Sub-Riemannian geometry and Lie groups
Part I
Seminar Notes, DMA-EPFL, 2001

M. Buliga

Institut Bernoulli
Bâtiment MA
École Polytechnique Fédérale de Lausanne
CH 1015 Lausanne, Switzerland
Marius.Buliga@epfl.ch

and

Institute of Mathematics, Romanian Academy
P.O. BOX 1-764, RO 70700
Bucureşti, Romania
Marius.Buliga@imar.ro

This version: October 31, 2002

Introduction

M. Gromov [13], pages 85–86:

"**3.15. Proposition:** Let (V, g) be a Riemannian manifold with g continuous. For each $v \in V$ the spaces $\lambda(V, v)$ Lipschitz converge as $\lambda \to \infty$ to the tangent space $(T_v V, 0)$ with its Euclidean metric g_v.

Proof $_+$**:** Start with a C^1 map $(R^n, 0) \to (V, v)$ whose differential is isometric at 0. The λ-scalings of this provide almost isometries between large balls in R^n and those in λV for $\lambda \to \infty$.

Remark: In fact we can *define* Riemannian manifolds as locally compact path metric spaces that satisfy the conclusion of Proposition 3.15. "

If so, Gromov's remark should apply to any sub-Riemannian manifold. Why then is the sub-Riemannian case so different from the Riemannian one? Here is a list of legitimate questions:

How can one define the manifold structure? Who are the tangent and cotangent bundles? What is the intrinsic differential calculus? Why are there abnormal geodesics if the Hamiltonian formalism on the cotangent bundle were complete? If the manifold is a compact Lie group does the tangent bundle carry a natural group structure? What are differential forms, de Rham cochain, and the variational complex? Consider the group of smooth volume preserving transformations. Why does this group have more invariants than the volume and what is the interpretation of these invariants?

The purpose of this working seminar was to explore as many as possible open questions from the list above. Special attention has been payed to the case of a Lie group with a left invariant distribution.

The seminar, organised by the author and Tudor Raţiu at the Mathematics Department, EPFL, started in November 2001.

The paper is by no means self contained. For any unproved result it is indicated the place where a complete proof can be found. The choice of the proofs is rather psychological: some of them have a geometrical or mixed geometric-analytical meaning (like the proof of Hopf-Rinow theorem), others help to understand that familiar reasoning applies in apparently unfamiliar situations (like in the preparations for the Pansu-Rademacher theorem). Important results (Ball-Box theorem for example) have elementary proof in particular situations; this might help to better understand their meaning and also their strength.

These pages covers the expository talks given by the author during this seminar. However, this is the first part of three, in preparation, dedicated to this subject. It covers, with mild modifications, an elementary introduction to the field.

The second part will deal with the applications of the Local-to-Global principle for moment maps arising in connection with Carnot groups. We prove, for example, that Kostant nonlinear convexity theorem 4.1 [18] can be seen as the sub-Riemannian version of the linear convexity theorem 8.2. *op. cit.* The key point is the redefinition of the tangent bundle of a Lie group endowed with a left invariant nonintegrable distribution, using a natural construction based on noncommutative derivative.

The third part is devoted to the study of some representations of groups of bi-

Lipschitz maps on Carnot groups. The main idea is that one can formulate an interesting rectifiability theory as a theory of (irreducible) representations of such groups.

1   From metric spaces to Carnot groups

1.1   Metric spaces

Definition 1.1 *(distance) A function $d : X \times X \to [0, +\infty)$ is a distance on X if:*

  (a) $d(x, y) = 0$ *if and only if $x = y$.*

  (b) *for any x, y $d(x, y) = d(y, x)$.*

  (c) *for any x, y, z $d(x, z) \leq d(x, y) + d(y, z)$.*

(X, d) *is called a metric space. The open ball with centre $x \in X$ and radius $r > 0$ is denoted by $B(x, r)$.*

If d ranges in $[0, +\infty]$ then it is called a pseudo-distance. The property of two points of being at finite distance is an equivalence relation. d is then a distance on each equivalence class (leaf).

Definition 1.2 *A map between metric spaces $f : X \to Y$ is Lipschitz if there is a positive constant C such that for any $x, y \in X$ we have*

$$d_Y(f(x), f(y)) \ \leq \ C \ d_X(x, y)$$

The least such constant is denoted by $Lip(f)$.

The dilatation, or metric derivative, of a map $f : X \to Y$ between metric spaces, in a point $u \in Y$ is

$$dil(f)(u) = \limsup_{\varepsilon \to 0} \ \sup \left\{ \frac{d_Y(f(v), f(w))}{d_X(v, w)} \ : \ v \neq w \ , \ v, w \in B(u, \varepsilon) \right\}$$

The distorsion of a map $f : X \to Y$ is

$$dis \ f \ = \ \sup \left\{ |\ d_Y(f(y), f(y')) - d_X(y, y')\ | \ : \ y, y' \in X \right\}$$

The name "metric derivative" is motivated by the fact that for any derivable function $f : R \to R^n$ the dilatation is $dil(f)(t) \ = \ |\ \dot{f}(t)\ |$.

A curve is a function $f : [a, b] \to X$. The image of a curve is called path. Length measures paths. Therefore length does not depends on the reparametrisation of the path and it is additive with respect to concatenation of paths.

In a metric space (X, d) one can measure the length of curves in several ways.

The length of a curve with L^1 dilatation $f : [a, b] \to X$ is

$$L(f) = \int_a^b dil(f)(t) \ dt$$

A different way to define a length of a curve is to consider its variation.

The curve f has bounded variation if the quantity

$$Var(f) \ = \ \sup \left\{ \sum_{i=1}^{n-1} d(f(t_i), f(t_{i+1})) \ : \ a \leq t_1 < ... < t_n \leq b \right\}$$

(called variation of f) is finite.

The variation of a curve and the length of a path , as defined previously, do not agree in general. To see this consider the following easy example: $f : [-1, 1] \rightarrow R^2$, $f(t) = (t, \ sign(t))$. We have $Var(f) = 4$ and $L(f([-1, 1]) = 2$. Another example: the Cantor staircase function is continuous, but not Lipschitz. It has variation equal to 1 and length of the graph equal to 2.

Nevertheless, for Lipschitz functions, the two definitions agree.

Theorem 1.3 *For each Lipschitz curve $f : [a, b] \rightarrow X$, we have $L(f) = Var(f)$.*

For the statement and the proof see the second part of Theorem 4.1.1., Ambrosio [2].

Proof. We prove the thesis by double inequality. f is continuous therefore $f([a, b])$ is a compact metric space. Let $\{x_n : n \in N\}$ be a dense sequence in $f([a, b])$. All the functions

$$t \mapsto \phi_n(t) = d(f(t), x_n)$$

are Lipschitz and $Lip(\phi_n) \leq Lip(f)$, because of the general property:

$$Lip(f \circ g) \leq Lip(f) \ Lip(g)$$

if f, g are Lipschitz. In the same way we see that :

$$dil(f)(t) = \sup \{dil(\phi_n)(t) : n \in N\}$$

We have then, for $t < s$ in $[a, b]$:

$$d(f(t), f(s)) = \sup \{| \ d(f(t), x_n) - d(f(s), x_n) \ | \ : n \in N\} \leq$$

$$\leq \int_s^t dil(f)(\tau) \ \mathrm{d}\tau$$

From the definition of the variation we get

$$Var(f) \leq L(f)$$

For the converse inequality let $\varepsilon > 0$ and $n \geq 2$ natural number such that $h = (b - a)/n < \varepsilon$. Set $t_i = a + ih$. Then

$$\frac{1}{h} \int_a^{b-\varepsilon} d(f(t+h), f(t)) \ \mathrm{d}t \leq \frac{1}{h} \sum_{i=0}^{n-2} d(f(\tau + t_{i+1}), f(\tau + t_i)) \ \mathrm{d}\tau \leq$$

$$\leq \frac{1}{h} \int_0^h Var(f) \ \mathrm{d}\tau = Var(f)$$

Fatou lemma and definition of dilatation lead us to the inequality

$$\int_a^{b-\varepsilon} dil(f)(t) \ \mathrm{d}t \leq Var(f)$$

This finishes the proof because ε is arbitrary. ∎

There is a third, more basic way to introduce the length of a curve in a metric space.

The length of the path $A = f([a, b])$ is by definition the one-dimensional Hausdorff measure of the path. The definition is the following:

$$l(A) \;=\; \lim_{\delta \to 0} \inf \left\{ \sum_{i \in I} diam \; E_i \; : \; diam \; E_i < \delta \;, \;\; A \subset \bigcup_{i \in I} E_i \right\}$$

The general Hausdorff measure is defined further:

Definition 1.4 *Let $k > 0$ and (X, d) be a metric space. The k-Hausdorff measure is defined by:*

$$\mathcal{H}^k(A) \;=\; \lim_{\delta \to 0} \inf \left\{ \sum_{i \in I} (diam \; E_i)^k \; : \; diam \; E_i < \delta \;, \;\; A \subset \bigcup_{i \in I} E_i \right\}$$

Any k-Hausdorff measure is an outer measure. The Hausdorff dimension of a set A is defined by:

$$\mathcal{H} - dim \; A \;=\; \inf \left\{ k > 0 \; : \; \mathcal{H}^k(A) \;=\; 0 \right\}$$

Variation of a curve and the Hausdorff measure of the associated path don't coincide. For example take a circle S^1 in R^n parametrised such that any point is covered two times. Then the variation is two times the length.

Theorem 1.5 *Suppose that $f : [a, b] \to X$ is an injective Lipschitz function and $A = f([a, b])$. Then $l(A) = Var(f)$.*

For the statement and the proof see Theorem 4.4.1., Ambrosio [2].

In order to give the proof of the theorem we need two things. The first is the geometrically obvious, but not straightforward to prove in this generality, Reparametrisation Theorem (theorem 4.2.1 Ambrosio [2]).

Theorem 1.6 *Any path $A \subset X$ with a Lipschitz parametrisation admits a reparametrisation $f : [a, b] \to A$ such that $dil(f)(t) = 1$ for almost any $t \in [a, b]$.*

The second result that we shall need is Lemma 4.4.1 Ambrosio [2].

Lemma 1.7 *If $f : [a, b] \to X$ is continuous then*

$$\mathcal{H}^1(f([a, b]) \;\leq\; d(f(a), f(b))$$

Proof.   Let us consider the Lipschitz function

$$\phi : X \to R , \quad \phi(x) = d(x, f(a))$$

It has the property $Lip(\phi) \leq 1$ therefore by the definition of Hausdorff measure we have

$$\mathcal{H}^1(\phi \circ f([a, b])) \leq \mathcal{H}^1(f([a, b])$$

On the left hand side of the inequality we have the Hausdorff measure on R, which coincides with the usual (outer) Lebesgue measure. Moreover $\phi \circ f([a, b]) = [0, \alpha]$, therefore we obtain

$$\mathcal{H}^1(\phi \circ f([a, b])) = \sup \{ d(f(t), f(a)) : t \in [a, b] \} \geq d(f(a), f(b))$$

■

The proof of theorem 1.6 follows.

Proof.   It is not restrictive to suppose that $A = f([a, b])$ can be parametrised by $f : [a, b] \to A$ such that $dil(f)(t) = 1$ for all $t \in [a, b]$. Due to theorem 1.3 and again the reparametrisation theorem, we can choose $[a, b] = [0, Var(f)]$.

For an arbitrary $\delta > 0$ we choose $n \in N$ such that $h = Var(f)/n < \delta$ and we divide the interval $[0, Var(f)]$ in intervals $J_i = [ih, (i + 1)h]$. The function f has Lipschitz constant equal to 1 therefore (see definition of Hausdorff measure and notations therein)

$$\mathcal{H}^1_\delta(A) \leq \sum_{i=0}^{n} diam \ (J_i) = Var(f)$$

δ is arbitrary therefore $\mathcal{H}^1(A) \leq Var(f)$. This is a general inequality which does not use the injectivity hypothesis.

We prove the converse inequality from injectivity hypothesis. Let us divide the interval $[a, b]$ by $a \leq t_0 < ... < t_n \leq b$. From lemma 1.7 and sub-additivity of Hausdorff measure we have:

$$\sum_{i=0}^{n-1} d(f(t_i), f(t_{i+1})) \leq \sum_{i=0}^{n-1} \mathcal{H}^1(f([t_i), t_{i+1}])) \leq H^1(A)$$

The partition of the interval was arbitrary, therefore $Var(f) \leq \mathcal{H}^1(A)$.   ■

In conclusion, in any metric space we can measure length of Lipschitz curves following one of the recipes from above, or we can measure length of paths with the one-dimensional Hausdorff measure. The lengths agree if the path admits a Lipschitz parametrisation.

We shall denote by l_d the length functional, defined only on Lipschitz curves, induced by the distance d.

The length induces a new distance d_l, say on any Lipschitz connected component of the space (X, d). The distance d_l is given by:

$$d_l(x, y) = \inf \ \{ l_d(f([a, b])) : f : [a, b] \to X \ \text{Lipschitz} ,$$

$$f(a) = x \ , \ f(b) = y\}$$

We have therefore two operators $d \mapsto l_d$ and $l \mapsto d_l$. Is one the inverse of another? The answer is no. This leads to the introduction of path metric spaces.

Definition 1.8 *A path metric space is a metric space (X, d) such that $d = d_l$.*

In terms of distances there is an easy criterion to decide if a metric space is path metric (Theorem 1.8., page 6-7, Gromov [13]).

Theorem 1.9 *A complete metric space is path metric if and only if (a) or (b) from above is true:*

(a) for any $x, y \in X$ and for any $\varepsilon > 0$ there is $z \in X$ such that

$$\max \left\{ d(x, z), d(z, y) \right\} \ \leq \ \frac{1}{2} d(x, y) + \varepsilon$$

(b) for any $x, y \in X$ and for any $r_1, r_2 > 0$, if $r_1 + r_2 \leq d(x, y)$ then

$$d(B(x, r_1), B(y, r_2)) \ \leq \ d(x, y) - r_1 - r_2$$

Proof.   We shall prove only that (a) implies that X is path metric. (b) implies (a) is straightforward, path metric space implies (b) likewise.

Set $\delta = d(x, y)$ and take a sequence $\varepsilon_k > 0$, with finite sum. We shall recursively define a function $z = z_\varepsilon$ from the dyadic numbers in $[0, 1]$ to X. $z_\varepsilon(1/2) \ = \ z_{1/2}$ is a point such that

$$\max \left\{ d(x, z_{1/2}), d(z_{1/2}, y) \right\} \ \leq \ \frac{1}{2} \delta(1 + \varepsilon_1)$$

Suppose now that all points $z_\varepsilon(p/2^n) \ = \ z_{p/2^n}$ were defined, for $p = 1, ..., 2^n - 1$. Then $z_{2p+1/2^{n+1}}$ is a point such that

$$\max \left\{ d(z_{p/2^n}, z_{2p+1/2^{n+1}}), d(z_{2p+1/2^{n+1}}, z_{p+1/2^n}) \right\} \ \leq \ \frac{\delta}{2^n} \prod_{k=1}^{n+1} (1 + \varepsilon_k)$$

Because (X, d) is a complete metric space it follows that z_ε can be prolonged to a Lipschitz curve c_ε, defined on the whole interval $[0, 1]$, such that $c(0) = x$, $c(1) = y$ and

$$d(x, y) \ \leq \ l(c_\varepsilon \ \leq \ d(x, y) \prod_{k \geq 1} (1 + \varepsilon_k)$$

But the product $\prod_{k \geq 1}(1 + \varepsilon_k)$ can be made arbitrarily close to 1, which proves the thesis.   ■

Path metric spaces are geometrically more interesting, because one can define geodesics. Let c be any curve; denote the length of the restriction of c to the interval t, t' by $l_c(t, t')$.

Definition 1.10 *A (local) geodesic is a curve $c : [a, b] \to X$ with the property that for any $t \in (a, b)$ there is a small $\varepsilon > 0$ such that $c : [t - \varepsilon, t + \varepsilon] \to X$ is length minimising. A global geodesic is a length minimising curve.*

Therefore in a path metric space a local geodesic has the property that in the neighbourhood of any of it's points the relation

$$d(c(t), c(t')) = l_c(t, t')$$

holds. Any global geodesic is also local geodesic.

Can one join any two points with a geodesic?

The abstract Hopf-Rinow theorem (Gromov [13], page 9) states that:

Theorem 1.11 *If (X, d) is a connected locally compact path metric space then each pair of points can be joined by a global geodesic.*

Proof.   It is sufficient to give the proof for compact path metric spaces. Given the points x, y, there is a sequence of curves f_h joining those points such that $l(f_h) \le d(x, y) + 1/h$. The sequence, if parametrised by arclength, is equicontinuous; by Arzela-Ascoli theorem one can extract a subsequence (denoted also f_h) which converges uniformly to f. By construction the length function is lower semicontinuous hence:

$$l(f) \le \liminf_{h \to \infty} l(f_h) \le d(x, y)$$

Therefore f is a length minimising curve joining x and y.   ■

The following remark of Gromov 1.13(b) [13] easily comes from the examination of the previous proof.

