Characterizations of differentiability
for h-convex functions in stratified groups

VALENTINO MAGNANI AND MATTEO SCIENZA

Abstract. Using the notion of h-subdifferential, we characterize both first and second order differentiability of h-convex functions in stratified groups. Besides some new results involving the h-subdifferential of h-convex functions, we show that at all h-differentiability points of an h-convex function, the existence of a second order expansion coincides with a suitable differentiability of its horizontal gradient.

Mathematics Subject Classification (2010): 32F17 (primary); 53C17, 26B05 (secondary).

1. Introduction

The purpose of this work is to establish new results for first and second order differentiability of h-convex functions from the viewpoint of Nonsmooth Calculus, namely, using a suitable notion of subdifferential.

Following notation and terminology of Section 2, we represent a stratified group $\mathbb{G}$ as a finite-dimensional Hilbert space, that is the direct sum of orthogonal subspaces $H_1, H_2, \ldots, H_\iota$ and that is equipped with a suitable polynomial operation. Here H_1 is the subspace of *horizontal directions* at the origin, that yields the first layer V_1 of the Lie algebra $\mathfrak{g}$, that is formed by the so-called left-invariant horizontal vector fields. More details and precise definitions can be found in Section 2.

According to [9], 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 restricted to all horizontal lines contained in Ω. More precisely, we say that $u : \Omega \longrightarrow \mathbb{R}$ is h-convex if for all $h \in H_1$ with $[0, h] = \{th : 0 \leq t \leq 1\}$ and $[0, h] \subset x^{-1} \cdot \Omega$, we have

$$u\big(x(th)\big) \leq (1 - t)u(x) + tu(xh) \tag{1.1}$$

for every $t \in [0, 1]$. This notion of h-convexity is local and does not require any assumption on Ω. It is not difficult to observe that smooth h-convex functions are characterized by an everywhere non-negative *horizontal Hessian*, see Definition 2.7. Throughout this work, all h-convex functions on an open subset of $\mathbb{G}$ will be assumed to be measurable, since this assumption implies their Lipschitz continuity on compact subsets, as proved in [23]. Measurability can be replaced by local boundedness from above, that also yields the Lipschitz continuity on compact subsets, [18].

We will also use the following local notion of h-subdifferential. We say that $p \in H_1$ is an *h-subdifferential* of $u : \Omega \longrightarrow \mathbb{R}$ at $x \in \Omega$ if

$$u(xh) \geq u(x) + \langle p, h \rangle \qquad (1.2)$$

whenever $h \in H_1$ and $[0, h] \subset x^{-1} \cdot \Omega$. The set of all h-subdifferentials of u at x is denoted by $\partial_H u(x)$. This defines the set-valued mapping $\partial_H u : \Omega \rightrightarrows H_1$. The scalar product $\langle \cdot, \cdot \rangle$ in the inequality (1.2) is the one fixed on $\mathbb{G}$.

Our starting point was the characterization of the second order differentiability of h-convex functions, that we establish in Theorem 1.5. In tackling this problem, we realized that it first requires different new tools involving h-subdifferentials, further nonsmooth notions of differentiability and the characterization of h-differentiability of h-convex functions through the h-subdifferential. We believe that these facts should play an important role in the development of a Nonsmooth Calculus for h-convex functions.

To present our results, we will start from notions of second order differentiability. We will say that a locally Lipschitz function $u : \Omega \to \mathbb{R}$ is *twice h-differentiable at x* if there exists the *horizontal gradient* $\nabla_H u(x)$ of u at x, and moreover there exists an h-linear map $A_x : \mathbb{G} \to H_1$ such that

$$\left\| \frac{\nabla_H u(xw) - \nabla_H u(x) - A_x(w)}{\|w\|} \right\|_{L^\infty(B_\delta, H_1)} \longrightarrow 0 \quad \text{as} \quad \delta \to 0^+. \qquad (1.3)$$

Horizontal gradients, h-linear maps and h-differentiability are introduced in Definition 2.5 and Definition 2.6. If the limit (1.3) holds, then we equivalently 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 $D_H^2 u(x)$, since it is uniquely defined. The notion of differentiability in the extended sense is well posed, since Lipschitz functions are almost everywhere h-differentiable, [21]. Differentiability in the extended sense in the Euclidean case has been introduced by Rockafellar, [26]. The next lemma establishes a precise characterization of this differentiability. In the sequel, we denote by $\mathbb{B}$ the closed unit ball in H_1 centered at the origin.

Lemma 1.1. *Let $u : \Omega \longrightarrow \mathbb{R}$ be h-convex, let $x \in \Omega$ and let $A_x : \mathbb{G} \to H_1$ be h-linear. We have that u is h-differentiable at x and satisfies* (1.3) *if and only if there exists $v \in H_1$ such that*

$$\partial_H u(xw) \subseteq v + A_x(w) + o(\|w\|)\mathbb{B} \qquad (1.4)$$

for all $w \in x^{-1}\Omega$. In particular, the validity of (1.4) *implies the h-differentiability of u at x, with v equal to the horizontal gradient $\nabla_H u(x)$.*

This result is an important tool to establish one implication in the characterization of the second order differentiability of h-convex functions, stated in Theorem 1.5. Inclusion (1.4) can be seen as a continuity of the subdifferential at those points where u is h-differentiable, joined with a first order expansion of the horizontal gradient. The delicate implication of Lemma 1.1 is that the extended differentiability of $\nabla_H u$ implies the inclusion (1.4). This is a consequence of the following theorem.

