Alma Mater Studiorum · Università di Bologna
in cotutela con Universität Bern

DOTTORATO DI RICERCA IN MATEMATICA
XXVIII Ciclo
Settore Concorsuale di afferenza: 01/A3
Settore Scientifico disciplinare: MAT/05

SINGULAR SETS OF GENERALIZED CONVEX FUNCTIONS

Presentata da: VALENTINA PENSO

Coordinatore Dottorato:
Chiar.ma Prof.ssa
GIOVANNA CITTI

Relatori:
Chiar.mi Proff.
ZOLTÁN BALOGH
BRUNO FRANCHI

Esame finale anno 2017

Contents

List of Symbols

c	cost function
$\Pi(\mu, \nu)$	set of all probability measure on $X \times Y$ such that relations (1.1) holds.
$I[\pi]$	total transportation cost associated to π
$\mathcal{T}_c(\mu, \nu)$	optimal transportation cost between μ and ν
$\mathcal{H}^k(E)$	k-dimensional Hausdorff measure of E
$\mathcal{S}^k(E)$	k-dimensional spherical Hausdorff measure of E
$\mathrm{dom}(\varphi)$	domain of φ
$\partial\varphi$	subdifferential of φ
φ^*	convex conjugate, or Legendre transform, of φ
ψ^c	c-transform of ψ
$\partial_c\psi$	c-subdifferential of ψ
$\partial^c\varphi$	c-superdifferential of φ
$\Sigma^c(\varphi)$	c-singular set of φ
$\Sigma^k(u)$	k-singular set of u
$T(S, x)$	contingent cone of S at the point x
$\mathrm{Tan}(S, x)$	vectore space spanned by $T(S, x)$

$p(A,x)$	porosity of A at the point x
$x \cdot y$	group product between x and y
τ_g	left translation by an element $g \in \mathbb{G}$
$\Gamma(T\mathbb{G})$	space of smooth vector fields on $\mathbb{G}$
$[X,Y]$	Lie bracket product between X and Y
δ_λ	homogeneous dilations in $\mathbb{G}$
Q	homogeneous dimension of a Carnot group $\mathbb{G}$
d_c	Carnot-Carathéodory metric
$\mathcal{H}^k_{d_c}$	k-dimensional Hausdorff measure associated with the Carnot-Caratéodory metric d_c
$\mathcal{S}^k_{d_c}$	k-dimensional spherical Hausdorff measure associated with the Carnot-Caratéodory metric d_c
$\mathcal{L}^n(E)$	n-dimensional Lebesgue measure of E
J	unit $(2n \times 2n)$-symplectic matrix
$\mathbb{H}^n$	n-th Heisenberg group
$\mathfrak{h}^n$	Lie algebra of the group $\mathbb{H}^n$
V_1	horizontal layer of $\mathfrak{h}^n$
$\mathbb{T}$	center of the group $\mathbb{H}^n$
d_K	distance associated to the Korányi norm
ξ_1	projection from $\mathbb{H}^n$ to V_1, via exponential coordinates
$P_{\mathbb{W}}$	intrinsic projection onto homogeneous subgroup $\mathbb{W}$ of $\mathbb{H}^n$
$\dim \mathbb{G}$	linear dimension of the subgroup $\mathbb{G}$

$\dim_H \mathbb{G}$	Hausdorff dimension of the subgroup $\mathbb{G}$
$\mathcal{G}(\mathbb{H}^n, d)$	intrinsic Grassmannian of the d-subgroups of $\mathbb{H}^n$
ρ	distance in the intrinsic Grassmannians
$G(2n, k)$	Euclidean Grassmannians of dimension k in $\mathbb{R}^{2n}$
d_G	distance in Euclidean Grassmannians
$G^H(2n, h)$	subset of $G(2n, h)$ defined in (4.5)
Π_W	Euclidean projection onto W
$G^V(2n, d)$	subset of $G(2n, d)$ defined in (4.6)
$C_{\mathbb{W},\mathbb{V}}(q, \alpha)$	intrinsic cone with base $\mathbb{W}$, axis $\mathbb{V}$, vertex q and opening α
$X(p_0, \mathbb{G}, \beta)$	intrinsic cone with axis $\mathbb{G}$, vertex p_0 and opening β
$\partial_H u$	horizontal gradient of u
$\Sigma_H^k(T)$	horizontal k-th singular set of T

Introduction

It is a classical fact that a convex function $u : \mathbb{R}^n \longrightarrow \mathbb{R}$ is locally Lipschitz continuous [58]. Therefore one can apply Rademacher's Theorem to deduce that a convex function u is almost everywhere differentiable. At this point a question arises naturally: what can we say about the points at which u is not differentiable? How many non differentiability points do we have?

We can study these questions by considering a generalization of differentiability. Let $u : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function and let $x \in \mathbb{R}^n$ be fixed. We call the *subdifferential* of u at the point x the (possibly empty) set defined as

$$\partial u(x) \;=\; \{y \in \mathbb{R}^n \,|\, u(z) \geq u(x) + \langle y, z - x \rangle \,, \forall\, z \in \mathbb{R}^n\} \,.$$

Clearly, if x is a differentiability point of u and $\partial u(x) \neq$ then $\partial u(x)$ is single valued and it holds that $\partial u(x) = \{\nabla u(x)\}$. Moreover, if u is a convex function, one can prove that, for every $x \in \mathbb{R}^n$, $\partial u(x)$ is non empty.

To better understand what is our task, we suggest now a simple example. Let $u : \mathbb{R}^n \longrightarrow \mathbb{R}$ be $u(x) = \|x\|_{\mathbb{R}^n}$[1]. It is a convex function, the subdifferential $\partial u(x)$ is non empty for every $x \in \mathbb{R}^n$. Trivially, u is differentiable in every $x \in \mathbb{R}^n \setminus \{0\}$. Moreover, one can easily check that $\partial u(0) = \overline{B(0,1)}$, the closed ball with center 0 and radius 1. To summarize, it holds that

$$\partial u(x) = \begin{cases} \overline{B(0,1)}, & \text{if } x = 0, \\ \frac{x}{\|x\|_{\mathbb{R}^n}}, & \text{if } x \neq 0. \end{cases}$$

[1]Throughout the dissertation, if there is not danger of misunderstanding, we will write shortly $|\cdot|$ to denote the Euclidean distance in $\mathbb{R}^n$.

In this example, we can see that $\dim \partial u(x) = n$, just for one point $x = 0$. What happens for a general cost function? Is it possible to classify points $x \in \mathbb{R}^n$ according to the dimension of $\partial u(x)$? In this regards, we define the *k-th singular set of u* as the set

$$\Sigma^k(u) = \{x \in \mathbb{R}^n \mid \dim(\partial u(x)) \geq k\}.$$

In 1992, Alberti, Ambrosio and Cannarsa [3] gave a Hausdorff dimension estimate of k-th singular sets of a convex function. More generally, they proved that the set $\Sigma^k(u)$ is countably $(n-k)$-rectifiable, for every $k = 0, ..., n$. This result has an important application in optimal mass transportation theory.

Let us briefly recall some facts related to this topic, as a motivation to our work. In 1781, Monge [75] formulated the following question: given two sets X and $Y \subset \mathbb{R}^n$ of equal volume, find the optimal volume-preserving map between them, where optimality is measured against a non negative cost function $c : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$. Roughtly speaking, one can see the set X as being filled with mass and $c(x, y)$ as being the cost per unit mass for moving material from $x \in X$ to $y \in Y$.

The optimal map minimizes the total cost of redistributing the mass of X to Y. More specifically, if μ and ν are probability measures supported on X and Y respectively, the goal is to minimize the functional

$$I[T] = \int_X c(x, T(x))\, d\mu(x),$$

over the set of all measurable maps T such that $T\#\mu = \nu$. Originally, Monge considered as cost function $c(x, y) = |x - y|$. An answer to this problem was first given by Sudakov in 1979 [90]. He showed that such a map exists.

It is really surprising that Monge's problem is a prototype for many other questions arising in differential geometry, infinite-dimensional linear programming, functional analysis, mathematical economics and in probability and statistics (see, for instance, [64] and [79]). The Academy of Paris offered a prize for its solution, which was claimed by Appell [11]. Moreover, Kantorovich received a Nobel prize in 1975 for related works in economics.

The approach of Kantorovich was slightly different from the original one. He formulated a more relaxed version of Monge's problem. His task was to minimize the functional

$$I[\pi] = \int_{X\times Y} c(x,y)\, d\pi(x,y),$$

over the set of all probability measures π on $X \times Y$ with marginals μ and ν.

Recently, two groups of authors studied Monge's problem for the cost function $c(x,y) = \frac{1}{2}|x-y|^2$. On one side Brenier and on the other Knott and Smith, Cuesta-Albertos, Matrán and Tuero-Díaz, Rüschendorf and Rachev, and Abdellaoui and Heinrich realized that, for the quadratic cost, an optimal map exists, that it is unique ([18], [17], [1], [27]). They also proved that it is represented by the gradient of a convex function ([63], [17], [87], [85], [17]).

Using Kantorovich's approach to Monge's problem, we are led to the study of singular sets of convex functions. Indeed, one can prove that a probability measure π on $X \times Y$ is optimal if and only if its support is a cyclical monotone set. This condition is equivalent to the fact that π is supported on the graph of the subdifferential of a convex function u. If the dimension of the set of those points x for which $\partial u(x)$ is not single valued is negligible in some sense, then one can conclude that, up to a "small" set, $\mathrm{supp}(\pi)$ is the graph of ∇u, which turns out to be the desired optimal map. Thanks to Alberti, Ambrosio and Cannarsa, we know that $\dim \Sigma^1(u) \leq n-1$; therefore, for the quadratic cost, Monge's problem is solved, when the probability measure μ does not give mass to sets of dimension at most $n-1$.

In a second step, it is quite natural to investigate the same problem for a more general cost function. In this direction, many authors gave answers according to the properties of c.

As already mentioned, optimal transport map for quadratic cost is characterized by the gradient of a convex function. Analogously, it turned out that, for a general continuous cost function, optimal map is related to a generalization of convexity, called c-convexity, and of subdifferentiability, called c-subdifferentiability, and their counterparts: c-concavity and c-superdifferentiability.

Under certain hypothesis on c, one can prove that π is a solution to Kantorovich's problem if and only if it is supported on a c-cyclical monotone set, which is the c-superdifferential of a c-concave map:

$$\partial^c\varphi(y) := \{x \in X \mid \forall\, s \in Y,\ \varphi(y) + c(x,y) \geq \varphi(s) + c(s,y)\}\,.$$

It is clear that the study of the set of points in which the c-superdifferential is not single valued has a central role in the pursuit of solutions to Monge's problem.

Many authors proved results in this direction. For instance, Rüschendorf [84], Knott and Smith [88] and Gangbo and McCann [51], [52]. In [52], the authors considered strictly convex, and also strictly concave, cost functions in $\mathbb{R}^n$. For the case of strictly convex cost functions in $\mathbb{R}^n$, also Ambrosio, Gigli and Savaré obtained some results using "approximate differentials" [8].

Finally, we recall the result of McCann [70] about uniqueness of a cyclically monotone transport in $\mathbb{R}^n$. In his work, McCann introduced a condition on the cost function c: he asked that the function $x \longmapsto c(x,y)$ is locally *semiconcave*, locally in y. We recall that a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is semiconcave if there exists a continuous function $\omega : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$, with $\omega(r) \longrightarrow 0$, as $r \to 0$, such that

$$u(x_t) \leq (1-t)u(x_0) + tu(x_1) + t(1-t)\omega(|x_0 - x_1|),$$

for every x_0 and $x_1 \in \mathbb{R}^n$, $t \in [0,1]$ and $x_t = tx_1 + (1-t)x_0$. With this new hypothesis, one can generalize Alberti, Ambrosio and Cannarsa result and prove that if $u : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a c-concave function, then the *c-singular set of u*, i.e. the set

$$\Sigma^c(u) = \{x \in \mathbb{R}^n \mid \#\partial^c u(x) > 1\}\,,$$

has Hausdorff dimension at most $n-1$.

In the first part of our dissertation, the main goal is to find a way to avoid the hypothesis of semiconcavity. We introduce a new assumption, called *gradient continuity condition*, on the cost function $c : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$:

Assumption. For any compact set $K \subset \mathbb{R}^n$, for any $\varepsilon > 0$ and for any $\eta > 0$, there exists $\delta = \delta(\varepsilon, \eta) > 0$ such that if $p_1, p_1', p_2, p_2' \in K$ and $x \in K$ are so that $|p_1 - p_2| > \eta$, $|p_1 - p_1'| < \delta$, $|p_2 - p_2'| < \delta$ and $x \in C(p_1) \cap C(p_1') \cap C(p_2) \cap C(p_2')$, then either

$$|\nabla_x c(x,p_1) - \nabla_x c(x,p_1')| < \varepsilon \tag{1}$$

or

$$|\nabla_x c(x, p_2) - \nabla_x c(x, p_2')| < \varepsilon. \tag{2}$$

This condition means that the map $y \longmapsto \nabla_x c(x, y)$ can not have too many discontinuities close to each other. This gradient continuity condition, in addition to more classical assumptions (Assumption 2.1.1 and Assumption 2.1.2, see Section 2.1), lets us prove the following:

Theorem 0.1. *Let $u : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a c-concave function. If the cost function $c : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ satisfies the assumptions introduced above, then*

$$\dim_H \Sigma^c(u) \leq n - 1.$$

The main ingredient of the proof is a result due to Mattila [67] about a dimension estimate for porous sets. Following this theory, we introduce a sufficient condition for a set to be porous and apply it to prove Theorem 0.1.

As an application, we can show that, under the hypothesis of Theorem 0.1, there exists a solution to Monge's problem:

Theorem 0.2. *Let μ and ν two probability measures in $\mathbb{R}^n$ such that μ does not give mass to $(n-1)$-dimensional sets. Then, under the same hypothesis of Theorem 0.1, there exists an optimal transport map $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ transporting μ to ν such that*

$$\int_{\mathbb{R}^n} c(x, T(x))\, d\mu(x) \leq \int_{\mathbb{R}^n} c\left(x, \tilde{T}(x)\right)\, d\mu(x),$$

for any measurable map $\tilde{T} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ such that $\tilde{T}\#\mu = \nu$. Moreover, there exists a c-concave function u such that $T(x) = \partial^c u(x)$, for μ-almost every $x \in \mathbb{R}^n$.

The classical literature about optimal mass transportation theory provides deep results also in settings different from the Euclidean one. For instance, the first optimal transport theorem on a Riemannian manifold was given by McCann [71]. Later, this result was extended by Bernard and Buffoni [15] to general Lagrangian cost functions. More recently, Shao and Fang [31] rewrote McCann's theorem in the formalism of Lie groups. They used this reformulation as a starting point to derive theorems of uniqueness and existence of the optimal transport on the path space over a Lie group.

At this point, we would like to cite the work of Ambrosio and Rigot [10], the paper which inspired the second part of this dissertation. Ambrosio and Rigot exended McCann's results to the subriemannian setting, in particular to the Heisenberg group $\mathbb{H}^n$. As cost function they chose the squared Carnot-Carathéodory metric or the squared Korányi metric. Their proof required a delicate analysis of minimizing geodesics and differentiability properties of the squared distances. Analogously to the classical cases, Ambrosio and Rigot proved that the optimal transport map can be represented by the total gradient of a c-concave (and locally Lipschitz) function.

Our goal is to generalize Ambrosio and Rigot result to a more general cost function $c : \mathbb{H}^n \times \mathbb{H}^n \longrightarrow \mathbb{R}$. Therefore, analogously to the strategy applied in the Euclidean case, it is interesting to study singularities of c-concave functions in $\mathbb{H}^n$. At the moment, we are still far from a deep understanding of this problem, either for the case of c equals to the squared Carnot-Carathéodory metric or the squared Korányi metric.

Let us come back to the Euclidean case and consider the quadratic cost $c(x, y) = \frac{1}{2}|x - y|^2$. With some simple computations, if $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a c-convex function, one can prove that there exists a convex function $u : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that $\varphi(x) = u(x) - \frac{1}{2}|x|^2$. In this way, we have a correspondance between c-convexity and classical convexity. Therefore, in this particular case of quadratic cost, the study of c-subdifferentials is reduced to the study of classical subdifferentials.

Is it possible to argue similarly in $\mathbb{H}^n$? We decided to do a first step in this direction by considering H-convex functions and their horizontal gradients, which are the generalizations, in the Heisenberg groups, of the notions of convexity and subdifferentiability ([28], [21], [20]).

Let us recall the definition of H-convex function. We say that a function $u : \mathbb{H}^n \longrightarrow \mathbb{R}$ is *H-convex* if

$$u\left(p \cdot \delta_\lambda(p^{-1} \cdot p')\right) \le u(p) + \lambda\left(u(p') - u(p)\right),$$

for every $p \in \mathbb{H}^n$, $p' \in H_p$ and $\lambda \in [0, 1]$. Its *horizontal subdifferential* at a point

$\eta_0 \in \mathbb{H}^n$ is the set

$$\partial_H u(\eta_0) \;=\; \{p \in V_1 \,|\, u(\eta) \geq u(\eta_0) + \langle p, \xi_1(\eta) - \xi_1(\eta_0)\rangle \,, \text{for every}\, \eta \in H_{\eta_0}\}\,.$$

It is very important to notice that the set valued map defined as $\eta \longmapsto \partial_H u(\eta)$ is H-monotone, i.e. it satisfies the following condition:

$$\langle \xi_1(\eta) - \xi_1(\eta')\; p - p'\rangle \;\geq\; 0,$$

for every η and $\eta' \in \mathbb{H}^n$ such that $\eta \in H_{\eta'}$ and for every $p \in \partial u(\eta)$ and $p' \in \partial u(\eta')$. Thanks to this last property, we decided to focus our attention to the study of H-monotone set valued maps and their singularities.

Using the recent paper [13], we conjecture a result which would like to be the subriemannian counterpart of the Euclidean result of Alberti and Ambrosio [2]. They proved that if $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a maximal monotone set valued map, then the k-th singular set of T,

$$\Sigma^k(T) \;=\; \{x \in \mathbb{R}^n \,|\, \dim T(x) \geq k\}$$

is at most $(n-k)$-rectifiable. The proof is based on this fact: the resolvent $(T+I)^{-1} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a Lipschitz function. Unfortunately, this is not the case in $\mathbb{H}^n$. Therefore, we needed to change completely the approach. Our conjecture is the following:

Conjecture 0.1. *Let* $T : \mathbb{H}^1 \rightrightarrows V_1$ *be a maximal* H*-monotone set valued map with* $\mathrm{dom}(T) = \mathbb{H}^n$*. Then, for every* $0 \leq k \leq 2n$*, the Hausdorff dimension of the set*

$$\Sigma^k(T) \;= \{p \in \mathbb{H}^n \,|\, \dim\left(T(p)\right) = k\}$$

is at most 2n+2-k.

The first step towards the proof of the conjecture is the main theorem of the second part of the dissertation. We restrict ourselves to study the Hausdorff dimension of the set of points $p \in \Sigma^k(T)$ such that $T(p)$ is contained, in some sense, in a horizontal subgroup of $\mathbb{H}^n$. This fact allows us to use the machinery about decomposition in homogeneous subgroups of $\mathbb{H}^n$ and intrinsic Lipschitz functions introduced by Franchi, Serapioni and Serra Cassano ([44], [45], [46], [48] and also the more recent paper [49]). We denote by $\Sigma^k_H(T)$ the restriction to those special points of $\Sigma^k(T)$. Since $T(p) \subset \mathbb{R}^{2n}$ is a convex set, there is an affine subspace of $\mathbb{R}^{2n}$ containing it. Let us denote by V_p the unique linear vector space generating that affine subspace. In this way we can define

$$\Sigma^k_H(T) \;=\; \left\{p \in \mathbb{H}^n \,|\, V_p \in G^H(2n,k)\right\},$$

where $G^H(2n,k)$ is the Grassmannian of horizontal k-dimensional subspaces of $\mathbb{R}^{2n}$ defined as

$$G^H(2n,k) \;=\; \{V \in G(2n,k) \,|\, \langle v, Ju\rangle = 0,\, \forall v,u \in V\}\,.$$

With these notions, we are able to prove the following:

Theorem 0.3. *Let $T : \mathbb{H}^n \rightrightarrows V_1$ be a maximal H-monotone set valued map with* $\mathrm{dom}(T) = \mathbb{H}^n$. *If $1 \le k \le n$, the Hausdorff dimension of $\Sigma^k_H(T)$ is at most $2n+2-k$.*

The structure of the dissertation is the following. The first part, *Singular set of c-convex functions*, is divided in two chapters. In Chapter 1, which is motivational for the second one, we recall the main results and definitions from optimal mass transportation theory in $\mathbb{R}^n$ endowed with the Euclidean metric. We state the problem as Monge and Kantorovich formulated it and we present the main results, existing in literature, about existence and uniqueness of the solution to Monge's problem, when c is a cost function with some general properties.

Chapter 2 is devoted to the exposition of the result obtained avoiding the classical property of semiconcavity on c. We introduce our new assumptions, we justify them with some examples and, finally, recalling the notion of porosity of a set, we give the proof of our result.

In the second part, *Singular set of H-convex functions*, we investigate H-monotone set valued maps and their singular sets. In particular, we recall that the horizontal gradient of an H-convex function is a maximal H-monotone set valued map.

In Chapter 3, we state the main properties about Lie groups and Carnot groups in particular. We emphasis on their most relevant peculiarities, such as homogeneous dilations, graded coordinates and the structure of left invariant vector fields. We will recall also their basic metric properties.

Chapter 4 is devoted to the study of the Heisenberg group and its homogeneous subgroups. We stated some results of Franchi, Serapioni and Serra Cassano, useful for our purpose, and we prove that the set of vertical homogeneous subgroups, which give rise to a decomposition of $\mathbb{H}^n$, is a compact metric space. The Chapter ends with a brief digression about intrinsic Lipschitz functions and the Hausdorff dimension of their graphs.

The last chapter, Chapter 5, contains the results about the dimension of singular set of H-monotone set valued map. We recall the Euclidean counterpart and we list the needed ingredient from [13]. Finally, before giving the proof of the main theorem, we prove that a set, which is locally contained in intrisic cones with axis a vertical group of algebraic dimension d, has Hausdorff dimension at most $d + 1$.

Part I

Singular sets of c-convex functions

Chapter 1

Optimal mass transport and c-convex functions

This first chapter is devoted to recall results about optimal transportation theory which already exist in literature. In particular, we want to focus our attention on some possible relations between Kantorovich's problem and Monge's problem. The two formulations are quite different but we will discover that, under some special assumptions, there is a precise relation between them.

In Section 1.1, we give the formulations of the optimal transport problem as Monge and Kantorovich did. We also propose some examples to highlight the differences between the two approaches.

In Section 1.2, we recall some notations and some basic facts about convex analysis. These results will be used in the remaining sections of this Chapter.

Section 1.3 is devoted to the study of solutions of Kantorovich's problem. It is here that we need to introduce a generalization of concavity and convexity: c-concavity and c-convexity.

In Section 1.4, we approach the study of solutions to Monge's problem, introducing the Kantorovich duality. This tool will be largely used in Section 1.7, where we look to the specific case of quadratic cost function.

The short Section 1.5 lets us link solutions of Kantorovich's problem to solutions of Monge's problem.

In Section 1.6, following Alberti, Ambrosio and Cannarsa [3], we recall some results about singular sets of semiconcave and semiconvex functions.

The Chapter ends with Section 1.8. Here we present, without giving proofs, some results which provide conditions for existence of an optimal tranport map. These conditions are exactly those we want to modify in order to improve classical results.

1.1 Formulation of the optimal mass transport problem

The first formulation of the optimal transportation problem is due to Monge, in 1781. In this very first section, we would like to give a quick and selfcontained introduction to this problem. We will focus on its formulation as Monge introduced it and also on the more recent (and relaxed) Kantorovich optimal mass trasportation problem. For a deeper excursus on this topic, we refer the reader to the first and second Chapters of [92].

Assume that we have two sets X and $Y \subset \mathbb{R}^n$ of equal volume. We can imagine that the set X is uniformly filled with mass normalized to 1. The task is to move this mass from the set X to the set Y. We can model the masses of X and Y with two probability measures μ and ν, respectively. We write $\mu \in P(X)$ and $\nu \in P(Y)$.

Performing this movement needs some efforts. These efforts are modelled by a *cost function* c, which is a non negative function from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}$. Informally, $c(x, y)$ represent the cost per unit mass for transporting material from $x \in X$ to $y \in Y$. Therefore, the main task is to minimize the total cost of redistributing the mass of X through Y.

Let us explain this problem in a more formal way. First, we have a question: how can we define a way of transportation? We define a *transference plan* as a probability measure on the cartesian product $X \times Y$. Clearly, we do not want to loose or gain mass, then we need that our transference plan in *admissible*. This means that π has marginals μ and ν, i.e.

$$\pi\,(A \times Y) \;=\; \mu(A), \quad \pi\,(X \times B) \;=\; \nu(B), \tag{1.1}$$

for every measurable sets $A \subset X$ and $B \subset Y$. An equivalent formulation can be the following

$$\int_{X\times Y} (\varphi(x) + \psi(y))\, d\pi(x, y) \;=\; \int_X \varphi(x) d\mu(x) + \int_Y \psi(y) d\nu(y),$$

for every couple of test functions $(\varphi, \psi) \in L^1(X, \mu) \times L^1(Y, \nu)$.

Notation 1.1.1. We denote by $\Pi(\mu, \nu)$ the set of all probability measure on $X \times Y$ such that relations (1.1) holds.

Remark 1.1.1. The set $\Pi(\mu, \nu)$ is always not empty, since the tensor product $\mu \otimes \nu \in \Pi(\mu, \nu)$.

Let us propose the formulation of the mass transportation problem as given by Kantorovich, who was the one who first tried to give an answer to Monge's question (see [62], [61]). We will refer to this problem as *Kantorovich's problem.*

$$\textit{Minimize the functional}\;\; I[\pi] \;=\; \int_{X\times Y} c(x, y) d\pi(x, y), \;\; \textit{for } \pi \in \Pi(\mu, \nu). \tag{1.2}$$

Let π be a given transference plan. We call $I[\pi]$ the *total transportation cost* associated to π. Notice that $I[\pi]$ is non negative and it could be infinite. The *optimal transportation cost* between μ and ν is the value

$$\mathcal{T}_c(\mu, \nu) \;=\; \inf_{\pi \in \Pi(\mu,\nu)} I[\pi].$$

If there exists $\pi \in \Pi(\mu, \nu)$ such that $\mathcal{T}_c(\mu, \nu) = I[\pi]$, we call π a *optimal transference plan.*

As already mentioned, the optimal mass transport problem was first formulated by Monge. His problem was a bit stronger than the one of Kantorovich. Indeed, it additionally requires that each point $x \in X$ is associated to a unique point $y \in Y$.

In other words, there is no splitting of mass: we are looking for a measurable map $T : X \longrightarrow Y$, which sends the mass in X to Y. We can translate this fact in terms of transference plans. This means that the plan π in (1.2) has the form

$$\pi_T = (Id \times T)\,\#\mu, \tag{1.3}$$

for a suitable measurable map $T : X \longrightarrow Y$. More specifically, it holds that

$$\int_{X\times Y} \zeta(x,y) d\pi_T(x,y) \;=\; \int_X \zeta(x, T(x)) d\mu(x),$$

for every non negative integrable test function ζ on $X \times Y$. Here, we introduce the subscript T in the notation of transference plan π to stress that it is related to the map T. In particular, the associated total transportation cost is

$$I[\pi_T] \;=\; \int_X c(x, T(x)) d\mu(x).$$

In Kantorovich's formulation, we were looking for an optimal transference plan among the probability measures belonging to $\Pi(\mu, \nu)$. Therefore, also in this case we need a condition under which π_T belongs to $\Pi(\mu, \nu)$. This condition can be recovered from equation (1.3). One easily obtains that

$$\int_X (\psi \circ T)(x) d\mu(x) \;=\; \int_Y \psi(y) d\nu(y),$$

for every non negative $\psi \in L^1(Y, \nu)$. When this condition is satisfied we write $\nu = \mathrm{T}\#\mu$, i.e. the probability measure ν is the push forward of μ through T.

We are now ready to state the *Monge's optimal mass transportation problem*:

$$\textit{Minimize the functional}\;\; I[T] \;=\; \int_X c\,(x, T(x))\, d\mu(x), \tag{1.4}$$

over the set of all measurable maps T such that $T\#\mu = \nu$.

We present now a very simple example which can clarify the nature of Monge optimal transport problem.

Example 1.1.1. Let $a \in \mathbb{R}^n$ be a fixed point. Let μ be a probability measure in $\mathbb{R}^n$ and let $\nu = \delta_a$, the Dirac mass supported at a. The goal is to move all the mass represented by μ to the point a in an optimal way. Obviuosly, we have a unique possibility: $T(x) = a$, for every $x \in \mathbb{R}^n$. Therefore,

$$\mathcal{T}_c(\mu, \delta_a) = \int_X c(x, a) d\mu(x).$$

As already said, it could happen that there exists a solution for the Kantorovich problem but not for the Monge problem. Here we propose an example, taken from [93], in which we have splitting of mass.

Example 1.1.2. Let $X = \mathbb{R}$ and also $Y = \mathbb{R}$. We define $\mu = \delta_0$ and $\nu = (\delta_1 + \delta_{-1})/2$ In this case the Monge problem has no solution, indeed there is no map such that $T\#\mu = \nu$. On the other hand, one can check that the optimal transport plan is given by $\frac{1}{2}\delta_0 \times \delta_{-1} + \frac{1}{2}\delta_0 \times \delta_1$.

With the formulation of the optimal mass transportation problem, we can address the study of the solutions. There are two main aspects that we should consider. The first one is the existence of a solution: do there exist minimizers for problems (1.2) or (1.4)? The second step is referred to the characterization of these minimizers. Can we say something about the nature of these solutions, depending on our knowledge about the probability measures μ and ν? As we will see in what follows, the answers to these major questions depend strictly on the properties of the cost function c.

Let us consider some examples in the case when $X = \mathbb{R}^n$ and $Y = \mathbb{R}^n$, the cost function is given by $c(x, y) = |x - y|^p$, for $0 < p < \infty$, and μ and ν are compactly supported probability measures.

Example 1.1.3. We assume that $p > 1$. This implies that c is strictly convex. It holds that, if μ and ν are absolutely continuous with respect to the Lebesgue measure, then there is a unique solution to the Kantorovich problem and it is also a solution of the Monge problem.

A step further can be done. We can weaken the condition on μ, asking that μ does not give mass to any set with finite $(n-1)$-dimensional Hausdorff measure[1].

[1]We recall that, given $E \subset \mathbb{R}^n$ and $k \geq 0$, the *k-dimensional Hausdorff* and *Spherical*

This fact will be a key point in Chapter 2 of this dissertation.

Example 1.1.4. If $p = 2$, we refer to c as a *quadratic cost*. In this case it turns out that an optimal map exists and it is unique. These facts are due to Brenier, Knott and Smith, Cuesta-Albertos, Matrán and Tuero-Díaz, Rüschendorf and Rachev and Abdellaoui and Heinich (see [18], [26], [17], [1], [27]). Moreover, the optimal transport map can be represented by the gradient of a suitable convex function on $\mathbb{R}^n$ (see [63], [18], [87], [85], [17]).

In Section 1.7, we will focus our attention to this case.

Example 1.1.5. There exist probability measures μ and ν on $\mathbb{R}^n$, which give mass to some sets of Hausdorff dimension at most $n-1$ and such that optimal transference plans in Kantorovich problem have to split mass. In this case it could happen that a solution to the Monge problem does not exist.

The questions about the existence, uniqueness and characterizations of a solution to the Kantorovich problem, and then of Monge problem, were for long time studied. It should be apparent that the solution would not be unique. This aspect was studied by Appel [11] and by Kantorovich [62]. About the existence of the solution we need to go back to the first who tried to give the solution to the Monge problem as originally formulated. Sudakov [90] tried to give a solution of the problem for the case when $c(x, y) = |x - y|$. Unfortunately the argument was not completely correct, as pointed out by Alberti, Kirchheim and Preiss. The interested reader can find the fixed proof in [5] and in Sections 1 and 8 of [9].

Hausdorff measures of E are defined, respectively, by

$$\mathcal{H}^k(E) := \liminf_{\delta \searrow 0} \left\{ \sum_{i=1}^{\infty} (\operatorname{diam} E_i)^k \,\middle|\, E \subset \bigcup_{i=0}^{\infty} E_i,\ \operatorname{diam} E_i < \delta \right\}$$

$$\mathcal{S}^k(E) := \liminf_{\delta \searrow 0} \left\{ \sum_{i=1}^{\infty} (\operatorname{diam} B_i)^k \,\middle|\, E \subset \bigcup_{i=0}^{\infty} B_i,\ \operatorname{diam} B_i < \delta,\ B_i \subset \mathbb{R}^n \text{ balls} \right\}.$$

1.2 Some facts from convex analysis

In this section, we interrupt our investigation of optimal transportation theory with the purpose of building a little background on convex analysis. The reader who wants to have a larger view on this topic is referred to [82] and [30].

Definition 1.2.1. *Let $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a proper[2] function. We say that φ is a* convex function *if it holds that*

$$\varphi(tx + (1-t)y) \;\leq\; t\varphi(x) + (1-t)\varphi(y),$$

for every x and $y \in \mathbb{R}^n$ and for every $t \in [0,1]$.

Remark 1.2.2. We define $\mathrm{dom}(\varphi)$ as the set of those points on which φ is finite. We point out that $\mathrm{dom}(\varphi)$ is a convex set. This implies that its boundary is a set of Hausdorff dimension at most $n-1$.

Proposition 1.2.1. *Let $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a proper convex function. Then φ is locally Lipschitz continuous on* $\mathrm{Int}(\mathrm{dom}(\varphi))$

Remark 1.2.3. Since a proper convex function φ is locally Lipschitz in the interior of its domain, one can apply Rademacher's Theorem and prove that $\nabla\varphi$ is well defined almost everywhere and locally bounded. Furthermore, it turns out that the set of points where $\nabla\varphi$ is not defined has dimension at most $n-1$ (see Section 1.6).

Now, consider just the points x at which $\nabla\varphi$ is well defined. Then it holds that, geometrically, the graph of φ lies above its tangent hyperplane atthe point x. More specifically, it is true that

$$\varphi(z) \;\geq\; \varphi(x) + \langle \nabla\varphi(x)\,,\,(z-x)\rangle\,,$$

for every $z \in \mathbb{R}^n$. In particular, this can be restated by saying that the map $\nabla\varphi$ is *monotone*, i.e.

$$\langle \nabla\varphi(x) - \nabla\varphi(z), x - z\rangle \;\geq\; 0,$$

[2]We say that a function $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$ is *proper* if for every $x \in \mathbb{R}^b$ $f(x) \neq -\infty$ and it is not identically $+\infty$.

for every differentiability points x and $z \in \mathbb{R}^n$.

How can we deal with the points at which a proper convex function is not differentiable? We introduce a generalized notion of gradient.

Definition 1.2.4. *Let $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a proper convex function. We define the* subdifferential *of φ at the point $x \in \mathbb{R}^n$ as the set*

$$\partial\varphi(x) \;=\; \{y \in \mathbb{R}^n \,|\, \varphi(z) \;\geq\; \varphi(x) + \langle y, z - x\rangle \,,\, \forall\, z \;\in \mathbb{R}^n\}\,.$$

Remark 1.2.5. Obviously, it is well defined the set valued map $x \longmapsto \partial\varphi(x)$ (see Definition 5.0.1), whose graph lies in $\mathbb{R}^n \times \mathbb{R}^n$. From its definition, it is quite simple to prove that this set valued map is monotone (see Definition 5.0.2). This means that

$$\langle y_1 - y_2, x_1 - x_2\rangle \;\geq\; 0,$$

for ever $y_1 \in \partial\varphi(x_1)$ and $y_2 \in \partial\varphi(x_2)$.

Remark 1.2.6. Using the Hahn-Banach separation Theorem, one can prove that, for every $x \in \text{Int}(\text{dom}(\varphi))$, $\partial\varphi(x)$ is non empty. Moreover, φ is differentiable at a point x if and only if $\partial\varphi(x)$ is a singleton, which is $\nabla\varphi(x)$.

Proposition 1.2.2. *Let $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a proper convex function. If φ is lower semicontinuous*[3]*, then the multivalued map $\partial\varphi$ is continuous, i.e. if $x_k \longrightarrow x$ and $v_k \in \partial\varphi(x_k) \longrightarrow y$, as $k \to \infty$, then $y \in \partial\varphi(x)$.*

We give now the definition of *convex conjugate functions* and we state some properties, very useful in what will follow.

Definition 1.2.7. *Let $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a proper function. We call its* convex conjugate function*, or the* Legendre transform*, the function defined as*

$$\varphi^*(y) \;=\; \sup_{x \in \mathbb{R}^n} \left(x \cdot y - \varphi(x)\right).$$

[3] A function F, defined on a metric space Ω, is called *lower semicontinuous* if, for any $x \in \Omega$,

$$F(x) \leq \liminf_{y \to x} F(y).$$

By definition, it holds that φ^* is a proper convex semicontinuous function. Moreover, the following inequality is valid

$$x \cdot y \leq \varphi(x) + \varphi^*(y),$$

for every x and $y \in \mathbb{R}^n$.

The notion of convex conjugate function gives rise to a characterization of the subdifferential of a convex function.

Proposition 1.2.3. *Let $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a proper convex lower semicontinuous function. Then, for every x and $y \in \mathbb{R}^n$, $x \cdot y = \varphi(x) + \varphi^*(y)$ if and only if $y \in \partial\varphi(x)$, or, equivalently, $x \in \partial\varphi^*(y)$.*

Proof. See [92], Proposition 2.4. □

We conclude this Section with a proposition about Legendre duality for convex function.

