Weak differentiability of BV functions on stratified groups

Luigi Ambrosio, Valentino Magnani
Scuola Normale Superiore
Piazza dei Cavalieri 7, 56126 Pisa, Italy
luigi@ambrosio.sns.it & v.magnani@sns.it

Abstract

We prove the almost everywhere approximate differentiability of functions with bounded variation in stratified groups and we study their approximate discontinuity set. We introduce functions of bounded higher order variation and obtain a weak version of Alexandroff differentiability theorem in this context. We present a nontrivial class of functions with second order bounded variation, arising from inf-convolution formula of a suitable "cost" function.

Contents

Introduction

Geometric Measure Theory in metric spaces has become a subject of increasing interest in these recent years, [4, 6, 7, 13, 32, 44, 57]. A particular attention has been devoted to the framework of Carnot-Carathéodory spaces and stratified groups, [24, 25, 26, 27, 33, 34, 37, 38, 39, 46, 48, 50, 51, 61].

Stratified groups, also known as Carnot groups, [21, 49], can be regarded as "intrinsic" tangent spaces of Carnot-Carathéodory manifolds, [9, 42, 45]. They naturally have additional homogeneity properties, e.g. the existence of a one parameter group of dilations which scale homogeneously with respect to a suitable left invariant distance, namely the Carnot-Carathéodory distance. Most of the classical analysis can be carried out in these groups and in Carnot-Carathéodory spaces. For instance, important results concerning the classical theory of Sobolev spaces, such as Poincaré inequalities, embedding theorems, representation formulas, trace theorems, compactness results and much more, have been extended to Carnot-Carathéodory spaces and to a general metric space setting [14, 22, 23, 29, 30, 31, 47]. In some respects, our study goes in the same direction of the previously mentioned results, dealing with the properties of functions whose "horizontal" distributional derivatives are measures, instead of L^p functions. Our approach takes advantage of some results about Sobolev spaces in Carnot-Carathéodory spaces, see Subsection 1.4, but the case of BV functions has also some distinctive features.

The notion of function with bounded variation has been extended to the context of Carnot-Carathéodory spaces in [11]. Recently, relevant rectifiability results concerning sets of H-finite perimeter in Heisenberg groups and general 2-step stratified groups have been achieved in [25], [27], extending the celebrated De Giorgi Rectifiability Theorem, [17]. In this context, it is rather natural to investigate what can be said about functions of horizontal bounded variation (in short, H-BV functions).

Our first basic result concerns weak differentiability of H-BV functions in stratified groups. Let $u : \Omega \longrightarrow \mathbb{R}$ be an H-BV function, where Ω denotes an open set of a stratified group $\mathbb{G}$. Then for a.e. $x \in \Omega$ there exist a horizontal vector $\nabla_H u(x)$ such that

$$\lim_{r\to 0^+} \frac{1}{r} \fint_{U_{x,r}} \left| u(y) - \tilde{u}(x) - \left\langle \nabla_H u(x), \ln(x^{-1}y) \right\rangle \right| \, dy = 0 \,. \tag{1}$$

Here $\nabla_H u$ is the density of the absolutely continuous part of the vector Radon measure $D_H u$ associated to u, the brackets $\langle \, , \rangle$ stand for the left invariant Riemannian scalar product considered in $\mathbb{G}$ and the ball $U_{x,r}$ of center x and radius r is built with respect to the Carnot-Carathéodory distance. The function $\ln : \mathbb{G} \longrightarrow \mathcal{G}$ is the inverse of the exponential map.

Moreover, we show that the approximate discontinuity set S_u (Definition 1.9) is contained in a countable union of essential boundaries of sets with H-finite perimeter. This result, together with those in [25], [27], allows us to conclude that in 2-step stratified groups the set S_u is countably rectifiable, in the sense of stratified groups.

These results extend to a more general setting some classical facts about BV functions on Euclidean spaces, see the historical note in [5], and [10], [20], [62].

We also introduce functions with H-finite higher order variation and we prove higher order differentiability results. In the Euclidean setting, functions of "Bounded Hessian" were introduced in [18], as those maps whose second order distributional derivatives are measures. Following a natural analogy, we define H-BV^k-functions as those functions whose horizontal distributional derivatives up to order k are measures.

For these maps we are able to prove an Alexandrov-type differentiability theorem: for a.e. $x \in \Omega$ there exists a polynomial $P_{[x]}$ with homogeneous degree less than or equal to k, such that

$$\lim_{r\to 0^+} \frac{1}{r^k} \fint_{U_{x,r}} |u - P_{[x]}| = 0\,. \tag{2}$$

The analogous Euclidean result can be found in Proposition 2.2 of [3].

Our method is based on two crucial estimates: first, in a suitable point x we estimate the difference $|u(y) - u(x)|$ utilizing the maximal function, see (23). Second, we use the representation formula (16) in the form (18) in order to obtain information on the behavior of $|D_H v|$, where $v = |D_H u|$. Notice also that the non-commutativity of the group is the source of additional difficulties, and our proof uses the Poincaré-Birkhoff-Witt Theorem. We point out that convex functions in Euclidean spaces are locally (H-)BV^2 maps, [1]. Thus, our approximate differentiability theorem (Theorem 3.9), together with standard L^∞ estimates for convex functions (see Theorem 1 of Section 6.3 in [19]) yield the classical Alexandrov differentiability Theorem (see Theorem 1 of Section 6.4 in [19]).

It is rather natural to look for a suitable notion of convexity for maps defined on stratified groups, this is motivated for instance by applications to PDE problems. This notion has been recently introduced in [15] and [36]. Notice that this intrinsic notion of convexity must be weaker than the classical Euclidean convexity. Here a remarkable open question that arises from the PDE context and has an interest in its own is to inquire whether the Alexandrov differentiability Theorem holds in a pointwise form for this class of convex functions. Precisely, what is still not clear is whether these maps are H-BV^2 in the sense of Definition 3.1. In fact, in [15] and [36] the above mentioned L^∞-estimates valid for Euclidean convex functions have been extended to the new class of maps introduced in the same papers. These estimates together with our Theorem 3.9 would imply the pointwise second order differentiability, as in the classical Alexandorv Theorem.

Acknowledgements. It is a pleasure to thank Fulvio Ricci for his precious comments.

1 Notation, definitions and basic tools

In this section we introduce the basic notions we are going to use throughout the paper. We start recalling some facts on the geometry of stratified groups.

1.1 Stratified groups

We consider a simply connected *stratified nilpotent Lie group* $\mathbb{G}$, whose Lie algebra $\mathcal{G}$ is the direct sum of subspaces V_i and the following commutator relations hold $[V_j, V_1] = V_{j+1}$, for any $j \geq 1$, with $V_i = \{0\}$ iff $i > \iota$. These groups are also called *Carnot groups*, [49]. The integer ι is called *step* of the group and the subspace V_1 is called the *horizontal space*. We denote left translations of the group as follows $l_x : \mathbb{G} \longrightarrow \mathbb{G}$, $l_x(y) = xy$. Via the differential of left translations, the horizontal space V_1 can be translated at any point x of $\mathbb{G}$, thus getting a subspace H_x of $T_x\mathbb{G}$. These subspaces are called *horizontal fibers* and the family of all H_x is called the *horizontal subbundle* of $\mathbb{G}$, denoted by H. We will write $H\Omega$ to indicate the horizontal subbundle on an open subset $\Omega \subset \mathbb{G}$. The graded structure allows us to define a one parameter group of dilations $\delta_r : \mathcal{G} \longrightarrow \mathcal{G}$, $r > 0$, which is defined as $\delta_r\Big(\sum_{j=1}^{\iota} v_j \Big) = \sum_{j=1}^{\iota} r^j v_j$, where $\sum_{j=1}^{\iota} v_j = v$ and $v_j \in V_j$ for each $j = 1, \dots \iota$. Since $\mathbb{G}$ is simply connected and nilpotent the the exponential map $\exp : \mathcal{G} \longrightarrow \mathbb{G}$ is a diffeomorphism, so that dilations can be canonically transposed on $\mathbb{G}$. We will use the same symbol to denote the dilations of the group. We denote by ln the inverse function of exp and we indicate by e the unit element of the group. The metric structure on a stratified group is built using what we call *graded metrics*. These are left invariant Riemannian metrics such that the spaces $\{V_i\}$ are orthogonal. Graded metrics are the "natural" choice that respects the algebraic structure of the group. The Riemannian scalar product induced by a graded metric will be denoted by $\langle\, , \rangle$. The symbol $|\cdot|$ will be used to indicate either the norm of a vector in the tangent bundle or the norm of a scalar number. By virtue of the left invariance of g we can construct a left invariant distance on $\mathbb{G}$ in such a way that it is 1-homogeneous with respect to dilations. To do this, we consider the class of *admissible paths*, e.g. absolutely continuous curves $\gamma : [a, b] \longrightarrow \mathbb{G}$, such that for a.e. $t \in [a, b]$ they satisfy $\gamma'(t) = \sum_{i=1}^{m} c_i(t) X_i\,(\gamma(t))$, where $\sum_{i=1}^{m} c_i^2(t) \leq 1$ and $(X_1, \dots, X_m)$ is an orthonormal frame of $H\Omega$. The conditions on commutators of H guarantee that any pair of points of $\mathbb{G}$ can be joined by an horizontal curve. Hence we can define the finite number

$$d(x, y) := \inf \Big\{ b - a \mid \gamma : [a, b] \longrightarrow \mathbb{G} \text{ is admissible and } \gamma(a) = x,\ \gamma(b) = y \Big\}$$

for any $x, y \in \mathbb{G}$. One can verify that d is a distance on $\mathbb{G}$, namely the *Carnot-Carathéodory* distance, see for instance [31]. This distance is continuous with respect to the topology of the group and has the following properties:

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

2. $d(\delta_r x, \delta_r y) = r\, d(x, y)$ for every $r > 0$.

Any continuous distance on $\mathbb{G}$ which has the above properties is called a *homogeneous distance.* We simply write $d(z)$ to denote the distance between z and e. All the homogeneous distances are bi-Lipschitz equivalent and induce the topology of the group. Throughout the paper we will utilize the Carnot-Carathéodory distance and we will denote it simply by d, if not otherwise stated. We denote by $Q = \sum_{j=1}^{l} j \dim(V_j)$ the Hausdorff dimension of $\mathbb{G}$ with respect to the Carnot-Carathéodory distance and by q the dimension of the Lie algebra. Notice that the Haar measure of $\mathbb{G}$, the Riemannian volume and the Q-dimensional Hausdorff measure on $\mathbb{G}$ coincide up to a dimensional factor. Furthermore, for a fixed basis of $\mathcal{G}$, we can read these measures on $\mathcal{G}$ as the q-dimensional Lebesgue measure up to a positive factor. We will use the symbol dx to denote the integration with respect to the Riemannian volume measure. We denote either by $|A|$ or $\mathrm{vol}(A)$ the Riemannian volume of a Borel set A. The symbol $⨍_A$ indicates the averaged integral and u_A denotes the average $⨍_A u$ of the map $u : A \longrightarrow \mathbb{R}$.

Definition 1.1 We denote by $U_{x,r}$ the open ball with center $x \in \mathbb{G}$ and radius $r > 0$ with respect to Carnot-Carathéodoy distance d. We will omit the center x if it coincides with the unit element of the group.

The notions of continuity and differentiability we will consider throughout the paper are clearly independent of the homogeneous metric we consider. So, for the sake of simplicity we will always consider the Carnot-Carathéodory distance.

1.2 H-BV functions

The notion of BV function has been generalized in the general framework of Carnot-Carathéodory spaces in [11] and it has been further studied in [24], [29]. A general notion of function of bounded variation has been also given in metric spaces, [44]. We particularize this notion to stratified groups in such a way that the variational measure of a BV function depends only on the restriction of the graded metric to the horizontal subbundle. Throughout the paper, Ω will denote an open subset of $\mathbb{G}$.

Definition 1.2 (Horizontal gradient) *Let $f : \Omega \longrightarrow \mathbb{R}$ be a differentiable function. The* horizontal gradient *of f at $x \in \Omega$ is*

$$\nabla_H f(x) := \sum_{i=1}^{m} X_i f(x)\, X_i \,,$$

where $(X_1, \ldots, X_m)$ is an orthonormal frame of $H\Omega$.

Notice that this definition is independent of the orthonormal frame $(X_1, \ldots, X_m)$.