Theorem 1.2. *Let $u : \Omega \to \mathbb{R}$ be h-convex. Then for every $x \in \Omega$, we have*

$$\bar{co}\left(\nabla_H^\star u(x)\right) = \partial_H u(x).$$ (1.5)

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 *reachable h-gradient* is given by

$$\nabla_H^\star u(x) = \left\{ p \in H_1 : x_k \to x, \nabla_H u(x_k) \text{ exists for all } k\text{'s and } \nabla_H u(x_k) \to p \right\}.$$ (1.6)

It is not difficult to prove that (1.5) immediately leads us to the characterization of h-convex functions by the property of having everywhere nonempty h-subdifferential, see Remark 3.7. The proof of equality (1.5) in the Euclidean case can be found for instance in [3]. Here we apply the Hahn-Banach's theorem inside the horizontal subspace H_1. However, the proof of Theorem 1.2 has a new difficulty. In fact, 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, as in [17]. The proof of Theorem 1.2 also uses the fact that the graph of the h-subdifferential mapping is closed. This follows from the following more general "set continuity" of the h-subdifferential, that does not require the h-differentiability of the function at the fixed point.

Proposition 1.3. *Let $\Omega \subset \mathbb{G}$ be an open set and let $u_i : \Omega \to \mathbb{R}$ denote a sequence of h-convex functions. Suppose that u_i uniformly converges on compact sets to an h-convex function $u : \Omega \to \mathbb{R}$. 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*

$$\partial_H u_i(x_i) \subseteq \partial_H u(x) + \epsilon \mathbb{B} \quad \text{for all } i \geq i_0.$$ (1.7)

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

$$\partial_H u_i(y) \subseteq \nabla_H u(y) + \epsilon \mathbb{B} \quad \text{for all } i \geq i_{\epsilon,K}, \text{ whenever } y \in K.$$ (1.8)

Proposition 1.3 joined with a nonsmooth mean value theorem, see Theorem 3.14, are used to prove the following first order characterization of h-differentiability.

Theorem 1.4 (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 $\nabla_H u(x) = p$.*

This theorem shows that (1.4) implies in particular the h-differentiability of u at x. The fact that h-differentiability implies the uniqueness of the h-subdifferential has been already observed in [9]. The opposite implication is more delicate, since in the Euclidean approach the use of the Hahn-Banach theorem requires the subadditivity of horizontal directional derivatives, that was not known before. We establish this property in Corollary 3.17, proving the formula

$$u'(x, h) = \max_{p \in \partial_H u(x)} \langle p, h \rangle, \qquad (1.9)$$

where $u'(x, h)$ is the horizontal directional derivative introduced in Definition 3.16. Nevertheless, in the proof of Theorem 1.4 we follow a different scheme, that does not use this subadditivity. We decompose the difference quotient of u into sums of difference quotients along horizontal directions. This fits with the general approach to differentiability in stratified groups, [21]. We also point at that Theorem 1.4 is new also in all Heisenberg groups, see [6] for other related results. The interesting point here is that the same scheme is one of the important features in the proof of one implication in the characterization of Theorem 1.5. This corresponds to the fact that twice h-differentiability implies the existence of a second order h-expansion.

The second order differentiability of h-convex functions is an interesting research area, where several questions are not yet understood. Since the works of Busemann and Feller, [5], and of Aleksandrov [2], there have been different methods to establish the a.e. second order differentiability of convex functions in Euclidean spaces. The functional analytic method by Reshetnyak, [22], relies on the fact that the gradient of a convex function has bounded variation. This scheme can be extended to stratified groups, provided that an h-convex function is H-BV^2 in the sense of [4]. This has been established by different authors for h-convex functions on Heisenberg groups and two step stratified groups [10, 12–14] and for k-convex functions with respect to Hörmander vector fields of step two, [28].

We refer to [10] for the presently known Aleksandrov-Busemann-Feller theorem in stratified groups: *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 $P_x : \mathbb{G} \longrightarrow \mathbb{R}$ whose homogeneous degree is less than or equal to two, such that

$$u(xw) = P_x(w) + o(\|w\|^2). \qquad (1.10)$$

Unfortunately, in higher step groups the fact that h-convex functions are H-BV^2 is not clear yet, hence the Aleksandrov-Busemann-Feller's theorem is an important open question for general stratified groups. On the other hand, the first proofs of this result in Euclidean spaces, [2,5] and also some of the subsequent proofs, did not use the bounded variation property of the gradient. For instance, Rockafellar's proof of

[24] relies on Mignot's a.e. differentiability of monotone functions, [20], where the crucial observation is that the subdifferential of a convex function is monotone. This approach to Aleksandrov's theorem constitutes a further motivation to develop the study of both first and second order Nonsmooth Calculus for h-convex functions. We are now in the position to state our main characterization.

Theorem 1.5 (Second order characterization). *Let* $u : \Omega \longrightarrow \mathbb{R}$ *be* h-convex *and let* $x \in \Omega$. *We have that* 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* $\nabla_{V_2} u(x) = \big(X_{m_1+1} u(x), \ldots, X_{m_2} u(x) \big)$ *of* u *at* x *along* V_2 *exists, where* $(X_{m_1+1}, \ldots, X_{m_2})$ *is an orthonormal basis of the second layer* V_2;
2. *if* P_x *is the second order* h-expansion of u at x, then

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

where $(\nabla_H^2)_{ij} = \frac{X_i X_j + X_j X_i}{2}$, $i, j = 1, \ldots, m_1$ *is the* horizontal Hessian operator;
3. *if* $D_H^2 u(x)$ *is the second order* h-differential of u at x and a_j^{li} *are the coefficients appearing in* (2.4), *then* $\big(D_H^2 u(x) \big)_{ij} = X_i X_j P_x$ *and we have*

$$\big(\nabla_H^2 P_x \big)_{ij} = \big(D_H^2 u(x) \big)_{ij} - \sum_{l=m_1+1}^{m_2} a_j^{li} \, X_l u(x). \tag{1.11}$$

Joining the Aleksandrov-Busemann-Feller's Theorem of [10] with our Theorem 1.5, we immediately achieve the following corollary.

