A MEASURE ZERO UNIVERSAL DIFFERENTIABILITY SET IN THE HEISENBERG GROUP

ANDREA PINAMONTI AND GARETH SPEIGHT

ABSTRACT. We show that the Heisenberg group $\mathbb{H}^n$ contains a measure zero set N such that every Lipschitz function $f\colon \mathbb{H}^n \to \mathbb{R}$ is Pansu differentiable at a point of N. The proof adapts the construction of small 'universal differentiability sets' in the Euclidean setting: we find a point of N and a horizontal direction where the directional derivative in horizontal directions is almost locally maximal, then deduce Pansu differentiability at such a point.

1. INTRODUCTION

Rademacher's theorem states that every Lipschitz function $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is differentiable almost everywhere with respect to Lebesgue measure. This result is classical but has many applications and has inspired much research. One direction of this research is the extension of Rademacher's theorem to more general spaces, while another involves finding points of differentiability in extremely small sets. In this article we investigate both directions, by constructing a measure zero 'universal differentiability set' in the Heisenberg group (Theorem 2.12), the most studied non-Euclidean Carnot group.

A Carnot group is a Lie group (smooth manifold which is also a group with smooth operations) whose Lie algebra (tangent space at the identity) admits a stratification. This stratification decomposes the Lie algebra as a direct sum of finitely many vector spaces; one of these consists of privileged 'horizontal directions' which generate the other directions using Lie brackets. The stratification allows one to define dilations on the group. Carnot groups have a natural metric, defined using lengths of horizontal curves, and Haar measure invariant under group operations.

Using the group translations and dilations, one can define Pansu differentiability of functions between Carnot groups. Pansu's theorem states that Lipschitz functions between Carnot groups are Pansu differentiable almost everywhere with respect to the Haar measure [23]. This can be applied to show that every Carnot group (other than Euclidean space itself) contains no subset of positive measure which bi-Lipschitz embeds into a Euclidean space. Carnot groups are a source of many questions in analysis and geometry [8, 17, 18, 22, 26, 28].

The Heisenberg group $\mathbb{H}^n$ (Definition 2.1) is the simplest non-Euclidean Carnot group, but is still not completely understood. For example, a geometric notion of intrinsic Lipschitz function between subgroups of $\mathbb{H}^n$ was introduced in [16] to study rectifiable sets [14, 15, 20] and minimal surfaces [5, 6, 21, 27]. Intrinsic Lipschitz functions need not be metric Lipschitz but for certain subgroups they are intrinsically differentiable almost everywhere [16]. The most general statement for subgroups is not yet known.

Differentiability and Rademacher's theorem are also studied for functions between Banach spaces. There are versions of Rademacher's theorem for Gâteaux differentiability of Lipschitz functions, but the case of the stronger Fréchet differentiability is not fully understood [19]. One of the main ideas in the present article is that (almost local) maximality of directional derivatives implies differentiability.

This was first used by Preiss [24] to find points of Fréchet differentiability of Lipschitz functions on Banach spaces with separable dual. However, [24] does not give an 'almost everywhere' type result. Indeed, it is not known if three real-valued Lipschitz functions on a separable Hilbert space have a common point of Fréchet differentiability.

Cheeger [7] gave a generalization of Rademacher's theorem for Lipschitz functions defined on metric spaces equipped with a doubling measure and satisfying a Poincaré inequality. This has inspired much research in the area of analysis on metric measure spaces. Bate [3] showed that Cheeger differentiability is strongly related to existence of many directional derivatives. However, in this context, the emphasis is on differentiability almost everywhere, rather than pointwise differentiability.

A rather different direction of research asks whether one can find points of differentiability in extremely small sets. In particular, we can ask if Rademacher's theorem is sharp: given a set $N \subset \mathbb{R}^n$ of Lebesgue measure zero, does there exist a Lipschitz function $f \colon \mathbb{R}^n \to \mathbb{R}^m$ which is differentiable at no point of N?

If $n \leq m$ the answer is yes: for $n = 1$ this is rather easy [29], while the general case is very difficult and combines ongoing work of multiple authors [2, 9]. This implies that in Rademacher's theorem, the Lebesgue measure cannot be replaced by a singular measure. The recent paper [1] proves a Rademacher type theorem for an arbitrary finite measure, but in this case the directions of differentiability at almost every point depend on the measure.

If $n > m$ the answer to our question is no: there are Lebesgue measure zero sets $N \subset \mathbb{R}^n$ such that every Lipschitz function $f \colon \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a point of N. The case $m = 1$ was a surprising corollary of the previously mentioned result in Banach spaces by Preiss [24]. The case $m > 1$ were resolved by combining tools from the Banach space theory with a technique for avoiding porous sets [25]. In all cases, maximizing directional derivatives had a crucial role.

Sets $N \subset \mathbb{R}^n$ containing a point of differentiability for every real-valued Lipschitz function are now called universal differentiability sets. The argument in [24] was greatly refined to show that $\mathbb{R}^n$, $n > 1$, contains universal differentiability sets which are compact and of Hausdorff dimension one [10, 11]. This was improved to obtain a set which even has Minkowski dimension one [12].

In the present article we show that one can adapt the ideas of [24] to the Heisenberg group. Our main result is Theorem 2.12 which asserts the following: there is a Lebesgue measure zero set $N \subset \mathbb{H}^n$ such that every Lipschitz function $f \colon \mathbb{H}^n \to \mathbb{R}$ is Pansu differentiable at a point of N. This illustrates both the flexibility of Preiss' argument and the geometry of $\mathbb{H}^n$, which is rather far from being a Banach space. We now give an overview of the paper and information about the proof of Theorem 2.12. See Section 2 for the relevant definitions.

In Section 3 we first introduce directional derivatives in horizontal directions for Lipschitz functions $f \colon \mathbb{H}^n \to \mathbb{R}$, and compare the supremum of directional derivatives with the Lipschitz constant. We then construct simple horizontal curves joining the origin to other points, and use these curves to study Pansu differentiability of the Carnot-Carathéodory distance. Finally we show that existence of a maximal horizontal directional derivative implies Pansu differentiability (Theorem 3.6). This is an adaptation of a similar statement in Banach spaces [13, Theorem 2.4]. We do not claim that such a maximal horizontal directional derivative exists; 'almost maximal' horizontal directional derivatives have a role in the rest of the article.

In Section 4 we define our 'universal differentiability set' N, which may be chosen as any Lebesgue null G_δ set containing all horizontal lines joining points of $\mathbb{Q}^{2n+1}$. We construct useful horizontal curves inside this set, which allow us to modify a

piece of a horizontal line to pass through a nearby point, without distorting the length or Lipschitz constant too much.

In Section 5 we first estimate how horizontal lines diverge in $\mathbb{H}^n$ and study some simple $\mathbb{H}$-linear maps. We then state Theorem 5.6, which is one of two theorems that taken together will prove Theorem 2.12. Theorem 5.6 states that if $x \in N$ and E is a horizontal direction at which the directional derivative $Ef(x)$ is 'almost maximal', then f is Pansu differentiable at x. This is an adaptation of [24, Theorem 4.1]. Intuitively, almost maximality means that competing directional derivatives must satisfy an estimate which bounds changes in difference quotients by changes in directional derivatives. Such a bound is useful later when one wants to construct an almost maximal directional derivative. To prove Theorem 5.6 one uses a contradiction argument. If f is not differentiable at x then there is some nearby point where the change in f is too large. One modifies the line along which the directional derivative is large to form an auxillary curve inside N which passes through the nearby point. On this curve we find a point and direction giving a larger directional derivative and satisfying the required bound on difference quotients. This gives a contradiction to almost maximality. New ideas in our adaptation to $\mathbb{H}^n$ include a restriction to horizontal directional derivatives and choosing an auxillary curve which is horizontal, with carefully estimated length and direction.

In Section 6 we adapt [10, Theorem 3.1] to show that one can actually construct an almost maximal directional derivative in the sense of Theorem 5.6. Proposition 6.1 states that if $f_0 \colon \mathbb{H}^n \to \mathbb{R}$ is Lipschitz then one can find $f \colon \mathbb{H}^n \to \mathbb{R}$ such that $f - f_0$ is $\mathbb{H}$-linear and f has an almost maximal directional derivative at a point of N. This is done by constructing sequences $x_n \in N$, E_n and f_n such that $f_n - f_0$ is $\mathbb{H}$-linear and the directional derivatives $E_n f_n(x_n)$ are closer and closer to being almost maximal compared to the allowed competitors. Then $x_n \to x_*$ and $f_n \to f$ for some x_* and f since at every stage the changes were small, while $E_n \to E$ for some E since $\mathbb{H}$-linear perturbations are added to make directional derivatives in directions close to E_n larger. We ensure $x_* \in N$ by using the fact that N is G_δ. Finally one must show that the directional derivative $Ef(x)$ exists and is almost maximal; to prove this it is crucial that at each stage we maximize over a constrainted set of points and directions. Our argument follows very closely that of [10], modified to use horizontal directions, $\mathbb{H}$-linear maps and Hölder equivalence of the Carnot-Carathéodory and Euclidean distance. Finally we observe that combining Theorem 5.6 and Proposition 6.1 gives Theorem 2.12.

One might ask if Theorem 2.12 can be extended to general Carnot groups, or to construct compact sets of small dimension as in the Euclidean theory. Since this would require further new ideas and the proof of Theorem 2.12 is already complicated, we do not address this here.

2. Preliminaries

In this section we recall the Heisenberg group, horizontal curves, the Carnot-Carathéodory distance and Pansu differentiability [4, 8, 17, 18, 22, 28].

Denote the Euclidean norm and inner product by $|\cdot|$ and $\langle\cdot,\cdot\rangle$ respectively. We represent points of $\mathbb{R}^{2n+1}$ as triples (a, b, c), where $a, b \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

Definition 2.1. The Heisenberg group $\mathbb{H}^n$ is $\mathbb{R}^{2n+1}$ equipped with the non commutative group law:

$$(a, b, c)(a', b', c') = (a + a', \, b + b', \, c + c' - 2(\langle a, b' \rangle - \langle b, a' \rangle)).$$

The identity element in $\mathbb{H}^n$ is 0 and inverses are given by $x^{-1} = -x$.

Definition 2.2. For $r > 0$ define dilations $\delta_r \colon \mathbb{H}^n \to \mathbb{H}^n$ by:

$$\delta_r(a, b, c) = (ra, rb, r^2 c),$$

where $a, b \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

Dilations $\delta_r \colon \mathbb{H}^n \to \mathbb{H}^n$ and the projection $p \colon \mathbb{H}^n \to \mathbb{R}^{2n}$ onto the first $2n$ coordinates are group homomorphisms, where $\mathbb{R}^{2n}$ is considered as a group with the operation of addition.

As sets there is no difference between $\mathbb{H}^n$ and $\mathbb{R}^{2n+1}$. Nevertheless, we sometimes think of elements of $\mathbb{H}^n$ as points and elements of $\mathbb{R}^{2n+1}$ as vectors. Let e_i denote the standard basis vectors of $\mathbb{R}^{2n+1}$ for $1 \leq i \leq 2n+1$. That is, e_i has all coordinates equal to 0 except for a 1 in the i'th coordinate. Note that, in much of the literature on Carnot groups, the notation $\partial / \partial x_i$ is used instead of e_i. We next define a distinguished family of 'horizontal' directions.

Definition 2.3. For $1 \leq i \leq n$ define vector fields on $\mathbb{H}^n$ by:

$$X_i(a, b, c) = e_i + 2b_i e_{2n+1}, \quad Y_i(a, b, c) = e_{i+n} - 2a_i e_{2n+1}.$$

Let $V = \mathrm{Span}\{X_i, Y_i : 1 \leq i \leq n\}$ and ω be the inner product norm on V making $\{X_i, Y_i : 1 \leq i \leq n\}$ an orthonormal basis. We say that the elements of V are horizontal vector fields or horizontal directions.

An easy calculation shows that if $E \in V$ then

$$x(tE(0)) = x + tE(x) \tag{2.1}$$

for any $x \in \mathbb{H}^n$ and $t \in \mathbb{R}$. That is, 'horizontal lines' are preserved by group translations. If $E \in V$ then $E(0)$ is a vector $v \in \mathbb{R}^{2n+1}$ with $v_{2n+1} = 0$. Conversely, for any such v there exists $E \in V$ such that $E(0) = v$. If $E \in V$ then $p(E(x))$ is independent of x, so we can unambiguously define $p(E) \in \mathbb{R}^{2n}$. The norm ω is then equivalently given by $\omega(E) = |p(E)|$.

We now use the horizontal directions to define horizontal curves and horizontal length in $\mathbb{H}^n$. Let I denote a subinterval of $\mathbb{R}$. Recall that a map $\gamma \colon I \to \mathbb{R}^{2n+1}$ is absolutely continuous if it is differentiable almost everywhere and $\gamma(t) - \gamma(s) = \int_s^t \gamma'$ whenever $s, t \in I$.

Definition 2.4. An absolutely continuous curve $\gamma \colon I \to \mathbb{H}^n$ is a horizontal curve if there is $h \colon I \to \mathbb{R}^{2n}$ such that for almost every $t \in I$:

$$\gamma'(t) = \sum_{i=1}^{n} (h_i(t) X_i(\gamma(t)) + h_{n+i}(t) Y_i(\gamma(t))).$$

Define the horizontal length of such a curve by:

$$L_{\mathbb{H}}(\gamma) = \int_I |h|.$$

Notice that in Definition 2.4 we have $|(p \circ \gamma)'(t)| = |h(t)|$ for almost every t, so $L_{\mathbb{H}}(\gamma)$ is computed by integrating $|(p \circ \gamma)'(t)|$. That is, $L_{\mathbb{H}}(\gamma) = L_{\mathbb{E}}(p \circ \gamma)$, where $L_{\mathbb{E}}$ is the Euclidean length of a curve in Euclidean space. It can be shown that left group translations preserve horizontal lengths of horizontal curves.

In the next lemma we recall that horizontal curves are lifts of curves in $\mathbb{R}^{2n}$.

Lemma 2.5. *An absolutely continuous curve $\gamma\colon I \to \mathbb{H}^n$ is a horizontal curve if and only if for almost every $t \in I$:*

$$\gamma'_{2n+1}(t) = 2\sum_{i=1}^{n}(\gamma'_i(t)\gamma_{n+i}(t) - \gamma'_{n+i}(t)\gamma_i(t)).$$

Proof. By definition, γ is horizontal if and only if there exists $h\colon I \to \mathbb{R}^{2n}$ such that for almost every $t \in I$:

$$\gamma'(t) = \sum_{i=1}^{n}(h_i(t)X_i(\gamma(t)) + h_{n+i}(t)Y_i(\gamma(t))).$$

Using Definition 2.3, the right hand side of this expression is exactly

$$(h_1(t),\ldots,h_{2n}(t), 2\sum_{i=1}^{n}(h_i(t)\gamma_{n+i}(t) - 2h_{n+i}(t)\gamma_i(t))).$$

By examining the initial coordinates we see $\gamma'_i(t) = h_i(t)$ for $1 \le i \le 2n$. Hence:

$$\gamma'_{2n+1}(t) = 2\sum_{i=1}^{n}(\gamma'_i(t)\gamma_{n+i}(t) - \gamma'_{n+i}(t)\gamma_i(t))$$

for almost every $t \in I$. $\qquad\square$

Any two points of $\mathbb{H}^n$ can be joined by a horizontal curve of finite horizontal length. This is a particular instance of Chow's Theorem in subriemannian geometry. We use this fact to define the Carnot-Carathéodory distance.

Definition 2.6. Define the Carnot-Carathéodory distance d on $\mathbb{H}^n$ by:

$$d(x,y) = \inf\{L_{\mathbb{H}}(\gamma)\colon \gamma \text{ is a horizontal curve joining } x \text{ to } y\}.$$

Denote $d(x) = d(x,0)$ and $B_{\mathbb{H}}(x,r) := \{y \in \mathbb{H}^n\colon d(x,y) < r\}$.

It is known that geodesics exist in the Heisenberg group. That is, the infimum in Definition 2.6 is actually a minimum. The Carnot-Carathéodory distance respects the group law and dilations [4, (5.10) and (5.11)] - for every $g, x, y \in \mathbb{H}^n$ and $r > 0$:

- $d(gx, gy) = d(x,y)$,
- $d(\delta_r(x), \delta_r(y)) = rd(x,y)$.

Notice $d(x,y) \ge |p(y) - p(x)|$, since the projection of a horizontal curve joining x to y is a curve in $\mathbb{R}^{2n}$ joining $p(x)$ to $p(y)$. Using the known fact that projections of geodesics in $\mathbb{H}^n$ are arcs of circles, it is also possible to show [8, page 32]:

$$d(x) \ge \sqrt{|x_{2n+1}|}. \tag{2.2}$$

The Carnot-Carathéodory distance and the Euclidean distance are topologically equivalent but not Lipschitz equivalent. However, they are Hölder equivalent on compact sets [4, Corollary 5.2.10 and Proposition 5.15.1].

Proposition 2.7. *Let $K \subset \mathbb{H}^n$ be a compact set. Then there exists a constant $C_{\mathrm{H}} = C_{\mathrm{Hölder}}(K) \ge 1$ such that for any $x, y \in K$:*

$$|x - y|/C_{\mathrm{H}} \le d(x,y) \le C_{\mathrm{H}}|x - y|^{\frac{1}{2}}.$$

The Koranyi distance is Lipschitz equivalent to the Carnot-Carathéodory distance and given by the formula:

$$d_K(x,y) = \|x^{-1}y\|_K, \text{ where } \|(a,b,c)\|_K = (|(a,b)|^4 + c^2)^{\frac{1}{4}}. \tag{2.3}$$

We use the Koranyi distance occasionally to simplify some calculations.

We now make some simple observations about the distances introduced. They behave very simply if we follow a fixed horizontal direction.

Lemma 2.8. *If $E \in V$ then:*

- $|E(0)| = \omega(E) = d(E(0))$,
- $d(x, x + tE(x)) = t\omega(E)$ *for any $x \in \mathbb{H}^n$ and $t \in \mathbb{R}$.*

Proof. The first equality is true because $\omega(E) = |p(E)|$, as noted earlier, and $|p(E)| = |E(0)|$, since $E(0) \in \mathbb{R}^{2n+1}$ has final coordinate equal to 0.

The inequality $d(E(0)) \leq \omega(E)$ is trivial since $t \mapsto tE(0)$, $t \in [0,1]$, is a horizontal curve joining 0 to $E(0)$ of horizontal length exactly $\omega(E)$.

Suppose γ is a horizontal curve joining 0 and $E(0)$. Then $L_{\mathbb{H}}(\gamma) = L_{\mathbb{E}}(p \circ \gamma)$ as noted in the discussion after Definition 2.4. Since $p \circ \gamma$ is a curve joining 0 and $p(E)$ we deduce:

$$L_{\mathbb{E}}(p \circ \gamma) \geq |p(E)| = \omega(E).$$

This holds for any horizontal curve γ joining 0 to $E(0)$, so $d(E(0)) \geq \omega(E)$.

