Journal of Convex Analysis
Volume 14 (2007), No. 4, 667–691

Non-Smooth Analysis in Infinite Dimensional Banach Homogenous Groups

Kevin Rogovin
University of Jyväskylä, Department of Mathematics and Statistics, P.O. Box 35 (MaD), 40014 Jyväskylä, Finland
kevin.rogovin@gmail.com

Received: April 19, 2004
Revised manuscript received: September 11, 2006

We develop a theory of non-smooth analysis in infinite dimensional Banach homogenous groups and extend the Fritz John necessary condition for minimizers of a Lipschitz function subject to Lipschitz constraints.

1. Introduction

The study of minimization is one of the cornerstone uses of calculus. Typically, to minimize a given function $f : \mathbb{R}^n \to \mathbb{R}$, one is often reduced to find the zeros of Df. This simple algorithm is the foundation for much of the calculus of variations. Indeed, given an open bounded domain Ω of $\mathbb{R}^n$ and boundary values $u \in W^{1,2}(\Omega)$, the harmonic function $h + u$ with $h \in W_0^{1,2}(\Omega)$ is the unique minimum to the functional

$$\mathcal{L}_u(\phi) = \int_\Omega |\nabla(u + \phi)|^2$$

across $W_0^{1,2}(\Omega)$. Many of the analytic properties of h follows from considering the derivative of $\mathcal{L}_u$. Indeed since h minimizes $\mathcal{L}_u$ we then see that

$$D\mathcal{L}_u(h) = 0$$

which implies that for each $\phi \in C_c^\infty(\Omega)$,

$$\int_\Omega \langle \nabla(u + h), \nabla\phi \rangle = 0.$$

Nonsmooth analysis deals with the case where the function $f : \mathbb{R}^n \to \mathbb{R}$ or the functional $\mathcal{L}$ are not continuously differentiable. Its main technique is to replace the derivative with a set of sub-derivatives. Indeed, following F. H. Clarke see [3, Chapter 2], if X is a Banach space and $\mathcal{L} : X \to \mathbb{R}$ is locally Lipschitz, one defines the generalized derivative of $\mathcal{L}$ at a point p as the set of all $\zeta \in X^*$ (where X^* is the space of continuous real valued linear maps) so that

$$\zeta(\cdot) \leq \mathcal{L}^\circ(p; \cdot)$$

where

$$\mathcal{L}^\circ(p; v) = \limsup_{q \to p, \lambda \downarrow 0} \frac{\mathcal{L}(q + \lambda v) - \mathcal{L}(q)}{\lambda}.$$

ISSN 0944-6532 / $ 2.50

It is well known that the the class of locally Lipschitz functions is much larger than the class of continuously differentiable functions. Further, starting with any such function f and performing composition, multiplication or addition with smooth or non-smooth functions, one can produce examples of locally Lipschitz functions which are not differentiable.

Recently there has been active research on homogenous groups also known as Carnot groups as introduced by Folland and Stein in [6]. We do not give a full list of the results in the area of analysis in Carnot groups, but we do mention the works [11], [2], [8], [1], [7], [10], [13] and [9]. A Carnot group $\mathcal{G}$ is topologically $\mathbb{R}^n$, but the group operation one employs is not the standard vector addition. This group operation is typically non-commutative. However, for each $\lambda > 0$ there is a non-homogenous dilation δ_λ which is a group isomorphism. The metric used on a Carnot group produces the same topology as the Euclidean metric, but has a much richer class of locally Lipschitz functions than the Euclidean case. Moreover, when doing analysis in Carnot groups of a real valued function f, one examines $\nabla_H f$ the horizontal gradient of f rather than the full gradient of f. The class of functions whose horizontal gradient exists and is continuous is denoted by C^1_{sub}. There are functions $f \in C^1_{\text{sub}}$ which are not locally Lipschitz with respect to the Euclidean metric, but to minimize such functions one just needs to find the zeros of the horizontal gradient. The horizontal gradient at a point p of a function f can be calculated by examining $D_H f(p)$ the horizontal derivative of f at p which is an element of $\mathcal{L}(\mathcal{G};\mathbb{R})$. The set $\mathcal{L}(\mathcal{G};\mathbb{R})$ is the set of all continuous maps $\phi : \mathcal{G} \to \mathbb{R}$ for which $\phi(p \cdot q) = \phi(p) + \phi(q)$ and $\phi \circ \delta_\lambda = \lambda\phi$, i.e., $\mathcal{L}(\mathcal{G};\mathbb{R})$ is the Carnot group analogue of the of space of continuous real valued linear maps, refer to Section 2 for the definition of δ_λ. When $f \in C^1_{\text{sub}}$, $D_H f(p)$ is the unique element in $\mathcal{L}(\mathcal{G};\mathbb{R})$ for which

$$D_H f(p)(g) = \lim_{\lambda \to 0} \frac{f(p \cdot \delta_\lambda(g)) - f(p)}{\lambda}$$

whenever $g \in \mathcal{G}$. With this in mind a natural way to extend the Clarke theory of non-smooth analysis to Carnot groups (see Section 2 for the definitions) is to say $\zeta \in \mathcal{L}(\mathcal{G};\mathbb{R})$ is a sub-derivative of a locally Lipschitz function u at p if

$$\zeta(\cdot) \le u^\circ(p;\cdot)$$

where

$$u^\circ(p;g) = \limsup_{q \to p, \lambda \downarrow 0} \frac{u(q\delta_\lambda(g)) - u(q)}{\lambda}$$

and we will write $\zeta \in \partial u(p)$. It is worth noting that if $u \in C^1_{\text{sub}}$, then $\partial u(p) = \{D_H u(p)\}$, see Corollary 3.16. We will define an infinite dimensional analogue of Carnot groups which we call Banach homogenous groups which will allow us, together with the theorems developed within, to examine minimization problems for which the standard tools of the calculus of variations do not apply. In the classical case of where the domain of u is a Banach space, the Hahn-Banach theorem assures us there are always sub-derivatives of a locally Lipschitz function u at a point p. Our first theorem, Theorem 3.11, allows us to use the Hahn-Banach theorem even in the setting of Banach homogenous groups. Our main theorem is an analogue of the Fritz John necessary conditions of minimizing a locally Lipschitz functions subject to locally Lipschitz constraints:

Theorem 1.1. *Let $\mathcal{G}$ be an infinite (or finite) dimensional Banach homogenous group and U an open subset of $\mathcal{G}$. Let $u : U \to \mathbb{R}$, $v : U \to \mathbb{R}^m$ and $w : U \to \mathbb{R}^n$ be locally*

Lipschitz functions (with $\mathcal{G}$ given the Carnot-Carathéodory metric). If $x_0 \in S$ so that $u(x_0) = \inf_S u$ where

$$S := \{x \in U \mid v(x) \leq \mathbf{0},\, w(x) = \mathbf{0}\},$$

then there exists $\alpha = (\lambda_0, \gamma, \lambda) \in \{0,1\} \times \mathbb{R}^m \times \mathbb{R}^n$ which is non-zero with $\gamma \geq \mathbf{0}$ and $\langle \gamma, v(x)\rangle = 0$ so that

$$0 \in \partial(\lambda_0 f + \langle \gamma, v\rangle + \langle \lambda, w\rangle)(x_0).$$

Here for vector $\gamma \in \mathbb{R}^m$ we write $\gamma \geq \mathbf{0}$ if each component of γ is non-negative.

One corollary of Theorem 1.1 is a Lagrange multiplier rule for C^1_{sub} functions on Carnot groups.

Corollary 1.2. *Let $u, h_1, \ldots, h_m$ be real valued C^1_{sub} functions defined on an open subset U of a Carnot group G so that for each $p \in S$ where $S = \cap_{j=1}^m h_j^{-1}(0)$, the vectors $\{\nabla_H h_j(p)\}_{j=1}^m$ are linearly independent. If $p_0 \in S$ so that $u(p_0) = \inf_S u$, then there exists $\lambda \in \mathbb{R}^m$ so that*

$$0 = \nabla_H(u + \langle \lambda, h\rangle)(p_0)$$

where $h = (h_1, h_2, \ldots, h_m)$.

The simplest example of a Banach homogenous group whose metric is different from the Euclidean metric is the first Heisenberg group $\mathcal{H}_1$. As a set $\mathcal{H}_1$ is $\mathbb{R}^3$, but the metric is so that the measure of a ball of radius R is comparable to R^4. Moreover the class of locally Lipschitz function with respect to the metric of $\mathcal{H}_1$ is much richer than the class of functions which are locally Lipschitz with respect to the Euclidean metric.

Corollary 1.2 in the Heisenberg group setting opens the window to calculating minimum of functions defined on surfaces which are not smooth, or even algebraic varieties. Indeed, F. Serra Cassano and B. Kirchheim in [12] have constructed a function $f : \mathbb{R}^2 \to \mathbb{R}$ whose graph is (locally) the zero set of a Lipschitz function on the Heisenberg group but has Euclidean Hausdorff dimension $2\frac{1}{2}$. In fact, the graph is (locally) the zero set of a C^1_{sub} function whose horizontal gradient is non-zero.

The paper is organized as follows. In Section 2 we state our basic definitions and give examples of Banach homogenous groups. In Section 3 we prove some of basic properties of these concepts. In Section 4 we present the proof to our main theorem, Theorem 1.1 and also give proofs for a mean value theorem and a chain rule for generalized derivatives of functions defined on Banach homogenous groups.

2. Definitions

Let $\mathcal{G} = V_1 \times V_2 \times \ldots \times V_n$ with each V_i a Fréchét space (i.e. each V_i is a complete metric space with a topological vector space structure) and V_1 is a Banach space. For each $1 \leq i \leq n$, let d_i denote the metric on V_i (note that $d_1(x,y) = ||x-y||_1$ where $||\cdot||_1$ is the norm of V_1) and set $\pi_i : \mathcal{G} \to V_i$ as projection. For an element $v \in \mathcal{G}$ we will write $v_i = \pi_i(v)$ and $v = (v_1, v_2, \ldots, v_n)$. We make $\mathcal{G}$ into a Fréchét space using the metric $d_0(v,w) = \sum d_i(v_i, w_i)$. We let $||\cdot||_1^*$ denote the norm on V_1^* produced by $||\cdot||_1$, i.e., for $\phi \in V_1^*$,

$$||\phi||_1^* = \sup_{x \in V_1,\, ||x||_1 = 1} |\phi(x)|.$$

Definition 2.1. We say that $\mathcal{G}$ with a continuous group structure that makes the identity 0 and a function $||\cdot||_{CC} : \mathcal{G} \to [0,\infty)$ is an *n-step Banach homogenous group* if it satisfies the following.

(1) The group operation has the form

$$x \cdot y = x + y + Q(x,y)$$

with $\pi_1 \circ Q = 0$, for $2 \le i \le n$, $\pi_i(Q(x,y)) = Q_i((x_j)_{j=1}^{i-1}, (y_j)_{j=1}^{i-1})$ and for each $x \in \mathcal{G}$ and $\lambda \in \mathbb{R}$ we require that $Q(x, \lambda x) = 0$.

(2) For each $\lambda > 0$ the map given by

$$\delta_\lambda(v_1, v_2, \ldots, v_n) = (\lambda v_1, \lambda^2 v_2, \ldots, \lambda^n v_n)$$

is a group isomorphism. Recall that we will write an element $v \in \mathcal{G}$ as $v = (v_1, v_2, \ldots, v_n)$ where $v_i = \pi_i(v)$.

(3) The function $||\cdot||_{CC} : \mathcal{G} \to [0,\infty)$ is a *gauge*, i.e.,
(i) for each $\lambda > 0$, $||\delta_\lambda(x)||_{CC} = \lambda ||x||_{CC}$,
(ii) $||x \cdot y||_{CC} \le ||x||_{CC} + ||y||_{CC}$,
(iii) $||-x||_{CC} = ||x||_{CC}$ and
(iv) $||x||_{CC} = 0$ if and only if $x = 0$.