Corollary 1.6. *Let* $\mathbb{G}$ *be a two step stratified group and let* $\Omega \subset \mathbb{G}$ *be an open subset. If* $u : \Omega \longrightarrow \mathbb{R}$ *is* h-convex, *then the second order* h-differential $D_H^2 u(x)$ *exists for a.e.* $x \in \Omega$. *Moreover, the properties* 1, 2 *and* 3 *of Theorem* 1.5 *hold.*

Let us point out that (1.11) in the case of commutative groups, namely, finite-dimensional Banach spaces, yields in particular Rockafellar's result on symmetry and nonnegativity of $D_H^2 u(x)$, see [26]. For h-convex functions on noncommutative stratified groups the symmetric part $\nabla_H^2 u(x)$ of $D_H^2 u(x)$ is non-negative and the loss of symmetry of $D_H^2 u(x)$ is natural, since it also takes into account the first order derivatives along directions of the second layer of the Lie algebra of $\mathbb{G}$.

2. Basic notions

A *stratified group* can be thought of as a graded vector space $\mathbb{G} = H_1 \oplus \cdots \oplus H_\iota$ equipped with a polynomial group operation given by the Backer-Campbell-Hausdorff formula and whose Lie algebra $\mathfrak{g}$ satisfies the following properties. There exist linear subspaces $V_1, V_2, \ldots, V_\iota$ of $\mathfrak{g}$ such that $\mathfrak{g} = V_1 \oplus \cdots \oplus V_\iota$, $V_j = \left[V_1, V_{j-1} \right]$, for all $j \leq 1$ and $V_j = \{0\}$ if and only if $j > \iota$. On $\mathbb{G}$ we can define a natural family of dilations $\delta_r : \mathbb{G} \to \mathbb{G}$ compatible with the group operation, see [11]. A scalar product on $\mathbb{G}$ will be understood, such 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, where $m_0 = 0$, then

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

We define $(X_{m_{s-1}+1}, \ldots, X_{m_s})$ to be the basis of V_s such that $X_j(0) = e_j$. Throughout, we fix a homogeneous distance d on $\mathbb{G}$, *i.e.* a continuous map $d : \mathbb{G} \times \mathbb{G} \to [0, +\infty)$ that makes $(\mathbb{G}, d)$ a metric space and it has the following properties

(1) $d(x, y) = d(ux, uy)$ for every $x, y, u \in \mathbb{G}$,
(2) $d(\delta_r x, \delta_r y) = r d(x, y)$ for every $r > 0$.

For every $w \in \mathbb{G}$, we denote by $\|w\|$ the homogeneous norm of w induced by the distance d by $\|w\| = d(0, w)$. Open balls with respect to d will be denoted by $B_{x,r}$. The following proposition is a well known fact, see for instance [19].

Proposition 2.1. *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 set $U \subset \mathbb{R}^\gamma$ such that*

$$\overline{B}_{0,1} \subset \left\{ \prod_{s=1}^{\gamma} a_s e_{i_s} \mid (a_1, \ldots, a_\gamma) \in U \right\}. \tag{2.1}$$

Remark 2.2. The inclusion (2.1) can be always established by a rescaling argument, once we know that $\left\{ \prod_{s=1}^{\gamma} a_s e_{i_s} \mid (a_1, \ldots, a_\gamma) \in U \right\}$ is a neighbourhood of the origin.

Definition 2.3. Let $U \subset \mathbb{R}^\gamma$ and $(i_1, \ldots, i_\gamma) \in \{1, \ldots, m_1\}^\gamma$ be as in Proposition 2.1. Thus, we define

$$W = \left\{ \prod_{s=1}^{\gamma} a_s e_{i_s} \mid (a_1, \ldots, a_\gamma) \in U \right\} \quad \text{and} \quad M = \sup_{y \in W} \|y\|.$$

Definition 2.4 (*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 $\lambda \in [0, 1]$.

We denote by H_x the left translation of H_1 by x, namely $H_x = x H_1$. 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 $\mathbb{G}$.

Definition 2.5. We say that a linear map $L : \mathbb{G} \longrightarrow \mathbb{R}$ is *h-linear* if $L(x) = L(\pi_1(x))$ for every $x \in \mathbb{G}$.

Definition 2.6. 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 it by $d_H u(x)$. Its associated vector with respect to the fixed scalar product is the *horizontal gradient*, denoted by $\nabla_H u(x)$.

Definition 2.7. Let $u : \Omega \longrightarrow \mathbb{R}$ be a $C^2(\Omega)$ function. We define the *horizontal Hessian* of u as follows $(\nabla_H^2 u)_{ij} = \left(\frac{X_i X_j u + X_j X_i u}{2} \right)_{ij}$, for all $i, j = 1, \dots, m_1$.

Definition 2.8. 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 $d(\alpha) = \sum_{i=1}^n d_i \alpha_i$, 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.

As in [11], given $a \in \mathbb{N}$, we shall denote by $\mathcal{P}_a$ the space of polynomials of homogeneous degree $\le a$. Moreover, by [11, Proposition 1.25], $\mathcal{P}_a$ is invariant under left translations. Given a multiindex $I = (i_1, \dots, i_n), 1 \le i_j \le m_1$, we set

$$X^I = X_{i_1} \cdots X_{i_n} \quad \text{and} \quad |I| = n.$$

Proposition 2.9 ([11, 1.30]). *Take $a \in \mathbb{N}$, and let μ be the dimension of $\mathcal{P}_a$. Then*

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

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

Remark 2.10. 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 multiindex $\alpha, |\alpha| \le 2$, then by the previous proposition P is uniquely determined.

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

Lemma 2.11. *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

$$P(x) = \langle \nabla_{V_2} P, x \rangle + \frac{1}{2} \langle \nabla_H^2 P x, x \rangle \, , \tag{2.2}$$

where $\langle \nabla_{V_2} P, x \rangle = \sum_{j=m_1+1}^{m_2} X_j P\, x_j$.

