UNIVERSITÀ DEGLI STUDI DI PISA

PH.D THESIS

SCIENTIFIC DISCIPLINE: MAT/05

DIFFERENTIABILITY PROPERTIES AND CHARACTERIZATIONS OF H-CONVEX FUNCTIONS

Candidate:
MATTEO SCIENZA

Advisor:
DR. VALENTINO MAGNANI

Contents

INTRODUCTION

The present thesis deals with a number of geometric properties of convex functions in a non-Euclidean framework. This setting is represented by the so-called Sub-Riemannian space, also called a Carnot-Carathéodory (CC) space, that can be thought of as a space where the metric structure is a constrained geometry and one can move only along a prescribed set of directions depending on the point. More precisely, a sub-Riemannian space is given by an open subset $\Omega \subset \mathbb{R}^n$ endowed with a family of vector fields

$$\mathcal{X} = \{X_1, \dots, X_m\}.$$

One also assumes that every two points $x, y \in \Omega$ can be joined by an absolutely continuous path $\gamma : [0, T] \to \Omega$, $T > 0$ such that for almost every $t \in [0, T]$

$$\dot{\gamma}(t) = \sum_{j=1}^{m} h_j(t) X_j(\gamma(t)), \quad \sum_{j=1}^{m} h_j(t)^2 \leq 1,$$

where $\{h_j\}$ are measurable functions. Such curve is called a *subunit* curve. The notion of Carnot-Carathéodory (CC) distance associated with the family $\mathcal{X}$ is defined as follows

$$d(x, y) = \inf\{T \geq 0 \text{ exists a subunit curve } \gamma : [0, T] \to \mathbb{R}^n, \text{ such that } \gamma(0) = x \text{ and } \gamma(T) = y\}. \tag{1}$$

By a classical result due to Chow [47], if the family $\mathcal{X}$ is bracket generating, namely, given a multi-index $I = (i_1, \dots, i_p)$, $1 \leq i_j \leq m$, setting $|I| = p$,

$$X_{[I]} = \left[X_{i_1}, \left[\dots, \left[X_{i_{p-1}}, X_{i_p}\right] \dots\right]\right],$$

and for every $x \in \mathbb{R}^n$ there exists an integer $r = r(x)$ such that

$$\operatorname{span}\{X_{[I]}(x) : |I| \leq r\} = \mathbb{R}^n, \tag{2}$$

then the CC distance $d(x, y)$ introduced above is finite for every $x, y \in \mathbb{R}^n$. Moreover, d is a metric that induces the Euclidean topology of $\mathbb{R}^n$.

In classical PDE theory, the bracket generating condition, also called Hörmander condition, is well known. It appeared first in [57], in this work the author proved the *hypoellipticity* of second order degenerate operators of the type

$$\mathcal{L} = -\sum_{i=1}^{m} X_i^2,$$

where X_i are the vector fields in $\mathcal{X}$ and condition (2) holds everywhere. We observe that there is a direct connection between the regularity of the CC distance d and the hypoellipticity of $\mathcal{L}$, we refer to [32] for further details. In general, one can associate a distance to every degenerate elliptic operator. For the operator $\mathcal{L}$ it coincides with the CC distance. Many contributions have appeared in this area. Let us mention for instance the work of Bony [16], Capogna, Danielli and Garofalo [21], Citti, Garofalo and Lanconelli [23], Fabes, Kenig and Franchi, Folland [33, 34], Franchi and Lanconelli [36, 37], Garofalo and Lanconelli [41], Jerison and Sánchez-Calle [59], Nagel, Ricci and Stein [82], Sánchez-Calle [94]. This list is certainly far from being complete.
Sub-Riemannian spaces, as the name means to suggest, are a natural generalization of the Riemannian ones. In fact, the CC metric can be introduced as a limit case of the Riemannian metric, see [22, 36, 37]. We point out that despite these Riemannian approximations, CC spaces are far from being Euclidean even locally, indeed the CC distance is not biLipschitz equivalent to the Euclidean one. Hence the extension of classical theory to this setting can also be a useful inspiration for further developments of analysis in general metric spaces.
As a special class of CC-spaces we have Carnot groups also known as *stratified groups*. They inherit all the geometric complexity of a non-Euclidean space along with a very rich structure for analytic and geometric investigations. A Carnot group of step r is a connected and simply-connected Lie group whose Lie algebra $\mathfrak{g}$ is stratified. This means that $\mathfrak{g}$ admits a decomposition as a vector space direct sum

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

with $[V_1, V_j] = V_{j+1}$, for $j = 1, \ldots, r-1$, and $\mathfrak{g}$ is r-nilpotent. We recall that the exponential map $exp : \mathfrak{g} \to \mathcal{G}$ is a global analytic diffeomorphism. It allows to define a set of analytic coordinate maps. Moreover a Carnot group is naturally equipped with a family of non-isotropic dilations defined as $\delta_\lambda(Z_j) = \lambda^j Z_j$, where $Z_j \in V_j$. According to notations and terminology of Section 1.2, we represent a stratified group $\mathbb{G}$ as a finite dimensional Hilbert space that is a direct sum of orthogonal subspaces H_1, H_2, ..., H_ι and that it is equipped with a suitable polynomial operation. Throughout this thesis H_1 will denote the subspace of *horizontal directions* at the origin. Let $\mathfrak{g} = V_1 \oplus \cdots \oplus V_\iota$ be the Lie algebra associated to $\mathbb{G}$. The left invariant vector fields of V_j are exactly the ones that at the origin take values in H_j. We call *horizontal* the vector fields in V_1. We also remark that Carnot groups arise as "tangent space" to sub-Riemannian manifolds with equiregular distributions, [9, 74, 77]. This goes back to the techniques used by Rothschild and Stein [93], Métivier [75] and Goodman [48] to approximate linear differential operators defined by means of vector fields satisfying the Hörmander condition with homogeneous left invariant vector fields. In view of this characterization Carnot groups can also be considered as a local model for general CC-spaces.

Motivated by some interesting questions in the theory of fully nonlinear equations and the role played by convexity in the Euclidean setting, D. Danielli, N. Garofalo and D.M. Nhieu introduced the notion of convexity in Carnot groups, for more details on these motivations, see [26, 44] and also [50, 51]. Convexity in Sub-Riemannian groups is a quite recent stream, that goes back to the works by Danielli, Garofalo and Nhieu [26], Lu, Manfredi and Stroffolini [64]. Contributions in the more general setting of Sub-Riemannian spaces can be found in Trudinger [99] and more recently in the paper of Bardi and Dragoni [8]. For Carnot groups different pointwise notions of convexity have been investigated in [26]. Among them, the most natural one turned out to be that of weakly h-convex function, in short, *h-convex function*. An h-convex function $u : \Omega \longrightarrow \mathbb{R}$ defined on an open set Ω of a stratified group $\mathbb{G}$ satisfies the property of being classically convex, when it is restricted to all horizontal lines contained in Ω. These are exactly the integral curves of the horizontal vector fields of $\mathbb{G}$. More precisely, we say that $u : \Omega \longrightarrow \mathbb{R}$ is h-convex if whenever $h \in H_1$ and $[0, h] \subset x^{-1} \cdot \Omega$, where $[0, h] = \{th : 0 \leq t \leq 1\}$, we have

$$u(x\delta_t h) \leq (1-t)u(x) + tu(xh). \tag{3}$$

We stress that this notion of convexity turns out to be "local" and it does not require any assumption on Ω. In fact, it is not difficult to observe that smooth h-convex functions are characterized by an everywhere nonnegative *horizontal Hessian*, see Definition 2.2.4. This fits with the approach of [64], where the authors introduce v-convex functions as upper semicontinuous functions, whose Hessian is nonnegative in the viscosity sense. Let us point out that the notions of v-convexity and of h-convexity are equivalent, [6], [102], [60], [71], [88]. It is worth to mention that other definitions of convexity can be introduced in this setting. For instance, the distributional convexity which we postpone to Chapter 4 or the geodetic convexity studied in [78]. This last notion, which is natural in the Riemannian setting, is useless in the sub-Riemannian case, since the classes of geodetically convex sets and functions are trivial in the Heisenberg group, [78].

All details and precise definitions related to convexity in stratified groups will be deferred to Section 2.1.

So far, in stratified groups the following first order regularity properties have been proved for h-convex functions. First, it has been proved in [26] that continuous h-convex functions are locally Lipschitz continuous with respect to the CC metric, then as a consequence of the celebrated Rademacher type differentiability theorem of Pansu [85], they are almost everywhere h-differentiable, see Theorem 1.2.25. The local Lipschitz continuity of h-convex functions has been first proved for h-convex functions bounded from above in [6, 71], and subsequently to all measurable h-convex function in [87]. In addition, a quantitative Lipschitz estimate holds, [26, 60, 64, 88]. More precisely one can control the supremum norm of an h-convex function on a ball of radius r by the integral mean of its absolute values on a ball with comparable radius, see Theorem 2.2.10. In Section 2.1 we will give a different proof of this result by a sub-mean formula proved by Bonfiglioli and Lanconelli in [12]. One of the main result in this thesis is that these $L^\infty - L^1$ estimates can be suitably extended to CC-spaces.

In order to introduce this result in Carnot Carathéodory spaces let us define the notion of $\mathcal{X}$-convexity, see [8]. Consider a family of smooth vector fields $\mathcal{X}$ which satisfy the Hörmander condition. All linear combinations of vector fields in $\mathcal{X}$ correspond to the so-called horizontal vector fields. We say that a function u is $\mathcal{X}$-convex if it is convex along the integral curves of horizontal vector fields, see Definition 5.3.1. In Theorem 6.1 of [8],

it is proved that $\mathcal{X}$-convexity, local boundedness and upper semicontinuity imply local Lipschitz continuity with respect to d, where the Carnot-Carathéodory distance d given by $\mathcal{X}$, is only assumed to yield the Euclidean topology. This result also gives L^∞-estimates for the horizontal derivatives Xu in terms of the L^∞-norm of u, where $X \in \mathcal{X}$. In the case $\mathcal{X}$ generates a Carnot group, these estimates take a quantitative form, see [26] and [60, 64].
The following result, proved in collaboration with V. Magnani [66], establishes that the previous estimates can be suitably extended to Carnot-Carathéodory spaces generated by a set $\mathcal{X}$ of Hörmander vector fields, see also Theorem 5.6.1

Theorem *For each $\mathcal{X}$-convex function $u : \Omega \to \mathbb{R}$, that is locally bounded from above, for every $x \in K$, where K is a compact set of $\mathbb{R}^n$ we have*

$$\sup_{B_{x,r}} |u| \quad \leq \quad C \fint_{B_{x,2r}} |u(w)|\, dw \tag{4}$$

$$|u(y) - u(z)| \quad \leq \quad C\, \frac{d(y,z)}{r} \fint_{B_{x,2r}} |u(w)|\, dw\,, \tag{5}$$

for every $0 < r < R$ and every $y, z \in B_{x,r}$. Here R and C are constants depending only on $\mathcal{X}$ and K.

Our approach to prove (4) and (5) differs from both the geometric approach of [26] and the PDEs approach of [60, 64]. In fact, we need both these aspects, according to the following scheme. We start from a $\mathcal{X}$-convex function $u : \Omega \to \mathbb{R}$ that is locally bounded from above. By a result of D. Morbidelli, [80], the Carnot-Carathéodory ball can be covered by suitable composition of flows of horizontal vector fields in a quantitative way, depending on the radius of the ball. This essentially allows us to apply the approach of [71] that relies on the one dimensional convexity of u along these flows, hence obtaining explicit Lipschitz estimates. It follows that u belongs to the anisotropic Sobolev space $W^{1,2}_{\mathcal{X},loc}(\Omega)$, see Section 1.1 for more information. The crucial step is to show that for every $x \in \Omega$ the $\mathcal{X}$-convex function u is a *weak subsolutions* on a small ball centered at x of a suitable "pointed sub-Laplacian"

$$\mathcal{L}_x = \sum_{j=1}^m Y_j^2\,,$$

that depends on x, see Theorem 5.5.3. Since the Lebesgue measure is locally doubling with respect to metric balls, the Poincaré inequality holds, and cut-off functions are available see [21], then the classical Moser iteration technique can be applied for weak subsolutions to the sub-Laplacian equation, hence getting in particular the local upper estimate, see also Corollary 5.5.4. The local lower estimate of u is reached using again the approximate exponential, obtaining the pointwise estimate

$$2^{N_x}\, u(x) - (2^{N_x} - 1) \sup_{B_{x,\bar{N}\delta}} u \leq \inf_{B_{x,b\delta}} u\,, \tag{6}$$

where N_x depends on x and it satisfies the uniform inequality $1 \leq N_x \leq \bar{N}$ on some compact set, see Lemma 5.3.2. This eventually leads us to the proof of (5). The estimate (6) is a straightforward consequence of Theorem 5.3.5 joined with Theorem 5.6.2. In sum, the geometric part of our method arises from a quantitative representation of the Carnot-Carathéodory ball by approximates exponentials and it leads us to the lower estimates.The PDEs part of our approach provides the upper estimates.

There are various challenging questions on h-convex functions in stratified groups, that are still far from being understood. The rich geometric structure of a Carnot group allows for a deeper study of its differentiability properties. One of the most important is certainly the validity of an Alexsandrov-Bakelman-Pucci estimate, that is still an intriguing open question already in the Heisenberg group and it was also one of the main motivations to study h-convexity in this framework, see [26] and [27]. On another side, we also have the second order differentiability of convex functions, namely, the classical Alexsandrov-Busemann-Feller's theorem. This is an important result in different areas of Analysis and Geometry. For instance, in the theory of fully nonlinear elliptic equations, this theorem plays an essential role in uniqueness theory, see Chapter 5 of [18]. Since the works of Busemann and Feller, [17], and of Alexsandrov [2], there have been different methods to establish this theorem in Euclidean spaces. The functional analytic method by Reshetnyak, [86], relies on the fact that the gradient of a convex function has bounded variation. This scheme can be extended to stratified groups, provided that one can prove that an h-convex function is BV_H^2 in the sense of [5]. This important fact has been established by different authors for h-convex functions on Heisenberg groups and two step stratified groups [50], [51], [44], [28] and also for k-convex functions with respect to two step Hörmander vector fields, [99]. Precisely, the main result of [28] gives us the following version of the Alexsandrov-Busemann-Feller theorem. *Let Ω be an open set of a two step stratified group and let $u : \Omega \longrightarrow \mathbb{R}$ be h-convex. Then u has at a.e. $x \in \Omega$ a second order h-expansion at x.* We say that $u : \Omega \longrightarrow \mathbb{R}$ has a *second order h-expansion* at $x \in \Omega$ if there exists a polynomial $P_x : \mathbb{G} \longrightarrow \mathbb{R}$ of homogeneous degree less than or equal to two, such that

$$u(xw) = P_x(w) + o(\|w\|^2) \qquad \text{as} w \to 0. \tag{7}$$

Unfortunately, it is still not clear whether h-convex functions are BV_H^2 in higher step groups and this makes the Alexsandrov-Busemann-Feller's theorem an important open issue for the higher step geometries of stratified groups.

On the other hand, the first proofs of this result in Euclidean spaces, [2], [17] and also some of the subsequent proofs did not use the bounded variation property of the gradient. For instance, Rockafellar's proof of [89] relies on Mignot's a.e. differentiability of monotone functions, [76], where the crucial observation is that the subdifferential of a convex function is a monotone function. This may suggest different approaches to Alexsandrov's theorem in stratified groups and constitutes our first motivation to study the properties of the h-subdifferential in Chapter 4. The notion of h-subdifferential has been introduced in [26] for h-convex functions. In analogy with the local notion of convexity mentioned above, we use "a local version" of this notion, that allows us to treat h-convex functions on arbitrary open sets.

We say that $p \in H_1$ is an *h-subdifferential* of $u : \Omega \longrightarrow \mathbb{R}$ at $x \in \Omega$ if whenever $h \in H_1$ and $[0,h] \subset x^{-1} \cdot \Omega$,

$$u(xh) \geq u(x) + \langle p, h \rangle \,. \tag{8}$$

We denote by $\partial_H u(x)$ the set of all h-subdifferentials of u at x and the corresponding set-valued mapping by $\partial_H u : \Omega \rightrightarrows H_1$. Here $\langle \cdot, \cdot \rangle$ in (8) is the scalar product fixed in $\mathbb{G}$, such that all subspaces H_j are orthogonal. Our starting point was the characterization of the second order differentiability of h-convex functions. In the Euclidean framework, this has been done by Rockafellar, where in Theorem 2.8 of [91] proves that *a convex function has a second order expansion at a fixed point if its gradient is differentiable at that point in the extended sense*. In Chapter 4, we extend such result to all Carnot groups. More precisely in Theorem 3.2.8 we prove that every h-convex function u has a second order h-expansion at $x \in \Omega$ if and only if

$$\partial_H u(xw) \subseteq v + A_x(w) + o(\|w\|)\mathbb{B}, \tag{9}$$

for all $w \in x^{-1}\Omega$, where $\mathbb{B}$ denotes the unit ball in H_1 and $A_x : \mathbb{G} \to H_1$ is a suitable linear map, see the joint work with V. Magnani, [67]. Notice that in order to prove the analogous of Rockafellar's theorem in stratified groups we have to establish a connection between the differentiability of the h-subdifferential (9) and the differentiability of the gradient. To this aim we have to develop a nonsmooth calculus for h-convex functions, that is the main object of Chapter 3. We start with a characterization of the h-subdifferential. Let $u : \Omega \to \mathbb{R}$ be h-convex. Then for every $x \in \Omega$ we have

$$\bar{co}\left(\nabla_H^{\star} u(x)\right) = \partial_H u(x) \,. \tag{10}$$

We denote by $co(E) \subset H_1$ the convex hull in H_1 of the subset $E \subset H_1$ and by $\bar{co}(E)$ its closure. The *h-reachable gradient* is given by

$$\nabla_H^{\star} u(x) = \left\{ p \in H_1 : \ x_k \to x, \ \nabla_H u(x_k) \text{ exists for all } k\text{'s and } \nabla_H u(x_k) \to p \ \right\}. \tag{11}$$

The proof of (10) cannot follow the Euclidean scheme. However, it is still possible to use the Hahn-Banach's theorem, when applied inside the horizontal subspace H_1, that has a linear structure. Another difficulty is that the group mollification does not commute with horizontal derivatives, hence the mollification argument of the Euclidean proof cannot be applied. We overcome this point by a Fubini type argument with respect to a semidirect factorization, following the approach of [68].

The uniqueness of the h-subdifferential as a consequence of h-differentiability has been already shown [27], see also [20] for the case of Heisenberg groups. To show the opposite implication we decompose the difference quotient of u into sums of difference quotients along horizontal directions. The same decomposition along horizontal directions have been first used by Pansu, [85]. The second ingredient is the following non-smooth mean value theorem.

Theorem *Let $u : \Omega \longrightarrow \mathbb{R}$ be an h-convex function. Then for every $x \in \Omega$ and every h such that $[0,h] \subseteq H_1 \cap x^{-1}\Omega$, there exists $t \in [0,1]$ and $p \in \partial_H u(x\delta_t h)$ such that $u(xh) - u(x) = \langle p, h \rangle$.*

This theorem, corresponding to Corollary 3.1.16, is also important in the proof of Theorem 3.2.8. In fact, it is an essential tool to establish that twice h-differentiability implies the existence of a second order *h*-expansion. In the Euclidean framework, the previous implication can be found in Theorem 7.10 of [1], where the Clarke's nonsmooth mean value theorem plays a key role.

Another natural notion of convexity introduced in the Euclidean setting is the distributional one: let Ω be an open set of $\mathbb{R}^n$ and let $T \in \mathcal{D}'(\Omega)$ be a distribution. We say that T is convex if its Hessian is positive definite in the sense of distributions. In Euclidean spaces, the first distributional characterization of convexity goes back to L. Schwartz in [95], who proved that a distribution in $\mathbb{R}$ is a convex function if and only if its second derivative is a non-negative Radon measure. Bakel'man showed that all second order distributional derivatives of a convex function in $\mathbb{R}^n$ are signed Radon measures, [7]. Subsequently, Reshetnyak established that a locally summable function is defined by a convex function if and only if its distributional Hessian is nonnegative, [86]. This characterization has been improved by Dudley, who proved that any distribution with nonnegative Hessian in the distributional sense is defined by a convex function, see Theorem 2.1 in [29].

Our aim is to extend this characterization in stratified groups. As it is clear from Dudley's proof, the geometry of the space plays an important role. Dudley's approach uses the following elementary fact: given a convex function u and a simplex S, then u attains its maximum on S at a vertex q. This allows to consider the set S', that is symmetric to S with respect to the vertex q. It follows that $u(y) \geq u(q)$ for every $y \in S'$, giving a local lower bound. This implies that in the case the sequence of convex functions diverges to $+\infty$ at some point, we get its uniform limit to $+\infty$, contradicting the distributional convergence. The preceding description constitutes the leading idea in Dudley's approach. Unfortunately this scheme is not completely reproducible in stratified groups due to the lack of existence of "simplexes" in general stratified groups. We refer to Section 4.2 for more details, where we prove Dudley's characterization for Heisenberg groups. This suggests to follow a different approach with respect to the Euclidean one, in order to get the full distributional characterization. It was already known that in all stratified groups every h-convex function has nonnegative *horizontal Hessian* in the distributional sense, see [26, 64], although the fact that any distribution with nonnegative horizontal Hessian is given by an h-convex function was not completely clear in the setting of general stratified groups. In Section 4.1 we give a full answer to this question, see the joint work with A. Bonfiglioli, E. Lanconelli and V. Magnani, [15]. More precisely, we have the following theorem

Theorem *Let $\mu \in \mathcal{D}'(\Omega)$ be a Radon measure, then μ is defined by an h-convex function if and only if it is given by an h-convex distribution.*

Our scheme is elementary, although it differs from the standard approach: we consider the group convolution of the measure μ, but instead of computing its horizontal Hessian by direct differentiation, we consider its distributional version. This respects the noncommutativity of the convolution operator. As a byproduct of the previous theorem if a function is in L^1_{loc} and h-convex in the distributional sense, then outside a negligible set it coincides with a locally Lipschitz continuous h-convex function.

This result also extends one of the characterizations given in [60], where the subharmonic theory used to prove the equivalence of various notions of h-convexity requires the upper semicontinuity of the function. To reach the complete distributional characterization of h-convexity, we combine Corollary 4.1.7 and Lemma 4.1.8. Using the fundamental solution of the sub-Laplacian Δ_H in stratified groups, [34], this lemma shows that an h-convex distribution T can be written as the sum of a Δ_H-harmonic function and a locally summable function. Since Δ_H-harmonic functions are smooth by Hörmander's theorem, [57], we conclude that T is given by a function in $L^1_{loc}(\Omega)$. Hence if $T \in \mathcal{D}'(\Omega)$ is h-convex, then T is defined by an h-convex function on Ω. Recall that all measurable h-convex functions are locally Lipschitz continuous, [87]. Thus, the previous result shows that the class of h-convex measurable functions coincides with that of h-convex distributions, that are locally Lipschitz continuous h-convex functions. Although we still do not know whether one can find h-convex functions in higher step groups that are nonmeasurable, these functions certainly would not be included in the previous families. This confirms that the natural notion of h-convexity in stratified groups should always include either measurability or local boundedness from above, that are indeed equivalent conditions.
As a last observation we wish to emphasize the importance of distributional characterization from Complex Analysis. In fact, the well known relationship between the Heisenberg group $\mathbb{H}^n$ and the unit sphere in $\mathbb{C}^{n+1}$ suggests an interesting analogy between plurisubharmonic functions, that are subharmonic on all one dimensional complex affine subspaces, and h-convex functions on Heisenberg groups, that are subharmonic on all horizontal lines, that are one dimensional real affine subspaces contained in the horizontal planes. Moreover, the horizontal planes in the Heisenberg groups $\mathbb{H}^n$ correspond to the complex subspaces of the real tangent spaces to the unit sphere in $\mathbb{C}^n$, in view of the identification of $\mathbb{H}^n$ by the holomorphic stereographic projections.

Thus, we may interpret h-convexity in Heisenberg groups as a kind of "real plurisubharmonicity" for mapping on the unit sphere of $\mathbb{C}^n$. In this setting, a distribution whose complex Hessian defines a positive current of bidegree (1,1) is precisely a plurisubharmonic function, see [25]. Since plurisubharmonicity has its real counterpart in h-convexity, as discussed above, we expect that in Heisenberg groups and in more general stratified groups the notion of h-convexity in the distributional sense should have further interesting developments.

The thesis is organized as follows. CHAPTER 1 presents Carnot-Carathéodory spaces and Carnot groups. We start from the basic properties of vector fields and Lie brackets, then we introduce CC spaces and the CC distance induced by a family of vector fields $\mathcal{X}$. In the second part of the chapter we introduce Carnot groups and we collect some definitions and elementary results related to differentiation in this setting. We conclude the chapter with a section on *homogeneous polynomials*.

In CHAPTER 2 we introduce convexity in sub-Riemannian spaces. This chapter is organized as follows. In the first section we introduce the notion of h-convexity in Carnot groups. In Section 2, we state some well known regularity results. We introduce the horizontal Hessian and the notion of v-convexity. Finally, we present a short proof of some $L^\infty - L^1$ estimates for h-convex functions proved in [26, 60]. In fact, these estimates easily follows by an integral representation formula proved in [14].

In CHAPTER 3 we introduce the notion of h-subdifferential in stratified groups and we characterize both first and second order differentiability of h-convex functions. Moreover we show that Alexsandrov's second order differentiability of h-convex functions is equivalent to a suitable differentiability of their horizontal gradient. In Section 3.1, we prove several basic results involving the h-subdifferentials. The set of h-subdifferentials of a function u at x will be denoted by $\partial_H u(x)$. As a first result we prove that $\partial_H u(x)$ is the convex envelope of the limit point set of $\nabla_H u$, see Theorem 3.1.4 and Theorem 3.1.8. The main result of this section is the nonsmooth mean value theorem for h-convex functions, Corollary 3.1.16. We mention also the first order characterization of h-convex functions proved in Theorem 3.1.20. In Section 3.2 we deal with second order analysis of h-convex functions. Here the main result is Theorem 3.2.8, where we prove the equivalence, for h-convex functions, of two notions of second order differentiability, namely Definition 3.2.1 and 3.2.5. In order to prove this we need some preliminary results. The most important one is the characterization of second order h-differentiability, Lemma 3.2.6.

The full characterization of h-convex distributions is the main topic of CHAPTER 4. In Section 4.1 as a first result we extend the Reshetnyak's characterization to all stratified groups, Theorem 4.1.6. As a byproduct we have that a locally summable h-convex function, in the distributional sense, outside a negligible set coincides with a locally Lipschitz continuous h-convex function, Corollary 4.1.7. The main result of this chapter is given in Theorem 4.1.9, where we prove the full distributional characterization. In Section 4.2 we give a different proof of Theorem 4.1.9 restricted to Heisenberg groups. In these groups it is possible to follow the original scheme of Dudley, [29]. This approach is more geometric, but unfortunately it cannot be extended to general stratified groups, see Theorem 4.2.16.

In CHAPTER 5 we consider the notion of convexity in general Sub-Riemannian spaces. The main result of the chapter is the quantitative Lipschitz estimate of Theorem 5.3.5, which is a consequence of the local $L^\infty - L^1$ inequality, see Theorem 5.0.17. In Section 5.1 we introduce two equivalent CC-distances. In Section 5.2 we define almost exponential maps an their properties, according to the work by D. Morbidelli [80]. The main result of Section 5.3 is Theorem 5.3.4. Here we prove that $\mathcal{X}$-convex functions bounded from below are also bounded form above, this follows as in [71]. In Section 5.4 we describe the Moser iteration technique in sub-Riemmanian setting. Although the subject is well known we give full proofs for completeness. Finally, in Section 5.5 we prove our main result, namely Theorem 5.0.17

Acknowledgements

.....

Basic Notations

$\mathbb{R}^n$	n-dimensional Euclidean space
$\mathbb{G}$	a Carnot group, p. 7
$\mathfrak{g}$	a stratified Lie algebra, p. 7
V_i	i layer of the stratified Lie algebra $\mathfrak{g}$, p. 7
exp	exponential map, p. 2
$e_a p$	almost exponential map, p 63
$\mathbb{H}^n$	n-Heisenberg group, p. 8
H_i	image of V_i under the exponential map, p. 8
H_x	horizontal plane at x, namely $x \cdot H_1$, p. 7
Ω	open set in $\mathbb{R}^n$
Ω_0	open and bounded set in $\mathbb{R}^n$
X_i	a smooth vector field on $\mathbb{R}^n$, p. 1
$\mathcal{X}$	family of vector fields in $\mathbb{R}^n$, p. 1
d	CC metric induced by a given family of vector fields, p. 3
$\\| \cdot \\|$	homogeneous norm in a Carnot group, p. 9
h-deg	homogeneous degree, p. 11
$\mathcal{P}_a$	space of polynomials of homogeneous degree less or equal than a, p. 12
$B_{x,r}$	open CC ball centered at $x \in \mathbb{R}^n$ with radius $r \geq 0$, p. 4
$D_{x,r}$	closed CC ball centered at $x \in \mathbb{R}^n$ with radius $r \geq 0$, p. 4
$B^{\rho}_{x,r}$	open ball with respect to the distace ρ centered at $x \in \mathbb{R}^n$ with radius $r \geq 0$
$B^{E}_{x,r}$	open Euclidean ball centered at $x \in \mathbb{R}^n$ with radius $r \geq 0$
∇	Euclidean gradient
∇_H	horizontal gradient in a Carnot group, p. 10
∇^2_H	horizontal symmetrized Hessian in a Carnot group, p. 21
$\mathcal{L}$	a sub-Laplacian, p. 15
Δ_H	Hörmander sub-Laplacian, p. 15
Γ	fundamental solution of Δ_H, p. 15
d_H	Δ_H-gauge, p. 15
$\mathcal{M}_r$ and M_r	respectively surface and solid mean, p. 16
μ^V_x	Δ_H-harmonic measure related to $V \subset G$ and $x \in V$, p. 17
$BV^2_H(\Omega)$	space of functions with bounded horizontal second variation in $\Omega \subset \mathbb{G}$, p. 11
$W^{1,p}_{\mathcal{X}}(\Omega)$	anisotropic Sobolev space associated to the family of vector fields $\mathcal{X}$, p. 4
$W^{1,p}_{\mathcal{X},0}(\Omega)$	closure of $C^\infty_0(\Omega)$ maps with respect to the norm of $W^{1,p}_{\mathcal{X}}$, p. 4

$\Gamma^k(\Omega)$	Follan-Stain class of k-differentiable functions, p. 11
$\langle x, y\rangle$	standard Euclidean inner product of $x, y \in \mathbb{R}^n$
$x \cdot y$	$x \cdot y = x + y + Q(x,y)$, product of $x, y \in \mathbb{G}$, p. 8
	with respect to a Carnot groups structure
δ_λ	dilations in a Carnot group, p. 7
Q	homogeneous dimension of a Carnot group, p. 8
$\mathcal{H}^k$	k-dimensional Hausdorff measure associated to the CC distance, p. 9
$\lvert E\rvert$	Lebesgue measure of measurable set $E \subset \mathbb{R}^n$
N	Heisenberg gauge, p. 21
$N_{\mathbb{G}}$	Carnot gauge, p. 21
$\nabla_H^\star$	h-reachable gradient, p. 27
co	convex hull, p. 27
$\Delta^2_{x,\tau}$	second h-differential quotient, p. 35
∂_H	h-subdifferential, p. 27
$\mathcal{L}_x$	pointed sub-Laplacian, p. 60

CHAPTER 1

MAIN NOTIONS

This chapter is an introduction to Carnot-Carathéodory spaces and Carnot groups. We start with the basic properties of vector fields and Lie brackets, then we introduce CC spaces and the CC distance induced by a family of vector fields $\mathcal{X}$. In the second part of the chapter we introduce Carnot groups and we collect some definitions and elementary results related to differentiation in this setting. We conclude the chapter with a section on *homogeneous polynomials*.

1. Sub-Riemannian spaces

After a brief introduction on vector fields and Lie brackets, we define Sub-Riemannian spaces and the CC-distance.

1.1. Vector fields and commutators. Let $\Omega \subset \mathbb{R}^n$ be an open set. A *vector field* X on Ω is a first order partial differential operator

$$X = \sum_{i=1}^{n} a_i(x)\partial_{x_i}$$

where a_i are smooth functions in Ω. Here, as usual, we identify the smooth section of the tangent bundle $(a_i(x), \ldots, a_n(x)) \in \mathbb{R}^n$ with the first order differential operator X. Throughout the thesis we will use both interpretations. Hence it makes sense to talk about the linear independence of a collection of vector fields. Let $\mathcal{X}$ be a family of m smooth vector fields $(X_1, \ldots, X_m)$ with $X_j = \sum_{i=1}^{n} a_j^i(x)\partial_{x_i}$. Fix a multi-index $I = (i_1, \ldots, i_n)$, $i_j \leq m$. We define the determinant

$$\lambda_I(x) = \det\left[X_{i_1}(x), \ldots, X_{i_n}(x)\right] = \det\left(a_{i_j}^k(x)\right).$$

If X and Y are two vector fields, then the *Lie bracket or commutator* is the operator

$$[X, Y] := XY - YX.$$

In coordinates, if $X = \sum_{i=1}^{n} a_i(x)\partial_{x_i}$ and $Y = \sum_{i=1}^{n} b_i(x)\partial_{x_i}$ then

$$[X, Y] = \sum_{k=1}^{n}\left(\sum_{i=1}^{n} a_i(x)\partial_{x_i} b_k(x) - b_i(x)\partial_{x_i} a_k(x)\right)\partial_{x_k}.$$

It is a well known fact that for every vector fields X, Y, Z the Jacoby identity holds:

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

The Lie algebra generated by $\mathcal{X}$, denoted by Lie $(X_1, \dots, X_m)$, is defined as the smallest vector space which contains $\mathcal{X}$ and it is closed under the bracket operation.

Definition 1.1.1 (Flow of a vector field) Let X be a smooth vector field of $\mathbb{R}^n$ and let $x \in \mathbb{R}^n$. We consider the Cauchy problem

$$\begin{cases} \dot{\gamma}(t) = X(\gamma(t)) \\ \gamma(0) = x \end{cases}$$

and denote it by $t \to \Phi^X(x,t)$. The mapping Φ^X defined on an open neighborhood of $\mathbb{R}^n \times \{0\}$ in $\mathbb{R}^{n+1}$ is the flow associated to X. The flow Φ^X will also define the local diffeomorphism $\Phi_t^X(\cdot) = \Phi^X(\cdot, t)$ on bounded open sets for t sufficiently small.

For every smooth vector field X on $\mathbb{R}^n$, we denote by $\exp(tX) = \Phi^X(0,t)$, where $t \in \mathbb{R}$ is sufficiently small. The following lemma is elementary, but provide a useful equivalent definition of the commutators of two vector fields.

Lemma 1.1.2 *Let X, Y be two smooth vector fields on $\mathbb{R}^n$ and suppose that $exp(\cdot X)$ and $exp(\cdot Y)$ are defined in a symmetric neighborhood of the origin $I \subset \mathbb{R}$. Then for every smooth function u on $\mathbb{R}^n$ and $x \in \mathbb{R}^n$ we have*

$$\left[\frac{d}{dt} Y(u \circ exp(-tX))\right](exp(tX)x) = [X,Y](u \circ exp(-tX))(exp(tX)x), \quad t \in I.$$

PROOF. By a direct computation we get

$$\begin{aligned} (12) \quad Y(\frac{d}{dt}(u \circ \exp(-tX)))(\exp(tX)x) + Y(u \circ \exp(-tX))(\frac{d}{dt}\exp(tX)x) = \\ - YX(u \circ \exp(-tX))(\exp(tX)x) + XY(u \circ \exp(-tX))(\exp(tX)x), \end{aligned}$$

and the proof is complete. □

Definition 1.1.3 Given a smooth vector field on $\mathbb{R}^n$, $X = \sum_{i=1}^n a_i(x)\partial_{x_i}$ we shall denote by X^* the formal adjoint to X in $L^2(\mathbb{R}^n)$, namely the operator which for all $\phi, \psi \in C_0^\infty(\mathbb{R}^n)$ satisfies

$$\int_{\mathbb{R}^n} \phi(x) X\psi(x) dx = \int_{\mathbb{R}^n} \psi X^* \phi(x) dx.$$

If X is a smooth vector field, the operator X^* can be written as

$$X^* = \sum_{i=1}^n -\partial_{x_j}(a_j(x)\cdot).$$

1.2. CC-distance. A Carnot-Carathéodory (CC) space is a manifold M with a metric induced by a fixed family of vector fields. Since we are interested in local properties we restrict our attention only to open and connected subset of $\mathbb{R}^n$. Let $\Omega \subset \mathbb{R}^n$ be an open set and let $\mathcal{X} = (X_1, \dots, X_m)$, with $m \leq n$ an X_i, $i = 1, \dots, m$ smooth vector fields on Ω. A detailed introduction to CC spaces can be found in [9], here we recall only the basic definitions.

Definition 1.1.4 A Lipschitz continuous map $\gamma : [0,T] \to \Omega$, $T \geq 0$, is *admissible* with respect to $\mathcal{X}$ if there exist measurable functions $(h_1, \ldots, h_m) : [0,T] \to \mathbb{R}^m$ such that
(i) $\gamma'(t) = \sum_{i=1}^m h_i(t) X_i(\gamma(t))$ for a.e. $t \in [0,T]$
(ii) $\|h\|_{L^\infty(0,T)} < +\infty$.
The curve γ is *subunit* if it is admissible and for a.e. $\in [0,T]$, $\sum_{j=1}^m h_j^2(t) \leq 1$.

Now we introduce the function $d : \Omega \times \Omega \to [0,+\infty]$, called *CC distance*, defined as

$$\begin{aligned} d_{CC}(x,y) = \inf\{T \geq 0 |\ &\text{there exists a subunit curve} \gamma : [0,T] \to \Omega, \\ &\text{such that } \gamma(0) = x,\ \gamma(T) = y\}. \end{aligned} \tag{13}$$

If the set on the right hand side of (13) is empty, then we set $d_{CC}(x,y) = +\infty$. In general it is false that the CC distance, with respect to $\mathcal{X}$, between two points will be finite. However, we have the following important fact, see for instance [9].

Theorem 1.1.5 *If $d(x,y) < +\infty$ for all $x, y \in \Omega$, then (Ω, d) is a metric space, called a* Carnot-Carathéodory *space.*

Given a multi-index $I = (i_1, \ldots, i_p)$, $1 \leq i_j \leq m$, we set $|I| = p$ and

$$X_{[I]} = \left[X_{i_1}, \left[\ldots, \left[X_{i_{p-1}}, X_{i_p}\right] \ldots\right]\right]. \tag{14}$$

Throughout the thesis we always assume that the family $\mathcal{X}$ of vector fields $X_1, \ldots, X_m$ satisfy the Hörmander condition, [57] : for any open bounded set $\Omega \subset \mathbb{R}^n$, and for every $x \in \mathbb{R}^n$ there exists an integer $r = r(\Omega)$ such that

$$\operatorname{span}\{X_{[I]}(x) : |I| \leq r\} = \mathbb{R}^n. \tag{15}$$

If condition (15) is satisfied the CC distance is finite, hence by Theorem 1.1.5 the space (Ω, d) is a metric space. This well know result was first proved by Chow in [47]. The metric space (Ω, d) under the assumption (15) is also called a *Sub-Riemannian* space.

Example 1.1.6 Let us fix $m < n$ and consider $\mathbb{R}^m \times \mathbb{R}^{n-m} = \mathbb{R}^n$. Denote by $(\bar{x}, x') \in \mathbb{R}^m \times \mathbb{R}^{n-m}$ the points in $\mathbb{R}^n$ and define n vector fields as follows

$$X_1 = \partial_{x_1}, \ldots, X_m = \partial_{x_m} \quad X_{m+1} = |\bar{x}|^{2k}\partial_{x_{m+1}}, \ldots, X_n = |\bar{x}|^{2k}\partial_{x_n},$$

where $|\cdot|$ denote the Euclidean norm and $k \in \mathbb{N}$, $k > 0$. Clearly the family $\mathcal{X} = (X_1, \ldots, X_n)$ satisfies the Hörmander condition in fact

$$[\underbrace{X_i, \cdots,}_{(k-1)-times} [X_i, X_{m+j}] \cdots] = (2k)!\partial_{x_{m+j}},$$

for every $0 < i \leq m$ and $0 < j \leq n - m$. The induced CC metric d on $\mathbb{R}^n$ is called the Grushin metric. This kind of sub-Riemannian metric has been introduced by Franchi and Lanconelli in [36, 37]. Notice that if k is a positive real number, the Hörmander condition is in general no longer verified by the family $X_1, \ldots, X_n$. However it is not difficult to prove that the Carnot-Carathéodory distance is well defined and finite also in this more general case, see for instance [79].