Proposition 1.12 *In a compact path metric space every free homotopy class is represented by a length minimising curve.*

1.2   Local to Global Principle

Let (X, d) be a connected locally compact and complete path metric space.

The natural notion of subspace in this category of metric spaces is the following one.

Definition 1.13 *A set $C \subset X$ is a closed convex set if the inclusion $(C, d_C) \subset (X, d)$ is an isometry, where d_C is the inner path metric distance:*

$$d_C(x, y) \;=\; \inf\; \{ l_d(f([a, b])) \; : \; f : [a, b] \to X \quad Lipschitz \; ,$$

$$f(a) = x \; , \;\; f(b) = y \; , \;\; f([a, b]) \subset C \; \}$$

We shall call a closed set weakly convex if any two points of the set can be joined by a local geodesic lying in the set.

A closed set it is therefore convex if any two points $x, y \in C$ can be joined by a global geodesic lying in C. Note that there might be other global geodesics joining two points in C which are not entirely in C.

The purpose of the remainder of this section is to state and prove the Local-to-Global Principle. The principle has been formulated for the first time by Hilgert, Neeb, Plank [15]. It has been used to prove convexity theorems; it's most widely known application concerns an alternative proof of the Atiyah-Guillemin-Sternberg theorem about convexity of the image of the moment map. We shall prove in this section a general version of the principle, true in path metric spaces. Applications of this principle will be given further, where we will meet also the moment map.

Definition 1.14 *A pointed closed convex set is a pair (C, x) with $x \in C \in X$ and C closed convex set with the property that there is a constant $K \geq 1$ such that for any $r > 0$ and any two points $y, y' \in C \cap B(x, r)$ can be joined by a global geodesic in $C \cap B(x, Kr)$.*

Let X be a connected, compact topological Hausdorff space, Y a connected locally compact path metric space and $f : X \to Y$ with the following properties:

LC1. f is proper and for any $x \in X$ there is a neighbourhood V_x of x such that $f^{-1}(f(y)) \cap V_x$ is connected for all $y \in V_x$.

LC2. for any $x \in X$ there is a pointed closed convex $(C_x, f(x))$ and a neighbourhood V_x of x such that $f : V_x \to C_x$ is open.

Theorem 1.15 *(Local-to-Global Principle). Under the hypotheses LC1, LC2, the image $f(X)$ is closed weakly convex in Y.*

The proof is inspired by the one from Sleewaegen [30], section 1.8, Theorem 1.8.1.

Proof.   The proof splits in two steps, one of topological nature and the other of metric character.

Step 1.   Let us consider on X the equivalence relation $x \approx y$ if $f(x) = f(y)$ and x, y are in the same connected component of $f^{-1}(f(x))$. Denote by $\tilde{X}$ the set of equivalence classes, endowed with the quotient topology, and by $\pi : X \to \tilde{X}$ the canonical projection. Let $\tilde{f} : \tilde{X} \to N$ be the map with the property $\tilde{f} \circ \pi = f$.

The property LC2 is hereditary: it is invariant to the choice of arbitrary small neighbourhoods V_x. We obtain the following proposition (see Lemmas 1.8.1, 1.8.2. Sleewaegen [30]) using topological reasoning and the hereditarity of LC2:

Lemma 1.16 *The space $\tilde{X}$ is Hausdorff, connected and compact. The function $\tilde{f}$ is proper, the preimage $\tilde{f}^{-1}(y)$ has a finite number of elements and it satisfies the improved version of LC2: for any $\tilde{x} \in \tilde{X}$ there are a neighbourhood $V_{\tilde{x}}$ of $\tilde{x}$, a pointed closed convex set $(C_{\tilde{x}}, \tilde{f}(\tilde{x}))$ and $r_{\tilde{x}} > 0$ such that the restriction $\tilde{f} : V_{\tilde{x}} \to C_{\tilde{x}} \cap B(f(\tilde{x}), r_{\tilde{x}})$ is a homeomorphism on the image.*

Step 2. Denote by l the length functional in the compact set $f(X)$ and by d_l the induced length distance. We begin by constructing a distance on $\tilde{X}$, which comes from the distance d_l and the function $\tilde{f}$.

For any two points $\tilde{x}, \tilde{y} \in \tilde{X}$ let

$$\tilde{d}(\tilde{x}, \tilde{y}) \; = \; \inf \left\{ l(\tilde{f} \circ \gamma) : \gamma : [a, b] \to \tilde{X} \text{ continuous}, \gamma(a) = \tilde{x} \, , \; \gamma(b) = \tilde{y} \right\}$$

We have the following immediate inequalities:

$$\tilde{d}(\tilde{x}, \tilde{y}) \; \geq \; d_l(\tilde{f}(\tilde{x}), \tilde{f}(\tilde{y})) \; \geq \; d(\tilde{f}(\tilde{x}), \tilde{f}(\tilde{y}))$$

In order to check that d_l is a distance it suffices to show that if $\tilde{d}(\tilde{x}, \tilde{y}) \; = \; 0$ then $\tilde{x} = \tilde{y}$. Indeed, we have $\tilde{f}(\tilde{x}) \; = \; \tilde{f}(\tilde{y})$ and there are two possibilities. If $\tilde{y} \in V_{\tilde{x}}$ (neighbourhood chosen as in the improved version of LC2) then $\tilde{y} = \tilde{x}$. Otherwise, if $\tilde{y} \notin V_{\tilde{x}}$ then we have arbitrary short closed curves in Y which start and end at $\tilde{f}(\tilde{x})$ but which go out of $C_{\tilde{x}} \cap B(f(\tilde{x}), r_{\tilde{x}})$, which is in contradiction with $r_{\tilde{x}} > 0$.

Use now Gromov Theorem 1.9 to prove that $(\tilde{X}, \tilde{d})$ is a compact path metric space. Hopf-Rinow Theorem 1.11 ensures us that any two points $\tilde{x}, \tilde{y} \in \tilde{X}$ can be joined by a global minimising geodesic. Moreover, such a geodesic projects on a global d_l geodesic in $f(X)$, because of the construction of the distance $\tilde{d}$.

Because $f(X) \; = \; \tilde{f}(\tilde{X})$ is compact, there is a $K \geq 1$ such that we can cover $f(X)$ with a finite number of sets of the form $C_{\tilde{x}^k} \cap B(\tilde{f}(\tilde{x}^k), r_k)$ and $C_{\tilde{x}^k} \cap B_N(\tilde{f}(\tilde{x}^k), Kr_k) \; \subset \; f(X)$. Then any d_l global geodesic is made by gluing a finite number of global geodesics in Y and the question is if such a curve is local geodesic in the neighbourhood of the gluing points. But the improved LC2 property ensures that. Indeed, take any two sufficiently closed points $\tilde{f}(\tilde{x})$ and $\tilde{f}(\tilde{y})$ on a d_l geodesic $\tilde{f} \circ \gamma$, with γ global $\tilde{d}$ geodesic. Then there is a neighbourhood $V_{\tilde{x}}$ of $\tilde{x}$ such that $\tilde{y} \in V_{\tilde{x}}$ and $\tilde{f} \circ \gamma$ is a Y global geodesic in $f(X)$ which joins $\tilde{f}(\tilde{x})$ and $\tilde{f}(\tilde{y})$. $\blacksquare$

The compactness assumption on the space X is not important. For example if the constant K in the definition of closed convex pointed set can be chosen independent of x in LC2 then the compactness assumption is no longer needed.

The classical Local-to-Global Principle is stated for vectorial target space Y, hence an arbitrary Euclidean space. Such spaces admit unique geodesic between any two different points, therefore the conclusion of the principle is that the image $f(X)$ is convex.

It seems hard to check that the function f satisfies LC2. In the case of the moment map this is a consequence of the existence of roots for some unitary representations. Note here the appearance of group theory.

1.3   Distances between metric spaces

The references for this section are Gromov [13], chapter 3, and Burago & al. [5] section 7.4. There are several definitions of distances between metric spaces. The very fertile idea of introducing such distances belongs to Gromov.

In order to introduce the Hausdorff distance between metric spaces, recall the Hausdorff distance between subsets of a metric space.

Definition 1.17 *For any set $A \subset X$ of a metric space and any $\varepsilon > 0$ set the ε neighbourhood of A to be*

$$A_\varepsilon = \cup_{x \in A} B(x, \varepsilon)$$

The Hausdorff distance between $A, B \subset X$ is defined as

$$d_H^X(A, B) = \inf \{\varepsilon > 0 : A \subset B_\varepsilon , \ B \subset A_\varepsilon\}$$

By considering all isometric embeddings of two metric spaces X, Y into an arbitrary metric space Z we obtain the Hausdorff distance between X, Y (Gromov [13] definition 3.4).

Definition 1.18 *The Hausdorff distance $d_H(X, Y)$ between metric spaces X Y is the infimum of the numbers*

$$d_H^Z(f(X), g(Y))$$

for all isometric embeddings $f : X \to Z$, $g : Y \to Z$ in a metric space Z.

If X, Y are compact then $d_H(X, Y) < +\infty$. Indeed, let Z be the disjoint union of X, Y and $M = \max\{diam(X), diam(Y)\}$. Define the distance on Z to be

$$d^Z(x, y) = \begin{cases} d^X(x, y) & x, y \in X \\ d^Y(x, y) & x, y \in Y \\ \frac{1}{2}M & \text{otherwise} \end{cases}$$

Then $d_H^Z(X, Y) < +\infty$.

The Hausdorff distance between isometric spaces equals 0. The converse is also true (Gromov *op. cit.* proposition 3.6) in the class of compact metric spaces.

Theorem 1.19 *If X, Y are compact metric spaces such that $d_H(X, Y) = 0$ then X, Y are isometric.*

For the proof of the theorem we need the Lipschitz distance (*op. cit. definition 3.1*) and a criterion for convergence of metric spaces in the Hausdorff distance (*op. cit. proposition 3.5*). We shall give the definition of Gromov for Lipschitz distance and the first part of the mentioned proposition.

Definition 1.20 *The Lipschitz distance $d_L(X, Y)$ between bi-Lipschitz homeomorphic metric spaces X, Y is the infimum of*

$$| \log \ dil(f) | + | \log \ dil(f^{-1}) |$$

for all $f : X \to Y$, bi-Lipschitz homeomorphisms.

Obviously, if $d_L(X, Y) = 0$ then X, Y are isometric. Indeed, by definition we have a sequence $f_n : X \to Y$ such that $dil(f_n) \to 1$ as $n \to \infty$. Extract an uniformly convergent subsequence; the limit is an isometry.

Definition 1.21 *A ε-net in the metric space X is a set $P \subset X$ such that $P_\varepsilon = X$. The separation of the net P is*

$$sep(P) \;=\; \inf \{d(x,y) \,:\, x \neq y \,,\, x,y \in P\}$$

A ε-isometry between X and Y is a function $f : X \to Y$ such that dis $f \leq \varepsilon$ and $f(X)$ is a ε net in Y.

The following proposition gives a connection between convergence of metric spaces in the Hausdorff distance and convergence of ε nets in the Lipschitz distance.

Definition 1.22 *A sequence of metric spaces (X_i) converges in the sense of Gromov-Hausdorff to the metric space X if $d_H(X, X_i) \to 0$ as $i \to \infty$. This means that there is a sequence $\eta_i \to 0$ and isometric embeddings $f_i : X \to Z_i$, $g_i : X_i \to Z_i$ such that*

$$d_H^{Z_i}(f_i(X), g_i(X_i)) \;<\; \eta_i$$

Proposition 1.23 *Let (X_i) be a sequence of metric spaces converging to X and (η_i) as in definition 1.22. Then for any $\varepsilon > 0$ and any ε-net $P \subset X$ with strictly positive separation there is a sequence $P_i \subset X_i$ such that*

1. *P_i is a $\varepsilon + 2\eta_i$ net in X_i,*

2. *$d_L(P_i, P) \to 0$ as $i \to \infty$, uniformly with respect to P, P_i.*

We shall not use further this proposition, given for the sake of completeness.

The next proposition is corollary 7.3.28 (b), Burago & al. [5]. The proof is adapted from Gromov, proof of proposition 3.5 (b).

Proposition 1.24 *If there exists a ε-isometry between X, Y then $d_H(X, Y) < 2\varepsilon$.*

Proof.   Let $f : X \to Y$ be the ε-isometry. On the disjoint union $Z = X \cup Y$ extend the distances d^X, d^Y in the following way. Define the distance between $x \in X$ and $y \in Y$ by

$$d(x,y) \;=\; \inf \{d^X(x,u) + d^Y(f(u),y) + \varepsilon\}$$

This gives a distance d^Z on Z. Check that $d_H^Z(X, Y) < 2\varepsilon$.   ■

The proof of theorem 1.19 follows as an application of previous propositions.

1.4   Metric tangent cones

The local geometry of a path metric space is described with the help of Gromov-Hausdorff convergence of pointed metric spaces.

This is definition 3.14 Gromov [13].

Definition 1.25 *The sequence of pointed metric spaces (X_n, x_n, d_n) converges in the sense of Gromov-Hausdorff to the pointed space (X, x, d_X) if for any $r > 0$, $\varepsilon > 0$ there is $n_0 \in N$ such that for all $n \geq n_0$ there exists $f_n : B_n(x_n, r) \subset X_n \to X$ such that:*

(1) $f_n(x_n) = x$,

(2) the distorsion of f_n is bounded by ε: dis $f_n < \varepsilon$,

(3) $B_X(x, r) \subset (f_n(B_n(x_n, r)))_\varepsilon$.

The Gromov-Hausdorff limit is defined up to isometry, in the class of compact metric spaces (Proposition 3.6 Gromov [13]) or in the class of locally compact cones. Here it is the definition of a cone.

Definition 1.26 *A pointed metric space* (X, x_0) *is called a cone if for any* $\lambda > 0$ *there is a map* $\delta_\lambda : X \to X$ *such that* $\delta_\lambda(x_0) = x_0$ *and for any* $x, y \in X$

$$d(\delta_\lambda(x), \delta_\lambda(y)) = \lambda\, d(x, y)$$

Such a map is called a dilatation with center x_0 *and coefficient* λ.

The local geometry of a metric space X in the neighbourhood of $x_0 \in X$ is described by the tangent space to (X, x_0) (if such object exists).

Definition 1.27 *The tangent space to* (X, x_0) *is the Gromov-Hausdorff limit*

$$(T_{x_0}, 0, d_{x_0}) = \lim_{\lambda \to \infty} (X, x_0, \lambda d)$$

Three remarks are in order:

1. The tangent space is obviously a cone. We shall see that in a large class of situations is also a group, hence a graded nilpotent group, called for short Carnot group.

2. The tangent space comes with a metric inside. This space is path metric (Proposition 3.8 Gromov [13]).

3. The tangent cone is defined up to isometry, therefore there is no way to use the metric tangent cone definition to construct a tangent bundle. For a modification of the definitions in this direction see Margulis, Mostow [21].

One can ask for a classification of metric spaces which admit everywhere the same (up to isometry) tangent cone. Such classification results exist, for example in the category of Lipschitz N manifold.

Definition 1.28 *Let* N *be a cone with a path metric distance. A Lipschitz* N *manifold is an (equivalence class of, or a maximal) atlas over the cone* N *such that the change of charts is locally bi-Lipschitz.*

We shall see in section 2.5 a rigidity (i.e. classification result) property of manifolds. Also, we shall prove in section 4.3 that a sub-Riemannian manifold (see definition 1.32) which is also a Lie group does not admit a manifold structure over the nilpotentisation of the defining distribution, such that the operation and the group exponential to be smooth. We will be forced therefore to modify the definition of a manifold in order to show that a group with a left invariant distribution has a manifold structure over the nilpotentisation. The modification will lead, surprising but natural, to (a generalisation of) the moment map.

1.5   Examples of path metric spaces

In this section are given examples of path metric spaces. These are:

- Riemannian manifolds

- Finsler manifolds

- Carnot-Carathéodory manifolds

Let X be a Riemannian manifold and $l(f)$ be the length of the curve f with respect to the metric on X, if f is piecewise C^1, otherwise $l(f) = +\infty$. This is, off course, a metric structure on the Riemannian X. The class of curves with *potential finite* length can be enlarged or restricted by using analytical arguments.

Change the metric on X by a Finsler metric, i.e. consider a continuous mapping $\Delta : TX \to R$ such that for any $x \in X$, $\lambda \in R$ and $v \in T_x X$

$$\Delta_x(\lambda v) = \mid \lambda \mid \Delta_x(v)$$

Then for any piecewise C^1 curve $c : [a, b] \to X$ define it's length by

$$l(c) = \int_a^b \Delta_{c(t)}(\dot{c}(t)) \; \mathrm{d}t$$

This give rise to a (non-oriented) length structure on X.

Instead of a Finsler metric Δ consider an oriented version obtained by imposing a milder condition upon Δ: to be positively one-homogeneous. That means: for any $x \in X$, $\lambda \geq 0$ and $v \in T_x X$

$$\Delta_x(\lambda v) = \lambda \Delta_x(v)$$

The length structure associated to Δ is now oriented.