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 $j = 1, \ldots, n$, where $a_j^l(x)$ are $(d_l - d_j)$-homogeneous polynomial. Since $\nabla_{V_2} P = (X_{m_1+1}P, \ldots, X_{m_2}P)$ 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 $j = m_1 + 1, \ldots, m_2$. Hence, it remains to prove that

$$\frac{c_{ij} + c_{ji}}{2} = \frac{X_i X_j P + X_j X_i P}{2} \tag{2.3}$$

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

$$X_j(x) = \partial_{x_j} + \sum_{l=m_1+1}^{m_2} \sum_{i=1}^{m_1} a_j^{li}\, x_i\, \partial_{x_l} + \sum_{l=m_2+1}^{n} a_j^l(x)\, \partial_{x_l} \tag{2.4}$$

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

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

that immediately yields

$$X_i X_j P = \frac{c_{ij} + c_{ji}}{2} + \sum_{l=m_1+1}^{m_2} X_l P\, a_j^{li} \, . \tag{2.5}$$

Finally, formula (2.3) follows by the equality $a_j^{li} = -a_i^{lj}$. This is in turn a consequence of the Baker-Campbell-Hausdorff formula for the second order bilinear terms. $\qquad\square$

Definition 2.12. Let us fix $x \in \Omega$ and $a \in \mathbb{N}$. Let $f : \Omega \longrightarrow \mathbb{R}$ be a continuous function whose distributional derivatives $X^I f$ are continuous functions in a neighborhood of x whenever $|I| \leq a$. 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 $|I| \leq a$.

Theorem 2.13 (Stratified Taylor Inequality [11, 1.42]). *For each positive integer k there exists $C_k > 0$ such that for any continuous function $f : \Omega \longrightarrow \mathbb{R}$ whose distributional derivatives $X^I f$ are continuous functions whenever $|I| \leq k$, we have*

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

for all $x, y \in \Omega$, 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 small $r > 0$, we have

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

where we have set $X^I = X_{i_1} \cdots X_{i_l}$ and $I = (i_1, \ldots, i_l) \in \{1, \ldots, m_1\}^l$.

Lemma 2.14. *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 $\lambda = \max_{\|w\|=1} |P^{(2)}(w)|$. If we consider $P(xh)$ as a function of $h \in \mathbb{G}$, then for all $h \in H_1$ there holds*

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

Proof. For every $1 \leq i, j \leq m_1$, 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 2.13 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

$$P(xh) - P(x) - \langle \nabla_H P(x), h \rangle = P_x^{(2)}(h). \tag{2.6}$$

By (2.6) 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 2.9. Finally we get

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

3. Properties of the h-subdifferential

This section is devoted to the proofs of our results concerning h-differentiability, h-subdifferential, converging sequences of h-convex functions and nonsmooth mean value theorems for h-convex functions. Recall that $\mathbb{B} \subset H_1$ is the closed ball centered at the origin of radius one with respect to the fixed scalar product on $\mathbb{G}$.

Lemma 3.1. *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

$$\lambda u(xh) \geq \lambda u(x) + \langle \lambda p, h \rangle$$
$$(1 - \lambda)u(xh) \geq (1 - \lambda)u(x) + \langle (1 - \lambda)q, h \rangle \,. \qquad \square$$

Remark 3.2. We recall from the introduction that our assumption of measurability for all h-convex functions yields their Lipschitz continuity on compact subsets, [23]. As a straightforward consequence of this fact, by definition of h-subdifferential it follows that for every $B_{x,r}$ with closure contained in Ω and for every h-convex function $u : \Omega \longrightarrow \mathbb{R}$, there exists a positive number $L > 0$, depending on $x \in \Omega$, $r > 0$ and u, such that

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

Remark 3.3. As already mentioned, an h-convex function $u : \Omega \longrightarrow \mathbb{R}$ that is h-differentiable at $x \in \Omega$ has unique h-subdifferential, hence $\partial_H u(x) = \{\nabla_H u(x)\}$, according to [9].

Recall that the symbol co to denote the linear convex envelope in H_1. Then our first important tool is the following theorem.

Theorem 3.4. *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*

$$\partial_H u(x) \subseteq \bar{co}\left(\nabla_H^\star u(x)\right) , \tag{3.1}$$

where $\nabla_H^\star u(x)$ is defined in (1.6).

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

$$\langle z, q \rangle > \alpha \qquad \forall z \in \nabla_H^\star u(x). \tag{3.2}$$

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 (3.2). 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 [17, Lemma 2.7] there exists a normal subgroup $N \subset \mathbb{G}$, such that $N \cap Q = \{e\}$ and $NQ = \mathbb{G}$. Moreover, by [17, Proposition 2.8], there exist ν_q and μ_N, respectively Haar measures on Q and N such that for every measurable set $A \subset \mathbb{G}$

$$\mu(A) = \int_N \nu_q(A_n) \, d\mu_N(n) \tag{3.3}$$

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 (3.3) 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 $B_{\bar{n},s}^N \cdot B_{\delta_{\bar{t}}q,\epsilon}^Q \subset B_{x,r}$, where $B_{\bar{n},s}^N$ and $B_{\delta_{\bar{t}}q,\epsilon}^Q$ are open balls respectively in N and Q. Fix a point $n \in B_{\bar{n},s}^N$ 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 and 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 $n_j \to \bar{n} \in B_{\bar{n},s}^N$ 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.

$$u(\bar{n}\delta_{t_1} q) - u(\bar{n}\delta_{t_2} q) \geq \frac{\alpha}{2}(t_1 - t_2) \qquad -\epsilon + \bar{t} < t_2 < t_1 < \epsilon + \bar{t}. \tag{3.4}$$

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 (3.4). $\square$

Joining Theorem 3.4 with [9, Theorem 9.2], we immediately get:

Corollary 3.5. *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*

$$\sup_{\substack{p \in \partial_H u(y) \\ y \in B_{x,r}}} |p| \le \frac{C}{r} \frac{1}{|B_{x,15r}|} \int_{B_{x,15r}} |u(y)| dy. \qquad (3.5)$$