Proposition 1.2.4. *Let $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a proper function. Then the following three statements are equivalent*

(i) *φ is convex lower semicontinuous;*

(ii) *$\varphi = \psi^*$ for some proper function ψ;*

(iii) *$\varphi^{**} = (\varphi^*)^* = \varphi$.*

Proof. See [92], Proposition 2.5. □

1.3 Optimality conditions for transference plans

Let us come back to the Kantorovich's problem and try to investigate a possible solution. We want to understand the structure of optimal plans. We start by considering a particular case.

Let $X = Y = \mathbb{R}^n$ and let $c(x, y) = \frac{1}{2}|x - y|^2$ be the quadratic cost function. We assume that μ and ν are two probability measures supported on two finite sets.

One can show that a tranference plan $\pi \in \Pi(\mu,\nu)$ is optimal if and only if it holds that

$$\sum_{i=1}^{N} \frac{1}{2}|x_i - y_i|^2 \leq \sum_{i=1}^{N} \frac{1}{2}|x_i - y_{\sigma(i)}|^2, \tag{1.5}$$

for every $N \in \mathbb{N}$, $(x_i, y_i) \in \mathrm{supp}(\pi)$ and for every permutation σ of the set $\{1, ..., N\}$.

Now, let us perform some computations in (1.5). Expanding the squares, we get

$$\sum_{i=1}^{N} \langle x_i, y_i \rangle \leq \sum_{i=1}^{N} \left\langle x_i, y_{\sigma(i)} \right\rangle, \tag{1.6}$$

for every $N \in \mathbb{N}$, $(x_i, y_i) \in \mathrm{supp}(\pi)$ and for every permutation σ of the set $\{1, ..., N\}$. Inequality (1.6) means that the support of the measure π is cyclically monotone:

Definition 1.3.1. *We say that a set $\Gamma \subset \mathbb{R}^n \times \mathbb{R}^n$ is* cyclically monotone *if, for every choice of points $(x_i, y_i) \in \Gamma$, for $i = 1, ..., N$, with $N \in \mathbb{N}$,*

$$\sum_{i=1}^{N} \langle y_i, x_{i+1} - x_i \rangle \leq 0.$$

This is a surprising property for a set. Indeed, the next theorem tells us that there is a correspondance between cyclically monotone sets and graphs of the subdifferential of convex and lower semicontinuous functions. In particular, this gives a characterization for an admissible tranference plan to be optimal.

Theorem 1.3.1. *Let $\Gamma \subset \mathbb{R}^n \times \mathbb{R}^n$ be a set. Γ is cyclically monotone if and only if there exists a convex and lower semicontinuous function $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ sucht that Γ is included in the subdifferential of φ.*

Proof. See [83], Theorem 12.25. □

Remark 1.3.2. We stress that the previous Theorem implies that the following statements are equivalent:

(i) $\pi \in \Pi(\mu,\nu)$ is optimal;

(ii) $\text{supp}(\pi)$ is cyclically monotone;

(iii) there exists a convex and lower semicontinuous function φ such that π is supported on the graph of the subdifferential of φ.

What happens for a much more general cost function? In the previous example, the notions of cyclical monotonicity, convexity and subdifferential play a key role, but they were linked to the special nature of the quadratic cost. Now we need to study some generalizations. From now on, if it is not specified, c is a generic non negative continuous cost function.

Definition 1.3.3. *Let $\Gamma \subset \mathbb{R}^n \times \mathbb{R}^n$ be a set. We say that Γ is* c-cyclically monotone *if, for every choice of $(x_i, y_i) \in \Gamma$, $i = 1, ..., N$, with $N \in \mathbb{N}$, it holds that*

$$\sum_{i=1}^{N} c(x_i, y_i) \le \sum_{i=1}^{N} c(x_i, y_{\sigma(i)}),$$

for every permutation σ of the set $\{1, ..., N\}$.

Definition 1.3.4. *Let X and Y be two subsets of $\mathbb{R}^n$. We say that a function $\psi : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ is* c-convex *if it is proper and there exists a function $\zeta : Y \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ such that, for every $x \in X$,*

$$\psi(x) = \sup_{y \in Y} (\zeta(y) - c(x, y)) .$$

Remark 1.3.5. Roughly speaking, a c-convex function is a function whose graph can be, locally, approximated from below with a tool whose shape is the negative of the cost function.

Definition 1.3.6. *We define also the* c-transform *of ψ as the function $\psi^c : Y \longrightarrow \mathbb{R} \cup \{+\infty\}$ defined as*

$$\psi^c(y) = \inf_{x \in X} (\psi(x) + c(x, y)) .$$

The functions ψ and ψ^c are said to be c-conjugate.

Definition 1.3.7. *Let $\psi : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a c-convex function and let $x \in X$ be a given point. We call the* c-subdifferential *of ψ at the point x the set*

$$\partial_c \psi(x) \,:=\, \{y \in Y \,|\, \psi^c(y) - \psi(x) = c(x, y)\}\,,$$

or, equivalently,

$$\partial_c \psi(x) \,:=\, \{y \in Y \,|\, \forall\, z \in X,\, \psi(x) + c(x, y) \leq \psi(z) + c(z, y)\}\,.$$

We have also the analogous definition for c-concavity.

Definition 1.3.8. *With the same notations of Definition 1.3.4, a function $\varphi : Y \longrightarrow \mathbb{R} \cup \{-\infty\}$ is a* c-concave function *if it is not identically $-\infty$ and there exists $\psi : X \longrightarrow \mathbb{R} \cup \{\pm\infty\}$ such that*

$$\varphi(y) \,=\, \inf_{x \in X} \left(c(x, y) - \psi(x)\right).$$

Its c-superdifferential *at a point $y \in Y$ is the set*

$$\partial^c \varphi(y) \,:=\, \{x \in X \,|\, \varphi(y) + \varphi^c(x) = c(x, y)\}\,,$$

or, equivalently,

$$\partial^c \varphi(y) \,:=\, \{x \in X \,|\, \forall\, s \in Y,\, c(x, y) - \varphi(y) \geq c(s, y) - \varphi(s)\}\,.$$

Remark 1.3.9. There are some relations between c-convexity and c-concavity. For instance, φ is c-concave if and only if $-\varphi$ is c-convex. Moreover, $\partial^c \varphi = \partial_c(-\varphi)$. The proofs of these facts are quite trivial and they need just some computations with the related definitions.

Proposition 1.3.2. *Let $\psi : X \longrightarrow \mathbb{R} \cup \{+\infty\}$. Then ψ is c-convex if and only if $\psi^{cc} = \psi$, where $\psi^{cc} = (\psi^c)^c$.*

Proof. See [93], Proposition 5.8. □

Let us consider a simple example which links this generalization with the classical notions.

Example 1.3.1. Let $X = Y = \mathbb{R}^n$ and let $c(x, y) = -x \cdot y$. In this case one can easily see that a set is c-cyclically monotone if and only if it is cyclically monotone. Moreover, a function is c-convex if and only if it is convex and lower semicontinuous. Its c-subdifferential is the classical subdifferential. Obviuosly, an analogous statement holds for c-concave functions and their c-superdifferentials.

We are now ready to state a fundamental theorem in optimal transport theory. The main part is that, under certain assumptions on the cost c, every c-cyclically monotone set can be obtained as the c-superdifferential of a c-concave function.

Theorem 1.3.3. *Let X and $Y \subset \mathbb{R}^n$ be two set and let $c : X \times Y \longrightarrow \mathbb{R}$ be non negative and continuous. Let μ and ν be two probability measures supported on X and Y, respectively. We assume that*

$$c(x, y) \leq a(x) + b(y),$$

for some $a \in L^1(X, \mu)$ and $b \in L^1(Y, \nu)$. If $\pi \in \Pi(\mu, \nu)$, then the following statements are equivalent:

(i) π is an optimal transference plan;

(ii) the set $\operatorname{supp}(\pi)$ is c-cyclically monotone;

(iii) there exists a c-concave function φ such that $\max\{0, \varphi\} \in L^1(X, \mu)$ and π is supported on the graph of $\partial^c \varphi$.

Proof. See [51], Theorem 2.7 and Corollary 2.8. □

Remark 1.3.10. It holds a statement stronger than Theorem 1.3.3. If $\operatorname{supp}(\pi)$ is contained in $\operatorname{graph}(\partial^c \varphi)$, for some optimal transference plan $\pi \in \Pi(\mu, \nu)$, then $\operatorname{supp}(\tilde{\pi}) \subset \operatorname{graph}(\partial^c \varphi)$, for every $\tilde{\pi} \in \Pi(\mu, \nu)$.

1.4 The Kantorovich duality

In what follows, we want to give a powerful tool, the Kantorovich duality, used in the study of the geometry of optimal transport maps. For sake of simplicity, we introduce a couple of pieces of notation.

Notation 1.4.1. Let X and Y be two subspaces of $\mathbb{R}^n$. We define Φ_c the set of all measurable functions $(\varphi, \psi) \in L^1(X, \mu) \times L^1(Y, \nu)$ satisfying

$$\varphi(x) + \psi(y) \leq c(x, y), \tag{1.7}$$

for μ-almost every $x \in X$ and for ν-almost every $y \in Y$.

Notation 1.4.2. If $\pi \in P(X \times Y)$ and $(\varphi, \psi) \in L^1(X, \mu) \times L^1(Y, \nu)$, we define

$$I[\pi] := \int_{X \times Y} c(x, y) d\pi(x, y) \quad \text{and} \quad J(\varphi, \psi) := \int_X \varphi(x) d\mu(x) + \int_Y \psi(y) d\nu(y).$$

Theorem 1.4.1 (Kantorovich duality)**.** *Let μ and ν two probability measures supported on $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^n$ respectively and let $c : X \times Y \longrightarrow \mathbb{R}^+$ be a lower semicontinuous cost function. Then*

$$\inf_{\Pi(\mu,\nu)} I[\pi] = \sup_{\Phi_c} J(\varphi, \psi). \tag{1.8}$$

Moreover, the infimum in the left-hand side of (1.8) *is attained.*

Remark 1.4.1. We point out that one can restrict the definition of Φ_c to functions (φ, ψ) which are continuous and bounded. This restriction does not change the value of the supremum in the right-hand side of (1.8).

We will skip the rigorous proof of Theorem 1.4.1. The interested reader is referred to the first Chapter of the book [92]. The proof is essentially based on a minimax principle due to Fenchel and Rockafellar (see [82]), based on Hahn-Banach Theorem.

For completeness, we aim to prove the easiest part of Kantorovich Duality Theorem.

Proposition 1.4.2. *Under the same assumption of Theorem 1.4.1, it holds that*

$$\sup_{\Phi_c} J(\varphi, \psi) \leq \inf_{\Pi(\mu,\nu)} I[\pi]. \tag{1.9}$$

Proof. Let $\pi \in \Pi(\mu, \nu)$ and $(\varphi, \psi) \in \Phi_c$ be fixed. By definition of $\Pi(\mu, \nu)$, we have that

$$J(\varphi, \psi) = \int_X \varphi(x) d\mu(x) + \int_Y \psi(y) d\nu(y) = \int_{X \times Y} (\varphi(x) + \psi(x)) \, d\pi(x, y).$$

Now, inequality (1.7) holds π-almost everywhere in $X \times Y$. Indeed, let $N_x \subset X$ and $N_y \subset Y$ be such that $\mu(N_x) = 0$ and $\nu(N_y) = 0$ and (1.7) holds on $N_x^c \times N_y^c$. We know that $\pi \in \Pi(\mu, \nu)$, this means that it has marginals μ and ν. Therefore, $\pi(N_x \times Y) = \mu(N_x) = 0$ and $\pi(X \times N_y) = \nu(N_y) = 0$, implying that $\pi\left(\left(N_x^c \times N_y^c\right)^c\right) = 0$.

Consequently, we have that

$$\int_{X\times Y} (\varphi(x) + \psi(y))d\pi(x, y) \leq \int_{X\times Y} c(x, y)d\pi(x, y) = I[\pi].$$

Taking the supremum on the left-hand side and the infimum on the right-hand side, we obtain inequality (1.9), concluding the proof of this simplified version of Kantorovich Duality Theorem. □

Let us conclude this section with the proof of a stronger version of the Kantorovich duality: we ask for the continuity of the cost function c.

Theorem 1.4.3. *Let μ and ν two probability measures supported on $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^n$, respectively. Let $c : X \times Y \longrightarrow \mathbb{R}$ be a non negative and continuous cost function. We assume that*

$$c(x, y) \leq a(x) + b(y),$$

for some $a \in L^1(X, \mu)$ and $b \in L^1(Y, \nu)$. Then

$$\sup_{\Phi_c} J(\varphi, \psi) = \inf_{\Pi(\mu,\nu)} I[\pi].$$

Moreover, the supremum is attained and the maximizing couple (φ, ψ) is of the form (φ, φ^c) for some c-concave function φ.

Proof. Let $\pi \in \Pi(\mu, \nu)$ and notice that if $(\varphi, \psi) \in \Phi_c$, then it holds that

$$\int_{X\times Y} c(x, y)d\pi(x, y) \geq \int_{X\times Y} (\varphi(x)+\psi(y))d\pi(x, y) = \int_X \varphi(x)d\mu(x)+\int_Y \psi(y)d\nu(y).$$

This implies that the minimum of the Kantorovich's problem is larger than or equal to the supremum of the dual problem.

We prove now the opposite inequality. Let $\pi \in \Pi(\mu, \nu)$ be an optimal transference plan and apply Theorem 1.3.3 to find a c-concave function φ such that π is supported on the graph of $\partial^c\varphi$, $\max\{0,\varphi\} \in L^1(X,\mu)$ and $\max\{0,\varphi^c\} \in L^1(Y,\nu)$. Then we have that

$$\int_{X\times Y} c(x,y)d\pi(x,y) = \int_{X\times Y} (\varphi(x)+\varphi^c(y))d\pi(x,y) = \int_X \varphi(x)d\mu(x)+\int_Y \varphi^c(y)d\nu(y),$$

and $\int_{X\times Y} c(x,y)d\pi(x,y) \in \mathbb{R}$. Thus, $\varphi \in L^1(X,\mu)$ and $\varphi^c \in L^1(Y,\nu)$, which shows that (φ,φ^c) is an admissible couple in the dual problem and gives the thesis. □

1.5 Existence of optimal maps and c-singular sets

In this short section, we want to motivate why the study of c-singular sets is so important. Our goal is to prove existence of a solution to Monge's problem passing through Kantorovich's problem, for which we know that a solution exists. As already pointed out, it is in general not true that there is a correspondance between optimal tranference plans and optimal transport maps. Furthermore, it could happen that there exists a solution to the Kantorovich's problem but none for the relative Monge's problem. We remind that the problem of existence of optimal transport maps consists in looking for optimal plans π which are induced by a map $T : X \longrightarrow Y$. In other words, we are looking for plans π such that $\pi = (Id \times T)\#\mu$, where T is some measurable map.

The question is: given μ and ν probability measures and a cost function c, is that true that at least one optimal plan π is induced by a map? To approach the answer, we have a proposition.

Proposition 1.5.1. *Let X and $Y \subset \mathbb{R}^n$ and let $\mu \in P(X)$ and $\nu \in P(Y)$. Let $\pi \in \Pi(\mu,\nu)$. Then π is induced by a measurable map $T : X \longrightarrow Y$ if and only if there exists a π-measurable set $\Gamma \subset X \times Y$ such that* $\mathrm{supp}(\pi) \subset \Gamma$ *and, for μ-a.e. $x \in X$, there exists a unique $y = T(x) \in Y$ such that $(x,y) \in \Gamma$.*

Proof. See [7], Lemma 1.20. □

By Theorem 1.3.3, we know that optimal transference plans are supported on c-cyclically monotone sets. Moreover, we know that these c-cyclically monotone sets are given by c-superdifferentials of c-concave functions. The main problem now is to understand "how often" the c-superdifferential of a c-concave function is single valued. It is here that we discover the importance of the study of the singular set of the c-superdifferential of a c-concave function.

Definition 1.5.1. *Let $X \subset \mathbb{R}^n$ be a set and let $\varphi : X \longrightarrow \mathbb{R} \cup \{-\infty\}$ be a c-concave function. We call the* c-singular set *of φ the set*

$$\Sigma^c(\varphi) \,=\, \{x \in X \,|\, \#\partial^c\varphi(x) > 1\}\,.$$

If we know that, under certain conditions on c, $\Sigma^c(\varphi)$ is "small enough" with respect to the measure μ, we can apply Proposition 1.5.1 to have existence of an optimal transport map associated to an optimal tranference plan, solution of a given Kantorovich's problem.

In the following sections, we will study the Hausdorff dimension of the singular set of semiconcave and semiconvex functions, which are a generalization of concave and convex functions. Then, we will consider again the particular case of the quadratic cost function and we will state the celebrated Brenier's theorem regarding the existence and the uniqueness of the optimal transport map. Finally, we will come back to the general case and introduce some classical results from the existing literature.

1.6 Study of singular sets of semiconvex and semiconcave functions

In this section we give a small introduction to the paper of Ambrosio, Alberti and Cannarsa [3], regarding an estimate for the Hausdorff dimension of the singular set of the subdifferential of semiconcave, or semiconvex, functions. For a detailed references, one can give a look also to the book [23], about the theory of semiconcave functions.

As preannounced, the main question is the following: if $u : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a convex function and $0 \leq k \leq n$ is a given integer, how can we estimate the dimension of the *k-singular set* of u

$$\Sigma^k(u) := \{x \in \mathbb{R}^n \mid \dim(\partial u(x)) \geq k\}? \tag{1.10}$$

It is clear that $\Sigma^0(u) = \mathbb{R}^n$ and that $\Sigma^n(u)$ is at most countable. For the case $k = 1$, the classical theory gives an answer to the question noting that ∇u has locally bounded first variation. This is true because the jump set of the function u is $\mathcal{H}^{n-1}$-rectifiable[4], or, equivalently, $\Sigma^1(u)$ can be covered with a sequence of C^1-hypersurface, up to a subset of $(n-1)$-dimensional Haudorff measure zero. The interested reader is referred to [50] and [81] for the complete proof.

In [3], the authors give an answer to this problem showing that the k-singular set of a semiconcave function has Hausdorff dimension at most $n-k$. In particular, they prove that $\Sigma^k(u)$ is countably $\mathcal{H}^{n-k}$-rectifiable. Their idea is to study the dimension of the contingent cone (see Definition 1.6.8) and give it an upper bound.

Let us start with some basic definitions.

Definition 1.6.1. *Let $\Omega \subset \mathbb{R}^n$ be an open and convex set and let $u : \Omega \longrightarrow \mathbb{R}$ be a function. We say that u is* semiconvex *if there exists a continuous function $\omega : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$, with $\omega(r) \to 0$, as $r \to 0$, such that*

$$u(x_t) \leq (1-t)u(x_0) + tu(x_1) + t(1-t)\omega(|x_0 - x_1|), \tag{1.11}$$

for every x_0 and $x_1 \in \Omega$, $t \in [0,1]$ and $x_t = tx_1 + (1-t)x_0$. We denote by $\omega_{u,\Omega}$ the least function ω for which (1.11) holds.

Remark 1.6.2. We point out that convex functions are particular cases of semiconvex functions. In this case the least ω is simply the identically zero function.

Remark 1.6.3. A similar definition for *semiconcavity* is obtained in an obvious way by reversing the sign of the inequality.

[4]We say that a set $S \subset \mathbb{R}^n$ is $\mathcal{H}^m$-rectifiable if there exists a countable family of C^1 hypersurfaces $\Gamma_h \subset \mathbb{R}^n$, $h \in \mathbb{N}$, of dimension m such that

$$\mathcal{H}^m\left(S \setminus \bigcup_{h=1}^{\infty} \Gamma_h\right) = 0.$$

We define the *subdifferential* of a semiconvex function, also in this more general case. It turns out that $\partial u(x)$ is a closed and convex set.

Definition 1.6.4. *Let $u : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ be a semiconvex function and let $x \in \Omega$ be fixed. We define the* subdifferential *of u at the point x the set*

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

Remark 1.6.5. One can check that, in the case of a convex function, this definition of subdifferential coincides with the classical one, given in Section 1.2.

In the next proposition we summarize the main properties of semiconvex functions.

Proposition 1.6.1. *Let $u : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ be a semiconvex function. Then the following statements are true:*

(i) u is locally Lipschitz continuous in Ω;

(ii) the set $\partial u(x)$ is non empty, convex and compact, for every $x \in \Omega$;

(iii) a point $p \in \partial u(x)$ if and only if, for every $y \in \Omega$,

$$u(y) - u(x) - \langle p, y - x \rangle \;\geq\; -|y - x|\omega_{u,\Omega}(|y - x|);$$

(iv) the map $x \longmapsto \partial u(x)$ is upper semicontinuous.

Proof. See, for instance, [3], Proposition 2.1. □

We consider a semiconvex function $u : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$. The goal is to study the dimension of the

Notation 1.6.1. Let $\alpha > 0$, we denote

$$\Sigma^k_\alpha(u) := \left\{ x \in \Sigma^k(u) \,|\, \exists B^k_\alpha \subset \partial u(x), \text{ with} \operatorname{diam}(B^k_\alpha) = 2\alpha \right\},$$

where B^k_α is a ball of dimension k and radius α.

Remark 1.6.6. It is clear that

$$\Sigma^k(u) = \bigcup_{p\in\mathbb{N}} \Sigma^k_{1/p}(u).$$

If one has that $\Sigma^k_\alpha(u)$, for every $\alpha > 0$, is countably $\mathcal{H}^{n-k}$-rectifiable, it follows also that $\Sigma^k(u)$ is countably $\mathcal{H}^{n-k}$-rectifiable too. Indeed, the previous union is countable.

Remark 1.6.7. The same idea of splitting the singular set will be also one of the ingredient of the proof of Theorem 2.3.1 of Section 2.3, where we prove an analogous result for c-convex functions and their c-singular sets.

The idea of Alberti, Ambrosio and Cannarsa is to estimate the dimension of the contingent cone. Then, it is time to introduce the notion and to give some results.

Definition 1.6.8. *Let $S \subset \mathbb{R}^n$ be an arbitrary set and let $x \in S$. We call the* contingent cone *of S at the point x the set*

$$T(S,x) := \left\{ r\theta \;\middle|\; r \geq 0,\ \lim_{h\to+\infty} \frac{x_h - x}{|x_h - x|} = \theta,\ \text{with}\, x_h \in S \setminus \{x\},\ x_h \to x \right\}.$$

We denote by $\mathrm{Tan}(S,x)$ *the vector space spanned by* $T(S,x)$.

Proposition 1.6.2. *Let $S \subset \mathbb{R}^n$ and let $x \in S$ be fixed. We assume that* $\mathrm{Tan}(S,x)$ *has dimension not larger than m, for every $x \in S$. Then S is countably $\mathcal{H}^m$-rectifiable.*

Proof. See [3], Theorem 3.1. □

Remark 1.6.9. The proof passes through a very classical result: if $D \subset \mathbb{R}^m$ and $f : D \longrightarrow \mathbb{R}^n$ is a Lipschitz continuous function, then $f(D)$ is countably $\mathcal{H}^m$-rectifiable. This fact is proved in [86], Lemma 1.1.

We are now approaching the main result of [3]. We need another preliminary step. In order to apply Proposition 1.6.2, we need to know that the contingent cone of $\Sigma^k_\alpha(u)$ has dimension at most $(n-k)$. First a notation.

Notation 1.6.2. Let S be an arbitrary subset of $\mathbb{R}^n$. We denote by $S^\perp$ the set

$$\{p \in \mathbb{H}^n \,|\, q \longmapsto \langle q, p\rangle \text{ is constant on } S\}\,.$$

Proposition 1.6.3. *Let $u : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ be a semiconvex function. Then the set $\Sigma^k_\alpha(u)$ is closed in Ω and*

$$\operatorname{Tan}\left(\Sigma^k_\alpha(u), x\right) \,\subset\, [\partial u(x)]^\perp\,,$$

for every $x \in \Sigma^k_\alpha(u) \setminus \Sigma^{k+1}(u)$. In particular, the dimension of $\operatorname{Tan}\left(\Sigma^k_\alpha(u), x\right)$ *is not larger than $(n-k)$, for any $x \in \Sigma^k_\alpha(u) \setminus \Sigma^{k+1}(u)$.*

Proof. See [3], Proposition 2.2. □

Remark 1.6.10. We point out that, from Proposition 1.6.3, it follows that

$$\operatorname{Tan}\left(\Sigma^n_\alpha(u), x\right) = \{0\},$$

for every $x \in \Sigma^n_\alpha(u)$. This implies that $\Sigma^n_\alpha(u)$ is a discrete set.

Putting together all the results stated above, one can prove the desired result.

Theorem 1.6.4. *Let $u : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ be a semiconvex and Lipschitz continuous function. Then, for every $0 \leq k \leq n$ integer number, the set $\Sigma^k(u)$ is countably $\mathcal{H}^{n-k}$-rectifiable.*

1.7 Optimal mass transport for quadratic cost function

In this section, our aim is to specialize Theorem 1.3.3 to the case of quadratic cost. Moreover, we want to give the statement of Brenier's Theorem, which provides a unique solution to the Monge problem for this specific case. It turns out that there exists a unique optimal tranference plan π and it is associated to the gradient of a suitable convex function. This gradient will be the desired optimal transport map.

We recall that, according to Theorem 1.3.3, under some assumptions, an admissible plan $\pi \in \Pi(\mu, \nu)$ is optimal if and only if it is supported on the graph of $\partial^c \varphi$, for some c-concave function φ. In the specific case of quadratic cost, one can simplify this fact, discovering that $\pi \in \Pi(\mu, \nu)$ is optimal if and only if it is supported on the graph of the subdifferential of a suitable convex function. This is what the following proposition says.

Proposition 1.7.1. *Let $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{-\infty\}$ be a function not identically equal to $-\infty$. Then φ is c-concave if and only if $\bar{\varphi}(x) := \frac{1}{2}|x|^2 - \varphi(x)$ is convex and lower semicontinuous. Moreover, $y \in \partial^c \varphi(x)$ if and only if $y \in \partial \bar{\varphi}(x)$.*

Proof. The proof follows by some simple computations. By definition of c-concavity, there exists a function $\psi : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$ such that

$$\varphi(x) \;=\; \inf_{y \in \mathbb{R}^n} \left(\frac{1}{2}|x-y|^2 - \psi(y) \right).$$

Expanding the square, this is equivalent to say that

$$\varphi(x) - \frac{1}{2}|x|^2 \;=\; \inf_{y \in \mathbb{R}^n} \left(-x \cdot y + \frac{1}{2}|y|^2 - \psi(y) \right),$$

or, equivalently,

$$\bar{\varphi}(x) \;=\; \sup_{y \in \mathbb{R}^n} \left(x \cdot y - \left(\frac{1}{2}|y|^2 - \psi(y) \right) \right),$$

which implies that $\bar{\varphi}$ is convex and lower semicontinuous.

Let us prove the second part of the statement. We have that $y \in \partial^c \varphi(x)$ if and only if

$$\varphi(x) = \varphi^c(y) \;=\; \frac{1}{2}|x-y|^2,$$

which is equivalent to

$$\varphi(x) - \frac{1}{2}|x|^2 \;=\; -x \cdot y + \frac{1}{2}|y|^2 - \varphi^c(y). \tag{1.12}$$

Now, it holds also that $y \in \partial^c \varphi(x)$ if and only if

$$\varphi(z) \;\leq\; \frac{1}{2}|z-y|^2 - \varphi^c(y),$$

for every $z \in \mathbb{R}^n$. Equivalently,

$$\varphi(z) - \frac{1}{2}|z|^2 \leq -y \cdot z + \frac{1}{2}|y|^2 - \varphi^c(y), \tag{1.13}$$

for every $z \in \mathbb{R}^n$. Finally, we have that (1.12) and (1.13), together, are equivalent to

$$\varphi(z) - \frac{1}{2}|z|^2 \leq \varphi(x) - \frac{1}{2}|x|^2 - \langle y, z - x \rangle ,$$

for every $z \in \mathbb{R}^n$. This last inequality means that $y \in \partial\bar{\varphi}(x)$ □

Let us now give a look at the existence of optimal transference plan and the associated optimal transport map. The main result that we propose here is the following: if μ and ν are two probability measures with finite second order moments (see Definition 1.7.1), a trasference plan is optimal if and only if it is supported on the subdifferential of a convex function. This fact was proved by Knott and Smith in [63]. Moreover, if we ask for more stronger assumptions on the two probability measures, Brenier's Theorem assures the uniqueness of such transference plan.

First of all, we recall that the total transportation cost, in this case, is given by

$$I[\pi] = \int_{\mathbb{R}^n \times \mathbb{R}^n} \frac{1}{2}|x - y|^2 d\pi(x, y),$$

and also that

$$J(\varphi, \psi) := \int_{\mathbb{R}^n} \varphi(x) d\mu(x) + \int_{\mathbb{R}^n} \psi(y) d\nu(y).$$

Definition 1.7.1. *Let μ and ν be two probability measures on $\mathbb{R}^n$. We say that they have* finite second order moments *if*

$$M_2 := \int_{\mathbb{R}^n} \frac{1}{2}|x|^2 d\mu(x) + \int_{\mathbb{R}^n} \frac{1}{2}|y|^2 d\nu(y) < +\infty.$$

Remark 1.7.2. This condition implies that the total trasportation cost is always finite on $\Pi(\mu, \nu)$. This fact is a simple consequence of the triangle inequality and of the fact that π has marginals μ and ν.

In order of justify Definition 1.7.1, we start with a proposition about the existence of a minimizer for the functional I. In the proof the condition about the second order moments of μ and ν plays a key role.

Proposition 1.7.2. *Let μ and ν be two probability measures with finite second order moments. Then there exists $\tilde{\pi} \in \Pi(\mu,\nu)$ such that*

$$I[\tilde{\pi}] = \inf_{\pi \in \Pi(\mu,\nu)} I[\pi].$$

Idea of the proof. First, we recall that $\Pi(\mu,\nu)$ is non empty. Moreover, by the finiteness of the second order moments of μ and ν, we have that $\pi \in \Pi(\mu,\nu)$ implies that

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} \left(|x|^2 + |y|^2\right) d\pi(x,y) < +\infty.$$

Therefore, one can deduce that $\Pi(\mu,\nu)$ is compact for the weak topology of probability measures.

Now, let $(\pi_k)_{k\in\mathbb{N}}$ be a minimizing sequence for I. We know that, eventually restricting to a subsequence, there exists a limit point $\pi_0 \in \Pi(\mu,\nu)$. We can write the cost function $c(x,y) = \frac{1}{2}|x-y|^2$ as the supremum of a non decreasing sequence of bounded functions, $(c_h)_{h\in\mathbb{N}}$. Thanks to the Monotone Convergence Theorem, we can conclude that

$$\begin{aligned}\int_{\mathbb{R}^n \times \mathbb{R}^n} c(x,y) d\pi_0(x,y) &= \lim_{h\to\infty} \int_{\mathbb{R}^n \times \mathbb{R}^n} c_h(x,y) d\pi_0(x,y) \\ &\leq \lim_{h\to\infty} \limsup_{k\to\infty} \int_{\mathbb{R}^n \times \mathbb{R}^n} c_h(x,y) d\pi_k(x,y) \\ &\leq \limsup_{k\to\infty} \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x,y) d\pi_k(x,y) = \inf_{\pi\in\Pi(\mu,\nu)} I[\pi].\end{aligned}$$

Hence, π_0 is a minimizer of I □

This Proposition is just a preliminary result of what we want. For the main statement (Theorem 1.7.4), we need to go back to the dual problem. For this particular case of quadratic cost function, we can have a more precise result. First, remember that a pair (φ,ψ) belongs to Φ_c, if

$$\varphi(x) + \psi(y) \leq \frac{1}{2}|x-y|^2,$$

for μ-almost every x and ν-almost every y in $\mathbb{R}^n$. Since we have a specific formula for c, we can rewrite this inequality in this way

$$x \cdot y \leq \left(\frac{1}{2}|x|^2 - \varphi(x)\right) + \left(\frac{1}{2}|y|^2 - \psi(y)\right).$$

It is natural to introduce a new pair of functions:

$$\tilde{\varphi}(x) = \frac{1}{2}|x|^2 - \varphi(x) \quad \text{and} \quad \tilde{\psi}(y) = \frac{1}{2}|y|^2 - \psi(y).$$

Now, using Definition 1.7.1, one has that

$$\inf_{\pi\in\Pi(\mu,\nu)} I[\pi] \;=\; M_2 - \sup_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^n\times\mathbb{R}^n} x\cdot y\, d\pi(x,y), \tag{1.14}$$

and also

$$\sup_{(\varphi,\phi)\in\Phi_c} J(\varphi,\psi) \;=\; M_2 - \inf_{(\tilde{\varphi},\tilde{\psi})\in\tilde{\Phi}_c} J(\tilde{\varphi},\tilde{\psi}), \tag{1.15}$$

where $\tilde{\Phi}_c$ is the set of all pairs $(\tilde{\varphi},\tilde{\psi}) \in L^1(\mathbb{R}^n,\mu)\times L^1(\mathbb{R}^n,\nu)$ such that

$$x\cdot y \le \tilde{\varphi}(x) + \tilde{\psi}(y), \tag{1.16}$$

for μ-almost every x and for ν-almost every y in $\mathbb{R}^n$. Now, let us combine (1.14) and (1.15) with the formula of the Kantorovich Duality Theorem (1.8). We obtain a restatement of the duality

$$\sup_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^n\times\mathbb{R}^n} x\cdot y\, d\pi(x,y) \;=\; \inf_{(\tilde{\varphi},\tilde{\psi})\in\tilde{\Phi}_c} J(\tilde{\varphi},\tilde{\psi}).$$

Consider then $(\tilde{\varphi},\tilde{\psi}) \in \tilde{\Phi}_c$. From inequality (1.16), it holds that, for ν-almost every $y\in\mathbb{R}^n$,

$$\tilde{\psi}(y) \;\ge\; \sup_{x\in\mathbb{R}^n}\left(x\cdot y - \tilde{\varphi}(x)\right) \;:=\; \tilde{\varphi}^*(y).$$

As a consequence, we have that

$$J(\tilde{\varphi},\tilde{\psi}) \;\ge\; J(\tilde{\varphi},\tilde{\varphi}^*). \tag{1.17}$$

Analogously, one can obtain that, for μ-almost every $x\in X$,

$$\tilde{\varphi}(x) \;\ge\; \sup_{y\in\mathbb{R}^n}\left(x\cdot y - \tilde{\varphi}^*(y)\right) \;:=\; \tilde{\varphi}^{**}(x),$$

and therefore

$$J(\tilde{\varphi},\tilde{\varphi}^*) \;\ge\; J(\tilde{\varphi}^{**},\tilde{\varphi}^*). \tag{1.18}$$

Now, we combine (1.17) and (1.18) and we get

$$\inf_{(\tilde{\varphi},\tilde{\psi})\in\tilde{\Phi}_c} J(\tilde{\varphi},\tilde{\psi}) \;\ge\; \inf_{\tilde{\varphi}\in L^1(\mathbb{R}^n,\mu)} J(\tilde{\varphi}^{**},\tilde{\varphi}^*).$$

If we admit that $(\tilde{\varphi}^{**}, \tilde{\varphi}^*) \in L^1(\mathbb{R}^n, \mu) \times L^1(\mathbb{R}^n, \nu)$, we have that $(\tilde{\varphi}^{**}, \tilde{\varphi}^*) \in \tilde{\Phi}_c$. This means that the infimum of J does not change if we restrict to a smaller subset of $\tilde{\Phi}_c$. The pairs of type $(\tilde{\varphi}^{**}, \tilde{\varphi}^*) \in \tilde{\Phi}_c$ are convex lower semicontinuous functions. Indeed, they are the supremum of a family of linear functions. Following the book of Villani (see [92]), we refer to this reduction of the set on which we look for minimizers as the *double convexification trick*.

With this background we are able to give the idea of the proof of the following proposition.

Proposition 1.7.3. *Let μ and ν be two probability measures with finite second order moments. Then there exists a pair (φ, φ^*) of lower semicontinuous proper conjugate functions on $\mathbb{R}^n$ such that*

$$\inf_{\tilde{\Phi}_c} J = J(\varphi, \varphi^*).$$

Idea of the proof. We use the double convexification trick. Let $(\varphi_k, \psi_k)_{k \in \mathbb{N}}$ be a minimizing sequence of J. We assume that (φ_k, ψ_k) are all pairs of convex conjugate functions. The first step is to prove that, eventually reducing to a subsequence, $\varphi_k \longrightarrow \varphi \in L^1(\mathbb{R}^n, \mu)$, $\psi_k \longrightarrow \psi \in L^1(\mathbb{R}^n, \nu)$, $(\varphi, \psi) \in \tilde{\Phi}_c$ and also that

$$J(\varphi, \psi) \leq \liminf_{k \to \infty} J(\varphi_k, \psi_k).$$

If one can prove these facts, then it follows that (φ, ψ) is an optimal pair. Applying the double convexification trick, we can get what we want. □

Remark 1.7.3. In the previous proof, it is important to keep in mind that we need also that $(\varphi_k, \psi_k) \in L^1(\mathbb{R}^n, \mu) \times L^1(\mathbb{R}^n, \nu)$.

We are now ready to state the main theorem about the existence and uniqueness of an optimal transference plan for the case with cost function $c(x, y) = \frac{1}{2}|x - y|^2$. This theorem summarize what we saw in this Section.