For the final statement we calculate as follows:

$$\begin{aligned}
d(x, x + tE(x)) &= d(x, x(tE(0))) \\
&= d(tE(0)) \\
&= t\omega(E).
\end{aligned}$$

$\square$

If $f \colon \mathbb{H}^n \to \mathbb{R}$ or $\gamma \colon \mathbb{R} \to \mathbb{H}^n$ we denote the Lipschitz constant (not necessarily finite) of f or γ with respect to d (in the domain or target respectively) by $\mathrm{Lip}_{\mathbb{H}}(f)$ and $\mathrm{Lip}_{\mathbb{H}}(\gamma)$. If we use the Euclidean distance then we use the notation $\mathrm{Lip}_{\mathbb{E}}(f)$ and $\mathrm{Lip}_{\mathbb{E}}(\gamma)$. Throughout this article 'Lipschitz' means with respect to the Carnot-Carathéodory distance if the domain or target is $\mathbb{H}^n$, unless otherwise stated. For horizontal curves we have the following relation between Lipschitz constants.

Lemma 2.9. *Suppose $\gamma \colon I \to \mathbb{H}^n$ is a horizontal curve. Then:*

$$\mathrm{Lip}_{\mathbb{H}}(\gamma) = \mathrm{Lip}_{\mathbb{E}}(p \circ \gamma).$$

Proof. Suppose $s, t \in I$ with $s < t$. Then:

$$|p(\gamma(t)) - p(\gamma(s))| \leq d(\gamma(t), \gamma(s)) \leq \mathrm{Lip}_{\mathbb{H}}(\gamma)|t - s|.$$

Hence $\mathrm{Lip}_{\mathbb{E}}(p \circ \gamma) \leq \mathrm{Lip}_{\mathbb{H}}(\gamma)$. For the opposite inequality, notice $\gamma|_{[s,t]}$ is a horizontal curve joining $\gamma(s)$ to $\gamma(t)$. Since $|(p \circ \gamma)'| \leq \mathrm{Lip}_{\mathbb{E}}(p \circ \gamma)$, we can estimate as follows:

$$\begin{aligned}
d(\gamma(s), \gamma(t)) &\leq L_{\mathbb{H}}(\gamma|_{[s,t]}) \\
&= \int_s^t |(p \circ \gamma)'| \\
&\leq \mathrm{Lip}_{\mathbb{E}}(p \circ \gamma)(t - s).
\end{aligned}$$

Hence $\mathrm{Lip}_{\mathbb{H}}(\gamma) \leq \mathrm{Lip}_{\mathbb{E}}(p \circ \gamma)$, which proves the lemma. $\square$

Lebesgue measure $\mathcal{L}^{2n+1}$ is the natural Haar measure on $\mathbb{H}^n$. It is compatible with group translations and dilations [4, page 44] - for every $g \in \mathbb{H}^n$, $r > 0$ and $A \subset \mathbb{H}^n$:

- $\mathcal{L}^{2n+1}(\{gx \colon x \in A\}) = \mathcal{L}^{2n+1}(A)$,
- $\mathcal{L}^{2n+1}(\delta_r(A)) = r^{2n+2}\mathcal{L}^{2n+1}(A)$.

We have defined group translations, dilations, distance and measure on $\mathbb{H}^n$. We can now introduce Pansu differentiability and state Pansu's theorem [23]. Intuitively, one replaces Euclidean objects by the corresponding ones in $\mathbb{H}^n$.

Definition 2.10. A function $L\colon \mathbb{H}^n \to \mathbb{R}$ is $\mathbb{H}$-linear if $L(xy) = L(x) + L(y)$ and $L(\delta_r(x)) = rL(x)$ for all $x, y \in \mathbb{H}^n$ and $r > 0$.

Let $f\colon \mathbb{H}^n \to \mathbb{R}$ and $x \in \mathbb{H}^n$. We say that f is Pansu differentiable at x if there is a $\mathbb{H}$-linear map $L\colon \mathbb{H}^n \to \mathbb{R}$ such that:

$$\lim_{y \to x} \frac{|f(y) - f(x) - L(x^{-1}y)|}{d(x, y)} = 0.$$

In this case we say that L is the Pansu derivative of f.

Clearly a $\mathbb{H}$-linear map is Pansu differentiable at every point. Pansu's theorem is the natural version of Rademacher's theorem in $\mathbb{H}^n$.

Theorem 2.11 (Pansu). *A Lipschitz function $f\colon \mathbb{H}^n \to \mathbb{R}$ is Pansu differentiable Lebesgue almost everywhere.*

We can now state formally our main result.

Theorem 2.12. *There is a Lebesgue measure zero set $N \subset \mathbb{H}^n$ such that every Lipschitz function $f\colon \mathbb{H}^n \to \mathbb{R}$ is Pansu differentiable at a point of N.*

3. Maximality of directional derivatives implies Pansu differentiability

We begin this section by defining directional derivatives in horizontal directions and comparing them to the Lipschitz constant (Lemma 3.3). We then construct simple horizontal curves connecting the origin to other points and estimate their horizontal length and direction (Lemma 3.4). We use these curves to investigate Pansu differentiability of the Carnot-Carathéodory distance (Lemma 3.5). Finally we prove that existence of a maximal horizontal directional derivative implies Pansu differentiability (Theorem 3.6).

Proposition 2.7 implies that curves in $\mathbb{H}^n$ which are Lipschitz with respect to the Carnot-Carathéodory distance are locally Lipschitz with respect to the Euclidean distance. Hence they are differentiable (in the usual sense) almost everywhere. We now show that derivatives of Lipschitz functions in horizontal directions can be defined by composing with any Lipschitz curve with tangent of the correct direction. Throughout this article C will denote a constant that may change from line to line but remains bounded.

Lemma 3.1. *Let $g, h\colon I \to \mathbb{H}^n$ be Lipschitz horizontal curves which are differentiable at $c \in I$ with $g(c) = h(c)$ and $g'(c) = h'(c)$. Let $f\colon \mathbb{H}^n \to \mathbb{R}$ be Lipschitz. If $(f \circ g)'(c)$ exists then $(f \circ h)'(c)$ exists and $(f \circ g)'(c) = (f \circ h)'(c)$.*

Proof. For convenience assume that $c = 0$ and I is a neighbourhood of 0. Using left group translations we may also assume that $g(0) = h(0) = 0$. Since f is Lipschitz, to prove the lemma it suffices to check that $d(g(t), h(t))/t \to 0$ as $t \to 0$. For this we use the Koranyi distance (2.3). An easy calculation gives:

$$d_K(g(t), h(t)) \leq C|h(t) - g(t)|$$
$$+ C\left| h_{2n+1}(t) - g_{2n+1}(t) + 2\sum_{i=1}^{n}(g_i(t)h_{n+i}(t) - g_{n+i}(t)h_i(t)) \right|^{\frac{1}{2}}.$$

Since $g(0) = h(0) = 0$ and $g'(0) = h'(0)$ we see:

$$|h(t) - g(t)|/t \to |h'(0) - g'(0)| = 0 \qquad \text{as } t \to 0. \tag{3.1}$$

Similarly for each $1 \leq i \leq n$:

$$(g_i(t)h_{n+i}(t) - g_{n+i}(t)h_i(t))/t^2 \to g_i'(0)h_{n+i}'(0) - g_{n+i}'(0)h_i'(0) = 0.$$

Hence it suffices to show that $(h_{2n+1}(t) - g_{2n+1}(t))/t^2 \to 0$ as $t \to 0$. Since h and g are Lipschitz we use Lemma 2.9 to see that for $1 \le i \le 2n$ and some constant C:

$$|h_i'(s) - g_i'(s)| \le \mathrm{Lip}_{\mathbb{E}}(p \circ h - p \circ g) = \mathrm{Lip}_{\mathbb{H}}(h - g) \le C.$$

Suppose without loss of generality that $t > 0$. We use Lemma 2.5 to estimate as follows:

$$\frac{|h_{2n+1}(t) - g_{2n+1}(t)|}{t^2} \le \frac{2}{t^2} \sum_{i=1}^{n} \int_0^t |(h_i' - g_i')(h_{n+i} - g_{n+i}) - (h_{n+i}' - g_{n+i}')(h_i - g_i)|$$

$$\le \frac{C}{t^2} \sum_{i=1}^{n} \int_0^t (|h_{n+i} - g_{n+i}| + |h_i - g_i|).$$

Let $\varepsilon > 0$. By (3.1) we know $|h(s) - g(s)| \le \varepsilon s$ for sufficiently small $s > 0$. Hence for sufficiently small $t > 0$:

$$\frac{|h_{2n+1}(t) - g_{2n+1}(t)|}{t^2} \le \frac{C}{t^2} \sum_{i=1}^{n} \int_0^t 2\varepsilon s \, \mathrm{d}s \le C\varepsilon,$$

for a new constant C depending on n. This shows $(h_{2n+1}(t) - g_{2n+1}(t))/t^2 \to 0$ as $t \to 0$, so concludes the proof. $\qquad\square$

Definition 3.2. Let $f\colon \mathbb{H}^n \to \mathbb{R}$ be a Lipschitz function and $E \in V$. Define $Ef(x) := (f \circ \gamma)'(t)$ whenever it exists, where γ is any Lipschitz horizontal curve with $\gamma(t) = x$ and $\gamma'(t) = E(x)$.

Notice that Lemma 3.1 implies Definition 3.2 makes sense. Suppose f is a Lipschitz function and γ is a Lipschitz horizontal curve. Then $f \circ \gamma\colon \mathbb{R} \to \mathbb{R}$ is Lipschitz, so differentiable almost everywhere. Hence Lipschitz functions have many directional derivatives in horizontal directions. We often use horizontal lines $\gamma(t) = x + tE(x)$ to calculate directional derivatives when they exist:

$$Ef(x) = \lim_{t \to 0} \frac{f(x + tE(x)) - f(x)}{t}.$$

Lipschitz constants with respect to the Euclidean distance are given by the supremum of directional derivatives over directions of Euclidean length 1. We now prove an analogue of this for the Carnot-Carathéodory distance.

Lemma 3.3. *Suppose $f\colon \mathbb{H}^n \to \mathbb{R}$ is Lipschitz. Then:*

$$\mathrm{Lip}_{\mathbb{H}}(f) = \sup\{|Ef(x)|\colon x \in \mathbb{H}^n, \ E \in V, \ \omega(E) = 1, \ Ef(x) \ exists\}.$$

Proof. Temporarily denote the right side of the above equality by $\mathrm{Lip}_{\mathrm{D}}(f)$. Fix $x, y \in \mathbb{H}^n$ and a Lipschitz horizontal curve $\gamma\colon [0, d(x,y)] \to \mathbb{H}^n$ joining x to y such that $|(p \circ \gamma)'(t)| = 1$ for almost every t. Let G be the set of $t \in [0, d(x,y)]$ for which:

- $(f \circ \gamma)'(t)$ exists,
- $\gamma'(t)$ exists,
- $\gamma'(t) \in \mathrm{Span}\{X_i(\gamma(t)), Y_i(\gamma(t))\colon 1 \le i \le n\}$,
- $|(p \circ \gamma)'(t)| = 1$.

Since γ is a horizontal curve and $f \circ \gamma$ is Lipschitz, we know that G has full measure. We estimate as follows:

$$|f(x) - f(y)| = \left| \int_0^{d(x,y)} (f \circ \gamma)' \right|$$

$$\le d(x,y) \sup\{|(f \circ \gamma)'(t)|\colon t \in G\}$$

$$\le d(x,y) \mathrm{Lip}_{\mathrm{D}}(f).$$

Here we used Definition 3.2: $\gamma'(t) \in \text{Span}\{X_i(\gamma(t)), Y_i(\gamma(t)) \colon 1 \le i \le n\}$ implies that there exists $E \in V$ with $E(\gamma(t)) = \gamma'(t)$, and $|(p \circ \gamma)'(t)| = 1$ then implies $\omega(E) = 1$. Hence $\text{Lip}_{\mathbb{H}}(f) \le \text{Lip}_D(f)$.

For the opposite inequality fix $x \in \mathbb{H}^n$ and $E \in V$ such that $\omega(E) = 1$ and $Ef(x)$ exists. Use Lemma 2.8 to estimate as follows:

$$\begin{aligned}
|Ef(x)| &= \left| \lim_{t \to 0} \frac{f(x + tE(x)) - f(x)}{t} \right| \\
&\le \limsup_{t \to 0} \frac{\text{Lip}_{\mathbb{H}}(f) d(x, x + tE(x))}{t} \\
&= \text{Lip}_{\mathbb{H}}(f).
\end{aligned}$$

Hence $\text{Lip}_D(f) \le \text{Lip}_{\mathbb{H}}(f)$ which concludes the proof. $\qquad\square$

The Carnot-Carathéodory distance d is invariant under left group translations. Hence to understand d it suffices to understand $d(x) = d(x, 0)$ for $x \in \mathbb{H}^n$. For this purpose we construct explicit Lipschitz horizontal curves joining 0 to points $x \in \mathbb{H}^n$. Our curves are simple concatenations of straight lines, but their Lipschitz constants and directions are sufficiently controlled for our applications.

Lemma 3.4. *Suppose $y \in \mathbb{H}^n$ with $p(y) \ne 0$. Write $y = (a, b, c)$ with $a, b \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Denote $L = |p(y)|$ and define $\gamma \colon [0, 1] \to \mathbb{H}^n$ by:*

$$\gamma(t) = \begin{cases} t\left(a - \frac{bc}{L^2}, b + \frac{ac}{L^2}, 0\right) & \text{if } 0 \le t \le 1/2, \\ \frac{1}{2}\left(a - \frac{bc}{L^2}, b + \frac{ac}{L^2}, 0\right) + \left(t - \frac{1}{2}\right)\left(a + \frac{bc}{L^2}, b - \frac{ac}{L^2}, 2c\right) & \text{if } 1/2 < t \le 1. \end{cases}$$

Then:

(1) *γ is a Lipschitz horizontal curve joining $(0, 0, 0) \in \mathbb{H}^n$ to $y = (a, b, c) \in \mathbb{H}^n$,*

(2) *$\text{Lip}_{\mathbb{H}}(\gamma) \le L\left(1 + \frac{c^2}{L^4} + \frac{4c^2}{L^2}\right)^{\frac{1}{2}}$,*

(3) *$\gamma'(t)$ exists and $|\gamma'(t) - (a, b, 0)| \le \frac{c}{L}(1 + 4L^2)^{\frac{1}{2}}$ for $t \in [0, 1] \setminus \{1/2\}$.*

We denote such a curve γ by γ_y.

Proof. Notice $\gamma(0) = 0$ and $\gamma(1) = (a, b, c)$. Recall that for $1 \le i \le n$:

$$X_i(a', b', c') = e_i + 2b_i' e_{2n+1}, \qquad Y_i(a', b', c') = e_{i+n} - 2a_i' e_{2n+1}.$$

For $t \in (0, 1/2)$ clearly

$$\gamma'(t) = \left(a - \frac{bc}{L^2}, b + \frac{ac}{L^2}, 0\right)$$

and an easy calculation shows that

$$\gamma'(t) = \sum_{i=1}^n \left(a_i - \frac{b_i c}{L^2}\right) X_i(\gamma(t)) + \left(b_i + \frac{a_i c}{L^2}\right) Y_i(\gamma(t)).$$

For $t \in (1/2, 1)$ we have:

$$\gamma'(t) = \left(a + \frac{bc}{L^2}, b - \frac{ac}{L^2}, 2c\right).$$

We verify that for $t \in (1/2, 1)$:

$$\gamma'(t) = \sum_{i=1}^n \left(a_i + \frac{b_i c}{L^2}\right) X_i(\gamma(t)) + \left(b_i - \frac{a_i c}{L^2}\right) Y_i(\gamma(t)).$$

Validity of this equality for the first $2n$ coordinates is clear. The final coordinate of the right side is given by:

$$\sum_{i=1}^{n} \left(a_i + \frac{b_i c}{L^2}\right)(2\gamma_{n+i}(t)) + \left(b_i - \frac{a_i c}{L^2}\right)(-2\gamma_i(t))$$

$$= \sum_{i=1}^{n} \left(a_i + \frac{b_i c}{L^2}\right)\left(\left(b_i + \frac{a_i c}{L^2}\right) + (2t-1)\left(b_i - \frac{a_i c}{L^2}\right)\right)$$

$$+ \left(b_i - \frac{a_i c}{L^2}\right)\left(-\left(a_i - \frac{b_i c}{L^2}\right) - (2t-1)\left(a_i + \frac{b_i c}{L^2}\right)\right)$$

$$= 2\sum_{i=1}^{n} \left(a_i + \frac{b_i c}{L^2}\right)\frac{a_i c}{L^2} + \left(b_i - \frac{a_i c}{L^2}\right)\frac{b_i c}{L^2}$$

$$= 2\sum_{i=1}^{n} \left(\frac{a_i^2 c}{L^2} + \frac{b_i^2 c}{L^2}\right)$$

$$= 2c.$$

Hence γ is a horizontal curve. Since γ is Lipschitz with respect to the Euclidean distance, Lemma 2.9 implies that γ is a Lipschitz horizontal curve. This proves (1).

A straightforward computation shows that:

$$|\gamma'(t)| = \begin{cases} L\left(1 + \frac{c^2}{L^4}\right)^{\frac{1}{2}} & \text{if } 0 \le t < 1/2, \\ L\left(1 + \frac{c^2}{L^4} + \frac{4c^2}{L^2}\right)^{\frac{1}{2}} & \text{if } 1/2 < t \le 1. \end{cases}$$

Using Lemma 2.9 this gives the desired estimate of the Lipschitz constant:

$$\text{Lip}_{\mathbb{H}}(\gamma) \le \text{Lip}_{\mathbb{E}}(\gamma) \le L\left(1 + \frac{c^2}{L^4} + \frac{4c^2}{L^2}\right)^{\frac{1}{2}}.$$

This proves (2). The estimate in (3) is also straightforward. $\qquad\square$

Next we will study the Carnot-Carathéodory distance near points of the form $u = E(0)$ for some $E \in V$.