(4) The left invariant metric $d_{CC}(x,y) := ||x^{-1} \cdot y||_{CC}$ is complete and the topologies induced by d_{CC} and d_0 are equivalent. We will call d_{CC} the Carnot-Carathéodory metric on $\mathcal{G}$. We will write $\mathbf{B}_{CC}(p,r)$ as the set of all those $x \in \mathcal{G}$ for which $d_{CC}(x,p) < r$.

Note that each Banach space $(V, ||\cdot||_V)$ with the group structure $x \cdot y = x + y$ with its norm is a one step Banach homogenous group.

For two Banach homogenous groups $\mathcal{G}_1$ and $\mathcal{G}_2$ we define

$$\mathcal{L}(\mathcal{G}_1;\mathcal{G}_2) := \{\phi : \mathcal{G}_1 \to \mathcal{G}_2 \mid \phi \text{ is a continuous group homomorphism so that for each } \lambda > 0, \phi \circ \delta_\lambda = \delta_\lambda \circ \phi\}.$$

We will say a function $f : U \to \mathbb{R}$ where U is an open subset of $\mathcal{G}$ is *locally Lipschitz*, if for each $p \in U$ there exists $r > 0$ and $K \in \mathbb{R}$ so that for each $x, y \in \mathbf{B}_{CC}(p,r)$,

$$|f(x) - f(y)| \le K d_{CC}(x,y).$$

Remark 2.2. The definition of a Banach homogenous group is motivated by the form of the group law of Carnot groups in exponential coordinates. Indeed in exponential coordinates (i.e., on the Lie algebra) the group law of a Carnot group is of the form

$$x \cdot y = x + y + Q(x,y)$$

where Q is a finite sum of iterated brackets generated by the Baker-Campbell-Hausdorff formula. Analogously to both Carnot groups and Banach spaces, a Banach homogenous group possesses the following properties.

(A) The fact that $Q(x, \lambda x) = 0$ implies that $x \cdot \lambda x = (\lambda + 1)x$, in particular for each $m \in \mathbb{Z}$ and $x \in \mathcal{G}$, $x^m = mx$. Thus, $x^{-1} = -x$ and $||x^{-1}||_{CC} = ||x||_{CC}$. Additionally, for each $x_1 \in V_1$, $\delta_{s+t}(x_1) = \delta_s(x_1) \cdot \delta_t(x_1)$. Moreover, since $Q_0 = 0$, for $g \in \mathcal{G}$ and $h \in \pi_1^{-1}(0)$, we see that $\pi_1(gh) = \pi_1(hg) = \pi_1(g)$. We also see that for each k, if $\pi_j(g) = \pi_j(h) = 0$ for $1 \le j \le k$, then $\pi_j(g \cdot h) = 0$ for $1 \le j \le k$ as well.

(B) For each $k \geq 1$, if $\pi_j(v) = 0$ for $1 \leq j < k$, then there exists $w^i \in V_i$ for $k \leq i \leq n$ so that $v = w^k \cdot w^{k+1} \cdot \ldots \cdot w^n$ and $w^k = \pi_k(v)$. Indeed, if $k = n$, then $v \in V_n$ and we are done. Inductively, if v is such that $\pi_j(v) = 0$ for $1 \leq j < k$, let $x = \pi_k(v)^{-1}v$. The group law (item (1) of Definition 2.1) then gives us that for $1 \leq j < k$,

$$\pi_j(x) = \pi_j(v) + Q_j(0,0) = 0$$

and

$$\pi_k(x) = \pi_k(-v + v) + Q_k(0,0) = 0.$$

Thus the induction hypothesis implies that $x = w^{k+1} \cdot \ldots \cdot w^n$ with $w^i \in V_i$. Thus, $v = v_k \cdot x = \pi_k(v) \cdot w^{k+1} \cdot \ldots \cdot w^n$ as needed. Also note that the above implies that if $v \in \mathcal{G}$ with $\pi_j(v) = 0$ for $1 \leq j < k$, then there exists $w^i \in V_i$ for $k \leq i \leq n$ so that $v = w^n \cdot w^{n-1} \cdot \ldots \cdot w^k$ and $w^k = \pi_k(v)$. Indeed, let v be as above and let $x = v^{-1}$. Then (A) gives us that $x = -v$ and thus $\pi_j(x) = 0$ for $1 \leq j < k$. Hence there exists $a^j \in V_j$ for $k \leq j \leq n$ so that $x = a^k \cdot a^{k+1} \cdot \ldots \cdot a^n$ with $a^k = \pi_k(x)$. Thus

$$v = x^{-1} = (a^k \cdot a^{k+1} \cdot \ldots \cdot a^n)^{-1} = (a^n)^{-1} \cdot (a^{n-1})^{-1} \cdot \ldots \cdot (a^k)^{-1}.$$

Letting $w^j = (a^j)^{-1} = -a^j$ we see that $w^j \in V_j$, $w^k = \pi_k(v)$ and $v = w^n \cdot w^{n-1} \cdot \ldots \cdot w^k$.

(C) In the case where each of the V_i's are Banach spaces and Q is continuously differentiable, Q and it's derivatives have polynomial growth. Indeed, since δ_λ is a group homomorphism, we see that for $2 \leq i \leq n$,

$$Q_i((\lambda^j x_j)_{j=1}^{i-1}, (\lambda^j y_j)_{j=1}^{i-1}) = \lambda^i Q_i((x_j)_{j=1}^{i-1}, (y_j)_{j=1}^{i-1}). \tag{1}$$

Recall that the Fréchét derivative of Q_i when it exists at a point $(x, y) \in \mathcal{G}$ denoted by $DQ_i(x, y)$ is the unique continuous linear map from $\mathcal{G} \times \mathcal{G}$ to V_i for which

$$\lim_{v,w \to 0} \frac{||Q_i(x + v, y + w) - Q_i(x, y) - DQ_i(x, y)(v, w)||_i}{||v||_0 + ||w||_0} = 0$$

where $|| \cdot ||_i$ is the norm on V_i and $||v||_0 = \sum_i ||v_i||_i$. Using equation (1) one has that for each $\lambda > 0$, $x, y \in \mathcal{G}$ and $v_j, w_j \in V_j$ (with $1 \leq j < i$),

$$DQ_i(\delta_\lambda(x), \delta_\lambda(y))(v_j, w_j) = \lambda^{i-j} DQ_i(x, y)(v_j, w_j).$$

In particular,

$$||DQ_i(\delta_\lambda(x), \delta_\lambda(y))|| \leq ||DQ_i(x, y)|| \cdot \sum_{j=1}^{i-1} \lambda^j$$

where $||DQ_i(x, y)||$ is the operator norm of $DQ_i(x, y)$. Since Q_i is continuously differentiable, there is a neighborhood of 0 for which $||DQ_i||$ is bounded by a constant M. Hence, using the above we see that $||DQ_i||$ has polynomial growth. If we let $\tau_x(y) = x \cdot y$, then the above gives a polynomial $\Phi : \mathbb{R}^2 \to \mathbb{R}$ so that $||D\tau_x(y)|| \leq \Phi(||x||_0, ||y||_0)$.

(D) If each of the V_i are Banach spaces, then we have a ball-box theorem on $\mathcal{G}$, i.e., there exists a constant $C > 0$ so that for each $r > 0$

$$\frac{1}{C}\mathbf{Box}(0, r) \subseteq \mathbf{B}_{CC}(0, r) \subseteq C\mathbf{Box}(0, r)$$

where for each $r > 0$ and $\lambda > 0$

$$\lambda\mathbf{Box}(0,r) := \{(v_1, v_2, \dots, v_n) \in \mathcal{G} \mid \text{for } 1 \le i \le n,\ ||v_i||_i < \lambda r^i\}.$$

Indeed, since the topologies produced by the metrics d_{CC} and d_0 are equivalent there exists $C > 0$ so that

$$\frac{1}{C}\mathbf{Box}(0,1) \subseteq \mathbf{B}_{CC}(0,1) \subseteq C\mathbf{Box}(0,1).$$

Let $v \in \mathbf{B}_{CC}(0,r)$ and set $w = \delta_{1/r}(v)$. Then $w \in \mathbf{B}_{CC}(0,1)$ and thus also $w \in C\mathbf{Box}(0,1)$. Hence for each i, $||w_i||_i \le C$. Now $v = \delta_r(w)$. Thus for each i, $||v_i|| = r^i||w_i|| \le Cr^i$, i.e., $v \in C\mathbf{Box}(0,r)$. Hence $\mathbf{B}_{CC}(0,r) \subseteq C\mathbf{Box}(0,r)$. On the other hand if $v \in \frac{1}{C}\mathbf{Box}(0,r)$, let $w = \delta_{1/r}(v)$. Then $||w_i||_i = \frac{1}{r^i}||v_i||_i \le \frac{1}{C}$, i.e., $w \in \frac{1}{C}\mathbf{Box}(0,1)$ and thus $w \in \mathbf{B}_{CC}(0,1)$. Now $v = \delta_r(w)$, thus $v \in \mathbf{B}_{CC}(0,r)$. Along similar lines by using a continuity argument, one can show that if V_i is a Banach space (without assuming weather or not the other components of $\mathcal{G}$ are Banach spaces) that there exists a constant $C > 0$ so that for each $v^i \in V_i$,

$$\frac{1}{C}||v^i||_i \le ||v^i||^i_{CC} \le C||v^i||_i.$$

(E) By homogeneity, a group homomorphism ζ between two Banach homogenous group $\mathcal{G}_1$ and $\mathcal{G}_2$ which commutes with δ_λ for each $\lambda > 0$ is continuous if and only if

$$||\phi||_{\mathcal{L}(\mathcal{G}_1;\mathcal{G}_2)} := \sup_{||g||_{CC} \le 1} ||\phi(g)||_{CC} < \infty.$$

For the case where $\mathcal{G}_2$ is $\mathbb{R}$ we make $\mathcal{L}(\mathcal{G};\mathbb{R})$ into a normed linear space by using the above norm. Additionally, for each Banach homogenous group $\mathcal{G}$, $\mathcal{L}(\mathcal{G};\mathcal{G})$ also becomes a group by composition for which each $\lambda > 0$ induces a group isomorphism, $\Delta_\lambda(\phi) = \delta_\lambda \circ \phi$. Additionally, for each $\phi_1, \phi_2 \in \mathcal{L}(\mathcal{G};\mathcal{G})$, $||\phi_1 \circ \phi_2||_{\mathcal{L}(\mathcal{G};\mathcal{G})} \le ||\phi_1||_{\mathcal{L}(\mathcal{G};\mathcal{G})} \cdot ||\phi_2||_{\mathcal{L}(\mathcal{G};\mathcal{G})}$. Moreover, a standard proof (see for instance [5]) to the Open Mapping theorem can be adapted to Banach homogenous groups to conclude the following Open Mapping theorem for Banach homogenous groups:

Theorem 2.3. *Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be Banach homogenous groups and $\zeta \in \mathcal{L}(\mathcal{G}_1;\mathcal{G}_2)$ so that ζ is onto. Then ζ is also an open map. In particular, if ζ is a continuous group isomorphism, then ζ^{-1} is continuous.*

One useful corollary of the above theorem is:

Theorem 2.4. *Let $\mathcal{G}$ with $||\cdot||_{CC}$ be a Banach homogenous group. If $|||\cdot|||_{CC}$ is another gauge on $\mathcal{G}$ which makes $\mathcal{G}$ into a Banach homogenous group for which there exists a constant $K > 0$ so that $||\cdot||_{CC} \le K|||\cdot|||_{CC}$, then there exists another constant $K' > 0$ so that $|||\cdot|||_{CC} \le K'||\cdot||_{CC}$.*

Some examples of Banach homogenous groups follow.