Since throughout the thesis we will need different CC-distances, first we introduce the notion of local equivalence. Following the terminology of [83], we introduce the following definition.

Definition 1.1.7 We say that two distances ρ_1 and ρ_2 in $\mathbb{R}^n$ are *equivalent*, if for every compact set $K \subset \mathbb{R}^n$, there exist $c_K \geq 1$, depending on K, such that

$$c_K^{-1}\rho_1(x,y) \leq \rho_2(x,y) \leq c_K\rho_1(x,y) \quad \text{for all} \quad x,y \in K.$$

One can prove that the distance (13) is equivalent to the following one, which was first introduced in [32].

Definition 1.1.8 For every $x, y \in \mathbb{R}^n$ we define

$$d(x,y) = \inf\{t > 0 : \text{ there exists } \gamma \in \Gamma_{x,y}(t)\}\,, \tag{16}$$

where $\Gamma_{x,y}(t)$ denotes the family of all absolutely continuous curves $\gamma : [0,t] \longrightarrow \mathbb{R}^n$ with $\gamma(0) = x$, $\gamma(t) = y$ and such that for a.e. $s \in [0,t]$ we have

$$\dot{\gamma}(s) = \sum_{j=1}^{m} a_j(s)X_j(\gamma(s)) \quad \text{and} \quad \max_{1\leq j\leq m} |a_j(s)| \leq 1\,.$$

Metric balls are defined using the following notation

$$B_{x,r} = \{z \in \mathbb{R}^n : d(z,x) < r\}, \quad D_{x,r} = \{z \in \mathbb{R}^n : d(z,x) \leq r\}$$

for any $r > 0$ and $x \in \mathbb{R}^n$.

Throughout the thesis we refer to d when speaking of Carnot-Charatheodory distance.

1.3. Sobolev spaces. Next, we introduce the anisotropic Sobolev space $W^{1,p}_{\mathcal{X}}$ with respect to the family $\mathcal{X}$. Throughout, for every open set $\Omega \subset \mathbb{R}^n$ we denote by $C_c^\infty(\Omega)$, the class of smooth functions with compact support.

Definition 1.1.9 Given an open set $\Omega \subset \mathbb{R}^n$, we define the *$\mathcal{X}$-Sobolev space* $W^{1,p}_{\mathcal{X}}(\Omega)$, with $1 \leq p \leq \infty$, as follows

$$W^{1,p}_{\mathcal{X}}(\Omega) = \left\{ f \in L^p(\Omega),\ X_j f \in L^p(\Omega),\ j = 1,\ldots,m \right\},$$

where $X_j u$ is the *distributional derivative* of $u \in L^1_{loc}(\Omega)$, namely

$$\langle X_i u, \phi\rangle = \int_\Omega u\ X_i^* \phi\, dx, \quad \phi \in C_0^\infty(\Omega),$$

and X_i^* is the formal adjoint of X_i, namely, $X_i^* = -X_i - \text{div} X_i$.

The linear space $W^{1,p}_{\mathcal{X}}(\Omega)$ is turned into a Banach space by the norm

$$\|f\|_{W^{1,p}_{\mathcal{X}}(\Omega)} := \|f\|_{L^p(\Omega)} + \sum_{j=1}^{m} \|X_i f\|_{L^p(\Omega)}\,.$$

The following Meyers-Serrin type theorem was proved in [38, 42] and holds for a more general class of vector fields than the smooth ones. We also remark that the method goes back to the work of Friedrichs [40].

Theorem 1.1.10 *Let $\mathcal{X} = (X_1, \ldots, X_m)$ be a system of locally Lipschitz vector fields and let $\Omega \subset \mathbb{R}^n$ be an open set. If $1 \leq p < \infty$ then*

$$C^\infty(\Omega) \cap W^{1,p}_{\mathcal{X}}(\Omega) \text{ is dense in } W^{1,p}_{\mathcal{X}}(\Omega).$$

In view of Theorem 1.1.10, the following definition is natural.

Definition 1.1.11 Let $\mathcal{X}$ be a family of Lipschitz continuous vector fields. If $1 \leq p < \infty$ then we define

$$W^{1,p}_{\mathcal{X},0}(\Omega) := \overline{C^\infty_0(\Omega)}^{W^{1,p}_{\mathcal{X}}}.$$

As in the Euclidean case Lipschitz continuous functions with respect to d are $W^{1,\infty}_{\mathcal{X}}$ functions, see [39, 43].

Theorem 1.1.12 *Let $\Omega \subset \mathbb{R}^n$ be a bounded, open set with $d_\Omega = \sup_{x,y\in\Omega} d(x,y) < \infty$. If for a given function $u : \Omega \to \mathbb{R}$ we have for some constant $C > 0$*

$$|u(x) - u(y)| \leq Cd(x,y), \quad \text{for } x, y \in \Omega,$$

then $u \in W^{1,\infty}_{\mathcal{X}}(\Omega)$.

2. Stratified groups

In this section we introduce a particular class of CC spaces, called Carnot groups. These groups are the main object of study in the thesis. First, we recall the definition of Lie group and Lie algebra.

Definition 1.2.1 A Lie group is a differentiable manifold $\mathbb{G}$ endowed with a differentiable group structure, namely the product $(x,y) \mapsto x \cdot y$ and the inversion $x \mapsto x^{-1}$ are smooth maps. We denote by 0 the identity of the group.

Definition 1.2.2 A vector field $X \in \Gamma(T\mathbb{G})$ is left invariant if

$$dl_g(X(0)) = X(g) \quad \text{for every} g \in \mathbb{G},$$

where the map l_g is the diffeomorphism $l_g : \mathbb{G} \to \mathbb{G}$ defined as $l_g(x) = gx$, $x \in \mathbb{G}$.

Notice that if we look at vector fields as differential operators, the left invariance of $X \in \Gamma(T\mathbb{G})$ is equivalent to the following

$$X(u \circ l_g) = (Xu) \circ l_g, \qquad \text{for any } u \in C^\infty(\mathbb{G}).$$

Let X_j left invariant vector fields on $\mathbb{G}$, given a multi-index $\alpha \in \mathbb{N}^p$ define X^α as $X^\alpha = X_1^{\alpha_1} \cdots X_p^{\alpha_p}$. Then for every smooth function u on $\mathbb{G}$ and every $x, g \in \mathbb{G}$, we have

$$X^\alpha(u(gx)) = (X^\alpha u)(gx). \tag{17}$$

By definition of left invariant vector field, $X_j(gx) = dl_g X_j(x)$ moreover given a smooth function u, $X_j u(x) = \langle du(x), X_j(x)\rangle$. Thus for every $j = 1, \ldots, m_1$ we have

$$\begin{aligned} X_j u(gx) &= \langle d(u \circ l_g)(x), X_j(x) \rangle \\ &= \langle du(gx) dl_x, X_j(g) \rangle \\ &= \langle du(gx), X_j(gx) \rangle = \left(X_j u \right)(gx). \end{aligned}$$

Clearly the previous equality implies (17).

Definition 1.2.3 We say that a finite dimensional vector space $\mathfrak{g}$ is a Lie algebra if there exists an antisymmetric bilinear map

$$\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}, \quad (X, Y) \to [X, Y],$$

such that the Jacobi identity holds, i.e. for every $X, Y, Z \in \mathfrak{g}$

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

A linear subspace $\mathfrak{a} \subset \mathfrak{g}$ is a Lie sub algebra of $\mathfrak{g}$ if $[X, Y] \in \mathfrak{a}$ for any $X, Y \in \mathfrak{a}$.

Let $\mathcal{G}$ be the set of all vector fields $X \in \Gamma(T\mathbb{G})$ which are left invariant. Since for every $X, Y \in \Gamma(T\mathbb{G})$, and every diffeomorphism $f : \mathbb{G} \to \mathbb{G}$, we have

$$f_\star[X, Y] = [f_\star X, f_\star Y],$$

it follows that $\mathcal{G}$ is a Lie subalgebra of $\Gamma(T\mathbb{G})$. Hence $\mathfrak{g} = \mathcal{G}$ is called the Lie algebra associated to $\mathbb{G}$.

Now we introduce the exponential map in Lie groups. Consider the following system of O.D.E.

$$\left\{ \begin{array}{lcl} \partial_t \phi(x, t) &=& X(\phi(x, t)) \\ \phi(x, 0) &=& x \end{array} \right. \tag{18}$$

where $X \in \mathcal{G}$. The flow ϕ is defined on all $\mathbb{R}$, this is a consequence of the left invariance of X. More precisely if $\phi(x, \cdot)$ is defined on some interval $[0, b]$ then also $\phi(x, \cdot) = x \cdot \phi(e, \cdot)$ is defined on $[0, b]$. Moreover

$$\phi(\phi(e, \frac{b}{2}), t) = \phi(e, \frac{b}{2})\phi(e, t) = \phi(e, t + \frac{b}{2}),$$

hence $\phi(e, \cdot)$ can be extended on $[0, \frac{3b}{2}]$. This argument can be repeated analogously on the left side.

Definition 1.2.4 For any $X \in \mathcal{G}$ we define the map $\mathrm{Exp} : \mathcal{G} \to \mathbb{G}$ as

$$\mathrm{Exp}(X) := \phi(e, 1),$$

where ϕ is the flow associated to the system (18).

Definition 1.2.5 Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{a}, \mathfrak{b}$ two subspaces. We denote by $[\mathfrak{a}, \mathfrak{b}]$ the subspace of $\mathfrak{g}$ generated by all linear combinations of elements $[X, Y]$, where $X \in \mathfrak{a}$ and $Y \in \mathfrak{b}$. For each $k \in \mathbb{N} \setminus \{0\}$ we define by induction the following subspaces

$$\mathfrak{g}^1 = \mathfrak{g}, \qquad \mathfrak{g}^{k+1} = [\mathfrak{g}^k, \mathfrak{g}].$$

The family $(\mathfrak{g}^k)_{k\geq 1}$ is called the descending central sequence of $\mathfrak{g}$. If there exists a positive integer s such that $g^{s+1} = 0$ we say that $\mathfrak{g}$ is a *nilpotent* Lie algebra of step s. A Lie group $\mathbb{G}$ is said to be *nilpotent* if its Lie algebra is nilpotent.

Theorem 1.2.6 (1.127 in [62]) *If $\mathbb{G}$ is a simply connected nilpotent Lie group with Lie algebra $\mathfrak{g}$, then the exponential map is a diffeomorphism of $\mathfrak{g}$ onto $\mathbb{G}$.*

As a consequence of the preceding theorem we can define the inverse map $\ln = \mathrm{Exp}^{-1}$ in simply connected nilpotent groups.

Now we state the important Baker-Campbell-Hausdorff formula, in the sequel called BCH formula, were a relation between vectors of the algebra and the product of their corresponding exponentials is established.

Theorem 1.2.7 (B.22 in [62]) *Let $\mathbb{G}$ be a Lie group with Lie algebra $\mathcal{G}$. Then for all X, Y sufficiently close to 0 in $\mathfrak{g}$, $Exp(A)Exp(B) = Exp(C)$, where*

$$C = A + B + C_2 + \ldots + C_n + \ldots$$

is a convergent series which $H_2 = \frac{1}{2}[X, Y]$ and H_n is finite linear combination of expressions $(adX_1)\cdots(adX_{n-1})X_n$ with each X_j equal to either X or Y. The particular linear combinations that occur may be taken to be independent of $\mathbb{G}$, as well as of X and Y.

Explicit computation shows that

$$C_3 = \frac{1}{12}\left[X,[X,Y]\right] + \frac{1}{12}\left[Y,[Y,X]\right], \qquad C_4 = -\frac{1}{24}\left[X,[Y,[X,Y]]\right]. \tag{19}$$

Definition 1.2.8 We say that a Lie algebra $\mathfrak{g}$ is graded if it can be decomposed as a direct sum of vector spaces as

$$\mathfrak{g} = V_1 \oplus \cdots \oplus V_s, \quad s \in \mathbb{N},$$

with $V_{i+1} \subset [V_i, V_1]$ for any $i \in \mathbb{N} \setminus \{0\}$, and $V_j = \{0\}$ for any $j > s$. A Lie group whose Lie algebra is graded is called graded group. The previous decomposition is called the grading of the group. If $\mathbb{G}$ is a simply connected group associated to the algebra $\mathfrak{g}$, we define for every $x \in \mathbb{G}$ the subspace of degree k at x as follows

$$H_x^k\mathbb{G} = \left\{X(x) \mid X \in V_j\right\} \subset T_x\mathbb{G}.$$

We denote by $H_k = \mathrm{Exp}V_k \subset \mathbb{G}$ the space of elements in $\mathbb{G}$ of degree $k = 1, \ldots, s$.

Remark 1.2.9 Notice that any graded group is nilpotent and the positive integer s is the step of the group. This follows because we assume that the group is finite dimensional.

Definition 1.2.10 *Let $\mathfrak{g}$ be a graded algebra. Then for every $t \leq 0$, we define the map $\delta_t : \mathfrak{g} \to \mathfrak{g}$ as*

$$\delta_t(v) = \sum_{i=1}^{s} r^i v_i,$$

where $v = \sum_{i=1}^{s} v_i$, $v_i \in V_i$. If $t < 0$ we define δ_t as

$$\delta_t(v) = -\delta_{|t|}(v).$$

Definition 1.2.11 A graded algebra $\mathfrak{g}$ with grading

$$\mathfrak{g} = V_1 \oplus \cdots \oplus V_s,\ s \in \mathbb{N},$$

is called stratified if for any $i \in \mathbb{N} \setminus \{0\}$ we have $V_{i+1} = [V_i, V_1]$, where $V_j = \{0\}$ for any $j > s$. A Lie group whose Lie algebra is stratified is called stratified group or Carnot group.

A stratified group can be also thought of as a graded vector space $\mathbb{G} = H_1 \oplus \cdots \oplus H_\iota$ with a polynomial group operation given by the Baker-Campbell-Hausdorff formula. The left invariant vector fields of V_j are exactly the ones that at the origin take values in H_j. Recall that the origin is exactly the unit element of the group. A scalar product on $\mathbb{G}$ will be understood, assuming that all subspaces H_j are orthogonal. We denote by $\pi_j : \mathbb{G} \longrightarrow H_j$ the associated orthogonal projections. For every $s = 1, \ldots \iota$, we fix a basis $(e_{m_{s-1}+1}, \ldots, e_{m_s})$ of H_s, then

$$\sum_{i=m_{s-1}+1}^{m_s} x_i e_i \in H_s \quad \text{and} \quad x = \sum_{s=1}^{\iota} \sum_{i=m_{s-1}+1}^{m_s} x_i\, e_i.$$

Throughout the thesis, for every Carnot group $\mathbb{G}$ we also fix

$$\left(X_{m_{s-1}+1}, \ldots, X_{m_s}\right) \tag{20}$$

as the basis of V_s such that, with respect to the coordinates (x_j), X_j is e_j.

We define the homogeneous dimension of $\mathbb{G}$ as

$$Q = \sum_{i=1}^{s} i \dim\left(V_i\right).$$

2.1. The Heisenberg groups. The first examples of non abelian stratified Lie groups are the so called Heisenberg groups $\mathbb{H}^n$, defined as follows.

Definition 1.2.12 A Lie algebra $\mathfrak{h}_{2n+1} = V_1 \oplus V_2$ with basis $(X_1, \ldots, X_n, Y_1, \ldots, Y_n)$, T respectively of V_1 and V_2, that satisfies relations

$$\left[X_i, X_j\right] = 0, \quad \left[Y_i, Y_j\right] = 0, \quad \left[X_j, Y_j\right] = T,$$

for every $j, i = 1, \ldots, n$ is called *Heisenberg algebra*. The *Heisenberg group* $\mathbb{H}^n$ is the simply connected nilpotent Lie group associated to $\mathfrak{h}_{2n+1}$.

Using exponential coordinates on $\mathbb{H}^n$ we identify the group as the algebra $\mathfrak{h}_{2n+1}$. The group operation is given by the Backer-Campbell-Hausdorff formula as

$$x \cdot y = x_1 + y_1 + x_2 + y_2 + \beta(x_1, y_1)$$

where $\beta : V_1 \times V_1 \to V_2$ is non-degenerate, bilinear and skew-symmetric. Moreover we have used the unique decomposition $x = x_1 + x_2$, $x_1 \in V_1$ and $x_2 \in V_2$.

Remark 1.2.13 Notice that, for every $t \in V_2$ and $x_1 \in V_1$, with $x_1 \neq 0$, there exists $y_1 \in V_1$ such that $\beta(x_1, y_1) = t$. In fact, β is a non-degenerate bilinear form.

2.2. Carnot groups as CC spaces. Let $\mathbb{G} = H_1 \oplus \cdots \oplus \mathbb{H}_\iota$ be a Carnot group, with stratified Lie algebra $\mathfrak{g} = V_1 \oplus \cdots \oplus V_\iota$, $\iota \geq 2$. Fix an orthonormal basis $(X_1, \ldots, X_m)$ of V_1 as in (20). Since the algebra $\mathfrak{g}$ is stratified the Hörmander condition (15) is clearly verified hence the basis $(X_1, \ldots, X_m)$ induces a CC metric d on $\mathbb{R}^n$. Moreover, since the fields inducing the metric space d are left invariant and 1-homogeneous one can prove the following remarkable facts:

$$d(\delta_\lambda x, \delta_\lambda y) = \lambda d(x,y)$$
$$d(g \cdot x, g \cdot y) = d(x,y),$$

for all $x, y, g \in \mathbb{G}$ and $\lambda > 0$.

Now we define a class of distances that are compatible with the geometry of stratified groups.

Definition 1.2.14 Let $\mathbb{G}$ be a graded group. A homogeneous distance on $\mathbb{G}$ is a continuous map $\eta : \mathbb{G} \times \mathbb{G} \to [0, +\infty[$ that makes $(\mathbb{G}, d)$ a metric spaces and has the following properties:
(i) $\eta(x,y) = \eta(ux, uy)$ for every $x, y, u \in \mathbb{G}$,
(ii) $\eta(\delta_t x, \delta_t y) = t\eta(x,y)$ for every $t > 0$.

We denote by $\eta(x)$ or $\|x\|$ the homogeneous norm of x induced by the distance η , i.e. $\|x\| := \eta(x, 0)$. As in [35], open balls with respect to the Carnot-Carathéodory distance d will be denoted by $B_{x,r}$. Balls of radius r centered ad 0 will be denoted simply by B_r. The symbols $D_{x,r}$ and D_r denote closed balls with analogous meanings.

Proposition 1.2.15 (2.3.37 in [72]) *Let η and δ be homogeneous distances on $\mathbb{G}$. Then there exist two positive constants C_1 and C_2 such that for any $x, y \in \mathbb{G}$ we have*

$$C_1 \delta(x,y) \leq \eta(x,y) \leq C_2 \delta(x,y).$$

Proposition 1.2.16 (2.3.39 in [72]) *Let d be the CC-distance induced by $(X_1, \ldots, X_1)$ of a Carnot group $\mathbb{G}$. Then d is an homogeneous distance.*

Carnot groups are nilpotent and so unimodular, thus the right and the left Haar measures coincide, up to constants. We shall denote by $\mathcal{H}^k$ the Hausdorff k-dimensional measure associated to the Carnot-Carathéodory distance on $\mathbb{G}$. The Hausdorff measure $\mathcal{H}^Q$ by the left translation and scaling invariance of the CC distance is an Haar measure on $\mathbb{G}$ and it will be also denoted by $\mathrm{vol}_\mathbb{G}$. Moreover, in exponential coordinates, this measure coincides with a constant multiple of the Lebesgue measure on $\mathbb{R}^n$. From these considerations it follows that

$$\mathrm{vol}_\mathbb{G}(\delta_\lambda(A)) = \lambda^Q \mathrm{vol}_\mathbb{G}(A),$$

for all Borel sets $A \subseteq \mathbb{G}$. In particular

$$\mathrm{vol}_\mathbb{G}(B_\rho(g)) = c\rho^Q \tag{21}$$

for some constant $c > 0$ independent of g.

The following proposition is a well known fact, see for instance Lemma 1.40 in [35].

Proposition 1.2.17 *Let $\mathbb{G}$ be a stratified group and let $(e_1, \ldots, e_{m_1})$ be an orthonormal basis of H_1. Then there exists a positive integer γ along with a vector of integers $(i_1, \ldots, i_\gamma) \in 1, \ldots, {m_1}^\gamma$ and* *a bounded open set $U \subset \mathbb{R}^\gamma$ such that*

$$D_{0,1} \subset \left\{ \prod_{s=1}^{\gamma} a_s e_{i_s} \,|\, (a_1, \ldots, a_\gamma) \in U \right\}. \tag{22}$$

Remark 1.2.18 The inclusion (22) can be always established by a rescaling argument, once we know that $\{\prod_{s=1}^{\gamma} a_s w_{i_s} \,|\, (a_s) \subset U\}$ is a neighborhood of the origin.

Definition 1.2.19 Let $U \subset \mathbb{R}^\gamma$ and $(i_1, \ldots, i_\gamma) \in \{1, \ldots, m_1\}^\gamma$ be as in Proposition 1.2.17. Thus, we define

$$W = \{ \prod_{s=1}^{\gamma} a_s w_{i_s} \,|\, (a_s) \subset U \}, \quad \text{and} \quad M = \sup_{y \in W} \|y\|. \tag{23}$$

Definition 1.2.20 Let γ be the positive integer defined in Proposition 1.2.17. Then we introduce the mapping $F : U \to \mathbb{G}$ as follows

$$F(a_1, \ldots, a_\gamma) = \prod_{s=1}^{N} a_s e_{i_s},$$

Notice that we have $F(\lambda a) = \delta_\lambda F(a)$ for all $\lambda > 0$ and $a \in \mathbb{R}^N$.

2.3. Calculus on stratified groups.

Definition 1.2.21 Let $L : \mathbb{G} \to \mathbb{R}$. We say that L is homogeneous if $\delta_t(Lx) = L(\delta_t x)$ for every $t > 0$.

Definition 1.2.22 We say that a map $L : \mathbb{G} \to \mathbb{R}$ is an horizontal linear map, briefly *h-linear* map, if it is an homogeneous Lie group homomorphism. We denote by $HL(\mathbb{G}, \mathbb{R})$ the space of all h-linear maps.

Remark 1.2.23 Notice that the previous definition is equivalent to the following one: a map $L : \mathbb{G} \longrightarrow \mathbb{R}$ is h-linear if and only if $L : \mathbb{G} \longrightarrow \mathbb{R}$ is linear and $L(x) = L(\pi_1(x))$, for every $x \in \mathbb{G}$.

Now we introduce the concept of h-differential for real valued functions defined on $\mathbb{G}$, see also the seminal paper of Pansu 1.2.25.

Definition 1.2.24 We say that $u : \Omega \longrightarrow \mathbb{R}$ is *h-differentiable* at $x \in \Omega$, if there exists an h-linear mapping $L : \mathbb{G} \longrightarrow \mathbb{R}$ such that $u(xz) = u(x) + L(z) + o(\|z\|)$. Notice that L is unique and we denote L by $d_H u(x)$, and its associated vector with respect to the scalar product, that we call *horizontal gradient* is denoted by $\nabla_H u(x)$.

Theorem 1.2.25 ([85]) *Let $\Omega \subset \mathbb{G}$ be an open set and let $u : \Omega \to \mathbb{R}$ be a Lipschitz function with respect to the CC distance. Then u is a.e. h-differentiable in Ω.*

Following [4] we introduce the X-derivative in a Carnot group $\mathbb{G}$. Given a vector field $X \in \Gamma(T\mathbb{G})$ we define the divergence $\mathrm{div}X$ in the sense of distributions as follows:

$$\int_{\mathbb{G}} Xu \, d\,\mathrm{vol}_{\mathbb{G}} = -\int_{\mathbb{G}} u \, \mathrm{div}X \, d\,\mathrm{vol}_{\mathbb{G}} \qquad \forall u \in C_c^\infty(\mathbb{G}). \tag{24}$$

Definition 1.2.26 Let $u \in L^1_{\rm loc}(\mathbb{G})$ and let $X \in \Gamma(T\mathbb{G})$ be divergence-free. We denote by Xu the distribution

$$\langle Xu, v\rangle := -\int_{\mathbb{G}} uXv\, d\,\mathrm{vol}_{\mathbb{G}}, \qquad v \in C^\infty_c(\mathbb{G}).$$

If $f \in L^1_{\rm loc}(\mathbb{G})$, we write $Xu = f$ if $\langle Xu, v\rangle = \int_{\mathbb{G}} vf\, d\,\mathrm{vol}_{\mathbb{G}}$ for all $v \in C^\infty_c(\mathbb{G})$. Analogously, if μ is a Radon measure on $\mathbb{G}$, we write $Xu = \mu$ if $\langle Xu, v\rangle = \int_{\mathbb{G}} v\, d\mu$ for all $v \in C^\infty_c(\mathbb{G})$.

Given $X \in \Gamma(T\mathbb{G})$ we denote by $\varphi_X : \mathbb{G} \times \mathbb{R} \to \mathbb{G}$ the flow of X, assuming that X is sufficiently smooth to ensure its global existence and uniqueness.

Theorem 1.2.27 ([4]) *Let $u \in L^1_{\rm loc}(\mathbb{G})$ be satisfying $Xu = 0$ in the sense of distributions. Then, for all $t \in \mathbb{R}$, $u = u \circ \Phi_X(\cdot, t)$ $vol_{\mathbb{G}}$-a.e. in $\mathbb{G}$.*

One can prove that all $X \in \mathfrak{g}$ are divergence free, using the invariance of the right Haar measure with respect to the flow of X (see Remark 2.13 of [4] for details).

Definition 1.2.28 Let $\phi \in C^\infty_c(\mathbb{G})$ be a nonnegative function, whose support is contained in the unit open ball of $\mathbb{G}$ with respect to the fixed homogeneous norm. For every $\varepsilon > 0$, we set $\phi_\varepsilon(x) = \varepsilon^{-Q}\phi(\delta_{\frac{1}{\varepsilon}} x)$. We say that $\{\phi_\varepsilon\}_{\varepsilon>0}$ is a *family of mollifiers*. For all $x \in \mathbb{G}$, we define the functions $\Phi_{x,\varepsilon} : \mathbb{G} \to \mathbb{R}$ as $\Phi_{x,\varepsilon}(y) = \phi_\varepsilon(xy^{-1})$.

Consider two measurable functions f, g on $\mathbb{G}$, their convolution is defined by

$$f \star g(x) = \int_{\mathbb{G}} f(y)g(y^{-1}x)dy, \tag{25}$$

provided the integral on the right hand side converges.
As in the Euclidean setting, it is useful to have a procedure which allows to regularize a function, see Proposition 1.20 [35].

Lemma 1.2.29 *Let $f \in L^1_{loc}(\Omega)$, and define $f_\varepsilon(x) := \phi_\varepsilon \star f(x)$. Then*
(i) $f_\varepsilon \in C^\infty(\Omega_\varepsilon)$,
(ii) $f_\varepsilon \to f$ uniformly in Ω as $\varepsilon \downarrow 0$ provided f is continuous.

Definition 1.2.30 Let $\Omega \subset \mathbb{G}$ be an open set. A function $u : \Omega \to \mathbb{R}$ is in the Folland-Stein class $\Gamma^k(\Omega)$ if it is continuous along with all its (distributional) derivatives $X^I u$, $I = (i_1, \ldots, i_k)$ and $i_l \in \{1, \ldots, m\}$.

If $u \in \Gamma^1(\Omega)$ then we define the *horizontal gradient* as

$$\nabla_H u := (X_1 u, \ldots, X_m u). \tag{26}$$

Definition 1.2.31 Let $u \in L^1(\Omega)$, we say that u has *h-bounded second variation* and write $u \in BV^2_H$ if the X-derivatives $X_i u$, $X_i X_j u$ are finite Radon measures for every $i, j = 1, \ldots, m$. If $u \in L^1_{loc}(\Omega)$ and $X_i u$, $X_i X_j u$ are Radon measures we write $u \in BV^2_{H,loc}(\Omega)$.

2.4. Homogeneous polynomials. In the previous sections we have already introduced polynomials of homogeneous order 1, namely h-linear maps, see Definition (1.2.22). Now we define polynomials on $\mathbb{G}$ of arbitrary homogeneous order and we prove some technical lemmas that we will need in Chapter 4.

Definition 1.2.32 We say that $P : \mathbb{G} \to \mathbb{R}$ is a *polynomial* on $\mathbb{G}$, if with respect to some fixed graded coordinates we have $P(x) = \sum_{\alpha \in \mathcal{A}} c_\alpha x^\alpha$, under the convention $x^\alpha = \prod_{i=1}^n x_i^{\alpha_i}$, and $x_j^0 = 1$, where $\mathcal{A} \subset \mathbb{N}^n$ is a finite set.

The *homogeneous degree* of P is the integer $h\text{-deg}(P) = \max\{d(\alpha),\ \alpha \in \mathcal{A}\}$, where $d(\alpha) = \sum d_i \alpha_i$, and $d_i = s$ if $m_{s-1} + 1 \le i \le m_s$.

By the previous definitions, any polynomial P can be decomposed into the sum of its *j-homogeneous parts*, denoted by $P^{(j)}$, hence

$$P = \sum_{0 \le j \le h\text{-deg}P} P^{(j)}.$$

A polynomial is *j-homogeneous* if it coincides with its j-homogeneous part.

There is a canonical way to realize abstract stratified Lie algebras as algebras of vector fields over $\mathbb{R}^n$. Here we state some general facts on Carnot groups, based on the B-C-H formula.

For a proof of the following Proposition see [96], Chapter 12.

Proposition 1.2.33 *In exponential coordinates the group product has the form*

$$x \cdot y = x + y + \mathcal{Q}(x, y) \qquad \forall x, y \in \mathbb{R}^n$$

where $\mathcal{Q} = (\mathcal{Q}_1, \ldots, \mathcal{Q}_n) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ *and each* $\mathcal{Q}_i$ *is a homogeneous polynomial of degree* α_i *with respect to the intrinsic dilations of* $\mathbb{G}$, *i.e.*

$$\mathcal{Q}_i(\delta_\lambda x, \delta_\lambda y) = \lambda^{\alpha_i} \mathcal{Q}_i(x, y) \qquad \forall x, y \in \mathbb{G}.$$

Moreover $\forall x, y \in \mathbb{G}$

$$\begin{array}{rcl} \mathcal{Q}_1(x,y) & = \ldots = & \mathcal{Q}_m(x,y) = 0, \\ Q_j(x,0) = Q_j(0,y) = 0 & \text{and} & Q_j(x,x) = Q_j(x,-x) = 0, \quad m < j \le n; \end{array}$$

and for $i > m$, Q_i *is a sum of terms each of which contains a factor* $(x_j y_l - x_l y_j)$ *for some* $1 < j, l < i$. *Finally, if* $m_{i-1} < j \le m_i$ *and* $2 \le i$, *then*

$$Q_j(x, y) = \mathcal{Q}_j(x_1, \ldots, x_{m_{i-1}}, y_1 \ldots, y_{m_{i-1}}).$$

Exponential coordinates characterize the left invariant vector fields X_j as vector fields on $\mathbb{R}^n$.

Proposition 1.2.34 *The vector fields* X_j *have polynomial coefficients, moreover if* $m_{l-1} < j \le m_l$ *and* $1 \le l \le s$,

$$X_j(x) = \partial_j + \sum_{i > m_l}^n q_{i,j}(x) \partial_i,$$

where $q_{i,j}(x) = \frac{\partial \mathcal{Q}_i}{\partial y_j}(x, y)|_{y=0}$, *so that if* $m_{l-1} < i \le m_l$ *then* $q_{i,j}(x) = q_{i,j}(x_1, \ldots, x_{m_{l-1}})$ *and* $q_{i,j}(0) = 0$.

As in [35], given $a \in \mathbb{N}$, we shall denote by $\mathcal{P}_a$ the space of polynomials of homogeneous degree $\le a$. Moreover, by Proposition 1.25 in [35], $\mathcal{P}_a$ is invariant under left translations. Given a multi-index $I = (i_1, \dots, i_n)$, $1 \le i_j \le m_1$ set

$$X^I = X_{i_1} \cdots X_{i_n}$$

and $|I| = n$.

Proposition 1.2.35 (1.30 in [35]) *Take $a \in \mathbb{N}$, and let μ be the dimension of $\mathcal{P}_a$. Then the map*

$$P \to (X^I P(0))_{|I| \le a},$$

is a linear isomorphism from $\mathcal{P}_a$ to $\mathbb{C}^\mu$.

Remark 1.2.36 Let P be a polynomial of homogeneous degree at most 2, and suppose that $P(0) = p_0$ and $X_i P(x) = l_i(x)$, for $i = 1, \dots, m_1$ where $l_i : \mathbb{G} \to \mathbb{R}$ are h-linear maps. Clearly we can compute $(X^\alpha P)(0)$ for each multi-index α, $|\alpha| \le 2$, then by the previous proposition P is uniquely determined.

We are interested to find the explicit isomorphism of the previous proposition in the case of real polynomials of homogeneous degree less than or equal to two.

Lemma 1.2.37 *Let P be a 2-homogeneous polynomial*

$$P(x) = \frac{1}{2} \sum_{1 \le i,j \le m_1} c_{ij}\, x_i x_j + \sum_{s=m_1+1}^{m_2} c_s\, x_s$$

Then the following formula holds

$$P(x) = \langle \nabla_{V_2} P, x \rangle + \frac{1}{2} \langle \nabla_H^2 P x, x \rangle, \tag{27}$$

here $\langle \nabla_{V_2} P, x \rangle = \sum_{j=m_1+1}^{m_2} X_j P\, x_j$, and $\nabla_H^2 = \sum_{i,j} \frac{1}{2} \{X_i X_j + X_j X_i\}$, see also Definition 2.2.4.

PROOF. Let us consider, with respect to the same system of graded coordinates, the left invariant vector fields

$$X_j = \partial_{x_j} + \sum_{l=m_{d_j}+1}^{n} a_j^l(x) \partial_{x_l}$$

for $j = 1, \dots, n$, where $a_j^l(x)$ are $(d_l - d_j)$-homogeneous polynomial.
Since $\nabla_{V_2} P = (X_{m_1+1} P, \dots, X_{m_2} P)$ and $\nabla^2 P$ are 0-homogeneous it follows that they are constant. The explicit expression of X_j immediately yields $X_j P = c_j$ for all $j = m_1 + 1, \dots, m_2$. Hence, it remains to prove that

$$\frac{c_{ij} + c_{j\,i}}{2} = \frac{X_i X_j P + X_j X_i P}{2} \tag{28}$$

for $1 \le i, j \le m_1$. First we observe that

$$X_j(x) = \partial_{x_j} + \sum_{l=m_1+1}^{m_2} \sum_{i=1}^{m_1} a_j^{li}\, x_i\, \partial_{x_l} + \sum_{l=m_2+1}^{n} a_j^l(x)\, \partial_{x_l} \tag{29}$$

since $a^l_j(x) = \sum_{i=1}^{m_1} a^{li}_j\, x_i$ is 1-homogeneous for $d_l = 2$ and $d_j = 1$. Taking into account the previous expression, we arrive at the following

$$X_j P(x) = \frac{1}{2}\sum_{i=1}^{m_1}(c_{ij} + c_{j\,i})\, x_i + \sum_{i=1}^{m_1}\sum_{l=m_1+1}^{m_2} X_l P\, a^{li}_j x_i$$

that immediately yields

$$X_i X_j P = \frac{c_{ij} + c_{j\,i}}{2} + \sum_{l=m_1+1}^{m_2} X_l P\ a^{li}_j\,. \tag{30}$$

Finally, formula (28) follows by the equality $a^{li}_j = -a^{lj}_i$. This is in turn a consequence of the Baker-Campbell-Hausdorff formula for the second order bilinear terms. □

Definition 1.2.38 Suppose $x \in \mathbb{G}$, $a \in \mathbb{N}$ and f a function whose (distributional) derivatives $X^I f$ are continuous functions in a neighborhood of x for $|I| \le a$. We define the *left Taylor polynomial* of f at x of homogeneous degree a as the unique $P \in \mathcal{P}_a$, such that $X^I P(0) = X^I f(x)$ for all $|I| \le a$.

Theorem 1.2.39 (Stratified Taylor Inequality, 1.42 in [35]) *For each positive integer k there is a constant C_k such that for all continuous function f whose distributional derivatives $X^I f$ are continuous functions and for all $x, y \in \Omega$,*

$$|f(xy) - P_x(y)| \le C_k \|y\|^k \eta(x, b^k\|y\|),$$

where P_x is the left Taylor polynomial of f at x of homogeneous degree k, b is a constant depending only on $\mathbb{G}$, and for $r > 0$,

$$\eta(x,r) = \sup_{\|z\|\le r, |I|=k} \left| X^I f(xz) - X^I f(x) \right|,$$

where $X^I = X_{i_1}\cdots X_{i_l}$, for a certain l dependent on I and $(i_1,\dots,i_l) \in \{1,\dots,m_1\}^l$.

Lemma 1.2.40 *Let $P : \mathbb{G} \to \mathbb{R}$ be a polynomial of homogeneous degree at most 2. Let $P^{(2)}(x)$ be the 2-homogeneous part of P, and define*

$$\lambda = \max_{\|w\|=1} |P^{(2)}(w)|.$$

Consider $P(xh)$ as a function of $h \in \mathbb{G}$, then we have the following inequality

$$P(xh) \ge P(x) + \langle \nabla_H P(x), h\rangle - \lambda \|h\|^2.$$

PROOF. For every $1 \le i, j \le m_1$, we have $X_i X_j P = X_i X_j (P(xh)) = c_{i,j}$ for every $x, h \in \mathbb{G}$. This is a consequence of the following general fact, given a smooth function u and X, a left invariant vector field on $\mathbb{G}$, then $X(u(xh)) = (Xu)(xh)$. Consider $P(xh)$ as a function of h, applying Theorem 1.2.39 we get a polynomial $P_x(h)$ such that

$$P(xh) = P_x(h) + o(\|h\|^2).$$

Notice that by the left translation invariance of $\mathcal{P}_2$, $P(xh)$ as a function of h is a polynomial of homogeneous degree at most 2, hence $P(xh) = P_x(h)$. Clearly $P_x^{(0)}(h) = P(x)$ and $P_x^{(1)}(h) = \langle \nabla_H P(x), h \rangle$, as a consequence

$$P(xh) - P(x) - \langle \nabla_H P(x), h \rangle = P_x^{(2)}(h). \tag{31}$$

By (31) and previous considerations it follows that

$$c_{i,j} = X_i X_j P(h) = X_i X_j P^{(2)}(xh) = X_i X_j P_x^{(2)}(h), \quad i,j = 1, \ldots, m_1.$$

Moreover all the other derivatives of $P_x^{(2)}$ are zero, thus we can conclude that $P_x^{(2)}(h) = P^{(2)}(h)$ by Proposition 1.2.35. Finally we get

$$\begin{aligned} P(xh) &= P(x) + \langle \nabla_H P(x), h \rangle + P^{(2)}(h) \\ &\geq P(x) + \langle \nabla_H P(x), h \rangle - \lambda \|h\|^2. \end{aligned}$$

□

2.5. Sub-Laplacians and sub-mean formulas on Carnot groups. In this section we present some basic facts on sub-Laplacians on Carnot groups, with a particular interest on the operator $\Delta_H = \sum_{i=1}^m X_i^2$, where $X_1, \ldots, X_m$ is the basis of V_1 defined in 20. Our reference for this section is the book [14] where this topic is covered in full generality. Some results presented here will be useful in Chapter 3 where we will give a simple proof of the $L^\infty - L^1$ estimates for h-convex functions using a sub-mean estimates of Δ_H-subsolutions, see Theorem 2.2.10 (a). Also in Chapter 5 we will need a representation formula with respect to the fundamental solution of Δ_H.