Lemma 3.5. *Fix $u_1, u_2 \in \mathbb{R}^n$ not both zero and let $u = (u_1, u_2, 0) \in \mathbb{H}^n$. Then:*
  (1) $d(uz) \ge d(u) + \langle z, u/d(u)\rangle$ *for any $z \in \mathbb{H}^n$,*
  (2) $d(uz) = d(u) + \langle z, u/d(u)\rangle + o(d(z))$ *as $z \to 0$. That is, the Pansu derivative of d at u is $L(x) := \langle x, u/d(u)\rangle$.*

Proof. We may assume that $d(u) = 1$ throughout the proof, since the general statement can be deduced using dilations. To prove (1) first recall $d(x) \ge |p(x)|$ for all $x \in \mathbb{H}^n$, while Lemma 2.8 shows $d(u) = |p(u)|$ for our particular choice of u. Clearly also $\langle p(z), p(u)\rangle = \langle z, u\rangle$ for such u. We use Pythagoras' theorem and $d(u) = |p(u)| = 1$ to estimate as follows:

$$d(uz) \ge |p(uz)|$$

$$= |p(u) + p(z)|$$

$$= |p(u)(1 + \langle p(z), p(u)\rangle) + (p(z) - \langle p(z), p(u)\rangle p(u))|$$

$$\ge |p(u)|(1 + \langle p(z), p(u)\rangle)$$

$$= |p(u)| + \langle p(z), p(u)\rangle$$

$$= d(u) + \langle z, u\rangle.$$

To prove (2) it suffices to show that $d(uz) \le d(u) + \langle z, u/d(u)\rangle + o(d(z))$ as $z \to 0$. Let $(a, b, c) = uz$ and $L = |p(uz)| = |p(u) + p(z)|$. Assume $d(z) \le 1/2$.

Using $|p(u)| = 1$ and $|p(z)| \leq d(z) \leq 1/2$, we see $1/2 \leq L \leq 2$. Using the formula for the group law and $|z_{2n+1}| \leq d(z)^2$ from (2.2), we can estimate c as follows:

$$
\begin{aligned}
|c| &\leq |z_{2n+1}| + 4|p(u)||p(z)| \\
&\leq d(z)^2 + 4d(z) \\
&\leq 5d(z).
\end{aligned}
$$

Lemma 3.4 and the definition of the Carnot-Carathéodory distance gives:

$$
\begin{aligned}
d(zw) &\leq L\Big(1 + \frac{c^2}{L^4} + \frac{4c^2}{L^2}\Big)^{\frac{1}{2}} \\
&\leq L(1 + 800d(z)^2)^{\frac{1}{2}} \\
&\leq L + o(d(z)).
\end{aligned}
$$

To conclude the proof of (2) we claim that $L \leq 1 + \langle z, u \rangle + o(d(z))$. Estimate as follows:

$$
\begin{aligned}
L &= |p(u) + p(z)| \\
&= |p(u)(1 + \langle p(z), p(u)\rangle) + (p(z) - \langle p(z), p(u)\rangle p(u))| \\
&= ((1 + \langle p(z), p(u)\rangle)^2 + |p(z) - \langle p(z), p(u)\rangle p(u)|^2)^{\frac{1}{2}} \\
&\leq ((1 + \langle p(z), p(u)\rangle)^2 + 4d(z)^2)^{\frac{1}{2}} \\
&= (1 + \langle p(z), p(u)\rangle)\Big(1 + \frac{4d(z)^2}{(1 + \langle p(z), p(u)\rangle)^2}\Big)^{\frac{1}{2}} \\
&\leq (1 + \langle z, u \rangle)\Big(1 + \frac{2d(z)^2}{(1 + \langle z, u \rangle)^2}\Big).
\end{aligned}
$$

The claim then follows since $d(z)/(1 + \langle z, u \rangle) \leq 2d(z) \to 0$ as $d(z) \to 0$. $\square$

We now use Lemma 3.5 to show that existence of a maximal horizontal directional derivative implies Pansu differentiability. This is an adaptation of [13, Theorem 2.4]. By Lemma 3.3, existence of a maximal horizontal directional derivative is equivalent to the agreement of a directional derivative with the Lipschitz constant.

Theorem 3.6. *Let $f : \mathbb{H}^n \to \mathbb{R}$ be Lipschitz, $x \in \mathbb{H}^n$ and $E \in V$ with $\omega(E) = 1$. Suppose $Ef(x)$ exists and $Ef(x) = \mathrm{Lip}_{\mathbb{H}}(f)$. Then f is Pansu differentiable at x with derivative $L(x) := \mathrm{Lip}_{\mathbb{H}}(f)\langle x, E(0)\rangle = \mathrm{Lip}_{\mathbb{H}}(f)\langle p(x), p(E)\rangle$.*

Proof. Let $0 < \varepsilon \leq 1/2$. Using Lemma 3.5, we can choose $0 < \alpha \leq \varepsilon$ such that whenever $d(z) \leq \alpha$:

$$
d(E(0)z) - d(E(0)) \leq \langle z, E(0)\rangle + \varepsilon d(z).
$$

Use existence of $Ef(x)$ to fix $\delta > 0$ such that whenever $|t| \leq \delta$:

$$
|f(x + tE(x)) - f(x) - tEf(x)| \leq \alpha^2 |t|.
$$

Suppose that $0 < d(w) \leq \alpha\delta$ and $t = \alpha^{-1}d(w)$. Then $0 < t \leq \delta$, $d(\delta_{t^{-1}}(w)) = \alpha$ and $2d(w) = 2\alpha t \leq t$. Recall that $\omega(E) = 1$ implies $d(E(0)) = 1$. We use also left

invariance of the Carnot-Carathéodory distance to estimate as follows:

$$\begin{aligned}
f(xw) - f(x) &= (f(xw) - f(x - tE(x))) + (f(x - tE(x)) - f(x)) \\
&\leq \mathrm{Lip}_{\mathbb{H}}(f)d(xw, x - tE(x)) - tEf(x) + \alpha^2 t \\
&= \mathrm{Lip}_{\mathbb{H}}(f)d(xw, x(-tE(0))) - tEf(x) + \alpha^2 t \\
&= \mathrm{Lip}_{\mathbb{H}}(f)d((tE(0))w) - t\mathrm{Lip}_{\mathbb{H}}(f) + \alpha^2 t \\
&= t\mathrm{Lip}_{\mathbb{H}}(f)(d(E(0)\delta_{t^{-1}}(w)) - d(E(0))) + \alpha^2 t \\
&\leq t\mathrm{Lip}_{\mathbb{H}}(f)(\langle \delta_{t^{-1}}(w), E(0)\rangle + \varepsilon d(\delta_{t^{-1}}(w))) + \alpha^2 t \\
&\leq \mathrm{Lip}_{\mathbb{H}}(f)\langle w, E(0)\rangle + \varepsilon \mathrm{Lip}_{\mathbb{H}}(f)d(w) + \alpha d(w) \\
&\leq \mathrm{Lip}_{\mathbb{H}}(f)\langle w, E(0)\rangle + \varepsilon(\mathrm{Lip}_{\mathbb{H}}(f) + 1)d(w).
\end{aligned}$$

For the opposite inequality we have:

$$\begin{aligned}
f(xw) - f(x) &= (f(xw) - f(x + tE(x))) + (f(x + tE(x)) - f(x)) \\
&\geq -\mathrm{Lip}_{\mathbb{H}}(f)d(xw, x + tE(x)) + tEf(x) - \alpha^2 t \\
&= -\mathrm{Lip}_{\mathbb{H}}(f)d(xw, x(tE(0))) + tEf(x) - \alpha^2 t \\
&= -\mathrm{Lip}_{\mathbb{H}}(f)d((-tE(0))w) + t\mathrm{Lip}_{\mathbb{H}}(f) - \alpha^2 t \\
&= -t\mathrm{Lip}_{\mathbb{H}}(f)(d((-E(0))\delta_{t^{-1}}(w)) - d(E(0))) - \alpha^2 t \\
&\geq -t\mathrm{Lip}_{\mathbb{H}}(f)(\langle \delta_{t^{-1}}(w), -E(0)\rangle + \varepsilon d(\delta_{t^{-1}}(w))) - \alpha^2 t \\
&\geq -\mathrm{Lip}_{\mathbb{H}}(f)\langle w, -E(0)\rangle - \varepsilon \mathrm{Lip}_{\mathbb{H}}(f)d(w) - \alpha^2 t \\
&\geq \mathrm{Lip}_{\mathbb{H}}(f)\langle w, E(0)\rangle - \varepsilon \mathrm{Lip}_{\mathbb{H}}(f)d(w) - \alpha d(w) \\
&= \mathrm{Lip}_{\mathbb{H}}(f)\langle w, E(0)\rangle - \varepsilon(\mathrm{Lip}_{\mathbb{H}}(f) + 1)d(w).
\end{aligned}$$

This shows that $d(w) \leq \alpha\delta$ implies:

$$|f(xw) - f(x) - \mathrm{Lip}_{\mathbb{H}}(f)\langle w, E(0)\rangle| \leq (\mathrm{Lip}_{\mathbb{H}}(f) + 1)\varepsilon d(w).$$

Hence f is Pansu differentiable at x with derivative $\mathrm{Lip}_{\mathbb{H}}(f)\langle \cdot, E(0)\rangle$. □

An arbitrary Lipschitz function may not have a maximal horizontal directional derivative as in Theorem 3.6. Construction of almost locally maximal horizontal directional derivatives inside a measure zero set, and deducing Pansu differentiability at such points, is the content of the rest of the paper.

4. The universal differentiability set and horizontal curves

In this section we identify our measure zero universal differentiability set (Lemma 4.1) and construct useful horizontal curves inside this set (Lemma 4.2). Recall that a set in a topological space is G_δ if it is a countable intersection of open sets.

Lemma 4.1. *There is a Lebesgue measure zero G_δ set $N \subset \mathbb{H}^n$ containing all straight lines which are also horizontal curves and join pairs of points of $\mathbb{Q}^{2n+1}$. Any such set contains the image of:*

(1) *the line $x + tE(x)$ whenever $x \in \mathbb{Q}^{2n+1}$ and $E \in V$ is a linear combination of $\{X_i, Y_i : 1 \leq i \leq n\}$ with rational coefficients,*

(2) *all curves of the form $x\gamma_y$ for $x, y \in \mathbb{Q}^{2n+1}$ with $p(y) \neq 0$, where γ_y is the curve constructed in Lemma 3.4.*

Proof. There exist open sets of of arbitrarily small measure containing a fixed line. By taking a union of countably many such open sets with decreasing measure, we find an open set of arbitrarily small measure containing all horizontal lines joining points of $\mathbb{Q}^{2n+1}$. By taking a countable intersection of such sets with measures converging to 0, we obtain the required measure zero G_δ set N.

The horizontal line $x + tE(x)$ joins x to $x + E(x)$ so its image is a subset of N whenever $x \in \mathbb{Q}^{2n+1}$ and $E \in V$ is a linear combination of $\{X_i, Y_i : 1 \leq i \leq n\}$ with rational coefficients. By examining the formula for γ_y in Lemma 3.4, we see each curve $x\gamma_y$, for $x, y \in \mathbb{Q}^n$ with $p(y) \neq 0$, is a union of two such horizontal lines. This proves the lemma. $\qquad \square$

To prove Theorem 5.6 (an almost maximal directional derivative implies Pansu differentiability) we will modify horizontal line segments (along which a Lipschitz function will have a large directional derivative) to pass through nearby points (which intuitively show non-Pansu differentiability at some point). In the next lemma we see how to do this without changing the length or direction of the line too much.

Lemma 4.2. *Given $\eta > 0$, there is $0 < \Delta(\eta) < 1/2$ and $C_{\mathrm{m}} = C_{\mathrm{modify}} \geq 1$ such that the following holds whenever $0 < \Delta < \Delta(\eta)$. Suppose:*

- $x, u \in \mathbb{H}^n$ *with* $d(u) \leq 1$,
- $E \in V$ *with* $\omega(E) = 1$,
- $0 < r < \Delta$ *and* $s := r/\Delta$.

Then there is a Lipschitz horizontal curve $g \colon \mathbb{R} \to \mathbb{H}^n$ such that:

(1) $g(t) = x + tE(x)$ *for* $|t| \geq s$,
(2) $g(\zeta) = x\delta_r(u)$, *where* $\zeta := r\langle u, E(0)\rangle$,
(3) $\mathrm{Lip}_{\mathbb{H}}(g) \leq 1 + \eta\Delta$,
(4) $g'(t)$ *exists and* $|(p \circ g)'(t) - p(E)| \leq C_{\mathrm{m}}\Delta$ *for* $t \in \mathbb{R}$ *outside a finite set.*

Suppose additionally $x, u \in \mathbb{Q}^{2n+1}$, E is a linear combination of $\{X_i, Y_i : 1 \leq i \leq n\}$ with rational coefficients and $r, s \in \mathbb{Q}$. Then g is a concatenation of curves from Lemma 4.1(1,2).

Proof. The distance d is invariant under left group translations and the group law is linear in the first $2n$ coordinates. Hence to prove (1)–(4) we can assume $x = 0$. From the proof it will be clear that this also suffices for the final statement.

For $|t| \geq s$ the curve $g(t)$ is explicitly defined by (1) and satisfies (3) and (4). To define $g(t)$ for $|t| < s$ we consider the two cases $-s < t \leq \zeta$ and $\zeta \leq t < s$. These are similar so we show how to define the curve for $-s < t \leq \zeta$. By using left group translations by $\pm sE(0)$ and reparameterizing the curve, it suffices to show the following claim.

Claim. Suppose $0 < \Delta < \eta/1632$. Then there exists a Lipschitz horizontal curve $\varphi \colon [0, s + \zeta] \to \mathbb{H}^n$ such that $\varphi(0) = 0$, $\varphi(s + \zeta) = (sE(0))\delta_r(u)$ and:

(A) $\mathrm{Lip}_{\mathbb{H}}(\varphi) \leq 1 + \eta\Delta$,
(B) $\varphi'(t)$ exists and $|(p \circ \varphi)'(t) - p(E)| \leq 184\Delta$ for $t \in [0, s + \zeta]$ outside a finite set.

Proof of Claim. Let $(a, b, c) := (sE(0))\delta_r(u)$ and $L := |(a, b)|$. Observe:

$$L = |sp(E) + rp(u)| = s|p(E) + \Delta p(u)|.$$

Our assumptions imply $|p(E)| = \omega(E) = 1$ and $|p(u)| \leq d(u) \leq 1$. Using also (2.2) gives $|u_{2n+1}| \leq 1$. Since $0 < \Delta < 1/2$ we deduce $s/2 \leq L \leq 2s$. Definition 2.1 and Definition 2.2 give:

$$|c| \leq r^2|u_{2n+1}| + 4rs \leq r^2 + 4rs \leq 5rs.$$

Lemma 3.4 provides a Lipschitz horizontal curve $\gamma \colon [0, 1] \to \mathbb{H}^n$ joining 0 to (a, b, c) such that:

(A') $\mathrm{Lip}_{\mathbb{H}}(\gamma) \leq L\left(1 + \frac{c^2}{L^4} + \frac{4c^2}{L^2}\right)^{\frac{1}{2}}$,
(B') $\gamma'(t)$ exists and $|\gamma'(t) - (a, b, 0)| \leq \frac{c}{L}(1 + 4L^2)^{\frac{1}{2}}$ for $t \in [0, 1] \setminus \{1/2\}$.

We verify the claim with $\varphi\colon [0, s+\zeta] \to \mathbb{H}^n$ defined by $\varphi(t) = \gamma(t/(s+\zeta))$. Notice that φ is a Lipschitz horizontal curve with $\varphi(0) = 0$ and $\varphi(s+\zeta) = (sE(0))\delta_r(u)$.

Proof of (A). We first develop the estimate (A'). For this we use our estimates of c and L, the inequality $s < 1$, and the equality $r = \Delta s$:

$$\mathrm{Lip}_{\mathbb{H}}(\gamma) \le L\Big(1 + \frac{c^2}{L^4} + \frac{4c^2}{L^2}\Big)^{\frac{1}{2}}$$

$$\le L\Big(1 + 400(r^2/s^2) + 400r^2\Big)^{\frac{1}{2}}$$

$$\le L\Big(1 + 800\Delta^2\Big)^{\frac{1}{2}}$$

$$\le L + 800s\Delta^2.$$

To estimate $L = |p((sE)(\delta_r(u)))|$ more carefully, first recall:

$$\zeta = r\langle u, E(0)\rangle = r\langle p(u), p(E)\rangle.$$

We use the orthogonal decomposition:

$$p((sE)\delta_r(u)) = (s+\zeta)p(E) + (rp(u) - \zeta p(E)). \qquad (4.1)$$

Using $d(u) \le 1$ and $\omega(E) \le 1$ gives $|\zeta| \le r \le s/2$ and $|rp(u) - \zeta p(E)| \le 2r$. We estimate as follows:

$$L = ((s+\zeta)^2 + |rp(u) - \zeta p(E)|^2)^{\frac{1}{2}}$$

$$\le ((s+\zeta)^2 + 4r^2)^{\frac{1}{2}}$$

$$= ((s+\zeta)^2 + 4\Delta^2 s^2)^{\frac{1}{2}}$$

$$= (s+\zeta)(1 + 4\Delta^2 s^2/(s+\zeta)^2)^{\frac{1}{2}}$$

$$\le (s+\zeta)(1 + 16\Delta^2)^{\frac{1}{2}}$$

$$\le (s+\zeta)(1 + 8\Delta^2)$$

$$\le s + \zeta + 16s\Delta^2.$$

Putting together these estimates gives:

$$\mathrm{Lip}_{\mathbb{H}}(\varphi) \le \mathrm{Lip}_{\mathbb{H}}(\gamma)/(s+\zeta)$$

$$\le (s + \zeta + 816s\Delta^2)/(s+\zeta)$$

$$\le 1 + 1632\Delta^2.$$

Hence $\mathrm{Lip}_{\mathbb{H}}(\varphi) \le 1 + \eta\Delta$ since $0 < \Delta < \eta/1632$. This proves (A).

Proof of (B). The decomposition (4.1) implies:

$$|p((sE(0))\delta_r(u)) - (s+\zeta)p(E)| \le 2r = 2\Delta s.$$

Since $E \in V$ and $(a, b, c) = (sE(0))\delta_r(u)$ we deduce:

$$|(a, b, 0) - (s+\zeta)E(0)| \le 2\Delta s.$$

Using (B') shows that for $t \in [0, 1] \setminus \{1/2\}$:

$$|\gamma'(t) - (s+\zeta)E(0)| \le 2\Delta s + \frac{c}{L}(1 + 4L^2)^{\frac{1}{2}}$$

$$\le 2\Delta s + 10\Delta s(1 + 16s^2)^{\frac{1}{2}}$$

$$\le 92\Delta s.$$

Hence for $t \in [0, s + \zeta] \setminus \{(s + \zeta)/2\}$:

$$|(p \circ \varphi)'(t) - p(E)| \leq |\varphi'(t) - E(0)|$$
$$\leq 92\Delta s/(s + \zeta)$$
$$\leq 184\Delta.$$

This verifies (B). The final statement of the lemma is clear from the construction.

