ISSN 2077–9879

Eurasian Mathematical Journal

2013, Volume 4, Number 2

Founded in 2010 by
the L.N. Gumilyov Eurasian National University
in cooperation with
the M.V. Lomonosov Moscow State University
the Peoples' Friendship University of Russia
the University of Padua

Supported by the ISAAC
(International Society for Analysis, its Applications and Computation)

Published by

the L.N. Gumilyov Eurasian National University
Astana, Kazakhstan

Aims and Scope

The Eurasian Mathematical Journal (EMJ) publishes carefully selected original research papers in all areas of mathematics written by mathematicians, principally from Europe and Asia. However papers by mathematicians from other continents are also welcome.

From time to time the EMJ will also publish survey papers.

The EMJ publishes 4 issues in a year.

The language of the paper must be English only.

The contents of EMJ are indexed in Mathematical Reviews, Zentralblatt Math (ZMATH), Referativnyi Zhurnal – Matematika, Math-Net.Ru.

Information for the Authors

Submission. Manuscripts should be written in LaTeX and should be submitted electronically in DVI, PostScript or PDF format to the EMJ Editorial Office via e-mail (eurasianmj@yandex.kz).

When the paper is accepted, the authors will be asked to send the tex-file of the paper to the Editorial Office.

The author who submitted an article for publication will be considered as a corresponding author. Authors may nominate a member of the Editorial Board whom they consider appropriate for the article. However, assignment to that particular editor is not guaranteed.

Copyright. When the paper is accepted, the copyright is automatically transferred to the EMJ. Manuscripts are accepted for review on the understanding that the same work has not been already published (except in the form of an abstract), that it is not under consideration for publication elsewhere, and that it has been approved by all authors.

Title page. The title page should start with the title of the paper and authors' names (no degrees). It should contain the Keywords (no more than 10), the Subject Classification (AMS Mathematics Subject Classification (2010) with primary (and secondary) subject classification codes), and the Abstract (no more than 150 words with minimal use of mathematical symbols).

Figures. Figures should be prepared in a digital form which is suitable for direct reproduction.

References. Bibliographical references should be listed alphabetically at the end of the article. The authors should consult the Mathematical Reviews for the standard abbreviations of journals' names.

Authors' data. The authors' affiliations, addresses and e-mail addresses should be placed after the References.

Proofs. The authors will receive proofs only once. The late return of proofs may result in the paper being published in a later issue.

Honorarium. An honorarium will be paid to the authors (at 50 $ per page, but no more than 2000 $ for a paper) and to the reviewers (at 50 $ per review). All papers will be edited with regard to usage of English. If severe editing is required, the honorarium may be reduced to 40 $ per page. Papers may be submitted in Russian. They will be translated into English, and in this case the honorarium will be 25 $ per page.

Offprints. The authors will receive offprints in electronic form.

Publication Ethics and Publication Malpractice

Manuscripts are accepted for review on the understanding that the same work has not been already published (except in the form of an abstract), that it is not under consideration for publication elsewhere, and that it has been approved by all authors. Only original papers are accepted for consideration. In particular, translations into English of papers already published in another language are not accepted. No other forms of scientific misconduct are allowed, such as plagiarism, falsification, incorrect interpretation of other works, incorrect citations, etc.

Web-page

The web-page of EMJ is www.emj.enu.kz. One can enter the web-page by typing Eurasian Mathematical Journal in any search engine (Google, Yandex, etc.). The archive of the web-page contains all papers published in EMJ (free access).

Subscription

For Institutions

- US$ 200 (or equivalent) for one volume (4 issues)
- US$ 60 (or equivalent) for one issue

For Individuals

- US$ 160 (or equivalent) for one volume (4 issues)
- US$ 50 (or equivalent) for one issue.

The price includes handling and postage.

The Subscription Form for subscribers can be obtained by e-mail:

eurasianmj@yandex.kz

The Eurasian Mathematical Journal (EMJ)
The Editorial Office
The L.N. Gumilyov Eurasian National University,
Main Building
Room 355
Tel.: +7-7172-709500 extension 31313
2 Mirzoyan Street
010008 Astana
Kazakhstan

EURASIAN MATHEMATICAL JOURNAL
ISSN 2077-9879
Volume 4, Number 2 (2013), 10 – 48

APPROXIMATE DIFFERENTIABILITY OF MAPPINGS OF CARNOT–CARATHÉODORY SPACES

S.G. Basalaev, S.K. Vodopyanov

Communicated by V.M. Goldshtein

Key words: approximate differentiability, Carnot–Carathéodory space.

AMS Mathematics Subject Classification: 53C17, 58C25, 28A75.

Abstract. We study the approximate differentiability of measurable mappings of Carnot–Carathéodory spaces. We show that the approximate differentiability almost everywhere is equivalent to the approximate differentiability along the basic horizontal vector fields almost everywhere. As a geometric tool we prove the generalization of Rashevsky–Chow theorem for C^1-smooth vector fields. The main result of the paper extends theorems on approximate differentiability proved by Stepanoff (1923, 1925) and Whitney (1951) for Euclidean spaces and by Vodopyanov (2000) for Carnot groups.

1 Introduction

In 1919 Rademacher proved a theorem that is the well-known result in the theory of functions of a real variable.

Theorem 1.1 ([23]). *If U is an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ is a Lipschitz mapping, then f is differentiable at almost all points of the set U.*

The result allows many enhancements and generalizations. Most natural is to consider an arbitrary measurable set as the domain of the function together with a weaker assumption on the function. The Stepanoff theorem is such a result.

Theorem 1.2 ([27]). *If $A \subset \mathbb{R}^n$ is a measurable set and a function $f : A \to \mathbb{R}^m$ satisfies the condition*

$$\varlimsup_{x \to a} \frac{|f(x) - f(a)|}{|x - a|} < \infty \quad \textit{for every point } a \in A, \tag{1.1}$$

then f is differentiable at almost all points of the set A.

The density of a measurable set $Y \subset \mathbb{R}^n$ at a point $x \in \mathbb{R}^n$ is the limit

$$\lim_{r \to +0} \frac{\mathcal{H}^n(Y \cap B(x,r))}{\mathcal{H}^n(B(x,r))},$$

in case it exists (here $\mathcal{H}^n$ is the n-dimensional Hausdorff measure).

It is known that almost all points of a measurable set Y are density points (i. e. the density of the set is equal to 1 at those points) and almost all points of the set $\mathbb{R}^n \setminus Y$ are the points of density 0.

A value $y \in \mathbb{R}^m$ is called *the approximate limit* of a function $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ at a density point $x_0 \in E$ $\big($denoted by $y = \operatorname{ap}\lim\limits_{x \to x_0} f(x)\big)$ if the set $E \setminus f^{-1}(W)$ has density 0 at the point x_0 for every neighborhood $W \subset \mathbb{R}^m$ of the point y. The approximate limit is unique [4].

The idea of the approximate limit is tightly related to the fundamental notion of the geometric measure theory: the notion of measurability. Namely, for a mapping of the Euclidean space to be measurable, it is necessary and sufficient that it is approximately continuous almost everywhere (see, for instance, [4]).

If we consider the convergence of the relation $\dfrac{f(x+tv)-f(x)}{t}$ to the value $L(v)$ of a linear mapping $L : \mathbb{R}^n \to \mathbb{R}^m$ in different topologies of the unit ball $B(0,1) \subset \mathbb{R}^n$ then we arrive at to different notions of differentiability. The convergence to L in the uniform topology $C(B(0,1))$ gives us the classical differentiability. The convergence to L in measure gives just the notion of approximate differentiability of the Euclidean space, see for instance [25].

With the approximate differential introduced by Stepanoff, the following result was obtained in his work:

Theorem 1.3 ([28]). *A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is approximately differentiable almost everywhere if and only if it has approximate derivatives with respect to each variable almost everywhere.*

It worth noting that if a mapping has a classical differential then it has an approximate one and these differentials coincide. Therefore, the approximate differential generalizes the concept of the classical differential.

With use of the approximate differential Theorem 1.2 can be further extended in the following direction. For doing this we apply a result of [4]:

Theorem 1.4. *If $A \subset \mathbb{R}^n$, $f : A \to \mathbb{R}^m$ and*

$$\operatorname{ap}\overline{\lim_{x\to a}} \frac{|f(x)-f(a)|}{|x-a|} < \infty \quad \text{for every point } a \in A, \tag{1.2}$$

then A is the union of a countable family of disjoint measurable sets A_i and a set of measure zero such that every restriction $f|_{A_i}$ is a Lipschitz mapping.

Hence, for a function f meeting condition (1.2), by Theorem 1.1, we have that every restriction $f|_{A_i}$ is differentiable almost everywhere in A_i. The density points of the sets A_i are also the density points of the set A. Therefore, one can conclude that the mapping f is approximately differentiable almost everywhere in A.

Condition (1.2) is the weakest one because it obviously holds for approximately differentiable function.

The final representation of the theorem is how it was stated by Whitney.

Theorem 1.5 ([40]). *Let the set $P \subset \mathbb{R}^n$ be measurable and bounded, $f : P \to \mathbb{R}^m$ be a measurable function. The following conditions are equivalent:*

1) *the mapping f is approximately differentiable almost everywhere in P;*

2) *the mapping f has approximate derivatives with respect to each variable almost everywhere in P;*

3) *there is a countable family of disjoint sets $Q_1, Q_2, \dots$ such that $|P \setminus \bigcup\limits_{i=1}^{\infty} Q_i| = 0$ and every restriction $f|_{Q_i}$ is a Lipschitz mapping;*

4) *for every $\varepsilon > 0$, there are a closed set $Q \subset P$ such that $|P \setminus Q| < \varepsilon$ and a C^1-smooth mapping $g : P \to \mathbb{R}^m$ such that $g = f$ in Q.*

An appropriate concept of differentiability for mappings of Carnot groups was introduced by P. Pansu in [20]. Now it is called the $\mathcal{P}$-differentiability. It was introduced in order to establish some results of the theory of quasiconformal mappings [20, 14]. Some classes of $\mathcal{P}$-differentiable mappings of Carnot groups were described in [38, 31, 16] with purpose to obtain some formulas of geometric measure theory and some crucial results of quasiconformal analysis [29, 39, 30, 32, 34, 21].

Later, in [33, 12] the concept of $\mathcal{P}$-differentiability was extended for mappings of Carnot–Carathéodory spaces for proving Rademacher and Stepanoff type theorems.

In this work we obtain a partial generalization of Theorem 1.5 for mappings of Carnot–Carathéodory spaces.

Theorem 1.6. *Let $\mathcal{M}$, $\widetilde{\mathcal{M}}$ be Carnot–Carathéodory spaces, $E \subset \mathcal{M}$ be a measurable subset of $\mathcal{M}$ and $f : E \to \widetilde{\mathcal{M}}$ be a measurable mapping. The following conditions are equivalent:*

1) *the mapping f is approximately differentiable almost everywhere in E;*

2) *the mapping f has approximate derivatives along the basic horizontal vector fields almost everywhere in E;*

3) *there is a sequence of disjoint sets $Q_1, Q_2, \dots$ such that $\bigl|E \setminus \bigcup\limits_{i=1}^{\infty} Q_i\bigr| = 0$ and every restriction $f|_{Q_i}$ is a Lipschitz mapping.*

The proof of Theorem 1.6 is a significant modification of the arguments of the work [31] where a similar result was proved for mappings of Carnot groups. In the proof we essentially use metric properties of the initial and nilpotentized vector fields established in [12, 9, 10, 6, 13].

2 Geometry of Carnot–Carathéodory spaces

We split our work in four sections. In the first one we give the basic notions and structures concerning Carnot–Carathéodory spaces. In Subsections 2.2 and 2.4 we have a look at different ways of specifying a metric and coordinate system in the Carnot–Carathéodory spaces. In Subsection 2.5 we build a special coordinate system of the second kind based on the compositions of the integral lines of the horizontal vector fields. As the consequence of this result we obtain Chow–Rashevsky theorem for C^1-smooth vector fields. We formulate also local approximation theorem for Carnot–Carathéodory metric.

In Section 2 we introduce definitions of measure, approximate limit, differentiability and approximate differentiability, and formulate necessary results obtained earlier.

The third section is devoted to the proof of the theorem on approximate differentiability. We state the theorem and show trivial implications. Then we formulate the key steps of the theorem. Main steps of its proof are carried out in separate lemmas. In this proof we make use of special coordinate system of the 2nd kind $(a_1, \dots, a_N) \mapsto \Phi_N(a_N) \circ \dots \circ \Phi_1(a_1)$ constructed in Subsection 2.5. First, in Subsection 4.1 we show that a function having approximate derivatives along the basic horizontal vector fields has approximate derivatives along the vector fields $Y_k(t)$ which generate the coordinate functions $\Phi_k(t) = \exp(Y_k(t))$. In the next subsection with use of this coordinate system we build a mapping of local Carnot groups and study its properties. Finally, in Subsection 4.3 we prove that this mapping is really the differential of the initial mapping.

As an application of our results, in the last section we prove an area formula for approximately differentiable mappings.

2.1 Carnot–Carathéodory spaces

Recall the definition of Carnot–Carathéodory space satisfying the condition of the equiregularity ([7, 19, 12]). Fix a connected Riemannian C^∞-manifold $\mathcal{M}$ of topological dimension N. The manifold $\mathcal{M}$ is called a *Carnot–Carathéodory space* if the tangent bundle $T\mathcal{M}$ has a filtration

$$H\mathcal{M} = H_1\mathcal{M} \subsetneq \dots \subsetneq H_i\mathcal{M} \subsetneq \dots \subsetneq H_M\mathcal{M} = T\mathcal{M}$$

by subbundles such that every point $g \in \mathcal{M}$ has a neighborhood $U(g) \subset \mathcal{M}$ equipped with a collection of C^1-smooth vector fields $X_1, \dots, X_N$, constituting a basis of $T_v\mathcal{M}$ in every point $v \in U(g)$ and meeting the following two properties. For every $v \in U(g)$,

(1) $H_i\mathcal{M}(v) = H_i(v) = \operatorname{span}\{X_1(v), \dots, X_{\dim H_i}(v)\}$ is a subspace of $T_v\mathcal{M}$ of a constant dimension $\dim H_i$, $i = 1, \dots, M$;

(2) $H_{j+1} = \operatorname{span}\{H_j, [H_1, H_j], [H_2, H_{j-1}], \dots, [H_k, H_{j+1-k}]\}$ where $k = \lfloor \frac{j+1}{2} \rfloor$, $j = 1, \dots, M-1$.

The subbundle $H\mathcal{M}$ is called *horizontal*.

The number M is called the *depth* of the manifold $\mathcal{M}$.

The *degree* $\deg X_k$ is defined as $\min\{m \mid X_k \in H_m\}$.

Remark 2.1. Condition (2) implies that we have the following "commutator table":

$$[X_i, X_j](v) = \sum_{k\colon \deg X_k \leq \deg X_i + \deg X_j} c_{ijk}(v) X_k(v). \tag{2.1}$$

Note, that (2.1) is weaker than condition (2) as it just implies $[H_i, H_j] \subseteq H_{i+j}$.

2.2 The coordinates of the 1st kind

In the sequel we denote by $B_e(a, r)$ an open Euclidean ball centered at the point $a \in \mathbb{R}^N$ and with radius r. From the theorems on smooth dependence of solutions of ordinary differential equations on a parameter it follows (see e. g. [1]) that the mapping

$$\theta_g : (x_1, \dots, x_N) \to \exp\Bigl(\sum_{i=1}^{N} x_i X_i\Bigr)(g), \quad \theta_g(0) = \theta_g(0, \dots, 0) = g,$$

is a C^1-smooth diffeomorphism of a ball $B_e(0, \varepsilon_g)$ in $\mathbb{R}^N$, where ε_g is a sufficiently small positive number, into the neighborhood O_g of the point $g \in \mathcal{M}$.

The collection of numbers $\{x_i\}$, $i = 1, \dots, N$, where $(x_1, \dots, x_N) = \theta_g^{-1} u \in B_e(0, \varepsilon_g)$, is called *the coordinates of the* 1*st kind* of the point $u = \exp\Bigl(\sum\limits_{i=1}^{N} x_i X_i\Bigr)(g)$.

The neighborhood $U(g_0)$ of the point g_0 can be chosen in such a way that $U(g_0) \subset \bigcap\limits_{g \in U(g_0)} O_g$. Then for every couple of points $u, g \in U(g_0)$ there is the unique N-tiple of numbers $(y_1, \dots, y_N)$ such that $u = \exp\Bigl(\sum\limits_{i=1}^{N} y_i X_i\Bigr)(g)$. For every couple of points u and g we define the non-negative quantity

$$d_\infty(u, g) = \max\bigl\{|y_i|^{1/\deg X_i} : i = 1, \dots, N\bigr\}.$$

An open ball in the quasidistance d_∞ of radius r centered at $g \in \mathcal{M}$ is denoted by $\operatorname{Box}(g, r)$.

2.3 Local geometry of Carnot–Carathéodory spaces

Using the normal coordinates θ_g^{-1}, we define the *dilation* $\Delta_\varepsilon^g : B(g, r) \to B(g, \varepsilon r)$, $0 < r \le r_g$: to an element $x = \exp\Bigl(\sum\limits_{i=1}^{N} x_i X_i\Bigr)(g)$ we assign

$$\Delta_\varepsilon^g x = \exp\Bigl(\sum_{i=1}^{N} x_i \varepsilon^{\deg X_i} X_i\Bigr)(g)$$

in the case when the right-hand side makes sense. The following theorem generalizes a result established under additional smoothness of vector fields in [17, 26, 7].