Theorem 1.7.4. *Let μ and ν two probability measures with finite second order moments. Then the following three statements are true.*

(i) **Knott-Smith optimality criterion.** *$\pi \in \Pi(\mu, \nu)$ is optimal if and only if there exists a convex lower semicontinuous function φ such that*

$$\mathrm{supp}(\pi) \subset \mathrm{graph}(\partial\varphi). \tag{1.19}$$

(ii) **Brenier's Theorem.** *If μ does not give mass to sets with Hausdorff dimension at most $(n-1)$, then there exists a unique optimal plan π which is*

$$\pi = (Id \times \nabla\varphi)\#\mu, \tag{1.20}$$

where $\nabla\varphi$ is the unique gradient of a convex function such that $\nabla\varphi\#\mu = \nu$.

(iii) *Under the assumption of (ii), $\nabla\varphi$ is the unique solution to the Monge trasportation problem*

$$\int_{\mathbb{R}^n} |x - \nabla\varphi(x)|^2 d\mu(x) = \inf_{T\#\mu=\nu} \int_{\mathbb{R}^n} |x - T(x)|^2 d\mu(x).$$

The proof of this Theorem is quite long and complicate. We direct the reader's attention to the papers of Brenier and of Knott and Smith (see [18], [63]). For a more selfcontained digression, one can refer to Chapter 2 of [92].

Here, for a better understanding, we make a small remark.

Remark 1.7.4. Inclusion (1.19) can be reformulated as follows: $y \in \partial\varphi(x)$ for π-almost every couple $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$. In Brenier's Theorem, the uniqueness of the convex lower semicontinuos function φ has to be intended as a uniqueness up to a μ-measure zero set.

In Section 1.6, we saw that the singular set of the subdifferential of a convex function has Hausdorff dimension at most $n-1$. In part (ii) of Theorem 1.7.4, it is requested that μ does not give mass to sets with Hausdorff dimension at most $n-1$. This implies that we can ignore those points where $\mathrm{supp}(\pi)$ can not be written as a graph, applying finally Proposition 1.5.1.

1.8 Optimal mass transport for general cost functions

Let X and Y be two subsets of $\mathbb{R}^n$ and let μ and ν be two probability measures on X and Y, respectively. We consider a non negative and continuous cost function $c : X \times Y \longrightarrow \mathbb{R}$.

We recall that, by Theorem 1.3.3, under some assumptions on c, it holds that an admissible tranference plan $\pi \in \Pi(\mu, \nu)$ is optimal if and only if it is supported on the graph of the c-superdifferential, $\partial^c\psi$, of a suitable c-concave function $\psi : X \longrightarrow \mathbb{R}$.

The main goal of this Chapter is to understand under what conditions the existence of an optimal plan implies the existence of an optimal map, solution of a Monge's problem. Similarly to what we did for the case of quadratic cost, we want to apply Proposition 1.5.1. This requires to know "how large" is the set $\Sigma^c(\psi)$. Obviously, this depends on the properties of the function c.

Following [93], we introduce certain possible assumptions on c, among which we will choose in the statements of the results we are going to present.

(Super) the function $x \longmapsto c(x, y)$ is everywhere superdifferentiable, for every $y \in \mathbb{R}^n$;

(Twist) if x, y and y' are such that $\nabla_x c(x, y) = \nabla_x c(x, y')$, then $y = y'$;

(Lip) the function $x \longmapsto c(x, y)$ is locally Lipschitz, uniformly in y;

(SC) the function $x \longmapsto c(x, y)$ is a semiconcave, uniformly in y;

(locLip) the function $x \longmapsto c(x, y)$ is locally Lipschitz, locally in y;

(locSC) the function $x \longmapsto c(x, y)$ is a locally semiconcave function, locally in y;

$(\mathbf{H\infty})_\mathbf{1}$ for any x and for any measurable set S for which $T(S, x)$ is not contained in a half-space there is a finite collection of elements $z_1, ..., z_k \in S$, and a small ball B containing x, such that for any y outside a compact set,

$$\inf_{w \in B} c(w, y) \geq \inf_{1 \leq j \leq k} c(z_j, y);$$

$(\mathbf{H}\infty)_2$ for any x and any neighbourhood U of x there is a small ball B containing x such that

$$\lim_{y\to\infty} \sup_{w\in B} \inf_{z\in U} [c(z,y) - c(w,y)] = -\infty.$$

Now, we list a serie of results concerning the differentiability of a c-convex function. All together, these results allow us to give an estimate for the Hausdorff dimension of the c-singular set of the c-superdifferential of a c-concave function. In this way one can apply Proposition 1.5.1 and obtain the existence of an optimal transport map.

Theorem 1.8.1. *Let $X \subset \mathbb{R}^n$ be a set. Let c be satisfying Assumption* $(\mathbf{H}\infty)$ *and let $\psi : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a c-convex function. Let us denote by Ω the interior of the set $\psi^{-1}(\mathbb{R})$. Then*

(i) $\psi^{-1}(\mathbb{R}) \setminus \Omega$ is a set of dimension at most $n-1$;

(ii) ψ is locally bounded and c-subdifferentiable everywhere in Ω;

(iii) if $K \subset \Omega$ is a compact set, then also $\partial_c\psi(K)$ is a compact set.

Proof. See [93], Theorem 10.24. □

Theorem 1.8.2. *Let $X \subset \mathbb{R}^n$ be a set. Let c be satisfying Assumption* (**Super**) *and let $\psi : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a c-convex function. Let $x \in X$ be a point such that $\partial_c\psi(x) \neq \emptyset$. Then ψ is subdifferentiable at x.*

Proof. See [93], Theorem 10.25. □

Theorem 1.8.3. *Let c be satisfying Assumptions* (**Super**), (**Twist**) *and* (**SC**). *Let $\psi : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a c-convex function. Then ψ is locally semiconvex and differentiable in the interior of its domain, Ω, apart from a set of dimension at most $n-1$.*

Proof. See [93], Theorem 10.26. □

Remark 1.8.1. The previous three statements, combined together, give a final result about the differentiability points of a c-convex function ψ. If c satisfies Assumptions (**Super**), (**Twist**), (**H**∞) and (**SC**), it follows that

$$\dim\{x \in \Omega \,|\, \#\partial_c\psi(x) \neq 1\} \,\leq\, n-1.$$

Thanks to remark 1.8.1, one has finally the following theorem.

Theorem 1.8.4. *Let X and Y two subsets of $\mathbb{R}^n$ and let $\mu \in P(X)$ and $\nu \in P(Y)$. Consider a cost function $c : X \times Y \longrightarrow \mathbb{R}^+$ such that*

(i) c is continuous;

(ii) c satisfies Assumption (**locSC**)*, i.e. c is locally semiconcave;*

(iii) c satisfies Assumption (**Twist**)*, i.e. c is injective;*

(iv) μ does not give mass to sets of dimension at most $n-1$.

Then there exists a unique optimal transference plan π between μ and ν and there exists a c-convex function $\psi : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ such that $\pi(\partial^c\psi)$.

Moreover, if we assume also that c satisfies Assumption (**SC**)*, one can define a continuous map $x \longmapsto T(x)$ on the set of differentiability points of ψ by the relation $T(x) \in \partial^c\psi$, and the* $\mathrm{supp}\nu = \overline{T(\mathrm{supp}\mu)}$

Remark 1.8.2. We point out that if one can say that the set of points where the function ψ is not differentiable is of dimension at most $n-1$, then we have the existence of a (unique) solution of the Monge problem. Actually, this is what Remark 1.8.1 says. One of the key ingredients is the semiconcavity of the cost function (Assumption (**SC**)). The goal of the next Chapter, in fact, is to find a way to avoid this request.

Chapter 2

Hausdorff dimension estimate of c-singular sets

In Section 1.8, we stated some conditions over the cost function c, which imply existence and uniqueness of the solution for the Monge-Kantorovich problem. This solution is given in term of the c-subdifferential of a c-convex function. In this chapter the main goal is to weaken these assumptions and preserve a result similar to Theorem 1.8.4.

Section 2.1 is devoted to introduce our new assumptions. Continuity, differentiability and twist condition are preserved but the request for c to be semiconcave is missed. Here we also discuss some examples.

Section 2.2 wants to be a very short recalling of the notion of porosity. We also state and prove a geometric condition for a set to be porous. This will be the central ingredient of the proof of the main theorem of Section 2.3, where we give an estimate for the size of the set of singular points of a c-concave function.

2.1 A general class of cost functions c

Our setting here is the Euclidean space $\mathbb{R}^n$. As already mentioned, we want to list three assumptions on the cost function. The first and the second are totally analogous to (**Super**) and (**Twist**). Indeed, Assumption (2.1.2) is a quantitative restatement of injectivity. The third one is new. It says that the mapping $y \longmapsto \nabla_x c(x,y)$ can not have too many discontinuities close to each other. This is the condition which replace the semiconcavity assumption on c.

Assumption 2.1.1 (Differentiability Condition)**.** *The cost function $c : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ is a continuous function. For every $p \in \mathbb{R}^n$, the function $x \longmapsto c(x,p)$ is differentiable almost everywhere. We denote by*

$$C(p) \;=\; \{x \in \mathbb{R}^n \,|\, \nabla_x c(x,p) \text{ exists}\}\,.$$

We assume that for every $x \in \mathbb{R}^n$, there exists $r = r(x) > 0$ such that for all $0 < r < r(x)$ there exists $\hat{x} \in B(x,r)$ with the property that $\hat{x} \in C(p)$ for all $p \in \mathbb{R}^n$. Furthermore, we assume that Taylor's formula holds in the sense that

$$\lim_{r\to 0} \frac{c(x,p) - c(\hat{x},p) - < \nabla_x c(\hat{x},p), (x-\hat{x}) >}{r} = 0. \tag{2.1}$$

Assumption 2.1.2 (Twist Condition)**.** *For any compact set $K \subset \mathbb{R}^n$ and for any $0 < \eta < 1$, there exists $0 < \xi = \xi(\eta) < 2$ such that if $p_1, p_2 \in K$ and $x \in K$ are so that $|p_1 - p_2| > \eta$ and $x \in C(p_1) \cap C(p_2)$, then*

$$|\nabla_x c(x,p_1) - \nabla_x c(x,p_2)| \;>\; \xi. \tag{2.2}$$

Assumption 2.1.3 (Gradient Continuity Condition)**.** *For any compact set $K \subset \mathbb{R}^n$, for any $\varepsilon > 0$ and for any $\eta > 0$, there exists $\delta = \delta(\varepsilon,\eta) > 0$ such that if $p_1, p_1', p_2, p_2' \in K$ and $x \in K$ are so that $|p_1 - p_2| > \eta$, $|p_1 - p_1'| < \delta$, $|p_2 - p_2'| < \delta$ and $x \in C(p_1) \cap C(p_1') \cap C(p_2) \cap C(p_2')$, then either*

$$|\nabla_x c(x,p_1) - \nabla_x c(x,p_1')| \;<\; \varepsilon \tag{2.3}$$

or

$$|\nabla_x c(x,p_2) - \nabla_x c(x,p_2')| \;<\; \varepsilon. \tag{2.4}$$

Remark 2.1.1. We point out that, in the first condition, the required differentiability at the point $\hat{x}$ does not necessarily imply the validity of Taylor's formula in the form that it is written in (2.1). The reason for this is that in Taylor's development formula the radius $r > 0$ depends on the choice of the development point $\hat{x}$. On the other hand, in our case the radius $r = r(x) > 0$ depends on x and not on $\hat{x}$ and this is why we need to require in addition the validity of (2.1). In practical applications however it is rather easy to check this condition as for typical cost functions the differentiability fails only on the diagonal $x = p$.

It is natural now to ask a question: is it really true that Assumption 2.1.3 is a more relaxed condition with respect to (**SC**)? To give an answer we present some examples.

Let us observe that the standard quadratic cost function: $c : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$, $c(x, y) = \frac{1}{2}|x - y|^2$ satisfies Assumptions 2.1.1, 2.1.2 and 2.1.3.

The next example, the case of the linear cost function $c(x, y) = |x - y|$, shows that the second assumption (Twist Condition) is necessary. In this case the first and the third assumption (and not the second one) are satisfied. Moreover, as a preliminary counterexample for Theorem 2.3.1, we also prove that there exists a c-concave function u whose singular set $\Sigma^c(u)$ is of full dimension.

Example 2.1.1. Consider the following cost function

$$\begin{aligned} c : \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto |x - y|. \end{aligned}$$

This function is continuous and, for every $y \in \mathbb{R}^n$ fixed, $x \longmapsto |x - y|$ is differentiable except for the point $x = y$. This shows that Assumption 2.1.1 is satisfied. Moreover, for $x \neq y$ we have that $\nabla_x c(x, y) = \frac{x-y}{|x-y|}$. Since the only discontinuity point of this map is at $x = y$ it easy to see that Assumption 2.1.3 is also satisfied. On the other hand we can observe that, $\nabla_x c(x, \cdot)$ is not injective, i.e. there exist x, y_1 and $y_2 \in \mathbb{R}^n$ such that $y_1 \neq y_2$ but

$$\nabla_x c(x, y_1) = \nabla_x c(x, p_2).$$

Let us fix $x = 0$ and consider $y \neq x$. We have that

$$\nabla_x c(0, y) = -\frac{y}{|y|}.$$

Now, if $y_1 = \lambda y_2$, for some $\lambda \in \mathbb{R}^+$, then

$$\nabla_x c(0, y_1) = \nabla_x c(0, \lambda y_2) = -\frac{\lambda y_2}{\lambda |y_2|} = \nabla_x c(0, y_2).$$

It is clear that this fact conflicts with Assumption 2.1.2.

Let us now indicate a c-concave function $u : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that $\dim \Sigma^c(u) = n$. Let $x_0 \in \mathbb{R}^n$ be fixed and consider the function

$$\psi(y) \;=\; |y - x_0|.$$

We write the respective c-concave function

$$u(x) \;=\; \inf_{y \in \mathbb{R}^n} \left(|x - y| - |y - x_0|\right). \tag{2.5}$$

By the triangle inequality, it holds that

$$|x - y| \;-\; |y - x_0| \;\geq\; -|x - x_0|.$$

This implies that the infimum in (2.5) is achieved when $y = x$. Therefore

$$u(x) \;=\; -|x - x_0|.$$

We prove now that, for every $x \in \mathbb{R}^n$, it holds that

$$\#\partial^c u(x) \;>\; 1.$$

Let us choose an arbitrary $x \in \mathbb{R}^n$. By definition of c-superdifferential, we have that $y \in \partial^c u(x)$ if and only if

$$|x - y| + |x - x_0| \;\leq\; |z - y| + |z - x_0|, \tag{2.6}$$

for every $z \in \mathbb{R}^n$.

Define now $y_t := x_0 + t(x - x_0)$ for $t \in \mathbb{R}$. We want to show that y_t satisfies (2.6) for every $t > 1$. Let us perform some calculations:

$$\begin{aligned} |x - y_t| + |x - x_0| &= |x - x_0 - t(x - x_0)| + |x - x_0| = t|x - x_0| \\ &= |t(x - x_0) + x_0 - x_0| = |y_t - x_0| \\ &\leq |y_t - z| + |z - x_0|, \end{aligned} \tag{2.7}$$

where the last inequality trivially follows from the triangle inequality and is true for each choice of $z \in \mathbb{R}^n$. This means that

$$\{y_t \in \mathbb{R}^n \,|\, y_t = x_0 + t(x - x_0), \text{ with } t > 1\} \subset \partial^c u(x),$$

and, since x was choosen arbitrarily, implies that

$$\Sigma^c(u) \;=\; \mathbb{R}^n.$$

The next example is a cost function that is simply the sum of the two previously considered cost functions. As we shall see all three Assumptions are satisfied and so the statement of Theorem 2.3.1 applies. On the other hand, it does not satisfy the classical Assumption (**SC**) about semiconcavity .

Example 2.1.2. Consider the cost function

$$\begin{aligned} c\,:\,\mathbb{R}^n\times\mathbb{R}^n &\longrightarrow \mathbb{R}\\ (x,y) &\longmapsto |x-y|+|x-y|^2. \end{aligned}$$

It is evident that c is a continuous function and that, for every fixed $y\in\mathbb{R}^n$, $x\longmapsto c(x,p)$ is differentiable in $\mathbb{R}^n\setminus\{y\}$.

Let us prove that c satisfies the Twist Condition. First of all we compute the gradient of the cost function with respect to the x variable:

$$\nabla_x c(x,y) \;=\; \frac{x-y}{|x-y|}+2(x-y). \tag{2.8}$$

In order to proceed we need the following:

Claim. Let v and w be two nonzero vectors in $\mathbb{R}^n$. Then

$$\left|\left(v+\frac{v}{|v|}\right)-\left(w+\frac{w}{|w|}\right)\right| \;\geq\; |v-w|.$$

Let us prove this Claim. For simplicity of notation, we set $a:=|v|$ and $b:=|w|$. Therefore, $a+1=\left|v+\frac{v}{|v|}\right|$ and $b+1=\left|w+\frac{w}{|w|}\right|$. Now, it is well known that

$$|v-w| \;=\; \sqrt{a^2+b^2-2ab\cos\gamma},$$

where γ is the angle between v and w on the plane spanned by the two vectors. Analogously, we can compute

$$\begin{aligned} \left|\left(v+\frac{v}{|v|}\right)-\left(w+\frac{w}{|w|}\right)\right| &= \sqrt{(a+1)^2+(b+1)^2-2(a+1)(b+1)\cos\gamma}\\ &= \sqrt{a^2+b^2-2ab\cos\gamma+2(1+a+b)(1-\cos\gamma)}\\ &\geq \sqrt{a^2+b^2-2ab\cos\gamma} \;=\; |v-w|, \end{aligned}$$

and this completes the proof of the claim.

Fix $0 < \eta < 1$ and $x \in \mathbb{R}^n$. We aim to show that there exists $\xi > 0$, dependent on η but not on x, such that, for every $y_1, y_2 \in \mathbb{R}^n$ so that $y_1 \neq x \neq y_2$ and $|y_1 - y_2| > \eta$, it follows

$$|\nabla_x c(x, y_1) - \nabla_x c(x, y_2)| > \xi.$$

It is here that we need to apply the Claim:

$$\begin{aligned} |\nabla_x c(x, y_1) - \nabla_x c(x, y_2)| &= \left| \left(\frac{x - y_1}{|x - y_1|} + 2(x - y_1) \right) - \left(\frac{x - y_2}{|x - y_2|} + 2(x - y_2) \right) \right| \\ &\geq |2(x - y_1) - 2(x - y_2)| = 2|y_1 - y_2| > 2\eta. \end{aligned}$$

If we choose $2\eta > \xi$, we can conclude that Assumption 2.1.2 is satisfied.

Let us check Assumption 2.1.3. Fix $\varepsilon > 0$ and $\eta > 0$. For our purposes, we can assume that $\varepsilon < 1$ and $\eta < 2$. Fix also $x \in \mathbb{R}^n$. We would like to find a suitable $\delta = \delta(\varepsilon, \eta) > 0$, not dependent on x, such that if y_1, y_1' and $y_2, y_2' \in \mathbb{R}^n$ are so that $|y_1 - y_1'| < \delta$, $|y_2 - y_2'| < \delta$, $|y_1 - y_2| > \eta$ and $x \notin \{y_1, y_1', y_2, y_2'\}$, then either

$$|\nabla_x c(x, y_1) - \nabla_x c(x, y_1')| < \varepsilon \tag{2.9}$$

or

$$|\nabla_x c(x, y_2) - \nabla_x c(x, y_2')| < \varepsilon. \tag{2.10}$$

We need to do some calculations, for $i = 1, 2$.

$$\begin{aligned} |\nabla_x c(x, y_i) - \nabla_x c(x, y_i')| &= \left| \left(\frac{x - y_i}{|x - y_i|} + 2(x - y_i) \right) - \left(\frac{x - y_i'}{|x - y_i'|} + 2(x - y_i') \right) \right| \\ &= \left| \frac{x - y_i}{|x - y_i|} - \frac{x - y_i'}{|x - y_i'|} - 2(y_i - y_i') \right|. \end{aligned}$$

Let us fix $\delta = \frac{1}{16}\eta\varepsilon$ and check that it is a good choice. We can assume that $|x - y_2| > \frac{\eta}{4}$ and $|x - y_2'| > \frac{\eta}{4}$. If these inequalities are not satisfied, then we can switch and consider the case $|x - y_1| > \frac{\eta}{4}$ and $|x - y_1'| > \frac{\eta}{4}$. Let us prove why this is true. Assume that $|x - y_2| > \frac{\eta}{4}$ and $|x - y_2'| \leq \frac{\eta}{4}$. The cases with $|x - y_2| \leq \frac{\eta}{4}$ and $|x - y_2'| > \frac{\eta}{4}$, or with $|x - y_2| \leq \frac{\eta}{4}$ and $|x - y_2'| \leq \frac{\eta}{4}$, can be treated in the same way. First, we notice that, since $\varepsilon < 1$,

$$|y_1' - y_2| \geq |y_1 - y_2| - |y_1 - y_1'| \geq \eta - \frac{1}{16}\eta\varepsilon > \frac{15}{16}\eta. \tag{2.11}$$

Moreover, we can also write that

$$|y_1' - y_2'| \geq |y_1' - y_2| - |y_2 - y_2'| \geq \frac{15}{16}\eta - \frac{1}{16}\eta\varepsilon = \frac{7}{8}\eta. \tag{2.12}$$

Extimates (2.12) allow us to conclude that

$$|x - y_1'| \geq |y_1' - y_2'| - |x - y_2'| \geq \frac{7}{8}\eta - \frac{\eta}{4} = \frac{5}{8}\eta > \frac{\eta}{4}.$$

On the other hand, it holds also that

$$|x - y_1| \geq |y_1 - y_2'| - |x - y_2'| \geq |y_1 - y_2| - |y_2' - y_2| - |x - y_2'| > \eta - \frac{1}{16}\eta - \frac{\eta}{4} > \frac{\eta}{4}.$$

This implies that it is not loss of generality if we assume that

$$|x - y_2| > \frac{\eta}{4} \quad \text{and} \quad |x - y_2'| > \frac{\eta}{4}. \tag{2.13}$$

We are now ready to prove that, with our choice of δ, (2.10) holds. First of all, one has

$$\begin{aligned} |\nabla_x c(x, y_2) - \nabla_x c(x, y_2')| &= \left| \frac{x - y_2}{|x - y_2|} - \frac{x - y_2'}{|x - y_2'|} - 2(y_2 - y_2') \right| \\ &\leq \left| \frac{x - y_2}{|x - y_2|} - \frac{x - y_2'}{|x - y_2'|} \right| + 2|y_2 - y_2'| \\ &< \left| \frac{x - y_2}{|x - y_2|} - \frac{x - y_2'}{|x - y_2'|} \right| + \frac{1}{8}\eta\varepsilon. \end{aligned} \tag{2.14}$$

Writing $r = |x - y_2|$ and $r' = |x - y_2'|$, there exist two unit vectors v and $v' \in \mathbb{R}^n$ such that $y_2 = x + rv$ and $y_2' = x + r'v'$. It holds that

$$\left| \frac{x - y_2}{|x - y_2|} - \frac{x - y_2'}{|x - y_2'|} \right| = |v - v'|.$$

Without loss of generality, we can assume that $r \leq r'$. We have

$$|y_2 - y_2'| = |rv - r'v'| = |(r - r')v + r'(v - v')| \geq r'|v - v'| - (r' - r).$$

This inequality implies the following extimate

$$\begin{aligned} |v - v'| &\leq \frac{r' - r}{r'} + \frac{|y_2 - y_2'|}{r'} \leq \frac{4}{\eta}((r' - r) + |y_2 - y_2'|) \\ &\leq \frac{4}{\eta}(|x - y_2'| - |x - y_2|) + |y_2 - y_2'|) \leq \frac{8}{\eta}|y_2 - y_2'| < \frac{\varepsilon}{2}. \end{aligned}$$

Let us come back to (2.14). We have that

$$|\nabla_x c(x, y_2) - \nabla_x c(x, y_2')| < \left| \frac{x - y_2}{|x - y_2|} - \frac{x - y_2'}{|x - y_2'|} \right| + \frac{1}{8}\eta\varepsilon < \frac{\varepsilon}{2} + \frac{1}{8}\eta\varepsilon < \varepsilon,$$

which is exactly what we wanted.

In our final example we illustrate applications to our result to the case of sub-Riemannian type singular metrics. For illustrative purposes we restrict ourselves to the case of the first Heisenberg group (for more details, the reader can see Section 4.1). Let us start by introducing the following:

Notation 2.1.1. We denote by $\xi_1 : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$, $\xi_1(x, y, t) = (x, y)$, the projection to the first components and by $\xi_2 : \mathbb{R}^3 \longrightarrow \mathbb{R}$, $\xi_2(x, y, t) = t$, the projection to the last variable.

Moreover, if we have two points $p = (x_1, y_1, t_1) \in \mathbb{R}^3$ and $q = (x_2, y_2, t_2) \in \mathbb{R}^3$, we introduce the Heisenberg group operation by

$$p^{-1} \cdot q = (x_2 - x_1, y_2 - y_1, t_2 - t_1 - 2(x_2 y_1 - x_1 y_2)). \tag{2.15}$$

Finally, we introduce this function

$$\begin{aligned} N : \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ (x, y, t) &\longmapsto \left(\left(x^2 + y^2\right)^2 + t^2 \right)^{\frac{1}{4}}, \end{aligned}$$

that is the so-called *Korànyi norm* on the first Heisenberg group $\mathbb{H}^1$.

Example 2.1.3. We define a cost function as follows

$$\begin{aligned} c : \mathbb{R}^3 \times \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ (p, q) &\longmapsto N^2(p^{-1} \cdot q) + \xi_2(q - p)^2. \end{aligned}$$

Our aim is to show that Theorem 2.3.1 can be applied for this cost function, which is not semiconcave. Clearly, the function $p \longmapsto c(p, q)$ is differentiable for every $p \in \mathbb{R}^3$ with $p \neq q$. This means that Assumption 2.1.1 is satisfied.

Let us prove that Assumption 2.1.2 is satisfied too. First, we show that the map

$$\begin{aligned} \mathbb{R}^3 \setminus \{q\} &\longrightarrow \mathbb{R}^3 \\ p &\longmapsto \nabla_q c(p, q) \end{aligned}$$

is injective for every $q \in \mathbb{R}^3$. Without loss of generality we can assume that $q = 0$. Hence, one has that, for $p = (x, y, t) \neq 0$,

$$\nabla_q c(p, 0) \;=\; -\frac{1}{2}\frac{1}{N^2(p)}\left(4(x^2+y^2)x - 4ty, 4(x^2+y^2)y + 4tx, 2t\right) + (0, 0, 2t). \tag{2.16}$$

Now, let $p_1 = (x_1, y_1, t_1)$ and $p_2 = (x_2, y_2, t_2) \in \mathbb{R}^3 \setminus \{0\}$ be such that

$$\nabla_q c(p_1, 0) \;=\; \nabla_q c(p_2, 0). \tag{2.17}$$

We want to prove that $p_1 = p_2$. From formula (2.16), it is easy to see that, for $i = 1, 2$,

$$\|\xi_1 (\nabla_q c(p_i, 0))\|^2_{\mathbb{R}^2} \;=\; 4\left(x_i^2 + y_i^2\right).$$

Hence, thanks to equality (2.17), it holds that $x_1^2 + y_1^2 = x_2^2 + y_2^2 := k$. Consider now the third component of $\nabla_q c(p_i, 0)$. If we show that the function

$$f(t) \;=\; \frac{t}{(k^2+t^2)^{\frac{1}{2}}} + 2t$$

is injective, then (2.17) implies that $p_1 = p_2$. This fact is clearly true, because f is strictly monotone, and it guarantees injectivity of the map $p \longmapsto \nabla_q c(p, 0)$ on $\mathbb{R}^3 \setminus \{0\}$.

Let us consider Assumption 2.1.2. We argue by contradiction and we assume that there exist a compact set $K \subset \mathbb{R}^3$ and $0 < \eta < 1$ such that for every $n \in \mathbb{N}$ there are $p_1^{(n)}$, $p_2^{(n)} \in K$ and $q_n \in K$ so that $\left|p_1^{(n)} - p_2^{(n)}\right| > \eta$, $q_n \in C\left(p_1^{(n)}\right) \cap C\left(p_2^{(n)}\right)$ and

$$\left|\nabla_q c\left(p_1^{(n)}, q_n\right) - \nabla_q c\left(p_2^{(n)}, q_n\right)\right| \leq \frac{1}{n}. \tag{2.18}$$

Now, since $\left\{p_1^{(n)}\right\}_{n\in\mathbb{N}}$, $\left\{p_2^{(n)}\right\}_{n\in\mathbb{N}} \subset K$ and $(q_n)_{n\in\mathbb{N}} \subset K$ and K is a compact set, eventually restricting to subsequences, we can assume that there exist p_1^∞, p_2^∞ and $q_0 \in K$ such that $p_1^{(n)} \longrightarrow p_1^\infty$, $p_2^{(n)} \longrightarrow p_2^\infty$ and $q_n \longrightarrow q_0$, as $n \to \infty$. Without loss of generality, we can assume that $q_0 = 0$. We need to consider two cases:

(i) $p_1^\infty \neq 0$ and also $p_2^\infty \neq 0$;

(ii) $p_1^\infty = 0$ or $p_2^\infty = 0$.

In the first case, it holds that, for $i = 1, 2$,

$$\nabla_q c(p_i^{(n)}, q_n) \longrightarrow \nabla_q c(p_i^\infty, 0),$$

as $n \to \infty$. Hence, it follows that

$$|\nabla_q c(p_1^\infty, 0) - \nabla_q c(p_2^\infty, 0)| = 0,$$

but we have a contradiction with the injectivity, because $|p_1^\infty - p_2^\infty| > \eta$.

Consider now the second case. Assume that $p_1^\infty = 0$, the other case is totally analogous. By restricting to subsequence if needed, we can assume that the limit $v = \lim_{n\to\infty} \nabla_q c\left(p_1^{(n)}, 0\right)$ exists. From explicit formula (2.16), one can easily see that

$$v = \lim_{n\to\infty} \nabla_q c\left(p_1^{(n)}, 0\right) \in \{0\} \times \{0\} \times [-1, 1]. \tag{2.19}$$

Again, we have two cases $p_2^\infty = (0, 0, t_2^\infty)$ or $p_2^\infty = (x_2^\infty, y_2^\infty, t_2^\infty) \notin \{0\} \times \{0\} \times \mathbb{R}$.

In the first case, we notice that

$$|\xi_2 (\nabla_q c(p_2^\infty, 0))| = \left| \frac{t_2}{|t_2|} + 4t_2 \right| > 1.$$

Therefore, $\nabla_q c(p_2^\infty, 0) \notin \{0\} \times \{0\} \times [-1, 1]$. This fact, and (2.19), provide a contradiction to (2.18) for large values of n.

On the other hand, when $p_2^\infty \neq (0, 0, t_2^\infty)$, we know that

$$\xi_1 (\nabla_q c (p_2^\infty, 0)) = 4 \left((x_2^\infty)^2 + (y_2^\infty)^2\right) \neq 0.$$

Therefore, also in this case $\nabla_q c(p_2^\infty, 0) \notin \{0\} \times \{0\} \times [-1, 1]$. Again, we have a contradiction with (2.18) and (2.19).

It remains to prove that this cost function satisfies Assumption 2.1.3. Let $K \subset \mathbb{R}^3$ be a compact set and let $\varepsilon > 0$ and $\eta > 0$ be fixed. We want to find $\delta = \delta(\varepsilon, \eta) > 0$ such that if $p_1, p_2, p_1', p_2', q \in K$ are so that $|p_1 - p_2| > \eta$, $|p_i - p_i'| < \delta$, for $i = 1, 2$, and $q \in C(p_1) \cap C(p_1') \cap C(p_2) \cap C(p_2')$, then one among (2.3) and (2.4) holds.

Set $\delta_1 \leq \frac{1}{16}\eta\varepsilon$. With this choice of δ_1, analogously to Example 2.1.2, we can assume that $|p_2 - q| \geq \frac{\eta}{4}$ and $|p_2' - q| \geq \frac{\eta}{4}$.

Let us denote

$$\Delta_\eta = \left\{(p,q) \in \mathbb{R}^3 \times \mathbb{R}^3 \,\middle|\, |p-q| < \frac{\eta}{4}\right\}$$

and consider the continuous function

$$\begin{aligned} F : K \times K \setminus \Delta_\eta &\longrightarrow \mathbb{R}^3 \\ (p,q) &\longmapsto \nabla_q c(p,q). \end{aligned}$$

Since F is a continuous function over a compact set, it is uniformly continuous. Therefore, there exists $\delta_2 = \delta_2(\varepsilon, \eta) > 0$ such that, if $|p - p'| < \delta_2$, then

$$|\nabla_q c(p,q) - \nabla_q c(p',q)| < \varepsilon.$$

Now, if we set $\delta = \min\{\delta_1, \delta_2\}$, Assumption 2.1.3 is satisfied.

2.2 Dimension estimate of porous sets

This section is essentially based on [67] and the book of the same author [68], in particular Chapter 11. We recall here an estimate of the Hausdorff dimension for porous sets (see Theorem 2.2.1) and we propose a sufficient condition for porosity. The proof is very simple and based only on some geometrical observations.

Roughly speaking, a porous set A is a "sparse" set: in the neighbourhood of each point of A, one can find a hole which does not touch the set.

Definition 2.2.1. *Let $A \subset \mathbb{R}^n$, $x \in \mathbb{R}^n$ and $r > 0$ be fixed. We set*

$$p(A,x,r) := \sup\{\rho > 0 \,|\, B(z,\rho) \subset B(x,r) \setminus A, \text{ for some } z \in \mathbb{R}^n\}$$

We call the (strong) porosity *of A at the point x the number*

$$p(A,x) := \liminf_{r \to 0} \frac{p(A,x,r)}{r}.$$

Remark 2.2.2. Notice that for any point $x \in A$ we have $0 \leq p(A, x, r) \leq \frac{r}{2}$. Consequently, it follows that $0 \leq p(A, x) \leq \frac{1}{2}$.

Theorem 2.2.1. *Let $p \in]0, \frac{1}{2}[$. Then there exists $d(p) \in [n-1, n]$ such that*

$$\lim_{p \to \frac{1}{2}} d(p) = n - 1$$

and $\dim A \leq d(p)$, for every $A \subset \mathbb{R}^n$ with the property that $p(A, x) \geq p$, for every $x \in A$.

Proof. See [68], Theorem 11.14. □

We can, now, give the proof of the preannounced geometric lemma, which will be a key ingredient in the proof of Theorem 2.3.1 and which generates a condition for a set to be porous, for the details of this fact we postpone to Remark 2.2.3.

Lemma 2.2.1. *Let $x \in \mathbb{R}^n$, $r > 0$, $w \in \mathbb{S}^{n-1}$ and $0 < \rho < \frac{1}{2}$ be fixed. If $x' \in B(z, (1/2 - \rho)r)$ where $z = x + \frac{1}{2}rw$, then*

$$\left\langle \frac{x' - x}{r}, w \right\rangle > \rho. \tag{2.20}$$

Proof. Let us write $x' \in B(z, (1/2 - \rho)r)$ in the form $x' = z + y$ where $|y| < (\frac{1}{2} - \rho)r$.

Since $z = x + \frac{1}{2}rw$ we obtain $x' = x + \frac{1}{2}rw + y$ for $y \in \mathbb{R}^n$, $|y| < (\frac{1}{2} - \rho)r$.

Observe that

$$\langle x' - x, w \rangle = \left\langle \frac{1}{2}rw + y, w \right\rangle = \frac{1}{2}r + \langle y, w \rangle \geq \frac{1}{2}r - |y| > \frac{1}{2}r - (\frac{1}{2} - \rho)r = \rho r > 0.$$

Dividing this inequality by $r > 0$ yields the claim.

□

Remark 2.2.3. Notice that this Lemma provides a sufficient condition for porosity. Let $A \subset \mathbb{R}^n$ and let $x \in A$ be fixed. If there exists $r_0 = r_0(x) > 0$ such that for all $0 < r < r_0(x)$ there exists a vector $w = w(x, r) \in \mathbb{S}^{n-1}$ such that for all $x' \in A \cap B(x, r)$ we have

$$\left\langle \frac{x' - x}{r}, w \right\rangle \leq \rho, \tag{2.21}$$

then according to the above Lemma $x' \notin B\left(z, \left(\frac{1}{2} - \rho\right) r\right)$, for every $0 < r < r_0$. This implies that $p(A, r, x) \geq \left(\frac{1}{2} - \rho\right) r$.

2.3 Dimension estimate of the singular set of a c-concave function

In this section we prove our main results of this first part of the dissertation. We make some assumptions on the cost function c and we give an estimate for the dimension of the c-singular set of a c-concave function. As already mentioned, we need to ask to c to satisfy an injectivity condition, the so-called *twist condition*, to be differentiable in some sense and to satisfy a continuity condition regarding its gradient $\nabla_x c$.

The section ends with an application of Theorem 2.3.1: under our assumptions, we have existence of a solution to Monge's problem.

Theorem 2.3.1. *Let $u : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a c-concave function. Suppose that c satisfies Assumption 2.1.1, 2.1.3 and 2.1.2. Then* $\dim \Sigma^c(u) \leq n-1$.

Proof of Theorem 2.3.1. Our goal is to show that, given a c-concave function $u : \mathbb{R}^n \longrightarrow \mathbb{R}$, the singular set $\Sigma(u)$ has dimension at most $n-1$.

Notice that we can restrict to the case of non isolated points. Indeed, the set of isolated points of $\Sigma(u)$ has zero dimension. For simplicity of notation, we denote by Σ the set $\{x \in \Sigma(u) \,|\, x \,\text{is not isolated}\}$.