Definition 1.2.41 Any operator

$$\mathcal{L} := \sum_{i=1}^m Y_i^2,$$

where $Y_1, \ldots, Y_m$ is a basis of the horizontal layer $\mathbb{V}_1$, is called a sub-Laplacian on $\mathbb{G}$. Let $(X_1, \ldots, X_m)$ be the basis of V_1 defined in (20), we set

$$\Delta_H = \sum_{i=1}^m X_i^2. \tag{32}$$

For every open set $\Omega \subset \mathbb{G}$, $\mathcal{D}(\Omega)$ corresponds to $C_c^\infty(\Omega)$ topologized in the standard way, where Ω is an open set of a stratified group thought of as a differentiable manifold. We denote by $\mathcal{D}'(\Omega)$ the topological vector space of distributions on Ω.
We say that a distribution $\tau \in \mathcal{D}'(\Omega)$ is homogeneous of degree α if for every $\phi \in \mathcal{D}(\Omega)$ and $r > 0$ we have $\langle \tau, \phi \circ \delta_r \rangle = r^{-\alpha - Q} \langle \tau, \phi \rangle$, where $Q = \sum_{i=1}^\iota i \dim V_i$ is the so called homogeneous dimension of $\mathbb{G}$.

Definition 1.2.42 ([34]) Let consider the sub-Laplacian $\Delta_H = \sum_{i=1}^m X_i^2$ on $\mathbb{G}$. A distribution Γ on Ω is a fundamental solution for Δ_H if :
(i) $\Gamma \in C^\infty(\mathbb{G} \setminus \{0\})$,
(ii) $\Gamma \in L^1_{loc}(\mathbb{R}^n)$ and $\Gamma(x) \to 0$ when $d(x) \to +\infty$,
(iii) $\Delta_H \Gamma = -\delta_0$ in the sense of distributions.

From the hypoellipticity of Δ_H, [57], one can infer the existence of a "local" fundamental solution of the operator Δ_H see Trèves [97]. Then by the homogeneity properties of Δ_H on can get the global result. Moreover one can prove that such fundamental solution is also unique, symmetric and δ_λ-homogeneous.

Theorem 1.2.43 ([34], Theorem 2.1) *Let $\mathbb{G}$ be a Carnot group of homogeneous dimension $Q > 2$, then there exists a fundamental solution Γ for Δ_H.*

Definition 1.2.44 (5.4.1 in [14]) We call Δ_H-gauge on $\mathbb{G}$ an homogeneous norm d_H smooth out of the origin and satisfying

$$\Delta_H(d_H^{2-Q}) = 0 \quad \text{in } \mathbb{G} \setminus \{0\}$$

The following Proposition relates a Δ_H-gauge to the fundamental solution of Δ_H.

Proposition 1.2.45 *Let Γ be the fundamental solution of Δ_H on $\mathbb{G}$. Then*

$$d_0(x) := \begin{cases} (\Gamma(x))^{\frac{1}{2-Q}} & \text{if } x \in \mathbb{G} \setminus \{0\}, \\ 0 & \text{if } x = 0 \end{cases}$$

is a Δ_H-gauge.

Moreover, also the reverse of Proposition 1.2.45 holds, in fact in Theorem 5.5.6 in [14] it is proved that if d_H is a Δ_H gauge on $\mathbb{G}$, then there exists a positive constant C such that $\Gamma = C d_H^{2-Q}$ is the fundamental solution of Δ_H.

Definition 1.2.46 Let d_H be the Δ_H-gauge. We set, for $x \in \mathbb{G} \setminus \{0\}$,

$$\Psi_H := |\nabla_H d_H|^2(x). \tag{33}$$

Moreover, for every $x, y \in \mathbb{G}$ with $x \neq y$, we define the functions

$$\Psi_H(x,y) := \Psi_H(x^{-1}y) \quad \text{and} \quad \mathcal{K}_H := \frac{|\nabla_H d_H|^2(x^{-1}y)}{|\nabla(d(x^{-1}\cdot))|(y)}. \tag{34}$$

Remark 1.2.47 We observe that Ψ_H is δ_λ-homogeneous of degree 0. This is a consequence of the following general fact: given an homogeneous vector field X of degree d_X and a function $f : \mathbb{G} \to \mathbb{R}$, homogeneous of degree d_f, then the function Xf is homogeneous of degree $d_f - d_X$. As a consequence, since ∇_H is 1-homogeneous as d_H, we deduce that $|\nabla_H d_H|$ is 0-homogeneous.

Definition 1.2.48 Let Ω be an open subset of $\mathbb{G}$, and $u \in C^2(\Omega)$. Then for every $x \in \Omega$ and $r > 0$ such that $D(x,r) \subset \Omega$, we define surface mean as

$$\mathcal{M}_r(u)(x) := \frac{(Q-2)\beta_d}{r^{Q-1}} \int_{\partial B(x,r)} \mathcal{K}_H(x,z) u(z) d\mathcal{H}^{n-1}(z), \tag{35}$$

and the solid mean as

$$M_r(u)(x) := \frac{m_d}{r^Q} \int_{B(x,r)} \Psi_H(x^{-1}y) u(y) dy, \tag{36}$$

where $\mathcal{K}_H$ and Ψ_H are defined respectively in (34) and (33). Here $\mathcal{H}^{n-1}$ is the $n-1$ dimensional Hausdorff measure associated to the Euclidean metric of $\mathbb{R}^n$.

Definition 1.2.49 If $\Omega \subset \mathbb{G}$ is an open set, we say that an USC function $u : \Omega \to [-\infty, \infty[$ satisfies the local surface (local solid) sub-mean property if, for every $x \in \Omega$, there exists $r_x > 0$ such that

$$u(x) \leq \mathcal{M}_r(u)(x) \qquad (u(x) \leq M_r(u)(x)) \qquad \text{for } 0 < r < r_x. \tag{37}$$

If (37) holds for any $r > 0$ such that $D_d(x, r) \subset \Omega$, we shall say that u satisfies the global surface (global solid) sub-mean property.

The next theorem shows that all the previous definitions are indeed equivalent for USC functions, hence in the sequel we will say that such functions have the sub-mean property.

Theorem 1.2.50 (8.1.3 in [14]) *Let $u : \Omega \to [-\infty, \infty[$ be an USC function. Then the following statements are equivalent:*
(i) u satisfies the local solid sub-mean property,
(ii) u satisfies the global solid sub-mean property,
(iii) u satisfies the local surface sub-mean property,
(iv) u satisfies the global surface sub-mean property.

For every open set $\Omega \subset \mathbb{G}$ we denote

$$\mathcal{H}(\Omega) := \{u \in C^\infty(\Omega) \ : \ \Delta_H u = 0\}.$$

A function $u \in \mathcal{H}(\Omega)$ will be called Δ_H-harmonic in Ω.

Definition 1.2.51 A bounded open set $V \subset \mathbb{G}$ will be called Δ_H-regular if the boundary value problem

$$\begin{cases} \Delta_H u = 0 & \text{in } V, \\ u_{\partial V} = \varphi \end{cases} \tag{38}$$

has a unique solution $u := H^V_\varphi$ for every continuous function $\varphi : \partial V \to \mathbb{R}$. We say that u solves (38) if u is $\mathcal{L}$-harmonic in V and

$$\lim_{x \to y} u(x) = \varphi(y) \quad \forall y \in \partial V.$$

As a consequence of the weak maximum principle for sub-Laplacians on Carnot groups, see Theorem 5.13.4 in [14], for every Δ_H-regular open set V there exists a Radon measure μ^V_x supported in ∂V such that

$$H^V_\varphi(x) = \int_{\partial V} \varphi(y) d\mu^V_x(y), \quad \forall \varphi \in C(\partial V). \tag{39}$$

We call μ^V_x the Δ_H-harmonic measure related to V and x. For further details on harmonic measures we refer to Chapter 7 of [14].

Definition 1.2.52 Let Ω be an open set of $\mathbb{G}$. A function $u : \Omega \to [-\infty, +\infty[$ will be called Δ_H-subharmonic in Ω if:
(i) u is USC and $u > -\infty$ in a dense subset of Ω,
(ii) for every Δ_H-regular open set V with closure $\bar{V} \subset \Omega$ and for every $x \in V$,

$$u(x) \leq \int_{\partial V} u(y) d\mu^V_x(y).$$

The family of the Δ_H-subharmonic functions in Ω will be denoted by $\underline{\mathcal{S}}(\Omega)$.

The next lemma characterize smooth Δ_H-subharmonic functions, see Proposition 7.2.5 in [14] for a proof.

Lemma 1.2.53 *Let u be a function in $C^2(\Omega)$. Then u is Δ_H-subharmonic if and only if*

$$\Delta_H u \geq 0 \qquad in\ \Omega.$$

We conclude this section with an useful result which relates sub-mean functions with Δ_H-subharmonic functions.

Lemma 1.2.54 (8.2.2 in [14]) *Let $\Omega \subset \mathbb{G}$ be an open set, and let $u : \Omega \to [-\infty, \infty[$ be an USC function, finite on a dense subset of Ω. Then $u \in \underline{\mathcal{S}}(\Omega)$ if and only if u is sub-mean.*

CHAPTER 2

Convexity in Stratified groups

The main topic of the thesis, namely convexity in sub-Riemannian spaces is addressed in this chapter. In the sub-Riemannian setting many regularity results analogous to the Euclidean ones have been proved for h-convex functions. In particular, it has been shown that continuous h-convex functions are locally Lipschitz continuous with respect to the CC distance. Hence, as a consequence of the Rademacher type differentiability theorem due to Pansu [85], it implies that continuous h-convex functions are almost everywhere differentiable. This result has been generalized first to h-convex functions bounded from above in [6, 71, 102], and subsequently to h-convex measurable functions by Rickly in [87]. Moreover some quantitative version of the previous result can be proved. In particular the $L^\infty - L^1$ estimates:

Theorem 2.2.10 *Let $u : \Omega \to \mathbb{R}$ be an h-convex function, then there exists a positive constant $C = C(\mathbb{G})$ such that for every ball $B_{x_0,r}$ one has*

$$(a)\ \sup_{B_{x_0,r}} |u| \le C \frac{1}{B_{x_0,8r}} \int_{B_{x_0,8r}} |u| dx,$$

$$(b)\ ess \sup_{B_{x_0,r}} |\nabla_H u| \le \frac{C}{r} \frac{1}{B_{x_0,24r}} \int_{B_{x_0,24r}} |u| dx \tag{40}$$

This theorem has been first proved in [26] and a similar result has been obtained in [64] also for v-convex functions, see Definition 2.2.11. In Section 2.1 we give a different proof of Theorem 2.2.10. Our approach is based on a sub-mean inequality, proved in [14], which holds for subharmonic functions on Carnot groups, for precise definitions we refer to Section 2.5.

1. H-convexity

Definition 2.1.1 (h-convex set) We say that a subset $C \subset \mathbb{G}$ is h-convex if for every $x, y \in C$ such that $x \in H_y$ we have $x\delta_\lambda(x^{-1}y) \in C$ for all $\lambda \in [0, 1]$.

In the Euclidean space convex sets are connected. This is not true for h-convex set, as proved in the following simple example.

Example 2.1.2 Let consider the Heisenberg group $\mathbb{H}^1$ and choose the two points

$$A = (0, 0, 0), \quad B = (0, 0, 1).$$

Those points lie on the vertical line $L = \{(0,0,t) \mid t \in \mathbb{R}\}$, hence each point of L lies on a different horizontal plane. Hence the set $\{A,\ B\}$ is h-convex and disconnected. More complicated examples are available, in fact M. Rickly proved in [87] the existence of a totally disconnected non measurable h-convex set in $\mathbb{H}^1$.

We denote by H_x the left translation of H_1 by x, namely $H_x = xH_1$. For each $h \in H_1$, we define the *horizontal segment* $\{th,\ t \in [0,1]\}$ through the short notation $[0,h]$. For any $x \in \mathbb{G}$, we set $x \cdot [0,h] = \{x\delta_t h,\ 0 \le t \le 1\}$ and throughout Ω denotes an open subset of $\mathbb{G}$.

We also remark that in groups of step 2 horizontal segments are indeed "affine segments". In fact if $\mathbb{G} = H_1 \oplus H_2$ is a step 2 group, $x \in \mathbb{G}$ and $h \in H_1$ then for every $t > 0$

$$x \cdot \delta_t h = x_1 + th + x_2 + Q_2(x, th), \tag{41}$$

where $Q_2 : \mathbb{G} \times \mathbb{G} \to \mathbb{G}$ is a bilinear 2-homogeneous map and $x_i = \pi_i(x)$, $i = 1,2$ see Chapter 1. Hence

$$x \cdot \delta_t h = x_1 + th + x_2 + tQ_2(x,h),$$

this implies that $x \cdot [0,h]$ is an "affine segment". This is no longer true for general stratified groups, for instance in a step 3 group (41) becomes

$$x \cdot \delta_t h = x_1 + th + x_2 + tQ_2(x,h) + x_3 + Q_3(x,th),$$

where $Q_3(x,th)$ may have second order terms with respect to t, in those cases $x \cdot [0,h]$ is an arch of parabola.

Definition 2.1.3 (h-convex function) We say that $u : \Omega \to \mathbb{R}$ is *h-convex* if for every $x, y \in \Omega$ such that $x \in H_y$ and $x \cdot [0, x^{-1}y] \subset \Omega$, we have

$$u\big(x\delta_\lambda(x^{-1}y)\big) \le \lambda u(y) + (1-\lambda)u(x), \qquad \forall \lambda \in [0,1]. \tag{42}$$

2. Regularity of H-convex functions

Notice that this notion of h-convexity is local and it does not require any assumption on Ω. Moreover, this definition requires that the restriction of the function u to all horizontal lines $t \to x\delta_t h$ with $h \in V_1$ is a one-dimensional convex function. Although the information of the behavior of an h-convex function is known only on a family of 1-dimensional manifolds, namely horizontal lines, the Hörmander condition on those directions yields a "global" information in terms of the Carnot-Carathéodory distance. In fact as an important property of h-convex functions, we have the following.

Theorem 2.2.1 (M. Rickly, [87]) *Every measurable h-convex function is locally Lipschitz.*

Throughout, all h-convex functions are assumed to be measurable, hence a.e. differentiable by the Rademacher-type theorem 1.2.25.

Remark 2.2.2 We observe that h-convexity is preserved under left translations and intrinsic dilations. We prove only the first assertion, the second one follows the same scheme. Let $u : \Omega \to \mathbb{R}$ be an h-convex function. Fix a point $x_0 \in \mathbb{G}$, and consider the function $\tau_{x_0}u(x) = u(x_0 x)$, defined in $x_0^{-1}\Omega$. Let $x, y \in x_0^{-1}\Omega$ be such that $x \in H_y$ and

$x \cdot [0, x^{-1}y] \subset x_0^{-1}\Omega$. Set $\bar{x} = x_0 x, \bar{y} = x_0 y$, and note that $\bar{x}, \bar{y} \in \Omega$. Moreover $\bar{x} \in H_{\bar{y}}$ by definition of horizontal planes. Hence by h-convexity of u

$$\begin{aligned}\tau_{x_0} u\big(x\delta_\lambda(x^{-1}y)\big) &= u\big(\bar{x}\delta_\lambda(\bar{x}^{-1}\bar{y})\big) \leq \lambda u(\bar{y}) + (1-\lambda)u(\bar{x}) \\ &= \lambda \tau_{x_0} u(y) + (1-\lambda)\tau_{x_0} u(x), \qquad \forall \lambda \in [0,1]\end{aligned}$$

this implies that $\tau_{x_0} u$ is h-convex.

As a first consequence of Definition 2.1.3 we obtain the following.

Proposition 2.2.3 ([26]) *Suppose that $u : \Omega \to \mathbb{R}$ is h-convex. If $u \in \Gamma^1(\Omega)$ one has for any $x \in \Omega$*

$$u(x') \geq u(x) + \langle \nabla_H u(x), x^{-1}x' \rangle, \quad x' \in H_x$$

¿From a geometric point of view we can interpret the previous basic result as follows: for every $x \in \Omega$, the graph of the h-convex function u restricted on H_x lies above the horizontal tangent plane. This will be a useful observation in Chapter 3 when we deal with h-subdifferentials.

Definition 2.2.4 Let $u : \Omega \longrightarrow \mathbb{R}$ be a $C^2(\Omega)$ function. We define the *symmetrized horizontal Hessian* of u as follows

$$(\nabla_H^2 u)_{ij} = \left(\frac{X_i X_j u + X_j X_i u}{2}\right)_{ij}, \quad i,j = 1, \ldots, m_1.$$

The following theorem gives a motivation for Definition 2.1.3: as in the Euclidean case, smooth h-convex functions are characterized by the positivity of their Hessian.

Theorem 2.2.5 (5.12 in [26]) *Let $\Omega \subset \mathbb{G}$ be an open set let u be in $\Gamma^2(\Omega)$. Then u is h-convex if and only if the symmetrized horizontal Hessian $\nabla_H^2 u(x)$ is positive semidefinite at every $x \in \Omega$.*

Example 2.2.6 We observe that in step 2 groups h-convexity is a weaker notion than the classical convexity (E-convexity) in $\mathbb{R}^n$, this follows form the fact that horizontal segments are "affine segments" as proved in (41), hence every E-convex set is h-convex. Indeed the inclusion is strict, see for instance Example 2.1.2. This is true also for h-convex functions. Let us consider the following homogeneous function

$$N(x,y,t) = \big((x^2+y^2)^2 + t^2\big)^{\frac{1}{4}}$$

on the first Heisenberg group $\mathbb{H}^1$. Clearly N is not E-convex, however its h-convexity has been proved in [26]. We recall that N is often called the gauge map in $\mathbb{H}^1$.
Given a Carnot group $\mathbb{G} = H_1 \oplus \cdots \oplus H_\iota$ we can define the analogous of the Heisenberg gauge as

$$N_{\mathbb{G}}(x) = \left(\sum_{i=1}^{\iota} |x_i|^{\frac{2\iota!}{i}}\right)^{\frac{1}{2\iota!}} \tag{43}$$

where $|\cdot|$ is the Euclidean norm on H_i, and $x_i \in H_i$, $i = 1, \ldots, \iota$. In [26] the authors raised the question whether $N_{\mathbb{G}}$ is h-convex for every Carnot groups.

Theorem 2.2.7 ([26, 64]) *Let $\Omega \subset \mathbb{G}$ be an open set, and let $u : \Omega \to \mathbb{R}$ be an h-convex function. Then the second order symmetric X-derivatives, namely*

$$\frac{X_i X_j u + X_j X_i u}{2}, \quad i, j = 1, \dots, m,$$

are nonnegative Radon measures.

Now we prove that we can approximate h-convex functions by smooth ones. We recall that the functions ϕ_ε are defined in Definition 1.2.28.

Lemma 2.2.8 *Let u be an h-convex function in an open set $\Omega \subset \mathbb{G}$. Then the functions $u_\varepsilon := \phi_\varepsilon \star u$ are h-convex and u_ε converges to u uniformly on compact set.*

PROOF. Let Ω' be an open bounded set such that $\overline{\Omega'} \subset \Omega$. There exists $\kappa > 0$ such that

$$\max_{x \in \overline{\Omega'}} d(y^{-1}x, x) < dist(\Omega', \mathbb{G} \setminus \Omega),$$

whenever $\|y\| \leq \kappa$. Then the convolution $u_\varepsilon = \int_\Omega \phi_\varepsilon(y) u(y^{-1}x) dy$ is smooth an well defined in Ω' for every $\varepsilon < \kappa$, notice that $u \in L^1(\Omega')$ since it is Lipschitz continuous. Moreover, by Lemma 1.2.29, we know that u_ε converges to u uniformly on Ω'. Consider an horizontal segment $x[0, h] \subset \Omega'$, then for every $t \in [0, 1]$ we have

$$u_\varepsilon(x\delta_t h) = \int_\Omega \phi_\varepsilon(y) u(y^{-1}x\delta_t h) dy.$$

Since h-convexity is invariant under left translation, see Remark 2.2.2, we get

$$\int_\Omega \phi_\varepsilon(y) u(y^{-1}x\delta_t h) dy \leq (1-t) \int_\Omega \phi_\varepsilon(y) u(y^{-1}x) dy + t \int_\Omega \phi_\varepsilon(y) u(y^{-1}xh),$$

or equivalently $u_\varepsilon(x\delta_t h) \leq (1-t) u_\varepsilon(x) + t u_\varepsilon(xh)$. This completes the proof. □

We already mentioned at the beginning of this section that measurable h-convex functions are locally Lipschitz continuous. Moreover, some quantitative versions of this fact hold for h-convex functions, [26, 60, 64].

Theorem 2.2.9 *Let $\Omega \subset \mathbb{G}$ be an open set and let $u : \Omega \to \mathbb{R}$ be an h-convex function. Then there exists a positive constant $C = C(\mathbb{G})$ such that the following estimate holds*

$$\|\nabla_H u\|_{L^\infty(B_{x_0,r})} \leq \frac{C}{r} \|u\|_{L^\infty(B_{x_0,3r})},$$

for any $x_0 \in \mathbb{G}$ and every $r > 0$.

As an improvement of the previous result we have the $L^\infty - L^1$ estimates, see for instance Theorem 9.2 in [26].

Theorem 2.2.10 *Let $u : \Omega \to \mathbb{R}$ be an h-convex function, then there exists a positive constant $C = C(\mathbb{G})$ such that for every ball $B_{x_0,r}$ one has*

$$(44) \quad \begin{aligned} (a)\ & \sup_{B_{x_0,r}} |u| \leq C \frac{1}{B_{x_0,8r}} \int_{B_{x_0,8r}} |u| dx, \\ (b)\ ess & \sup_{B_{x_0,r}} |\nabla_H u| \leq \frac{C}{r} \frac{1}{B_{x_0,24r}} \int_{B_{x_0,24r}} |u| dx \end{aligned}$$

In Subsection 2.1 we give a new proof of (44) (a), see also [26, 60, 64].

We conclude this section with the definition of v-convexity, this notions has been introduced by Lu, Manfredi and Stroffolini in [64]. It turns out that it is equivalent to h-convexity, see [60, 64, 71, 88, 102].

Definition 2.2.11 (v-convex function) An upper semicontinuous function $u : \Omega \to \mathbb{R}$ is v-convex if

$$\nabla_H^2 u \geq 0, \quad \text{in the viscosity sense,}$$

namely for every $\phi \in C^2(\Omega)$ such that $u - \phi$ has a maximum in x_0, we have

$$\nabla_H^2 \phi(x_0) \geq 0.$$

Theorem 2.2.12 *Let $u : \Omega \to \mathbb{R}$ be an upper semicontinuous function. Then u is h-convex if and only if it is v-convex.*

2.1. $L^1 - L^\infty$ inequality in Carnot group. Here we present a simple proof of the inequality (44) (a) on Carnot groups. We remark that the proof involves methods from potential theory, more precisely the key ingredient is the sub-mean formula for sub-harmonic functions proved in [14], see Chapter 1.

Throughout the section we denote by $USC(\Omega)$ the class of upper semicontinuous functions in the open set $\Omega \subset \mathbb{G}$.

Lemma 2.2.13 *Let $\Omega \subset \mathbb{G}$ be an open set. There exists a constant $C = C(\mathbb{G})$, such that for every sub-mean function $u \in USC(\Omega)$, see (37), for every $x_0 \in \Omega$ and all $r > 0$ with $D_{x_0,2r} \subset \Omega$ we have*

$$\sup_{x \in B_{x_0,r}} u(x) \leq C \fint_{B_{x_0,2r}} |u(y)| dy. \tag{45}$$

PROOF. This is an easy consequence of the fact that Ψ_H, defined in (33), is δ_λ-homogeneous of degree 0 and positive on $\mathbb{G}$. Let $\bar{x} \in B(x_0, r)$, such that $u(\bar{x}) \geq \frac{1}{2} \sup_{B_{x_0,r}} u$. Since u is sub-mean we have

$$\begin{aligned} u(\bar{x}) \leq M_r(u)(\bar{x}) &= \frac{m_d}{r^Q} \int_{B_{\bar{x},r}} \Psi_H(\bar{x}^{-1} y) u(y) dy \\ &\leq \left(\sup_{\mathbb{G} \setminus \{0\}} \Psi_H \right) \frac{m_d}{r^Q} \int_{B_{\bar{x},r}} |u(y)| dy. \end{aligned}$$

Clearly we have that $B_{\bar{x},r} \subset B_{x_0,2r}$, hence by the previous inequality

$$\sup_{x \in B_{x_0,r}} u(x) \leq 2u(\bar{x}) \leq C \fint_{B_{x_0,2r}} |u(y)| dy.$$

□

In order to estimate $\sup_{B_{x_0,r}} |u|$ we state a lemma, see for instance [71, 88], where it is proved that an h-convex function, locally bounded above is bounded from below.

Lemma 2.2.14 *Let $u : \Omega \to \mathbb{R}$ be an h-convex function. Suppose that u is locally bounded above. Then there exists a positive constant C only depending on $\mathbb{G}$ such that for every $x_0 \in \Omega$ and $r > 0$, $D_{x_0,4r} \subset \Omega$ we have*

(46) $$u(x_1) \geq -C(\mathbb{G}) \sup_{B_{x_0,4r}} u, \quad \forall x_1 \in B_{x_0,r}.$$

PROOF OF THEOREM 2.2.10. Let ϕ_ε be as in Definition 1.2.28, then consider the smooth function $u_\varepsilon = \phi_\varepsilon \star u$, see Lemma 1.2.29. It follows, by Lemma 2.2.8, that u_ε is an h-convex function, hence as a consequence of Theorem 2.2.5 we have $\Delta_H u_\varepsilon \geq 0$ in Ω_ε. Lemma 1.2.53 implies that $u_\varepsilon \in \underline{\mathcal{S}}$. As a consequence of Lemma 1.2.54 u_ε has the sub-mean property, hence we can apply Lemma 2.2.13, getting a constant $C > 0$ (independent of u_ε) such that

$$\sup_{x \in B_{x_0,r}} u_\varepsilon(x) \leq C \fint_{B_{x_0,2r}} |u_\varepsilon(y)| dy.$$

Since $u_\varepsilon \to u$ uniformly on compact sets, the previous inequality can be extended to all continuous h-convex functions. Now by Lemma 2.2.14 and the previous estimate, it follows that for every $x_1 \in B_{x_0,r}$

$$u(x_1) \geq -C_1 \sup_{B_{x_0,4r}} u \geq -C_2 \fint_{B_{x_0,8r}} |u(y)| dy$$

where C_1 and C_2 are suitable positive constant depending only on $\mathbb{G}$. Hence taking the infimum over all $x_1 \in B_{x_0,r}$ in the left hand side of the previous inequality we get

$$\inf_{B_{x_0,r}} u \geq -C_2 \fint_{B_{x_0,8r}} |u(y)| dy.$$

This concludes the proof of (44) (a). The estimate of the horizontal gradient, namely (44) (b), follows from Theorem 2.2.9 and (a). □

CHAPTER 3

H-SUBDIFFERENTIALS AND APPLICATIONS

This chapter is based on the paper [67], a joint work with V. Magnani. In the Euclidean framework R. T. Rockafellar in Theorem 2.8 of [91] proves that *a convex function has a second order expansion at a fixed point if its gradient is differentiable at that point in the extended sense.* Extending this characterization to the framework of stratified groups is the aim of this section, that requires the development of some basic tools.

We translate this notion in stratified groups. Let us fix an orthonormal basis of V_1, $(X_1, \dots, X_{m_1})$. A locally Lipschitz function $u : \Omega \to \mathbb{R}$ is *twice h-differentiable at* x if there exists the *horizontal gradient* of u at x, namely $\nabla_H u(x) = (X_1 u(x), \dots, X_{m_1} u(x))$ and moreover there exists a linear map $A_x : \mathbb{G} \to H_1$ such that

$$\left\| \frac{\nabla_H u(xw) - \nabla_H u(x) - A_x(w)}{\|w\|} \right\|_{L^\infty(B_\delta, H_1)} \longrightarrow 0 \quad \text{as} \quad \delta \to 0^+ . \tag{47}$$

If (47) holds we also say that $\nabla_H u$ is *h-differentiable at* x *in the extended sense*. This notion makes sense, since Lipschitz functions are almost everywhere h-differentiable, by Pansu's result [85]. We are now in the position to state the main result of this paper.

Theorem 3.2.8 *Let $u : \Omega \longrightarrow \mathbb{R}$ be h-convex and let x be a point in Ω. Then u has a second order h-expansion at x if and only if it is twice h-differentiable at x.*

The proof of Theorem 3.2.8 needs several basic results involving the h-subdifferential. Since we expect that these results should play a role in the potential development of a nonsmooth calculus for h-convex functions, we wish to emphasize some of them. To this aim, we first prove the following

Lemma 3.2.6 *Let $u : \Omega \longrightarrow \mathbb{R}$ be h-convex. Then u is twice h-differentiable at x if and only if there exist an h-linear mapping $A_x : \mathbb{G} \to H_1$ and $v \in H_1$ such that*

$$\partial_H u(xw) \subseteq v + A_x(w) + o(\|w\|)\mathbb{B} \tag{48}$$

for all $w \in x^{-1}\Omega$, where $\mathbb{B}$ denote the unit ball in H_1. In particular, if (48) *holds, then $v = \nabla_H u(x)$.*

At first sight, extended differentiability in the sense of (48) seems stronger than (47), that implies a convergence up to a negligible set, where $\nabla_H u$ is not defined. In fact, the delicate point is to prove that extended differentiability implies (48). This is a consequence of the following characterization of the h-subdifferential.

Theorem 3.1.8 *Let $u : \Omega \to \mathbb{R}$ be h-convex. Then for every $x \in \Omega$ we have*

$$\bar{co}\left(\nabla_H^\star u(x)\right) = \partial_H u(x) . \tag{49}$$

We denote by $co(E) \subset H_1$ the convex hull in H_1 of the subset $E \subset H_1$ and by $\bar{co}(E)$ its closure. The *h-reachable gradient* is given by

$$\nabla_H^\star u(x) = \Big\{ p \in H_1 : \, x_k \to x, \, \nabla_H u(x_k) \text{ exists for all } k\text{'s and } \nabla_H u(x_k) \to p \Big\}. \tag{50}$$

The proof of equality (49) in the Euclidean case can be found for instance in [3]. There are two main features in the proof of Theorem 3.1.8, with respect to the Euclidean one. First, it is still possible to use the Hahn-Banach's theorem, when applied inside the horizontal subspace H_1, that has a linear structure. Second, the group mollification does not commute with horizontal derivatives, hence the mollification argument of the Euclidean proof cannot be applied. We overcome this point by a Fubini type argument with respect to a semidirect factorization, following the approach of [69].

The h-differentiability of u from validity of (48) is a consequence of the following

Theorem 3.1.20 *Let $u : \Omega \longrightarrow \mathbb{R}$ be h-convex. Then u is h-differentiable at x if and only if $\partial_H u(x) = \{p\}$ and in this case $\nabla_H u(x) = \{p\}$.*

The uniqueness of the h-subdifferential as a consequence of h-differentiability has been already shown [27], see also [20] for the case of Heisenberg groups. To show the opposite implication we decompose the difference quotient of u into sums of difference quotients along horizontal directions. The same decomposition along horizontal directions have been first used by Pansu, [85]. The second ingredient is the following

Theorem 3.1.16 *Let $u : \Omega \longrightarrow \mathbb{R}$ be an h-convex function. Then for every $x \in \Omega$ and every h such that $[0, h] \subseteq H_1 \cap x^{-1}\Omega$, there exists $t \in [0, 1]$ and $p \in \partial_H u(x\delta_t h)$ such that $u(xh) - u(x) = \langle p, h \rangle$.*

This theorem is also important in the proof of Theorem 3.2.8. In fact, it is an essential tool to establish that twice h-differentiability implies the existence of a second order *h*-expansion. This implication again requires Pansu's approach to differentiability and in addition a nonsmooth mean value theorem for functions of the form $U + P$, where U is h-convex and P is a polynomial of homogeneous degree at most two. This slightly more general version of Theorem 3.1.16 is given in Theorem 3.1.15, where the h-subdifferential is replaced by the more general λ-subdifferential, see Definition 3.1.12. In the Euclidean framework, a short proof of the previous result can be found in Theorem 7.10 of [1], where the Clarke's nonsmooth mean value theorem plays a key role.

In this connection, we wish to emphasize the intriguing open question on the validity of a nonsmooth mean value theorem for Lipschitz functions in stratified groups. In the Euclidean framework, this theorem holds using the notion of Clarke's differential. This notion of differential relies on subadditivity of "limsup directional derivatives", that allows in turn to apply Hahn-Banach's theorem, see [24]. The obvious extension of this notion to stratified groups does not work and the analogous obstacle comes up considering h-convex functions, where horizontal directional derivatives always exist, see Definition 3.1.18. It is curious to notice that our nonsmooth mean value theorem implies this subadditivity, see Corollary 3.1.19, whereas in the Euclidean framework subadditivity eventually leads to the nonsmooth mean value theorem.

1. Properties of the h-subdifferential

This section is devoted to the proof of our results concerning h-differentiability, h-subdifferentials, converging sequences of h-convex functions and nonsmooth mean value theorems. Recall that $\mathbb{B}$ is the the unit ball in H_1 centered at the origin with respect to the fixed scalar product on $\mathbb{G}$. We say that $p \in H_1$ is an *h-subdifferential* of $u : \Omega \longrightarrow \mathbb{R}$ at $x \in \Omega$ if whenever $h \in H_1$ and $[0, h] \subset x^{-1} \cdot \Omega$,

$$u(xh) \geq u(x) + \langle p, h \rangle \,. \tag{51}$$

We denote by $\partial_H u(x)$ the set of all h-subdifferentials of u at x and the corresponding set-valued mapping by $\partial_H u : \Omega \rightrightarrows H_1$. Here $\langle \cdot, \cdot \rangle$ in (51) is the scalar product in $\mathbb{G}$.

Lemma 3.1.1 *Let $\Omega \subset \mathbb{G}$ be an open set and let $u : \Omega \to \mathbb{R}$ be a continuous function. Then the set $\partial_H u(x) \subset H_1$ is convex.*

PROOF. Let p and q be in $\partial_H u(x)$ and choose $\lambda \in [0, 1]$, we need to prove that $\lambda p + (1 - \lambda) q \in \partial_H u(x)$. This follows from adding the two inequalities

$$\begin{array}{rcl} \lambda u(xh) & \geq & \lambda u(x) + \langle \lambda p, h \rangle \\ (1-\lambda) u(xh) & \geq & (1-\lambda) u(x) + \langle (1-\lambda) q, h \rangle \,. \end{array}$$

□

Remark 3.1.2 Let u be an h-convex function in Ω, then our assumption of measurability yields by Theorem 2.2.1 the locally Lipschitz continuity of u. Hence as a straightforward consequence of the definition of h-subdifferential, for every $B_{x,r} \subseteq \Omega$, there exists a positive number L depending on $x \in \Omega$, $r > 0$ and u, such that

$$\partial_H u(y) \subseteq L\mathbb{B} \quad \text{for every } y \in B_{x,r}. \tag{52}$$

Remark 3.1.3 As already mentioned, an h-convex function $u : \Omega \to \mathbb{R}$ that is h-differentiable at $x \in \Omega$ has unique subdifferential, hence $\partial_H u(x) = \{\nabla_H u(x)\}$ according to Proposition 2.2.3.

We denote by $co(E) \subset H_1$ the convex hull in H_1 of the subset $E \subset H_1$ and by $\bar{co}(E)$ its closure. The *h-reachable gradient* is given by

$$\nabla_H^\star u(x) = \Big\{ p \in H_1 : \, x_k \to x, \, \nabla_H u(x_k) \text{ exists for all } k\text{'s and } \nabla_H u(x_k) \to p \Big\} \,. \tag{53}$$

Theorem 3.1.4 *Let $\Omega \subset \mathbb{G}$ be an open set, and let $u : \Omega \to \mathbb{R}$ be h-convex. Then for every $x \in \Omega$ we have*

$$\partial_H u(x) \subseteq \bar{co} \left(\nabla_H^\star u(x) \right) \,, \tag{54}$$

where $\nabla_H^\star u(x)$ is defined in (53).