Example 2.5. All Carnot groups are Banach homogenous groups. Indeed, let G be a Carnot group whose Lie algebra $\mathcal{G}$ admits the stratification $\mathcal{G} = V_1 \oplus V_2 \oplus \ldots \oplus V_r$, i.e.,

$\text{span}[V_1, V_i] = V_{i+1}$ for $1 \leq i \leq n-1$ and $[V_1, V_n] = 0$. Let $\langle \cdot, \cdot \rangle_0$ be an inner product on $\mathcal{G}$ and define a Riemannian metric, $\langle \cdot, \cdot \rangle_x$ on G by letting

$$\langle D\tau_x(0)(v), D\tau_x(0)(w) \rangle_x = \langle v, w \rangle_0$$

for each $v, w \in T_0G$ where $\tau_x(p) = x \cdot p$, i.e., the above inner product is the left-invariant inner product produced by $\langle \cdot, \cdot \rangle_0$. Given an interval I of $\mathbb{R}$, we say a piecewise smooth curve $\gamma : I \to G$ is horizontal if for a.e. $t \in I$, $\gamma'(t) \in D\tau_{\gamma(t)}(0)(V_1)$. Define the Carnot-Carathéodory metric on G as

$$d_{CC}(p, q) = \inf\{l(\gamma) \,|\, \gamma \text{ is a horizontal curve connecting } p \text{ to } q\}$$

where

$$l(\gamma) = \int \langle \gamma'(t), \gamma'(t) \rangle_{\gamma(t)}^{1/2} \, dt.$$

By Chow's theorem, $d_{CC}(p, q)$ is finite for each pair $p, q \in G$. Since G and $\mathcal{G}$ are diffeomorphic through the exponential map we can identify G and $\mathcal{G}$. The Baker-Campbell-Hausdorff formula gives that on $\mathcal{G}$, the group law is given by

$$x \cdot y = x + y + Q(x, y)$$

where Q is a finite sum of iterated brackets. Hence Q satisfies item (1) of Definition 2.1. Additionally, for each $\lambda > 0$ we have that the map $\delta_\lambda : \mathcal{G} \to \mathcal{G}$ given by $\pi_i(\delta_\lambda(v)) = \lambda^i \pi_i(v)$ is a group homomorphism where $\pi_i : \mathcal{G} \to V_i$ is projection. The function $||x||_{CC} = d_{CC}(0, x)$ is a gauge because d_{CC} is a left invariant metric with $d_{CC}(-x, 0) = d_{CC}(x, 0)$ and $d_{CC}(\delta_\lambda(x), 0) = \lambda d_{CC}(x, 0)$ for all $x \in \mathcal{G}$ and $\lambda > 0$.

Example 2.6. A less traditional example is an infinite dimensional Heisnenberg group. Fix an infinite dimensional real Hilbert space $(H_\mathbb{R}, \langle \cdot, \cdot \rangle_\mathbb{R})$ and let $(H_\mathbb{C}, \langle \cdot, \cdot \rangle_\mathbb{C})$ be the complexification of $\mathbb{H}_\mathbb{R}$, i.e., $H_\mathbb{C} = \mathbb{H}_\mathbb{R} \times \mathbb{H}_\mathbb{R}$ and

$$\langle (x_1, y_1), (x_2, y_2) \rangle_\mathbb{C} = \langle x_1, x_2 \rangle_\mathbb{R} + \langle y_1, y_2 \rangle_\mathbb{R} - i \langle x_1, y_2 \rangle_\mathbb{R} + i \langle y_1, x_2 \rangle_\mathbb{R} .$$

Let $\mathcal{G} = H_\mathbb{C} \times \mathbb{R}$ with the group operation

$$(x, s) \cdot (y, t) = (x + y, s + t + \frac{1}{2} \text{Im} \langle x, y \rangle_\mathbb{C}).$$

Analogous to the Heisenberg setting, we define the horizontal plane at a point $p \in \mathcal{G}$ as

$$\mathcal{H}_p\mathcal{G} := D\tau_p(0)(H_\mathbb{C} \times 0)$$

where $\tau_p(x) := p \cdot x$. Given and interval I of $\mathbb{R}$, we say a piecewise smooth curve $\gamma : I \to \mathcal{G}$ where I is an interval of $\mathbb{R}$ is horizontal if for a.e. t, $\gamma'(t) \in \mathcal{H}_{\gamma(t)}\mathcal{G}$. Let $\langle \cdot, \cdot \rangle_0$ be the real inner-product on $\mathcal{G}$ produced when viewing $\mathcal{G}$ as $H_\mathbb{R}^2 \times \mathbb{R}$ and $|| \cdot ||_0$ be the norm produced by $\langle \cdot, \cdot \rangle_0$. For each $p \in \mathcal{G}$ we create a new inner product as

$$\langle D\tau_p(0)(v), D\tau_p(0)(w) \rangle_p = \langle v, w \rangle_0 .$$

For $(x, s) \in \mathcal{G}$, let V be a one dimensional complex vector subspace of $H_\mathbb{C}$ containing x, i.e., if $x \neq 0$, then $V = \text{span}_\mathbb{C}(x)$. The set $V \times \mathbb{R}$ with the above group action is isomorphic to the first Heisenberg group $\mathcal{H}_1$ through the map $\iota : \mathcal{H}_1 \to V \times \mathbb{R}$ defined by

$$\iota(c, t) = (cz, t) \tag{2}$$

where $z \in V$ is fixed with $\langle z, z \rangle_{\mathbb{C}} = 1$. Moreover, if a smooth curve $\gamma : I \to \mathcal{H}_1$ is horizontal, then the curve $\iota \circ \gamma$ is horizontal in $\mathcal{G}$. Hence by Chow's theorem on the first Heisenberg group, we see that for each $p \in \mathcal{G}$ there exists a piecewise horizontal smooth curve $\gamma : I \to \mathcal{G}$ connecting p to $(0, 0)$. This allows us to define a Carnot-Carathéodory metric on $\mathcal{G}$, d_{CC} as

$$d_{CC}(g, h) := \inf\{l(\gamma) \,|\, \gamma \text{ is a horizontal curve connecting } g \text{ to } h\}$$

where

$$l(\gamma) := \int_I \langle \gamma'(t), \gamma'(t) \rangle_{\gamma(t)} \, dt.$$

Just as in the case of a Carnot group, the above metric is left invariant because a curve γ is horizontal if and only if for each $p \in \mathcal{G}$ and $\lambda > 0$, the curves $\tau_p \circ \gamma$ and $\delta_\lambda \circ \gamma$ are horizontal. Hence the function $||x||_{CC} = d_{CC}(x, 0)$ is a gauge and for each pair of points $p, q \in \mathcal{G}$, $d_{CC}(p, q) = ||p^{-1}q||_{CC}$.

Since $D\tau_{-(z,t)}(q)(w, s) = (w, s) + (0, \frac{1}{2}\text{Im}\,\langle -z, w \rangle_{\mathbb{C}})$ we see that for each $v \in \mathcal{G}$ and $p \in \mathcal{G}$

$$\begin{aligned} \langle v, v \rangle_0 &= \langle D\tau_{-p}(p)(D\tau_p(0)(v)), D\tau_{-p}(p)(D\tau_p(0)(v)) \rangle_0 \\ &\leq ||D\tau_{-p}(p)||^2 \langle D\tau_p(0)(v)), D\tau_p(0)(v)) \rangle_0 \\ &\leq (1 + ||p||_0)^2 \langle v, v \rangle_p . \end{aligned}$$

In particular, for each $R > 0$ and each pair of points p and q with $||p||_0 < R$ and $||q||_0 < R$ we have that $d_{CC}(p, q) \geq \frac{1}{C(R)}||p - q||_0$. We claim that the metric d_{CC} still produces the same topology. Indeed, let $p = (x, t) \in \mathcal{G}$ and let V be a one complex dimensional subspace of $H_{\mathbb{C}}$ containing x endowed with the inner-product from $\langle \cdot, \cdot \rangle_0$. Then the subgroup $V \times \mathbb{R}$ is isomorphic to $\mathcal{H}_1$ by the map $\iota : \mathcal{H}_1 \to V \times \mathbb{R}$ given by (2) so that $d_{CC}(\iota(p), 0) \leq d_{CC}(p, 0)$. Now the Carnot-Carathéodory metric on $\mathcal{H}_1$ satisfies $d_{CC}((z, t), 0) \leq C(||z|| + |t|^{1/2})$. In particular, $||(x, t)||_{CC} \leq C(||x||_0 + |t|^{1/2})$. Thus, $d_{CC}((x, s), (y, t)) \leq C(||x - y||_0 + |t - s - \frac{1}{2}\text{Im}\,\langle x, y \rangle_{\mathbb{C}}|^{1/2})$ which goes to zero as $||(x - y, t - s)||_0 \to 0$, i.e., d_{CC} is a complete metric and produces the same topology as the Hilbert space structure. We conclude that the function $||p||_{CC} := d_{CC}(p, 0)$ will make $\mathcal{G}$ a Banach homogenous group. Note that if we took $H_{\mathbb{R}} = \mathbb{R}^n$, then $\mathcal{G}$ would be the n-th Heisenberg group, $\mathcal{H}_n$.

Following the notation found in Example 2.6 we define the following for a Banach homogenous group $\mathcal{G} = V_1 \times V_2 \times \ldots V_n$ where each V_i is a Hilbert space.

Definition 2.7. Let $\mathcal{G} = V_1 \times \ldots, \times V_n$ be a Banach homogenous group where each V_i is a Hilbert space with inner-product $\langle \cdot, \cdot \rangle_i$ and the group law is smooth. Let $\langle v, w \rangle_0 = \sum \langle v_i, w_i \rangle_i$. For each $p \in \mathcal{G}$ create the inner-product $\langle \cdot, \cdot \rangle_p$ as

$$\langle D\tau_p(0)(v), D\tau_p(0)(w) \rangle_p = \langle v, w \rangle_0 ,$$

i.e.,

$$\langle v, w \rangle_p = \langle D\tau_{-p}(p)(v), D\tau_{-p}(p)(w) \rangle_0$$

where $\tau_p(x) = p \cdot x$. This inner-product is the left invariant inner-product produced by $\langle \cdot, \cdot \rangle_0$. For each $p \in \mathcal{G}$ we let $\mathcal{H}_p\mathcal{G}$ be the *horizontal space at* p defined by

$$\mathcal{H}_p\mathcal{G} = D\tau_p(0)(V_1 \times 0 \times \ldots \times 0).$$

For each $p \in \mathcal{G}$ we let $\pi_p : \mathcal{G} \to \mathcal{H}_p\mathcal{G}$ be orthogonal projection onto $\mathcal{H}_p\mathcal{G}$ with respect to inner-product $\langle \cdot, \cdot \rangle_p$, i.e.,

$$\pi_p = D\tau_p(0) \circ \pi_1 \circ D\tau_p(0)^*$$

where π_1 is orthogonal projection onto V_1 with respect to the inner-product $\langle \cdot, \cdot \rangle_0$ and $D\tau_p(0)^*$ denotes the adjoint of $D\tau_p(0)$. For a function $u : \mathcal{G} \to \mathbb{R}$, we say u has a *horizontal gradient* at p if the function $u_p : V_1 \to \mathbb{R}$

$$u_p(v) = u(\tau_p(v))$$

has a Fréchét derivative at $v = 0$, i.e., there exists $\phi \in V_1^*$ so that

$$\lim_{v \to 0, v \in V_1} \frac{|u_p(v) - u_p(0) - \phi(v)|}{\langle v, v \rangle_1^{1/2}} = 0.$$

We set $\nabla_H u(p) = D\tau_p(0)(w_p)$ where w_p is the unique element in V_1 so that $\phi(\cdot) = \langle \cdot, w_p \rangle_1$. We say $u \in C^1_{\text{sub}}$ if the map $p \to \nabla_H u(p)$ is continuous as a map between $\mathcal{G}$ and V_1 when V_1 is given the topology generated by the norm $|| \cdot ||_1$. Note that if u is C^1, then u is also C^1_{sub} and

$$\begin{aligned} \nabla_H u(p) &= D\tau_p(0)(\pi_1(\nabla(u \circ \tau_p)(0))) \\ &= \pi_p(\nabla u(p)). \end{aligned}$$

Definition 2.8. For the case where not all of V_i are Hilbert spaces we can still define the *horizontal derivative* $D_H u$ of a real valued function u as follows. We say a function u defined on a open neighborhood of p has a horizontal derivative at $p \in \mathcal{G}$ if the function $u_p : U \to \mathbb{R}$ where $U \subseteq V_1$ is an open neighborhood 0 in V_1

$$u_p(x) = u(p \cdot x)$$

has a Fréchét derivative at $x = 0$, i.e., there exists $\phi \in V_1^*$ so that

$$\lim_{x \to 0} \frac{|u_p(x) - u_p(0) - \phi(x)|}{||x||_1} = 0$$

where the limit is across all $x \in V_1$ going to zero and $|| \cdot ||_1$ is the norm on V_1. We set $D_H u(p) = \phi$. We will say $u \in C^1_{\text{sub}}$ if for each p, u has a horizontal derivative and the function $p \to D_H u(p)$ is continuous as a map between $\mathcal{G}$ and V_1^* when V_1^* is given the topology generated by the norm $|| \cdot ||_1^*$. Note that if each of the V_i are Hilbert spaces and the group law is smooth, then

$$D_H u(p)(v_1) = \langle \nabla_H u(p), D\tau_p(v_1) \rangle_p$$

where $\nabla_H u(p)$ and $\langle \cdot, \cdot \rangle_p$ are as in Definition 2.7. Moreover if each of the V_i are Banach spaces and the group law is smooth, then

$$D_H u(p) = D(u \circ \tau_p)(0) \circ \pi_1$$

where π_1 is projection onto V_1. In particular, if u is smooth, then u is also C^1_{sub}.

