$$\{f > t\} := \{x \in \mathbb{R}^n | \ f_i(x) > t_i \ i = 1, ..., m\}$$

and the sub-level sets

$$\{f < t\} := \{x \in \mathbb{R}^n | \ f_i(x) < t_i \ i = 1, ..., m\},$$

are indecomposable, for $\mathcal{H}^m$-a.e. $t \in \mathbb{R}^m$.

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ a BV function $f = \begin{pmatrix} f_1 \\ ... \\ f_m \end{pmatrix}$.

For $i = 1, ..., m$, every f_i is a BV function from $\mathbb{R}^n$ to $\mathbb{R}$ so that Theorem 6.5 applies. Therefore, for every $i = 1, ..., m$, one has the decomposition in BV monotone functions $f_i = \sum_{j \in J_i} f_i^j$.

Note that, if $g : \mathbb{R}^n \to \mathbb{R}$ is a BV monotone function, the function $\begin{pmatrix} 0 \\ ... \\ g \\ ... \\ 0 \end{pmatrix}$ is a BV monotone function too, from $\mathbb{R}^n$ to $\mathbb{R}^m$, in the sense of Definition 6.17.

It follows that we can decompose f in that way

$$f = \sum_{j \in J_1} \begin{pmatrix} f_1^j \\ 0 \\ ... \\ 0 \end{pmatrix} + ... + \sum_{j \in J_m} \begin{pmatrix} 0 \\ ... \\ 0 \\ f_m^j \end{pmatrix}.$$

However, this decomposition does not preserve the total variation of f and one can only say that

$$|Df| \leq \sum_{j \in J_1} \begin{pmatrix} |Df_1^j| \\ 0 \\ ... \\ 0 \end{pmatrix} + ... + \sum_{j \in J_m} \begin{pmatrix} 0 \\ ... \\ 0 \\ |Df_m^j| \end{pmatrix}.$$

We give now a counterexample in the case of Lipschitz function from $\mathbb{R}^2$ to $\mathbb{R}^2$. In this situation we extend the Definition 6.1.

Definition 6.18. A function $f : \mathbb{R}^2 \to \mathbb{R}^2$, which belongs to $[\mathrm{Lip}(\mathbb{R}^2)]^2$, is said to be *monotone* if the level sets $\{f = t\} = \{x \in \mathbb{R}^2 | \ f(x) = t\}$ are connected for every $t \in \mathbb{R}^2$.

We observe that if $f : \mathbb{R}^n \to \mathbb{R}^n$ Lipschitz is a monotone operator, then its level sets are closed convex. Hence the requirement to preserve the connectedness of the level sets is weaker than being a monotone operator.

From the Area Formula

$$\int_{\mathbb{R}^2} \mathcal{H}^0(f^{-1}(t)) d\mathcal{H}^2(t) = \int_{\mathbb{R}^2} \det(\nabla f(x)) dx$$

one can say that $f^{-1}(t)$ is finite for $\mathcal{H}^2$-a.e. $t \in \mathbb{R}^2$, i.e. $f^{-1}(t) = \{x_1(t), ..., x_{q(t)}(t)\}$. Therefore there exists a measurable selection $h : \mathbb{R}^2 \to \mathbb{R}^2$ such that $h(t) \in f^{-1}(t)$ for all $t \in \mathbb{R}^2$.

Note that the graph

$$G(f) = \{(x, f(x)) | \ x \in \mathbb{R}^2\}$$

is closed, thus for Theorem 5.8.11 of [36],

$$G(f) = \bigcup_{i \in I} \{(h_i(t), t) | \ t \in \mathbb{R}^2\},$$

where every h_i is a Borel function and I a countable set.

Define, for every $x \in A_i := h_i(\mathbb{R}^2)$, the function $f_i(x) := h_i^{-1}(x)$.
Being A_i the set where h_i is invertible, $f_i : A_i \to \mathbb{R}^2$ is well defined and, in its domain, it is a Lipschitz function with constant equal to the one of f. One also has $f = f_i$ in A_i.
Due to the injectivity of f_i, for all $t \in f_i(\mathbb{R}^2)$ there exists a unique $x \in A_i$ such that $\{f_i = t\} = \{x\}$, which is a connected set. Therefore, for every $i \in I$ f_i is a Lipschitz monotone function in A_i.

Thus, we can decompose $f = \sum_{i \in I} f_i$. This decomposition in sum of Lipschitz monotone functions f_i preserves total variation as desired $|Df| = \sum_{i \in I} |Df_i|$. However, these functions are not defined on the all $\mathbb{R}^2$ but only on the sets $A_i \subseteq \mathbb{R}^2$ for which we just know measurability.