A particular case is the one of a Finsler-Minkowski metric. Suppose first that (X, g) is Riemannian (g is the metric). Take $Q_x \subset T_x X$ a convex bounded set and denote by χ_x the characteristic function

$$\chi_x(v) \;\; = \;\; \begin{cases} 0 & \text{if } v \in Q_x \\ +\infty & \text{otherwise} \end{cases}$$

Suppose that Q_x contains a ball centered in the origin. Define now the Finsler-Minkowski metric

$$\Delta_x(v) \;\; = \;\; \sup \{g(v, u) - \chi_x(u) \; : \; u \in T_x X\}$$

that is Δ_x is the polar of χ_x. From the hypothesis, there are positive constants c, c' such that ($\mid v \mid_x = \sqrt{g(v, v)}$)

$$c \mid v \mid_x \leq \Delta_x(v) \leq c' \mid v \mid_x$$

The length structure associated to Δ is in general oriented. It becomes non oriented if the set Q_x is symmetric with respect to the origin.

In almost all examples we have a Riemannian manifold and a length structure constructed from a map Δ. We shall suppose now that the manifold is complete with respect to the Riemannian distance.

Proposition 1.29 *The Hopf-Rinow theorem holds on a Finsler-Minkowski manifold, provided that it is complete with respect to the initial Riemannian distance and that Δ is convex.*

A more particular case is the following:

Proposition 1.30 *Let G be a Lie group and U be a left invariant Finsler-Minkowski metric on G. Then (G, d_U) is complete and locally compact.*

Proof.   The exponential map exp is a diffeomorphism from $\mathcal{V}(0) \ni U \subset g$ to $\exp U$. From the hypothesis we get that $\exp_{|_U}$ is d_U Lipschitz. The exponential is also open. G is locally a linear group, hence from the hypothesis there is a constant c such that

$$d_U(e, \exp v) \geq c\|v\|$$

Therefore the metric topology given by d_U is the same as the manifold topology, hence locally compact.

Take now a Cauchy sequence (x_h). For a sufficiently small $\varepsilon > 0$ there is $N = N(\varepsilon)$ such that $x_N^{-1} x_h$ lies in a compact neighbourhood of the identity element e for any $h \geq N$. One can extract a convergent subsequence from it and the proof finishes.   ∎

As a corollary of propositions 1.29 and 1.30 we have the following concrete Hopf-Rinow theorem:

Corollary 1.31 *On a Lie group endowed with a convex left (or right) invariant Finsler-Minkowski metric any two points can be joined by a geodesic.*

Another class of examples is provided by sub-Riemannian manifolds.

Definition 1.32 *A sub-Riemannian (SR) manifold is a triple (M, H, g), where M is a connected manifold, H is a subbundle of TM, named horizontal bundle or distribution, and g is a metric (inner-product) on the horizontal bundle.*

A horizontal curve is a continuous, almost everywhere differentiable curve, whose tangent lies in the horizontal bundle.

The length of a horizontal curve $c : [a, b] \to M$ is

$$l(c) \; = \; \int_a^b g(c(t), \dot{c}(t)) \; dt$$

The SR manifold is called a Carnot-Carathéodory (CC) space if any two points can be joined by a finite length horizontal curve.

A CC space is a path metric space with the Carnot-Carathéodory distance induced by the length l:

$$d(x, y) \; = \; \inf \{l(c) \, : \, c : [a, b] \to M \; , \; c(a) = x \; , \; c(b) = y\}$$

The particular case that will be interesting for us is the following one. Consider a connected real Lie group and a vector space $V \subset g$, which generates the whole algebra g by Lie brackets. Then we shall see that the left translations of V provide a non-integrable distribution V on G (easy form of Chow theorem). For any (left or right invariant) metric defined on D we have an associated Carnot-Carathéodory distance.

2 Carnot groups

2.1 Structure of Carnot groups

Definition 2.1 *A Carnot (or stratified nilpotent) group is a connected simply connected group N with a distinguished vectorspace V_1 such that the Lie algebra of the group has the direct sum decomposition:*

$$n = \sum_{i=1}^{m} V_i , \quad V_{i+1} = [V_1, V_i]$$

The number m is the step of the group. The number

$$Q = \sum_{i=1}^{m} i \ dim V_i$$

is called the homogeneous dimension of the group.

Because the group is nilpotent and simply connected, the exponential mapping is a diffeomorphism. We shall identify the group with the algebra, if is not locally otherwise stated.

The structure that we obtain is a set N endowed with a Lie bracket and a group multiplication operation.

The Baker-Campbell-Hausdorff formula shows that the group operation is polynomial. It is easy to see that the Lebesgue measure on the algebra is (by identification) a bi-invariant measure.

We give further examples of such groups:

(1.) R^n with addition is the only commutative Carnot group.

(2.) The Heisenberg group is the first non-trivial example. This is the group $H(n) = R^{2n} \times R$ with the operation:

$$(x, \bar{x})(y, \bar{y}) = (x + y, \bar{x} + \bar{y} + \frac{1}{2}\omega(x, y))$$

where ω is the standard symplectic form on R^{2n}. The Lie bracket is

$$[(x, \bar{x}), (y, \bar{y})] = (0, \omega(x, y))$$

The direct sum decomposition of (the algebra of the) group is:

$$H(n) = V + Z , \quad V = R^{2n} \times \{0\} , \quad Z = \{0\} \times R$$

Z is the center of the algebra, the group has step 2 and homogeneous dimension $2n + 2$.

(3.) H-type groups. These are two step nilpotent Lie groups N endowed with an inner product $(\cdot, \cdot)$, such that the following *orthogonal* direct sum decomposition occurs:

$$N = V + Z$$

Z is the center of the Lie algebra. Define now the function

$$J : Z \to End(V) \;, \;\; (J_z x, x') \; = \; (z, [x, x'])$$

The group N is of H-type if for any $z \in Z$ we have

$$J_z \circ J_z \; = \; - \mid z \mid^2 \; I$$

From the Baker-Campbell-Hausdorff formula we see that the group operation is

$$(x, z)(x', z') \; = \; (x + x', z + z' + \frac{1}{2}[x, x'])$$

These groups appear naturally as the nilpotent part in the Iwasawa decomposition of a semisimple real group of rank one. (see [6])

(4.)   The last example is the group of $n \times n$ upper triangular matrices, which is nilpotent of step $n - 1$. This example is important because any Carnot group is isomorphic with a subgroup of a group of upper triangular matrices.

Any Carnot group admits a one-parameter family of dilatations. For any $\varepsilon > 0$, the associated dilatation is:

$$x \; = \; \sum_{i=1}^{m} x_i \; \mapsto \; \delta_\varepsilon x \; = \; \sum_{i=1}^{m} \varepsilon^i x_i$$

Any such dilatation is a group morphism and a Lie algebra morphism.

In fact the class of Carnot groups is characterised by the existence of dilatations.

Proposition 2.2 *Suppose that the Lie algebra* $\mathfrak{g}$ *admits an one parameter group* $\varepsilon \in (0, +\infty) \mapsto \delta_\varepsilon$ *of simultaneously diagonalisable Lie algebra isomorphisms. Then* $\mathfrak{g}$ *is the algebra of a Carnot group.*

Proof.   The hypothesis means that there is a direct sum decomposition of

$$\mathfrak{g} \; = \; \oplus_{i=1}^{m} V_i$$

such that for any $\varepsilon > 0$ we have

$$x \; = \; \sum_{i=1}^{m} x_i \; \mapsto \; \delta_\varepsilon x \; = \; \sum_{i=1}^{m} f_i(\varepsilon) x_i$$

with f_i continuous, moreover

$$\delta_\varepsilon \circ \delta_\mu \; = \; \delta_{\varepsilon\mu}$$

for any $\varepsilon, \mu > 0$ and $\delta_1 \; = \; id$. From this we get $f_i(\varepsilon) \; = \; \varepsilon^{\alpha_i}$. Each δ_ε is also Lie algebra morphism, therefore ($x_i \in V_i$, $x_j \in V_j$)

$$\delta_\varepsilon [x_i, x_j] \; = \; [\delta_\varepsilon x_i, \delta_\varepsilon x_j] \; = \; \varepsilon^{\alpha_i + \alpha_j}[x_i, x_j]$$

We conclude that $\alpha_i + \alpha_j$ is also an eigenvalue, unless $[x_i, x_j] = 0$. The direct sum decomposition of $\mathfrak{g}$ is finite therefore $\alpha_i \; = \; i\alpha$ and, more important, $[V_i, V_j] \; = \; V_{i+j}$.

In conclusion $\mathfrak{g}$ is the Lie algebra of a Carnot group and δ_ε are the dilatations, up to a scale factor α. ∎

We can always find inner products on N such that the decomposition $N = \sum_{i=1}^m V_i$ is an orthogonal sum.

We shall endow the group N with a structure of a sub-Riemannian manifold now. For this take the distribution obtained from left translates of the space V_1. The metric on that distribution is obtained by left translation of the inner product restricted to V_1.

If V_1 Lie generates (the algebra) N then any element $x \in N$ can be written as a product of elements from V_1. An useful lemma is:

Lemma 2.3 *Let N be a Carnot group and $X_1, ..., X_p$ an orthonormal basis for V_1. Then there is a a natural number M and a function $g : \{1, ..., M\} \rightarrow \{1, ..., p\}$ such that any element $x \in N$ can be written as:*

$$x = \prod_{i=1}^M \exp(t_i X_{g(i)}) \tag{2.1.1}$$

Moreover, if x is sufficiently close (in Euclidean norm) to 0 then each t_i can be chosen such that $\mid t_i \mid \leq C\|x\|^{1/m}$

Proof.   This is a slight reformulation of Lemma 1.40, Folland, Stein [8]. We have used Proposition 2.8 applied for the homogeneous norm $|\cdot|_1$ (see below). ∎

This means that there is a horizontal curve joining any two points. Hence the distance

$$d(x, y) = \inf\left\{\int_a^b \|c^{-1}\dot{c}\| \ \mathrm{d}t \ : \ c(a) = x, \ c(b) = y, \ c^{-1}\dot{c} \in V_1\right\}$$

is finite for any two $x, y \in N$. The distance is obviously left invariant.

A Carnot group N can be therefore described in the following way.

Proposition 2.4 (N, d) *is a path metric cone and also a group such that left translations are isometries. Also, dilatations are group isomorphisms.*

Associate to the CC distance d the function:

$$\mid x \mid_d = d(0, x)$$

This function has the following properties:

(a) the set $\{x \in G \ : \mid x \mid_d = 1\}$ does not contain 0.

(b) $\mid x^{-1} \mid_d = \mid x \mid_d$ for any $x \in G$.

(c) $\mid \delta_\varepsilon x \mid_d = \varepsilon \mid x \mid_d$ for any $x \in G$ and $\varepsilon > 0$.

(d) for any $x, y \in N$ we have

$$\mid xy \mid_d \leq \mid x \mid_d + \mid y \mid_d$$

Proposition 2.5 *Balls with respect to the distance d are open and there is a basis for the topology on N (as a manifold) made by d balls.*

Proof. We have to prove that d induces the same topology on N as the manifold topology, which is the standard topology induced by Euclidean distance on N.

Set d_R to be the Riemannian distance induced by a metric which extends the metric on D and makes the direct sum $N = \oplus_i V_i$ orthogonal. We shall have then the inequality

$$d_R(x,y) \leq d(x,y)$$

for all $x, y \in N$.

The function $| \cdot |_d$ is also continuous. Indeed, from Lemma 2.3 we see that

$$| x |_d \leq \sum_{i=1}^{M} | t_i |$$

with $t_i \to 0$ when $\|x\| \to 0$.

These two facts prove the thesis. ∎

The map $| \cdot |_d$ looks like a norm. However it is intrinsically defined and hard to work with. This is the reason for the introduction of homogeneous norms.

Definition 2.6 *A continuous function from $x \mapsto | x |$ from G to $[0, +\infty)$ is a homogeneous norm if*

(a) the set $\{x \in G : | x |= 1\}$ does not contain 0.

(b) $| x^{-1} |=| x |$ for any $x \in G$.

(c) $| \delta_\varepsilon x |= \varepsilon | x |$ for any $x \in G$ and $\varepsilon > 0$.

Homogeneous norms exist. An example is:

$$| x |_1 = \sum_{i=1}^{m} | x_i |^{1/i}$$

Proposition 2.7 *Any two homogeneous norms are equivalent. Let $| \cdot |$ be a homogeneous norm. Then the set $\{x : | x |= 1\}$ is compact.*

There is a constant $C > 0$ such that for any $x, y \in N$ we have:

$$| xy | \leq C (| x | + | y |)$$

Proof. We show first that there are constants $c, C > 0$ such that for any $x \in N$ we have

$$c | x |_1 \leq | x | \leq C | x |$$

For this consider the compact set $B = \{x : | x |_1= 1\}$. The norm $| \cdot |_1$ has a minimum c and a maximum C on the set B ($0 < c \leq C$). Let now $x \in N$, $x \neq 0$. Then $x = \delta_\varepsilon y$ with $\varepsilon = | x |_1$. The homogeneity of the norms show that $| y |_1= 1$ and the proof is done.

A consequence of the equivalence of the norms is that for any norm $| \cdot |$ the set $\{x : \ | \, x \, | = 1\}$ is compact.

For the third assertion remark that the set in N^2 of all (x, y) such that $| \, x \, | + | \, y \, | \leq 1$ is compact. The function $(x, y) \mapsto | \, x + y \, |$ attains a maximum C on this set. Use again the homogeneity of the norm to finish the proof. ∎

Let us denote by $\| \cdot \|$ the Euclidean norm on N. The following estimate is easily obtained using homogeneity, as in the previous proposition.

Proposition 2.8 *Let $| \cdot |$ be a homogeneous norm. Then there are constants $c, C > 0$ such that for any $x \in N$, $| \, x \, | < 1$, we have*

$$c\|x\| \ \leq \ | \, x \, | \ \leq \ C\|x\|^{1/m}$$

The balls in CC-distance look roughly like boxes (as it is the case with the Euclidean balls). A box is a set

$$Box(r) \ = \ \left\{ x \ = \ \sum_{i=1}^{m} x_i \ : \ \|x_i\| \leq r^i \right\}$$

Proposition 2.9 *("Ball-Box theorem") There are positive constants c, C such that for any $r > 0$ we have:*
$$Box(cr) \ \subset \ B(0, r) \ \subset \ Box(Cr)$$

Proof.   The box $Box(r)$ is the ball with radius r with respect to the homogeneous norm
$$| \, x \, |_{\infty} = \ \max \left\{ \|x_i\|^{1/i} \right\}$$
By Proposition 2.7 the norm $| \cdot |_{\infty}$ is equivalent with $| \cdot |_d$, which proves the thesis. ∎

A consequence of the Ball-Box Theorem is

Theorem 2.10 *(Hopf-Rinow theorem for Carnot groups) Any two points in a Carnot group can be joined by a geodesic.*

Proof.   Take two arbitrary points in N. Because the distance is left invariant we can choose $x = 0$. Because dilatations change the distances by a constant factor, we can suppose that $y \in B(0, 1)$.

A previous proposition shows that the topology generated by a homogeneous norm and the topology generated by the Euclidean norm are the same. Therefore the space is metrically complete. The Ball-Box Theorem implies that the ball $B(0, 1)$ is compact. We are in the assumptions of the abstract Hopf-Rinow theorem 1.11. ∎

An easy computation shows that

$$\mathcal{L}(B(0, \delta_\varepsilon r)) \;=\; \varepsilon^Q \mathcal{L}(B(0, r))$$

As a consequence the Lebesgue measure is absolutely continuous with respect to the Hausdorff measure $\mathcal{H}^Q$. Because of the invariance with respect to the group operation it follows that the Lebesgue measure is a multiple of the mentioned Hausdorff measure.

The following theorem is Theorem 2. from Mitchell [23], in the particular case of Carnot groups.

Theorem 2.11 *The ball $B(0, 1)$ has Hausdorff dimension Q.*

Proof. We know that the volume of a ball with radius ε is $c\varepsilon^Q$.

Consider a maximal filling of $B(0, 1)$ with balls of radius ε. There are N_ε such balls in the filling; an upper bound for this number is:

$$N_\varepsilon \;\le\; 1/\varepsilon^Q$$

The set of concentric balls of radius 2ε cover $B(0, 1)$; each of these balls has diameter smaller than 4ε, so the Hausdorff α measure of $B(0, 1)$ is smaller than

$$\lim_{\varepsilon \to 0} N_\varepsilon (2\varepsilon)^\alpha$$

which is 0 if $\alpha > Q$. Therefore the Hausdorff dimension is smaller than Q.

Conversely, given any covering of $B(0, 1)$ by sets of diameter $\le \varepsilon$, there is an associated covering with balls of the same diameter; the number M_ε of this balls has a lower bound:

$$M_\varepsilon \;\ge\; 1/\varepsilon^Q$$

thus there is a lower bound

$$\sum_{cover} \varepsilon^\alpha \;\ge\; \varepsilon^\alpha / \varepsilon^Q$$

which shows that if $\alpha < Q$ then $\mathcal{H}^\alpha(B(0, 1)) = \infty$. Therefore the Hausdorff dimension of the ball is greater than Q. ∎

We collect the important facts discovered until now: let N be a Carnot group endowed with the left invariant distribution generated by V_1 and with an Euclidean norm on V_1.