3. Basics

Throughout we let $\mathcal{G} = V_1 \times V_2 \times \ldots \times V_n$ be a Banach homogenous group. Recall that we say a real valued function f defined on an open subset $U \subseteq \mathcal{G}$ is locally Lipschitz if for each $p \in U$, there exists a constant K and an $\epsilon > 0$ so that for $x, y \in \mathbf{B}_{CC}(p, \epsilon)$, $|f(x) - f(y)| \leq K d_{CC}(x, y)$.

Proposition 3.1. *If $L \in \mathcal{L}(\mathcal{G}; \mathbb{R})$, then L restricted to V_1 is an element of V_1^* and $L = L \circ \pi_1$. Moreover, given $\phi \in V_1^*$, $\phi \circ \pi_1 \in \mathcal{L}(\mathcal{G}; \mathbb{R})$. In particular, for a map $u : \mathcal{G} \to \mathbb{R}$, $D_H u(p) \in \mathcal{L}(\mathcal{G}; \mathbb{R})$ whenever $D_H u(p)$ is defined.*

Proof. We first claim that for each $i > 1$ and $v_i \in V_i$, $L(v_i) = 0$. Indeed, $\delta_2(v_i) = 2^i v_i$ and for each $m \in \mathbb{N}$, $(v_i)^m = m v_i$. Thus

$$2L(v_i) = L(\delta_2(v_i)) = L(2^i v_i) = L((v_i)^{2^i}) = 2^i L(v_i)$$

which implies that $L(v_i) = 0$. By Remark 2.2(B), we have for each $v \in \mathcal{G}$ there exist $w^2 \in V_2$, $w^3 \in V_3$, $\ldots$, $w^n \in V_n$ so that $v = \pi_1(v) \cdot w^2 \cdot \ldots \cdot w^n$. Hence

$$\begin{aligned} L(v) &= L(\pi_1(v) \cdot w^2 \cdot w^3 \cdot \ldots \cdot w^n) \\ &= L(\pi_1(v)) + L(w^2) + L(w^3) + \ldots + L(w^n) \\ &= L(\pi_1(v)). \end{aligned}$$

Moreover, for $x, y \in V_1$,

$$L(x + y) = L(\pi_1(x \cdot y)) = L(x \cdot y) = L(x) + L(y)$$

and for $\lambda > 0$,

$$L(\lambda x) = L(\delta_\lambda(x)) = \lambda L(x).$$

Since $x^{-1} = -x$, the above implies that $L : V_1 \to \mathbb{R}$ is linear. That L restricted to V_1 is continuous follows from the fact that L is continuous on $\mathcal{G}$.

On the other hand, given $\phi \in V_1^*$, the map $\phi \circ \pi_1$ is continuous on $\mathcal{G}$ because $\pi_1 : \mathcal{G} \to V_1$ is continuous. For each $x, y \in \mathcal{G}$,

$$\phi(\pi_1(x \cdot y)) = \phi(\pi_1(x) + \pi_1(y)) = \phi(\pi_1(x)) + \phi(\pi_1(y))$$

and for each $\lambda > 0$,

$$\phi(\pi_1(\delta_\lambda(x))) = \phi(\lambda \pi_1(x)) = \lambda \phi(\pi_1(x))$$

which implies that $\phi \circ \pi_1 \in \mathcal{L}(\mathcal{G}; \mathbb{R})$ as needed. □

Corollary 3.2. *Let $\zeta \in \mathcal{L}(\mathcal{G}; \mathbb{R})$. Then for each $g \in \pi_1^{-1}(0)$ and $h \in \mathcal{G}$,*

$$\zeta(gh) = \zeta(hg) = \zeta(h).$$

In light of Proposition 3.1, we will say a sequence ζ_n of elements in $\mathcal{L}(\mathcal{G}; \mathbb{R})$ converges weak* to an element $\zeta \in \mathcal{L}(\mathcal{G}; \mathbb{R})$ if for each $g \in \mathcal{G}$, $\zeta_n(g) \to \zeta(g)$, which is equivalent to requiring that $\zeta_n \to \zeta$ weak* as elements of V_1^*.

Definition 3.3. We call a function $\rho : \mathcal{G} \to \mathbb{R}$ a *bounded signed gauge* if the following three conditions are satisfied.

(1) For each $g, h \in \mathcal{G}$, $\rho(gh) \leq \rho(g) + \rho(h)$.
(2) For each $g \in \mathcal{G}$ and $\lambda > 0$, $\rho(\delta_\lambda(g)) = \lambda\rho(g)$.
(3) There exists a constant $K > 0$ so that for each $g \in \mathcal{G}$ we have $\rho(g) \leq K||g||_{CC}$.

Remark 3.4. By using the fact that $d_{CC}(g,h) = ||h^{-1}g||_{CC} = ||g^{-1}h||_{CC}$, one can easily derive that if (1)-(3) hold with K, then for each $g, h \in \mathcal{G}$,

$$|\rho(g) - \rho(h)| \leq K d_{CC}(g,h),$$

i.e., ρ is K-Lipschitz. Moreover, if a function $\rho : \mathcal{G} \to \mathbb{R}$ satisfies (1) and (2) of above, then homogeneity implies that ρ satisfies (3) if and only if ρ is continuous.

Definition 3.5. Let u be a real valued locally Lipschitz function defined on a neighborhood of $p \in \mathcal{G}$. Following Clarke, we define the *generalized directional derivative of u at p in the (right) direction g* as

$$u^\circ(p; g) = \limsup_{q \to p, \lambda \downarrow 0} \frac{u(q\delta_\lambda(g)) - u(q)}{\lambda}.$$

Proposition 3.6. *If u is C^1_{sub} on a neighborhood of $p \in \mathcal{G}$, then $u^\circ(p; g) = D_H u(p)(g)$ whenever $g \in V_1$ where $D_H u(p)$ is the horizontal derivative of u at p as defined in Definition* 2.8. *Moreover, if v is Lipschitz on a neighborhood of p, then $(u + v)^\circ(p; g) = u^\circ(p; g) + v^\circ(p; g)$ for each $g \in V_1$.*

Proof. Let $g \in V_1$. We first claim that if u is C^1_{sub}, then for q close to p the map

$$f_q(t) = u(q \cdot \delta_t(g))$$

is C^1 on $(0, \epsilon)$ for some $\epsilon > 0$. Since u is C^1_{sub}, for p close to q, the map $u_q : V_1 \to \mathbb{R}$ defined as $u_q(v_1) = u(q \cdot v_1)$ has a Fréchét derivative $Du_q(0) \in V_1^*$ at $v_1 = 0$, hence for each $g \in V_1$,

$$\lim_{\lambda \to 0} \frac{u_q(\lambda g) - u_q(0)}{\lambda} = Du_q(0)(g).$$

Now for $t > 0$,

$$\begin{aligned}
\lim_{h \to 0} \frac{f_q(t+h) - f_q(t)}{h} &= \lim_{h \to 0} \frac{u(q\delta_{t+h}(g)) - u(q\delta_t(g)}{h} \\
&= \lim_{h \to 0} \frac{u(q\delta_t(g) \cdot \delta_h(g)) - u(q\delta_t(g))}{h} \\
&= \lim_{h \to 0} \frac{u_{q\delta_t(g)}(hg) - u_{q\delta_t(g)}(0)}{h} \\
&= Du_{q\delta_t(g)}(0)(g) \\
&= D_H u(q\delta_t(g))(g)
\end{aligned}$$

which because $D_H u$ is continuous, is a continuous function in t. In particular, f_q is C^1. For each q near p and $\lambda > 0$ near 0, the Mean Value theorem produces a $t(\lambda, q) \in (0, \lambda)$ so that

$$\frac{f_q(\lambda) - f_q(0)}{\lambda} = f'(t(\lambda, q)).$$

Hence,

$$\begin{aligned} u^\circ(p;g) &= \limsup_{q\to p,\lambda\downarrow 0} \frac{u(q\delta_\lambda(g)) - u(q)}{\lambda} \\ &= \limsup_{q\to p,\lambda\downarrow 0} \frac{f_q(\lambda) - f_q(0)}{\lambda} \\ &= \limsup_{q\to p,\lambda\downarrow 0} f'(t(\lambda,q)) \\ &= \limsup_{q\to p,\lambda\downarrow 0} D_H u(q\delta_{t(\lambda,q)}(g))(g) \\ &= D_H u(p)(g) \end{aligned}$$

where the last line comes from the fact that $D_H u$ is continuous. The above also shows us that

$$u^\circ(p;g) = \lim_{q\to p,\lambda\downarrow 0} \frac{u(q\delta_\lambda(g)) - u(q)}{\lambda}$$

which gives us for any real valued locally Lipschitz function v defined on a neighborhood of p that $(u+v)^\circ(p;g) = u^\circ(p;g) + v^\circ(p;g)$ for $g \in V_1$. □

Proposition 3.7. *Let u be a real valued K-Lipschitz function defined on a neighborhood of p. We then have the following.*

(1) *The function $g \to u^\circ(p;g)$ is a bounded signed gauge with constant K.*
(2) *If $p_n \to p$ and $g_n \to g$, then*

$$\limsup_{n\to\infty} u^\circ(p_n;g_n) \le u^\circ(p;g).$$

Proof. We first show that $\rho(g) = u^\circ(p;g)$ is a bounded signed gauge with constant K. Indeed, for each pair $g,h \in \mathcal{G}$, we have

$$\begin{aligned} u^\circ(p;gh) &= \limsup_{q\to p,\lambda\downarrow 0} \frac{u(q\delta_\lambda(gh)) - u(q)}{\lambda} \\ &= \limsup_{q\to p,\lambda\downarrow 0} \left(\frac{u(q\delta_\lambda(gh)) - u(q\delta_\lambda(g))}{\lambda} + \frac{u(q\delta_\lambda(g)) - u(q)}{\lambda} \right) \\ &\le \limsup_{q\to p,\lambda\downarrow 0} \frac{u(q\delta_\lambda(g)\delta_\lambda(h)) - u(q\delta_\lambda(g))}{\lambda} + \limsup_{q\to p,\lambda\downarrow 0} \frac{u(q\delta_\lambda(g)) - u(q)}{\lambda} \\ &\le u^\circ(p;h) + u^\circ(p;g). \end{aligned}$$

The last line follows from the fact that if $\lambda \downarrow 0$ and $q \to p$, then $q\delta_\lambda(g) \to p$. Also, for each $\mu > 0$ we have

$$\begin{aligned} u^\circ(p;\delta_\mu(g)) &= \limsup_{q\to p,\lambda\downarrow 0} \frac{u(q\delta_\lambda(\delta_\mu(g))) - u(q)}{\lambda} \\ &= \limsup_{q\to p,\lambda\downarrow 0} \frac{u(q\delta_{(\lambda\cdot\mu)}(g)) - u(q)}{\lambda} \\ &= \mu \cdot \limsup_{q\to p,\lambda\downarrow 0} \frac{u(q\delta_{(\lambda\cdot\mu)}(g)) - u(q)}{\lambda\cdot\mu} \\ &= \mu \cdot u^\circ(p;g). \end{aligned}$$