$\square$

5. ALMOST MAXIMALITY OF DIRECTIONAL DERIVATIVES IMPLIES PANSU DIFFERENTIABILITY

In this section we first estimate how Lipschitz horizontal curves with the same starting point and moving in similar directions stay close together (Lemma 5.1). We then give simple properties of the maps $x \mapsto \langle x, E(0) \rangle$ (Lemma 5.2) and quote a mean value estimate (Lemma 5.3) by Preiss [24, Lemma 3.4]. Finally we show that existence of an almost locally maximal horizontal directional derivative implies Pansu differentiability (Theorem 5.6).

Lemma 5.1. *Given $S > 0$, there is a constant $C_{\mathrm{a}} = C_{\mathrm{angle}}(S) \geq 1$ for which the following is true. Suppose:*

- *$g, h \colon I \to \mathbb{H}^n$ are Lipschitz horizontal curves with $\mathrm{Lip}_{\mathbb{H}}(g), \mathrm{Lip}_{\mathbb{H}}(h) \leq S$,*
- *$g(c) = h(c)$ for some $c \in I$,*
- *there exists $0 \leq A \leq 1$ such that $|(p \circ g)'(t) - (p \circ h)'(t)| \leq A$ for almost every $t \in I$.*

Then $d(g(t), h(t)) \leq C_{\mathrm{a}}\sqrt{A}|t - c|$ for every $t \in I$.

Proof. Assume $c = 0 \in I$ and, using left group translations, $g(0) = h(0) = 0$. We estimate using the equivalent Koranyi distance (2.3):

$$d(g(t), h(t)) \leq C d_K(g(t), h(t)) \tag{5.1}$$

$$\leq C \sum_{i=1}^{2n} |h_i(t) - g_i(t)|$$

$$+ C \left| h_{2n+1}(t) - g_{2n+1}(t) + 2 \sum_{i=1}^{n} (g_i(t)h_{n+i}(t) - g_{n+i}(t)h_i(t)) \right|^{\frac{1}{2}}.$$

Let $1 \leq j \leq 2n$. Using $|(p \circ g)'(t) - (p \circ h)'(t)| \leq A$ almost everywhere implies $|h_j(t) - g_j(t)| \leq A|t|$ for every $t \in I$. Lemma 2.9 and $\mathrm{Lip}_{\mathbb{H}}(g) \leq S$ give the inequality $\mathrm{Lip}_{\mathbb{E}}(g_j) \leq S$. Using also $g(0) = 0$ then gives $|g_j(t)| \leq S|t|$ for $t \in I$. For $1 \leq i \leq n$ and $t \in I$:

$$|g_i(t)h_{n+i}(t) - g_{n+i}(t)h_i(t)| = |g_i(t)(h_{n+i}(t) - g_{n+i}(t)) + g_{n+i}(t)(g_i(t) - h_i(t))|$$
$$\leq S|t||h_{n+i}(t) - g_{n+i}(t)| + S|t||g_i(t) - h_i(t)|$$
$$\leq SAt^2.$$

We estimate the final term using Lemma 2.5:

$$|h_{2n+1}(t) - g_{2n+1}(t)| \leq 2 \sum_{i=1}^{n} \int_0^t |(h_i' - g_i')(h_{n+i} - g_{n+i}) - (h_{n+i}' - g_{n+i}')(h_i - g_i)|$$

$$\leq 4A^2 \sum_{i=1}^{n} \int_0^t s \, \mathrm{d}s$$

$$= 2nA^2t^2.$$

Combining our estimates of each term in (5.1) gives $d(g(t), h(t)) \leq C\sqrt{A}|t|$ for $t \in I$, where C is a constant depending on S.

$\square$

We will use the maps $x \mapsto \langle x, E(0) \rangle$ for $E \in V$ both as Pansu derivatives and as perturbations to construct an almost maximal directional derivative in the proof of Proposition 6.1. We now give simple properties of these maps.

Lemma 5.2. *Suppose $E \in V$ with $\omega(E) = 1$ and let $L \colon \mathbb{H}^n \to \mathbb{R}$ be the function $L(x) = \langle x, E(0) \rangle$. Then:*
  (1) *L is $\mathbb{H}$-linear and $\mathrm{Lip}_{\mathbb{H}}(L) = 1$,*
  (2) *for $x \in \mathbb{H}^n$ and $\widetilde{E} \in V$:*
$$\widetilde{E} L(x) = L(\widetilde{E}(0)) = \langle p(\widetilde{E}), p(E) \rangle.$$

Proof. Since $E \in V$ we know that the final coordinate of $E(0)$ is 0. Suppose $x, y \in \mathbb{H}^n$ and $r > 0$. In the first $2n$ coordinates the group product and dilations are Euclidean, hence:
$$L(x) + L(y) = \langle x + y, E(0) \rangle = \langle xy, E(0) \rangle = L(xy)$$
and
$$L(\delta_r(x)) = \langle \delta_r(x), E(0) \rangle = r \langle x, E(0) \rangle = rL(x).$$
This shows that L is $\mathbb{H}$-linear.

It follows from $\mathbb{H}$-linearity that $L(x) - L(y) = L(xy^{-1})$. To bound the Lipschitz constant from above it is enough to use $\omega(E) = 1$ to observe:
$$|\langle x, E(0) \rangle| \leq |p(x)| \leq d(x).$$
Conversely, $L(E(0)) - L(0) = \langle E(0), E(0) \rangle = 1$, so equals $d(E(0))$ by Lemma 2.8. This proves (1).

To prove (2) we observe:
$$\begin{aligned}
\widetilde{E} L(x) &= \lim_{t \to 0} \frac{L(x + t\widetilde{E}(x)) - L(x)}{t} \\
&= \lim_{t \to 0} \left\langle \frac{x + t\widetilde{E}(x) - x}{t}, E(0) \right\rangle \\
&= \langle \widetilde{E}(x), E(0) \rangle \\
&= \langle \widetilde{E}(0), E(0) \rangle.
\end{aligned}$$
$\square$

A key feature of the special pairs used to define almost maximal directional derivatives will be that changes in difference quotients are bounded by changes in directional derivatives. We use the following lemma [24, Lemma 3.4].

Lemma 5.3. *Suppose $|\zeta| < s < \rho$, $0 < v < 1/32$, $\sigma > 0$ and $L > 0$ are real numbers. Let $\varphi, \psi \colon \mathbb{R} \to \mathbb{R}$ satisfy $\mathrm{Lip}_{\mathbb{E}}(\varphi) + \mathrm{Lip}_{\mathbb{E}}(\psi) \leq L$, $\varphi(t) = \psi(t)$ for $|t| \geq s$ and $\varphi(\zeta) \neq \psi(\zeta)$. Suppose, moreover, that $\psi'(0)$ exists and that*
$$|\psi(t) - \psi(0) - t\psi'(0)| \leq \sigma L|t|$$
whenever $|t| \leq \rho$,
$$\rho \geq s\sqrt{(sL)/(v|\varphi(\zeta) - \psi(\zeta)|)},$$
and
$$\sigma \leq v^3 \left(\frac{\varphi(\zeta) - \psi(\zeta)}{sL} \right)^2.$$
Then there is $\tau \in (-s, s) \setminus \{\zeta\}$ such that $\varphi'(\tau)$ exists,
$$\varphi'(\tau) \geq \psi'(0) + v|\varphi(\zeta) - \psi(\zeta)|/s,$$
and
$$|(\varphi(\tau + t) - \varphi(\tau)) - (\psi(t) - \psi(0))| \leq 4(1 + 20v)\sqrt{(\varphi'(\tau) - \psi'(0))L}|t|$$

for every $t \in \mathbb{R}$.

Remark 5.4. By examining the proof of Lemma 5.3 in [24] it is easy to see that τ can additionally be chosen outside a given Lebesgue measure zero subset of $\mathbb{R}$. A stronger observation, that τ can be chosen outside a set of sufficiently small yet positive measure, is used in [24] to prove [24, Theorem 6.3].

We can now prove that existence of an almost locally maximal horizontal directional derivative implies Pansu differentiability. The argument is based on that of [24, Theorem 4.1], but we use our analysis of the Carnot-Carathéodory distance and use exclusively horizontal curves and directional derivatives in horizontal directions.

Notation 5.5. Fix a Lebesgue null G_δ set $N \subset \mathbb{H}^n$ as in Lemma 4.1 for the remainder of the article. For any Lipschitz function $f \colon \mathbb{H}^n \to \mathbb{R}$ define:
$$D^f := \{(x, E) \in N \times V \colon \omega(E) = 1,\ Ef(x) \text{ exists}\}.$$

Theorem 5.6. *Let $f \colon \mathbb{H}^n \to \mathbb{R}$ be a Lipschitz function with $\mathrm{Lip}_{\mathbb{H}}(f) \leq 1/2$. Suppose $(x_*, E_*) \in D^f$. Let M denote the set of pairs $(x, E) \in D^f$ such that $Ef(x) \geq E_*f(x_*)$ and*
$$|(f(x + tE_*(x)) - f(x)) - (f(x_* + tE_*(x_*)) - f(x_*))|$$
$$\leq 6|t|((Ef(x) - E_*f(x_*))\mathrm{Lip}_{\mathbb{H}}(f))^{\frac{1}{4}}$$
for every $t \in (-1, 1)$. If
$$\lim_{\delta \downarrow 0} \sup\{Ef(x) \colon (x, E) \in M \text{ and } d(x, x_*) \leq \delta\} \leq E_*f(x_*)$$
then f is Pansu differentiable at x_* with Pansu derivative
$$L(x) = E_*f(x_*)\langle x, E_*(0)\rangle = E_*f(x_*)\langle p(x), p(E_*)\rangle.$$

Remark 5.7. Since we will apply Lemma 5.3, it may seem more intuitive to instead bound $|(f(x + tE(x)) - f(x)) - (f(x_* + tE_*(x_*)) - f(x_*))|$ in the statement of Theorem 5.6. The precise form in Theorem 5.6 will be useful when we construct an almost locally maximal horizontal directional derivative in Proposition 6.1.

We will prove Theorem 5.6 by contradiction. We first use Lemma 4.2 to modify the line $x_* + tE_*(x_*)$ to form a Lipschitz horizontal curve g in N which passes through a nearby point showing non-Pansu differentiability at x_*. We then apply Lemma 5.3 with $\varphi = f \circ g$ to obtain a large directional derivative along g and estimates for difference quotients in the new direction. We then develop these estimates to show that the new point and direction form a pair in M. This shows that there is a nearby pair in M giving a larger directional derivative than (x_*, E_*), a contradiction.

Proof of Theorem 5.6. We can assume $\mathrm{Lip}_{\mathbb{H}}(f) > 0$ since otherwise the statement is trivial. Let $\varepsilon > 0$ and fix various parameters as follows.

Parameters. Choose:
1. $0 < v < 1/32$ such that $4(1 + 20v)\sqrt{(2 + v)/(1 - v)} + v < 6$,
2. $\eta = \varepsilon v^3/3200$,
3. $\Delta(\eta/2)$, C_{m} and $C_{\mathrm{a}} = C_{\mathrm{angle}}(2)$ using Lemma 4.2 and Lemma 5.1,
4. rational $0 < \Delta < \min\{\eta v^2,\ \Delta(\eta/2),\ \varepsilon v^5/(8C_{\mathrm{m}}^2 C_{\mathrm{a}}^4 \mathrm{Lip}_{\mathbb{H}}(f)^3)\}$,
5. $\sigma = 9\varepsilon^2 v^5 \Delta^2/256$,
6. $0 < \rho < 1$ such that
$$|f(x_* + tE_*(x_*)) - f(x_*) - tE_*f(x_*)| \leq \sigma \mathrm{Lip}_{\mathbb{H}}(f)|t| \qquad (5.2)$$
for every $|t| \leq \rho$,

(7) $0 < \delta < \rho\sqrt{3\varepsilon v\Delta^3}/4$ such that

$$Ef(x) < E_*f(x_*) + \varepsilon v\Delta/2$$

whenever $(x, E) \in M$ and $d(x, x_*) \leq 4\delta(1 + 1/\Delta)$.

To prove Pansu differentiability of f at x_* we will show:

$$|f(x_*\delta_t(h)) - f(x_*) - tE_*f(x_*)\langle h, E_*(0)\rangle| \leq \varepsilon t$$

whenever $d(h) \leq 1$ and $0 < t < \delta$. Suppose this is not true. Then there exists $u \in \mathbb{Q}^{2n+1}$ with $d(u) \leq 1$ and rational $0 < r < \delta$ such that:

$$|f(x_*\delta_r(u)) - f(x_*) - rE_*f(x_*)\langle u, E_*(0)\rangle| > \varepsilon r. \tag{5.3}$$

Let $s = r/\Delta \in \mathbb{Q}$. We next construct Lipschitz horizontal curves g and h for which we can apply Lemma 5.3 with $\varphi := f \circ g$ and $\psi := f \circ h$.

Construction of g. To ensure that the image of g is a subset of the set N, we first introduce rational approximations to x_* and E_*.

Since the Carnot-Carathéodory and Euclidean distances are topologically equivalent, $\mathbb{Q}^{2n+1}$ is dense in $\mathbb{R}^{2n+1}$ with respect to the distance d. The set

$$\{E \in V : \omega(E) = 1, E \text{ a rational linear combination of } X_i, Y_i, 1 \leq i \leq n\}$$

is dense in $\{E \in V : \omega(E) = 1\}$ with respect to the norm ω. To see this, suppose $E \in V$ satisfies $\omega(E) = 1$. The Euclidean sphere $\mathbb{S}^{2n-1} \subset \mathbb{R}^{2n}$ contains a dense set S of points with rational coordinates. This fact is well known, e.g. one can use stereographic projection. Let $q = (q_1, \ldots, q_{2n}) \in \mathbb{S}^{2n-1}$ be the coefficients of E in the basis $\{X_i, Y_i : 1 \leq i \leq n\}$. Take $\widetilde{q} \in S$ such that $|q - \widetilde{q}|$ is small and define the rational approximation of E as the linear combination of $\{X_i, Y_i : 1 \leq i \leq n\}$ with coefficients $\widetilde{q}_i$.

Define

$$A_1 = (\eta\Delta/C_{\mathrm{a}})^2 \tag{5.4}$$

and

$$A_2 = \left(6 - \left(4(1 + 20v)\left(\frac{2+v}{1-v}\right)^{\frac{1}{2}} + v\right)\right)^2 \frac{(\varepsilon v\Delta/2)\mathrm{Lip}_{\mathbb{H}}(f))^{\frac{1}{2}}}{C_{\mathrm{a}}^2\mathrm{Lip}_{\mathbb{H}}(f)^2}. \tag{5.5}$$

Notice $A_1, A_2 > 0$ using, in particular, our choice of v. Choose $\widetilde{x}_* \in \mathbb{Q}^{2n+1}$ and $\widetilde{E}_* \in V$ with $\omega(\widetilde{E}_*) = 1$, a rational linear combination of $\{X_i, Y_i : 1 \leq i \leq n\}$, sufficiently close to x_* and E_* to ensure:

$$d(\widetilde{x}_*\delta_r(u), x_*) \leq 2r, \tag{5.6}$$

$$d(\widetilde{x}_*\delta_r(u), x_*\delta_r(u)) \leq \sigma r, \tag{5.7}$$

$$\omega(\widetilde{E}_* - E_*) \leq \min\{\sigma, C_{\mathrm{m}}\Delta, A_1, A_2\}, \tag{5.8}$$

which is possible since all terms on the right side of the above inequalities are strictly positive. Additionally, we choose $\widetilde{x}_*$ and $\widetilde{E}_*$ sufficiently close to x_* and E_* so that if $a, b \in \mathbb{R}^n$ and $c \in \mathbb{R}$ are defined by

$$(a, b, c) = (x_* + (s/2)E_*(x_*))^{-1}(\widetilde{x}_* + s\widetilde{E}_*(\widetilde{x}_*)),$$

then:

$$(a^2 + b^2)^{\frac{1}{2}}\left(1 + \frac{c^2}{(a^2+b^2)^2} + \frac{4c^2}{(a^2+b^2)}\right)^{\frac{1}{2}} \leq \frac{s}{2}\left(1 + \frac{\eta\Delta}{2}\right) \tag{5.9}$$

and

$$\frac{c}{(a^2+b^2)^{\frac{1}{2}}}(1 + 4(a^2+b^2))^{\frac{1}{2}} \leq \frac{s}{2}\min\{A_1, A_2\}. \tag{5.10}$$

To show these requirements can be satisfied, we remark that

$$\begin{aligned}
(x_* + (s/2)E_*(x_*))^{-1}(x_* + sE_*(x_*)) &= (x_* \exp((s/2)E_*))^{-1}(x_* \exp(sE_*)) \\
&= \exp(-(s/2)E_*)x_*^{-1}x_* \exp(sE_*) \\
&= \exp((s/2)E_*) \\
&= (s/2)E_*(0),
\end{aligned}$$

which has final coordinate zero and satisfies $|p((s/2)E_*(0))| = s/2$. We also choose $\widetilde{x}_*$ and $\widetilde{E}_*$ so that (5.9) and (5.10) hold if instead:

$$(a, b, c) = (\widetilde{x}_* - s\widetilde{E}_*(\widetilde{x}_*))^{-1}(x_* - (s/2)E_*(x_*)),$$

with a similar justification.

Note $0 < r < \Delta$ and recall $s = r/\Delta$ is rational. To construct g we apply Lemma 4.2 with the following parameters:

- η, r, Δ and u as defined above in (5.3),
- $x = \widetilde{x}_*$ and $E = \widetilde{E}_*$.

This gives a Lipschitz horizontal curve $g \colon \mathbb{R} \to \mathbb{H}^n$ with the following properties:

(1) $g(t) = \widetilde{x}_* + t\widetilde{E}_*(x_*)$ for $|t| \geq s$,
(2) $g(\zeta) = \widetilde{x}_*\delta_r(u)$, where $\zeta := r\langle u, \widetilde{E}_*(0)\rangle$,
(3) $\mathrm{Lip}_{\mathbb{H}}(g) \leq 1 + \eta\Delta/2$,
(4) $g'(t)$ exists and $|(p \circ g)'(t) - p(\widetilde{E}_*)| \leq C_{\mathrm{m}}\Delta$ for $t \in \mathbb{R}$ outside a finite set.

Since the relevant quantities were rational and the set N was chosen using Lemma 4.1, we also know that the image of g is contained in N.

Construction of h. Denote $q_1 = x_* + (s/2)E_*(x_*)$ and $q_2 = \widetilde{x}_* + s\widetilde{E}_*(\widetilde{x}_*)$. Applying Lemma 3.4 with $(a, b, c) = q_1^{-1}q_2$ and using inequalities (5.9) and (5.10) gives $\widetilde{h}_1 \colon [0, 1] \to \mathbb{H}^n$ such that:

(1) $\widetilde{h}_1$ is a Lipschitz horizontal curve joining 0 to $q_1^{-1}q_2$,
(2) $\mathrm{Lip}_{\mathbb{H}}(\widetilde{h}_1) \leq (s/2)(1 + \eta\Delta/2)$,
(3) $\widetilde{h}_1'(t)$ exists and satisfies

$$|(p \circ \widetilde{h}_1)'(t) - (s/2)p(E_*)| \leq (s/2)\min\{A_1, A_2\}$$

for $t \in [0, 1] \setminus \{1/2\}$.