(a) If V_1 Lie-generates the whole Lie algebra of N then any two points can be joined by a horizontal path.

(b) The metric topology of N is the same as Euclidean topology.

(c) The ball $B(0, r)$ looks roughly like the box $\left\{ x \;=\; \sum_{i=1}^m x_i \;:\; \|x_i\| \le r^i \right\}$.

(d) the Hausdorff measure $\mathcal{H}^Q$ is group invariant and the Hausdorff dimension of a ball is Q.

(e) the Hopf-Rinow Theorem applies.

(f) there is a one-parameter group of dilatations, where a dilatation is an isomorphism δ_ε of N which transforms the distance d in εd.

2.2   Pansu differentiability

A Carnot group has it's own concept of differentiability, introduced by Pansu [25].

In Euclidean spaces, given $f : R^n \to R^m$ and a fixed point $x \in R^n$, one considers the difference function:

$$X \in B(0,1) \subset R^n \; \mapsto \; \frac{f(x + tX) - f(x)}{t} \in R^m$$

The convergence of the difference function as $t \to 0$ in the uniform convergence gives rise to the concept of differentiability in it's classical sense. The same convergence, but in measure, leads to approximate differentiability. Other topologies might be considered (see Vodop'yanov [33]).

In the frame of Carnot groups the difference function can be written using only dilatations and the group operation. Indeed, for any function between Carnot groups $f : G \to P$, for any fixed point $x \in G$ and $\varepsilon > 0$ the finite difference function is defined by the formula:

$$X \in B(1) \subset G \; \mapsto \; \delta_\varepsilon^{-1} \left(f(x)^{-1} f\left(x \delta_\varepsilon X \right) \right) \in P$$

In the expression of the finite difference function enters δ_ε^{-1} and δ_ε, which are dilatations in P, respectively G.

Pansu's differentiability is obtained from uniform convergence of the difference function when $\varepsilon \to 0$.

The derivative of a function $f : G \to P$ is linear in the sense explained further. For simplicity we shall consider only the case $G = P$. In this way we don't have to use a heavy notation for the dilatations.

Definition 2.12 *Let N be a Carnot group. The function $F : N \to N$ is linear if*

(a) F is a group morphism,

(b) for any $\varepsilon > 0$ $F \circ \delta_\varepsilon \; = \; \delta_\varepsilon \circ F$.

We shall denote by $HL(N)$ the group of invertible linear maps of N, called the linear group of N.

The condition (b) means that F, seen as an algebra morphism, preserves the grading of N.

The definition of Pansu differentiability follows:

Definition 2.13 *Let $f : N \to N$ and $x \in N$. We say that f is (Pansu) differentiable in the point x if there is a linear function $Df(x) : N \to N$ such that*

$$\sup \left\{ d(F_\varepsilon(y), Df(x)y) \; : \; y \in B(0,1) \right\}$$

converges to 0 when $\varepsilon \to 0$. The functions F_ε are the finite difference functions, defined by

$$F_t(y) \; = \; \delta_t^{-1} \left(f(x)^{-1} f(x \delta_t y) \right)$$

The definition says that f is differentiable at x if the sequence of finite differences F_t uniformly converges to a linear map when t tends to 0.

We are interested to see how this differential looks like. For any $f : N \to N$, $x, y \in N$ $Df(x)y$ means

$$Df(x)y = \lim_{t \to 0} \delta_t^{-1} \left(f(x)^{-1} f(x \delta_t y) \right)$$

provided that the limit exists.

Proposition 2.14 *Let $f : N \to N$, $y, z \in N$ such that:*

(i) *$Df(x)y$, $Df(x)z$ exist for any $x \in N$.*

(ii) *The map $x \mapsto Df(x)z$ is continuous.*

(iii) *$x \mapsto \delta_t^{-1} \left(f(x)^{-1} f(x \delta_t z) \right)$ converges uniformly to $Df(x)z$.*

Then for any $a, b > 0$ and any $w = \delta_a x \delta_b y$ the limit $Df(x)w$ exists and we have:

$$Df(x)\delta_a x \delta_b y = \delta_a Df(x)y \, \delta_b Df(x)z$$

Proof.   The proof is standard. Remark that if $Df(x)y$ exists then for any $a > 0$ the limit $Df(x)\delta_a y$ exists and

$$Df(x)\delta_a y = \delta_a Df(x)y$$

It is not restrictive therefore to suppose that $w = yz$, such that $\mid y \mid_d = 1$.

We write:

$$\delta_t^{-1} \left(f(x)^{-1} f(x \delta_t w) \right) = (1)(2)(3)$$

where

$$(1) = \delta_t^{-1} \left(f(x)^{-1} f(x \delta_t y) \right)$$

$$(2) = \delta_t^{-1} \left(f(x \delta_t y)^{-1} f(x \delta_t y \delta_t z) \right) \left(Df(x \delta_t y)z \right)^{-1}$$

$$(3) = Df(x \delta_t y)z$$

When t tends to 0 (i) implies that (1) tends to $Df(x)y$, (ii) implies that (3) tends to $Df(x)z$ and (iii) implies that (2) goes to 0. The proof is finished.   ∎

If f is Lipschitz then the previous proposition holds almost everywhere. In Proposition 3.2, Pansu [25] there is a more general statement:

Proposition 2.15 *Let $f : N \to N$ have finite dilatation and suppose that for almost any $x \in N$ the limits $Df(x)y$ and $Df(x)z$ exist. Then for almost any x and for any $w = \delta_a y \, \delta_b z$ the limit $Df(x)w$ exists and we have:*

$$Df(x)\delta_a x \delta_b y = \delta_a Df(x)y \, \delta_b Df(x)z$$

Proof. The idea is to use Proposition 2.14. We shall suppose first that f is Lipschitz. Then any finite difference function F_t is also Lipschitz.

Recall the general Egorov theorem:

Lemma 2.16 *(Egorov Theorem) Let μ be a finite measure on X and $(f_n)_n$ a sequence of measurable functions which converges μ almost everywhere (a.e.) to f. Then for any $\varepsilon > 0$ there exists a measurable set X_ε such that $\mu(X \setminus X_\varepsilon) < \varepsilon$ and f_n converges uniformly to f on X_ε.*

Proof. It is not restrictive to suppose that f_n converges pointwise to f. Define, for any $q, p \in \mathbf{N}$, the set:

$$X_{q,p} \ = \ \{x \in X \ : \ \ | f_n(x) - f(x) | \ < \ 1/p \ , \ \forall n \geq q\}$$

For fixed p the sequence of sets $X_{q,p}$ is increasing and at the limit it fills the space:

$$\bigcup_{q \in \mathbf{N}} X_{q,p} \ = \ X$$

Therefore for any $\varepsilon > 0$ and any p there is a $q(p)$ such that $\mu(X \setminus X_{q(p)p}) \ < \ \varepsilon/2^p$.

Define then

$$X_\varepsilon \ = \ \bigcap_{p \in N} X_{q(p)p}$$

and check that it satisfies the conclusion. ∎

We can improve the conclusion of Egorov Theorem by claiming that if μ is a Borel measure then each X_ε can be chosen to be open.

In our situation we can take X to be a ball in the group N and μ to be the Q Hausdorff measure, which is Borel. Then we are able to apply Proposition 2.14 on X_ε. When we tend ε to 0 we obtain the claim (for f Lipschitz). ∎

A consequence of this proposition is Corollaire 3.3, Pansu [25].

Corollary 2.17 *If $f : N \to N$ has finite dilatation and $X_1, ..., X_K$ is a basis for the distribution V_1 such that*

$$s \mapsto f(x\delta_s X_i)$$

is differentiable in $s = 0$ for almost any $x \in N$, then f is differentiable almost everywhere and the differential is linear, i.e. if

$$y \ = \ \prod_{i=1}^{M} \delta_{t_i} X_{g(i)}$$

then

$$Df(x)y \ = \ \prod_{i=1}^{M} \delta_{t_i} X_{g(i)}(f)(x)$$

With the definition of (Pansu) derivative at hand, it is natural to introduce the class of C^1 maps from N to M.

Definition 2.18 *Let N, M be Carnot groups. A function $f : N \to M$ is of class C^1 if it is Pansu derivable everywhere and the derivative is continuous as a function $Df : N \times N \to M \times M$. Here $N \times N$ and $M \times M$ are endowed with the product topology.*

Remark 2.19 *For example left translations $L_x : N \to N$, $L_x(y) = xy$ are C^1 but right translations $R_x : N \to N$, $L_x(y) = yx$ are not even derivable.*

We can introduce the class of C^1 N manifolds.

Definition 2.20 *Let N be a Carnot group. A C^1 N manifold is an (equivalence class of, or a maximal) atlas over the group N such that the change of charts is locally bi-Lipschitz.*

2.3   Rademacher theorem

A very important result is theorem 2, Pansu [25], which contains the Rademacher theorem for Carnot groups.

Theorem 2.21 *Let $f : M \to N$ be a Lipschitz function between Carnot groups. Then f is differentiable almost everywhere.*

The proof is based on the corollary 2.17 and the technique of development of a curve, Pansu sections 4.3 - 4.6 [25].

Let $c : [0,1] \to N$ be a Lipschitz curve such that $c(0) = 0$. To any division $\Sigma : 0 = t_0 < .... < t_n = 1$ of the interval $[0,1]$ is a associated the element $\sigma_\Sigma \in N$ (the algebra) given by:

$$\sigma_\Sigma = \sum_{k=0}^{n} c(t_k)^{-1} c(t_{k+1})$$

Lemma (18) Pansu [26] implies the existence of a constant $C > 0$ such that

$$\|\sigma_\Sigma - c(1)\| \leq C \left(\sum_{k=0}^{n} \|c(t_k)^{-1} c(t_{k+1})\| \right)^2 \qquad (2.3.2)$$

Take now a finer division Σ' and look at the interval $[t_k, t_{k+1}]$ divided further by Σ' like this:

$$t_k = t'_l < ... < t'_m = t_{k+1}$$

The estimate (2.3.2) applied for each interval $[t_k, t_{k+1}]$ lead us to the inequality:

$$\|\sigma_{\Sigma'} - \sigma_\Sigma\| \leq \sum_{k=0}^{n} d(c(t_k), c(t_{k+1}))^2$$

The curve has finite length (being Lipschitz) therefore the right hand side of the previous inequality tends to 0 with the norm of the division. Set

$$\sigma(s) \;=\; \lim_{\|\Sigma\|\to 0} \sigma_\Sigma(s)$$

where $\sigma_\Sigma(s)$ is relative to the curve c restricted to the interval $[0, s]$. The curve such defined is called the development of the curve c. It is easy to see that σ has the same length (measured with the Euclidean norm $\| \cdot \|$) as c. If we parametrise c with the length then we have the estimate:

$$\|\sigma(s)\, c(s)\| \;\leq\; Cs^2 \tag{2.3.3}$$

This shows that σ is a Lipschitz curve (with respect to the Euclidean distance). Indeed, prove first that σ has finite dilatation in almost any point, using (2.3.3) and the fact that c is Lipschitz. Then show that the dilatation is majorised by the Lipschitz constant of c. By the classical Rademacher theorem σ is almost everywhere derivable.

Remark 2.22 *The development of a curve can be done in an arbitrary Lie connected group, endowed with a left invariant distribution which generates the algebra. One should add some logarithms, because in the case of Carnot groups, we have identified the group with the algebra. The inequality (2.3.3) still holds.*

Conversely, given a curve σ in the algebra N, we can perform the inverse operation to development (called multiplicative integral by Pansu; we shall call it "lift"). Indeed, to any division Σ of the interval $[0, s]$ we associate the point

$$c_\Sigma(s) \;=\; \prod_{k=0}^{n} \left(\sigma(t_{k+1}) \;-\; \sigma(t_k) \right)$$

Define then

$$c(s) \;=\; \lim_{\|\Sigma\|\to 0} c_\Sigma(s)$$

Remark that if $\sigma([0,1]) \subset V_1$ and it is almost everywhere differentiable then c is a horizontal curve.

Fix $s \in [0, 1)$ and parametrise c by the length. Apply the inequality (2.3.3) to the curve $t \mapsto c(s)^{-1}c(s + t)$. We get the fact that the vertical part of $\sigma(s + t) - \sigma(s)$ is controlled by s^2. Therefore, if c is Lipschitz then σ is included in V_1.

Denote the i multiple bracket $[x, [x, ...[x, y]...]$ by $[x, y]_i$. For any division Σ of the interval $[0, s]$ set:

$$A_\Sigma^i(s) \;=\; \sum_{k=0}^{n} [\sigma(t_k), \sigma(t_{k+1})]_i$$

As before, one can show that when the norm of the division tends to 0, A_Σ^i converges. Denote by

$$A^i(s) \;=\; \int_0^s [\sigma, d\sigma]_i \;=\; \lim_{\|\Sigma\|\to 0} A_\Sigma^i(s)$$

the i-area function. The estimate corresponding to (2.3.3) is

$$\|A^i(s)\| \ \leq \ Cs^{i+1} \tag{2.3.4}$$

What is the signification of $A^i(s)$? The answer is simple, based on the well known formula of derivation of left translations in a group. Define, in a neighbourhood of 0 in the Lie algebra $\mathfrak{g}$ of the Lie group G, the operation

$$X \stackrel{\mathfrak{g}}{\cdot} Y \ = \ \log_G (\exp_G(X) \exp_G(Y))$$

The left translation by X is the function $L_X(Y) \ = \ X \stackrel{\mathfrak{g}}{\cdot} Y$. It is known that

$$DL_X(0)(Z) \ = \ \sum_{i=0} \frac{1}{(i+1)!} [X, Z]_i$$

It follows that if the group G is Carnot then $[X, Z]_i$ measures the infinitesimal variation of the V_i component of $L_X(Y)$, for $Y = 0$. Otherwise said,

$$\frac{d}{dt} \sigma_i(t) \ = \ \frac{1}{(i+1)!} \lim_{\varepsilon \to 0} \varepsilon^{-i} \int_t^{t+\varepsilon} [\sigma, d\sigma]_i$$

Because σ is Lipschitz, the left hand side exists for all i for a.e. t. The estimate (2.3.4) tells us that the right hand side equals 0. This implies the following proposition (Pansu, 4.1 [25]).

Proposition 2.23 *If c is Lipschitz then c is differentiable almost everywhere.*

Proof of theorem 2.21. The proposition 2.23 implies that we are in the hypothesis of corollary 2.17. ∎

Pansu-Rademacher theorem has been improved for Lipschitz functions $f : A \subset M \to N$, where M, N are Carnot groups and A is just measurable, in Vodop'yanov, Ukhlov [34]. Their technique differs from Pansu. It resembles with the one used in Margulis, Mostow [20], where it is proven that any quasi-conformal map from a Carnot-Carathéodory manifold to another is a.e. differentiable, without using Rademacher theorem. Same result for quasi-conformal maps on Carnot groups has been first proved by Koranyi, Reimann [17]. Finally Magnani [19] reproved a.e. differentiability of Lipschitz functions on Carnot groups, defined on measurable sets, continuing Pansu technique.

For a review of other connected results, such as approximate differentiability, differentiability in Sobolev topology, see the excellent Vodop'yanov [33].

2.4 Area formulas

The subject of this subsection is the change of variable formula. In order to prove it one needs more involved (though basic) knowledge on metric measure spaces. We shall just sketch the proofs, delaying for the second part of these notes the rigorous proofs.

Definition 2.24 *A metric measure (or mm) space is a metric space (X, d) endowed with a Borel measure μ satisfying the doubling condition, i.e. there is a positive constant C such that*

$$\mu(B(x, 2r)) \;\leq\; C\mu(B(x, r))$$

for any $x \in X$, $r > 0$.

A Carnot group is a metric measure space, endowed with the invariant volume measure (which is the Lebesgue measure). Indeed, this is a consequence of the existence of dilatations.

In any mm space the Vitali covering theorem holds:

Theorem 2.25 *(Vitali) Let (X, d, μ) be a mm space, $A \subset X$ and $\mathcal{F}$ a family of closed sets in X which covers A. If for any $x \in A$ and $\varepsilon > 0$ there exists $V \in \mathcal{F}$ such that $x \in V$ and diam $V < \varepsilon$ then one can extract from $\mathcal{F}$ a countable disjoint family $(V_i)_{i \in N}$ such that*

$$\mu(A \setminus \cup_{i \in N} V_i) \;=\; 0$$

Vitali covering theorem is the only technical ingredient needed to prove the following result (compare with Margulis, Mostow [20], lemma 2.3). But before this we need a definition.