Since u is K-Lipschitz on a neighborhood of p and

$$d_{CC}(q\delta_\lambda(g), q) = ||q^{-1}q\delta_\lambda(g)||_{CC} = \lambda||g||_{CC}$$

we get that $u^\circ(p; g) \leq K||g||_{CC}$. To show item (*2*), let $p_n \to p$ and $g_n \to g$. For each n, let $0 < \lambda_n < 2^{-n}$ and $q_n \in \mathcal{G}$ with $d_{CC}(q_n, p_n) < 2^{-n}$ so that

$$u^\circ(p_n; g_n) \leq 2^{-n} + \frac{u(q_n\delta_{\lambda_n}(g_n)) - u(q_n)}{\lambda_n}.$$

Now $q_n \to p$ and $\lambda_n \downarrow 0$, hence

$$\begin{aligned} u^\circ(p; g) &\geq \limsup_{n\to\infty} \frac{u(q_n\delta_{\lambda_n}(g)) - u(q_n)}{\lambda_n} \\ &= \limsup_{n\to\infty} \frac{u(q_n\delta_{\lambda_n}(g_n)) - u(q_n)}{\lambda_n} - \frac{u(q_n\delta_{\lambda_n}(g_n)) - u(q_n\delta_{\lambda_n}(g))}{\lambda_n} \\ &\geq \limsup_{n\to\infty} u^\circ(p_n; g_n) - 2^{-n} - Kd_{CC}(g_n, g) \\ &= \limsup_{n\to\infty} u^\circ(p_n; g_n) \end{aligned}$$

as needed, completing the proof. □

Analogous to the definitions of non-smooth analysis in Banach spaces (see Clarke [3, Chapter 2]), we define the generalized derivative of a real valued locally Lipschitz function as follows.

Definition 3.8. Let u be a real valued Lipschitz function defined on a neighborhood of p. We call $\zeta \in \mathcal{L}(\mathcal{G}; \mathbb{R})$ a *generalized derivative of* u *at* p if for each $h \in \mathcal{G}$,

$$\zeta(h) \leq u^\circ(p; h)$$

and we write $\zeta \in \partial u(p)$.

Remark 3.9. From the definition we conclude the following.

(1) If u is locally Lipschitz, then for each constant k, $(u + k)^\circ(p; g) = u^\circ(p; g)$. Hence, for for each constant k we have $\partial(u + k)(p) = \partial u(p)$.

(2) Let u and v be Lipschitz functions on a neighborhood of $p \in \mathcal{G}$. If for each $g \in \mathcal{G}$, $u^\circ(p; g) \leq v^\circ(p; g)$, then $\partial u(p) \subseteq \partial v(p)$.

(3) If u is a real valued K-Lipschitz function defined on a neighborhood of p and $\zeta \in \partial u(p)$, then $||\zeta||_{\mathcal{L}(\mathcal{G};\mathbb{R})} \leq K$. Indeed, for each $g \in \mathcal{G}$,

$$|\zeta(g)| \leq |u^\circ(p; g)| \leq \limsup_{q\to p, \lambda\downarrow 0} \frac{|u(q\delta_\lambda(g)) - u(q)|}{\lambda} \leq K||g||_{\mathcal{L}(\mathcal{G};\mathbb{R})}.$$

(4) If $\alpha > 0$, then $(\alpha u)^\circ(p; g) = \alpha u^\circ(p; g)$ whenever u is locally Lipschitz. Hence, for $\alpha > 0$ we have $\partial(\alpha u)(p) = \alpha\partial u(p)$. Additionally, a direct computation yields that $u^\circ(p; g) = (-u)^\circ(p; g^{-1})$ which together with the above implies that for each $\alpha \in \mathbb{R}$, $\partial(\alpha u)(p) = \alpha\partial u(p)$.

Proposition 3.10. *Let u be a real valued K-Lipschitz function defined on a neighborhood of $p \in \mathcal{G}$. If $p_n \to p$ and $\zeta_n \to \zeta$ weak* so that $\zeta_n \in \partial u(p_n)$ for each n, then $\zeta \in \partial u(p)$. In particular, $\partial u(p)$ is a weak* closed subset of V_1^*.*

Proof. Fix $h \in \mathcal{G}$. We then have that for each n,

$$\zeta_n(h) \leq u^\circ(p_n; h).$$

Proposition 3.7 then gives us that

$$\zeta(h) = \limsup_{n\to\infty} \zeta_n(h) \leq \limsup_{n\to\infty} u^\circ(p_n; h) \leq u^\circ(p; h)$$

as needed. □

Proposition 3.1 shows us that for $\zeta \in \mathcal{L}(\mathcal{G}; \mathbb{R})$, $\zeta(g) = \zeta(\pi_1(g))$ for all $g \in \mathcal{G}$. Hence, if $\zeta \in \partial u(p)$, then for each $g \in \mathcal{G}$,

$$\zeta(g) \leq \inf_{h\in\pi_1^{-1}(0)} u^\circ(p; gh).$$

Our next proposition when coupled with the Hahn-Banach theorem shows us that $\partial u(p)$ is never empty whenever u is Lipschitz on a neighborhood of p.

Theorem 3.11. *Let $\rho : \mathcal{G} \to \mathbb{R}$ be a bounded signed gauge with constant K. Then the function*

$$f(v) := \inf_{v_2\in V_2,\ldots,v_n\in V_n} \rho(v, v_2, \ldots, v_n)$$

is a Minkowski functional on V_1 with $f(v) \leq K||v||_{CC}$, where $||\cdot||_{CC}$ is the gauge on $\mathcal{G}$. In particular by Remark 2.2(D)*, there is a constant C independent of ρ so that for each $v_1 \in V_1$, $f(v_1) \leq CK||v_1||$ where $||\cdot||_1$ is the norm of V_1. Moreover, f satisfies the relation*

$$f(v) = \rho(v, 0, \ldots, 0).$$

Proof. We first do the proof when $\mathcal{G}$ has step 2. The step n case we outline at the end of the proof.

Once we have shown that f is finite valued, it is a straightforward calculation to check that f is a Minkowski functional on V_1. For each $g \in V_2$, define ρ_g as $\rho_g(h) = \rho(gh)$. We note that

$$|\rho_g(h) - \rho_g(k)| = |\rho(gh) - \rho(gk)| \leq K d_{CC}(gh, gk) = K d_{CC}(g, h)$$

i.e., for each g, ρ_g is K-Lipschitz. Hence, $f = \inf_{g\in V_2} \rho_g$ which is an infimum of K-Lipschitz functions. Thus to show that f is finite valued it suffices to show that $f(0)$ is not $-\infty$. Indeed, let $\theta(t) = \rho(0, t)$ and note that $f(0) = \inf_{V_2} \theta$. We will now show that $\theta \geq 0$ on V_2. Since ρ is a bounded signed gauge we have

$$\theta(2t) = \rho(0, 2t) = \sqrt{2}\rho(0, t) = \theta(t)\sqrt{2}$$

on the other hand

$$\theta(2t) = \rho(0, 2t) = \rho((0, t) \cdot (0, t)) \leq 2\rho(0, t) = 2\theta(t)$$

which implies that $\theta(t)\sqrt{2} \leq 2\theta(t)$. Hence, $\theta(t) \geq 0$ for all $t \in V_2$ which implies that f is a finite valued function.

We will now show that $f(z) = \rho(z, 0)$ for each $z \in V_1$. For each $z \in V_1$ and $\epsilon > 0$ define the non-empty closed set

$$S_z^\epsilon := \{t \in V_2 \mid \rho(z, t) \leq \epsilon + f(z)\}.$$

To finish, we need to show that $0 \in \cap_{\epsilon>0} S_z^\epsilon$. Since ρ is a bounded signed gauge, we have for each $\lambda > 0$,

$$\lambda\rho(z, t) = \rho(\lambda z, \lambda^2 t).$$

Hence, $t \in S_z^\epsilon$ if and only if for each $\lambda > 0$, $\lambda^2 t \in S_{\lambda z}^{\lambda\epsilon}$. Let $z \in V_1$ and $t \in S_z^\epsilon$. Since ρ is a bounded signed gauge, using Remark 2.2(A) gives us that for each $m \in \mathbb{N}$,

$$\rho(mz, mt) = \rho((z, t)^m) \leq m\rho(z, t) \leq m(\epsilon + f(z)) = f(mz) + m\epsilon.$$

I.e., if $t \in S_z^\epsilon$, then $mt \in S_{mz}^{m\epsilon}$. Combining with the fact that for each $w \in V_1$, $s \in S_w^\epsilon$ implies that $\lambda^2 s \in S_{\lambda w}^{\lambda\epsilon}$ and letting $\lambda = 1/m$ we see that if $t \in S_z^\epsilon$, then for each $m \in \mathbb{N}$, $t/m \in S_z^\epsilon$. However, S_z^ϵ is a non-empty closed set, thus $0 \in S_z^\epsilon$ for each $\epsilon > 0$, which completes the proof for the step 2 case.

For the general case note that by Remark 2.2(A) and (B) that $f(v) = \inf_{g\in\pi_1^{-1}(0)} \rho_g(v)$ is an infimum of K-Lipschitz functions. The first part is to show that $f(0)$ is not $-\infty$. First consider the function on V_n as $\theta_n(v_n) = \rho(v_n)$. Just as in the proof of the step 2 case, one can use the fact that ρ is sub-additive and that it commutes with dilations to conclude that $\inf_{V_n} \theta_n \geq 0$ whenever $n > 1$. Now define the function θ_{n-1} on $\mathcal{G}$ as

$$\theta_{n-1}(g) = \inf_{v_n\in V_n} \rho_{v_n}(g).$$

Since the above is an infimum of Lipschitz functions, and $\inf_{V_n} \theta_n \geq 0$, the above function is again a finite valued Lipschitz function. Moreover because of the nature of the group structure (i.e., items(A) and (B) of Remark 2.2) we see that θ_{n-1} is also a bounded signed gauge. If $n - 1 > 1$ then the same proof as in the step 2 case will imply that $\inf_{V_{n-1}} \theta_{n-1} \geq 0$. Continue this process until one reaches V_1, where the produced function will then be θ_1. It is straightforward to check that for each i, θ_i is a bounded signed gauge and $\theta_i(g) \leq K d_{CC}(0, g)$. Also the group structure on $\mathcal{G}$ insures that $\theta_1(v) = \theta_1(\pi_1(v))$. Hence, for $v_1 \in V_1$, $f(v_1) = \theta_1(v_1)$. The fact that θ_1 is sub-additive and commutes with dilations implies that f is a Minkowski functional defined on V_1. The fact that $f(v) = \rho(v, 0, \ldots, 0)$ follows from an analogues argument as in the step 2 case. □

An immediate consequence of the above proposition is the following

Corollary 3.12. *Let u be a real valued K-Lipschitz function defined on a neighborhood of $p \in \mathcal{G}$. Then $\zeta \in \partial u(p)$ if and only if for each $g \in V_1$, $\zeta(g) \leq u^\circ(p; g)$.*

Corollary 3.13. *$(\mathcal{L}(\mathcal{G}; \mathbb{R}), ||\cdot||_{\mathcal{L}(\mathcal{G};\mathbb{R})})$ is homeomorphicily isomorphic to V_1^* via the map*

$$\iota(\phi) = \phi \circ \pi_1.$$

In particular, there exists a constant $C > 0$ so that for each $r > 0$,

$$\overline{\mathbf{B}}_{V_1^*}(0, \frac{r}{C}) \subseteq \overline{\mathbf{B}}_{\mathcal{L}(\mathcal{G};\mathbb{R})}(0, r) \subseteq \overline{\mathbf{B}}_{V_1^*}(0, Cr)$$

where $\overline{\mathbf{B}}_{\mathcal{L}(\mathcal{G};\mathbb{R})}(0, r)$ is the closed ball of radius r under the norm $||\cdot||_{\mathcal{L}(\mathcal{G};\mathbb{R})}$ and $\overline{\mathbf{B}}_{V_1^*}(0, r)$ is the close ball of radius r under the norm $||\cdot||_1^*$ which is the norm in V_1^* produced by the norm on V_1, $||\cdot||_1$.

