UNIVERSAL DIFFERENTIABILITY SETS IN CARNOT GROUPS OF ARBITRARILY HIGH STEP

ANDREA PINAMONTI AND GARETH SPEIGHT

ABSTRACT. We show that any Carnot group $\mathbb{G}$ with sufficiently many *deformable directions* contains a measure zero set N such that every Lipschitz map $f \colon \mathbb{G} \to \mathbb{R}$ is differentiable at some point of N. We also prove that model filiform groups satisfy this condition, extending some previous results to a class of Carnot groups of arbitrarily high step. Essential to our work is the question of whether the existence of an (almost) maximal directional derivative $Ef(x)$ in a Carnot group implies the differentiability of a Lipschitz map f at x. We show that such an implication is valid in model Filiform groups for directions that are outside a one-dimensional subspace of horizontal directions. Conversely, we show that this implication fails for every horizontal direction in the free Carnot group of step three and rank two.

1. INTRODUCTION

Rademacher's theorem asserts that every Lipschitz map $f \colon \mathbb{R}^n \to \mathbb{R}^m$ is differentiable almost everywhere with respect to the Lebesgue measure. This important result has been extended to many other spaces and measures [2, 5, 18, 27]. It is also interesting to consider whether Rademacher's theorem admits a converse: given a Lebesgue null set $N \subset \mathbb{R}^n$, does there exist a Lipschitz map $f \colon \mathbb{R}^n \to \mathbb{R}^m$ which is differentiable at no point of N? The answer to this question is yes if and only if $n \leq m$ and combines the work of several authors [38, 30, 31, 3, 8]. In the case where $n > m = 1$, the results in [9, 10, 12] provide a stronger result: there is a compact set of Hausdorff dimension one in $\mathbb{R}^n$ which contains some point of differentiability of any Lipschitz map $f \colon \mathbb{R}^n \to \mathbb{R}$. Such a set may even be chosen with upper Minkowski dimension one [12]. Sets containing a point of differentiability for any real-valued Lipschitz map are called *universal differentiability sets*. We refer the reader to [32] and the references therein for more discussion of such sets.

The present paper continues the investigation of universal differentiability sets in Carnot groups which was started in [28, 17], see also the survey [29]. We recall that a Carnot group (Definition 2.1) is a simply connected Lie group whose Lie algebra $\mathfrak{g}$ admits a stratification, i.e. it admits a decomposition $\mathfrak{g} = V_1 \oplus \cdots \oplus V_s$ where $V_{i+1} = [V_1, V_i]$ for $i = 1, \ldots, s-1$. The subspace V_1 is called the horizontal layer while s is the step of the Carnot group and to some extent indicates its complexity (Carnot groups of step one are simply Euclidean spaces). Carnot groups have a rich geometric structure adapted to the horizontal layer, including translations, dilations, Carnot-Carathéodory (CC) distance, and a Haar measure [1, 7, 15, 23]. In the last two decades Carnot groups have been studied in connection with several

different areas of mathematics, such as PDE, differential geometry, control theory, geometric measure theory, mathematical finance and robotics. Their rich structure allows one to define differentiability of maps between Carnot groups (Definition 2.8). Pansu's theorem states that every Lipschitz map is differentiable almost everywhere with respect to the Haar measure [27]. This is a generalization of Rademacher's theorem to Carnot groups.

In [28], it was shown that Heisenberg groups contain measure zero universal differentiability sets. Heisenberg groups are the most frequently studied non-Euclidean Carnot groups and have step two. In [17] this result was extended to give a measure zero and Hausdorff dimension one universal differentiability set in any step two Carnot group. The present paper extends these results and the associated techniques to higher step Carnot groups satisfying a precise geometric condition, namely having sufficiently many *deformable directions* (Definition 2.15). This is a geometric condition expressing that, roughly speaking, horizontal lines can be nicely modified to pass through nearby points, without changing too much their length or their direction. This condition applies in particular to model filiform groups (Definition 2.22), which can have arbitrarily high step despite their relatively simple Lie brackets. Model filiform groups have been previously investigated in connection with non-rigidity of Carnot groups [25], quasiconformal mappings between Carnot groups [36, 37] and geometric control theory [4].

Before describing more carefully the results of this paper, we briefly discuss the techniques involved in constructing universal differentiability sets. We believe these are of independent interest as they only depend on the geometry of the space involved. In [30, 28, 17], the key technique for constructing measure zero universal differentiability sets builds upon the idea that existence of a *maximal directional derivative* for a Lipschitz map suffices for its differentiability. In Euclidean spaces, this observation takes the following form: if $f\colon \mathbb{R}^n \to \mathbb{R}$ is Lipschitz and $|f'(x,v)| = \mathrm{Lip}(f)$ for some direction $v \in \mathbb{R}^n$ with $|v| = 1$, then f is differentiable at x, see [13]. However, a general Lipschitz map may not have such a maximal directional derivative. In [30] it was shown that any Lipschitz map $f\colon \mathbb{R}^n \to \mathbb{R}$ admits a linear perturbation that has an *almost maximal directional derivative* at some point x in a direction v. They also show that almost maximality suffices for differentiability and the point x can be chosen inside a measure zero set N that is independent of f. Combining the two facts we have that N is a universal differentiability set of measure zero.

In [28, 17], the present authors and E. Le Donne showed that if $f\colon \mathbb{G} \to \mathbb{R}$ is a Lipschitz map on a step two Carnot group and $Ef(x)$ is a maximal directional derivative (Definition 2.13), then f is differentiable at x in the sense of Pansu, see Definition 2.8). Moreover, in step two Carnot groups, it was also shown that almost maximality of a directional derivative suffices for differentiability. Generalizing the Euclidean techniques one can then construct a measure zero and Hausdorff dimension one universal differentiability set. Moreover, for each horizontal direction E in an arbitrary Carnot group, differentiability of the CC distance at $\exp(E)$ is equivalent to validity of the following implication: maximality of $Ef(x)$ for Lipschitz $f\colon \mathbb{G} \to \mathbb{R}$ implies differentiability of f at x (Proposition 2.14). However, in the Engel group, which represents the simplest step 3 Carnot group, neither of the

above properties hold. The counterexample is simply given by the horizontal direction X_2, since the CC distance fails to be differentiable at $\exp(X_2)$. It is then clear that the geometry of the space impacts the differentiability of its Lipschitz maps.

The reason why maximality implies differentiability (Proposition 2.14(2)) fails for the direction X_2 in the Engel group is that horizontal lines in the direction X_2 cannot be modified to pass through nearby points without increasing too much their length. If 'maximality implies differentiability' fails, then so does the stronger implication 'almost maximality implies differentiability'. This stronger implication depends upon the possibility of modifying horizontal lines with some controlled bounds on both their length *and* their direction. This stronger modification is useful because in 'almost maximality' the directional derivative is maximal only compared to directional derivatives coming from pairs of points and directions which satisfy estimates expressed using difference quotients of the Lipschitz map. A direction is deformable if suitable deformations of horizontal lines are possible. All horizontal directions in step two Carnot groups are deformable. This was proved in [17], though the word deformable was not used there. In the present paper we show that in model filiform groups $\mathbb{E}_n$, any horizontal direction other than $\pm X_2$ is deformable (Theorem 4.4). We also show that $\pm X_2$ are deformable in $\mathbb{E}_n$ if and only if $n = 2$ or $n = 3$ (Corollary 3.4).

A set N in a Carnot group $\mathbb{G}$ is a universal differentiability set if every Lipschitz map $f \colon \mathbb{G} \to \mathbb{R}$ is differentiable at some point of N (Definition 2.10). We say that a set has CC Hausdorff dimension one if it has Hausdorff dimension one with respect to the CC metric. Our main result is the following.

Theorem 1.1. *Let $\mathbb{G} \neq \mathbb{R}$ be any Carnot group that has a ball of uniformly deformable directions (see Assumptions 6.1). Then $\mathbb{G}$ contains a universal differentiability set $N \subset \mathbb{G}$ of CC Hausdorff dimension one (in particular measure zero).*

In particular, all model filiform groups $\mathbb{E}_n$ for $n \geq 2$ contain a CC Hausdorff dimension one universal differentiability set.

A ball of uniformly deformable directions is needed in Theorem 1.1 because one constructs the measure zero UDS using countably many horizontal curves which are dense in some sense. To prove that 'almost maximality implies differentiability' (Theorem 6.6) one needs to approximate the almost maximal direction with a sequence of deformable ones. If we do not require that the UDS has measure zero, then only one deformable direction is needed to show that almost maximality implies differentiability.

Notice that Theorem 1.1 applies to the Engel group $\mathbb{E}_4$, which was the problematic group in [17]. Hence one may ask whether Theorem 1.1 holds without assuming Assumptions 6.1. Our second result shows that, unless one fundamentally changes the techniques used, the class of Carnot groups must indeed be restricted.

Theorem 1.2. *In the free Carnot group $\mathbb{F}_{2,3}$ with rank two and step three, the CC distance is not differentiable at $\exp(E)$ for any horizontal direction $E \in V_1$.*

Consequently, 'maximality implies differentiability' fails in $\mathbb{F}_{2,3}$ for every horizontal direction.

This improves in a strong way upon [17], where it was shown that maximality implies differentiability fails for one direction in the Engel group. It would be interesting to know whether $\mathbb{F}_{2,3}$ contains a measure zero UDS or instead the opposite result holds: for every null set $N \subset \mathbb{F}_{2,3}$, does there exists a Lipschitz map

$f\colon \mathbb{F}_{2,3} \to \mathbb{R}$ which is differentiable at no point of N? At present we do not know the answer to this question.

The notion of deformability introduced in the present paper seems to share some analogy with the property of not being an abnormal curve, see [35] for the definition of abnormal curves. For example, it is known that in $\mathbb{F}_{2,3}$ the set of all abnormal curves coincides with the set of horizontal lines [1]. This phenomenon could explain why 'maximality implies differentiability' fails in $\mathbb{F}_{2,3}$ for every horizontal direction. A similar characterization holds in model filiform groups, which admit only one abnormal curve which is given by the $\pm X_2$ direction [1]. However the picture is far from being clear. For example in $\mathbb{F}_{3,2}$, where 'maximality implies differentiability' holds for every horizontal direction [17], every horizontal line is an abnormal curve and viceversa, see [16, Proposition 3.11, Theorem 3.14] and [26]. We plan to investigate this possible relation in future works.

We now describe the structure of the paper. In Section 2 we recall the necessary background on Carnot groups and differentiability. In Section 3 we investigate the differentiability of the CC distance. We show that, if E is a deformable direction, then the CC distance is differentiable at $\exp(E)$ (Proposition 3.1) and that in any model filiform group $\mathbb{E}_n$, with $n \geq 4$, the CC distance is not differentiable at $\exp(\pm X_2)$ (Proposition 3.3). We eventually prove Theorem 1.2. In Section 4 we prove Lemma 4.1 and Lemma 4.2 which allow us to construct suitable horizontal curves in model filiform groups. These are then used to show that in every model filiform group all horizontal directions other than $\pm X_2$ are deformable (Theorem 4.4). In Section 5 we prove an estimate for distances between piecewise linear curves with similar directions (Lemma 5.2). In Section 6 we consider Carnot groups $\mathbb{G}$ that contain a ball of uniformly deformable directions with parameters. With this assumption we construct a universal differentiability set (Lemma 6.2) and prove that almost maximality implies differentiability if the direction belongs to the given ball (Theorem 6.6). In Section 7 we show that any Lipschitz map $f\colon \mathbb{G} \to \mathbb{R}$ admits a group linear perturbation which has an almost maximal direction derivative at some point x in some horizontal direction E (Proposition 7.1). Moreover, the point x can be found inside a given measure zero G_δ set and the direction E can be found close to a starting direction E_0. A proof of Theorem 1.1 is given by combining Theorem 6.6 and Proposition 7.1.

2. Preliminaries

In this section we recall concepts which will be important throughout the paper.

2.1. **Basic notions in Carnot groups.**

Definition 2.1. A *Carnot group* $\mathbb{G}$ of *step s* is a simply connected Lie group whose Lie algebra $\mathfrak{g}$ admits a decomposition as a direct sum of subspaces of the form

$$\mathfrak{g} = V_1 \oplus V_2 \oplus \cdots \oplus V_s$$

such that $V_i = [V_1, V_{i-1}]$ for any $i = 2, \ldots, s$, and $[V_1, V_s] = 0$. The subspace V_1 is called the *horizontal layer* and its elements are called *horizontal left invariant vector fields*. The *rank* of $\mathbb{G}$ is $\dim(V_1)$.

The exponential mapping $\exp\colon \mathfrak{g} \to \mathbb{G}$ is a diffeomorphism. Given a basis $X_1, \ldots, X_n$ of $\mathfrak{g}$ adapted to the stratification, any $x \in \mathbb{G}$ can be written in a unique way as
$$x = \exp(x_1 X_1 + \ldots + x_n X_n).$$
We identify x with $(x_1, \ldots, x_n) \in \mathbb{R}^n$ and hence $\mathbb{G}$ with $\mathbb{R}^n$. This is known as *exponential coordinates of the first kind*. To compute the group law in these coordinates, one uses the equality
$$\exp(X)\exp(Y) = \exp(X \diamond Y) \quad \text{for all } X, Y \in \mathfrak{g}.$$
Here $\diamond$ is defined by the Baker-Campbell-Hausdorff (BCH) formula
$$X \diamond Y = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}([X, [X, Y]] + [Y, [Y, X]]) + \ldots, \qquad (2.1)$$
where higher order terms are nested commutators of X and Y [34], see e.g. [4, Theorem 2.2.13].

Unless otherwise stated, $\mathbb{G}$ will be a Carnot group of step s and rank r with $\dim(\mathfrak{g}) = n$ which is represented in exponential coordinates of the first kind.

We say that a curve $\gamma\colon [a, b] \to \mathbb{G}$ is absolutely continuous if it is absolutely continuous as a curve into $\mathbb{R}^n$. Fix a basis $X_1, \ldots, X_r$ of V_1 and an inner product norm ω on V_1 making the chosen basis orthonormal.

Definition 2.2. An absolutely continuous curve $\gamma\colon [a, b] \to \mathbb{G}$ is *horizontal* if there exist $u_1, \ldots, u_r \in L^1[a, b]$ such that
$$\gamma'(t) = \sum_{j=1}^{r} u_j(t) X_j(\gamma(t)) \quad \text{for almost every } t \in [a, b].$$
The *length* of such a curve is $L_{\mathbb{G}}(\gamma) := \int_a^b |u|$.

Since $\mathbb{G}$ is identified with $\mathbb{R}^n$ as a manifold, its tangent spaces are also naturally identified with $\mathbb{R}^n$. We say that a vector $v \in \mathbb{R}^n$ is *horizontal* at $p \in \mathbb{G}$ if $v = E(p)$ for some $E \in V_1$. Thus a curve γ is horizontal if and only if $\gamma'(t)$ is horizontal at $\gamma(t)$ for almost every t. All the curves of the form $t \mapsto p \exp(tV)$ for some $p \in \mathbb{G}$ and $E \in V_1$ are horizontal and they will be called *horizontal lines*.

Chow's theorem [6] asserts that any two points in a Carnot group can be connected by a horizontal curve. Hence the following definition gives a metric on $\mathbb{G}$.

Definition 2.3. The *Carnot-Carathéodory (CC) distance* between any two points $x, y \in \mathbb{G}$ is defined by
$$d(x, y) := \inf\{L_{\mathbb{G}}(\gamma)\colon \gamma \text{ is a horizontal curve joining } x \text{ and } y\}.$$
We also use the notation $d(x) := d(x, 0)$ for $x \in \mathbb{G}$.

Left group translations preserve lengths of horizontal curves. This implies
$$d(gx, gy) = d(x, y) \quad \text{for every } g, x, y \in \mathbb{G}.$$
Even though the CC distance and the Euclidean distance are not Lipschitz equivalent, they induce the same topology. Hence $\mathbb{Q}^n$ is dense in $\mathbb{R}^n$ with respect to the

CC distance. The following proposition will be useful to compare the two distances [24], see also [4, Corollary 5.2.10 and Proposition 5.15.1].

Proposition 2.4. *Let $\mathbb{G}$ be a Carnot group of step s and $K \subset \mathbb{G}$ be a compact set. Then there exists a constant $C_{\mathrm{H}} \geq 1$ depending on K such that*

$$C_{\mathrm{H}}^{-1}|x - y| \leq d(x,y) \leq C_{\mathrm{H}}|x - y|^{1/s} \qquad \text{for all } x,y \in K.$$

We will also need the following estimate for the CC distance [14, Lemma 2.13].

Proposition 2.5. *Let $\mathbb{G}$ be a Carnot group of step s. Then there is a constant $C_D \geq 1$ such that*

$$d(x^{-1}yx) \leq C_D \left(d(y) + d(x)^{\frac{1}{s}} d(y)^{\frac{s-1}{s}} + d(x)^{\frac{s-1}{s}} d(y)^{\frac{1}{s}} \right) \quad \text{for } x,y \in \mathbb{G}.$$

Definition 2.6. For any $\lambda > 0$, we define the *dilation* $\delta_\lambda \colon \mathbb{G} \to \mathbb{G}$ in coordinates by

$$\delta_\lambda(x_1, \ldots, x_n) = (\lambda^{\alpha_1} x_1, \ldots, \lambda^{\alpha_n} x_n)$$

where $\alpha_i \in \mathbb{N}$ is the homogeneity of the variable x_i, which is defined by

$$\alpha_j = i \qquad \text{whenever} \qquad h_{i-1} + 1 < j \leq h_i,$$

where $h_i := \dim(V_1) + \ldots + \dim(V_i)$ for $i \geq 1$ and $h_0 := 0$. For our purposes, it will be enough to know that $\alpha_1 = \cdots = \alpha_r = 1$, where $r = \dim(V_1)$.

Dilations are group homomorphisms of $\mathbb{G}$ and they satisfy

$$d(\delta_\lambda(x), \delta_\lambda(y)) = \lambda d(x,y) \quad \text{for every } x,y \in \mathbb{G} \text{ and } \lambda > 0.$$

We will also use the fact that $\delta_\lambda(\exp(E)) = \exp(\lambda E)$ for every $\lambda > 0$ and $E \in V_1$.