Define $h_1 \colon [s/2, s] \to \mathbb{H}^n$ by:

$$h_1(t) = q_1\widetilde{h}_1((2/s)(t - s/2)).$$

Then h_1 is a Lipschitz horizontal curve joining q_1 to q_2 with

$$\mathrm{Lip}_{\mathbb{H}}(h_1) \leq 1 + \eta\Delta/2.$$

The derivative $h_1'(t)$ exists and satisfies

$$|(p \circ h_1)'(t) - p(E_*)| \leq \min\{A_1, A_2\}$$

for $t \in [s/2, s] \setminus \{3s/4\}$, since left translations act linearly in the first $2n$ coordinates.

Similarly, there is a Lipschitz horizontal curve $h_2 \colon [-s, -s/2] \to \mathbb{H}^n$ joining $\widetilde{x}_* - s\widetilde{E}_*(\widetilde{x}_*)$ to $x_* - (s/2)E_*(x_*)$ satisfying

$$\mathrm{Lip}_{\mathbb{H}}(h_2) \leq 1 + \eta\Delta/2$$

and

$$|(p \circ h_2)'(t) - p(E_*)| \leq \min\{A_1, A_2\}$$

for $t \in [-s, -s/2] \setminus \{-3s/4\}$.

Define a Lipschitz horizontal curve $h \colon \mathbb{R} \to \mathbb{H}^n$ by:

$$h(t) = \begin{cases} \widetilde{x}_* + t\widetilde{E}_*(\widetilde{x}_*) & \text{if } |t| \geq s, \\ x_* + tE_*(x_*) & \text{if } |t| \leq s/2, \\ h_1(t) & \text{if } s/2 < t < s, \\ h_2(t) & \text{if } -s < t < -s/2. \end{cases}$$

Using the inequalities $\mathrm{Lip}_{\mathbb{H}}(h_1), \mathrm{Lip}_{\mathbb{H}}(h_2) \leq 1 + \eta\Delta/2$ and $\omega(E_*), \omega(\widetilde{E}_*) \leq 1$ gives:

$$\mathrm{Lip}_{\mathbb{H}}(h) \leq 1 + \eta\Delta/2.$$

Also $h'(t)$ exists for $t \in \mathbb{R} \setminus \{\pm 3s/4, \pm s/2, \pm s\}$ and

$$|(p \circ h)'(t) - p(E_*)| \leq \min\{A_1, A_2\},$$

using the corresponding bounds for h_1, h_2 and $\omega(E_* - \widetilde{E}_*) \leq \min\{A_1, A_2\}$ from (5.8).

Application of Lemma 5.3. We now prove that the assumptions of Lemma 5.3 hold with $L := (2 + \eta\Delta)\mathrm{Lip}_{\mathbb{H}}(f)$, $\varphi := f \circ g$ and $\psi := f \circ h$. The inequalities $|\zeta| < s < \rho$, $0 < v < 1/32$ and the equality $\varphi(t) = \psi(t)$ for $|t| \geq s$ are clear. Using $\mathrm{Lip}_{\mathbb{H}}(g), \mathrm{Lip}_{\mathbb{H}}(h) \leq 1 + \eta\Delta/2$ gives $\mathrm{Lip}_{\mathbb{E}}(\varphi) + \mathrm{Lip}_{\mathbb{E}}(\psi) \leq L$.

Notice that (5.7) implies:

$$|f(\widetilde{x}_* \delta_r(u)) - f(x_* \delta_r(u))| \leq \sigma r \mathrm{Lip}_{\mathbb{H}}(f).$$

Since $|\zeta| \leq r \leq \rho$, we may substitute $t = \zeta$ in (5.2) to obtain

$$|f(x_* + \zeta E_*(x_*)) - f(x_*) - \zeta E_* f(x_*)| \leq \sigma \mathrm{Lip}_{\mathbb{H}}(f)|\zeta|$$
$$\leq \sigma r \mathrm{Lip}_{\mathbb{H}}(f).$$

Next note that (5.8) implies $|\widetilde{E}_*(0) - E_*(0)| \leq \sigma$. We use also $\zeta = r\langle u, \widetilde{E}_*(0)\rangle$ to estimate as follows:

$$|\zeta E_* f(x_*) - r\langle u, E_*(0)\rangle E_* f(x_*)| = r|E_* f(x_*)||\langle u, \widetilde{E}_*(0) - E_*(0)\rangle|$$
$$\leq r \mathrm{Lip}_{\mathbb{H}}(f)|\widetilde{E}_*(0) - E_*(0)|$$
$$\leq \sigma r \mathrm{Lip}_{\mathbb{H}}(f).$$

Hence we obtain,

$$|f(x_* + \zeta E_*(x_*)) - f(x_*) - r\langle u, E_*(0)\rangle E_* f(x_*)| \leq 2\sigma r \mathrm{Lip}_{\mathbb{H}}(f). \tag{5.11}$$

Notice $|\zeta| \leq r = \Delta s \leq s/2$, so the definition of the h gives $h(\zeta) = x_* + \zeta E_*(x_*)$. The definition of g gives $g(\zeta) = \widetilde{x}_* \delta_r(u)$. Using also (5.3) and (5.11), we can estimate as follows:

$$\begin{aligned} |\varphi(\zeta) - \psi(\zeta)| &= |f(g(\zeta)) - f(h(\zeta))| \\ &= |f(\widetilde{x}_* \delta_r(u)) - f(x_* + \zeta E_*(x_*))| \\ &\geq |f(x_* \delta_r(u)) - f(x_* + \zeta E_*(x_*))| - |f(\widetilde{x}_* \delta_r(u)) - f(x_* \delta_r(u))| \\ &\geq |f(x_* \delta_r(u)) - f(x_*) - rE_* f(x_*)\langle u, E_*(0)\rangle| \\ &\quad - |f(x_* + \zeta E_*(x_*)) - f(x_*) - rE_* f(x_*)\langle u, E_*(0)\rangle| \\ &\quad - \sigma r \mathrm{Lip}_{\mathbb{H}}(f) \\ &\geq \varepsilon r - 2\sigma r \mathrm{Lip}_{\mathbb{H}}(f) - \sigma r \mathrm{Lip}_{\mathbb{H}}(f) \\ &= \varepsilon r - 3\sigma r \mathrm{Lip}_{\mathbb{H}}(f) \\ &\geq 3\varepsilon r/4. \end{aligned} \tag{5.12}$$

In particular, $\varphi(\zeta) \neq \psi(\zeta)$.

We next check that $\psi'(0)$ exists and

$$|\psi(t) - \psi(0) - t\psi'(0)| \leq \sigma L|t| \qquad (5.13)$$

whenever $|t| \leq \rho$. Notice $\psi'(0)$ exists and equals $E_* f(x_*)$, since $E_* f(x_*)$ exists and $\psi(t) = f(x_* + tE_*(x_*))$ for $|t| \leq s/2$. Recall $|(p \circ h)' - p(E_*)| \leq A_1$ and also the definition $A_1 = (\eta\Delta/C_{\mathrm{a}})^2$ from (5.4). Since $h(0) = x_*$, Lemma 5.1 implies that

$$d(x_* + tE_*(x_*), h(t)) \leq C_{\mathrm{a}}\sqrt{A_1}|t| \leq \eta\Delta|t|.$$

Hence, using also (5.2) and $L = (2 + \eta\Delta)\mathrm{Lip}_{\mathbb{H}}(f)$,

$$\begin{aligned}
|\psi(t) - \psi(0) - t\psi'(0)| &\leq |f(x_* + tE_*(x_*)) - f(x_*) - tE_* f(x_*)| \\
&\quad + |f(x_* + tE_*(x_*)) - f(h(t))| \\
&\leq \sigma\mathrm{Lip}_{\mathbb{H}}(f)|t| + \mathrm{Lip}_{\mathbb{H}}(f)\eta\Delta|t| \\
&\leq \sigma L|t|.
\end{aligned}$$

Recall that $\mathrm{Lip}_{\mathbb{H}}(f) \leq 1/2$, which implies $L \leq 4$. Using also $r < \delta$, $s = r/\Delta$, (5.12) and the definition of δ in Parameters (7) we deduce:

$$\begin{aligned}
s\sqrt{sL/(v|\varphi(\zeta) - \psi(\zeta)|)} &\leq 4s\sqrt{s/(3\varepsilon rv)} \\
&= 4r/\sqrt{3\varepsilon v\Delta^3} \\
&\leq 4\delta/\sqrt{3\varepsilon v\Delta^3} \\
&\leq \rho.
\end{aligned}$$

Finally we use (5.12), $L \leq 4$ and the definition of σ in Parameters (5) to observe:

$$\begin{aligned}
v^3(|\varphi(\zeta) - \psi(\zeta)|/(sL))^2 &\geq v^3(3\varepsilon r/16s)^2 \\
&= 9\varepsilon^2 v^3\Delta^2/256 \\
&\geq \sigma.
\end{aligned}$$

We may now apply Lemma 5.3. We obtain $\tau \in (-s, s) \setminus \{\zeta\}$ such that $\varphi'(\tau)$ exists,

$$\varphi'(\tau) \geq \psi'(0) + v|\varphi(\zeta) - \psi(\zeta)|/s, \qquad (5.14)$$

and

$$|(\varphi(\tau + t) - \varphi(\tau)) - (\psi(t) - \psi(0))| \leq 4(1 + 20v)\sqrt{(\varphi'(\tau) - \psi'(0))L}|t| \qquad (5.15)$$

for every $t \in \mathbb{R}$. Since g is a horizontal curve, we may use Remark 5.4 to additionally choose τ such that $g'(\tau)$ exists and is in $\mathrm{Span}\{X_i(g(\tau)), Y_i(g(\tau)) \colon 1 \leq i \leq n\}$.

Conclusion. Let $x = g(\tau) \in N$ and choose $E \in V$ with $E(g(\tau)) = g'(\tau)/|p(g'(\tau))|$, which implies that $\omega(E) = 1$. We will transform (5.14) and (5.15) into

$$Ef(x) \geq E_* f(x_*) + \varepsilon v\Delta/2 \qquad (5.16)$$

and

$$(x, E) \in M. \qquad (5.17)$$

We first observe that this suffices to conclude the proof. Recall $d(\widetilde{x}_*\delta_r(u), x_*) \leq 2r$ from (5.6). Using $g(\tau) = x$ and $g(\zeta) = \widetilde{x}_*\delta_r(u)$ gives:

$$\begin{aligned}
d(x, x_*) &\leq d(g(\tau), g(\zeta)) + d(\widetilde{x}_*\delta_r(u), x_*) \\
&\leq \mathrm{Lip}_{\mathbb{H}}(g)|\tau - \zeta| + 2r \\
&\leq 4(s + r) \\
&= 4r(1 + 1/\Delta) \\
&\leq 4\delta(1 + 1/\Delta).
\end{aligned}$$

Since $x \in N$, combining this with (5.16) and (5.17) contradicts the choice of δ in Parameters (7). This contradiction forces us to conclude that (5.3) is false, finishing the proof.

Proof of (5.16). Using (5.12) and (5.14) we see:

$$\varphi'(\tau) - \psi'(0) \geq 3\varepsilon v r / 4s = 3\varepsilon v \Delta / 4. \tag{5.18}$$

Notice that $\varphi'(\tau) = Ef(x)|p(g'(\tau))|$ using Definition 3.2 and our choice of E. Since $\omega(E) = 1$ implies $|Ef(x)| \leq \mathrm{Lip}_{\mathbb{H}}(f)$, we deduce $|\varphi'(\tau)|/|p(g'(\tau))| \leq \mathrm{Lip}_{\mathbb{H}}(f)$. Also $|p(g'(\tau))| \leq \mathrm{Lip}_{\mathbb{H}}(g) \leq 1 + \eta\Delta$. Using $\psi'(0) = E_* f(x_*)$ and (5.18) gives:

$$Ef(x) - E_* f(x_*) - (1 - v)(\varphi'(\tau) - \psi'(0))$$
$$= v(\varphi'(\tau) - \psi'(0)) + (1 - |p(g'(\tau))|)\varphi'(\tau)/|p(g'(\tau))|$$
$$\geq 3\varepsilon v^2 \Delta / 4 - \eta \Delta \mathrm{Lip}_{\mathbb{H}}(f)$$
$$\geq 0,$$

where in the last inequality we used $\mathrm{Lip}_{\mathbb{H}}(f) \leq 1/2$ and $\eta \leq 3\varepsilon v^2/2$ from Parameters (2). From this we use $0 < v < 1/32$ and (5.18) again to deduce:

$$Ef(x) - E_* f(x_*) \geq (1 - v)(\varphi'(\tau) - \psi'(0)) \geq \varepsilon v \Delta / 2 \tag{5.19}$$

which proves (5.16).

Proof of (5.17). Recall that $|(p \circ g)'(t) - p(\widetilde{E}_*)| \leq C_{\mathrm{m}}\Delta$ for all but finitely many t, from our construction of g. Using (5.8), this implies $|(p \circ g)'(t) - p(E_*)| \leq 2C_{\mathrm{m}}\Delta$ for all but finitely many t. Since $x = g(\tau)$, we can apply Lemma 5.1 to obtain

$$d(g(\tau + t), x + tE_*(x)) \leq C_{\mathrm{a}}\sqrt{2C_{\mathrm{m}}\Delta}|t|$$

for $t \in \mathbb{R}$. By (5.19) we have $\Delta \leq 2(Ef(x) - E_* f(x_*))/(\varepsilon v)$. Using also the definition of Δ from Parameters (4), we deduce that for $t \in \mathbb{R}$:

$$|(f(x + tE_*(x)) - f(x)) - (f(g(\tau + t)) - f(g(\tau)))|$$
$$= |f(x + tE_*(x)) - f(g(\tau + t))|$$
$$\leq \mathrm{Lip}_{\mathbb{H}}(f)d(g(\tau + t), x + tE_*(x))$$
$$\leq C_{\mathrm{a}}\sqrt{2C_{\mathrm{m}}\Delta}\mathrm{Lip}_{\mathbb{H}}(f)|t|$$
$$\leq C_{\mathrm{a}}\sqrt{2C_{\mathrm{m}}}\mathrm{Lip}_{\mathbb{H}}(f)|t|\Delta^{\frac{1}{4}}\Big(\frac{2(Ef(x) - E_* f(x_*))}{\varepsilon v}\Big)^{\frac{1}{4}}$$
$$\leq v|t|\big((Ef(x) - E_* f(x_*))\mathrm{Lip}_{\mathbb{H}}(f)\big)^{\frac{1}{4}}\Big(\frac{8C_{\mathrm{m}}^2 C_{\mathrm{a}}^4 \Delta \mathrm{Lip}_{\mathbb{H}}(f)^3}{\varepsilon v^5}\Big)^{\frac{1}{4}}$$
$$\leq v|t|\big((Ef(x) - E_* f(x_*))\mathrm{Lip}_{\mathbb{H}}(f)\big)^{\frac{1}{4}}. \tag{5.20}$$

Combining (5.15), (5.19) and $L = (2 + \eta\Delta)\mathrm{Lip}_{\mathbb{H}}(f) \leq (2 + v)\mathrm{Lip}_{\mathbb{H}}(f)$ gives:

$$|(\varphi(\tau + t) - \varphi(\tau)) - (\psi(t) - \psi(0))|$$
$$\leq 4(1 + 20v)|t|\Big(\frac{(2 + v)\mathrm{Lip}_{\mathbb{H}}(f)(Ef(x) - E_* f(x_*))}{1 - v}\Big)^{\frac{1}{2}} \tag{5.21}$$

for $t \in \mathbb{R}$. Using $\mathrm{Lip}_{\mathbb{H}}(f) \leq 1/2$ gives the simple bound:

$$((Ef(x) - E_* f(x_*))\mathrm{Lip}_{\mathbb{H}}(f))^{\frac{1}{2}} \leq ((Ef(x) - E_* f(x_*))\mathrm{Lip}_{\mathbb{H}}(f))^{\frac{1}{4}}$$

since both sides are less than 1. Hence adding (5.20) and (5.21) and using the definition $\varphi = f \circ g$ gives:

$$|f(x + tE_*(x) - f(x)) - (\psi(t) - \psi(0))|$$
$$\leq \Big(4(1 + 20v)\Big(\frac{2 + v}{1 - v}\Big)^{\frac{1}{2}} + v\Big)|t|((Ef(x) - E_* f(x_*))\mathrm{Lip}_{\mathbb{H}}(f))^{\frac{1}{4}} \tag{5.22}$$

for $t \in \mathbb{R}$.

Recall $\psi = f \circ h$ and $h(0) = x_*$. Using the inequality $|(p \circ h)' - p(E_*)| \leq A_2$, Lemma 5.1, our definition of A_2 in (5.5) and (5.16), we can estimate as follows:

$$
\begin{aligned}
|(\psi(t) - \psi(0)) &- (f(x_* + tE_*(x_*)) - f(x_*))| \\
&= |f(h(t)) - f(x_* + tE_*(x_*))| \\
&\leq \mathrm{Lip}_{\mathbb{H}}(f) d(h(t), x_* + tE_*(x_*)) \\
&\leq \mathrm{Lip}_{\mathbb{H}}(f) C_{\mathrm{a}} \sqrt{A_2} |t| \\
&= \Big(6 - \Big(4(1 + 20v) \Big(\frac{2+v}{1-v} \Big)^{\frac{1}{2}} + v \Big) \Big) |t| (\varepsilon v \Delta / 2 \, \mathrm{Lip}_{\mathbb{H}}(f))^{\frac{1}{4}} \qquad (5.23) \\
&\leq \Big(6 - \Big(4(1 + 20v) \Big(\frac{2+v}{1-v} \Big)^{\frac{1}{2}} + v \Big) \Big) |t| ((Ef(x) - E_* f(x_*)) \mathrm{Lip}_{\mathbb{H}}(f))^{\frac{1}{4}} \quad (5.24)
\end{aligned}
$$

for all $t \in \mathbb{R}$. Adding (5.22) and (5.23) gives:

$$
\begin{aligned}
|(f(x + tE_*(x)) - f(x)) &- (f(x_* + tE_*(x_*)) - f(x_*))| \\
&\leq 6|t| \big((Ef(x) - E_* f(x_*)) \mathrm{Lip}_{\mathbb{H}}(f) \big)^{\frac{1}{4}}
\end{aligned}
$$

for $t \in \mathbb{R}$. This implies (5.17), hence proving the theorem. $\qquad \square$

6. Construction of an almost maximal directional derivative

The main result of this section is Proposition 6.1, which is an adaptation of [10, Theorem 3.1] to $\mathbb{H}^n$. It shows that given a Lipschitz function $f_0 \colon \mathbb{H}^n \to \mathbb{R}$, there is a Lipschitz function $f \colon \mathbb{H}^n \to \mathbb{R}$ such that $f - f_0$ is $\mathbb{H}$-linear and f has an almost locally maximal horizontal directional derivative in the sense of Theorem 5.6. We will conclude that any Lipschitz function f_0 is Pansu differentiable at a point of N, proving Theorem 2.12. Our argument follows very closely that of [10], modified to use horizontal directions, $\mathbb{H}$-linear maps and Hölder equivalence of the Carnot-Carathéodory and Euclidean distance.