Theorem 2.1. *Let g be a point in the Carnot–Carathéodory space $\mathcal{M}$. The following statements hold*:

(1) *Coefficients*

$$\widehat{c}_{ijk} = \begin{cases} c_{ijk}(g), & \textit{if } \deg X_i + \deg X_j = \deg X_k; \\ 0 & \textit{otherwise}; \end{cases}$$

where $c_{ijk}(\cdot)$ are the functions from commutator table (2.1)*, define the structure of nilpotent graded Lie algebra on $T_g\mathcal{M}$.*

(2) *There are vector fields $\{\widehat{X}_i^g\}$ with the initial conditions $\widehat{X}_i^g(g) = X_i(g)$, $i = 1, \dots, N$, taking place in* $\operatorname{Box}(g, r_g)$ *that constitute a basis of the nilpotent graded Lie algebra $V(g)$ with the following "commutator table"*:

$$[\widehat{X}_i^g, \widehat{X}_j^g] = \sum_{k=1}^{N} \widehat{c}_{ijk} \widehat{X}_k^g = \sum_{\deg X_k = \deg X_i + \deg X_j} c_{ijk}(g) \widehat{X}_k^g. \tag{2.2}$$

(3) *For* $x \in \operatorname{Box}(g, r_g)$ *consider the vector fields*

$$X_i^\varepsilon(x) = (\Delta^g_{\varepsilon^{-1}})_* \varepsilon^{\deg X_i} X_i(\Delta^g_\varepsilon x), \quad i = 1, \dots, N.$$

Then the following equality holds

$$X_i^\varepsilon(x) = \widehat{X}_i^g(x) + \sum_{j=1}^N a_{ij}(x) \widehat{X}_j^g(x) \tag{2.3}$$

where $a_{ij}(x) = o(\varepsilon^{\max\{0, \deg X_j - \deg X_i\}})$ *for* $x \in \operatorname{Box}(g, r_g)$ *as* $\varepsilon \to 0$.

Moreover, given a compact set $\mathcal{K} \subset \mathcal{M}$ *there exists* $r > 0$ *such that relation* (2.3) *holds for all* $g \in \mathcal{K}$ *with* $x \in \operatorname{Box}(g, r)$ *and* $o(\cdot)$ *is uniform in* g *belonging to* $\mathcal{K}$ *as* $\varepsilon \to 0$.

The first statement of the theorem is proved in [12]. The second follows from the classical Lie theorem [15, 22]. The third statement is obtained in [10] for $C^{1,\alpha}$-smooth vector fields and in [6] for C^1-smooth vector fields.

Equality (2.3) implies Gromov's nilpotentization theorem with respect to the coordinates of the first kind. Notice that for the first time it was formulated in [7, p. 130] in the coordinates of the second kind.

Theorem 2.2 ([10, 6]). *The uniform convergence* $X_i^\varepsilon \to \widehat{X}_i^g$ *as* $\varepsilon \to 0$, $i = 1, \dots, N$, *holds at the points of* $\operatorname{Box}(g, r_g)$ *and this convergence is uniform in* g *belonging to some compact neighborhood.*

The Lie algebra in Theorem 2.1 can be constructed as a graded nilpotent Lie algebra V' of vector fields $(\widehat{X}_j^g)'$ in $\mathbb{R}^N$, $j = 1, \dots, N$, such that the exponential mapping $(x_1, \dots, x_N) \mapsto \exp\Big(\sum\limits_{i=1}^N x_i (\widehat{X}_j^g)'\Big)(0)$ equals identity [22, 2].

The connected simply connected Lie group $\mathbb{G}_g\mathcal{M}$ with the nilpotent graded Lie algebra V' is called the *nilpotent tangent cone* of the Carnot–Carathéodory space $\mathcal{M}$ at the point $g \in \mathcal{M}$. Ccondition (2) in the definition of Carnot–Carathéodory space provides that $\mathbb{G}_g\mathcal{M}$ is a *Carnot group*, i. e. if we denote $V_k = \operatorname{span}\{(\widehat{X}_i^g)' : \deg X_i = k\}$ then

$$V' = V_1 \oplus V_2 \oplus \dots \oplus V_M, \quad \begin{array}{l}[V_1, V_k] = V_{k+1}, \quad k = 1, \dots, M-1, \\ [V_1, V_M] = \{0\}.\end{array}$$

By means of the exponential map we can push-forward the vector fields $(\widehat{X}_j^g)'$ onto some neighborhood of $g \in \mathcal{M}$ for obtaining the vector fields $\widehat{X}_j^g(\theta_g(x)) = D\theta_g(x)\langle(\widehat{X}_j^g)'\rangle$.

To the Carnot group $\mathbb{G}_g\mathcal{M}$ corresponds a *local Carnot group* $\mathcal{G}^g$ with the nilpotent Lie algebra with the basic vector fields $\widehat{X}_1^g, \dots, \widehat{X}_N^g$. Define it so that the mapping θ_g is a *local group isomorphism* between some neighborhoods of the identity elements of the groups $\mathbb{G}_g\mathcal{M}$ and $\mathcal{G}^g$. The group operation for the elements $x = \exp\Big(\sum\limits_{i=1}^N x_i \widehat{X}_i^g\Big)(g) \in \mathcal{G}^g$

and $y = \exp\Big(\sum\limits_{i=1}^N y_i \widehat{X}_i^g\Big)(g) \in \mathcal{G}^g$ is defined by means of local group isomorphism:

$$
\begin{aligned}
x \cdot y &= \exp\Big(\sum_{i=1}^N y_i \widehat{X}_i^g\Big) \circ \exp\Big(\sum_{i=1}^N x_i \widehat{X}_i^g\Big)(g) \\
&= \theta_g \circ \exp\Big(\sum_{i=1}^N y_i (\widehat{X}^g)_i'\Big) \circ \exp\Big(\sum_{i=1}^N x_i (\widehat{X}^g)_i'\Big)(0).
\end{aligned}
$$

Define the one-parameter dilation group δ_t^g on $\mathcal{G}^g$:
to an element $x = \exp\Big(\sum\limits_{i=1}^N x_i \widehat{X}_i^g\Big)(g) \in \mathcal{G}^g$ we assign

$$
\delta_t^g x = \exp\Big(\sum_{i=1}^N x_i t^{\deg X_i} \widehat{X}_i^g\Big)(g) \in \mathcal{G}^g, \quad t \in (0, t(x)).
$$

The relation $\delta_t^g x \cdot \delta_\tau^g x = \delta_{t\tau}^g x$ is defined for t, τ such that $t, \tau, t\tau \in (0, t(x))$.

We extend the definition of δ_t^g on negative t, setting $\delta_t^g x = \delta_{|t|}^g (x^{-1})$ for $t < 0$.

Since the local Carnot group $\mathcal{G}^g$ itself is a Carnot–Carathéodory space with the collection of vector fields $\{\widehat{X}_j^g\}$, it is endowed with the quasidistance $d_\infty^g(x, y)$.

Throughout the paper we use the following properties.

Property 2.1 ([12]). *Geometric properties of the local Carnot group:*

(1) *The mapping δ_t^g is a group automorphism: for all elements $x, y \in \mathcal{G}^g$ and numbers $t \in (0, \min\{t(x), t(y), t(x \cdot y)\})$ we have $\delta_t^g x \cdot \delta_t^g y = \delta_t^g (x \cdot y)$.*

(2) *The function $\mathcal{G}^g \ni x \to d_\infty^g(g, x)$ is a local homogeneous norm on $\mathcal{G}^g$, i. e., it meets the following conditions:*

(a) *$d_\infty^g(g, x) \geq 0$ for $x \in \mathcal{G}^g$ and $d_\infty^g(g, x) = 0$ if and only if $x = g$;*

(b) *$d_\infty^g(g, \delta_t^g x) = t d_\infty^g(g, x)$ for every $t \in (0, t(x))$;*

(c) *$d_\infty^g(g, x \cdot y) \leq Q_1\big(d_\infty^g(g, x) + d_\infty^g(g, y)\big)$ for all x, y, $x \cdot y \in \mathcal{G}^g$. The constant Q_1 is bounded with respect to g in some compact set in $\mathcal{M}$.*

(3) *The quantity $d_\infty^g(a, b) = d_\infty^g(g, b^{-1} \cdot a)$ is a left invariant distance on $\mathcal{G}^g$: $d_\infty^g(x \cdot a, x \cdot b) = d_\infty^g(a, b)$ for all a, b, $x \in \mathcal{G}^g$ for which the left- and right-hand sides of the equality make sense.*

Property 2.2 ([12]). *Let $g \in \mathcal{M}$. Then*

$$
\exp\Big(\sum_{i=1}^N a_i X_i\Big)(g) = \exp\Big(\sum_{i=1}^N a_i \widehat{X}_i^g\Big)(g)
$$

for all $|a_i| < r_g$, $i = 1, \ldots, N$.

Observe, that the latter implies $d_\infty^g(g, x) = d_\infty(g, x)$.

Proposition 2.1 ([12, 13]). *The quantity d_∞ is a quasimetric in the sense of* [19] *that is the following relations hold for all points of the neighborhood $U(g_0)$:*
1) *$d_\infty(u, g) \geq 0$, $d_\infty(u, g) = 0$ if and only if $u = g$;*

2) $d_\infty(u,g)=d_\infty(g,u)$;
3) *there is a constant* $Q\ge 1$ *such that, for every triple of points* u, w, $v\in U(g_0)$, *we have*

$$d_\infty(u,v)\le Q(d_\infty(u,w)+d_\infty(w,v)).$$

An essential distinction between the geometry of a sub-Riemannian space and the geometry of a Riemannian space is that the metrics of the initial space and of the nilpotent tangent cone are not bi-Lipschitz equivalent. Therefore, in studying the questions of the local behavior of the geometric objects, it is important to know estimates of the deviation of one metric from another.

Theorem 2.3 ([13, Theorem 8]). *Assume that* g, $w_0\in U(g_0)$ *satisfy* $d_\infty(g,w_0)=C\varepsilon$. *For a fixed* $L\in\mathbb{N}$, *consider the points*

$$\widehat{w}_j^\varepsilon=\exp\Big(\sum_{i=1}^N w_{i,j}\varepsilon^{\deg X_j}\widehat{X}_j^g\Big)(\widehat{w}_{j-1}^\varepsilon),\quad w_j^\varepsilon=\exp\Big(\sum_{i=1}^N w_{i,j}\varepsilon^{\deg X_j}X_j\Big)(w_{j-1}^\varepsilon),$$

$\widehat{w}_0^\varepsilon=w_0^\varepsilon=w_0$, $j=1,\dots,L$. *Then*

$$\max\{d_\infty^g(\widehat{w}_L^\varepsilon,w_L^\varepsilon),d_\infty(\widehat{w}_L^\varepsilon,w_L^\varepsilon)\}=o(\varepsilon)\qquad as\ \varepsilon\to 0,$$

where $o(\varepsilon)$ *is uniform in* g, $w_0\in U(g_0)$ *and* $\{w_{i,j}\}$, $i=1,\dots,N$, $j=1,\dots,L$, *in some compact neighborhood of* 0 *and* $\varepsilon>0$.

Theorem 2.4 ([13, Theorem 6]). *Consider points* $g\in\mathcal{M}$ *and* u, $v\in\operatorname{Box}(g,\varepsilon)$, *where* $\varepsilon\in(0,r_g)$. *Then*

$$|d_\infty^g(u,v)-d_\infty(u,v)|=o(\varepsilon)\qquad as\ \varepsilon\to 0,$$

where $o(\varepsilon)$ *is uniform in* u, $v\in\operatorname{Box}(g,\varepsilon)$ *and* g *belonging to some compact set.*

2.4 The coordinates of the 2nd kind

In the neighborhood of a point g_0 consider the same family of the basic vector fields $\{X_1,\dots,X_{\dim H_1},X_{\dim H_1+1},\dots,X_N\}$ as in the definition of the coordinates of the first kind. It is known that the mapping

$$(a_1,\dots,a_N)\mapsto\exp(a_NX_N)\circ\dots\circ\exp(a_1X_1)(g)\tag{2.4}$$

is a C^1-diffeomorphism of some neighborhood $B_e(0,\varepsilon)\subset\mathbb{R}^N$ to a neighborhood $V(g)$ of g (the so-called *coordinates of the second kind*). Similarly to the case of the coordinates of the first kind we can choose a neighborhood $U(g_0)$ such that $U(g_0)\subset\bigcap\limits_{g\in U(g_0)}V(g)$.

For the points $u,g\in U(g_0)$, $u=\exp(a_NX_N)\circ\dots\circ\exp(a_1X_1)(g)$, by means of the coordinates of the 2nd kind we can define the quantity

$$d_2(u,g)=\max\big\{|a_i|^{1/\deg X_i}:i=1,\dots,N\big\}.$$

Next we show that the quantity $d_2(u,g)$ is comparable with the quasimetric $d_\infty(u,g)$ in a neighborhood $U(g_0)$ i. e.

$$c_1d_\infty(u,g)\le d_2(u,g)\le c_2d_\infty(u,g)\tag{2.5}$$

for all points u, $g\in U(g_0)$ and positive constants c_1 and c_2 independent of u, $g\in U(g_0)$.

Remark 2.2. For Carnot groups the equivalence of d_∞ and d_2 is known (see, for instance, [5]). This means that if d^g_∞ and d^g_2 are quasimetrics in the local Carnot group $\mathcal{G}^g$, $g \in \mathcal{M}$, then there are constants c^g_1 and c^g_2 such that

$$c^g_1 d^g_\infty(u,v) \le d^g_2(u,v) \le c^g_2 d^g_\infty(u,v) \tag{2.6}$$

for all u, $v \in \mathcal{G}^g$.

Proposition 2.2. *There are constants c_1 and c_2 such that inequalities* (2.5) *hold for all points u, g in some neighborhood $U(g_0)$ in which quasimetrics d_∞ and d_2 are defined.*

Proof. Let u, $g \in U(g_0)$ be arbitrary points and $d_2(u,g) = r$. Assuming that $y_0 = g$, $y_1 = \exp(a_1X_1)(y_0), \dots, y_N = \exp(a_NX_N)(y_{N-1})$ by the generalized triangle inequality (see Proposition 2.1) we have the following relations

$$\begin{aligned} d_\infty(u,g) &\le Q^{N-1}\Big(\sum_{i=1}^N d_\infty(y_k,y_{k-1})\Big) \\ &\le Q^{N-1}\Big(\sum_{i=1}^N |a_i|^{\frac{1}{\deg X_i}}\Big) \le NQ^{N-1}r = NQ^{N-1}d_2(u,g). \end{aligned} \tag{2.7}$$

Thus the left inequality in (2.5) is proved with $c_1 = (NQ^{N-1})^{-1}$.

Next, suggest that the right inequality in (2.5) does not hold in some closed ball $\overline{\mathrm{Box}}(g_0, 2r_0)$. Then there are sequences of points x_n, $y_n \in \overline{\mathrm{Box}}(g_0,r_0)$ converging to the same point $x_0 \in \overline{\mathrm{Box}}(g_0,r_0)$, such that

$$\varepsilon_n = d_2(x_n,y_n) \ge n\, d_\infty(x_n,y_n),$$

where $\varepsilon_n \to 0$ as $n \to \infty$ (otherwise the right inequality in (2.5) would be fulfilled in $\overline{\mathrm{Box}}(g_0,r_0)$). Define on $\overline{\mathrm{Box}}(g_0,r_0)$ dilations $\mathfrak{D}^g_t$ and $\widehat{\mathfrak{D}}^g_t$ as follows: to an element $x = \exp(x_NX_N) \circ \dots \circ \exp(x_1X_1)(g) \in \overline{\mathrm{Box}}(g_0,r_0)$ assign

$$\mathfrak{D}^g_t x = \exp(x_N t^{\deg X_N} X_N) \circ \dots \circ \exp(x_1 t X_1)(g)$$

and to an element $\hat{x} = \exp(x_N\widehat{X}^g_N) \circ \dots \circ \exp(x_1\widehat{X}^g_1)(g) \in \overline{\mathrm{Box}}(g_0,r_0) \cap \mathcal{G}^g$ assign

$$\widehat{\mathfrak{D}}^g_t \hat{x} = \exp(x_N t^{\deg X_N} \widehat{X}^g_N) \circ \dots \circ \exp(x_1 t \widehat{X}^g_1)(g).$$

Observe that $d_2(g, \mathfrak{D}^g_t x) = td_2(g,x)$ and $d^g_2(g, \widehat{\mathfrak{D}}^g_t x) = td^g_2(g,x)$. Let

$$0 < \delta = \sup\{t > 0 : \mathfrak{D}^g_t x, \widehat{\mathfrak{D}}^g_t x \in \overline{\mathrm{Box}}(g_0,2r_0) \text{ for all } x,g \in \overline{\mathrm{Box}}(g_0,r_0)\}.$$

Then $\mathfrak{D}^{x_n}_{\delta/\varepsilon_n} y_n \in \overline{\mathrm{Box}}(g_0,2r_0)$ and

$$d_2(x_n, \mathfrak{D}^{x_n}_{\delta/\varepsilon_n} y_n) = \frac{\delta}{\varepsilon_n} d_2(x_n,y_n) = \delta > 0. \tag{2.8}$$

Represent y_n in coordinates of the 2nd kind as $y_n = \exp(y_{nN}X_N) \circ \dots \circ \exp(y_{n1}X_1)(x_n)$ and define

$$z_n = \exp(y_{nN}\widehat{X}^g_N) \circ \dots \circ \exp(y_{n1}\widehat{X}^g_1)(x_n).$$

Since $d_\infty(x_n, y_n) = d_\infty^{x_n}(x_n, y_n) \le \frac{\varepsilon_n}{n}$, by (2.6) it follows

$$d_2^{x_n}(x_n, y_n) \le c_2^{x_n} d_\infty^{x_n}(x_n, y_n) \le c_2^{x_n} \frac{\varepsilon_n}{n} = O\Big(\frac{\varepsilon_n}{n}\Big)$$

where $O(\cdot)$ is uniform in $\overline{\mathrm{Box}}(g_0, r_0)$. This means that in the representation

$$y_n = \exp(v_{nN} \widehat{X}_N^g) \circ \dots \circ \exp(v_{n1} \widehat{X}_1^g)(x_n)$$

the coordinates v_j meet the property $|v_{nj}|^{\deg X_j} = O(\frac{\varepsilon_n}{n})$. Then we can apply Theorem 2.3 to the points y_n and z_n and derive that $d_\infty^{x_n}(y_n, z_n) = o(\frac{\varepsilon_n}{n})$. Consequently,

$$d_\infty^{x_n}(x_n, z_n) \le C(d_\infty^{x_n}(x_n, y_n) + d_\infty^{x_n}(y_n, z_n)) = O\Big(\frac{\varepsilon_n}{n}\Big) + o\Big(\frac{\varepsilon_n}{n}\Big) = O\Big(\frac{\varepsilon_n}{n}\Big).$$

By Theorem 2.3 it also follows $d_\infty^{x_n}(\mathfrak{D}_{\delta/\varepsilon_n}^{x_n} y_n, \widehat{\mathfrak{D}}_{\delta/\varepsilon_n}^{x_n} z_n) = o(\frac{1}{n})$. Therefore,

$$\begin{aligned}
d_2^{x_n}(x_n, \mathfrak{D}_{\delta/\varepsilon_n}^{x_n} y_n) &\le C_1\big(d_2^{x_n}(x_n, \widehat{\mathfrak{D}}_{\delta/\varepsilon_n}^{x_n} z_n) + d_2^{x_n}(\widehat{\mathfrak{D}}_{\delta/\varepsilon_n}^{x_n} z_n, \mathfrak{D}_{\delta/\varepsilon_n}^{x_n} y_n)\big) \\
&= C_1\Big(\frac{\delta}{\varepsilon_n} d_2^{x_n}(x_n, z_n) + d_2^{x_n}(\widehat{\mathfrak{D}}_{\delta/\varepsilon_n}^{x_n} z_n, \mathfrak{D}_{\delta/\varepsilon_n}^{x_n} y_n)\Big) \\
&\le C_2\Big(\frac{\delta}{\varepsilon_n} d_\infty^{x_n}(x_n, z_n) + d_\infty^{x_n}(\widehat{\mathfrak{D}}_{\delta/\varepsilon_n}^{x_n} z_n, \mathfrak{D}_{\delta/\varepsilon_n}^{x_n} y_n)\Big) \\
&= O\Big(\frac{1}{n}\Big) + o\Big(\frac{1}{n}\Big) = O\Big(\frac{1}{n}\Big) \to 0 \quad \text{as } n \to \infty,
\end{aligned}$$

where C_1, $C_2 < \infty$ are bounded, all $O(\cdot)$ are uniform in $\overline{\mathrm{Box}}(g_0, r_0)$.