Proof. Since $||\cdot||_{CC}$ is a gauge, Theorem 3.11 implies that for each $(v_1, v_2, \ldots, v_n) \in \mathcal{G}$,

$$||(v_1, v_2, \ldots, v_n)||_{CC} \geq ||(v_1, 0, \ldots, 0)||_{CC}.$$

Now, for each $\zeta \in \mathcal{L}(\mathcal{G}; \mathbb{R})$,

$$\zeta(v) = \zeta(\pi_1(v)).$$

Hence for each $\zeta \in \mathcal{L}(\mathcal{G}; \mathbb{R})$ we see that

$$||\zeta||_{\mathcal{L}(\mathcal{G};\mathbb{R})} = \sup_{v \in V_1, ||v||_{CC} \leq 1} |\zeta(v_1)|.$$

By Remark 2.2(D), for $v \in V_1$ we have that $||v||_{CC}$ is comparable to $||v||_1$, thus $||\zeta||_{\mathcal{L}(\mathcal{G};\mathbb{R})}$ is comparable to

$$||\zeta||_{\mathcal{L}(\mathcal{G};\mathbb{R})'} = \sup_{v \in V_1, ||v||_1 \leq 1} |\zeta(v)|.$$

Proposition 3.1 shows us that each element of $\mathcal{L}(\mathcal{G}; \mathbb{R})$ is really just an element of V_1^*. Combined with the above we then conclude that $\mathcal{L}(\mathcal{G}; \mathbb{R})$ with $||\cdot||_{\mathcal{L}(\mathcal{G};\mathbb{R})'}$ is homeomorphicly isomorphic to V_1^* equipped with the operator norm. □

By using Theorem 3.11 and the Hahn-Banach theorem we obtain the following Corollary:

Corollary 3.14. *Let u be a real valued K-Lipschitz function defined on a neighborhood of $p \in \mathcal{G}$. Then $\partial u(p)$ is a non-empty weak* closed convex bounded subset of V_1^*. Moreover, for each $v \in V_1$,*

$$u^\circ(p; v) = \sup_{\zeta \in \partial u(p)} \zeta(v).$$

In particular, for each pair of real valued Lipschitz functions defined on a neighborhood of p, $\partial u(p) \subseteq \partial v(p)$ if and only if for each $g \in V_1$, $u^\circ(p; g) \leq v^\circ(p; g)$.

Proof. Proposition 3.10 gives us that $\partial u(p)$ is a weak* closed subset of V_1^*. The definition of $\partial u(p)$ guarantees that it is convex. Moreover, since u is K-Lipschitz near p, we see that for each $g \in \mathcal{G}$, $u^\circ(p; g) \leq K||g||_{CC}$ which implies that for all $\phi \in \partial u(p)$ and $g \in \mathcal{G}$, $|\phi(g)| \leq K||g||_{CC}$, i.e., $||\phi||_{\mathcal{L}(\mathcal{G};\mathbb{R})} \leq K$ which with Corollary 3.13, implies that $\partial u(p)$ is a bounded subset of V_1^*. Now, by Theorem 3.11 the function $\rho(v) = u^\circ(p; v)$ is a Minkowski functional on V_1. Moreover since u is K-Lipschitz we have by Remark 2.2(D) that for each $v \in V_1$

$$|\rho(v)| \leq K||v||_{CC} \leq CK||v||_1.$$

Given $v \in V_1$ define the linear function ϕ on $\text{span}_{\mathbb{R}}(v)$ as $\phi(\lambda v) = \lambda\rho(v)$. Since ρ is a Minkowski functional, $\phi \leq \rho$ on $\text{span}_{\mathbb{R}}(v)$. We now employ the Hahn-Banach theorem to extend ϕ to all of V_1 so that $\phi \leq \rho$ on V_1. Since $|\rho(v)| \leq CK||v||_1$, we have that ϕ is continuous, in particular, $\phi \in V_1^*$. Thus $\phi \circ \pi_1 \in \mathcal{L}(\mathcal{G}; \mathbb{R})$ and for $g \in V_1$, $\phi \circ \pi_1(g) \leq u^\circ(p; g)$ which implies that $\phi \circ \pi_1 \in \partial u(p)$. Moreover, $\phi \circ \pi_1(v) = \phi(v) = u^\circ(p; v)$. Hence for all $v \in V_1$ we have

$$u^\circ(p; v) \leq \sup_{\zeta \in \partial u(p)} \zeta(v).$$

That

$$u^\circ(p; z) \geq \sup_{\zeta \in \partial u(p)} \zeta(z)$$

follows from the definition of $\partial u(p)$. □

Corollary 3.15. *Let u and v be Lipschitz functions defined on a neighborhood of p. We then have that*

$$\partial(u+v)(p) \subseteq \partial u(p) + \partial v(p).$$

Proof. Suppose there exists $\zeta \in \partial(u+v)(p)$ which is not an element of the weak* closed convex set $\partial u(p) + \partial v(p)$. Since $\partial u(p)$ and $\partial v(p)$ are convex weak* closed subsets of V_1^*, a standard separation theorem gives us a $c \in \mathbb{R}$ and a $z \in V_1$ so that

$$\zeta(z) > c \geq \zeta_1(z) + \zeta_2(z)$$

whenever $\zeta_1 \in \partial u(p)$ and $\zeta_2 \in \partial v(p)$. The definition of $\partial(u+v)(p)$ then implies that

$$(u+v)^\circ(p; z) > c \geq \zeta_1(z) + \zeta_2(z)$$

whenever $\zeta_1 \in \partial u(p)$ and $\zeta_2 \in \partial v(p)$. Easily, $(u+v)^\circ(p;h) \leq u^\circ(p;h) + v^\circ(p;h)$. Hence we have

$$\zeta_1(z) - u^\circ(p;z) + \zeta_2(z) - v^\circ(p;z) < 0$$

for each $\zeta_1 \in \partial u(p)$ and $\zeta_2 \in \partial v(p)$. Corollary 3.14 (using the fact that both $\partial u(p)$ and $\partial v(p)$ are weak* closed bounded convex subsets of V_1^*) implies there exists $\zeta_1 \in \partial u(p)$ and $\zeta_2 \in \partial v(p)$ so that $\zeta_1(z) = u^\circ(p;z)$ and $\zeta_2(z) = v^\circ(p;z)$. Hence the above implies that $0 < 0$, a contradiction. □

Corollary 3.16. *Let $u \in C^1_{sub}$ be a real valued locally Lipschitz function defined on a neighborhood of $p \in \mathcal{G}$. Then $\partial u(p) = \{D_H u(p) \circ \pi_1\}$ where $\pi_1 : \mathcal{G} \to V_1$ is projection onto V_1. Moreover, for each locally Lipschitz real valued function v defined on a neighborhood of p, $\partial(u+v)(p) = \partial u(p) + \partial v(p)$.*

Proof. By Proposition 3.6, we see that for $g \in V_1$, $u^\circ(p;g) = D_H u(g)$. Thus, $D_H u(p) \circ \pi_1 \in \partial u(p)$. If $\phi \in \partial u(p)$ we then have that $\phi \in V_1^*$ and $\phi \leq D_H u(p)$. Since both ϕ and $D_H(p)$ are linear, we see that $\phi = D_H u(p)$. By Corollary 3.15 and that both $\partial u(p)$ and $\partial(-u)(p)$ are singletons, we see that $\partial(u+v)(p) = \partial u(p) + \partial v(p)$. □

One interesting corollary of the proof of Theorem 3.11 is that one can create a simpler Banach homogenous group from a given one:

Corollary 3.17. *Let $\mathcal{G} = V_1 \times V_2 \times \ldots \times V_n$ be an n-step Banach homogenous group with $n \geq 2$ under the gauge $\rho_{\mathcal{G}}$. Then $\mathcal{H} = V_1 \times V_2 \times \ldots \times V_{n-1}$ is an $(n-1)$-step Banach homogenous group under the gauge $\rho_{\mathcal{H}}$:*

$$\rho_{\mathcal{H}}(v_1, \ldots, v_{n-1}) = \rho_{\mathcal{G}}(v_1, \ldots, v_{n-1}, 0).$$

Proof. The proof of Theorem 3.11 shows us that $\rho_{\mathcal{H}}$ is a bounded signed gauge on $\mathcal{G}$ since

$$\begin{aligned} \rho_{\mathcal{G}}(v_1, \ldots, v_{n-1}, 0) &= \inf_{v_n \in V_n} \rho_{\mathcal{G}}(v_1, \ldots, v_{n-1}, v_n) \\ &= \inf_{v_n \in V_n} \rho_{\mathcal{G}}((v_1, \ldots, v_{n-1}, 0) \cdot (0, \ldots, 0, v_n)). \end{aligned}$$

□

A quick calculation shows that if p is a local minimum or maximum of a locally Lipschitz function u, then $u^\circ(p;h) \geq 0$ for each $h \in \mathcal{G}$. Hence we have the following proposition:

Proposition 3.18. *If $p \in \mathcal{G}$ is a local minimum or local maximum of a locally Lipschitz function u, then $0 \in \partial u(p)$.*

The next proposition is useful in the proof of our main theorem, Theorem 1.1.

Proposition 3.19. *Let $u : \mathcal{G} \to (-\infty, \infty]$ be a lower semi-continuous function which is bounded below and not identically equal to infinity. If $p_0 \in \mathcal{G}$ and $\epsilon > 0$ so that*

$$u(p_0) < \epsilon + \inf_{\mathcal{G}} u,$$

then there exists a $p_\epsilon \in \mathcal{G}$ so that the function

$$u_\epsilon(x) = u(x) + \sqrt{\epsilon} d_{CC}(x, p_\epsilon)$$

achieves its infimum uniquely at $x = p_\epsilon$ and $d_{CC}(p_\epsilon, x) < \sqrt{\epsilon}$. In particular (by Corollary 3.13*, Remark* 3.9(3) *and Corollary* 3.15*), there exists a constant $C > 0$ so that if u is Lipschitz on a neighborhood of p_ϵ, then for x near p_ϵ,*

$$\partial u_\epsilon(x) \subseteq \partial u(x) + \mathbf{B}_{V_1^*}(0, C\sqrt{\epsilon})$$

and in particular,

$$0 \in \partial u(p_\epsilon) + \mathbf{B}_{V_1^*}(0, C\sqrt{\epsilon})$$

where $\overline{\mathbf{B}}_{V_1^*}(0, r)$ is the closed ball of radius r in V_1^* .

Proof. This proposition is actually just an application of the following minimization principle, [4, Theorem 1.1].

Theorem 3.20. *Let X be a complete metric space with metric d_X and $f : X \to (-\infty, \infty]$ be a lower semi-continuous function which is not identically equal to infinity and bounded from below. If $x_0 \in X$ and $\epsilon > 0$ so that $f(x_0) < \epsilon + \inf_X f$, then for each $\lambda > 0$, there exist $x_\lambda \in X$ with $d_X(x_0, x_\lambda) < \lambda$ so that*

(1) $f(x_\lambda) \leq f(x_0)$ and
(2) for each $x \neq x_\lambda$, $f(x) > f(x_\lambda) - \frac{\epsilon}{\lambda} d_X(x, x_\lambda)$,

i.e., the function

$$f_\lambda(x) = f(x) + \frac{\epsilon}{\lambda} d_X(x, x_\lambda)$$

attains its infimum at exactly one point, x_λ.

We prove the Proposition by using the above theorem and letting $\lambda = \sqrt{\epsilon}$ and noting that for each fixed $y \in \mathcal{G}$ the function $x \to d_{CC}(x, y)$ is 1-Lipschitz. □

4. Main Results

Equipped with the material of Sections 2 and 3, we can now prove our main theorem, Theorem 1.1 and also prove analogues of a mean value theorem and a chain rule in the context of non-smooth analysis on Banach homogenous groups. It is worth noting that these proofs are almost straight adaptions of the proofs of the analogous theorems in Banach spaces. The key results needed for these proofs to work in Banach homogenous group are Theorem 3.11 and its corollaries.