Carnot groups have a Haar measure which is unique up to scalars. When $\mathbb{G}$ is represented in first exponential coordinates as $\mathbb{R}^n$, the Haar measure is simply the Lebesgue measure $\mathcal{L}^n$, which satisfies

$$\mathcal{L}^n(gA) = \mathcal{L}^n(A) \qquad \text{and} \qquad \mathcal{L}^n(\delta_\lambda(A)) = \lambda^Q \mathcal{L}^n(A)$$

for every $g \in \mathbb{G}$, $\lambda > 0$ and $A \subset \mathbb{G}$ measurable. Here $Q := \sum_{i=1}^s i \dim(V_i)$ is the *homogeneous dimension* of $\mathbb{G}$, which is also the Hausdorff dimension of $\mathbb{G}$ with respect to the CC metric.

2.2. Differentiability in Carnot groups.

Definition 2.7. Let $f \colon \mathbb{G} \to \mathbb{R}$ be a Lipschitz function, $x \in \mathbb{G}$ and $E \in V_1$. The *directional derivative of f at x in direction E* is defined by

$$Ef(x) := \lim_{t \to 0} \frac{f(x \exp(tE)) - f(x)}{t},$$

whenever the limit exists.

Pansu defined the notion of differentiability in Carnot groups and proved a Rademacher theorem for maps between general Carnot groups [27]. We will only be concerned with the case where the target is $\mathbb{R}$.

Definition 2.8. A function $L \colon \mathbb{G} \to \mathbb{R}$ is $\mathbb{G}$-*linear* if $L(xy) = L(x) + L(y)$ and $L(\delta_r(x)) = rL(x)$ for all $x,y \in \mathbb{G}$ and $r > 0$.

Let $f \colon \mathbb{G} \to \mathbb{R}$ and $x \in \mathbb{G}$. We say that f is *differentiable at x* if there is a $\mathbb{G}$-linear map $L \colon \mathbb{G} \to \mathbb{R}$ 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 differential* of f at x.

Theorem 2.9 (Pansu). *Every Lipschitz function $f \colon \mathbb{G} \to \mathbb{R}$ is differentiable almost everywhere with respect to the Haar measure on $\mathbb{G}$.*

Note that Theorem 2.9 also holds for Carnot group targets [27] and even for suitable infinite dimensional targets [21, 22].

Definition 2.10. A set $N \subset \mathbb{G}$ is called a *universal differentiability set (UDS)* if every Lipschitz map $f \colon \mathbb{G} \to \mathbb{R}$ is differentiable at a point of N.

Theorem 2.9 implies that every positive measure subset of $\mathbb{G}$ is a UDS [19]. In Euclidean space $\mathbb{R}^n$ for $n > 1$, where the group law is simply addition and the step is 1, measure zero UDS exist and they can be made compact and of Hausdorff and Minkowski dimension one [30, 10, 12]. All step 2 Carnot groups contain a measure zero UDS of Hausdorff dimension one with respect to the CC metric [28, 17]. Note that the Hausdorff dimension of any UDS must be at least one [17].

Define the horizontal projection $p \colon \mathbb{G} \to \mathbb{R}^r$ by $p(x) := (x_1, \ldots, x_r)$. We now recall some relevant results from [17].

Lemma 2.11. *Let $u = \exp(E)$ for some $E \in V_1$. Then*

$$d(uz) \geq d(u) + \langle p(z), p(u)/d(u) \rangle \qquad \text{for any } z \in \mathbb{G}.$$

Moreover, if the CC distance $d \colon \mathbb{G} \to \mathbb{R}$ is differentiable at $u = \exp(E)$, then its Pansu differential at u takes the form $z \mapsto \langle p(z), p(u)/d(u) \rangle$.

Recall that ω is an inner product norm on V_1 making the basis $X_1, \ldots, X_r$ of V_1 orthonormal. We have the following connection between directional derivatives and the Lipschitz constant of a Lipschitz map.

Lemma 2.12. *Let $f \colon \mathbb{G} \to \mathbb{R}$ be a Lipschitz map. Then*

$$\mathrm{Lip}(f) = \sup\{|Ef(x)| \colon x \in \mathbb{G}, \ E \in V_1, \ \omega(E) = 1, \ Ef(x) \text{ exists}\}.$$

This justifies the definition of maximal directional derivative.

Definition 2.13. Let $f \colon \mathbb{G} \to \mathbb{R}$ be Lipschitz and let $E \in V_1$ with $\omega(E) = 1$. We say that a directional derivative $Ef(x)$ is *maximal* if $|Ef(x)| = \mathrm{Lip}(f)$.

In Euclidean spaces (and Banach spaces with a differentiable norm), maximality of a directional derivative suffices for differentiability. The following proposition from [17] gives a condition for 'maximality implies differentiability' using the differentiability of the CC distance.

Proposition 2.14. *Let $E \in V_1$ with $\omega(E) = 1$. Then the following are equivalent.*
  (1) *The CC distance d is differentiable at $\exp(E)$.*
  (2) *The following implication holds: whenever $f \colon \mathbb{G} \to \mathbb{R}$ is Lipschitz and $Ef(x)$ is maximal for some $x \in \mathbb{G}$, then f is differentiable at x.*

Known constructions of measure zero UDS rely upon a stronger implication, namely that the existence of an *almost maximal* directional derivative suffices for differentiability. To investigate this stronger implication, deformability as defined below will be important. First recall that two horizontal curves $f_1 \colon [a, b] \to \mathbb{G}$ and $f_2 \colon [b, c] \to \mathbb{G}$ with $f_1(b) = f_2(b)$ can be joined to form a horizontal curve $f \colon [a, c] \to \mathbb{G}$ given by $f(t) = f_1(t)$ if $a \leq t \leq b$ and $f(t) = f_2(t)$ if $b \leq t \leq c$. Similarly one can join any finite number of horizontal curves provided the end of each curve agrees with the start of the subsequent curve.

Definition 2.15. We say that $E \in V_1$ with $\omega(E) = 1$ is *deformable* if there exist C_E, N_E and a map $\Delta_E \colon (0, \infty) \to (0, \infty)$ such that the following condition holds.

For every $0 < s < 1, \eta \in (0, \infty), 0 < \Delta < \Delta_E(\eta)$ and $u \in \mathbb{G}$ with $d(u) \le 1$, there is a Lipschitz horizontal curve $g \colon \mathbb{R} \to \mathbb{G}$ formed by joining at most N_E horizontal lines such that

(1) $g(t) = \exp(tE)$ for $|t| \ge s$,
(2) $g(\zeta) = \delta_{\Delta s}(u)$, where $\zeta := \langle \delta_{\Delta s}(u), E(0) \rangle$,
(3) $\mathrm{Lip}_{\mathbb{G}}(g) \le 1 + \eta \Delta$,
(4) $|(p \circ g)'(t) - p(E)| \le C_E \Delta$ for all but finitely many $t \in \mathbb{R}$.

Remark 2.16. Consider the restriction of the curve g from Definition 2.15 to the interval $[-s, \zeta]$. By applying left translations and reparameterizing, we obtain a curve $\varphi \colon [0, s + \zeta] \to \mathbb{G}$ with $\varphi(0) = 0$, $\varphi(s + \zeta) = \exp(sE)\delta_{\Delta s}(u)$, and satisfying conditions 3 and 4 of Definition 2.15.

2.3. **Useful facts about exponential coordinates of the first kind.** In the first r coordinates, the group operation and dilations δ_λ read as

$$p(xy) = p(x) + p(y) \quad \text{and} \quad p(\delta_\lambda(x)) = \lambda p(x) \quad \text{for } x, y \in \mathbb{G} \text{ and } \lambda > 0.$$

Let $e_1, \ldots, e_n$ be the standard basis vectors of $\mathbb{R}^n$. If $1 \le j \le r$, the basis element X_j of $V_1 \subset \mathfrak{g}$ can be written as

$$X_j(x) = e_j + \sum_{i > r}^{n} q_{i,j}(x)e_i, \tag{2.2}$$

where $q_{i,j}$ are homogeneous polynomials, in particular, $q_{i,j}(0) = 0$. Using (2.2), it follows that $\exp(E) = E(0)$ for any $E \in V_1$. Thus points $u = \exp(E)$ for some $E \in V_1$ are exactly those of the form $u = (u_h, 0)$ for some $u_h \in \mathbb{R}^r$ and therefore $\exp(E) = (p(\exp(E)), 0)$. If $E \in V_1$, it follows from (2.2) that $p(E(x))$ is independent of $x \in \mathbb{G}$. Hence one can unambiguously define $p(E) \in \mathbb{R}^r$ for every $E \in V_1$. The inner product norm ω is equivalently given by $\omega(E) = |p(E)|$.

From Definition 2.2 and (2.2) we notice that $L_{\mathbb{G}}(\gamma) = L_{\mathbb{E}}(p \circ \gamma)$, where $L_{\mathbb{E}}$ is the Euclidean length of a curve in $\mathbb{R}^r$. This implies that

$$d(x, y) \ge |p(y) - p(x)| \quad \text{for all} \quad x, y \in \mathbb{G}.$$

Lemma 2.17 and Lemma 2.18 below give useful facts about length and distance in coordinates. They can be proved exactly as in [28, Lemma 2.8 and Lemma 2.9].

Lemma 2.17. *If $E \in V_1$ then the following facts hold.*

(1) $|E(0)| = \omega(E) = d(E(0))$,
(2) $d(x, x \exp(tE)) = t\omega(E)$ for any $x \in \mathbb{G}$ and $t \in \mathbb{R}$.

Lemma 2.18. *Suppose that $\gamma \colon I \to \mathbb{G}$ is a horizontal curve. Then*

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

The following lemma gives easy facts about the simplest $\mathbb{G}$-linear maps. It can be easily proved, e.g. as in [28, Lemma 5.2].

Lemma 2.19. *Suppose $E \in V_1$ with $\omega(E) = 1$ and let $L \colon \mathbb{G} \to \mathbb{R}$ be the function defined by $L(x) = \langle x, E(0) \rangle$. Then the following facts hold.*

(1) *L is $\mathbb{G}$-linear and $\mathrm{Lip}_{\mathbb{G}}(L) = 1$,*
(2) *for every $x \in \mathbb{G}$ and every $\tilde{E} \in V_1$ one has*

$$\tilde{E}L(x) = L(\tilde{E}(0)) = \langle p(\tilde{E}), p(E) \rangle.$$

2.4. **Free Carnot groups and model filiform groups.** Recall that a homomorphism between Lie algebras is a linear map that preserves the Lie bracket, while isomorphisms are bijective homomorphisms. Free-nilpotent Lie algebras are then defined as follows (e.g. [4, Definition 14.1.1]).

Definition 2.20. Let $r \geq 2$ and $s \geq 1$ be integers. We say that $\mathcal{F}_{r,s}$ is the *free-nilpotent Lie algebra* with r *generators* $x_1, \ldots, x_r$ of *step* s if:

(1) $\mathcal{F}_{r,s}$ is a Lie algebra generated by elements $x_1, \ldots, x_r$,
(2) $\mathcal{F}_{r,s}$ is nilpotent of step s (i.e., nested Lie brackets of length $s + 1$ are 0),
(3) for every Lie algebra $\mathfrak{g}$ that is nilpotent of step s and for every map $\Phi \colon \{x_1, \ldots, x_r\} \to \mathfrak{g}$, there is a unique homomorphism of Lie algebras $\tilde{\Phi} \colon \mathcal{F}_{r,s} \to \mathfrak{g}$ that extends Φ.

We next define free Carnot groups, e.g. [4, Definition 14.1.3].

Definition 2.21. The *free Carnot group* with rank r and step s is the Carnot group whose Lie algebra is isomorphic to the free-nilpotent Lie algebra $\mathcal{F}_{r,s}$. We denote it by $\mathbb{F}_{r,s}$.

By saying that two Carnot groups are isomorphic we simply mean that they are isomorphic as Lie groups, with an isomorphism that preserves the stratification. Since Carnot groups are simply connected Lie groups, any homomorphism ϕ between their Lie algebras lifts to a Lie group homomorphism F between the Carnot groups satisfying $dF = \phi$.

Intuitively, model filiform groups are the Carnot groups with the simplest Lie brackets possible while still having arbitrarily large step. The formal definition is as follows.

Definition 2.22. Let $n \geq 2$. The *model filiform group of step* $n - 1$ is the Carnot group $\mathbb{E}_n$ whose Lie algebra $\mathcal{E}_n$ admits a basis $X_1, \ldots, X_n$ for which the only non-vanishing bracket relations are given by $[X_i, X_1] = X_{i+1}$ for $1 < i < n$.

The stratification of $\mathcal{E}_n$ is $\mathcal{E}_n = V_1 \oplus \cdots \oplus V_{n-1}$ with $V_1 = \text{Span}\{X_1, X_2\}$ and $V_i = \text{Span}\{X_{i-1}\}$ for $1 < i < n$.

Throughout the paper, we will view the model Filiform group $\mathbb{E}_n$ in first exponential coordinates as $\mathbb{R}^n$ with group operation obtained from the Lie brackets by the BCH formula (2.1).

3. Differentiability of the CC distance in Carnot groups

In this section we investigate the differentiability of the CC distance at endpoints of horizontal vectors. By Proposition 2.14, this is equivalent to the implication 'maximality implies differentiability'.

3.1. **Deformability implies differentiability of the CC distance.** We first observe that if E is a deformable direction then the CC distance is differentiable at $\exp(E)$. At present we do not know whether the converse holds.

Proposition 3.1. *Suppose $\mathbb{G}$ is a Carnot group and let $E \in V_1$ with $\omega(E) = 1$ be deformable. Then the CC distance is differentiable at $\exp(E)$.*

Proof. First notice that Lemma 2.11 gives $d(\exp(E)z) \geq 1 + \langle \exp(E), z \rangle$ for any $z \in \mathbb{G}$. Hence it suffices to derive a suitable upper bound for $d(\exp(E)z)$.

Let $\eta > 0$ and $\Delta_E \colon (0, \infty) \to (0, \infty)$ be as in Definition 2.15. Suppose $z \in \mathbb{G}$ satisfies $d(z) < \min(\Delta_E(\eta)/2, \, 1)$ and let $s = 1/2$. Then we may choose $u \in \mathbb{G}$ with $d(u) = 1$ and $0 < \Delta < \Delta_E(\eta)$ so that $z = \delta_{\Delta s}(u)$. By applying Definition 2.15 with this choice of η, s, u, Δ we find a Lipschitz horizontal curve $g \colon \mathbb{R} \to \mathbb{G}$ satisfying $\mathrm{Lip}(g) \le 1 + \eta\Delta$, $g(-1) = \exp(-E)$ and $g(\zeta) = z$, where $\zeta = \langle \exp(E), z \rangle$. Since $|\zeta| \le 1$, we now estimate as follows:

$$d(\exp(E)z) = d(\exp(-E), z) = d(g(-1), g(\zeta))$$
$$\le (1 + \eta\Delta)|1 + \zeta|$$
$$\le 1 + \langle \exp(E), z \rangle + 4\eta d(z).$$

This shows that given $\eta > 0$ we have $d(\exp(E)z) \le 1 + \langle \exp(E), z \rangle + 4\eta d(z)$ for all sufficiently small z. This completes the proof. $\qquad\square$

3.2. **A strong example of non-differentiability of the CC distance.** The CC distance in the Engel group (which is the model filiform group $\mathbb{E}_4$) is not differentiable at $\exp(X_2)$ [17]. In other words, the implication 'maximality implies differentiability' fails for the direction X_2 in the Engel group. We now derive some consequences of this result for other Carnot groups.

Fix two Carnot groups $\mathbb{G}$ and $\mathbb{H}$ of rank r which have horizontal layers V and W with the following property. There exist bases $\mathbf{X} = (X_1, \ldots, X_r)$ and $\mathbf{Y} = (Y_1, \ldots, Y_r)$ of V and W respectively, together with a Lie group homomorphism $F \colon \mathbb{G} \to \mathbb{H}$ such that $F_*(X_i) = Y_i$ for $1 \le i \le r$.

Equip $\mathbb{G}$ and $\mathbb{H}$ with the CC metrics $d_{\mathbb{G}}$ and $d_{\mathbb{H}}$ induced by the bases $\mathbf{X}$ and $\mathbf{Y}$ respectively. We view both $\mathbb{G}$ and $\mathbb{H}$ in exponential coordinates of the first kind and let $p_{\mathbb{G}} \colon \mathbb{G} \to \mathbb{R}^r$ and $p_{\mathbb{H}} \colon \mathbb{H} \to \mathbb{R}^r$ denote the horizontal projections. For any $u \in \mathbb{G}$ we have $p_{\mathbb{H}}(F(u)) = p_{\mathbb{G}}(u)$. Also if $u = (u_h, 0) \in \mathbb{G}$ for some $u_h \in \mathbb{R}^r$, then $F(u) = (u_h, 0) \in \mathbb{H}$.

The following proposition was proven in [17].

Proposition 3.2. *Suppose the CC distance in $\mathbb{G}$ is differentiable at $\exp(E)$ for some $E \in V_1$. Then the CC distance in $\mathbb{H}$ is differentiable at $\exp(F_*(E))$.*

Our first result about non-differentiability of the CC distance is the following.

Proposition 3.3. *In any model filiform group $\mathbb{E}_n$ with $n \ge 4$, the CC distance is not differentiable at $\exp(X_2)$.*

Proof. Let $X_1, \ldots, X_n$ be a basis of the Lie algebra $\mathcal{E}_n$ such that the only non-vanishing bracket relations are given by $[X_i, X_1] = X_{i+1}$ for $1 < i < n$. Let $Y_1, \ldots, Y_4$ denote a similar basis for the Lie algebra $\mathcal{E}_4$. Define a linear map $\Phi \colon \mathcal{E}_n \to \mathcal{E}_4$ by $\Phi(X_i) = Y_i$ for $1 \le i \le 4$ and $\Phi(X_i) = 0$ for $i > 4$. It is easy to see that Φ is a Lie algebra homomorphism, hence lifts to a Lie group homomorphism $F \colon \mathbb{E}_n \to \mathbb{E}_4$ satisfying $F_*(X_i) = Y_i$ for $1 \le i \le 2$. By [17, Theorem 4.2], the CC distance in $\mathbb{E}_4$ is not differentiable at $\exp(Y_2)$. Hence, by Proposition 3.2, the CC distance in $\mathbb{E}_n$ is not differentiable at $\exp(X_2)$. $\qquad\square$

Recall that $\mathbb{E}_2$ is just $\mathbb{R}^2$ and $\mathbb{E}_3$ is a Carnot group of step 2. Combining the results of [17] with Proposition 3.1 and Proposition 3.3 for $n \ge 4$, we obtain the following corollary.