Definition 2.26 *Let (X, d, μ) be a mm space and σ a measure on X. The lower spherical density of the measure σ at the point $x \in X$ is*

$$\frac{d\sigma}{d\mu}^-(x) \;=\; \liminf_{\varepsilon \to 0} \frac{\sigma(B(x, \varepsilon))}{\mu(B(x, \varepsilon))}$$

If (Y, σ) is measure space and $f : X \to Y$ is measurable then the pull-back of σ by f is the measure $f^*(\sigma)$ on X defined by:

$$f^*(\sigma)(A) \;=\; \sigma(f(A))$$

The Jacobian of f in $x \in X$ is the spherical density of $f^*(\sigma)$ in x:

$$Jf(x) \;=\; \frac{df^*(\sigma)}{d\mu}^-(x) \;=\; \liminf_{\varepsilon \to 0} \frac{\sigma(f(B(x, \varepsilon)))}{\mu(B(x, \varepsilon))}$$

Proposition 2.27 *Let (X, d, μ) be a mm space, (Y, σ) a finite measure space and $f : X \to Y$ be an injective measurable map. We have then the inequality:*

$$\int_X Jf(x) \; d\mu \;\leq\; \sigma(f(X)) \tag{2.4.5}$$

In order to establish the change of variables (or area formula) for f bi-Lipschitz between Carnot groups one has to prove first that in (2.4.5) we have equality. This is done by showing that $\mu(Z) = 0$, where Z is the set of zeroes of Jf. Further, one has to use the Rademacher theorem 2.21 in order to show that the Jacobian equals the (modulus of the) determinant of the derivative, almost everywhere. Indeed, the idea is to use the uniform convergence in the definition of Pansu derivative. We arrive to the change of variables theorem.

Theorem 2.28 *Let $f : A \subset M \to N$ be a bi-Lipschitz function from a measurable subset A of a Carnot group M to another Carnot group N. We have then:*

$$\int_A | \det Df(x) | \ dx \ = \ vol \ f(A) \tag{2.4.6}$$

A corollary of the area formula is the fact that bi-Lipschitz maps have Lusin property.

Corollary 2.29 *Let $f : A \subset M \to N$ be a bi-Lipschitz function from a measurable subset A of a Carnot group M to another Carnot group N. If $B \subset A$ is a set of 0 measure then $f(B)$ has 0 measure.*

2.5   Rigidity phenomenon

As an application to the previous sections we shall prove the following well known rigidity result.

Theorem 2.30 *Let X, Y be Lipschitz manifolds over the Carnot groups M, respectively N. If there is no subgroup of N isomorphic with M then there is no bi-Lipschitz embedding of X in Y.*

Proof.   Suppose that such an embedding exists. This implies (by choosing charts for X, Y) that there is a bi-Lipschitz embedding $f : A \subset M \to N$ with A open. Rademacher theorem 2.21 and corollary 2.29 imply that for almost any $x \in A \ Df(x)$ exists and it is an injective morphism from M to N. This contradicts the hypothesis. ∎

For example there is no bi-Lipschitz function from an open set in a Heisenberg group to an Euclidean space R^n. Otherwise stated, the Heisenberg group (or any other non-commutative Carnot group), is purely unrectifiable. This calls for an intrinsic definition of rectifiability. The subject will be treated in the second part of these notes, where we shall propose the following point of view: rectifiability is a theory of irreducible representations of the group of bi-Lipschitz homeomorphisms.

Let us give another example of rigidity, connected to Sussmann [31], section 8, example of abundance of abnormal geodesics on four-dimensional sub-Riemannian manifolds.

Consider a connected Lie group G with Lie algebra $\mathfrak{g}$ generated by X_1, X_2, X_3, X_4, such that the following conditions hold:

(i) $X_3 \ = \ [X_1, X_2], X_4 \ = \ [X_1, X_3]$,

(ii) $[X_2, X_3] \ = \ \alpha X_1 \ + \ \beta X_2 \ + \ \gamma X_3$ such that $\beta \neq 0$

Close examination shows that there is a three-dimensional family of Lie algebras satisfying these bracket conditions. Set $D \ = \ span \ \{X_1, X_2\}$ and endow G with the left invariant distribution generated by D and with the metric obtained by left translation of a metric on D such that X_1, X_2 are orthonormal. Sussmann proves that $t \mapsto \exp_G(tX_2)$

is an abnormal geodesic (see the second part of these notes or Sussmann [31], or Montgomery [24] and the references therein).

In this section we shall look to the nilpotentisation of the group G (section 4.1). The algebra $\mathfrak{g}$ admits the filtration

$$V^1 \; = \; span \; \{X_1, X_2\} \; \subset \; V^2 \; = \; span \; \{X_1, X_2, X_3\} \; \subset \; \mathfrak{g}$$

The nilpotentisation is $\mathfrak{g}$ endowed with the following bracket:

$$[X_1, X_2]_N \; = \; X_3 \; , \quad [X_1, X_3]_N \; = \; X_4$$

all the other brackets being zero. The dilatations in this Carnot algebra are

$$\delta_\varepsilon\left(\sum a_i X_i\right) \; = \; \varepsilon a_1 X_1 \; + \; \varepsilon a_2 X_2 \; + \; \varepsilon^2 a_3 X_3 \; + \; \varepsilon^3 a_4 X_4$$

Denote by N the Carnot group which has the algebra $(\mathfrak{g}, [\cdot, \cdot]_N)$. The group operation on N will be denoted by $\overset{\text{n}}{\cdot}$.

The linear group $HL(N)$ is the class of linear invertible maps which commute with the dilatations and preserve the nilpotent bracket. We give without proof the description of this group (for a similar proof see proposition 3.1).

Proposition 2.31 *In the basis* $\{X_1, ..., X_4\}$ *the linear group* $HL(N)$ *is made by all linear transformations represented by the matrices of the form*

$$\begin{pmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ 0 & 0 & a_{11}a_{22} & 0 \\ 0 & 0 & 0 & a_{11}^2 a_{22} \end{pmatrix}$$

for all $a_{11}, a_{22} \neq 0$.

Set $H \; = \; span \; \{X_1, X_3, X_4\}$. This is a group isomorphic with the Heisenberg group $H(1)$, with respect to the nilpotent bracket. Define now the family of horizontal curves

$$c_z(t) \; = \; z \overset{\text{n}}{\cdot} (t X_2)$$

for any $z \in H$. Any such curve is transversal to H.

Proposition 2.32 *Let* $\phi : \mathfrak{g} \to \mathfrak{g}$ *be a bi-Lipschitz map on the Carnot group* N. *Then for almost every* $z \in H$ *(with respect to the Lebesgue measure on* H*) there is* $z' \in H$ *such that*

$$\phi(c_z(R)) \; \subset \; c_{z'}(R)$$

Otherwise stated the family of curves c_z *is preserved by* N *bi-Lipschitz maps.*

Proof.   Let M be the family of all $z \in H$ such that ϕ is not a.e. differentiable in the points $c_z(t)$. Then M has to be negligible in H, otherwise we contradict Rademacher theorem 2.21. Take now $z \in H \setminus M$. Then for almost any $t \in R$ ϕ is differentiable in $c_z(t)$ and $\phi \circ c_z$ is differentiable in t. From the chain rule it follows that

$$\frac{d}{dt}\left(\phi \circ c_z\right)(t) \;=\; D\phi(c_z(t))\frac{d}{dt}c_z(t)$$

(all derivatives are Pansu). But

$$\frac{d}{dt}c_z(t) \;=\; X_2$$

and from proposition 2.31 it follows that

$$\frac{d}{dt}\left(\phi \circ c_z\right)(t) \;=\; \lambda(z,t)X_2$$

Indeed, $D\phi(x) \in HL(N)$ and the direction X_2 is preserved by any linear $F \in HL(N)$. Because the curve $\phi \circ c_z$ is almost everywhere tangent to X_2 it follows that it is a curve in the family c_z ■

Let us turn back to theorem 2.30 and remark that it can be improved. With the notation used before H is a subgroup of N and H is isomorphic with the Heisenberg group $H(1)$. Nevertheless the isomorphism does not commute with dilatations, or better, there is no isomorphism $f : H(1) \to N$ commuting with dilatations:

$$f \circ \delta_\varepsilon^{H(1)} \;=\; \delta_\varepsilon^N$$

Therefore there can be no bi-Lipschitz embedding of $H(1)$ in N. This is because linear maps are not only group morphisms, but also commute with dilatations.

Definition 2.33 *Let N, M be Carnot groups. We write $N \leq M$ if there is an injective group morphism $f : N \to M$ which commutes with dilatations.*

Theorem 2.30 improves like this:

Theorem 2.34 *Let X,Y be Lipschitz manifolds over the Carnot groups M, respectively N. If $N \not\leq M$ then there is no bi-Lipschitz embedding of X in Y.*

Rigidity in the sense of this section manifests in subtler ways. The purpose of Pansu paper [25] was to extend a result of Mostow [22], called Mostow rigidity. Although it is straightforward now to explain what Mostow rigidity means and how it can be proven, it is beyond the purposes of these notes.

3   The Heisenberg group

Let us rest a little bit and look closer to an example. The Heisenberg group is the most simple non commutative Carnot group. We shall apply the achievements of the previous chapter to this group.

3.1 The group

The Heisenberg group $H(n) = R^{2n+1}$ is a 2-step nilpotent group with the operation:

$$(x, \bar{x})(y, \bar{y}) = (x + y, \bar{x} + \bar{y} + \frac{1}{2}\omega(x, y))$$

where ω is the standard symplectic form on R^{2n}. We shall identify the Lie algebra with the Lie group. The bracket is

$$[(x, \bar{x}), (y, \bar{y})] = (0, \omega(x, y))$$

The Heisenberg algebra is generated by

$$V = R^{2n} \times \{0\}$$

and we have the relations $V + [V, V] = H(n)$, $\{0\} \times R = [V, V] = Z(H(n))$.
    The dilatations on $H(n)$ are

$$\delta_\varepsilon(x, \bar{x}) = (\varepsilon x, \varepsilon^2 \bar{x})$$

We shall denote by $HL(H(n))$ the group of invertible linear transformations and by $SL(H(n))$ the subgroup of volume preserving ones.

Proposition 3.1 *We have the isomorphisms*

$$HL(H(n)) \approx CSp(n) , \quad SL(H(n)) \approx Sp(n)$$

Proof. By direct computation. We are looking first for the algebra isomorphisms of $H(n)$. Let the matrix

$$\begin{pmatrix} A & b \\ c & a \end{pmatrix}$$

represent such a morphism, with $A \in gl(2n, R)$, $b, c \in R^{2n}$ and $a \in R$. The bracket preserving condition reads: for any $(x, \bar{x}), (y, \bar{y}) \in H(n)$ we have

$$(0, \omega(Ax + \bar{x}b, Ay + \bar{y}b)) = (\omega(x, y)b, a\omega(x, y))$$

We find therefore $b = 0$ and $\omega(Ax, Ay) = a\omega(x, y)$, so $A \in CSp(n)$ and $a \geq 0$, $a^n = \det A$.

    The preservation of the grading gives $c = 0$. The volume preserving condition means $a^{n+1} = 1$ hence $a = 1$ and $A \in Sp(n)$.   ∎

Get acquainted with Pansu differential Derivative of a curve: Let us see which are the smooth (i.e. derivable) curves. Consider $\tilde{c} : [0,1] \to H(n)$, $t \in (0,1)$ and $\varepsilon > 0$ sufficiently small. Then the finite difference function associated to c, t, ε is

$$C_\varepsilon(t)(z) = \delta_\varepsilon^{-1}\left(\varepsilon\tilde{c}(t)^{-1}\tilde{c}(t + \varepsilon z)\right)$$

After a short computation we obtain:

$$C_\varepsilon(t)(z) = \left(\frac{c(t + \varepsilon z) - c(t)}{\varepsilon}, \frac{\bar{c}(t + \varepsilon z) - \bar{c}(t)}{\varepsilon^2} - \frac{1}{2}\omega(c(t), \frac{c(t + \varepsilon z) - c(t)}{\varepsilon^2})\right)$$

When $\varepsilon \to 0$ we see that the finite difference function converges if:

$$\dot{\bar{c}}(t) = \frac{1}{2}\omega(c(t), \dot{c}(t))$$

Hence the curve has to be horizontal; in this case we see that

$$Dc(t)z = z(\dot{c}(t), 0)$$

This is almost the tangent to the curve. The tangent is obtained by taking $z = 1$ and the left translation of $Dc(t)1$ by $c(t)$.

The horizontality condition implies that, given a curve $t \mapsto c(t) \in R^{2n}$, there is only one horizontal curve $t \mapsto (c(t), \bar{c}(t))$, such that $\bar{c}(0) = 0$. This curve is called the lift of c.

Derivative of a functional: Take now $f : H(n) \to R$ and compute its Pansu derivative. The finite difference function is

$$F_\varepsilon(x, \bar{x})(y, \bar{y}) = \left(f(x + \varepsilon y, \bar{x} + \frac{\varepsilon}{2}\omega(x, y) + \varepsilon^2\bar{y}) - f(x, \bar{x})\right)/\varepsilon$$

Suppose that f is classically derivable. Then it is Pansu derivable and this derivative has the expression:

$$Df(x, \bar{x})(y, \bar{y}) = \frac{\partial f}{\partial x}(x, \bar{x})y + \frac{1}{2}\omega(x, y)\,\frac{\partial f}{\partial \bar{x}}(x, \bar{x})$$

Not any Pansu derivable functional is derivable in the classical sense. As an example, check that the square of any homogeneous norm is Pansu derivable everywhere, but not derivable everywhere in the classical sense.

3.2 Lifts of symplectic diffeomorphisms

In this section we are interested in the group of volume preserving diffeomorphisms of $H(n)$, with certain classical regularity. We establish connections between volume preserving diffeomorphisms of $H(n)$ and symplectomorphisms of R^{2n}.

Volume preserving diffeomorphisms

Definition 3.2 *$Diff^2(H(n), vol)$ is the group of volume preserving diffeomorphisms $\tilde{\phi}$ of $H(n)$ such that $\tilde{\phi}$ and it's inverse have (classical) regularity C^2. In the same way we define $Sympl^2(R^{2n})$ to be the group of C^2 symplectomorphisms of R^{2n}.*

Theorem 3.3 *We have the isomorphism of groups*

$$Diff^2(H(n), vol) \approx Sympl^2(R^{2n}) \times R$$

given by the mapping

$$\tilde{f} = (f, \bar{f}) \in Diff^2(H(n), vol) \mapsto \left(f \in Sympl^2(R^{2n}), \bar{f}(0,0) \right)$$

The inverse of this isomorphism has the expression

$$\left(f \in Sympl^2(R^{2n}), a \in R \right) \mapsto \tilde{f} = (f, \bar{f}) \in Diff^2(H(n), vol)$$

$$\tilde{f}(x, \bar{x}) = (f(x), \bar{x} + F(x))$$

where $F(0) = a$ and $dF = f^*\lambda - \lambda$.

Proof. Let $\tilde{f} = (f, \bar{f}) : H(n) \to H(n)$ be an element of the group $Diff^2(H(n), vol)$. We shall compute:

$$D\tilde{f}((x, \bar{x}))(y, \bar{y}) = \lim_{\varepsilon \to 0} \delta_{\varepsilon^{-1}} \left(\left(\tilde{f}(x, \bar{x}) \right)^{-1} \tilde{f} \left((x, \bar{x}) \delta_\varepsilon(y, \bar{y}) \right) \right)$$

We know that $D\tilde{f}(x, \bar{x})$ has to be a linear mapping.