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

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

Proof of Proposition 1.3. 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 (3.5) 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 $\operatorname{dist}(p_{j_k}, \partial_H u(x_{j_k})) \ge \epsilon$. 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 us select a sequence of compact sets (K_l) and a subsequence (j_l) such that $p_{j_l} \to q$ and $\|u_{i_l} - u\|_{L^\infty(K_l)} < \frac{1}{l}$. Recall that $p_{j_l} \in \partial_H u_{j_l}(x_{j_l})$. It follows that $u_{j_l}(x_{j_l} h) \ge u_{j_l}(x_{j_l}) + \langle p_{j_l}, h \rangle$, whenever $[0, h] \subseteq H_1 \cap x_{j_l}^{-1} \Omega$. By the uniform convergence of (u_i), for any integer l sufficiently large, we get

$$u(x_{j_l} h) \ge u(x_{j_l}) - \frac{1}{l} + \langle p_{j_l}, h \rangle, \qquad (3.6)$$

for all $h \in H_1$ such that $[0, h] \subseteq H_1 \cap x_{j_l}^{-1} \cdot K_l$. Now, we fix any $\bar{h} \in H_1$ such that $[0, \bar{h}] \subseteq x^{-1} \cdot \Omega$. Therefore there exists $l_0 \in \mathbb{N}$ such that for every $l \in \mathbb{N}$ greater than l_0, we have $[0, \bar{h}] \subset x^{-1} \cdot K_l$ and there exists $\rho > 0$ such that $I(x \cdot [0, \bar{h}], \rho) \subset K_l$. By the continuity of left translations, we get $l_1 \in \mathbb{N}$ greater than l_0, such that

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

for all $l \ge l_1$. It follows that $[0, \bar{h}] \subseteq x_{j_l}^{-1} K_l$ for all $l \ge l_1$ and this allows us to apply (3.6) for $h = \bar{h}$ and pass to the limit with respect to l, getting

$$u(x\bar{h}) \ge u(x) + \langle q, \bar{h} \rangle. \qquad (3.7)$$

The arbitrary choice of $\bar{h}$ implies that $q \in \partial_H u(x)$, giving a contradiction.

Let us now suppose in addition that u is everywhere h-differentiable. Again, by contradiction, there exist a compact set $D \subset \Omega$, $\epsilon > 0$ and a subsequence (j_l) such that for all $l, x_{j_l} \in D$ we have

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

Therefore 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, $x_{j_l} \to \bar{x} \in D$ and $p_{j_l} \to \bar{p}$. By h-differentiability at $\bar{x}$ and Remark 3.3, taking into account the first part of this proposition, for every $\delta > 0$ there exists $l' \in \mathbb{N}$ such that

$$\partial_H u_{j_l}(x_{j_l}) \subset \nabla_H u(\bar{x}) + \delta \mathbb{B}$$

for every $l > l'$. Thus, we achieve $\epsilon \leq \mathrm{dist}(p_{j_l}, \partial_H u(x_{j_l})) \leq 2\delta$ for all $l > l'$. If we choose $\delta = \frac{\epsilon}{4}$, then reach a contradiction, concluding the proof. $\qquad\square$

Taking the constant sequence of h-convex functions in the previous proposition and taking into account (1.7), we immediately reach the following simple consequence.

Corollary 3.6. *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 1.2.

Proof of Theorem 1.2. By virtue of Theorem 3.4, we only have to prove the inclusion

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

By Corollary 3.6, 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. $\qquad\square$

Remark 3.7. The almost everywhere h-differentiability of an h-convex function u implies that $\nabla_H^* u(x) \neq \emptyset$ for all $x \in \Omega$. Thus, in view of formula (1.5), it follows 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 observed in [6] for h-convex functions on Heisenberg groups. The opposite implication in general stratified groups can be found in [9] for h-convex domains. This implication holds also for h-convex functions on general open sets. In fact, the everywhere existence of an h-subdifferential implies the everywhere existence of the one dimensional classical subdifferential along horizontal lines. Then the one dimensional characterization of convexity through the subdifferential gives the classical convexity along horizontal lines and this proves the h-convexity.

Definition 3.8. Let $\Omega \subset \mathbb{G}$ be open, let $u : \Omega \to \mathbb{R}$ and let $x \in \Omega$. We define the set

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

Remark 3.9. 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, that in the Heisenberg group has been proved in [6]. 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 $[0, h] \subseteq (x^{-1} \cdot \Omega) \cap H_1$. 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 3.10. 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 3.11. *Consider a function $u = U + P$ in Ω, where U is h-convex and P is a polynomial with h-deg $P \leq 2$. Denote by $P^{(2)}$ the 2-homogeneous part of P. Thus, defining $\lambda = \max\limits_{\|w\|=1} |P^{(2)}(w)|$, it follows that $\partial_H^\lambda u(x) \supseteq \partial_H U(x) + \nabla_H P(x)$.*

Proof. Recall that by Lemma 2.14, for every $x, h \in \mathbb{G}$ we have

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

If $p \in \partial_H U(x)$, then

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

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

Proposition 3.12. *Let $\Omega \subset \mathbb{G}$ be open, let $U : \Omega \to \mathbb{R}$ be an h-convex function and let V be in $C_H^1(\Omega)$. We define the function u be equal to $U + V$ and fix $\lambda \geq 0$. It follows that 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_H^1 smoothness of V gives

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

hence $p - \nabla_H V(x) \in J_U^{1,-}(x)$. Since U is h-convex, in view of Remark 3.9 we have that $p - \nabla_H V(x) \in \partial_H U(x)$. This concludes the proof. $\qquad\square$

The following theorem extends the classical nonsmooth mean value theorem to stratified groups.

Theorem 3.13. *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 $p \in \partial_H^\lambda u(x\delta_t h)$ 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 $p \in \partial_H^\lambda u(x\delta_t h)$. By Proposition 1.3, 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 3.11, $\partial_H^\lambda u(x) \supseteq \partial_H U(x) + \nabla_H P(x)$ thus the previous inclusion implies that

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

then letting $k \to \infty$ we get that $p \in \partial_H^\lambda u(x\delta_t h)$. $\qquad\square$

As a consequence of the previous result, we immediately get the following theorem.

Theorem 3.14 (Nonsmooth mean value 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* $u(xh) - u(x) = \langle p, h \rangle$.

Proof. It suffices to apply Theorem 3.13 with $P = 0$ and $\lambda = 0$. □

Remark 3.15. In the literature a nonsmooth mean value can be found for Lipschitz mappings on Banach homogeneous groups, that clearly include stratified groups, [27]. Unfortunately, this work does not imply our Theorem 3.14, since it uses the notion of Clarke generalized gradient for Lipschitz mappings adapted to homogeneous groups.