The fact that it is possible to extend these functions to $\mathbb{R}^2$ requires an additional property of the function f. Clearly every f_i can be extended to $\overline{A_i}$ preserving its Lipschitzianity [1].

Fix an $i \in I$. We have $\mathbb{R}^2 \setminus \overline{A_i} = \bigcup_{j\in J} O_j$ where the O_j are connected open sets. The extension of f_i on the all $\mathbb{R}^2$ must preserve monotonicity and the total variation of f_i. For this reason and due to the fact that we already know that $|Df| = \sum_{i\in I} |Df_i|$, the function f_i must be constant on the O_j with positive measure.
Therefore, to preserve the Lipschitzianity, f_i must be constant on ∂O_j. Thus, for every $j \in J$ such that O_j has positive measure, there must be a t_j for which $\mathcal{H}^1(\{f_i = t_j\}) > 0$.
Note that, if for every $j \in J$ the sets O_j have zero measure, the function f_i is the only one in the decomposition and is already monotone, therefore the only interesting case is when there exists at least a $j \in J$ where the corresponding set O_j has positive measure.

Thus one must have

$$\mathcal{H}^1(\{f = \bar{t}\ \}) \geq \mathcal{H}^1(\{f_i = \bar{t}\ \}) > 0$$

for at least a $\bar{t} \in \mathbb{R}^2$. The condition $\mathcal{H}^1(\{f = \bar{t}\ \}) > 0$ for at least a $\bar{t} \in \mathbb{R}^2$ is a necessary condition for the decomposition of a function in that particular way.

Example 6.19. Taken a Lipschitz function $f : \mathbb{R}^2 \to \mathbb{R}^2$ we have seen that a necessary condition for its decomposition is

$$\mathcal{H}^1(\{f = \bar{t}\ \}) > 0$$

for at least a $\bar{t} \in \mathbb{R}^2$.

However, not all Lipschitz functions from $\mathbb{R}^2$ to $\mathbb{R}^2$ have this particular property. For example consider

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \quad f(x) = \begin{pmatrix} 1 - \cos(\frac{\pi x_1}{2}) \\ 1 - \cos(\frac{\pi x_2}{2}) \end{pmatrix}.$$

For this function the level sets $\{f = t\}$ have zero length for every $t \in \mathbb{R}^2$. Thus any decomposition with the properties desired is impossible.

[1] Thanks to Kirszbraun's theorem, see Theorem 2.10.43 in [27], every f_i can be extended to a Lipschitz function of the all $\mathbb{R}^2$. However, for our purpose, it is sufficient to consider the basic Lipschitz extension to the closure.

Acknowledgements

This thesis has been realized under the supervision of Stefano Bianchini. I am really grateful to him for proposing me the topics of research presented here, for his precious help and his enriching discussions made during these years of my PhD.

I would like to thank Giovanni Alberti and Gianluca Crippa for their useful suggestions on the main topics of this thesis.

Deep gratitude is devoted to my parents and my sister, they always supported and encouraged me with love.

Kindest thanks are to the friends that followed me from a distance and to the ones that I had the pleasure to meet along these years in Trieste and that shared with me this beautiful experience.

In SISSA I have met colleagues that little by little became friends. Since some of them are like brothers and sisters to me I really need to thank this big family.

Bibliography

[1] G. Alberti and L. Ambrosio. A geometrical approach to monotone functions in $\mathbb{R}^n$. *Math. Z.*, 2:259–316, 1999.

[2] G. Alberti, S. Bianchini, and G. Crippa. Structure of level sets and Sard-type properties of Lipschitz maps: results and counterexamples. *Preprint*, 2011.

[3] L. Ambrosio, V. Caselles, S. Masnou, and J.-M. Morel. Connected components of sets of finite perimeter and applications to image processing. *J. Eur. Math. Soc. (JEMS)*, 3:39–92, 2001.

[4] L. Ambrosio and C. De Lellis. A note on admissible solutions of 1D scalar conservation laws and 2D Hamilton-Jacobi equations. *J. Hyperbolic Differ. Equ.*, 31 (4):813–826, 2004.

[5] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems.* Oxford University Press, 2000.

[6] M. Bardi and L.C. Evans. On Hopf's formulas for solutions of Hamilton-Jacobi equations. *Nonlinear Anal.*, 8:1373–1381, 1984.

[7] E.N. Barron, P. Cannarsa, R. Jensen, and C. Sinestrari. Regularity of Hamilton-Jacobi equations when forward is backward. *Indiana Univ. Math. J.*, 48:385–409, 1999.

[8] S. Bianchini. On the Euler-Lagrance equation for a variational problem. *DCDS-A*, 17:449–480, 2007.