Hence we come to a contradiction with (2.8), and, therefore, the right inequality in (2.5) is proved. □

Corollary 2.1. *The quantity d_2 is a quasimetric in the sense of* [19], *i. e. the following conditions hold for the points of the neighborhood $U(g_0)$:*
1) $d_2(u, g) \ge 0$, $d_2(u, g) = 0$ *if and only if* $u = g$;
2) $d_2(u, g) \le c_1^{-1} c_2 d_2(g, u)$, *where the constants c_1 and c_2 are the ones from the proposition 2.2*;
3) *there is a constant $Q_2 \ge 1$ such that for every triple of the points u, w, $v \in U(g_0)$ we have*

$$d_2(u, v) \le Q_2(d_2(u, w) + d_2(w, v)),$$

where $Q_2 = c_1^{-1} c_2 Q$ and Q is a constant in the generalized triangle inequality for d_∞;
(4) $d_2(u, v)$ *is continuous with respect to the first variable.*

Proof. We prove for example the second property: $d_2(u, g) \le c_2 d_\infty(u, g) = c_2 d_\infty(g, u) \le c_1^{-1} c_2 d_2(g, u)$. The third property can be proved using the same procedure. The last property follows from the continuous dependence of solutions of ODE on the initial data. □

2.5 Special coordinate system of the 2nd kind and Rashevsky–Chow Theorem

The goal of this section is to modify the coordinate system of the 2nd kind

$$(t_1, \dots, t_N) \mapsto \exp(t_N X_N) \circ \dots \circ \exp(t_1 X_1)(g)$$

in the following way. We prove that exponents of nonhorizontal vector fields X_k, $k = \dim H_1 + 1, \dots, N$, can be replaced by compositions of exponents of some family of horizontal vector fields $X_1, \dots, X_{\dim H_1}$ and the resulting mapping still covers a neighborhood of g. For Carnot groups this property is known as the following statement.

Lemma 2.1 ([5]). *Let $\mathbb{G} = (\mathbb{R}^N, \cdot)$ be a Carnot group and let vector fields $Y_1, \dots, Y_n$ be a basis of horizontal subspace V_1 of its Lie algebra. Then every point $v \in \mathbb{G}$ can be represented as*

$$v = \prod_{k=1}^{L} \exp(a_k Y_{i_k})(0)$$

where $1 \le i_k \le n$, $|a_k| \le c_1 \|v\|_\infty$, constants L and c_1 are independent of v.

Lemma 2.2. *Fix $g \in \mathcal{M}$. There exists a mapping $\widehat{\Phi}_g : B_e(0, \varepsilon) \to \mathcal{G}^g$ defined as*

$$\widehat{\Phi}_g : (t_1, \dots, t_N) \mapsto \widehat{\Phi}_N(t_N) \circ \dots \circ \widehat{\Phi}_{\dim H_1 + 1}(t_{\dim H_1 + 1}) \circ \exp(\widehat{X}^g_{\dim H_1}) \circ \dots \circ \exp(\widehat{X}^g_1)(g) \quad (2.9)$$

which is a homeomorphism of a ball $B_e(0, \varepsilon)$ onto the neighborhood $V(g) \subset \mathcal{G}^g$ of a point g with the mappings $\widehat{\Phi}_k$ satisfying

$$\widehat{\Phi}_k(t)(\cdot) = \begin{cases} \exp(a_{L,k} t \widehat{X}^g_{L,k}) \circ \dots \circ \exp(a_{1,k} t \widehat{X}^g_{1,k})(\cdot), & t \ge 0, \\ \exp(a_{1,k} t \widehat{X}^g_{1,k}) \circ \dots \circ \exp(a_{L,k} t \widehat{X}^g_{L,k})(\cdot), & t < 0, \end{cases}$$

where $|a_{i,k}| \le c_1$ for all $k = \dim H_1 + 1, \dots, N$, $i = 1, \dots, L$, every $\widehat{X}^g_{i,k}$ is from $\{\widehat{X}^g_1, \dots, \widehat{X}^g_{\dim H_1}\}$.

Proof. Consider coordinate system of the 2nd kind on the nilpotent tangent cone $\mathbb{G}_g \mathcal{M}$.

$$\Theta_g(t_1, \dots, t_N) = \exp(t_N (\widehat{X}^g_N)') \circ \dots \circ \exp(t_1 (\widehat{X}^g_1)')(0).$$

The mapping Θ_g is a diffeomorphism of $\mathbb{R}^N$. For every nonhorizontal vector field $(\widehat{X}^g_k)'$ fix the decomposition given by Lemma 2.1

$$\exp((\widehat{X}^g_k)')(0) = \exp(a_{L,k} (\widehat{X}^g_{L,k})') \circ \dots \circ \exp(a_{1,k} (\widehat{X}^g_{1,k})')(0).$$

Here $|a_{i,k}| < c_1$ for all $i = 1, \dots, L$, $k = \dim H_1 + 1, \dots, N$, and every $(\widehat{X}^g_{i,k})'$ is in the set $\{(\widehat{X}^g_1)', \dots, (\widehat{X}^g_{\dim H_1})'\}$. Applying the dilation δ^g to this decomposition we obtain the following representation

$$
\begin{aligned}
\delta^g_t \exp((\widehat{X}^g_k)')(0) &= \exp(t^{\deg X_k}(\widehat{X}^g_k)')(0)\\
&= \exp(a_{L,k}t(\widehat{X}^g_{L,k})')\circ\dots\circ\exp(a_{1,k}t(\widehat{X}^g_{1,k})')(0), \qquad t\geq 0,\\
\delta^g_t \exp((\widehat{X}^g_k)')(0) &= \exp(-|t|^{\deg X_k}(\widehat{X}^g_k)')(0)\\
&= \exp(a_{1,k}t(\widehat{X}^g_{1,k})')\circ\dots\circ\exp(a_{L,k}t(\widehat{X}^g_{L,k})')(0), \qquad t< 0.
\end{aligned}
\tag{2.10}
$$

Since the vector fields $(\widehat{X}^g_k)'$ are left-invariant, representation (2.10) holds also if we replace 0 by arbitrary $x\in\mathbb{G}_g\mathbb{M}$.

Next, we push-forward representation (2.10) using the local group isomorphism θ_g. Define the mappings $\widehat{\Phi}_k:[-\varepsilon,\varepsilon]\times\mathrm{Box}(g,\varepsilon)\to\mathcal{G}^g$ by

$$
\widehat{\Phi}_k(t)(w)=\begin{cases}\exp(a_{L,k}t\widehat{X}^g_{L,k})\circ\dots\circ\exp(a_{1,k}t\widehat{X}^g_{1,k})(w), & t\geq 0,\\ \exp(a_{1,k}t\widehat{X}^g_{1,k})\circ\dots\circ\exp(a_{L,k}t\widehat{X}^g_{L,k})(w), & t<0\end{cases}
\tag{2.11}
$$

where, by definition,

$$
\exp(a\widehat{X}^g_i)\circ\exp(b\widehat{X}^g_j)=\theta_g\circ\exp(a(\widehat{X}^g_i)')\circ\exp(b(\widehat{X}^g_j)')\circ\theta_g^{-1}
$$

and $\varepsilon>0$ is sufficiently small so that (2.11) makes sense for all $k=\dim H_1+1,\dots,N$, $t\in[-\varepsilon,\varepsilon]$ and $w\in\mathrm{Box}(g,\varepsilon)$.

Consider a mapping $\widehat{\Phi}_g$ defined as in (2.9). Since, by construction,

$$
\widehat{\Phi}_g(t_1,\dots,t_N)=\theta_g\circ\Theta_g(t_1^{\deg X_1},\dots,t_N^{\deg X_N}),
$$

the mapping $\widehat{\Phi}_g$ is a homeomorphism of a ball $B_e(0,\varepsilon)\subset\mathbb{R}^N$ onto the neighborhood $V(g)\subset\mathcal{M}\cap\mathcal{G}^g$. The lemma is proved. □

For every point $g\in U(g_0)$ define the mappings $\Phi_k:[-\varepsilon,\varepsilon]\to\mathcal{M}$ by

$$
\Phi_k(t)(\cdot)=\begin{cases}\exp(a_{L,k}tX_{L,k})\circ\dots\circ\exp(a_{1,k}tX_{1,k})(\cdot), & t\geq 0,\\ \exp(a_{1,k}tX_{1,k})\circ\dots\circ\exp(a_{L,k}tX_{L,k})(\cdot), & t<0,\end{cases}
\tag{2.12}
$$

where the coefficients $a_{i,k}$, $i=1,\dots,L$, $k=\dim H_1+1,\dots,N$, are taken from the representation (2.10). Define also the mapping $\Phi_g:B_e(0,\varepsilon)\to\mathcal{M}$ by

$$
\begin{aligned}
\Phi_g:(t_1,\dots,t_N)\mapsto{}&\Phi_N(t_N)\circ\dots\circ\Phi_{\dim H_1+1}(t_{\dim H_1+1})\\
&\circ\exp(t_{\dim H_1}X_{\dim H_1})\circ\dots\circ\exp(t_1X_1)(g).
\end{aligned}
\tag{2.13}
$$

Next, we prove that Φ_g is the desired mapping, i. e. there is a neighborhood $V(g)$ such that $V(g)\subset\Phi(B_e(0,\varepsilon))$.

Theorem 2.5. *Fix a point $g_0\in\mathcal{M}$. Let $X_1,\dots,X_{\dim H_1}$ be a basis in H_1. Then there is a neighborhood $U(g_0)$ such that for every point $g\in U(g_0)$ an element $v\in U(g_0)$ can be represented as*

$$
v=\exp(a_LX_{j_L})\circ\dots\circ\exp(a_2X_{j_2})\circ\exp(a_1X_{j_1})(g),
\tag{2.14}
$$

where $1\leq j_i\leq\dim H_1$, $i=1,\dots,L$, $L\in\mathbb{N}$, $|a_i|\leq c_2\,d_\infty(g,v)$, and constants L and c_2 are independent of g and v.

Proof. Fix $g_0 \in \mathcal{M}$. Let $\widehat{\Phi}_k(t)(\cdot)$ and $\Phi_k(t)(\cdot)$ be defined as in (2.10) and (2.12). By Theorem 2.3 we have

$$d_\infty\big(\widehat{\Phi}_k(t)(w), \Phi_k(t)(w)\big) = o(t) \quad \text{as } t \to 0$$

where $o(t)$ is uniform with respect to g, w in a compact neighborhood $U(g_0)$.

Let $B_e(0,r)$ be an Euclidean ball in $\mathbb{R}^N$ and mappings $\widehat{\Phi}_g$ and $\Phi_g : B_e(0,r) \to \mathcal{M}$ be defined as in (2.9) and (2.13). Observe that both mappings are continuous and that $d_\infty(\Phi_g(x), \widehat{\Phi}_g(x)) = o(r)$ as $r \to 0$ where $o(r)$ is uniform in $g \in U(g_0)$ and $x \in B_e(0,r)$. Moreover, $\widehat{\Phi}_g$ is a homeomorphism of $B_e(0,r)$ onto a neighborhood $V(g) \in \mathcal{M} \cap \mathcal{G}^g$.

Define $\psi = \Phi_g \circ \widehat{\Phi}_g^{-1}$. The mapping $\psi : V(g) \to \mathcal{M}$ is continuous and $d_\infty(v, \psi(v)) = o(d_\infty(g,v))$ as $v \to g$ where $o(\cdot)$ is uniform in $g, v \in U(g_0)$. Choose $\varepsilon_0 > 0$ such that $d_\infty(v,\psi(v)) \leq \frac{\varepsilon}{2Q}$ for every $v \in \overline{\mathrm{Box}(g,\varepsilon)}$, $0 < \varepsilon \leq \varepsilon_0$ and $g \in U(g_0)$, where $Q \geq 1$ is a constant in the generalized triangle inequality for d_∞. Next, we prove that $\psi(\mathrm{Box}(g,\varepsilon))$ is a neighborhood of g.

Consider the homotopy $\psi_t(v) = \delta^v_{1-t}\psi(v)$, $t \in [0,1]$. It is clear that $\psi_0(v) = \psi(v)$ and $\psi_1(v) = v$. Fix a point $w \in \mathrm{Box}(g, \frac{\varepsilon}{2Q})$. Then for every $v \in \partial\mathrm{Box}(g,\varepsilon)$ we have

$$\varepsilon = d_\infty(g,v) \leq Q\big(d_\infty(g,w) + d_\infty(w,v)\big) < \frac{\varepsilon}{2} + Q d_\infty(w,v).$$

Hence, $d_\infty(w,v) > \frac{\varepsilon}{2Q}$. On the other side, for all $v \in \partial\mathrm{Box}(g,\varepsilon)$ we also have

$$\begin{aligned} d_\infty(\psi_t(v), v) &= d_\infty(\delta^v_{1-t}\psi(v), v) \\ &= d^v_\infty(\delta^v_{1-t}\psi(v), v) = (1-t) d^v_\infty(\psi(v), v) \\ &\leq d^v_\infty(\psi(v), v) = d_\infty(\psi(v), v) \leq \frac{\varepsilon}{2Q}. \end{aligned}$$

Consequently, $w \notin \psi\big(\partial\mathrm{Box}(g,\varepsilon)\big)$ for all $t \in [0,1]$. Therefore, the topological degree of ψ_t at w is invariant for all $t \in [0,1]$. Since

$$\deg(w, \mathrm{Box}(g,\varepsilon), \psi) = \deg(w, \mathrm{Box}(g,\varepsilon), \psi_1) = \deg(w, \mathrm{Box}(g,\varepsilon), \psi_0) = 1,$$

we conclude $w \in \psi\big(\mathrm{Box}(g,\varepsilon)\big)$. In other words $\mathrm{Box}(g, \frac{\varepsilon}{2Q}) \subset \Phi_g(\mathrm{Box}_e(0,\varepsilon))$, where $\mathrm{Box}_e(0,\varepsilon) = \{x \in \mathbb{R}^N : |x_i| < \varepsilon, i = 1, \dots, N\}$ is an Euclidean cube.

Let $U(g_0)$ be a neighborhood of g_0 small enough that

$$U(g_0) \subset \bigcap_{g \in U(g_0)} \mathrm{Box}(g, \tfrac{\varepsilon_0}{2Q}).$$

Let $\varepsilon = d_\infty(g,v)$ where $g, v \in U(g_0)$. Then there exists an N-tiple of numbers $(t_1, \dots, t_N)$ such that $|t_i| < 2Q\varepsilon$ and $v = \Phi_g(t_1, \dots, t_N)$. This completes the proof. $\square$

An absolutely continuous curve $\gamma : [0,T] \to \mathcal{M}$ is said to be *horizontal* if $\dot\gamma(t) \in H_{\gamma(t)}\mathcal{M}$ for almost all $t \in [0,T]$.

As an immediate consequence of Theorem 2.5 we obtain the following generalization of Rashevsky–Chow theorem [24, 3, 12]. For C^1-smooth fields $X_1, \dots, X_N$ this statement is new.

Theorem 2.6. 1) *Let $g \in \mathcal{M}$. There exists a neighborhood U of a point g such that every pair of points u, $v \in U$ in a Carnot–Carathéodory space $\mathcal{M}$ can be joined by an absolutely continuous horizontal curve γ constituted of at most L segments of integral lines of basic horizontal fields where L is independent of the choice of points x, $y \in U$.*

2) *Every pair of points u, v in a connected Carnot–Carathéodory space $\mathcal{M}$ can be joined by an absolutely continuous horizontal curve γ constituted of finite number of segments of integral lines of basic horizontal fields.*

2.6 Carnot–Carathéodory metric and Ball-Box Theorem

The *Carnot–Carathéodory distance* between two points x, $y \in \mathcal{M}$ is defined by

$$\begin{aligned} d_{cc}(x,y) = \inf\{T > 0 : \text{there exists a horizontal path } \gamma : [0,T] \to \mathcal{M}, \\ \gamma(0) = x, \gamma(T) = y, |\dot{\gamma}(t)| \leq 1\}. \end{aligned}$$

Theorem 2.6 guarantees that $d_{cc}(x,y) < \infty$ for all x, $y \in \mathcal{M}$. An open ball in Carnot–Carathéodory metric of radius r centered at x is denoted as $B_{cc}(x,r)$.

The following statement is called the local approximation theorem. It was formulated in [7, p. 135] for "sufficiently smooth vector fields". It was proved in [37] for $C^{1,\alpha}$-smooth vector fields but the same arguments work for the case of C^1-smooth vector fields since they are based on the property (2.3) [13, Theorem 7].

Theorem 2.7 ([37, 13]). *Let $g \in \mathcal{M}$. Then for every two points u, $v \in B_{cc}(g,\varepsilon)$ we have*

$$|d_{cc}(u,v) - d^g_{cc}(u,v)| = o(\varepsilon) \quad \text{as } \varepsilon \to 0$$

where $o(\varepsilon)$ is uniform in u, $v \in B(g,\varepsilon)$ and g belonging to some compact set.

As a corollary we obtain a comparison of metric d_{cc} and quasimetric d_∞, and Ball-Box theorem.