After a short computation we see that we have to pass to the limit $\varepsilon \to 0$ in the following expressions (representing the two components of $D\tilde{f}((x, \bar{x}))(y, \bar{y})$):

$$\frac{1}{\varepsilon} \left(f \left(x + \varepsilon y, \bar{x} + \varepsilon^2 \bar{y} + \frac{\varepsilon}{2} \omega(x, y) \right) - f(x, \bar{x}) \right) \tag{3.2.1}$$

$$\frac{1}{\varepsilon^2} \left(\bar{f} \left(x + \varepsilon y, \bar{x} + \varepsilon^2 \bar{y} + \frac{\varepsilon}{2} \omega(x, y) \right) - \bar{f}(x, \bar{x}) - \tag{3.2.2}$$

$$- \frac{1}{2} \omega \left(f(x, \bar{x}), f \left(x + \varepsilon y, \bar{x} + \varepsilon^2 \bar{y} + \frac{\varepsilon}{2} \omega(x, y) \right) \right) \right)$$

The first component (3.2.1) tends to

$$\frac{\partial f}{\partial x}(x, \bar{x}) y + \frac{1}{2} \frac{\partial f}{\partial \bar{x}}(x, \bar{x}) \omega(x, y)$$

The terms of order ε must cancel in the second component (3.2.2). We obtain the cancellation condition (we shall omit from now on the argument $(x, \bar{x})$ from all functions):

$$\frac{1}{2} \omega(x, y) \frac{\partial \bar{f}}{\partial \bar{x}} - \frac{1}{2} \omega(f, \frac{\partial f}{\partial x} y) - \frac{1}{4} \omega(x, y) \omega(f, \frac{\partial f}{\partial \bar{x}}) + \frac{\partial \bar{f}}{\partial x} \cdot y = 0 \tag{3.2.3}$$

The second component tends to

$$\frac{\partial \bar{f}}{\partial \bar{x}} \bar{y} - \frac{1}{2} \omega(f, \frac{\partial f}{\partial \bar{x}}) \bar{y}$$

The group morphism $D\tilde{f}(x, \bar{x})$ is represented by the matrix:

$$d\tilde{f}(x, \bar{x}) \;=\; \begin{pmatrix} \frac{\partial f}{\partial x} + \frac{1}{2} \frac{\partial f}{\partial \bar{x}} \otimes Jx & 0 \\ 0 & \frac{\partial \bar{f}}{\partial \bar{x}} - \frac{1}{2} \omega(f, \frac{\partial f}{\partial \bar{x}}) \end{pmatrix} \tag{3.2.4}$$

We shall remember now that $\tilde{f}$ is volume preserving. According to proposition 3.1, this means:

$$\frac{\partial f}{\partial x} + \frac{1}{2} \frac{\partial f}{\partial \bar{x}} \otimes Jx \;\in Sp(n) \tag{3.2.5}$$

$$\frac{\partial \bar{f}}{\partial \bar{x}} - \frac{1}{2} \omega(f, \frac{\partial f}{\partial \bar{x}}) = 1 \tag{3.2.6}$$

The cancellation condition (3.2.3) and relation (3.2.6) give

$$\frac{\partial \bar{f}}{\partial x} y \;=\; \frac{1}{2} \omega(f, \frac{\partial f}{\partial x} y) \;-\; \frac{1}{2} \omega(x, y) \tag{3.2.7}$$

These conditions describe completely the class of volume preserving diffeomorphisms of $H(n)$. Conditions (3.2.6) and (3.2.7) are in fact differential equations for the function $\bar{f}$ when f is given. However, there is a compatibility condition in terms of f which has to be fulfilled for (3.2.7) to have a solution $\bar{f}$. Let us look closer to (3.2.7). We can see the symplectic form ω as a closed 2-form. Let λ be a 1-form such that $d\lambda = \omega$. If we take the (regular) differential with respect to x in (3.2.7) we quickly obtain the compatibility condition

$$\frac{\partial f}{\partial x} \;\in\; Sp(n) \tag{3.2.8}$$

and (3.2.7) takes the form:

$$d\bar{f} \;=\; f^* \lambda \;-\; \lambda \tag{3.2.9}$$

(all functions seen as functions of x only).

Conditions (3.2.8) and (3.2.5) imply: there is a scalar function $\mu = \mu(x, \bar{x})$ such that

$$\frac{\partial f}{\partial \bar{x}} \;=\; \mu \, Jx$$

Let us see what we have until now:

$$\frac{\partial f}{\partial x} \;\in\; Sp(n) \tag{3.2.10}$$

$$\frac{\partial \bar{f}}{\partial x} \;=\; \frac{1}{2} \left[\left(\frac{\partial f}{\partial x} \right)^T Jf \;-\; Jx \right] \tag{3.2.11}$$

$$\frac{\partial \bar{f}}{\partial \bar{x}} \;=\; 1 + \frac{1}{2}\omega(f, \frac{\partial f}{\partial \bar{x}}) \tag{3.2.12}$$

$$\frac{\partial f}{\partial \bar{x}} \;=\; \mu \; Jx \tag{3.2.13}$$

Now, differentiate (3.2.11) with respect to $\bar{x}$ and use (3.2.13). In the same time differentiate (3.2.12) with respect to x. From the equality

$$\frac{\partial^2 \bar{f}}{\partial x \partial \bar{x}} \;=\; \frac{\partial^2 \bar{f}}{\partial \bar{x} \partial x}$$

we shall obtain by straightforward computation $\mu = 0$.   ■

3.3   Hamilton's equations

For any element $\phi \in Sympl^2(R^{2n})$ we define the lift $\tilde{\phi}$ to be the image of $(\phi, 0)$ by the isomorphism described in the theorem 3.3.

Let $A \subset R^{2n}$ be a set. $Sympl^2(A)_c$ is the group of symplectomorphisms with regularity C^2, with compact support in A.

Definition 3.4 *For any flow $t \mapsto \phi_t \in Sympl^2(A)_c$ denote by $\phi^h(\cdot, x))$ the horizontal flow in $H(n)$ obtained by the lift of all curves $t \mapsto \phi(t, x)$ and by $\tilde{\phi}(\cdot, t)$ the flow obtained by the lift of all ϕ_t. The vertical flow is defined by the expression*

$$\phi^v \;=\; \tilde{\phi}^{-1} \circ \phi^h \tag{3.3.14}$$

Relation (3.3.14) can be seen as Hamilton equation.

Proposition 3.5 *Let $t \in [0,1] \mapsto \phi^v_t$ be a curve of diffeomorphisms of $H(n)$ satisfying the equation:*

$$\frac{d}{dt}\phi^v_t(x, \bar{x}) \;=\; (0, H(t, x)) \;\;,\;\; \phi^v_0 \;=\; id_{H(n)} \tag{3.3.15}$$

Then the flow $t \mapsto \phi_t$ which satisfies (3.3.14) and $\phi_0 = id_{R^{2n}}$ is the Hamiltonian flow generated by H.

Conversely, for any Hamiltonian flow $t \mapsto \phi_t$, generated by H, the vertical flow $t \mapsto \phi^v_t$ satisfies the equation (3.3.15).

Proof.   Write the lifts $\tilde{\phi}_t$ and ϕ^h_t, compute then the differential of the quantity $\dot{\tilde{\phi}}_t - \dot{\phi}^h_t$ and show that it equals the differential of H.   ■

Flows of volume preserving diffeomorphisms   We want to know if there is any nontrivial smooth (according to Pansu differentiability) flow of volume preserving diffeomorphisms.

Proposition 3.6 *Suppose that $t \mapsto \tilde{\phi}_t \in Diff^2(H(n), vol)$ is a flow such that*

- *is C^2 in the classical sense with respect to (x, t),*

- *is horizontal, that is $t \mapsto \tilde{\phi}_t(x)$ is a horizontal curve for any x.*

Then the flow is constant.

Proof.   By direct computation, involving second order derivatives. Indeed, let $\tilde{\phi}_t(x, \bar{x}) = (\phi_t(x), \bar{x} + F_t(x))$. From the condition $\tilde{\phi}_t \in Diff^2(H(n), vol)$ we obtain

$$\frac{\partial F_t}{\partial x} y \;=\; \frac{1}{2}\omega(\phi_t(x), \frac{\partial \phi_t}{\partial x}(x)y) \;-\; \frac{1}{2}\omega(x, y)$$

and from the hypothesis that $t \mapsto \tilde{\phi}_t(x)$ is a horizontal curve for any x we get

$$\frac{dF_t}{dt}(x) \;=\; \frac{1}{2}\omega(\phi_t(x), \dot{\phi}_t(x))$$

Equal now the derivative of the RHS of the first relation with respect to t with the derivative of the RHS of the second relation with respect to x. We get the equality, for any $y \in R^{2n}$:

$$0 \;=\; \frac{1}{2}\omega(\frac{\partial \phi_t}{\partial x}(x)y, \dot{\phi}_t(x))$$

therefore $\dot{\phi}_t(x) \;=\; 0$.   ■

   One should expect such a result to be true, based on two remarks. The first, general remark: take a flow of left translations in a Carnot group, that is a flow $t \mapsto \phi_t(x) \;=\; x_t x$. We can see directly that each ϕ_t is smooth, because the distribution is left invariant. But the flow is not horizontal, because the distribution is not right invariant. The second, particular remark: any flow which satisfies the hypothesis of proposition 3.6 corresponds to a Hamiltonian flow with null Hamiltonian function, hence the flow is constant.

   At a first glance it is disappointing to see that the group of volume preserving diffeomorphisms contains no smooth paths according to the intrinsic calculus on Carnot groups. But this makes the richness of such groups of homeomorphisms, as we shall see.

3.4   Volume preserving bi-Lipschitz maps

We shall work with the following groups of homeomorphisms.

Definition 3.7 *The group $Hom(H(n), vol, Lip)$ is formed by all locally bi-Lipschitz, volume preserving homeomorphisms of $H(n)$, which have the form:*

$$\tilde{\phi}(x, \bar{x}) \;=\; (\phi(x), \bar{x} + F(x))$$

   The group $Sympl(R^{2n}, Lip)$ of locally bi-Lipschitz symplectomorphisms of R^{2n}, in the sense that for a.e. $x \in R^{2n}$ the derivative $D\phi(x)$ (which exists by classical Rademacher theorem) is symplectic.

   Given $A \subset R^{2n}$, we denote by $Hom(H(n), vol, Lip)(A)$ the group of maps $\tilde{\phi}$ which belong to $Hom(H(n), vol, Lip)$ such that ϕ has compact support in A (i.e. it differs from identity on a compact set relative to A).

   The group $Sympl(R^{2n}, Lip)(A)$ is defined in an analogous way.

   Remark that any element $\tilde{\phi} \in Hom(H(n), vol, Lip)$ preserves the "vertical" left invariant distribution

$$(x\bar{x}) \;\mapsto\; (x, \bar{x})Z(H(n))$$

for a.e. $(x, \bar{x}) \in H(n)$.

   We shall prove that any locally bi-Lipschitz volume preserving homeomorphism of $H(n)$ belongs to $Hom(H(n), vol, Lip)$.

Theorem 3.8 *Take any $\tilde{\phi}$ locally bi-Lipschitz volume preserving homeomorphism of $H(n)$. Then $\tilde{\phi}$ has the form:*

$$\tilde{\phi}(x, \bar{x}) \;=\; (\phi(x), \bar{x} \;+\; F(x))$$

Moreover

$$\phi \in Sympl(R^{2n}, Lip)$$

$F : R^{2n} \to R$ is Lipschitz and for almost any point $(x, \bar{x}) \in H(n)$ we have:

$$DF(x)y \;=\; \frac{1}{2}\omega(\phi(x), D\phi(x)y) \;-\; \frac{1}{2}\omega(x, y)$$

Proof. Set $\tilde{\phi}(\tilde{x}) \;=\; (\phi(\tilde{x}), \bar{\phi}(\tilde{x}))$. Also denote by

$$|\,(y, \bar{y})\,|^2 = |\,y\,|^2 \;+\; |\,\bar{y}\,|$$

   By theorem 2.21 $\tilde{\phi}$ is almost everywhere derivable and the derivative can be written in the particular form:

$$\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}$$

with $A \;=\; A(x, \bar{x}) \;\in Sp(n, R)$.

   For a.e. $x \in \mathbb{R}^{2n}$ the function $\tilde{\phi}$ is derivable. The derivative has the form:

$$D\tilde{\phi}(x, \bar{x})(y, \bar{y}) \;=\; (Ay, \bar{y})$$

where by definition

$$D\tilde{\phi}(x, \bar{x})(y, \bar{y}) \;=\; \lim_{\varepsilon \to 0} \delta_\varepsilon^{-1} \left(\tilde{\phi}(x, \bar{x})^{-1} \; \phi((x, \bar{x})\delta_\varepsilon(y, \bar{y})) \right)$$

Let us write down what Pansu derivability means.

For the $\mathbb{R}^{2n}$ component we have: the limit

$$\lim_{\varepsilon \to 0} \mid \frac{1}{\varepsilon}[\phi((x,\bar{x})\delta_\varepsilon(y,\bar{y})) - \phi(x,\bar{x})] - A \mid = 0 \qquad (3.4.16)$$

is uniform with respect to $\tilde{y} \in B(0,1)$. We shall rescale all for the ball $B(0,R) \in H(n)$. Relation (3.4.16) means: for any $\lambda > 0$ there exists $\varepsilon_0(\lambda) > 0$ such that for any $R > 0$ and any $\varepsilon > \varepsilon_0/R$ and $\tilde{y} \in H(n)$ such that $\mid \tilde{y} \mid < R$ we have

$$\mid \frac{1}{\varepsilon}\left(\phi(\tilde{x}\delta_\varepsilon\tilde{y}) - \phi(\tilde{x})\right) - Ay \mid < \lambda R \qquad (3.4.17)$$

Suppose that $\tilde{x} = (x,\bar{x})$ satisfies $\mid x \mid \leq R$ with $R > 1$. Fix now $\lambda > 0$ and consider arbitrary $y \in R^{2n}$, $\mid y \mid < 1$ and $\mu > 0$. Set

$$\bar{y}_\mu = \frac{1}{2\mu}\omega(y,x)$$

There exists a constant $C > 0$ such that

$$\mid (y,\bar{y}_\mu) \mid^2 \leq 1 + \frac{C}{\mu}(1+R)^2$$

For $\varepsilon' < \varepsilon_0(\lambda)$ choose

$$\varepsilon = \mu = \frac{1}{4}\left(\sqrt{C^2(1+R)^4 + 4\varepsilon'}\ C(1+R)^2\right)$$

Then the inequality (3.4.17) is true for R replaced by $\left(1 + \frac{C}{\mu}(1+R)^2\right)^{1/2}$, $\tilde{y} = (y,\bar{y}_\mu)$, with arbitrary $\mid y \mid < 1$. We get: for any $\lambda > 0$ there is $\varepsilon(\lambda) = \varepsilon_0(\lambda)/2 > 0$ such that for any $\varepsilon < \varepsilon(\lambda)$ and any $\mid y \mid < 1$ we have

$$\mid \frac{1}{\varepsilon}\left(\phi(x+\varepsilon y,\bar{x}) - \phi(x,\bar{x})\right) - Ay \mid < 2\lambda$$

Therefore $x \mapsto \phi(x,\bar{x})$ is derivable a.e. with (locally) bounded derivative. It is therefore (locally) Lipschitz.

Let us now return to (3.4.16) and choose $\bar{y} = 0$. We get: the map

$$y \mapsto \phi(x+y,\bar{x}+\frac{1}{2}\omega(x,y))$$

is derivable in $y = 0$ and the derivative equals A. But we have just proven that the derivative of $\phi(\cdot,\bar{x})$ in x equals A. We get: ϕ is constant with respect to $\bar{x}$.

We have proven until now that

$$\tilde{\phi}(x,\bar{x}) = (\phi(x),\bar{\phi}(x,\bar{x}))$$

with $\phi \in Sympl(R^{2n}, Lip)$.

Let us use this information in the definition of the center component of the Pansu derivative. We have:

$$\mid \frac{1}{\varepsilon^2}\left(\bar{\phi}(\tilde{x}\delta_\varepsilon\tilde{y}) - \bar{\phi}(\tilde{x}) - \frac{1}{2}\omega(\phi(x),\phi(x+\varepsilon y))\right) - \bar{y} - \qquad (3.4.18)$$

$$-\frac{1}{2}\omega(\frac{1}{\varepsilon}(\phi(x+\varepsilon y)-\phi(x)), Ay)\mid \to 0$$

as $\varepsilon \to 0$, uniformly with respect to $\mid \tilde{y} \mid < 1$. From previously proved facts we obtain that (3.4.18) is equivalent to

$$\mid \frac{1}{\varepsilon^2}\left(\bar{\phi}(\tilde{x}\delta_\varepsilon \tilde{y})-\bar{\phi}(\tilde{x})-\frac{1}{2}\omega(\phi(x),\phi(x+\varepsilon y)-\phi(x))\right)-\bar{y}\mid \to 0$$

as $\varepsilon \to 0$, uniformly with respect to $\tilde{y}$. We quickly get that $\bar{\phi}(x,\bar{x}) = \bar{x} + F(x)$.
   All in all the Pansu derivability (center component) reads

$$\lim_{\varepsilon \to 0}\mid \frac{1}{\varepsilon^2}[F(x+\varepsilon y)-F(x)]+\frac{1}{2\varepsilon}\omega(x,y)-\frac{1}{2}\omega(\phi(x),\frac{1}{\varepsilon^2}[\phi(x+\varepsilon y)-\phi(x)])\mid = 0 \quad (3.4.19)$$

This means that for almost any $x \in \mathbb{R}^{2n}$

1 the function ϕ is derivable and the derivative is equal to $A = A(x)$,

2 the function F is derivable and is connected to ϕ by the relation from the conclusion of the proposition.

   We have seen that $\tilde{\phi}$ is locally Lipschitz implies that $x \in \mathbb{R}^{2n} \mapsto A(x)$ is locally bounded, therefore ϕ is locally Lipschitz. In the same way we obtain that F is locally Lipschitz.   ∎
   The flows in the group $Hom(H(n),vol,Lip)$ are defined further.

Definition 3.9 *A flow in the group $Hom(H(n),vol,Lip)$ is a curve*

$$t \mapsto \tilde{\phi}_t \in Hom(H(n),vol,Lip)$$

such that for a.e. $x \in R^{2n}$ the curve $t \mapsto \phi_t(x) \in R^{2n}$ is (locally) Lipschitz.
   For any flow we can define the horizontal lift of this flow like this: a.e. curve $t \mapsto \phi_t(x)$ lifts to a horizontal curve $t \mapsto \phi_t^h(x,0)$. Define then

$$\phi_t^h(x,\bar{x}) = \phi_t^h(x,0)(0,\bar{x})$$

   If the flow is smooth (in the classical sense) and $\tilde{\phi}_t \in Diff^2(H(n),vol)$ then this lift is the same as the one described in definition 3.4. We can define now the vertical flow by the formula (3.3.14), that is

$$\phi^v = \tilde{\phi}^{-1}\circ \phi^h$$

   There is an analog of proposition 3.6. In the proof we shall need lemma 3.11, which comes after.