PROOF. Suppose that there exists $p \in \partial_H u(x)$ such that $p \notin \bar{co} \left(\nabla_H^\star u(x) \right)$. We can assume that $p = 0$, otherwise one considers $v(x) = u(x) - \langle p, \pi_1(x) \rangle$, that is still h-convex. Since $\bar{co} \left(\nabla_H^\star u(x) \right)$ is a closed convex subset of H_1, the Hahn-Banach separation theorem can be applied to this set and the origin, hence there exists $q \in H_1$, $d(0,q) = 1$, and $\alpha > 0$ such that

$$\langle z, q \rangle > \alpha \qquad \forall z \in \nabla_H^\star u(x). \tag{55}$$

We claim the existence of $r > 0$ such that $B_{x,r} \subset \Omega$ and $\langle \nabla_H u(y), q \rangle > \frac{\alpha}{2}$ for every $y \in B_{x,r}$ where u is h-differentiable. By contradiction, suppose there exist sequences $r_j \to 0$ and $y_j \in B_{x,r_j}$ such that $\langle \nabla_H u(y_j), q \rangle \leq \frac{\alpha}{2}$, then possibly passing to a subsequence we have $y_j \to x$ and $\nabla_H u(y_j) \to z \in \nabla_H^\star u(x)$, with $\langle z, q \rangle \leq \frac{\alpha}{2}$ and this conflicts with (55). Denote by r the positive number having the previous property. Let Q be the set $Q = \{\delta_t q \,:\, t \in \mathbb{R}\}$ and consider μ the Haar measure on $\mathbb{G}$. By Lemma 2.7 in [68] there exists a normal subgroup $N \subset \mathbb{G}$, such that $N \cap Q = \{e\}$ and $NQ = \mathbb{G}$. Moreover, by Proposition 2.8 in [68], there exist ν_q and μ_N, respectively Haar measures on Q and N such that for every measurable set $A \subset \mathbb{G}$

$$\mu(A) = \int_N \nu_q(A_n)\, d\mu_N(n) \tag{56}$$

where $A_n = \{h \in Q \,:\, nh \in A\}$. Let P be the set of h-differentiable points of u, which has full measure in Ω. ¿From (56) it follows that for μ_N-a.e. $n \in N$, $\nu_Q(Q \setminus n^{-1}P) = 0$. Then for μ_N-a.e. $n \in N$, $n\delta_t q \in P$ for a.e. $t \in \mathbb{R}$. Let $\bar{n} \in N$ and $\delta_{\bar{t}} q \in Q$ be respectively the unique elements in N and Q such that $x = \bar{n}\delta_{\bar{t}} q$. Let $\epsilon > 0$ and $s > 0$ be such that $B^N_{\bar{n},s} \cdot B^Q_{\delta_{\bar{t}} q,\epsilon} \subset B_{x,r}$, where $B^N_{\bar{n},s}$ and $B^Q_{\delta_{\bar{t}} q,\epsilon}$ are open balls respectively in N and Q. Fix a point $n \in B^N_{\bar{n},s}$ where $u(nh)$ is ν_q-a.e. differentiable and consider the convex function $v(t) = u(n\delta_t q)$, for ν_q-a.e. $\delta_t q$, $t \in (-\epsilon + \bar{t}, \epsilon + \bar{t})$ we have

$$v'(t) = \langle \nabla_H u(n\delta_t q), q \rangle > \frac{\alpha}{2}.$$

Integrating the previous inequality, taking into account the Lipschitz regularity of v we get

$$v(t_1) - v(t_2) = u(n\delta_{t_1} q) - u(n\delta_{t_2} q) > \frac{\alpha}{2}(t_1 - t_2)$$

where $-\epsilon + \bar{t} < t_2 < t_1 < \epsilon + \bar{t}$. Now let $n_j \to \bar{n} \in B^N_{\bar{n},s}$ be such that $n_j h$ is a differentiable point of the map $h \to u(n_j h)$ for every j and ν_q-a.e. h, by the previous considerations we have

$$u(n_j\delta_{t_1} q) - u(n_j\delta_{t_2} q) > \frac{\alpha}{2}(t_1 - t_2) \qquad -\epsilon + \bar{t} < t_2 < t_1 < \epsilon + \bar{t}$$

finally we can pass to the limit in j and get the strict monotonicity of $u(\bar{n}\delta_t q)$ i.e.

$$u(\bar{n}\delta_{t_1} q) - u(\bar{n}\delta_{t_2} q) \geq \frac{\alpha}{2}(t_1 - t_2) \qquad -\epsilon + \bar{t} < t_2 < t_1 < \epsilon + \bar{t}. \tag{57}$$

Recall that $0 \in \partial_H u(x)$, i.e. $u(xh) \geq u(x)$ whenever $[0,h] \subseteq H_1 \cap x^{-1}\Omega$. Thus, $u(\bar{n}\delta_t q) \geq u(\bar{n}\delta_{\bar{t}} q)$ for all $t \in (\bar{t} - \epsilon, \bar{t} + \epsilon)$, in contrast with the monotonicity (57). □

Joining Theorem 3.1.4 with Theorem 2.2.10, we immediately get

Corollary 3.1.5 *Let $u : \Omega \to \mathbb{R}$ be an h-convex function. There exists $C = C(\mathbb{G}) > 0$ such that for every ball $B(x,r) \subset \mathbb{G}$ one has*

$$\sup_{\substack{p\in\partial_H u(y)\\ y\in B_{x,r}}} |p| \le \frac{C}{r}\frac{1}{|B_{x,15r}|}\int_{B_{x,15r}} |u(y)|dy. \tag{58}$$

Given a set $E \subset \mathbb{G}$ and $\rho > 0$, by $I(E,\rho)$, we denote the open set

$$I(E,\rho) = \{x \in \mathbb{G},\ d(x,E) < \rho\}.$$

Proposition 3.1.6 *Let $\Omega \subset \mathbb{G}$ be an open set ad $u_i : \Omega \to \mathbb{R}$ be a sequence of h-convex functions. Suppose that u_i uniformly converges on compact sets to an h-convex function u. Let x be a point in Ω and let (x_i) be a sequence in Ω converging to x. Then for every $\epsilon > 0$, there exists $i_0 \in \mathbb{N}$ such that*

$$\partial_H u_i(x_i) \subseteq \partial_H u(x) + \epsilon\mathbb{B} \quad \textit{for all} \ \ i \ge i_0. \tag{59}$$

In addition, if u is everywhere h-differentiable in Ω, then for every compact set $K \subset \Omega$ and every $\epsilon > 0$, there exist $i_{\epsilon,K}$ such that

$$\partial_H u_i(y) \subseteq \nabla_H u(y) + \epsilon\mathbb{B} \quad \textit{for all} \ \ i \ge i_{\epsilon,K},\ \textit{whenever}\ y \in K. \tag{60}$$

PROOF. We argue by contradiction in both cases, hence we suppose that there exist $\epsilon > 0$ and a subsequence $p_{i_k} \in \partial_H u_{i_k}(x_{i_k})$ such that for every $p \in \partial_H u(x)$ we have $|p_{i_k} - p| > \epsilon$. By estimate (58) one easily observes that the sets $\partial_H u_i(x_i)$ are equibounded, thus possibly passing to a subsequence, $p_{i_k} \to q$ and $\mathrm{dist}(p_{j_k}, \partial_H u(x_{j_k})) \ge \epsilon$. Define a monotone family of compact sets

$$K_\tau = \left\{x \in D_\tau \ :\ d(x,\Omega^c) \ge \frac{1}{\tau}\right\},$$

such that $\bigcup_{\tau>0} K_\tau = \Omega$. Let j_l be a subsequence such that $p_{j_l} \to q$ and $\|u_{i_l} - u\|_{L^\infty(K_l)} < \frac{1}{l}$. Recall that $p_{j_l} \in \partial_H u_{j_l}(x_{j_l})$, then

$$u_{j_l}(x_{j_l}h) \ge u_{j_l}(x_{j_l}) + \langle p_{j_l}, h\rangle \qquad \text{whenever } [0,h] \subseteq H_1 \cap x_{j_l}^{-1}\Omega.$$

By uniform convergence for l sufficiently large, we get

$$u(x_{i_l}h) \ge u(x_{i_l}) - \frac{2}{l} + \langle p_{i_l}, h\rangle \qquad \text{whenever } [0,h] \subseteq H_1 \cap x_{i_l}^{-1}K_l. \tag{61}$$

Take $[0,h] \subseteq (x^{-1}\Omega) \cap H_1$, then there exists l_0 such that for every $l > l_0$, $[0,h] \subset x^{-1}K_l \cap H_1$. Since Ω is an open set there exists $\rho > 0$ such that $I(x \cdot [0,h], \rho) \subset K_l$. By continuity of left translation there exists $j(\rho)$ such that for every $j_l > j(\rho)$,

$$x_{j_l} \cdot [0,h] \subseteq I(x \cdot [0,h], \rho),$$

hence $[0,h] \subseteq x_{j_l}^{-1}K_l$. Then (61) holds with h and passing to the limit in l we get

$$u(xh) \ge u(x) + \langle q, h\rangle, \tag{62}$$

thus $q \in \partial_H u(x)$, getting a contradiction. Now suppose that u is everywhere h-differentiable. Again, by contradiction there exist a compact set $Z \subset \Omega$, $\epsilon > 0$ and a subsequence j_l such that for all l, $x_{j_l} \in Z$ we have

$$\partial_H u_{j_l}(x_{j_l}) \nsubseteq \partial_H u(x_{j_l}) + \epsilon \mathbb{B}.$$

Then, we can find $p_{j_l} \in \partial_H u_{j_l}(x_{j_l})$ such that $\mathrm{dist}(p_{j_l}, \partial_H u(x_{j_l})) \geq \epsilon$, for all $l > 0$. As before, we can suppose that, possibly passing to a subsequence, $x_{j_l} \to \bar{x} \in W$ and $p_{j_l} \to \bar{p}$. By h-differentiability at $\bar{x}$ and Remark 3.1.3, taking into account the first part of this proposition, we get that for every $\eta > 0$ there exists $j_{l'}$ such that

$$\begin{aligned} \partial_H u_{j_l}(x_{j_l}) &\subset \nabla_H u(\bar{x}) + \eta \mathbb{B} \\ \partial_H u(x_{j_l}) &\subset \nabla_H u(\bar{x}) + \eta \mathbb{B}, \qquad \forall j_l > j_l'. \end{aligned}$$

¿From the previous inclusions, it follows that

$$\epsilon \leq \mathrm{dist}(p_{j_k}, \partial_H u(x_{j_k})) \leq 2\eta.$$

If we choose $\eta = \frac{\epsilon}{4}$, then reach a contradiction, concluding the proof. □

Taking the constant sequence $u_i = u$ in the previous proposition and taking into account (59), we immediately reach the following

Corollary 3.1.7 *Let Ω be an open set of $\mathbb{G}$ and let $u : \Omega \to \mathbb{R}$ be an h-convex function, then $\partial_H u : \Omega \to \mathcal{P}(H_1)$ has closed graph.*

The previous corollary allows us to complete the proof of Theorem...

Theorem 3.1.8 *Let $u : \Omega \to \mathbb{R}$ be h-convex. Then for every $x \in \Omega$ we have*

(63) $$\bar{co}\,(\nabla_H^\star u(x)) = \partial_H u(x)\,.$$

PROOF. By virtue of Theorem 3.1.4, we only have to prove the inclusion

$$\bar{co}\,(\nabla_H^\star u(x)) \subseteq \partial_H u(x).$$

By Corollary 3.1.7, the set-valued map $\partial_H u$ has closed graph and $\partial_H u(y) = \{\nabla_H u(y)\}$ at any h-differentiable point y of u. This immediately yields

$$\nabla_H^\star u(x) \subseteq \partial_H u(x).$$

Moreover $\partial_H u(x)$ is a convex set in H_1 for every $x \in \mathbb{G}$, then our claim follows. □

Remark 3.1.9 The a.e. h-differentiability of an h-convex function u implies that $\nabla_H^* u(x) \neq \emptyset$ for all $x \in \Omega$. Then (63) implies that $\partial_H u(x) \neq \emptyset$ for all $x \in \Omega$. We have then shown that any h-convex function has everywhere nonempty h-subdifferential. This fact was first proved in [20] for h-convex functions on Heisenberg groups. The opposite implication can be found in [26] for h-convex domains. The same implication holds for h-convex functions on open sets, since the everywhere h-subdifferentiability implies the everywhere Euclidean subdifferentiability along horizontal lines. Then the Euclidean characterization of convexity through the subdifferential gives the Euclidean convexity along horizontal lines, that coincides with the notion of h-convexity.

Definition 3.1.10 Let $\Omega \subset \mathbb{G}$ be an open subset and let u be a real valued function in Ω. Then we define the first order *sub jet* of u at $x \in \Omega$ as

$$J_u^{1,-}(x) = \left\{ p \in H_1 \ : \ u(xh) \geq u(x) + \langle p, h \rangle + o(\|h\|), \text{ if } [0,h] \subset H_1 \cap x^{-1}\Omega \right\}$$

Remark 3.1.11 Let u be an h-convex function in Ω. Then u is h-subdifferentiable at x if and only if $J_u^{1,-}(x) \neq \emptyset$. Moreover $J_u^{1,-}(x) = \partial_H u(x)$. For the reader's sake we give the proof of this property, in the Heisenberg group it has been proved in [20]. The inclusion $J_u^{1,-}(x) \supseteq \partial_H u(x)$ follows by definition. Now let p be in $J_u^{1,-}(x)$, and fix $[0,h] \subseteq x^{-1}\Omega \cap H_1$. Then p satisfies

$$u(x\delta_t h) \geq u(x) + \langle p, th \rangle + o(\|th\|).$$

By h-convexity of u, $tu(xh) + (1-t)u(x) \geq u(x\delta_t h)$ which implies

$$u(xh) \geq u(x) + \langle p, h \rangle + \frac{o(\|th\|)}{t}.$$

Now the claim follows letting $t \to 0$.

Definition 3.1.12 Let $\Omega \subset \mathbb{G}$ an open subset and consider $u : \Omega \to \mathbb{R}$. Given $\lambda \geq 0$ we define the λ-subdifferential of u at $x \in \Omega$ as

$$\partial_H^\lambda u(x) = \left\{ p \in H_1 \ : \ u(xh) \geq u(x) + \langle p, h \rangle - \lambda \|h\|^2, \text{ whenever } [0,h] \subseteq H_1 \cap x^{-1}\Omega \right\}.$$

Notice that $\partial_H^0 u(x)$ coincides with the h-subdifferential $\partial_H u(x)$.

Lemma 3.1.13 *Consider a function $u = U + P$ in Ω. Let U be h-convex and let P be a polynomial with h-deg$P \leq 2$, denote by $P^{(2)}$ the 2-homogeneous part of P. Define $\lambda = \max_{\|w\|=1} |P^{(2)}(w)|$, then*

$$\partial_H^\lambda u(x) \supseteq \partial_H U(x) + \nabla_H P(x).$$

PROOF. Recall that by Lemma 1.2.40, for every $x, h \in \mathbb{G}$ we have

$$P(xh) \geq P(x) + \langle \nabla_H P(x), h \rangle - \lambda \|h\|^2.$$

Let p be in $\partial_H U(x)$ then

$$U(xh) + P(xh) \geq U(x) + P(x) + \langle p + \nabla_H P(x), h \rangle - \lambda \|h\|^2,$$

whenever $[0,h] \subseteq x^{-1}\Omega \cap H_1$. This implies that $p + \nabla_H P(x) \in \partial_H^\lambda u(x)$. □

Proposition 3.1.14 *Let Ω be a subset of $\mathbb{G}$. Let U and V be respectively an h-convex function and a $C_H^1(\Omega)$ function, we define the map u as $u = U + V$. Fix $\lambda \geq 0$, then for every $x \in \Omega$ we have*

$$\partial_H^\lambda u(x) \subseteq \partial_H U(x) + \nabla_H V(x).$$

PROOF. Let p be in $\partial_H^\lambda u(x)$ and consider any $[0,h] \subseteq H_1 \cap x^{-1}\Omega$. We have

$$u(xh) \geq u(x) + \langle p, h \rangle - \lambda \|h\|^2,$$

that can be written as follows

$$U(xh) + V(xh) \geq U(x) + V(x) + \langle \nabla_H V(x), h \rangle + \langle p - \nabla_H V(x), h \rangle - \lambda \|h\|^2.$$

Thus, the C^1_H smoothness of V gives

$$U(xh) \geq U(x) + \langle p - \nabla_H V(x), h \rangle + o(\|h\|)\,,$$

hence $p - \nabla_H V(x) \in J^{1,-}_U(x)$. Since U is h-convex, in view of Remark 3.1.11 we have that $p - \nabla_H V(x) \in \partial_H U(x)$. This concludes the proof. □

In the following theorem we extend the classical non-smooth mean value theorem to stratified groups.

Theorem 3.1.15 *Let U be h-convex and let P be a polynomial, with h-deg $P \leq 2$ and $\lambda = \max_{\|w\|=1} |P^{(2)}(w)|$. We define the function u as $u = U + P$. Then for every $x \in \Omega$ and every h such that $[0,h] \subseteq H_1 \cap x^{-1}\Omega$, there exist $t \in [0,1]$ and $p \in \partial^\lambda_H u(x\delta_t h)$ such that*

$$u(xh) - u(x) = \langle p, h \rangle\,.$$

PROOF. Let U_i be a sequence of $C^\infty(\Omega)$ h-convex functions, converging to U uniformly on compact sets. Define $u_i = U_i + P$. For such functions the mean value theorem holds i.e. there exists $t_j \in [0,1]$ such that

$$u_i(xh) - u_i(x) = \langle \nabla_H u_i(x\delta_{t_i} h), h \rangle\,, \quad [0,h] \subset H_1 \cap x^{-1}\Omega.$$

Possibly passing to a subsequence we have $t_i \to t$ and $\nabla_H u_i(x\delta_{t_i} h) \to p$, thus by the uniform convergence

$$u(xh) - u(x) = \langle p, h \rangle\,.$$

Our claim follows if we prove that $p \in \partial^\lambda_H u(x\delta_t h)$. By Proposition 3.1.6, for every $k > 0$ there exists i_k such that

$$\nabla_H U_i(x\delta_{t_i} h) = \partial_H U_i(x\delta_{t_i} h) \subseteq \partial_H U(x\delta_t h) + \frac{1}{k}\mathbb{B}, \qquad \forall i \geq i_k$$

Moreover, possibly choosing a larger i_k, we have

$$\nabla_H U_i(x\delta_{t_i} h) + \nabla_H P(x\delta_{t_i} h) \subseteq \partial_H U(x\delta_t h) + \nabla_H P(x\delta_t h) + \frac{2}{k}\mathbb{B}, \qquad \forall i \geq i_k$$

By Lemma 3.1.13, $\partial^\lambda_H u(x) \supseteq \partial_H U(x) + \nabla_H P(x)$ thus the previous inclusion implies that

$$\nabla_H u_i(x\delta_{t_i} h) = \nabla_H U_i(x\delta_{t_i} h) + \nabla_H P(x\delta_{t_i} h) \subseteq \partial^\lambda_H u(x\delta_t h) + \frac{2}{k}\mathbb{B}, \qquad \forall i \geq i_k$$

then letting $k \to \infty$ we get that $p \in \partial^\lambda_H u(x\delta_t h)$. □

As an immediate consequence of the previous result, we have the following.

Corollary 3.1.16 *Let $u : \Omega \to \mathbb{R}$ be an h-convex function. Then for every $x \in \Omega$ and every h such that $[0,h] \subset H_1 \cap x^{-1}\Omega$, there exists $t \in [0,1]$ and $p \in \partial_H u(x\delta_t h)$ such that $u(xh) - u(x) = \langle p, h \rangle$.*

PROOF. It suffices to apply Theorem 3.1.15 with $P = 0$ and $\lambda = 0$. □

Remark 3.1.17 In the literature a nonsmooth mean value theorem can be found for Lipschitz mappings on Banach homogeneous groups, that clearly include stratified group, [92]. Unfortunately, this work does not imply our Theorem 3.1.16, since it uses the notion of Clarke generalized gradient for Lipschitz mappings adapted to homogeneous groups.

Definition 3.1.18 Let $\Omega \subset \mathbb{G}$ be an open set, and let $u : \Omega \to \mathbb{R}$ be a function. Take $h \in H_1$. The *horizontal directional derivative* of u at x, along h, is given by the limit

$$\lim_{\lambda \to 0^+} \frac{u(x\delta_\lambda h) - u(x)}{\lambda},$$

whenever it exists. We denote this derivative by $u'(x,h)$.

Corollary 3.1.19 *Let u be an h-convex function in Ω. Then for every $x \in \Omega$ and $h \in H_1$ the horizontal directional derivative $u'(x,h)$ exists and satisfies*

$$u'(x,h) = \max_{p \in \partial_H u(x)} \langle p, h \rangle, \tag{64}$$

hence it is subadditive with respect to the variable h.

PROOF. The h-convexity of u implies the existence of $u'(x,h)$ for any $x \in \Omega$ and $h \in H_1$. Let $p_0 \in \partial_H u(x)$ be such that $\langle p_0, h \rangle = \max_{p \in \partial_H u(x)} \langle p, h \rangle$. By definition of $\partial_H u(x)$,

$$u(x\delta_\lambda h) \geq u(x) + \langle p_0, \lambda h \rangle, \qquad \text{whenever } [0, \lambda h] \subset x^{-1}\Omega \cap H_1.$$

Then we easily get that

$$\lim_{\lambda \to 0^+} \frac{u(x\delta_\lambda h) - u(x)}{\lambda} \geq \langle p_0, h \rangle.$$

Notice that, for λ small enough, $[0, \lambda h] \subset x^{-1}\Omega \cap H_1$, hence we can apply Theorem 3.1.15. Then for every λ there exist $c(\lambda) \in [0,1]$ and $p(\lambda) \in \partial_H u(x\delta_{c(\lambda)\lambda} h)$ such that

$$\frac{u(x\delta_\lambda h) - u(x)}{\lambda} = \langle p(\lambda), h \rangle.$$

Now fix a sequence $\lambda_i \to 0$ such that $p(\lambda_i) \to \bar{p}$, then by the closure property of the subdifferential we get $\bar{p} \in \partial_H u(x)$. Moreover, the existence of the following limit gives

$$\lim_{\lambda \to 0^+} \frac{u(x\delta_\lambda h) - u(x)}{\lambda} = \langle \bar{p}, h \rangle \leq \max_{p \in \partial_H u(x)} \langle p, h \rangle,$$

concluding the proof. □

Theorem 3.1.20 (First order characterization) *Let $u : \Omega \longrightarrow \mathbb{R}$ be h-convex. Then u is h-differentiable at x if and only if $\partial_H u(x) = \{p\}$ and in this case $\nabla_H u(x) = \{p\}$.*

PROOF. Uniqueness of the h-subdifferential under h-differentiability has been already established in [26], see Remark 3.1.3. Let assume now that the h-subdifferential p of u at x is unique. Let U, W and M be as in Definition 1.2.19. Thus for any $w \in \mathbb{G}$ with $\|w\| = 1$ we have $w = \prod_{s=1}^{\gamma} a_s e_{i_s}$, for some $(a_1, \ldots, a_\gamma) \in U$. We fix $r > 0$ such that $B_{0,r} \subset x^{-1}\Omega$ and define the h-convex function

$$g(y) = u(xy) - u(x) - \langle p, y \rangle, \qquad y \in x^{-1}\Omega.$$

We choose $\rho_0 > 0$ such that $\rho_0 M < r$. Thus, whenever $0 < \rho < \rho_0$, by Theorem 3.1.16 and the generating property, we have

$$g(\delta_\rho w) = \sum_{s=1}^{\gamma} \langle p_s, \rho a_s e_{i_s} \rangle - \langle p, \rho a_s e_{i_s} \rangle$$

where $p_s \in \partial_H u \left(x\delta_\rho (\prod_{k=1}^{s-1} a_k e_{i_k}) \delta_{t_s} \delta_\rho a_s e_{i_s} \right)$ with $t_s \in [0, 1]$. By Proposition 3.1.6, for every $\epsilon > 0$ there exists ρ_0 such that

$$\partial_H u \left(x\delta_\rho (\prod_{k=1}^{s-1} a_s e_{i_k}) \delta_{t_s} \delta_\rho a_s e_{i_s} \right) \subseteq \partial_H u(x) + \epsilon \mathbb{B} = \{p\} + \epsilon \mathbb{B} \qquad \forall \rho < \rho_0, \quad s = 1, \ldots, \gamma.$$

Thus $|g(\delta_\rho w)| \leq C\gamma\epsilon\rho$, where C is independent on $(a_1, \ldots, a_\gamma)$, since W is a bounded set. This implies that $\dfrac{|g(\delta_\rho w)|}{\rho}$ uniformly converges to zero with respect to $w \in W$ as $\rho \to 0^+$. □

2. Second order differentiability

The aim of this section is to prove the characterization of the second order differentiability of h-convex functions.

Definition 3.2.1 We mean that $u : \Omega \to \mathbb{R}$ has a *second order h-expansion* at $x \in \Omega$ if there exists a polynomial $P_x : \mathbb{G} \to \mathbb{R}$ whose homogeneous degree is less than or equal to two, such that

$$u(xw) = P_x(w) + o(\|w\|^2). \tag{65}$$

Let us begin this section with the following simple fact.

Lemma 3.2.2 *Let $\Omega \subset \mathbb{G}$ be an open set, let $u : \Omega \to \mathbb{R}$ be a function. If u has a second order expansion at $x \in \Omega$, then u is h-differentiable at x and*

$$P_x^{(1)}(w) = \langle \nabla_H u(x), w \rangle. \tag{66}$$

PROOF. If (65) holds for u at $x \in \Omega$,then we can rewrite (65) as

$$u(xw) - P_x^{(0)}(w) - P_x^{(1)}(w) = P_x^{(2)}(w) + o(\|w\|^2).$$

Clearly $P_x^{(0)}(w) = u(x)$ and $P_x^{(1)}(w)$ is an h-linear map. Thus we achieve

$$|u(xw) - u(x) - P_x^{(1)}(w)| = o(\|w\|),$$

and the h-differentiability of u follows. In view of the uniqueness of the h-differential we get (66), concluding the proof. □

As in [91], we introduce the difference quotients of convex functions.

Definition 3.2.3 (Difference quotients, [91]) Let $u : \Omega \to \mathbb{R}$ be h-convex and assume that it is h-differentiable at x. For every $\tau > 0$ define the *first order h-quotient* at x

$$u_{x,\tau}(w) = \tau^{-1}\{u(x\delta_\tau w) - u(x)\}$$

and *second h-differential quotient* at x

$$\Delta^2_{x,\tau}u(w) = \frac{u(x\delta_\tau w) - u(x) - \tau\,\langle \nabla_H u(x), w\rangle}{\tau^2}\,. \tag{67}$$

assuming in addition that u is h-differentiable at x. At this h-differentiability point, the h-difference quotient of the subdifferential mapping is given by the set-valued mapping

$$\Delta_{x,\tau}\partial_H u : w \rightrightarrows \frac{\partial_H u(x\delta_\tau w) - \nabla_H u(x)}{\tau}. \tag{68}$$

Remark 3.2.4 Notice that $\Delta^2_{x,\tau}u$ can be written as

$$\Delta^2_{x,\tau}u(w) = \tau^{-1}\left[u_{x,\tau}(w) - \langle \nabla_H u(x), w\rangle\right]$$

where $u_{x,\tau}$ is clearly h-convex. Moreover if we take the subdifferential of $\Delta^2_{x,\tau}u$ we get

$$\begin{aligned} \partial_H\left[\Delta^2_{x,\tau}u(w)\right] &= \tau^{-1}\left\{\partial_H u_{x,\tau}(w) - \nabla_H u(x)\right\} \\ &= \tau^{-1}\left\{\partial_H u(x\delta_\tau w) - \nabla_H u(x)\right\} \\ &= \Delta_{x,\tau}\partial_H u(w). \end{aligned} \tag{69}$$

where the equality $\partial_H u_{x,\tau}(w) = \partial_H u(x\delta_\tau w)$ follows from the definition of $u_{x,\tau}$.

Definition 3.2.5 A locally Lipschitz function $u : \Omega \to \mathbb{R}$ is *twice h-differentiable at x* if there exists the *horizontal gradient* of u at x, and moreover there exists a linear map $A_x : \mathbb{G} \to H_1$ such that

$$\left\| \frac{\nabla_H u(xw) - \nabla_H u(x) - A_x(w)}{\|w\|} \right\|_{L^\infty(B_\delta, H_1)} \longrightarrow 0 \quad \text{as} \quad \delta \to 0^+\,, \tag{70}$$

If (70) holds we also say that $\nabla_H u$ is *h-differentiable at x in the extended sense*. We call A_x the *second order h-differential* of u at x and denote it by $D^2_H u(x)$, since it is uniquely defined. The notion of differentiability in the extended sense is well posed, since Lipschitz functions are almost everywhere h-differentiable, [85]. Differentiability in the extended sense in the Euclidean case has been introduced by Rockafellar, [91]. The next lemma establishes a precise characterization of this differentiability.

Lemma 3.2.6 *Let $u : \Omega \longrightarrow \mathbb{R}$ be h-convex. Then u is twice h-differentiable at x if and only if there exist an h-linear mapping $A_x : \mathbb{G} \to H_1$ and $v \in H_1$ such that*

$$\partial_H u(xw) \subseteq v + A_x(w) + o(\|w\|)\mathbb{B} \tag{71}$$

for all $w \in x^{-1}\Omega$, where $\mathbb{B}$ denote the unit ball in H_1. In particular, if (71) *holds, then $v = \nabla_H u(x)$.*

PROOF. Choosing $w = 0$ we get $\partial_H u(x) = \{v\}$, thus by Theorem 3.1.20, u is h-differentiable at x, moreover $v = \nabla_H u(x)$. The twice h-differentiability immediately follows from (71), taking its restriction to all h-differentiable points. For the converse implication, we rewrite expansion (70) as follows, for all $\epsilon > 0$ there exists $\rho > 0$ such that

$$\left| \frac{\nabla_H u(xh) - \nabla_H u(x) - A_x(h)}{\|h\|} \right| \leq \epsilon \qquad \|h\| < \rho. \tag{72}$$

for all $h \in x^{-1}\Omega$ such that u is h-differentiable at xh. By (53), for any $w \in x^{-1}\Omega \cap B_{0,\rho}$, taking into account (72), we get

$$\left|\frac{p - \nabla_H u(x) - A_x(w)}{\|w\|}\right| \leq \epsilon \qquad \text{for all } p \in \nabla_H^\star u(xw).$$

In an equivalent form, we have

$$\nabla_H^\star u(xw) \subseteq \nabla_H u(x) + A_x(w) + \epsilon\|w\|\mathbb{B}. \tag{73}$$

Moreover, the set on the right is convex, hence Theorem 3.1.8 yields

$$\partial_H u(xw) = \bar{co}\left(\nabla_H^\star u(xw)\right) \subseteq \nabla_H u(x) + A_x(w) + o(\|w\|)\mathbb{B}. \tag{74}$$

This leads us to the conclusion. □

The previous lemma is an important tool to establish one implication of the characterization of second order differentiability of h-convex functions, stated in Theorem 3.2.8. It can be seen as a "set inclusion continuity" of the subdifferential joined with a first order expansion of the horizontal gradient, at those points where it exists.

Corollary 3.2.7 *If $u : \Omega \to \mathbb{R}$ is h-convex, then it is twice h-differentiable at x if and only if for any compact set $D \Subset \Omega$ and for all $\epsilon > 0$, there exists $\delta > 0$ such that for all $w \in W$ and $\tau \in (0, \delta)$ we have*

$$\Delta_{x,\tau}\partial_H u(w) - A_x(w) \subseteq \epsilon\mathbb{B}. \tag{75}$$

PROOF. Let u be twice h-differentiable at x, fix a bounded set $D \Subset \Omega$ and $\epsilon > 0$. We set $\mu_D = \max_{w\in D}\|w\|$. Then there is $\rho(\epsilon, D, \Omega) > 0$ such that

$$\partial_H u(xw) \subset \nabla_H u(x) + A_x(w) + \frac{\|w\|\epsilon}{\mu_D}\mathbb{B},$$

whenever $\|w\| < \rho(\varepsilon, D, \Omega)$ and $B_{x,\rho(\varepsilon,D,\Omega)} \subset \Omega$. We consider $w = \delta_\tau h$, where $h \in D$, and $0 < \tau < \frac{\rho(\varepsilon,D,\Omega)}{\mu_D}$. It follows that

$$\partial_H u(x\delta_\tau h) \subset \nabla_H u(x) + \tau A_x(h) + \epsilon\tau\mathbb{B}$$

which is equivalent to (75). Conversely, let $S = \{w \in \mathbb{G} \,:\, \|w\| = 1\}$ be a compact set and fix $\varepsilon > 0$. Then there exists $\delta > 0$ such that (75) holds whenever $0 < \tau < \delta$. Thus, we have

$$\frac{\partial_H u(x\delta_\tau w) - \nabla_H u(x)}{\tau} - A_x(w) \subseteq \epsilon\mathbb{B}.$$

In other words, whenever $0 < \|h\| < \delta$, we have

$$\partial_H u(xh) \subseteq \nabla_H u(x) + A_x(h) + \epsilon\|h\|\mathbb{B},$$

that establishes the twice h-differentiability of u at x.

□

Finally we can prove our main result.

Theorem 3.2.8 (Second order characterization) *Let $u : \Omega \longrightarrow \mathbb{R}$ be h-convex and let x be a point in Ω. Then u has a second order h-expansion at x if and only if it is twice h-differentiable at x. In addition, in this case the following facts hold*

(1) *the gradient* $\nabla_{V_2} u(x) = \big(X_{m_1+1}u(x), \dots, X_{m_2}u(x)\big)$ *of* u *at* x *along* V_2 *exists, where* $(X_{m_1+1}, \dots, X_{m_2})$ *is an orthonormal basis of the second layer* V_2,

(2) *denoting by* P_x *the second order h-expansion of* u *at* x, *we have*

$$P_x(w) = u(x) + \big\langle \big(\nabla_H u(x) + \nabla_{V_2} u(x)\big), w \big\rangle + \frac{1}{2}\, \langle \nabla^2_H P_x\, w, w \rangle,$$

where $(\nabla^2_H)_{ij} = \frac{X_iX_j + X_jX_i}{2}$, $i, j = 1, \dots, m_1$ *is the* horizontal Hessian,

(3) *denoting by* A_x *the h-differential of* $\nabla_H u$ *in the extended sense at* x, *then its connection with* P_x *is given by the formula*

$$\big(\nabla^2_H P_x\big)_{ij} = (A_x)^i_j - \sum_{l=m_1+1}^{m_2} a^{li}_j \; X_l u(x)\,,$$

where a^{li}_j *only depend on the coordinates of the group and appear in* (29), *the horizontal Hessian* $\nabla^2_H P_x$ *is nonnegative and* $X_iX_jP_x = (A_x)^i_j$.

PROOF. Let us assume that u has a second order h-expansion at x. By Proposition 3.2.2, u is h-differentiable at x, then $P^{(0)}_x(w) = u(x)$ and $P^(_x1)(w) = \langle \nabla_H u(x), w\rangle$, where P_x is the polynomial associated to the second order h-expansion. Define $\phi(w) := P^{(2)}_x(w)$ and notice that $\nabla_H P^{(2)}_x(w)$ is an h-linear map, since it is a polynomial of homogeneous degree 1. The second order h-expansion yields

$$\Delta^2_{x,\tau} u(w) - \phi(w) = \frac{u(x\delta_\tau w) - P^{(0)}_x(\delta_\tau w) - P^{(1)}_x(\delta_\tau w) - P^{(2)}_x(\delta_\tau w)}{\tau^2} = \frac{o(\|\delta_\tau w\|^2)}{\tau^2}. \tag{76}$$

As a consequence, $\Delta^2_{x,\tau} u$ uniformly converges to ϕ on compact sets as $\tau \to 0^+$. Moreover $\Delta^2_{x,\tau} u$ is h-convex, then so is ϕ. Applying Theorem 3.1.6, we can establish that for every compact set $D \subset \Omega$ and $\epsilon > 0$ there exists $\delta > 0$ such that

$$\partial_H U_{x,\tau}(w) \subseteq \nabla_H \phi(w) + \epsilon \mathbb{B}, \quad \text{for all} \quad w \in D \quad \text{and} \quad \tau \in (0, \delta).$$

Notice that (69), gives $\partial_H U_{x,\tau}(w) = \Delta_{x,\tau} \partial_H u(w)$, hence

$$\Delta_{x,\tau} \partial_H u(w) \subseteq \nabla_H P^{(2)}_x(w) + \epsilon \mathbb{B}.$$

As a result, we have $\Delta_{x,\tau} \partial_H u(w) - \nabla_H P^{(2)}_x(w) \subset \epsilon \mathbb{B}$ whenever $w \in D$ and $0 < \tau < \delta$. By Proposition 3.2.7, u is twice h-differentiable. Furthermore, $\nabla_H P^{(2)}_x$ is the second order h-differential $D^2_H u(x)$ of u at x.

We now assume that u is twice h-differentiable at x, where $D^2_H u(x)$ denotes the second order h-differential of u at x. By Lemma 3.2.6 we have

$$\nabla_H u(xw) = \nabla_H u(x) + D^2_H u(w) + o(\|w\|).$$

where $D^2_H u(x)$ is regarded as an h-linear mapping. Let U, W and M be as introduced in Definition (1.2.19). We define

$$v(w) = u(xw) - u(x) - P_x(w)$$

for every $w \in x^{-1} \cdot \Omega$ and P_x is the unique polynomial, with h-deg$P \leq 2$ that satisfies the condition $P_x(0) = 0$ and

$$\nabla_H P_x(w) = \nabla_H u(x) + D^2_H u(x) w. \tag{77}$$

This is a consequence of Remark 1.2.36. Let $r > 0$ be such that $B_{0,r} \subset x^{-1}\Omega$. Let ρ_0 be such that $\rho_0 M < r$ and choose w such that $\|w\| = 1$. We consider for every $0 < \rho < \rho_0$ and $\delta_\rho w = \prod_{s=1}^{\gamma} a_s e_{i_s}$, for some $(a_1, \dots, a_\gamma) \in U$. Then $v(\delta_\rho w) = v(\delta_\rho w) - v(0)$ can be written as

$$v(\delta_\rho w) = \sum_{s=1}^{\gamma} v(\prod_{l=1}^{s} \delta_\rho a_{i_l} e_{i_l}) - v(\prod_{l=1}^{s-1} \delta_\rho a_{i_l} e_{i_l}).$$

Observe that v is an h-convex function plus a polynomial of homogeneous degree less than or equal to two. By Theorem 3.1.15 applied to horizontal directions e_{i_s} we get

$$v(\delta_\rho w) = \sum_{i=1}^{\gamma} \left\langle p_s, \delta_\rho a_s e_{i_s} \right\rangle$$

with $p_s \in \partial^\lambda_H v \left(x\delta_\rho (\prod_{k=1}^{s-1} a_k e_{i_k}) \delta_{t_s} \delta_\rho a_s e_{i_s} \right)$, $\lambda = \max_{\|h\|=1} |P^{(2)}_x(h)|$, where $t_s \in [0,1]$ and $\lambda = \max_{\|h\|=1} |P^{(2)}_x(h)|$. Moreover, by Proposition 3.1.14 we know that

$$p_s + \nabla P_x \left(\delta_\rho (\prod_{k=1}^{s-1} a_k e_{i_k}) \delta_{t_s} \delta_\rho a_s e_{i_s} \right) \in \partial_H u \left(x\delta_\rho (\prod_{k=1}^{s-1} a_k e_{i_k}) \delta_{t_s} \delta_\rho a_s e_{i_s} \right). \tag{78}$$

The expansion (71) for the h-subdifferential of u implies that

$$\begin{aligned} \partial_H u \left(x\delta_\rho (\prod_{k=1}^{s-1} a_k e_{i_k}) \delta_{t_s} \delta_\rho a_s e_{i_s} \right) \quad \subset \quad & \nabla_H u(x) + A_x \left(\delta_\rho (\prod_{k=1}^{s-1} a_k e_{i_k}) \delta_{t_s} \delta_\rho a_s e_{i_s} \right) \\ & + o \left(|\delta_\rho (\prod_{k=1}^{s-1} a_k e_{i_k}) \delta_{t_s} \delta_\rho a_s e_{i_s}| \right) \mathbb{B}, \end{aligned} \tag{79}$$

Thus, by formula (77), taking into account (78) and (79), we get that

$$|p_s| = o \left(|\delta_\rho (\prod_{k=1}^{s-1} a_k e_{i_k}) \delta_{t_s} \delta_\rho a_s e_{i_s}| \right) = o(\rho).$$

As a consequence, $|v(\delta_\rho w)| = o(\rho^2)$. This concludes the proof of our characterization. Next we have to prove the claims (1), (2) and (3). The first one follows considering the restriction of (76) to directions $z \in H_2$, with $|z| = 1$, and taking into account (27), hence getting the uniform limit

$$\frac{u(x \cdot \exp(t^2 Z)) - u(x) - t^2 \langle \nabla_{V_2} P^{(2)}_x, z \rangle}{t^2} \longrightarrow 0$$

as $t \to 0^+$, where Z is the unique left invariant vector field such that $Z(0) = z$. In fact, we have used the equality

$$x\delta_t z = x \cdot \delta_t \exp(Z) = x \cdot \exp(t^2 Z).$$

In particular, we have $\nabla_{V_2} u(x) = \nabla_{V_2} P$. Taking into account Lemma 3.2.2 and formula (27), then claim (2) follows. Now, with respect to the fixed basis $(e_1, \dots, e_n)$ of $\mathbb{G}$, we have coefficients $(D^2_H u(x))_{ij}$ such that

$$D^2_H u(x) w = \sum_{i,j=1}^{m_1} (D^2_H u(x))_{ij}\, w_i\, e_j\,,$$

therefore (77) yields $\nabla_H P^{(2)}_x(w) = D^2_H u(x) w$. For any $j = 1, \dots, m_1$, we have

$$X_j P^{(2)}_x(w) = \sum_{i=1}^{m_1} (D^2_H u(x))_{ij}\, w_i\,,$$

then formula (30) gives

$$X_i X_j P^{(2)}_x = (D^2_H u(x))_{ij} = \left(\nabla^2_H P^{(2)}_x\right)_{ij} + \sum_{l=m_1+1}^{m_2} X_l u(x)\; a^{li}_j\,.$$

As a result, we get

$$(\nabla^2_H P^{(2)}_x)_{ij} = (D^2_H u(x))_{ij} - \sum_{l=m_1+1}^{m_2} X_l u(x)\; a^{li}_j\,,$$

that coincides with the formula of claim (3). Finally, we observe that $P^{(2)}_x$ is the uniform limit on compact sets of the h-convex functions $\Delta^2_{x,\tau} u(w)$. This implies that $P^{(2)}_x$ is also h-convex and then its symmetrized horizontal Hessian is nonnegative. □

CHAPTER 4

H-CONVEX DISTRIBUTIONS IN STRATIFIED GROUPS

In this chapter, we address the question of the characterization of h-convex distributions. The results in the first part of the chapter have been obtained in a joint work with A. Bonfiglioli, E. Lanconelli and V. Magnani, [15]. In all stratified groups, every h-convex function has nonnegative *horizontal Hessian* in the distributional sense, as observed in [26] and [64]. Surprisingly, the converse to this fact, namely, establishing whether a distribution with nonnegative horizontal Hessian is given by an h-convex function was not addressed yet. This chapter gives a full answer to this question. The first result in this direction extends the Reshetnyak's characterization to all stratified groups.

Theorem 4.1.6 *If $\mu \in \mathcal{D}'(\Omega)$ is a Radon measure, then μ is defined by an h-convex function if and only if it is an h-convex distribution.*

Our scheme is elementary, although it differs from the standard approach: we consider the group convolution of the measure μ, but instead of computing its horizontal Hessian by direct differentiation, we consider its distributional version. This respects the noncommutativity of the convolution operator. As a byproduct of Theorem 4.1.6, we have the following important corollary.

Corollary 4.1.7 *If $u \in L^1_{loc}(\Omega)$ is h-convex in the distributional sense, then outside a negligible set it coincides with a locally Lipschitz continuous h-convex function on Ω.*

This result also extends one of the characterizations given in [60], where the subharmonic theory used to prove the equivalence of various notions of h-convexity requires that the function be also upper semicontinuous. To reach the complete distributional characterization of h-convexity, we combine Corollary 4.1.7 and Lemma 4.1.8. Using the fundamental solution of the sub-Laplacian Δ_H in stratified groups, [34], this lemma shows that an h-convex distribution T can be written as the sum of a Δ_H-harmonic function and a locally summable function. Since Δ_H-harmonic functions are smooth by Hörmander's theorem, [57], we conclude that T is given by a function in $L^1_{loc}(\Omega)$.

We should also remark that this approach seems to be new also in the classical context. In fact, Dudley's proof uses a special geometric construction that cannot be extended to general stratified groups. On the other hand, the rich geometric properties of Heisenberg groups allow to carry out the Dudley's scheme in these groups, this is done in Section 2. We are now arrived to our main result of the chapter.