Theorem 2.8 ([13, Theorem 11]). *Let $g \in \mathcal{M}$. There exists a compact neighborhood $U(g) \subset \mathcal{M}$ and constants $0 < C_1 \leq C_2 < \infty$ independent of u, $v \in U(g)$ such that*

$$C_1 d_\infty(u,v) \leq d_{cc}(u,v) \leq C_2 d_\infty(u,v) \tag{2.15}$$

for all u, $v \in U(g)$.

The following statement was proved for sufficiently smooth vector fields in [19, 7], for $C^{1,\alpha}$-smooth vector fields, $\alpha \in (0,1]$, in [12] and for C^1-smooth vector fields in [13].

Corollary 2.2 (Ball-Box theorem [13]). *Given a compact neighborhood $U \in \mathcal{M}$, there exist constants $0 < C_1 \leq C_2 < \infty$ and $r_0 > 0$ independent of $x \in U$ such that*

$$\operatorname{Box}(x, C_1 r) \subset B_{cc}(x,r) \subset \operatorname{Box}(x, C_2 r)$$

for all $r \in (0, r_0)$ and $x \in U$.

3 Approximate limit and differentiability

3.1 Hausdorff measure

The (spherical) k-dimensional *Hausdorff measure* of the set E with respect to the metric d_{cc} is the quantity

$$\mathcal{H}^k(E) = \lim_{\varepsilon\to 0+} \inf\Big\{\sum_i r_i^k : E \subset \bigcup_i B_{cc}(x_i, r_i), r_i < \varepsilon\Big\}.$$

Theorem 3.1 ([18, 12]). *The Hausdorff dimension of $\mathcal{M}$ with respect to d_{cc} is equal to*

$$\nu = \sum_{k=1}^{N} \deg X_k = \sum_{i=1}^{M} i(\dim H_i - \dim H_{i-1})$$

where $\dim H_0 = 0$*.*

Ball-Box theorem implies the double property of measure.

Proposition 3.1. *We have*

$$\mathcal{H}^\nu(B_{cc}(x, 2r)) \le C\mathcal{H}^\nu(B_{cc}(x, r))$$

where $C < \infty$ is bounded in $r \in (0, r_0]$ and x belonging to some compact part $V \subset \mathcal{M}$.

3.2 Approximate limit and its properties

The *density* of a set Y at $x \in \mathcal{M}$ is a limit

$$\lim_{r\to+0} \frac{\mathcal{H}^\nu(B_{cc}(x,r)\cap Y)}{\mathcal{H}^\nu(B_{cc}(x,r))}$$

if it exists at x (where ν is the Hausdorff dimension of the space $\mathcal{M}$).

Let $E \subset \mathcal{M}$ be a measurable set and $f : E \to \mathbb{M}$ be a mapping to a metric space $\mathbb{M}$.

A point $y \in \mathbb{M}$ is called the *approximate limit* of the mapping f at the point $g \in E$ of density 1 and is denoted by $y = \operatorname{ap}\lim\limits_{x\to g} f(x)$ if the density of set $E \setminus f^{-1}(W)$ at g equals zero for every neighborhood W of the point y.

In the case $\mathbb{M} = \overline{\mathbb{R}}$ we also define the *approximate upper limit* of the function f at the point $g \in E$, denoted by $\operatorname{ap}\overline{\lim\limits_{x\to g}} f(x)$, as the greatest lower bound of the set of all numbers s for which the density of the set $\{z \in \mathcal{M} : f(z) > s\}$ at the point g equals zero. By definition, $\operatorname{ap}\underline{\lim\limits_{x\to g}} f(x) = -\operatorname{ap}\overline{\lim\limits_{x\to g}}(-f(x))$ is the *approximate lower limit*. It is easy to verify that $\operatorname{ap}\underline{\lim\limits_{x\to g}} f(x) \le \operatorname{ap}\overline{\lim\limits_{x\to g}} f(x)$ and that $\operatorname{ap}\lim\limits_{x\to g} f(x)$ exists if and only if $\operatorname{ap}\underline{\lim\limits_{x\to g}} f(x) = \operatorname{ap}\overline{\lim\limits_{x\to g}} f(x)$.

Next we state several properties regarding measurability and approximate limit which we need in further arguments.

Property 3.1. *Let S be a $\mathcal{H}^\nu \times \mathcal{H}^{\tilde\nu}$-measurable set in $\mathcal{M} \times \widetilde{\mathcal{M}}$ and z_0 be a fixed point in $\widetilde{\mathcal{M}}$. For every $\varepsilon > 0$ and $\delta > 0$ define T as a set of the points x for which*

$$\mathcal{H}^{\tilde\nu}\{z : (x,z) \in S, \widetilde{d}_{cc}(z_0,z) \le r\} \le \varepsilon r^{\tilde\nu} \quad \textit{for all } 0 < r < \delta.$$

Then the set T is measurable.

Really, for any $r > 0$, a set

$$S_r = S \cap \{(x,z) : \widetilde{d}_{cc}(z_0,z) \le r\} = S \cap (\mathcal{M} \times \overline{\widetilde{B}_{cc}(z_0,r)})$$

is $\mathcal{H}^\nu \times \mathcal{H}^{\tilde\nu}$-measurable. By Tonelli–Fubini theorem the set $\{z : (x,z) \in S_r\}$ is $\mathcal{H}^{\tilde\nu}$-measurable for $\mathcal{H}^\nu$-almost all x and

$$\iint\limits_{\mathcal{M}\times\widetilde{\mathcal{M}}} \chi_{S_r}(x,z)\,dx\,dz = \int\limits_{\mathcal{M}}\int\limits_{\widetilde{\mathcal{M}}} \chi_{S_r}(x,z)\,dz\,dx = \int\limits_{\mathcal{M}} \mathcal{H}^{\tilde\nu}\{z : (x,z) \in S_r\}\,dx.$$

Consequently, the mapping

$$\varphi : x \mapsto \int\limits_{\widetilde{\mathcal{M}}} \chi_{S_r}(x,z)\,dz = \mathcal{H}^{\tilde\nu}\{z : (x,z) \in S_r\}$$

is $\mathcal{H}^\nu$-measurable. Then we have

$$T = \bigcap_{r\in(0,\delta)\cap\mathbb{Q}} \{x : \varphi(x) \le \varepsilon r^\nu\}$$

where $\mathbb{Q}$ denotes the set of rational numbers. It remains only to note that every set $\{x : \varphi(x) \le \varepsilon r^\nu\}$ is $\mathcal{H}^\nu$-measurable.

Property 3.2. *If $\sigma : \mathcal{M} \times \widetilde{\mathcal{M}} \to \overline{\mathbb{R}}$ is $\mathcal{H}^\nu \times \mathcal{H}^{\tilde\nu}$-measurable real-valued mapping and z_0 is a point in $\widetilde{\mathcal{M}}$ then*

$$\operatorname{ap}\overline{\lim_{z\to z_0}}\,\sigma(x,z) \quad \textit{and} \quad \operatorname{ap}\underline{\lim_{z\to z_0}}\,\sigma(x,z)$$

are $\mathcal{H}^\nu$-measurable mappings of argument x.

First, notice that

$$\{x \in \mathcal{M} : \operatorname{ap}\overline{\lim_{z\to z_0}}\,\sigma(x,z) \le \tau\} = \bigcap_{t>\tau} A_t = \bigcap_{n=1}^{\infty} A_{\tau+\frac{1}{n}},$$

where A_t is a set of the points $x \in \mathcal{M}$ for which the set $\{z \in \widetilde{\mathcal{M}} : \sigma(x,z) > t\}$ has the density zero at z_0. We have to make sure that A_t is measurable. In order to do this we apply Property 3.1 to the set

$$S_t = \{(x,z) \in \mathcal{M} \times \widetilde{\mathcal{M}} : \sigma(x,z) > t\}$$

and derive that the set $T_t(m,k)$ of the points $x \in \mathcal{M}$ for which

$$\mathcal{H}^\nu\{z : (x,z) \in S_t,\ \widetilde{d}_{cc}(z_0,z) \le r\} \le \frac{r^\nu}{m} \ \text{ for all } 0 < r < k^{-1},$$

is measurable for all positive integers m and k. It remains only to observe that

$$A_t = \bigcap_{m=1}^{\infty}\bigcap_{k=1}^{\infty} T_t(m,k).$$

3.3 Differentiability in the sub-Riemannian geometry

Fix $E \subset \mathbb{R}$ and a limit point $s \in E$. The mapping $\gamma : E \to \mathcal{M}$ has *sub-Riemannian derivative* at the point s if there is an element $a \in \mathcal{G}^{\gamma(s)}$ such that

$$d_{cc}^{\gamma(s)}(\gamma(s+t), \delta_t^{\gamma(s)} a) = o(t) \text{ as } t \to 0,\ s+t \in E. \tag{3.1}$$

We use the notation $a = \frac{d}{dt\,sub}\gamma(t+s)|_{t=0}$. A derivative is called *horizontal* if $a \in \exp(H_{\gamma(s)}\mathcal{M})$, i. e.

$$a = \exp\Big(\sum_{j=1}^{\dim H_1} \alpha_j \widehat{X}_j^{\gamma(s)}\Big)(\gamma(s)) = \exp\Big(\sum_{j=1}^{\dim H_1} \alpha_j X_j\Big)(\gamma(s))$$

for certain $\alpha_j \in \mathbb{R}$.

Recall that $\gamma : E \subset \mathbb{R} \to \mathcal{M}$ is called a *Lipschitz* mapping if there is a constant $C > 0$ such that the inequality

$$d_{cc}(\gamma(x), \gamma(y)) \leq C|x-y|$$

holds for all $x, y \in E$.

Theorem 3.2 ([33]). *Every Lipschitz mapping $\gamma : E \to \mathcal{M}$, where the set $E \subset \mathbb{R}$ is closed, has horizontal derivative almost everywhere in E.*

The mapping $f : E \subset \mathcal{M} \to \widetilde{\mathcal{M}}$ of two Carnot–Carathéodory spaces is called [35] differentiable at the point $g \in E$ if there is horizontal homomorphism $L : \mathcal{G}^g \to \mathcal{G}^{f(g)}$ of the local Carnot groups such that

$$\widetilde{d}_{cc}^{f(g)}(f(v), L(v)) = o(d_{cc}^g(g, v)) \text{ as } E \cap \mathcal{G}^g \ni v \to g. \tag{3.2}$$

Recall that *the horizontal homomorphism* of Carnot groups is a homomorphism $L : \mathbb{G} \to \widetilde{\mathbb{G}}$ such that $DL(0)(H\mathbb{G}) \subset H\widetilde{\mathbb{G}}$.

Local approximation theorem (Theorem 2.7) gives an opportunity to use both metrics of the initial space and of local Carnot group in the definition (3.2). Indeed, by Theorem 2.7 we have

$$\widetilde{d}_{cc}(f(v), L(v)) = \widetilde{d}_{cc}^{f(g)}(f(v), L(v)) + o\big(\widetilde{d}_{cc}^{f(g)}(f(g), f(v))\big) + o\big(\widetilde{d}_{cc}^{f(g)}(f(g), L(v))\big).$$

Using the triangle inequality

$$\widetilde{d}_{cc}(f(g), f(v)) \leq \widetilde{d}_{cc}(f(g), L(v)) + \widetilde{d}_{cc}(L(v), f(v)),$$

and homogeneity of L we obtain

$$\begin{aligned} \widetilde{d}_{cc}(f(v), L(v)) &= \widetilde{d}_{cc}^{f(g)}(f(v), L(v)) + o\big(\widetilde{d}_{cc}^{f(g)}(f(g), f(v))\big) + o\big(\widetilde{d}_{cc}^{f(g)}(f(g), L(v))\big) \\ &= [1+o(1)]\widetilde{d}_{cc}^{f(g)}(f(v), L(v)) + o\Big(d_{cc}^g(v, g) \sup_{u:\, d_{cc}^g(u,g)=1} \widetilde{d}_{cc}^{f(g)}(f(g), L(u))\Big) \\ &= o(d_{cc}^g(g, v)) = o(d_{cc}(g, v)). \end{aligned} \tag{3.3}$$

The homomorphism $L : \mathcal{G}^g \to \mathcal{G}^{f(g)}$ satisfying (3.2) is called *the differential* of the mapping f and is denoted by $D_g f$. One can show that if g is the density point then the differential is unique. Moreover, it is easy to verify that differential commutes with the one-parameter dilation group:

$$\tilde{\delta}_t^{f(g)} \circ D_g f = D_g f \circ \delta_t^g. \tag{3.4}$$

If $v \in \mathcal{G}^g$ and $\delta_t^g v \in \mathcal{G}^g$ then, by (3.4), we have

$$\begin{aligned}\widetilde{d}_{cc}^{f(g)}(f(\delta_t^g v), \tilde{\delta}_t^{f(g)} D_g f(v)) &= \widetilde{d}_{cc}^{f(g)}(f(\delta_t^g v), D_g f(\delta_t^g v)) \\ &= o(d_{cc}^g(g, \delta_t^g v)) = d_{cc}^g(g, v) o(t), \end{aligned} \tag{3.5}$$

i. e. element $D_g f(v)$ is a derivative of the curve $\gamma(t) = f(\delta_t^g v)$ at $t = 0$.

By *the derivative* of the mapping f along the horizontal vector field X at the point g we mean the derivative of the curve

$$\gamma(t) = f(\delta_t^g \exp \widehat{X}^g(g)) = f(\exp tX(g))$$

for $t = 0$. We use the notation $Xf(g)$ to denote this derivative. To be more precise we have to write $\widetilde{\exp} Xf(g)$ since usually $Xf(g)$ is the Riemannian derivative $\frac{d}{dt} f(\exp(tX)(g))\big|_{t=0}$. To simplify notations we will use $Xf(g)$ for the sub-Riemannian derivative except of the cases when the opposite is stated explicitly.

The mapping $f : E \subset \mathcal{M} \to \widetilde{\mathcal{M}}$ of two Carnot–Carathéodory spaces is called a *Lipschitz* mapping if there is a constant $C > 0$ such that the inequality

$$\widetilde{d}_{cc}(f(x), f(y)) \leq C d_{cc}(x, y)$$

holds for all $x, y \in E$.

In the work [33] there were generalized the classical Rademacher [23] and Stepanoff [27] theorems to the case of Carnot–Carathéodory spaces.

Theorem 3.3 ([33, Theorem 4.1]). *Let E be a set in $\mathcal{M}$ and let $f : E \to \widetilde{\mathcal{M}}$ be a Lipschitz mapping. Then f is differentiable almost everywhere in E and the differential is unique.*

Theorem 3.4 ([33, Theorem 5.1]). *Let E be a set in $\mathcal{M}$ and let a mapping $f : E \to \widetilde{\mathcal{M}}$ satisfy the condition*

$$\varlimsup_{x \to a, x \in E} \frac{\widetilde{d}_{cc}(f(a), f(x))}{d_{cc}(a, x)} < \infty$$

for almost all $a \in E$. Then f is differentiable almost everywhere in E and the differential is unique.

Here we will write an alternative proof of Theorems 3.3 and 3.4 using the theorem on approximate differentiability.

3.4 Approximate differentiability

Now we replace a regular limit in (3.1) by the approximate one. This leads us to definition of an approximate (horizontal) derivative as an element $a \in \exp H\mathcal{G}^{\gamma(s)}$ such that

$$\operatorname{ap}\lim_{t\to 0} \frac{d_{cc}^{\gamma(s)}(\gamma(s+t), \delta_t^{\gamma(s)}a)}{|t|} = 0,$$

i. e. the set

$$\{t \in (-r, r) : d_{cc}^{\gamma(s)}(\gamma(s+t), \delta_t^{\gamma(s)}a) > |t|\varepsilon\}$$

has density zero at the point $t = 0$ for an arbitrary $\varepsilon > 0$.

Similarly an approximate differential is the horizontal homomorphism $L : \mathcal{G}^g \to \mathcal{G}^{f(g)}$ of the local Carnot groups such that

$$\operatorname{ap}\lim_{v\to g} \frac{\widetilde{d}_{cc}^{f(g)}(f(v), L(v))}{d_{cc}^g(g, v)} = 0,$$

i. e. a set

$$\{v \in B_{cc}(g, r) \cap \mathcal{G}^g : \widetilde{d}_{cc}^{f(g)}(f(v), L(v)) > d_{cc}^g(g, v)\varepsilon\}$$

has $\mathcal{H}^\nu$-density zero at the point $v = g$ for any $\varepsilon > 0$. We denote such homomorphism as $\operatorname{ap} D_g f$.

Using the notion of an approximate differential we can generalize Theorem 3.4 in the following direction.

Theorem 3.5. *Let E be a set in $\mathcal{M}$ and let $f : E \to \widetilde{\mathcal{M}}$ meet the condition*

$$\operatorname{ap}\overline{\lim_{x\to g}} \frac{\widetilde{d}_{cc}(f(g), f(x))}{d_{cc}(g, x)} < \infty. \tag{3.6}$$

Then f is approximately differentiable almost everywhere in E.

For proving Theorem 3.5 we need the following statement.

Theorem 3.6. *Let E be a measurable subset in $\mathcal{M}$ and $f : E \to \widetilde{\mathcal{M}}$ be a measurable mapping enjoying (3.6) for all points $g \in E$. Then there is a sequence of disjoint sets $E_0, E_1, \dots,$ such that $E = E_0 \cup \bigcup\limits_{i=1}^{\infty} E_i$, $\mathcal{H}^\nu(E_0) = 0$ and every restriction $f|_{E_i}$, $i = 1, 2, \dots,$ is a Lipschitz mapping.*

Proof. Since our considerations are local, we focus our arguments on the case when $E \subset U$ where U is an open subset in $\mathcal{M}$. Consider a sequence of sets

$$U_m = \{x \in U : d_{cc}(x, \partial U) \geq 2m^{-1}\}, \quad m \in \mathbb{N}.$$

Each U_m is closed and $\bigcup\limits_{m=1}^{\infty} U_m = U$. For all distinct points u and v of U the relation

$$h(u, v) = \frac{\mathcal{H}^\nu(B_{cc}(u, d_{cc}(u, v)) \cap B_{cc}(v, d_{cc}(u, v)))}{d_{cc}(u, v)^\nu}, \quad u \neq v,$$

is a continuous real-valued function. For every m define a constant

$$\gamma_m = \inf\{h(u,v) : u, v \in U_m,\ d_{cc}(u,v) \leq m^{-1}\}.$$

Let $d_{cc}(u,v) = l$. By definition of d_{cc} for an arbitrary number $\varepsilon > 0$ there exists a piecewise smooth path $\gamma : [0, l+\varepsilon] \to \mathcal{M}$ such that $\gamma(0) = u$, $\gamma(l+\varepsilon) = v$ and $|\dot{\gamma}| \leq 1$. Let $w = \gamma(\frac{l+\varepsilon}{2})$. Then $d_{cc}(u,w) \leq \frac{l+\varepsilon}{2}$ and $d_{cc}(v,w) \leq \frac{l+\varepsilon}{2}$. Consequently, $B_{cc}(w, \frac{l-\varepsilon}{2}) \subset B_{cc}(u,l)$ and $B_{cc}(w, \frac{l-\varepsilon}{2}) \subset B_{cc}(v,l)$. Hence,

$$h(u,v) \geq \frac{\mathcal{H}^\nu\big(B_{cc}(w, \frac{l-\varepsilon}{2})\big)}{l^\nu} \geq \frac{C_1(\frac{l-\varepsilon}{2})^\nu}{l^\nu} > 0,$$

where $C_1 > 0$ is a constant from Ball–Box theorem. Since $\varepsilon > 0$ is arbitrary, we infer $\gamma_m \geq C_1 2^{-\nu} > 0$.