Definition 1.3 (Horizontal vector fields) The space of smooth sections of $H\Omega$ is denoted by $\Gamma(H\Omega)$. The space $\Gamma_c(H\Omega)$ denotes all the elements of $\Gamma(H\Omega)$ with compact support contained in Ω. Elements of $\Gamma(H\Omega)$ are called *horizontal vector fields.*

Definition 1.4 (H-BV functions) We say that a function $u \in L^1(\Omega)$ is a function of H-*bounded variation* (in short, an H-BV function) if

$$|D_H u|(\Omega) := \sup\left\{ \int_\Omega u \operatorname{div}\phi \, dx \,\middle|\, \phi \in \Gamma_c(H\Omega),\ |\phi| \leq 1 \right\} < \infty\,,$$

where the symbol div denotes the Riemannian divergence. We denote respectively by $BV_H(\Omega)$ and $BV_{loc,H}(\Omega)$ the space of all functions of H-bounded variation and of locally H-bounded variation.

Notice that our definition of H-BV function does not involve any frame of vector fields and by Proposition 1.7 below, the associated variational measure only depends on the restriction of the graded metric to the horizontal subbundle. The same proposition guarantees that our definition can be equivalently stated using horizontal and orthonormal vector fields of $H\Omega$, so it is consistent with the known definitions given in Carnot-Carathéodory spaces, [11], [24], [25], [26], [29]. Indeed, the lack of a homogeneous structure in these spaces forces the use of a particular frame of vector fields. However, following [24], a fixed frame of vector fields induces a nonnegative matrix $A(x)$ (which should be interpreted as a degenerate Riemannian metric) and defines the space $BV_A(\Omega)$ in a way similar to ours.

By Riesz Representation Theorem we get the existence of a nonnegative Radon measure $|D_H u|$ and a Borel section ν of $H\Omega$ such that $|\nu| = 1$ and for any horizontal vector field $\phi \in \Gamma_c(H\Omega)$ we have

$$\int_\Omega u \operatorname{div}\phi = -\int_\Omega \langle \phi, \nu \rangle \, d|D_H u|\,. \tag{3}$$

Some remarks here are in order, since the canonical Riesz theorem deals with linear operators on spaces of continuous functions. In this case the space is $\Gamma_c(H\Omega)$ and we have used the scalar product in each fiber of the tangent spaces (indeed, strictly speaking ν should be thought of as a section of the cotangent bundle). Using local coordinates it is not hard to prove the extension of Riesz theorem we used. The "vector" measure $\nu\,|D_H u|$, acting on bounded Borel sections ϕ of $H\Omega$ as in (3) is denoted by $D_H u$. Splitting $|D_H u|$ in absolutely continuous part $|D_H u|^a$ and singular part $|D_H u|^s$ with respect to the volume measure, we have the Radon-Nikodým decomposition $D_H u = D_H^a u + D_H^s u$, with $D_H^a u = \nu |D_H u|^a$, $D_H^s u = \nu |D_H u|^s$. We denote by $\nabla_H u$ the density of $D_H^a u$ with respect to the volume measure $\mathcal{H}^Q$. Note that

$$\nabla_H u = \frac{\nu\,|D_H u|^a}{\mathrm{vol}}$$

and therefore the Borel map $\nabla_H u$ is a section of $H\Omega$.

Remark 1.5 For a.e. $x \in \Omega$ we have

$$\lim_{r \to 0^+} \frac{|D_H^s u|(U_{x,r})}{r^Q} = 0\,.$$

Indeed, notice that from Radon-Nikodým Theorem we get a Borel subset $N \subset \Omega$ such that $|N| = 0$ and $|D_H^s u|(N^c) = 0$. Therefore, if we had a measurable subset $A \subset \Omega$, with $|A| > 0$ and

$$\limsup_{r \to 0^+} \frac{|D_H^s u|(U_{x,r})}{|U_1|\, r^Q} > 0\,,$$

for any $x \in A$ we would get $A' \subset A$ and $\lambda > 0$ such that $|D_H^s u|(A') \geq \lambda\, |A'| > 0$, see for instance Theorem 2.10.17 and Theorem 2.10.18 of [20]. Hence

$$|D_H^s u|(A' \setminus N) \geq \lambda |A' \setminus N| > 0\,,$$

which contradicts $|D_H^s u|(N^c) = 0$.

Definition 1.6 We define $n_j = \dim V_j$ for any $j = 1, \ldots, \iota$, $m_0 = 0$ and $m_i = \sum_{j=1}^{i} n_j$ for any $i = 1, \ldots, \iota$. We say that $(W_1, \ldots, W_q)$ of $\mathcal{G}$ is an *adapted basis*, if

$$(W_{m_{j-1}+1}, \ldots, W_{m_j}) \tag{4}$$

is a basis of V_j for any $j = 1, \ldots, \iota$. It is easy to realize that any graded metric induces an adapted and orthonormal basis of $\mathcal{G}$.

The following proposition guarantees that the intrinsic notion of H-BV function fits the one given by vector fields. Its proof follows computing the Riemannian divergence $\mathrm{div}\phi$ expressed as $\mathrm{Tr} D\phi$, where D is the standard Riemannian connection, and taking into account that the Riemannian metric is graded.

Proposition 1.7 *For every orthonormal basis $(X_1, \ldots, X_m)$ of $H\Omega$ we have*

$$\mathrm{div}\, \phi = \sum_{i=1}^{m} X_i \phi^i\,,$$

where $\phi \in \Gamma(H\Omega)$ and $\phi = \sum_{i=1}^{m} \phi^i\, X_i$.

PROOF. We complete the horizontal orthonormal frame $(X_1 \ldots, X_m)$ to an orthonormal adapted basis $(X_1 \ldots, X_m, Y_{m+1}, \ldots, Y_q)$, so we are considering a graded metric. By definition of Riemannian divergence we have

$$\mathrm{div}\ \phi = \mathrm{Tr} D\phi = \sum_{i=1}^{m} g(D_{X_i}\phi, X_i) + \sum_{i=m+1}^{q} g(D_{Y_i}\phi, Y_i)$$

where D is the Riemannian connection. We choose $\phi \in \Gamma(H\Omega)$, with the representation $\phi = \sum_{i=1}^m \phi^i X_i$ for some smooth functions ϕ^i. By the properties of the Riemannian connection (using the summation convention) we have

$$g(D_{X_i}\phi, X_i) = g(X_i\phi^l\, X_l + \phi^l D_{X_i}X_l, X_i) = X_i\phi^i + \phi^l g(D_{X_i}X_l, X_i)\,,$$

$$g(D_{X_i}X_l, X_i) = g([X_i, X_l], X_i) + g(D_{X_l}X_i, X_i) = 0\,.$$

The last equation holds because $[X_i, X_j] \in V_2$ is orthogonal to $X_i \in V_1$ and

$$2g(D_{X_l}X_i, X_i) = X_l\,(g(X_i, X_i)) = 0\,.$$

Reasoning as above we get

$$g(D_{Y_i}\phi, Y_i) = g(Y_i\phi^l\, X_l + \phi^l D_{Y_i}X_l, Y_i) = \phi^l\, g(D_{Y_i}X_l, Y_i) = 0\,,$$

and this completes the proof. $\square$

Definition 1.8 We say that a Borel set $E \subset \Omega$ has *H-finite perimeter* in Ω if

$$P_H(E, \Omega) := \sup\left\{ \int_E \mathrm{div}\phi\; dx \,\middle|\, \phi \in \Gamma_c(H\Omega),\ |\phi| \leq 1 \right\} < \infty\,.$$

If $\Omega = \mathbb{G}$ we simply say that E has H-finite perimeter.

By the previous discussion, $P_H(E, A) = |D_H \mathbf{1}_E|(A)$ is the restriction to open sets A of a finite Borel measure in Ω. It is clear that if E has H-finite perimeter in Ω and $\mathbf{1}_E \in L^1(\Omega)$, then $\mathbf{1}_E \in BV_H(\Omega)$ and $|D_H\mathbf{1}_E|(F) = P_H(E, F)$, for any Borel set $F \subset \Omega$.

1.3 Approximately regular functions

Here we introduce some weak notions of limit and differential for Borel functions on stratified groups.

Definition 1.9 (Approximate limit) We say that a function $u \in L^1_{loc}(\Omega, \mathbb{R}^m)$ has an *approximate limit* $\lambda \in \mathbb{R}^m$ at $x \in \Omega$ if

$$\lim_{r \to 0^+} \fint_{U_{x,r}} |u(y) - \lambda|\, dy = 0\,.$$

If u does not have an approximate limit at x we say that x is an *approximate discontinuity point* and we denote by S_u the Borel set of all these points, namely the *approximate discontinuity set.*

It is clear that the approximate limit is uniquely defined and that it does not depend on the representative element of u; it will be denoted by $\tilde{u}(x)$. We call the points in $\Omega \setminus S_u$ *approximate continuity points* of u. Since stratified groups are doubling spaces we have that S_u is negligible and $u(x) = \tilde{u}(x)$ for a.e. $x \in \Omega$, see for instance Theorem 2.9.8 in [20] or Theorem 14.15 in [30].

Definition 1.10 We say that $x \in \Omega$ is a *density point* of a Borel set E if

$$\lim_{r \to 0^+} \frac{|U_{x,r} \setminus E|}{|U_{x,r}|} = 0\,.$$

We denote by $\mathcal{I}(E)$ the set of all density points of A. The *essential boundary* of E is defined as $\partial^* E = \Omega \setminus \Big(\mathcal{I}(E) \cup \mathcal{I}(\Omega \setminus E)\Big)$.

Note that density points of A are approximate continuity points of the locally summable map $\mathbf{1}_A$, therefore the set $A \setminus \mathcal{I}(A)$ is negligible with respect to the volume measure. There is a weaker, and more canonical, definition of approximate limit (see for instance [20]). Let us consider a Borel function $u : \Omega \longrightarrow \mathbb{R}$, $x \in \Omega$ and $\lambda \in \mathbb{R}$. We say that λ is the approximate limit of u at x if for any $\varepsilon > 0$ we have $x \in \mathcal{I}(\{z \in \Omega \mid |u(z) - \lambda| < \varepsilon\})$. The approximate limit λ is uniquely defined and it is denoted by $\operatorname{ap}\lim_{z \to x} u(z)$. Note that $x \in \Omega \setminus S_u$ implies $\operatorname{ap}\lim_{z\to x} u(z) = \tilde{u}(x)$, but the converse is not true in general, therefore we use the same word (but a different notation) for the two concepts. Moreover, for locally bounded functions u there is a complete equivalence: $\operatorname{ap}\lim_{z\to x} u(z) = \lambda$ implies $x \in \Omega \setminus S_u$ and $\lambda = \tilde{u}(x)$.

Definition 1.11 (G-Linear maps) We say that $L : \mathbb{G} \longrightarrow \mathbb{R}$ is a G-*linear map* if it is a group homomorphism such that

$$L(\delta_r z) = r\, L(z)$$

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

Remark 1.12 It is not difficult to see that G-linear maps are indeed linear on the space $\mathcal{G}$, see [37], so via the Riemannian metric we can represent any G-linear map $L : \mathbb{G} \longrightarrow \mathbb{R}$ with a unique vector of $v \in \mathcal{G}$ as $L(x) = \langle v, \ln x\rangle$ for any $x \in \mathbb{G}$. Furthermore, by the homogeneity property of L, the vector v is horizontal. Conversely, for any $v \in H$ the map $x \longrightarrow \langle v, \ln x\rangle$ is G-linear.

Throughout the paper we will use the notation v^* to indicate the map $x \longrightarrow \langle v, \ln x\rangle$ for any $x \in \mathbb{G}$. Then, all G-linear maps are representable as v^*, for some $v \in H$.

Definition 1.13 We define the *homogeneous norm* of a G-linear map $L : \mathbb{G} \longrightarrow \mathbb{R}$

$$\|L\| = \sup_{d(z)=1} |L(z)|\,.$$

This implies that for any G-linear map $L : \mathbb{G} \longrightarrow \mathbb{R}$ we have

$$|L(z)| \le \|L\|\, d(z)\,, \tag{5}$$

for any $z \in \mathbb{G}$.