Definition 3.16. 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 $\lim_{\lambda \to 0^+} \left(u(x\delta_\lambda h) - u(x) \right)\lambda^{-1}$, whenever it exists. We denote this derivative by $u'(x, h)$.

Corollary 3.17. *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*

$$u'(x, h) = \max_{p \in \partial_H u(x)} \langle p, h \rangle, \tag{3.8}$$

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 $\partial_H u(x)$,

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

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 3.13. 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. □

Proof of Theorem 1.4. Uniqueness of the h-subdifferential under h-differentiability has already been established in [9], see Remark 3.3. Let us assume now that the h-subdifferential p of u at x is unique. Let U, W and M be as in Definition 2.3. Thus, for any $w \in \mathbb{G}$ with $\|w\| = 1$, we have $w = \prod_{s=1}^{\gamma} a_s e_{i_s}$ for some $(a_1, \dots, a_\gamma) \in U$. We fix $r > 0$ such that $B_{0,r} \subset x^{-1}\Omega$ and define the h-convex function

$$g(y) = u(xy) - u(x) - \langle p, y \rangle$$

for any $y \in x^{-1}\Omega$. We choose $\rho_0 > 0$ such that $\rho_0 M < r$. Thus, whenever $0 < \rho < \rho_0$, by Theorem 3.14 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 $p_s \in \partial_H u \big(x \delta_\rho (\prod_{k=1}^{s-1} a_k e_{i_k}) \delta_{t_s} \delta_\rho a_s e_{i_s}\big)$ with $t_s \in [0, 1]$. By Proposition 1.3, for every $\epsilon > 0$ there exists ρ_0 such that

$$\partial_H u \left(x \delta_\rho \left(\prod_{k=1}^{s-1} a_s e_{i_k} \right) \delta_{t_s} \delta_\rho a_s e_{i_s} \right) \subseteq \partial_H u(x) + \epsilon \mathbb{B} = \{p\} + \epsilon \mathbb{B}$$

for all $0 < \rho < \rho_0$ and $s = 1, \dots, \gamma$. Thus $|g(\delta_\rho w)| \leq C\gamma\epsilon\rho$, where C is independent on $(a_1, \dots, a_\gamma)$, since W is a bounded set. This implies that $\frac{|g(\delta_\rho w)|}{\rho}$ uniformly converges to zero with respect to $w \in W$ as $\rho \to 0^+$. $\quad\square$

4. Second order differentiability

The aim of this section is to prove the characterization of the second order differentiability of h-convex functions, stated in Theorem 1.5. Let us begin with the following simple fact.

Proposition 4.1. *Let $\Omega \subset \mathbb{G}$ be an open set and let $u : \Omega \to \mathbb{R}$. If u has a second order h-expansion P_x at $x \in \Omega$, then u is h-differentiable at x*

$$P_x^{(1)}(w) = \langle \nabla_H u(x), w \rangle . \tag{4.1}$$

Proof. If (1.10) holds, then we can rewrite this condition 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 (4.1), concluding the proof. $\quad\square$

Definition 4.2 (Difference quotients). 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 the *second order h-quotient* at x

$$\Delta^2_{x,\tau} u(w) = \frac{u(x\delta_\tau w) - u(x) - \tau \langle \nabla_H u(x), w \rangle}{\tau^2} \tag{4.2}$$

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

$$\Delta_{x,\tau} \partial_H u(w) = \frac{\partial_H u(x\delta_\tau w) - \nabla_H u(x)}{\tau}. \tag{4.3}$$

Remark 4.3. 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

$$
\begin{aligned}
\partial_H \left[\Delta^2_{x,\tau} u(w) \right] &= \tau^{-1} \{ \partial_H u_{x,\tau}(w) - \nabla_H u(x) \} \\
&= \tau^{-1} \{ \partial_H u(x\delta_\tau w) - \nabla_H u(x) \} \\
&= \Delta_{x,\tau} \partial_H u(w),
\end{aligned}
\tag{4.4}
$$

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

Proof of Lemma 1.1. Choosing $w = 0$ we get $\partial_H u(x) = \{v\}$, thus by Theorem 1.4, u is h-differentiable at x, moreover $v = \nabla_H u(x)$. The twice h-differentiability immediately follows from (1.4), taking its restriction to all points of h-differentiability. For the converse implication, we rewrite expansion (1.3) as follows, for all $\epsilon > 0$ there exists $\rho > 0$ such that

$$\left| \frac{\nabla_H u(xh) - \nabla_H u(x) - A_x(h)}{\|h\|} \right| \leq \epsilon \qquad \|h\| < \rho, \tag{4.5}$$

for all $h \in x^{-1}\Omega$ such that u is h-differentiable at xh. By (1.6), for any $w \in x^{-1}\Omega \cap B_{0,\rho}$, taking into account (4.5), we get

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

In an equivalent form, we have

$$\nabla^\star_H u(xw) \subseteq \nabla_H u(x) + A_x(w) + \epsilon \|w\| \mathbb{B}. \tag{4.6}$$

Moreover, the set on the right is convex, hence Theorem 1.2 yields

$$\partial_H u(xw) = \bar{co}\left(\nabla_H^{\star} u(xw)\right) \subseteq \nabla_H u(x) + A_x(w) + o(\|w\|)\mathbb{B}. \quad (4.7)$$

This concludes the proof. □

Proposition 4.4. *If $u : \Omega \longrightarrow \mathbb{R}$ 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 $w \in D$ and $0 < \tau < \delta$ we have*

$$\Delta_{x,\tau}\partial_H u(w) - A_x(w) \subseteq \epsilon\mathbb{B}. \quad (4.8)$$

Proof. Let u be twice h-differentiable at x, fix a compact 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 $\|w\| < \rho(\epsilon, D, \Omega)$ and $B_{x,\rho(\epsilon,D,\Omega)} \subset \Omega$. We consider $w = \delta_\tau h$, where $h \in D$, and $0 < \tau < \frac{\rho(\epsilon,D,\Omega)}{\mu_D}$. 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 (4.8). Conversely, let $S = \{w \in \mathbb{G} : \|w\| = 1\}$ be a compact set and fix $\epsilon > 0$. Then there exists $\delta > 0$ such that (4.8) 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 have enough tools to establish the characterization of second order differentiability of h-convex functions.