For every $m \in \mathbb{N}$, let E^m be a set of all density points of $E \cap (U_m \setminus U_{m-1})$ (assuming $U_0 = \emptyset$). The sequence E^m is a disjoint family and $\mathcal{H}^\nu(E \setminus \bigcup\limits_{m=1}^{\infty} E_m) = 0$.

For $k \in \mathbb{N}$, $u \in E$, $0 < r < m^{-1}$ define

$$Q_k^m(u,r) = B_{cc}(u,r) \cap \{x : x \notin E^m \text{ or } \widetilde{d}_{cc}(f(x), f(u)) > k\, d_{cc}(x,u)\}$$

and also define

$$B_k^m = E \cap \Big\{u : \mathcal{H}^\nu(Q_k^m(u,r)) < \gamma_m \frac{r^\nu}{2} \text{ for all } 0 < r < \min\{k^{-1}, m^{-1}\}\Big\}.$$

By Property 3.1, all B_k^m are measurable and $E^m = \bigcup\limits_{k=1}^{\infty} B_k^m$. Next, if $u, v \in B_k^m$ and $r = d_{cc}(u,v) < \min\{k^{-1}, m^{-1}\}$ we have

$$\mathcal{H}^\nu(Q_k^m(u,r) \cup Q_k^m(v,r)) < \gamma_m r^\nu \leq \mathcal{H}^\nu(B_{cc}(u,r) \cap B_{cc}(v,r)).$$

Hence we can choose a point

$$x \in (B_{cc}(u,r) \cap B_{cc}(v,r)) \setminus (Q_k^m(u,r) \cup Q_k^m(v,r)).$$

For this point

$$\begin{aligned}\widetilde{d}_{cc}(f(u), f(v)) &\leq \widetilde{d}_{cc}(f(u), f(x)) + \widetilde{d}_{cc}(f(x), f(v))\\ &\leq k d_{cc}(u,x) + k d_{cc}(x,v) \leq 2kr = 2k d_{cc}(u,v).\end{aligned}$$

Consequently, representing B_k^m as an union of a countable family of measurable sets $B_{k,j}^m$, whose diameters are less than $\min\{k^{-1}, m^{-1}\}$, we see that every restriction $f|_{B_{k,j}^m}$ is a Lipschitz mapping. □

Proof of Theorem 3.5. By Theorem 3.6 the domain of f is the union of a countable family of disjoint sets E_i (up to the set of measure 0) such that every restriction $f|_{E_i}$ is a Lipschitz mapping. By Theorem 3.3 every $f|_{E_i}$ is differentiable almost everywhere in E_i. For the density points of E_i this is equivalent to the approximate differentiability in E. □

4 Theorem on approximate differentiability

Now we have all necessary tools for formulating and proving the main result.

Theorem 4.1. *Let $E \subset \mathcal{M}$ be a measurable subset of the Carnot–Carathéodory space $\mathcal{M}$ and let $f : E \to \widetilde{\mathcal{M}}$ be a measurable mapping. The following statements are equivalent:*

1) *The mapping f is approximately differentiable almost everywhere in E.*

2) *The mapping f has approximate derivatives $\operatorname{ap} X_j f$ along the basic horizontal vector fields $X_1, \dots, X_{\dim H_1}$ almost everywhere in E.*

3) *There is a sequence of disjoint sets $Q_1, Q_2, \dots$ such that $\mathcal{H}^\nu(E \setminus \bigcup\limits_{i=1}^{\infty} Q_i) = 0$ and every restriction $f|_{Q_i}$ is a Lipschitz mapping.*

Proof of the implication 1) $\Rightarrow$ 3). Let $g \in \mathcal{M}$ be a density point of E and Let f be approximately differentiable in g. Fix a point v in a set

$$C_\varepsilon(g) = \{v \in B_{cc}(g, r_g) \cap \mathcal{G}^g : \widetilde{d}_{cc}(f(v), \operatorname{ap} D_g f(v)) < \varepsilon d_{cc}(g, v)\}, \quad \varepsilon > 0.$$

By Theorem 2.7 we have

$$\begin{aligned}\widetilde{d}_{cc}^{f(g)}(f(v), \operatorname{ap} D_g f(v)) &\le \widetilde{d}_{cc}(f(v), \operatorname{ap} D_g f(v))[1 + o(1)] \\ &< d_{cc}(v, g)[\varepsilon + o(\varepsilon)] = d_{cc}^g(v, g)[\varepsilon + o(\varepsilon)].\end{aligned}$$

From the definition of an approximate differential it follows that $\mathcal{H}^\nu$-density of the set $B_{cc}(g, r_g) \setminus C_\varepsilon(g)$ equals zero for any $\varepsilon > 0$. In other words

$$\operatorname{ap}\lim_{v \to g} \frac{\widetilde{d}_{cc}(f(v), D_g f(v))}{d_{cc}(g, v)} = 0.$$

Therefore,

$$\begin{aligned}&\operatorname{ap}\overline{\lim_{v \to g}} \frac{\widetilde{d}_{cc}(f(g), f(v))}{d_{cc}(g, v)} \\ &\le \operatorname{ap}\overline{\lim_{v \to g}} \frac{\widetilde{d}_{cc}(f(g), D_g f(v))}{d_{cc}(g, v)} + \operatorname{ap}\overline{\lim_{v \to g}} \frac{\widetilde{d}_{cc}(D_g f(v), f(v))}{d_{cc}(g, v)} \\ &= \overline{\lim_{v \to g}} \frac{\widetilde{d}_{cc}(f(g), D_g f(v))}{d_{cc}^g(g, v)} + 0 \\ &\le \overline{\lim_{v \to g}} \frac{\widetilde{d}_{cc}^{f(g)}(f(g), D_g f(v))[1 + o(1)]}{d_{cc}^g(g, v)} \\ &= [1 + o(1)] \sup_{v :\, d_{cc}^g(v, g) = 1} \widetilde{d}_{cc}^{f(g)}(f(g), D_g f(v)) < \infty\end{aligned}$$

for almost all $g \in E$. Hence, the conditions of Theorem 3.6 are fulfilled. □

The implication 3) $\Rightarrow$ 2) is proved as Corollary 4.1 in the next subsection.

The implication 2) $\Rightarrow$ 1) is a direct corollary of the following crucial

Theorem 4.2. *Let $f : \mathcal{M} \to \widetilde{\mathcal{M}}$ be a measurable mapping of Carnot–Carathéodory spaces. Then*

$$A_j = \operatorname{dom} \operatorname{ap} X_j f \ \text{ is a measurable set,}$$
$$\operatorname{ap} X_j f : A_j \to \widetilde{\exp}(H\widetilde{\mathcal{M}}) \ \text{ is a measurable mapping in } A_j,$$

for all $j = 1, \dots, \dim H_1$, and f is approximately differentiable almost everywhere on the set $A = \bigcap\limits_{j=1}^{\dim H_1} A_j$. Moreover, if $g \in A$ is a point of an approximate differentiability of the mapping f and in the neighborhood of g we have representation from Theorem 2.5

$$v = \exp(a_L X_{j_L}) \circ \dots \circ \exp(a_1 X_{j_1})(g)$$

where $1 \le j_i \le \dim H_1$, $i = 1, \dots, L$, $L \in \mathbb{N}$, then

$$\operatorname{ap} D_g f(v) = \prod_{i=1}^{L} \delta_{a_i}^{f(g)} \operatorname{ap} X_{j_i} f(g) \in \mathcal{G}^{f(g)}.$$

We follow the proof in [31] where the similar result was established for Carnot groups (which in turn was inspired by the proof [4] of the similar theorem for mappings of Euclidean spaces). The essential steps of the proof are carried out in separate lemmas which are proved below and the proof of the theorem itself is located in the subsection 4.3 just after proofs of lemmas.

4.1 Approximate derivatives

Lemma 4.1. *Let $E \subset \mathcal{M}$ be a measurable set and $f : E \to \widetilde{\mathcal{M}}$ be a measurable mapping. Then*

$$A_j = \{x \in E : \operatorname{ap} \varlimsup_{t \to 0} \frac{\widetilde{d}_{cc}(\, f(x),\ f(\exp tX_j(x))\,)}{|t|} < \infty\} \ \text{ is measurable;}$$
$$\operatorname{ap} X_j f : A_j \to \widetilde{\mathcal{M}} \ \text{ is defined almost everywhere and is measurable;}$$
$$\operatorname{ap} X_j f(g) \in \widetilde{\exp}(H_g\widetilde{\mathcal{M}}) \ \text{ for almost all } g \in A_j$$

for every $j = 1, \dots, \dim H_1$.

Proof. Fix $j \in \{1, \dots, \dim H_1\}$. A mapping

$$t \mapsto |t|^{-1} \widetilde{d}_{cc}(\, f(x), f(\exp tX_j(x))\,)$$

is measurable and by Property 3.2 the set A_j is measurable. For every $x \in E$ define A_x as a set of real numbers t such that $\exp tX_j(x) \in A_j$. In the case $A_x \neq \emptyset$ define also the mapping $h : A_x \to \widetilde{\mathcal{M}}$ by the rule $h(t) = f(\exp tX_j(x))$.

If $y = \exp tX_j(x)$, $t \in A_x$, we have

$$\begin{aligned}
&\operatorname{ap}\overline{\lim_{\tau\to 0}} \frac{\widetilde{d}_{cc}(h(t), h(t+\tau))}{|\tau|}\\
&= \operatorname{ap}\overline{\lim_{\tau\to 0}} \frac{\widetilde{d}_{cc}(f(\exp tX_j(x)), f(\exp(t+\tau)X_j(x)))}{|\tau|}\\
&= \operatorname{ap}\overline{\lim_{\tau\to 0}} \frac{\widetilde{d}_{cc}(f(\exp tX_j(x)), f(\exp \tau X_j(\exp tX_j(x))))}{|\tau|}\\
&= \operatorname{ap}\overline{\lim_{\tau\to 0}} \frac{\widetilde{d}_{cc}(f(y), f(\exp \tau X_j(y)))}{|\tau|} < \infty.
\end{aligned}$$

Hence, h meets the conditions of Theorem 3.6. Therefore, $A_x = B_0 \cup \bigcup\limits_{i=1}^{\infty} B_i$, where $\mathcal{H}^\nu(B_0) = 0$, all B_i, $i = 1, \dots, \infty$, are measurable and restriction of h to every B_i is a Lipschitz mapping. If $h : B_i \to \widetilde{\mathcal{M}}$ is one of these restrictions then by Theorem 3.2 the sub-Riemannian derivative

$$\frac{d}{d\tau}_{sub} h(t+\tau)\Big|_{\substack{\tau=0\\ t+\tau\in B_i}} \in \widetilde{\exp} H_{h(t)}\widetilde{\mathcal{M}}$$

exists for almost all t. If t is a density point for the set B_i then

$$\begin{aligned}
\frac{d}{d\tau}_{sub} h(t+\tau)\Big|_{\substack{\tau=0\\ t+\tau\in B_i}} &= \operatorname{ap} \frac{d}{d\tau}_{sub} h(t+\tau)\Big|_{\tau=0}\\
&= \operatorname{ap} \frac{d}{d\tau}_{sub} f(\exp \tau X_j(y))\Big|_{\tau=0} = \operatorname{ap} X_j f(y).
\end{aligned}$$

Thus, $\operatorname{ap} X_j f(y)$ exists at $\{y = \exp tX_j(x) : t \in A_x\}$ for almost all $t \in A_x$. This provides existence of the derivative $\operatorname{ap} X_j f$ almost everywhere in A_j. $\square$

Corollary 4.1. *A Lipschitz mapping f has approximate derivatives $\operatorname{ap} X_j f$ along the horizontal vector fields X_j almost everywhere and $\operatorname{ap} X_j f(g) \in \widetilde{\exp}(H_g\mathcal{M})$ for almost all $g \in \operatorname{dom} f$.*

Remark 4.1. Note that if $\operatorname{ap} X_j f(g)$ defined at $g \in \mathcal{M}$ then $\operatorname{ap}(aX_j) f(g)$ is also defined for all real numbers a. Moreover

$$\operatorname{ap}(aX_j)f(g) = \tilde{\delta}_a^{f(g)} \operatorname{ap} X_j f(g).$$

Let the coordinate system (2.13) be defined in a neighborhood of a point $g \in \mathcal{M}$. Consider a curve

$$\Gamma_k(g;t) = \Phi_k(t)(g). \tag{4.1}$$

We say that the mapping f is *approximately differentiable* along the curve $\Gamma_k(g;t)$ at $t = 0$ if there is an element $a \in \mathcal{G}^{f(g)} \cap \widetilde{\mathcal{M}}$ such that

$$\frac{1}{r^{\deg X_k}} \mathcal{H}^{\deg X_k}\Big\{t \in (-r,r) : \frac{\widetilde{d}_{cc}^{f(g)}(f\circ\Gamma_k(g;t), \tilde{\delta}_t^{f(g)} a)}{d_{cc}^g(g, \Gamma_k(g;t))} > \varepsilon\Big\} \to 0 \quad as \quad r \to 0.$$

We denote this derivative by $a = \operatorname{ap} d_{sub}(f\circ\Gamma_k)(g)$. If $k = 1, \dots, \dim H_1$, this definition coincides with the definition of the approximate derivative from Subsection 3.4.

Lemma 4.2. *Let $E \subset \mathcal{M}$ be a bounded measurable set and $f : E \to \widetilde{\mathcal{M}}$ be a measurable mapping. Let also the coordinate system* (2.13) *be defined in a neighborhood of a point $g \in U$ with functions Φ_k satisfying* (2.12). *Then the mapping f is approximately differentiable along the curve $\Gamma_k(g;t)$ defined by* (4.1), *$k = \dim H_1 + 1, \dots, N$, at $t = 0$ almost everywhere in $A = \bigcap\limits_{j=1}^{\dim H_1} \operatorname{dom} \operatorname{ap} X_j f$. Moreover, the approximate derivative can be written as*

$$\begin{aligned} \operatorname{ap} d_{sub}(f \circ \Gamma_k)(g) &= \operatorname{ap}(s_{L_k} \widehat{X}^g_{j_{L_k}}) f \circ \dots \circ \operatorname{ap}(s_1 \widehat{X}^g_{j_1}) f(g) \\ &= \operatorname{ap}(s_1 \widehat{X}^g_{j_1}) f(g) \cdot \ldots \cdot \operatorname{ap}(s_{L_k} \widehat{X}^g_{j_{L_k}}) f(g) \in \mathcal{G}^{f(g)} \end{aligned} \tag{4.2}$$

almost everywhere. Here $L_k \le L$ and $s_i = \pm 1$ are from the representation (2.12). *Also the following estimate*

$$\begin{aligned} &\widetilde{d}^{f(g)}_{cc}\big(f(g), \operatorname{ap} d_{sub}(f \circ \Gamma_k)(g)\big) \\ &\qquad \le L_k \max\{\widetilde{d}_{cc}\big(f(g), \operatorname{ap} X_j f(g)\big) : j = 1, \dots, \dim H_1\} \end{aligned} \tag{4.3}$$

holds for all $k = \dim H_1 + 1, \dots, N$.

A sketch of the proof:

At the first step we apply Luzin's and Egorov's theorems to a bounded set A and obtain a set $A' \subset A$ that differs from A on a set of a measure small enough and on which the limit $\operatorname{ap}\lim\limits_{t\to 0} \widetilde{\delta}^{f(g)}_{t^{-1}} f(\exp tX_j)(g)$ converges to $\operatorname{ap} X_j f(g)$ uniformly.

Next we assure that the set of real numbers t, for which the relation (4.2) does not hold, is negligible.

At last, we prove that $\operatorname{ap}\lim\limits_{t\to 0} \widetilde{\delta}^{f(g)}_{t^{-1}} f \circ \Gamma_k(g;t)$ converges uniformly to (4.2) in A'.

Proof. By Lemma 4.1 the sets $A_j = \operatorname{dom} \operatorname{ap} X_j f \subset E$ are measurable and the mappings $\operatorname{ap} X_j f$ are measurable in A_j for all $j = 1, \dots, \dim H_1$.

We have $\mathcal{H}^\nu(A_j) \le \mathcal{H}^\nu(E) < \infty$. Fix $\varepsilon > 0$. Applying Luzin's theorem we find a closed set $E' \subset A$ such that $\mathcal{H}^\nu(A \setminus E') < \varepsilon/2$ and all $\operatorname{ap} X_j f$ are uniformly continuous in E'.

Consider a sequence of functions $\{\varphi^j_n : E' \to \mathbb{R}\}_{n\in\mathbb{N}}$ defined as

$$\varphi^j_n(g) = \sup_{|t|<\frac{1}{n}} \frac{\widetilde{d}^{f(g)}_{cc}\big(f(\exp(tX_j)(g)), \widetilde{\delta}^{f(g)}_t \operatorname{ap} X_j f(g)\big)}{|t|}, \quad j = 1, \dots, \dim H_1.$$

Since $\varphi^j_n(g) \underset{ap}{\to} 0$ as $n \to \infty$, by Egorov's theorem we obtain a measurable set $E'' \subset E'$ such that $\mathcal{H}^\nu(E' \setminus E'') < \varepsilon/2$ and $\varphi^j_n(g) \to 0$ as $n \to \infty$ uniformly on E''. Therefore, the limits

$$\operatorname{ap}\lim_{t\to 0} \frac{\widetilde{d}^{f(g)}_{cc}\big(f(\exp(tX_j)(g)), \widetilde{\delta}^{f(g)}_t \operatorname{ap} X_j f(g)\big)}{|t|} = 0$$

converge uniformly on E'' for all $j = 1, \dots, \dim H_1$.