Definition 1.14 (Differential) Consider a Borel set $A \subset \Omega$, $x \in A \cap \mathcal{I}(A)$ and $u : A \longrightarrow \mathbb{R}$. We say that u is *differentiable* at x if there exists a G-linear map $L : \mathbb{G} \longrightarrow \mathbb{R}$ such that

$$\lim_{A \ni z \to x} \frac{|u(z) - u(x) - L(x^{-1}z)|}{d(z,x)} = 0\,. \tag{6}$$

We denote the differential of u at x by $d_H u(x)$.

The uniqueness of L easily follows by the fact that x is a density point of A.

Definition 1.15 (Approximate differential) Consider $u \in L^1_{loc}(\Omega)$ and a point $x \in \Omega \setminus S_u$. We say that u is *approximately differentiable* at x if there exists a G-linear map $L : \mathbb{G} \longrightarrow \mathbb{R}$ such that

$$\lim_{r \to 0^+} ⨍_{U_{x,r}} \frac{|u(z) - \tilde{u}(x) - L(x^{-1}z)|}{r}\, dz = 0\,. \tag{7}$$

The map L is uniquely defined, it is denoted by $d_H u(x)$ and it is called the *approximate differential* of u at x.

Remark 1.16 We have used the same symbol to denote both differentials in Definitions 1.14 and 1.15. This slight abuse of notation is justified by the fact that differentiability implies approximate differentiability. An even weaker notion of approximate differentiability can be given in the spirit of [20], saying that the approximate differential of a map $u : A \subset \mathbb{G} \longrightarrow \mathbb{R}$ at $x \in \mathcal{I}(A)$ is the unique G-linear map $L : \mathbb{G} \longrightarrow \mathbb{R}$ such that

$$\operatorname{ap}\lim_{y \to x} \frac{u(y) - u(x) - L(x^{-1}y)}{d(x,y)} = 0\,. \tag{8}$$

We point out that the approximate differentiability implies the existence of the approximate limit (8), as it will be proved in Proposition 2.1, but already in the Euclidean case the converse is not true, see for instance Remark 3.66 of [5].

An extension of the classical Rademacher's Theorem holds, see [37], [49], [61].

Theorem 1.17 *If A is a Borel subset of $\mathbb{G}$ and $u : A \longrightarrow \mathbb{R}$ is a Lipschitz map, then u is differentiable at a.e. point of A.*

1.4 Some general facts

In this subsection we recall some important general theorems we will use throughout the paper. We start introducing the coarea formula for H-BV functions, see for instance [24], [29].

Theorem 1.18 (Coarea formula) *For any $u \in BV_H(\Omega)$, the following formula holds*

$$|D_H u|(\Omega) = \int_{\mathbb{R}} P_H\left(\{x \in \Omega \mid u(x) > t\}, \Omega\right) dt\,. \tag{9}$$

A crucial tool in the analysis on sub-Riemannain groups is the Poincaré inequality. This theorem holds for general vector fields that satisfy the so-called *Hörmander condition*, see [31].

Theorem 1.19 (Poincaré inequality) *There exists a constant $C > 0$ such that for any C^∞ smooth map $w : \Omega \longrightarrow \mathbb{R}$ and any ball $U_{x,r} \Subset \Omega$, we have*

$$\int_{U_{x,r}} |w(z) - w_{U_{x,r}}|\, dz \leq C\, r\, |D_H w|(U_{x,r})\,. \tag{10}$$

Now, we state an important theorem about the smooth approximation of H-BV functions, see either Theorem 2.2.2 of [24] or Theorem 1.14 of [29].

Theorem 1.20 (Smooth approximation) *Let $u : \Omega \longrightarrow \mathbb{R}$ an H-BV function. Then there exists a sequence (u_k) of C^∞-smooth functions such that*

1. *$u_k \longrightarrow u$ in $L^1(\Omega)$;*
2. *$|D_H u_k|(\Omega) \longrightarrow |D_H u|(\Omega)$.*

In view of (10) and Theorem 1.20 we obtain the following theorem.

Theorem 1.21 *Let $w : \Omega \longrightarrow \mathbb{R}$ be a locally H-BV function. Then for any ball $U_{x,r} \Subset \Omega$ we have*

$$\int_{U_{x,r}} |w(z) - w_{U_{x,r}}|\, dz \leq C\, r\, |D_H w|(U_{x,r})\,. \tag{11}$$

An important consequence of (11) is the local isoperimetric inequality for sets of H-finite perimeter.

Theorem 1.22 (Isoperimetric estimate) *Let E be a set of H-finite perimeter. Then for any $U_{x,r} \subset \mathbb{G}$ we have*

$$\min\{|U_{x,r} \cap E|, |U_{x,r} \setminus E|\} \leq \; C\, r\, P_H(E, U_{x,r})\,. \tag{12}$$

It is a general fact that the Poincaré inequality (10) implies a Sobolev-Poincaré inequality, see for instance Theorem 2 of [22] or Theorem 1.15 (II) of [29]. This inequality can be extended to H-BV functions via Theorem 1.20.

Theorem 1.23 *Let $w : \Omega \longrightarrow \mathbb{R}$ be a locally H-BV function. Then*

$$\left(\fint_{U_{x,r}} |w(z) - w_{U_{x,r}}|^{1^*}\right)^{1/1^*} \leq C\, r\, \frac{|D_H w|(U_{x,r})}{|U_{x,r}|}\,, \tag{13}$$

for any $U_{x,r} \Subset \Omega$, where $1^* = Q/(Q-1)$.

The following theorem is a consequence of Theorem 1.28 and Theorem 1.15 of [29].

Theorem 1.24 (Compact embedding) *Let U denote a Carnot-Caratheodory ball of $\mathbb{G}$. Then for any $q \in [1, 1^*[$ the inclusion $BV_H(U) \hookrightarrow L^q(U)$ is compact.*

Proposition 1.25 *Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a Lipschitz map which vanishes at the origin and let $u \in [BV_H(\Omega)]^n$. Then $f \circ u : \Omega \longrightarrow \mathbb{R}$ is a H-BV function and*

$$|D_H(f \circ u)| \leq L \sum_{l=1}^{n} |D_H u^l|\,, \tag{14}$$

where L is the Lipschitz constant of f.

Definition 1.26 (Maximal operator) We consider a nonnegative Radon measure ν in Ω. For each $r > 0$ the *restricted maximal function* of ν is defined as follows

$$M_r\nu(x) := \sup\left\{\frac{\nu(U_{x,t})}{|U_{x,t}|}:\ 0 < t < r,\, U_{x,t} \subset \Omega\right\} \qquad x \in \Omega\,.$$

The *maximal function* of ν is defined as $M\nu(x) = \sup_{r>0} M_r\nu(x)$. If the measure ν is induced by a locally integrable function $f : \Omega \longrightarrow \mathbb{R}$, we define analogously

$$M_r f(x) := \sup\left\{ ⨍_{U_{x,t}} |f(y)|\, dy :\ 0 < t < r,\, U_{x,t} \subset \Omega\right\}$$

and $Mf(x) = \sup_{r>0} M_r f(x)$.

It is well known that the maximal operator is (1,1)-weakly continuous, i.e. there exists a constant $C > 0$ such that

$$|\{x \in E \mid M\nu(x) > t\}| \leq \frac{C}{t}\,\nu(E)\,, \tag{15}$$

for any Borel set $E \subset \Omega$ and any $t > 0$, see for instance [8]. Inequality (15) implies that if ν is a finite measure, then $M\nu$ is finite a.e. in Ω. Following the terminology of authors in [23], a general "representation formula" holds in metric spaces of homogeneous type which satisfy the Poincaré inequality, [23]. In stratified groups the same formula was first proved in [35]. We state this result in the following theorem.

Theorem 1.27 (Representation formula) *There exists a dimensional constant $C > 0$ such that*

$$|\tilde{w}(x) - w_{U_{x,r}}| \leq C \int_{U_{x,r}} \frac{1}{d(z,y)^{Q-1}}\, d|D_H w|(z). \tag{16}$$

for any $w \in BV_H(U_{x,r})$ and $x \notin S_w$.

Remark 1.28 By virtue of Fubini's Theorem we have

$$\int_{U_{x,r}} \frac{1}{d(z,y)^{Q-1}}\, d|D_H w|(z) = (Q-1)\int_0^{+\infty} \frac{|D_H w|(U_{x,r}\cap U_{x,t})}{t^Q}\, dt\,,$$

so that (16) becomes

$$|\tilde{w}(x) - w_{U_{x,r}}| \le C\left[(Q-1)\int_0^r \frac{|D_H w|(U_{x,t})}{t^Q} dt + \frac{|D_H w|(U_{x,r})}{r^{(Q-1)}}\right]. \tag{17}$$

Furthermore, in the case $\tilde{w}(x)=0$ the monotonicity of the right hand side of (16) (and of (17) as well) with respect to r implies

$$|M_r w(x)| \le C\left[(Q-1)\int_0^r \frac{|D_H w|(U_{x,t})}{t^Q} dt + \frac{|D_H w|(U_{x,r})}{r^{(Q-1)}}\right]. \tag{18}$$

2 First order differentiability

In this section we prove the approximate differentiability of H-BV functions and we give an estimate on the size of the approximate discontinuity set S_u. Our strategy will be to prove first that BV functions are a.e. differentiable in a weaker sense; then, a bootstrap argument based on Poincaré inequality leads us to the approximate differentiability. We mention that the validity of the following inequality

$$\int_{U_{x,r}} \frac{|u(x)-u(z)|}{d(x,z)}\, dz \le C \int_0^1 \frac{|D_H u|(U_{x,\lambda r t})}{t^Q}\, dt$$

with $\lambda, C > 0$ absolute constants, implies a slightly stronger approximate differentiability via classical methods described in [5]. In the case of the Heisenberg group this approach is followed in [56]. However, on arbitrary stratified groups the above inequality seems to be an open question.

In the following propositions we consider weaker differentiability properties.

Proposition 2.1 *Let $u:\Omega \longrightarrow \mathbb{R}$ be a Borel map. Then the following statements are equivalent:*

1. *for a.e. $x\in\Omega$ there exists a G-linear map $L_x:\mathbb{G}\longrightarrow\mathbb{R}$ such that*

$$\operatorname{ap}\lim_{y\to x} \frac{u(y)-u(x)-L_x(x^{-1}y)}{d(x,y)} = 0\,, \tag{19}$$

2. *u is countably Lipschitz up to a negligible set, i.e. there exists a countable family of Borel subsets $\{A_i \mid A_i\subset\Omega,\ i\in\mathbb{N}\}$ such that for each $i\in\mathbb{N}$ the restriction $u_{|A_i}$ is a Lipschitz map and we have*

$$\left|\Omega\setminus\bigcup_{i\in\mathbb{N}} A_i\right| = 0\,.$$

Furthermore 1. and 2. hold if u is approximately differentiable a.e. in Ω.