Our first order of business is to prove Theorem 1.1: Let $\mathcal{G}$ be a Banach homogenous group and W an open subset of $\mathcal{G}$. Let $u : W \to \mathbb{R}$, $v : W \to \mathbb{R}^l$ and $w : W \to \mathbb{R}^k$ all be locally Lipschitz. Let

$$S := \{p \in W \mid v(p) \leq \mathbf{0}, w(p) = \mathbf{0}\}$$

where we write $\mathbf{a} = (a_1, a_2, \ldots, a_l) \leq \mathbf{0}$ if and only if for each j, $a_j \leq 0$. If $p_0 \in S$ so that $u(p_0) = \inf_S u$, then there exists $\alpha = (\lambda_0, \gamma, \lambda) \in \mathbb{R} \times \mathbb{R}^l \times \mathbb{R}^m$ with

(1) $\alpha \neq \mathbf{0}$,
(2) $\lambda_0 \in \{0, 1\}$,
(3) $\gamma \geq \mathbf{0}$ and
(4) $\langle v(p_0), \gamma \rangle = 0$

so that

$$0 \in \partial(\lambda_0 u + \langle \gamma, v \rangle + \langle \lambda, w \rangle)(p_0).$$

Proof of Theorem 1.1. Our proof closely mirrors the proof of the Fritz John necessary conditions found in [3, pp. 100–101]. Since u,v and w are locally Lipschitz there exists $r > 0$ so that u,v and w are Lipschitz on $V = \{x \in \mathcal{G} \mid d_{CC}(x, p_0) \leq r\}$. Define the compact set

$$A := \{\alpha = (\lambda_0, \gamma, \lambda) \in \mathbb{R} \times \mathbb{R}^l \times \mathbb{R}^m \mid \lambda_0 \geq 0, \gamma \geq \mathbf{0}, ||\alpha|| = 1\}$$

where $|| \cdot ||$ is the Euclidean norm on $\mathbb{R}^{1+l+m}$.

For each $\epsilon > 0$ define the maps $U_\epsilon : V \to \mathbb{R}^{1+m+l}$ and $\theta_\epsilon : V \to \mathbb{R}$ by

$$U_\epsilon(p) := (u(p) - u(p_0) + \epsilon, v(p), w(p)),$$
$$\theta_\epsilon(p) := \sup_{\alpha \in A} \langle \alpha, U_\epsilon(p) \rangle .$$

Because u, v and w are Lipschitz on V, there exists $K > 0$ so that U_ϵ and θ_ϵ are K-Lipschitz on V.

We first claim that $\theta_\epsilon \geq 0$ on V. Indeed, let $p \in V$. If $p \in S$, then $u(p_0) \leq u(p)$, hence

$$\begin{aligned} \theta_\epsilon(p) &\geq \langle (1, \mathbf{0}, \mathbf{0}), U_\epsilon(p) \rangle \\ &= u(p) - u(p_0) + \epsilon \geq \epsilon. \end{aligned}$$

For $p \notin S$, either $w(p) \neq \mathbf{0}$ or there exists a j so that $v_j(p) > 0$. Let $\lambda = w(p)$ and $\gamma_i = v_i(p)$ if $v_i(p) > 0$ and $\gamma_i = 0$ if $v_i(p) \leq 0$. Then the vector $\beta = (0, \gamma, \lambda)$ is non-zero and $\alpha = \frac{1}{||\beta||}\beta$ is an element of the set A. Thus

$$\theta_\epsilon(p) \geq \langle \alpha, U_\epsilon(p) \rangle = ||\beta|| > 0$$

as needed. Hence we have shown that on V, $\theta_\epsilon > 0$. That $\theta_\epsilon(p_0) = \epsilon$ follows from the fact that u restricted to S is minimized at p_0. Extend the function θ_ϵ as infinity outside of V. Because V is closed, we see that $\theta_\epsilon : \mathcal{G} \to (-\infty, \infty]$ is lower semi-continuous. We now invoke Proposition 3.19 to conclude that there exists a $p_\epsilon \in \mathbf{B}_{CC}(p_0, 2\sqrt{\epsilon})$ so that the function

$$\Psi_\epsilon(p) = \theta_\epsilon(p) + 2\sqrt{\epsilon} d_{CC}(p, p_\epsilon)$$

attains its unique minimum at $p = p_\epsilon$. Since V contains an open neighborhood of p_0, for ϵ small enough, $p_\epsilon \in \text{int}(V)$. Hence,

$$0 \in \partial\theta_\epsilon(p_\epsilon) + \overline{\mathbf{B}}_{V_1^*}(0, 2C\sqrt{\epsilon})$$

where C is as in Corollary 3.13 and $\overline{\mathbf{B}}_{V_1^*}(0, r)$ is the closed ball of radius r in V_1^*. The next lemma will give us a better handle on the set $\partial\theta_\epsilon(p_\epsilon)$.

Lemma 4.1. *Let U be an open subset of $\mathcal{G}$ and $F : U \to \mathbb{R}^N$ be locally Lipschitz and $B \subset \mathbb{R}^N$ be compact. Define the locally Lipschitz function $G(p) := \sup_{\beta\in B} \langle F(p), \beta\rangle$. If there exists a unique $\beta \in B$ so that $G(q_0) = \langle \beta, F(q_0)\rangle$, then*

$$\partial G(q_0) \subseteq \partial \langle F(\cdot), \beta\rangle (q_0).$$

Proof. By Remark 3.9(2), it suffices to show that for each $g \in \mathcal{G}$,

$$G^\circ(q_0; g) \leq \langle F(\cdot), \beta\rangle^\circ (q_0; g).$$

Let $q_n \to q_0$ and $\lambda_n \downarrow 0$ so that

$$G^\circ(q_0; g) = \lim_{n\to\infty} \frac{G(q_n\delta_{\lambda_n}(g)) - G(q_n)}{\lambda_n}.$$

For each n, since B is compact there exists $\beta_n \in B$ so that

$$G(q_n\delta_{\lambda_n}(g)) = \langle \beta_n, F(q_n\delta_{\lambda_n}(g))\rangle .$$

Since B is compact, we can find a convergent subsequence of $\{\beta_n\}$ that converges to an element $\alpha \in B$. Passing to that subsequence (without relabelling) we then have

$$\begin{aligned} G^\circ(q_0; g) &= \lim_{n\to\infty} \frac{G(q_n\delta_{\lambda_n}(g)) - G(q_n)}{\lambda_n} \\ &\leq \limsup_{n\to\infty} \frac{\langle \beta_n, F(q_n\delta_{\lambda_n}(g)) - F(q_n)\rangle}{\lambda_n} \\ &\leq \limsup_{n\to\infty} \left(\frac{\langle \alpha, F(q_n\delta_{\lambda_n}(g)) - F(q_n)\rangle}{\lambda_n} + K||\beta_n - \alpha|| \right) \end{aligned}$$

where F is K-Lipschitz on a neighborhood of q_0. Hence we have

$$G^\circ(q_0; g) \leq \limsup_{n\to\infty} \frac{\langle \alpha, F(q_n\delta_{\lambda_n}(g)) - F(q_n)\rangle}{\lambda_n} \leq \langle F(\cdot), \alpha\rangle^\circ (q_0; g).$$

We now claim that $\alpha = \beta$. Indeed we have

$$\begin{aligned} G(q_0) &= \lim_{n\to\infty} G(q_n\delta_{\lambda_n}(g)) \\ &= \lim_{n\to\infty} \langle \beta_n, F(q_n\delta_{\lambda_n}(g))\rangle \\ &= \langle \alpha, F(q_0)\rangle , \end{aligned}$$

but β was the unique point in B so that $G(q_0) = \langle \beta, F(q_0)\rangle$. Hence $\alpha = \beta$ completing the proof to the lemma. □

With the above lemma in mind, we are now going to show that there is a unique $\alpha_\epsilon \in A$ so that $\theta_\epsilon(p_\epsilon) = \langle \alpha_\epsilon, U_\epsilon(p_\epsilon)\rangle$. If $p_\epsilon \in S$, then $\alpha_\epsilon = (1, \mathbf{0}, \mathbf{0})$ is that unique element. For $p_\epsilon \notin S$, let

$$\begin{aligned}\lambda_0 &= \max(0, u(p_\epsilon) - u(p) + \epsilon)\\ \gamma_i &= \max(0, v_i(p_\epsilon))\\ \lambda &= w(p_\epsilon),\end{aligned}$$

$\beta = (\lambda_0, \gamma, \lambda)$ and $\alpha = \frac{1}{||\beta||}\beta$. Then α is the unique element of A that does the job. Note that in each of these two cases, $\langle \gamma, v(p_\epsilon)\rangle \geq 0$. Putting things together, we have for each $\epsilon > 0$ an $\alpha_\epsilon \in A$ so that

$$0 \in \partial \left\langle \alpha_\epsilon, U_\epsilon(\cdot)\right\rangle (p_\epsilon) + \overline{\mathbf{B}}_{V_1^*}(0, 2C\sqrt{\epsilon}).$$

Let $\epsilon_i \downarrow 0$ and for each i let $\alpha_i = \alpha_{\epsilon_i}$, $f_i = f_{\epsilon_i}$ and $p_i = p_{\epsilon_i}$. By passing to a subsequence, there exists $\alpha \in A$ so that $\alpha_i \to \alpha$. Note that

$$\langle \alpha_i, U_i\rangle = \langle \alpha, U_i\rangle + \langle \alpha_i - \alpha, U_i\rangle$$

which with Remark 3.9(1) and the previous implies that

$$0 \in \partial(\langle \alpha, U_i\rangle)(p_i) + \overline{\mathbf{B}}_{V_1^*}(0, 2C\sqrt{\epsilon_i})$$

Let $U = (u, v, w)$. Since $U_i - U$ is a constant function, we then have

$$0 \in \partial(\langle \alpha, U\rangle)(p_i) + \overline{\mathbf{B}}_{V_1^*}(0, 2C\sqrt{\epsilon_i})$$

for each i. Letting $||\cdot||_1^*$ denote the norm on V_1^* produced by $||\cdot||_1$, we see that for each i, there exists $\zeta_i \in \partial(\langle \alpha, U\rangle)(p_i)$ and $\omega_i \in V_1^*$ with $||\omega_i||_1^* \leq 2C\sqrt{\epsilon_i}$ so that $0 = \zeta_i + \omega_i$. Because $\epsilon_i \to 0$, we conclude that $||\zeta_i||_1^* \to 0$. In particular, $\zeta_i \to 0$ weak* in V_1^*. Applying Proposition 3.10 gives us that

$$0 \in \partial(\langle \alpha, U\rangle)(p_0)$$

with $\alpha = (\lambda_0, \gamma, \lambda) \in A$. Since $\alpha \in A$, $\lambda_0 \geq 0$ and $\gamma \geq \mathbf{0}$. Also for each i, we have that $\langle v(p_i), \gamma_i\rangle \geq 0$. Since v is continuous and $\gamma_i \to \gamma$ we have that $\langle v(p_0), \gamma\rangle \geq 0$. Which together with $v(p_0) \leq \mathbf{0}$ implies that $\langle v(p_0), \gamma\rangle = 0$. If $\lambda_0 = 0$ then we are done, otherwise replace α by α/λ_0. Using Remark 3.9(4) gives

$$0 \in \partial(\lambda_0 u + \langle \gamma, v\rangle + \langle \lambda, w\rangle)(p_0)$$

with

(1) $(\lambda_0, \gamma, \lambda) \neq \mathbf{0}$,
(2) $\lambda_0 \in \{0, 1\}$,
(3) $\gamma \geq \mathbf{0}$ and
(4) $\langle v(x), \gamma\rangle = 0$

completing the proof to the theorem. □

Corollary 3.16 with the above theorem gives us the following Corollary which is an exact analogue of Lagrange multiplier in the setting of Carnot groups.