Recall the measure zero G_δ set N and the notation D^f fixed in Notation 5.5. In particular, the statement $(x, E) \in D^f$ implies that $x \in N$. Note that if $f - f_0$ is $\mathbb{H}$-linear then $D^f = D^{f_0}$ and also the functions f and f_0 have the same points of Pansu differentiability.

Proposition 6.1. *Suppose $f_0 : \mathbb{H}^n \to \mathbb{R}$ is a Lipschitz function, $(x_0, E_0) \in D^{f_0}$ and $\delta_0, \mu, K > 0$. Then there is a Lipschitz function $f : \mathbb{H}^n \to \mathbb{R}$ such that $f - f_0$ is $\mathbb{H}$-linear with $\mathrm{Lip}_{\mathbb{H}}(f - f_0) \leq \mu$, and a pair $(x_*, E_*) \in D^f$ with $d(x_*, x_0) \leq \delta_0$ such that $E_* f(x_*) > 0$ is almost locally maximal in the following sense.*

For any $\varepsilon > 0$ there is $\delta_\varepsilon > 0$ such that whenever $(x, E) \in D^f$ satisfies both:

(1) *$d(x, x_*) \leq \delta_\varepsilon$, $Ef(x) \geq E_* f(x_*)$,*
(2) *for any $t \in (-1, 1)$:*

$$
\begin{aligned}
|(f(x + tE_*(x)) - f(x)) &- (f(x_* + tE_*(x_*)) - f(x_*))| \\
&\leq K|t|(Ef(x) - E_* f(x_*))^{\frac{1}{4}},
\end{aligned}
$$

then:

$$
Ef(x) < E_* f(x_*) + \varepsilon.
$$

We use the remainder of this section to prove Proposition 6.1. Fix parameters $f_0, x_0, E_0, \delta_0, \mu, K$ as given in the statement of the theorem.

Assumptions 6.2. Without loss of generality, we make the following assumptions:
- $K \geq 8$, since increasing K makes the statement of Proposition 6.1 stronger,
- $\mathrm{Lip}_{\mathbb{H}}(f_0) \leq 1/2$, after multiplying f_0 by a positive constant and possibly increasing K,

- $E_0 f(x_0) \geq 0$, by replacing E_0 by $-E_0$ if necessary.

We prove Proposition 6.1 by using Algorithm 6.4 below to construct pairs (x_n, E_n) and Lipschitz functions f_n, satisfying various constraints, such that $E_n f(x_n)$ is closer and closer to maximal. We then show that the limits (x_*, E_*) and f have the properties stated in Proposition 6.1. Algorithm 6.4 is an adaptation of [10, Algorithm 3.2]. We use the following notation to repeatedly find better pairs.

Notation 6.3. Suppose $h : \mathbb{H}^n \to \mathbb{R}$ is Lipschitz, the pairs (x, E) and (x', E') belong to D^h, and $\sigma \geq 0$. We write:

$$(x, E) \leq_{(h,\sigma)} (x', E')$$

if $Eh(x) \leq E'h(x')$ and for all $t \in (-1, 1)$:

$$|(h(x' + tE(x')) - h(x')) - (h(x + tE(x)) - h(x))|$$
$$\leq K(\sigma + (E'f(x') - Ef(x))^{\frac{1}{4}}|t|.$$

In the language of Notation 6.3, Proposition 6.1(2) means $(x_*, E_*) \leq_{(f,0)} (x, E)$.

In Algorithm 6.4 we introduce parameters satisfying various estimates, but the most important factor is the order in which the parameters are chosen. We use the following constants:

- $C_a = C_{\text{angle}}(1) \geq 1$ chosen by applying Lemma 5.1 with $S = 1$,
- $C_H = C_{\text{Hölder}} \geq 1$ denotes the constant in Proposition 2.7 for the compact set $\overline{B_{\mathbb{H}}(x_0, 2 + \delta_0)} \subset \mathbb{H}^n$,
- $C_V \geq 1$ such that $\text{Lip}_{\mathbb{E}}(E) \leq C_V$ whenever $E \in V$ and $\omega(E) = 1$. This is possible since $V = \text{Span}\{X_i, Y_i : 1 \leq i \leq n\}$ and $\{X_i, Y_i : 1 \leq i \leq n\}$ are Lipschitz functions $\mathbb{R}^{2n+1} \to \mathbb{R}^{2n+1}$ with respect to the Euclidean distance.

Since N is G_δ we can fix open sets $U_k \subset \mathbb{H}^n$ such that $N = \cap_{k=0}^{\infty} U_k$. We may assume that $U_0 = \mathbb{H}^n$.

Algorithm 6.4. *Recall* f_0, x_0, E_0 *and* δ_0 *from the hypotheses of Proposition 6.1. Let* $\sigma_0 := 2$ *and* $t_0 := \min\{1/4, \mu/2\}$.

Suppose that $m \geq 1$ *and the parameters* $f_{m-1}, x_{m-1}, E_{m-1}, \sigma_{m-1}, t_{m-1}, \delta_{m-1}$ *have already been defined. Then we can choose:*

(1) $f_m(x) := f_{m-1}(x) + t_{m-1}\langle x, E_{m-1}(0)\rangle$,
(2) $\sigma_m \in (0, \sigma_{m-1}/4)$,
(3) $t_m \in (0, \min\{t_{m-1}/2, \sigma_{m-1}/(4m)\})$,
(4) $\lambda_m \in (0, t_m\sigma_m^4/(2C_a^4))$,
(5) D_m *to be the set of pairs* $(x, E) \in D^{f_m} = D^{f_0}$ *such that* $d(x, x_{m-1}) < \delta_{m-1}$ *and*

$$(x_{m-1}, E_{m-1}) \leq_{(f_m, \sigma_{m-1}-\varepsilon)} (x, E)$$

for some $\varepsilon \in (0, \sigma_{m-1})$,
(6) $(x_m, E_m) \in D_m$ *such that* $Ef_m(x) \leq E_m f_m(x_m) + \lambda_m$ *for every pair* $(x, E) \in D_m$,
(7) $\varepsilon_m \in (0, \sigma_{m-1})$ *such that* $(x_{m-1}, E_{m-1}) \leq_{(f_m, \sigma_{m-1}-\varepsilon_m)} (x_m, E_m)$,
(8) $\delta_m \in (0, (\delta_{m-1} - d(x_m, x_{m-1}))/2)$ *such that* $\overline{B_{\mathbb{H}}(x_m, \delta_m)} \subset U_m$ *and for all* $|t| < C_H^2(1 + C_V)^{\frac{1}{2}}\delta_m^{\frac{1}{2}}/\varepsilon_m$:

$$|(f_m(x_m + tE_m(x_m)) - f_m(x_m)) - (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}))|$$
$$\leq (E_m f_m(x_m) - E_{m-1} f_m(x_{m-1}) + \sigma_{m-1})|t|.$$

Proof. Clearly one can make choices satisfying (1)–(5).

For (6) first notice that $(x_{m-1}, E_{m-1}) \in D_m$ and hence $D_m \neq \emptyset$. By Lemma 5.2, the functions $f_m \colon \mathbb{H}^n \to \mathbb{R}$ are Lipschitz and

$$f_m(x) = f_0(x) + \Big\langle x, \sum_{k=0}^{m-1} t_k E_k(0) \Big\rangle. \tag{6.1}$$

Using $\mathrm{Lip}_{\mathbb{H}}(f_0) \le 1/2$, $t_{k+1} \le t_k/2$, $t_0 \le 1/4$ and Lemma 5.2 gives $\mathrm{Lip}_{\mathbb{H}}(f_m) \le 1$. Lemma 3.3 implies $|Ef_m(x)| \le \omega(E)\mathrm{Lip}_{\mathbb{H}}(f_m)$. Hence $\sup_{(x,E) \in D_m} Ef_m(x) \le 1$, so we can choose $(x_m, E_m) \in D_m$ as in (6).

The definition of D_m in (5) implies that one can choose ε_m as in (7).

Notice that (6) and the definition of D_m in (5) imply that $x_m \in N \subset U_m$, $d(x_m, x_{m-1}) < \delta_{m-1}$ and $E_m f_m(x_m) \ge E_{m-1} f_m(x_{m-1})$. Therefore we can use existence of the directional derivatives of f_m to choose δ_m as in (8). $\qquad\square$

We record for later use that $\mathrm{Lip}_{\mathbb{H}}(f_m) \le 1$ for all $m \ge 1$. We next show that several parameters in Algorithm 6.4 converge to 0 and the balls $B_{\mathbb{H}}(x_m, \delta_m)$ form a decreasing sequence.

Lemma 6.5. *The sequences $\sigma_m, t_m, \lambda_m, \delta_m, \varepsilon_m$ converge to 0. For every $m \ge 1$ the following inclusion holds:*

$$\overline{B_{\mathbb{H}}(x_m, \delta_m)} \subset B_{\mathbb{H}}(x_{m-1}, \delta_{m-1}).$$

Proof. Algorithm 6.4(2) and $\sigma_0 = 2$ gives $0 < \sigma_m \le 2/4^m$ so $\sigma_m \to 0$. Combining this with Algorithm 6.4(3,4,7,8) shows the other sequences converge to 0. Let $x \in \overline{B_{\mathbb{H}}(x_m, \delta_m)}$. Then Algorithm 6.4(6,8) gives:

$$
\begin{aligned}
d(x, x_{m-1}) &\le \delta_m + d(x_m, x_{m-1}) \\
&< \delta_{m-1}/2 + d(x_m, x_{m-1})/2 \\
&< \delta_{m-1}.
\end{aligned}
$$

This shows the desired inclusion. $\qquad\square$

Define $\varepsilon'_m > 0$ by:

$$\varepsilon'_m := \min\{\varepsilon_m/2, \, \sigma_{m-1}/2\}. \tag{6.2}$$

We next show that the sets D_m of special pairs form a decreasing sequence. This is an adaptation of [10, Lemma 3.3].

Lemma 6.6. *The following statements hold:*

(1) *If $m \ge 1$ and $(x, E) \in D_{m+1}$ then:*

$$(x_{m-1}, E_{m-1}) \le_{(f_m, \sigma_m - \varepsilon'_m)} (x, E),$$

(2) *If $m \ge 1$ then $D_{m+1} \subset D_m$,*
(3) *If $m \ge 0$ and $(x, E) \in D_{m+1}$ then $d(E(0), E_m(0)) \le \sigma_m$.*

Proof. If $m = 0$ then (3) holds since:

$$d(E(0), E_0(0)) \le d(E(0), 0) + d(0, E_0(0)) \le 2 = \sigma_0.$$

It is enough to check that whenever $m \ge 1$ and (3) holds for $m - 1$, then (1), (2) and (3) hold for m. Fix $m \ge 1$ and assume that (3) holds for $m - 1$:

$$d(E(0), E_{m-1}(0)) \le \sigma_{m-1} \quad \text{for all} \quad (x, E) \in D_m.$$

Proof of (1). Algorithm 6.4(6) states that $(x_m, E_m) \in D_m$. Hence:

$$d(E_m(0), E_{m-1}(0)) \le \sigma_{m-1}. \tag{6.3}$$

Let $(x, E) \in D_{m+1}$. In particular, we have $Ef_{m+1}(x) \ge E_m f_{m+1}(x_m)$ by Algorithm 6.4(5). Notice that, since $\omega(E_m) = \omega(E) = 1$, we have $\langle E_m(0), E_m(0) \rangle = 1$ and

$\langle E(0), E_m(0)\rangle \le 1$. Let $A := E f_m(x) - E_m f_m(x_m)$. Using Lemma 5.2 and the inequality $E f_{m+1}(x) \ge E_m f_{m+1}(x_m)$ gives:

$$A = E f_{m+1}(x) - E_m f_{m+1}(x_m) - t_m \langle E(0), E_m(0)\rangle + t_m \ge 0.$$

Using Algorithm 6.4(5) again, together with the above inequality, gives:

$$E f_m(x) \ge E_m f_m(x_m) \ge E_{m-1} f_m(x_{m-1}).$$

In particular, $E f_m(x) \ge E_{m-1} f_m(x_{m-1})$ proves the first part of the statement $(x_{m-1}, E_{m-1}) \le_{(f_m, \sigma_m - \varepsilon'_m)} (x, E)$.

Let $B := E f_m(x) - E_{m-1} f_m(x_{m-1}) \ge 0$. Lemma 3.3 and $\mathrm{Lip}_{\mathbb{H}}(f_m) \le 1$ imply that $0 \le A, B \le 2$. Using these inequalities and $K \ge 8$ gives:

$$\begin{aligned}
K(B^{\frac{1}{4}} - A^{\frac{1}{4}}) &\ge (B^{\frac{3}{4}} + B^{\frac{1}{2}} A^{\frac{1}{4}} + B^{\frac{1}{4}} A^{\frac{1}{2}} + A^{\frac{3}{4}})(B^{\frac{1}{4}} - A^{\frac{1}{4}}) \\
&= B - A \\
&= E_m f_m(x_m) - E_{m-1} f_m(x_{m-1}).
\end{aligned} \tag{6.4}$$

Since $A \ge E f_{m+1}(x) - E_m f_{m+1}(x_m)$, (6.4) implies:

$$\begin{aligned}
E_m f_m(x_m) - E_{m-1} f_m(x_{m-1}) &+ K(E f_{m+1}(x) - E_m f_{m+1}(x_m))^{\frac{1}{4}} \\
&\le K B^{\frac{1}{4}}.
\end{aligned} \tag{6.5}$$

To prove the second part of $(x_{m-1}, E_{m-1}) \le_{(f_m, \sigma_m - \varepsilon'_m)} (x, E)$ we need to estimate:

$$|(f_m(x + t E_{m-1}(x)) - f_m(x)) - (f_m(x_{m-1} + t E_{m-1}(x_{m-1})) - f_m(x_{m-1}))|. \tag{6.6}$$

We consider two cases, depending on whether t is small or large.

Suppose $|t| < C_{\mathrm{H}}^2 (1 + C_V)^{\frac{1}{2}} \delta_m^{\frac{1}{2}} / \varepsilon_m$. Estimate (6.6) as follows:

$$\begin{aligned}
|(f_m(x + t E_{m-1}(x)) &- f_m(x)) - (f_m(x_{m-1} + t E_{m-1}(x_{m-1})) - f_m(x_{m-1}))| \\
\le\ &|(f_m(x + t E_m(x)) - f_m(x)) - (f_m(x_m + t E_m(x_m)) - f_m(x_m))| \\
&+ |(f_m(x_m + t E_m(x_m)) - f_m(x_m)) \\
&\quad - (f_m(x_{m-1} + t E_{m-1}(x_{m-1})) - f_m(x_{m-1}))| \\
&+ |f_m(x + t E_{m-1}(x)) - f_m(x + t E_m(x))|.
\end{aligned} \tag{6.7}$$

We consider the three terms on the right side of (6.7) separately.

Firstly, Algorithm 6.4(1) and Lemma 5.2 give:

$$\begin{aligned}
(f_m(x + t E_m(x)) - f_m(x)) &- (f_m(x_m + t E_m(x_m)) - f_m(x_m)) \tag{6.8} \\
= (f_{m+1}(x + t E_m(x)) &- f_{m+1}(x)) - (f_{m+1}(x_m + t E_m(x_m)) - f_{m+1}(x_m)) \\
&- t_m \langle x + t E_m(x), E_m(0)\rangle + t_m \langle x, E_m(0)\rangle \\
&+ t_m \langle x_m + t E_m(x_m), E_m(0)\rangle - t_m \langle x_m, E_m(0)\rangle \\
= (f_{m+1}(x + t E_m(x)) &- f_{m+1}(x)) - (f_{m+1}(x_m + t E_m(x_m)) - f_{m+1}(x_m)).
\end{aligned}$$

Using (6.8) and $(x, E) \in D_{m+1}$ then gives:

$$\begin{aligned}
|(f_m(x + t E_m(x)) - f_m(x)) &- (f_m(x_m + t E_m(x_m)) - f_m(x_m))| \\
&\le K(\sigma_m + (E f_{m+1}(x) - E_m f_{m+1}(x_m))^{\frac{1}{4}})|t|.
\end{aligned} \tag{6.9}$$

For the second term in (6.7) we recall that, for the values of t we are considering, Algorithm 6.4(8) states:

$$\begin{aligned}
|(f_m(x_m + t E_m(x_m)) - f_m(x_m)) &- (f_m(x_{m-1} + t E_{m-1}(x_{m-1})) - f_m(x_{m-1}))| \\
&\le (E_m f_m(x_m) - E_{m-1} f_m(x_{m-1}) + \sigma_{m-1})|t|.
\end{aligned}$$

The final term in (6.7) is estimated using $\mathrm{Lip}_{\mathbb{H}}(f_m) \leq 1$, (2.1) and (6.3):

$$\begin{aligned}|f_m(x + tE_{m-1}(x)) - f_m(x + tE_m(x))| &\leq d(x + tE_{m-1}(x), x + tE_m(x)) \\ &= d(tE_{m-1}(0), tE_m(0)) \\ &\leq \sigma_{m-1}|t|.\end{aligned}$$

Adding the three estimates and using (6.5) then (6.2) and Algorithm 6.4(2) allows us to develop (6.7):

$$\begin{aligned}|(f_m(x &+ tE_{m-1}(x)) - f_m(x)) - (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}))| \\ &\leq K(\sigma_m + (Ef_{m+1}(x) - E_m f_{m+1}(x_m))^{\frac{1}{4}})|t| \\ &\quad + (E_m f_m(x_m) - E_{m-1}f_m(x_{m-1}) + \sigma_{m-1})|t| \\ &\quad + \sigma_{m-1}|t| \\ &\leq K(\sigma_{m-1} - \varepsilon'_m + (Ef_m(x) - E_{m-1}f_m(x_{m-1}))^{\frac{1}{4}})|t|.\end{aligned}$$

This gives the required estimate of (6.6) for small t.