PROOF. We start proving that property 1 is implied by the approximate differentiability. Assume that u is approximately differentiable at $x \in \Omega$ with approximate differential $d_H u(x)$. Let us fix $\varepsilon > 0$ and consider the set

$$E_{x,\rho} = \left\{ z \in U_{x,\rho} \Big| |u(z) - u(x) - d_H u(x)(x^{-1}z)| > \varepsilon\, d(x,z) \right\} .$$

In order to get (19) we have to prove that

$$\lim_{\rho \to 0^+} |E_{x,\rho}|\, \rho^{-Q} = 0 . \tag{20}$$

Let us define the maps $T_{x,\rho}(z) = \delta_{1/\rho}(x^{-1}z)$ and

$$R_{x,\rho}(z) = \frac{|u(x\delta_\rho z) - u(x) - d_H u(x)(\delta_\rho z)|}{\rho} ,$$

observing that

$$T_{x,\rho}(E_{x,\rho}) = \left\{ y \in U_1 \Big| \; R_{x,\rho}(y) > \varepsilon\, d(y) \right\} := A_\rho .$$

Hence we have $|E_{x,\rho}|\rho^{-Q} = |A_\rho|$ and (20) follows if we prove that

$$\lim_{\rho \to 0^+} |A_\rho| = 0 . \tag{21}$$

By hypothesis, making a change of variable we get

$$\lim_{\rho \to 0^+} \fint_{U_{x,\rho}} \frac{|u(z) - u(x) - d_H u(x)(x^{-1}z)|}{\rho}\, dz = \lim_{\rho \to 0^+} \fint_{U_1} R_{x,\rho}(z)\, dz = 0 . \tag{22}$$

For each $t \in]0,1[$ we have

$$\int_{A_\rho \setminus U_t} R_{x,\rho} \geq \; |A_\rho \setminus U_t|\, \varepsilon\, t ,$$

so in view of (22) we obtain $|A_\rho \setminus U_t| \longrightarrow 0$ as $\rho \to 0^+$. It follows that

$$\limsup_{\rho \to 0^+} |A_\rho| \leq \limsup_{\rho \to 0} |A_\rho \setminus U_t| + |U_t| = |U_t| .$$

Finally, letting $t \to 0$ equation (21) follows, so statement 1 is proved. The fact that statement 1 implies statement 2 can be proved arguing as in Theorem 3.1.8 of [20], see also Theorem 6 in [60]. Now, let us prove that statement 2 implies 1. By Theorem 1.17 we know that $u_{|A_i}$ is a.e. differentiable. Let us indicate by $\mathcal{D}_u(A_i)$ the subset of $\mathcal{I}(A_i)$ where $u_{|A_i}$ is differentiable in A_i. Clearly, we have

$$\Big| \Omega \setminus \bigcup_{i \in \mathbb{N}} \mathcal{D}_u(A_i) \Big| = 0 .$$

Consider $x \in \mathcal{D}_u(A_i)$ and choose $\varepsilon > 0$. Then there exists $\delta > 0$ such that for any $z \in A_i \cap U_{x,\delta}$ we get

$$R(z) = \frac{|u(z) - u(x) - L(x^{-1}z)|}{d(z,x)} < \varepsilon\,,$$

with $L = du_{|A_i}(x)$. From the last inequality it follows that

$$U_{x,r} \cap \{z \in \Omega \mid R(z) \geq \varepsilon\} \subset U_{x,r} \setminus A_i$$

for any $r \leq \delta$. Hence we get

$$\limsup_{r \to 0^+} \frac{|U_{x,r} \cap \{z \in \Omega \mid R(z) \geq \varepsilon\}|}{|U_{x,r}|} \leq \limsup_{r \to 0^+} \frac{|U_{x,r} \setminus A_i|}{|U_{x,r}|} = 0\,,$$

in view of the fact that x is a density point of A_i. □

Theorem 2.2 *Let $u : \Omega \longrightarrow \mathbb{R}$ be a locally H-BV function. Then, u is approximately differentiable a.e. in Ω and the differential corresponds to the density of the absolutely continuous part of $D_H u$, i.e. $d_H u(x) = \nabla_H u(x)^*$ for a.e. $x \in \Omega$.*

Proof. We first prove that u is countably Lipschitz up to a negligible set. Let us fix $t > 0$ and define the open subset

$$\Omega_t = \{z \in \Omega \mid \mathrm{dist}(z, \Omega^c) > t\}\,.$$

We want to prove that u is countably Lipschitz on Ω_t. We cover Ω_t with a countable union of open balls $\{P_j \mid j \in \mathbb{N}\}$ with center in Ω_t and radius $t/4$. Let us consider $j \in \mathbb{N}$ and two approximate continuity points $x, y \in P_j$. By the well known technique of the "telescopic estimate" (see for instance Theorem 3.2 of [30]) we obtain

$$|\tilde{u}(x) - \tilde{u}(y)| \leq c\, d(x,y)\ \left(M_{2d(x,y)}|D_H u|(x) + M_{2d(x,y)}|D_H u|(y)\right) \tag{23}$$

with $c = (2^{Q+2} + 2)\, C$. Now, let us consider the decomposition

$$P_j = N_j \cup \left(\bigcup_{l \in \mathbb{N}} E_{jl}\right)$$

where E_{jl} is the Borel set of all approximate continuity points $z \in P_j$ such that $M|D_H u|(z) \leq l$ and $N_j = S_u \cup \{z \in P_j \mid M|D_H u|(z) = +\infty\}$. Then N_j is a negligible set and by (23) it follows that

$$|\tilde{u}(x) - \tilde{u}(y)| \leq 2\,c\,l\ d(x,y) \qquad \forall x,\, y \in E_{jl}$$

and any $j, l \in \mathbb{N}$. This gives the countably Lipschitz property of u in Ω_t. Observing that Ω is a countable union of $\Omega_{1/k}$, with $k \in \mathbb{N} \setminus \{0\}$ we obtain the countably Lipschitz property of u in Ω. In view of Proposition 2.1 the countably Lipschitz property yields the existence of a G-linear map $L_x : \mathbb{G} \longrightarrow \mathbb{R}$ such that for a.e. $x \in \Omega$ we have

$$\operatorname{ap}\lim_{y\to x} \frac{u(y)-\tilde{u}(x)-L_x(x^{-1}y)}{d(x,y)} = 0\,. \tag{24}$$

In order to prove the a.e. approximate differentiability, we select a point $x \in \Omega \setminus (S_u \cup S_{\nabla_H u})$ such that (24) holds and

$$\lim_{r\to 0^+} \frac{|D_H^s u|(U_{x,r})}{r^Q} = 0\,. \tag{25}$$

In view of Remark 1.5 the set of points which do not satisfy (25) is negligible, so the set of selected points with all the above properties has full measure in Ω. We fix $\varepsilon > 0$ and consider the set

$$F_{x,r} = \left\{ y \in U_{x,r} \mid |u(y)-\tilde{u}(x)-L_x(x^{-1}y)| > \varepsilon d(x,y) \right\}\,,$$

observing that

$$Z_{x,r} := \delta_{1/r}(x^{-1}F_{x,r}) = \left\{ z \in U_1 \mid \frac{|u(x\delta_r z)-\tilde{u}(x)-L_x(\delta_r z)|}{r} > \varepsilon \right\}\,. \tag{26}$$

In view of (25) we have that $|F_{x,r}|\, r^{-Q} \longrightarrow 0$ as $r \to 0^+$, therefore

$$|Z_{x,r}| = |\delta_{1/r}(x^{-1}F_{x,r})| = r^{-Q}|x^{-1}F_{x,r}| = r^{-Q}|F_{x,r}| \longrightarrow 0 \quad \text{as} \quad r \to 0^+\,. \tag{27}$$

Now, we consider the difference $S_x = \nabla_H u(x)^* - L_x$ and define the maps

$$v(y) = u(y) - \tilde{u}(x) - \nabla_H u(x)^*(x^{-1}y)\,,$$
$$w_{x,r}(z) = \frac{v(x\delta_r z) + S_x(\delta_r z)}{r} = v_{x,r}(z) + S_x(z)\,,$$

observing that $\tilde{v}(x) = 0$, $|D_H v_{x,r}|(U_1) \longrightarrow 0$ as $r \to 0^+$ and

$$Z_{x,r} = \{ z \in U_1 \mid |w_{x,r}(z)| > \varepsilon \}\,.$$

Thus, by (27) it follows that $w_{x,r} \to 0$ in measure as $r \to 0^+$. Since $v_{x,r}$ is an H-BV function, we can apply Poincaré inequality (11), getting

$$\int_{U_1} |v_{x,r}(z) - m_{x,r}|\, dz \le C\, |D_H v_{x,r}|(U_1) \longrightarrow 0 \quad \text{as} \quad r \to 0^+\,, \tag{28}$$

where $m_{x,r} = ⨍_{U_1} v_{x,r}$. Then, we obtain

$$\int_{U_1} |w_{x,r}(z) - m_{x,r} - S_x(z)|\, dz \longrightarrow 0 \quad \text{as} \quad r \to 0^+\,.$$

It follows that $m_{x,r} + S_x$ converges to zero in measure on U_1 as $r \to 0^+$. This easily implies that $m_{x,r} \to 0$ and $S_x = 0$. So, $\nabla_H u(x)^* = L_x$ and in view of (28) we get

$$\frac{1}{r} ⨍_{U_{x,r}} |v(z)|\, dz = ⨍_{U_1} |v_{x,r}(z)|\, dz \longrightarrow 0 \quad \text{as} \quad r \to 0^+\,,$$

which proves the approximate differentiability of u at x with $d_H u(x) = \nabla_H u(x)^*$. $\square$

2.1 Size of S_u

In this subsection we prove that the approximate discontinuity set of an H-BV function is contained in a countable union of essential boundaries of sets with H-finite perimeter, up to a $\mathcal{H}^{Q-1}$-negligible set, see inequality (31). Thus, whenever a rectifiability theorem for sets of H-finite perimeter in stratified groups holds, we immediately get the countably rectifiability of S_u for any H-BV function u (here the notion of rectifiability must be properly understood in intrinsic terms, see [25]). For instance, in the Heisenberg group a rectifiability theorem holds, i.e. the reduced boundary of any set of H-finite perimeter is contained in a countable union of $\mathbb{H}$-regular surfaces up to $\mathcal{H}^{Q-1}$-negligible sets, [25]. Recently this result has been extended to any 2-step stratified group, [27].

The following two lemmas are crucial to prove that the approximate singular set of an H-BV function is countably $\mathcal{H}^{Q-1}$-finite. They are a version of Lemma 3.74 and Lemma 3.75 of [5] adapted for stratified groups. We give the proof of them in order to emphasize the main steps, where relevant theorems of Analysis for stratified groups are needed. Furthermore, Lemma 3.74 in [5] is proved using the Besicovitch Covering Theorem. Due to the fact that this theorem may fail in a general stratified group, we show another simpler way to prove it, adopting the Vitali Covering Theorem for doubling spaces.

Lemma 2.3 *Let (E_h) be a sequence of measurable subsets of Ω, such that $|E_h| \longrightarrow 0$ and $P_H(E_h,\Omega) \longrightarrow 0$ as $h \to \infty$. Then, for any $\alpha > 0$ we have*

$$\mathcal{H}^{Q-1}\left(\bigcap_{h=1}^{\infty}\{x\in\Omega \mid \theta^*_Q(E_h,x)\ge\alpha\}\right)=0\,.$$

PROOF. Let us fix $\delta > 0$ and $\alpha \in]0,1[$. We consider a Borel set $E \subset \Omega$ such that $|E| < |U_1|\,\alpha\,\delta^Q/2$ and define

$$E^\alpha=\{x\in\Omega \mid \theta^*_Q(E,x)\ge\alpha\}\,.$$

For any $x \in E^\alpha$ the estimate

$$\frac{|U_{x,\delta}\cap E|}{|U_{x,\delta}|}\le\frac{|E|}{|U_1|\,\delta^Q}<\frac{\alpha}{2}$$

implies the existence of a radius $r_x \in]0,\delta[$ such that $|U_{x,r_x}\cap E| = \alpha|U_{x,r_x}|/2$. Thus, in view of (12) we get

$$\frac{\alpha}{2}\,|U_1|\,r_x^Q=|U_{x,r_x}\cap E|\le C\,r_x\,P_H(E,U_{x,r_x}). \tag{29}$$

Now, let us consider an open subset $\Omega' \Subset \Omega$, with $0 < \delta < \mathrm{dist}(\Omega',\partial\Omega)$ and $|E| \le |U_1|\,\alpha\,\delta^Q/2$. Using a well known covering theorem for the family $\{U_{x,r_x} \mid x\in\Omega'\cap E^\alpha\}$ (Corollary 2.8.5 in [20]), we get a countable disjoint subfamily

$$\{B_j \mid B_j=U_{x_j,r_{x_j}},\, j\in\mathbb{N}\}$$

such that $\Omega' \cap E^\alpha \subset \bigcup_{h=1}^\infty 5B_j$, where $5B_j$ is the ball of center x_j and radius $5r_{x_j}$. Therefore, the estimate (29) implies

$$\mathcal{H}^{Q-1}_{10\delta}(\Omega' \cap E^\alpha) \leq 5^{Q-1} \sum_{i=1}^\infty r_{x_i}^{Q-1} \leq \frac{2C5^{Q-1}}{\alpha\,|U_1|} \sum_{i=1}^\infty P_H(E, B_i) \leq \frac{C'}{\alpha} P_H(E, \Omega)\,.$$

We fix the sequence $\delta_i = (2|E_i|/|U_1|\alpha)^{1/Q}$, observing that $\delta_i \leq \delta$ for i large, hence

$$\mathcal{H}^{Q-1}_{10\delta_i}\left(\Omega' \cap \bigcap_{h=1}^\infty E_h^\alpha\right) \leq \frac{C'}{\alpha} P_H(E_i, \Omega)\,.$$

Thus, letting first $\delta_i \to 0^+$ and then $\Omega' \uparrow \Omega$ the conclusion follows. □

Lemma 2.4 *Let $u : \Omega \longrightarrow \mathbb{R}$ be an H-BV function. Then, the set*

$$L = \left\{ x \in \Omega \mid \limsup_{r \to 0^+} \fint_{U_{x,r}} |u(y)|^{1^*}\, dy = \infty \right\}$$

is $\mathcal{H}^{Q-1}$-negligible, where $1^* = Q/(Q-1)$.