Theorem 4.1.9 *Let Ω be an open set of $\mathbb{G}$. If $T \in \mathcal{D}'(\Omega)$ is h-convex, then T is defined by an h-convex function on Ω.*

Recall that all measurable h-convex functions are locally Lipschitz continuous, [87]. Thus, Theorem 4.1.9 shows that the class of h-convex measurable functions coincides with that of h-convex distributions, that are locally Lipschitz continuous h-convex functions. Although we still do not know whether one can find h-convex functions in higher step groups that are nonmeasurable, these functions certainly would not be included in the previous families. This confirms that the natural notion of h-convexity in stratified groups should always include either measurability or local boundedness from above, that are indeed equivalent conditions.

It should be eventually apparent that the main point of this note is to show how the well developed theory of subharmonic functions in stratified groups allows us to use simple arguments to establish new results in the realm of h-convex functions.

In Section 2 we follows the Dudley approach to give a more geometric proof Theorem 4.1.9, unfortunately this holds only in Heisenberg groups.

Moreover, the same scheme to prove Theorem 4.1.6, joined with Lemma 4.2.3 leads us to a more general distributional approximation theorem.

Theorem 4.2.4 *Let $\Omega \subset \mathbb{G}$ be an open bounded subset, let $T \in \mathcal{D}'(\Omega)$ be such that $D^2_H T \geq 0$ and let $\delta > 0$ be such that $\Omega_{-\delta}$ is nonempty. It follows that there exists $C(\Omega) > 0$ such that the convolutions $\langle T, \Phi_{x,\varepsilon} \rangle$ are smooth h-convex functions on $\Omega_{-\delta}$ for all $0 < C(\Omega)\varepsilon^{1/\iota} < \delta$ and converge to T in $\mathcal{D}'(\Omega_{-\delta})$, where $\iota > 1$ is the step of $\mathbb{G}$.*

It is worth to stress that this theorem shows in particular that smooth convolutions of h-convex distributions are smooth h-convex functions. We also notice that in the case T is given by a measurable h-convex function, the smooth convolution $\langle T, \Phi_{x,\varepsilon} \rangle$ defined in Lemma 4.2.3 uniformly converges to u on compact sets, since measurable h-convex functions are locally Lipschitz continuous, [87]. This fact was already observed in [60], since in this case the h-convex function u is in particular upper semicontinuous.

Concerning the hypotheses on Ω in Theorem 4.2.4, the requirement that Ω be bounded only depends on the noncommutativity of the group, as explained in Remark 4.2.6. It is also clear that one can easily modify the conclusion of Theorem 4.2.4 in the case of an arbitrary open set, according to Corollary 4.2.5.

It is now natural to see the role of Theorem 4.2.4, in the problem of showing that h-convex distributions are given by h-convex functions. Unfortunately, the fact that an h-convex distribution T is the distributional limit of smooth h-convex functions is not enough to conclude that T is given by an h-convex function. To establish this result, we need some compactness, namely, we have to establish a closure theorem. This is a delicate issue, that requires more information on the geometry on the group. The second part of this work is devoted to the proof of this closure theorem in all Heisenberg groups $\mathbb{H}^n$.

Theorem 4.2.15 *Let Ω be an open set of $\mathbb{H}^n$. If $u_n : \Omega \to \mathbb{R}$ is a sequence of h-convex functions converging to $T \in \mathcal{D}'(\Omega)$ in the distributional sense, then T is defined by an h-convex function.*

This theorem in Euclidean spaces has been proved by Dudley, see Theorem 2.1 in [29]. As it is clear from his proof, the geometry of the space plays an important role. Dudley's approach uses the following elementary fact: given a convex function u and a simplex S, then u attains its maximum on S at a vertex q. This allows to consider the set S', that is symmetric to S with respect to the vertex q. It follows that $u(y) \geq u(q)$ for every $y \in S'$. Thus, in the case the sequence of convex functions diverges to $+\infty$ at some point, we get its uniform limit to $+\infty$, contradicting the distributional convergence. This constitutes the leading idea in Dudley's approach.

Our first step is to replace simplexes with *finitely h-convex* sets in $\mathbb{H}^n$, namely, finite sets whose h-convex closure has nonempty interior. The existence of these sets is proved by Rickly in [87] for step two groups. He also shows that this result is false in higher step groups, as the Engel group. However, this h-convex closure is not as manageable as a simplex in the Euclidean space. This forces us to establish some technical lemmas to extract the geometric properties of this set. In Proposition 4.2.12, we modify the finitely h-convex set A_0, getting the new finite set A_1, such that at any vertex $p \in A_1$ the horizontal plane H_p through p contains an interior point of the h-convex closure $\mathcal{C}(A_0)$. However, this lemma can give a uniform lower bound on the sequence only on an open piece of a horizontal plane, since h-convexity is only defined on horizontal planes. The lower bound on an open set is obtained by moving the base point p along a horizontal direction, making the horizontal plane rotate around this direction, see Lemma 4.2.14. This leads to the proof of the uniform upper bound on the sequence. Notice that here we have used the fact that we are working in Heisenberg groups. By a technique similar to the one used for Theorem 3.17 of [71], we also get the uniform lower bound on the sequence. Finally, by the L^∞ estimates of h-convex functions with respect to their L^1 norm, see for instance Theorem 9.2 of [26], we reach the wished compactness.

After this result, in Heisenberg groups the hypothesis in Theorem 4.1.6 that the h-convex distribution is a Radon measure can be removed, getting the complete distributional characterization of h-convexity. Precisely, we have the following theorem.

Theorem 4.2.16 *Let Ω be an open set of $\mathbb{H}^n$. If $T \in \mathcal{D}'(\Omega)$ is h-convex, then T is defined by an h-convex function in Ω.*

1. H-convex distributions

Remark 4.1.1 Let $(X_1, \ldots, X_m)$ denote an orthonormal basis of the first layer V_1 and let $T \in \mathcal{D}'(\Omega)$. The vector fields X_j have formal adjoint $X_j^\star = -X_j$, this justifies the following definition of X_jT,

$$\langle X_jT, \phi\rangle = -\langle T, X_j\phi\rangle, \quad \phi \in \mathcal{D}(\Omega).$$

Notice that X_jT is a distribution. Since T is a distribution, for any compact set $K \subset \Omega$, there exists $C_K > 0$ and $N > 0$ such that

$$|\langle T, \phi\rangle| \leq C_K \|\phi\|_N, \quad \text{for all } \phi \in \mathcal{D}(K),$$

where $\|\phi\|_N = \sup_{y\in\Omega}\{|D^\alpha\phi(y)|, \alpha \in 1, \ldots, n^k, k = 1, \ldots, N\}$. Therefore,

$$|\langle X_jT, \phi\rangle| = |\langle T, X_j\phi\rangle| \leq C_K\|X_j\phi\|_N \leq C_K'\|\phi\|_{N+1},$$

for every $\phi \in \mathcal{D}(K)$ and this proves the claim. Since X_j in general do not commute, the order of the iterated differential operators X_j is important. We have

$$\langle X_{i_1} \dots X_{i_k} T, \phi \rangle = (-1)^k \langle T, X_{i_k} \dots X_{i_1} \phi \rangle, \quad \phi \in \mathcal{D}(\Omega).$$

Definition 4.1.2 Let $T \in \mathcal{D}'(\Omega)$. The *distributional Hessian of* T is the matrix valued distribution $\langle D_H^2 T,\ \psi \rangle := \langle T,\ \nabla_H^2 \psi \rangle$ of entries

$$\left\langle \frac{1}{2} \left(X_j X_i + X_i X_j \right) T,\ \psi \right\rangle := \left\langle T,\ \frac{1}{2} \left(X_i X_j + X_j X_i \right) \psi \right\rangle$$

for every $i, j = 1, \dots, m$ and every $\psi \in C_c^\infty(\Omega)$.

Definition 4.1.3 We say that the distributional Hessian of $T \in \mathcal{D}'(\Omega)$ is *nonnegative* if for every nonnegative test function $\psi \in C_c^\infty(\Omega)$ the matrix $\langle T,\ \nabla_H^2 \psi \rangle$ is nonnegative. In this case we write $D_H^2 T \geq 0$ and say that the distribution T is *h-convex*.

In the smooth case, a simple computation shows that a nonnegative horizontal Hessian characterizes h-convexity, see for instance [71].

Definition 4.1.4 Let $\phi \in C_c^\infty(\mathbb{G})$ be a nonnegative function, whose support is contained in the unit open ball of $\mathbb{G}$ with respect to the fixed homogeneous norm. For every $\varepsilon > 0$, we set $\phi_\varepsilon(x) = \varepsilon^{-Q} \phi(\delta_{\frac{1}{\varepsilon}} x)$. We say that $\{\phi_\varepsilon\}_{\varepsilon > 0}$ is a *family of mollifiers*. For all $x \in \mathbb{G}$, we define the functions $\Phi_{x,\varepsilon} : \mathbb{G} \to \mathbb{R}$ as $\Phi_{x,\varepsilon}(y) = \phi_\varepsilon(xy^{-1})$.

Lemma 4.1.5 (Remark 3.10 of [70]) *Let $\mathbb{G}$ be a stratified group of step ι. Let $w, h \in \mathbb{G}$ and let $\nu > 0$ be such that $d(w), d(h) \leq \nu$. Then there exists a constant $C(\nu)$ only depending on $\mathbb{G}$ such that $d(w^{-1} h w) \leq C(\nu) d(h)^{\frac{1}{\iota}}$.*

We will use the notation $\Omega_{-r} = \{x \in \Omega : \operatorname{dist}(x, \Omega^c) > r\}$ for any $r > 0$.

Theorem 4.1.6 *If $\mu \in \mathcal{D}'(\Omega)$ is a Radon measure, then μ is defined by an h-convex function if and only if it is an h-convex distribution.*

PROOF. We first suppose that Ω is bounded. Let $h > 0$ be such that Ω_{-h} is nonempty. Let

$$\mu_\varepsilon(x) = \int_\Omega \phi_\varepsilon(xy^{-1})\, d\mu(y)$$

is well defined on Ω for all $\varepsilon > 0$. Then for any $\psi \in \mathcal{D}(\Omega_{-h})$, we get

$$\begin{aligned}
\int_{\Omega_{-h}} \nabla_H^2 \psi(x)\ \mu_\varepsilon(x)\, dx &= \int_\Omega \left(\int_{\Omega_{-h}} \phi_\varepsilon(xy^{-1})\ \nabla_H^2 \psi(x)\, dx \right) d\mu(y) \\
(\, x = zy\,) \quad &= \int_\Omega \left(\int_{D_\varepsilon} \phi_\varepsilon(z)\ \nabla_H^2 \psi(zy)\, dz \right) d\mu(y) \\
&= \int_{D_\varepsilon} \phi_\varepsilon(z) \left(\int_\Omega\ \nabla_H^2 [\psi(zy)]\, d\mu(y) \right) dz \\
&\geq\ 0\,,
\end{aligned}$$

since $y \to \psi(zy)$ is smooth and compactly supported in Ω. In fact, let $\omega \in \Omega^c$, then applying Lemma 4.1.5, with $\nu = C(\Omega) = \sup_{x\in\Omega} d(x)$, we have

$$d(y,\omega) \geq d(zy,\omega) - d(y^{-1}zy) > h - C(\Omega)\varepsilon^{\frac{1}{\iota}} > 0.$$

for all $\varepsilon^{\frac{1}{\iota}} < \frac{h}{C(\Omega)}$. Then it follows from Proposition 5.1 in [71] that μ_ε is h-convex. Furthermore, for every compact set K contained in Ω_{-h} we have the uniform estimate

$$\int_K |\mu_\varepsilon(x)|\, dx \leq |\mu|(K)| < +\infty\,.$$

Then the L^∞-estimates in Theorem 9.2 of [26] joined with the classical Ascoli-Arzelà compactness theorem imply the existence of a continuous h-convex function $\tilde{u}$ on Ω_{-h} that a.e. coincides with u. The arbitrary choice of $h > 0$ concludes the proof in the case of Ω bounded. For a general open set Ω we have to consider for every $M > 0$ the set $\Omega \cap B_M$, the proof proceed as in the previous case, then by the arbitrary choice of $M > 0$ we can conclude. □

As a byproduct of Theorem 4.1.6, we have the following important corollary.

Corollary 4.1.7 *If $u \in L^1_{loc}(\Omega)$ is h-convex in the distributional sense, then outside a negligible set it coincides with a locally Lipschitz continuous h-convex function on Ω.*

Lemma 4.1.8 *Let $\Omega \subset \mathbb{G}$ be a open set and let $\Omega_1 \subset \Omega$ be a bounded open set such that $\bar{\Omega}_1 \subset \Omega$. If $T \in \mathcal{D}'(\Omega)$ satisfies $\Delta_H T \geq 0$, then its restriction to Ω_1 is given by function in $L^1_{loc}(\Omega_1)$.*

PROOF. Since $\Delta_H T \geq 0$, we know that there exists a nonnegative Radon measure μ on Ω such that $\Delta_H T = \mu$. Let Γ be as in Definition 1.2.42, and consider the function

$$v(x) = -\int_{\Omega_1} \Gamma(y^{-1}x)d\mu(y).$$

Since Γ is locally integrable on $\mathbb{G}$, for every compact set $K \subset \Omega_1$ we have

$$\int_K |v(x)|dx \leq \int_{\Omega_1}\int_K |\Gamma(y^{-1}x)|dxd\mu(y) \leq \mu(\Omega_1) \sup_{y\in\bar{\Omega}_1} \int_K |\Gamma(y^{-1}x)|dx < +\infty.$$

Let us show that v satisfies the distributional equality $\Delta_H v = \mu_{|\Omega_1}$. In fact, for every $\phi \in C^\infty_c(\Omega_1)$ we have

$$\langle \Delta_H v, \phi\rangle = -\int_{\Omega_1}\int_{\Omega_1} \Gamma(y^{-1}x)\Delta_H\phi(x)dx\, d\mu(y).$$

Thus, being Γ the fundamental solution for Δ_H, we get

$$\langle \Delta_H v, \phi\rangle = \int_{\Omega_1} \phi(y)d\mu(y).$$

Hence $\Delta_H (T - v) = 0$ in $\mathcal{D}'(\Omega_1)$. Since Δ_H is hypoelliptic, [57], the function $T - v$ coincides with a smooth Δ_H-harmonic function h on Ω_1, up to a negligible set. Finally we can conclude that T is represented by an L^1_{loc} function on Ω_1. □

We wish to point out that this lemma follows the same lines to prove representation formulas for upper semicontinuous subharmonic functions, see Theorem 9.4.4 in [14]. Combining Lemma 4.1.8 and Corollary 4.1.7, we establish the proof of our main result.

Theorem 4.1.9 *Let Ω be an open set of $\mathbb{G}$. If $T \in \mathcal{D}'(\Omega)$ is h-convex, then T is defined by an h-convex function on Ω.*

PROOF. Let $\bar{\Omega}_n \subset \Omega$, $n \in \mathbb{N}$ be an increasing sequence of open and bounded sets. By hypothesis $D^2_H T \geq 0$ and in particular $\Delta_H T \geq 0$, hence Lemma 4.1.8 implies that T is represented by an L^1_{loc} function on Ω_n. Finally we can conclude, by Corollary 4.1.7, that T is defined by an h-convex function on Ω_n. Since $\bigcup_{n\in\mathbb{N}} \Omega_n = \Omega$, the theorem follows. □

2. Distributional h-convexity in Heisenberg groups: a classical approach

In this section we give a different proof of Theorem 4.1.9 for Heisenberg groups. The proof follows the ideas of Dudley in [29], and the main point is to prove a closure theorem for convex function in $\mathcal{D}'$. We extend this result to h-convex functions in Heisenberg groups, however the proof follows a different approach to the Euclidean one.

First, we prove two technical lemmas, which holds in general Carnot groups. There are three common ways of viewing the elements of the Lie algebra $\mathfrak{g}$ of $\mathbb{G}$: (i) as tangent vectors at the origin, (ii) as left-invariant vector fields or (iii) as right-invariant vector fields. Throughout the thesis we have always used definition (ii), see Chapter 1. However in this section we shall need also (iii). We will use the notations $\mathfrak{g}_L$ and $\mathfrak{g}_R$ respectively for $\mathfrak{g}$ seen as in definition (ii) and (iii). We observe that the map Δ which sends $X \in \mathfrak{g}_L$ to the unique $\tilde{X} \in \mathfrak{g}_R$ which agrees with X at the origin is an anti-isomorphism. As a consequence if one applies the Barker-Campbell-Hausdorff formula to $\mathfrak{g}_R$, one obtains a different, but isomorphic, group law for $\mathbb{G}$. Let $Y_1, \dots, Y_m$ the basis of $\mathfrak{g}_R$ obtained as image of $X_1, \dots, X_m$ through Δ.

Proposition 4.2.1 (Corollary 1.44 in [35]) *Let $f : \mathbb{G} \to \mathbb{R}^n$ be a smooth function. Then for every $k > 0$ there exists a constant $C(k) > 0$ such that*

$$|f(xy) - P_x(y)| \leq C(k)|y|^{k+1} \sup_{|z|<\beta^{k+1}|y|,\ d(I)=k+1} |X^I f(xz)|,$$

where $P_x(y)$ is the right Taylor polynomial of f of homogeneous order k at x, and β is a constant depending only on $\mathbb{G}$.

Throughout the section we define the functions $\Phi_{x,\varepsilon} : \mathbb{G} \longrightarrow \mathbb{R}$ as $\Phi_{x,\varepsilon}(y) = \phi_\varepsilon(xy^{-1})$ for all $x, y \in \mathbb{G}$ and $\varepsilon > 0$. Where the functions ϕ_ε, $\varepsilon > 0$ are defined in Proposition 1.2.28.

Lemma 4.2.2 *Let $x, z \in \mathbb{G}$ then $\Phi_{x,\varepsilon}$ converges to $\Phi_{z,\varepsilon}$ in $\mathcal{D}(\mathbb{G})$ as $x \to z$. Moreover, let $(e_1, \dots, e_m)$ be a basis of H_1, as in Section 2, then*

$$\frac{1}{r}(\Phi_{x,\varepsilon}(\cdot(-re_j)) - \Phi_{x,\varepsilon}) \to -X_j\Phi_{x,\varepsilon}(\cdot) \quad \text{in } \mathcal{D}(\mathbb{G}),$$

for every $j = 1, \dots, m$.

PROOF. Let $Y_1, \dots, Y_n$ be a basis of right invariant vector fields, with homogeneous degree d_j, $j = 1, \dots, n$, generated by the horizontal vector fields $Y_1, \dots, Y_m$ defined above.

By Theorem 1.33 in [35], there exists $\bar{C} > 0$ and $\beta > 0$ such that for all $x, y \in \mathbb{G}$

$$|\phi_\varepsilon(xz^{-1}zy) - \phi_\varepsilon(zy)| \leq \bar{C} \sum_{j=1}^{n} |xz^{-1}|^{d_j} \sup_{|w| \leq \beta |xz^{-1}|} |Y_j f(wy)|, \tag{80}$$

Hence $\Phi_{x,\varepsilon}$ converges to $\Phi_{z,\varepsilon}$ uniformly on every compact set. Moreover, observe that we can apply (80) to the functions $X^I\phi_\varepsilon(xz^{-1}zy) - X^I\phi_\varepsilon(zy)$, where $X^I = X_{i_1} \dots X_{i_k}$ and $I \subset \{1, \dots, m\}^k$, $k > 0$. This implies the uniform convergence on compact set of $X^I\phi_\varepsilon(xy)$ to $X^I\phi_\varepsilon(zy)$, for every I and as a consequence the convergence in $\mathcal{D}(\mathbb{G})$. Now we prove the second part of the lemma. Let $f(y,r)$ be the smooth function $f(y,r) = (\Phi_{x,\varepsilon}(y(-re_j)) - \Phi_{x,\varepsilon})/r$. Then for any compact set $K \subset \mathbb{G}$, by Proposition 4.2.1 with $k = 1$, there exists C_K, such that

$$\sup_{y \in K} |f(r,y) + X_j\Phi_{x,\varepsilon}(y)| \leq C_K r, \quad \text{for } r < \frac{\delta}{2b^{k+1}}. \tag{81}$$

Notice that $C_K = C(K, X^2 f)$, more precisely

$$C_K = C(1) \sup_{z \in \Omega_{-\frac{\delta}{2}},\, h,l=1,\dots,m} |X_h X_l f(r,z)|,$$

where $C(1)$ is the constant that appears in Proposition 4.2.1 for $k = 1$. Observe that

$$Y_i f(r,y) = \frac{(Y_i\Phi_{x,\varepsilon})(y(-re_j)) - (Y_i\Phi_{x,\varepsilon})(y)}{r}.$$

Fix Y_i and apply estimates (81) to $Y_i f$, then we get

$$\sup_{y \in K} |Y_i f(r,y) + X_j Y_i \Phi_{x,\varepsilon}(y)| \leq C(K, X^2 Y_i f) r, \quad \text{for } r < \frac{\delta}{2b^{k+1}}$$

Given two vector fields X, Y respectively left invariant and right invariant we have $[X, Y] = 0$. Recall that for every smooth function u on $\mathbb{G}$ the following elementary formula holds

$$\frac{d}{dt} Y(u\exp(-tX))(\exp(tX)x) = [X, Y](u\exp(-tX))(\exp(tX)x). \tag{82}$$

We also recall that

$$\exp(tX)(x_0) = x \cdot \operatorname{Exp}(tX) \qquad \exp(tY)(x_0) = \operatorname{Exp}(tY) \cdot x.$$

Hence

$$\begin{aligned} Y(u\exp(-tX))(\exp(tX)x) &= Y(u\exp(-tX))(x \cdot \operatorname{Exp}(tX)) \\ &= Y(u\exp(-tX))(R_{\operatorname{Exp}(tX)}(x)) \\ &= Y(u \circ \exp(-tX) \circ R_{\operatorname{Exp}(tX)})(x) \end{aligned}$$

Let us consider the function $u \circ \exp(-tX) \circ R_{\operatorname{Exp}(tX)}(\cdot)$. Then we obtain

$$\begin{aligned} u \circ \exp(-tX) \circ R_{\operatorname{Exp}(tX)}(y) &= u(\exp(-tX)(y \cdot \operatorname{Exp}(tX)) \\ &= u(y \cdot \operatorname{Exp}(tX) \cdot \operatorname{Exp}(-tX)) \\ &= u(y). \end{aligned}$$

Thus the right left hand side of (82) does not depend on t, as a consequence $[X, Y] = 0$ and the two vector fields commute. Then we get

$$\sup_{y \in K} |Y_i f(r,y) + Y_i X_j \Phi_{x,\varepsilon}(y)| \leq C(K, X^2 Y_i f) r.$$

Then by the same observation, for every $I \subset 1, \dots, m^k$, $k > 0$, we get

$$\sup_{y \in K} |Y^I f(r,y) + Y^I X_j \Phi_{x,\varepsilon}(y)| \leq C(K, X^2 Y^I f) r, , \quad \text{for } r < \frac{\delta}{2b^{k+1}}$$

Hence we have the convergence of $f(r,y)$ to $-X_j \Phi_{x,\varepsilon}(y)$ in $\mathcal{D}(\mathbb{G})$, □

Lemma 4.2.3 *Let $\Omega \subset \mathbb{G}$ be an open bounded set and let $u \in \mathcal{D}'(\Omega)$. Let $\{\phi_\varepsilon\}$ be a family of mollifiers. We choose any $\psi \in C_c^\infty(\Omega)$ and define $\delta > 0$ as the distance between supp ψ and Ω^c. Then there exists a constant $C(\Omega) > 0$ such that for all $\varepsilon > 0$, $C(\Omega)\epsilon^{\frac{1}{\imath}} < \delta$, we have supp $\psi \subset \Omega_{-\varepsilon}$, the function $\Omega_{-\delta} \ni x \to \langle u, \Phi_{x,\varepsilon}\rangle$ is continuous and*

$$\int_{\Omega_{-\delta}} \psi(x)\, \langle u, \Phi_{x,\varepsilon}\rangle\, dx = \left\langle u, \int_{\Omega_{-\delta}} \psi(x)\, \Phi_{x,\varepsilon}\, dx \right\rangle, \tag{83}$$

where $\int_{\Omega_{-\delta}} \psi(x)\Phi_{x,\varepsilon}\, dx$ denotes the function $y \longrightarrow \int_{\Omega_{-\delta}} \psi(x)\, \phi_\varepsilon(xy^{-1})\, dx$.

PROOF. Let consider $z \in \Omega_{-\varepsilon}$. It follows, by Lemma 4.2.2, that $\Phi_{z,\varepsilon} \to \Phi_{x,\varepsilon}$ in $\mathcal{D}(\Omega)$ as $z \to x$. Hence $x \to \langle u, \Phi_{x,\varepsilon}\rangle$ is continuous on $\Omega_{-\varepsilon}$. Moreover we notice that the function $y \longrightarrow \int_{\Omega_{-\delta}} \psi(x)\, \phi_\varepsilon(xy^{-1})\, dx$ is compactly supported in Ω. In fact, for every $\omega \in \Omega^c$, every $y \in \Omega$ and every $x \in \Omega_{-\delta}$, we get

$$d(y,\omega) \geq d(x,\omega) - d(x,y) \geq \delta - d(y^{-1}xy^{-1}y).$$

Hence applying Lemma 4.1.5 with $\nu = C(\Omega) = \sup_{x \in \Omega} d(x)$, and $h = xy^{-1}$ we get

$$d(y,\omega) \geq \delta - C(\Omega)\epsilon^{\frac{1}{\imath}} > 0.$$

With respect to a fixed scalar product on $\mathbb{G}$, we consider the family $\mathcal{Q}_n$ of closed cubes in $\mathbb{G}$ of side e^{-n}, whose interior parts are disjoint and such that $\mathbb{G} = \bigcup_{Q \in \mathcal{Q}_n} Q$. We choose a fixed $x_Q \in Q$ for any $Q \in \mathcal{Q}_n$ and then obtain

$$\left| \int_{\Omega_{-\delta}} \psi(x)\langle u, \Phi_{x,\varepsilon}\rangle\, dx - \sum_{Q \in \mathcal{Q}_n} \psi(x_Q)\langle u, \Phi_{x_Q,\varepsilon}\rangle |Q \cap \Omega_{-\delta}| \right| \leq h\, \mathcal{L}^q\big((\text{supp}\,\psi)_1\big) \tag{84}$$

for all $n \geq n_h$. Notice that the uniform continuity of $x \to \psi(x)\langle u, \Phi_{x,\varepsilon}\rangle$ on supp ψ implies (84). This shows that

$$\int_{\Omega_{-\delta}} \psi(x)\langle u, \Phi_{x,\varepsilon}\rangle\, dx \quad = \quad \lim_{n\to\infty} \sum_{Q \in \mathcal{Q}_n} \psi(x_Q)\langle u, \Phi_{x_Q,\varepsilon}\rangle |Q \cap \Omega_{-\delta}|\,.$$

Since we have finite sums, by linearity of distributions we also have

$$\sum_{Q \in \mathcal{Q}_n} \psi(x_Q)\langle u, \Phi_{x_Q,\varepsilon}\rangle |Q \cap \Omega_{-\delta}| \quad = \quad \Big\langle u, \sum_{Q \in \mathcal{Q}_n} \psi(x_Q)\Phi_{x_Q,\varepsilon}\, |Q \cap \Omega_{-\delta}| \Big\rangle\,.$$

Arguing as before, one also observes the convergence

$$\sum_{Q\in\mathcal{Q}_n} \psi(x_Q)\Phi_{x_Q,\varepsilon}\,|Q\cap\Omega_{-\delta}| \longrightarrow \int_{\Omega_{-\delta}} \psi(x)\,\Phi_{x,\varepsilon}\,dx \quad \text{in} \quad \mathcal{D}(\Omega_{-\delta})$$

as $n\to\infty$. This leads us to the limit

$$\int_{\Omega_{-\delta}} \psi(x)\langle u,\Phi_{x,\varepsilon}\rangle\,dx \;=\; \lim_{n\to\infty}\langle u, \sum_{Q\in\mathcal{Q}_n} \psi(x_Q)\Phi_{x_Q,\varepsilon}\,|Q\cap\Omega_{-\delta}|\rangle$$

that concludes the proof. □

Many thanks are due to Giovanni Alberti for a stimulating conversation about the lemma above. As a consequence of Lemma 4.2.3 we get a more general distributional approximation theorem.

Theorem 4.2.4 *Let $\Omega\subset\mathbb{G}$ be an open bounded subset, let $T\in\mathcal{D}'(\Omega)$ be such that $D^2_H T\geq 0$ and let $\delta>0$ be such that $\Omega_{-\delta}$ is nonempty. It follows that there exists $C(\Omega)>0$ such that the convolutions $\langle T,\Phi_{x,\varepsilon}\rangle$ are smooth h-convex functions on $\Omega_{-\delta}$ for all $0<C(\Omega)\varepsilon^{1/\iota}<\delta$ and converge to T in $\mathcal{D}'(\Omega_{-\delta})$, where $\iota>1$ is the step of $\mathbb{G}$.*

PROOF. Let $\psi\in\mathcal{D}(\Omega)$ and let $\delta=\operatorname{dist}(\operatorname{supp}\psi,\Omega^c)>0$. Let $\{\phi_\varepsilon\}_{\varepsilon>0}$ be a family of mollifiers and set $\Phi_{x,\varepsilon}(y)=\phi_\varepsilon(xy^{-1})$. We consider the family of functions

$$T_\varepsilon(x)=\langle T,\Phi_{x,\varepsilon}\rangle \quad \text{on} \quad \Omega_{-\delta}$$

for all $0<C(\Omega)\varepsilon^{\frac{1}{s}}<\delta$, where $C(\Omega)>0$ is defined as in Lemma 4.2.3. It follows that T_ε are smooth functions. Le $e_1,\dots,e_m$ be a basis of H_1, as in Section 2. Then by definition of T_ε we get

$$\frac{1}{r}(T_\varepsilon(x(re_j))-T_\varepsilon(x))=\langle T,\frac{1}{r}(\Phi_{x(re_j),\varepsilon}-\Phi_{x,\varepsilon})\rangle.$$

for every $j=1,\dots,m$ and $r>0$. Since

$$\Phi_{x(re_j),\varepsilon}-\Phi_{x,\varepsilon}=\Phi_{x,\varepsilon}(\cdot(-re_j))-\Phi_{x,\varepsilon},$$

Lemma 4.2.2 implies that

$$X_jT_\varepsilon(x)=-\langle T,X_j\Phi_{x,\varepsilon}\rangle.$$

Iterating this procedure one easily obtains that

$$X^IT_\varepsilon(x)=(-1)^{|I|}\langle T,X^{-I}\Phi_{x,\varepsilon}\rangle,$$

for every $X^I=X_{i_1}\dots X_{i_k}$, $X^{-I}=X_{i_k}\dots X_{i_1}$, where $I\subset\{1,\dots,m\}^k$, $k>0$. Hence T_ε is a smooth function on $\Omega_{-\delta}$. In view of Lemma 4.2.3, we get

$$\int_\Omega \nabla^2_H\psi(x)\;T_\varepsilon(x)\,dx=\left\langle T,\int_\Omega \nabla^2_H\psi(x)\,\Phi_{x,\varepsilon}\,dx\right\rangle.$$

Notice that as in the previous lemma the function $y\to\int_\Omega \nabla^2_H\psi(x)\,\Phi_{x,\varepsilon}\,dx$ has compact support in Ω. The change of variable $x=zy$ gives

$$\int_{\Omega_{-\delta}} \nabla^2_H\psi(x)\,\Phi_{x,\varepsilon}(y)\,dx=\int_{D_\varepsilon}(\nabla^2_H\psi)(zy)\;\phi_\varepsilon(z)\,dz\,.$$

We denote by $l_z : \mathbb{G} \longrightarrow \mathbb{G}$ the left translation $l_z(w) = zw$. Then the left invariance of the family of vector fields X_i, $i = 1, \dots, m$, yields

$$\int_{D_\varepsilon} (\nabla^2_H \psi)\big(l_z(y)\big)\ \phi_\varepsilon(z)\, dz = \nabla^2_H \bigg(\int_{D_\varepsilon} \phi_\varepsilon(z)\ (\psi \circ l_z)(y)\, dz \bigg).$$

Since $y \to \int_{D_\varepsilon} \psi(zy)\, \phi_\varepsilon(z)\, dz$ is smooth and compactly supported in Ω and taking into account the previous equalities and our hypothesis $D^2_H T \geq 0$, we have

$$\int_\Omega \nabla^2_H \psi(x)\ T_\varepsilon(x)\, dx = \langle T, \nabla^2_H \int_{D_\varepsilon} \phi_\varepsilon(z)\ (\psi \circ l_z)\ dz \rangle \geq 0\,.$$

It remains to prove that $T_\epsilon \to T$ in $\mathcal{D}'(\Omega_{-\delta})$. In view of Lemma 4.2.3, we have

$$\langle T_\epsilon, \psi \rangle = \langle T, \int_{\Omega_{-\delta}} \psi(x) \Phi_{x,\epsilon} dx \rangle,$$

for all $\psi \in \mathcal{D}(\Omega_{-\delta})$. Recall that the functions

$$y \to \int_{\Omega_{-\delta}} \psi(x) \phi_\epsilon(xy^{-1}) dx$$

are smooth and compactly supported in Ω. Moreover, if we denote by

$$\psi_\epsilon(y) = \int_{\Omega_{-\delta}} \psi(x) \phi_\epsilon(xy^{-1}) dx,$$

we have that $\psi_\epsilon \to \psi$ in $\mathcal{D}(\Omega)$, for $\epsilon \to 0$. This implies that $T_\epsilon \to T$ in $\mathcal{D}'(\Omega_{-\delta})$. □

A simple consequence of Theorem 4.2.4 is the following result.

Corollary 4.2.5 *For any $M > 0$ and any $\delta > 0$, there exists $C_M > 0$, only depending on M, Ω and $\mathbb{G}$ and B_M, such that the smooth h-convex function $x \to \langle T, \Phi_{x,\varepsilon} \rangle$ is defined in $(\Omega \cap B_M)_{-\delta}$ whenever $0 < C_M\, \varepsilon^{\frac{1}{\iota}} < \delta$ and converges to u in $\mathcal{D}'\big((\Omega \cap B_M)_{-\delta}\big)$.*

Remark 4.2.6 We wish to show that in general the estimate of $d(xy)$ with respect to $d(yx)$ depends on both $d(x)$ and $d(y)$. For every $x, y \in \mathbb{G}$, the product is given by $x \cdot y = x_1 + y_1 + x_2 + y_2 + \beta(x_1, y_1)$, where $x_i, y_i \in H_i$ for $i = 1, 2$ and $\beta : V_1 \times V_1 \to V_2$ is non-degenerate, bilinear and skew-symmetric. Let $c \geq 1$ such that $|\beta(x_1, y_1)| \leq c|x||y|$ for all $x, y \in \mathbb{G}$, then $d(x) = \max\{|x_1|, \sqrt{|x_2/c|}\}$, defines a homogeneous norm on $\mathbb{G}$. By contradiction, we assume that there exists $C > 0$ such that

$$d(xyx^{-1}) \leq C\, d(y)^{\frac{1}{2}}$$

and C does not depend on y. Thus, for every $\lambda > 0$, we get

$$\lambda d(\delta_{\frac{1}{\lambda}}(x) y \delta_{\frac{1}{\lambda}}(x^{-1})) = d(x \delta_\lambda(y) x^{-1}) \leq C \lambda^{\frac{1}{2}} d(y)^{\frac{1}{2}},$$

that yields a contradiction as $\lambda \to +\infty$. It follows that C must depend on $d(y)$. Again, by contradiction, we assume that C is independent from x. Here we use the explicit expression of both d and the group operation, then we get

$$\left| \frac{y_2 + 2\beta(x_1, y_1)}{c} \right|^{\frac{1}{2}} \leq d(xyx^{-1}) \leq C d(y)^{\frac{1}{2}}. \tag{85}$$

Since β is nondegenerate, we can select a sequence $(x_{1n})_n$ such that $\beta(x_{1n}, y_1) \to +\infty$, that conflicts with (85).

In Lemma 4.2.3 and Theorem 4.2.4 in order to define the functions $x \to < T, \Phi_{x,\varepsilon} >$ on $\Omega_{-\delta}$, we need the condition $C\epsilon^{\frac{1}{\iota}} < \delta$. The constant C appears in the estimate of $d(x,y)$ in terms of $d(y,x)$ for $y \in \Omega$ and $x \in \Omega_{-\delta}$. By the previous discussion C must depend on $\mathrm{diam}(\Omega)$, hence the hypothesis of boundedness of Ω can not be removed.

Now we state some definitions and properties introduced in [87]. The *h-convex closure* $\mathcal{C}(A)$ of a subset $A \subset \mathbb{G}$ is the smallest h-convex set containing A. A stratified group $\mathbb{G}$ is *finitely h-convex* if it contains a finite subset $F \subset \mathbb{G}$ whose h-convex closure $\mathcal{C}(F)$ has non-empty interior.

Theorem 4.2.7 ([87]) *Any stratified group of step two is finitely h-convex.*

Notice that Theorem 4.2.7 does not hold in general stratified groups. In fact Rickly, in [87], proved that the Engel group $\mathbb{E}$ is not finitely h-convex. This is the main obstacle to the extension of Dudley's characterization in general stratified groups.

Definition 4.2.8 Denote by Γ the set of integral curves $\gamma : \mathbb{R} \to \mathbb{G}$ of left invariant horizontal vector fields on $\mathbb{G}$. Given $S \subset \mathbb{G}$ we define

$$\begin{aligned} H(S) &= \{\gamma(t) \,|\gamma \in \Gamma,\ t \in [0,1],\ \gamma(0),\ \gamma(1) \in S\}, \\ H^0(S) &= S, \qquad H^{k+1}(S) := H(H^k(S)), \\ H^\infty(S) &= \textstyle\bigcup_{k=1}^\infty H^k(S). \end{aligned}$$

The following corollary of Theorem 4.2.7 allows us to work with a finite h-convex combination of F, i.e. $H^k(F)$ for a certain $k \in \mathbb{N}$, instead the whole of $\mathcal{C}(F)$. For the reader's convenience we give also the proof taken form [87].

Corollary 4.2.9 Let $F \subset \mathbb{G}$ be a finitely h-convex set. Then there exists $k_0 = k_0(\mathbb{G}, F)$ such that $H^{k_0}(F)$ has nonempty interior.

PROOF. Clearly $H^\infty(F)$ is h-convex, by construction. Given an h-convex set $K \supseteq F$, it follows that if $H^l(F) \subset K$ then $H^{l+1}(F) \subset K$. This implies that $H^\infty(F) \subset K$ and by definition of $\mathcal{C}(F)$, $H^\infty(F) = \mathcal{C}(F)$. Moreover a direct computation gives that $H(F)$ is compact if F is. The compactness property of H and the theorem of Baire imply that $H^k(F)$, has non-empty interior for a certain $k_0 = k_0(\mathbb{G}, F) \in \mathbb{N}$. □

In the sequel, open and closed balls will refer to a fixed homogeneous distance of the Heisenberg group $\mathbb{H}^n$.

Definition 4.2.10 We define the set $A_0 = \{q_1, \dots, q_\alpha\}$ in $\mathbb{H}^n$ and the positive integer k_0 given by Theorem 4.2.7 and Remark 4.2.9, such that $A := H^{k_0}(A_0)$ has non-empty interior.