[9] S. Bianchini, C. De Lellis, and R. Robyr. SBV regularity for Hamilton-Jacobi equations in $\mathbb{R}^n$. *To appear in Arch. Ration. Mech. Anal.*, 2010.

[10] S. Bianchini and M. Gloyer. On the Euler-Lagrange equation for a variational problem: the general case II. *Math. Z.*, 265:889–923, 2010.

[11] S. Bianchini and D. Tonon. A decomposition theorem for *BV* functions. *Commun. Pure Appl. Anal. (CPAA)*, 10(6):1549–1566, 2011.

[12] S. Bianchini and D. Tonon. SBV-like regularity for Hamilton-Jacobi equations with a convex Hamiltonian. *Submitted to J. Math. Analysis Apps.*, 2011.

[13] S. Bianchini and D. Tonon. SBV regularity for Hamilton-Jacobi equations with Hamiltonian depending on (t, x). *Submitted to SIAM-SIMA*, 2011.

[14] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4), 1991.

[15] Y. Brenier. L^2 formulation of multidimensional scalar conservation laws. *Arch. Ration. Mech. Anal.*, 193(1), 2009.

[16] Y. Brenier. On the Hydrostatic and Darcy limits of the convective Navier-Stokes Equations. *Chin. Ann. Math.*, 30 B(6), 2009.

[17] Y. Brenier. Optimal transport, convection, magnetic relaxation and generalized Boussinesq equations. *J. Nonlinear Sci*, 19(5), 2009.

[18] Y. Brenier and E. Grenier. Sticky particles and scalar conservation laws. *SIAM J. Numer. Anal.*, 35(6), 1998.

[19] P. Cannarsa, A. Mennucci, and C. Sinestrari. Regularity results for solutions of a class of Hamilton-Jacobi equations. *Arch. Ration. Mech. Anal.*, 140:197–223, 1997.

[20] P. Cannarsa and C. Sinestrari. *Viscosity solutions of Hamilton-Jacobi equations and optimal control problems.* Birkhäuser, Boston, 2004.

[21] P. Cannarsa and H.M. Soner. On the singularities of the viscosity solutions to Hamilton-Jacobi equations. *Indiana Univ. Math. J.*, 36:501–524, 1987.

[22] M.G. Crandall, L.C. Evans, and P.-L. Lions. Some properties of viscosity solutions of Hamilton-Jacobi equations. *Trans. Amer. Math. Soc.*, 282:487–502, 1984.

[23] M.G. Crandall and P.-L. Lions. Viscosity solutions of Hamilton-Jacobi equations. *Trans. Amer. Math. Soc.*, 277:1–42, 1983.

[24] C. De Lellis. Hyperbolic equations and SBV fucntions. *Journes quations aux drives partielles*, 6:10, 2010.

[25] R. Engelking. *General Topology.* PWN, Warsaw, 1977.

[26] L. C. Evans. *Partial differential equations.* American Mathematical Society, RI, 2010.

[27] H. Federer. *Geometric measure theory.* Springer, 1969.

[28] W.H. Fleming. The Cauchy problem for a nonlinear first order partial differential equation. *J. Differential Equations*, 5:515–530, 1969.

[29] W.H. Fleming and R.W. Rishel. *Deterministic and stochastic optimal control.* Springer, New York, 1975.

[30] W.H. Fleming and H.M. Soner. *Controlled Markov process and viscosity solutions.* Springer, Berlin, 1993.

[31] P Hajlasz and J. Malý. Approximation in Soblev spaces of nonlinear expressions involving the gradient. *Ark. Mat.*, 40 (2):245–274, 2002.

[32] P.-L. Lions. *Generalized solutions of Hamilton-Jacobi equations.* Pitman, Boston, 1982.

[33] J. J. Manfredi. Weakly monotone functions. *J. Geom. Anal.*, 4(2):393–402, 1994.

[34] J. Mossino. *Inégalités isopérimétriques et applications en physique.* Hermann, Paris, 1984.

[35] R. Robyr. SBV regularity of entropy solutions for a class of genuinely nonlinear scalar balance laws with non-convex flux function. *J. Hyperbolic Differ. Equ.*, 5:449–475, 2008.

[36] S. M. Srivastava. *A course on Borel sets.* Springer, 1998.

[37] D. Tonon. Some applications of the SBV regularity theorem for entropy solutions of 1d scalar conservation laws to convection theory and sticky particles. *Accepted by Riv. Mat. Univ. Parma*, 2010.

[38] A. Vasseur. *Contributions à l'approche cinétique des systèmes de lois de conservation hyperboliques.* Unpublished Doctoral Thesis, Paris, 1984.