Corollary 3.4. *In the model filiform group $\mathbb{E}_n$, the directions $\pm X_2$ are deformable for $n = 2$ and $n = 3$. They are not deformable for $n \ge 4$.*

Our second result addressing non-differentiability of the CC distance gives an example of a Carnot group where the CC distance fails to be differentiable at the endpoint of every horizontal vector. This is Theorem 1.2.

Proof of Theorem 1.2. Fix an orthonormal basis X_1, X_2 of the horizontal layer of $\mathbb{F}_{2,3}$. It suffices to show that the CC distance in $\mathbb{F}_{2,3}$ is not differentiable at $\exp(E)$ whenever $E = aX_1 + bX_2$ with $a^2 + b^2 = 1$. Let Y_1, Y_2 be a basis of the horizontal layer in the Engel group $\mathbb{E}_4$, where the CC distance defined using Y_1, Y_2 is not differentiable at $\exp(Y_2)$.

Define $W_1 = bY_1 + aY_2$, $W_2 = -aY_1 + bY_2$. Notice W_1, W_2 are orthonormal with respect to the inner product induced by Y_1, Y_2. Hence the CC distance in $\mathbb{E}_4$ obtained from W_1, W_2 is the same as the CC distance obtained from Y_1, Y_2. Now define a linear map Φ from the horizontal layer of $\mathbb{F}_{2,3}$ to the horizontal layer of $\mathbb{E}_4$ by $\Phi(X_i) = W_i$ for $i = 1, 2$. Using the definition of free Lie algebra and lifting, we obtain a Lie group homomorphism $F \colon \mathbb{F}_{2,3} \to \mathbb{E}_4$ such that $F_*(X_i) = W_i$ for $i = 1, 2$. Since $a^2 + b^2 = 1$, we have:

$$F_*(aX_1 + bX_2) = aW_1 + bW_2 = Y_2.$$

Since the CC distance in $\mathbb{E}_4$ is not differentiable at $\exp(Y_2)$, we deduce using Proposition 3.2 that the CC distance in $\mathbb{F}_{2,3}$ cannot be differentiable at $aX_1 + bX_2$. $\square$

4. Deformability in model filiform groups

In this section we work in $\mathbb{E}_n$ for some $n \geq 3$. Our goal is to prove that every horizontal direction in $\mathbb{E}_n$ except possibly $\pm X_2$ is deformable. For simplicity of notation, in this section we identify the model filiform group $\mathbb{E}_n$ with its Lie algebra $\mathcal{E}_n$. Hence for $E \in \mathcal{E}_n$ we will simply write E instead of $\exp(E)$.

4.1. Construction of horizontal curves in filiform groups. We start by proving two lemmas that show how a horizontal line can be perturbed to reach a nearby point. The first lemma shows how to reach a point whose n'th coordinate is known, with small errors in the other vertical coordinates. Given $A \in \mathbb{R}$, we will use the notation

$$E := X_1 + AX_2 \qquad \text{and} \qquad E' := AX_1 - X_2. \tag{4.1}$$

Note that E and E' are orthogonal with respect to ω, for any $A \in \mathbb{R}$.

Lemma 4.1. *For all $A \in \mathbb{R}$ there exist polynomials $P_i(x)$ for $3 \leq i \leq n$ depending on n such that*

- *each $P_i(x)$ is divisible by x^2*
- *the coefficients of each $P_i(x)$ are polynomials in A*

and the following holds.

There exist $\sigma_i \in \{\pm 1\}$ for $1 \leq i \leq 2^{n-2}$ independent of A such that

$$\prod_{i=1}^{2^{n-2}} \frac{1}{2^{n-2}}(E + \sigma_i \eta E') = E + C_n(A^2 + 1)\eta X_n + \sum_{i=3}^{n} P_i(\eta)X_i,$$

where $C_n \neq 0$ is a constant depending on n.

Proof. We define products $p_k(\eta)$ inductively by

$$p_1(\eta) := (E + \eta E')(E - \eta E') \tag{4.2}$$

and

$$p_{k+1}(\eta) := p_k(\eta)p_k(-\eta) \tag{4.3}$$

for all $k \geq 1$. Choose $\eta_i \in \{\pm\eta\}$ for $1 \leq i \leq 2^{n-2}$ such that

$$p_{n-2}(\eta) = \prod_{i=1}^{2^{n-2}} (E + \eta_i E').$$

To establish the lemma, it suffices to prove there exist $C = C_n \neq 0$ and polynomials $P_i(x)$ as in the statement such that

$$p_{n-2}(\eta) = 2^{n-2}E + 2^{n-2}C(A^2 + 1)\eta X_n + \sum_{i=3}^{n} P_i(\eta)X_i. \tag{4.4}$$

To do so we prove by induction that, for every $1 \leq k \leq n - 2$, p_k has the form

$$p_k(\eta) = 2^k E + \sum_{i=4}^{n} S_i(\eta)X_i + \eta \sum_{i=k+2}^{n} \lambda_i X_i, \tag{4.5}$$

where λ_i are constants with $\lambda_{k+2} \neq 0$ and $S_i(x)$ are polynomials divisible by x^2.

To begin proving (4.5), notice that Definition 2.22, (2.1), (4.1) and (4.2) give

$$p_1(\eta) = 2E - (A^2 + 1)\eta X_3 + \sum_{i=4}^{n} Q_i(\eta)X_i, \tag{4.6}$$

where the polynomials $Q_i(x)$ are divisible by x. Hence (4.5) holds for $k = 1$.

Next we suppose that p_k satisfies (4.5) for some $1 \leq k \leq n - 3$; we will show that p_{k+1} has the desired form too. Firstly, the BCH formula (2.1) gives

$$p_{k+1}(\eta) = p_k(\eta)p_k(-\eta)$$
$$= p_k(\eta) + p_k(-\eta) + \frac{1}{2}[p_k(\eta), p_k(-\eta)]$$
$$+ \text{ brackets of length} \geq 3. \tag{4.7}$$

A simple computation using (4.5) yields

$$p_k(\eta) + p_k(-\eta) = 2^{k+1}E + \sum_{i=4}^{n} (S_i(\eta) + S_i(-\eta))X_i \tag{4.8}$$

and

$$\frac{1}{2}[p_k(\eta), p_k(-\eta)] = 2^k \eta \sum_{i=k+2}^{n-1} \lambda_i X_{i+1} + 2^{k-1} \sum_{i=4}^{n-1} (S_i(\eta) - S_i(-\eta))X_{i+1}. \tag{4.9}$$

Note that $(S_i(\eta) - S_i(-\eta))$ is divisible by η^2 because each polynomial $S_i(x)$ is divisible by x^2. The coefficient of X_{k+3} in the first term on the right hand side of (4.9) is $2^k \lambda_{k+2}\eta$, with $\lambda_{k+2} \neq 0$ coming from the induction hypothesis. The other terms in (4.7) are linear combinations of nested commutators of length greater or equal than three, i.e.

$$[Z_1, [Z_2, \ldots, [Z_{M-1}, Z_M]] \ldots]$$

where $Z_i \in \{p_k(\eta), p_k(-\eta)\}$ and $M \geq 3$. By the definition of Z_i we get that each of the previous commutators is a constant multiple of

$$[X_1, [X_1, \ldots, [Z_{M-1}, Z_M]] \ldots]$$
$$= 2^{k+1} \eta \sum_{i=k+2}^{n+1-M} \lambda_i X_{i+M-1} + 2^k \sum_{i=4}^{n+1-M} (S_i(\eta) - S_i(-\eta)) X_{i+M-1}.$$

The leading term in the first sum above is a multiple of X_j with $j \geq k+4$, while the second sum consists of terms with coefficients divisible by η^2. Combining this with (4.8) and (4.9) shows that p_{k+1} satisfies (4.5) with k replaced by $k+1$.

It follows by induction that p_k has the desired form for every k. Evaluating (4.5) at $k = n-2$ gives (4.4). This proves the lemma. $\qquad \square$

We now use the Lemma 4.1 and an induction argument on the dimension of the filiform group to show how a horizontal line can be perturbed to reach a nearby point, without changing too much its length or direction.

Lemma 4.2. *For every $A_0 \in \mathbb{R}$, there are numbers*

$$\varepsilon = \varepsilon(A_0, n) > 0, \qquad K = K(A_0, n) > 0, \qquad N = N(n) \in \mathbb{N}$$

so that the following holds for $A \in (A_0 - \varepsilon, A_0 + \varepsilon)$ and $a_2, \ldots, a_n \in (-\varepsilon, \varepsilon)$.

There exist $\theta_1, \ldots, \theta_N \in \mathbb{R}$ which depend smoothly on $a_2, \ldots, a_n$ with $|\theta_i| \leq K \max_{j \geq 2} |a_j|$ for $1 \leq i \leq N$ and

$$\prod_{i=1}^N \frac{1}{N}(E + \theta_i E') = E + a_2 E' + a_3 X_3 + \cdots + a_n X_n. \tag{4.10}$$

Proof. We prove the lemma by induction on n. If $n = 2$ the result is clear as $\mathbb{E}_2$ is simply Euclidean space $\mathbb{R}^2$ as a Carnot group. Suppose the statement holds in $\mathbb{E}_{n-1}$; we will show that it also holds for $\mathbb{E}_n$.

By the induction hypothesis, there exist

$$\varepsilon = \varepsilon(A_0, n-1) > 0, \qquad K = K(A_0, n-1) > 0, \qquad N = N(n-1) \in \mathbb{N}$$

such that the following holds in $\mathbb{E}_{n-1}$. For any choice of $A \in (A_0 - \varepsilon, A_0 + \varepsilon)$ and $a_2, \ldots, a_{n-1} \in (-\varepsilon, \varepsilon)$, there exist $\theta_1, \ldots, \theta_N$ which depend smoothly on $a_2, \ldots, a_{n-1}$ with $|\theta_i| \leq K \max_{j \geq 2} |a_j|$ for $1 \leq i \leq N$ and

$$\prod_{i=1}^N \frac{1}{N}(E + \theta_i E') = E + a_2 E' + a_3 X_3 + \cdots + a_{n-1} X_{n-1} \qquad \text{in } \mathbb{E}_{n-1}. \tag{4.11}$$

We now lift the above equation to $\mathbb{E}_n$. In other words, we consider E and E' as elements of the Lie algebra of $\mathbb{E}_n$ in the natural way. All calculations when computing the product using the BCH formula (2.1) remain the same, except for $[X_{n-1}, X_1]$ which will be equal to X_n rather than 0. An easy calculation shows

$$[E + \theta_1 E', E + \theta_2 E'] = (1 + A^2)(\theta_2 - \theta_1) X_3. \tag{4.12}$$

From now on we work in $\mathbb{E}_n$. By the BCH formula (2.1), (4.11) and (4.12):

$$p_1 := \prod_{i=1}^{N} \frac{1}{N}(E + \theta_i E')$$

$$= E + a_2 E' + a_3 X_3 + \cdots + a_{n-1} X_{n-1} + L(\theta_1, \ldots, \theta_N) X_n$$

$$= (1 + Aa_2) X_1 + (A - a_2) X_2 + a_3 X_3 + \cdots + a_{n-1} X_{n-1} + L(\theta_1, \ldots, \theta_N) X_n,$$

where $L(\theta_1, \ldots, \theta_N)$ is a polynomial in $\theta_1, \ldots, \theta_N$ (with coefficients depending on A) with no constant term due to (4.12). By using Lemma 4.1, for any $\eta \in \mathbb{R}$ we can choose $\eta_i \in \{\pm\eta\}$ for $1 \le i \le 2^{n-2}$ with sign depending on i but not η such that

$$p_2 := \prod_{i=1}^{2^{n-2}} \frac{1}{2^{n-2}}(E + \eta_i E')$$

$$= E + C(A^2 + 1)\eta X_n + \sum_{i=3}^{n} P_i(\eta) X_i$$

$$= X_1 + A X_2 + C(A^2 + 1)\eta X_n + \sum_{i=3}^{n} P_i(\eta) X_i,$$

where $P_i(x)$ are polynomials divisible by x^2 and $C = C_n \ne 0$.

We now analyze $p_1 p_2$. First notice that

$$p_1 + p_2 = (2 + Aa_2) X_1 + (2A - a_2) X_2 + a_3 X_3 + \ldots + a_{n-1} X_{n-1}$$

$$+ (L(\theta_1, \ldots, \theta_N) + C(A^2 + 1)\eta) X_n + \sum_{i=3}^{n} P_i(\eta) X_i$$

and

$$[p_1, p_2] = -(A^2 + 1) a_2 X_3 + a_3 X_4 + \ldots + a_{n-1} X_n$$

$$- \sum_{i=4}^{n} (1 + Aa_2) P_{i-1}(\eta) X_i.$$

By using the BCH formula (2.1), the coefficients of $E, E', X_3, \ldots, X_n$ in $p_1 p_2$ are given by

E	2
E'	a_2
X_3	$a_3 + F_3(a_2) + Q_3(\eta)$
$\ldots$	$\ldots$
X_i	$a_i + F_i(a_2, \ldots, a_{i-1}) + Q_i(\eta, a_2)$
$\ldots$	$\ldots$
X_n	$L(\theta_1, \ldots, \theta_N) + C(A^2 + 1)\eta + F_n(a_2, \ldots, a_{n-1}) + Q_n(\eta, a_2),$

where, for $i = 3, \ldots, n$,

- $F_i(a_2, \ldots, a_{i-1})$ is a polynomial with no constant term whose coefficients depend smoothly on A,
- $Q_i(\eta, a_2)$ is a polynomial divisible by η^2 whose coefficients depend smoothly on A.

Define $\Phi_A \colon (-\varepsilon, \varepsilon)^n \to \mathbb{R}^{n-1}$ to be the function of $a_2, \ldots, a_{n-1}, \eta$ whose coordinates are given by the coefficients of $E', X_3, X_4, \ldots, X_n$ in the above table. Recall that $\varepsilon = \varepsilon(A_0, n-1)$ was chosen using the induction hypothesis, which implies $\theta_1, \ldots, \theta_N$ depend smoothly on $a_2, \ldots, a_{n-1}$ whenever $(a_2, \ldots, a_{n-1}, \eta) \in (-\varepsilon, \varepsilon)^n$. Notice that $\Phi_A(0) = 0$ and the equality

$$\delta_{1/2}(p_1 p_2) = E + b_2 E' + b_3 X_3 + \ldots + b_n X_n \tag{4.13}$$

is equivalent to

$$\Phi_A(a_2, \ldots, a_{n-1}, \eta) = (2b_2, 2^2 b_3, \ldots, 2^i b_i, \ldots, 2^n b_n). \tag{4.14}$$

Claim 4.3. *There exists $\tilde{\varepsilon} = \tilde{\varepsilon}(A_0, n) > 0$ and $\tilde{K} = \tilde{K}(A_0, n) > 0$ such that for all $A \in (A_0 - \tilde{\varepsilon}, A_0 + \tilde{\varepsilon})$ and $b_2, \ldots, b_n \in (-\tilde{\varepsilon}, \tilde{\varepsilon})$, the equation (4.14) can be solved for $a_2, \ldots, a_{n-1}, \eta$. One can choose the solutions so that:*

(1) *$a_2, \ldots, a_{n-1}, \eta$ depend smoothly on $b_2, \ldots, b_n$,*
(2) *$|a_i|, |\eta| \leq \tilde{K} \max_{j \geq 1} |b_j|$.*

Proof of Claim. Notice first that $(\partial Q_i / \partial \eta)(0) = 0$ for each i, since Q_i is divisible by η^2. Hence $T_A := \Phi_A'(0)$ is a lower triangular matrix with determinant $C(A^2 + 1) \neq 0$. By the inverse function theorem, Φ_A is invertible with a C^1 inverse in a neighborhood of 0. In other words for each A and n, given $b_2, \ldots b_n$ sufficiently small, there exist $a_2, \ldots, a_{n-1}, \eta$ which depend smoothly on $b_2, \ldots, b_n$ such that (4.14) holds. We must show that one can use a uniform neighborhood for all $A \in (A_0 - \tilde{\varepsilon}, A_0 + \tilde{\varepsilon})$, where $\tilde{\varepsilon} > 0$ is sufficiently small and depends on A_0 and n. To establish such a neighborhood and the desired bounds on $|a_i|, |\eta|$, we briefly study the proof of the inverse function theorem from [33, 9.2.4 Theorem].

Define $\lambda(A) > 0$ by $1/\lambda(A) = 2\|T_A^{-1}\|_{\mathrm{op}}$. As the determinant of T_A is $C(A^2 + 1)$ and the entries of the adjoint of T_A are linear combinations of products of entries of T_A, there exists $\lambda > 0$ depending on A_0 and n such that $\lambda(A) > \lambda$ whenever $A \in (A_0 - \varepsilon, A_0 + \varepsilon)$. By the mean value theorem, there is $0 < \tilde{\varepsilon} < \varepsilon$ depending on A_0 and n such that $\|\Phi_A'(x) - T_A\|_{\mathrm{op}} < \lambda$ whenever $\|x\| < \tilde{\varepsilon}$ and $A \in (A_0 - \tilde{\varepsilon}, A_0 + \tilde{\varepsilon}) =: I$. It follows from [33] that for $A \in I$, the restricted map $\Phi_A \colon B(0, \tilde{\varepsilon}) \to \Phi_A(B(0, \tilde{\varepsilon}))$ is bijective with C^1 inverse. Since the entries of $(\Phi_A^{-1})'(x)$ are bounded for every $A \in I$ and every $x \in B(0, \tilde{\varepsilon})$, it follows that Φ_A is bi-Lipschitz for $A \in I$, with bi-Lipschitz constants depending on A_0. Hence $\Phi_A(B(0, \tilde{\varepsilon}))$ contains a ball $B(0, \tilde{\varepsilon}/\tilde{K})$, where $\tilde{K} \geq 1$ is some constant depending on n and A_0. Replacing $\tilde{\varepsilon}$ by a slightly smaller constant (still depending only on A_0 and n) gives the desired neighborhood for the inversion.