Suppose $C_{\mathrm{H}}^2(1 + C_V)^{\frac{1}{2}}\delta_m^{\frac{1}{2}}/\varepsilon_m \leq |t| < 1$. In particular, this implies:

$$\delta_m \leq \varepsilon_m^2 t^2/C_H^4(1 + C_V) \leq \varepsilon_m|t|. \tag{6.10}$$

The last inequality above used that

$$\varepsilon_m|t|/C_H^4(1 + C_V) \leq \varepsilon_m/C_H^4(1 + C_V) \leq 1,$$

using $\varepsilon_m \leq 2$ and $C_H, C_V \geq 1$.

Estimate (6.6) as follows:

$$\begin{aligned}|(f_m(x &+ tE_{m-1}(x)) - f_m(x)) - (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}))| \\ &\leq |(f_m(x_m + tE_{m-1}(x_m)) - f_m(x_m)) \\ &\quad - (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}))| \\ &\quad + |f_m(x) - f_m(x_m)| \\ &\quad + |f_m(x + tE_{m-1}(x)) - f_m(x_m + tE_{m-1}(x_m))|.\end{aligned}$$

The estimate of the first term is given by Algorithm 6.4(7). This states:

$$(x_{m-1}, E_{m-1}) \leq_{(f_m, \sigma_{m-1} - \varepsilon_m)} (x_m, E_m),$$

which gives the inequality:

$$\begin{aligned}|(f_m(x_m &+ tE_{m-1}(x_m)) - f_m(x_m)) - (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}))| \\ &\leq K(\sigma_{m-1} - \varepsilon_m + (E_m f_m(x_m) - E_{m-1}f_m(x_{m-1}))^{\frac{1}{4}})|t|.\end{aligned}$$

The estimate of the second term uses $\mathrm{Lip}_{\mathbb{H}}(f_m) \leq 1$ and (6.10):

$$|f_m(x) - f_m(x_m)| \leq d(x, x_m) \leq \delta_m \leq \varepsilon_m|t| \leq K\varepsilon_m|t|/8.$$

We estimate the final term using Proposition 2.7 to compare the Carnot-Carathéodory and the Euclidean distance. Notice that $x + tE_{m-1}(x)$ and $x_m + tE_{m-1}(x_m)$ belong to $\overline{B_{\mathbb{H}}(x_0, 2 + \delta_0)}$. Hence we can use the constants C_{H} and C_V defined immediately

before Algorithm 6.4. For our current values of t we can estimate as follows:

$$|f_m(x + tE_{m-1}(x)) - f_m(x_m + tE_{m-1}(x_m))|$$
$$\leq d(x + tE_{m-1}(x), x_m + tE_{m-1}(x_m))$$
$$\leq C_{\mathrm{H}}|x + tE_{m-1}(x) - x_m - tE_{m-1}(x_m)|^{\frac{1}{2}}$$
$$\leq C_{\mathrm{H}}(1 + C_V)^{\frac{1}{2}}|x - x_m|^{\frac{1}{2}}$$
$$\leq C_{\mathrm{H}}^2(1 + C_V)^{\frac{1}{2}}d(x, x_m)^{\frac{1}{2}}$$
$$\leq C_{\mathrm{H}}^2(1 + C_V)^{\frac{1}{2}}\delta_m^{\frac{1}{2}}$$
$$\leq \varepsilon_m|t|$$
$$\leq K\varepsilon_m|t|/8.$$

Combine the estimates of the three terms and use $A \geq 0$ to obtain:

$$|(f_m(x + tE_{m-1}(x)) - f_m(x)) - (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}))|$$
$$\leq K(\sigma_{m-1} - \varepsilon_m/2 + (E_m f_m(x_m) - E_{m-1}f_m(x_{m-1}))^{\frac{1}{4}})$$
$$\leq K(\sigma_{m-1} - \varepsilon_m' + (Ef_m(x) - E_{m-1}f_m(x_{m-1}))^{\frac{1}{4}}).$$

This gives the correct estimate of (6.6) for large t. Combining the two cases proves (1) for m.

Proof of (2). Suppose $(x, E) \in D_{m+1}$. Then $(x, E) \in D^{f_{m+1}} = D^{f_m}$ and Lemma 6.5 implies $d(x, x_{m-1}) < \delta_{m-1}$. Combining this with (1) gives $(x, E) \in D_m$. This proves (2) for m.

Proof of (3). Suppose $(x, E) \in D_{m+1}$. Then $E_m f_{m+1}(x_m) \leq Ef_{m+1}(x)$ using Algorithm 6.4(5). Equivalently, by Algorithm 6.4(1):

$$E_m f_m(x_m) + t_m\langle E_m(0), E_m(0)\rangle \leq Ef_m(x) + t_m\langle E(0), E_m(0)\rangle.$$

Also $(x, E) \in D_m$ by (2) above, so Algorithm 6.4(6) implies:

$$Ef_m(x) \leq E_m f_m(x_m) + \lambda_m.$$

Combining the two inequalities above gives $t_m \leq t_m\langle E(0), E_m(0)\rangle + \lambda_m$. Rearranging, this implies:

$$\langle E(0), E_m(0)\rangle \geq 1 - \lambda_m/t_m.$$

Lemma 5.1 applied to $g(t) := tE(0)$ and $h(t) := tE_m(0)$ gives:

$$d(E(0), E_m(0)) \leq C_{\mathrm{a}}|E(0) - E_m(0)|^{\frac{1}{2}}$$
$$= C_{\mathrm{a}}(2 - 2\langle E(0), E_m(0)\rangle)^{\frac{1}{4}}$$
$$\leq C_{\mathrm{a}}(2\lambda_m/t_m)^{\frac{1}{4}}$$
$$\leq \sigma_m$$

by Algorithm 6.4(4). This proves (3) for m. □

We next study the convergence of (x_m, E_m) and f_m. We show that the directional derivatives converge to a directional derivative of the limiting function, and the limit of (x_m, E_m) belongs to D_m for every m. This is our adaptation of [10, Lemma 3.4].

Lemma 6.7. *The following statements hold:*

    (1) $f_m \to f$ *pointwise, where* $f : \mathbb{H}^n \to \mathbb{R}$ *is Lipschitz and* $\mathrm{Lip}_{\mathbb{H}}(f) \leq 1$,
    (2) $f - f_m$ *is* $\mathbb{H}$-*linear and* $\mathrm{Lip}_{\mathbb{H}}(f - f_m) \leq 2t_m$ *for* $m \geq 0$,
    (3) $x_m \to x_* \in N$ *and* $E_m(0) \to E_*(0)$ *for some* $E_* \in V$ *with* $\omega(E_*) = 1$. *For* $m \geq 0$:

$$d(x_*, x_m) < \delta_m \quad and \quad d(E_*(0), E_m(0)) \leq \sigma_m,$$

(4) $E_* f(x_*)$ *exists, is strictly positive and* $E_m f_m(x_m) \uparrow E_* f(x_*)$,

(5) $(x_{m-1}, E_{m-1}) \leq_{(f_m, \sigma_{m-1} - \varepsilon'_m)} (x_*, E_*)$ *for* $m \geq 1$,

(6) $(x_*, E_*) \in D_m$ *for* $m \geq 1$.

Proof. We prove each statement individually.

Proof of (1). Algorithm 6.4(1) gives $f_m(x) = f_0(x) + \langle x, \sum_{k=0}^{m-1} t_k E_k(0) \rangle$. Define $f : \mathbb{H}^n \to \mathbb{R}$ by

$$f(x) := f_0(x) + \left\langle x, \sum_{k=0}^{\infty} t_k E_k(0) \right\rangle. \tag{6.11}$$

Notice $|f(x) - f_m(x)| \leq |x| \sum_{k=m}^{\infty} t_k |E_k(0)|$. Hence $f_m \to f$ pointwise and, since $\mathrm{Lip}_{\mathbb{H}}(f_m) \leq 1$, we deduce $\mathrm{Lip}_{\mathbb{H}}(f) \leq 1$.

Proof of (2). Lemma 5.2 shows that $f - f_m$ is $\mathbb{H}$-linear. Using also Algorithm 6.4(3) shows that for every $m \geq 0$:

$$\mathrm{Lip}_{\mathbb{H}}(f - f_m) \leq \sum_{k=m}^{\infty} t_k \leq t_m \sum_{k=m}^{\infty} \frac{1}{2^{k-m}} \leq 2t_m.$$

Proof of (3). Let $q \geq m \geq 0$. The definition of D_{q+1} in Algorithm 6.4(5) shows that $(x_q, E_q) \in D_{q+1}$. Hence Lemma 6.6(2,3) imply that $(x_q, E_q) \in D_{m+1}$, and consequently:

$$d(E_q(0), E_m(0)) \leq \sigma_m. \tag{6.12}$$

Since $(x_q, E_q) \in D_{m+1}$, Algorithm 6.4(5) implies:

$$d(x_q, x_m) < \delta_m. \tag{6.13}$$

Since $\sigma_m, \delta_m \to 0$ we see that $(x_m)_{m=1}^{\infty}$ and $(E_m(0))_{m=1}^{\infty}$ are Cauchy sequences, so converge to some $x_* \in \mathbb{H}^n$ and $v \in \mathbb{H}^n$. Since $E_m \in V$ and $\omega(E_m) = 1$, we know $|p(v)| = 1$ and $v_{2n+1} = 0$. Using group translations, we can extend v to a vector field $E_* \in V$ with $\omega(E_*) = 1$ and $E_*(0) = v$. Letting $q \to \infty$ in (6.12) and (6.13) implies $d(E_*(0), E_m(0)) \leq \sigma_m$ and $d(x_*, x_m) \leq \delta_m$. Lemma 6.5 then gives the strict inequality $d(x_*, x_m) < \delta_m$.

We now know that $x_* \in \overline{B_{\mathbb{H}}(x_m, \delta_m)}$ for every $m \geq 1$. Recall that $N = \cap_{m=0}^{\infty} U_m$ for open sets $U_m \subset \mathbb{H}^n$, and Algorithm 6.4(8) states $\overline{B_{\mathbb{H}}(x_m, \delta_m)} \subset U_m$. Hence $x_* \in N$.

Proof of (4). As in the proof of (3) we have $(x_q, E_q) \in D_{m+1}$ for every $q \geq m \geq 0$. Therefore, using Lemma 6.6(1), $q \geq m \geq 1$ implies:

$$(x_{m-1}, E_{m-1}) \leq_{(f_m, \sigma_{m-1} - \varepsilon'_m)} (x_q, E_q). \tag{6.14}$$

Applying Algorithm 6.4(1) and (6.14) (with m and q replaced by $q + 1$) gives:

$$E_q f_q(x_q) < E_q f_{q+1}(x_q) \leq E_{q+1} f_{q+1}(x_{q+1}) \quad \text{for} \quad q \geq 0. \tag{6.15}$$

Hence $(E_q f_q(x_q))_{q=0}^{\infty}$ is strictly increasing and positive as $E_0 f_0(x_0) \geq 0$.

Recall $\mathrm{Lip}_{\mathbb{H}}(f_q) \leq 1$ for $q \geq 1$. Hence, by Lemma 3.3, the sequence $(E_q f_q(x_q))_{q=1}^{\infty}$ is bounded above by 1. Consequently $E_q f_q(x_q) \to L$ for some $0 < L \leq 1$. Inequality (6.15) implies that also $E_q f_{q+1}(x_q) \to L$. Further:

$$E_q f(x_q) = E_q f_q(x_q) + E_q(f - f_q)(x_q)$$

and $|E_q(f - f_q)(x_q)| \leq \mathrm{Lip}_{\mathbb{H}}(f - f_q) \leq 2t_q \to 0$. Hence also $E_q f(x_q) \to L$.

Let $q \geq m \geq 0$ and consider:

$$s_{m,q} := E_q f_m(x_q) - E_{m-1} f_m(x_{m-1}).$$

By (6.14) we have $s_{m,q} \geq 0$. Letting $q \to \infty$, writing $f_m = f + (f_m - f)$, and using $\mathbb{H}$-linearity of $f_m - f$ and $E_q f(x_q) \to L$ implies:

$$s_{m,q} \to s_m := (f_m - f)(E_*(0)) + L - E_{m-1} f_m(x_{m-1}) \geq 0. \qquad (6.16)$$

Also $s_m \to 0$ as $m \to \infty$ since $\mathrm{Lip}_{\mathbb{H}}(f_m - f) \leq 2t_m$ and $E_{m-1} f_m(x_{m-1}) \to L$. Using (6.14) shows that for $t \in (-1,1)$:

$$|(f_m(x_q + tE_{m-1}(x_q)) - f_m(x_q)) - (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}))|$$
$$\leq K(\sigma_{m-1} - \varepsilon'_m + (s_{m,q})^{\frac{1}{4}})|t|. \qquad (6.17)$$

Letting $q \to \infty$ in (6.17) shows that for $t \in (-1,1)$:

$$|(f_m(x_* + tE_{m-1}(x_*)) - f_m(x_*)) - (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}))|$$
$$\leq K(\sigma_{m-1} - \varepsilon'_m + (s_m)^{\frac{1}{4}})|t|. \qquad (6.18)$$

Using $\mathrm{Lip}_{\mathbb{H}}(f) \leq 1$ and $d(E_*(0), E_{m-1}(0)) \leq \sigma_{m-1}$ from (3) of the present Lemma gives:

$$|f(x_* + tE_*(x_*)) - f(x_* + tE_{m-1}(x_*))| \leq d(x_*(tE_*(0)), x_*(tE_{m-1}(0)))$$
$$\leq \sigma_{m-1}|t|.$$

Since $f - f_m$ is $\mathbb{H}$-linear we can use $\mathrm{Lip}_{\mathbb{H}}(f - f_m) \leq 2t_m$ to estimate:

$$|(f - f_m)(x_* + tE_{m-1}(x_*)) - (f - f_m)(x_*)| = |(f - f_m)(tE_{m-1}(0))|$$
$$\leq t\mathrm{Lip}_{\mathbb{H}}(f - f_m)$$
$$\leq 2t_m|t|.$$

Combining the previous three inequalities shows that for $t \in (-1,1)$:

$$|(f(x_* + tE_*(x_*)) - f(x_*)) - (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}))|$$
$$\leq |(f_m(x_* + tE_{m-1}(x_*)) - f_m(x_*))$$
$$\quad - (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}))|$$
$$\quad + |f(x_* + tE_*(x_*)) - f(x_* + tE_{m-1}(x_*))|$$
$$\quad + |(f - f_m)(x_* + tE_{m-1}(x_*)) - (f - f_m)(x_*)|$$
$$\leq (K(\sigma_{m-1} - \varepsilon'_m + (s_m)^{\frac{1}{4}}) + \sigma_{m-1} + 2t_m)|t|.$$

Fix $\varepsilon > 0$ and choose $m \geq 1$ such that:

$$K(\sigma_{m-1} - \varepsilon'_m + (s_m)^{\frac{1}{4}}) + \sigma_{m-1} + 2t_m \leq \varepsilon/3$$

and

$$|E_{m-1} f_m(x_{m-1}) - L| \leq \varepsilon/3.$$

Using the definition of $E_{m-1} f_m(x_{m-1})$, fix $0 < \delta < 1$ such that for $|t| < \delta$:

$$|f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}) - tE_{m-1} f_m(x_{m-1})| \leq \varepsilon|t|/3.$$

Hence for $|t| < \delta$:

$$|f(x_* + tE_*(x_*)) - f(x_*) - tL|$$
$$\leq |(f(x_* + tE_*(x_*)) - f(x_*)) - (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}))|$$
$$\quad + |f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}) - tE_{m-1} f_m(x_{m-1})|$$
$$\quad + |E_{m-1} f_m(x_{m-1}) - L||t|$$
$$\leq \varepsilon|t|.$$

This proves that $E_* f(x_*)$ exists and is equal to L. We already saw that $(E_q f_q(x_q))_{q=1}^{\infty}$ is a strictly increasing sequence of positive numbers. This proves (4).

Proof of (5). The definition of L and Lemma 5.2 implies:

$$E_* f_m(x_*) = L + E_*(f_m - f)(x_*) = L + (f_m - f)(E_*(0)).$$

Using (6.16) shows $s_m = E_* f_m(x_*) - E_{m-1} f_m(x_{m-1}) \geq 0$. Substituting this in (6.18) gives (5).

Proof of (6). Property (6) is a consequence of (3), (4) and (5). $\qquad\square$

We now prove that the limit directional derivative $E_* f(x_*)$ is almost locally maximal in horizontal directions. This is our adaptation of [10, Lemma 3.5].

Lemma 6.8. *For all $\varepsilon > 0$ there is $\delta_\varepsilon > 0$ such that if $(x, E) \in D^f$ satisfies $d(x_*, x) \leq \delta_\varepsilon$ and $(x_*, E_*) \leq_{(f,0)} (x, E)$ then:*

$$Ef(x) < E_* f(x_*) + \varepsilon.$$

Proof. Fix $\varepsilon > 0$. Use Lemma 6.5 to choose $m \geq 1$ such that:

$$m \geq 4/\varepsilon^{\frac{3}{4}} \quad \text{and} \quad \lambda_m, t_m \leq \varepsilon/4. \tag{6.19}$$

Recall $\varepsilon'_m = \min\{\varepsilon_m/2, \sigma_{m-1}/2\}$. Using Lemma 6.7(3), fix $\delta_\varepsilon > 0$ such that $\delta_\varepsilon < \delta_{m-1} - d(x_*, x_{m-1})$ such that for every $|t| < C_{\mathrm{H}}^2 (1 + C_V)^{\frac{1}{2}} \delta_\varepsilon^{\frac{1}{2}} / \varepsilon'_m$:

$$|(f_m(x_* + t E_*(x_*)) - f_m(x_*)) - (f_m(x_{m-1} + t E_{m-1}(x_{m-1})) - f_m(x_{m-1}))|$$
$$\leq (E_* f_m(x_*) - E_{m-1} f_m(x_{m-1}) + \sigma_{m-1})|t|. \tag{6.20}$$

Such δ_ε exists since Lemma 6.7(5) implies $E_* f_m(x_*) \geq E_{m-1} f_m(x_{m-1})$.

We argue by contradiction. Suppose that $(x, E) \in D^f$ satisfies $d(x_*, x) \leq \delta_\varepsilon$, $(x_*, E_*) \leq_{(f,0)} (x, E)$ and $Ef(x) \geq E_* f(x_*) + \varepsilon$. We plan to show that $(x, E) \in D_m$. We first observe that this gives a contradiction. Indeed, Algorithm 6.4(6) and the monotone convergence $E_m f_m(x_m) \uparrow E_* f(x_*)$ would then imply:

$$Ef_m(x) \leq E_m f_m(x_m) + \lambda_m \leq E_* f(x_*) + \lambda_m.$$

From Lemma 6.7(2) and (6.19) we deduce:

$$Ef(x) - E_* f(x_*) = (Ef_m(x) - E_* f(x_*)) + E(f - f_m)(x)$$
$$\leq \lambda_m + 2t_m$$
$$\leq 3\varepsilon/4.$$

This contradicts the assumption that $Ef(x) \geq E_* f(x_*) + \varepsilon$.