The proof will be divided in several steps, with the purpose of simplifying the set to which we apply our argument.

Step 1. Let us define, for every ν and $N \in \mathbb{N}$,

$$\Sigma_{\nu,N} = \left\{ x \in \overline{B(0,N)} \,\middle|\, \exists\, p, q \in \overline{B(0,N)} \cap \partial^c u(x),\, d(p,q) > \frac{1}{\nu} \right\}. \tag{2.22}$$

Therefore, we can assert that

$$\Sigma = \bigcup_{\nu,N \in \mathbb{N}} \Sigma_{\nu,N}. \tag{2.23}$$

Let us check that (2.23) is true. If $x \in \bigcup_{\nu,N\in\mathbb{N}} \Sigma_{\nu,N}$, there are ν and $N \in \mathbb{N}$ such that $p, q \in \overline{B(0,N)} \cap \partial^c u(x)$ with $d(p,q) > \frac{1}{\nu}$. This implies that $p \neq q$ so $\#\partial^c u(x) > 1$.

Now the opposite inclusion should be proved. Let $x \in \Sigma$. Since $\#\partial^c u(x) > 1$, there are $p, q \in \partial^c u(x)$ such that $d := d(p,q) > 0$. We choose $\bar{N} = \lceil \max\{|p|, |q|\} + 1 \rceil$ and $\bar{\nu} \in \mathbb{N}$ such that $\bar{\nu} > \frac{1}{d}$. With this choice, we can conclude that $x \in \Sigma_{\bar{\nu},\bar{N}}$.

Decomposition (2.23) will help us to reach our goal. Indeed, it is clear that

$$\dim \Sigma \;=\; \dim \left(\bigcup_{\nu,N} \Sigma_{\nu,N} \right) \;=\; \sup\{\dim \Sigma_{\nu,N} \,|\, \nu, N \in \mathbb{N}\}.$$

The decomposition built in this first step allows us to focus our attention on a set of the form

$$\Sigma_{\nu,N} \;=\; \left\{ x \in \overline{B(0,N)} \;\middle|\; \partial^c u(x) \subset \overline{B(0,N)} \text{ and } \operatorname{diam} \partial^c u(x) > \frac{1}{\nu} \right\}.$$

We shall namely prove that $\dim \Sigma_{\nu,N} \leq n-1$ for general ν and $N \in \mathbb{N}$ fixed.

Step 2. From now on, let $\nu \in \mathbb{N}$ and $N \in \mathbb{N}$ be fixed. The proof of the Theorem is a direct consequence of the following

Claim. Let $0 < \rho < \frac{1}{2}$. Then there exists a finite family of sets $\left\{\Sigma^i_{\nu,N}\right\}_{i=1}^{\sigma}$ such that

- $\Sigma_{\nu,N} \;=\; \bigcup_{i=1}^{\sigma} \Sigma^i_{\nu,N}$;
- $p\left(\Sigma^i_{\nu,N}, x\right) \;>\; \frac{1}{2} - \rho$, for every $x \in \Sigma^i_{\nu,N}$.

Let us postpone the proof of this Claim for now, and see why it implies the theorem. By Theorem 2.2.1, if Claim 2.3 is true, it follows that, for every $i \in \{1, ..., \sigma\}$,

$$\dim \Sigma^i_{\nu,N} \;\leq\; d\left(\frac{1}{2} - \rho\right),$$

with the property that

$$\lim_{\rho \to 0} d\left(\frac{1}{2} - \rho\right) \;=\; n - 1.$$

This implies that, for $\varepsilon > 0$, there exists $\rho = \rho(\varepsilon) > 0$, sufficiently small, such that $\dim \Sigma^i_{\nu,N} \leq n - 1 + \varepsilon$, which clearly entails that

$$\dim \Sigma_{\nu,N} \;\leq\; n - 1 + \varepsilon.$$

Letting $\varepsilon \longrightarrow 0$, the Claim follows. The remaining steps are devoted to proving Claim 2.3. From now on, let $0 < \rho < \frac{1}{2}$ be fixed.

Step 3. The task of this step is to obtain a finite decomposition for the set $\Sigma_{\nu,N}$, proving the first part of Claim 2.3.

Our efforts will focused in building a δ-covering for the ball $\overline{B(0,N)}$, where the c-superdifferential lives. The question now is: how do we choose the radius $\delta > 0$, depending on ρ?

If we set $\eta = \frac{1}{2\nu}$, we can choose $\xi > 0$, dependent on η therefore on ν, such that inequality (2.2) holds. We set $\varepsilon = \frac{1}{2}\xi\rho$ and then select the relative $\delta = \delta(\rho,\nu) > 0$, for which Assumption 2.1.3 is satisfied. Without loss of generality, we assume that $\delta < \frac{1}{8\nu}$.

We consider a δ-covering $\mathcal{B} := \{B(q_i,\delta)\}_{i=1}^{M}$ of $\overline{B(0,N)}$. Clearly, we can assume that it is a finite family because $\overline{B(0,N)}$ is a compact set. We then define

$$\Sigma_{\nu,N}^{(j,l)} := \left\{ x \in \Sigma_{\nu,N} \,\middle|\, \partial^c u(x) \cap B(q_j,\delta) \neq \emptyset,\ \partial^c u(x) \cap B(q_l,\delta) \neq \emptyset,\ \mathrm{dist}(B(q_j,\delta), B(q_l,\delta)) > \frac{1}{4\nu} \right\}.$$

Let us prove that

$$\Sigma_{\nu,N} = \bigcup_{j,l} \Sigma_{\nu,N}^{(j,l)}. \tag{2.24}$$

If $x \in \Sigma_{\nu,N}$, then there exist p_1 and $p_2 \in \partial^c u(x)$ such that $|p_1 - p_2| > \frac{1}{2\nu}$. Since p_1 and $p_2 \in \partial^c u(x) \subset \overline{B(0,N)}$, there are $B(q_j,\delta)$ and $B(q_l,\delta) \in \mathcal{B}$ such that $p_1 \in B(q_j,\delta)$ and $p_2 \in B(q_l,\delta)$. Then $x \in \Sigma_{\nu,N}^{(j,l)}$, because

$$\mathrm{dist}\left(B(q_k,\delta), B(q_l,\delta)\right) \geq d(p_1,p_2) - 2\delta > \frac{1}{4\nu}.$$

Moreover, the union in (2.24) is finite; indeed the family $\mathcal{B}$ has a finite number of elements. This concludes the proof of the first part of Claim 2.3.

Step 4. Let us now consider the second part of Claim 2.3: the aim is to show that, for every $j,l \in \{1,...,M\}$ fixed,

$$p\left(\Sigma_{\nu,N}^{(j,l)}, x\right) \geq \frac{1}{2} - \rho, \tag{2.25}$$

for every $x \in \Sigma_{\nu,N}^{(j,l)}$.

In this step we show that for a point $x \in \Sigma_{\nu,N}^{(j,l)}$ there exists $r_0 = r_0(x) > 0$ such that for all $0 < r < r_0(x)$ we find a unit vector $w = w(x,r) \in \mathbb{S}^{n-1}$ such that if $x' \in \Sigma_{\nu,N}^{(j,l)} \cap B(x,r)$, then

$$\left\langle \frac{x'-x}{r}\,,\, w \right\rangle \,<\, \rho. \tag{2.26}$$

Let $x \in \Sigma_{\nu,N}^{(j,l)}$ be fixed. By definition of $\Sigma_{\nu,N}^{(j,l)}$, we can select $p_1 \in \partial^c u(x) \cap B(q_j,\delta)$ and $p_2 \in \partial^c u(x) \cap B(q_l,\delta)$. Saying that p_1 and $p_2 \in \partial^c u(x)$ means that the following two inequalities hold

$$c(x,p_1) - u(x) \,\le\, c(z,p_1) - u(z), \tag{2.27}$$

$$c(x,p_2) - u(x) \,\le\, c(z,p_2) - u(z), \tag{2.28}$$

for every $z \in \mathbb{R}^n$.

By our differentiability Assumption 2.1.1 we find a $r_0 = r_0(x) > 0$ such that for all $0 < r < r_0(x)$ there exists a point $\hat{x} \in B(x,r)$ that is a differentiability point for the function $x \to c(x,p)$ for any $p \in \mathbb{R}^n$. Without loss of generality we assume that for the point $\hat{x}$, the second option in our Continuity Assumption 2.1.2 (i.e. (2.4)) holds.

We set

$$w := \frac{\nabla_x c(\hat{x},p_2) - \nabla_x c(\hat{x},p_1)}{|\nabla_x c(\hat{x},p_2) - \nabla_x c(\hat{x},p_1)|}.$$

In the following we shall prove the estimate (2.26) for this choice of w. In order to do that, consider a point $x' \in \Sigma_{\nu,N}^{(j,l)} \cap B(x,r)$. Hence there exist $p_1' \in \partial^c u(x') \cap B(q_j,\delta)$ and $p_2' \in \partial^c u(x') \cap B(q_l,\delta)$. Therefore, again by the definition of c-superdifferential, one has

$$c(x',p_1') - u(x') \,\le\, c(z,p_1') - u(z), \tag{2.29}$$

$$c(x',p_2') - u(x') \,\le\, c(z,p_2') - u(z), \tag{2.30}$$

for every $z \in \mathbb{R}^n$.

We focus our attention to inequalities (2.27) and (2.29). In the first we set $z = x'$ and in the second $z = x$. Consequently, they read as follows

$$c(x,p_1) - c(x',p_1) \,\le\, u(x) - u(x'). \tag{2.31}$$

In a similar way, using inequalities (2.28) and (2.30) we obtain

$$c(x, p_2') - c(x', p_2') \ \geq \ u(x) - u(x'). \tag{2.32}$$

Inequalities (2.31) and (2.32) combined give

$$c(x, p_1) - c(x', p_1) \ \leq \ c(x, p_2') - c(x', p_2'). \tag{2.33}$$

Using the fact that $x', \hat{x} \in B(x, r)$ and thus $x, x' \in B(\hat{x}, 2r)$ we can apply Taylor's formula at the point $\hat{x}$. We have

$$c(x, p_1) \ = \ c\left(\hat{x}, p_1\right) \ + \ \left\langle \nabla_x c(\hat{x}, p_1)\, , \, x - \hat{x} \right\rangle \ + \ o(r), \tag{2.34}$$

$$c(x', p_1) \ = \ c\left(\hat{x}, p_1\right) \ + \ \left\langle \nabla_x c(\hat{x}, p_1)\, , \, x' - \hat{x} \right\rangle \ + \ o(r). \tag{2.35}$$

We subtract equation (2.34) to (2.35) in order to get

$$c(x, p_1) - c(x', p_1) \ = \ \left\langle \nabla_x c(\hat{x}, p_1)\, , \, x - x' \right\rangle \ + \ o(r). \tag{2.36}$$

With a similar argument, one has also

$$c(x, p_2') - c(x', p_2') \ = \ \left\langle \nabla_x c(\hat{x}, p_2')\, , \, x - x' \right\rangle \ + \ o(r). \tag{2.37}$$

Now, we combine (2.36) and (2.37) with inequality (2.33) and obtain

$$\left\langle \nabla_x c(\hat{x}, p_1)\, , \, x - x' \right\rangle \ + \ o(r) \ \leq \ \left\langle \nabla_x c(\hat{x}, p_2')\, , \, x - x' \right\rangle \ + \ o(r),$$

or, equivalently,

$$\langle \nabla_x c(\hat{x}, p_2') - \nabla_x c(\hat{x}, p_1), x' - x \rangle \leq o(r),$$

which can be written as

$$\langle \nabla_x c(\hat{x}, p_2) - \nabla_x c(\hat{x}, p_1), x' - x \rangle \leq \langle \nabla_x c(\hat{x}, p_2) - \nabla_x c(\hat{x}, p_2'), x' - x \rangle + o(r).$$

Dividing this relation by $|\nabla_x c(\hat{x}, p_2) - \nabla_x c(\hat{x}, p_1)| \cdot r$ yields:

$$\left\langle \frac{\nabla_x c(\hat{x}, p_2) - \nabla_x c(\hat{x}, p_1)}{|\nabla_x c(\hat{x}, p_2) - \nabla_x c(\hat{x}, p_1)|}, \frac{x' - x}{r} \right\rangle \leq \left\langle \frac{\nabla_x c(\hat{x}, p_2) - \nabla_x c(\hat{x}, p_2')}{|\nabla_x c(\hat{x}, p_2) - \nabla_x c(\hat{x}, p_1)|}, \frac{x' - x}{r} \right\rangle + o(1).$$

Note that $|\nabla_x c(\hat{x}, p_2) - \nabla_x c(\hat{x}, p_2')| \leq \epsilon$ and $|\nabla_x c(\hat{x}, p_2) - \nabla_x c(\hat{x}, p_1)| > \xi$ while $\frac{\epsilon}{\xi} < \frac{\varrho}{2}$ by our choices of parameters made at the beginning of Step 3.

This implies that the first term on the right side of the above estimate is less than $\frac{\varrho}{2}$ since the second term converges to 0 as $r \to 0$. We obtain that the second term is also less than $\frac{\varrho}{2}$ for $r > 0$ small enough. Consequently we obtain that

$$\left\langle \frac{\nabla_x c(\hat{x}, p_2) - \nabla_x c(\hat{x}, p_1)}{|\nabla_x c(\hat{x}, p_2) - \nabla_x c(\hat{x}, p_1)|}, \frac{x' - x}{r} \right\rangle = \left\langle w, \frac{x' - x}{r} \right\rangle \leq \rho,$$

for $r > 0$ small enough, as required.

□

As an application, we state a theorem about the existence of an optimal transport map. Thanks to our Theorem 2.3.1, we are able to weaken the assumptions on the cost function c.

Theorem 2.3.2. *Let us assume that $c : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ satisfies the Assumptions 2.1.1, 2.1.2, 2.1.3. Let μ and ν be two Borel regular probability measures such that μ does not give mass to $(n-1)$-dimensional sets. Then there is an optimal transport map $T : \mathbb{R}^n \to \mathbb{R}^n$ transporting μ to ν such that*

$$\int_{\mathbb{R}^n} c(x, T(x)) d\mu(x) \leq \int_{\mathbb{R}^n} c(x, \tilde{T}(x)) d\mu(x),$$

for any measurable map $\tilde{T} : \mathbb{R}^n \to \mathbb{R}^n$ such that $\tilde{T}_\sharp \mu = \nu$. Moreover, there exists a c-concave function φ such that $\mathrm{T}(x) = \partial^c \varphi(x)$, for μ-almost every $x \in \mathbb{R}^n$.

Proof. Based on the general theory of optimal mass transportation recalled in the first Chapter (see also [92, 93]), there exists an optimal transport plan π with marginals μ and ν, $\pi \in \Pi(\mu, \nu)$, supported on the graph of the c-superdifferential $\partial^c \varphi$ of some c-concave function φ. If the multivalued map $x \longmapsto \partial^c \varphi(x)$ is single valued for μ-almost every $x \in \mathbb{R}^n$, then this will give rise to an optimal transport map defined by $T(x) = \partial^c \varphi(x)$ for μ-almost every $x \in \mathbb{R}^n$.

According to the above consideration, we only need to check that $x \longmapsto \partial^c \varphi(x)$ is single valued for μ almost every x (Proposition 1.5.1). Clearly, $x \longmapsto \partial^c \varphi(x)$ is single valued outside the singular set $\Sigma^c(\varphi)$. By Theorem 2.3.1 $\dim \Sigma^c(\varphi) \leq n - 1$. By our assumption on the measure μ, we have $\mu(\Sigma^c(\varphi)) = 0$ and thus the claim of the theorem follows. □

Part II

Singular sets of H-convex functions

Chapter 3

An introduction to Carnot groups

In this chapter we will build some background on Carnot groups (see Definition 3.3.4), which are a special case of Lie groups (see Definition 3.1.1).

In the first section, we will introduce the definition of Lie group and its Lie algebra and we will state some basic facts.

In the second section, we provide a brief exposition of general features concerning the exponential map. We will conclude by recalling the celebrated Baker-Campbell-Hausdorff formula.

Section 3.3 is devoted to the study of Carnot groups. In particular, we will recall that every Carnot group $\mathbb{G}$ is diffeomorphic to some $\mathbb{R}^n$.

The chapter will end with a couple of sections devoted to the study of the Carnot-Carathéodory metric. First, we introduce this metric generally in $\mathbb{R}^n$, then we will specialize to the Carnot-Carathéodory metric on Carnot groups.

3.1 Lie Groups and Lie Algebras

We recall some notations and results about Lie groups and their Lie algebras. Let us start with the definition of Lie group (for a comprehensive treatment and for references to the extensive literature on the subject one may refer to the books [91] and [16]).

Definition 3.1.1. *A* Lie group *is a differentiable manifold* $\mathbb{G}$ *endowed with a differentiable group structure. This means that the product* $(x, y) \mapsto x \cdot y$ *and the inverse* $x \mapsto x^{-1}$ *are smooth maps.*

Definition 3.1.2. *A* Lie subgroup *of* $\mathbb{G}$ *is an embedded submanifold of* $\mathbb{G}$ *which is also a subgroup of* $\mathbb{G}$*.*

Definition 3.1.3. *Let* $\mathbb{G}$ *and* $\mathbb{H}$ *be Lie groups and let* $k \in \mathbb{N}$*. A* Lie homomorphism *from* $\mathbb{G}$ *to* $\mathbb{H}$ *is a* C^k*-map*

$$\varphi : \mathbb{G} \longrightarrow \mathbb{H}$$

that is also a group homomorphism.

Remark 3.1.4. A map $\varphi : \mathbb{G} \to \mathbb{H}$ is called a *Lie isomorphism* if it is a Lie homomorphism and also its inverse is a Lie homomorphism.

To understand the objects we are working with, let us treat a simple example of Lie groups on Euclidean spaces: the first Heisenberg group. We postpone the general case and main properties until Section 4.1 of the next chapter.

Example 3.1.1. We consider $\mathbb{R}^3$ identified to $\mathbb{C} \times \mathbb{R}$ and use the notation

$$p = (x, y, t) = (z, t) \in \mathbb{C} \times \mathbb{R}.$$

We give to $\mathbb{C} \times \mathbb{R}$ a Lie group structure with group law:

$$(z\,,\, t) \cdot (w\,,\, s) = (z + w\,,\, t + s + 2\Im(z \cdot \bar{w})\,)\,.$$

It is not difficult to check that the identity is 0 and that the inverse is given by $(z, t)^{-1} = (-z, -t)$. We call the Lie Group $\mathbb{H}^1 = (\mathbb{R}^3, \cdot)$ the *first Heisenberg Group.*

Definition 3.1.5. *Fixed* $g \in \mathbb{G}$*, we denote by*

$$\begin{aligned} \tau_g : \mathbb{G} &\longrightarrow \mathbb{G} \\ x &\longmapsto g \cdot x \end{aligned}$$

the left translation *by* g *on* $\mathbb{G}$*.*

Let us give the definition of smooth left invariant vector fields. Indeed, we aim to study the Lie algebra associated to a Lie group.

Definition 3.1.6. *A smooth vector field* $X \in \Gamma(T\mathbb{G})$ *is said to be* left invariant *if, for every* $g \in \mathbb{G}$,

$$d\tau_g X \;=\; X \circ \tau_g, \tag{3.1}$$

where $d\tau_g : T\mathbb{G} \to T\mathbb{G}$ *is the derivative map of the left translation* τ_g.

Since τ_g is a Lie isomorphism (a diffeomorphism more generally), notice that $d\tau_g X$ is well defined as vector field. The condition (3.1) is equivalent to the following one:

$$d_x\tau_g\left(X(x)\right) \;=\; X(g\cdot x),$$

for every $g, x \in \mathbb{G}$. If we apply the previous formula to the identity of $\mathbb{G}$, we obtain

$$d_e\tau_g\left(X(e)\right) \;=\; X(g),$$

for every $g \in \mathbb{G}$. Moreover, the condition of left invariance can be rewritten in this way

$$X\left(f\circ\tau_g\right)(x) \;=\; Xf\circ\tau_g(x),$$

for all $x, g \in \mathbb{G}$ and for all smooth function f on $\mathbb{G}$.

Definition 3.1.7. *Let* $\mathbb{G}$ *be a Lie group. We call the* Lie algebra *of* $\mathbb{G}$*, and write* $\mathfrak{g}$*, the set of all smooth left invariant vector fields on* $\mathbb{G}$.

Proposition 3.1.1. $\mathfrak{g}$ *is a* Lie algebra[1] *under the Lie Bracket product defined as*

$$[\,X, Y]\,f \;=\; XYf \;-\; YXf,$$

for all $X, Y \in \mathfrak{g}$*, and for all* $f \in C^k(\mathbb{G})$.

Remark 3.1.8. The dimension of $\mathfrak{g}$ as vector space equals that of $\mathbb{G}$. Indeed, $\mathfrak{g}$ is canonically isomorphic to $T_e\mathbb{G}$ via the identification of X and $X(e)$.

[1] We recall that a vector space $\mathfrak{g}$ is a *Lie algebra* if there is a bilinear and antisymmetric map $[\cdot,\cdot] : \mathfrak{g}\times\mathfrak{g} \to \mathfrak{g}$ which satisfies the Jacobi identity

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0,$$

for all $X, Y, Z \in \mathfrak{g}$.

Example 3.1.2. Return to the first Heisenberg group $\mathbb{H}^1$. It is not difficult to show that the vector fields

$$X_1 = \partial_x + 2y\,\partial_t$$
$$X_2 = \partial_y - 2x\,\partial_t$$

are left invariant with respect to the group law.

3.2 The Exponential Map

Given a Lie group $\mathbb{G}$, we defined its Lie algebra. The exponential map gives rise to a canonical way to associate each element of $\mathfrak{g}$ to a point of $\mathbb{G}$. In this section we recall the definition and the main properties. We start with a proposition (for more details we refer the reader to [16]):

Proposition 3.2.1. *The left invariant vector fields on a Lie group $\mathbb{G}$ are complete.*

Given $g \in \mathbb{G}$ and $X \in \mathfrak{g}$, let us consider the solution of the following Cauchy problem:

$$\begin{cases} \dot{\gamma}_g(t) = X(\gamma_g(t)) \\ \gamma_g(0) = g. \end{cases}$$

Remark 3.2.1. Notice that, by Proposition 3.2.1, the integral curve γ_g is defined for each $t \in \mathbb{R}$.

In the following, we set

$$\exp_X(t) := \gamma_e(t).$$

Thanks to this notation, we can construct a canonical map which, with each vector field in $\mathfrak{g}$, associates a point of $\mathbb{G}$. We consider once again the integral curve of a fixed left invariant vector field X, we stop at time $t = 1$, that point will be the element of $\mathbb{G}$ associated with X:

Definition 3.2.2. *Let $\mathbb{G}$ be a Lie group with Lie algebra $\mathfrak{g}$, we set*

$$\begin{aligned} \exp : \mathfrak{g} &\longrightarrow \mathbb{G} \\ X &\longmapsto \exp(X) := \exp_X(1). \end{aligned}$$

This map is called exponential map related to the Lie group $\mathbb{G}$.

In the following proposition, we summarize some of the main properties of the exponential map and of integral curves more generally.

Proposition 3.2.2. *Let $\mathbb{G}$ be a Lie group and $\mathfrak{g}$ its Lie algebra. Then*

(i) The exponential map is a smooth map;

(ii) For every $X \in \mathfrak{g}$ and for any $t, s \in \mathbb{R}$, $\exp_X(t+s) = \exp_X(t)\exp_X(s)$;

(iii) The derivative map of the exponential map $d\exp : T_0\mathfrak{g} \to T_e\mathbb{G}$ is the identity map, under the canonical identification of both $T_0\mathfrak{g}$ and $T_e\mathbb{G}$ with $\mathfrak{g}$;

(iv) The exponential map is a local diffeomorphism from some neighborhood of 0 in $\mathfrak{g}$ to a neighborhood of e in $\mathbb{G}$.

Proposition 3.2.3. *If $\mathbb{G}$ is a nilpotent,*[2] *simply connected Lie group, then the exponential map is a global diffeomorphism of $\mathfrak{g}$ onto $\mathbb{G}$. Moreover, if $\mathbb{H}$ is a Lie subgroup of $\mathbb{G}$, and $\mathfrak{h}$ is its Lie algebra, then $\mathbb{H} = \exp\mathfrak{h}$.*

Consider now two vector fields $X, Y \in \mathfrak{g}$, we aim to reconstruct the group law of the Lie group associated with $\mathfrak{g}$; we define $C(X,Y) \in \mathfrak{g}$ setting

$$\exp(C(X,Y)) = \exp(X) \cdot \exp(Y).$$

It is possible to compute explicitly $C(X,Y)$. We start with some notations: let $\alpha = (\alpha_1, ..., \alpha_l)$ be a multiindex of non-negative integers, we define

$$|\alpha| := \alpha_1 + ... + \alpha_l$$
$$\alpha! := \alpha_1! \cdot ... \cdot \alpha_l!,$$

we say that l is the *length of the multiindex* α. Let $\beta = (\beta_1, ..., \beta_l)$ another multiindex, with the same legth as α, such that $\alpha_l + \beta_l \geq 1$. We set

$$C_{\alpha,\beta}(X,Y) := \begin{cases} (adX)^{\alpha_1}(adY)^{\beta_1} \cdot ... \cdot (adX)^{\alpha_l}(adY)^{\beta_l - 1}Y, & \text{if } \beta_l > 0 \\ (adX)^{\alpha_1}(adY)^{\beta_1} \cdot ... \cdot (adX)^{\alpha_l - 1}X, & \text{if } \beta_l = 0, \end{cases}$$

[2] A Lie group is *nilpotent of step r* if its Lie algebra is nilpotent of step r, that is, defined the *descending central serie* of $\mathfrak{g}$, $\mathfrak{g}^{(1)} = \mathfrak{g}$ and $\mathfrak{g}^{(k+1)} = [\mathfrak{g}^{(k)}, \mathfrak{g}]$, for $k > 1$, there exists $r \in \mathbb{N}$ such that $\mathfrak{g}^{(r+1)} = 0$ and $\mathfrak{g}^{(k)} \neq \{0\}$ if $k \leq r$.

where $(adX)(Y) := [X, Y]$. Then the *Baker-Campbell-Hausdorff formula* states that

$$C(X,Y) := \sum_{l=1}^{\infty} \frac{(-1)^{l+1}}{l} \sum_{\substack{\alpha,\beta \\ \alpha_l+\beta_l \geq 1}} \frac{1}{\alpha!\beta!(\alpha+\beta)} C_{\alpha,\beta}(X,Y), \tag{3.2}$$

whenever the serie at the right hand side makes sense. Moreover, it is clear that 3.2 holds in nilpotent Lie groups.

3.3 Carnot Groups

In this section we will approach to the setting of our studies. As already mentioned, we are interested in Heisenberg groups, which form a particular family of Carnot groups. Therefore, we need a little background about them (for more details we refer, once more, the reader to [16]). We start with a definition:

Definition 3.3.1. *A Lie algebra* $\mathfrak{g}$ *is called* stratified *if it admits a* stratification*, i.e. there exists* $V_1, ..., V_r \subset \mathfrak{g}$ *subspaces such that*

$$\mathfrak{g} = V_1 \oplus ... \oplus V_r,$$

where

$$V_j = [V_1, V_{j-1}], \text{ for } j = 2, ..., r$$
$$[V_1, V_r] = \{0\}.$$

Remark 3.3.2. It is clear that V_r is contained in the center of $\mathfrak{g}$. We point out also that V_1 generates the whole Lie algebra. Because of its major role, we will call it *horizontal layer.*

Definition 3.3.3. *A group* $\mathbb{G}$ *is called* stratified *if its Lie algebra* $\mathfrak{g}$ *admits a stratification. Moreover, if the dimension of* $\mathbb{G}$ *is finite, then it is nilpotent of step* r*, exactly the number of subspaces in the stratification of* $\mathfrak{g}$*.*

From the definition of stratified Lie algebra, we can construct on $\mathfrak{g}$ a one parameter group of Lie homomorphisms, called *dilations* and denoted by $\{\delta_\lambda\}_{\lambda \geq 0}$. We fix $\lambda \geq 0$ and define, for $X \in V_j$:

$$\delta_\lambda X = \lambda^j X,$$

and then we extend this map over the entire $\mathfrak{g}$. Moreover, if $\lambda < 0$, we set

$$\delta_\lambda X \;=\; -\delta_{|\lambda|} X.$$

Proposition 3.3.1. *The following properties hold*

(i) $\delta_{\lambda\mu} = \delta_\lambda \circ \delta_\mu$*;*

(ii) $\delta_\lambda ([X, Y]) = [\delta_\lambda X, \delta_\mu Y]$*;*

(iii) $\delta_\lambda (C(X, Y)) = C (\delta_\lambda X, \delta_\mu Y)$*,*

for any λ, μ *and for any* $X, Y \in \mathfrak{g}$*.*

Definition 3.3.4. *A* Carnot group $\mathbb{G}$ *is a finite dimensional, connected, simply connected Lie group, whose Lie algebra admits a stratification. If* r *is the step of the stratification, we say that* $\mathbb{G}$ *is of* step r*.*

Remark 3.3.5. We should stress that a Carnot group can admit more than one stratification. For example, consider again the first Heisenberg group $\mathbb{H}^1$. Its Lie algebra $\mathfrak{h}$ admits the following stratifications:

$$\begin{aligned}
&\mathrm{span}\{X_1, X_2\} \oplus \mathrm{span}\{[X_1, X_2]\}\\
&\mathrm{span}\{X_1 - 3[X_1, X_2], X_2\} \oplus \mathrm{span}\{[X_1, X_2]\}\\
&\mathrm{span}\{X_1 + X_2, 3X_1 + [X_1, X_2]\} \oplus \mathrm{span}\{[X_1, X_2]\}.
\end{aligned}$$

Definition 3.3.6. *Let* $\mathbb{G}$ *be a Carnot group with Lie algebra* $\mathfrak{g}$*. Let* $\mathcal{V} = (V_1, ..., V_r)$ *be a fixed stratification of* $\mathfrak{g}$*. We say that a basis* $\mathcal{B}$ *of* $\mathfrak{g}$ *is* adapted to $\mathcal{V}$ *if*

$$\mathcal{B} \;=\; \left(E_1^{(1)}, ..., E_{m_1}^{(1)}; ....; E_1^{(r)}, ..., E_{m_r}^{(r)}\right),$$

where, for $i = 1, ..., r$*, we have* $m_i := \dim (V_i)$ *and* $\left(E_1^{(i)}, ..., E_{m_i}^{(i)}\right)$ *is a basis for* V_i*.*

Notation 3.3.1. We say that $\mathbb{G}$ has m generators, where $m := \dim (V_1)$.

In Remark 3.3.5, we saw that a Lie algebra of a Carnot group could admit more than one stratification. In the following proposition we point out that the main algebraic aspects of a Carnot group do not depend on the choice of the stratification:

Proposition 3.3.2. *Let $\mathbb{G}$ be a Carnot group and $\mathfrak{g}$ its Lie algebra. Let $(V_1, ..., V_r)$ and $(W_1, ..., W_r)$ be two stratifications of $\mathfrak{g}$. Then $r = s$ and $\dim V_i = \dim W_i$ for every $i = 1, ..., r$. Moreover, the real number*

$$Q := \sum_{i=1}^{r} i \cdot \dim V_i$$

depends only on the stratified nature of $\mathbb{G}$ and not on the particular stratification. Q is called homogeneous dimension of $\mathbb{G}$.

We conclude the section introducing on Carnot groups the so-called *exponential coordinates*. Let $(X_1, ..., X_n)$ a basis for the Lie algebra of $\mathbb{G}$, $\mathfrak{g}$. As usual, for general manifolds, in particular for Lie groups, we can write uniquely two vector fields in coordinates, setting $X = \sum_{i=1}^{n} x_i X_i$ and $Y = \sum_{i=1}^{n} y_i X_i$.[3] This fact permits us to give the following

Definition 3.3.7. *A system of* exponential coordinates *associated with $X_1, ..., X_n$ is the map*

$$\begin{aligned} \Psi : \mathbb{R}^n &\longrightarrow \mathbb{G} \\ (x_1, ..., x_n) &\mapsto \exp\left(\sum_{i=1}^{n} x_i X_i\right). \end{aligned} \tag{3.3}$$

We endow $\mathbb{R}^n$ with a group law, so that Ψ is a group isomorphism, that means $x \cdot y = z$ if and only if, using (3.2),

$$\sum_{i=1}^{n} z_i X_i = C\left(\sum_{i=1}^{n} x_i X_i, \sum_{i=1}^{n} y_i X_i\right).$$

With this group law, $\mathbb{R}^n$ is a Lie group whose Lie algebra is isomorphic to $\mathfrak{g}$. Now, $\mathbb{G}$ and $\mathbb{R}^n$ are both nilpotent, connected and simply connected, so, by Proposition 3.2.3, Ψ is also a diffeomorphism. From now on, we identify abstract Carnot groups with Carnot groups on $\mathbb{R}^n$. We will refer to coordinates (3.3) as *graded exponential coordinates.*

[3]The reader should keep in mind that Carnot Groups are connected, simply connected and nilpotent. Then the exponential map, being a global diffeomorphism (Proposition 3.2.3), provides a global chart for the manifold.

As reminded before, the exponential map is a global diffeomorphism, then also its inverse in well defined. This allows us to introduce a one parameter group of automorphisms on $\mathbb{G}$. Using for simplicity the same notation of the algebras case, we define the *dilations* on $\mathbb{G}$, and write $\{\delta_\lambda\}$, as follows

$$\delta_\lambda(x) \;:=\; \exp\left(\delta_\lambda(\exp^{-1}(x))\right),$$

for every $x \in \mathbb{G}$. With the same notation as in 3.3.6, if i is an index such that

$$m_1 + ... + m_{d_{i-1}} \;<\; i \;\leq\; m_1 + ... + m_{d_i},$$

for some $1 \leq d_i < k$, the coordinate x_i of $x = (x_1, ..., x_n) \in \mathbb{G}$ is said to have degree d_i. With this definition, group dilations $\delta_\lambda : \mathbb{G} \to \mathbb{G}$ can be written as

$$\delta_\lambda(x) \;=\; \left(\lambda^{d_1} x_1, \lambda^{d_2} x_2, ..., \lambda^{d_n} x_n\right).$$

Using Proposition 3.3.1, one can prove the following properties:

(i) $\delta_{\lambda\mu} \;=\; \delta_\lambda \cdot \delta_\mu$;

(ii) $\delta_\lambda(x\,y) = \delta_\lambda(x) \cdot \delta_\mu(y)$.

Using the notions introduced in Section 3.1, since the exponential map is a diffeomorphism from $\mathfrak{g}$ and $\mathbb{G}$, it follows, for each $x, y \in \mathbb{G}$,

$$x \cdot y \;=\; \exp\left(C(X,Y)\right) := P(x,y),$$

where X and $Y \in \mathfrak{g}$ are such that $\exp(X) = x$ and $\exp(Y) = y$. From this formula, one can prove some facts about the group law:

Proposition 3.3.3. *There exists a polynomial vector function*

$$Q \;:\; \mathbb{G} \times \mathbb{G} \;\longrightarrow\; \mathbb{R}^n = \mathbb{R}^{m_1} \oplus ... \oplus \mathbb{R}^{m_r},$$

where $Q(x,y) = (Q_1(x,y), ..., Q_r(x,y))$, *and* $Q_i(x,y) = \left(Q_1^{(i)}(x,y), ..., Q_{m_i}^{(i)}(x,y)\right)$, *for all* $i = 1, ..., r$, *such that*

$$x \cdot y \;=\; x + y + Q(x,y) \;=\; P(x,y).$$

Lemma 3.3.1. *The following properties hold:*

(i) for all $x, y \in \mathbb{R}^n$, $\lambda > 0$, $P(\delta_\lambda(x), \delta_\lambda(y)) = \lambda(P(x, y))$;

(ii) for all $x \in \mathbb{R}^n$, $P(x, 0) = 0$;

(iii) for all $x, y \in \mathbb{G}$, $Q_j(x, y) = 0$, *for* $j = 1, ..., m_1$, *and* $Q_j(x, 0) = Q_j(0, x) = Q_j(x, x) = Q_j(x, x^{-1}) = 0$, *for* $j \geq m_1 + 1$.

3.4 Carnot-Carathéodory Metric

We give to $\mathbb{R}^n$ the so-called *Carnot-Carathéodory metric*, induced by a family of vector fields which satisfies certain conditions. For a deeper investigation of these facts, the reader could refer herself or himself to [77].

After doing that, in Section 3.5, we will discover that, because of their peculiarities, Carnot groups can be naturally equipped with a Carnot-Carathéodory metric.

Let us consider a family of locally Lipschitz continuous vector fields on an open set $\Omega \subseteq \mathbb{R}^n$

$$X_j(x) = \sum_{i=1}^{n} a_{ij}(x)\partial_i, \quad j = 1, ..., m.$$

As usual, we call *horizontal fiber at the point* x, and write $H_x\mathbb{R}^n$, the subspace of $T_x\mathbb{R}^n$ generated by $X_1(x), ..., X_m(x)$. $H\mathbb{R}^n$ will be the *horizontal subbundle* of $T\mathbb{R}^n$.

Notation 3.4.1. We denote by

$$\mathcal{A} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

the matrix whose columns are the coefficients of the vector fields $\mathbb{X} := (X_1, ..., X_m)$.