Proof. In view of Proposition 1.25 we can assume that $u \geq 0$ (replacing u by $|u|$). We define the set

$$D = \left\{ y \in \Omega \mid \limsup_{r \to 0^+} \frac{|D_H u|(U_{x,r})}{r^{Q-1}} = \infty \right\},$$

observing that by Theorem 2.10.17 and Theorem 2.10.18 of [20] and the fact that $|D_H u|(\Omega) < \infty$, we have $\mathcal{H}^{Q-1}(D) = 0$. For any integer $h \in \mathbb{N}$ we can choose $t_h \in]h, h+1[$ such that

$$P_H(E_{t_h}, \Omega) \leq \int_h^{h+1} P_H(E_t, \Omega)\, dt\,,$$

where $E_t = \{x \in \Omega \mid u(x) > t\}$, for each $t \geq 0$. By (9) we have

$$\sum_{h=0}^\infty P_H(E_{t_h}, \Omega) \leq \int_0^\infty P_H(E_t, \Omega) dt = |D_H u|(\Omega) < \infty\,.$$

Then, we apply Lemma 2.3 to the sequence (E_{t_h}) with $\alpha = 1$, getting

$$\mathcal{H}^{Q-1}\Big(\bigcap_{h=0}^\infty F_h\Big) = 0\,,$$

where we have defined $F_h = \{x \in \Omega \mid \theta^*_Q(E_{t_h}, x) = 1\}$. We want to prove that $L \subset D \cup \bigcap_{h=0}^\infty F_h$. In order to do that, we consider $x \notin D \cup \bigcap_{h=0}^\infty F_h$ and we prove that $x \notin L$. We define the constants $c_{x,r}$ to be the mean value of u on $U_{x,r}$ and apply the Sobolev-Poincaré inequality (13) obtaining

$$\fint_{U_{x,r}} |u(z) - c_{x,r}|^{1^*}\, dz \le C \left(\frac{|D_H u|(U_{x,r})}{r^{Q-1}} \right)^{1^*} . \tag{30}$$

Notice that if $\limsup_{r\to 0^+} c_{x,r} < \infty$ then (30) implies $x \notin L$. Then, reasoning by contradiction, suppose that there exists a sequence c_{x,r_j} such that $r_j \to 0^+$ and $c_{x,r_j} \to \infty$ as $j \to \infty$. We define the function $v_j(y) = u(x\delta_{r_j} y) - c_{x,r_j}$, observing that $|D_H v_j|(U_1) = |D_H u|(U_{x,r_j})\, r_j^{1-Q}$. Since the sequence $|D_H v_j|(U_1)$, $j \in \mathbb{N}$, is bounded, Theorem 1.24 implies the convergence a.e. of (v_j) to a function $w \in L^1(U_1)$, possibly extracting a subsequence. As a consequence, $u(x\delta_{r_j} y) \to +\infty$ as $j \to \infty$ for a.e. $y \in U_1$, and therefore

$$|U_1| = \lim_{j\to\infty} |\{z \in U_1 \mid u(x\delta_{r_j} z) > t_h\}| = \lim_{j\to\infty} \frac{|\{y \in U_{x,r_j} \mid u(y) > t_h\}|}{r_j^Q} .$$

This implies $x \in \bigcap_{h=1}^\infty F_h$, contradicting the initial assumption. □

Theorem 2.5 *Let $u : \Omega \longrightarrow \mathbb{R}$ be an H-BV function. Then the approximate discontinuity set S_u is a countable union of sets with finite $\mathcal{H}^{Q-1}$ measure.*

Proof. We define $E_t = \{x \in \Omega \mid u(x) > t\}$ for $t \in \mathbb{R}$. By coarea formula (9) the set of numbers $t \in \mathbb{R}$ such that $P_H(E_t, \Omega) < \infty$ has full measure in $\mathbb{R}$, then it is possible to consider a countable dense subset $D \subset \mathbb{R}$ such that $P_H(E_t, \Omega) < \infty$ for any $t \in D$. Notice that from general results about sets of finite perimeter in Ahlfors metric spaces, see Theorem 4.2 in [4], we have that $\mathcal{H}^{Q-1}(E_t) < \infty$ for any $t \in D$. So, in view of Lemma 2.4 it suffices to prove the following inclusion

$$S_u \setminus L \subset \bigcup_{t\in D} \partial^* E_t , \tag{31}$$

where $L = \{x \in \Omega \mid \limsup_{r\to 0^+} \fint_{U_{x,r}} |u(y)|^{1^*}\, dy = \infty\}$. Let us consider a point $x \notin \bigcup_{t\in D} \partial^* E_t \cup L$. Then, for any positive t we have

$$\limsup_{r\to 0^+} \frac{|E_t|}{|U_{x,r}|} \le \frac{1}{t} \limsup_{r\to 0^+} \fint_{U_{x,r}} |u| , \tag{32}$$

hence, for any $t \in D$ sufficiently large such that the right hand side of (32) is less than one, it must be $\theta_Q^*(E_t, x) = 0$. Analogously, for $t \in D \cap (-\infty, 0)$ with $|t|$ large enough we have $\theta_Q^*(E_t^c, x) = 0$, so $\theta_Q^*(E_t, x) = 1$. This means that

$$\tau = \sup\{t \in D \mid \theta_Q^*(E_t, x) = 1\}$$

is a real number. Since D is dense in $\mathbb{R}$ and $t \longrightarrow |E_t|$ is a decreasing map it follows that $\theta_Q^*(E_t, x) = 0$ for any $t > \tau$ and $\theta_Q^*(E_t^c, x) = 0$ for any $t < \tau$. By virtue of this fact it follows that for any $\varepsilon > 0$ we have $|F_\varepsilon \cap U_{x,r}| = o(r^Q)$, where $F_\varepsilon = \{y \in \Omega \mid |u(y) - \tau| > \varepsilon\}$. Finally

$$\limsup_{r\to 0^+} \fint_{U_{x,r}} |u(y) - \tau|\, dy \le \varepsilon + \limsup_{r\to 0^+} \frac{1}{|U_{x,r}|} \int_{F_\varepsilon} |u(y) - \tau|\, dy$$

$$\le \varepsilon + \limsup_{r\to 0^+} \left(\frac{|F_\varepsilon|}{|U_{x,r}|}\right)^{1/Q} \left(\fint_{U_{x,r}} |u(y) - \tau|^{1^*} dy\right)^{1/1^*} = \varepsilon\,.$$

Letting $\varepsilon \to 0^+$, we obtain that $x \notin S_u$, so the inclusion (31) is proved. $\square$

3 Higher order differentiability of H-BV^k functions

In this section we study the differentiability properties of maps with higher order H-bounded variation. The method to accomplish this study is substantially different from that one employed for H-BV functions. Particularly interesting is the case of maps with second H-bounded variation, in view of potential applications to the theory of convex functions on stratified groups (see [36]).

We begin with the definition of high order H-BV function.

Definition 3.1 Let us fix an orthonormal frame $(X_1, \ldots, X_m)$ of $H\Omega$. By induction on $k \ge 2$ and taking into account the definition of H-BV with $k = 1$, we say that a Borel map $u : \Omega \longrightarrow \mathbb{R}$ has *H-bounded k-variation* (in short, H-BV^k) if for any $i = 1, \ldots, m$ the distributional derivatives $X_i u$ are representable by functions with H-bounded $(k-1)$-variation. We denote by $BV^k_H(\Omega)$ the space of all H-BV^k functions.

Remark 3.2 The notion of H-BV^k function does not depend on the choice of the orthonormal frame $(X_1, \ldots, X_m)$.

Now, we review some basics about polynomials on stratified groups. We will refer to Chapter 1.C of [21].

Definition 3.3 We say that a function $P : \mathbb{G} \longrightarrow \mathbb{R}$ is a *polynomial on* $\mathbb{G}$ if the composition $P\circ \exp$ is a polynomial on $\mathcal{G}$.

In order to define the homogeneous degree of a polynomial $P : \mathbb{G} \longrightarrow \mathbb{R}$ we need to fix an adapted basis $(W_1, \ldots, W_q)$ and its dual $(\eta_1, \ldots, \eta_q)$. Clearly the degree $d_i \in \mathbb{N} \setminus \{0\}$ of W_i is determined by the unique relation $W_i \in V_{d_i}$. Let us consider the coordinate system $(x_1, \ldots, x_q)$ on $\mathbb{G}$, where $x_i = \eta_i \circ \exp^{-1}$, for any $i = 1, \ldots, q$ and observe that it generates the algebra of polynomials on $\mathbb{G}$, i.e. every polynomial P on $\mathbb{G}$ can be represented as $P = \sum_\alpha c_\alpha x^I$, where $\alpha \in \mathbb{N}^q$, $x^\alpha = x_1^{i_1} \cdots x_q^{i_q}$ and only a finite number of coefficients $c_\alpha \in \mathbb{R}$ do not vanish.

Definition 3.4 In the above notation we define the *homogeneous degree* of a polynomial $P : \mathbb{G} \longrightarrow \mathbb{R}$ to be

$$\text{h-deg}(P) = \max\{d(\alpha) \mid c_\alpha \neq 0\},$$

where $d(\alpha) = \sum_{k=1}^{q} d_k i_k$. We denote by $\mathcal{P}_{H,k}(\mathbb{G})$ the space of polynomials of homogeneous degree less than or equal to k.

For instance, in the Heisenberg group $\mathbb{H}^1$ with standard coordinates (x, y, t), the polynomial $P(x, y, t) = t^2 - x^3$ has homogeneous degree equal to 4.

Remark 3.5 The definition of homogeneous degree is independent of the adapted basis (W_j). In fact, the affine transformation between two adapted bases $A : \mathcal{G} \longrightarrow \mathcal{G}$ has the property $A(V_i) = V_i$ for any $i = 1, \dots, \iota$, so the homogeneous degree of a polynomial is preserved under the transformation A.

The Poincaré-Birkhoff-Witt Theorem (shortly PBW Theorem) states that for any basis $(W_1, W_2, \dots W_q)$ of $\mathcal{G}$ regarded as frame of first order differential operators, the algebra of left invariant differential operators on $\mathbb{G}$ has a basis formed by the following ordered terms

$$W^\alpha = W_1^{i_1} \cdot \dots \cdot W_q^{i_q},$$

where $\alpha = (i_1, \dots, i_q)$ varies in $\mathbb{N}^q$, see p.21 of [21]. Analogously as for polynomials we define the degree of a left invariant differential operator $Z = \sum_\alpha c_\alpha W^\alpha$ as

$$\text{h-deg}(Z) = \max\{d(\alpha) \mid c_\alpha \neq 0\},$$

where $d(\alpha) = \sum_{k=1}^{q} d_k i_k$. The space $\mathcal{A}_k(\mathbb{G})$ represents the space of left invariant differential operators of homogeneous degree less than or equal to k. This analogy between polynomials and differential operators is not only formal, as the following proposition shows.

Proposition 3.6 *There exists an isomorphism* $L : \mathcal{P}_{H,k}(\mathbb{G}) \longrightarrow \mathcal{A}_k(\mathbb{G})$*, given by*

$$L(P) = \sum_{d(\alpha) \leq k} W^\alpha P(0)\, W^\alpha.$$

For the proof of this fact we refer the reader to Proposition 1.30 of [21].

In order to deal with higher order differentiability theorems we make some preliminary considerations. Let us consider a basis $\{W^\alpha \mid d(\alpha) \leq k\}$ of $\mathcal{A}_k(\mathbb{G})$ and $u \in BV_H^k(\Omega)$, where $(W_1, \dots, W_q)$ is an adapted basis of $\mathcal{G}$. We denote $W_i = X_i$, with $i = 1, \dots, m$, where $(X_1, \dots, X_m)$ is a fixed horizontal orthonormal frame. Our aim is to find out a polynomial $P : \mathbb{G} \longrightarrow \mathbb{R}$ which approximates u at a fixed point $x \in \Omega$ with order k. In view of the last proposition it is natural to look for a substitute of homogeneous derivatives W^α of u at x, with $d(\alpha) \leq k$. Our first observation is that due to the stratification of $\mathcal{G}$ the operators W^{α} with $d(\alpha) \leq l$ are linear combinations of operators $X_{\gamma_1} \cdots X_{\gamma_l}$ with $1 \leq \gamma_i \leq m$ and $l \leq k$. Therefore the distributional derivatives $D_W^{\alpha} u$ are measures whenever $d(\alpha) \leq k$. So, taking into account the preceeding observation and the fact that vector fields W_i have vanishing divergence, we state the following definition.