Proposition 3.10 *Let $t \mapsto \tilde{\phi}_t \in Hom(H(n),vol,Lip)$ be a curve such that the function $t \mapsto \Phi(\tilde{x},t) = (\tilde{\phi}_t(\tilde{x}),t)$ is locally Lipschitz from $H(n) \times R$ to itself. Then $t \mapsto \tilde{\phi}_t$ is a constant curve.*

Proof.   By Rademacher theorem 2.15 for the group $H(n) \times R$ we obtain that Φ is almost everywhere derivable. Use now lemma 3.11 to deduce the claim.   ■

A short preparation is needed in order to state the lemma 3.11. Let N be a non-commutative Carnot group. We shall look at the group $N \times R$ with the group operation defined component wise. This is also a Carnot group. Indeed, consider the family of dilatations

$$\delta_\varepsilon(x,t) \;=\; (\delta_\varepsilon(x), \varepsilon t)$$

which gives to $N \times R$ the structure of a CC group. The left invariant distribution on the group which generates the distance is (the left translation of) $W_1 = V_1 \times R$.

Lemma 3.11 *Let N be a noncommutative Carnot group which admits the orthogonal decomposition*

$$N \;=\; V_1 + [N,N]$$

and satisfies the condition

$$V_1 \cap Z(N) \;=\; 0$$

The group of linear transformations of $N \times R$ is then

$$HL(N \times R) \;=\; \left\{ \begin{pmatrix} A & 0 \\ c & d \end{pmatrix} \; : A \in HL(N) \;,\; c \in V_1 \;,\; d \in R \right\}$$

Proof.   We shall proceed as in the proof of proposition 3.1. We are looking first at the Lie algebra isomorphisms of $N \times R$, with general form

$$\begin{pmatrix} A & b \\ c & d \end{pmatrix}$$

We obtain the conditions:

(i) c orthogonal on $[N,N]$,

(ii) b commutes with the image of A: $[b, Ay] = 0$, for any $y \in N$,

(iii) A is an algebra isomorphism of N.

From (ii), (iii) we deduce that b is in the center of N and from (i) we see that $c \in V_1$.

We want now the isomorphism to commute with dilatations. This condition gives:

(iv) $b \in V_1$,

(v) A commutes with the dilatations of N.

(iii) and (v) imply that $A \in HL(N)$ and (iv) that $b = 0$.   ■

Hamiltonian diffeomorphisms: more structure   In this section we look closer to the structure of the group of volume preserving homeomorphisms of the Heisenberg group.

Let $Hom^h(H(n), vol, Lip)(A)$ be the group of time one homeomorphisms $t \mapsto \phi^h$, for all curves $t \mapsto \phi_t \in Sympl(R^{2n}, Lip)(A)$ such that $(x, t) \mapsto (\phi_t(x), t)$ is locally Lipschitz.

The elements of this group are also volume preserving, but they are not smooth with respect to the Pansu derivative.

Definition 3.12 *The group $Hom(H(n), vol)(A)$ contains all maps $\tilde{\phi}$ which have a.e. the form:*

$$\tilde{\phi}(x, \bar{x}) \;=\; (\phi(x), \bar{x} + F(x))$$

where $\phi \in Sympl(R^{2n}, Lip)(A)$ and $F : R^{2n} \to R$ is locally Lipschitz and constant outside a compact set included in the closure of A (for short: with compact support in A).

This group contains three privileged subgroups:

- $Hom(H(n), vol, Lip)(A)$,

- $Hom^h(H(n), vol, Lip)(A)$

- and $Hom^v(H(n), vol, Lip)(A)$.

The last is the group of vertical homeomorphisms, any of which has the form:

$$\phi^v(x, \bar{x}) \;=\; (x, \bar{x} + F(x))$$

with F locally Lipschitz, with compact support in A.

Take a one parameter subgroup $t \mapsto \tilde{\phi}_t \in Hom(H(n), vol, Lip)(A)$, in the sense of the definition 3.9. We know that it cannot be smooth as a curve in in the group $Hom(H(n), vol, Lip)(A)$, but we also know that there are vertical and horizontal flows $t \mapsto \phi^v_t$, ϕ^h_t such that we have the decomposition $\tilde{\phi}_t \circ \phi^v_t \;=\; \phi^h_t$. Unfortunately none of the flows $t \mapsto \phi^v_t$, ϕ^h_t are one parameter groups.

There is more structure here that it seems. Consider the class

$$HAM(H(n))(A) = Hom(H(n), vol, Lip)(A) \times Hom^v(H(n), vol, Lip)(A)$$

For any pair in this class we shall use the notation $(\tilde{\phi}, \phi^v)$ This class forms a group with the (semidirect product) operation:

$$(\tilde{\phi}, \phi^v)(\tilde{\psi}, \psi^v) \;=\; (\tilde{\phi} \circ \tilde{\psi}, \phi^v \circ \tilde{\phi} \circ \psi^v \circ \tilde{\phi}^{-1})$$

Proposition 3.13 *If $t \mapsto \tilde{\phi}_t \in Hom(H(n), vol, Lip)(A)$ is an one parameter group then $t \mapsto (\tilde{\phi}_t, \phi^v_t) \in HAM(H(n))$ is an one parameter group.*

Proof.   The check is left to the reader. Use definition and proposition 1, chapter 5, Hofer & Zehnder [16], page 144.   ∎

We can (indirectly) put a distribution on the group $HAM(H(n))(A)$ by specifying the class of horizontal curves. Such a curve $t \mapsto (\tilde{\phi}_t, \phi_t^v)$ in the group $HAM(H(n))(A)$ projects in the first component $t \mapsto \tilde{\phi}_t$ to a flow in $Hom(H(n), vol, Lip)(A)$. Define for any $(\tilde{\phi}, \phi^v) \in HAM(H(n))(A)$ the function

$$\phi^h \;=\; \tilde{\phi} \circ \phi^v$$

and say that the curve

$$t \mapsto (\tilde{\phi}_t, \phi_t^v) \in HAM(H(n))(A)$$

is horizontal if

$$t \mapsto \phi_t^h(x, \bar{x})$$

is horizontal for any $(x, \bar{x}) \in H(n)$.

We introduce the following length function for horizontal curves:

$$L\left(t \mapsto (\tilde{\phi}_t, \phi_t^v)\right) \;=\; \int_0^1 \|\dot{\phi}_t^v\|_{L^\infty(A)} \; dt$$

With the help of the length we are able to endow the group $HAM(H(n))(A)$ as a path metric space. The distance is defined by:

$$dist\left((\tilde{\phi}_I, \phi_I^v), (\tilde{\phi}_{II}, \phi_{II}^v)\right) \;=\; \inf L\left(t \mapsto (\tilde{\phi}_t, \phi_t^v)\right)$$

over all horizontal curves $t \mapsto (\tilde{\phi}_t, \phi_t^v)$ such that

$$(\tilde{\phi}_0, \phi_0^v) \;=\; (\tilde{\phi}_I, \phi_I^v)$$

$$(\tilde{\phi}_1, \phi_1^v) \;=\; (\tilde{\phi}_{II}, \phi_{II}^v)$$

The distance is not defined for any two points in $HAM(H(n))(A)$. In principle the distance can be degenerated.

The group $HAM(H(n))(A)$ acts on AZ (where Z is the center of $H(n)$) by:

$$(\tilde{\phi}, \phi^v)(x, \bar{x}) \;=\; \phi^v \circ \tilde{\phi}(x, \bar{x})$$

Remark that the group

$$Hom^v(H(n), vol, Lip)(A) \;\equiv\; \{id\} \times Hom^v(H(n), vol, Lip)(A)$$

is normal in $HAM(H(n))(A)$. We denote by

$$HAM(H(n))(A)/Hom^v(H(n), vol, Lip)(A)$$

the factor group. This is isomorphic with the group $Sympl(R^{2n}, Lip)(A)$.

Consider now the natural action of $Hom^v(H(n), vol, Lip)(A)$ on AZ. The space of orbits $AZ/Hom^v(H(n), vol, Lip)(A)$ is nothing but A, with the Euclidean metric.

Factorize now the action of $HAM(H(n))(A)$ on AZ, by the group

$$Hom^v(H(n), vol, Lip)(A)$$

We obtain the action of $HAM(H(n))(A)/Hom^v(H(n), vol, Lip)(A)$ on

$$AZ/Hom^v(H(n), vol, Lip)(A)$$

Elementary examination shows this is the action of the symplectomorphisms group on R^{2n}.

All in all we have the following (for the definition of the Hofer distance see next section):

Theorem 3.14 *The action of $HAM(H(n))(A)$ on AZ descends after reduction with the group $Hom^v(H(n), vol, Lip)$ to the action of symplectomorphisms group with compact support in A.*

The distance dist on $HAM(H(n))$ descends to the Hofer distance on the connected component of identity of the symplectomorphisms group.

3.5 Symplectomorphisms, capacities and Hofer distance

Symplectic capacities are invariants under the action of the symplectomorphisms group. Hofer geometry is the geometry of the group of Hamiltonian diffeomorphisms, with respect to the Hofer distance. For an introduction into the subject see Hofer, Zehnder [16] chapters 2,3 and 5, and Polterovich [28], chapters 1,2.

A symplectic capacity is a map which associates to any symplectic manifold (M, ω) a number $c(M, \omega) \in [0, +\infty]$. Symplectic capacities are special cases of conformal symplectic invariants, described by:

A1. Monotonicity: $c(M, \omega) \leq c(N, \tau)$ if there is a symplectic embedding from M to N,

A2. Conformality: $c(M, \varepsilon\omega) =| \varepsilon | c(M, \omega)$ for any $\alpha \in R$, $\alpha \neq 0$.

We can see a conformal symplectic invariant from another point of view. Take a symplectic manifold (M, ω) and consider the invariant defined over the class of Borel sets $B(M)$, (seen as embedded submanifolds). In the particular case of R^{2n} with the standard symplectic form, an invariant is a function $c : B(R^{2n}) \to [0, +\infty]$ such that:

B1. Monotonicity: $c(M) \leq c(N)$ if there is a symplectomorphism ϕ such that $\phi(M) \subset N$,

B2. Conformality: $c(\varepsilon M) = \varepsilon^2 c(M)$ for any $\varepsilon \in R$.

An invariant is nontrivial if it takes finite values on sets with infinite volume, like cylinders:

$$Z(R) = \left\{ x \in R^{2n} : x_1^2 + x_2^2 < R \right\}$$

There exist highly nontrivial invariants, as the following theorem shows:

Theorem 3.15 *(Gromov's squeezing theorem) The ball $B(r)$ can be symplectically embedded in the cylinder $Z(R)$ if and only if $r \leq R$.*

This theorem permits to define the invariant:

$$c(A) \;=\; \sup \left\{ R^2 : \exists \phi(B(R)) \subset A \right\}$$

called Gromov's capacity.

Another important invariant is Hofer-Zehnder capacity. In order to introduce this we need the notion of a Hamiltonian flow.

A flow of symplectomorphisms $t \mapsto \phi_t$ is Hamiltonian if there is a function $H : M \times R \to R$ such that for any time t and place x we have

$$\omega(\dot{\phi}_t(x), v) \;=\; dH(\phi_t(x), t)v$$

for any $v \in T_{\phi_t(x)} M$.

Let $H(R^{2n})$ be the set of compactly supported Hamiltonians. Given a set $A \subset R^{2n}$, the class of admissible Hamiltonians is $H(A)$, made by all compactly supported maps in A such that the generated Hamiltonian flow does not have closed orbits of periods smaller than 1. Then the Hofer-Zehnder capacity is defined by:

$$hz(A) \;=\; \sup \left\{ \|H\|_\infty : H \in H(A) \right\}$$

Let us denote by $Ham(A)$ the class of Hamiltonian diffeomorphisms compactly supported in A. A Hamiltonian diffeomorphism is the time one value of a Hamiltonian flow. In the case which interest us, that is R^{2n}, $Ham(A)$ is the connected component of the identity in the group of compactly supported symplectomorphisms.

A curve of Hamiltonian diffeomorphisms (with compact support) is a Hamiltonian flow. For any such curve $t \mapsto c(t)$ we shall denote by $t \mapsto H_c(t, \cdot)$ the associated Hamiltonian function (with compact support).

On the group of Hamiltonian diffeomorphisms there is a bi-invariant distance introduced by Hofer. This is given by the expression:

$$d_H(\phi, \psi) \;=\; \inf \left\{ \int_0^1 \|H_c(t)\|_{\infty, R^{2n}} \; dt : c : [0,1] \to Ham(R^{2n}) \right\} \qquad (3.5.20)$$

It is easy to check that d is indeed bi-invariant and it satisfies the triangle property. It is a deep result that d is non-degenerate, that is $d(id, \phi) \;=\; 0$ implies $\phi = id$.

With the help of the Hofer distance one can define another symplectic invariant, called displacement energy. For a set $A \subset R^{2n}$ the displacement energy is:

$$de(A) \;=\; \inf \left\{ d_H(id, \phi) \; : \; \phi \in Ham(R^{2n}) \; , \; \phi(A) \cap A = \emptyset \right\}$$

A displacement energy can be associated to any group of transformations endowed with a bi-invariant distance (see Eliashberg & Polterovich [7], section 1.3). The fact that the displacement energy is a nontrivial invariant is equivalent with the non-degeneracy of the Hofer distance. Section 2. Hofer & Zehnder [16] is dedicated to the implications of the existence of a non-trivial capacity. All in all the non-degeneracy of the Hofer distance (proved for example in the Section 5. Hofer & Zehnder [16]) is the cornerstone of symplectic rigidity theory.

3.6 Hausdorff dimension and Hofer distance

We shall give in this section a metric proof of the non-degeneracy of the Hofer distance.

A flow of symplectomorphisms $t \mapsto \phi_t$ with compact support in A is Hamiltonian if it is the projection onto R^{2n} of a flow in $Hom(H(n), vol, Lip)(A)$.

Consider such a flow which joins identity id with ϕ. Take the lift of the flow $t \in [0,1] \mapsto \tilde{\phi}_t \in Hom(H(n), vol, Lip)(A)$.

Proposition 3.16 *The curve $t \mapsto \tilde{\phi}_t(x, 0)$ has Hausdorff dimension 2 and measure*

$$\mathcal{H}^2\left(t \mapsto \tilde{\phi}_t(x, \bar{x})\right) = \int_0^1 |\, H_t(\phi_t(x)) \,| \;\; dt \tag{3.6.21}$$

Proof. The curve is not horizontal. The tangent has a vertical part (equal to the Hamiltonian). Use the definition of the (spherical) Hausdorff measure $\mathcal{H}^2$ and the Ball-Box theorem 2.9 to obtain the formula (3.6.21). ∎

We shall prove now that the Hofer distance (3.5.20) is non-degenerate. For given ϕ with compact support in A and generating function F, look at the one parameter family:

$$\tilde{\phi}^a(x, \bar{x}) = (\phi(x), \bar{x} + F(x) + a)$$

The volume of the cylinder

$$\{(x, z) : x \in A, \; z \text{ between } 0 \text{ and } F(x) + a\}$$

attains the minimum $V(\phi, A)$ for an $a_0 \in R$.

Given an arbitrary flow $t \in [0,1] \mapsto \tilde{\phi}_t \in Hom(H(n), vol, Lip)(A)$ such that $\tilde{\phi}_0 = id$ and $\tilde{\phi}_1 = \tilde{\phi}^a$, we have the following inequality:

$$V(\phi, A) \leq \int_A \mathcal{H}^2\left(\left(t \mapsto \tilde{\phi}_t(x, 0)\right)\right) \; dx \tag{3.6.22}$$

Indeed, the family of curves $t \mapsto \tilde{\phi}_t(x, 0)$ provides a foliation of the set

$$\left\{\tilde{\phi}_t(x, 0) : x \in A\right\}$$

The volume of this set is lesser than the RHS of inequality (3.6.22):

$$C \, Vol\left(\left\{\tilde{\phi}_t(x, 0) : x \in A\right\}\right) \leq \int_A \mathcal{H}^2\left(t \mapsto \tilde{\phi}_t(x, 0)\right) \; dx \tag{3.6.23}$$

where $C > 0$ is an universal positive constant (that is it depends only on the algebraic structure of $H(n)$). This is a statement which is local in nature; it is simply saying that the area of the parallelogram defined by the vectors $\mathbf{a}$, $\mathbf{b}$ is smaller than the product of the norms $\|\mathbf{a}\| \|\mathbf{b}\|$. We give further a proof of this inequality.