Next, since Φ_A^{-1} is Lipschitz with some Lipschitz constant $\tilde{K}$ depending on A_0 and n we have

$$|\Phi_A^{-1}(2b_2, \ldots, 2^n b_n)| \leq \tilde{K} |(2b_2, \ldots, 2^n b_n)|,$$

which yields

$$|(a_2, \ldots, a_{n-1}, \eta)| \leq 2^n \tilde{K} |(b_2, \ldots, b_n)|.$$

This gives the desired bounds on $|a_i|$ and $|\eta|$. $\qquad\square$

Since θ_i and η_i depend smoothly on $a_2, \ldots, a_{n-1}$ and η, Claim 4.3(1) ensures that θ_i and η_i depend smoothly on $b_2, \ldots, b_n$. Since $|\theta_i| \leq K \max_{j \geq 2} |a_j|$, Claim 4.3(2) implies that $|\theta_i|, |\eta| \leq K \tilde{K} \max_{j \geq 1} |b_j|$. To conclude, since (4.13) and (4.14) are equivalent, it suffices to check that $\delta_{1/2}(p_1 p_2)$ is of the form of the left hand side

of (4.10). Indeed, since the BCH formula (2.1) implies $XX = 2X$ for any element X of the Lie algebra, we have

$$p_1 p_2 = \prod_{i=1}^{N} \frac{1}{N}(E + \theta_i E') \prod_{i=1}^{2^{n-2}} \frac{1}{2^{n-2}}(E + \eta_i E')$$

$$= \prod_{i=1}^{N2^{n-2}} \frac{1}{N2^{n-2}}(E + \tilde{\theta}_i E') \prod_{i=1}^{N2^{n-2}} \frac{1}{N2^{n-2}}(E + \tilde{\eta}_i E'),$$

where the sequence $\tilde{\theta}_i$ repeats each term of θ_i for 2^{n-2} times and the sequence $\tilde{\eta}_i$ repeats each term of η_i for N times. Hence we can write

$$\delta_{1/2}(p_1 p_2) = \prod_{i=1}^{N2^{n-1}} \frac{1}{N2^{n-1}}(E + \Theta_i E'), \qquad (4.15)$$

where the terms of the sequence Θ_i consist of the terms of $\tilde{\theta}_i$ followed by the terms of $\tilde{\eta}_i$. Notice that (4.15) is of the form given in (4.10). This completes the proof. $\square$

4.2. Deformability in filiform groups.

We now use the horizontal curves built in Lemma 4.2 to show that every horizontal direction in $\mathbb{E}_n$, except possibly $\pm X_2$, is deformable. Recall that by Corollary 3.4, the directions $\pm X_2$ are deformable if and only if $n = 2$ or $n = 3$.

Theorem 4.4. *For $n \geq 3$, every $E \in \mathcal{E}_n$ with $\omega(E) = 1$, except for possibly $\pm X_2$, is deformable. Moreover, the parameters C_E, N_E, Δ_E related to the deformability of E can be chosen so that any $\tilde{E}$ sufficiently close to E is also deformable with the same parameters*

$$C_{\tilde{E}} = C_E, \qquad N_{\tilde{E}} = N_E, \qquad \Delta_{\tilde{E}}(\eta) = \Delta_E(\eta).$$

Proof. Fix $E \in \mathcal{E}_n$ with $\omega(E) = 1$ and $E \neq \pm X_2$. We need to show that for some C_E, N_E and $\Delta_E \colon (0, \infty) \to (0, \infty)$ the following holds. Given $0 < s < 1$, $\eta \in (0, \infty)$, $u \in \mathbb{E}_n$ with $d(u) \leq 1$ and $0 < \Delta < \Delta_E(\eta)$, there is a Lipschitz horizontal curve $g \colon \mathbb{R} \to \mathbb{E}_n$, formed by joining N_E horizontal lines, such that

(1) $g(t) = \exp(tE)$ for $|t| \geq s$,
(2) $g(\zeta) = \delta_{\Delta s}(u)$, where $\zeta := \langle \delta_{\Delta s}(u), E(0) \rangle$,
(3) $\text{Lip}(g) \leq 1 + \eta \Delta$,
(4) $|(p \circ g)'(t) - p(E)| \leq C_E \Delta$ for all but finitely many $t \in \mathbb{R}$.

Moreover, the same parameters C_E, N_E, Δ_E should work for any direction sufficiently close to E.

Notice that for $|t| \geq s$ the curve is explicitly defined by (1) and satisfies (3) and (4). Hence our task is to extend $g(t)$ for $-s < t < \zeta$ and $\zeta < t < s$. Since the two cases are similar, we show how to handle $-s < t < \zeta$. Up to left translations and reparameterizations of the curve, it suffices to verify the following equivalent claim.

Claim 4.5. *There exist C_E, N_E and $\Delta_E \colon (0, \infty) \to (0, \infty)$ such that the following holds.*

Given any $0 < s < 1$, $\eta > 0$, $u \in \mathbb{E}_n$ with $d(u) \leq 1$ and $0 < \Delta < \Delta_E(\eta)$, there is a Lipschitz horizontal curve $\varphi \colon [0, s+\zeta] \to \mathbb{E}_n$, where $\zeta := \langle \delta_{\Delta s}(u), E(0) \rangle$, formed by joining at most N_E horizontal lines, such that $\varphi(0) = 0$, $\varphi(s+\zeta) = \exp(sE)\delta_{\Delta s}(u)$, and

(A) $\text{Lip}(\varphi) \leq 1 + \eta \Delta$,

(B) $|(p \circ \varphi)'(t) - p(E)| \leq C_E \Delta$ *for all but finitely many* $t \in \mathbb{R}$.

Moreover, the same parameters $C_E, N_E, \Delta_E(\eta)$ *work for any direction sufficiently close to* E.

Proof of Claim. Since $E \neq \pm X_2$, we can write $E = aX_1 + bX_2$, where $a^2 + b^2 = 1$ and $a \neq 0$. Without loss of generality, up to changing the direction of the curve, we can assume $a > 0$. Let $u := u_1 X_1 + u_2 X_2 + \cdots + u_n X_n$. Identifying $\mathbb{E}_n$ and $\mathcal{E}_n$, since $E \in V_1$ we have

$$(sE)\delta_{\Delta s}(u) = \delta_s(E\delta_\Delta(u)),$$

and a simple computation gives

$$E + \delta_\Delta(u) = (a + \Delta u_1)X_1 + (b + \Delta u_2)X_2 + \Delta^2 u_3 X_3 + \cdots + \Delta^{n-1} u_n X_n,$$

$$[E, \delta_\Delta(u)] = \Delta(bu_1 - au_2)X_3 - a\Delta^2 u_3 X_4 - \cdots - a\Delta^{n-2} u_{n-1} X_n.$$

By the BCH formula (2.1), it is then clear that $E\delta_\Delta(u)$ has the form

$$E\delta_\Delta(u) = (a + \Delta u_1)X_1 + (b + \Delta u_2)X_2 + \sum_{i=3}^{n} \eta_i X_i,$$

where η_i satisfy $|\eta_i| \leq \tilde{Q}\Delta$ for a constant $\tilde{Q}$ depending only on n. Next we write

$$E\delta_\Delta(u) = \delta_{a + \Delta u_1}\left(X_1 + \frac{b + \Delta u_2}{a + \Delta u_1}X_2 + \sum_{i=3}^{n} \frac{\eta_i}{(a + \Delta u_1)^{i-1}}X_i\right),$$

and we define

$$A := \frac{b + \Delta u_2}{a + \Delta u_1}, \qquad a_2 := 0, \qquad a_i := \frac{\eta_i}{(a + \Delta u_1)^{i-1}} \text{ for } 3 \leq i \leq n.$$

Let $A_0 := b/a$ and let $\varepsilon = \varepsilon(A_0, n)$, $K = K(A_0, n)$ and $N = N(n)$ be defined according to Lemma 4.2. By making Δ sufficiently small, we can ensure that $A \in (A_0 - \varepsilon, A_0 + \varepsilon)$ and $|a_j| < \varepsilon$. If we consider a direction $\tilde{E} = \tilde{a}X_1 + \tilde{b}X_2$ with $\tilde{a}$ and $\tilde{b}$ sufficiently close to a and b (bound depending on a, b, ε), then we can ensure that if

$$\tilde{A} := \frac{\tilde{b} + \Delta u_2}{\tilde{a} + \Delta u_1}, \qquad \tilde{a}_2 := 0, \qquad \tilde{a}_i := \frac{\eta_i}{(a + \Delta u_1)^{i-1}} \text{ for } 3 \leq i \leq n,$$

then $\tilde{A} \in (A_0 - \varepsilon, A_0 + \varepsilon)$ and $|\tilde{a}_j| < \varepsilon$. Hence, in what follows, everything which applies to the direction E will also apply to every direction $\tilde{E}$ sufficiently close to E with the same parameters.

Applying Lemma 4.2 with $E_0 = X_1 + AX_2$ and $E_0' = AX_1 - X_2$ gives smooth functions $\theta_1, \ldots, \theta_N$ satisfying $|\theta_i| \leq K \max_{j \geq 2} |a_j|$ such that

$$\prod_{i=1}^{N} \frac{1}{N}(E_0 + \theta_i E_0') = E_0 + a_2 E_0' + a_3 X_3 + \cdots + a_n X_n. \qquad (4.16)$$

By definition of a_j, it follows that $|\theta_i| \leq Q\Delta$ for some constant Q depending on E, provided that Δ is small compared to a. Using the definitions of E_0, E_0' and

dilating both sides of (4.16) by $a + \Delta u_1$ gives

$$\prod_{i=1}^{N} \frac{1}{N} \Big((a + \Delta u_1)X_1 + (b + \Delta u_2)X_2 + \theta_i((b + \Delta u_2)X_1 - (a + \Delta u_1)X_2) \Big)$$
$$= (a + \Delta u_1)X_1 + (b + \Delta u_2)X_2 + \sum_{i=3}^{n} \eta_i X_i$$
$$= E\delta_\Delta(u).$$

Then, dilating both sides by s and recalling that

$$\zeta = \langle \delta_{\Delta s}(u), E(0) \rangle = \Delta s(au_1 + bu_2),$$

we get

$$\prod_{i=1}^{N} \frac{s + \zeta}{N} \frac{s}{s + \zeta} \Big((a + \Delta u_1)X_1 + (b + \Delta u_2)X_2 + \theta_i((b + \Delta u_2)X_1 - (a + \Delta u_1)X_2) \Big)$$
$$= (sE)\delta_{\Delta s}(u).$$

Define the horizontal curve $\varphi \colon [0, s + \zeta] \to \mathbb{E}_n$ by $\varphi(0) = 0$ and

$$\varphi'(t) = \frac{s}{s + \zeta} \Big((a + \Delta u_1)X_1 + (b + \Delta u_2)X_2 + \theta_i((b + \Delta u_2)X_1 - (a + \Delta u_1)X_2) \Big)$$

whenever

$$t \in I_i := \left(\frac{(i-1)(s + \zeta)}{N}, \frac{i(s + \zeta)}{N} \right), \qquad 1 \le i \le N.$$

Then $\varphi(0) = 0, \varphi(s + \zeta) = (sE)\delta_{\Delta s}(u)$ and φ is a Lipschitz horizontal curve formed from joining N horizontal lines. It remains to check that conditions (A) and (B) hold.

To verify (A), notice that by Lemma 2.18 it suffices to bound $|(p \circ \varphi)'|$. Recall that $|\theta_i| \le Q\Delta$ for $1 \le i \le N$ and $(1 + x)^{1/2} \le 1 + x/2$ for $x \ge -1$. For any $t \in I_i$ and any sufficiently small Δ one has

$$|(p \circ \varphi)'| = \frac{s}{s + \zeta} \left| \Big((a + \Delta u_1) + \theta_i(b + \Delta u_2), (b + \Delta u_2) - \theta_i(a + \Delta u_1) \Big) \right|$$
$$= \frac{1}{1 + \Delta(au_1 + bu_2)} \Big(1 + 2\Delta(au_1 + bu_2) + \Delta^2(u_1^2 + u_2^2) \Big)^{1/2} \Big(1 + \theta_i^2 \Big)^{1/2}$$
$$\le \frac{1}{1 + \Delta(au_1 + bu_2)} \Big(1 + \Delta(au_1 + bu_2) + \Delta^2(u_1^2 + u_2^2)/2 \Big) \Big(1 + \theta_i^2/2 \Big)$$
$$\le 1 + \eta\Delta,$$

where in the last inequality we used that $|\theta_i| \le Q\Delta$ and made Δ sufficiently small relative to η. This proves (A).

To verify (B), we estimate as follows. For any $t \in I_i$ and any sufficiently small Δ one has

$$
\begin{aligned}
|(p \circ \varphi)'(t) - p(E)| &\leq C\Delta + \left| \frac{s}{s+\zeta} - 1 \right| \\
&= C\Delta + \left| \frac{\zeta}{s+\zeta} \right| \\
&= C\Delta + \left| \frac{\Delta(au_1 + bu_2)}{1 + \Delta(au_1 + bu_2)} \right| \\
&\leq C\Delta.
\end{aligned}
$$

This shows that (B) holds. Since our conclusions also hold for any direction sufficiently close to E, the claim is proved. $\qquad\square$

As described earlier, Claim 4.5 suffices to prove the theorem. $\qquad\square$

Since deformability implies differentiability of the CC distance by Proposition 3.1, combining Proposition 3.3 and Theorem 4.4 shows that a horizontal direction in a model filiform group $\mathbb{E}_n$ for $n \geq 4$ is deformable if and only if it is not $\pm X_2$.

5. Distances between piecewise linear curves

In this section we prove a simple estimate for the distance between piecewise linear curves with similar directions in a general Carnot group $\mathbb{G}$. This will be needed to prove 'almost maximality' implies differentiability. We first recall the following useful lemma [20, Lemma 3.7].

Lemma 5.1. *Let $\mathbb{G}$ be a Carnot group of step s. Given any $\nu > 0$, there exists a constant $K_\nu > 0$ with the following property. If $N \in \mathbb{N}$ and $A_j, B_j \in \mathbb{G}$ defined for $j = 1, \ldots, N$ satisfy $d(B_j B_{j+1} \cdots B_N) \leq \nu$ and $d(A_j, B_j) \leq \nu$ for $j = 1, \ldots, N$, then it holds that*

$$
d(A_1 A_2 \cdots A_N, B_1 B_2 \cdots B_N) \leq K_\nu \sum_{j=1}^{N} d(A_j, B_j)^{1/s}.
$$

Our estimate for the distance between curves is given by the following lemma.

Lemma 5.2. *Let $\mathbb{G}$ be a Carnot group of step s. Then there is a constant $C_a \geq 1$ depending on $\mathbb{G}$ for which the following is true.*

Suppose $E \in V_1$ with $\omega(E) \leq 1$, $0 \leq D \leq 1$ and $N \in \mathbb{N}$. Let $g \colon (-R, R) \to \mathbb{G}$ be a Lipschitz horizontal curve with $g(0) = 0$ satisfying the following conditions:

(1) g is formed by joining of at most N horizontal lines,
(2) $|(p \circ g)'(t) - p(E)| \leq D$ whenever $(p \circ g)'(t)$ exists.

Then $d(g(t), \exp(tE)) \leq C_a N D^{1/s^2} |t|$ for every $t \in (-R, R)$.

Proof. Fix $t \geq 0$ without loss of generality. We may write

$$
g(t) = \exp(t_1 E_1) \exp(t_2 E_2) \cdots \exp(t_N E_N),
$$

where $t = t_1 + t_2 + \ldots + t_N$ with $t_i \geq 0$ and $E_j \in V_1$ with $|p(E_j) - p(E)| \leq D$ for $1 \leq j \leq N$. We intend to apply Lemma 5.1 to estimate

$$
\frac{d(g(t), \exp(tE))}{|t|} = d\left(\exp\left(\frac{t_1 E_1}{t}\right) \cdots \exp\left(\frac{t_N E_N}{t}\right), \exp\left(\frac{t_1 E}{t}\right) \cdots \exp\left(\frac{t_N E}{t}\right) \right).
$$

We first check that the hypotheses of Lemma 5.1 hold with the choice

$$A_j := \exp((t_j/t)E_j), \qquad B_j := \exp((t_j/t)E), \qquad \nu := 3.$$

We first notice

$$d(B_j B_{j+1} \cdots B_N) = d\left(\exp\left(\frac{(t_j + \ldots + t_N)E}{t} \right) \right)$$
$$= \frac{(t_j + \ldots + t_N)}{t} d(\exp(E))$$
$$\leq 1.$$

Second, we can estimate as follows

$$d(A_j, B_j) = (t_j/t)d(\exp(E_j), \exp(E))$$
$$\leq d(\exp(E_j)) + d(\exp(E))$$
$$\leq 3.$$

We can then combine Lemma 5.1 with Proposition 2.4 to get

$$\frac{d(g(t), \exp(tE))}{|t|} \leq K_3 \sum_{j=1}^{N} d(\exp((t_j/t)E_j), \exp((t_j/t)E))^{1/s}$$
$$\leq K_3 \sum_{j=1}^{N} d(\exp(E_j), \exp(E))^{1/s}$$
$$\leq C \sum_{j=1}^{N} |\exp(E_j) - \exp(E)|^{1/s^2}$$
$$\leq C \sum_{j=1}^{N} |p(E_j) - p(E)|^{1/s^2}$$
$$\leq C N D^{1/s^2}.$$

The proof is complete noticing that $C \geq 1$ is a constant depending only on $\mathbb{G}$. $\square$

6. Almost maximal directional derivatives and the UDS

In this section we fix a Carnot group $\mathbb{G}$ satisfying the following condition.

Assumptions 6.1. Assume $\mathbb{G} \neq \mathbb{R}$. We say that $\mathbb{G}$ admits a ball of uniformly deformable directions if there exists an open ball $B \subset V_1$ of directions such that every $E \in B$ is deformable with the same parameters C_B, N_B and Δ_B.

We will show that every Carnot group $\mathbb{G}$ which admits a ball of uniformly deformable directions contains a CC-Hausdorff dimension one (hence measure zero) set N so that almost maximality of a directional derivative $Ef(x)$ implies differentiability if $x \in N$ and $E \in B$. Combining this with Proposition 7.1 will lead to a proof of Theorem 1.1. By Theorem 4.4, all model filiform groups $\mathbb{E}_n$ with $n \geq 2$ admit a ball of uniformly deformable directions. In particular, Theorem 1.1 applies to Carnot groups of arbitrarily high step.

The Carnot group $\mathbb{G}$ will be identified with $\mathbb{R}^n$ by means of exponential coordinates of the first kind. Let $B_{\mathbb{Q}}$ denote the set of $E \in B$ with $\omega(E) = 1$ which are a rational linear combination of the basis vectors $X_1, \ldots, X_r$ of V_1. Note that $B_{\mathbb{Q}}$ is dense in B since the Euclidean sphere contains a dense set of points with rational

coordinates. The construction of our universal differentiability set is given by the following lemma.