Proof of Theorem 1.5. Let us assume that u has a second order h-expansion at x. By Proposition 4.1, u is h-differentiable at x, then $P_x^{(0)}(w) = u(x)$ and $P_x^{(1)}(w) = \langle \nabla_H u(x), w \rangle$, where P_x is the polynomial associated to the second order h-expansion. Define $\phi(w) := P_x^{(2)}(w)$ and notice that $\nabla_H P_x^{(2)}(w)$ is an h-linear map, since it is a polynomial of homogeneous degree 1. The second order h-expansion yields

$$\begin{aligned}
\Delta_{x,\tau}^2 u(w) - \phi(w) &= \frac{u(x\delta_\tau w) - P_x^{(0)}(\delta_\tau w) - P_x^{(1)}(\delta_\tau w) - P_x^{(2)}(\delta_\tau w)}{\tau^2} \\
&= \frac{o(\|\delta_\tau w\|^2)}{\tau^2}.
\end{aligned} \quad (4.9)$$

As a consequence, $\Delta_{x,\tau}^2 u$ uniformly converges to ϕ on compacts sets as $\tau \to 0^+$. Moreover $\Delta_{x,\tau}^2 u$ is h-convex, then so is ϕ. Applying Proposition 1.3, we can establish that for every compact set $D \subset \Omega$ and $\epsilon > 0$ there exists $\delta > 0$ such that

$$\partial_H \Delta_{x,\tau}^2 u(w) \subseteq \nabla_H \phi(w) + \epsilon \mathbb{B}, \quad \text{for all} \quad w \in D \quad \text{and} \quad \tau \in (0, \delta).$$

Notice that (4.4) gives $\partial_H \Delta_{x,\tau}^2 u(w) = \Delta_{x,\tau} \partial_H u(w)$, hence

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

As a result, we have $\Delta_{x,\tau} \partial_H u(w) - \nabla_H P_x^{(2)}(w) \subset \epsilon \mathbb{B}$ whenever $w \in D$ and $0 < \tau < \delta$. By Proposition 4.4, u is twice h-differentiable. Furthermore, $\nabla_H P_x^{(2)}$ is the second order h-differential $D_H^2 u(x)$ of u at x.

We now assume that u is twice h-differentiable at x, where $D_H^2 u(x)$ denotes the second order h-differential of u at x. By Lemma 1.1, we have

$$\nabla_H u(xw) = \nabla_H u(x) + D_H^2 u(x)w + o(\|w\|),$$

where $D_H^2 u(x)$ is regarded as an h-linear mapping. Let U, W and M be as introduced in Definition 2.3. 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 $h\text{-deg} P \leq 2$ that satisfies the condition $P_x(0) = 0$ and

$$\nabla_H P_x(w) = \nabla_H u(x) + D_H^2 u(x)w. \tag{4.10}$$

This is as a consequence of Remark 2.10. Let $r > 0$ be such that $B_{0,r} \subset x^{-1} \cdot \Omega$. Let $\rho_0 > 0$ be such that $\rho_0 M < r$ and choose w such that $\|w\| = 1$. We consider $0 < \rho < \rho_0$ and $\delta_\rho w = \prod_{s=1}^{\gamma} \rho a_s e_{i_s}$, for some $(a_1, \ldots, a_\gamma) \in U$. Then $v(\delta_\rho w) = v(\delta_\rho w) - v(0)$ can be written as

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

Observe that v is an h-convex function plus a polynomial of homogeneous degree less than or equal to two. By Theorem 3.13 applied to horizontal directions e_{i_s}, we get

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

with $p_s \in \partial_H^\lambda v\big(x\big(\prod_{k=1}^{s-1} \rho\, a_k e_{i_k}\big)(t_s \rho\, a_s e_{i_s})\big)$, where $t_s \in [0, 1]$ and

$$\lambda = \max_{\|h\|=1} |P_x^{(2)}(h)|.$$

Moreover, by Proposition 3.12 we know that

$$p_s + \nabla P_x \left(\left(\prod_{k=1}^{s-1} \rho\, a_k e_{i_k} \right) (t_s \rho\, a_s e_{i_s}) \right) \in \partial_H u \left(x \left(\prod_{k=1}^{s-1} \rho\, a_k e_{i_k} \right) (t_s \rho\, a_s e_{i_s}) \right). \quad (4.11)$$

The expansion (1.4) for the h-subdifferential of u implies that

$$\begin{aligned}
\partial_H u &\left(x \left(\prod_{k=1}^{s-1} \rho\, a_k e_{i_k} \right) (t_s \rho\, a_s e_{i_s}) \right) \subset \nabla_H u(x) \\
&+ A_x \left(\left(\prod_{k=1}^{s-1} \rho\, a_k e_{i_k} \right) (t_s \rho\, a_s e_{i_s}) \right) + o\left(\left| \left(\prod_{k=1}^{s-1} \rho\, a_k e_{i_k} \right) (t_s \rho\, a_s e_{i_s}) \right| \right) \mathbb{B}.
\end{aligned} \quad (4.12)$$

Thus, by formula (4.10), taking into account (4.11) and (4.12), we get that

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

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 (4.9) to directions $z \in H_2$, with $|w| = 1$, and taking into account (2.2), hence getting the uniform limit

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

as $t \to 0^+$, where z varies in a compact neighborhood of zero in H_2 and 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 Proposition 4.1 and formula (2.2), then claim (2) follows. Now, with respect to the fixed basis $(e_1, \ldots, e_n)$ of $\mathbb{G}$, we have the coefficients $\left(D_H^2 u(x) \right)_{ij}$ such that