Definition 3.7 Let $u \in BV_H^k(\Omega)$. For any $\alpha \in \mathbb{N}^q$, $d(\alpha) \leq k$, we consider the following multi-index Radon measures $D_W^{\alpha} u$ defined as

$$\int_{\Omega} \phi \, dD_W^{\alpha} u = (-1)^{|\alpha|} \int_{\Omega} u \ W_l^{\alpha_l} \cdots W_1^{\alpha_1} \phi \qquad \forall \phi \in C_c^{\infty}(\Omega).$$

By Radon-Nikodým Theorem we have $D_W^{\alpha} u = (D_W^{\alpha} u)^a + (D_W^{\alpha} u)^s$, where the addenda are respectively the absolutely continuous part and the singular part of the measure $D_W^{\alpha} u$ with respect to the volume measure. We define the *weak mixed derivatives* as the summable maps $\nabla_W^{\alpha} u$ such that

$$(D_W^{\alpha} u)^a = \nabla_W^{\alpha} u \ \mathrm{vol}\,.$$

Our substitute for the α-derivative of u is $\tilde{u}_{W^{\alpha}}(x)$, which is the approximate limit of $\nabla_W^{\alpha} u$ at points $x \in \Omega \setminus S_{\nabla_W^{\alpha} u}$. Now, let us consider the differential operator $X_{\gamma_1} \cdots X_{\gamma_l} u$ where $1 \leq \gamma_i \leq m$ and $1 \leq l \leq k$. By virtue of PBW Theorem, there exist coefficients $\{c_{\gamma}^{\alpha}\}$ such that

$$X_{\gamma_1} \cdots X_{\gamma_l} u = \sum_{j=1}^{N_k} c_{\gamma,\alpha} \ W^{\alpha} u\,, \tag{33}$$

where $N_k = \dim\left(\mathcal{A}_k(\mathbb{G})\right)$.

Definition 3.8 Let $u \in BV_H^k(\Omega)$. Utilizing the above notation, we denote by u_{γ} the density of the absolutely continuous part of the measure $X_{\gamma_1} \cdots X_{\gamma_l} u$, where $\gamma \in \{1, \ldots, m\}^l$ and $l \leq k$.

Decomposing the singular and absolutely continuous part of both the measures in (33), we obtain the following equality of summable maps

$$u_{\gamma} = \sum_{j=1}^{N_k} c_{\gamma,\alpha} \ \nabla_W^{\alpha} u\,. \tag{34}$$

The next theorem is the main result of this section and can be regarded as a weak extension of Alexandrov differentiability theorem to the setting of non-Riemannian geometries.

Theorem 3.9 (Alexandrov) *Let $u \in BV^2_H(\Omega)$. Then for a.e. $x \in \Omega$ there exists a polynomial $P_{[x]}$ with h-$\deg(P_{[x]}) \leq 2$, such that*

$$\lim_{r\to 0^+} \frac{1}{r^2} \fint_{U_{x,r}} |u - P_{[x]}| = 0 \tag{35}$$

PROOF. First of all, we fix a point $x \notin \bigcup_{d(\alpha)\leq 2} S_{\nabla^\alpha_W u}$ such that (34) and the limit

$$\lim_{r\to 0^+} \frac{|\,(X_{\gamma_1}X_{\gamma_2}u)^s\,|(U_{x,r})}{r^Q} = 0 \tag{36}$$

hold for every $\alpha \in \mathbb{N}^q$ and $\gamma \in \{1,\dots,m\}^l$, with $d(\alpha) \leq 2$, $l = 1, 2$. By the previous discussion and Remark 1.5, the set of points where these conditions do not occur is negligible. Due to Proposition 3.6, there exists a unique polynomial $P_{[x]} = P$ which satisfies the condition $W^\alpha P(x) = \tilde u_{W^\alpha}(x)$, whenever $d(\alpha) \leq 2$. Now, let us define $w = u - P$. By relation (34) we observe that $\tilde w_\gamma(x) = 0$ for any $\gamma \in \{1,\dots,m\}^l$, $l = 0, 1, 2$. This means that

$$\tilde w(x) = 0\,, \quad \tilde w_i(x) = 0\,, \quad \tilde w_{ij}(x) = 0 \tag{37}$$

for any $i, j = 1, \dots, m$. We consider the summable map

$$v = |D_H w| = \Big(\sum_{i=1}^m w_i^2\Big)^{1/2}.$$

By Proposition 1.25 it follows that

$$|D_H v| \leq \sum_{i=1}^m |D_H w_i|\,,$$

hence conditions (36) and (37) yield

$$|D_H v|(U_{x,r}) = o(r^Q)\,. \tag{38}$$

We can fix $r_0 > 0$ small enough such that $U_{x,4r_0} \subset \Omega$, so we will consider all $r \in]0, r_0[$. By the standard telescopic estimate (23), for a.e. $y \in U_{x,r}$ we have

$$|\tilde w(y)| \leq C\,[M_{2r}v(x) + M_{2r}v(y)]\,d(x,y)\,,$$

therefore, taking the average over $U_{x,r}$ and dividing by r^2 we obtain

$$\frac{1}{r^2}\fint_{U_{x,r}} |w(y)|\,dy \leq C\left(\frac{M_{2r}v(x)}{r} + \frac{1}{r}\fint_{U_{x,r}} M_{2r}v(y)\,dy\right).$$

Thus, in order to prove (35) we show that the maps

$$a(r) = r^{-1}M_{2r}v(x)\,, \qquad b(r) = r^{-1}\fint_{U_{x,r}} M_{2r}v(y)\,dy$$

go to zero as $r \to 0^+$. Since also $\tilde{v}(x) = 0$, inequality (18) gives

$$|M_r v(x)| \le C \left[(Q-1) \int_0^r \frac{|D_H v|(U_{x,t})}{t^Q} dt + \frac{|D_H v|(U_{x,r})}{r^{(Q-1)}} \right] .$$

By (38) and the last estimate we get that $a(r) \to 0$ as $r \to 0^+$. Let us consider the estimate

$$b(r) \le \frac{1}{r} \fint_{U_{x,r}} |M_{2r} v(y) - \tilde{v}(y)| \, dy + \frac{1}{r} \fint_{U_{x,r}} |v(y)| \, dy \,,$$

observing that

$$\frac{1}{r} \fint_{U_{x,r}} |v(y)| \, dy \le r^{-1} M_r v(x) \le a(r) \longrightarrow 0 \quad \text{as} \quad r \to 0^+ .$$

In view of inequality

$$|M_{2r} v(y) - \tilde{v}(y)| \le M_{2r}[v - \tilde{v}(y)](y) \,, \tag{39}$$

and applying inequality (18) to the map $z \longrightarrow v(z) - \tilde{v}(y)$ we get

$$M_{2r}[v - \tilde{v}(y)](y) \le C \left[(Q-1) \int_0^{2r} \frac{|D_H v|(U_{y,t})}{t^Q} dt + \frac{|D_H v|(U_{y,2r})}{(2r)^{(Q-1)}} \right] . \tag{40}$$

Thus, estimates (39) and (40) yield

$$\frac{1}{r} \fint_{U_{x,r}} |M_{2r} v(y) - \tilde{v}(y)| dy \le \frac{C}{r} \fint_{U_{x,r}} \left[(Q-1) \int_0^{2r} \frac{|D_H v|(U_{y,t})}{t^Q} dt + \frac{|D_H v|(U_{y,2r})}{(2r)^{(Q-1)}} \right] dy \,.$$

Now, in order to get the thesis, we have to prove that both terms

$$\alpha(r) = \frac{1}{r} \fint_{U_{x,r}} \int_0^{2r} \left(\frac{|D_H v|(U_{y,t})}{t^Q} dt \right) dy \,, \quad \beta(r) = \frac{1}{r} \fint_{U_{x,r}} \frac{|D_H v|(U_{y,2r})}{(2r)^{(Q-1)}} dy$$

are infinitesimal as $r \to 0^+$. By Fubini's Theorem we have

$$\begin{aligned} \alpha(r) &= \frac{r^{-1}}{|U_{x,r}|} \int_0^{2r} \frac{dt}{t^Q} \int_{U_{x,r}} \left(\int_{U_{x,3r}} \mathbf{1}_{U_{y,t}}(z) \, d|D_H v|(z) \right) dy \\ &= \frac{r^{-1}}{|U_{x,r}|} \int_0^{2r} \frac{dt}{t^Q} \int_{U_{x,3r}} |U_{x,r} \cap U_{z,t}| \, d|D_H v|(z) \\ &= \frac{|U_1| r^{-1}}{|U_{x,r}|} \int_0^{2r} \int_{U_{x,3r}} \frac{|U_{x,r} \cap U_{z,t}|}{|U_{z,t}|} \, d|D_H v|(z) \le 3^Q \, 2 \, \frac{|D_H v|(U_{x,3r})}{(3r)^Q} . \end{aligned}$$

By (38) the last term goes to zero as $r \to 0$, so $\lim_{r \to 0} \alpha(r) = 0$. Similarly, we have

$$\beta(r) = \frac{1}{2^{Q-1} r^Q |U_{x,r}|} \int_{U_{x,r}} \left(\int_{U_{x,3r}} \mathbf{1}_{U_{y,2r}}(z) \, d|D_H v|(z) \right) dy$$

$$= \frac{1}{2^{Q-1}r^Q}\int_{U_{x,3r}} \frac{|U_{z,2r}\cap U_{x,r}|}{|U_{x,r}|}\, d|D_H v|(z) \le \frac{3^Q}{2^{Q-1}}\frac{|D_H v|(U_{x,3r})}{(3r)^Q}\,.$$

Again, utilizing (38) on the last term we get $\lim_{r\to 0^+}\beta(r) = 0$, so the thesis follows. □

The arguments used for second order differentiability of H-BV^2 functions can be extended with some additional efforts to higher order differentiability.

Theorem 3.10 *Let $u \in BV_H^k(\Omega)$ and $1 \le l \le k$. Then for a.e. $x \in \Omega$ there exists a polynomial $P_{[x]}$, with h-deg$(P_{[x]}) \le l$, such that*

$$\lim_{r\to 0^+} \frac{1}{r^l} \fint_{U_{x,r}} |u - P_{[x]}| = 0. \tag{41}$$

PROOF. We prove the theorem by induction on $k \ge 2$. Theorem 2.2 and Theorem 3.9 give us the validity of induction hypothesis for $k = 2$. Now, let us consider $u \in BV_H^k(\Omega)$ with $k \ge 3$. Clearly we have $X_iX_ju \in BV_H^{k-2}$ for any $i,j = 1,\dots,m$. By induction hypothesis for a.e. $x \in \Omega$ there exist polynomials $R_{[x,ij]}$, with h-deg$(R_{[x,ij]}) \le k-2$, such that

$$\fint_{U_{x,r}} |u_{ij} - R_{[x,ij]}| = o(r^{k-2}) \tag{42}$$

and

$$W^\beta R_{[x,ij]}(x) = \tilde u_{ijW^\beta}(x)\,, \quad \text{whenever } d(\beta) \le k-2\,. \tag{43}$$

Moreover, for a.e. $x \in \Omega$ there exists a polynomial $P_{[x]}$, with h-deg$(P_{[x]}) \le k$, such that

$$W^\alpha P_{[x]}(x) = \tilde u_{W^\alpha}(x)\,, \quad \text{whenever } d(\alpha) \le k\,. \tag{44}$$

The PBW Theorem yields the distributional relations

$$W^\beta X_iX_j = \sum_{d(\alpha)\le k} c^\beta_{ij,\alpha} W^\alpha \tag{45}$$

for any $i,j = 1,\dots,m$ and $\beta \in \mathbb{N}^q$ with $d(\beta) \le k-2$. Thus, relations (43), (44), (45) and the following equality

$$(W^\beta X_iX_ju)^a = (W^\beta u_{ij})^a = \tilde u_{ijW^\beta}$$

imply

$$W^\beta R_{[x,ij]}(x) = \sum_{d(\alpha)\le k} c^\beta_{ij,\alpha}\tilde u_{W^\alpha}(x) = \sum_{d(\alpha)\le k} c^\beta_{ij,\alpha} W^\alpha P_{[x]}(x) = W^\beta X_iX_jP(x)\,,$$

whenever $d(\beta) \le k-2$. Thus, Proposition 3.6 yields $R_{[x,ij]} = X_iX_jP$.