Lemma 6.2. *For each choice of $E \in B_{\mathbb{Q}}$, $u \in \mathbb{Q}^n$ with $d(u) \leq 1$ and rationals*

$$0 < s < 1, \qquad \eta > 0, \qquad 0 < \Delta < \Delta_B(\eta),$$

let $\gamma_{E,u,s,\Delta,\eta}$ denote a curve granted by Definition 2.15 with parameters C_B, N_B, $\Delta_B(\eta)$.

Let L be the countable union of images of all translated curves $x\gamma_{E,u,s,\Delta,\eta}$, where $x \in \mathbb{Q}^n$ and E, u, s, Δ, η are as above. Then there is a G_δ set $N \subset \mathbb{G}$ containing L which has Hausdorff dimension one with respect to the CC metric.

The proof of Lemma 6.2 is essentially the same as that of [17, Lemma 5.4]. We also recall the following mean value type lemma for future use [30, Lemma 3.4].

Lemma 6.3. *Suppose $|\zeta| < s < \rho$, $0 < v < 1/32$, $\sigma > 0$ and $L > 0$ are real numbers and let $\varphi, \psi \colon \mathbb{R} \to \mathbb{R}$ be Lipschitz maps satisfying $\mathrm{Lip}_{\mathbb{E}}(\varphi) + \mathrm{Lip}_{\mathbb{E}}(\psi) \leq L$, $\varphi(t) = \psi(t)$ 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

$$\sigma \leq v^3 \Big(\frac{\varphi(\zeta) - \psi(\zeta)}{sL} \Big)^2.$$

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 6.4. By examining the proof of Lemma 6.3 in [30]. one can see that τ can additionally be chosen outside a given Lebesgue measure zero subset of $\mathbb{R}$.

From now on we fix a set $N \subset \mathbb{G}$ as given by Lemma 6.2.

Notation 6.5. For any Lipschitz function $f : \mathbb{G} \to \mathbb{R}$, define:

$$D^f := \{(x, E) \in N \times V_1 : \omega(E) = 1, \, Ef(x) \text{ exists}\}.$$

Theorem 6.6. *Let $\mathbb{G}$ be a Carnot group of step s which admit a ball of uniformly deformable directions (Assumptions 6.1).*

Let $f \colon \mathbb{G} \to \mathbb{R}$ be Lipschitz with $\mathrm{Lip}_{\mathbb{G}}(f) \leq 1/2$ and suppose $(x_*, E_*) \in D^f$ with $E_* \in B$. Let M denote the set of pairs $(x, E) \in D^f$ such that $Ef(x) \geq E_*f(x_*)$ and for every $t \in (-1, 1)$:

$$|(f(x\exp(tE_*)) - f(x)) - (f(x_*\exp(tE_*)) - f(x_*))|$$

$$\leq 6|t|((Ef(x) - E_*f(x_*))\mathrm{Lip}_{\mathbb{G}}(f))^{1/2s^2}.$$

If

$$\limsup_{\delta \downarrow 0}\{Ef(x) \colon (x, E) \in M \text{ and } d(x, x_*) \leq \delta\} \leq E_*f(x_*),$$

then f is differentiable at x_* and its Pansu differential is given by

$$L(x) := E_*f(x_*)\langle x, E_*(0)\rangle = E_*f(x_*)\langle p(x), p(E_*)\rangle.$$

Proof. We assume $\mathrm{Lip}_{\mathbb{G}}(f) > 0$, since otherwise the statement is trivial. Fix the following parameters:

(1) $\varepsilon > 0$ rational,
(2) $0 < v < 1/32$ rational such that $4(1 + 20v)\sqrt{(2+v)/(1-v)} + v < 6$,
(3) $\eta = \varepsilon v^3/3200$,
(4) $\Delta_B(\eta/2)$, C_B and C_a according to Lemma 5.2 and Assumptions 6.1,
(5) rational $0 < \Delta < \min\{\eta v^2, \Delta_B(\eta/2), \Upsilon\}$, where

$$\Upsilon := \frac{\varepsilon v^{2s^2+1}}{8C_B^2 C_a^{2s^2} N_B^{2s^2} \mathrm{Lip}_{\mathbb{G}}(f)^{2s^2-1}},$$

(6) $\sigma = 9\varepsilon^2 v^5 \Delta^2/256$,
(7) $0 < \rho < 1$ such that

$$|f(x_* \exp(tE_*)) - f(x_*) - tE_* f(x_*)| \le \sigma \mathrm{Lip}_{\mathbb{G}}(f)|t| \quad \text{for } |t| \le \rho, \qquad (6.1)$$

(8) $0 < \delta < \rho\sqrt{3\varepsilon v \Delta^3}/4$ such that

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

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

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

$$|f(x_* \delta_t(h)) - f(x_*) - tE_* f(x_*)\langle u, E_*(0)\rangle| \le \varepsilon t \qquad \text{for } d(u) \le 1, \ 0 < t < \delta.$$

Suppose this is not true. Then there exist $u \in \mathbb{Q}^n$ with $d(u) \le 1$ and rational $0 < r < \delta$ such that

$$|f(x_* \delta_r(u)) - f(x_*) - rE_* f(x_*)\langle u, E_*(0)\rangle| > \varepsilon r. \qquad (6.2)$$

Let $\mathrm{s} = r/\Delta \in \mathbb{Q}$. To contradict (6.2), we first construct Lipschitz horizontal curves g and h for which we can apply Lemma 6.3 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_*. Let

$$A_1 := \left(\frac{\eta \Delta}{C_a(N_B + 2)}\right)^{s^2} \qquad (6.3)$$

and

$$A_2 := \left(6 - \left(4(1 + 20v)\left(\frac{2+v}{1-v}\right)^{1/2} + v\right)\right)^{s^2} \frac{(\varepsilon v \Delta \mathrm{Lip}_{\mathbb{G}}(f)/2)^{1/2}}{C_a^{s^2}(N_B + 2)^{s^2} \mathrm{Lip}_{\mathbb{G}}(f)^{s^2}}. \qquad (6.4)$$

Notice that $A_1, A_2 > 0$. We choose $\tilde{x}_* \in \mathbb{Q}^n$ and $\tilde{E}_* \in B_{\mathbb{Q}}$ sufficiently close to x_* and E_* respectively to ensure:

$$d(\tilde{x}_* \delta_r(u), x_*) \le 2r, \qquad (6.5)$$

$$d(\tilde{x}_* \delta_r(u), x_* \delta_r(u)) \le \sigma r, \qquad (6.6)$$

$$\omega(\tilde{E}_* - E_*) \le \min\{\sigma, C_B \Delta, A_1, A_2\}. \qquad (6.7)$$

Recall that $0 < r < \Delta$ and $\mathrm{s} = r/\Delta$ are rational and that $0 < \mathrm{s} < 1$. To construct g we first apply Definition 2.15 with $E = \tilde{E}_*$ and parameters $\eta, \mathrm{s}, \Delta, \delta_r(u)$ and u as defined above in (6.2). We then left translate this curve by $\tilde{x}_*$. This gives a Lipschitz horizontal curve $g: \mathbb{R} \to \mathbb{G}$ which is formed by joining at most N_B horizontal lines such that

- $g(t) = \tilde{x}_* \exp(t\tilde{E}_*)$ for $|t| \ge \mathrm{s}$,

- $g(\zeta) = \tilde{x}_* \delta_r(u)$, where $\zeta := r\langle u, \tilde{E}_*(0)\rangle$,
- $\mathrm{Lip}_{\mathbb{G}}(g) \leq 1 + \eta\Delta$,
- for all but finitely many $t \in \mathbb{R}$, $g'(t)$ exists and $|(p \circ g)'(t) - p(\tilde{E}_*)| \leq C_B\Delta$ for $t \in \mathbb{R}$.

Since all the relevant quantities are chosen to be rational and N is built according to Lemma 6.2, it follows that the image of g is contained in N.

Construction of h.

Claim 6.7. *There exists a Lipschitz horizontal curve $h \colon \mathbb{R} \to \mathbb{G}$ such that*

$$h(t) = \begin{cases} \tilde{x}_* \exp(t\tilde{E}_*) & \text{if } |t| \geq \mathrm{s}, \\ x_* \exp(tE_*) & \text{if } |t| \leq \mathrm{s}/2, \text{and} \end{cases}$$

in each of the regions $(\mathrm{s}/2, \mathrm{s})$ and $(-\mathrm{s}, -\mathrm{s}/2)$, h is formed by joining at most N_B horizontal lines. Moreover

- $\mathrm{Lip}_{\mathbb{G}}(h) \leq 1 + \eta\Delta/2$,
- *for all but finitely many $t \in \mathbb{R}$, $h'(t)$ and satisfies the bound $|(p \circ h)'(t) - p(E_*)| \leq \min\{A_1, A_2\}$.*

Proof of Claim. Up to a left translation, we may start by assuming that $x_* = 0$. Clearly $h(t)$ is defined explicitly and satisfies the required conditions for $|t| \leq \mathrm{s}/2$ and $|t| \geq \mathrm{s}$. We now show how to extend h in $(\mathrm{s}/2, \mathrm{s})$. The extension to $(-\mathrm{s}, -\mathrm{s}/2)$ is essentially the same.

Recall $\Delta_B(1)$ and C_B from Assumptions 6.1. Choose $0 < \Gamma < \Delta(1)$ satisfying

$$(1 + \Gamma)^2 \leq 1 + \eta\Delta/2$$

and

$$C_B\Gamma(1 + \Gamma) + \Gamma \leq \min\{A_1, A_2\}.$$

Define $\lambda = \mathrm{s}\Gamma/2 < \Gamma$. Choose $v \in \mathbb{G}$ with $d(v) \leq 1$ such that

$$\delta_\lambda(v) = \exp(-\mathrm{s}E_*)\tilde{x}_* \exp(\mathrm{s}\tilde{E}_*).$$

This is possible if the rational approximation introduced earlier is chosen correctly; note that the rational approximation was introduced after all quantities upon which λ depends. We now apply Remark 2.16 with

- $\eta = 1$ and Δ replaced by Γ,
- s replaced by $\mathrm{s}/2$,
- u replaced by v,
- ζ replaced by $\tilde{\zeta} := \langle \delta_{\mathrm{s}\Gamma/2}(v), E_*(0)\rangle$.

We obtain a Lipschitz horizontal curve $\varphi \colon [0, \mathrm{s}/2 + \tilde{\zeta}] \to \mathbb{G}$ that is formed by joining at most N_B horizontal lines such that

- $\varphi(0) = 0$,
- $\varphi(\mathrm{s}/2 + \tilde{\zeta}) = \exp(-(\mathrm{s}/2)E_*)\tilde{x}_* \exp(\mathrm{s}\tilde{E}_*)$,
- $\mathrm{Lip}_{\mathbb{G}}(\varphi) \leq 1 + \Gamma$,
- $\varphi'(t)$ exists and $|(p \circ \varphi)'(t) - p(E_*)| \leq C_B\Gamma$ for all except finitely many $t \in [0, \mathrm{s}/2 + \tilde{\zeta}]$.

Then $\tilde{\varphi} \colon [0, 1] \to \mathbb{G}$ defined by $\tilde{\varphi}(t) = \varphi((\mathrm{s}/2 + \tilde{\zeta})t)$ is a Lipschitz horizontal curve such that

- $\tilde{\varphi}(0) = 0$,
- $\tilde{\varphi}(1) = \exp(-(\mathrm{s}/2)E_*)\tilde{x}_* \exp(\mathrm{s}\tilde{E}_*)$,

- $\mathrm{Lip}_{\mathbb{G}}(\tilde{\varphi}) \leq (1 + \Gamma)(\mathrm{s}/2 + \tilde{\zeta})$,
- $\tilde{\varphi}'(t)$ exists and $|(p \circ \tilde{\varphi})'(t) - (\mathrm{s}/2 + \tilde{\zeta})p(E_*)| \leq C_B\Gamma(\mathrm{s}/2 + \tilde{\zeta})$ for all but finitely many $t \in [0, 1]$.

Define $h_1 : [\mathrm{s}/2, \mathrm{s}] \to \mathbb{G}$ by

$$h_1(t) = \exp((\mathrm{s}/2)E_*)\tilde{\varphi}((2/\mathrm{s})(t - \mathrm{s}/2)).$$

Then h_1 is a Lipschitz horizontal curve which satisfies $h_1(\mathrm{s}/2) = \exp((\mathrm{s}/2)E_*)$ and $h_1(\mathrm{s}) = \tilde{x}_* \exp(\mathrm{s}\tilde{E}_*)$. Note that $|p(v)| \leq d(v) \leq 1$ implies $|\tilde{\zeta}| \leq \lambda$. Hence we have

$$\begin{aligned}
\mathrm{Lip}_{\mathbb{G}}(h_1) &\leq \frac{2(1 + \Gamma)(\mathrm{s}/2 + \tilde{\zeta})}{\mathrm{s}} \\
&\leq \frac{2(1 + \Gamma)(\mathrm{s}/2 + \lambda)}{\mathrm{s}} \\
&\leq (1 + \Gamma)^2 \\
&\leq 1 + \eta\Delta/2.
\end{aligned}$$

Then, for all but finitely many $t \in [\mathrm{s}/2, \mathrm{s}]$

$$|(p \circ h_1)'(t) - (1 + 2\tilde{\zeta}/\mathrm{s})p(E_*)| \leq C_B\Gamma(1 + 2\tilde{\zeta}/\mathrm{s}),$$

and this implies

$$\begin{aligned}
|(p \circ h_1)'(t) - p(E_*)| &\leq C_B\Gamma(1 + 2\tilde{\zeta}/\mathrm{s}) + 2|\tilde{\zeta}|/\mathrm{s} \\
&\leq C_B\Gamma(1 + \Gamma) + \Gamma \\
&\leq \min\{A_1, A_2\}.
\end{aligned}$$

Defining $h(t) := h_1(t)$ for any $t \in [\mathrm{s}/2, \mathrm{s}]$ we obtain the desired properties. The extension of h in $[-\mathrm{s}, -\mathrm{s}/2]$ is similar. $\qquad\square$

Application of Lemma 6.3.

We now prove that the hypotheses of Lemma 6.3 hold with $L := (2 + \eta\Delta)\mathrm{Lip}_{\mathbb{G}}(f)$, $\varphi := f \circ g$ and $\psi := f \circ h$. The inequalities $|\zeta| < \mathrm{s} < \rho$, $0 < v < 1/32$ and the equality $\varphi(t) = \psi(t)$ for $|t| \geq \mathrm{s}$ are clear. Since $\mathrm{Lip}_{\mathbb{G}}(g), \mathrm{Lip}_{\mathbb{G}}(h) \leq 1 + \eta\Delta/2$, we have $\mathrm{Lip}_{\mathbb{E}}(\varphi) + \mathrm{Lip}_{\mathbb{E}}(\psi) \leq L$.

Notice that (6.6) implies

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

Since $|\zeta| \leq r \leq \rho$, we may evaluate (6.1) at $t = \zeta$ to obtain

$$\begin{aligned}
|f(x_* \exp(\zeta E_*)) - f(x_*) - \zeta E_* f(x_*)| &\leq \sigma\mathrm{Lip}_{\mathbb{G}}(f)|\zeta| \\
&\leq \sigma r\mathrm{Lip}_{\mathbb{G}}(f).
\end{aligned}$$

Next, note that (6.7) implies $|\tilde{E}_*(0) - E_*(0)| \leq \sigma$. Recalling that $\zeta = r\langle u, \tilde{E}_*(0)\rangle$ we can estimate as follows:

$$\begin{aligned}
|\zeta E_* f(x_*) - r\langle u, E_*(0)\rangle E_* f(x_*)| &= r|E_* f(x_*)||\langle u, \tilde{E}_*(0) - E_*(0)\rangle| \\
&\leq r\mathrm{Lip}_{\mathbb{G}}(f)|\tilde{E}_*(0) - E_*(0)| \\
&\leq \sigma r\mathrm{Lip}_{\mathbb{G}}(f).
\end{aligned}$$

Hence we obtain,

$$|f(x_* \exp(\zeta E_*)) - f(x_*) - r\langle u, E_*(0)\rangle E_* f(x_*)| \leq 2\sigma r\mathrm{Lip}_{\mathbb{G}}(f). \qquad (6.8)$$

Since $|\zeta| \leq r = \Delta s \leq s/2$ we have $h(\zeta) = x_* \exp(\zeta E_*)$. The definition of g gives $g(\zeta) = \tilde{x}_* \delta_r(u)$. Using also (6.2) and (6.8), we can estimate as follows:

$$
\begin{aligned}
|\varphi(\zeta) - \psi(\zeta)| &= |f(g(\zeta)) - f(h(\zeta))| \\
&= |f(\tilde{x}_* \delta_r(u)) - f(x_* \exp(\zeta E_*))| \\
&\geq |f(x_* \delta_r(u)) - f(x_* \exp(\zeta E_*))| - |f(\tilde{x}_* \delta_r(u)) - f(x_* \delta_r(u))| \\
&\geq |f(x_* \delta_r(u)) - f(x_*) - rE_* f(x_*) \langle u, E_*(0) \rangle| \\
&\quad - |f(x_* \exp(\zeta E_*)) - f(x_*) - rE_* f(x_*) \langle u, E_*(0) \rangle| \\
&\quad - \sigma r \mathrm{Lip}_{\mathbb{G}}(f) \\
&\geq \varepsilon r - 2\sigma r \mathrm{Lip}_{\mathbb{G}}(f) - \sigma r \mathrm{Lip}_{\mathbb{G}}(f) \\
&= \varepsilon r - 3\sigma r \mathrm{Lip}_{\mathbb{G}}(f) \\
&\geq 3\varepsilon r/4. \tag{6.9}
\end{aligned}
$$

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

The derivative $\psi'(0)$ exists and equals $E_* f(x_*)$, since $\psi(t) = f(x_* \exp(tE_*))$ for every $|t| \leq s/2$. We next check that

$$
|\psi(t) - \psi(0) - t\psi'(0)| \leq \sigma L|t| \quad \text{for } |t| \leq \rho. \tag{6.10}
$$

Recall that $h(0) = x_*$, $|(p \circ h)' - p(E_*)| \leq A_1$ (see (6.3) for the definition of A_1) and h is formed by joining at most $N_B + 2$ horizontal lines. Hence Lemma 5.2 implies that

$$
d(x_* \exp(tE_*), h(t)) \leq C_a(N_B + 2) A_1^{1/s^2} |t| \leq \eta \Delta |t| \quad \text{for every } t \in \mathbb{R}.
$$

Hence, using also (6.1) and $L = (2 + \eta\Delta)\mathrm{Lip}_{\mathbb{G}}(f)$, one has

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

Since $\mathrm{Lip}_{\mathbb{G}}(f) \leq 1/2$ we have $L \leq 4$. By using $r < \delta$, $s = r/\Delta$, (6.9) and the definition of r, δ, Δ and s 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 (6.9), $L \leq 4$ and the definition of σ to get

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

We can now apply Lemma 6.3. We obtain $\tau \in (-s, s) \setminus \{\zeta\}$ such that $\varphi'(\tau)$ exists and satisfies

$$
\varphi'(\tau) \geq \psi'(0) + v|\varphi(\zeta) - \psi(\zeta)|/s, \tag{6.11}
$$

and for every $t \in \mathbb{R}$:

$$|(\varphi(\tau + t) - \varphi(\tau)) - (\psi(t) - \psi(0))| \leq 4(1 + 20v)\sqrt{(\varphi'(\tau) - \psi'(0))L}|t| \qquad (6.12)$$

Since g is a horizontal curve, we may use Remark 6.4 to additionally choose τ such that $g'(\tau)$ exists and is in $\text{Span}\{X_i(g(\tau)) \colon 1 \leq i \leq r\}$.