Definition 3.4.1. *We say that a Lipschitz continuous curve* $\gamma : [0, T] \to \Omega$ *is* $\mathbb{X}$-admissible *if there exists a measurable vector function* $h = (h_1, ..., h_m) : [0, T] \to \Omega$ *such that* *(i)* $\dot{\gamma}(t) = \sum_{j=1}^{m} h_j(t) X_j(\gamma(t))$, *for a.e.* $t \in [0,T]$*;*

(ii) $|h| \in L^{\infty}([0,T])$.

The curve γ *is called* $\mathbb{X}$-subunit *if it is* $\mathbb{X}$*-admissible and* $\|h\|_{L^\infty} \leq 1$*, for a.e.* $t \in [0,T]$.

Definition 3.4.2. *We define* $d_c : \Omega \times \Omega \longrightarrow [0,\infty]$ *as follows*

$$d_c(x,y) \;=\; \inf\{T \geq 0 \,|\, \exists\, \gamma : [0,T] \to \Omega \text{ subunit curve } : \; \gamma(0) = x, \, \gamma(T) = y\}.$$

If there exists no $\mathbb{X}$*-subunit curve in* Ω *which joins* x *to* y*, then we write* $d_c(x,y) = \infty$.

Definition 3.4.3. *We say that* $\Omega \subseteq \mathbb{R}^n$ *is* $\mathbb{X}$-connected *if for all* $x, y \in \Omega$*, there is a* $\mathbb{X}$*-subunit curve joining* x *to* y.

Theorem 3.4.1. *If* $d_c(x,y) < \infty$ *for all* $x, y \in \Omega$*, then* (Ω, d) *is a metric space. We call* d_c *the* Carnot-Carathéodory metric *on* Ω *(CC-metric for short).*

Notation 3.4.2. We can define, as usual, the metric balls with respect to the Carnot-Carathéodory metric setting, for $r > 0$,

$$B(x,r) \;:=\; \{y \in \mathbb{R}^n \,|\, d_c(x,y) < r\}.$$

We point out that the metric d is finite on $\mathbb{R}^n$, and in general we can assume that the identity map between $(\mathbb{R}^n, d)$ and $(\mathbb{R}^n, |\cdot|)$ is a homeomorphism. This condition is satisfied when d is the CC-metric associated with a family of smooth vector fields $X_1, ..., X_m$ which satisfy the *Hörmander condition*:

$$\text{rank}\,(\mathcal{L}(X_1, ..., X_m))\,(x) \;=\; n \tag{3.4}$$

for all $x \in \mathbb{R}^n$. With $\mathcal{L}(X_1, ..., X_m)$ we denote the Lie algebra generated by $X_1, ..., X_m$. Geometrically, condition (3.4) means that the vector fields $X_1, ..., X_m$ and their iterated brackets generate the whole tangent space at every point.

3.5 CC-Metric on Carnot Groups

Let us consider a Carnot group $\mathbb{G}$ with Lie algebra $\mathfrak{g}$. We know that we can represent $\mathbb{G}$ by $\mathbb{R}^n$, endowed with a Carnot structure, through a system of exponential coordinates, associated with a basis adapted to a stratification of $\mathfrak{g}$. Using the same notations as above, let $\mathfrak{g} = V_1 \oplus ... \oplus V_r$, $r \geq 2$, $m = \dim V_1$ and fix a basis $\mathbb{X} = (X_1, ..., X_m)$ of V_1. From the definition of stratified algebra, V_1 generates the whole $\mathfrak{g}$ as an algebra. Hence, $X_1, ..., X_m$ satisfy Hörmander's condition, inducing a Carnot-Carathéodory metric on $\mathbb{G}$.

Proposition 3.5.1. *For all $x, y, z \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}^+$, the following properties hold:*

(i) $d_c(\tau_z(x), \tau_z(y)) = d_c(x, y)$;

(ii) $d_c(\delta_\lambda(x), \delta_\lambda(y)) = \lambda d_c(x, y)$.

Let us recall some remarks about measures and metrics. If we denote $\mathcal{H}^k_{d_c}$ and $\mathcal{S}^k_{d_c}$ the k-dimensional Hausdorff and Spherical Hausdorff measures associated with the Carnot-Caratéodory metric d_c, then one can check that

(i) $\mathcal{H}^k_{d_c}(x \cdot E) = \mathcal{H}^k_{d_c}(E)$,

(ii) $\mathcal{H}^k_{d_c}(\delta_\lambda E) = \lambda^k \mathcal{H}^k_{d_c}(E)$,

for every Lebesgue measurable set $E \subset \mathbb{R}^n$ and for all $x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}^+$. The same formulae hold for $\mathcal{S}^k_{d_c}$. The homogeneous dimension Q of $(\mathbb{R}^n, \cdot)$ is the Hausdorff dimension of $\mathbb{R}^n$ with respect to the CC-distance.

Moreover, we recall that the n-dimensional Lebesgue measure $\mathcal{L}^n$ is the Haar measure of the group. Therefore, the translation and dilation conditions read as follows:

Proposition 3.5.2. *Let $E \subset \mathbb{R}^n$ be a Lebesgue measurable set. Then, for all $x \in \mathbb{R}^n$ and $\lambda \geq 0$,*

(i) $\mathcal{L}^n(x \cdot E) = \mathcal{L}^n(E \cdot x) = \mathcal{L}^n(E)$;

(ii) $\mathcal{L}^n(\delta_\lambda E) = \lambda^Q \mathcal{L}^n(E)$.

In particular $\mathcal{L}^n(B(x,r)) = r^Q \mathcal{L}^n(B(x,1))$.

We conclude this short digression about Carnot groups introducing on them a *homogeneous distance*, i.e. a metric which "conserve" the intrinsic aspects of the groups: invariances by left translations and by dilations.

Definition 3.5.1. *We say that a metric* ρ *on* $\mathbb{G}$ *is a* homogeneous distance *if, for all* $x, y, z \in \mathbb{G}$ *and* $\lambda \in \mathbb{R}^+$

(i) $\rho(x,y) = \rho(\tau_z(x), \tau_z(y))$*;*

(ii) $\rho(\delta_\lambda(x), \delta_\lambda(y)) = \lambda\rho(x,y)$*.*

Notice that Proposition 3.5.1 says that the Carnot-Carathéodory metric is a homogeneous metric. We can construct other examples of homogeneous metrics. We start by defining the following quasi-metric

$$d_\infty(x,y) \;=\; \|y^{-1} \cdot x\|, \tag{3.5}$$

where $\|\cdot\|$ is a homogeneous norm. For example we can choose

$$\|x\|_\infty \;=\; \sum_{i=1}^{n} |x_i|^{\frac{1}{d_i}}$$

or

$$\|x\| \;=\; \max_i \left\{ \varepsilon_i |x_i|^{\frac{1}{d_i}} \right\},$$

where the ε_i's are suitable positive constants which depend on the group structure and which let d_∞ be a distance on the group.

Remark 3.5.2. We point out that it is always possible to find the right ε_i's such that d_∞ satisfies the triangle inequality (see [24]).

Remark 3.5.3. We notice that these homogeneous metrics induce the same topology as the Carnot-Carathéodory one.

Chapter 4

The Heisenberg group and its subgroups

This chapter is devoted to the study of the Heisenberg group $\mathbb{H}^n$, the main setting of our future investigations.

In the first section, keeping in mind the background built in the previous chapter about Carnot groups, we recall the main algebraic and geometric properties of $\mathbb{H}^n$. In the following sections, our goals is to give a sufficient understanding about homogeneous subgroups of $\mathbb{H}^n$. We first introduce a metric structure on the set of some special subgroups and, then, we conclude the chapter giving a look at the notion of intrinsic Lipschitz functions within the Heisenberg group and their major properties.

4.1 Heisenberg Groups

In this section we study some peculiarities of the Heisenberg group, which is the most simple non Abelian Carnot group and the setting of our future investigations. We start by recalling the definition and some general properties.

Notation 4.1.1. We denote by $p = (z, t)$ a point in $\mathbb{C}^n \times \mathbb{R}$, where $z = (z_1, ..., z_n) \in \mathbb{C}^n$ and $t \in \mathbb{R}$. If $z_j = x_j + iy_j$, we write $z = (x_1, ..., x_n, y_1, ..., y_n)$, with $x_j, y_j \in \mathbb{R}$, for $j = 1, ..., n$.

Let us consider in $\mathbb{C}^n \times \mathbb{R}$ the following composition law

$$(z,t)\cdot(z',t') \;=\; \left(z+z', t+t'+2\,\Im(z\cdot\bar{z}')\right), \tag{4.1}$$

where $z \cdot \bar{z}'$ denotes the usual Hermitian product in $\mathbb{C}^n$:

$$z\cdot\bar{z}' = \sum_{j=1}^{n}(x_j+iy_j)(x_j'-iy_j').$$

Remark 4.1.1. If we identify $\mathbb{C}^n$ with $\mathbb{R}^{2n}$, we can rewrite the operation law (4.1) in the following way:

$$(z,t)\cdot(z',t') \;=\; \left(z+z', t+t'+2\,\langle Jz, z'\rangle\right),$$

where J is the unit $(2n\times 2n)$-symplectic matrix and $\langle\cdot,\cdot\rangle$ denotes the usual inner product in $\mathbb{R}^{2n}$.

It is not difficult to verify that $(\mathbb{R}^{2n+1},\cdot)$ is a Lie group, whose identity is the origin of $\mathbb{R}^{2n+1}$ and the inverse is $(z,t)^{-1} = (-z,-t)$. We call this Lie group the *n-th Heisenberg group*, and we write $\mathbb{H}^n := (\mathbb{R}^{2n+1},\cdot)$.

The Heisenberg group $\mathbb{H}^n$ is the Lie group associated with the $(2n+1)$-dimensional real Lie algebra $\mathfrak{h}^n$ generated by $\{X_1,...,X_n,Y_1,...,Y_n,T\}$, that satisfies the relations

$$[X_i,X_j] \;=\; 0,\;\; [Y_i,Y_j] \;=\; 0,\;\; [X_j,Y_j] \;=\; T,$$

for every $i,j=1,...,n$. By the Jacobi's identity, we get that $[X_i,T]=[Y_i,T]=0$, for each $i=1,...,n$. This means that $\mathfrak{h}^n$ is a nilpotent Lie algebra. It is also clear that its center is span$\{T\}$.

Let us denote

$$V_1 \;=\; \mathrm{span}\{X_1,...,X_n,Y_1,...,Y_n\} \;\text{ and }\; V_2 \;=\; \mathrm{span}\{T\}.$$

Then the Heisenberg algebra is stratified of step 2 with stratification

$$\mathfrak{h}^n \;=\; V_1\oplus V_2.$$ [1]

[1] Using exponential coordinates, one can prove that $\mathbb{H}^n$ is the unique simply connected, nilpotent Lie group associated with $\mathfrak{h}^n$.

Remark 4.1.2. By the structure of $\mathfrak{h}^n$, we can say that the center of the group $\mathbb{H}^n$ is

$$\mathbb{T} = \{(z,t) \in \mathbb{R}^{2n+1} \,|\, z = 0\},$$

and the homogeneous dilations are, for $\lambda \in \mathbb{R}^+$,

$$\begin{aligned} \delta_\lambda \,:\, \mathbb{R}^{2n+1} &\longrightarrow \mathbb{R}^{2n+1} \\ (z,t) &\longmapsto (\lambda z, \lambda^2 t). \end{aligned}$$

We can realize the Heisenberg Lie algebra $\mathfrak{h}^n$ as an algebra of left invariant differential operators on $\mathbb{R}^{2n+1}$. For example, one can set $T = 4\partial_t$ and, consequently,

$$X_j \;=\; \partial_{x_j} + 2\, y_j \partial_t, \quad Y_j \;=\; \partial_{y_j} - 2\, x_j \partial_t,$$

for $j = 1, ..., n$. With this identification between vector fields and first order differetial operators, $X_1, ..., X_n, Y_1, ..., Y_n$ generate a vector bundle on $\mathbb{H}^n$, called *horizontal bundle* $H\mathbb{H}^n$. The horizontal bundle is a subbundle of the tangent bundle $T\mathbb{H}^n$. By definition of vector bundle, we know that we can identify canonically each fiber of $H\mathbb{H}^n$ with a vector subspace of $\mathbb{R}^{2n+1}$, so each section φ of $H\mathbb{H}^n$ can be identified with a map $\varphi : \mathbb{H}^n \longrightarrow \mathbb{R}^{2n+1}$. We denote by H_p the fiber of $H\mathbb{H}^n$ at a point $p \in \mathbb{H}^n$. On $\mathbb{H}^n$ is defined a Sub-Riemannian structure: we can endow each fiber H_p with a scalar product, denoted by $\langle \cdot, \cdot \rangle_p$, and the associated norm $|\cdot|_p$ that make the basis of H_p, $X_1(p), ..., X_n(p), Y_1(p), ..., Y_n(p)$, orthonormal; in other words, if we consider $v = \sum_{i=1}^n (v_i X_i(p) + v_{n+i} Y_i(p))$ and $w = \sum_{i=1}^n (w_i X_i(p) + w_{n+i} Y_i(p))$ vectors of H_p, then $\langle v, w \rangle_p := \sum_{i=1}^n (v_i \cdot w_i + v_{n+i} \cdot w_{n+i})$ and $|v|_p^2 := \langle v, v \rangle_p$.

We introduce the *Korányi norm*: if $p = (z,t) \in \mathbb{H}^n$,

$$\|p\| \;=\; \sqrt[4]{\|z\|_{\mathbb{R}^{2n}}^4 \,+\, |t|^2}.$$

If it is not specified, through this thesis we will use this homogeneous norm. To verify that $d_K(x,y) = \|y^{-1} \cdot x\|$ is a metric, when $\| \cdot \|$ is the Korányi norm, one needs to prove the triangle inequality

$$d_K(x,y) \;\leq\; d_K(x,z) + d_K(z,y). \tag{4.2}$$

This can be done by a direct computation.

First, we can replace $z^{-1} \cdot x$ with x and $y^{-1} \cdot z$ with y; then it is sufficient to prove (4.2) in the case when $z = e$ and to show that

$$\|x \cdot y\| \ \le \ \|x\| + \|y\|.$$

Writing $x = (z, t)$ and $y = (w, s)$ and using the group law (4.1), we find

$$\begin{aligned} \|x \cdot y\|^4 \ &= \ \|v + w\|^4_{\mathbb{R}^{2n}} + \left(t + w + 2\,\Im(v \cdot \bar{w})\right)^2 \\ &= \ \left|\|v + w\|^2_{\mathbb{R}^{2n}} + 2i\,(t + w + 2\,\Im(v \cdot \bar{w}))\right|^2 \\ &= \ \left|\|v\|^2_{\mathbb{R}^{2n}} + 2it + \bar{v} \cdot w + \|w\|^2_{\mathbb{R}^{2n}} + 2is\right|^2 \\ &\le \ \left(\|x\|^2 + 2\|v\|^2_{\mathbb{R}^{2n}}\|w\|^2_{\mathbb{R}^{2n}} + \|y\|^2\right)^2 \\ &\le \ (\|x\| + \|y\|)^4 . \end{aligned}$$

We conclude this section with a small observation, very useful in the future. Thanks to the exponential map, we introduced in every Carnot group the so-called exponential coordinates. In the Heisenberg group, this implies that a vector field $\sum_{i=1}^n (x_i X_i + y_i Y_i) + tT$ is identified with the point $(x_1, ..., x_n, y_1, ..., y_n, t) \in \mathbb{H}^n$. On the other hand, we can consider the inverse of the exponential map $\xi : \mathbb{H}^n \longrightarrow \mathfrak{h}^n$. Since $\mathfrak{h}^n = V_1 \oplus V_2$, the map ξ has two components: $\xi_1 : \mathbb{H}^n \longrightarrow V_1$ and $\xi_2 : \mathbb{H}^n \longrightarrow V_2$. Since we can identify V_1 with the Euclidean space $\mathbb{R}^{2n}$, it holds that $\xi_1(p) = \xi_1(x, y, t) = (x, y) \in \mathbb{R}^{2n}$.

4.2 Decomposition in complementary subgroups

In this section we study the main properties of homogeneous subgroups (see Definition 4.2.1) of $\mathbb{H}^n$. In Section 4.4, we are interested in intrinsic Lipschitz graphs. Roughly speaking, we are interested in the graphs of some special functions whose graphs lie in the Heisenberg group with some Lipschitz-type property. We can compare this notion with the Euclidean case, in which we decompose $\mathbb{R}^n$ in the cartesian product of two subspaces. In this setting, it is very simple to imagine a map acting between these two subgroups. In $\mathbb{H}^n$, on the other side, we need more conditions for a “nice” decomposition. This notion of decomposition in homogeneous subgroups was first introduced in [44], in [46], in [48] and also in [12].

We start with the definition of homogeneous subgroup, which are subgroups invariant under group dilations.

Definition 4.2.1. *We say that a subgroup* $\mathbb{G}$ *of* $\mathbb{H}^n$ *is a* homogeneous Lie subgroup *if, for all* $g \in \mathbb{G}$ *and* $\lambda > 0$*,* $\delta_\lambda(g) \in \mathbb{G}$*.*

We point out that Definition 4.2.1 can be stated also for a general Carnot group of step k. In this case, one can prove that each homogeneous subgroup is necessarily a graded subgroup with step at most k, but in general it is not a Carnot group.

Definition 4.2.2. *We say that* $\mathbb{H}^n$ *is a* semidirect product *of homogeneous subgroups* $\mathbb{W}$ *and* $\mathbb{V}$*, and we write* $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$*, if* $\mathbb{W} = \exp(\mathfrak{w})$ *and* $\mathbb{V} = \exp(\mathfrak{v})$*, where* $\mathfrak{w}$ *and* $\mathfrak{v}$ *are homogeneous subalgebras of* $\mathfrak{h}^n$ *such that*

(i) $\mathfrak{h}^n = \mathfrak{w} \oplus \mathfrak{v}$*;*

(ii) $\mathfrak{w}$ *is an ideal*[2] *in* $\mathfrak{h}^n$*.*

We will say that $\mathbb{W}$ *and* $\mathbb{V}$ *are* complementary subgroups *in* $\mathbb{H}^n$*.*

Remark 4.2.3. If $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$, then $\mathbb{W} \cap \mathbb{V} = \{e\}$. From (ii), it follows also that $\mathbb{W}$ is a normal subgroup[3] of $\mathbb{H}^n$. It is a very classical fact (see [91]) that, in connected Lie groups, there is a bijective correspondance between normal subgroups and ideals of the Lie algebra.

Example 4.2.1. A simple example of a semidirect product is given by $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$, where

$$\mathbb{V} = \{(x_1, 0, ..., 0) \,|\, x_1 \in \mathbb{R}\}$$

[2] Let A be a Lie algebra. We say that a subset B of A is an *ideal* of A if it is a vector subspace of A and $[x, y] \in B$, for every $x \in B$ and $y \in A$.

[3] Let G be a group. We say that a subgroup H of G is *normal* if $g \cdot h \cdot g^{-1} \in H$, for every $g \in G$ and $h \in H$.

and

$$\mathbb{W} = \{(0, x_2, ..., x_n, y_1, ..., y_n, t) \,|\, x_i, y_j, t \in \mathbb{R},\ i = 2, ..., n, j = 1, ..., n\}.$$

Proposition 4.2.1. *Let $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$ be as in Definition 4.2.2. Then each $q \in \mathbb{H}^n$ has unique components $q_{\mathbb{W}} \in \mathbb{W}$ and $q_{\mathbb{V}} \in \mathbb{V}$ such that $q = q_{\mathbb{W}} \cdot q_{\mathbb{V}}$.*

Proof. Let us assume that $q \in \mathbb{H}^n$ admits two decompositions: $q = q_{\mathbb{W}} \cdot q_{\mathbb{V}}$ and $q = q'_{\mathbb{W}} \cdot q'_{\mathbb{V}}$. Then, since e is the unique common element of $\mathbb{W}$ and $\mathbb{V}$,

$$(q'_{\mathbb{W}})^{-1} \cdot q_{\mathbb{W}} \ = \ q'_{\mathbb{V}} \cdot (q_{\mathbb{V}})^{-1} \ = \ e.$$

Thus, $q'_{\mathbb{W}} = q_{\mathbb{W}}$ and $q'_{\mathbb{V}} = q_{\mathbb{V}}$. □

Proposition 4.2.2. *Let $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$ be a semidirect product as in Definition 4.2.2. Then the maps*

$$\begin{aligned} P_{\mathbb{W}} \,:\, \mathbb{H}^n &\longrightarrow \mathbb{W} \\ q &\longmapsto q_{\mathbb{W}} \end{aligned}$$

and

$$\begin{aligned} P_{\mathbb{V}} \,:\, \mathbb{H}^n &\longrightarrow \mathbb{V} \\ q &\longmapsto q_{\mathbb{V}} \end{aligned}$$

are continuous.

Proof. See [12], Proposition 3.4. □

Proposition 4.2.3. *All homogeneous subgroups of $\mathbb{H}^n$ are either horizontal, that are contained in $\{(z, t) \in \mathbb{H}^n \,|\, t = 0\}$ or vertical, that are containing the subgroup $\mathbb{T}$.*

A horizontal subgroup $\mathbb{V}$ has linear dimension k, with $1 \leq k \leq n$; moreover, $\mathbb{V}$ is algebraically isomorphic and isometric to $\mathbb{R}^k$.

A vertical subgroup $\mathbb{W}$ can have any algebraic dimension d, with $1 \leq d \leq 2n+1$, and its metric dimension is $d + 1$.

From the previous Proposition, we have directly:

Proposition 4.2.4. *All possible pairs* $\mathbb{W}$ *and* $\mathbb{V}$ *of complementary subgroups of* $\mathbb{H}^n$ *are of the type*

(i) $\mathbb{V}$ *horizontal of linear dimension* k, $1 \leq k \leq n$,

(ii) $\mathbb{W}$ *normal of metric dimension* $2n + 2 - k$.

Proof. See [12], Proposition 3.21. □

Notation 4.2.1. Let $\mathbb{G}$ be a homogeneous subgroup of $\mathbb{H}^n$. We use the symbol $\dim \mathbb{G}$ to indicate the linear dimension of the subgroup $\mathbb{G}$, i.e. the dimension of its Lie algebra, and the symbol $\dim_H \mathbb{G}$ for its Hausdorff dimension, or metric dimension.

Remark 4.2.4. We aim to highlight that any horizontal subgroup $\mathbb{V}$ has a complementary normal subgroup $\mathbb{W}$; also the converse is true for normal subgroups $\mathbb{W}$ with linear dimension larger than n. On the contrary, normal subgroups of dimension less than or equal to n do not have complementary subgroups. For example the center $\mathbb{T}$ does not have a complementary subgroup.

We conclude this digression about homogeneous subgroups of $\mathbb{H}^n$ with a couple of results, which will be used in a proof concerning the Hausdorff dimension of the graph of an intrinsic Lipschitz function (Proposition 4.4.7). First a remark.

Remark 4.2.5. A homogeneous subgroup of $\mathbb{H}^n$, $\mathbb{G}$, can be endowed with a norm, which is the restriction to $\mathbb{G}$ of the Heisenberg norm.

Proposition 4.2.5. *Let* $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$ *be as in Proposition 4.2.4. Then there exists a positive constant* $C = C(\mathbb{W}, \mathbb{V})$ *such that*

$$C\ (\|q_{\mathbb{V}}\| + \|q_{\mathbb{W}}\|) \leq \|q\| \leq (\|q_{\mathbb{V}}\| + \|q_{\mathbb{W}}\|)\,. \tag{4.3}$$

Moreover, $(q^{-1})_{\mathbb{V}} = (q_{\mathbb{V}})^{-1}$, $(q^{-1})_{\mathbb{W}} = q_{\mathbb{V}}^{-1} \cdot (q_{\mathbb{W}})^{-1} \cdot q_{\mathbb{V}}$, $(p \cdot q)_{\mathbb{V}} = p_{\mathbb{V}} \cdot q_{\mathbb{V}}$, *and* $(p \cdot q)_{\mathbb{W}} = p_{\mathbb{W}} \cdot p_{\mathbb{V}} \cdot q_{\mathbb{W}} \cdot p_{\mathbb{V}}^{-1}$.

Proof. See [48], Proposition 3.2. □

Remark 4.2.6. By the previous Proposition, it follows that $P_{\mathbb{V}}$ is a group homomorphism from $\mathbb{H}^n$ to $\mathbb{V}$; while, in general, $P_{\mathbb{W}}$ is not a group homomorphism from $\mathbb{H}^n$ to $\mathbb{W}$. Moreover, one can notice that $P_{\mathbb{V}} : \mathbb{H}^n \longrightarrow \mathbb{V}$ is a Lipschitz map. Indeed, let $p = (z, t)$ and $q = (z', t') \in \mathbb{H}^n$. Since $\mathbb{V}$ is isometric to $\mathbb{R}^k$,

$$\|P_{\mathbb{V}}(p)^{-1} \cdot P_{\mathbb{V}}(q)\| \;=\; \|P_{\mathbb{V}}(q) - P_{\mathbb{V}}(p)\|.$$

Then it follows

$$\|P_{\mathbb{V}}(q) - P_{\mathbb{V}}(p)\| \;\leq\; \|z' - z\|_{\mathbb{R}^{2n}} \;\leq\; \|p^{-1} \cdot q\|.$$

On the contrary, $P_{\mathbb{W}} : \mathbb{H}^n \longrightarrow \mathbb{W}$, in general, is not a Lipschitz map. For example, consider the first Heisenberg group $\mathbb{H}^1$ decomposed in complementary subgroups as $\mathbb{H}^1 = \mathbb{W} \cdot \mathbb{V}$, where $\mathbb{W} = \{(0, y, t) \,|\, y, t \in \mathbb{R}\}$ and $\mathbb{V} = \{(x, 0, 0) \,|\, x \in \mathbb{R}\}$. Now, for $\varepsilon \in \mathbb{R}^+$, let $p = (1, 0, 0)$ and $q = \left(0, \varepsilon, \frac{\varepsilon}{2}\right)$. Then $p_{\mathbb{W}} = (0, 0, 0)$ and $q_{\mathbb{W}} = (0, \varepsilon, \frac{\varepsilon}{2})$. Thus,

$$\|P_{\mathbb{W}}(q)^{-1} \cdot P_{\mathbb{W}}(p)\| \;=\; \|q_{\mathbb{W}}\| \;\approx\; \varepsilon^{\frac{1}{2}}$$
$$\|p^{-1} \cdot q\| \;=\; \|(0, \varepsilon, 0)\| \;\approx\; \varepsilon$$

Proposition 4.2.6. *Let $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$ be as in Proposition 4.2.4 with $\dim(\mathfrak{v}) = k$, and let $p \in \mathbb{H}^n$ be fixed. Then there exists a positive constant $C = C(\mathbb{W}, \mathbb{V})$ such that, for $B(p, r) \subset \mathbb{H}^n$,*

$$\mathcal{L}^{2n+1-k}\left(P_{\mathbb{W}}(B(p, r))\right) \;=\; C(\mathbb{W}, \mathbb{V})\, r^{2n+2-k}. \tag{4.4}$$

Proof. See [48], Lemma 4.3.

4.3 The intrinsic Grassmannian

With the notion of decomposition in homogeneous subgroups of $\mathbb{H}^n$, we can make a step further. Following [69], we endow the set of all vertical subgroups, which give rise to an admissible decomposition, with a metric. Our metric is slightly different from the one constructed in [69]. It does not involve the Hausdorff distance and it is totally analogous to the metric of the classical space of Grassmannians in $\mathbb{R}^n$. We will also prove that this metric space is compact.

Definition 4.3.1. *Let d be a non negative integer. We say that a subgroup of algebraic dimension d (a d-subgroup) of $\mathbb{H}^n$, $\mathbb{W}$, belongs to the* intrinsic Grassmannian of the d-subgroups $\mathcal{G}(\mathbb{H}^n, d)$ *if there exists a $(2n+1-d)$-subgroup $\mathbb{V}$ such that $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$. The* intrinsic Grassmannian *is defined as*

$$\mathcal{G} = \bigcup_{d=0}^{2n+1} \mathcal{G}(\mathbb{H}^n, d).$$

Definition 4.3.2. *Let $\mathbb{W}_1$ and $\mathbb{W}_2 \in \mathcal{G}(\mathbb{H}^n, d)$. We define*

$$\rho(\mathbb{W}_1, \mathbb{W}_2) = \sup_{\|p\| \leq 1} \|(P_{\mathbb{W}_1}(p))^{-1} \cdot P_{\mathbb{W}_2}(p)\|.$$

Proposition 4.3.1. *$(\mathcal{G}(\mathbb{H}^n, d), \rho)$ is a metric space.*

Proof. We need to check that ρ satisfies the properties of metrics. Let $\mathbb{W} \in \mathcal{G}(\mathbb{H}^n, d)$ be fixed. Then

$$\rho(\mathbb{W}, \mathbb{W}) = \sup_{\|p\| \leq 1} \|(P_{\mathbb{W}}(p))^{-1} \cdot P_{\mathbb{W}}(p)\| = \|e\| = 0.$$

Consider now $\mathbb{W}_1$ and $\mathbb{W}_2 \in \mathcal{G}(\mathbb{H}^n, d)$. Since the gauge is a distance, it holds that

$$\rho(\mathbb{W}_1, \mathbb{W}_2) = \sup_{\|p\| \leq 1} \|(P_{\mathbb{W}_1}(p))^{-1} \cdot P_{\mathbb{W}_2}(p)\| = \sup_{\|p\| \leq 1} \|(P_{\mathbb{W}_2}(p))^{-1} \cdot P_{\mathbb{W}_1}(p)\| = \rho(\mathbb{W}_2, \mathbb{W}_1).$$

Finally, we need to check the triangle inequality. Let us take $\mathbb{W}_1$, $\mathbb{W}_2$ and $\mathbb{W}_3 \in \mathcal{G}(\mathbb{H}^n, d)$ and consider

$$\begin{aligned}
\|(P_{\mathbb{W}_1}(p))^{-1} \cdot P_{\mathbb{W}_3}(p)\| &= \|(P_{\mathbb{W}_1}(p))^{-1} \cdot P_{\mathbb{W}_2}(p) \cdot (P_{\mathbb{W}_2}(p))^{-1} \cdot P_{\mathbb{W}_3}(p)\| \\
&\leq \|(P_{\mathbb{W}_1}(p))^{-1} \cdot P_{\mathbb{W}_2}(p)\| + \|(P_{\mathbb{W}_2}(p))^{-1} \cdot P_{\mathbb{W}_3}(p)\| \\
&\leq \sup_{\|p\| \leq 1} \|(P_{\mathbb{W}_1}(p))^{-1} \cdot P_{\mathbb{W}_2}(p)\| + \sup_{\|p\| \leq 1} \|(P_{\mathbb{W}_2}(p))^{-1} \cdot P_{\mathbb{W}_3}(p)\| \\
&= \rho(\mathbb{W}_1, \mathbb{W}_2) + \rho(\mathbb{W}_2, \mathbb{W}_3),
\end{aligned}$$

for every $p \in \mathbb{H}^n$ with $\|p\| \leq 1$. Therefore, taking the supremum, we get that

$$\rho(\mathbb{W}_1, \mathbb{W}_3) \leq \rho(\mathbb{W}_1, \mathbb{W}_2) + \rho(\mathbb{W}_2, \mathbb{W}_3).$$

□

We prove a proposition which will be a key ingredient in a proof concerning the dimension of the singular set of an H-monotone operator (see Definition 5.1.7 and Theorem 5.2.3). Reducing to the classical case, we are able to prove that $(\mathcal{G}(\mathbb{H}^n, d), \rho)$ is a compact metric space. To prove this fact we need some simple preliminary results concerning the Euclidean Grassmannians.

Notation 4.3.1. Let us denote the space of the Euclidean Grassmannians of dimension $k \in \mathbb{N}$ in $\mathbb{R}^{2n}$ by $G(2n, k)$.

We endow the space $G(2n, k)$ with the usual metric

$$d_G(W_1, W_2) = \sup_{\|\bar{p}\|_{\mathbb{R}^{2n}} \leq 1} \|\Pi_{W_1}(\bar{p}) - \Pi_{W_2}(\bar{p})\|_{\mathbb{R}^{2n}},$$

where $\Pi_{W_i} : \mathbb{R}^{2n} \longrightarrow W_i$ is the Euclidean projection.

Remark 4.3.3. It is well known that $(G(2n, k), d_G)$ is a compact metric space.

With these considerations, we can prove a couple of lemmas useful in the proof of the main result of this section.

Lemma 4.3.1. *Let $1 \leq h \leq n$ and let*

$$G^H(2n, h) := \left\{ V \subset \mathbb{R}^{2n} \;\middle|\; \dim V = h,\ \langle x, Jy \rangle = 0,\ \forall x, y \in V \right\}. \tag{4.5}$$

Then $G^H(2n, h)$ is a closed subspace of $G(2n, h)$.

Proof. Let us consider a sequence $(V_k)_{k \in \mathbb{N}} \subset G^H(2n, h)$ and assume that there exists $V_0 \in G(2n, h)$ such that

$$d_G(V_k, V_0) \longrightarrow 0,$$

as $k \to \infty$. We want to show that $V_0 \in G^H(2n, h)$.

Let x and $y \in V_0$ be two arbitrary points. We want to prove that $\langle x, Jy \rangle = 0$. Without loss of generality, we can normalize and consider $\hat{x} := \frac{x}{\|x\|}$ and $\hat{y} := \frac{y}{\|y\|}$.

By definition of the metric d_G, we can notice that

$$\|\Pi_{V_k}(\hat{x}) - \Pi_{V_0}(\hat{x})\|_{\mathbb{R}^{2n}} \leq \sup_{\|\bar{p}\|_{\mathbb{R}^{2n}} \leq 1} \|\Pi_{V_k}(\bar{p}) - \Pi_{V_0}(\bar{p})\|_{\mathbb{R}^{2n}}.$$

If we denote by $\hat{x}_k := \Pi_{V_k}(\hat{x}) \in V_k$, it follows that

$$\|\hat{x}_k - \hat{x}\|_{\mathbb{R}^{2n}} \longrightarrow 0,$$

as $k \to \infty$. An analogous argument gives that there exists a sequence $\hat{y}_k := \Pi_{V_k}(\hat{y})$ such that

$$\|\hat{y}_k - \hat{y}\|_{\mathbb{R}^{2n}} \longrightarrow 0,$$

as $k \to \infty$.

Now, since $V_k \in G^H(2n,k)$, for every $k \in \mathbb{N}$, it holds that

$$\langle \hat{x}_k, J\hat{y}_k \rangle = 0,$$

for every $k \in \mathbb{N}$. We can pass to the limit, letting $k \to \infty$. We obtain that

$$\langle \hat{x}, J\hat{y} \rangle = 0,$$

and this implies that $V_0 \in G^H(2n,h)$. □

Remark 4.3.4. We highlight that every $V \in G^H(2n,h)$ can be seen as a horizontal subgroup of $\mathbb{H}^n$ of linear dimension h.

Lemma 4.3.2. *Let $n \le d \le 2n-1$ and let*

$$\begin{aligned} G^V(2n,d) := \big\{ W \subset \mathbb{R}^{2n} \,\big|\, & \dim W = d, \\ & \exists\, V \in G^H(2n, 2n-d), \text{ such that } \mathbb{R}^{2n} = W \oplus V \big\}. \end{aligned} \tag{4.6}$$

Then $G^V(2n,d)$ is a closed subspace of $G(2n,d)$.

Proof. Let $(W_k)_{k\in\mathbb{N}}$ be a sequence in $G^V(2n,d)$ and assume that there exists $W_0 \in G(2n,d)$ such that

$$d_G(W_k, W_0) \longrightarrow 0,$$

as $k \to \infty$. We want to prove that $W_0 \in G^V(2n,d)$, i.e. that there exists $V_0 \in G^H(2n, 2n-d)$ such that $\mathbb{R}^{2n} = W_0 \oplus V_0$. In other words, let us consider the orthogonal complement V_0 of W_0 such that $\mathbb{R}^n = W_0 \oplus V_0$. We want to show that $V_0 \in G^H(2n, 2n-d)$. Let $V_k \in G^H(2n, 2n-d)$ be such that $\mathbb{R}^{2n} = W_k \oplus V_k$.

Now, let us performe some calculations:

$$\begin{aligned}
d_G(V_k, V_0) &= \sup_{\|\bar p\|_{\mathbb{R}^{2n}} \le 1} \|\Pi_{V_k}(\bar p) - \Pi_{V_0}(\bar p)\|_{\mathbb{R}^{2n}} \\
&= \sup_{\|\bar p\|_{\mathbb{R}^{2n}} \le 1} \|\Pi_{V_k}(\bar p) + \Pi_{W_k}(\bar p) - \Pi_{W_k}(\bar p) + \Pi_{W_0}(\bar p) - \Pi_{W_0}(\bar p) - \Pi_{V_0}(\bar p)\|_{\mathbb{R}^{2n}} \\
&= \sup_{\|\bar p\|_{\mathbb{R}^{2n}} \le 1} \|\bar p - \Pi_{W_k}(\bar p) + \Pi_{W_0}(\bar p) - \bar p\|_{\mathbb{R}^{2n}} \\
&= \sup_{\|\bar p\|_{\mathbb{R}^{2n}} \le 1} \|\Pi_{W_k}(\bar p) - \Pi_{W_0}(\bar p)\|_{\mathbb{R}^{2n}} \\
&= d_G(W_k, W_0) \longrightarrow 0,
\end{aligned}$$

as $k \to \infty$. Then, thanks to Lemma 4.3.1, it holds that $V_0 \in G^H(2n, 2n-d)$, and the proof is complete. □

Remark 4.3.5. In the previous Lemma, we proved that $G^V(2n, d)$, with $n \le d \le 2n-1$, is a closed subspace of $G(2n, d)$, which is compact. This implies that also $G^V(2n, d)$ is a compact metric space.