Lemma 4.2.11 *Let $p \in \mathbb{H}^n \setminus H_1$, let $0 < t < 1$ and let $\epsilon > 0$ be such that $B_{p,\varepsilon} \cap H_1 = \emptyset$. Then there exist $q, q' \in B_{p,\epsilon}$, with $q \neq q'$, along with $h \in qH_1 \cap H_1$ and $h' \in q'H_1 \cap H_1$ such that $th + (1-t)h' = 0$.*

PROOF. Let $p = p_1 + p_2$, where $p_i \in H_i$ with $i = 1, 2$ and $p_2 \neq 0$. It is not restrictive to assume $p_1 \neq 0$, since one can possibly consider a different center and a smaller radius of a ball contained in $B_{p,\varepsilon}$. Since β is non-degenerate, there exists $\bar{p}_1 \in H_1$ such that

$$\beta(p_1, \bar{p}_1) + p_2 = 0. \tag{86}$$

We choose a sufficiently small $\varepsilon' > 0$ such that, defining $w = \varepsilon' \bar{p}_1$ and defining $w' \in H_1$ such that $tw + (1-t)w' = 0$, we have $\|p^{-1}(p+w)\|, \|p^{-1}(p+w')\| < \varepsilon$. We set $q = p + w$ and $q' = p + w'$ and observe that

$$\begin{aligned} q\left(\bar{p}_1 + \lambda(p_1 + w)\right) &= p_1 + \bar{p}_1 + w + \lambda(p_1 + w) \in U \subset H_1 \\ q'\left(\bar{p}_1 + \mu(p_1 + w')\right) &= p_1 + \bar{p}_1 + w' + \mu(p_1 + w') \in U \subset H_1 \end{aligned}$$

for any $\lambda, \mu \in \mathbb{R}$, with $U = \operatorname{span}\{p_1, \bar{p}_1\}$. Since $p_1 + w$ and $p_1 + w'$ are linearly independent and belong to U, there exist $\bar{\lambda}, \bar{\mu} \in \mathbb{R}$ such that

$$t\big(p_1 + \bar{p}_1 + w + \bar{\lambda}(p_1 + w)\big) + (1-t)\big(p_1 + \bar{p}_1 + w' + \bar{\mu}(p_1 + w')\big) = 0\,.$$

This concludes the proof. □

Proposition 4.2.12 (Construction of the "good" set) *Let $A_0 = \{q_1, \ldots, q_\alpha\} \subset \mathbb{H}^n$ and $k_0 \in \mathbb{N}$ be as in Definition 4.2.10. Let $B \subset H^{k_0}(A_0)$ be an open ball. Then we can construct*

$$A_1 = \{q_{11}, q_{12}, \ldots, q_{\alpha 1}, q_{\alpha 2}\},$$

such that $q_{i1} = q_{i2} = q_i$ whenever $q_i H_1 \cap B \neq \emptyset$. If $q_i H_1 \cap B = \emptyset$, then we have

$$q_{i1} \neq q_{i2}, \quad q_{ij} \in q_i H_1, \quad q_{ij} H_1 \cap B \neq \emptyset \quad \text{for} \quad j = 1, 2, \tag{87}$$

and there exists $t \in (0,1)$ such that $tq_{i1} + (1-t)q_{i2} = q_i$.

PROOF. We have to consider the case $q_i H_1 \cap B = \emptyset$, hence we apply Lemma 4.2.11 to a small open ball contained in $q_i^{-1}B$. Then this ball does not intersect H_1. Then there exist distinct elements $b_1, b_2 \in q_i^{-1}B$ and also $h_1, h_2 \in H_1$ such that

$$h_1 \in b_1 H_1 \cap H_1, \quad h_2 \in b_2 H_1 \cap H_1, \quad th_1 + (1-t)h_2 = 0 \quad \text{and} \quad 0 < t < 1. \tag{88}$$

This implies that $h_j H_1 \cap q_i^{-1} B \neq \emptyset$ for $j = 1, 2$. Since $b_1, b_2 \notin H_1$, it follows that $h_1, h_2 \neq 0$ and the equality of (88) implies that $h_1 \neq h_2$. This allows us to define two distinct elements $q_{i1} = q_i h_1$ and $q_{i2} = q_i h_2$ that satisfy (87). Finally, a direct computation gives $tq_{i1} + (1-t)q_{i2} = q_i$, concluding the proof. □

Remark 4.2.13 The proof of Proposition 4.2.12 also implies that $A_0 \subset H(A_1)$.

The next geometric lemma will be important in the proof of the closure theorem.

Lemma 4.2.14 *Let $y, \bar{q} \in \mathbb{H}^n$ such that $y \in \bar{q}H_1$ and consider a ball $B_{y,r}$ for some $r > 0$. If $y = \bar{q}\bar{h}^{-1}$, then for every $\sigma, \sigma_0, \bar{\varepsilon} > 0$ there exist $\varepsilon_0 \in (0, \bar{\varepsilon})$, an $\varepsilon > 0$ and a bounded open neighborhood $V \subset H_1 \cap B_{\bar{h},\sigma_0}$ of $\bar{h}$ such that*

$$P = \{(\bar{q}\delta_s \bar{h})(\delta_l h) : s \in [0, \varepsilon_0], l \in [0, \varepsilon], h \in V\}$$

has nonempty interior and in addition

$$P \subset \left\{\bar{q}(\delta_s \bar{h})(\delta_l h) \mid s \in [0, \varepsilon_0],\, l \in [0, \varepsilon],\, h \in H_1,\, \bar{q}(\delta_s \bar{h})h^{-1} \in B_{y,r}\right\} \cap B_{\bar{q},\sigma}.$$

PROOF. Since $y = \bar{q}\bar{h}^{-1} \in B_{y,r}$ there exists an open bounded set U in H_1 and $\delta > 0$ such that $\bar{h} \in U$ and for all $0 < s < \delta$ and $h \in U$, we have $\bar{q}(\delta_s\bar{h})h^{-1} \in B_{y,r}$. Let us fix a basis $(\bar{h}, h_1, \ldots, h_{2n-1})$ of H_1 and consider

$$G(s,t,\tau_1,\ldots,\tau_{2n-1}) = (\delta_s\bar{h})h = (s\bar{h})\left(t\bar{h} + \sum_{j=1}^{2n-1}\tau_j h_j\right).$$

The open bounded set U of H_1 gives an open bounded set Ω of $\{0\}\times\mathbb{R}^{2n}$ with respect to the basis $(\bar{h}, h_1, \ldots, h_{2n-1})$ of H_1 and the canonical basis $(\bar{e}_2, \ldots, \bar{e}_{2n+1})$ of $\{0\}\times\mathbb{R}^{2n}$, hence

$$G\big((0,\delta)\times\Omega\big) \subset \left\{(\delta_s\bar{h})h \mid s > 0,\ h \in H_1,\ \bar{q}(\delta_s\bar{h})h^{-1} \in B_{y,r}\right\}.$$

We now consider the open set $\Omega_0 = \{lz : z \in \Omega,\ l > 0\}$ and observe that

$$G\big((0,\delta)\times\Omega_0\big) \subset \left\{(\delta_s\bar{h})(\delta_l h) \mid s, l > 0,\ h \in H_1,\ \bar{q}(\delta_s\bar{h})h^{-1} \in B_{y,r}\right\}.$$

With respect to the canonical basis $(\bar{e}_1, \ldots, \bar{e}_{2n+1})$ of $\mathbb{R}^{2n+1}$, we consider the explicit form of G as follows

$$G\left(s\bar{e}_1 + t\bar{e}_2 + \sum_{j=1}^{2n-1}\tau_j\bar{e}_{j+2}\right) = (s+t)\bar{h} + \sum_{j=1}^{2n-1}\tau_j h_j + \sum_{j=1}^{2n-1} s\,\tau_j\,\beta(\bar{h},h_j)\,e_{2n+1}.$$

There exists $1 \le j_0 \le 2n-1$ such that $\beta(\bar{h}, h_{j_0}) \neq 0$, then for all $\varepsilon, \varepsilon' > 0$ and for all $j = 1, \ldots, 2n-1$, the vectors

$$\begin{cases} \partial_s G(\varepsilon\bar{e}_2 + \varepsilon\varepsilon'\bar{e}_{j_0+2}) &= \bar{h} + \varepsilon\varepsilon'\,\beta(\bar{h},h_{j_0})e_{2n+1} \\ \partial_t G(\varepsilon\bar{e}_2 + \varepsilon\varepsilon'\bar{e}_{j_0+2}) &= \bar{h} \\ \partial_{\tau_j} G(\varepsilon\bar{e}_2 + \varepsilon\varepsilon'\bar{e}_{j_0+2}) &= h_j \end{cases}$$

are linearly independent. Let us fix $\varepsilon' > 0$ small, such that $\bar{e}_2 + \varepsilon'\bar{e}_{j_0+2} \in \Omega$ and

$$G(\bar{e}_2 + \varepsilon'\bar{e}_{j_0+2}) = \bar{h} + \varepsilon' h_{j_0} \in H_1 \cap B_{\bar{h},\sigma_0/2}.$$

Then we choose an open neighborhood $\Omega_1 \subset \Omega$ of $\bar{e}_2 + \varepsilon'\bar{e}_{j_0+2}$ and $\varepsilon > 0$ such that

$$G(lu) \in B_{\sigma/2} \quad \text{for all} \quad (l,u) \in [0,\varepsilon]\times\Omega_1 \quad \text{and} \quad G(\Omega_1) \subset H_1 \cap B_{\bar{h},\sigma_0}.$$

Since $dG\big(\frac{\varepsilon}{2}(\bar{e}_2 + \varepsilon'\bar{e}_{j_0+2})\big)$ is invertible, we choose a possibly smaller neighborhood $\Omega_2 \subset \Omega_1$ of $\bar{e}_2 + \varepsilon'\bar{e}_{j_0+2}$, an open interval $I_{\varepsilon/2}$ with $\varepsilon/2 \in I_{\varepsilon/2} \subset (0,\varepsilon)$ and a number

$$0 < \varepsilon_0 < \min\{\delta, \bar{\varepsilon}\}$$

such that $G\big(\{s\bar{e}_1 + lu : s \in [0,\varepsilon_0],\ l \in [0,\varepsilon],\ u \in \Omega_1\}\big) \subset B_\sigma$ and $G(I_{\varepsilon_0/2}\times I_{\varepsilon/2}\Omega_2)$ has nonempty interior in $\mathbb{H}^n$, for a suitably small open interval $I_{\varepsilon_0/2} \subset (0,\varepsilon_0)$ containing $\varepsilon_0/2$, where $I_{\varepsilon/2}\Omega_2 = \{tu : (t,u) \in I_{\varepsilon/2}\times\Omega_2\}$. This concludes the proof. □

Theorem 4.2.15 *Let Ω be an open set of $\mathbb{H}^n$. If $u_n : \Omega \to \mathbb{R}$ is a sequence of h-convex functions converging to $T \in \mathcal{D}'(\Omega)$ in the distributional sense, then T is defined by an h-convex function.*

PROOF. In the sequel, possible subsequences of u_n will be denoted by the same symbol. We have first to prove that $\sup_n u_n(x) < +\infty$ for all $x \in \Omega$. By contradiction, suppose this is not the case for some $p \in \Omega$. Then up to subsequences, we can assume that $u_n(p) \to +\infty$. Since h-convexity is preserved under left translations and so is the distributional convergence, it is not restrictive to assume that $p = 0$. Let A_0 and $A = H^{k_0}(A)$ be as in Definition 4.2.10 and let A_1 be as in Proposition 4.2.12. In view of Remark 4.2.13, we have $A = H^{k_0}(A_0) \subset H^{k_0+1}(A_1)$. The fact that both $H^{k_0+1}(A_1)$ and $H^{k_0}(A)$ have nonempty interior part is a "spanning property" of both A_0 and A_1 that is clearly invariant under left translations and group dilations. Thus, we can allow the origin to be an element of A_0 and also assume that the diameter of A_1 is sufficiently small. Then so can be taken the diameter of $H^{k_0+2}(A_1)$, hence we have $R > 0$ such that $A_0 \cup A_1 \cup H^{k_0+2}(A_1) \subset B_R \subset B_{3R} \subset \Omega$. By h-convexity, any u_n restricted to A attains its maximum at a point of A_0. Then up to subsequences, we have $q_{i_0} \in A_0$ such that

$$\max_{x \in A} u_n(x) = u_n(q_{i_0})\,.$$

Then $u_n(q_{i_0}) \geq u_n(0) \longrightarrow +\infty$ as $n \to +\infty$. By Proposition 4.2.12, we have that either $q_{i_0}H_1$ contains an interior point of A or both $q_{i_01}H_1$ and $q_{i_02}H_1$ do, with $q_{i_01}, q_{i_02} \in A_1$. From Proposition 4.2.12, we also know that q_{i_01} and q_{i_02} lie in the same horizontal line passing through q_{i_0}. Then h-convexity implies that, up to extracting a subsequence, $u_n(q_{i_0j_0}) \geq u_n(q_{i_0})$ for all n, where $j_0 \in \{1,2\}$.

As a result, in both of the previous cases we have an interior point y of A and an element $\bar q \in A_0 \cup A_1$ such that $y \in \bar q H_1$ and $u_n(\bar q) \to +\infty$ as $n \to +\infty$. Then we have $\bar h \in H_1$ such that $y = \bar q \bar h^{-1}$ and also $\bar q \delta_t \bar h^{-1} \in \Omega$ for all $0 \leq t \leq 1$. In fact, in the event $\bar q \in A_0$, the condition $\bar q, \bar q \bar h^{-1} \in H^{k_0}(A_0)$, implies

$$\bar q \delta_t \bar h^{-1} \in H^{k_0+1}(A_0) \subset H^{k_0+2}(A_1) \subset B_R \subset B_{3R} \subset \Omega \quad \text{for all} \quad 0 \leq t \leq 1.$$

In the remaining case $\bar q \in A_1$, we have $\bar q, \bar q \bar h^{-1} \in H^{k_0+1}(A_1)$ gives

$$\bar q \delta_t \bar h^{-1} \in H^{k_0+2}(A_1) \subset B_R \subset B_{3R} \subset \Omega \quad \text{for all} \quad 0 \leq t \leq 1\,.$$

Hence h-convexity and the fact that $u_n(\bar q \bar h^{-1}) \leq u_n(\bar q)$ give

$$u_n(\bar q) \leq u_n(\bar q \delta_s \bar h) \quad \text{for all} \quad s \geq 0$$

such that $\bar q \delta_l \bar h \in \Omega$ for all $l \in [0, s]$. Now, we choose $r > 0$ such that $B_{y,r}$ is contained in the interior of A and fix $\sigma > 0$ such that $B_{\bar q,\sigma} \subset \Omega$. We set $\bar\varepsilon > 0$ sufficiently small, such that $\bar q \delta_s \bar h \in B_R$ for all $s \in [0, \bar\varepsilon]$ and $\sigma > 0$ such that $B_{\bar q,\sigma} \subset B_R$. In view of Lemma 4.2.14, we find $\varepsilon, \varepsilon_0 > 0$, with $0 < \varepsilon_0 < \bar\varepsilon$ and an open neighborhood $V \subset H_1$ of $\bar h$ such that

$$P = \{(\bar q \delta_s \bar h)(\delta_l h) : s \in [0, \varepsilon_0], l \in [0, \varepsilon], h \in V\} \subset B_{\bar q,\sigma}$$

has nonempty interior and $\bar q(\delta_s \bar h)h^{-1} \in B_{y,r} \subset B_R$ whenever $(\bar q \delta_s \bar h)(\delta_l h) \in P$, hence

$$\|h\| \leq \|q \delta_s \bar h h^{-1}\| + \|q \delta_s \bar h\| < 2R\,.$$

It follows that for all $0 \le t \le 1$, $0 \le s \le \varepsilon_0$ and $h \in V$, we have $\bar{q}\delta_s\bar{h}\delta_t h^{-1} \in B_{3R} \subset \Omega$. By h-convexity, since

$$u_n\big(\bar{q}(\delta_s\bar{h})h^{-1}\big) = u_n\big(\bar{q}\,(s\bar{h})(-h)\big) \le u_n(\bar{q}) \le u(\bar{q}\delta_s\bar{h}),$$

it follows that for all $0 \le l \le \varepsilon$, we have

$$u_n(\bar{q}) \le u_n(\bar{q}\delta_s\bar{h}) \le u_n\big(\bar{q}(\delta_s\bar{h})\delta_l h\big).$$

Thus, u_n uniformly converges to $+\infty$ on the interior part of P, that is nonempty. This contradicts the distributional convergence u_n.

We have now to prove that $\inf_n u_n(x) > -\infty$ for all $x \in \Omega$. Again, by contradiction we assume that there exists $p \in \Omega$ such that, up to subsequences, $u_n(p) \to -\infty$. As before, invariance with respect to left translations allows us to assume that $p = 0$. We consider again A and A_0 as in Definition 4.2.10, where up to translating and rescaling A_0, we can assume that 0 is an interior point of A and $B_r \subset A \subset \Omega$, for some $r > 0$. By Proposition 1.2.17, we can find $r_0 > 0$ sufficiently small such that

$$F([-2r_0, 2r_0]^N) = U_2 \subset B_r$$

is a compact neighborhood of 0 and F is introduced in Definition 1.2.20. Up to subsequences, we have $\bar{q} \in A_0$ such that $\max_{A_0} u_n = u_n(\bar{q})$ and the previous part of the proof gives $u_n(\bar{q}) \le M < +\infty$ for all n, hence $\sup_A u_n \le \max_{A_0} u \le M$. We have

$$u_n(F(a)) = u_n\Big(\big(\prod_{j=1}^{N-1} a_j h_{i_j}\big)\big((t_N\sigma_N r_0 + (1-t_N)0)h_{i_N}\big)\Big),$$

where $\sigma_j = \operatorname{sign}(a_j)$ if $a_j \ne 0$, $\sigma_j = 1$ otherwise and $a_j = t_j\sigma_j 2r_0$ for all $j = 1, \dots, N$. We consider all $a \in [-r_0, r_0]^N$, hence $0 \le t_j \le 1/2$ for all $j = 1, \dots, N$. The h-convexity of u_n yields

$$\begin{aligned}
u_n(F(a)) \;&\le\; t_N\, u_n\Big(\big(\prod_{j=1}^{N-1} a_j h_{i_j}\big)\big(\sigma_N r_0 h_{i_N}\big)\Big) + (1-t_N)\, u_n\Big(\prod_{j=1}^{N-1} a_j h_{i_j}\Big) \\
&\le\; M + (1-t_N)\, u_n\Big(\prod_{j=1}^{N-1} a_j h_{i_j}\Big) \\
&\le\; N\,M + \Big(\prod_{l=0}^{N-1} (1-t_{N-l})\Big)\, u_n(0)
\end{aligned}$$

where the last inequality follows by iteration of the first one. For n large enough $u_n(0) < 0$, hence the previous estimate gives

$$u_n(F(a)) \le N\,M + \frac{1}{2^N}\, u_n(0) \to -\infty \quad \text{as} \quad n \to +\infty.$$

Then u_n uniformly goes to $-\infty$ on the neighborhood $\delta_{1/2}U_2$ of the origin, giving a contradiction to distributional convergence. Thus, it follows that $\inf_n u_n(x) > -\infty$ for all $x \in \Omega$. To conclude the proof, it suffices to show that any $p \in \Omega$ has a neighborhood where, up to extracting a subsequence, u_n uniformly converges. As in the previous steps, it is not restrictive to assume that $p = 0$, then taking A and A_0 from Definition 4.2.10, assuming that 0 is an interior point of A and $B_r \subset A \subset \Omega$, for some $r > 0$. In particular, $\sup_{B_r} u_n \leq \max_{A_0} u_n = u_n(\bar{q}) \leq M$ for all n, up to subsequences. As before, let $r_0 > 0$ be sufficiently small such that

$$F([-r_0, r_0]^N) = U_1 \subset B_r.$$

Then arguing as in the proof of Theorem 3.17 of [71], one easily gets the lower bound for u_n on the neighborhood U_1. For the reader's sake, we add a few details. In fact,

$$\mu_{1n} = 2u_n(0) - M \leq 2u_n(0) - u_n(-a_1 h_{i_1}) \leq u(a_1 h_{i_1})$$

that in turn yields

$$\mu_{2n} = 2\mu_{1n} - M \leq 2u_n(a_1 h_{i_1}) - u_n\big((a_1 h_{i_1})(-a_2 h_{i_2})\big) \leq u_n\big((a_1 h_{i_1})(a_2 h_{i_2})\big)$$

and further iterating this argument, one ends up with

$$\mu_{Nn} = 2\mu_{(N-1)n} - M \leq u_n\big(F(a)\big) \quad \text{for all} \quad a \in [-r_0, r_0]^N .$$

where $\mu_{jn} = 2\mu_{(j-1)n} - M$ for any $j = 2, \ldots, N$. Since $\inf_n u_n(0) > -\infty$, we have $M' > 0$ such that $\inf_{U_1} u_n \geq -M'$ for all n, therefore $\sup_n \sup_{U_0} |u_n| \leq \max\{M, M'\}$. Finally, by standard arguments, the L^∞ estimates of Theorem 9.2 in [26], along with Ascoli-Arzelà compactness theorem yield a uniformly convergent subsequence on a possibly smaller compact neighborhood of the origin. □

Theorem 4.2.16 *Let Ω be an open set of $\mathbb{H}^n$. If $T \in \mathcal{D}'(\Omega)$ is h-convex, then T is defined by an h-convex function in Ω.*

PROOF. Let us fix $M, \delta > 0$ such that $(\Omega \cap B_M)_{-\delta}$ is non empty. Then by Corollary 4.2.5, there exists $C_M > 0$ such that for $C_M \epsilon^{\frac{1}{\iota}} < \delta$, the functions T_ε are defined in $(\Omega \cap B_M)_{-\delta}$ and converge to T in $\mathcal{D}'\left((\Omega \cap B_M)_{-\delta}\right)$. By Theorem 4.2.15 we deduce that T, restricted to $(\Omega \cap B_M)_{-\delta}$, is defined by an h-convex function. The arbitrary choice of M and δ concludes the proof. □

CHAPTER 5

CONVEXITY IN CARNOT-CARATHÉODORY SPACES

Introduction

Motivated by the study of comparison principles for Monge-Ampere type equations with respect to Hörmander vector fields, Bardi and Dragoni have introduced the notion of $\mathcal{X}$-convexity that corresponds to the one dimensional convexity along integral curves of horizontal vector fields, [8]. Another natural type of convexity in this setting analogous to the one discussed for Carnot groups is the the v-convexity. Let us fix a system $\mathcal{X}$ of vector fields, then an upper semicontinuous functions $u : \Omega \to \mathbb{R}$, defined on an open set $\Omega \subset \mathbb{R}^n$, is v-convex if

$$\nabla^2_{\mathcal{X}} u \geq 0 \quad \text{in the viscosity sense,} \tag{89}$$

where the entries of this *horizontal Hessian* in the case u is smooth are exactly the symmetrized second order derivatives $(X_i X_j u + X_j X_i u)/2$ for all $i, j = 1, \dots, m$. The main result of [8] is that *in the class of upper semicontinuous functions, v-convexity and $\mathcal{X}$-convexity do coincide*, where the vector fields of $\mathcal{X}$ are only assumed to be of class C^2. When these vector fields generate a Carnot group structure, the previous characterization can be found in a number of previous works, [6], [60], [71], [88], [102]. In Theorem 6.1 of [8], it is also proved that $\mathcal{X}$-convexity, local boundedness and upper semicontinuity imply local Lipschitz continuity with respect to d, where the Carnot-Carathéodory distance d given by $\mathcal{X}$, is only assumed to yield the Euclidean topology. This result also gives L^∞-estimates for the horizontal derivatives Xu in terms of the L^∞-norm of u, where $X \in \mathcal{X}$. In the case $\mathcal{X}$ generates a Carnot group, these estimates take a quantitative form, see the central results of [26] and [60, 64].
The following theorem establishes that the previous estimates can be suitably extended to Carnot-Carathéodory spaces generated by a set $\mathcal{X}$ of Hörmander vector fields, see also Theorem 5.6.1.

Theorem 5.0.17 *Let $\mathcal{X} = \{X_1, \dots, X_m\}$ be a set of Hörmander vector fields, let $\Omega \subset \mathbb{R}^n$ be open and let $K \subset \Omega$ be compact. Then there exist $C > 0$ and $R > 0$, depending on K, such that each $\mathcal{X}$-convex function $u : \Omega \to \mathbb{R}$, that is locally bounded from above, for every $x \in K$ satisfies the following estimates*

$$\sup_{B_{x,r}} |u| \leq C \fint_{B_{x,2r}} |u(w)|\, dw \tag{90}$$

$$|u(y) - u(z)| \leq C\, \frac{d(y,z)}{r} \fint_{B_{x,2r}} |u(w)|\, dw\,, \tag{91}$$

for every $0 < r < R$ *and every* $y, z \in B_{x,r}$.

Clearly, the constant $C > 0$ cannot be chosen independent of K as in the case of Carnot groups, since general Carnot-Carathéodory spaces need not have either a group operation or dilations and the *doubling dimension* may change from point to point. This should suggest that the estimates of Theorem 5.0.17 are somehow sharp.

Our approach to prove (90) and (91) differs from both the geometric approach of [26] and the PDEs approach of [60, 64]. In fact, we need both these aspects, according to the following scheme. We start from a $\mathcal{X}$-convex function $u : \Omega \to \mathbb{R}$ that is locally bounded from above. By a result of D. Morbidelli, [80], the Carnot-Carathéodory ball can be covered by suitable compositions of flows of horizontal vector fields in a quantitative way, depending on the radius of the ball. This essentially allows us to apply the approach of [71] that relies on the one dimensional convexity of u along these flows, hence obtaining explicit Lipschitz estimates. It follows that u belongs to the anisotropic Sobolev space $W^{1,2}_{\mathcal{X},loc}(\Omega)$, see Section 1.1 for more information. The crucial step is to show that for every $x \in \Omega$ the $\mathcal{X}$-convex function u is a *weak subsolutions* of a suitable "pointed sub-Laplacian"

$$\mathcal{L}_x = \sum_{j=1}^{m} Y_j^2 \,,$$

that is constructed around x, see Theorem 5.5.3. Since the Lebesgue measure is locally doubling with respect to metric balls and the Poincaré inequality holds, the classical Moser iteration technique holds for weak subsolutions to the sub-Laplacian equation, hence getting the classical inequality

$$\sup_{B_{y,\frac{r}{2}}} u \le \kappa_x \fint_{B_{y,r}} |u(z)| dz \tag{92}$$

for $0 < r < \sigma_x$ and $y \in B_{x,\delta_x}$, where the positive constants κ_x, σ_x and $\delta_x > 0$ depend on x, see Section 5.5 for more information and in particular Corollary 5.5.4. The lower estimate of u is reached using again the approximate exponential, hence obtaining the following pointwise estimate

$$2^{N_x}\, u(x) - (2^{N_x} - 1) \sup_{B_{x,\bar{N}\delta}} u \le \inf_{B_{x,b\delta}} u \,, \tag{93}$$

where N_x depends on x and it satisfies the uniform inequality $1 \le N_x \le \bar{N}$ on some compact set, see Lemma 5.3.2. This eventually leads us to the proof of (90). The estimate (91) is a straightforward consequence of Theorem 5.3.5 joined with Theorem 5.6.2. In sum, the geometric part of our method arises from a quantitative representation of the Carnot-Carathéodory ball by approximates exponentials and it leads us to the lower estimates. The PDEs part of our approach provides the upper estimates.
The results in this chapter have been obtained in a joint work with V. Magnani, [66].

1. Equivalent CC-distances

We begin this section with the definition of a distance introduced by Franchi and Lanconelli in [36]. Throughout the chapter we consider a family $\mathcal{X} = (X_1, \dots, X_m)$ of smooth vector fields on $\mathbb{R}^n$ satisfying the Hrmander condition (15).

Definition 5.1.1 Let $\Gamma^c_{x,y}(t)$ be the family of all absolutely continuous curves $\gamma : [0,t] \longrightarrow \mathbb{R}^n$ with $\gamma(0) = x$, $\gamma(t) = y$, such that for a.e. $s \in [0,t]$ we have

$$\dot\gamma(s) = \sum_{j=1}^m a_j(s) X_j(\gamma(s)) \quad \text{and} \quad (a_1, \dots, a_m) \in \{\pm e_1, \dots, \pm e_m\},$$

where the curve $(a_1, \dots, a_m)$ is piecewise constant on $[0,t]$ and $(e_1, \dots, e_m)$ is the canonical basis of $\mathbb{R}^m$. Thus, we define the distance

$$\rho(x,y) = \inf\{t > 0 : \text{ there exists } \gamma \in \Gamma^c_{x,y}(t)\}. \tag{94}$$

Notice that these distances are well defined, since the Hörmander condition implies the connectivity with respect to absolutely continuous curves that are piecewise equal to the flow $t \to \Phi^X_t(x)$ of some $X \in \mathcal{X}$ and $x \in \mathbb{R}^n$.

Remark 5.1.2 Let us consider $X \in \mathcal{X}$ and $t, \tau \in \mathbb{R}$, by definition of d (1.1.8) and ρ, we have

$$\max\{d(\Phi^X_t(x), \Phi^X_\tau(x)), \rho(\Phi^X_t(x), \Phi^X_\tau(x))\} \le |t - \tau|$$

for any $x \in \mathbb{R}^n$, whenever the flows are defined for times t and τ.

Remark 5.1.3 Let $\mathcal{X}$ be the family of smooth Hörmander vector fields $X_1, \dots, X_m$ introduced above. Then by a rescaling argument, one can easily check that there holds

$$d(x,y) = \inf\left\{\delta > 0 : \text{there exists } \gamma \in \Gamma^\delta_{x,y}\right\}, \tag{95}$$

where $\Gamma^\delta_{x,y}(\mathcal{X})$ is the family of absolutely continuous curves $\gamma : [0,1] \to \mathbb{R}^n$ such that $\gamma(0) = x$, $\gamma(1) = y$ and for a.e. $t \in [0,1]$ we have

$$\dot\gamma(t) = \sum_{j=1}^m a_j(t) X_j(\gamma(t)) \quad \text{and} \quad \max_{1 \le j \le m} |a_j(t)| < \delta,$$

where d is introduced in Definition 1.1.8.

Lemma 5.1.4 *Let d and d_1 two CC-distances associated to the families of smooth Hörmander vector fields $\mathcal{X} = \{X_1, \dots, X_m\}$ and $\mathcal{X}_1 = \{Y_1, \dots, Y_m\}$, respectively. Let $\{i_1, j_1, \dots, j_{m-1}\} = \{1, 2, \dots, m\}$ and assume that $Y_j = X_j$ for all $j \ne i_1$ and $Y_{i_1} = X_{i_1} + X_{j_1}$. Then we have $4^{-1} d \le d_1 \le 4d$.*

PROOF. We can use for d and d_1 the equivalent definition stated in Remark 5.1.3. Taking this into account, we fix a compact set $K \subset \mathbb{R}^n$ and choose any $x_1, x_2 \in K$, setting $d(x_1, x_2) = \delta/2$, for some $\delta > 0$. Then there exists an absolutely continuous curve $\gamma : [0,1] \to \mathbb{R}^n$ belonging to $\Gamma^\delta_{x,y}(\mathcal{X})$. Clearly, we observe that

$$\dot\gamma = a_{i_1} Y_{i_1}(\gamma) + (a_{j_1} - a_{i_1}) Y_{j_1}(\gamma) + \sum_{s=2}^{m-1} a_{j_s} Y_{j_s}(\gamma),$$

hence $\gamma \in \Gamma^{2\delta}_{x,y}(\mathcal{X}_1)$, then $d_1(x,y) \leq 2\delta = 4\,d(x,y)$. In analogous way we get $d(x_1,x_2) \leq 4\,d_1(x_1,x_2)$, concluding the proof. □

Remark 5.1.5 As a consequence of the Hörmander condition, for every bounded set $A \subset \mathbb{R}^n$ we have a positive integer r such that (15) is satisfied for $r' = r$ and all $x \in A$.

We say that a function $u \in W^{1,2}_{\mathcal{X}}(\Omega)$ is an *$\mathcal{L}$-weak subsolution* of

$$\mathcal{L}u = \sum_{i=1}^{m} X_i^2 u = 0, \tag{96}$$

if for every nonnegative $\eta \in W^{1,2}_{\mathcal{X},0}(\Omega)$, we have $\displaystyle\sum_{i=1}^{m} \int_\Omega X_i u X_i^* \eta dx \geq 0$.

Lemma 5.1.6 *Let Ω' be an open set compactly contained in Ω and let $X \in \mathcal{X}$. There exists $T > 0$ such that the map Φ^X is well defined on $\Omega' \times (-2T, 2T)$ and for every $t \in (-2T, 2T)$, the mapping $\Phi^X(\cdot, t) : \Omega' \to \mathbb{R}^n$ is bi-Lipschitz onto its image with inverse $\Phi^X(\cdot, -t)$. The Jacobian J_X of Φ^X satisfies*

$$J_X(x,t) = 1 + \tilde{J}_X(x,t) \quad \textit{and} \quad |\tilde{J}_X(x,t)| \leq C|t|$$

for all $x \in \Omega'$ and $|t| < 2T$, where $C > 0$ is independent of x and t.

The proof of this lemma can be achieved by ODEs methods [55], see also [39] for the general case of a Lipschitz vector field.

Theorem 5.1.7 *Every Lipschitz function on an open set $\Omega \subset \mathbb{R}^n$ belongs to $W^{1,\infty}_{\mathcal{X}}(\Omega)$.*

The proof of Theorem 5.1.7 can be found from either Proposition 2.9 of [39] or Theorem 1.3 of [43]. From either these papers or the arguments of Theorem 11.7 of [53], it is also not difficult to deduce the following proposition.

Proposition 5.1.8 *Let $u : \Omega \to \mathbb{R}$ be a Lipschitz function. Let X be a vector field of $\mathcal{X}$ and fix $x \in \Omega$. Let $\Phi^X_t(x)$ be the flow of X starting at x. Then the directional derivative $\frac{d}{dt} u(\Phi^X_t(x))|_{t=0}$ exists almost everywhere and it coincides with the distributional derivative Xu.*

2. Almost exponentials and CC-distances

In this section, we will introduce almost exponential mappings and their properties, following the notations of D. Morbidelli [80]. We define

$$\begin{aligned} X^{(1)} &= \{X_1, \ldots, X_m\}, \\ X^{(2)} &= \{X_{[i_1,i_2]},\ 1 \leq i_1 < i_2 \leq m\} \end{aligned}$$

and so on, in such a manner that elements of $X^{(k)}$ are the commutators of length k. We denote by $Y_1, \ldots, Y_q$ an enumeration of all the elements of $X^{(1)}, \ldots, X^{(r)}$, where r is an integer large enough to ensure that $Y_1, \ldots, Y_q$ span $\mathbb{R}^n$ at each point of a fixed bounded open set $\Omega \subset \mathbb{R}^n$, see Remark 5.1.5. We call r the *local spanning step* and q the *local spanning number* of $\mathcal{X}$, to underly that they depend on Ω. It may be worth to stress that the Lie algebra spanned by $\mathcal{X}$ at some point need not be nilpotent, although the local spanning step is finite.

If Y_i is an element of $X^{(j)}$, we say Y_i has formal degree $d_i := d(Y_i) = j$. Let $I = (i_1, \ldots, i_n) \in \{1, 2, \ldots, q\}^n$ be a multi-index and define from [83] the functions

$$\lambda_I(x) = \det [Y_{i_1}(x), \ldots, Y_{i_n}(x)] \quad \text{and} \quad \|h\|_I = \max_{j=1,\ldots,n} |h_j|^{\frac{1}{d(Y_{i_j})}}.$$

As a consequence of the choice of $(Y_1, \ldots, Y_q)$, we have that for every $x \in \Omega$ there exists $I \in \{1, 2, \ldots, q\}^n$ with $\lambda_I(x) \neq 0$. We denote by $d(I)$ the integer $d_{i_1} + \ldots + d_{i_n}$, where $d_{i_k} = d(Y_{i_k})$.

Definition 5.2.1 Let $X, S \in \mathcal{X}$ and consider the mappings Φ_t^X and Φ_t^S, that coincide with Φ_1^{tX} and Φ_1^{tS}, respectively. Thus, for t sufficiently small, we can define the *local exponentials* $\exp(tX) := \Phi_1^{tX}$ and $\exp(tS) := \Phi_1^{tS}$, along with the *local product*

$$\exp(tX)\exp(tS) = \Phi_1^{tX} \circ \Phi_1^{tS}.$$

Let $S_1, \ldots, S_l$ be vector fields belonging to the family $\mathcal{X}$. Therefore, for every $a \in \mathbb{R}$ sufficiently small, we can define

$$\begin{aligned} &C_1(a, S_1) = \exp(aS_1), \\ &C_2(a, S_1, S_2) = \exp(-aS_2)\exp(-aS_1)\exp(aS_2)\exp(aS_1), \\ &C_l(a, S_1, \ldots, S_l) = C_{l-1}(a; S_2, \ldots, S_l)^{-1}\exp(-aS_1)C_{l-1}(a; S_2, \ldots, S_l)\exp(aS_1). \end{aligned}$$

According to (14) of [80], for $\sigma \in \mathbb{R}$ sufficiently small we define

$$\mathrm{e}_{\mathrm{ap}}^{\sigma S_{[(1,\ldots,l)]}} = \begin{cases} C_l(\sigma^{\frac{1}{l}}, S_1, \ldots, S_l), & \sigma > 0, \\ C_l(|\sigma|^{\frac{1}{l}}, S_1, \ldots, S_l)^{-1}, & \sigma < 0. \end{cases} \tag{97}$$

Following (16) of [80], given a multi-index $I = (i_1, \ldots, i_n)$, $1 \leq i_j \leq q$ and $h \in \mathbb{R}^n$ small enough, we also set

$$E_I(x, h) = \mathrm{e}_{\mathrm{ap}}^{h_1 Y_{i_1}} \cdots \mathrm{e}_{\mathrm{ap}}^{h_n Y_{i_n}}(x). \tag{98}$$

The next theorem, that is contained in Theorem 3.1 of [80], shows that almost exponential maps give a good representation of the Carnot-Carathéodory balls.

Theorem 5.2.2 *If $\Omega \subset \mathbb{R}^n$ is an open bounded set with local spanning number q and $K \subset \Omega$ is a compact set, then there exist $\delta_0 > 0$ and positive numbers a and b, $b < a < 1$, so that, given any $I \in \{1, \ldots, q\}^n$ such that*

$$|\lambda_I(x)|\delta^{d(I)} \geq \frac{1}{2} \max_{J \in \{1,\ldots,q\}^n} |\lambda_J(x)|\delta^{d(J)}, \tag{99}$$

for $x \in K$ and $0 < \delta < \delta_0$, it follows that $B_{x,b\delta} \subset E_{I,x}(\{h \in \mathbb{R}^n : \|h\|_I < a\delta\}) \subset B_{x,\delta}$.

We recall Definition 1.1.7 form Chapter 1.

Definition 5.2.3 *We say that two distances ρ_1 and ρ_2 in $\mathbb{R}^n$ are* equivalent, *if for every compact set $K \subset \mathbb{R}^n$, there exist $c_K \geq 1$, depending on K, such that*

$$c_K^{-1}\rho_1(x, y) \leq \rho_2(x, y) \leq c_K \rho_1(x, y) \quad \textit{for all} \quad x, y \in K.$$

Remark 5.2.4 We have stated Theorem 5.2.2 using only metric balls with respect to the distance d. In fact, in [80] the same symbol denotes the same distance, with a different definition, see Remark 5.1.3. Up to a change of the constant $b > 0$ in Theorem 3.1 of [80], we can replace the distance denoted by "ρ" in [80] with d. In fact, these two distances are equivalent, due to Theorem 4 of [83], joined with our Remark 5.1.3.

We fix a multi-index $I = (i_1, \dots, i_n)$ and for each Y_{i_k} we have a multi index

$$J_{i_k} = (j_1^{i_k}, j_2^{i_k}, \dots, j_{d_{i_k}}^{i_k}) \quad \text{such that} \quad Y_{i_k} = X_{[J_{i_k}]},$$

where d_{i_k} is the formal degree of Y_{i_k}. We notice that $1 \le j_s^{i_k} \le m$ for all $1 \le s \le d_{i_k}$ and $d_{i_k} \le r$, where r is the local spanning step of $\mathcal{X}$. By definition of e_{ap} we get

$$\mathrm{e}_{\mathrm{ap}}^{hY_{i_k}} = \begin{cases} \prod_{s=1}^{N_{i_k}} \exp(\sigma_s h^{\frac{1}{d_{i_k}}} X_s^{i_k}) & h \ge 0, \\ \prod_{s=1}^{N_{i_k}} \exp(-\sigma_{N_{i_k}+1-s} |h|^{\frac{1}{d_{i_k}}} X_{N_{i_k}+1-s}^{i_k}) & h < 0. \end{cases} \tag{100}$$

where $\sigma_s \in \{-1, 1\}$, N_{i_k} is the length of $\mathrm{e}_{\mathrm{ap}}^{hY_{i_k}}$ and $X_1^{i_k}$, $X_2^{i_k}, \dots, X_{N_{i_k}}^{i_k}$ is a suitable possibly iterated choice among the vectors $X_{j_1^{i_k}}, X_{j_2^{i_k}}, \dots, X_{j_{d_{i_k}}^{i_k}}$. A simple calculation gives $N_{i_k} = 2^{d_{i_k}} - 2 + 2^{d_{i_k}-1}$. We define $N(I) = \sum_{k=1}^n 2N_{i_k}$ along with the mapping $G_{I,x} : \mathbb{R}^N \to \mathbb{R}^n$, that is

$$G_{I,x}(w) = \prod_{k=1}^{n} \left\{ \prod_{s=1}^{N_{i_k}} \exp(w_{k,s,2} X_{N_{i_k}+1-s}^{i_k}) \prod_{s=1}^{N_{i_k}} \exp(w_{k,s,1} X_s^{i_k}) \right\} (x). \tag{101}$$

In the definition of $G_{I,x}$, we use the product to indicate the composition of flows according to the order that starts from the right. The variable w denotes the vector

$$(w_{1,1,1}, w_{1,2,1}, \dots, w_{1,N_{i_1},1}, w_{1,1,2}, w_{1,2,2}, \dots, w_{1,N_{i_1},2}, \dots, w_{n,1,2}, \dots, w_{n,N_{i_n},2})$$

belonging to $\mathbb{R}^{N(I)}$. The integer $N(I)$ is locally uniformly bounded from above, since every multi-index $I = (i_1, \dots, i_n)$ of Theorem 5.2.2 depends on x and satisfies $N_{i_k} \le 2^r - 2 + 2^{r-1}$, where r is the local spanning step of $\mathcal{X}$, depending on the fixed bounded open set Ω. Therefore we have a local upper bound $\bar{N}$ defined as follows

$$\bar{N} = 2n(2^{r+1} - 2 + 2^{r-1}) \tag{102}$$

and clearly $N(I) \le \bar{N}$, where $\bar{N}$ is independent of I.