Conclusion.

Let $x := g(\tau) \in N$ and choose $E \in V_1$ with $E(g(\tau)) = g'(\tau)/|p(g'(\tau))|$, which implies that $\omega(E) = 1$. From (6.11) and (6.12) we will obtain

$$Ef(x) \geq E_*f(x_*) + \varepsilon v\Delta/2, \qquad (6.13)$$

$$(x, E) \in M. \qquad (6.14)$$

We first observe that this suffices to conclude the proof. Indeed, by (6.5) and since $g(\zeta) = \tilde{x}_*\delta_r(u)$ one has

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

Since $x \in N$, combining this with (6.13) and (6.14) contradicts the choice of δ. This forces us to conclude that (6.2) is false, finishing the proof.

Proof of (6.13).

Using (6.9) and (6.11) we have that

$$\varphi'(\tau) - \psi'(0) \geq 3\varepsilon vr/4\mathsf{s} = 3\varepsilon v\Delta/4. \qquad (6.15)$$

Notice that, by the definition of E, by Definition 2.7, and the fact that g is a concatenation of horizontal lines, we have $\varphi'(\tau) = Ef(x)|p(g'(\tau))|$. Since $\omega(E) = 1$, we deduce that $|\varphi'(\tau)|/|p(g'(\tau))| \leq \text{Lip}_{\mathbb{G}}(f)$. Similarly $|p(g'(\tau))| \leq \text{Lip}_{\mathbb{G}}(g) \leq 1 + \eta\Delta$. Since $\psi'(0) = E_*f(x_*)$, by (6.15) we have

$$\begin{aligned}
Ef(x) &- E_*f(x_*) - (1 - v)(\varphi'(\tau) - \psi'(0)) \\
&= v(\varphi'(\tau) - \psi'(0)) + (1 - |p(g'(\tau))|)\varphi'(\tau)/|p(g'(\tau))| \\
&\geq 3\varepsilon v^2\Delta/4 - \eta\Delta\text{Lip}_{\mathbb{G}}(f) \\
&\geq 0.
\end{aligned}$$

In the last inequality we used $\text{Lip}_{\mathbb{G}}(f) \leq 1/2$ and $\eta \leq 3\varepsilon v^2/2$. From this we use $0 < v < 1/32$ and (6.15) again to deduce

$$Ef(x) - E_*f(x_*) \geq (1 - v)(\varphi'(\tau) - \psi'(0)) \geq \varepsilon v\Delta/2, \qquad (6.16)$$

which proves (6.13).

Proof of (6.14).

Recall that $|(p \circ g)'(t) - p(\tilde{E}_*)| \leq C_B\Delta$ for all but finitely many t. Using (6.7), this implies $|(p \circ g)'(t) - p(E_*)| \leq 2C_B\Delta$ for all but finitely many t. Since $x = g(\tau)$ and g is formed by joining at most N_B horizontal lines, we can apply Lemma 5.2 to obtain

$$d(g(\tau + t), x\exp(tE_*)) \leq C_a N_B(2C_B\Delta)^{1/s^2}|t| \qquad \text{for every } t \in \mathbb{R}.$$

By (6.16) we have $\Delta \leq 2(Ef(x) - E_*f(x_*))/(\varepsilon v)$. Combining this fact with the definition of Δ, we deduce that

$$
\begin{aligned}
&|(f(x\exp(tE_*)) - f(x)) - (f(g(\tau+t)) - f(g(\tau))))| \\
&\qquad = |f(x\exp(tE_*)) - f(g(\tau+t))| \\
&\qquad \leq \mathrm{Lip}_{\mathbb{G}}(f)d(g(\tau+t), x\exp(tE_*)) \\
&\qquad \leq C_a N_B (2C_B\Delta)^{1/s^2}\mathrm{Lip}_{\mathbb{G}}(f)|t| \\
&\qquad \leq C_a N_B (2C_B\sqrt{\Delta})^{1/s^2}\mathrm{Lip}_{\mathbb{G}}(f)|t|\Big(\frac{2(Ef(x) - E_*f(x_*))}{\varepsilon v}\Big)^{\frac{1}{2s^2}} \\
&\qquad \leq v|t|\big((Ef(x) - E_*f(x_*))\mathrm{Lip}_{\mathbb{G}}(f)\big)^{\frac{1}{2s^2}}\Big(\frac{8C_B^2 C_a^{2s^2} N_B^{2s^2}\Delta\mathrm{Lip}_{\mathbb{G}}(f)^{2s^2-1}}{\varepsilon v^{2s^2+1}}\Big)^{\frac{1}{2s^2}} \\
&\qquad \leq v|t|\big((Ef(x) - E_*f(x_*))\mathrm{Lip}_{\mathbb{G}}(f)\big)^{\frac{1}{2s^2}} \quad \text{for } t\in\mathbb{R}. \qquad (6.17)
\end{aligned}
$$

Combining (6.12), (6.16) and $L = (2+\eta\Delta)\mathrm{Lip}_{\mathbb{G}}(f) \leq (2+v)\mathrm{Lip}_{\mathbb{G}}(f)$ gives

$$
\begin{aligned}
&|(\varphi(\tau+t) - \varphi(\tau)) - (\psi(t) - \psi(0))| \\
&\qquad \leq 4(1+20v)|t|\Big(\frac{(2+v)\mathrm{Lip}_{\mathbb{G}}(f)(Ef(x) - E_*f(x_*))}{1-v}\Big)^{\frac{1}{2}} \quad \text{for } t\in\mathbb{R}. \quad (6.18)
\end{aligned}
$$

Since $\mathrm{Lip}_{\mathbb{G}}(f) \leq 1/2$, we easily get

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

since both sides are less than 1. Hence adding (6.17) and (6.18) and using the definition $\varphi = f\circ g$ gives for $t\in\mathbb{R}$:

$$
\begin{aligned}
&|f(x\exp(tE_*)) - f(x)) - (\psi(t) - \psi(0))| \\
&\qquad \leq \Big(4(1+20v)\Big(\frac{2+v}{1-v}\Big)^{\frac{1}{2}} + v\Big)|t|\big((Ef(x) - E_*f(x_*))\mathrm{Lip}_{\mathbb{G}}(f)\big)^{\frac{1}{2s^2}}. \quad (6.19)
\end{aligned}
$$

Recall that $\psi = f\circ h$ and that h is a concatenation of at most N_B+2 horizontal lines such that $h(0) = x_*$ and the inequality $|(p\circ h)' - p(E_*)| \leq A_2$ holds for all but finitely many $t\in\mathbb{R}$. Then, by Lemma 5.2, (6.4) and (6.13), we have

$$
\begin{aligned}
&|(\psi(t) - \psi(0)) - (f(x_*\exp(tE_*)) - f(x_*))| \\
&\qquad = |f(h(t)) - f(x_*\exp(tE_*))| \\
&\qquad \leq \mathrm{Lip}_{\mathbb{G}}(f)d(h(t), x_*\exp(tE_*)) \\
&\qquad \leq \mathrm{Lip}_{\mathbb{G}}(f)C_a(N_B+2)A_2^{1/s^2}|t| \\
&\qquad = \Big(6 - \Big(4(1+20v)\Big(\frac{2+v}{1-v}\Big)^{\frac{1}{2}} + v\Big)\Big)|t|(\varepsilon v\Delta/2)\mathrm{Lip}_{\mathbb{G}}(f))^{\frac{1}{2s^2}} \\
&\qquad \leq \Big(6 - \Big(4(1+20v)\Big(\frac{2+v}{1-v}\Big)^{\frac{1}{2}} + v\Big)\Big)|t|\big((Ef(x) - E_*f(x_*))\mathrm{Lip}_{\mathbb{G}}(f)\big)^{\frac{1}{2s^2}} \quad \text{for } t\in\mathbb{R}. \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (6.20)
\end{aligned}
$$

Adding (6.19) and (6.20) gives for every $t\in\mathbb{R}$

$$
\begin{aligned}
&|(f(x\exp(tE_*)) - f(x)) - (f(x_*\exp(tE_*)) - f(x_*))| \\
&\qquad \leq 6|t|\big((Ef(x) - E_*f(x_*))\mathrm{Lip}_{\mathbb{G}}(f)\big)^{\frac{1}{2s^2}}.
\end{aligned}
$$

This implies (6.14), hence proving the theorem. $\qquad\square$

7. Construction of an almost maximal directional derivative

Assume $\mathbb{G}$ is a Carnot group of step s, rank r and topological dimension n. Fix a G_δ set $N \subset \mathbb{G}$. The main result of this section is Proposition 7.1, which is an adaptation of [9, Theorem 3.1] and of [28, Theorem 6.1] to $\mathbb{G}$. It shows that given a Lipschitz function $f_0 \colon \mathbb{G} \to \mathbb{R}$, there is a Lipschitz function $f \colon \mathbb{G} \to \mathbb{R}$ such that $f - f_0$ is $\mathbb{G}$-linear and f has an almost locally maximal horizontal directional derivative at a point of N.

Recall Notation 6.5. Note that if $f - f_0$ is $\mathbb{G}$-linear then $D^f = D^{f_0}$ and also the functions f and f_0 have the same points of Pansu differentiability.

Proposition 7.1. *Suppose $f_0 : \mathbb{G} \to \mathbb{R}$ is a Lipschitz function, $(x_0, E_0) \in D^{f_0}$ and $\delta_0, \mu, \tau, K > 0$. Then there is a Lipschitz function $f : \mathbb{G} \to \mathbb{R}$ such that $f - f_0$ is $\mathbb{G}$-linear with $\mathrm{Lip}_{\mathbb{G}}(f - f_0) \leq \mu$, and a pair $(x_*, E_*) \in D^f$ with $d(x_*, x_0) \leq \delta_0$ and $\omega(E_* - E_0) \leq \tau$ 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) *$d(x, x_*) \leq \delta_\varepsilon$, $Ef(x) \geq E_* f(x_*)$ and*
(2) *for any $t \in (-1, 1)$*

$$|(f(x \exp(tE_*)) - f(x)) - (f(x_* \exp(tE_*)) - f(x_*))|$$
$$\leq K|t|(Ef(x) - E_* f(x_*))^{\frac{1}{2s^2}},$$

then

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

We use the remainder of this section to prove Proposition 7.1. We recall the following constants:

- $C_{\mathrm{a}} \geq 1$ chosen as in Lemma 5.2,
- $C_D \geq 1$ as in Proposition 2.5.

Fix $f_0, x_0, E_0, \delta_0, \tau, \mu, K$ as given in the statement of Proposition 7.1 and define $t_0 := \min\{1/4, \mu/2\}$.

Assumptions 7.2. Without loss of generality, we make the following assumptions:

- $K \geq 4s^2$, since increasing K makes the statement of Proposition 7.1 stronger,
- $\mathrm{Lip}_{\mathbb{G}}(f_0) \leq \min\{1/2, t_0\tau^2/32\}$, 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 7.1 using a technique similar to the one implemented in [28, Theorem 6.1], namely by using Algorithm 7.4 below to construct a sequence of Lipschitz functions (f_m) and a sequence of pairs (x_m, E_m) in D^{f_m} so that $E_m f(x_m)$ converges to an almost locally maximal directional derivative for f. More precisely, we show the limits (x_*, E_*) and f have the properties stated in Proposition 7.1.

Notation 7.3. Suppose $h : \mathbb{G} \to \mathbb{R}$ 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 $t \in (-1, 1)$

$$|(h(x' \exp(tE)) - h(x')) - (h(x \exp(tE)) - h(x))|$$
$$\leq K(\sigma + (E'f(x') - Ef(x))^{\frac{1}{2s^2}})|t|.$$

In the language of Notation 7.3, Proposition 7.1(2) means $(x_*, E_*) \leq_{(f,0)} (x, E)$.

Since N is G_δ we can fix open sets $U_k \subset \mathbb{G}$ such that $N = \cap_{k=0}^\infty U_k$. We may assume that $U_0 = \mathbb{G}$. We point out that, in Algorithm 7.4 below, the order in which the parameters are chosen plays a crucial role in what follows.

Algorithm 7.4. *Let f_0, x_0, E_0, τ and δ_0 be as in the assumptions of Proposition 7.1. Let $\sigma_0 := 2$ and $t_0 := \min\{1/4, \mu/2\}$.*

Then we can recursively define

(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) $t_m \in (0, \min\{t_{m-1}/2, \sigma_{m-1}/(s^2 m)\})$,

(4) $\lambda_m \in (0, \min\{t_m \sigma_m^{2s^2}/(2C_a^{2s^2}), t_m \tau^2/2^{2m+3}\})$,

(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 $\overline{B_{\mathbb{G}}(x_m, \delta_m)} \subset U_m$ and for all $|t| < C_D \delta_m^{\frac{1}{s}}/\varepsilon_m$*

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

Proof. Clearly one can make choices satisfying (1)–(5). For (6)–(8) we can proceed exactly as in [28, Proof of Algorithm 6.4] using Lemma 2.19 and Lemma 2.12 instead of [28, Lemma 5.2] and [28, Lemma 3.3], respectively.

□

We omit the proof of the following Lemma since it is exactly the same as the one of [28, Lemma 6.5] for the Heisenberg group.

Lemma 7.5. *The sequences $\sigma_m, t_m, \lambda_m, \delta_m, \varepsilon_m$ converge to 0, and for every $m \geq 1$ the inclusion*

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

holds.

We record for later use that $\mathrm{Lip}_{\mathbb{G}}(f_m) \leq 1$ for all $m \geq 1$ and we define $\varepsilon'_m > 0$ by letting

$$\varepsilon'_m := \min\{\varepsilon_m/2, \sigma_{m-1}/2\}. \tag{7.1}$$

We next show that the sets D_m form a decreasing sequence. This is an adaptation of [9, Lemma 3.3].

Lemma 7.6. *The following statements hold.*

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

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

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

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

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

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

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

Proof of (1). Algorithm 7.4(6) states that $(x_m, E_m) \in D_m$ and hence

$$d(E_m(0), E_{m-1}(0)) \leq \sigma_{m-1}. \tag{7.2}$$

Let $(x, E) \in D_{m+1}$. In particular, by Algorithm 7.4(5) we have $E f_{m+1}(x) \geq E_m f_{m+1}(x_m)$. Notice that, since $\omega(E_m) = \omega(E) = 1$, we have $\langle E_m(0), E_m(0) \rangle = 1$ and $\langle E(0), E_m(0) \rangle \leq 1$. Let $A := E f_m(x) - E_m f_m(x_m)$. Lemma 2.19 and the inequality $E f_{m+1}(x) \geq E_m f_{m+1}(x_m)$ give

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

Combining again Algorithm 7.4(5) with the above inequality, gives

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

In particular, $E f_m(x) \geq E_{m-1} f_m(x_{m-1})$ which is the first requirement for $(x_{m-1}, E_{m-1}) \leq_{(f_m, \sigma_m - \varepsilon'_m)} (x, E)$.

Let $B := E f_m(x) - E_{m-1} f_m(x_{m-1}) \geq 0$. Lemma 2.12 and $\text{Lip}_{\mathbb{G}}(f_m) \leq 1$ implies that $0 \leq A \leq B \leq 2$. Using these inequalities and $K \geq 4s^2$ gives

$$K(B^{\frac{1}{2s^2}} - A^{\frac{1}{2s^2}}) \geq (B^{\frac{2s^2-1}{2s^2}} + B^{\frac{2s^2-2}{2s^2}} A^{\frac{1}{2s^2}} + \ldots + B^{\frac{1}{2s^2}} A^{\frac{2s^2-2}{2s^2}} + A^{\frac{2s^2-1}{2s^2}})(B^{\frac{1}{2s^2}} - A^{\frac{1}{2s^2}})$$

$$= B - A$$

$$= E_m f_m(x_m) - E_{m-1} f_m(x_{m-1}). \tag{7.3}$$

Since $A \geq E f_{m+1}(x) - E_m f_{m+1}(x_m)$, (7.3) implies that

$$E_m f_m(x_m) - E_{m-1} f_m(x_{m-1}) + K(E f_{m+1}(x) - E_m f_{m+1}(x_m))^{\frac{1}{2s^2}}$$

$$\leq K B^{\frac{1}{2s^2}}. \tag{7.4}$$

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

$$|(f_m(x \exp(t E_{m-1})) - f_m(x)) - (f_m(x_{m-1} \exp(t E_{m-1})) - f_m(x_{m-1}))|. \tag{7.5}$$

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

Suppose $|t| < 3C_D \delta_m^{\frac{1}{s}}/\varepsilon_m$. Estimate (7.5) as follows

$$|(f_m(x \exp(tE_{m-1})) - f_m(x)) - (f_m(x_{m-1} \exp(tE_{m-1})) - f_m(x_{m-1}))|$$
$$\leq |(f_m(x \exp(tE_m)) - f_m(x)) - (f_m(x_m \exp(tE_m)) - f_m(x_m))|$$
$$+ |(f_m(x_m \exp(tE_m)) - f_m(x_m))$$
$$- (f_m(x_{m-1} \exp(tE_{m-1})) - f_m(x_{m-1}))|$$
$$+ |f_m(x \exp(tE_{m-1})) - f_m(x \exp(tE_m))|. \tag{7.6}$$

We consider the three terms on the right hand side of (7.6) separately.