Proposition 4.3.2. *Let $n+1 \le d \le 2n$. The metric space $(\mathcal{G}(\mathbb{H}^n, d), \rho)$ is compact.*

Proof. The strategy of the proof is to realize the metric space $(\mathcal{G}(\mathbb{H}^n, d), \rho)$ as the image of a compact metric space through a continuous and surjective map.

For simplicity, we fix a notation. We denote by $p = (\bar p, t)$ a point in $\mathbb{R}^{2n+1}$, where $\bar p \in \mathbb{R}^{2n}$ and $t \in \mathbb{R}$.

Now, we define the map

$$\begin{aligned}
f : G^V(2n, d-1) &\longrightarrow \mathcal{G}(\mathbb{H}^n, d) \\
W &\longmapsto W \oplus \mathbb{R}.
\end{aligned} \tag{4.7}$$

By definition of vertical homogeneous subgroups, it is clear that for every $\mathbb{W} \in \mathcal{G}(\mathbb{H}^n, d)$, there exists $W \in G^V(2n, d-1)$ such that $\mathbb{W} = W \oplus \mathbb{R}$. This implies that f is surjective.

It is now necessary to prove continuity of f. In order to do this, some considerations about projections in $\mathbb{H}^n$ are needed. Let $p = (\bar p, t) \in \mathbb{H}^n$ be fixed. We know that

$$(\bar p, t) = P_{\mathbb{W}}(\bar p, t) \cdot P_{\mathbb{W}^\perp}(\bar p, t), \tag{4.8}$$

where $\mathbb{W}^\perp$ is the unique homogeneous subgroup of $\mathbb{H}^n$ in $\mathcal{G}(\mathbb{H}^n, 2n+2-d)$ such that $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{W}^\perp$ is a decomposition in complementary subgroups. Moreover, it holds also that

$$\bar{p} \,=\, \Pi_W(\bar{p}) + \Pi_{W^\perp}(\bar{p}),$$

where $W \in G^V(2n, d-1)$ is such that $f(W) = \mathbb{W}$ and $W^\perp \in G^H(2n, 2n+1-d)$ satisfies $\mathbb{R}^{2n} = W \oplus W^\perp$.

Now, since $\mathbb{W}^\perp$ is a horizontal homogeneous subgroup, it holds that

$$P_{\mathbb{W}^\perp}(\bar{p}, t) \,=\, (\Pi_{W^\perp}(\bar{p}), 0)\,. \tag{4.9}$$

On the other hand, there exists a function $h : \mathbb{H}^n \longrightarrow \mathbb{R}$ such that

$$P_{\mathbb{W}}(\bar{p}, t) \,=\, (\Pi_W(\bar{p}), h(\bar{p}, t))\,. \tag{4.10}$$

Using equations (4.9) and (4.10), we rewrite relation (4.8):

$$\begin{aligned} p \,=\, (\bar{p}, t) \,&=\, P_{\mathbb{W}}(\bar{p}, t) \cdot P_{\mathbb{W}^\perp}(\bar{p}, t) \\ &=\, (\Pi_W(\bar{p}), h(\bar{p}, t)) \cdot (\Pi_{W^\perp}(\bar{p}), 0) \\ &=\, (\Pi_W(\bar{p}) + \Pi_{W^\perp}(\bar{p}), h(\bar{p}, t) + 2\,\langle \Pi_W(\bar{p}), J\Pi_{W^\perp}(\bar{p})\rangle)\,. \end{aligned}$$

From this chain of equalities, we get an explicit formula for the function h:

$$h(\bar{p}, t) \,=\, t - 2\,\langle \Pi_W(\bar{p}), J\Pi_{W^\perp}(\bar{p})\rangle\,,$$

which helps us to write $P_{\mathbb{W}}(\bar{p}, t) = (\Pi_W(\bar{p}), t - 2\,\langle \Pi_W(\bar{p}), J\Pi_{W^\perp}(\bar{p})\rangle)$.

With this observation, we can come to the last part of the proof. Let $p = (\bar{p}, t) \in \mathbb{H}^n$ be such that $\|p\| \leq 1$ and let $\mathbb{W}_1$ and $\mathbb{W}_2 \in \mathcal{G}(\mathbb{H}^n, d)$ and perform some calculations. We will use the properties of the symplectic matrix J.

$$
\begin{aligned}
&(P_{\mathbb{W}_1}(\bar{p},t))^{-1}\cdot P_{\mathbb{W}_2}(\bar{p},t)\\
&\quad= \left(\Pi_{W_1}(\bar{p}), t-2\left\langle \Pi_{W_1}(\bar{p}), J\Pi_{W_1^\perp}(\bar{p})\right\rangle\right)^{-1}\cdot\left(\Pi_{W_2}(\bar{p}), t-2\left\langle \Pi_{W_2}(\bar{p}), J\Pi_{W_2^\perp}(\bar{p})\right\rangle\right)\\
&\quad= \left(\Pi_{W_2}(\bar{p})-\Pi_{W_1}(\bar{p}), 2\left(\left\langle \Pi_{W_1}(\bar{p}), J\Pi_{W_1^\perp}(\bar{p})\right\rangle - \left\langle \Pi_{W_2}(\bar{p}), J\Pi_{W_2^\perp}(\bar{p})\right\rangle\right)\right.\\
&\qquad\qquad \left. +2\left\langle \Pi_{W_1}(\bar{p}), J\Pi_{W_2^\perp}(\bar{p})\right\rangle\right)\\
&\quad= \left(\Pi_{W_2}(\bar{p})-\Pi_{W_1}(\bar{p}), 2\left(\langle \Pi_{W_1}(\bar{p}), J(\bar{p}-\Pi_{W_1}(\bar{p}))\rangle - \langle \Pi_{W_2}(\bar{p}), J(\bar{p}-\Pi_{W_2}(\bar{p}))\rangle\right)\right.\\
&\qquad\qquad \left. +2\langle \Pi_{W_1}(\bar{p})-\Pi_{W_2}(\bar{p}), J\Pi_{W_2}(\bar{p})\rangle\right)\\
&\quad= \left(\Pi_{W_2}(\bar{p})-\Pi_{W_1}(\bar{p}), 2\left(\langle \Pi_{W_1}(\bar{p}), J\bar{p}\rangle - \langle \Pi_{W_2}(\bar{p}), J\bar{p}\rangle\right)\right.\\
&\qquad\qquad \left. +2\langle \Pi_{W_1}(\bar{p})-\Pi_{W_2}(\bar{p}), J\Pi_{W_2}(\bar{p})\rangle\right)\\
&\quad= \left(\Pi_{W_2}(\bar{p})-\Pi_{W_1}(\bar{p}), 2\langle \Pi_{W_1}(\bar{p})-\Pi_{W_2}(\bar{p}), J\bar{p}\rangle + 2\langle \Pi_{W_1}(\bar{p})-\Pi_{W_2}(\bar{p}), J\Pi_{W_2}(\bar{p})\rangle\right).
\end{aligned}
$$

Now, let us consider the gauge norm of $(P_{\mathbb{W}_1}(\bar{p},t))^{-1}\cdot P_{\mathbb{W}_2}(\bar{p},t)$, which can be estimated using the previous computations and the Cauchy-Schwarz inequality,

$$
\begin{aligned}
&\left\|(P_{\mathbb{W}_1}(\bar{p},t))^{-1}\cdot P_{\mathbb{W}_2}(\bar{p},t)\right\|\\
&\quad= \left(\|\Pi_{W_1}(\bar{p})-\Pi_{W_2}(\bar{p})\|_{\mathbb{R}^{2n}}^4 + |2\langle \Pi_{W_1}(\bar{p})-\Pi_{W_2}(\bar{p}), J\bar{p}\rangle + 2\langle \Pi_{W_1}(\bar{p})-\Pi_{W_2}(\bar{p}), J\Pi_{W_2}(\bar{p})\rangle|^2\right)^{\frac{1}{4}}\\
&\quad= \left(\|\Pi_{W_1}(\bar{p})-\Pi_{W_2}(\bar{p})\|_{\mathbb{R}^{2n}}^4 + 16\,\|\Pi_{W_1}(\bar{p})-\Pi_{W_2}(\bar{p})\|^2\right)^{\frac{1}{4}}\\
&\quad\le \max\left\{d_G(W_1,W_2), 2d_G(W_1,W_2)^{\frac{1}{2}}\right\}.
\end{aligned}
$$

Since the previous inequality holds for every $p\in\mathbb{H}^n$, with $\|p\|\le 1$, we take the supremum and conclude that

$$\rho(\mathbb{W}_1,\mathbb{W}_2) = \rho(f(W_1), f(W_2)) \le \max\left\{d_G(W_1,W_2), 2d_G(W_1,W_2)^{\frac{1}{2}}\right\},$$

which implies that f is continuous, more precisely f is 1/2-Hölder continuous. □

4.4 Intrinsic cones and Lipschitz graphs

Think for a moment about what it means that a function $f:\mathbb{R}^k\longrightarrow\mathbb{R}^h$ is Lipschitz continuous: there exists a constant $L>0$ such that, for every x and $y\in\mathbb{R}^k$,

$$\|f(x)-f(y)\|_{\mathbb{R}^h} \le L\,\|x-y\|_{\mathbb{R}^k}.$$

Geometrically, this means that there exists a cone, with opening L, whose vertex can be translated along the graph, so that the graph always remains locally outside the cone.

In the case of intrinsic Lipschitz functions within $\mathbb{H}^n$ the idea is the same. With the notion of decomposition of $\mathbb{H}^n$ in the semidirect product of homogeneous subgroups in mind, we can consider a notion of *intrinsic cone*. The interested readers are referred to the papers [39], [69] and [48]. The definition of intrinsic cone is the principal ingredient of the notion of intrinsic Lipschitz map (see Definition 4.4.13).

Definition 4.4.1. *Let $\mathbb{H}^n$ be the semidirect product of two subgroups $\mathbb{W}$ and $\mathbb{V}$. Let $q \in \mathbb{H}^n$ and $\alpha \in \mathbb{R}^+$ be fixed. We call* intrinsic closed cone *with base $\mathbb{W}$, axis $\mathbb{V}$, vertex q and opening α*

$$C_{\mathbb{W},\mathbb{V}}(q,\alpha) := q \cdot C_{\mathbb{W},\mathbb{V}}(e,\alpha),$$

where

$$C_{\mathbb{W},\mathbb{V}}(e,\alpha) := \{p \in \mathbb{H}^n \mid \|p_{\mathbb{W}}\| \leq \alpha \|p_{\mathbb{V}}\|\}.$$

The next proposition justifies the adjective "intrinsic". Indeed, intrinsic cones are invariant under group dilations.

Proposition 4.4.1. *Let $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$ be a semidirect product, $\lambda \in \mathbb{R}^+$ and $0 < \alpha < \beta$. Then the following statements hold:*

(i) $C_{\mathbb{W},\mathbb{V}}(e,0) = \mathbb{V}$;

(ii) $C_{\mathbb{W},\mathbb{V}}(q,\alpha) \subset C_{\mathbb{W},\mathbb{V}}(q,\beta)$;

(iii) $\delta_\lambda \left(C_{\mathbb{W},\mathbb{V}}(e,\alpha)\right) = C_{\mathbb{W},\mathbb{V}}(e,\alpha)$.

Proof. Since (i) and (ii) are trivial, we prove just (iii). Because of the uniqueness of the components in the decomposition in semirect product, we have that $(\delta_\lambda p)_{\mathbb{W}} =$

$\delta_\lambda(p_{\mathbb{W}})$ and also $(\delta_\lambda p)_{\mathbb{V}} = \delta_\lambda(p_{\mathbb{V}})$. Hence,

$$\begin{aligned}\delta_\lambda(C_{\mathbb{W},\mathbb{V}}(e,\alpha)) &= \delta_\lambda\{p \in \mathbb{H}^n \mid \|p_{\mathbb{W}}\| \le \alpha \|p_{\mathbb{V}}\|\} \\ &= \{\delta_\lambda(p) \in \mathbb{H}^n \mid \|\delta_\lambda(p)_{\mathbb{W}}\| \le \alpha \|\delta_\lambda(p)_{\mathbb{V}}\|\} \\ &= \{\delta_\lambda(p) \in \mathbb{H}^n \mid \lambda \|p_{\mathbb{W}}\| \le \alpha\, \lambda \|p_{\mathbb{V}}\|\} \\ &= \{p \in \mathbb{H}^n \mid \|p_{\mathbb{W}}\| \le \alpha \|p_{\mathbb{V}}\|\} = C_{\mathbb{W},\mathbb{V}}(e,\alpha).\end{aligned}$$

□

Here it is a slightly different definition of intrinsic cone, introduced in [69] and [39]. The two notions are equivalent, as one can see in Remark 4.4.4, but depending on what it is more convenient, in what follows we will prefer one over the other.

Definition 4.4.2. *Let $\mathbb{G}$ be a subgroup of dimension k of $\mathbb{H}^n$. We define the* intrinsic cone *$X(p_0, \mathbb{G}, \beta)$, with axis $\mathbb{G}$, vertex p_0 and opening $\beta \le 1$, the set*

$$X(p_0, \mathbb{G}, \beta) := \left\{p \in \mathbb{H}^n \;\middle|\; \operatorname{dist}\left(p_0^{-1} \cdot p, \mathbb{G}\right) \le \beta d(p, p_0)\right\}.$$

Remark 4.4.3. Also for this type of cones, a result similar to Proposition 4.4.1 holds. Indeed, Definition 4.4.2 is invariant by group translations. This means that

$$X(p_0, \mathbb{G}, \beta) = p_0 \cdot X(e, \mathbb{G}, \beta).$$

Moreover, $X(e, \mathbb{G}, \beta)$ is homogeneous, i.e. if $\lambda \in \mathbb{R}^+$, then

$$\delta_\lambda\left(X(e, \mathbb{G}, \beta)\right) = X(e, \mathbb{G}, \beta).$$

Remark 4.4.4. Let $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$ be a decomposition in complementary subgroups. Then, for every $0 < \beta \le 1$, there exists $\alpha \ge 1$, depending on $\beta, \mathbb{W}$ and $\mathbb{V}$, such that

$$C_{\mathbb{W},\mathbb{V}}(p_0, 1/\alpha) \subset X(p_0, \mathbb{V}, \beta) \subset C_{\mathbb{W},\mathbb{V}}(p_0, \alpha).$$

The proof of this fact can be found in [39].

The following lemma provides a relation between cones with axis $\mathbb{V}$ and $\mathbb{W}$, where $\mathbb{V}$ and $\mathbb{W}$ are complementary subgroups of $\mathbb{H}^n$. In the Euclidean case, whenever we have a cone, the complementary space is a cone too. In this setting, the same holds. We call *complementary cones* pair of cones built how it follows.

Lemma 4.4.1. *Let* $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$ *be a decomposition in complementary subgroups, with* $\dim \mathbb{V} = k \leq n$*. If* $0 < \alpha_0 < 1$ *is fixed, then there exists* $0 < \beta_0 < 1$*, dependent on* α_0*, such that*

$$\mathbb{H}^n \setminus X(e, \mathbb{V}, \alpha_0) \subseteq X(e, \mathbb{W}, \beta_0).$$

Proof. For simplicity of calculations, we can assume that

$$\mathbb{V} \,=\, \{(x_1, ..., x_k, 0, ..., 0) \,\in \mathbb{H}^n \,|\, x_i \in \mathbb{R}, i = 1, ..., k\,\}$$

and

$$\mathbb{W} = \{(0, ..., 0, x_{k+1}, ..., x_n, y_1, ..., y_n, t) \in \mathbb{H}^n \,|\, x_j, y_i, t \in \mathbb{R}, i = 1, ..., n, j = k{+}1, ..., n\}.$$

Let us denote by $E := \mathbb{H}^n \setminus X(e, \mathbb{V}, \alpha_0)$ and $\pi(E) = \left\{\delta_{1/\|p\|}(p) \in \mathbb{H}^n \,|\, p \in E\right\}$. It is clear that $\pi(E)$ is contained in the unit sphere of $\mathbb{H}^n$, denoted by $\mathbb{S}$. Since $E \cap X(e, \mathbb{V}, \alpha_0) = \emptyset$, it holds that $\pi(E) \cap X(e, \mathbb{V}, \alpha_0) = \emptyset$.

Now, $\mathbb{V} \cap \mathbb{S} = \mathbb{S}^{k-1}$, where $\mathbb{S}^{k-1}$ is the unit sphere of $\mathbb{R}^k \simeq \operatorname{span}\{e_1, ..., e_k\}$, with respect to the Euclidean metric. Therefore, there exists $\eta = \eta(\alpha_0) > 0$ such that

$$\mathbb{S} \cap \left(\mathbb{S}^{k-1}\right)_\eta \subseteq \mathbb{S} \cap X(e, \mathbb{V}, \alpha_0), \tag{4.11}$$

where $\left(\mathbb{S}^{k-1}\right)_\eta = \left\{q \in \mathbb{H}^n \,|\, \operatorname{dist}(q, \mathbb{S}^{k-1}) \leq \eta\right\}$ is the η-neighbourhood of $\mathbb{S}^{k-1}$.

Thanks to convexity of the gauge ball, inclusion (4.11) implies that there exists $\delta = \delta(\eta) > 0$ such that, for every $\hat{p} = (\hat{x}_1, ..., \hat{x}_k, \hat{x}_{k+1}, ..., \hat{x}_n, \hat{y}_1, ..., \hat{y}_n, \hat{t}) \in \pi(E)$, $\|(\hat{x}_1, ..., \hat{x}_k)\|_{\mathbb{R}^k} < 1 - \delta$. We highlight that $\delta > 0$ depends on η, therefore on α_0, but not on the point $\hat{p}$.

Let $p \in E$ be arbitrary and set $\beta_0 = 1 - \delta$. We aim to show that $p \in X(e, \mathbb{W}, \beta_0)$. Define $\hat{p} = \delta_{1/\|p\|}(p)$. Clearly, $\hat{p} \in \pi(E)$ as it is true that

$$\operatorname{dist}(\hat{p}, \mathbb{W}) \,=\, \min_{q \in \mathbb{W}} \|q^{-1} \cdot \hat{p}\| \,\leq\, \|(\hat{x}_1, ..., \hat{x}_k)\|_{\mathbb{R}^k} \,<\, \beta_0.$$

Hence, by homogeneity of $\mathbb{W}$, we have that

$$\operatorname{dist}(p, \mathbb{W}) \leq \beta_0 \|p\|,$$

which implies that $p \in X(e, \mathbb{W}, \beta_0)$, giving that $E \subseteq X(e, \mathbb{W}, \beta_0)$. □

Remark 4.4.5. Let $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$ be a decomposition in complementary subgroups, with $\dim \mathbb{V} = k \leq n$. Thanks to Lemma 4.4.1, we know that if $0 < \alpha_0 < 1$, then there exists $0 < \beta_0 < 1$, dependent on α_0, such that

$$\mathbb{H}^n \setminus X(e, \mathbb{V}, \alpha_0) \subseteq X(e, \mathbb{W}, \beta_0).$$

Now, from Remark 4.4.4, since $0 < \beta_0 < 1$, there exists $\alpha \geq 1$, dependent on β_0, $\mathbb{W}$ and $\mathbb{V}$, such that

$$X(e, \mathbb{W}, \beta_0) \subseteq C_{\mathbb{V},\mathbb{W}}(e, \alpha). \tag{4.12}$$

Analogously, since $0 < \alpha < 1$, there exists $\beta \geq 1$, such that

$$X(e, \mathbb{V}, \alpha_0) \subseteq C_{\mathbb{W},\mathbb{V}}(e, \beta).$$

Therefore, one has that

$$\mathbb{H}^n \setminus C_{\mathbb{W},\mathbb{V}}(e, \beta) \subseteq \mathbb{H}^n \setminus X(e, \mathbb{V}, \alpha_0). \tag{4.13}$$

We combine (4.12) and (4.13) to obtain finally

$$\mathbb{H}^n \setminus C_{\mathbb{W},\mathbb{V}}(e, \beta) \subset C_{\mathbb{V},\mathbb{W}}(e, \alpha).$$

The question now is: what can we say about two cones whose axis are vertical subgroups which are near to each other? We consider two elements of the intrinsic Grassmannian and we assume that their distance is very small. It holds that there exist two openings, depending on the distance, for which we have an inclusion between the cones.

For simplicity, first, we fix once for all a notation.

Notation 4.4.1. Let $n + 1 \leq d \leq 2n$. If $\mathbb{W} \in \mathcal{G}(\mathbb{H}^n, d)$, we denote by $\mathbb{W}^\perp$ the horizontal subgroup of $\mathbb{H}^n$ such that $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{W}^\perp$ is a decomposition in homogeneous subgroups.

Lemma 4.4.2. *Let $n+1 \leq d \leq 2n$ and let $\alpha > 0$. If $\delta < \frac{1}{2(1+\alpha)}$ and $\mathbb{W}_1$ and $\mathbb{W}_2 \in \mathcal{G}(\mathbb{H}^n, d)$ are such that $\rho(\mathbb{W}_1, \mathbb{W}_2) < \delta$, it holds that*

$$C_{\mathbb{W}_1^\perp, \mathbb{W}_1}(e, \alpha) \subseteq C_{\mathbb{W}_2^\perp, \mathbb{W}_2}(e, 2(1+\alpha)).$$

Proof. By homogeneity of the intrinsic cones, it is sufficient to prove the following inclusion

$$C_{\mathbb{W}_1^\perp, \mathbb{W}_1}(e, \alpha) \cap \partial B(e,1) \subseteq C_{\mathbb{W}_2^\perp, \mathbb{W}_2}(e, 2(1+\alpha)) \cap \partial B(e,1).$$

Let $q \in \partial B(e,1)$ be such that $q \in C_{\mathbb{W}_1^\perp, \mathbb{W}_1}(e, \alpha)$. This means that

$$\left\| P_{\mathbb{W}_1^\perp}(q) \right\| \leq \alpha \left\| P_{\mathbb{W}_1}(q) \right\|. \tag{4.14}$$

We want to show that $q \in C_{\mathbb{W}_2^\perp, \mathbb{W}_2}(e, 2(1+\alpha))$, i.e.

$$\left\| P_{\mathbb{W}_2^\perp}(q) \right\| \leq 2(1+\alpha) \left\| P_{\mathbb{W}_2}(q) \right\|.$$

First, we notice that, since $\mathbb{H}^n = \mathbb{W}_1 \cdot \mathbb{W}_1^\perp = \mathbb{W}_2 \cdot \mathbb{W}_2^\perp$, it holds that

$$q = P_{\mathbb{W}_1}(q) \cdot P_{\mathbb{W}_1^\perp}(q) = P_{\mathbb{W}_2}(q) \cdot P_{\mathbb{W}_2^\perp}(q).$$

Therefore,

$$\left\| P_{\mathbb{W}_2}(q) \cdot P_{\mathbb{W}_1}(q)^{-1} \right\| = \left\| P_{\mathbb{W}_1^\perp}(q) \cdot P_{\mathbb{W}_2^\perp}(q)^{-1} \right\| < \delta. \tag{4.15}$$

Now, using the triangle inequality and relations (4.14) and (4.15), we compute

$$\begin{aligned}
\left\| P_{\mathbb{W}_2^\perp}(q) \right\| &= \left\| P_{\mathbb{W}_2^\perp}(q) \cdot P_{\mathbb{W}_1^\perp}(q)^{-1} \cdot P_{\mathbb{W}_1^\perp}(q) \right\| \leq \left\| P_{\mathbb{W}_1^\perp}(q) \right\| + \delta \\
&\leq \alpha \left\| P_{\mathbb{W}_1}(q) \right\| + \delta = \alpha \left\| P_{\mathbb{W}_1}(q) \cdot P_{\mathbb{W}_2}(q)^{-1} \cdot P_{\mathbb{W}_2}(q) \right\| + \delta \\
&\leq \alpha \left\| P_{\mathbb{W}_2}(q) \right\| + \delta(1+\alpha) < \alpha \left\| P_{\mathbb{W}_2}(q) \right\| + \frac{1}{2}.
\end{aligned}$$

Using the fact that $p \in \partial B(e,1)$, we can make a step further, we have that

$$\begin{aligned}
1 = \|q\| &= \left\| P_{\mathbb{W}_2}(q) \cdot P_{\mathbb{W}_2^\perp}(q) \right\| \leq \left\| P_{\mathbb{W}_2}(q) \right\| + \left\| P_{\mathbb{W}_2^\perp}(q) \right\| \\
&< \left\| P_{\mathbb{W}_2}(q) \right\| + \alpha \left\| P_{\mathbb{W}_2}(q) \right\| + \frac{1}{2},
\end{aligned}$$

which implies that

$$\left\| P_{\mathbb{W}_2}(q) \right\| > \frac{1}{2(1+\alpha)}.$$

On the other hand, we have also that

$$\left\| P_{\mathbb{W}_2^\perp}(q) \right\| \le 1 = 2(1+\alpha)\frac{1}{2(1+\alpha)} < 2(1+\alpha)\left\| P_{\mathbb{W}_2}(q) \right\|,$$

which is the inequality we were looking for. □

We are ready to consider intrinsic Lipschitz graphs. First we give the definition of intrinsic graph and some major properties.

Definition 4.4.6. *Let $\mathbb{W}$ and $\mathbb{V}$ be homogeneous subgroups of $\mathbb{H}^n$, with $\mathbb{H}^n = \mathbb{W}\cdot\mathbb{V}$ a semidirect product. We say that $S \subset \mathbb{H}^n$ is a* (left) graph *over $\mathbb{W}$ along $\mathbb{V}$ (or from $\mathbb{W}$ to $\mathbb{V}$) if*

$$S \cap (\xi \cdot \mathbb{V})$$

contains at most one point for all $\xi \in \mathbb{W}$.

Remark 4.4.7. An equivalent definition is the following: we say that $S \subset \mathbb{H}^n$ is a (left) graph from $\mathbb{W}$ to $\mathbb{V}$ if there exists a function $f : \mathcal{E} \subset \mathbb{W} \longrightarrow \mathbb{V}$ such that

$$S = \{\xi \cdot f(\xi) \mid \xi \in \mathcal{E}\}.$$

In this case we write $S = \mathrm{graph}(f)$.

Remark 4.4.8. In this definition, it does not play any role which subgroup of the decomposition is horizontal or vertical.

In [48], the authors explain that they call "intrinsic" these graphs because their properties are defined only in terms of the group structure of $\mathbb{H}^n$. The following propositions explain that if we dilate a graph we have again a graph and that the same holds if we left-translate it. For the left translation, we have to split in two cases. The first refers to graphs of functions acting from a vertical subgroup to a horizontal one, the second refers to the opposite case.

Proposition 4.4.2. *Let $\mathbb{H}^n = \mathbb{W}\cdot\mathbb{V}$ be a semidirect product of $\mathbb{H}^n$ and let S be a graph from $\mathbb{W}$ to $\mathbb{V}$. Then, for all $\lambda \in \mathbb{R}^+$, $\delta_\lambda(S)$ is a graph.*

Proof. Since S is a graph, there exists a function $f : \mathcal{E} \subset \mathbb{W} \longrightarrow \mathbb{V}$ such that $S = \{\, \xi \cdot f(\xi) \,|\, \xi \in \mathcal{E} \,\}$. Let us define

$$f_\lambda := \delta_\lambda \circ f \circ \delta_{\frac{1}{\lambda}} : \delta_\lambda\mathcal{E} \subset \mathbb{W} \longrightarrow \mathbb{V}.$$

Therefore, $\delta_\lambda S := \mathrm{graph}(f_\lambda)$. Indeed

$$\begin{aligned}\delta_\lambda(S) &= \{\, \delta_\lambda(\xi \cdot f(\xi)) \,|\, \xi \in \mathcal{E} \,\} \\ &= \{\, \delta_\lambda(\xi) \cdot \delta_\lambda(f(\xi)) \,|\, \xi \in \mathcal{E} \,\}.\end{aligned}$$

Setting $\eta := \delta_\lambda(\xi)$, $\eta \in \delta_\lambda(\mathcal{E})$, then

$$\left\{\, \eta \cdot \delta_\lambda f\left(\delta_{\frac{1}{\lambda}}(\eta)\right) \,\middle|\, \eta \in \delta_\lambda(\mathcal{E}) \,\right\} = \{\, \eta \cdot f_\lambda(\eta) \,|\, \eta \in \delta_\lambda(\mathcal{E}) \,\} = \mathrm{graph}(f_\lambda).$$

□

Proposition 4.4.3. *Let $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$ be as in Proposition 4.2.4. Let $S \subset \mathbb{H}^n$ be a left graph such that*

$$S = \{\, \xi \cdot f(\xi) \,|\, \xi \in \mathcal{E} \subset \mathbb{W} \,\},$$

with $f : \mathcal{E} \subset \mathbb{W} \longrightarrow \mathbb{V}$. Then, for every $q \in \mathbb{H}^n$, there are

$$\begin{aligned}&\mathcal{E}_q = \{\, q \cdot \xi \cdot (q_{\mathbb{V}})^{-1} \,|\, \xi \in \mathcal{E} \,\} \text{ and} \\ &f_q : \mathcal{E}_q \longrightarrow \mathbb{V},\ f_q(\eta) = q_{\mathbb{V}} \cdot f(q_{\mathbb{V}}^{-1} \cdot q_{\mathbb{W}}^{-1} \cdot \eta \cdot q_{\mathbb{V}}),\end{aligned}$$

such that

$$q \cdot S = \mathrm{graph}(f_q) = \{\, \eta \cdot f_q(\eta) \,|\, \eta \in \mathcal{E}_q \,\}.$$

Proof. See [12], Proposition 3.6. □

Proposition 4.4.4. *Let $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$ be as in Proposition 4.2.4. Let $S \subset \mathbb{H}^n$ be a left graph such that*

$$S = \{\, f(\xi) \cdot \xi \,|\, \xi \in \mathcal{A} \subset \mathbb{V} \,\},$$

with $f : \mathcal{A} \subset \mathbb{V} \longrightarrow \mathbb{W}$. Then, for every $q \in \mathbb{H}^n$, there are

$$\begin{aligned}&\mathcal{A}_q = \{\, q \cdot \xi \,|\, \xi \in \mathcal{A} \,\} \text{ and} \\ &f_q : \mathcal{A}_q \longrightarrow \mathbb{V},\ f_q(\eta) = \eta^{-1} \cdot q_{\mathbb{W}} \cdot \eta \cdot f(q_{\mathbb{V}}^{-1} \cdot \eta),\end{aligned}$$

such that

$$q \cdot S = \mathrm{graph}(f_q) = \{\, \eta \cdot f_q(\eta) \,|\, \eta \in \mathcal{A}_q \,\}.$$

Proof. See [12], Proposition 3.6. □

Remark 4.4.9. Let $f : \mathcal{A} \subset \mathbb{V} \longrightarrow \mathbb{W}$ be such that $S = \{\, \xi \cdot f(\xi) \,|\, \xi \in \mathcal{A} \subset \mathbb{V} \,\}$ is a left graph in $\mathbb{H}^n$. Then S is also a Euclidean graph over $\mathbb{V}$. Indeed, recalling that $\mathbb{V}$ is isometric and isomorphic to $\mathbb{R}^k$, for some $0 < k < 2n+1$, we can identify $\mathbb{V}$ with a k-dimensional vector subspace of $\mathbb{R}^{2n+1}$. On the contrary, if $S = \text{graph}(f)$, where $f : \mathcal{E} \subset \mathbb{W} \longrightarrow \mathbb{V}$, then, in general, S is not a Euclidean graph.

An example is given in [44]: consider the semidirect product $\mathbb{H}^1 = \mathbb{W} \cdot \mathbb{V}$, where $\mathbb{W} = \{\, (0, y, t) \in \mathbb{H}^1 \,|\, y, t \in \mathbb{R} \,\}$ and $\mathbb{V} = \{\, (x, 0, 0) \in \mathbb{H}^1 \,|\, x \in \mathbb{R} \,\}$. Let us fix $\frac{1}{2} < \alpha < 1$ and take $f : \mathbb{W} \longrightarrow \mathbb{V}$ defined as

$$f(0, y, t) \;=\; (|t|^\alpha, 0, 0).$$

It is clear that $\text{graph}(f) = S$ is not an Euclidean graph near the origin:

$$S \;=\; \{\, \xi \cdot f(\xi) \,|\, \xi \in \mathbb{W} \,\} \;=\; \{\, (|t|^\alpha, y, t + 2y|t|^\alpha) \in \mathbb{H}^1 \,|\, t, y \in \mathbb{R} \,\}.$$

Let us give the definition of intrinsic Lipschitz function. This first version is more analytic. In a second moment, we will use the notion of intrinsic cones to give a more geometrical definition.

Definition 4.4.10. *Let $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$ be a semidirect product. We say that*

$$f \;:\; \mathcal{E} \subset \mathbb{W} \;\longrightarrow\; \mathbb{V}$$

is an intrinsic Lipschitz continuous function *if there exists a positive constant L such that, for all $q \in \text{graph}(f)$,*

$$\|f_{q^{-1}}(x)\| \;\leq\; L\,\|x\|, \tag{4.16}$$

for each $x \in \mathcal{E}_{q^{-1}}$. As usual, we call the intrinsic Lipschitz constant *of f the infimum of the numbers L such that* (4.16) *holds.*

Remark 4.4.11. The same definition holds also for $f : \mathcal{E} \subset \mathbb{V} \longrightarrow \mathbb{W}$.

Remark 4.4.12. Let $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$ be as in Proposition 4.2.4. Using Propositions 4.4.3 and 4.4.4, we can specify the two cases:

(i) $f \;:\; \mathbb{W} \;\longrightarrow\; \mathbb{V}$ is said to be an intrinsic Lipschitz function, if there exists a positive constant L such that, for all $\xi, \bar{\xi} \in \mathbb{W}$,

$$\|f(\xi)^{-1} \cdot f(\bar{\xi})\| \;\leq\; L\,\|f(\xi)^{-1} \cdot \xi^{-1} \cdot \bar{\xi} \cdot f(\xi)\|;$$

(ii) $f : \mathbb{V} \longrightarrow \mathbb{W}$ is said to be an intrinsic Lipschitz function, if there exists a positive constant L such that, for all $\eta, \bar{\eta} \in \mathbb{V}$,

$$\|\bar{\eta}^{-1} \cdot \eta \cdot f(\eta)^{-1} \cdot \eta^{-1} \cdot \bar{\eta} \cdot f(\bar{\eta})\| \leq L \, \|\eta^{-1} \cdot \bar{\eta}\|.$$

Our aim now is to give a more geometrical definition of intrinsic Lipschitz continuity: if $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$ is a semidirect product and $f : \mathbb{W} \longrightarrow \mathbb{V}$, we will say that f is intrinsic Lipschitz continuous if, at every point $p \in \text{graph}(f)$, there is an intrinsic closed cone, with vertex p and axis $\mathbb{V}$, intersecting $\text{graph}(f)$ only in p.

Definition 4.4.13. *Let $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$ be a semidirect product. We say that*

$$f : \mathcal{E} \subset \mathbb{W} \longrightarrow \mathbb{V}$$

is intrinsic Lipschitz continuous *in $\mathcal{E}$, if there exists a positive constant L such that, for every $q \in \text{graph}(f)$,*

$$C_{\mathbb{W},\mathbb{V}}\left(q, \frac{1}{L}\right) \cap \text{graph}(f) = \{q\}. \tag{4.17}$$

As usual we call the Lipschitz constant *of f in $\mathcal{E}$ the infimum of the numbers L such that* (4.17) *holds.*

We prove now the equivalence between the two definitions of intrinsic Lipschitz continuity:

Proposition 4.4.5. *Let $\mathbb{H}^n = \mathbb{W} \cdot \mathbb{V}$ be as in Proposition 4.2.4. A function $f : \mathbb{W} \longrightarrow \mathbb{V}$ is intrinsic Lipschitz continuous according to Definition 4.4.10, with Lipschitz constant L, if and only if , for each $q \in \text{graph}(f)$ and for all α such that $0 \leq \alpha < \frac{1}{L}$,*

$$C_{\mathbb{W},\mathbb{V}}(q, \alpha) \cap \text{graph}(f) = \{q\}.$$

Proof. If $q \in \text{graph}(f)$,

$$C_{\mathbb{W},\mathbb{V}}(e, \alpha) \cap \text{graph}(f_{q^{-1}}) = \{e\},$$

hence,

$$\tau_q \left(C_{\mathbb{W},\mathbb{V}}(e, \alpha) \cap \text{graph}(f_{q^{-1}})\right) = \{q\}.$$

On the other hand,

$$\begin{aligned}\tau_q\left(C_{\mathbb{W},\mathbb{V}}(e,\alpha)\,\cap\,\text{graph}(f_{q^{-1}})\right) &= \tau_q\left(C_{\mathbb{W},\mathbb{V}}(e,\alpha)\right)\,\cap\,\tau_q\left(\tau_{q^{-1}}\text{graph}(f)\right)\\ &= C_{\mathbb{W},\mathbb{V}}(q,\alpha)\,\cap\,\text{graph}(f).\end{aligned}$$

□

Proposition 4.4.6. *Let $\mathbb{H}^n = \mathbb{W}\cdot\mathbb{V}$ be as in Proposition 4.2.4. Then*

(i) $f : \mathcal{E} \subset \mathbb{V} \longrightarrow \mathbb{W}$ is intrinsic Lipschitz continuous in $\mathcal{E}$, if and only if the parametrization map

$$\Phi_f \,:\, \mathcal{E} \,\longrightarrow\, \mathbb{H}^n,$$

defined as $\Phi_f(v) = v\cdot f(v)$, is metric Lipschitz continuous

(ii) $f : \mathcal{E} \subset \mathbb{V} \longrightarrow \mathbb{W}$ is intrinsic Lipschitz continuous in $\mathcal{E}$, if and only if there is a positive constant L such that, for all $\eta, \bar{\eta} \in \mathcal{E}$,

$$\|f(\xi)^{-1}\cdot f(\bar{\xi})\| \,\leq\, L\,\|f(\xi)^{-1}\cdot \xi^{-1}\cdot \bar{\xi}^{-1}\cdot f(\xi)\|.$$

Proof. See [48], Proposition 4.6. □

Proposition 4.4.7. *Let $\mathbb{H}^n = \mathbb{W}\cdot\mathbb{V}$ be a decomposition in homogeneous subgroups and let $n+1 \leq d \leq 2n$ be the linear dimension of $\mathbb{W}$. If $f : \mathcal{A} \subset \mathbb{W} \longrightarrow \mathbb{V}$ is intrinsic L-Lipschitz in $\mathcal{A}$ and $\mathcal{A}$ is relatively open in $\mathbb{W}$, then* graph(f) *has Hausdorff dimension $d+1$ and there is a geometric constant $c = c(\mathbb{W},\mathbb{V}) > 0$ such that, for all $p \in \mathbb{H}^n$ and $R > 0$,*

$$\mathcal{S}^{d+1}_{d_c}(\text{graph}(f)\cap B(p,R)) \,\leq\, c(1+L)^{d+1}R^{d+1}. \tag{4.18}$$

Proof. See [49], Theorem 3.9. □

Remark 4.4.14. An analogous result holds for intrinsic Lipschitz functions of type $f : \mathbb{V} \longrightarrow \mathbb{W}$.