For every positive integer m and for all $x \in E$, $r > 0$ define a set

$$T_j^m(x,r) = \Big\{t \in (-r,r) : \widetilde{d}_{cc}^{f(x)}\big(f(\exp tX_j(x)), \tilde{\delta}_t^{f(g)} \operatorname{ap} X_j f(x)\big) > \frac{|t|}{m}\Big\}.$$

For all positive integers p we introduce

$$B_j^m(p) = A_j \cap \Big\{x \in E : \mathcal{H}^1[T_j^m(x,r)] \le \frac{r}{m} \text{ for all } 0 < r < p^{-1}\Big\}.$$

By Property 3.1 the sets $B_j^m(p)$ are measurable for all $j = 1, \dots, \dim H_1$. We have also $\bigcup\limits_{p=1}^{\infty} B_j^m(p) = A_j$.

Moreover, $B_j^m(p) \subset B_j^m(p+1)$. Hence, we can choose a sequence of numbers $p_1, p_2, \dots$ such that $\mathcal{H}^\nu(E'' \setminus B_j^m(p_m)) < \frac{\varepsilon}{2^m}$ holds. Therefore,

$$\mathcal{H}^\nu(E'' \setminus F) < \varepsilon{\cdot}\dim H_1, \qquad \text{where } F = \bigcap_{j=1}^{\dim H_1} \bigcap_{m=1}^{\infty} B_j^m(p_m).$$

Next, for all $x \in F$, $r > 0$ define a set

$$Z_j(x,r) = \{y = \exp tX_j(x) : |t| < r \text{ and } y \notin F\}, \quad j = 1, \dots, \dim H_1.$$

For all positive integers m and q define the sets

$$C_j^m(q) = F \cap \Big\{x \in E : \mathcal{H}^1[Z_j(x,r)] \le \frac{r}{2^m} \text{ for all } 0 < r < q^{-1}\Big\}.$$

By Property 3.1 all $C_j^m(q)$ are measurable. Also $\mathcal{H}^\nu\Big(F \setminus \bigcup\limits_{q=1}^{\infty} C_j^m(q)\Big) = 0$.

Moreover, $C_j^m(q) \subset C_j^m(q+1)$. Hence, we can choose a sequence of numbers $q_1, q_2, \dots$ such that $\mathcal{H}^\nu(F \setminus C_j^m(q_m)) < \frac{\varepsilon}{2^m}$ holds. Therefore,

$$\mathcal{H}^\nu(F \setminus F_1) < m\varepsilon, \qquad \text{where } F_1 = \bigcap_{j=1}^{\dim H_1} \bigcap_{n=1}^{\infty} C_j^n(q_n).$$

Next, we prove that the function f is approximately differentiable along the curve $\Gamma_k(g;t)$ uniformly in F_1 and the mapping $g \mapsto \operatorname{ap}\frac{d}{dt}_{sub} f(\Gamma_k(g;t))\big|_{t=0}$ is uniformly continuous in F_1.

Fix $m \in \mathbb{N}$, $0 < r < \min\{p_m^{-1}, q_m^{-1}\}$ and a density point $g \in F_1$. Denote

$$\begin{aligned} u_1(t) &= \exp(ts_1X_{j_1})(g), \\ u_i(t) &= \exp(ts_iX_{j_i})(u_{i-1}(t)), \quad i = 2, \dots, L_k. \end{aligned}$$

Then $u_{L_k}(t) = \Gamma_k(g;t)$. Define the set $S^m \subset (-r,r)$ as follows:

$$\begin{aligned} t \in S^m, \ \text{if } & s_1 t \in T_{j_1}^m(g,r), \\ \text{or } & s_i t \in T_{j_i}^m(u_{i-1}(t),r), \\ \text{or } & u_1(t) \in Z_{j_1}(g,r), \\ \text{or } & u_i(t) \in Z_{j_i}(u_{i-1}(t),r), \quad i = 2, \dots, L_k. \end{aligned}$$

Since $\mathcal{H}^1[T^m_{j_1}(g,r)] \leq \frac{r}{m}$, $\mathcal{H}^1[Z_{j_1}(g,r)] \leq \frac{r}{m}$ and since $\mathcal{H}^1[T^m_{j_i}(u_{i-1}(t),r)] \leq \frac{r}{m}$, $\mathcal{H}^1[Z_{j_i}(u_{i-1}(t),r)] \leq \frac{r}{m}$ if $u_{i-1}(t) \in F_1$, $i = 2,\dots,L_k$, we have

$$\mathcal{H}^1(S^m) \leq 2L_k\frac{r}{m}.$$

Now we estimate $\mathcal{H}^{\deg X_k}$-measure of the set S^m. Fix arbitrary numbers

$$\delta > 0 \text{ and } \Lambda > 2L_k\frac{r}{m}. \tag{4.4}$$

Cover the set S^m with a countable family of intervals (a_ξ, b_ξ) so that

$$b_\xi - a_\xi < \delta, \quad \sum_\xi (b_\xi - a_\xi) < \Lambda.$$

Then

$$|b_\xi - a_\xi|^{\deg X_k} < \delta(2r)^{\deg X_k - 1}, \quad \sum_\xi |b_\xi - a_\xi|^{\deg X_k} < \Lambda(2r)^{\deg X_k - 1}.$$

Since δ and Λ are arbitrary numbers of (4.4), we have

$$\mathcal{H}^{\deg X_k}(S^m) \leq 2^{\deg X_k} L_k \frac{r^{\deg X_k}}{m}.$$

Now we show that the expression (4.2) is really the derivative of the composition $f\circ\Gamma_k$. For the points $u,v\in F_1$ we have

$$\begin{aligned}
&\widetilde{d}^{f(g)}_{cc}\big(f(\exp(ts_iX_i)(v)),\, \mathrm{ap}(ts_iX_i)f(v)\big) \leq \varphi_i(t),\\
&\widetilde{d}^{f(g)}_{cc}\big(\mathrm{ap}(ts_iX_i)f(u),\, \mathrm{ap}(ts_iX_i)f(v)\big) \leq t\omega_i(d_{cc}(u,v)),
\end{aligned}$$

where $\frac{\varphi_i(t)}{t} \to 0$ as $t\to 0$ uniformly for $v\in F_1$ and $\omega_i(t)$ are moduli of continuity of the mappings $\mathrm{ap}(s_iX_i)f(\cdot)$ in F_1, $i=1,\dots,\dim H_1$.

If $|t|<r$ and $t\notin S^m$ we obtain

$$\begin{aligned}
&\widetilde{d}^{f(g)}_{cc}\big(f\circ u_1(t),\, \tilde{\delta}^{f(g)}_t \mathrm{ap}(s_1X_1)f(g)\big)\\
&\quad = \widetilde{d}^{f(g)}_{cc}\big(f\circ \exp(ts_1X_1)(g),\, \mathrm{ap}(ts_1X_1)f(g)\big)\\
&\quad \leq \varphi_1(t) = C_1(t).
\end{aligned}$$

Further, by induction:

$$\begin{aligned}
&\widetilde{d}^{f(g)}_{cc}\Big(f\circ u_j(t),\, \tilde{\delta}^{f(g)}_t \prod_{i=1}^{j}\mathrm{ap}(s_iX_i)f(g)\Big)\\
&= \widetilde{d}^{f(g)}_{cc}\Big(f\circ \exp(ts_jX_j)(u_{j-1}(t)),\, \mathrm{ap}(ts_jX_j)f\circ\prod_{i=1}^{j-1}\mathrm{ap}(ts_iX_i)f(g)\Big)\\
&\leq \widetilde{d}^{f(g)}_{cc}\big(f\circ \exp(ts_jX_j)(u_{j-1}(t)),\, \mathrm{ap}(ts_jX_j)f(u_{j-1}(t))\big)\\
&\qquad + \widetilde{d}^{f(g)}_{cc}\Big(\mathrm{ap}(ts_jX_j)f(u_{j-1}(t)),\, \mathrm{ap}(ts_jX_j)f\circ\prod_{i=1}^{j-1}\mathrm{ap}(ts_iX_i)f(g)\Big)\\
&\leq \varphi_j(t) + t\omega_j\big(C_{j-1}(t)\big) = C_j(t),
\end{aligned}$$

where $\frac{C_j(t)}{t} \to 0$ as $t \to 0$ uniformly for $g \in F_1$.

Therefore, for $t \in (-r, r) \setminus S^m$ we have an evaluation

$$\widetilde{d}_{cc}^{f(g)}\Big(f(\Gamma_k(g;t)),\, \tilde{\delta}_t^{f(g)} \prod_{i=1}^{L_k} \mathrm{ap}(s_i X_i) f(g)\Big) = o(t),$$

i. e. the equality

$$\mathrm{ap}\frac{d}{dt}_{sub} f(\Gamma_k(g;t)) = \prod_{i=1}^{L_k} \mathrm{ap}(s_i X_i) f(g)$$

holds for $g \in F_1$. Since r, m, ε are arbitrary the latter takes place almost everywhere in E. The inequality (4.3) follows from (4.2) and the generalized triangle inequality. □

Remark 4.2. Consider in the previous lemma the curve $\Gamma'_k(g;t) = \Gamma_k(g;\lambda t)$, $\lambda \in \mathbb{R} \setminus \{0\}$, instead of $\Gamma_k(g;t)$. The following representation takes place

$$\Gamma'_k(g;t) = \exp(\lambda t s_{L_k} X_{j_{L_k}}) \circ \cdots \circ \exp(\lambda t s_1 X_{j_1})(g),$$

where $s_i = \pm 1$, $1 \leq j_i \leq \dim H_1$. Then if there is $\mathrm{ap}\, d_{sub}(f \circ \Gamma_k)(g)$ defined at the point $g \in \mathcal{M}$ the derivative $\mathrm{ap}\, d_{sub}(f \circ \Gamma'_k)(g)$ is also defined and we have

$$\begin{aligned}\mathrm{ap}\, d_{sub}(f \circ \Gamma'_k)(g) &= \prod_{i=1}^{L_k} \mathrm{ap}(\lambda s_i X_i) f(g) \\ &= \tilde{\delta}_\lambda^{f(g)} \prod_{i=1}^{L_k} \mathrm{ap}(s_i X_i) f(g) = \tilde{\delta}_\lambda^{f(g)}\, \mathrm{ap}\, d_{sub}(f \circ \Gamma_k)(g). \qquad (4.5)\end{aligned}$$

4.2 Construction and properties of a mapping of local groups

Consider the system of the coordinates of the second kind (2.9) in a neighborhood $V(g) \subset \mathcal{G}^g$ of g. Define a mapping $L_g : V(g) \to \mathcal{G}^{f(g)}$ as follows:

$$L_g : \hat{v} = \widehat{\Phi}_g(t_1, \ldots, t_N) \mapsto \prod_{k=1}^{N} \tilde{\delta}_{t_k}^{f(g)}\, \mathrm{ap}\, d_{sub}(f \circ \Gamma_k)(g). \qquad (4.6)$$

Declare some properties of this mapping.

Property 4.1. *The mapping L_g is continuous.*

It follows directly from (4.6).

Property 4.2. $\tilde{\delta}_\lambda^{f(g)} \circ L_g = L_g \circ \delta_\lambda^g$.

Really, for $\hat{v} = \widehat{\Phi}_g(t_1, \ldots, t_N)$ we have

$$\begin{aligned}\delta_\lambda^g \hat{v} &= \delta_\lambda^g \widehat{\Phi}_g(t_1, \ldots, t_N) \\ &= \delta_\lambda^g \widehat{\Phi}_N(t_N) \circ \cdots \circ \widehat{\Phi}_{\dim H_1 + 1}(t_{\dim H_1 + 1}) \\ &\qquad \circ \exp(t_{\dim H_1} \widehat{X}^g_{\dim H_1}) \circ \cdots \circ \exp(t_1 \widehat{X}^g_1)(g) \\ &= \widehat{\Phi}_N(\lambda t_N) \circ \cdots \circ \widehat{\Phi}_{\dim H_1 + 1}(\lambda t_{\dim H_1 + 1}) \\ &\qquad \circ \exp(\lambda t_{\dim H_1} \widehat{X}^g_{\dim H_1}) \circ \cdots \circ \exp(\lambda t_1 \widehat{X}^g_1)(g) \\ &= \widehat{\Phi}_g(\lambda t_1, \ldots, \lambda t_N).\end{aligned}$$

Then, taking into account (4.5), we get

$$\begin{aligned} L_g(\delta^g_\lambda \hat{v}) &= \prod_{k=1}^{N} \tilde{\delta}^{f(g)}_{\lambda t_k} \operatorname{ap} d_{sub}(f \circ \Gamma_k)(g) \\ &= \tilde{\delta}^{f(g)}_{\lambda} \prod_{k=1}^{N} \tilde{\delta}^{f(g)}_{t_k} \operatorname{ap} d_{sub}(f \circ \Gamma_k)(g) = \tilde{\delta}^{f(g)}_{\lambda} L_g(\hat{v}). \end{aligned}$$

Property 4.3. *The mapping L_g is bounded.*

By Property 4.2 the mapping L_g is homogeneous, so

$$\|L_g\| = \sup_{v \neq g} \frac{\widetilde{d}^{f(g)}_{cc}(L_g(g), L_g(v))}{d^g_{cc}(g, v)} = \sup_{d^g_{cc}(g,v)=1} \widetilde{d}^{f(g)}_{cc}(L_g(g), L_g(v)).$$

The latter is finite because of continuity of L_g.

Property 4.4. *Let u, $v \in \mathcal{G}^g$ be such that $d^g_{cc}(u, v) = o(d^g_{cc}(g, u))$ as $u \to g$. Then*

$$\widetilde{d}^{f(g)}_{cc}(L_g(u), L_g(v)) = o(d^g_{cc}(g, u)).$$

Let $\omega(t)$ be a modulus of continuity of the mapping $L_g : B_{cc}(g, 2) \to \mathcal{G}^{f(g)}$. Then if we define $r = \max\{d^g_{cc}(g, u), d^g_{cc}(g, v)\}$ by Property 4.2 we have

$$\begin{aligned} \widetilde{d}^{f(g)}_{cc}(L_g(u), L_g(v)) &= O(r)\, \widetilde{d}^{f(g)}_{cc}(L_g(\delta^g_{r^{-1}} u), L_g(\delta^g_{r^{-1}} v)) \\ &\leq O(r)\, \omega\Big(\frac{d^g_{cc}(u, v)}{r}\Big) = r\, o(1) \quad \text{as } r \to 0. \end{aligned}$$

Lemma 4.3. *Let $E \subset \mathcal{M}$ be a bounded measurable set and let $f : E \to \widetilde{\mathcal{M}}$ be a measurable mapping. Let the coordinate system of the 2nd kind* (2.13) *be defined in a neighborhood of $g \in \mathcal{M}$. Then the mapping f is approximately differentiable along the curves $\Gamma_k(g; t)$, $k = 1, \dots, N$, almost everywhere in $A = \bigcap\limits_{j=1}^{\dim H_1} \operatorname{dom} \operatorname{ap} X_j f$ and the equality*

$$\operatorname*{ap\,lim}_{v \to g} \frac{\widetilde{d}^{f(g)}_{cc}(f(v), L_g(v))}{d^g_{cc}(g, v)} = 0 \tag{4.7}$$

holds for almost all $g \in A$, where L_g is the mapping defined by the formula (4.6).

Proof. By Lemma 4.1 all sets $A_j = \operatorname{dom} \operatorname{ap} X_j f$ are measurable and by Lemma 4.2 f is approximately differentiable along the curves Γ_k, $k = 1, \dots, N$, almost everywhere in A.

Fix $\varepsilon > 0$. By Luzin's theorem there is a measurable set $E' \subset A$ such that $\mathcal{H}^\nu(A \setminus E') < \varepsilon/2$ and the mapping $E' \ni x \mapsto \operatorname{ap} d_{sub}(f \circ \Gamma_k)(x)$ is uniformly continuous for all $k = 1, \dots, N$.

Consider a sequence of functions $\{\varphi^k_n : E' \to \mathbb{R}\}_{n \in \mathbb{N}}$ defined as

$$\varphi^k_n(g) = \sup_{|t| < \frac{1}{n}} \frac{\widetilde{d}^{f(g)}_{cc}\big(f(\Gamma_k(v; t)),\ \tilde{\delta}^{f(v)}_t \operatorname{ap} d_{sub}(f \circ \Gamma_k)(v)\big)}{|t|}, \quad k = 1, \dots, N.$$

We have $\varphi_n^k(g) \underset{ap}{\to} 0$ as $n \to \infty$ in E'. By Egorov's theorem there is $E'' \subset E'$ such that $\mathcal{H}^\nu(E' \setminus E'') < \varepsilon/2$ and $\varphi_n^k(g) \to 0$ as $n \to \infty$ uniformly on E''.

For every positive integer m and for all $x \in E$, $r > 0$, we define the set

$$T_k^m(x,r) = \Big\{ t \in (-r,r) : \widetilde{d}_{cc}^{f(x)}\big(f(\Gamma_k(x;t)), \tilde{\delta}_t^{f(x)} \operatorname{ap} d_{sub}(f \circ \Gamma_k)(x)\big) > \frac{|t|}{m} \Big\}.$$

For all positive integers p, we define the set

$$B_k^m(p) = A \cap \Big\{ x \in E : \mathcal{H}^{\deg X_k}[T_k^m(x,r)] \leq \frac{r^{\deg X_k}}{m} \text{ for all } r \in (0,p^{-1}) \Big\}.$$

In the case $k > 1$ we also define $Z_k^m(x,r;p)$, as the set of the points $z = (z_1,\dots,z_{k-1},0,\dots,0) \in \mathbb{R}^N$ such that $z \in B(0,r)$ and $\Phi_x(z) \notin B_k^m(p)$. Finally, for every positive integer q, we define the set

$$C_k^m(p,q) = B_k^m \cap \Big\{ x \in E : \mathcal{H}^{h_{k-1}}[Z_k^m(x,r;p)] \leq \frac{r^{h_{k-1}}}{m} \text{ for all } r \in (0,q^{-1}) \Big\}$$

where $h_k = \sum\limits_{i=1}^k \deg X_i$.