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

therefore (4.10) yields $\nabla_H P_x^{(2)}(w) = D_H^2 u(x) w$. For any $j = 1, \ldots, m_1$, we have

$$X_j P_x^{(2)}(w) = \sum_{i=1}^{m_1} (A_x)_{ij}\, w_i,$$

then formula (2.5) gives

$$X_i X_j P_x^{(2)} = \left(D_H^2 u(x)\right)_{ij} = (\nabla_H^2 P_x^{(2)})_{ij} + \sum_{l=m_1+1}^{m_2} X_l u(x)\, a_j^{li}.$$

As a result, we get

$$(\nabla_H^2 P_x^{(2)})_{ij} = (A_x)_{ij} - \sum_{l=m_1+1}^{m_2} X_l u(x)\, a_j^{li},$$

that coincides with the formula of claim (3). Finally, we recall that $P_x^{(2)}$ is the uniform limit on compact sets of the h-convex functions $\Delta_{x,\tau}^2 u$. This implies that $P_x^{(2)}$ is also h-convex and then its horizontal Hessian is non-negative. $\qquad\square$

References

[1] G. ALBERTI and L. AMBROSIO, *A geometrical approach to monotone functions in $\mathbb{R}^n$*, Math. Z. **230** (1999), 259–316.

[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** (1939), 3–35 (in Russian).

[3] L. AMBROSIO and N. DANCER, *Calculus of variations and partial differential equations*, In: "Topics on Geometrical Evolution Problems and Degree Theory", G. Buttazzo, A. Marino and M. K. V. Murthy (eds.), Papers from the Summer School held in Pisa, September 1996, Springer-Verlag, Berlin, 2000, 5–93.

[4] L. AMBROSIO and V. MAGNANI, *Weak differentiability of BV functions on stratified groups*, Math. Z. **245** (2003), 123–153.

[5] H. BUSEMANN and W. FELLER, *Krümmungseigenschaften konvexer Flächen*, Acta Math. **66** (1935), 1-47.

[6] A. CALOGERO and R. PINI, *Horizontal normal map on the Heisenberg group*, J. Nonlinear Convex Anal. **12** (2011), 287–307.

[7] F. H. CLARKE, "Optimization and Nonsmooth Analysis", Classics in Appl. Math., Vol. 5, SIAM, Philadelphia, PA, 1990.

[8] D. DANIELLI, N. GAROFALO and 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** (2003), 3487–3498.

[9] D. DANIELLI, N. GAROFALO and D. M. NHIEU, *Notions of convexity in Carnot groups*, Comm. Anal. Geom. **11** (2003), 263–341.

[10] D. DANIELLI, N. GAROFALO, D. M. NHIEU and F. TOURNIER *The theorem of Busemann-Feller-Alexandrov in Carnot groups*, Comm. Anal. Geom. **12** (2004), 853–886.

[11] G. B. FOLLAND and E. M. STEIN, "Hardy Spaces on Homogeneous Groups", Princeton University Press, 1982.

[12] N. GAROFALO and F. TOURNIER, *New properties of convex functions in the Heisenberg group*, Trans. Amer. Math. Soc. **358** (2005), 2011–2055.

[13] C. E. GUTIÉRREZ and A. MONTANARI, *Maximum and comparison principle for convex functions on the Heisenberg group*, Comm. Partial Differential Equations **29** (2004), 1305–1334.

[14] C. E. GUTIÉRREZ and A. MONTANARI, *On the second order derivatives of convex functions on the Heisenberg group*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **2** (2004), 349–366.

[15] P. JUUTINEN, G. LU, J. J. MANFREDI and B. STROFFOLINI, *Convex functions on Carnot groups*, Rev. Mat. Iberoam. **23** (2007), 191–200.

[16] G. LU, J. J. MANFREDI and B. STROFFOLINI, *Convex functions on the Heisenberg group*, Calc. Var. Partial Differential Equations **19** (2004), 1–22.

[17] V. MAGNANI, *Contact equations, Lipschitz extensions and isoperimetric inequalities*, Calc. Var. Partial Differential Equations **39** (2010), 233–271.

[18] V. MAGNANI, *Lipschitz continuity, Aleksandrov theorem and characterizations for H-convex functions*, Math. Ann. **334** (2006), 199–233.

[19] V. MAGNANI, "Elements of Geometric Measure Theory on Sub-Riemannian Groups", Scuola Normale Superiore Pisa, 2002.

[20] F. MIGNOT, *Contrôle optimal dans les inéquations variationelles elliptiques*, J. Funct. Anal. **22** (1976), 130–185.

[21] P. PANSU, *Métriques de Carnot-Carathéodory et quasiisométries des espaces symétriques de rang un*, Ann. Math. **129** (1989), 1–60.

[22] YU. G. RESHETNYAK, *Generalized derivatives and differentiability almost everywhere*, Math. USSR-Sb. **4** (1968), 293–302.

[23] M. RICKLY, *First-order regularity of convex functions on Carnot groups*, J. Geom. Anal. **16** (2006), 679–702.

[24] R. T. ROCKAFELLAR, *Maximal monotone relations and the second derivatives of convex functions*, Annal. Inst. H. Poincaré, Anal. Non Linéaire **2** (1985), 167–184.

[25] R. T. ROCKAFELLAR and R. J. WETS, "Variational Analysis", Springer, 1997.

[26] R. T. ROCKAFELLAR, *Second-order convex analysis*, J. Nonlinear Convex Anal. **1** (2000), 1–16.

[27] K. ROGOVIN, *Non-smooth analysis in infinite-dimensional Banach homogenous groups*, J. Convex Anal. **14** (2007), 667–691.

[28] N. S. TRUDINGER, *On Hessian measures for non-commuting vector fields*, Pure Appl. Math. Q. **2** (2006), 147–161.

[29] C. WANG, *Viscosity convex functions on Carnot groups*, Proc. Amer. Math. Soc. **133** (2005), 1247–1253.

Dipartimento di Matematica
Largo Bruno Pontecorvo 5
I-56127 Pisa, Italia
magnani@dm.unipi.it
scienza@mail.dm.unipi.it