Now, let us define $w = u - P$ and $v = |D_H w|$, obtaining the following inequalities of measures

$$|D_H v| \le \sum_{i=1}^{m} |D_H X_i w| \le \sum_{i,j=1}^{m} |X_j X_i w| \,. \tag{46}$$

By the fact that $u \in BV_H^k(\Omega)$, with $k \ge 3$, the distributional derivatives $X_j X_i w$ are represented by integrable functions w_{ij}. So, equality $R_{[x,ij]} = X_i X_j P$ and the inductive formula (42) yield

$$\frac{|X_j X_i w|(U_{x,r})}{|U_{x,r}|} = \fint_{U_{x,r}} |w_{ij}| = \fint_{U_{x,r}} |u_{ij} - R_{[x,ij]}| = o(r^{k-2}),$$

hence (46) implies

$$|D_H v|(U_{x,r}) = o(r^{Q+k-2})\,. \tag{47}$$

Now, the rest of the proof proceeds analogously to Theorem 3.9, replacing property (38) with (47). This last observation leads us to the conclusion. □

4 Some examples of H-BV^2 functions

In this section we present a class of H-BV^2 functions arising from the inf-convolution of the so-called gauge distance in the Heisenberg group $\mathbb{H}^n$.

We begin with some elementary remarks about distributional derivatives along vector fields. In the following preliminary considerations the set Ω will be an open subset of $\mathbb{R}^q$ with the Euclidean metric.

Let $X : \Omega \longrightarrow \mathbb{R}^q$ be a locally Lipschitz vector field; then, the following chain rule

$$D_X(h \circ u) = h'(u) D_X u \tag{48}$$

holds whenever $h : \mathbb{R} \to \mathbb{R}$ is continuously differentiable, $u : \Omega \longrightarrow \mathbb{R}$ is continuous and $D_X u$ is representable by a Radon measure in Ω as follows:

$$\int_\Omega u \; X^* \varphi \; d\mathrm{vol} = \int_\Omega \varphi \; dD_X u \qquad \forall \varphi \in C_c^\infty(\Omega),$$

where $X^* = -X - \mathrm{div} X$ is the formal adjoint of X. Analogously, the product rule

$$D_X(uv) = v D_X u + u D_X v \tag{49}$$

holds whenever $u : \Omega \longrightarrow \mathbb{R}$ is locally integrable (or locally bounded) and $D_X u$ is representable in Ω by a Radon measure, $v : \Omega \longrightarrow \mathbb{R}$ is continuous and $D_X v$ is representable in Ω by a locally bounded (or locally summable) function. The proofs of (48) and (49) can be achieved by approximations of the following type.

Proposition 4.1 *Let $u \in \Xi$, where Ξ is either $C(\Omega)$, $L^1_{loc}(\Omega)$ or $L^\infty_{loc}(\Omega)$, respectively. Then there exists a sequence of smooth functions (u_l) such that*

$$|D_X u_l|(\Omega) \leq |D_X u|(\Omega) + \frac{1}{l} \tag{50}$$

and either (u_l) uniformly converges to u on compact sets, or it converges to u in $L^1_{loc}(\Omega)$, or it is locally uniformly locally bounded, respectively.

The estimate (50) is proved in [24], [29]. One considers a locally finite open cover $\{A_i\}$, where $A_i = \Omega_{i+1} \setminus \overline{\Omega}_{i-1}$ and

$$\Omega_i = \left\{ x \in \Omega \middle| , |x| < i, \ \mathrm{dist}(x, \Omega^c) > \frac{1}{i+1} \right\}$$

for any $i \in \mathbb{N}$, with $\Omega_{-1} = \emptyset$. A smooth partition of unity $\{\psi_i\}$ is defined with respect to $\{A_i\}$, hence the candidate to be the approximating functions is as follows

$$u_l := \sum_{i=0}^{\infty} (u\psi_i) * \phi_{\varepsilon_i}$$

with $\varepsilon_i = \varepsilon_i(l)$ small enough. Since $\sup_i \varepsilon_i(l)$ tends to 0 as $l \to \infty$ all L^p convergence properties of the sequence follow directly from this representation. Notice also that when $D_X u << \mathrm{vol}$, we get the $L^1_{loc}(\Omega)$ convergence of the densities of $D_X u_l$ to the density of $D_X u$, see either Proposition 1.2.2 of [24] or Theorem A.2 of [29].

The proof of (48) follows by approximation of u with the sequence (u_l) of Proposition 4.1, so that $D_X u_l$ weakly converges to $D_X u$ in the topology of Radon measures and u_l converges to u uniformly on compact sets of Ω. The proof of (49) is similar and requires either the L^1_{loc} convergence of u_l to u when $u \in L^1_{loc}$, or the additional uniform local bound, when $u \in L^\infty_{loc}$, and the L^1_{loc} convergence of densities $D_X v_l$ to $D_X v$, when $D_X v \in L^1_{loc}$, or the additional uniform local bound, when $D_X v \in L^\infty(\Omega)$, together with the uniform convergence of v_l to v on compact sets of Ω.

In the sequel the "minimum function" between two real valued maps u and v we will be denoted by $(u \wedge v)(x) = \min\{u(x), v(x)\}$. Radon measures will be also interpreted as linear functional on continuous compactly supported test functions. Hence the notation $\mu_1 \leq \mu_2$, where μ_1 and μ_2 are Radon measures on Ω, means that for any nonnegative continuous compactly supported test function φ on Ω, we have

$$\int_\Omega \varphi \, d\mu_1 \leq \int_\Omega \varphi \, d\mu_2 \, .$$

Lemma 4.2 *Let u, $v : \Omega \longrightarrow \mathbb{R}$ be continuous functions, $\gamma \in \mathbb{R}$ and let $X : \Omega \longrightarrow \mathbb{R}^q$ be a locally Lipschitz vector field. Then*

$$D_X u \leq \gamma \, vol \quad D_X v \leq \gamma \, vol \qquad \Longrightarrow \qquad D_X(u \wedge v) \leq \gamma \, vol. \tag{51}$$

If $D_X u$ and $D_X v$ are representable by $L^\infty_{loc}(\Omega)$ functions, then

$$D_{XX} u \leq \gamma \, vol, \quad D_{XX} v \leq \gamma \, vol \qquad \Longrightarrow \qquad D_{XX}(u \wedge v) \leq \gamma \, vol. \tag{52}$$

PROOF. In order to show (51) it suffices to approximate $u \wedge v$ by $u + h_\epsilon(u-v)$, where $h_\epsilon \in C^\infty(\mathbb{R})$, $-1 \le h'_\epsilon \le 0$, $h_\epsilon(t) \to -t^+$ uniformly as $\epsilon \to 0^+$. Indeed, the chain rule (48) gives

$$D_X(u + h_\epsilon(u-v)) = \big(1 + h'_\epsilon(u-v)\big)\, D_X u - h'_\epsilon(u-v) D_X v \le \gamma\,\mathrm{vol}.$$

The implication (52) follows by the same argument, noticing that the functions h_ϵ can be chosen to be concave. We have

$$\begin{aligned}
&D_{XX}(u + h_\epsilon(u-v))\\
&= \big(1 + h'_\epsilon(u-v)\big)\, D_{XX} u - h'_\epsilon(u-v) D_{XX} v + h''_\epsilon(u-v)(D_X u - D_X v)^2\\
&\le \big(1 + h'_\epsilon(u-v)\big)\, D_{XX} u - h'_\epsilon(u-v) D_{XX} v \le \gamma\,\mathrm{vol}.
\end{aligned}$$

□

Now we particularize our study to $\mathbb{H}^n$ (we recall that $\mathbb{H}^n$ is isomorphic to $\mathbb{R}^{2n+1}$). To denote elements of $\mathbb{H}^n$ we consider the coordinates $(\xi, t) = (\xi_1, \ldots, \xi_{2n}, t)$. The following family of vector fields

$$X_i = \partial_{\xi_i} + 2\xi_{n+i}\partial_t, \quad Y_i = \partial_{\xi_{n+i}} - 2\xi_i\partial_t, \qquad i = 1, \ldots, n \tag{53}$$

can be considered as an horizontal orthonormal frame of $H\mathbb{H}^n$, so

$$\nabla_H u = \sum_{i=1}^n X_i u\, X_i + Y_i u\, Y_i$$

whenever u is smooth. The only nontrivial bracket relations are

$$[X_i, Y_i] = -4\, Z = -4\,\partial_t, \qquad i = 1, \ldots, n\,.$$

Via the Baker-Campbell-Hausdorff formula our vector fields induce the following group operation

$$x\, x' = \left(\xi + \xi', t + t' + 2\sum_{i=1}^n \xi_{n+i}\xi'_i - \xi_i\xi'_{n+i}\right).$$

Now for any element $x = (\xi, t) \in \mathbb{H}^n$ we define the following gauge norm

$$\|(\xi, t)\| = \sqrt[4]{|\xi|^4 + t^2}\,.$$

A non-trivial fact is that $d(x, y) = \|x^{-1}y\|$ yields a left invariant distance on $\mathbb{H}^n$, see [34]. In the following we define $c(x, y) = d(x, y)^2$ and we consider a function u arising from the inf-convolution of c. Precisely, we assume that there exists a bounded family $\{y_i\}_{i\in I} \subset \mathbb{H}^n$ and $t_i \in \mathbb{R}$ such that

$$u(x) = \inf_{i\in I} c(x, y_i) + t_i \qquad \forall x \in \mathbb{H}^n. \tag{54}$$

Inf-convolution formulas of this type appear in several fields, for instance in the representation theory of viscosity solutions, in the related field of dynamic programming (see for instance [12], [40]) and in the theory of optimal transportation problems. In the latter theory, functions representable as in (54) are called *c-concave* (see [53], [54]). In these theories it is well known that in many situations the function u inherits from c a one-sided estimate on the second distributional derivative; for instance, this is the case when $c(x,y) = h(x-y)$ and $h : \mathbb{R}^q \longrightarrow \mathbb{R}$ is a $C^{1,1}_{loc}$ function (see for instance [28]). In the following theorem we extend this result to the Heisenberg setting, thus getting a non-trivial class of examples of H-BV^2-functions.