By Property 3.1, the sets $B_k^m(p)$, $C_k^m(p,q)$ are measurable and

$$A = \bigcup_{p=1}^\infty B_k^m(p),$$

$$\mathcal{H}^\nu\Big(B_k^m(p) \setminus \bigcup_{q=1}^\infty C_k^m(p,q)\Big) = 0 \qquad \text{for all } k = 1,\dots,N \text{ and } m \in \mathbb{N}.$$

Moreover, $B_k^m(p) \subset B_k^m(p+1)$, $C_k^m(p,q) \subset C_k^m(p,q+1)$. Hence, we can choose sequences of numbers $p_1, p_2, \dots$ and $q_1, q_2, \dots$ such that

$$\mathcal{H}^\nu(E'' \setminus B_k^m(p_m)) < \frac{\varepsilon}{2^m},$$
$$\mathcal{H}^\nu(E'' \cap B_k^m(p_m) \setminus C_k^m(p_m,q_m)) < \frac{\varepsilon}{2^m}$$

for all $k = 1,\dots,N$ and for every m. Then

$$\mathcal{H}^\nu(E'' \setminus F) < 2N\varepsilon \qquad \text{where } F = \bigcap_{k=1}^N \bigcap_{m=1}^\infty C_k^m(p_m,q_m).$$

Next we show that the limit (4.7) converges uniformly in F. Really, we have uniform estimates:

$$\widetilde{d}_{cc}^{f(g)}\big(f(\Gamma_k(v;t)),\, \tilde{\delta}_t^{f(v)} \operatorname{ap} d_{sub}(f \circ \Gamma_k)(v)\big) \leq \varphi_k(t),$$
$$\widetilde{d}_{cc}^{f(g)}\big(\tilde{\delta}_t^{f(u)} \operatorname{ap} d_{sub}(f \circ \Gamma_k)(u),\, \tilde{\delta}_t^{f(v)} \operatorname{ap} d_{sub}(f \circ \Gamma_k)(v)\big) \leq t\omega_k(d_{cc}(u,v))$$

for all $k = 1,\dots,N$, $u,v \in F$, where $\frac{\varphi(t)}{t} \to 0$ as $t \to 0$ uniformly for $v \in F$, $\omega_k(\cdot)$ is a modulus of the continuity of the mapping $\operatorname{ap} d_{sub}(f \circ \Gamma_k)$, $k = 1,\dots,N$.

Fix a density point $g \in F$, $m \in \mathbb{N}$ and $0 < r < \min\{p_m^{-1}, q_m^{-1}\}$. For every $k = 1, \dots, N$ define $S_k \subset \mathbb{R}^N$ as the set of the points $(t_1, \dots, t_N) \in B(0, r)$ such that

$$\begin{aligned}&\text{either } k > 1 \text{ and } (t_1, \dots, t_{k-1}) \in Z_k^m(g, r; p_m),\\ &\text{or } t_k \in T_k^m(\Phi_g(t_1, \dots, t_{k-1}, 0, \dots, 0), r).\end{aligned}$$

Since $\mathcal{H}^{h_{k-1}}[Z_k^m(g, r; p_m)] \le \frac{r^{h_{k-1}}}{m}$ and since $\mathcal{H}^{\deg X_k}[T_k^m(x, r)] \le \frac{r^{\deg X_k}}{m}$ if $x = \Phi_g(t_1, \dots, t_{k-1}, 0, \dots, 0) \in B_k^m(p_m)$, we have

$$\mathcal{H}^\nu(S_k) \le C_1 \frac{r^{h_{k-1}}}{m} r^{\nu - h_{k-1}} + C_2 \frac{r^{\deg X_k}}{m} r^{\nu - \deg X_k} \le C_3 \frac{r^\nu}{m}.$$

If we use the notation $W = \bigcup\limits_{k=1}^N S_k$ then $\mathcal{H}^\nu(W) \le C_4 \frac{r^\nu}{m}$. Denote

$$\begin{aligned}&u_1 = \Gamma_1(g; t_1),\\ &u_k = \Gamma_k(u_{k-1}; t_k) \quad \text{for all } k = 2, \dots, N.\end{aligned}$$

Now, if $v \in F \setminus W$ and $u_N(t) \in F \setminus W$, we have

$$\widetilde{d}_{cc}^{f(g)}\big(f(\Gamma_1(g; t_1)), \tilde{\delta}_{t_1}^{f(g)} \operatorname{ap} d_{sub}(f \circ \Gamma_1)(g)\big) \le \varphi_1(t_1) = C_1(|t_1|),$$

and then, by induction,

$$\begin{aligned}&\widetilde{d}_{cc}^{f(g)}\Big(f(u_k), \prod_{l=1}^k \tilde{\delta}_{t_l}^{f(g)} \operatorname{ap} d_{sub}(f \circ \Gamma_l)(g)\Big)\\ &\le \widetilde{d}_{cc}^{f(g)}\big(f(\Gamma_k(u_{k-1}; t_k)), \tilde{\delta}_{t_k}^{f(u_{k-1})} \operatorname{ap} d_{sub}(f \circ \Gamma_k)(u_{k-1})\big)\\ &\quad + \widetilde{d}_{cc}^{f(g)}\Big(\tilde{\delta}_{t_k}^{f(u_{k-1})} \operatorname{ap} d_{sub}(f \circ \Gamma_k)(u_{k-1}), \prod_{l=1}^k \tilde{\delta}_{t_l}^{f(g)} \operatorname{ap} d_{sub}(f \circ \Gamma_l)(g)\Big)\\ &\le \varphi_k(t_k) + |t_k| \omega_k(C_{k-1}(|t_1| + \dots + |t_{k-1}|)) = C_k(|t_1| + \dots + |t_k|),\end{aligned}$$

where $\max\{|t_1|, \dots, |t_k|\}^{-1} C_k(|t_1| + \dots + |t_k|) \to 0$ as $t \to 0$ uniformly for $g \in F$.

Denoting $\hat{v} = \widehat{\Phi}_g(t_1, \dots, t_N)$ we finally obtain

$$\operatorname{ap}\lim_{v \to g} \frac{\widetilde{d}_{cc}^{f(g)}(f(v), L_g(\hat{v}))}{d_{cc}^g(g, v)} = 0.$$

If $v = \Phi_g(t_1, \dots, t_N) \in F \cap \mathcal{G}^g$ then $d_{cc}^g(v, \hat{v}) = o(d_{cc}^g(g, v))$ as $v \to g$ by Theorem 2.3. Hence, using Property 4.4 of the mapping L_g we have

$$\widetilde{d}_{cc}^{f(g)}(f(v), L_g(v)) \le \widetilde{d}_{cc}^{f(g)}(f(v), L_g(\hat{v})) + \widetilde{d}_{cc}^{f(g)}(L_g(v), L_g(\hat{v})) = o(d_{cc}^g(g, v))$$

as $v \to g$. Since r, m, ε are arbitrary we have

$$\operatorname{ap}\lim_{v \to g} \frac{\widetilde{d}_{cc}^{f(g)}(f(v), L_g(v))}{d_{cc}^g(g, v)} = 0$$

for almost all $g \in A$. $\square$

4.3 Proof of theorem on approximate differentiability

Lemma 4.4. *Let $E \subset \mathcal{M}$ be a measurable set, $f : E \to \mathcal{M}$ be a measurable mapping, g be a density point of E and let*

$$\operatorname{ap}\lim_{v\to g} \frac{\widetilde{d}_{cc}^{f(g)}(f(v), L_g(v))}{d^g_{cc}(g,v)} = 0, \tag{4.8}$$

where $L_g : \mathcal{G}^g \cap \mathcal{M} \to \mathcal{G}^{f(g)}$ enjoys Properties 4.1 – 4.4. If there are $\eta > 0$, $0 < K < \infty$ such that

$$\widetilde{d}_{cc}(f(u), f(v)) < K d_{cc}(u,v)$$

for all $u, v \in B(g,\eta)$, then there exists the uniform limit

$$\lim_{v\to g} \frac{\widetilde{d}_{cc}^{f(g)}(f(v), L_g(v))}{d^g_{cc}(g,v)} = 0. \tag{4.9}$$

Proof. Let $\omega(t)$ be a modulus of continuity of $L_g : B(g,2) \cap \mathcal{G}^g \to \mathcal{G}^{f(g)}$. Then if $d^g_{cc}(u,v) < d^g_{cc}(g,v) < \eta$, by Property 4.2, we have

$$\begin{aligned}\widetilde{d}_{cc}^{f(g)}\big(L(u), L(v)\big) &= d^g_{cc}(g,v)\, \widetilde{d}_{cc}^{f(g)}\big(L(\delta^g_{d^g_{cc}(g,v)^{-1}}u), L(\delta^g_{d^g_{cc}(g,v)^{-1}}v)\big)\\ &\le d^g_{cc}(g,v)\,\omega\big(d^g_{cc}(\delta^g_{d^g_{cc}(g,v)^{-1}}u, \delta^g_{d^g_{cc}(g,v)^{-1}}v)\big) = d^g_{cc}(g,v)\,\omega\Big(\frac{d^g_{cc}(u,v)}{d^g_{cc}(g,v)}\Big).\end{aligned}$$

Suppose $0 < \varepsilon < 1$. Fulfillment of the condition (4.8) means there exists $\delta > 0$ such that, for the set

$$W = \{z \in E : \widetilde{d}_{cc}^{f(g)}(f(z), L_g(z)) < \varepsilon d^g_{cc}(g,z)\}$$

we have $\mathcal{H}^\nu(B(g,r)\setminus W) < r^\nu \varepsilon^\nu$ for any $0 < r < \delta$. If we take $x \in B(g,\delta(1-\varepsilon)) \cap E$ and $r = d^g_{cc}(g,x)/(1-\varepsilon)$ then $B(x, r\varepsilon) \subset B(g,r)$. It follows $B(x,r\varepsilon) \cap W \neq \emptyset$, hence, we can choose $z \in B(x,r\varepsilon) \cap E$. By Theorem 2.4 we have $d_{cc}(x,z) = o(d_{cc}(g,x)) = o(d^g_{cc}(g,x))$ and

$$\begin{aligned}\widetilde{d}_{cc}^{f(g)}(f(x), f(z)) = \widetilde{d}_{cc}(f(x), f(z)) + o\big(\widetilde{d}_{cc}(f(g), f(x))\big)&\\ = \widetilde{d}_{cc}(f(x), f(z)) + o(d^g_{cc}(g,x)),&\end{aligned}$$

where all $o(\cdot)$ are uniform. Thus, we infer

$$\begin{aligned}&\widetilde{d}_{cc}^{f(g)}(L_g(x), f(x))\\ &\quad\le \widetilde{d}_{cc}^{f(g)}(L_g(x), L_g(z)) + \widetilde{d}_{cc}^{f(g)}(L_g(z), f(z)) + \widetilde{d}_{cc}^{f(g)}(f(z), f(x))\\ &\quad\le d^g_{cc}(g,x)\,\omega\Big(\frac{d^g_{cc}(x,z)}{d^g_{cc}(g,x)}\Big) + \varepsilon\, d^g_{cc}(g,z) + \widetilde{d}_{cc}(f(x), f(z)) + o(d^g_{cc}(g,x))\\ &\quad\le d^g_{cc}(g,x)\,\omega(1) + \varepsilon\, d^g_{cc}(g,x) + \varepsilon\, d^g_{cc}(x,z) + K d_{cc}(x,z) + o(d^g_{cc}(g,x))\\ &\quad\le d^g_{cc}(g,x)\big(\omega(1) + \varepsilon + (\varepsilon + K + 1) o(1)\big)\end{aligned}$$

where all $o(\cdot)$ are uniform. □

Remark 4.3. If we prove that the mapping L_g is the approximate differential of f then from Lemma 4.4 it follows that the Lipschitz mapping is differentiable almost everywhere since the claims of Lemmas 4.1, 4.2, 4.3 and 4.4 hold almost everywhere in $\operatorname{dom} f$. This gives us an alternative proof of Theorem 3.4.

Now we have all necessary tools to complete the proof of Theorem 4.2.

Proof of Theorem 4.2. 1ST STEP. In the conditions of Theorem 4.2 the claims of Lemmas 4.1, 4.2 and 4.3 hold. In particular $A_j = \operatorname{dom} \operatorname{ap} X_j f$ is a measurable set, $j = 1, \dots, \dim H_1$, f is approximately differentiable along the curves $\Gamma_k(g; t)$ at $t = 0$, $k = 1, \dots, N$ almost everywhere in the set $A = \bigcap\limits_{i=1}^{\dim H_1} A_j$ and relations (4.2) and (4.7) hold.

If (4.7) holds at the point $g \in A$ then, in view of structure of L_g (4.6), estimate (4.3) implies

$$\begin{aligned}
&\operatorname{ap} \varlimsup_{v \to g} \frac{\widetilde{d}_{cc}(f(g), f(v))}{d_{cc}(g, v)} \\
&\le \operatorname{ap} \varlimsup_{v \to g} \frac{\widetilde{d}_{cc}^{f(g)}(f(g), L_g(v)) + \widetilde{d}_{cc}^{f(g)}(L_g(v), f(v))\big) + o\big(\widetilde{d}_{cc}^{f(g)}(L_g(v), f(v))\big)}{d_{cc}^{g}(g, v)} \\
&\le C \sup_{d^g_{cc}(g,v) \le 1} \Big(\prod_{k=1}^{N} \tilde{\delta}_{t_k}^{f(g)} \operatorname{ap} d_{sub}(f \circ \Gamma_k)(g) \Big) < \infty.
\end{aligned} \tag{4.10}$$

Hence, the left hand side of (4.10) is finite almost everywhere in A. Applying Theorem 3.6, we obtain a countable family of measurable sets covering A up to the set of measure 0 such that the restriction of f to each of them is a Lipschitz mapping.

Let E be one set of this countable family and let $L_g : \mathcal{G}^g \cap \mathcal{M} \to \mathcal{G}^{f(g)}$ be defined at almost all points of $E \subset A$. For proving the theorem it remains to verify that L_g is a homomorphism of the Lie groups. In particular, we need to prove that given two points $\hat{u}, \hat{v} \in \mathcal{G}^g$ we have

$$L_g(\hat{u} \cdot \hat{v}) = L_g(\hat{u}) \cdot L_g(\hat{v}). \tag{4.11}$$

2ND STEP. Let $g \in E$ be a density point where (4.7) holds and suppose $B_{cc}(g, r_g) \subset \mathcal{G}^g$. Then given $\hat{v} \in B_{cc}(g, r_g)$, $t \in [-r_g, r_g]$ there exists $v_t' = v_t'(g) \in E$, such that $d_{cc}^g(\delta_t^g \hat{v}, v_t') = o(t)$. By Lemma 4.4, we have

$$\operatorname{ap} \lim_{t \to 0} \frac{\widetilde{d}_{cc}^{f(g)}(f(v_t'), L_g(v_t'))}{t} = \lim_{t \to 0} \frac{\widetilde{d}_{cc}^{f(g)}(f(v_t'), L_g(v_t'))}{t} = 0.$$

Then, using Property 4.4 of the mapping L_g, we derive

$$\begin{aligned}
\widetilde{d}_{cc}^{f(g)}(f(v_t'), L_g(\delta_t^g \hat{v})) &\le \widetilde{d}_{cc}^{f(g)}(f(v_t'), L_g(v_t')) + \widetilde{d}_{cc}^{f(g)}(L_g(v_t'), L_g(\delta_t^g \hat{v})) \\
&= o(d_{cc}^g(g, v_t')) + o(d_{cc}^g(g, \delta_t^g \hat{v})) = o(t) \quad \text{as } t \to 0.
\end{aligned}$$

Next, consider two points $\hat{u}$, $\hat{v} \in B_{cc}(g, r_g/2)$ and their product $\hat{u} \cdot \hat{v}$. If $\hat{u} = \widehat{\Phi}_g(s_1, \dots, s_N)$ and $\hat{v} = \widehat{\Phi}_g(r_1, \dots, r_N)$ then define by induction

$$\begin{aligned}
&u_1(t)(\cdot) = \Phi_1(ts_1)(\cdot);\\
&u_k(t)(\cdot) = \Phi_k(ts_k) \circ u_{k-1}(t)(\cdot), \quad k = 2, \dots, N;\\
&v_1(t)(\cdot) = \Phi_1(tr_1)(\cdot);\\
&v_k(t)(\cdot) = \Phi_k(tr_k) \circ v_{k-1}(t)(\cdot), \quad k = 2, \dots, N.
\end{aligned}$$

From the structure of functions $\Phi_k(\cdot)$ and from Theorem 2.3 it follows

$$\begin{aligned}
&d^g_{cc}(u_N(t)(g), \delta^g_t \hat{u}) = o(t),\\
&d^g_{cc}(v_N(t)(g), \delta^g_t \hat{v}) = o(t),\\
&d^g_{cc}(v_N(t) \circ u_N(t)(g), \delta^g_t(\hat{u} \cdot \hat{v})) = o(t) \quad \text{as } t \to 0.
\end{aligned}$$

As long as g is a density point of E we can find $w'_k(t)$, $k = 1, \dots, 2N$, such that $d^g_{cc}\big(u_k(t)(g), w'_k(t)\big) = o(t)$ and $d^g_{cc}\big(v_k(t)(u_N(t)(g)), w'_{N+k}(t)\big) = o(t)$ as $t \to 0$, $k = 1, \dots, N$. By the same arguments as above we conclude that

$$\widetilde{d}^{f(g)}_{cc}(f(w'_{2N}(t)), L_g(\delta^g_t[\hat{u} \cdot \hat{v}])) = o(t) \quad \text{as } t \to 0.$$

All we need is to verify that

$$\widetilde{d}^{f(g)}_{cc}(f(w'_{2N}(t)), L_g(\delta^g_t \hat{u}) \cdot L_g(\delta^g_t \hat{v})) = o(t) \quad \text{as } t \to 0. \tag{4.12}$$

3RD STEP. For proving (4.12) we assume $\mathcal{H}^\nu(E) < \infty$ and restrict the set E applying Egorov's and Luzin's theorems.

Recall that the mapping $x \mapsto \operatorname{ap} d_{sub} f \circ \Gamma_k(x)$, defined in E, is measurable. By Lemma 4.4 we get

$$\lim_{t \to 0} \widetilde{d}^{f(x)}_{cc}(f \circ \Phi_k(t)(x), \delta^{f(x)}_t \operatorname{ap} d_{sub}(f \circ \Gamma_k)(x)) = 0 \tag{4.13}$$

for every density point $x \in E$ as $t \to 0$, $\Phi_k(t)(x) \in E$.