Firstly, Algorithm 7.4(1) and Lemma 2.19 give

$$(f_m(x \exp(tE_m)) - f_m(x)) - (f_m(x_m \exp(tE_m)) - f_m(x_m)) \tag{7.7}$$
$$= (f_{m+1}(x \exp(tE_m)) - f_{m+1}(x)) - (f_{m+1}(x_m \exp(tE_m)) - f_{m+1}(x_m))$$
$$- t_m \langle x \exp(tE_m), E_m(0) \rangle + t_m \langle x, E_m(0) \rangle$$
$$+ t_m \langle x_m \exp(tE_m), E_m(0) \rangle - t_m \langle x_m, E_m(0) \rangle$$
$$= (f_{m+1}(x \exp(tE_m)) - f_{m+1}(x)) - (f_{m+1}(x_m \exp(tE_m)) - f_{m+1}(x_m)).$$

Since $(x, E) \in D_{m+1}$, using (7.7) gives

$$|(f_m(x \exp(tE_m)) - f_m(x)) - (f_m(x_m \exp(tE_m)) - f_m(x_m))|$$
$$\leq K(\sigma_m + (Ef_{m+1}(x) - E_m f_{m+1}(x_m))^{\frac{1}{2s^2}})|t|. \tag{7.8}$$

For the second term in (7.6) we recall that, for the values of t we are considering, Algorithm 7.4(8) states that

$$|(f_m(x_m \exp(tE_m)) - f_m(x_m)) - (f_m(x_{m-1} \exp(tE_{m-1})) - f_m(x_{m-1}))| \tag{7.9}$$
$$\leq (E_m f_m(x_m) - E_{m-1} f_m(x_{m-1}) + \sigma_{m-1})|t|.$$

The final term in (7.6) is estimated using $\mathrm{Lip}_{\mathbb{G}}(f_m) \leq 1$ and (7.2):

$$|f_m(x \exp(tE_{m-1})) - f_m(x \exp(tE_m))| \leq d(x \exp(tE_{m-1}), x \exp(tE_m)) \tag{7.10}$$
$$= d(tE_{m-1}(0), tE_m(0))$$
$$\leq \sigma_{m-1}|t|.$$

Adding (7.8), (7.9) and (7.10), then using (7.4), (7.1) and Algorithm 7.4(2), gives

$$|(f_m(x \exp(tE_{m-1})) - f_m(x)) - (f_m(x_{m-1} \exp(tE_{m-1})) - f_m(x_{m-1}))|$$
$$\leq K(\sigma_m + (Ef_{m+1}(x) - E_m f_{m+1}(x_m))^{\frac{1}{2s^2}})|t|$$
$$+ (E_m f_m(x_m) - E_{m-1} f_m(x_{m-1}) + \sigma_{m-1})|t|$$
$$+ \sigma_{m-1}|t|$$
$$\leq K(\sigma_{m-1} - \varepsilon_m' + (Ef_m(x) - E_{m-1} f_m(x_{m-1}))^{\frac{1}{2s^2}})|t|,$$

which gives the required estimate of (7.5) for all t with $|t| < 3C_D \delta_m^{\frac{1}{s}}/\varepsilon_m$.

Suppose $3C_D \delta_m^{\frac{1}{s}}/\varepsilon_m \leq |t| < 1$. In particular, this implies

$$\delta_m \leq \varepsilon_m^s t^s / 3C_D \leq \varepsilon_m |t|, \tag{7.11}$$

where in the last inequality above we used that

$$\varepsilon_m |t|/3C_D \leq \varepsilon_m/3C_D \leq 1,$$

which follows from $\varepsilon_m \leq 2$ and $C_D \geq 1$.

We estimate (7.5) as follows:

$$|(f_m(x\exp(tE_{m-1})) - f_m(x)) - (f_m(x_{m-1}\exp(tE_{m-1})) - f_m(x_{m-1}))|$$
$$\leq |(f_m(x_m\exp(tE_{m-1})) - f_m(x_m))$$
$$- (f_m(x_{m-1}\exp(tE_{m-1})) - f_m(x_{m-1}))|$$
$$+ |f_m(x) - f_m(x_m)|$$
$$+ |f_m(x\exp(tE_{m-1})) - f_m(x_m\exp(tE_{m-1}))|,$$

and again we separately consider the three terms on the right hand side.

By Algorithm 7.4(7) we have

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

which gives

$$|(f_m(x_m\exp(tE_{m-1})) - f_m(x_m)) - (f_m(x_{m-1}\exp(tE_{m-1})) - f_m(x_{m-1}))|$$
$$\leq K(\sigma_{m-1} - \varepsilon_m + (E_m f_m(x_m) - E_{m-1}f_m(x_{m-1}))^{\frac{1}{2s^2}})|t|. \qquad (7.12)$$

For the estimate of the second term we use $\mathrm{Lip}_{\mathbb{G}}(f_m) \leq 1$ and (7.11) to get

$$|f_m(x) - f_m(x_m)| \leq d(x, x_m) \leq \delta_m \leq \varepsilon_m|t| \leq K\varepsilon_m|t|/(4s^2). \qquad (7.13)$$

Notice that $x\exp(tE_{m-1})$ and $x_m\exp(tE_{m-1})$ belong to $\overline{B_{\mathbb{G}}(x_0, 2+\delta_0)}$. Using Proposition 2.4 and Proposition 2.5, recalling that $\delta_m < 1$, we get

$$|f_m(x\exp(tE_{m-1})) - f_m(x_m\exp(tE_{m-1}))|$$
$$\leq d(x\exp(tE_{m-1}), x_m\exp(tE_{m-1}))$$
$$= d(\exp(tE_{m-1})^{-1}x_m^{-1}x\exp(tE_{m-1}))$$
$$\leq C_D(d(x_m, x) + t^{\frac{1}{s}}d(x_m, x)^{\frac{s-1}{s}} + t^{\frac{s-1}{s}}d(x_m, x)^{\frac{1}{s}})$$
$$\leq C_D(\delta_m + t^{\frac{1}{s}}\delta_m^{\frac{s-1}{s}} + t^{\frac{s-1}{s}}\delta_m^{\frac{1}{s}})$$
$$\leq 3C_D\delta_m^{\frac{1}{s}} \leq \varepsilon_m|t| \leq K\varepsilon_m|t|/(4s^2). \qquad (7.14)$$

Combine (7.12), (7.13) and (7.14) to obtain

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

which gives the required estimate of (7.5) for all t satisfying $3C_D\delta_m^{\frac{1}{s}}/\varepsilon_m \leq |t| < 1$. This completes the proof of (1).

Proof of (2). Suppose $(x, E) \in D_{m+1}$. Then $(x, E) \in D^{f_{m+1}} = D^{f_m}$ and Lemma 7.5 implies that $d(x, x_{m-1}) < \delta_{m-1}$. Combining this with (1) gives $(x, E) \in D_m$. This completes the proof of (2).

Proof of (3). Suppose $(x, E) \in D_{m+1}$. Then by Algorithm 7.4(5) we have $E_m f_{m+1}(x_m) \leq Ef_{m+1}(x)$. Moreover, by Algorithm 7.4(1) we have

$$E_m f_m(x_m) + t_m\langle E_m(0), E_m(0)\rangle \leq Ef_m(x) + t_m\langle E(0), E_m(0)\rangle. \qquad (7.15)$$

By (2), we have also that $(x, E) \in D_m$, so Algorithm 7.4(6) implies

$$Ef_m(x) \leq E_m f_m(x_m) + \lambda_m. \qquad (7.16)$$

Combining (7.15) and (7.16) gives $t_m \leq t_m \langle E(0), E_m(0) \rangle + \lambda_m$, which, up to rearrangements, implies

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

Therefore one has

$$|p(E) - p(E_m)| = |E(0) - E_m(0)| = (2 - 2\langle E(0), E_m(0) \rangle)^{\frac{1}{2}} \leq (2\lambda_m/t_m)^{\frac{1}{2}}. \quad (7.17)$$

Combining Algorithm 7.4(5) and Lemma 5.2 with $g(t) := \exp(tE_m)$, $N = 1$ and $D = (2\lambda_m/t_m)^{\frac{1}{2}}$, we get

$$d(E(0), E_m(0)) = d(\exp(E), \exp(E_m)) \leq C_{\mathrm{a}}(2\lambda_m/t_m)^{\frac{1}{2s^2}}$$
$$\leq \sigma_m,$$

which proves (3). $\qquad \square$

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 an adaptation of [9, Lemma 3.4].

Lemma 7.7. *The following statements hold:*
   (1) $f_m \to f$ *pointwise, where* $f : \mathbb{G} \to \mathbb{R}$ *is Lipschitz and* $\mathrm{Lip}_{\mathbb{G}}(f) \leq 1$,
   (2) $f - f_m$ *is* $\mathbb{G}$-*linear and* $\mathrm{Lip}_{\mathbb{G}}(f - f_m) \leq 2t_m$ *for* $m \geq 0$,
   (3) *There exist* $x_* \in N$ *and* $E_* \in V$ *with* $\omega(E_*) = 1$ *such that for* $m \geq 0$ *we have*
$$d(x_*, x_m) < \delta_m, \quad \text{and} \quad d(E_*(0), E_m(0)) \leq \sigma_m.$$
   (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* $m \geq 1$,
   (7) $\omega(E_* - E_0) < \tau$.

Proof. We prove each statement individually.

Proof of (1). Algorithm 7.4(1) gives $f_m(x) = f_0(x) + \langle x, \sum_{k=0}^{m-1} t_k E_k(0) \rangle$. Define $f : \mathbb{G} \to \mathbb{R}$ by

$$f(x) := f_0(x) + \Big\langle x, \sum_{k=0}^{\infty} t_k E_k(0) \Big\rangle. \quad (7.18)$$

Notice $|f(x) - f_m(x)| \leq |x| \sum_{k=m}^{\infty} t_k |E_k(0)|$. Hence by Algorithm 7.4(3) $f_m \to f$ pointwise and, since $\mathrm{Lip}_{\mathbb{G}}(f_m) \leq 1$, we deduce $\mathrm{Lip}_{\mathbb{G}}(f) \leq 1$.

Proof of (2). Lemma 2.19 shows that $f - f_m$ is $\mathbb{G}$-linear. Moreover, by Algorithm 7.4(3), we have that for every $m \geq 0$

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

Proof of (3). Let $q \geq m \geq 0$. The definition of D_{q+1} in Algorithm 7.4(5) shows that $(x_q, E_q) \in D_{q+1}$. Hence points 2 and 3 of Lemma 7.6 imply that $(x_q, E_q) \in D_{m+1}$, and consequently

$$d(E_q(0), E_m(0)) \leq \sigma_m. \quad (7.19)$$

Since $(x_q, E_q) \in D_{m+1}$, Algorithm 7.4(5) implies

$$d(x_q, x_m) < \delta_m. \quad (7.20)$$

Since, by Lemma 7.5, $\sigma_m, \delta_m \to 0$ the sequences $(x_m)_{m=1}^\infty$ and $(E_m(0))_{m=1}^\infty$ are Cauchy, and therefore they converge to some $x_* \in \mathbb{G}$ and $v \in \mathbb{G}$, respectively. Since $E_m \in V_1$ and $\omega(E_m) = 1$, we know that $|p(v)| = 1$ and $v = (p(v), 0)$. Using group translations, we can extend v to a vector field $E_* \in V_1$ with $\omega(E_*) = 1$ and $E_*(0) = v$. Letting $q \to \infty$ in (7.19) and (7.20) implies that $d(E_*(0), E_m(0)) \le \sigma_m$ and $d(x_*, x_m) \le \delta_m$. Combining Lemma 7.5 and the fact that $\delta_m < \delta_{m-1}/2$, we have the strict inequality $d(x_*, x_m) < \delta_m$.

We now know that $x_* \in \overline{B_\mathbb{G}(x_m, \delta_m)}$ for every $m \ge 1$. Recall that $N = \cap_{m=0}^\infty U_m$ for open sets $U_m \subset \mathbb{G}$, and Algorithm 7.4(8) states that $\overline{B_\mathbb{G}(x_m, \delta_m)} \subset U_m$. 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 $q \ge m \ge 0$. Therefore, by Lemma 7.6(1), for every $q \ge m \ge 1$ we have

$$(x_{m-1}, E_{m-1}) \le_{(f_m, \sigma_{m-1} - \varepsilon'_m)} (x_q, E_q). \tag{7.21}$$

Algorithm 7.4(1) and (7.21) (with m and q replaced by $q+1$) give

$$E_q f_q(x_q) < E_q f_{q+1}(x_q) \le E_{q+1} f_{q+1}(x_{q+1}) \quad \text{for every} \quad q \ge 0. \tag{7.22}$$

Hence, since $E_0 f_0(x_0) \ge 0$, the sequence $(E_q f_q(x_q))_{q=0}^\infty$ is strictly increasing and positive.

Since $\mathrm{Lip}_\mathbb{G}(f_q) \le 1$ for every $q \ge 1$, by Lemma 2.12, 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 $0 < L \le 1$. Inequality (7.22) implies that also $E_q f_{q+1}(x_q) \to L$, and, moreover, one has

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

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

Let $q \ge m \ge 0$ and consider

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

By (7.21) we have that $s_{m,q} \ge 0$. Letting $q \to \infty$, writing $f_m = f + (f_m - f)$, and using the $\mathbb{G}$-linearity of $f_m - f$ one gets

$$s_{m,q} \to s_m := (f_m - f)(E_*(0)) + L - E_{m-1} f_m(x_{m-1}) \ge 0. \tag{7.23}$$

Since $\mathrm{Lip}_\mathbb{G}(f_m - f) \le 2t_m$ and $E_{m-1} f_m(x_{m-1}) \to L$, also $s_m \to 0$ as $m \to \infty$. (7.21) implies that

$$|(f_m(x_q \exp(tE_{m-1})) - f_m(x_q)) - (f_m(x_{m-1} \exp(tE_{m-1})) - f_m(x_{m-1}))|$$
$$\le K(\sigma_{m-1} - \varepsilon'_m + (s_{m,q})^{\frac{1}{2s^2}})|t| \quad \text{for } t \in (-1, 1). \tag{7.24}$$

Letting $q \to \infty$ in (7.24) shows that

$$|(f_m(x_* \exp(tE_{m-1})) - f_m(x_*)) - (f_m(x_{m-1} \exp(tE_{m-1})) - f_m(x_{m-1}))|$$
$$\le K(\sigma_{m-1} - \varepsilon'_m + (s_m)^{\frac{1}{2s^2}})|t| \quad \text{for } t \in (-1, 1). \tag{7.25}$$

Since $\mathrm{Lip}_\mathbb{G}(f) \le 1$ and $d(E_*(0), E_{m-1}(0)) \le \sigma_{m-1}$, we obtain

$$|f(x_* \exp(tE_*)) - f(x_* \exp(tE_{m-1}))| \le d(x_*(tE_*(0)), x_*(tE_{m-1}(0)))$$
$$\le \sigma_{m-1}|t|. \tag{7.26}$$

Since $f - f_m$ is $\mathbb{G}$-linear and $\mathrm{Lip}_{\mathbb{G}}(f - f_m) \leq 2t_m$ we can estimate

$$
\begin{aligned}
|(f - f_m)(x_* \exp(tE_{m-1})) - (f - f_m)(x_*)| &= |(f - f_m)(\exp(tE_{m-1}))| \\
&= |(f - f_m)(\delta_t(\exp(E_{m-1})))| \\
&\leq t\mathrm{Lip}_{\mathbb{G}}(f - f_m) \\
&\leq 2t_m|t|. \quad\quad (7.27)
\end{aligned}
$$

Combining (7.25), (7.26) and (7.27) shows that for $t \in (-1, 1)$:

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

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

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

and

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

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

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

Hence, for every $|t| < \delta$

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

This proves that $E_*f(x_*)$ exists and is equal to L. We have already seen 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 2.19 imply

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

Using (7.23) shows $s_m = E_* f_m(x_*) - E_{m-1} f_m(x_{m-1}) \geq 0$. Substituting this in (7.25) gives (5).

Proof of (6). Property (6) is a consequence of (3), (4) and (5).

Proof of (7). We start by estimating $\omega(E_1 - E_0)$. By Algorithm 7.4(6), for every $(x, E) \in D_1$ we have

$$
Ef_1(x) \leq E_1 f_1(x_1) + \lambda_1,
$$

where

$$
f_1(x) = f_0(x) + t_0 \langle x, E_0(0) \rangle. \quad\quad (7.28)
$$

Clearly
$$E_0 f_1(x_0) = E_0 f_0(x_0) + t_0$$
and
$$E_1 f_1(x_1) = E_1 f_0(x_1) + t_0 \langle E_1(0), E_0(0) \rangle.$$
By Algorithm 7.4(5), $(x_0, E_0) \in D_1$ and therefore
$$E_0 f_1(x_0) \leq E_1 f_1(x_1) + \lambda_1. \tag{7.29}$$
A simple calculation using (7.29) then gives
$$\omega(E_1, E_0) = \langle E_1(0), E_0(0) \rangle \geq 1 - \frac{\lambda_1}{t_0} + \frac{E_0 f_0(x_0) - E_1 f_0(x_1)}{t_0}. \tag{7.30}$$
Since ω is an inner product norm we can estimate as follows:
$$
\begin{aligned}
\omega(E_1 - E_0) &= \left(\omega(E_1)^2 + \omega(E_0)^2 - 2\omega(E_1, E_0) \right)^{\frac{1}{2}} \\
&= (2 - 2\omega(E_1, E_0))^{\frac{1}{2}} \\
&\leq \left(\frac{2\lambda_1}{t_0} + \frac{2|E_0 f_0(x_0) - E_1 f_0(x_1)|}{t_0} \right)^{\frac{1}{2}} \\
&\leq \left(\frac{2\lambda_1}{t_0} + \frac{4\mathrm{Lip}_{\mathbb{G}}(f_0)}{t_0} \right)^{\frac{1}{2}} \\
&\leq \frac{\tau}{2},
\end{aligned}
$$
where in the last inequality above we used the estimate on $\mathrm{Lip}_{\mathbb{G}}(f_0)$ in Assumption 7.2 and the estimate on λ_1 in Algorithm 7.4(3). Next, as proved in (7.17), for every $m \geq 1$ and $(x, E) \in D_{m+1}$ we have
$$\omega(E - E_m) \leq (2\lambda_m / t_m)^{\frac{1}{2}}. \tag{7.31}$$
Using the estimate in Algorithm 7.4(4), this implies that for every $m \geq 1$:
$$\omega(E_{m+1} - E_m) < \frac{\tau}{2^{m+1}}. \tag{7.32}$$
Therefore,
$$
\begin{aligned}
\omega(E_* - E_0) = \lim_{m \to \infty} \omega(E_m - E_0) &\leq \sum_{m=2}^{\infty} \omega(E_m - E_{m-1}) + \omega(E_1 - E_0) \\
&< \tau \sum_{m=2}^{\infty} \frac{1}{2^m} + \frac{\tau}{2} \\
&= \tau.
\end{aligned}
$$

$\square$

We now prove that the limit directional derivative $E_* f(x_*)$ is almost locally maximal in horizontal directions. This is an adaptation of [9, Lemma 3.5].