Corollary 4.2. *Let $u, h_1, \dots, h_m$ be real valued locally Lipschitz C^1_{sub} functions defined on a Carnot group G so that for each $p \in S$ where $S = \cap_{j=1}^m h_j^{-1}(0)$, the vectors $\{\nabla_H h_j(p)\}_{j=1}^m$ are linearly independent. If $p_0 \in S$ so that $u(p_0) = \inf_S u$, then there exists $\lambda \in \mathbb{R}^m$ so that*

$$0 = \nabla_H(u + \langle \lambda, h \rangle)(p_0)$$

where $h = (h_1, h_2, \dots, h_m)$.

The next theorem is a mean value theorem for Lipschitz functions on Banach homogenous groups.

Theorem 4.3. *Let $f : \mathcal{G} \to \mathbb{R}$ be locally Lipschitz. For each $p \in \mathcal{G}$ and $g \in V_1$ there exists $0 < t < 1$ and $\zeta \in \partial(f)(p\delta_t(g))$ so that*

$$f(pg) - f(p) = \zeta(g).$$

Proof. Define $\phi(t) = p \cdot (tg)$, and let $\psi = f \circ \phi$. Since g is horizontal we have that $\delta_t(g) = tg$ for $t > 0$ and $\delta_{-t}(g) = -tg$ for $t < 0$. We can see that $\psi : \mathbb{R} \to \mathbb{R}$ is a locally Lipschitz function. We first claim that

$$\partial\psi(t) \subseteq \{\zeta(g) \,|\, \zeta \in \partial(f)(\phi(t))\}.$$

Indeed, since both sets in question are closed, bounded, convex subsets of $\mathbb{R}$, it suffices to check that for $v = 1$ and $v = -1$ that

$$\sup\{\alpha v \mid \alpha \in \partial\psi(t)\} \le \sup\{v\zeta(g) \,|\, \zeta \in \partial(f)(\phi(t))\}.$$

Since g is horizontal, by Corollary 3.14 we have that

$$f^\circ(p\delta_t(g); vg) = \sup_{\zeta \in \partial(f)(p\delta_t(g))} \zeta(vg).$$

Moreover, since $\mathbb{R}$ is a step one Carnot group, the same Corollary gives us that

$$\psi^\circ(t; v) = \sup\{\alpha v \mid \alpha \in \partial\psi(t)\}.$$

Hence it suffices to check that for $v = 1, -1$ that

$$\psi^\circ(t; v) \le f^\circ(p\delta_t(g); vg).$$

Indeed we see that

$$\begin{aligned} \psi^\circ(t; v) &= \limsup_{s \to t, \lambda \downarrow 0} \frac{\psi(s + \lambda v) - \psi(s)}{\lambda} \\ &= \limsup_{s \to t, \lambda \downarrow 0} \frac{f(p\delta_{s+\lambda}(vg)) - f(p\delta_s(vg))}{\lambda} \\ &\le f^\circ(p\delta_t(g); vg) \end{aligned}$$

which completes the proof of the claim. To finish the proof of the theorem, define

$$\theta(t) = f(p\delta_t(g)) + t(f(p) - f(pg)) = \psi(t) - t(f(p) - f(pg)).$$

Since θ is continuous and $\theta(0) = \theta(1)$ there exists $0 < t < 1$ so that θ achieves a local maximum or a local minimum at t. Thus, $0 \in \partial\theta(t)$, which implies that

$$0 \in \partial\psi(t) + \{f(p) - f(pg)\}$$

which then gives us that

$$f(pg) - f(p) \in \partial\psi(t) \subseteq \{\zeta(g) \,|\, \zeta \in \partial(f)(p\delta_t(g))\},$$

completing our proof. □

Remark 4.4. Theorem 4.3 is false if g is not assumed to be horizontal. As an example let $\mathcal{G}$ be $\mathcal{H}_1$ and $f(z,t) = t$. Let $p = (\mathbf{0}, 0)$ and $g = (\mathbf{0}, 1)$. Then for every $\zeta \in \mathcal{L}(\mathcal{H}_1; \mathbb{R})$, $\zeta(g) = 0$ but $f(pg) - f(p) = 1$.

We close this section with a chain rule.

Theorem 4.5. *Let $\mathcal{G}$ be as in Theorem 4.3 and $F : \mathcal{G} \to \mathbb{R}^k$ be a locally Lipschitz map (with respect to the Carnot-Carathéodory metric on $\mathcal{G}$ and the Euclidean metric on $\mathbb{R}^k$) and $\psi : \mathbb{R}^k \to \mathbb{R}$ also be locally Lipschitz. Then we have that*

$$\partial(\psi \circ F)(x_0) \subseteq \overline{co}\{\zeta \,|\, \zeta \in \partial(\gamma \circ F)(x_0) \text{ for some } \gamma \in \partial(\psi)(F(x_0))\}$$

where $\overline{co}(S)$ denotes the closure of the convex hull of S.

Proof. Define the set

$$S = \{\zeta \,|\, \zeta \in \partial(\gamma \circ F)(x_0) \text{ for some } \gamma \in \partial(\psi)(F(x_0))\}.$$

We are to show that $\partial(\psi \circ F)(x_0) \subseteq \overline{\mathrm{co}}(S)$. Define a function $\rho : V_1 \to \mathbb{R}$ as

$$\rho(v) = \sup_{\zeta \in \overline{\mathrm{co}}(S)} \zeta(v).$$

Because S is bounded, ρ is finite valued. Easily, ρ is a Minkowski functional as well. We first claim that $(\psi \circ F)^\circ(x_0; v) \leq \rho(v)$ for all $v \in V_1$. Indeed, let $v_1 \in V_1$ and $x_i \to x_0$ and $\lambda_i \downarrow 0$ so that

$$(\psi \circ F)^\circ(x_0; v) = \lim_{i\to\infty} \frac{\psi(F(x_i\delta_{\lambda_i}(v))) - \psi(F(x_i))}{\lambda_i}.$$

Let $a_i = F(x_i\delta_{\lambda_i}(v))$ and $b_i = F(x_i)$. Applying Theorem 4.3 to the one step Carnot group $\mathbb{R}^k$ on the function ψ at the point a_i on the horizontal element $b_i - a_i$ produces a $0 < s_i < 1$ and a $\gamma_i \in \partial\psi(a_i + s_i(b_i - a_i))$ so that

$$\gamma_i(b_i - a_i) = \psi(b_i) - \psi(a_i).$$

Since F is continuous and ψ is locally Lipschitz, the γ_i are uniformly bounded in norm. Thus we can extract a subsequence of γ_i that converges to an element γ in the space $\mathcal{L}(\mathbb{R}^k; \mathbb{R})$ endowed with the operator norm. By Proposition 3.10 and the fact that both

a_i and b_i converge to $F(x_0)$ we get that $\gamma \in \partial(\psi)(F(x_0))$. Without bothering to relabel, we can conclude that

$$\begin{aligned}(\psi \circ F)^\circ(x_0; v) &= \lim_{i\to\infty} \gamma_i \left(\frac{F(x_i\delta_{\lambda_i}(v)) - F(x_i)}{\lambda_i} \right) \\ &= \lim_{i\to\infty} \gamma \left(\frac{F(x_i\delta_{\lambda_i}(v)) - F(x_i)}{\lambda_i} \right) \\ &= \lim_{i\to\infty} \frac{(\gamma \circ F)(x_i\delta_{\lambda_i}(v)) - (\gamma \circ F)(x_i)}{\lambda_i}.\end{aligned}$$

In the above we needed that F was Lipschitz near x_0 and that $||\gamma_i - \gamma|| \to 0$. We now apply Theorem 4.3 to the function $\gamma \circ F$ at the point x_i and the horizontal element $\delta_{\lambda_i}(v)$ to produce $0 < t_i < 1$ and $\zeta_i \in \partial(\gamma \circ F)(x_i\delta_{t_i\lambda_i}(v))$ so that

$$\frac{(\gamma \circ F)(x_i\delta_{\lambda_i}(v)) - (\gamma \circ F)(x_i)}{\lambda_i} = \zeta_i(v).$$

Taking advantage of the fact that $\gamma \circ F$ is locally Lipschitz gives us that the ζ_i's are uniformly bounded in V_1^*. Hence, there exists a subsequence which converges weak* to an element ζ. Since $x_i \to x_0$, Proposition 3.10 implies that $\zeta \in \partial(\gamma \circ F)(x_0)$. Eschewing relabelling, we then have that

$$(\psi \circ F)^\circ(x_0; v) = \lim_{i\to\infty} \zeta_i(v) = \zeta(v)$$

which implies that $(\psi \circ F)^\circ(x_0; v) \le \rho(v)$. To finish, let $\phi \in \partial(\psi \circ F)(x_0)$. Then by definition and combined with the previous gives us that $\phi(v) \le \sup_{\zeta \in \overline{\text{co}}(S)} \zeta(v)$ for each $v \in V_1$ which implies that $\phi \in \overline{\text{co}}(S)$. □

5. Closing Remarks

Our definition of a Banach homogenous group does not guarantee that the components of a Banach homogenous group are coupled. As a consequence functions can be C^1_{sub} but not be locally Lipschitz. Indeed, let $V_1 = \mathbb{R}$ and $V_2 = \mathbb{R}$. We let the group law be given by the usual vector addition of $\mathbb{R}^2$, i.e.,

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1 + x_2, y_1 + y_2).$$

We define the dilation by $\lambda > 0$ as

$$\delta_\lambda(x, y) = (\lambda x, \lambda^2 y).$$

Finally we define the gauge as

$$||(x, y)|| = |x| + \sqrt{|y|}.$$

It is straightforward to check that the above is sub-additive with respect to vector addition. However, the class of Lipschitz functions with respect to the metric created by our gauge is just those functions $f : \mathbb{R}^2 \to \mathbb{R}$ which are Lipschitz in the first argument and $\frac{1}{2}$-Hölder continuous in the second argument. The class of C^1_{sub} functions are those for which the partial derivative in the first argument exists at every point and is continuous. For example the function

$$f(x, y) = x + \chi_{\mathbb{Q}}(y),$$

where $\chi_{\mathbb{Q}}$ is the characteristic function of the rational numbers, is C^1_{sub}, but not even continuous.

References

[1] A. Bellaïche, J. Risler (eds.): Sub-Riemannian Geometry, Progress in Mathematics 144, Birkhäuser, Basel (1996).

[2] Z. M. Balogh, J. T. Tyson: Polar coordinates in Carnot groups, Math Z. 241(4) (2002) 697–730.

[3] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, P. R. Wolenski: Nonsmooth Analysis and Control Theory, Springer, New York (1998).

[4] I. Ekeland: On the variational principle, J. Math. Anal. Appl. 47 (1974) 324–353.

[5] G. B. Folland: Real Analysis: Modern Techniques and their Applications, John Wiley & Sons, New York (1984).

[6] G. B Folland, E. M. Stein: Hardy Spaces on Homogeneous Groups, Princeton University Press, Princeton (1982).

[7] B. Franchi, R. Serapioni, F. Serra Cassano: Rectifiability and perimeter in the Heisenberg group, Math. Ann. 321(3) (2001) 479–531.

[8] G. Lu, J. Manfredi, B. Stroffolini: Convex functions on the Heisenberg group, Calc. Var. Partial Differ. Equ. 19(1) (2004) 1–22.

[9] N. Garofalo, D.-M. Nhieu: Isoperimetric and sobolev inequalities for Carnot-Carathéodory spaces and the existence of minimal surfaces, Commun. Pure Appl. Math. 49(10) (1996) 1081–1144.

[10] N. Garofalo, D.-M. Nhieu: Lipschitz continuity, global smooth approximations and extension theorems for Sobolev functions in Carnot-Carathéodory spaces, J. Anal. Math. 74 (1998) 67–97.

[11] P. Pansu: Métriques de Carnot-Carathéodory et quasiisométries des espaces symétriques de rang un, Ann. Math. (2) 129 (1989) 1–60.

[12] F. Serra Cassano, B. Kirchheim: Rectifiability and parameterization of intrinsic regular surfaces in the Heisenberg group, preprint.

[13] N. Th. Varopoulos, L. Saloff-Coste, T. Coulhon: Analysis and Geometry on Groups, Cambridge University Press, Cambridge (1992).