First, by Luzin's theorem there is a closed set $E_1 \subset E$ such that $\mathcal{H}^\nu(E \setminus E_1) < \varepsilon/3$ and

(a) all the mappings $x \mapsto \operatorname{ap} d_{sub} f \circ \Gamma_k(x)$ are uniformly continuous in E_1, $k = 1, \dots, N$.

Next, by Egorov's theorem there is a measurable set $E_2 \subset E_1$ such that $\mathcal{H}^\nu(E_1 \setminus E_2) < \varepsilon/3$ and

(b) the limit (4.13) converges uniformly on E_2, $k = 1, \dots, N$.

Now we consider a family of measurable functions

$$E_2 \ni x \to \psi_t(x) = \frac{\mathcal{H}^\nu(B_{cc}(x, t) \setminus E)}{\mathcal{H}^\nu(B_{cc}(x, t))}.$$

We have that $\lim\limits_{t \to 0} \psi_t(x) = 0$ at almost all points of $x \in E_2$. By Egorov's theorem there exists a measurable set $E_3 \subset E_2$ such that $\mathcal{H}^\nu(E_2 \setminus E_3) < \varepsilon/3$ and the limit

(c) $\lim\limits_{t\to 0}\psi_t(x)=0$ is uniform in E_3.

Property (c) allows us to repeat the arguments of the 2nd step with all $o(\cdot)$ uniform in E_3. Therefore, if $x\in E_3$ we have

$$\begin{aligned}
&\widetilde{d}^{f(x)}_{cc}\big(f(w'_1(t)(x)),\delta^{f(x)}_{t\sigma_1}\operatorname{ap}d_{sub}(f\circ\Gamma_1)(x)\big)=o(t),\\
&\widetilde{d}^{f(x)}_{cc}\big(f(w'_k(t)(w'_{k-1}(t))(x)),\delta^{f(w'_{k-1}(t))}_{t\sigma_k}\operatorname{ap}d_{sub}(f\circ\Gamma_k)(w'_{k-1}(t)))\big)=o(t),\\
&\widetilde{d}^{f(x)}_{cc}\big(f(w'_{N+1}(t)(w'_N(t))(x)),\delta^{f(w'_N(t))}_{t\tau_1}\operatorname{ap}d_{sub}(f\circ\Gamma_1)(w'_N(t)))\big)=o(t),\\
&\widetilde{d}^{f(x)}_{cc}\big(f(w'_{N+k}(t)(w'_{N+k-1}(t))(x)),\delta^{f(w'_{N+k-1}(t))}_{t\tau_k}\operatorname{ap}d_{sub}(f\circ\Gamma_k)(w'_{N+k-1}(t)))\big)=o(t)
\end{aligned}$$

as $t\to 0$, $k=2,\dots,N$, and all $o(\cdot)$ are uniform with respect to $x\in E_3$. Here the coefficients σ_k and τ_k are defined from (4.6) for the points $\hat{u}$ and $\hat{v}$ respectively. Then, by properties (a) and (b), the relation

$$\begin{aligned}
\widetilde{d}^{f(x)}_{cc}\Big(f(w'_{2N}(t)(x)),\prod_{k=1}^{N}\delta^{f(x)}_{t\sigma_k}\operatorname{ap}d_{sub}(f\circ\Gamma_k)(x)\cdot\prod_{k=1}^{N}\delta^{f(x)}_{t\tau_k}\operatorname{ap}d_{sub}(f\circ\Gamma_k)(x)\Big)&\\
=\widetilde{d}^{f(x)}_{cc}\big(f(w'_{2N}(t)(x)),\delta^{f(x)}_t L_x(\hat{u})\cdot\delta^{f(x)}_t L_x(\hat{v})\big)=o(t)&
\end{aligned}$$

is uniform with respect to $x\in E_3$. Finally,

$$\begin{aligned}
&t\,\widetilde{d}^{f(x)}_{cc}\big(L_x(\hat{u}\cdot\hat{v}),L_x(\hat{u})\cdot L_x(\hat{v})\big)\\
&\quad=\widetilde{d}^{f(x)}_{cc}\big(\delta^{f(x)}_t L_x(\hat{u}\cdot\hat{v}),\delta^{f(x)}_t L_x(\hat{u})\cdot\delta^{f(x)}_t L_x(\hat{v})\big)\\
&\quad\le\widetilde{d}^{f(x)}_{cc}\big(\delta^{f(x)}_t L_x(\hat{u}\cdot\hat{v}),f(w'_{2N}(t)(x))\big)\\
&\qquad+\widetilde{d}^{f(x)}_{cc}\big(f(w'_{2N}(t)(x)),\delta^{f(x)}_t L_x(\hat{u})\cdot\delta^{f(x)}_t L_x(\hat{v})\big)=o(t)
\end{aligned}$$

and (4.10) is proved for $x\in E_3$. Since ε is an arbitrary positive number, the Theorem is proved. $\square$

5 Application: an area formula

Suppose that $x=\exp\Big(\sum\limits_{i=1}^{N}x_iX_i\Big)(g)$. Define a quantity

$$d_\rho(g,x)=\max\Big\{\Big(\sum_{j=1}^{\dim H_1}|x_j|^2\Big)^{\frac{1}{2}},\Big(\sum_{j=\dim H_1+1}^{\dim H_2}|x_j|^2\Big)^{\frac{1}{4}},\dots,\Big(\sum_{j=\dim H_{M-1}+1}^{N}|x_j|^2\Big)^{\frac{1}{2M}}\Big\}. \tag{5.1}$$

It is easy to see that d_ρ is locally equivalent to d_∞. Since we have already proved that d_∞ and d_{cc} are locally equivalent, the following statement also holds.

Proposition 5.1. *Let $g \in \mathcal{M}$. There is a compact neighborhood $U(g) \subset \mathcal{M}$ such that*

$$c_1 d_{cc}(u,v) \le d_\rho(u,v) \le c_2 d_{cc}(u,v)$$

for all u, v in $U(g)$, where constants $0 < c_1 \le c_2 < \infty$ independent of $u, v \in U(g)$.

Corollary 5.1. *Quantity d_ρ is a quasimetric.*

Denote an open ball in the quasimetric d_ρ of radius r with center in x as $\operatorname{Box}_\rho(x,r)$. Define the (spherical) Hausdorff measure of a set E with respect to metric d_ρ as

$$\mathcal{H}^k_\rho(E) = \lim_{\varepsilon\to 0+} \inf\Big\{\sum_i r_i^k : E \subset \bigcup_i \operatorname{Box}_\rho(x_i, r_i), r_i < \varepsilon\Big\}.$$

For Lipschitz mappings of Carnot–Carathéodory mappings the following area formula holds.

Theorem 5.1 ([11]). *Suppose $E \subset \mathcal{M}$ is a measurable set, and the mapping $\varphi : E \to \widetilde{\mathcal{M}}$ is Lipschitz with respect to sub-Riemannian quasimetrics d_ρ and $\widetilde{d}_\rho$. Then the area formula*

$$\int\limits_E f(x)\mathcal{J}^{SR}(\varphi,x)\,d\mathcal{H}^\nu_\rho(x) = \int\limits_{\varphi(E)} \sum_{x\colon x\in\varphi^{-1}(y)} f(x)\,d\mathcal{H}^\nu_\rho(y) \tag{5.2}$$

holds, where $f : F \to \mathbb{M}$ (here $\mathbb{M}$ is an arbitrary Banach space) is such that function $f(x)\mathcal{J}^{SR}(\varphi,x)$ is integrable, and

$$\mathcal{J}^{SR}(\varphi,x) = \sqrt{\det(D\varphi(x)^* D\varphi(x))} \tag{5.3}$$

is the sub-Riemannian Jacobian of φ at x.

As an immediate corollary of 5.1 and 4.2 we obtain the following result.

Theorem 5.2. *Suppose $E \subset \mathcal{M}$ is a measurable set, and the mapping $\varphi : E \to \widetilde{\mathcal{M}}$ is approximately differentiable almost everywhere. Then the area formula*

$$\int\limits_E f(x)\operatorname{ap}\mathcal{J}^{SR}(\varphi,x)\,d\mathcal{H}^\nu_\rho(x) = \int\limits_{\widetilde{\mathcal{M}}} \sum_{x\colon x\in\varphi^{-1}(y)\setminus\Sigma} f(x)\,d\mathcal{H}^\nu_\rho(y)$$

holds, where $f : F \to \mathbb{M}$ (here $\mathbb{M}$ is an arbitrary Banach space) is such that function $f(x)\operatorname{ap}\mathcal{J}^{SR}(\varphi,x)$ is integrable, $\mathcal{H}^\nu_\rho(\Sigma) = 0$ and

$$\operatorname{ap}\mathcal{J}^{SR}(\varphi,x) = \sqrt{\det(\operatorname{ap}D\varphi(x)^* \operatorname{ap}D\varphi(x))}$$

is the approximate sub-Riemannian Jacobian of φ at x.

Proof. By Theorem 4.2, there is a sequence of disjoint sets Σ, $E_1, E_2, \dots$ such that $E = \Sigma \cup \bigcup\limits_{i=1}^{\infty} E_i$, $\mathcal{H}^{\nu}_{\rho}(\Sigma) = 0$ and every restriction $\varphi|_{E_i}$ is a Lipschitz mapping. Then, by Theorem 5.1, we have

$$\int\limits_{E} f(x)\,\mathrm{ap}\,\mathcal{J}^{SR}(\varphi, x) d\mathcal{H}^{\nu}_{\rho}(x) = \sum_{i=1}^{\infty} \int\limits_{E_i} f(x)\mathcal{J}^{SR}(\varphi, x) d\mathcal{H}^{\nu}_{\rho}(x)$$

$$= \sum_{i=1}^{\infty} \int\limits_{\widetilde{\mathcal{M}}} \sum_{x:\, x\in\varphi^{-1}(y)\cap E_i} f(x) d\mathcal{H}^{\nu}_{\rho}(y) = \int\limits_{\widetilde{\mathcal{M}}} \sum_{x:\, x\in\varphi^{-1}(y)\setminus\Sigma} f(x) d\mathcal{H}^{\nu}_{\rho}(y).$$

$\square$

Remark 5.1. Since Ball-Box theorem holds, the Hausdorff measure, constructed by the metric d_{cc}, is absolutely continuous with respect to the Hausdorff measure, constructed by the distance d_ρ, and vice versa. Therefore we have

$$d\mathcal{H}^{\nu}_{\rho}(x) = \mathcal{D}_{\rho,cc}(x)\, d\mathcal{H}^{\nu}_{cc}(x), \quad x \in \mathcal{M},$$

where $0 < \alpha \leq \mathcal{D}_{\rho,cc}(x) \leq \beta < \infty$ is measurable function. So, the area formula of Theorem 5.2 can be written also for the Hausdorff measure, constructed by the distance d_ρ.

Acknowledgments

The research was partially supported by the Russian Foundation for Basic Research (grants 10–01–00662-a and 11–01–00819-a), the State Maintenance Program for Young Russian Scientists and the Leading Scientific Schools of the Russian Federation (grant NSh-921.2012.1), and the Federal Target Grant "Scientific and educational personnel of innovation Russia" for 2009-2013 (Agreements no. 8206 and 8212).

References

[1] V.I. Arnol'd, *Ordinary Differential Equations*, Springer-Verlag Berlin Heidelberg, 1992.

[2] A. Bonfiglioli, E. Lanconelli, F. Uguzonni, *Stratified Lie Groups and Potential Theory for their Sub-Laplasians*, Springer-Verlag Berlin Heidelberg, 2007.

[3] W.L. Chow, *Über systeme von linearen partiallen differentialgleichungen erster ordnung*, Math. Ann. 117 (1939), 98–105.

[4] H. Federer, *Geometric measure theory*, series Die Grundlehren der mathematischen Wissenschaften, Band 153, New York: Springer-Verlag New York Inc. (1969).

[5] G. B. Folland, E. M. Stein, *Hardy spaces on homogeneous groups*, Princeton Univ. Press, Princeton, 1982. (Math notes; 28)

[6] A.V. Greshnov, *A proof of Gromov Theorem on homogeneous nilpotent approximation for C^1-smooth vector fields*, Mathematicheskie Trudy, 15 (2012), no. 2, 72–88.

[7] M. Gromov, *Carnot–Carathéodory spaces seen from within*, Sub-Riemannian Geometry, Prog. Math., vol. 144, Birkhäuser, Basel, 1996, 72–323.

[8] L.Hörmander, *Hypoelliptic second order differential equations*, Acta Math. 119 (1967), 147–171.

[9] M. Karmanova, *A new approach to investigation of Carnot–Carathéodory geometry*, Geom. Func. Anal. 21, no. 6 (2011), 1358–1374.

[10] M. Karmanova, *Convergence of scaled vector fields and local approximation theorem on Carnot–Carathéodory spaces and applications*, Dokl. AN 440 (2011), no. 6, 736–742.

[11] M. Karmanova, *The Area Formula for Lipschitz Mappings of Carnot–Carathéodory Spaces*, Izvestiya: Mathematics (submitted), `arXiv:1110.5483v1 [math.MG]`

[12] M. Karmanova, S. Vodopyanov, *Geometry of Carnot–Carathéodory spaces, differentiability and coarea formula*, Analysis and Mathematical Physics. Birkhäuser 2009, 284–387.

[13] M. Karmanova, S. Vodopyanov, *On Local Approximation Theorem on Equiregular Carnot–Carathéodory spaces*, Proceedings of "INDAM Meeting on Geometric Control and Sub-Riemannian Geometry", Cortona, May 2012. Springer INdAM Series 2013 (to appear).

[14] A. Koranyi, H.M. Reimann, *Foundations for the theory of quasiconformal mappings on the Heisenberg group*, Adv. Math., 111 (1995), 1–87.

[15] S. Lie, F. Engel, *Theorie der Transformationsgruppen*, Bd 1-3, Lpz., 1888–93.

[16] V. Magnani, *Differentiability and area formula on stratified Lie groups*, Houston J. Math. 27 (2001), no. 2, 297–323.

[17] G. Metivier, *Fonction spectrale et valeurs propres d'une classe d'opérateurs non elliptiques*, Commun. Partial Differential Equations 1 (1976), 467–519.

[18] J. Mitchell, *On Carnot–Carathéodory metrics*, J. Differential Geometry 21 (1985), 35–45.

[19] A. Nagel, E. M. Stein, S. Waigner, *Balls and metrics defined by vector fields. I: Basic properties*, Acta Math. (2) 155 (1985), 103–147.

[20] P. Pansu, *Metriqués de Carnot–Carathéodory et quasiisométries des espaces symétriques de rang un*, Ann. of Math. 119 (1989), 1–60.

[21] S. Pauls, *A notion of rectifiability modeled on Carnot groups*, Indiana Univ. Math. J. 53 (2004), 49–82.

[22] M.M. Postnikov, *Lectures in Geometry. Semester V: Lie Groups and Lie Algebras,* Moscow, "Nauka", 1982.

[23] H. Rademacher, *Über partielle und totalle Differenzierbarkeit von Funktionen mehrerer Variablen und über die Transformation der Doppelintegrale*, I, Math. Ann. 79 (1919), 340–359.

[24] P.K. Rashevsky, *Any two points of a totally nonholonomic space may be connected by an admissable line*, Uch. Zap. Ped. Inst. im. Liebknechta, Ser. Phys. Math. 2 (1938), 83–94.

[25] Yu.G. Reshetnyak, *Generalized derivatives and differentiability almost everywhere*, Math. USSR-Sb. 4 (1968), 293–302.

[26] L.P. Rotchild, E. S. Stein, *Hypoelliptic differential operators and nilpotent groups*, Acta Math. 137 (1976), 247–320.

[27] W. Stepanoff, *Über totale Differenzierbarkeit*, Math. Ann. 90 (1923), 318–320.

[28] W. Stepanoff, *Sur les conditions de l'existence de la différentielle totale*, Matem. Sborn. Moscow 32 (1925), 511–527.

[29] S.K. Vodop'yanov, *Monotone functions and quasiconformal mappings on Carnot groups*, Siberian Math. J. 37 (1996), no. 6, 1113–1136.

[30] S.K. Vodop'yanov, *Mappings with bounded distortion and with finite distortion on Carnot groups,* Siberian Math. J. 40 (1999), no. 4, 644–677.

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

[32] S.K. Vodop'yanov, *Differentiability of mappings in the geometry of Carnot manifolds*, Siberian Math. J. 48 (2007), no. 2, 197–213.

[33] S.K. Vodopyanov, *Geometry of Carnot–Carathéodory Spaces and Differentiability of Mappings*, Contemporary Mathematics 424 (2007), 247–302.

[34] S.K. Vodopyanov, *Foundations of the theory of mappings with bounded distortion on Carnot groups*, The interaction of analysis and geometry, 303–344, Contemp. Math., 424, Amer. Math. Soc., Providence, RI, 2007.

[35] S.K. Vodop'yanov, A. V. Greshnov, *On the differentiability of mappings of Carnot–Carathéodory spaces*, Doklady Math. 67 (2003), no. 2, 246–251.

[36] S.K. Vodopyanov, M. B. Karmanova, *Local Geometry of Carnot Manifolds Under Minimal Smoothness*, Doklady Math. 75 (2) (2007), 240–246.

[37] S.K. Vodop'yanov, M. B. Karmanova, *Local approximation theorem on Carnot manifolds under minimal smoothness*, Dokl. AN 427 (2009), no. 6, 731–736.

[38] S.K. Vodop'yanov, A. D. Ukhlov, *Approximately differentiable tranformations and the change of variables on nilpotent groups*, Siberian Math. J. 37 (1996), no. 1, 62–78.

[39] S.K. Vodop'yanov, A. D. Ukhlov, *Sobolev spaces and (P, Q)-quasiconformal mappings of Carnot groups*, Siberian Math. J. 39 (1998), no. 4, 665–682.

[40] H. Whitney, *On totally differentiable and smooth functions*, Pacific J. Math. 1 (1951), 143–159.

Basalaev Sergei Genad'evich
Novosibirsk State University
E-mail: sbasalaev@gmail.com

Vodopyanov Sergei Konstantinovich
Sobolev Institute of Mathematics of the Siberian Branch of the Russian Academy of Sciences
E-mail: vodopis@math.nsc.ru

Received: 27.09.2010