Lemma 7.8. *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 $\varepsilon > 0$. By Lemma 7.5 we choose $m \geq 1$ such that

$$m \geq 4/\varepsilon^{\frac{2s^2-1}{2s^2}} \quad \text{and} \quad \lambda_m, t_m \leq \varepsilon/4. \tag{7.33}$$

Recall that $\varepsilon'_m = \min\{\varepsilon_m/2,\ \sigma_{m-1}/2\}$. Using Lemma 7.7(3) and 7.7(6) , fix $\delta_\varepsilon > 0$ such that

$$\delta_\varepsilon < \delta_{m-1} - d(x_*, x_{m-1})$$

such that for every $|t| < 3D\delta_\varepsilon^{\frac{1}{s}}/\varepsilon'_m$

$$|(f_m(x_* \exp(tE_*)) - f_m(x_*)) - (f_m(x_{m-1}\exp(tE_{m-1})) - f_m(x_{m-1}))|$$
$$\leq (E_* f_m(x_*) - E_{m-1}f_m(x_{m-1}) + \sigma_{m-1})|t|. \tag{7.34}$$

Such δ_ε exists since, by Lemma 7.7(5), we have $E_* f_m(x_*) \geq E_{m-1}f_m(x_{m-1})$.

We argue by contradiction and we 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 7.4(6) and the monotone convergence $E_m f_m(x_m) \uparrow E_* f(x_*)$ would then imply

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

By Lemma 7.7(2) and (7.33) we would deduce that

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

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

Proof that $(x, E) \in D_m$. Since $f - f_m$ is $\mathbb{G}$-linear we have $D^f = D^{f_m}$ and therefore $(x, E) \in D^{f_m}$. Next observe that

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

Hence, it suffices to show that $(x_{m-1}, E_{m-1}) \leq_{(f_m, \sigma_{m-1} - \varepsilon'_m/2)} (x, E)$. Lemma 2.12 implies

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

Hence by definition of (x, E) and by (7.33) we have

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

Lemma 7.7(6) states that $(x_*, E_*) \in D_m$, which implies $E_{m-1}f_m(x_{m-1}) \leq E_* f_m(x_*)$ and 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 requirement of $(x_{m-1}, E_{m-1}) \leq_{(f_m, \sigma_{m-1} - \varepsilon'_m/2)} (x, E)$.

We next deduce several inequalities from our hypotheses. Let

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

By definition of (x, E) we have $A \geq \varepsilon$, while the inequalities above give $0 \leq B \leq C$. By Lemma 2.12 we have that A, B, $C \leq 2$. Recalling the factorization

$$A - B = \left(B^{\frac{2s^2-1}{2s^2}} + B^{\frac{s^2-1}{s^2}} A^{\frac{1}{2s^2}} + \ldots + B^{\frac{1}{2s^2}} A^{\frac{s^2-1}{s^2}} + A^{\frac{2s^2-1}{2s^2}} \right) \left(A^{\frac{1}{2s^2}} - B^{\frac{1}{2s^2}} \right) \tag{7.35}$$

and using Lemma 7.7(2), (7.33) and Algorithm 7.4(3), we obtain

$$\begin{aligned} A^{\frac{1}{2s^2}} - B^{\frac{1}{2s^2}} &\leq (A - B)/\varepsilon^{\frac{2s^2-1}{2s^2}} \\ &= (E(f - f_m)(x) - E_*(f - f_m)(x_*))/\varepsilon^{\frac{2s^2-1}{2s^2}} \\ &\leq 4t_m/\varepsilon^{\frac{2s^2-1}{2s^2}} \\ &\leq m t_m \\ &\leq \sigma_{m-1}/s^2. \end{aligned}$$

Since B, $C \leq 2$ and $K \geq 4s^2$ we have

$$B^{\frac{2s^2-1}{2s^2}} + B^{\frac{s^2-1}{s^2}} A^{\frac{1}{2s^2}} + \ldots + B^{\frac{1}{2s^2}} A^{\frac{s^2-1}{s^2}} + A^{\frac{2s^2-1}{2s^2}} \leq 4s^2 \leq K.$$

Hence using (7.35) with A replaced by C gives

$$KC^{\frac{1}{2s^2}} - KB^{\frac{1}{2s^2}} \geq C - B = E_* f_m(x_*) - E_{m-1} f_m(x_{m-1}).$$

Combining our estimates we eventually find

$$\begin{aligned} E_* f_m(x_*) &- E_{m-1} f_m(x_{m-1}) + K(Ef(x) - E_* f(x_*))^{\frac{1}{2s^2}} \\ &= E_* f_m(x_*) - E_{m-1} f_m(x_{m-1}) + KA^{\frac{1}{2s^2}} \\ &\leq KC^{\frac{1}{2s^2}} - KB^{\frac{1}{2s^2}} + K(B^{\frac{1}{2s^2}} + \sigma_{m-1}/s^2) \\ &= K((Ef_m(x) - E_{m-1} f_m(x_{m-1}))^{\frac{1}{2s^2}} + \sigma_{m-1}/s^2). \end{aligned} \tag{7.36}$$

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

$$|(f_m(x \exp(tE_{m-1})) - f_m(x)) - (f_m(x_{m-1} \exp(tE_{m-1})) - f_m(x_{m-1}))|. \tag{7.37}$$

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

Suppose $|t| \leq 3C_D \delta_{\varepsilon}^{\frac{1}{s}}/\varepsilon'_m$. To estimate (7.37) we use the inequality

$$\begin{aligned} |(f_m(x \exp(tE_{m-1})) &- f_m(x)) - (f_m(x_{m-1} \exp(tE_{m-1})) - f_m(x_{m-1})| \\ &\leq |(f_m(x \exp(tE_*)) - f_m(x)) - (f_m(x_* \exp(tE_*)) - f_m(x_*))| \\ &\quad + |(f_m(x_* \exp(tE_*)) - f_m(x_*)) \\ &\qquad - (f_m(x_{m-1} \exp(tE_{m-1})) - f_m(x_{m-1}))| \\ &\quad + |f_m(x \exp(tE_{m-1})) - f_m(x \exp(tE_*))|. \end{aligned} \tag{7.38}$$

Since $(x_*, E_*) \leq_{(f,0)} (x, E)$, by Lemma 2.19 and $\mathbb{G}$-linearity of $f_m - f$ we can estimate the first term in (7.38) by

$$\begin{aligned} |(f_m(x \exp(tE_*)) &- f_m(x)) - (f_m(x_* \exp(tE_*)) - f_m(x_*))| \\ &\leq |f(x \exp(tE_*)) - f(x)) - (f(x_* \exp(tE_*)) - f(x_*))| \\ &\leq K(Ef(x) - E_* f(x_*))^{\frac{1}{2s^2}} |t|. \end{aligned} \tag{7.39}$$

Since t is small, by (7.34) we get

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

Lemma 7.7 implies that the third term of (7.38) is bounded above by $\sigma_{m-1}|t|$. By combining the estimates of each term and using (7.36) we get

$$|(f_m(x \exp(tE_{m-1})) - f_m(x)) - (f_m(x_{m-1}\exp(tE_{m-1})) - f_m(x_{m-1}))|$$
$$\leq (K(Ef(x) - E_* f(x_*))^{\frac{1}{2s^2}} + E_* f_m(x_*) - E_{m-1}f_m(x_{m-1}) + 2\sigma_{m-1})|t|$$
$$\leq (K((Ef_m(x) - E_{m-1}f_m(x_{m-1}))^{\frac{1}{2s^2}} + \sigma_{m-1}/s^2) + 2\sigma_{m-1})|t|$$
$$\leq K(\sigma_{m-1} - \varepsilon'_m/2 + (Ef_m(x) - E_{m-1}f_m(x_{m-1}))^{\frac{1}{2s^2}})|t|, \tag{7.40}$$

where we have used $\varepsilon'_m \leq \sigma_{m-1}/2$ and $K \geq 4s^2$ in the final line. This gives the required estimate of (7.37) for small t.

Suppose $3C_D \delta_\varepsilon^{\frac{1}{s}}/\varepsilon'_m \leq |t| \leq 1$. To estimate (7.37) we use the inequality

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

Lemma 7.7(5) shows that the first term on the right hand side is bounded above by

$$K(\sigma_{m-1} - \varepsilon'_m + (E_* f_m(x_*) - E_{m-1}f_m(x_{m-1}))^{\frac{1}{2s^2}})|t|.$$

The second term is bounded by $d(x_*, x) \leq \delta_\varepsilon \leq \varepsilon'_m|t| \leq K\varepsilon'_m|t|/s^2$. For the third term, we use Proposition 2.5 with $x = \exp(tE_{m-1})$ and $y = x_*^{-1}x$ to get

$$|f_m(x \exp(tE_{m-1})) - f_m(x_* \exp(tE_{m-1}))|$$
$$\leq d(x \exp(tE_{m-1}), x_* \exp(tE_{m-1}))$$
$$\leq C_D \left(d(x_*, x) + t^{\frac{1}{s}}d(x_*, x)^{\frac{s-1}{s}} + t^{\frac{s-1}{s}}d(x_*, x)^{\frac{1}{s}}\right)$$
$$\leq C_D(\delta_m + \delta_m^{\frac{s-1}{s}} + \delta_m^{\frac{1}{s}})$$
$$\leq 3C_D \delta_m^{\frac{1}{s}}$$
$$\leq \varepsilon'_m|t|$$
$$\leq K\varepsilon'_m|t|/(4s^2).$$

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

$$|(f_m(x \exp(tE_{m-1})) - f_m(x)) - (f_m(x_{m-1}\exp(tE_{m-1})) - f_m(x_{m-1}))|$$
$$\leq K(\sigma_{m-1} - \varepsilon'_m/2 + (Ef_m(x) - E_{m-1}f_m(x_{m-1}))^{\frac{1}{2s^2}})|t|.$$

This gives the required estimate of (7.37) for large t and therefore

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

which concludes the proof. $\square$

We conclude this section proving Proposition 7.1 and Theorem 1.1.

Proof of Proposition 7.1. Proposition 7.1 easily follows from Lemma 7.7 and Lemma 7.8. Indeed, Lemma 7.7 states that there is $f\colon \mathbb{G} \to \mathbb{R}$ Lipschitz such that $f - f_0$ is linear and $\mathrm{Lip}_{\mathbb{G}}(f - f_0) \leq 2t_0 \leq \mu$. It also states there is $(x_*, E_*) \in D^f$ satisfying $d(x_*, x_0) < \delta_0$ and $E_* f(x_*) > 0$. Lemma 7.8 then shows that $E_* f(x_*)$ is almost locally maximal in the sense of Proposition 7.1. $\qquad\square$

Proof of Theorem 1.1. Let $B \subset V_1$ be a ball of directions as in Assumption 6.1. Let $f_0\colon \mathbb{G} \to \mathbb{R}$ 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 without loss of generality that $\mathrm{Lip}_{\mathbb{G}}(f_0) \leq 1/4$.

Fix an arbitrary pair $(x_0, E_0) \in D^{f_0}$ and apply Proposition 7.1 with $\delta_0 = 1$, $\mu = 1/4$ and $K = 4s^2$. This gives a Lipschitz function $f\colon \mathbb{G} \to \mathbb{R}$ such that $f - f_0$ is $\mathbb{G}$-linear with $\mathrm{Lip}_{\mathbb{G}}(f - f_0) \leq 1/4$ and a pair $(x_*, E_*) \in D^f$ with $x_* \in N$ and $E_* f(x_*) > 0$ which 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) $d(x, x_*) \leq \delta_\varepsilon$, $Ef(x) \geq E_* f(x_*)$, and
(2) for any $t \in (-1, 1)$:

$$|(f(x \exp(tE_*)) - f(x)) - (f(x_* \exp(tE_*)) - f(x_*))|$$
$$\leq 4s^2 |t| (Ef(x) - E_* f(x_*))^{\frac{1}{2s^2}},$$

then

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

Since $\mathrm{Lip}_{\mathbb{G}}(f_0) \leq 1/4$ and $\mathrm{Lip}_{\mathbb{G}}(f - f_0) \leq 1/4$ we have $\mathrm{Lip}_{\mathbb{G}}(f) \leq 1/2$. Notice that (x_*, E_*) is also almost locally maximal in the sense of Theorem 6.6, since the restriction on pairs above is weaker than that in Theorem 6.6. Hence Theorem 6.6 implies that f is Pansu differentiable at $x_* \in N$. Since a $\mathbb{G}$-linear function is Pansu differentiable everywhere, it follows f_0 is Pansu differentiable at x_*. This proves Theorem 1.1. $\qquad\square$

References

[1] Agrachev, A., Barilari, D., Boscain, U. *Introduction to Riemannian and Sub-Riemannian geometry*, available at https://webusers.imj-prg.fr/ davide.barilari/Notes.php

[2] Alberti, G., Marchese, A.: *On the differentiability of Lipschitz functions with respect to measures in Euclidean spaces*, Geometric and Functional Analysis 26(1) (2016), 1–66.

[3] Alberti, G., Csornyei, M., Preiss, D.: *Differentiability of Lipschitz functions, structure of null sets, and other problems*, Proceedings of the International Congress of Mathematicians III (2010), 1379–1394.

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

[5] Cheeger, J.: *Differentiability of Lipschitz functions on metric measure spaces*, Geometric and Functional Analysis 9(3) (1999), 428–517.

[6] Chow, W. L.: *Uber Systeme von linearen partiellen Differentialgleichungen erster Ordnung*, Mathematische Annalen 117 98–105.

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

[8] Csornyei, M., Jones, P.: *Product formulas for measures and applications to analysis and geometry*, announcement of result is in slides available at: http://www.math.sunysb.edu/Videos/dfest/PDFs/38-Jones.pdf.

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

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

[11] Doré, M., Maleva, O.: *A universal differentiability set in Banach spaces with separable dual*, Journal of Functional Analysis 261 (2011), 1674–1710.

[12] Dymond, M., Maleva, O.: *Differentiability inside sets with upper Minkowski dimension one*, Michigan Mathematical Journal 65 (2016), 613–636.

[13] Fitzpatrick, S.: *Differentiation of real-valued functions and continuity of metric projections*, Proceedings of the American Mathematical Society 91(4) (1984), 544–548.

[14] Franchi, B., Serapioni, R.: *Intrinsic Lipschitz graphs within Carnot groups*, Journal of Geometric Analysis, 26(3) (2016), 1946–1994.

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

[16] Le Donne, E., Montgomery, R., Ottazzi, A., Pansu, P., Vittone, D.: *Sard Property for the endpoint map on some Carnot groups*, Annales de l'Institut Henri Poincaré, Analyse Non Lineairé 33(6) (2016), 1639–1666.

[17] Le Donne, E., Pinamonti, A., Speight, G.: *Universal Differentiability Sets and Maximal Directional Derivatives in Carnot Groups*, Journal de Mathématiques Pures et Appliquées 121 (2019), 83–112.

[18] 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).

[19] Magnani, V.: *Differentiability and area formula on stratified Lie groups*, Houston Journal of Mathematics 27(2) (2001), 297–323.

[20] Magnani, V.: *Towards differential calculus in stratified groups*, Journal of the Australian Mathematical Society 95 (2013), 76–128.

[21] Magnani, V., Rajala, T.: *Radon-Nikodym property and area formula for Banach homogeneous group targets*, International Mathematics Research Notices 23 (2014), 6399–6430.

[22] Magnani, V., Pinamonti, A., Speight, G.: *Differentiability for Lipschitz maps from stratified groups to Banach homogeneous groups*, preprint available at arXiv:1706.01782.

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

[24] Nagel, A., Stein, E., Wainger, S.: *Balls and metrics defined by vector fields I: Basic properties*, Acta Mathematica 155 (1985), 103–147.

[25] Ottazzi, A.: *A sufficient condition for nonrigidity of Carnot groups*, Mathematische Zeitschrift 259(3) (2008), 617–629.

[26] Ottazzi, A., Vittone, D.: *On the codimension of the abnormal set in step two Carnot groups*, to appear in ESAIM: Control, Optimisation and Calculus of Variations, preprint available at arXiv:1709.02854.

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

[28] Pinamonti, A., Speight, G.: *A Measure Zero Universal Differentiability Set in the Heisenberg Group*, Mathematische Annalen 368, no. 1-2, (2017) 233–278.

[29] Pinamonti, A., Speight, G.: *A Measure Zero UDS in the Heisenberg Group*, Bruno Pini Mathematical Analysis Seminar, Vol. 7 (2016) 85–96.

[30] Preiss, D.: *Differentiability of Lipschitz functions on Banach spaces*, Journal of Functional Analysis 91(2) (1990), 312–345.

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

[32] Preiss, D., Maleva, O.: *Cone unrectifiable sets and non-differentiability of Lipschitz functions*, to appear in Israel Journal of Mathematics, preprint available at arXiv:1709.04233.

[33] Rudin, W.: *Principles of mathematical analysis, third edition*, International Series in Pure and Applied Mathematics, McGraw-Hill Book Co. (1976).

[34] Varadarajan, V.S.: *Lie Groups, Lie Algebras, and Their Representations*, Graduate Textbooks in Mathematics 102, Springer-Verlag, 1984.

[35] Vittone, D.: *The regularity problem for sub-Riemannian geodesics*, pages 193–226 in *Geometric Measure Theory and Real Analysis*, publications of the Scuola Normale Superiore 17.

[36] Warhurst, B.: *Contact and quasiconformal mappings on real model filiform groups*, Bulletin of the Australian Mathematical Society 68(2) (2003), 329–343.

[37] Xie, X.: *Quasi-conformal maps on model filiform groups*, Michigan Mathematical Journal 64(1) (2015), 169–202.

[38] Zahorski, Z.: *Sur l'ensemble des points de non-dérivabilité d'une fonction continue*, Bulletin de la Société Mathématique de France 74 (1946), 147–178.

(Andrea Pinamonti) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TRENTO, VIA SOMMARIVE 14, 38050 POVO (TRENTO), ITALY
    E-mail address, Andrea Pinamonti: `Andrea.Pinamonti@unitn.it`

(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`