Definition 5.2.5 For every $N \in \mathbb{N} \setminus \{0\}$, we set $\|w\|_N = \max_{k=1,\dots,N} |w_k|$, for every $w \in \mathbb{R}^N$. The corresponding open ball is defined as follows

$$S_{N,\delta} = \{w \in \mathbb{R}^N : \|w\|_N < \delta\}.$$

From standard theorems on ODEs, one can establish the following fact.

Proposition 5.2.6 *If $K \subset \Omega$ is a compact set and $N \in \mathbb{N}$ is positive, then there exists $\delta_1 > 0$ only depending on K, Ω and $\mathcal{X}$ such that for every $0 < \delta \le \delta_1$ and every $x \in K$ we have $B_{x,N\delta_1} \subset \Omega$ and for every integers $1 \le j_1, \ldots, j_N \le m$, the composition*

$$\Big(\exp(w_N X_{j_N}) \cdots \exp(w_2 X_{j_2}) \exp(w_1 X_{j_1})\Big)(x)$$

is well defined and contained in $B_{x,N\delta}$ for all $w \in S_{N,\delta}$.

The previous proposition immediately leads us to the following consequence.

Corollary 5.2.7 *Let Ω be an open bounded set with local spanning number q and local spanning step r. If $K \subset \Omega$ is a compact set, then there exist $\delta_1 > 0$ such that for every $x \in K$, every $0 < \delta \le \delta_1$ and every multi-index $I \in \{1, 2, \ldots, q\}^n$, the mapping $G_{I,x}$ introduced in* (101) *is well defined on $S_{N(I),\delta}$ and*

$$G_{I,x}(S_{N(I),\delta}) \subset B_{x,\bar{N}\delta} \subset D_{x,\bar{N}\delta_1} \subset \Omega\,,$$

where $\bar{N}$ is defined in (102).

For any of the above multi-indexes $I = (i_1, \ldots, i_n)$, we introduce the function $F_{I,x} : \mathbb{R}^n \to \mathbb{R}^{N(I)}$ as follows

$$F_{I,x}(h_1, \ldots, h_n) = (\sigma_{1,1}\delta_1(h_1)h_1^{\frac{1}{d_{i_1}}}, \ldots, \sigma_{1,N_{i_1}}\delta_1(h_1)h_1^{\frac{1}{d_{i_1}}}, -\sigma_{1,N_{i_1}}\delta_2(h_1)|h_1|^{\frac{1}{d_{i_1}}},$$
$$\ldots, -\sigma_{1,1}\delta_2(h_1)|h_1|^{\frac{1}{d_{i_1}}}, \ldots, \sigma_{n,1}\delta_1(h_n)h_n^{\frac{1}{d_{i_n}}} \ldots, \sigma_{n,N_{i_n}}\delta_1(h_n)h_n^{\frac{1}{d_{i_n}}}, \ldots)$$

where $\sigma_{k,j} \in \{-1, 1\}$, $k = 1, \ldots, n$ and $j = 1, \ldots, N_{i_k}$. More precisely, we have

$$F_{I,x}(h) = \sum_{k=1}^{n} \Big\{ \sum_{s=1}^{N_{i_k}} \sigma_{k,s}\, \delta_1(h_k)\, h_k^{1/d_{i_k}}\, e_{k,s,1} - \sum_{s=1}^{N_{i_k}} \sigma_{k,N_{i_k}+1-s}\, \delta_2(h_k)\, |h_k|^{1/d_{i_k}}\, e_{k,s,2} \Big\}\,, \tag{103}$$

where we have introduced the canonical basis

$$\big\{e_{k,s,i} : 1 \le k \le n,\ 1 \le s \le N_{i_k},\ i = 1, 2\big\}$$

of $\mathbb{R}^{N(I)}$ and the functions

$$\delta_1(x) = \begin{cases} 1 & x \ge 0 \\ 0 & x < 0 \end{cases} \quad \text{and} \quad \delta_2(x) = \begin{cases} 0 & x \ge 0 \\ 1 & x < 0 \end{cases}.$$

Remark 5.2.8 From the definitions of $G_{I,x}$ and $F_{I,x}$, it is straightforward to observe that $E_{I,x} = G_{I,x} \circ F_{I,x}$ on a sufficiently small neighborhood of the origin in $\mathbb{R}^n$.

Theorem 5.2.9 *If $\Omega \subset \mathbb{R}^n$ is an open bounded set with local spanning number q and $K \subset \Omega$ is compact, then there exist $\delta_0 > 0$ and positive numbers a and b, $b < a < 1$, so that for any $x \in K$ and $0 < \delta < \delta_0$ and any $I \in \{1, \ldots, q\}^n$ with*

$$|\lambda_I(x)|\delta^{d(I)} \ge \frac{1}{2} \max_{J \in \{1,\ldots,q\}^n} |\lambda_J(x)|\delta^{d(J)}\,, \tag{104}$$

we have $B_{x,b\delta} \subset G_{I,x}(\{w \in \mathbb{R}^{N(I)} : \|w\|_{N(I)} < a\delta\}) \subset B_{x,\bar{N}\delta} \subset D_{x,\bar{N}\delta_0} \subset \Omega$.

PROOF. From Theorem 5.2.2, we get the existence of $\delta_0, a, b > 0$, with $b < a < 1$ such that for every $x \in K$, $0 < \delta < \delta_0$ and $I \in \{1, \ldots, q\}^n$ satisfying (104), we have the inclusion

$$B_{x,b\delta} \subset E_{I,x}(\{h \in \mathbb{R}^n : \|h\|_I < a\delta\}).$$

This proves the validity of this inclusion, since for every $x \in K$ and $0 < \delta < \delta_0$ the existence of I satisfying (104) is trivial. From formula (103), we have

$$\|F_{I,x}(h)\|_{N(I)} = \|h\|_I \quad \text{for all} \quad h \in \mathbb{R}^n. \tag{105}$$

Remark 5.2.8 implies that $E_{I,x}(h) = G_{I,x} \circ F_{I,x}(h)$ for all $h \in \mathbb{R}^n$, possibly small, such that $G_{I,x}$, introduced in (101), is well defined on $F_{I,x}(h)$. In view of Corollary 5.2.7, it is not restrictive to choose $\delta_0 > 0$ possibly smaller, such that $G_{I,x}$ is well defined on

$$S_{N(I),\delta_0} \quad \text{and} \quad G_{I,x}(S_{N(I),\delta}) \subset B_{x,\bar{N}\delta} \subset D_{x,\bar{N}\delta_0} \subset \Omega. \tag{106}$$

Taking into account (105), we have $F_{I,x}(\{h \in \mathbb{R}^n : \|h\|_I < a\delta\}) \subset S_{N(I),\delta}$, that leads us to the following inclusions

$$B_{x,b\delta} \subset E_{I,x}(\{h \in \mathbb{R}^n : \|h\|_I < a\delta\}) \subset G_{I,x}\big(S_{N(I),\delta}\big) \subset B_{x,\bar{N}\delta} \tag{107}$$

concluding the proof. □

According to [83], for $x \in \mathbb{R}^n$, we set

$$\Lambda(x,\delta) = \sum_{I \in \{1,2,\ldots,q\}^n} |\lambda_I(x)|\, \delta^{d(I)}.$$

From Theorem 1 of [83], we get the following important fact.

Theorem 5.2.10 *For every $K \subset \mathbb{R}^n$ compact, there exist $\delta_0 > 0$ and positive constants C_1 and C_2, depending on K, so that for all $x \in K$ and every $0 < \delta < \delta_0$ we have*

$$C_1 \le \frac{|B_{x,\delta}|}{\Lambda(x,\delta)} \le C_2.$$

The point of this theorem is that it gives the doubling property of metric balls, as pointed out in [83]. In fact, Λ is a polynomial with respect to δ, that only depends on the enumeration of vector fields $Y_1, \ldots, Y_q$ on some fixed open bounded set Ω. Thus, we have the following corollary.

Corollary 5.2.11 *For every compact set $K \subset \mathbb{R}^n$ there exist positive constants C and r_0, depending on K, such that for every $x \in K$ and every $0 < r < r_0$, we have*

$$|B_{x,2r}| \le C\, |B_{x,r}|.$$

3. Boundedness from above implies Lipschitz continuity

The following notion of convexity was first introduced by Bardi and Dragoni in [8].

Definition 5.3.1 *Let Ω be an open set of $\mathbb{R}^n$ and let $\mathcal{X} = \{X_1, \ldots, X_m\}$ be a set of vector fields defined on $\mathbb{R}^n$. Then $u : \Omega \to \mathbb{R}$ is $\mathcal{X}$-convex, if the composition $u \circ \gamma$ is convex whenever $\gamma : I \to \Omega$ satisfies $\gamma' = \sum_{i=1}^m \alpha_i X_i \circ \gamma$ on the open interval I and α_i are arbitrary real numbers.*

Lemma 5.3.2 *Let $u : \Omega \to \mathbb{R}$ be a $\mathcal{X}$-convex function on an open set $\Omega \subset \mathbb{R}^n$ and let K be a compact set. Then there exist $\delta_0 > 0$, $0 < b < 1$ and an integer $\bar{N}$ only depending on K and $\mathcal{X}$ such that for every $x \in K$, there exists an integer $1 \le N_x \le \bar{N}$ such that for every $0 < \delta < \delta_0$ we have $D_{x,\bar{N}\delta_0} \subset \Omega$ and*

$$2^{N_x}\, u(x) - (2^{N_x} - 1) \sup_{B_{x,\bar{N}\delta}} u \le \inf_{B_{x,b\delta}} u\,. \tag{108}$$

PROOF. Let Ω' be an open bounded set containing K such that $\overline{\Omega'} \subset \Omega$, let r be the local spanning step and q be the local spanning number with local spanning frame $Y_1, \ldots, Y_q$ on Ω'. We apply Theorem 5.2.9 to both K and Ω', getting an integer $\bar{N}$ and positive number $\delta_0 > 0$, $0 < b < a < 1$, depending on K, Ω' and $\mathcal{X}$, having the properties stated in this theorem. Thus, we choose any $x \in K$ and $0 < \delta < \delta_0$, so that we can find a multi-index $I \in \{1, \ldots, q\}^n$ such that (104) holds. Theorem 5.2.9 implies that

$$B_{x,b\delta} \subset G_{I,x}(S_{N(I),a\delta}) \subset B_{x,\bar{N}\delta} \subset D_{x,\bar{N}\delta_0} \subset \Omega'$$

where $G_{I,x}$ is defined in (101). In particular, the closure $\overline{S}_{N(I),x}$ satisfies

$$G_{I,x}(\overline{S}_{N(I),a\delta}) \subset \Omega'.$$

Let us consider the scalar function $\varphi(w) = u \circ G_{I,x}(w)$, that is well defined for all $w \in \overline{S}_{N(I),a\delta}$. By definition of $\mathcal{X}$-convexity, we have

$$\mu_1 = 2\varphi(0) - \sup_{B_{x,\bar{N}\delta}} u \le 2\varphi(0) - \varphi(-w_1, 0, \ldots, 0) \le \varphi(w_1, 0, \ldots, 0)\,,$$

whenever $|w_1| \le a\delta$. Notice that $\mu_1 = 2\,u(x) - \sup_{B_{x,\bar{N}\delta}} u$. Of course, in the case $\sup_{B_{x,\bar{N}\delta}} u = +\infty$, then the inequalities (110) become trivial. For each $w_1 \in [-a\delta, a\delta]$, the function

$$[-a\delta, a\delta] \ni s \mapsto \varphi(w_1, s, 0, \ldots, 0),$$

is convex with respect to s, hence arguing as before we get

$$\mu_2 = 2\mu_1 - \sup_{B_{x,\bar{N}\delta}} u \le \varphi(w_1, s, 0, \ldots, 0)).$$

whenever $|s| \le a\delta$. We can repeat this argument up to $N(I)$ times, achieving

$$\mu_{N(I)} \le u \circ G_{I,x_0}(w) \quad \text{for every} \quad w \in \overline{S}_{N(I),a\delta}, \tag{109}$$

where $\mu_j = 2\mu_{j-1} - \sup_{B_{x,\bar{N}\delta}} u$ for $j = 1, \ldots, N(I)$. In particular, we have

$$\mu_{N(I)} = 2^{N(I)}u(x) - \Big(\sum_{j=0}^{N(I)-1} 2^j\Big)M = 2^{N(I)}u(x) - 2^{N(I)}M + M$$

with $M = \sup_{B_{x,\bar{N}\delta}} u$. In sum, we have proved that there exist $\delta_0 > 0$, $0 < b < 1$ and an integer $\bar{N}$ only depending on K and $\mathcal{X}$ such that for every $x \in K$, we can provide an integer $1 \le N_x \le \bar{N}$, depending on x, such that for every $0 < \delta < \delta_0$ we have $D_{x,\bar{N}\delta_0} \subset \Omega$ and (108) holds. □

Corollary 5.3.3 *Under the assumptions of Lemma 5.3.2, we have*

$$\inf_{B_{x,b\delta}} u \geq \begin{cases} 2\,u(x) - (2^{\bar N}-1)\sup_{B_{x,\bar N\delta}} u & \text{if} \quad u(x)\geq 0 \\ 2^{\bar N}\,u(x) - (2^{\bar N}-1)\sup_{B_{x,\bar N\delta}} u & \text{if} \quad u(x)<0 \text{ and } \sup_{B_{x,\bar N\delta}} u \geq 0 \\ 2^{\bar N}\,u(x) - \sup_{B_{x,\bar N\delta}} u & \text{if} \quad \sup_{B_{x,\bar N\delta}} u < 0 \end{cases} . \tag{110}$$

The previous corollary immediately leads us to another consequence.

Corollary 5.3.4 *Every $\mathcal{X}$-convex function that is locally bounded from above on an open set is also locally bounded from below.*

The proof of the next theorem follows the scheme of Lemma 3.1 in [71]. In the sequel, we will use the distance function $\operatorname{dist}_d(A,x) = \inf_{a\in A} d(a,x)$ for every $A \subset \mathbb{R}^n$.

Theorem 5.3.5 *Let $u : \Omega \to \mathbb{R}$ be a $\mathcal{X}$-convex function. If u is locally bounded from above, then it is locally Lipschitz continuous. More precisely, for every $\alpha_1, \alpha_2 > 0$ such that $\alpha_1 + \alpha_2 < \operatorname{dist}_d(K, w)$ and every $x, y \in K$, we have*

$$|u(x) - u(y)| \leq \frac{2\rho(x,y)}{\min\{\alpha_1,\alpha_2\}} \sup_{K_{\alpha_1+\alpha_2}} |u| \tag{111}$$

where $K_{\alpha_1+\alpha_2} = \{z \in \mathbb{R}^n : \operatorname{dist}_d(K,z) \leq \alpha_1+\alpha_2\} \subset \Omega$.

PROOF. First of all, from Corollary 5.3.4 it follows that u is locally bounded. Let us choose $0 < D < \operatorname{dist}_d(K, \Omega^c)$ and consider the compact set

$$K_D = \{z \in \mathbb{R}^n : \operatorname{dist}_d(K,z) \leq D\},$$

that is clearly contained in Ω. Choose any $\alpha > 0$ such that $D + \alpha < \operatorname{dist}_d(K, \Omega^c)$. Therefore for every $x \in K_D$ and $X \in \mathcal{X}$, we have

$$\operatorname{dist}_d\big(K_D, \Phi^X(x,t)\big) \leq d(\Phi^X(x,t), x) \leq |t| \leq \alpha$$

hence $\Phi^X(x,t) \in K_{D+\alpha} = \{z \in \mathbb{R}^n : \operatorname{dist}_d(K,z) \leq D+\alpha\} \subset \Omega$ for all $|t| \leq \alpha$. Hence Φ^X is defined on $K_D \times [-\alpha, \alpha]$ and it is contained in the larger compact set $K_{D+\alpha} \subset \Omega$. Let us fix $x, y \in K$ such that $\rho(x,y) < D$. Let $\varepsilon > 0$ be arbitrary chosen such that $\rho(x,y) + \varepsilon < D$. Thus, by definition of ρ, there exists $\rho(x,y) < \bar t < \rho(x,y) + \varepsilon$ and $\gamma \in \Gamma^c_{x,y}(\bar t)$ such that $t_0 = 0 < t_1 < \cdots < t_\nu = \bar t$ and

$$\gamma(t) = \Phi^{X_{j_k}}\big(\gamma(t_{k-1}), t - t_{k-1}\big) \tag{112}$$

for all $t \in [t_{k-1}, t_k]$ and $k = 1, \ldots, \nu$, where $1 \leq j_1, \ldots, j_\nu \leq m$. We have that

$$d(\gamma(t), x) \leq \rho(\gamma(t), x) \leq t \leq \bar t < D,$$

therefore the whole curve γ is contained in K_D and any restriction $\gamma|_{[t_{k-1},t_k]}$ can be smoothly extended on $[t_{k-1} - \alpha, t_k + \alpha]$ preserving the same form (112). Since u is locally bounded, we set

$$M = \sup_{w \in K_{D+\alpha}} |u(w)| < +\infty.$$

As a result, the $\mathcal{X}$-convexity of u implies that the difference quotient

$$\frac{|u(\gamma(t_k)) - u(\gamma(t_{k-1}))|}{|t_k - t_{k-1}|}$$

is not greater than the

$$\max\left\{|u(\Phi^{X_{j_k}}(\gamma(t_{k-1}), t_k + \alpha - t_{k-1})) - u(\gamma(t_k))|\,\alpha^{-1},\ |u(\Phi^{X_{j_k}}(\gamma(t_{k-1}), -\alpha)) - u(\gamma(t_{k-1}))|\alpha^{-1}\right\}.$$

This yields proves that

$$\frac{|u(\gamma(t_k)) - u(\gamma(t_{k-1}))|}{|t_k - t_{k-1}|} \le \frac{2M}{\alpha}.$$

It follows that

$$|u(y) - u(x)| \le \sum_{k=1}^{\nu} |u(\gamma(t_k)) - u(\gamma(t_{k-1})| \le \frac{2M}{T}\sum_{k=1}^{\nu}(t_k - t_{k-1}) < \frac{2M}{\alpha}(\rho(x,y) + \varepsilon),$$

with an arbitrary choice of $\varepsilon > 0$. In the case $\rho(x,y) \ge D$, we immediately have $|u(x) - u(y)| \le 2M\rho(x,y)/D$, that leads to the inequality

$$|u(x) - u(y)| \le \frac{2\rho(x,y)}{\min\{D,\alpha\}} \sup_{K_{D+\alpha}} |u|$$

for every $x, y \in K$, where $D, \alpha > 0$ satisfy $D + \alpha < \mathrm{dist}_d(K, \Omega^c)$. □

The following proposition has been pointed out to us by D. Morbidelli. It is essentially contained in his work, being a consequence of Theorem 3.1 of [80].

Proposition 5.3.6 *The distances d and ρ introduced in Definition 1.1.8 are equivalent.*

Remark 5.3.7 Notice that the inequality $d \le \rho$ is trivial. As a consequence of the previous proposition, $\mathcal{X}$-convex functions that are locally bounded are also locally Lipschitz continuous with respect to d and any other equivalent distance, according to the notion of equivalence given in Definition 1.1.7

4. Moser iteration technique

We begin this section stating the Sobolev embedding theorem for the anisotropic Sobolev spaces, proved by Capogna, Danielli and Garofalo in [21]. As a consequence of the following theorem we are able to use the Moser iteration technique on $\mathcal{L}$-weak subelliptic functions.

Theorem 5.4.1 (2.3 in [21]) *Let $\Omega_0 \subset \mathbb{R}^n$ be a bounded open set and denote by Q the homogeneous dimension relative to Ω_0. Let $1 < p < Q$. Then there exists $C > 0$ and $R_0 > 0$ such that for any $x \in \Omega_0$ and B_R, with $R < R_0$, we have*

$$\left(\frac{1}{|B_R|}\int_{B_R} |u|^{\kappa p} dx\right)^{\frac{1}{\kappa p}} \le CR\left(\frac{1}{|B_R|}\int_{B_R} |D_{\mathcal{X}}u|^p dx\right)^{\frac{1}{p}}, \tag{113}$$

for any $u \in W^{1,p}_{\mathcal{X},0}(B_R)$. Here $1 \le \kappa \le \frac{Q}{Q-p}$.

The following technical lemmas are well known, see for instance Lemma 8.15 in [45] and Theorem 7.8 in [46].

Lemma 5.4.2 *Let $\phi : [0, T] \to \mathbb{R}$ be a non-negative bounded function. Suppose that for $0 \le \rho < R \le T$ we have*

$$\phi(\rho) \le A(R-\rho)^{-\alpha} + \epsilon\phi(R)$$

for some $A, \alpha > 0$, $0 \le \epsilon < 1$. Then there exists a constant $c = c(\alpha, \epsilon)$ such that for $0 \le \rho < R \le T$ we have

$$\phi(\rho) \le cA(R-\rho)^{-\alpha}.$$

Lemma 5.4.3 *Let $f \in C^1(\mathbb{R})$ be such that $f' \in L^\infty(\mathbb{R})$, and consider $u \in W^{1,2}_{\mathcal{X}}(\Omega)$, $\Omega \subset \mathbb{R}^n$ bounded. Then $f \circ u \in W^{1,2}_{\mathcal{X}}(\Omega)$ and $X(f \circ u) = f'(u)Xu$, for every $X \in \mathcal{X}$.*

PROOF. By Theorem 1.1.10, there exists a sequence $u_m \in C^1(\Omega)$ such that u_m and Xu_m converge to u and Xu respectively in $L^2(\Omega)$. Then integrating in Ω we have

$$\int_\Omega |f(u_m) - f(u)|^2 dx \le \sup |f'|^2 \int_\Omega |u_m - u|^2 dx \to 0,$$

as m goes to infinity. Moreover, we get

$$\frac{1}{2}\int_\Omega |f'(u_m)Xu_m - f'(u)Xu|^2 dx \le \sup |f'|^2 \int_\Omega |Xu_m - Xu|^2 dx + \int_\Omega |f'(u_m) - f'(u)|^2 |Xu|^2 dx. \tag{114}$$

Possibly passing to a subsequence u_m converges a.e. to u in Ω. Moreover, since f' is continuous $f'(u_m)$ converges to $f'(u)$ a.e.. Hence the second term in the right hand side of (114) tends to zero by dominated convergence. As a consequence $f(u_m)$ and $f'(u_m)Xu_m$ tend to $f(u)$ and $f'(u)Xu$ respectively. Therefore $X(f \circ u) = f'(u)Xu$ and the proof is complete. □

The following lemma is essentially Lemma 1 of [81].

Lemma 5.4.4 ([81]) *Let $f : \mathbb{R} \to \mathbb{R}$ be a non-negative, convex, monotone increasing function. Let $\Omega \subset \mathbb{R}^n$ be an open bounded set and let $u \in W^{1,2}_{\mathcal{X}}(\Omega)$ be a weak subsolution of* (96). *If $f \circ u \in L^2_{loc}(\Omega)$, then $f \circ u$ is a weak sub-solution.*

PROOF. Assume that $f \in C^2(\mathbb{R})$ and for some $M > 0$ we have $f''(t) = 0$ if $|t| > M$. Let $\eta \in C_0^\infty(\Omega)$, $\eta \ge 0$ and define the non-negative function

$$\zeta(x) := f'(u(x))\eta(x).$$

Since $f' \in L^\infty$, from Lemma 5.4.3, it follows that $f \circ u \in W^{1,2}_{\mathcal{X}}(\Omega)$ and

$$X_i f \circ u = f'(u)X_i u, \quad i = 1, \dots, m.$$

The same conclusions hold for $f' \circ u$, hence $\zeta \in W^{1,2}_{\mathcal{X},0}(\Omega)$. Then we have

$$X_i u X_i^* \zeta = X_i(f \circ u)X_i^* \eta - f''(u)X_i u X_i u \eta \le X_i(f \circ u)X_i^*\eta.$$

Hence we get

$$0 \le \sum_{i=1}^m \int_\Omega X_i u X_i^* \zeta dx \le \sum_{i=1}^m \int_\Omega X_i(f \circ u)X_i^* \eta dx.$$

The general case follows by approximations. Let $v \in W^{1,2}_{\mathcal{X}}(\Omega)$ be a nonnegative subsolution of (96). For any $0 < \rho < R < 1$, take $\eta \in C_0^\infty(B_{x_0,R})$, with $\eta = 1$ on $B_{x_0,\rho}$ and

$|D_{\mathcal{X}}\eta| \le \frac{2}{R-\rho}$. The existence of such a function can be found in Lemma 3.2 [21], see also 7.12 [63]. Consider the admissible test function $v\eta^2 \ge 0$, then we get

$$\begin{aligned} 0 \ge \sum_{i=1}^{m} \int_{\Omega} X_i v(-X_i^*)(v\eta^2) &= \sum_{i=1}^{m} \int_{B_{x_0,R}} (X_i v)^2 \eta^2 dx + 2\sum_{i=1}^{m} \int_{B_{x_0,R}} X_i v X_i \eta v \eta dx \\ &\quad - \sum_{i=1}^{m} \int_{B_{x_0,R}} X_i u v \eta^2 \left(\sum_{j=1}^{n} \partial_j a_i^j(x) \right) dx. \end{aligned}$$

Since a_j^i are smooth functions there exists $C = C(\Omega)$ such that

$$\sum_{i=1}^{m} \int_{B_{x_0,R}} (X_i v)^2 \eta^2 dx \le 2\sum_{i=1}^{m} \int_{B_{x_0,R}} |X_i v| \, |X\eta| \, v\eta dx + C\sum_{i=1}^{m} \int_{B_{x_0,R}} |X_i v| v\eta^2 dx$$

Now using Young inequality $ab \le \frac{\varepsilon a^2}{2} + \frac{b^2}{2\varepsilon}$, we get

$$c_1 \sum_{i=1}^{m} \int_{B_{x_0,R}} (X_i v)^2 \eta^2 dx \le c_2 \sum_{i=1}^{m} \int_{B_{x_0,R}} v^2 \left(|X_i \eta|^2 + \eta^2 \right) dx.$$

Therefore we have a constant $c_3 > 0$ such that

$$\sum_{i=1}^{m} \int_{B_{x_0,R}} (X_i v)^2 \eta^2 dx \le c_3 \sum_{i=1}^{m} \int_{B_{x_0,R}} v^2 \left((X_i \eta)^2 + \eta^2 \right) dx.$$

Now note that $(X_i(\eta v))^2 \le 2\,|X_i v|^2\,\eta^2 + 2v^2\,|X_i\eta|^2$, then

$$\sum_{i=1}^{m} \int_{B_{x_0,R}} (X_i(\eta v))^2 dx \le c_4 \sum_{i=1}^{m} \int_{B_{x_0,R}} \left\{ (X_i\eta)^2 + \eta^2 \right\} v^2 dx. \tag{115}$$

Let $\phi \in C_c^\infty((0,1))$, $\phi \ge 0$, $\int_{(0,1)} \phi(x)dx = 1$, and consider $f_m = f^m \star \phi_{\frac{1}{m}}$. Where f^m is a convex functions, monotone increasing functions, such that $f^m = f$ on $|x| < m$ and $f^m(x) \le f(x)$ is a piecewise linear function on $|x| \ge m$. Clearly f_m is a smooth convex function, moreover $f^m(s - \varepsilon t) \le f^m(s)$, for every $t, \varepsilon > 0$ and $s \in \mathbb{R}$, by monotonicity of f^m. This yields that $f_m \le f^m \le f$. Notice that $f_m \to f$, $f_m'(u) \to f'(u)$, where $f'(u)$ exists. Then set $v_m = f_m(u)$ and $v = f(u)$, by the previous computations v_m is a subsolution of (96), moreover one has by (115)

$$\sum_{i=1}^{m} \int_{B_{x_0,R}} (X_i(\eta v_m))^2 dx \le c_4 \sum_{i=1}^{m} \int_{B_{x_0,R}} \left\{ (X_i\eta)^2 + \eta^2 \right\} v_m^2 dx.$$

Hence, by the estimates $v_m \le f$ we can write

$$\sum_{i=1}^{m} \int_{B_{x_0,R}} (X_i(\eta v_m))^2 dx \le \sum_{i=1}^{m} \int_{B_{x_0,R}} \left\{ (X_i\eta)^2 + \eta^2 \right\} v^2 dx,$$

on the other hand, by Fatou's theorem

$$\sum_{i=1}^{m} \int_{B_{x_0,R}} (X_i(\eta v))^2 dx \le \liminf_{m\to\infty} \sum_{i=1}^{m} \int_{B_{x_0,R}} (X_i(\eta v_m))^2 dx \le c_4 \int_{B_{x_0,R}} \left\{ (X_i\eta)^2 + \eta^2 \right\} v^2 dx,$$

hence $f \circ u \in W^{1,2}_{\mathcal{X}}(\Omega)$. Since the sequence v_m is bounded in $W^{1,2}_{\mathcal{X}}$, possibly passing to a subsequence we can suppose that for every $i = 1, \dots, m$, $X_i v_m$ weakly converges to $g_i \in L^2(\Omega)$. Moreover by Lebesque convergence theorem we have for every $\phi \in C_0^\infty(\Omega)$

$$\int_\Omega v_m X_i^\star \phi dx \to \int_\Omega v X_i^\star \phi dx = \int_\Omega X_i v \phi dx,$$

as a consequence $X_i v = g_i$. This implies that for every $\zeta \in W^{1,2}_{\mathcal{X},0}(\Omega)$,

$$0 \leq \sum_{i=1}^m \int_\Omega X_i(v_m) X_i^\star \zeta dx \to \sum_{i=1}^m \int_\Omega X_i(v) X_i^\star \zeta dx,$$

and the proof is complete. □

The proof of the following theorem is standard: it follows the celebrated Moser iteration technique for weak solutions to elliptic equations in divergence form [81], that applies to very general frameworks, including Carnot-Carathéodory spaces. There is a plenty of works in this area, so we limit ourselves to mention just a few of them, [37] [63], [54], [21]. A further discussion of this topic can be found for instance in [53]. We present a proof that is an adaptation of Proposition 8.19 in [45].

Theorem 5.4.5 *Let $\Omega \subset \mathbb{R}^n$ be an open bounded set, and let $\mathcal{X}$ be a family of smooth Hörmander vector fields and let $p > 0$. Thus, there exists $r_0 > 0$, depending on Ω and $\mathcal{X}$, and there exists $\kappa \geq 1$, depending on p, Ω and $\mathcal{X}$, such that whenever $u \in W^{1,2}_{\mathcal{X}}(\Omega)$ is a weak $\mathcal{L}$-subsolution to* (96), *we have*

$$\operatorname*{esssup}_{B_{x,\frac{r}{2}}} u \leq \kappa \left(\fint_{B_{x,r}} |u(y)|^p dy \right)^{\frac{1}{p}}, \tag{116}$$

for every $x \in \Omega$ such that $0 < r \leq \min\{r_0, \operatorname{dist}(\Omega^c, x)\}$.

PROOF. Without loss of generality we can consider only nonnegative sub-solutions, in fact if u is a sub-solution of (96) then $u^+ = \max\{0, u\}$ is a sub-solution. Moreover

$$\sup_{B_{x_0,\frac{R}{2}}} u \leq \sup_{B_{x_0,\frac{R}{2}}} u^+ \leq k_1 \left(\fint_{B_{x_0,R}} (u^+) dx \right) \leq k_1 \left(\fint_{B_{x_0,R}} |u(x)| dx \right).$$

For any $0 < \rho < R < 1$, take $\eta \in C_0^\infty(B_{x_0,R})$, with $\eta = 1$ on $B_{x_0,\rho}$ and $|D_{\mathcal{X}}\eta| \leq \frac{2}{R-\rho}$. Choose $\beta \geq 0$, and consider the test function $u_m^\beta u \eta^2 \geq 0$, where

$$u_m(x) = \min\left(u(x), m\right).$$

The function $u_m^\beta u\eta^2$ is in $W^{1,2}_{\mathcal{X},0}$ as one can easily prove. Using this function as a test function in (96) we get

$$\begin{aligned} 0 &\geq \sum_{i=1}^m \int_\Omega X_i u(-X_i^*)(u_m^\beta u\eta^2) \\ &= \beta\sum_{i=1}^m \int_{B_{x_0,R}} (X_i u_m)^2 u_m^\beta \eta^2 dx + \sum_{i=1}^m \int_{B_{x_0,R}} (X_i u)^2 u_m^\beta \eta^2 dx + 2\sum_{i=1}^m \int_{B_{x_0,R}} X_i u X_i \eta u_m^\beta u\eta dx \\ &\quad - \sum_{i=1}^m \int_{B_{x_0,R}} X_i u u_m^\beta u\eta^2 \left(\sum_{j=1}^n \partial_j a_i^j(x)\right) dx. \end{aligned}$$

Since a_j^i are smooth functions there exists $C = C(\Omega_0)$ such that

$$\begin{aligned} \beta\sum_{i=1}^m \int_{B_{x_0,R}} (X_i u_m)^2 u_m^\beta \eta^2 dx + \sum_{i=1}^m \int_{B_{x_0,R}} (X_i u)^2 u_m^\beta \eta^2 dx &\leq 2\sum_{i=1}^m \int_{B_{x_0,R}} |X_i u|\,|X\eta|\, u_m^\beta u\eta dx \\ &\quad + C\sum_{i=1}^m \int_{B_{x_0,R}} |X_i u| u_m^\beta u\eta^2 dx \end{aligned}$$

We should emphasize that later on we will begin the iteration with $\beta = 0$. Now using Young inequality $ab \leq \frac{\varepsilon a^2}{2} + \frac{b^2}{2\varepsilon}$, we get

$$\beta\sum_{i=1}^m \int_{B_{x_0,R}} (X_i u_m)^2 u_m^\beta \eta^2 dx + c_1\sum_{i=1}^m \int_{B_{x_0,R}} (X_i u)^2 u_m^\beta \eta^2 dx \leq c_2\sum_{i=1}^m \int_{B_{x_0,R}} u_m^\beta u^2 \left((X_i\eta)^2 + \eta^2\right) dx$$

Set $w = u_m^{\frac{\beta}{2}} u$, and note that

$$|X_i w|^2 \leq (2+\beta)\{\beta u_m^\beta (X_i u_m)^2 + u_m^\beta (X_i u)^2\}$$

Therefore we have

$$\sum_{i=1}^m \int_{B_{x_0,R}} (X_i w)^2 \eta^2 dx \leq c_3(1+\beta)\sum_{i=1}^m \int_{B_{x_0,R}} ((X_i\eta)^2 + \eta^2) w^2 dx.$$

Now note that $(X_i(\eta w))^2 \leq 2(X_i w)^2\eta^2 + 2w^2(X_i\eta)^2$, then

$$\sum_{i=1}^m \int_{B_{x_0,R}} (X_i(\eta w))^2 dx \leq c_4(1+\beta)\sum_{i=1}^m \int_{B_{x_0,R}} \left\{(X_i\eta)^2 + \eta^2\right\} w^2 dx,$$

where c_4 is independent of β. By Sobolev inequality (113) we have

(117)

$$\begin{aligned} \left(\int_{B_{x_0,\rho}} w^{2^*} dx\right)^{\frac{2}{2^*}} \leq \left(\int_{B_{x_0,R}} (w\eta)^{2^*} dx\right)^{\frac{2}{2^*}} &\leq CR|B_{x_0,R}|^{Q\left(\frac{Q-2}{Q}-1\right)+1} \int_{B_{x_0,R}} \sum_{i=1}^m \left\{X_i\left(\eta w\right)\right\}^2 dx \\ &\leq C_1 R^{Q\left(\frac{Q-2}{Q}-1\right)+2} \int_{B_{x_0,R}} \sum_{i=1}^m \left\{X_i\left(\eta w\right)\right\}^2 dx, \end{aligned}$$

since $Q\left(\frac{Q-2}{Q}-1\right)+2=-Q\frac{2}{Q}+2=0$ we get

$$\begin{aligned}\left(\int_{B_{x_0,\rho}} w^{2^*}dx\right)^{\frac{2}{2^*}} &\leq C\int_{B_{x_0,R}}\sum_{i=1}^{m}\left\{X_i\left(\eta w\right)\right\}^2dx\\ &\leq \frac{c_5(1+\beta)}{(R-\rho)^2}\int_{B_{x_0,R}} w^2dx.\end{aligned}$$

Now set $\lambda:=\frac{2^*}{2}$, and recalling the definition of w, we have

$$\left(\int_{B_{x_0,\rho}} u^{2\lambda}u_m^{\beta\lambda}dx\right)^{\frac{1}{\lambda}} \leq \frac{c_5(1+\beta)}{(R-\rho)^2}\int_{B_{x_0,R}} u^2u_m^{\beta}dx. \tag{118}$$

Hence by the definition of u_m, we get

$$\left(\int_{B_{x_0,\rho}} u_m^{(\beta+2)\lambda}dx\right)^{\frac{1}{\lambda}} \leq \frac{c_5(1+\beta)}{(R-\rho)^2}\int_{B_{x_0,R}} u^{\beta+2}dx,$$

provided the integral in the right-hand side is bounded. By letting $m\to+\infty$ we conclude that

$$\left(\int_{B_{x_0,\rho}} u^{(\beta+2)\lambda}dx\right)^{\frac{1}{\lambda}} \leq \frac{c_5(1+\beta)}{(R-\rho)^2}\int_{B_{x_0,R}} u^{\beta+2}dx, \tag{119}$$

Since $u^{(\beta+2)\lambda}$ is still a sub-solution, thanks to estimate (126) we can apply Lemma 5.4.4, and iterate (126). We start the iteration with $\beta=0$,

$$\sigma_i:=\beta_i+2=2\lambda^i,\qquad R_i:=\rho+\frac{R-\rho}{2^i},\qquad (R_i-R_{i+1})^2=\frac{(R-\rho)^2}{2^{2i+2}}.$$

Then by (119) we have

$$\begin{aligned}\left(\int_{B_{x_0,R_{i+1}}} u^{\sigma_{i+1}}dx\right)^{\frac{1}{\lambda^{i+1}}} &\leq \left(\frac{c_5(\sigma_i-1)}{2^{-2i-2}(R-\rho)^2}\right)^{\frac{1}{\lambda^i}}\left(\int_{B_{x_0,R_i}} u^{\sigma_i}dx\right)^{\frac{1}{\lambda^i}}\\ &\leq \prod_{k=0}^{i}\left(\frac{c_5(\sigma_i-1)}{2^{-2k-2}(R-\rho)^2}\right)^{\frac{1}{\lambda^k}}\int_{B_{x_0,R}} u^2dx.\end{aligned}$$

Since

$$\begin{aligned}\log\left(\prod_{k=0}^{i}\left(\frac{c_4(\sigma_i+1)}{2^{-2k-2}}\right)^{\frac{1}{\lambda^k}}\right) &= \sum_{k=0}^{i}\frac{1}{\lambda^k}\left\{(2k+2)\log(2)+\log(2c_5(2\lambda^k-1))\right\}<+\infty\\ \sum_{k=0}^{\infty}\frac{1}{\lambda^k} &= \frac{Q}{2}\end{aligned}$$

we have $\lim_{i\to\infty}\prod_{k=0}^{i}\left(\frac{c_4}{2^{-2k-2}(R-\rho)^2}\right)^{\frac{1}{\lambda^k}} \le c_5(R-\rho)^{-\frac{Q}{2}}$. Therefore

$$\left(\int_{B_{x_0,\rho}} u^{\sigma_i}dx\right)^{\frac{1}{\sigma_i}} \le \left(\int_{B_{x_0,R_i}} u^{\sigma_i}dx\right)^{\frac{1}{\sigma_i}} \le c_6(R-\rho)^{-\frac{Q}{2}}\left(\int_{B_{x_0,R}} u^2dx\right)^{\frac{1}{2}}.$$

Letting $i\to\infty$ the left hand side converges to $\sup_{B_{x_0,\rho}} u$, hence

$$\sup_{B_{x_0,\rho}} u \le c_6(R-\rho)^{-\frac{Q}{2}}\left(\int_{B_{x_0,R}} u^2dx\right)^{\frac{1}{2}}.$$

Set $\rho=\frac{R}{2}$ and let $p<2$, then we have

$$\sup_{B_{x_0,\frac{R}{2}}} u \le c_6(\frac{R}{2})^{-\frac{Q}{2}}\left(\sup_{B_{x_0,R}} u\right)^{1-\frac{p}{2}}\left(\int_{B_{x_0,R}} u^p dx\right)^{\frac{1}{2}}.$$

By Young's inequality $ab\le\frac{a^{p_1}}{p_1}+\frac{b^{p_2}}{p_2}$, valid whenever $a,b\ge 0$, $\frac{1}{p_1}+\frac{1}{p_2}=1$. We choose $\frac{1}{p_1}=1-\frac{p}{2}$, $p_2:=\frac{2}{p}$,

$$\sup_{B_{x_0,\frac{R}{2}}} u \le \frac{1}{2}\sup_{B_{x_0,R}} u + c_7(\frac{R}{2})^{-\frac{Q}{p}}\left(\int_{B_{x_0,R}} u^p dx\right)^{\frac{1}{p}}.$$

If we set $\phi(\rho)=\sup_{B_d(x_0,\rho)} u$ the conclusion follows by Lemma 5.4.2. □

5. $\mathcal{L}$-weak subsolutions and upper integral estimates

The point of this section is to show that locally bounded above $\mathcal{X}$-convex functions are $\mathcal{L}$-weak subsolutions of (96), where $\mathcal{X}=\{X_1,\dots,X_m\}$ is a family of Hörmander vector fields. This will enable us to apply Theorem 5.4.5.