Proof that $(x, E) \in D_m$. Notice that $(x, E) \in D^{f_m}$ since $(x, E) \in D^f$ and $D^f = D^{f_m}$ because $f - f_m$ is $\mathbb{H}$-linear. Next observe:

$$d(x, x_{m-1}) \leq d(x, x_*) + d(x_*, x_{m-1}) < \delta_{m-1}.$$

Hence it suffices to show that $(x_{m-1}, E_{m-1}) \leq_{(f_m, \sigma_{m-1} - \varepsilon'_m/2)} (x, E)$. Lemma 3.3 implies:

$$|E(f - f_m)(x)|, |E_*(f - f_m)(x_*)| \leq \mathrm{Lip}_{\mathbb{H}}(f - f_m).$$

Hence using (6.19) and our definition of (x, E), we deduce:

$$Ef_m(x) - E_* f_m(x_*) \geq Ef(x) - E_* f(x_*) - 2\mathrm{Lip}_{\mathbb{H}}(f_m - f)$$
$$\geq \varepsilon - 4t_m \geq 0.$$

Lemma 6.7(6) states that $(x_*, E_*) \in D_m$, which implies $E_{m-1} f_m(x_{m-1}) \leq E_* f_m(x_*)$. Hence:

$$Ef_m(x) \geq E_* f_m(x_*) \geq E_{m-1} f_m(x_{m-1}).$$

In particular, the inequality $Ef_m(x) \geq E_{m-1} f_m(x_{m-1})$ proves the first half of the statement $(x_{m-1}, E_{m-1}) \leq_{(f_m, \sigma_{m-1} - \varepsilon'_m/2)} (x, E)$.

We next deduce several inequalities from our hypotheses. Denote:

- $A := Ef(x) - E_* f(x_*)$,

- $B := Ef_m(x) - E_* f_m(x_*)$,
- $C := Ef_m(x) - E_{m-1} f_m(x_{m-1})$.

Our definition of (x, E) states $A \geq \varepsilon$, while the inequalities above give $0 \leq B \leq C$. Also A, B, $C \leq 2$ by Lemma 3.3. Recall the factorization:

$$(A^{\frac{1}{4}} - B^{\frac{1}{4}})(B^{\frac{3}{4}} + B^{\frac{1}{2}} A^{\frac{1}{4}} + B^{\frac{1}{4}} A^{\frac{1}{2}} + A^{\frac{3}{4}}) = A - B. \qquad (6.21)$$

Using Lemma 6.7(2), (6.19) and Algorithm 6.4(3), we obtain:

$$
\begin{aligned}
A^{\frac{1}{4}} - B^{\frac{1}{4}} &\leq (A - B)/\varepsilon^{\frac{3}{4}} \\
&= (E(f - f_m)(x) - E_*(f - f_m)(x_*))/\varepsilon^{\frac{3}{4}} \\
&\leq 4t_m/\varepsilon^{\frac{3}{4}} \\
&\leq mt_m \\
&\leq \sigma_{m-1}/4.
\end{aligned}
$$

Since B, $C \leq 2$ and $K \geq 8$ we have:

$$B^{\frac{3}{4}} + B^{\frac{1}{2}} A^{\frac{1}{4}} + B^{\frac{1}{4}} A^{\frac{1}{2}} + A^{\frac{3}{4}} \leq 8 \leq K.$$

Hence using (6.21) with A replaced by C gives:

$$KC^{\frac{1}{4}} - KB^{\frac{1}{4}} \geq C - B = E_* f_m(x_*) - E_{m-1} f_m(x_{m-1}).$$

Combining our estimates gives:

$$
\begin{aligned}
E_* f_m(x_*) &- E_{m-1} f_m(x_{m-1}) + K(Ef(x) - E_* f(x_*))^{\frac{1}{4}} \\
&= E_* f_m(x_*) - E_{m-1} f_m(x_{m-1}) + KA^{\frac{1}{4}} \\
&\leq KC^{\frac{1}{4}} - KB^{\frac{1}{4}} + K(B^{\frac{1}{4}} + \sigma_{m-1}/4) \\
&= K((Ef_m(x) - E_{m-1} f_m(x_{m-1}))^{\frac{1}{4}} + \sigma_{m-1}/4). \qquad (6.22)
\end{aligned}
$$

We can now prove the second half of $(x_{m-1}, E_{m-1}) \leq_{(f_m, \sigma_{m-1} - \varepsilon'_m/2)} (x, E)$. We need to estimate:

$$|(f_m(x + tE_{m-1}(x)) - f_m(x)) - (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}))|. \quad (6.23)$$

We consider two cases, depending on whether t is small or large.

Suppose $|t| \leq C_{\mathrm{H}}^2 (1 + C_V)^{\frac{1}{2}} \delta_\varepsilon^{\frac{1}{2}} / \varepsilon'_m$. To estimate (6.23) we use the inequality:

$$
\begin{aligned}
|(f_m(x &+ tE_{m-1}(x)) - f_m(x)) - (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1})| \\
&\leq |(f_m(x + tE_*(x)) - f_m(x)) - (f_m(x_* + tE_*(x_*)) - f_m(x_*))| \\
&\quad + |(f_m(x_* + tE_*(x_*)) - f_m(x_*)) \\
&\qquad - (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}))| \\
&\quad + |f_m(x + tE_{m-1}(x)) - f_m(x + tE_*(x))|. \qquad (6.24)
\end{aligned}
$$

Using Lemma 5.2, the hypothesis $(x_*, E_*) \leq_{(f,0)} (x, E)$ and $\mathbb{H}$-linearity of $f_m - f$, we can estimate the first term in (6.24):

$$
\begin{aligned}
|(f_m(x + tE_*(x_*)) &- f_m(x)) - (f_m(x_* + tE_*(x_*)) - f_m(x_*))| \\
&= |f(x + tE_*(x)) - f(x)) - (f(x_* + tE_*(x_*)) - f(x_*))| \\
&\leq K(Ef(x) - E_* f(x_*))^{\frac{1}{4}} |t|. \qquad (6.25)
\end{aligned}
$$

Using (6.20) and the assumption that t is small bounds the second term in (6.24) by:

$$(E_* f_m(x_*) - E_{m-1} f_m(x_{m-1}) + \sigma_{m-1})|t|.$$

Lemma 6.7 implies that the third term of (6.24) is bounded above by $\sigma_{m-1}|t|$. By combining the estimates of each term and using (6.22) we develop (6.24):

$$|(f_m(x + tE_{m-1}(x)) - f_m(x)) - (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}))|$$

$$\leq (K(Ef(x) - E_*f(x_*))^{\frac{1}{4}} + E_*f_m(x_*) - E_{m-1}f_m(x_{m-1}) + 2\sigma_{m-1})|t|$$

$$\leq (K((Ef_m(x) - E_{m-1}f_m(x_{m-1}))^{\frac{1}{4}} + \sigma_{m-1}/4) + 2\sigma_{m-1})|t|$$

$$\leq K(\sigma_{m-1} - \varepsilon_m'/2 + (Ef_m(x) - E_{m-1}f_m(x_{m-1}))^{\frac{1}{4}})|t|, \tag{6.26}$$

using $\varepsilon_m' \leq \sigma_{m-1}/2$ and $K \geq 8$ in the final line. This gives the correct estimate of (6.23) for small t.

Suppose $C_H^2(1+C_V)^{\frac{1}{2}}\delta_\varepsilon^{\frac{1}{2}}/\varepsilon_m' \leq |t| \leq 1$. To estimate (6.23) we use the inequality:

$$|(f_m(x + tE_{m-1}(x)) - f_m(x)) - (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}))|$$

$$\leq |(f_m(x_* + tE_{m-1}(x_*)) - f_m(x_*))$$

$$- (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}))|$$

$$+ |f_m(x_*) - f_m(x)| + |f_m(x + tE_{m-1}(x)) - f_m(x_* + tE_{m-1}(x_*))|.$$

Lemma 6.7(5) shows that the first term is bounded above by:

$$K(\sigma_{m-1} - \varepsilon_m' + (E_*f_m(x_*) - E_{m-1}f_m(x_{m-1}))^{\frac{1}{4}})|t|.$$

The second term is bounded by $d(x_*, x) \leq \delta_\varepsilon \leq \varepsilon_m'|t| \leq K\varepsilon_m'|t|/4$. For the third term we use Proposition 2.7 to relate the Carnot-Carathéodory distance and the Euclidean distance. Notice $x + tE_{m-1}(x), x_* + tE_{m-1}(x_*) \in \overline{B_H(x_0, 2 + \delta_0)}$. Hence we can use the constants C_H and C_V fixed before Algorithm 6.4:

$$|f_m(x + tE_{m-1}(x)) - f_m(x_* + tE_{m-1}(x_*))|$$

$$\leq C_H|x + tE_{m-1}(x) - x_* - tE_{m-1}(x_*)|^{\frac{1}{2}}$$

$$\leq C_H(1 + C_V)^{\frac{1}{2}}|x_* - x|^{\frac{1}{2}}$$

$$\leq C_H^2(1 + C_V)^{\frac{1}{2}}d(x_*, x)^{\frac{1}{2}}$$

$$\leq C_H^2(1 + C_V)^{\frac{1}{2}}\delta_\varepsilon^{\frac{1}{2}}$$

$$\leq \varepsilon_m'|t|$$

$$\leq K\varepsilon_m'|t|/4.$$

Putting together the three estimates and using $E_*f_m(x_*) \leq Ef_m(x)$ gives:

$$|(f_m(x + tE_{m-1}(x)) - f_m(x)) - (f_m(x_{m-1} + tE_{m-1}(x_{m-1})) - f_m(x_{m-1}))|$$

$$\leq K(\sigma_{m-1} - \varepsilon_m'/2 + (Ef_m(x) - E_{m-1}f_m(x_{m-1}))^{\frac{1}{4}})|t|.$$

This gives the correct estimate of (6.23) for large t.

Combining the two cases estimates (6.23) for any $t \in (-1, 1)$. Hence:

$$(x_{m-1}, E_{m-1}) \leq_{(f_m, \sigma_{m-1} - \varepsilon_m'/2)} (x, E).$$

This concludes the proof. $\qquad\square$

We can now conclude by proving Proposition 6.1 and hence, using also Theorem 5.6, prove Theorem 2.12.

Proof of Proposition 6.1. Lemma 6.7 and Lemma 6.8 prove Proposition 6.1. Indeed, Lemma 6.7 states that there is $f : \mathbb{H}^n \to \mathbb{R}$ Lipschitz such that $f - f_0$ is linear and $\text{Lip}_\mathbb{H}(f - f_0) \leq 2t_0 \leq \mu$. It also states that there is $(x_*, E_*) \in D^f$ satisfying, among other properties, $d(x_*, x_0) < \delta_0$ and $E_*f(x_*) > 0$. Lemma 6.8 then shows that $E_*f(x_*)$ is almost locally maximal in the sense of Proposition 6.1. $\qquad\square$

Proof of Theorem 2.12. Let $f_0 \colon \mathbb{H}^n \to \mathbb{R}$ be a Lipschitz function. Multiplying f_0 by a non-zero constant does not change the set of points where it is Pansu differentiable. Hence we can assume $\mathrm{Lip}_{\mathbb{H}}(f_0) \leq 1/4$. Fix an arbitrary pair $(x_0, E_0) \in D^{f_0}$.

Apply Proposition 6.1 with $\delta_0 = 1$, $\mu = 1/4$ and $K = 8$. This gives a Lipschitz function $f \colon \mathbb{H}^n \to \mathbb{R}$ such that $f - f_0$ is $\mathbb{H}$-linear with $\mathrm{Lip}_{\mathbb{H}}(f - f_0) \leq 1/4$ and a pair $(x_*, E_*) \in D^f$, in particular $x_* \in N$, such that $E_* f(x_*) > 0$ is almost locally maximal in the following sense.

For any $\varepsilon > 0$ there is $\delta_\varepsilon > 0$ such that whenever $(x, E) \in D^f$ satisfies both:

(1) $d(x, x_*) \leq \delta_\varepsilon$, $E f(x) \geq E_* f(x_*)$,
(2) for any $t \in (-1, 1)$:
$$|(f(x + tE_*(x)) - f(x)) - (f(x_* + tE_*(x_*)) - f(x_*))|$$
$$\leq 8|t|(Ef(x) - E_* f(x_*))^{\frac{1}{4}},$$

then:
$$Ef(x) < E_* f(x_*) + \varepsilon.$$

Combining $\mathrm{Lip}_{\mathbb{H}}(f_0) \leq 1/4$ and $\mathrm{Lip}_{\mathbb{H}}(f - f_0) \leq 1/4$ gives $\mathrm{Lip}_{\mathbb{H}}(f) \leq 1/2$. Notice that (x_*, E_*) is also almost locally maximal in the sense of Theorem 5.6, since the restriction on pairs above is weaker than that in Theorem 5.6. Hence Theorem 5.6 implies that f is Pansu differentiable at $x_* \in N$. A $\mathbb{H}$-linear function is Pansu differentiable everywhere. Consequently f_0 is Pansu differentiable at x_*, proving Theorem 2.12. $\square$

References

[1] Alberti, G., Marchese, A.: *On the differentiability of Lipschitz functions with respect to measures in Euclidean spaces*, to appear in Geom. Funct. Anal., preprint available at http://cvgmt.sns.it/paper/2500/.

[2] Alberti, G., Csornyei, M., Preiss, D.: *Differentiability of Lipschitz functions, structure of null sets, and other problems*, Proc. Int. Congress Math. III (2010), 1379–1394.

[3] Bate, D.: *Structure of measures in Lipschitz differentiability spaces*, J. Amer. Math. Soc. 28 (2015), 421–482.

[4] Bonfiglioli, A., Lanconelli, E., Uguzzoni, F.: *Stratified Lie Groups and Potential Theory for Their Sub-Laplacians*, Springer Monographs in Mathematics (2007).

[5] Citti, G., Manfredini, M., Pinamonti, A., Serra Cassano, F.: *Smooth approximation for intrinsic Lipschitz functions in the Heisenberg group*, Calc. Var. & PDEs 49(3-4) (2014), 1279–1308.

[6] Citti, G., Manfredini, M., Pinamonti, A., Serra Cassano, F.: *Poincaré-type inequality for Lipschitz continuous vector fields*, J. Math. Pures Appl. (9) 105 (2016), no. 3, 265–292.

[7] Cheeger, J.: *Differentiability of Lipschitz functions on metric measure spaces*, Geom. Funct. Anal. 9(3) (1999), 428–517.

[8] Capogna, L., Danielli, D., Pauls, S., Tyson, J.: *An introduction to the Heisenberg group and the sub-Riemannian isoperimetric problem*, Birkhauser Progress in Mathematics 259 (2007).

[9] Csornyei, M., Jones, P.: *Product formulas for measures and applications to analysis and geometry*, announced result: http://www.math.sunysb.edu/Videos/dfest/PDFs/38-Jones.pdf.

[10] Doré, M., Maleva, O.: *A compact null set containing a differentiability point of every Lipschitz function*, Math. Ann. 351(3) (2011), 633–663.

[11] Doré, M., Maleva, O.: *A compact universal differentiability set with Hausdorff dimension one*, Israel J. Math. 191(2) (2012), 889–900.

[12] Dymond, M., Maleva, O.: *Differentiability inside sets with upper Minkowski dimension one*, to appear in Michigan Math J., preprint available at arXiv:1305.3154.

[13] Fitzpatrick, S.: *Differentiation of real-valued functions and continuity of metric projections*, Proc. Amer. Math. Soc. 91(4) (1984), 544–548.

[14] Franchi, B., Serapioni, R., Serra Cassano, F.: *Rectifiability and perimeter in the Heisenberg group*, Math. Ann. 321 (2001), 479–531.

[15] Franchi, B., Serapioni, R., Serra Cassano, F.: *Regular submanifolds, graphs and area formula in Heisenberg groups*, Adv. Math. 211(1) (2007), 152–203.

[16] Franchi, B., Serapioni, R., Serra Cassano, F.: *Differentiability of intrinsic Lipschitz Functions within Heisenberg Groups*, J. Geom. Anal. 21(4) (2011), 1044–1084.

[17] Gromov, M.: *Carnot-Carathéodory spaces seen from within*, Progress in Mathematics 144 (1996), 79–323.

[18] Le Donne, E.: *Lecture notes on sub-Riemannian geometry*, notes available at https://sites.google.com/site/enricoledonne/.

[19] Lindenstrauss, J., Preiss, D., Tiser, J.: *Fréchet differentiability of Lipschitz functions and porous sets in Banach spaces*, Annals of Mathematics Studies 179, Princeton University Press (2012).

[20] Magnani, V.: *Towards differential calculus in stratified groups*, J. Aust. Math. Soc. 95(1) (2013), 76–128.

[21] Monti, R.: *Minimal surfaces and harmonic functions in the Heisenberg group*, Nonlinear Analysis: Theory, Methods & Applications 126 (2015), 378–393.

[22] Montgomery, R.: *A tour of subriemannian geometries, their geodesics and applications*, American Mathematical Society, Mathematical Surveys and Monographs 91 (2006).

[23] Pansu, P.: *Metriques de Carnot-Carathéodory et quasiisometries des espaces symetriques de rang un*, Annals of Mathematics 129(1) (1989), 1–60.

[24] Preiss, D.: *Differentiability of Lipschitz functions on Banach spaces*, J. Funct. Anal. 91(2) (1990), 312–345.

[25] Preiss, D., Speight, G.: *Differentiability of Lipschitz functions in Lebesgue null sets*, Inven. Math. 199(2) (2015), 517–559.

[26] Speight, G.: *Lusin Approximation and Horizontal Curves in Carnot Groups*, to appear in Revista Matemática Iberoamericana, preprint available at arXiv:1412.2531.

[27] Serra Cassano, F., Vittone, D.: *Graphs of bounded variation, existence and local boundedness of non-parametric minimal surfaces in the Heisenberg group*, Adv. Calc. Var. 7(4) (2014), 409–492.

[28] Vittone, D.: *The regularity problem for sub-Riemannian geodesics*, notes available at http://cvgmt.sns.it/paper/2416/.

[29] Zahorski, Z.: *Sur l'ensemble des points de non-derivabilite d'une fonction continue*, Bull. Soc. Math. France 74 (1946), 147–178.

(Andrea Pinamonti) Dipartimento di Matematica, Universita di Bologna, Piazza di Porta San Donato 5, Bologna 40126, Italy

E-mail address, Andrea Pinamonti: `Andrea.Pinamonti@gmail.com`

(Gareth Speight) Department of Mathematical Sciences, University of Cincinnati, 2815 Commons Way, Cincinnati, OH 45221, United States

E-mail address, Gareth Speight: `Gareth.Speight@uc.edu`