Chapter 5

Hausdorff dimension estimate of singular sets of H-monotone maps

The main goal of this chapter is to study the singular set of an H-monotone map (see Definition 5.1.7).

For a better understanding, let us go back to the Euclidean case. In this setting, the problem was largely studied by Alberti and Ambrosio in [2]. In this paper, the authors studied the singular set of a monotone map in $\mathbb{R}^n$. What is a monotone map? First we need the definition of set valued maps.

Definition 5.0.1. *We say that a function $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a* set valued map *if, for every point $x \in \mathbb{R}^n$, $T(x)$ is a subset of $\mathbb{R}^n$. In this case we write $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$.*

Notation 5.0.1. Let T be a set valued map. For every $x \in \mathbb{R}^n$, we denote

- the *domain* of T the set $\mathrm{dom}(T) = \{x \in \mathbb{R}^n \,|\, T(x) \neq \emptyset\}$;
- the *image* of T is the set $\mathrm{im}(T) = \{y \in \mathbb{R}^n \,|\, \text{there exists}\, x \in \mathbb{R}^n \,\text{such that}\, y \in T(x)\}$;
- the *graph* of T is the set $\mathrm{gr}(T) = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n \,|\, y \in T(x)\}$;
- the *inverse* of T is the set valued map $[T^{-1}](y) = \{x \in \mathbb{R}^n \,|\, x \in T(y)\}$.

If S is another set valued map, we write $S \subset T$ when the graph of S is contained in the graph of T.

Definition 5.0.2. *Let $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a set valued map. We say that T is* monotone *if*

$$\langle y_1 - y_2\,,\, x_1 - x_2 \rangle \,\geq\, 0,$$

for every $x_i \in \mathbb{R}^n$, $y_i \in T(x_i)$, with $i = 1, 2$.

Remark 5.0.3. We point out that for this definition it is needed just a notion of scalar product. In fact, one can define a monotone set valued map also on a general Hilbert space.

Definition 5.0.4. *Let $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a monotone set valued map. We say that T is* maximal *if it is maximal with respect to inclusion in the class of monotone functions.*

Remark 5.0.5. The previous definition means the following: if S is a monotone set valued map such that $T \subset S$, then necessarily $T = S$.

Remark 5.0.6. We point out that T is monotone if and only if T^{-1} is monotone.

Remark 5.0.7. Let $u : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a convex function. We defined in Section 1.2 the notion of subdifferential and we pointed out that the map $x \longmapsto \partial u(x)$ is a monotone set valued map.

Proposition 5.0.1. *Let $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a monotone set valued map. We denote by $I : \mathbb{R}^n \to \mathbb{R}^n$ the identity map. Then the following statements are true:*

(i) if T is maximal, then $\mathrm{gr}(T)$ is closed and $T(x)$ is closed and convex for every $x \in \mathbb{R}^n$;

(ii) T is maximal if and only if $(T + I)$ is surjective or, equivalently, if and only if the domain of $(T + I)^{-1}$ is $\mathbb{R}^n$;

(iii) $(T + I)$ and $(T + I)^{-1}$ are monotone and $(T + I)^{-1}$ is a 1-Lipschitz function;

(iv) T is upper semicontinuous, i.e. if $x_n \longrightarrow x$, $y_n \longrightarrow y$, as $n \to \infty$, and $y_n \in T(x_n)$, then $y \in T(x)$.

Proof. See [2], Proposition 1.2 and Corollary 1.3. □

As announced before, we are interested in the study of the singular sets of a monotone set valued map. With singular set, we mean the set of those points where the function is not univalued. More specifically, we consider the set of those points which are sent to a set of a given dimension.

Definition 5.0.8. *Let $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a maximal monotone set valued map and let $k = 1, ..., n$. We call the* k-th singular set *of T the set*

$$\Sigma^k(T) \;=\; \{x \in \mathbb{R}^n \mid \dim T(x) \geq \; k\} \, .$$

Theorem 5.0.2. *Let $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a maximal monotone set valued map. Then the Hausdorff dimension of the set $\Sigma^k(T)$ is at most $(n-k)$-rectifiable.*

Proof. Since $\Sigma^k(T) = \Sigma^k(T{+}I)$, we can move our investigation to the set $\Sigma^k(T{+}I)$. Let S be a countable and dense subset of $\mathbb{R}^n$ and F be a countable and dense subset of $G(n, n-k)$, the Grassmann manifold of $(n-k)$-planes in $\mathbb{R}$. It holds that

$$\Sigma^k(T+I) \;\subseteq\; (T+I)^{-1}\left(\bigcup_{y\in S,\, P\in F} y+P\right) \;=\; \bigcup_{y\in S,\, P\in F} (T+I)^{-1}\,(y+P). \tag{5.1}$$

Let us show why this is true. First, notice that by (ii) of Proposition 5.0.1 we have that $(T + I)$ is surjective. If $x \in \Sigma^k(T + I)$, then $\dim(T + I)(x) \geq k$, therefore there exists a closed and convex set B_k of dimension larger than or equal to k contained in $(T + I)(x)$. Now, there exist $y \in S$ and $P \in F$ such that $B_k \cap (y + P) \neq \emptyset$. Therefore, $(T + I)(x) \cap (y + P) \neq \emptyset$. In particular, we can say that $x \in (T + I)^{-1} \cap (y + P)$.

Since the union in (5.1) is countable, it is enough to show that $(T{+}I)^{-1}(y{+}P)$ is countably $\mathcal{H}^{n-k}$-rectifiable for any $y \in S$ and $P \in F$. This follows from Proposition 5.0.1: we know that $(T + I)^{-1}$ is a Lipschitz function and the image of a $(n-k)$-plane through a Lipschitz function is countably $\mathcal{H}^{n-k}$-rectifiable (one can find the proof of this fact, for example, in [86]). □

As we can see in the proof of Theorem 5.0.2, the main ingredient is the fact that the resolvent $(T + I)^{-1}$ is a Lipschitz continuous function. In the Heisenberg case this is not true in general (see [21] and [13]). Therefore, we are forced to use another strategy.

In the first section, following [13], we give a short introduction to the theory of H-monotone set valued maps and we will recall some important results about upper semicontinuity for such operators. This property, analogous to the Euclidean one, will be central in the proof of the main result of this Chapter (see Theorem 5.2.1 and, also, Theorem 5.2.2).

5.1 Maximal H-monotone maps in $\mathbb{H}^n$

We start by defining the domain and the graph of a set valued map $T : \mathbb{H}^n \rightrightarrows V_1$. We recall that V_1 denotes the horizontal layer of the Lie algebra $\mathfrak{h}^n$ of the Heisenberg group $\mathbb{H}^n$. The space V_1 can be identified with $\mathbb{R}^{2n}$.

Definition 5.1.1. *Let $T : \mathbb{H}^n \rightrightarrows V_1$ be a set valued map. We call the* effective domain *of T, and we write* dom(T)*, the set $\{p \in \mathbb{H}^n \,|\, T(p) \neq \emptyset\}$. Morover, the* graph *of T, denoted by* gr(T)*, is the set $\{(p, v) \in \mathbb{H}^n \times V_1 \,|\, p \in \mathrm{dom}(T),\ v \in T(p)\}$.*

Definition 5.1.2. *Let $T : \mathbb{H}^n \rightrightarrows V_1$ be a set valued map. We say that T is* closed-valued *if, for every $p \in \mathrm{dom}(T)$, $T(p)$ is a closed subset of V_1. Analogously, we say that T is compact-valued if, for every $p \in \mathbb{H}^n$, $T(p)$ is a compact subset of V_1.*

Definition 5.1.3. *Let $T : \mathbb{H}^n \rightrightarrows V_1$ be a set valued map. We say that T is* upper semicontinuous *at a point $p \in \mathbb{H}^n$ if, for every $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$T(p') \subseteq T(p) + B_{\mathbb{R}^{2n}}(0, \varepsilon),$$

for every $p' \in \mathbb{H}^n$ with $d_c(p, p') < \delta$.

Remark 5.1.4. We point out that this Definition is totally equivalent to the one given in (iv) of Proposition 5.0.1. In fact, this notion does not depend on the structure of the Heisenberg group. It is needed just a metric space structure.

Remark 5.1.5. If the map T is compact-valued, the upper semicontinuity property can be restated as follows: if $p_k \longrightarrow p$ and $v_k \in T(p_k)$, then there exists a subsequence $\{v_{n_k}\}_{k\in\mathbb{N}}$ such that $v_{n_k} \longrightarrow v \in T(p)$, as $k \to \infty$.

Definition 5.1.6. *Let $T : \mathbb{H}^n \rightrightarrows V_1$ be a set valued map. We say that T is* closed *if* gr(T) *is a closed subset of $\mathbb{H}^n \times V_1$.*

Next propositions and remarks give some fine properties for set valued maps on $\mathbb{H}^n$. The reader interested to the proofs is referred to the book [4], where these facts are discussed in a very general setting. In fact, all the needed machinery does not depend on the Heisenberg structure.

Proposition 5.1.1. *Let $T : \mathbb{H}^n \rightrightarrows V_1$ be a set valued map. Then the following two statements are true:*

(i) if T is upper semicontinuous and closed-valued, then it is closed;

(ii) if T is closed and $\mathrm{im}(T)$ *is compact, then T is upper semicontinuous.*

Proof. The proof can be done following the line of [4], Theorem 16.12. □

Proposition 5.1.2. *Let $T : \mathbb{H}^n \rightrightarrows V_1$ be a set valued map. If T is compact-valued and upper semicontinuous, then $T(K) \subset V_1$ is compact for every $K \subset \mathbb{H}^n$ compact set.*

Proof. The proof can be done following the line of [4], Lemma 17.8. □

We are now ready to give the definition of H-monotone map in the Heisenberg group. Due to the special structure of $\mathbb{H}^n$, it is natural to introduce a definition dependent on the horizontal planes.

Definition 5.1.7. *Let $T : \mathbb{H}^n \rightrightarrows V_1$ be a set valued map. We say that T is* H-monotone *if*

$$\langle \xi_1(p) - \xi_1(p'), v - v' \rangle \ \geq \ 0,$$

for every p and p' such that $p \in H_{p'}$ and for every $v \in T(p)$ and $v' \in T(p')$.

Remark 5.1.8. Definition 5.1.7 is equivalent to say that

$$\langle \xi_1(p) - \xi_1 \left(p \cdot \exp(tw)\right), v - v' \rangle \ \geq \ 0,$$

for every $p \in \mathbb{H}^n$, for every $v \in T(p)$, for every $t \in \mathbb{R}$, for every $w \in V_1$ and for every $v' \in T(p \cdot \exp(tw))$. This means that we have information about monotonicity of T only in the horizontal directions.

Definition 5.1.9. *We say that T is* maximal H-monotone *if whenever there is a H-monotone map $S : \mathbb{H}^n \rightrightarrows V_1$ such that* $\mathrm{gr}(T) \subseteq \mathrm{gr}(S)$*, then $T = S$.*

Now, we give a look to an important example. Analogously to the definition of the subdifferential of convex functions, one can give the definition of *horizontal gradient* of a *H-convex function*. It turns out that that the horizontal gradient is a H-monotone set valued map.

Example 5.1.1. We consider first a *H-convex set* $\Omega \subseteq \mathbb{H}^n$, i.e. a set such that, for every $p \in \Omega$ and for every $p' \in H_p \cap \Omega$,

$$p \cdot \delta_\lambda(p^{-1} \cdot p') \,\in\, \Omega,$$

for every $\lambda \in [0, 1]$.

Let $u : \Omega \longrightarrow \mathbb{R}$ be a function, with $\Omega \subseteq \mathbb{H}^n$ a H-convex set. We say that u is *H-convex* if

$$u\left(p \cdot \delta_\lambda(p^{-1} \cdot p')\right) \,\leq\, u(p) + \lambda\left(u(p') - u(p)\right),$$

for every $p' \in \Omega \cap H_p$ and $\lambda \in [0, 1]$.

We define the *horizontal gradient* of u at the point $\eta_0 \in \mathbb{H}^n$ the set (possibly empty)

$$\partial_H u(\eta_0) \,=\, \{p \in V_1 \,|\, u(\eta) \geq u(\eta_0) + \langle p, \xi_1(\eta) - \xi_1(\eta_0)\rangle \,, \text{for every}\, \eta \in H_{\eta_0}\}\,.$$

As a characterization, in [20], analogously to what happens in the classical case, the authors show that $\partial_H u(\eta) \neq \emptyset$, for every $\eta \in \mathbb{H}^n$, if and only if u is a H-convex function.

We define now the set valued map

$$\begin{aligned} \partial_H u : \mathbb{H}^n &\longrightarrow V_1 \\ \eta &\longmapsto \partial_H u(\eta). \end{aligned}$$

If u is H-convex, then it follows that $\partial_H u$ is a H-monotone map. Let us prove this fact. Consider η_1 and $\eta_2 \in \mathbb{H}^n$ such that $\eta_2 \in H_{\eta_1}$ and select $p_1 \in \partial_H u(\eta_1)$ and $p_2 \in \partial_H u(\eta_2)$. By definінition of horizontal gradient, the following inequalities are true

$$\begin{aligned} u(\eta) &\geq u(\eta_1) + \langle p_1\,,\, \xi_1(\eta) - \xi_1(\eta_1)\rangle\,, \\ u(\eta) &\geq u(\eta_2) + \langle p_2\,,\, \xi_1(\eta) - \xi_1(\eta_2)\rangle\,, \end{aligned}$$

for every $\eta \in \mathbb{H}^n$. Let us set in the first $\eta = \eta_2$ and in the second $\eta = \eta_1$. Adding together the new inequalities, we can deduce that

$$\langle p_1 \, , \, \xi_1(\eta_2) - \xi_1(\eta_1) \rangle \; \leq \; \langle p_2 \, , \, \xi_1(\eta_2) - \xi_1(\eta_1) \rangle \, ,$$

which clearly gives the definition of H-monotonicity.

Following [13], we give some propositions about maximal H-monotone maps. These results will be used in the study of the dimension of the singular set of an H-monotone map in Section 5.2.

Proposition 5.1.3. *Let $T : \mathbb{H}^n \rightrightarrows V_1$ be a maximal H-monotone map. Then $T(\eta)$ is closed and convex for every $\eta \in \mathbb{H}^n$. Moreover, if $\mathrm{dom}(T) = \mathbb{H}^n$, then T is also compact-valued.*

Proof. See [13], Proposition 2.2. □

From Proposition 5.1.3, together with Proposition 5.1.2, one can prove the following corollary.

Corollary 5.1.4. *Let $T : \mathbb{H}^n \rightrightarrows V_1$ be an upper semicontinuous maximal H-monotone map with $\mathrm{dom}(T) = \mathbb{H}^n$. Then T is closed and $T(K) \subset V_1$ is compact for every $K \subset \mathbb{H}^n$ compact. Moreover, it is locally bounded.*

Corollary 5.1.4 is key to prove the following theorems, useful in what will follow in the next section.

Theorem 5.1.5. *Let $T : \mathbb{H}^n \rightrightarrows V_1$ be a maximal H-monotone map with $\mathrm{dom}(T) = \mathbb{H}^n$. Then T is locally bounded if and only if T is upper semicontinuous.*

Proof. See [13], Theorem 2.2. □

Theorem 5.1.6. *Let $T : \mathbb{H}^n \rightrightarrows V_1$ be a maximal H-monotone map with $\mathrm{dom}(T) = \mathbb{H}^n$. Then T is locally bounded.*

Proof. See [13], Theorem 1.1. □

Remark 5.1.10. We point out that, from Theorem 5.1.5 and Theorem 5.1.6, one can deduce that if T is maximal H-monotone with $\mathrm{dom}(T) = \mathbb{H}^n$ then it is upper semicontinuous.

5.2 Dimension estimate of the singular set of a H-monotone map

In the previous Section we gave the definition of H-monotone set valued map. Now it is time to "measure" how much an H-monotone map can assume set values: we want to study how many points $p \in \mathbb{H}^n$ are mapped to a subset of V_1 of a given size. Using the same terminology introduced for the Euclidean case, we will call the sets of these points *singular sets*. Here it is the definition.

Definition 5.2.1. *We call the k-th* singular set*, or the* singular set of order k*, of a set valued map $T : \mathbb{H}^n \rightrightarrows V_1$ the set*

$$\Sigma^k(T) \;=\; \{p \in \mathbb{H}^n \mid \dim(T(p)) \;=\; k\}\,, \tag{5.2}$$

for $k = 0, 1, ..., 2n$.

Conjecture 5.2.1. *Let $T : \mathbb{H}^n \rightrightarrows V_1$ be a maximal H-monotone operator with* $\mathrm{dom}(T) = \mathbb{H}^n$*. Then, for every $0 \leq k \leq 2n$, the Hausdorff dimension of $\Sigma^k(T)$ is smaller than or equal to $2n + 2 - k$, i.e.*

$$\dim_H\left(\Sigma^k(T)\right) \;\leq\; 2n + 2 - k. \tag{5.3}$$

Our goal is to approach Conjecture 5.2.1. For the moment, we are able to give the proof of a preliminary result. Taking into account the peculiarities of the Heisenberg group, we study the size just of some particular subsets of $\Sigma^k(T)$.

First, we need some preparatory definitions and lemmas.

Definition 5.2.2. *Let $k \in \mathbb{N}$. If $v_1, ..., v_k$ are k linearly independent unit vectors of $\mathbb{R}^{2n}$ and $r > 0$, we call k-*simplex*, denoted by $\Delta\,(v_1, ..., v_k, r)$, the set*

$$\Delta\,(v_1, ..., v_k, r) \;:=\; \left\{ \sum_{i=1}^{k} \lambda_i v_i \in \mathbb{R}^{2n} \;\middle|\; \sum_{i=1}^{k} |\lambda_i| \leq r\,,\; \lambda_i \in \mathbb{R} \right\}.$$

Definition 5.2.3. *Let $T : \mathbb{H}^n \rightrightarrows V_1$ be a maximal H-monotone set valued map and let $p \in \mathbb{H}^n$ be fixed. If $\dim(T(p)) = k$, then there exist $\eta \in \mathbb{R}^{2n}$, $r_1 > 0$ and $v_1, ..., v_k \in \mathbb{R}^{2n}$, linearly independent unit vectors, such that*

$$\eta + \Delta\,(v_1, ..., v_k, r_1) \;\subset\; T(p). \tag{5.4}$$

We define the k-dimensional vector space $V_p = \mathrm{span}\{v_1, ..., v_k\} \subseteq \mathbb{R}^{2n}$.

Remark 5.2.4. We underline that inclusion (5.4) is true because $T(p)$ is convex, for every $p \in \mathbb{H}^n$. This was the statement of Proposition 5.1.3.

Lemma 5.2.1. *With the same assumptions and notations introduced in Definition 5.2.3, V_p does not depend on the choice of η, r_1 and $v_1, ..., v_k$.*

Proof. We argue by contradiction. Let $\eta \in \mathbb{R}^{2n}$, $r_1 > 0$ and let $v_1, ..., v_k \in \mathbb{R}^{2n}$ be linearly independent unit vectors such that

$$\eta + \Delta(v_1, ..., v_k, r_1) \subset T(p).$$

Let $V_p = \mathrm{span}\{v_1, ..., v_k\}$. We assume also that there exist $\xi \in \mathbb{R}^{2n}$, $r_2 > 0$, and $w_1, ..., w_k \in \mathbb{R}^{2n}$, linearly independent unit vectors, such that

$$\xi + \Delta(w_1, ..., w_k, r_2) \subset T(p).$$

Let $W_p := \mathrm{span}\{w_1, ..., w_k\}$ and assume that $W_p \neq V_p$. Since W_p and V_p are two k-dimensional vector subspaces of $\mathbb{R}^{2n}$, it holds that there is at least a vector $w \in W$ such that $w, v_1, ..., v_k$ are linearly independent. Without loss of generality, we can assume that $w = w_1$.

Now, since $\xi + \Delta(w_1, ..., w_k, r_2) \subset T(p)$, it holds that for every $\lambda \in \mathbb{R}$, with $|\lambda| < r_2$,

$$\xi + \lambda w_1 \in T(p).$$

Moreover, there exists $\tilde{\lambda} \in \mathbb{R}$, with $\left|\tilde{\lambda}\right| < \frac{r_2}{2}$, such that $\xi + \tilde{\lambda} w_1 \notin \eta + V_p$. Let us prove why this is true. Assume, by contradiction, that for every $\tilde{\lambda} \in \mathbb{R}$, with $\left|\tilde{\lambda}\right| < \frac{r_2}{2}$, it holds that $\xi + \tilde{\lambda} w_1 \in \eta + V_p$. Select $\tilde{\lambda}_1$ and $\tilde{\lambda}_2 \in \mathbb{R}$, with $|\tilde{\lambda}_1|, |\tilde{\lambda}_2| < \frac{r_2}{2}$ and $\tilde{\lambda}_1 \neq \tilde{\lambda}_2$, such that

$$\xi + \tilde{\lambda}_1 w_1 \in \eta + V_p \quad \text{and} \quad \xi + \tilde{\lambda}_2 w_2 \in \eta + V_p.$$

This is equivalent to say that we can find $\alpha_1, ..., \alpha_k \in \mathbb{R}$ and $\alpha'_1, ..., \alpha'_k \in \mathbb{R}$ such that

$$\tilde{\lambda}_1 w_1 = (\eta - \xi) + \sum_{i=1}^{k} \alpha_i v_i \quad \text{and} \quad \tilde{\lambda}_2 w_1 = (\eta - \xi) + \sum_{i=1}^{k} \alpha'_i v_i.$$

Subtracting the second equation to the first one, we get

$$(\lambda_1 - \lambda_2)\, w_1 \;=\; \sum_{i=1}^{k} (\alpha_i - \alpha_i')\, v_i.$$

Therefore, $w_1 \in V_p$, which is clearly not true.

The desired contradiction will be achieved in showing that $\dim T(p) \geq k+1$. To do that, we shall find a $(k+1)$-dimensional simplex included in $T(p)$.

Now, consider $\tilde{\eta} = \frac{1}{2}\eta + \frac{1}{2}\left(\xi + \tilde{\lambda} w_1\right)$ and let $r_3 = \min\left\{\frac{r_1}{4}, \frac{r_2}{4}\right\}$. If we show that

$$\tilde{\eta} + \Delta(v_1, ..., v_k, w_1, r_3) \;\subset\; T(p), \tag{5.5}$$

then we have the desired contradiction.

Let us prove inclusion (5.5). Consider an arbitrary point

$$\tilde{\eta} + \sum_{i=1}^{k} \lambda_i v_i + \lambda_{k+1} w_1 \in \tilde{\eta} + \Delta(v_1, ..., v_k, w_1, r_3),$$

where $\sum_{i=1}^{k+1} |\lambda_i| < r_3$. We perform some calculations:

$$\begin{aligned}
\tilde{\eta} + \sum_{i=1}^{k} \lambda_i v_i + \lambda_{k+1} w_1 \;&=\; \frac{1}{2}\eta + \frac{1}{2}\left(\xi + \tilde{\lambda} w_1\right) + \sum_{i=1}^{k} \lambda_i v_i + \lambda_{k+1} w_1 \\
&=\; \frac{1}{2}\left(\eta + \sum_{i=1}^{k} 2\lambda_i v_i\right) + \frac{1}{2}\left(\xi + \left(\tilde{\lambda} + 2\lambda_{k+1}\right) w_1\right) \\
&:=\; \frac{1}{2}\overline{\eta} + \frac{1}{2}\overline{\xi}.
\end{aligned}$$

Let us show that $\overline{\eta}$ and $\overline{\xi} \in T(p)$. By definition of k-simplex, $\overline{\eta} \in \eta + \Delta(v_1, ..., v_k, r_1) \subset T(p)$ if and only if $\sum_{i=1}^{k} |2\lambda_i| < r_1$. This is true, indeed

$$\sum_{i=1}^{k} |2\lambda_i| \;=\; 2\sum_{i=1}^{k} |\lambda_i| \;\leq\; 2\sum_{i=1}^{k+1} |\lambda_i| \;<\; 2r_3 \;\leq\; r_1.$$

We prove now that $\overline{\xi} \in T(p)$. We need to check that $\left|\tilde{\lambda} + 2\lambda_{k+1}\right| < r_2$. This condition is satisfied, indeed

$$\left|\tilde{\lambda} + 2\lambda_{k+1}\right| \;\leq\; \left|\tilde{\lambda}\right| + 2\left|\lambda_{k+1}\right| \;<\; \frac{r_2}{2} + 2r_3 \;<\; r_2.$$

Now, since $T(p)$ is convex, it holds that

$$\tilde{\eta} + \sum_{i=1}^{k} \lambda_i v_i + \lambda_{k+1} w_1 \;=\; \frac{1}{2}\overline{\eta} + \frac{1}{2}\overline{\xi} \;\in\; T(p),$$

and this concludes the proof. □

We give now the definition of *horizontal k-th singular set*. Roughly speaking, this is the set of those points in the domain of a maximal H-monotone set valued map which are mapped to a set included in a k-dimensional horizontal subgroup of $\mathbb{H}^n$.

Definition 5.2.5. *We define, for $1 \leq k \leq n$, the* horizontal k-th singular set *of T the set*

$$\Sigma^k_H(T) \;=\; \left\{p \in \Sigma^k(T) \,|\, V_p \in G^H(2n, k)\right\}.$$

Remark 5.2.6. We point out that it is very important to assume $k \leq n$, because horizontal subgroups of $\mathbb{H}^n$ can have dimension at most n (see Proposition 4.2.3).

Remark 5.2.7. We epoint out that, in general, $\Sigma^k_H(T) \neq \Sigma^k(T)$. Let us give an example which shows this fact. Consider the Euclidean convex function (which is also H-convex) $u : \mathbb{H}^2 \longrightarrow \mathbb{R}^4$ given by $u(x_1, x_2, y_1, y_2, t) \;=\; (x_1^2 + y_1^2)^{\frac{1}{2}}$ and compute its horizontal subdifferential at a point $p = (x_1, x_2, y_1, y_2, t) \in \mathbb{H}^2$:

$$\partial_H u(p) = \begin{cases} B_1, & \text{if } p = (0, x_2, 0, y_2, t), \\ \frac{1}{\left(x_1^2+y_1^2\right)^{\frac{1}{2}}} \left(x_1, 0, y_1, 0\right), & \text{if } p \neq (0, x_2, 0, y_2, t), \end{cases}$$

where $B_1 \;=\; \{(x_1, 0, y_1, 0) \in \mathbb{R}^4 \,|\, x_1^2 + y_1^2 \leq 1\}$. Since $\dim B_1 \;=\; 2$, it holds that $\Sigma^2(\partial_H u) = \{(0, x_2, 0, y_2, t) \in \mathbb{H}^2 \,|\, x_2, y_2, t \in \mathbb{R}\}$, which has Hausdorff dimension 4. On the other hand, we notice that there is no horizontal subgroup of $\mathbb{H}^2$ which contains B_1. Therefore, $\Sigma^2_H(\partial_H u) = \emptyset$.

We are now ready to give the statement of the main result of the second part of the dissertation.

Theorem 5.2.2. *Let $T : \mathbb{H}^n \rightrightarrows V_1$ be a maximal H-monotone set valued map with* $\mathrm{dom}(T) = \mathbb{H}^n$. *If $1 \le k \le n$, then*

$$\dim_H \Sigma_H^k(T) \;\le\; 2n + 2 - k.$$

Remark 5.2.8. If $k = 1$, then $\Sigma^1(T) = \Sigma_H^1(T)$. This is true because a 1-dimensional vector subspace of $\mathbb{R}^{2n}$ can be always seen as a 1-dimensional horizontal subgroup of $\mathbb{H}^n$. Therefore, thanks to Theorem 5.2.2, Conjecture 5.2.1 is proved for this case.

Before giving the proof of Theorem 5.2.2, we need a couple of preliminary results. The first result provides a condition to control the Hausdorff dimension of a given set. The main idea, which passes through the compactness of intrinsic Grassmannians, is taken from [6].

Theorem 5.2.3. *Let $E \subset \mathbb{H}^n$ be a Borel set such that, for every $p \in E$, there exist $r_p > 0$, $\alpha_p > 0$ and $\mathbb{W}_p \in \mathcal{G}(\mathbb{H}^n, d)$, with $n + 1 \le d \le 2n$, such that*

$$E \cap B(p, r_p) \;\subset\; C_{\mathbb{W}_p^\perp, \mathbb{W}_p}(p, \alpha_p). \tag{5.6}$$

Then E has Hausdorff dimension smaller than or equal to $d + 1$.

Remark 5.2.9. We recall that if $\mathbb{W}_p \in \mathcal{G}(\mathbb{H}^n, d)$, then $\dim_H \mathbb{W}_p = d + 1$.

Proof of Theorem 5.2.3. First of all, for every $k \in \mathbb{N} \cup \{0\}$, we define the set

$$E^k \;:=\; \{p \in E \,|\, k < \alpha_p \le k + 1\}\,.$$

It is clear that

$$E \;=\; \bigcup_{k \in \mathbb{N}} E^k. \tag{5.7}$$

We point out that the union in (5.7) is countable. Therefore, if we prove that $\dim_H E^k \le d + 1$, for every $k \in \mathbb{N}$, the assertion will follow.

Let us fix $k \in \mathbb{N}$ and consider E^k. For simplicity of notations, in what follows, we skip the superscript k.

We define $\alpha := \sup_{p \in E} \alpha_p$. We highlight that α is finite. This is true because of the decomposition in (5.7). Let $\delta > 0$ be such that $\delta < \frac{1}{2(1+\alpha)}$.

By Proposition 4.3.2, we know that $(\mathcal{G}(\mathbb{H}^n, d), \rho)$ is compact. This implies that there exists a set $\mathcal{G} := \{\mathbb{W}_1, ..., \mathbb{W}_N\} \subset \mathcal{G}(\mathbb{H}^n, d)$ such that, for every $\mathbb{W} \in \mathcal{G}(\mathbb{H}^n, d)$, there is $\mathbb{W}_i \in \mathcal{G}$ such that

$$\rho(\mathbb{W}_i, \mathbb{W}) \, < \, \delta.$$

This observation allows us to split the set E in the following way:

$$E \, = \, \bigcup_{i=1}^{N} E_i, \tag{5.8}$$

where $E_i = \{p \in E \,|\, \rho(\mathbb{W}_p, \mathbb{W}_i) < \delta\}$. Since the union in (5.8) is finite, our assertion follows if we show that $\dim_H E_i \leq d + 1$.

Let us fix a point $p \in E_i$. By Lemma 4.4.2, we have that

$$C_{\mathbb{W}_p^\perp, \mathbb{W}_p}(p, \alpha) \, \subset \, C_{\mathbb{W}_i^\perp, \mathbb{W}_i}(p, 2(1 + \alpha)).$$

Consequently, keeping in mind the hypothesis, one has that

$$E_i \cap B(p, r_p) \, \subset \, C_{\mathbb{W}_i^\perp, \mathbb{W}_i}(p, 2(1 + \alpha)).$$

Now, by Remark 4.4.5, it holds that there exists $\beta > 0$ such that

$$E_i \cap B(p, r_p) \, \cap \, C_{\mathbb{W}_i, \mathbb{W}_i^\perp}(p, \beta) \, = \, \{p\}. \tag{5.9}$$

Our goal is to apply Proposition 4.4.7, stated in Section 4.4 and show that E_i is contained in the graph of an intrinsic Lipschitz function $\varphi : F \subseteq \mathbb{W}_i \longrightarrow \mathbb{W}_i^\perp$.

We have now a claim.

Claim. For every p_1 and $p_2 \in E_i$ such that $p_1 \neq p_2$, it holds that $P_{\mathbb{W}_i}(p_1) \neq P_{\mathbb{W}_i}(p_2)$.

The proof of this claim is by contradiction. Let p_1 and $p_2 \in E_i$, sufficiently near to each other, be such that $p_1 \neq p_2$ and $P_{\mathbb{W}_i}(p_1) = P_{\mathbb{W}_i}(p_2)$. We will provide the desired contradiction by showing that $p_2 \in C_{\mathbb{W}_i, \mathbb{W}_i^\perp}(p_1, \beta)$. This can not be possible, because of (5.9). Since $P_{\mathbb{W}_i}(p_1) = P_{\mathbb{W}_i}(p_2) =: q$ and $\mathbb{H}^n = \mathbb{W}_i \cdot \mathbb{W}_i^\perp$, there exist q_1 and $q_2 \in \mathbb{W}_i^\perp$ such that $p_1 = q \cdot q_1$ and $p_2 = q \cdot q_2$. From these relations, we can write

$$p_1^{-1} \cdot p_2 \, = \, q_1^{-1} \cdot q^{-1} \cdot q \cdot q_2 \, = \, q_1^{-1} \cdot q_2.$$

Since $q_1^{-1} \cdot q_2 \in \mathbb{W}_i^\perp$, it holds that $P_{\mathbb{W}_i}(p_1^{-1} \cdot p_2) = e$.

Now, by definition of intrinsic cone, $p_2 \in C_{\mathbb{W}_i,\mathbb{W}_i^\perp}(p_1, \beta)$ if and only if $p_1^{-1} \cdot p_2 \in C_{\mathbb{W}_i,\mathbb{W}_i^\perp}(e, \beta)$, more specifically, if and only if

$$\|P_{\mathbb{W}_i}(p_1^{-1} \cdot p_2)\| \,\leq\, \beta\|P_{\mathbb{W}_i^\perp}(p_1^{-1} \cdot p_2)\|. \tag{5.10}$$

Since the left-hand side is zero, the inequality (5.10) is true. We have the desired contradiction and the claim is proved.

Let us set $F_i := P_{\mathbb{W}_i}(E) \subset \mathbb{W}_i$. The previous Claim implies that the map $P_{\mathbb{W}_i}|_{E_i} : E_i \longrightarrow F_i$ is injective. Therefore, it is well defined the map $\varphi = (P_{\mathbb{W}_i}|_{E_i})^{-1}$, whose graph is $F_i \times E_i$.

Keeping in mind the definition of intrinsic Lipschitz function, since $E_i = \text{graph}(\varphi)$, it holds that, locally,

$$\text{graph}(\varphi) \cap C_{\mathbb{W}_i,\mathbb{W}_i^\perp}(p, \beta) \,=\, \{p\},$$

for every $p \in \text{graph}(\varphi)$. Therefore, $\varphi : F_i \subset \mathbb{W}_i \longrightarrow \mathbb{W}_i^\perp$ is an intrinsic Lipschitz function.

The proof of the theorem is to the end; indeed we finally call into play Proposition 4.4.7 and we can conclude that $\dim_H(E_i) \leq \dim_H(\text{graph}(\varphi)) \leq d+1$. $\square$

Lemma 5.2.2. *Let $T : \mathbb{H}^n \rightrightarrows V_1$ be a maximal H-monotone set valued map with* $\text{dom}(T) = \mathbb{H}^n$. *Let $h \in \mathbb{N}$ and define*

$$\Sigma_H^{k,h}(T) \,:=\, \left\{ p \in \Sigma_H^k(T) \;\middle|\; \exists\, \eta + \Delta\left(v_1, ..., v_k, \frac{1}{h}\right) \subset T(p) \right\}. \tag{5.11}$$

Let $p_0 \in \Sigma_H^{k,h}(T)$ and let $\{p_m\}_{m\in\mathbb{N}} \subset \Sigma_H^{k,h}(T)$ be such that $p_m \longrightarrow p_0$, as $m \to \infty$. Then there exist a subsequence of $\{p_m\}_{m\in\mathbb{N}}$, denoted in the same way, $\eta_0 \in \mathbb{R}^{2n}$ and $v_0^1, ..., v_0^k$ linearly independent unit vectors such that:

- *(i)* $\text{span}\{v_0^1, ..., v_0^k\} \in G^H(2n, k)$*;*
- *(ii)* $\eta_0 + \Delta\left(v_0^1, ..., v_0^k, \frac{1}{h}\right) \subset T(p_0)$*;*
- *(iii) for each $w_0 \in \eta_0 + \Delta\left(v_0^1, ..., v_0^k, \frac{1}{h}\right)$, there exists $w_m \in T(p_m)$, for every $m \in \mathbb{N}$, such that $w_m \longrightarrow w_0$, as $m \to \infty$.*

Proof. Let us select $\{v_m^1, ..., v_m^k\}$ an orthonormal basis of V_{p_m}, for every $m \in \mathbb{N}$. Fix $i \in \{1, ..., k\}$. It holds that $\|v_m^i\|_{\mathbb{R}^{2n}} = 1$, for every $m \in \mathbb{N}$. Therefore, up to a subsequence, there exists $v_0^i \in \mathbb{R}^{2n}$ such that $v_m^i \longrightarrow v_0^i$, as $m \to \infty$.