Theorem 4.3 *Let $\Omega \subset \mathbb{H}^n$ be a bounded open set. The function u defined in (54) is locally Lipschitz and belongs to $BV^2_H(\Omega)$.*

PROOF. Since the family $\{y_i\}_{i\in I}$ is bounded it is easy to check that $\{c(\cdot, y_i)\}$ is uniformly locally Lipschitz in Ω, therefore u is a Lipschitz function in Ω. Notice also that, since $\mathbb{H}^n$ is separable, we can assume I to be finite or countable with no loss of generality. The essential fact leading to the H-BV^2 property consists in the estimates on $\mathbb{H}^n \setminus \{e\}$

$$\max_{i,j=1,\dots,n} \Big\{ |X_iX_jc(\cdot,e)|, |X_iY_jc(\cdot,e)|, |Y_iY_jc(\cdot,e)|, |Y_jX_ic(\cdot,e)| \Big\} \le 10 \tag{55}$$

These inequalities can be obtained by explicit calculation. We define $\sigma(\xi,t) = |\xi|^4 + t^2$ and note that $c\left((\xi,t),e\right) = \sqrt{\sigma(\xi,t)} = \eta(\xi,t)$. Now we write the second horizontal derivatives in terms of the functions σ. We obtain

$$X_jX_i\eta = \frac{X_jX_i\sigma}{2\sqrt{\sigma}} - \frac{X_j\sigma\, X_i\sigma}{4\sqrt{\sigma^3}}, \quad Y_jY_i\eta = \frac{Y_jY_i\sigma}{2\sqrt{\sigma}} - \frac{Y_j\sigma\, Y_i\sigma}{4\sqrt{\sigma^3}}, \tag{56}$$

$$Y_jX_i\eta = \frac{Y_jX_i\sigma}{2\sqrt{\sigma}} - \frac{Y_j\sigma\, X_i\sigma}{4\sqrt{\sigma^3}}, \quad X_jY_i\eta = \frac{X_jY_i\sigma}{2\sqrt{\sigma}} - \frac{X_j\sigma\, Y_i\sigma}{4\sqrt{\sigma^3}} \tag{57}$$

for any $i,j = 1,\dots,n$. Expressions (53) and a direct calculation yield

$$\max_{i,j=1,\dots,n} \Big\{ |X_iX_j\sigma|, |Y_jX_i\sigma|, |Y_jY_i\sigma| \Big\} \le 12\sqrt{\sigma} \tag{58}$$

$$\max_{i=1,\dots,n} \Big\{ |X_i\sigma|, |Y_i\sigma| \Big\} \le 4\sqrt[4]{\sigma^3}\,. \tag{59}$$

Applying estimates (58) and (59) to formulae (56) and (57) the estimate (55) follows. Notice that the identity $c(z,y) = c(y^{-1}z,e)$ yields

$$T[c(\cdot,y)](z) = T[c(y^{-1}\cdot,e)](z) = Tc(\cdot,e)(y^{-1}z)$$

for any left invariant vector field T. It follows that (55) holds replacing the unit element e with any element $y \in \mathbb{H}^n$. Furthermore, horizontal vector fields P of the first layer V_1 have the representation $P = \sum_{i=1}^n a_iX_i + b_iY_i$. We also require that the

Riemannian norm of P is less than or equal to one, hence we have $\sum_{i=1}^n a_i^2 + b_i^2 \leq 1$. Then, another direct calculation using the expression of P and the estimate (55) yield

$$|PPc(\cdot, y)| \leq \gamma \quad \text{on } \mathbb{H}^n \setminus \{y\},$$

for any $y \in \mathbb{H}^n$, where $\gamma = 20n^2$. By Lemma 4.2 we obtain that

$$D_{PP}u \leq \gamma \text{ vol}$$

first for finite families and then, by a limiting argument, for countable families. In particular $D_{PP}u$ is representable in Ω by a Radon measure for any P of the above form. By polarization identity, taking $P = (X_i \pm X_j)/\sqrt{2}$, $P = (X_i \pm Y_j)/\sqrt{2}$ and $P = (Y_i \pm Y_j)/\sqrt{2}$, respectively, we obtain that $D_{X_iX_j}u$, $D_{Y_iY_j}u$ $D_{X_iY_j+Y_jX_i}u$ are Radon measures for any $i, j = 1, \ldots, n$. In particular $D_{X_iY_j}u$ is a measure whenever $i \neq j$. With the previous notation, we also see that whenever $(\xi, t) \neq 0$ we have

$$|Zc((\xi, t), e)| = |\partial_t \eta(\xi, t)| = \left| \frac{t}{\sqrt{|\xi|^4 + t^2}} \right| \leq 1$$

Again, Lemma 4.2 implies that $D_Z u$ is representable in Ω by a Radon measure (actually absolutely continuous with respect to the volume measure vol). Finally, from the relation $D_{X_iY_i+Y_iX_i} + D_Z = 2D_{X_iY_i}$ we conclude that $D_{X_iY_i}u$ is a Radon measure. □

Remark 4.4 In the case when $c(x, y) = d(x, y)^2$ is the square of the Carnot-Carathéodory distance in $\mathbb{H}^1$, one can proceed in the same way using the explicit formula for d computed in [46]. By a direct long calculation it is still possible to prove the existence of upper bounds on $D_{PP}c(\cdot, y)$ and $D_Z c(\cdot, y)$, when $P = \sum_{i=1}^n aX + bY$, [41], [52]. This suffices to obtain Theorem 4.3 with this cost function.

References

[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)

[3] M.Amar, V.De Cicco, *Relaxation of quasi-convex integrals of arbitrary order*, Proc. Roy. Soc. Edinburgh Sect. A, **124**, n.5, 927-946, (1994)

[4] L.Ambrosio, *Some fine properties of sets of finite perimeter in Ahlfors regular metric measure spaces*, Adv. Math., **159**, 51-67, (2001)

[5] L.Ambrosio, N.Fusco, D.Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford University Press, (2000).

[6] L.Ambrosio, B.Kirchheim, *Rectifiable sets in metric and Banach spaces*, Math. Ann. **318**, 527-555, (2000)

[7] L.Ambrosio, B.Kirchheim, *Currents in metric spaces*, Acta Math., **185**, 1-80, (2000)

[8] L.Ambrosio, P.Tilli, *Selected Topics on "Analysis in Metric Spaces"*, Quaderno S.N.S. (2000)

[9] A.Bellaïche, *The Tangent space in sub-Riemannian geometry*, in *Subriemannian Geometry*, Progress in Mathematics, **144**. ed. by A.Bellaiche and J.Risler, Birkhäuser Verlag, Basel, 1996.

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

[11] L.Capogna, D.Danielli, N.Garofalo, *The geometric Sobolev embedding for vector fields and the isoperimetric inequality*, Comm. Anal. Geom. **2**, n.2, 203-215, (1994)

[12] I.Capuzzo Dolcetta, M.Bardi, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi equations*, Birkhäuser, (1997).

[13] J.Cheeger, *Differentiability of Lipschitz functions on metric measure spaces*, Geometric and Functional Analysis, **9**, 428-517, (1999)

[14] D.Danielli, N.Garofalo, D.M. Nhieu, *Traces Inequalities for Carnot-Carathéodory Spaces and Applications*, Ann. Scuola. Norm. Sup., **27**, 195-252, (1998)

[15] D.Danielli, N.Garofalo, D.M. Nhieu, *Notions of convexity in Carnot groups*, forthcoming

[16] G.David, S.Semmes, *Fractured Fractals and Broken Dreams. Self-Similar Geometry through Metric and Measure*, Oxford University Press, 1997.

[17] E.De Giorgi, *Nuovi teoremi relativi alle misure (r-1) dimensionali in uno spazio ad r dimensioni*, Ricerche Mat. **4**, 95-113, (1955)

[18] F.Demengel, *Fonctions a hessien borne*, Ann. Inst. Fourier, Grenoble, **34**, n.2, 155-190, (1984)

[19] L.C.Evans, R.F.Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, (1992)

[20] H.Federer, *Geometric Measure Theory*, Springer, (1969)

[21] G.B. Folland, E.M. Stein, *Hardy Spaces on Homogeneous groups*, Princeton University Press, 1982

[22] B.Franchi, G.Lu, R.L.Wheeden, *Representation formulas and weighted Poincaré inequalities for Hörmander vector fields*, Ann. Inst. Fourier, Grenoble, **45**, 577-604, (1995)

[23] B.Franchi, G.Lu, R.L.Wheeden, *A relationship between Poincaré inequality and Representation formulas in spaces of Homogeneous type*, Int. Math. Res. Not., n.1, (1996)

[24] B.Franchi, R.Serapioni, F.Serra Cassano, *Meyers-Serrin type theorems and relaxation of variational integrals depending on vector fields*, Houston Jour. Math., **22**, 859-889, (1996)

[25] B.Franchi, R.Serapioni, F.Serra Cassano, *Rectifiability and Perimeter in the Heisenberg group*, Math. Ann. **321**, n.3, (2001)

[26] B.Franchi, R.Serapioni, F.Serra Cassano, *Regular hypersurfaces, intrinsic perimeter and implicit function theorem in Carnot groups*, to appear in Comm. Anal. Geom.

[27] B.Franchi, R.Serapioni, F.Serra Cassano, *On the structure of finite perimeter sets in step 2 Carnot groups*, forthcoming

[28] W.Gangbo, R.J.McCann, *The geometry of optimal transportation*, Acta Math., **177** 113-166, (1996)

[29] N.Garofalo, D.M.Nhieu, *Isoperimetric and Sobolev Inequalities for Carnot-Carathéodory Spaces and the Existence of Minimal Surfaces*, Comm. Pure Appl. Math. **49**, 1081-1144 (1996)

[30] P.Hajlasz, P.Koskela, *Sobolev met Poincaré*, Mem. A.M.S. **145**, (2000)

[31] D.Jerison, *The Poincaré inequality for vector fields satisfying Hörmander's condition*, Duke Math. Jour., **53**, n.2, 503-523, (1986)

[32] B.Kirchheim, *Rectifiable metric spaces: local structure and regularity of the Hausdorff measure*, Proc. AMS, **121**, 113-123, (1994).

[33] B.Kirchheim, V.Magnani, *A counterexample to metric differentiability*, to appear in Proc. Ed. Math. Soc.

[34] A.Korányi, H.M.Reimann, *Quasiconformal mappings on the Heisenberg group*, Invent. Math. **80**, 309-338, (1985)

[35] G.Lu, *Weighted Poincaré and Sobolev inequalities for vector fields satisfying Hörmander's condition and applications*, Rev. Mat. Iberoamericana, **8**, n. 3, 367-439, (1992)

[36] G.Lu, J.J.Manfredi, B.Stroffolini, *Convex functions in Carnot groups*, preprint

[37] V.Magnani, *Differentiability and Area formula on stratified Lie groups*, Houston Jour. Math., **27**, n.2, 297-323, (2001)

[38] V.Magnani, *On a general coarea inequality and applications*, Ann. Acad. Sci. Fenn. Math., **27**, 121-140, (2002)

[39] V.Magnani, *A blow-up theorem for regular hypersurfaces on nilpotent groups*, to appear on Manuscripta Math.

[40] J.J.Manfredi, B.Stroffolini, *A version of the Hopf-Lax formula in the Heisenberg group*, to appear in Comm. P.D.E.

[41] J.J.Manfredi, B.Stroffolini, *Personal communication*

[42] G.A.Margulis, G.D.Mostow, *Some remarks on the definition of tangent cones in a Carnot-Carathéodory space*, J. Analyse Math. **80**, 299-317 (2000)

[43] P.Mattila, *Geometry of sets and measures in Euclidean spaces*, Cambridge U.P. (1995).

[44] M.Miranda, *Functions of bounded variation on "good" metric measure spaces*, preprint

[45] J.Mitchell, *On Carnot-Carathéodory metrics*, J.Diff. Geom. **21**, 35-45, (1985)

[46] R.Monti, *Some properties of Carnot-Carathéodory balls in the Heisenberg group*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Mem. s.9 vol.11, 155-167, (2000)

[47] R.Monti, D.Morbidelli, *Some trace theorems for vector fields*, Math. Zeit., **239**, n.4, 747-776, (2002)

[48] R.Monti, F.Serra Cassano, *Surface measures in Carnot-Carathédory spaces*, Calc. Var., **13**, n.3, 339-376, (2001)

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

[50] S.D.Pauls, *The Large Scale Geometry of Nilpotent Lie groups*, to appear on Comm. Anal. Geom.,

[51] S.D.Pauls, *A notion of rectifiability modeled on Carnot groups*, preprint

[52] S.Rigot, *Personal communication*

[53] L.Rüschendorf: *On c-optimal random variables,* Statistics & Probability Letters, **27** (267-270, (1996)

[54] S.T.Rachev, L.Rüschendorf, *Mass transportation problems.* Probability and its applications, Springer, (1998).

[55] V.S.Varadarajan, *Lie groups, Lie algebras and their representation*, Springer-Verlag, New York, (1984).

[56] Y.Q.Song, X.P.Yang, *BV functions in the Heisenberg group*, preprint

[57] S.Semmes, *Fractal Geometry, "Decent Calculus", Structure among Geometries*, preprint

[58] R.S.Strichartz, *Sub-Riemannian geometry*, J. Diff. Geom., **24**, 221-263, (1986). Corrections: J. Diff. Geom., **30**, 595-596, (1989).

[59] N.Th.Varopoulos, L.Saloff-Coste, T.Coulhon, *Analysis and Geometry on Groups*, Cambridge University Press, Cambridge, 1992.

[60] S.K.Vodop'yanov, A.D.Ukhlov, *Approximately differentiable transformations and change of variables on nilpotent groups*, Sib. Math. J., **37**, No.1, 62-78, (1996)

[61] S.K.Vodop'yanov, *$\mathcal{P}$-differentiability on Carnot groups in different topologies and related topics*, Proc.on Analysis and Geometry, 603-670, Sobolev Institute Press, Novosibirsk, (2000).

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

[63] W.P.Ziemer, *Weakly differentiable functions*, Springer, (1989)