In the proof of Theorem 5.5.3, we will use the following basic fact.

Lemma 5.5.1 *Let X be a vector field on $\mathbb{R}^n$, let $z\in\mathbb{R}^n$ be such that $X(z)\neq 0$ and let π be a hyperplane of $\mathbb{R}^n$ transversal to $X(z)$ and passing through z. There exists an open neighborhood A of z in π, $\tau>0$ and an open neighborhood U of z in $\mathbb{R}^n$ such that the restriction of the flow Φ^X to $A\times(-\tau,\tau)$ is a diffeomorphism onto U. Moreover, for every fixed system of coordinates $(\xi_1,\dots,\xi_{n-1})$ on π, denoting by ϕ the previous restriction with respect to these coordinates and by J_ϕ its Jacobian, we get*

$$\operatorname{div}X(x)=\frac{\partial_t J_\phi}{J_\phi}\circ\phi^{-1}(x)\quad \textit{for all}\quad x\in U. \tag{120}$$

Remark 5.5.2 From the definition of commutator and the fact that the family $\mathcal{X}$ satisfies the Hörmander condition, it is clear that for each $z\in\mathbb{R}^n$, there exists $X\in\mathcal{X}$ such that $X(z)\neq 0$.

Theorem 5.5.3 *Let $\Omega \subset \mathbb{R}^n$ be an open set, let $x_0 \in \Omega$ and let $u : \Omega \to \mathbb{R}$ be a $\mathcal{X}$-convex function that is locally bounded from above. There exist $\delta_0 > 0$ and a family of vector fields $\mathcal{X}_1 = \{Y_1, \ldots, Y_m\}$, with $Y_i = \sum_{j=1}^m a_{ij} X_j$, and $a_{ij} \in \{0,1\}$, both depending on x_0, such that $B_{x_0,\delta_0} \subset \Omega$ and u is a weak subsolution of the equation*

$$\sum_{i=1}^{m} Y_i^2 v = 0 \quad on \quad B_{x_0,\delta_0}\,. \tag{121}$$

PROOF. As observed in Remark 5.5.2, since $\mathcal{X}$ is a family of Hörmander vector fields, we must have some $j_1 \in \{1, 2, \ldots, m\}$ such that $X_{j_1}(x_0) \neq 0$. Thus, for each $i = 1, \ldots, m$, we define $Y_i = X_i$ if $X_i(x_0) \neq 0$ and $Y_i = X_i + X_{j_1}$ otherwise, so that all Y_i do not vanish on x_0. In view of Lemma 5.5.1, for each $i = 1, \ldots, m$ we can find an open bounded neighborhood U_i of x_0, that is compactly contained in Ω, an open bounded set $A_i \subset \mathbb{R}^{n-1}$, $\tau_i > 0$ and a diffeomorphism $\phi_i : S_i \longrightarrow U_i$, with $S_i = A_i \times (-\tau_i, \tau_i)$, ϕ_i is the restriction of the flow of Y_i and then it satisfies (120). We can find $\delta_0 > 0$ such that B_{x_0,δ_0} is compactly contained in U_i for all $i = 1, \ldots, m$. Let us choose any $\varphi \in C_c^\infty(B_{x_0,\delta_0})$ with $\varphi \geq 0$. Our claim follows if we prove that

$$\sum_{i=1}^{m} \int_{B_{x_0,\delta_0}} Y_i u(x)\; Y_i^* \varphi(x)\, dx \geq 0. \tag{122}$$

We will prove a stronger fact, namely, the validity of

$$\int_{B_{x_0,\delta_0}} Y_i u(x)\; Y_i^* \varphi(x)\, dx \geq 0 \quad \text{for all} \quad i = 1, \ldots, m\,.$$

By definition of $\mathcal{X}$-convexity, we have that $u(\phi_i(\omega, \cdot))$ is convex on the interval where it is defined for all $i = 1, \ldots, m$. By Theorem 5.3.5, u is locally Lipschitz continuous with respect to d. Iterating Lemma 5.1.4, no more than $m-1$ times, and observing that $\mathcal{X}_1 = \{Y_1, \ldots, Y_m\}$ is also a family of Hörmander vector fields, its associated distance d_1 is equivalent to d, that is obtained from $\mathcal{X}$. Theorem 5.1.7 and Proposition 5.1.8 imply that $u \in W^{1,\infty}_{\mathcal{X},loc}(\Omega)$ and the pointwise derivative

$$\partial_{Y_i} u(x) = \frac{d}{dt} u(\Phi^{Y_i}(x,t))|_{t=0}$$

exists for almost every $x \in \Omega$ and coincides with the distributional derivative $Y_i u$, up to a negligible set. In particular, there exists $L > 0$ such that $|Y_i u| \leq L$ almost everywhere in U_i, where $Y_i u$ is the distributional derivative of u along Y_i. Since ϕ_i sends negligible sets into negligible sets, we have that

$$\frac{\partial}{\partial s} u(\phi_i(\omega, s))|_{s=t} = \partial_{Y_i} u(\phi(\omega, t)) = Y_i u(\phi_i(\omega, t)) \tag{123}$$

for almost every $(\omega, t) \in S_i$. There exist $0 < t_i < \tau_i$ such that $\phi(A_i \times (-t_i, t_i)) = U_i'$ still contains B_{x_0,δ_0}, hence for $\varepsilon > 0$ sufficiently small, we can consider

$$(u \circ \phi_i)_\varepsilon(\omega, t) = \int_{-\tau_i}^{\tau_i} (u \circ \phi_i)(\omega, s))\, \nu_\varepsilon(t-s) ds,$$

for all $t \in (-t_i, t_i)$, where ν_ε are one dimensional mollifiers. Since $(u \circ \phi_i)(\omega, \cdot)$ is convex on $(-\tau_i, \tau_i)$ it is also locally Lipschitz, with distributional derivative. It follows that

$$\frac{\partial}{\partial t}(u \circ \phi_i)_\varepsilon(\omega, t) = (\partial_{Y_i} u \circ \phi_i)_\varepsilon(\omega, t)$$

for all $\omega \in A_i$ and $t \in (-t_i, t_i)$. Due to (123), applying Fubini's theorem it follows that for almost every $\omega \in A_i$ the pointwise derivative $\partial_{Y_i} u(\omega, t)$ equals the distributional derivative $Y_i u(\omega, t)$ for almost every $t \in (-\tau_i, \tau_i)$, that is precisely represented almost everywhere. As a consequence, we have

$$\frac{\partial}{\partial t}(u \circ \phi_i)_\varepsilon(\omega, t) = (\partial_{Y_i} u \circ \phi_i)_\varepsilon(\omega, t) = \big((Y_i u) \circ \phi_i\big)_\varepsilon(\omega, t) \tag{124}$$

for almost every $\omega \in A_i$ and every $t \in (-t_i, t_i)$. Since $(u \circ \phi)_\varepsilon(\omega, \cdot)$ is smooth and convex for all $\omega \in A_i$, we achieve

$$\int_{S_i'} \frac{\partial^2}{\partial t^2}(u \circ \phi_i)_\varepsilon(\omega, t)\ \varphi(\phi(\omega, t)) J_{\phi_i}(\omega, t)\, d\omega dt \geq 0$$

where $S_i' = A_i \times (-t_i, t_i)$.

Integrating by parts, it follows that the previous nonnegative integral equals the following one

$$-\int_{S_i'} \frac{\partial}{\partial t}(u \circ \phi_i)_\varepsilon(\omega, t) \frac{\partial}{\partial t}\left\{\varphi(\phi_i(\omega, t)) J_{\phi_i}\right\} d\omega dt$$

that can be written as follows

$$-\int_{S_i'} \left(\frac{\partial}{\partial t}(u \circ \phi_i)_\varepsilon(\omega, t) \frac{\partial}{\partial t}\varphi(\phi_i(\omega, t)) J_{\phi_i} + \frac{\partial}{\partial t}(u \circ \phi_i)_\varepsilon(\omega, t)\ (\varphi \circ \phi_i)(\omega, t)) \frac{\partial}{\partial t} J_{\phi_i} \right) d\omega dt$$

Clearly, we have $\frac{\partial}{\partial t}(\varphi \circ \phi_i)(\omega, t) = (Y_i \varphi)(\phi_i(\omega, t))$, hence by Lemma 5.5.1, we obtain

$$-\int_{S_i'} \frac{\partial}{\partial t}(u \circ \phi_i)_\varepsilon(\omega, t) \Big((Y_i \varphi)(\phi(\omega, t)) + (\mathrm{div} Y_i \circ \phi_i)(\omega, t)\ (\varphi \circ \phi_i)(\omega, t)) \Big) J_{\phi_i} d\omega dt \geq 0.$$

We can then pass to the limit as $\varepsilon \to 0^+$, taking into account that $Y_i u \in L^\infty(U_i)$ and that both (123) and (124) hold, getting

$$-\int_{\phi_i^{-1}(U_i')} (Y_i u) \circ \phi_i\ \left\{ (Y_i \varphi \circ \phi_i + (\mathrm{div} Y_i \circ \phi_i)\ \varphi \circ \phi_i \right\}\ J_{\phi_i}\ d\omega dt \geq 0.$$

By a change of variables towards the former coordinates, we have

$$-\int_{U_i'} Y_i u(x) \left\{ (Y_i \varphi)(x) + \mathrm{div} Y_i(x)\ \varphi(x) \right\} dx = \int_{B_{x_0, \delta_0}} Y_i u(x)\ Y_i^* \varphi(x)\, dx \geq 0\,,$$

that establishes our claim. □

As a consequence of both Theorem 5.4.5 and Theorem 5.5.3, we get the following consequence.

Corollary 5.5.4 *Let $\Omega \subset \mathbb{R}^n$ be open and let $p > 0$. If $x \in \Omega$, then there exist $\sigma_x, \delta_x > 0$ and $\kappa_x \geq 1$, depending on x, Ω, p and $\mathcal{X}$, such that $B_{x,\delta_x} \subset \Omega$, $\sigma_x \leq \delta_x/2$ and whenever $u : \Omega \longrightarrow \mathbb{R}$ is $\mathcal{X}$-convex and locally bounded from above, for all $y \in B_{x,\delta_x/2}$ and $0 < r \leq \sigma_x$, we have*

$$\sup_{B_{y,\frac{r}{2}}} u \leq \kappa_x \left(\fint_{B_{y,r}} |u(z)|^p dz \right)^{\frac{1}{p}} . \tag{125}$$

PROOF. Let $x \in \Omega$ and and consider the corresponding $\delta_x > 0$ given by Theorem 5.5.3, such that $B_{x,\delta_x} \subset \Omega$ and u is a weak subsolution of (121) where the vector fields Y_j depend on x. In view of Theorem 5.4.5 applied to the open bounded set B_{x,δ_x}, we get some constants $\kappa_x \geq 1$ and $r_x > 0$, depending on B_{x,δ_x}, p, and the vector fields Y_j, such that there holds

$$\operatorname*{ess\,sup}_{B_{y,\frac{r}{2}}} u \leq \kappa_x \left(\fint_{B_{y,r}} |u(z)|^p dz \right)^{\frac{1}{p}} , \tag{126}$$

for all $0 < r \leq \min\{r_x, \operatorname{dist}(B^c_{x,\delta_x}, y)\}$. Since for all $y \in B_{x,\delta_x/2}$, we have

$$\operatorname{dist}(B^c_{x,\delta_x}, y) \geq \delta_x/2,$$

setting $\sigma_x = \min\{r_x, \frac{\delta_x}{2}\}$, then (126) holds for all $0 < r \leq \sigma_x$ and all $y \in B_{x,\delta_x/2}$. □

Remark 5.5.5 Notice that we do not need to use the essential supremum in (125), since $\mathcal{X}$-convex functions that are locally bounded from above are locally Lipschitz continuous, due to Theorem 5.3.5.

As a consequence of Corollary 5.5.4, we can easily establish the following result.

Theorem 5.5.6 *Let $\Omega \subset \mathbb{R}^n$ be open, let $p > 0$ and let $K \subset \Omega$ be compact. Then there exists $\sigma > 0$ and $\kappa \geq 1$, depending on K, Ω, $\mathcal{X}$ and p, such that for every $\mathcal{X}$-convex function $u : \Omega \longrightarrow \mathbb{R}$ that is locally bounded from above and for every $x \in K$, we have $B_{x,\sigma} \subset \Omega$ and there holds*

$$\sup_{B_{x,\frac{r}{2}}} u \leq \kappa \left(\fint_{B_{x,r}} |u(z)|^p dy \right)^{\frac{1}{p}} \quad \textit{for all} \quad 0 < r \leq \sigma. \tag{127}$$

6. Integral estimates for $\mathcal{X}$-convex functions

Theorem 5.6.1 *Let $\Omega \subset \mathbb{R}^n$ be open, let $K \subset \Omega$ be compact and let $u : \Omega \longrightarrow \mathbb{R}$ be a $\mathcal{X}$-convex function that is locally bounded from above. Then there exists $C_0 > 0$, $b_0 > 0$ and $N_0 > 1$, depending on K, such that for every $x \in K$ there holds*

$$\sup_{B_{x,r}} |u| \leq C_0 \fint_{B_{x,N_0 r}} |u(z)|\, dz$$

whenever $0 < r < b_0$ and $K_0 = \{z \in \mathbb{R}^n : \operatorname{dist}(K, z) \leq N_0\, b_0\} \subset \Omega$.

PROOF. By Lemma 5.3.2, we have $\delta_0 > 0$, $0 < b < 1$ and a positive integer $\bar{N}$ such that for every $y \in K$, we have $D_{y,\bar{N}\delta_0} \subset \Omega$ and there exists with $1 \le N_y \le \bar{N}$ such that

$$2^{N_y} u(y) - (2^{N_y} - 1) \sup_{B_{y,\bar{N}\delta}} u \le \inf_{B_{y,b\delta}} u \tag{128}$$

for all $0 < \delta < \delta_0$. Let us consider $x \in K$ and any $0 < \delta' < b\delta_0/4$, observing that there exists $x' \in B_{x,\delta'}$ such that

$$u(x') \ge -\!\!\!\!\!\!\int_{B_{x,\delta'}} |u(z)|\, dz\,.$$

We clearly have $\inf_{B_{x,\delta'}} u \ge \inf_{B_{x',2\delta'}} u$, hence for some $1 \le N_{x'} \le \bar{N}$, we can apply the estimate (128) at x', getting

$$\inf_{B_{x,\delta'}} u \ge 2^{N_{x'}} u(x') - (2^{N_{x'}} - 1) \sup_{B_{x',\bar{N}\frac{2\delta'}{b}}} u.$$

From the previous inequalities, it follows that

$$\inf_{B_{x,\delta'}} u \ge -2^{\bar{N}} -\!\!\!\!\!\!\int_{B_{x,\delta'}} |u(z)|\, dz - (2^{N_{x'}} - 1) \sup_{B_{x,\bar{N}\frac{4\delta'}{b}}} u.$$

Theorem 5.5.6 provides $\sigma > 0$ and $\kappa \ge 1$ such that, up to choose $\delta_0 > 0$ possibly smaller, such that $\bar{N}\delta_0 < \sigma/2$, hence $\bar{N}\frac{8\delta'}{b} < \sigma$ and it follows that

$$\inf_{B_{x,\delta'}} u \ge -2^{\bar{N}} -\!\!\!\!\!\!\int_{B_{x,\delta'}} |u(z)|\, dz - (2^{\bar{N}} - 1)\, \kappa -\!\!\!\!\!\!\int_{B_{x,\bar{N}\frac{8\delta'}{b}}} |u(z)|\, dz.$$

As a consequence of Corollary 5.2.11, we have $Q_0 > 0$ and $r_0 > 0$ such that

$$|B_{x,\bar{N}\frac{8\delta'}{b}}| \le 2^{Q_0} \left(\bar{N}\frac{8}{b}\right)^{Q_0} |B_{x,\delta'}|,$$

up to making δ_0 further smaller, namely, satisfying $2\bar{N}\delta_0 < r_0$. It follows that

$$\inf_{B_{x,\delta'}} u \ge -2^{\bar{N}} \left[\kappa + (16)^{Q_0} \left(\frac{\bar{N}}{b}\right)^{Q_0}\right] -\!\!\!\!\!\!\int_{B_{x,\bar{N}\frac{8\delta'}{b}}} |u(z)|\, dz$$

and also

$$\sup_{B_{x,\delta'}} u \le \kappa\, 2^{Q_0} \left(\bar{N}\frac{4}{b}\right)^{Q_0} -\!\!\!\!\!\!\int_{B_{x,\bar{N}\frac{8\delta'}{b}}} |u(z)|\, dz\,,$$

that yield a constant $C_0 > 0$ depending on K, such that

$$\sup_{B_{x,r}} |u| \le C_0 -\!\!\!\!\!\!\int_{B_{x,N_0 r}} |u(z)|\, dz$$

for every $0 < r < b_0$ and every $x \in K$, with $b_0 = b\delta_0/4$ and $N_0 = \bar{N}\frac{8}{b} > 1$. By the previous requirements on δ_0, being $N_0 b_0 = 2\delta_0 \bar{N}$, we also have

$$K_0 = \{z \in \mathbb{R}^n : \operatorname{dist}(K, z) \le N_0 b_0\} \subset \Omega,$$

reaching the conclusion of the proof. □

Theorem 5.6.2 *Let $\Omega \subset \mathbb{R}^n$ be open, let $K \subset \Omega$ be compact and let $\lambda > 1$. Then there exist $\bar{C} > 0$ and $\bar{Q} > 0$, depending on K and there there exists $\bar{r} > 0$, depending on both K and λ, such that for every $x \in K$ and every $0 < r < \bar{r}$, each $\mathcal{X}$-convex function $u : \Omega \longrightarrow \mathbb{R}$, that is locally bounded from above satisfies the following estimate*

$$\sup_{B_{x,r}} |u| \leq \bar{C} \left(\frac{\lambda+1}{\lambda-1}\right)^{\bar{Q}} \fint_{B_{x,\lambda r}} |u(z)|\, dz\,. \tag{129}$$

PROOF. We fix any $\beta > 0$ such that $K_1 = \{z \in \mathbb{R}^n : \text{dist}(K, z) \leq \beta\} \subset \Omega$ and apply Theorem 5.6.1 to K_1, getting the corresponding positive constants C_1, b_1 and $N_1 > 1$. We have in particular

$$\{z \in \mathbb{R}^n : \text{dist}(K_1, z) \leq N_1\, b_1\} \subset \Omega.$$

Taking $0 < r < \beta/\lambda$, we have $B_{x,\lambda r} \subset K_1$ for all $x \in K$ and fixing $a = (\lambda - 1)/N_1$, it follows that for $0 < r < r_1$ and $r_1 = \min\{b_1/a, \beta/\lambda\}$, the following inequality

$$\sup_{B_{y,ar}} |u| \leq C_1 \fint_{B_{y,N_1 ar}} |u(z)|\, dz$$

holds for all $y \in K_1$. Now, let us fix $x \in K$. Thus, whenever $0 < r < r_1$ we can cover the compact set $D_{x,r}$ with a finite number of balls $B_{x_j,ar}$ centered at points of $D_{x,r}$, hence there exists $x_{j_0} \in D_{x,r}$ such that

$$\sup_{B_{x,r}} |u| \leq \sup_{B_{x_{j_0},ar}} |u|\,.$$

Since $x_{j_0} \in K_1$ and $ar < b_1$, Theorem 5.6.1 implies that

$$\sup_{B_{x_{j_0},ar}} |u| \leq C_1 \fint_{B_{x_{j_0},N_1 ar}} |u(z)|\, dz = C_1 \fint_{B_{x_{j_0},(\lambda-1)r}} |u(z)|\, dz\,.$$

As a result, we have proved that

$$\sup_{B_{x,r}} |u| \leq C_1 \frac{|B_{x,\lambda r}|}{|B_{x_{j_0},(\lambda-1)r}|} \fint_{B_{x,\lambda r}} |u(z)|\, dz \leq C_1 \frac{|B_{x_{j_0},(\lambda+1)r}|}{|B_{x_{j_0},(\lambda-1)r}|} \fint_{B_{x,\lambda r}} |u(z)|\, dz$$

for all $0 < r < r_1$, where r_1 also depends on λ. Finally, we apply Corollary 5.2.11 to K_0, getting $r_2 > 0$ and $\bar{Q} > 0$ such that for all $0 < r < \min\{r_1, r_2/\lambda + 1\}$ our claim (129) holds with $\bar{C} = C_1\, 2^{\bar{Q}}$. □

Bibliography

[1] G.ALBERTI, L.AMBROSIO, *A geometrical approach to monotone functions in $\mathbb{R}^n$*, Math. Z., **230**, 259-316, (1999).

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

[3] L. AMBROSIO, N. DANCER, *Calculus of variations and partial differential equations. Topics on geometrical evolution problems and degree theory*, Papers from the Summer School held in Pisa, September 1996. Edited by G. Buttazzo, A. Marino and M. K. V. Murthy, Springer-Verlag, Berlin, (2000).

[4] L. AMBROSIO, B. KLEINER AND E. LE DONNE, *Rectifiability of Sets of Finite Perimeter in Carnot Groups: Existence of a Tangent Hyperplane*, J. Geom. Anal., **19**, 509-540, (2009).

[5] L.AMBROSIO, V.MAGNANI, *Weak differentiability of BV functions on stratified groups*, Math. Z., **245**, 123-153, (2003).

[6] Z.M.BALOGH, M.RICKLY, *Regularity of convex functions on Heisenberg groups*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **2**, n.4, 847-868, (2003).

[7] I.YA.BAKEL'MAN, *Geometric methods of solution of elliptic equations*, (in Russian), Nauka, Moscow, (1965).

[8] M. BARDI, F. DRAGONI, *Convexity and semiconvexity along vector fields*, Calc. Var. Partial Differential Equations, **42**, 405-427, (2011).

[9] A. BELLA¨CHE, J. RIESLER (EDITORS) *Sub-Riemannian Geometry*, Progress in Mathematics, 144, Birkhäuser, Basel, (1996).

[10] G. BIANCHI, A. COLESANTI, C. PUCCI, *On the second order differentiability of convex surfaces*, Geom. Dedicata, **60**, 39-48, (1996).

[11] T. BIESKE, J. GONG, *The P-Laplace equation on a class of grushin-type space*, Proc. Am. Math. Soc., **134**, 12, 3585-3594, (2006).

[12] A. BONFIGLIOLI, E. LANCONELLI, *Subharmonic functions on Carnot groups*, Math. Ann. **325**, 97-122, (2003).

[13] A. BONFIGLIOLI, E. LANCONELLI, *Subharmonic functions in sub-Riemannian settings*, J. Eur. Math. Soc., in press.

[14] A. BONFIGLIOLI, E. LANCONELLI, F. UGUZZONI, *Stratified Lie groups and Potential Theory for their sub-Laplacians*, Springer-Verlag, (2007).

[15] A. BONFIGLIOLI, E. LANCONELLI, V. MAGNANI, M. SCIENZA, *H-convex distributions in stratified groups*, to appear Proc. Am. Math. Soc. (2012).

[16] J.M. BONY, *Principe du maximum, inegalité de Harnack et unicité du problème de Cauchy*

pour les opérateurs elliptiques dégénérés, Ann. Inst. Fourier, **19**, 1, 277-304, (1969).

[17] H. BUSEMANN, W. FELLER, *Krümmungseigenschaften konvexer Flächen*, Acta Math., **66**, 1-47, (1935).

[18] L.A.CAFFARELLI, X.CABRÉ, *Fully nonlinear elliptic equations*, AMS Colloquium Publications, **43**, AMS, Providence, RI, (1995).

[19] L. CAFFARELLI, L. NIRENBERG, J. SPRUCK, *The Dirichlet problem for nonlinear second order elliptic equations III, Functions of the eigenvalues of the Hessian*, Acta Math. **155**, 261-301, (1985).

[20] A. CALOGERO, R. PINI, *Horizontal Normal Map on the Heisenberg group*, J. Nonlinear Convex Anal., **12**, 287-307, (2011).

[21] L. CAPOGNA, D. DANIELLI, N. GAROFALO, *An embedding theorem and the Harnack inequality for nonlinear sub-elliptic equations*, Comm. Partial Diff. Equations,**18**, 1765-1794, (1993).

[22] L. CAPOGNA, D. DANIELLI, S.D. PAULS, J.T. TAYSON, *An introduction to the Heisenberg group and the sub-Riemannian isoperimetric problem*, Prog. in Math., **259**, Brikhäuser Verlag, Basel, (2007).

[23] G. CITTI, N. GAROLFALO, E. LANCONELLI, *Harnack's inequality for sum of squares of vector fields plus a potential*, Amer. J. Math., **115**, 3, 699-734, (1993).

[24] F. H. CLARKE, *Optimization and nonsmooth analysis*, SIAM, (1990).

[25] J.P.DEMAILLY, *Complex Analytic and Differential Geometry*, draft version available at `http://www-fourier.ujf-grenoble.fr/~demailly/books.html`.

[26] D.DANIELLI, N.GAROFALO, D.M. NHIEU, *Notions of convexity in Carnot groups*; Comm. Anal. Geom. **11**, n.2, 263-341, (2003).

[27] D.DANIELLI, N.GAROFALO, D.M. NHIEU, *On the best possible character of the L^Q norm in some a priori estimates for non-divergence form equations in Carnot groups*, Proc. Amer. Math. Soc. **131**, n.11, 3487-3498, (2003).

[28] D.DANIELLI, N.GAROFALO, D.M. NHIEU, F.TOURNIER *The theorem of Busemann-Feller-Alexandrov in Carnot groups*, Comm. Anal. Geom. **12**, n.4, 853-886, (2004)

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

[30] L. C. EVANS, R. F. GARIEPY, *Measure Theory and Fine Properties of Functions*, CRC Press, (1992).

[31] E.B. FABES, C.E.KENIG, R.SERAPIONI, *The local regularity of solutions of degenerate elliptic equations*, Comm. Partial Differential Equations, **7**, 1, 77-116, (1982).

[32] C. FEFFERMAN, D.H. PHONG, *Subelliptic eigenvalue problems*, Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II, 590-606, (1981).

[33] G.B. FOLLAND, *A fundamental solution for a subelliptic operator*, Bull. Amer. Math. Soc.,**79**, 373-376, (1973).

[34] G.B.FOLLAND, *Subelliptic estimates and function spaces on nilpotent Lie groups*, Ark. Mat., **13** , 161-207, (1975).

[35] G.B.FOLLAND, E.M. STEIN, *Hardy Spaces on Homogeneous groups*, Princeton University Press, (1982).

[36] B.FRANCHI, E.LANCONELLI, *Une métrique associée à une classe dopérateurs elliptiques dégénérés*, Rend. Sem. Mat. Univ. Politec. Torino, Special Issue, 105114, (1982).

[37] B.FRANCHI, E.LANCONELLI, *Hölder regularity theorem for a class of non uniformly elliptic operators with measurable coefficients*, Ann. Scuola Norm. Sup. Pisa, **10**, 523541, (1983).

[38] B.FRANCHI, R.SERAPIONI, F.SERRA CASSANO, *Meyer-Serrin type theorems and relaxation of variational integrals depending on vector fields*, Houston J. Math., **22**, 859-890, (1996).

[39] B.FRANCHI, R.SERAPIONI, F.SERRA CASSANO, *Approximation and Imbedding Theorems for Weighted Sobolev Spaces Associated with Lipschitz Continuous Vector Fields*, Bollettino U.M.I.,**7**, 11-B, 83117, (1997).

[40] K.O. FRIEDRICHS, *The identity of weak and strong extensions of differential operators*, Trans. Amer. Math. Soc., **55**, 132151, (1944).

[41] N. GAROFALO, E. LANCONELLI, *Existence and nonexistence results for semilienar equations on the Heisenberg group*, Indiana Univ. Math. J., **41**, 1, 71-98, (1992).

[42] N.GAROFALO, D.M.NHIEU, *Isoperimetric and Sobolev inequalities for Carnot-Carathéodory spaces and the existence of minimal surfaces*, Comm. Pure Appl. Math., **49**, 10811144, (1996).

[43] N. GAROFALO , D.M. NHIEU, *Lipschitz continuity, global smooth approximations and extension theorems for Sobolev functions in Carnot-Carathéodory spaces*, J. Anal. Math., **74**, 6797, (1998).

[44] N.GAROFALO, F.TOURNIER, *New properties of convex functions in the Heisenberg group*, Trans. Am. Math. Soc., **358**, n.5, 2011-2055, (2005).

[45] M. GIAQUINTA, L. MARTINAZZI, *An introduction to the regularity theory for elliptic systems, harmonic maps and minimal graphs*, Edizioni della Normale, Pisa, (2005).

[46] D. GILBARG, N. TRUDINGER, *Elliptic partial differential equations of second order* Grundlehren der Mathematischen Wissenschaften, Vol. 224. Springer-Verlag, Berlin-New York, (1977).

[47] W.L. CHOW, *Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung*, Math. Ann., **117**, 85-115, (1940).

[48] R. GOODMAN, *Nilpotent Lie groups: structure and applications to analysis*, Lecture Notes in Math., **562**, Springer-Verlag, (1976).

[49] C. E. GUTIÉRREZ, E. LANCONELLI, *Classical, viscosity and average solutions for PDE's with nonnegative characteristic form*, Rend. Mat. Acc. Lincei, **15**, 17-28, (2004).

[50] C.E. GUTIÉRREZ, A. MONTANARI, *Maximum and comparison principle for convex functions on the Heisenberg group*, Comm. Partial Differential Equations, **29** ,no. 9-10, 1305-1334, (2004).

[51] C.E. GUTIÉRREZ, A. MONTANARI, *On the second order derivatives of convex functions on the Heisenberg group*, Ann. Sc. Norm. Super. Pisa Cl. Sci. 5, 2, 349-366, (2004).

[52] M. HALL, *A basis for free Lie rings and higher commutators in free groups*, Proc. Amer. Math. Soc., **1**, 575-581, (1950).

[53] P.HAJLASZ, P.KOSKELA, *Sobolev met Poincaré*, Mem. Amer. Math. Soc., **145**, (2000).

[54] J. HEINONEN, T. KILPELÄINEN, O. MARTIO, *Nonlinear Potential Theory of Degenerate Elliptic Equations*, Oxford Universtiy Press, (1993).

[55] P. HARTMAN, *Ordinary differential equations*, Birkhauser, (1982).

[56] M. HERVÉ, R.M. HERVÉ, *Les fonctions surharmoniques dans l'axiomatique de M. Brelot*

associeé à un operateur elliptique dégénéré, Ann. Inst. Fourier, **22**, 2, 131-145, (1972).
[57] L. HÖRMANDER, *Hypoelliptic second order differential equations* Acta Math., **119**, 147-171, (1967).
[58] N. JACOBSON. *Lie Algebras*, Interscience tractats in pure and applied mathematics,**10**, (1962).
[59] D.S. JERISON, A. SÁNCEZ-CALLE, *Estimates for the heat kernel for a sum of squares of vector fields*, Indiana Univ. Math. J., **35**, 4, 835-854, (1986).
[60] P. JUUTINEN, G. LU, J.J. MANFREDI, B. STROFFOLINI, *Convex functions on Carnot groups*, Rev. Mat. Iberoam. **23**, no. 1, 191-200, (2007).
[61] J. KINNUNEN, O. MARTIO, *Nonlinear potential theory on metric spaces*, Illinois J. Math., **46**, 3, 857-883, (2002).
[62] A. KNAPP, *Lie groups beyond an introduction*, Progress in mathematics, **140** , Birchäuser, (2002).
[63] G. LU, *Weighted Poincaré and Sobolev inequalities for vector fields satisfying Hörmander's condition and applications*, Rev. Mat. Iber., **8**, 3, 376-349, (1992).
[64] G.LU, J.MANFREDI, B.STROFFOLINI, *Convex functions on the Heisenberg group*, Calc. Var. Partial Differential Equations **19**, n.1, 1-22, (2004).
[65] E. LANCONELLI, *Stime sub-ellittiche e metriche Riemanniane singolari*, Seminario di Analisi Matematica, Dip. di Matematica, Univ. di Bologna, A.A. 1982-83.
[66] V. MAGNANI, M. SCIENZA, *Regularity estimates for convex functions in Carnot-Carathéodory spaces* , preprint, (2012).
[67] V. MAGNANI, M. SCIENZA, *Characterizations of differentiability for h-convex functions in stratified groups*, To Appear, Ann. Scuola Norm. Sup. Pisa Cl. Sci, (2012).
[68] V. MAGNANI, *Contact equations, Lipschitz extensions and isoperimetric inequalities* , Calc. Var. Partial Differential Equations, **39**, 233-271, (2010)
[69] V. MAGNANI, *Area implies coarea*, Indiana Univ. Math. J., accepted, (2010).
[70] V. MAGNANI *Towards differential calculus in stratified groups*, to appear J. Aust. Math. Soc.
[71] V. MAGNANI, *Lipschitz continuity, Aleksandrov theorem and characterizations for H-convex functions*, Math. Ann. **334**, 199–233, (2006).
[72] V. MAGNANI, *Elements of Geometric Measure Theory on Sub-Riemannian Groups*, Scuola Normale Superiore Pisa, (2002).
[73] J.J. MANFREDI, *Notes for the Course: Nonlinear Subelliptic Equations on Carnot Groups*, III School on Analysis and Geometry in Metric Spaces, Trento, May 2003. Available at: http://www.pitt.edu/ manfredi/papers/fullynonlsubtrentofinal.pdf.
[74] G.A.MARGULIS, G.D.MOSTOW, *Some remarks on the definition of tangent cones in a Carnot-Carathéodory space*, J. Anal. Math. **80** , 299-317, (2000).
[75] G. MÉTIVIER, FONCTION SPECTRALE ET VALEURS PROPRES DUNE CLASSE DOPÉRATEURS NON ELLIPTIQUES, Commun. Partial Differ. Equations, **1**, 467-519, (1976).
[76] F. MIGNOT, *Contrôle optimal dans les inéquations variationelles elliptiques*, J. Funct. Anal. **22**, 130-185, (1976).
[77] J.MITCHELL, *On Carnot-Caratheodory metrics*, J.Differ. Geom. 21, 35-45, (1985).

[78] R. MONTI, M. RICKLY, *Geodetically convex sets in the Heisenberg group*, J. Convex Anal., **12** , 187-196, (2005).

[79] R. MONTI, *Distances, boundaries and surface measures in Carnot-Carathéodory spaces*, PhD Thesis, Universià di Trento, (2001).

[80] A. MORBIDELLI, *Fractional Sobolev norms and structure of Carnot-Caratheodory balls for Hörmander vector fields*, Studia Math. **139**, no.3 , 213-144, (2000).

[81] J. MOSER, *A new proof of De Giorgi's theorem concerning the regularity problem for elliptic differential equations*, Comm. Pure Appl. Math., bf 13, 457-468, 1960.

[82] A. NAGEL, F. RICCI, E.M. STEIN, FUNDAMENTAL SOLUTIONS AND HARMONIC ANALYSIS ON NILPOTENT GROUPS, Bull. Am. Math. Soc., **23**, 1, 139-145, (1999).

[83] A. NAGEL, E.M. STEIN, S. WAINGER, *Balls and metrics defined by vector fields I: Basic properties*, Acta Math., **155**, 103-147, (1985).

[84] P. NEGRINI, V. SCORNAZZANI, *Winer criterion for a class of degenerate elliptic operators*, J. Diff. Equation, **66**, 2, 152-164, (1987).

[85] P. PANSU, *Métriques de Carnot-Carathéodory quasiisométries des espaces symétriques de rang un*, Ann. Math., **129**, 1-60, (1989).

[86] YU.G.RESHETNYAK, *Generalized derivatives and differentiability almost everywhere*, Mat. Sb. **75**, 323-334 (in Russian), Math. USSR-Sb. **4**, 293-302 (English translation), (1968).

[87] M. RICKLY, *First-order regularity of convex functions on Carnot groups*, J. Geom. Anal. **16**, n.4, 679-702 (2006).

[88] M. RICKLY, *On questions of existence and regularity related to notions of convexity in Carnot groups*, PhD thesis, (2005).

[89] R. T. ROCKAFELLAR, *Maximal monotone relations and the second derivatives of convex functions*, Ann. Inst. H. Poincaré, Analyse non linéaire, **2**, 167-184, (1985).

[90] R. T. ROCKAFELLAR, R. J. WETS, *Variational Analysis*, Springer, (1997).

[91] R. T. ROCKAFELLAR, *Second Order Convex Analysis*, NonLinear and Convex An. **1**, 1-16, (2000).

[92] K. ROGOVIN, *Non-smooth analysis in infinite dimensional Banach homogeneous groups*, J. Convex Anal., **14**, 4, 667-691, (2007)

[93] L.ROTHSCHILD, E.M.STEIN, *Hypoelliptic differential operators and nilpotent groups*, Acta Math., **137**, 247-320, (1976).

[94] A. SÁNCHEZ-CALLE, *Fundamental solutions and geometry of the sum of squares of vector fields*, Invent. Math., **78**, 143-160, (1984).

[95] L. SCHWARTZ, *Théorie des distributions*, Hermann Paris, (1966).

[96] E.M. STEIN , *Singular integrals and differentiability properties of functions*, Princeton Univ. Press (1970).

[97] F. TRÉVES, *Topological vector spaces, distributions and kernels*, Academic Press, (1967).

[98] N.S. TRUDINGER, *On Hessian measures for non-commuting vector fields*, Pure Appl. Math. Q. **2**, n. 1, part 1, 147-161, (2006).

[99] N.S. TRUDINGER, *On Hessian measures for non-commuting vector fields*, Pure Appl. Math. Q. **2**, n. 1, part 1, 147-161, (2006).

[100] N. S. TRUDINGER, X. WANG, *Hessian measures I*, Topol. Methods Nonlinear Anal. 10, 225-239, (1997).

[101] N.S. TRUDINGER, X. WANG, *Hessian measures II*, Ann. of Math. **150**, 579-604, (1999).
[102] C. WANG, *Viscosity convex functions on Carnot groups*, Proc. Amer. Math. Soc., **133**, 1247-1253 (2005).
[103] X. WANG, *The k-Hessian equation*, Geometric analysis and PDEs, Lecture Notes in Math., **1977**, Springer, Dordrecht, (2009).