We prove that $v_0^1, ..., v_0^k$ are linearly independent unit vectors. Let i and $j \in \{1, ..., k\}$. We know that $\langle v_m^i, v_m^j \rangle = \delta_{ij}$[1], for every $m \in \mathbb{N}$. Passing to the limit for $m \to \infty$, we conclude that $\left\langle v_0^i, v_0^j \right\rangle = \delta_{ij}$, for every i and $j \in \{1, ..., k\}$. A second step is to prove that $\text{span}\{v_0^1, ..., v_0^k\} \in G^H(2n, k)$. Since $\text{span}\{v_m^1, ..., v_m^k\} \in G^H(2n, k)$, it holds that $\langle v_m^i, v_m^j \rangle = 0$, for every $m \in \mathbb{N}$ and for every $i \neq j$. Again, if $m \to \infty$, we obtain $\left\langle v_0^i, v_0^j \right\rangle = 0$, for every $i \neq j$.

Now, since $p_m \in \Sigma_H^{h,k}(T)$, there exists $\eta_m \in T(p_m)$ such that

$$\eta_m + \Delta\left(v_m^1, ..., v_m^k, \frac{1}{h}\right) \subset T(p_m).$$

By hypothesis, we know that T is a maximal H-monotone set valued map. Moreover, $p_m \longrightarrow p_0$, as $m \to \infty$. Therefore, we can apply Remark 5.1.5 and Proposition 5.1.3. It follows that there exists $\eta_0 \in T(p_0)$ such that, eventually restricting to a subsequence, $\eta_m \longrightarrow \eta_0$, as $m \to \infty$.

It remains to show that

$$\eta_0 + \Delta\left(v_0^1, ..., v_0^k, \frac{1}{h}\right) \subset T(p_0).$$

We argue by contradiction. Let $\eta_0 + \sum_{i=1}^k \tilde{\lambda}_i v_0^i \notin T(p_0)$, for some $\tilde{\lambda}_1, ..., \tilde{\lambda}_k \in \mathbb{R}$ such that $\sum_{i=1}^k |\tilde{\lambda}_i| < \frac{1}{h}$. By definition of $\Sigma_H^{h,k}(T)$, it holds that

$$w_m := \eta_m + \sum_{i=1}^k \tilde{\lambda}_i v_m^i \in T(p_m).$$

Again by upper semicontinuity of the set valued map T, it holds that there exists $w_0 \in T(p_0)$ such that $w_m \longrightarrow w_0$, as $m \to \infty$.

[1]With the simbol δ_{ij} we denote the *Kronecker's delta*

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

On the other hand, one has that

$$w_m = \eta_m + \sum_{i=1}^{k} \tilde{\lambda}_i v_m^i \longrightarrow \eta_0 + \sum_{i=1}^{k} \tilde{\lambda}_i v_0^i,$$

as $m \to \infty$. Hence, by uniqueness of the limit, we can conclude that

$$w_0 = \eta_0 + \sum_{i=1}^{k} \tilde{\lambda}_i v_0^i \in T(p_0),$$

which provides the desired contradiction. □

Proof of Theorem 5.2.2. First of all, we consider the set defined in (5.11), for every $h \in \mathbb{N}$. It holds that

$$\Sigma_H^k(T) = \bigcup_{h \in \mathbb{N}} \Sigma_H^{k,h}(T). \tag{5.12}$$

Let us check why this is true. First, by definition, $\Sigma_H^{k,h}(T) \subset \Sigma_H^k(T)$, for every $h \in \mathbb{N}$.

For the opposite inclusion, let $p \in \Sigma_H^k(T)$. We want to show that there exists $h \in \mathbb{N}$ such that $p \in \Sigma_H^{k,h}(T)$. We know that $\dim T(p) = k$ and that $T(p)$ is a convex set. This implies that there exists a k-dimensional ball B contained in $T(p)$. Let $r > 0$ be the radius of B and η its center. There exists $\tilde{h} \in \mathbb{N}$ such that $\frac{1}{\tilde{h}} \leq r$.

Let $v_1, ..., v_k$ be linearly independent unit vectors such that $V_p = \text{span}\{v_1, ..., v_k\}$. It follows that $\eta_p + \Delta\left(v_1, ..., v_k, \frac{1}{\tilde{h}}\right) \subset B \subset T(p)$. Therefore $p \in \Sigma_H^{k,\tilde{h}}(T)$.

Since the union in (5.12) is countable, we can reduce our investigation to $\Sigma_H^{k,h}(T)$, for every $h \in \mathbb{N}$. Let us choose an arbitrary $h \in \mathbb{N}$. We aim to prove that

$$\dim \Sigma_H^{k,h}(T) \leq 2n + 2 - k.$$

Our strategy is to apply Theorem 5.2.3: we need to show that, for every $p \in \Sigma_H^{k,h}(T)$, there exist a radius $r_p > 0$, an opening $\alpha_p > 0$ and a vertical homogeneous subgroup $\mathbb{W}_p \in \mathcal{G}(\mathbb{H}^n, 2n + 1 - k)$ such that

$$\Sigma_H^{k,h}(T) \cap B(p, r_p) \subset C_{\mathbb{V}_p, \mathbb{W}_p}(p, \alpha_p),$$

where $\mathbb{V}_p$ is the unique horizontal subgroup of $\mathbb{H}^n$ of dimension k such that $\mathbb{H}^n = \mathbb{W}_p \cdot \mathbb{V}_p$ is a decomposition in complementary subgroups.

Without loss of generality, we can assume that $0 \in \Sigma_H^{k,h}(T)$ and we restrict our investigation to 0.

Since $0 \in \Sigma_H^{k,h}(T)$, there exists $V_0 \in G^H(2n, k)$ defined as in Definition 5.2.3. Let us call it $V_0 = \mathbb{V}_0$. This is equivalent to say that there exist a point $\eta_0 \in T(0)$ and a k-simplex $\Delta\left(v_0^1, ..., v_0^k, \frac{1}{h}\right)$ such that $\eta_0 + \Delta\left(v_0^1, ..., v_0^k, \frac{1}{h}\right) \subset T(0)$. This fact implies that there is a k-dimensional disk $D_0 \subset \mathbb{V}_0$, centered at zero, such that $\eta_0 + D_0 \subset T(0)$.

Let $\mathbb{W}_0$ be the vertical subgroup of $\mathbb{H}^n$ of linear dimension $2n+1-k$ such that $\mathbb{H}^n = \mathbb{W}_0 \cdot \mathbb{V}_0$ is a decomposition in complementary subgroups.

For simplicity of calculations, we can assume that $\mathbb{V}_0 = \{(v_1, ..., v_k, 0, ..., 0) \in \mathbb{H}^n \,|\, v_i \in \mathbb{R},\, i = 1, ..., k\}$ and $\mathbb{W}_0 = \{(0, ..., 0, w_{k+1}, ..., w_{2n}, t) \in \mathbb{H}^n \,|\, w_j \in \mathbb{R}, t \in \mathbb{R},\, j = k+1, ..., 2n\}$.

We have now the following:

Claim. There exist $r_0 > 0$ and $\beta_0 > 0$ such that

$$\Sigma_H^{k,h}(T) \cap B(0, r) \cap C_{\mathbb{W}_0, \mathbb{V}_0}(0, \beta) \;=\; \{0\},$$

for every $0 < r < r_0$ and $0 < \beta < \beta_0$.

The proof of the Claim is by contradiction. Let us assume that for every $\beta > 0$ and for every $r > 0$ there exists $p \neq 0$ such that

$$p \;\in\; \Sigma_H^{k,h}(T) \cap B(0, r) \cap C_{\mathbb{W}_0, \mathbb{V}_0}(0, \beta).$$

Let us choose $r = \frac{1}{m}$ and $\beta = \frac{1}{m}$, for $m \in \mathbb{N}$. Then, for every $m \in \mathbb{N}$, there exists $p^{(m)} \neq 0$ such that

$$p^{(m)} \;\in\; \Sigma_H^{k,h}(T) \cap B\left(0, \frac{1}{m}\right) \cap C_{\mathbb{W}_0, \mathbb{V}_0}\left(0, \frac{1}{m}\right).$$

It is clear that

$$\lim_{m \to \infty} p^{(m)} \;=\; 0.$$

Moreover, since $p^{(m)} \in \Sigma_H^{k,h}(T)$, there exists $\mathbb{V}_{p^{(m)}} \in G^H(2n,k)$, here denoted by $\mathbb{V}_m$, defined as in Definition 5.2.3.

Now, let us choose arbitrarily a vector $w_0 \in \eta_0 + D_0$. Let $v_0 \in D_0$ be such that $w_0 = \eta_0 + v_0$. Thanks to Lemma 5.2.2, we can select a subsequence of $\left\{p^{(m)}\right\}_{m\in\mathbb{N}}$, still denoted in the same way, such that we can find $w_m \in T\left(p^{(m)}\right)$, for every $m \in \mathbb{N}$, in order to have $w_m \longrightarrow w_0$, as $m \to \infty$.

By definition of intrinsic cone, we have that

$$\begin{aligned}C_{\mathbb{W}_0,\mathbb{V}_0}\left(0,\frac{1}{m}\right) = \{p = (p_1,...,p_{2n},p_{2n+1}) \in \mathbb{H}^n |\\ (p_{k+1}^2 + ... + p_{2n}^2)^2 + \left(p_{2n+1} - 2\sum_{i=1}^{k} p_i \cdot p_{n+i}\right)^2 \leq \frac{1}{m^4}(p_1^2 + ... + p_k^2)^2\right\}.\end{aligned} \tag{5.13}$$

Therefore, since $p^{(m)} \in C_{\mathbb{W}_0,\mathbb{V}_0}\left(0,\frac{1}{m}\right)$, one can deduce that

$$\left(\left(p_{k+1}^{(m)}\right)^2 + ... + \left(p_{2n}^{(m)}\right)^2\right)^2 \leq \frac{1}{m^4}\left(\left(p_1^{(m)}\right)^2 + ... + \left(p_k^{(m)}\right)^2\right)^2.$$

This implies that there exists $\alpha = (\alpha_1,...,\alpha_k,0,...,0)$, with $\|\alpha\|_{\mathbb{R}^{2n}} = 1$ for which there is an index $i \in \{1,...,k\}$ such that $\alpha_i \neq 0$ and

$$\lim_{m\to\infty} \frac{\xi_1\left(p^{(m)}\right)}{\|\xi_1\left(p^{(m)}\right)\|_{\mathbb{R}^{2n}}} = \alpha.$$

More specifically, we have that

$$\begin{aligned}\lim_{m\to\infty} \frac{p_i^{(m)}}{\|\xi_1\left(p^{(m)}\right)\|_{\mathbb{R}^{2n}}} &= \alpha_i, \quad \text{for } i = 1,...,k,\\ \lim_{m\to\infty} \frac{p_j^{(m)}}{\|\xi_1\left(p^{(m)}\right)\|_{\mathbb{R}^{2n}}} &= 0, \quad \text{for } j = k+1,...,2n.\end{aligned}$$

Without loss of generality, we can assume that $\alpha_i \neq 0$. Therefore, eventually restricting to a subsequence, $p^{(m)} \neq 0$, for every $m \in \mathbb{N}$.

Consider now $\tilde{p}^{(m)} \in H_{p^{(m)}} \cap H_0$, for every $m \in \mathbb{N}$. By definition of horizontal plane at the point $p^{(m)} = \left(p_1^{(m)},...,p_{2n}^{(m)},p_{2n+1}^{(m)}\right) \in \mathbb{H}^n$, there exists $(u_1,...,u_{2n},0) \in$

$\mathbb{H}^n$ such that

$$\begin{aligned}\tilde{p}^{(m)} &= \left(\tilde{p}_1^{(m)}, ..., \tilde{p}_{2n}^{(m)}, \tilde{p}_{2n+1}^{(m)}\right) \\ &= \left(p_1^{(m)} + u_1, ..., p_{2n}^{(m)} + u_{2n}, p_{2n+1}^{(m)} + 2\sum_{i=1}^{n} p_i^{(m)} u_{n+i} - 2\sum_{j=1}^{n} p_{n+j}^{(m)} u_j\right).\end{aligned}$$

Since we asked for $\tilde{p}^{(m)} \in H_0$, in order to satisfy condition

$$p_{2n+1}^{(m)} + 2\sum_{i=1}^{n} p_i^{(m)} u_{n+i} - 2\sum_{j=1}^{n} p_{n+j}^{(m)} u_j = 0,$$

we choose

$$u_i = \begin{cases} -\frac{p_i^{(m)}}{2}, & \text{if } i \neq n+1,\ i = 1, ..., 2n, \\ -\frac{p_{2n+1}^{(m)}}{2p_1^{(m)}} - \frac{p_{n+1}^{(m)}}{2}, & \text{if } i = n+1. \end{cases} \tag{5.14}$$

Consider now $\tilde{w}^{(m)} \in T\left(\tilde{p}^{(m)}\right)$. Since $\tilde{p}^{(m)} \longrightarrow 0$, as $m \to \infty$, by upper semicontinuity of T, eventually restricting to a subsequence, there exists $\tilde{w}_0 \in T(0)$, such that $\tilde{w}^{(m)} \longrightarrow \tilde{w}_0$, as $m \to \infty$. By convexity of $T(0)$, there exists $\tilde{v}_0 \in \mathbb{V}_0$ such that $\tilde{w}_0 = \eta_0 + \tilde{v}_0$.

Now, we use the definition of H-monotonicity of the set valued map T, evaluated at the points $\tilde{p}^{(m)}$ and 0. It holds that

$$\left\langle \xi_1\left(\tilde{p}^{(m)}\right),\ \tilde{w}^{(m)} - w_0 \right\rangle \geq 0.$$

Writing explicity in coordinates, we have

$$\left\langle \left(\frac{p_1^{(m)}}{2}, ..., \frac{p_n^{(m)}}{2}, -\frac{p_{2n+1}^{(m)}}{2p_1^{(m)}} + \frac{p_{n+1}^{(m)}}{2}, \frac{p_{n+2}^{(m)}}{2}, ..., \frac{p_{2n}^{(m)}}{2},\right), \tilde{w}^{(m)} - w_0 \right\rangle \geq 0.$$

Dividing by $\left\|\xi_1\left(p^{(m)}\right)\right\|_{\mathbb{R}^{2n}}$, one has

$$\frac{1}{2\left\|\xi_1\left(p^{(m)}\right)\right\|_{\mathbb{R}^{2n}}} \left\langle \left(p_1^{(m)}, ..., p_n^{(m)}, -\frac{p_{2n+1}^{(m)}}{p_1^{(m)}} + p_{n+1}^{(m)}, p_{n+2}^{(m)}, ..., p_{2n}^{(m)}\right), \tilde{w}^{(m)} - w_0 \right\rangle \geq 0.$$

The task now is to compute the limit for $m \longrightarrow \infty$. We already know that

$$\begin{aligned}\lim_{m\to\infty} \frac{p_i^{(m)}}{\left\|\xi_1\left(p^{(m)}\right)\right\|_{\mathbb{R}^{2n}}} &= \alpha_i, \quad \text{for } i = 1, ..., k, \\ \lim_{m\to\infty} \frac{p_j^{(m)}}{\left\|\xi_1\left(p^{(m)}\right)\right\|_{\mathbb{R}^{2n}}} &= 0, \quad \text{for } j = k+1, ..., 2n.\end{aligned} \tag{5.15}$$

Hence, it remains to show that

$$\lim_{m\to\infty} \frac{p_{2n+1}^{(m)}}{\|\xi_1(p^{(m)})\|_{\mathbb{R}^{2n}}\, p_1^{(m)}} = 0.$$

By definition of intrinsic cone (see (5.13)), it holds that

$$\left| p_{2n+1}^{(m)} - 2\sum_{i=1}^{k} p_1^{(m)} \cdot p_{n+i}^{(m)} \right| \le \frac{1}{m^2}\left(\left(p_1^{(m)}\right)^2 + ... + \left(p_k^{(m)}\right)^2 \right).$$

Therefore, we can estimate

$$\begin{aligned}
\frac{p_{2n+1}^{(m)}}{\|\xi_1(p^{(m)})\|_{\mathbb{R}^{2n}}\, p_1^{(m)}} &\le \frac{1}{m^2}\frac{\left(\left(p_1^{(m)}\right)^2 + ... + \left(p_k^{(m)}\right)^2 \right)}{\|\xi_1(p^{(m)})\|_{\mathbb{R}^{2n}}\, p_1^{(m)}} + \frac{2\sum_{i=1}^{k} p_1^{(m)} \cdot p_{n+i}^{(m)}}{\|\xi_1(p^{(m)})\|_{\mathbb{R}^{2n}}\, p_1^{(m)}} \\
&= \frac{1}{m^2}\frac{\left\|\xi_1\left(p^{(m)}\right)\right\|_{\mathbb{R}^{2n}}^2}{\|\xi_1(p^{(m)})\|_{\mathbb{R}^{2n}}\, p_1^{(m)}} + \frac{2\sum_{i=1}^{k} p_1^{(m)} \cdot p_{n+i}^{(m)}}{\|\xi_1(p^{(m)})\|_{\mathbb{R}^{2n}}\, p_1^{(m)}} \frac{\left\|\xi_1\left(p^{(m)}\right)\right\|_{\mathbb{R}^{2n}}}{\left\|\xi_1\left(p^{(m)}\right)\right\|_{\mathbb{R}^{2n}}} \\
&= \frac{1}{m^2}\frac{\left\|\xi_1\left(p^{(m)}\right)\right\|_{\mathbb{R}^{2n}}}{p_1^{(m)}} + 2\frac{\left\|\xi_1\left(p^{(m)}\right)\right\|_{\mathbb{R}^{2n}}}{p_1^{(m)}} \sum_{i=1}^{k} \frac{p_i^{(m)}}{\|\xi_1(p^{(m)})\|_{\mathbb{R}^{2n}}} \frac{p_{n+i}^{(m)}}{\|\xi_1(p^{(m)})\|_{\mathbb{R}^{2n}}} \\
&\longrightarrow 0,
\end{aligned}$$

as $m \to \infty$. The final limit follows from (5.15). Analogously, one can estimate from below in the following way

$$\frac{p_{2n+1}^{(m)}}{\|\xi_1(p^{(m)})\|_{\mathbb{R}^{2n}}\, p_1^{(m)}} \ge -\frac{1}{m^2}\frac{\left(\left(p_1^{(m)}\right)^2 + ... + \left(p_k^{(m)}\right)^2 \right)}{\|\xi_1(p^{(m)})\|_{\mathbb{R}^{2n}}\, p_1^{(m)}} + \frac{2\sum_{i=1}^{k} p_1^{(m)} \cdot p_{n+i}^{(m)}}{\|\xi_1(p^{(m)})\|_{\mathbb{R}^{2n}}\, p_1^{(m)}} \longrightarrow 0,$$

as $m \to \infty$. Thus, we can conclude that

$$\left\langle \left(\frac{\alpha_1}{2}, ..., \frac{\alpha_k}{2}, 0, ..., 0\right), \tilde{w}_0 - w_0 \right\rangle \ge 0. \tag{5.16}$$

Consider now H-monotonicity at the points $\tilde{p}^{(m)}$ and $p^{(m)}$. It holds that

$$\left\langle \xi_1\left(\tilde{p}^{(m)}\right) - \xi_1\left(p^{(m)}\right), \tilde{w}^{(m)} - w_m \right\rangle \ge 0.$$

Repeating the previous argument, we obtain

$$\left\langle \left(-\frac{\alpha_1}{2}, ..., -\frac{\alpha_k}{2}, 0, ..., 0\right), \tilde{w}_0 - w_0 \right\rangle \ge 0. \tag{5.17}$$

Combining together inequalities (5.16) and (5.17), we have

$$\left\langle \left(\frac{\alpha_1}{2}, ..., \frac{\alpha_k}{2}, 0, ..., 0 \right), \tilde{w}_0 - w_0 \right\rangle = \left\langle \left(\frac{\alpha_1}{2}, ..., \frac{\alpha_k}{2}, 0, ..., 0 \right), (\eta_0 + \tilde{v}_0) - (\eta_0 + v_0) \right\rangle = 0.$$

This equality provides a contradiction. Let us see why. We recall that $\alpha \in \mathbb{V}_0$, $\tilde{v}_0 \in \mathbb{V}_0$ is fixed and that we chose v_0 arbitrarly in the k-dimensional disk $D_0 \subset \mathbb{V}_0$. Therefore, equality

$$\left\langle \left(\frac{\alpha_1}{2}, ..., \frac{\alpha_k}{2}, 0, ..., 0 \right), \tilde{v}_0 \right\rangle = \left\langle \left(\frac{\alpha_1}{2}, ..., \frac{\alpha_k}{2}, 0, ..., 0 \right), v_0 \right\rangle,$$

is verified if and only if α is the zero vector, which is a contradiction. The Claim in finally proved.

The proof of the Theorem is now to the end. Thanks to the Claim, we now know that there exist $r_0 > 0$ and $\beta_0 > 0$ such that

$$\Sigma_H^{k,h}(T) \cap B(0, r) \cap C_{\mathbb{W}_0, \mathbb{V}_0}(0, \beta) = \{0\},$$

for every $0 < r < r_0$ and for every $0 < \beta < \beta_0$. Using Remark 4.4.5, it holds that there exists $\alpha_0 > 0$ such that

$$\Sigma_H^{k,h}(T) \cap B(0, r) \subset C_{\mathbb{V}_0, \mathbb{W}_0}(0, \alpha),$$

for every $0 < r < r_0$. With this observation the proof is finished. Indeed, we can apply Theorem 5.2.3, which guarantees that $\dim_H \Sigma_H^{k,h}(T) \leq 2n + 2 - k$. $\square$

Bibliography

[1] T. Abdellaoui and H. Heinich, *Sur la distance de deux lois dans le cas vectoriel*, C. R. Acad. Sci. Paris Sér. I Math., 319 (1994), pp. 397–400.

[2] G. Alberti and L. Ambrosio, *A geometrical approach to monotone functions in* $\mathbb{R}^n$, Math. Z., 230 (1999), pp. 259–316.

[3] G. Alberti, L. Ambrosio, and P. Cannarsa, *On the singularities of convex functions*, Manuscripta Math., 76 (1992), pp. 421–435.

[4] C. D. Aliprantis and K. C. Border, *Infinite-dimensional analysis*, Springer-Verlag, Berlin, second ed., 1999. A hitchhiker's guide.

[5] L. Ambrosio, *Lecture notes on optimal transport problems*, in Mathematical aspects of evolving interfaces (Funchal, 2000), vol. 1812 of Lecture Notes in Math., Springer, Berlin, 2003, pp. 1–52.

[6] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 2000.

[7] L. Ambrosio and N. Gigli, *A user's guide to optimal transport*, in Modelling and optimisation of flows on networks, vol. 2062 of Lecture Notes in Math., Springer, Heidelberg, 2013, pp. 1–155.

[8] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, second ed., 2008.

[9] L. Ambrosio and A. Pratelli, *Existence and stability results in the L^1 theory of optimal transportation*, in Optimal transportation and applications (Martina Franca, 2001), vol. 1813 of Lecture Notes in Math., Springer, Berlin, 2003, pp. 123–160.

[10] L. Ambrosio and S. Rigot, *Optimal mass transportation in the Heisenberg group*, J. Funct. Anal., 208 (2004), pp. 261–301.

[11] P. Appell, *Mémoire sur les déblais et les remblais des systèmes continue ou discontinues*, Mémoires présentés par divers Savants à l'académie des Sciences de l'Institut de France, 29 (1887), pp. 1–208.

[12] G. Arena and R. Serapioni, *Intrinsic regular submanifolds in Heisenberg groups are differentiable graphs*, Calc. Var. Partial Differential Equations, 35 (2009), pp. 517–536.

[13] Z. M. Balogh, A. Calogero, and R. Pini, *On the local boundedness of maximal H-monotone operators*, Nonlinear Anal., 148 (2017), pp. 88–105.

[14] A. Bellaïche, *The tangent space in sub-Riemannian geometry*, in Sub-Riemannian geometry, vol. 144 of Progr. Math., Birkhäuser, Basel, 1996, pp. 1–78.

[15] P. Bernard and B. Buffoni, *Optimal mass transportation and Mather theory*, J. Eur. Math. Soc. (JEMS), 9 (2007), pp. 85–121.

[16] A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni, *Stratified Lie groups and potential theory for their sub-Laplacians*, Springer Monographs in Mathematics, Springer, Berlin, 2007.

[17] Y. Brenier, *Décomposition polaire et réarrangement monotone des champs de vecteurs*, C. R. Acad. Sci. Paris Sér. I Math., 305 (1987), pp. 805–808.

[18] Y. Brenier, *Factorisation polaire et réarrangement des fonctions à valeurs vectorielles*, in Nonlinear partial differential equations and their applications.

Collège de France Seminar, Vol. XI (Paris, 1989–1991), vol. 299 of Pitman Res. Notes Math. Ser., Longman Sci. Tech., Harlow, 1994, pp. 52–67.

[19] L. A. Caffarelli, *The regularity of mappings with a convex potential*, J. Amer. Math. Soc., 5 (1992), pp. 99–104.

[20] A. Calogero and R. Pini, *Horizontal normal map on the Heisenberg group*, J. Nonlinear Convex Anal., 12 (2011), pp. 287–307.

[21] ——, *c horizontal convexity on Carnot groups*, J. Convex Anal., 19 (2012), pp. 541–567.

[22] ——, *On Minty's theorem in the Heisenberg group*, Nonlinear Anal., 104 (2014), pp. 12–20.

[23] P. Cannarsa and C. Sinestrari, *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*, Progress in Nonlinear Differential Equations and their Applications, 58, Birkhäuser Boston, Inc., Boston, MA, 2004.

[24] L. Capogna, D. Danielli, S. D. Pauls, and J. T. Tyson, *An introduction to the Heisenberg group and the sub-Riemannian isoperimetric problem*, vol. 259 of Progress in Mathematics, Birkhäuser Verlag, Basel, 2007.

[25] T. Champion and L. De Pascale, *The Monge problem in* $\mathbb{R}^d$, Duke Math. J., 157 (2011), pp. 551–572.

[26] J. A. Cuesta and C. Matrán, *Notes on the Wasserstein metric in Hilbert spaces*, Ann. Probab., 17 (1989), pp. 1264–1276.

[27] J. A. Cuesta-Albertos, C. Matrán-Bea, and A. Tuero-Diaz, *On lower bounds for the* L^2*-Wasserstein metric in a Hilbert space*, J. Theoret. Probab., 9 (1996), pp. 263–283.

[28] D. Danielli, N. Garofalo, and D.-M. Nhieu, *Notions of convexity in Carnot groups*, Comm. Anal. Geom., 11 (2003), pp. 263–341.

[29] L. De Pascale and S. Rigot, *Monge's transport problem in the Heisenberg group*, Adv. Calc. Var., 4 (2011), pp. 195–227.

[30] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, Textbooks in Mathematics, CRC Press, Boca Raton, FL, revised ed., 2015.

[31] S. Fang and J. Shao, *Optimal transport maps for Monge-Kantorovich problem on loop groups*, J. Funct. Anal., 248 (2007), pp. 225–257.

[32] H. Federer, *Geometric measure theory*, Die Grundlehren der mathematischen Wissenschaften, Band 153, Springer-Verlag New York Inc., New York, 1969.

[33] A. Figalli and N. Juillet, *Absolute continuity of Wasserstein geodesics in the Heisenberg group*, J. Funct. Anal., 255 (2008), pp. 133–141.

[34] A. Figalli, Y.-H. Kim, and R. J. McCann, *Hölder continuity and injectivity of optimal maps*, Arch. Ration. Mech. Anal., 209 (2013), pp. 747–795.

[35] A. Figalli and L. Rifford, *Mass transportation on sub-Riemannian manifolds*, Geom. Funct. Anal., 20 (2010), pp. 124–159.

[36] ——, *Mass transportation on sub-Riemannian manifolds*, Geom. Funct. Anal., 20 (2010), pp. 124–159.

[37] A. Figalli, L. Rifford, and C. Villani, *Necessary and sufficient conditions for continuity of optimal transport maps on Riemannian manifolds*, Tohoku Math. J. (2), 63 (2011), pp. 855–876.

[38] G. B. Folland and E. M. Stein, *Hardy spaces on homogeneous groups*, vol. 28 of Mathematical Notes, Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1982.

[39] B. Franchi, M. Marchi, and R. P. Serapioni, *Differentiability and approximate differentiability for intrinsic Lipschitz functions in Carnot groups and a Rademacher theorem*, Anal. Geom. Metr. Spaces, 2 (2014), pp. 258–281.

[40] B. Franchi, V. Penso, and R. Serapioni, *Remarks on Lipschitz domains in Carnot groups*, in Geometric control theory and sub-Riemannian geometry, vol. 5 of Springer INdAM Ser., Springer, Cham, 2014, pp. 153–166.

[41] B. Franchi, R. Serapioni, and F. Serra Cassano, *Rectifiability and perimeter in the Heisenberg group*, Math. Ann., 321 (2001), pp. 479–531.

[42] ——, *Rectifiability and perimeter in step 2 groups*, in Proceedings of EQUADIFF, 10 (Prague, 2001), vol. 127, 2002, pp. 219–228.

[43] ——, *On the structure of finite perimeter sets in step 2 Carnot groups*, J. Geom. Anal., 13 (2003), pp. 421–466.

[44] ——, *Regular hypersurfaces, intrinsic perimeter and implicit function theorem in Carnot groups*, Comm. Anal. Geom., 11 (2003), pp. 909–944.

[45] ——, *Intrinsic submanifolds, graphs and currents in Heisenberg groups*, in Lecture notes of Seminario Interdisciplinare di Matematica. Vol. IV, Lect. Notes Semin. Interdiscip. Mat., IV, S.I.M. Dep. Mat. Univ. Basilicata, Potenza, 2005, pp. 23–38.

[46] ——, *Intrinsic Lipschitz graphs in Heisenberg groups*, J. Nonlinear Convex Anal., 7 (2006), pp. 423–441.

[47] ——, *Regular submanifolds, graphs and area formula in Heisenberg groups*, Adv. Math., 211 (2007), pp. 152–203.

[48] ——, *Differentiability of intrinsic Lipschitz functions within Heisenberg groups*, J. Geom. Anal., 21 (2011), pp. 1044–1084.

[49] B. Franchi and R. P. Serapioni, *Intrinsic Lipschitz Graphs Within Carnot Groups*, J. Geom. Anal., 26 (2016), pp. 1946–1994.

[50] J. H. G. Fu, *Monge-Ampère functions. I, II*, Indiana Univ. Math. J., 38 (1989), pp. 745–771, 773–789.

[51] W. Gangbo and R. J. McCann, *Optimal maps in Monge's mass transport problem*, C. R. Acad. Sci. Paris Sér. I Math., 321 (1995), pp. 1653–1658.

[52] ——, *The geometry of optimal transportation*, Acta Math., 177 (1996), pp. 113–161.

[53] N. Garofalo and F. Tournier, *New properties of convex functions in the Heisenberg group*, Trans. Amer. Math. Soc., 358 (2006), pp. 2011–2055 (electronic).

[54] M. Gromov, *Carnot-Carathéodory spaces seen from within*, in Sub-Riemannian geometry, vol. 144 of Progr. Math., Birkhäuser, Basel, 1996, pp. 79–323.

[55] C. E. Gutiérrez and Q. Huang, *The refractor problem in reshaping light beams*, Arch. Ration. Mech. Anal., 193 (2009), pp. 423–443.

[56] ——, *The near field refractor*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 31 (2014), pp. 655–684.

[57] J. Heinonen, *Lectures on Lipschitz analysis*, vol. 100 of Report. University of Jyväskylä Department of Mathematics and Statistics, University of Jyväskylä, Jyväskylä, 2005.

[58] E. Hewitt and K. Stromberg, *Real and abstract analysis*, Springer-Verlag, New York-Heidelberg, 1975. A modern treatment of the theory of functions of a real variable, Third printing, Graduate Texts in Mathematics, No. 25.

[59] L. Hörmander, *Hypoelliptic second order differential equations*, Acta Math., 119 (1967), pp. 147–171.

[60] N. Juillet, *On displacement interpolation of measures involved in Brenier's theorem*, Proc. Amer. Math. Soc., 139 (2011), pp. 3623–3632.

[61] L. V. Kantorovich, *On a problem of Monge*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI), 312 (2004), pp. 15–16.

[62] ——, *On mass transportation*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI), 312 (2004), pp. 11–14.

[63] M. Knott and C. S. Smith, *On the optimal mapping of distributions*, J. Optim. Theory Appl., 43 (1984), pp. 39–49.

[64] V. L. Levin, *General Monge-Kantorovich problem and its applications in measure theory and mathematical economics*, in Functional analysis, optimization, and mathematical economics, Oxford Univ. Press, New York, 1990, pp. 141–176.

[65] X.-N. Ma, N. S. Trudinger, and X.-J. Wang, *Regularity of potential functions of the optimal transportation problem*, Arch. Ration. Mech. Anal., 177 (2005), pp. 151–183.

[66] V. Magnani, *Elements of geometric measure theory on sub-Riemannian groups*, Scuola Normale Superiore, Pisa, 2002.

[67] P. Mattila, *Distribution of sets and measures along planes*, J. London Math. Soc. (2), 38 (1988), pp. 125–132.

[68] ——, *Geometry of sets and measures in Euclidean spaces*, vol. 44 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 1995. Fractals and rectifiability.

[69] P. Mattila, R. Serapioni, and F. Serra Cassano, *Characterizations of intrinsic rectifiability in Heisenberg groups*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5), 9 (2010), pp. 687–723.

[70] R. J. McCann, *Existence and uniqueness of monotone measure-preserving maps*, Duke Math. J., 80 (1995), pp. 309–323.

[71] ——, *Polar factorization of maps on Riemannian manifolds*, Geom. Funct. Anal., 11 (2001), pp. 589–608.

[72] ——, *A glimpse into the differential topology and geometry of optimal transport*, Discrete Contin. Dyn. Syst., 34 (2014), pp. 1605–1621.

[73] R. J. McCann, B. Pass, and M. Warren, *Rectifiability of optimal transportation plans*, Canad. J. Math., 64 (2012), pp. 924–934.

[74] J. Mitchell, *On Carnot-Carathéodory metrics*, J. Differential Geom., 21 (1985), pp. 35–45.

[75] G. Monge, *Mémoire sur la théorie des déblais et des remblais*, Histoire de l'Académie Royale des Sciences de Paris, avec les Mémoires de Mathémathique et Physique pour la même année, (1781), pp. 666–704.

[76] R. Montgomery, *A tour of subriemannian geometries, their geodesics and applications*, vol. 91 of Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 2002.

[77] R. Monti, *Some properties of Carnot-Carathéodory balls in the Heisenberg group*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl., 11 (2000), pp. 155–167 (2001).

[78] P. Pansu, *Métriques de Carnot-Carathéodory et quasiisométries des espaces symétriques de rang un*, Ann. of Math. (2), 129 (1989), pp. 1–60.

[79] S. T. Rachev, *The Monge-Kantorovich problem on mass transfer and its applications in stochastics*, Teor. Veroyatnost. i Primenen., 29 (1984), pp. 625–653.

[80] T. Rajala, *Large porosity and dimension of sets in metric spaces*, Ann. Acad. Sci. Fenn. Math., 34 (2009), pp. 565–581.

[81] J. G. Rešetnjak, *Generalized derivatives and differentiability almost everywhere*, Mat. Sb. (N.S.), 75(117) (1968), pp. 323–334.

[82] R. T. Rockafellar, *Convex analysis*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original, Princeton Paperbacks.

[83] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*, vol. 317 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 1998.

[84] L. Rüschendorf, *Fréchet-bounds and their applications*, in Advances in probability distributions with given marginals (Rome, 1990), vol. 67 of Math. Appl., Kluwer Acad. Publ., Dordrecht, 1991, pp. 151–187.

[85] L. Rüschendorf and S. T. Rachev, *A characterization of random variables with minimum L^2-distance*, J. Multivariate Anal., 32 (1990), pp. 48–54.

[86] L. Simon, *Lectures on geometric measure theory*, vol. 3 of Proceedings of the Centre for Mathematical Analysis, Australian National University, Australian National University, Centre for Mathematical Analysis, Canberra, 1983.

[87] C. S. Smith and M. Knott, *Note on the optimal transportation of distributions*, J. Optim. Theory Appl., 52 (1987), pp. 323–329.

[88] ——, *On Hoeffding-Fréchet bounds and cyclic monotone relations*, J. Multivariate Anal., 40 (1992), pp. 328–334.

[89] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, vol. 43 of Princeton Mathematical Series, Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

[90] V. N. Sudakov, *Geometric problems in the theory of infinite-dimensional probability distributions*, Proc. Steklov Inst. Math., (1979), pp. i–v, 1–178. Cover to cover translation of Trudy Mat. Inst. Steklov **1**41 (1976).

[91] V. S. Varadarajan, *Lie groups, Lie algebras, and their representations*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1974. Prentice-Hall Series in Modern Analysis.

[92] C. Villani, *Topics in optimal transportation*, vol. 58 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2003.

[93] ——, *Optimal transport*, vol. 338 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 2009. Old and new.