Let $\tilde{A} \subset H(n)$ be a set with finite Hausdorff measure $Q = 2n + 2$ which contains a set $\tilde{B} = B \times \{0\}$ such that there is a flow

$$t \in [0,1] \mapsto \phi_t : H(n) \to H(n)$$

with the properties:

(a) $\phi_0(x) = x$ for any $x \in B$,

(b) for almost any t ϕ_t preserves the Hausdorff measure Q,

(c) for any $t \in [0,1]$ ϕ_t is bi-Lipschitz and the Lipschitz constants of ϕ_t, ϕ_t^{-1} are upper bounded by a positive constant $Const$,

(d) $\tilde{A} = \{\phi_t(x) : t \in [0,1] , x \in B\}$,

(e) for $\mathcal{H}^{2n}$ almost any $x \in B$ the fiber $f_x = \{\phi_t(x) : t \in [0,1]\}$ has finite Hausdorff dimension $P = 2$.

Then

$$C \, \mathcal{H}^Q(\tilde{A}) \leq \int_B \mathcal{H}^2(f_x) \, \mathrm{d}(x)$$

Indeed, we start from the isodiametric inequality in the Heisenberg group:

$$\left(diam \; \tilde{D}\right)^Q \geq c \, \mathcal{H}^Q(\tilde{D})$$

We can rescale the Hausdorff measure Q in order to have $c = 1$. Consider also the projection $\pi : H(n) \to R^{2n}$, $\pi(x, \bar{x}) = x$. We can rescale the Lebesgue measure on R^{2n} such that the volume of the projection of the ball $B^{CC}(\tilde{x}, 1)$ equals 1. Here we have used the notation: $B^{CC}(\tilde{x}, R)$ is the ball in the Carnot-Carathéodory distance in $H(n)$, and $B^E(x, R)$ is the Euclidean ball of radius R in R^{2n}.

Fix ε 0 small and cover B with Euclidean balls of radius ε $B^E(x_j, \varepsilon)$, $j \in J_\varepsilon$, in an efficient way, i.e. such that $J_\varepsilon \varepsilon^{2n}$ tends to $vol(B)$ when $\varepsilon \to 0$. We want now to cover $\tilde{A}$ with sets of the form

$$\tilde{\phi}_{t_k}(B^{CC}((x_j, 0), \varepsilon))$$

where k depends on x_j. How many sets do we need?

From the isodiametric inequality and the fact that $\tilde{\phi}_t$ is volume preserving, we see that

$$diam \; \tilde{\phi}_t(B^{CC}((x_j, 0), \varepsilon)) \geq \varepsilon$$

Put this together with the fact that the fiber f_{x_j} has Hausdorff dimension $P = 2$ and get that we need about $\mathcal{H}^P(f_{x_j})/\varepsilon^P$ such sets to cover the fiber f_{x_j}. The sum of volumes of the cover is greater then the volume of A. Tend ε to 0 and obtain the claimed inequality (3.6.23). The constant C comes from the rescalings of the measures. Without rescalling we have:

$$C = \frac{vol(B^{CC}(0, 1))}{vol(B^E(0, 1))}$$

We go back now to the main stream of the proof and remark that by continuity of the flow the volume of the set

$$\left\{ \tilde{\phi}_t(x, 0) : x \in A \right\}$$

is greater than $V(\phi, A)$.

For any curve of the flow we have the uniform obvious estimate:

$$\mathcal{H}^2\left(\left(t \mapsto \tilde{\phi}_t(x,0)\right) \leq \int_0^1 \|H_t(\phi_t(x))\|_{A,\infty} \, dt \tag{3.6.24}$$

Put together inequalities (3.6.22), (3.6.24) and get

$$C \ V(\phi, A) \leq \int_0^1 \|H_t(\phi_t(x))\|_{A,\infty} \, dt \tag{3.6.25}$$

Use the definition of the Hofer distance (3.5.20) to obtain the inequality

$$C \ V(\phi, A) \leq \ vol(A) \ d_H(id, \phi) \tag{3.6.26}$$

This proves the non-degeneracy of the Hofer distance, because if the RHS of (3.6.26) equals 0 then $V(\phi, A)$ is 0, which means that the generating function of ϕ is almost everywhere constant, therefore ϕ is the identity everywhere in R^{2n}.

We close with the translation of the inequality (3.6.25) in symplectic terms.

Proposition 3.17 *Let ϕ be a Hamiltonian diffeomorphism with compact support in A and F its generating function, that is $dF = \phi^*\lambda - \lambda$, where $d\lambda = \omega$. Consider a Hamiltonian flow $t \mapsto H_t$, with compact support in A, such that the time one map equals ϕ. Then the following inequality holds:*

$$C \ \inf\left\{\int_A | \ F(x) - c \ | \ dx : c \in R\right\} \leq \ vol(A) \int_0^1 \|H_t\|_{\infty,A} \, dt$$

with $C > 0$ an universal constant equal to

$$C = \frac{vol(B^{CC}(0,1))}{vol(B^E(0,1)}$$

3.7   Invariants of volume preserving maps

Theorem 3.8 gives the strucure of a volume preserving locally bi-Lipschitz homeomorphism. We have denoted by $Hom(H(n), vol, Lip)$ the class of these maps. Any $\tilde{\phi} \in Hom(H(n), vol, Lip)$ can be written as

$$\tilde{\phi}(x, \bar{x}) = (\phi(x), \bar{x} + F(x))$$

where $\phi \in Sympl(R^{2n}, Lip)$ is a locally bi-Lipschitz symplectomorphism of R^{2n}. For a set $B \subset R^{2n}$ we denote by $vol(B)$ it's Lebesgue measure. For a set $C \subset R \ l(C)$ will be it's length (one dimensional Lebesgue measure). An immediate consequence of the structure theorem 3.8 is given further.

Corollary 3.18 *Let $\tilde{A} \subset H(n)$ be a $\mathcal{H}^{2n+2}$ measurable set such that $\mathcal{H}^{2n}(A) < +\infty$, where A denotes the orthogonal projection of $\tilde{A}$ on R^{2n}. Then for any $\tilde{\phi} \in Hom(H(n), vol, Lip)$ we have:*

$$w(\tilde{A}) = \frac{\mathcal{H}^{2n+2}(\tilde{A})}{vol(A)} = \frac{\mathcal{H}^{2n+2}(\tilde{\phi}(\tilde{A}))}{vol(\phi(A))}$$

We call $w(\tilde{A})$ the width of $\tilde{A}$.

More general, for any $i \geq 1$ define the i-height of $\tilde{A}$ to be

$$\left(h_i(\tilde{A})\right)^i \;=\; \frac{1}{vol(A)^i} \int_A \left(l(\{\bar{x} \in R : (x, \bar{x}) \in \tilde{A}\})\right)^{1/i} \; dx$$

Then $h_i(\tilde{A}) \;=\; h_i(\tilde{\phi}(\tilde{A}))$.

Proof. Any symplectomorphism is volume preserving. The thesis follows from this and the structure theorem 3.8. ■

Suppose $\tilde{A}$ is open, bounded and classically smooth. In the case of the Heisenberg group $H(1)$ Pansu [27] proved the following isoperimetric inequality:

$$\mathcal{H}^4(\tilde{A}) \;\leq\; C\left(\mathcal{H}^3(\partial\tilde{A})\right)^{\frac{4}{3}}$$

The isoperimetric inequality is a very important subject in analysis in metric spaces. It has been intensively studied in it's disguised form (Sobolev inequalities) in many papers. To have a grasp of what is hidden here the reader can consult Varopoulos & al. [32], Hajlasz, Koskela [14] and Garofalo, Nhieu [10]. In order to give a decent exposition of this subject we would need a good notion of perimeter. In this first part we don't enter much into the measure theoretic aspects though.

Let us work with objects which have classical regularity, as in Gromov [12] section 2.3. There we find a more general isoperimetric inequality, true in particular in any Carnot group, provided that $\tilde{A}$ is open, bounded, with (classically) smooth boundary. If we denote by Q the homogeneous dimension of the Carnot group N (which equals its Hausdorff dimension) then Gromov's isoperimetric inequality is:

$$\mathcal{H}^Q(\tilde{A}) \;\leq\; C\left(\mathcal{H}^{Q-1}(\partial\tilde{A})\right)^{\frac{Q}{Q-1}}$$

Recall that for Heisenberg group $H(n)$ the homogeneous dimension is $Q = 2n + 2$.

We should use further the group $Diff^2(H(n), vol)$ instead of $Hom(H(n), vol, Lip)$, because we work with classical regularity . We kindly ask the reader to believe that a good definition of the perimeter of the set $\tilde{A}$ allows us to use the group $Hom(H(n), vol, Lip)$. We shall denote the perimeter function by Per. In "smooth" situations $Per(\partial\tilde{A}) \;=\; \mathcal{H}^{Q-1}(\partial\tilde{A})$.

We can define then the isoperimeter of $\tilde{A}$ to be:

$$isop(\tilde{A})^{2n+1} \;=\; \inf \frac{\left(Per\; \partial\tilde{\phi}(\tilde{A})\right)^{2n+2}}{\left(\mathcal{H}^{2n}(A)\right)^{2n+1}}$$

with respect to all $\tilde{\phi} \in Hom(H(n), vol, Lip)$. This is another invariant of the group of volume preserving locally bi-Lipschitz maps.

Yet another invariant is the isodiameter. This is defined as

$$isod(\tilde{A}) \;=\; \inf \frac{(diam\; \tilde{\phi}(\tilde{A}))^{2n+2}}{\mathcal{H}^{2n}(A)}$$

with respect to all $\tilde{\phi} \in Hom(H(n), vol, Lip)$. Because the diameter of $\tilde{A}$ is greater than the square root of the ∞ height of $\tilde{A}$, it follows that

$$isod(\tilde{A}) \geq \frac{h_\infty(\tilde{A})^{n+1} \, w(\tilde{A})}{\mathcal{H}^{2n+2}(\tilde{A})}$$

All these invariants scale with dilatations like this (denote generically by inv any of the mentioned invariants):

$$inv \, \delta_\varepsilon \tilde{A} \; = \; \varepsilon^2 \tilde{A}$$

The groupoid $\mathcal{H}om(H(n), vol, Lip)$ with objects open sets $H(n)$ and arrows volume preserving locally bi-Lipschitz maps between then, can be considered instead of the group $Hom(H(n), vol, Lip)$. Then the heights are no longer invariants. Indeed, set $\tilde{A}$ to be the disjoint union of two isometric sets. Then we can keep one of these sets fixed and put on the top of it the other one, using only translations. The width will change. We can define other invariants though, using the heights and the number of connected components, for example.

By using lifts of symplectic flows one can define symplectic invariants. Conversely, with any symplectic capacity comes an invariant of $Hom(H(n), vol, Lip)$. Indeed, let c be a capacity. Then $\tilde{c}(\tilde{A}) \; = \; c(A)$ is an invariant.

Is there any general definition of a $Hom(H(n), vol, Lip)$ invariant which has as particular realization heights and symplectic capacities?

How about the invariants of the group $Hom(N, vol, Lip)$, where N is a general Carnot group? We leave this for further study.

4 Sub-Riemannian Lie groups

In this section we shall look closer to the last example of section 1.5. Let G be a real connected Lie group with Lie algebra $\mathfrak{g}$ and $D \subset \mathfrak{g}$ a vector space which generates the algebra. This means that the sequence

$$V_1 \; = \; D \; , \; V^{i+1} \; = \; V^i + [D, V^i]$$

provides a filtration of $\mathfrak{g}$:

$$V^1 \; \subset \; V^2 \; \subset \; ... \, V^m \; = \; \mathfrak{g} \tag{4.0.1}$$

4.1 Nilpotentisation

The filtration has the straightforward property: if $x \in V^i$ and $y \in V^j$ then $[x, y] \in V^{i+j}$ (where $V^k = \mathfrak{g}$ for all $k \geq m$ and $V^0 = \{0\}$). This allows to construct the Lie algebra

$$\mathfrak{n}(\mathfrak{g}, D) \; = \; \oplus_{i=1}^m V_i \; , \; V_i \; = \; V^i / V^{i-1}$$

with Lie bracket $[\hat{x}, \hat{y}]_n \; = \; \hat{[x, y]}$, where $\hat{x} = x + V^{i-1}$ if $x \in V^i \setminus V^{i-1}$.

Proposition 4.1 $\mathfrak{n}(\mathfrak{g}, D)$ *with distinguished space* $V_1 \; = \; D$ *is the Lie algebra of a Carnot group of step* m*. It is called the nilpotentisation of the filtration* (4.0.1)*.*

Remark that $\mathfrak{n}(\mathfrak{g}, D)$ and $\mathfrak{g}$ have the same dimension, therefore they are isomorphic as vector spaces.

We set the distribution induced by D to be

$$D_x \ = \ TL_x D$$

where $x \in G$ is arbitrary and $L_x : G \to G$, $L_x y \ = \ xy$ is the left translation by x. We shall use the same notation D for the induced distribution.

Let $\{X_1, ..., X_p\}$ be a basis of the vector space D. We shall build a basis of $\mathfrak{g}$ which will give a vector spaces isomorphism with $\mathfrak{n}(\mathfrak{g}, D)$.

A word with letters $A \ = \ \{X_1, ..., X_p\}$ is a string $X_{h(1)}...X_{h(s)}$ where $h : \{1, ..., s\} \to \{1, ..., p\}$. The set of words forms the dictionary $Dict(A)$, ordered lexicographically. We set the function $Bracket : Dict(A) \to \mathfrak{g}$ to be

$$Bracket(X_{h(1)}...X_{h(s)}) \ = \ [X_{h(1)}, [X_{h(2)}, [..., X_{h(s)}]...]$$

For any $x \in Bracket(Dict(A))$ let $\hat{x} \in Dict(A)$ be the least word such that $Bracket(\hat{x}) \ = \ x$. The collection of all these words is denoted by $\hat{g}$.

The length $l(x) \ = \ length(\hat{x})$ is well defined. The dictionary $Dict(A)$ admit a filtration made by the length of words function. In the same way the function l gives the filtration

$$V^1 \cap Bracket(Dict(A)) \ \subset \ V^2 \cap Bracket(Dict(A)) \ \subset \ ... \ Bracket(Dict(A))$$

Choose now, in the lexicographic order in $\hat{g}$, a set $\hat{B}$ such that $B \ = \ Bracket(\hat{B})$ is a basis for $\mathfrak{g}$. Any element X in this basis can be written as

$$X \ = \ [X_{h(1)}, [X_{h(2)}, [..., X_{h(s)}]...]$$

such that $l(X) \ = \ s$ or equivalently $X \in V^s \setminus V^{s-1}$.

It is obvious that the map

$$X \in B \ \mapsto \ \tilde{X} \ = \ X + V^{l(X)-1} \in \mathfrak{n}(\mathfrak{g}, D)$$

is a bijection and that $\tilde{B} \ = \ \left\{ \tilde{X}_j : j \ = \ 1, ..., dim \ \mathfrak{g} \right\}$ is a basis for $\mathfrak{n}(\mathfrak{g}, D)$. We can identify then $\mathfrak{g}$ with $\mathfrak{n}(\mathfrak{g}, D)$ by the identification $X_j \ = \ \tilde{X}_j$.

Equivalently we can define the nilpotent Lie bracket $[\cdot, \cdot]_n$ directly on $\mathfrak{g}$, with the use of the dilatations on $\mathfrak{n}(\mathfrak{g}, D)$.

Instead of the filtration (4.0.1) let us start with a direct sum decomposition of $\mathfrak{g}$

$$\mathfrak{g} \ = \ \oplus_{i=1}^m W_i \ , \ \ V^i \ = \ \oplus_{j=1}^i W_j \tag{4.1.2}$$

such that $[V^i, V^j] \subset V^{i+j}$. The chain V^i form a filtration like (4.0.1).

We set

$$\delta_\varepsilon(x) \ = \ \sum_{i=1}^m \varepsilon^i x_i$$

for any $\varepsilon > 0$ and $x \in \mathfrak{g}$, which decomposes according to (4.1.2) as $x \ = \ \sum_{i=1}^m x_i$.

Proposition 4.2 *The limit*

$$[x,y]_n = \lim_{\varepsilon \to 0} \delta_\varepsilon^{-1}[\delta_\varepsilon x, \delta_\varepsilon y] \tag{4.1.3}$$

exists for any $x, y \in \mathfrak{g}$ and $(\mathfrak{g}, [\cdot, \cdot]_n)$ is the Lie algebra of a Carnot group with dilatations δ_ε.

Proof. Let $x = \sum_{j=1}^{i} x_j$ and $y = \sum_{k=1}^{l} y_k$. Then

$$[\delta_\varepsilon x, \delta - \varepsilon y] = \sum_{s=2}^{i+l} \varepsilon^s \sum_{j+k=s} \sum_{p=1}^{s} [x_j, y_k]_p$$

We apply δ_ε^{-1} to this equality and we obtain:

$$\delta_\varepsilon^{-1}[\delta_\varepsilon x, \delta - \varepsilon y] = \sum_{s=2}^{i+l} \sum_{j+k=s} \sum_{p=1}^{s} \varepsilon^{s-p}[x_j, y_k]_p$$

When ε tends to 0 the expression converges to the limit

$$[x,y]_n = \sum_{s=2}^{i+l} \sum_{j+k=s} [x_j, y_k]_s \tag{4.1.4}$$

For any $\varepsilon > 0$ the expression

$$[x,y]^\varepsilon = \delta_\varepsilon^{-1}[\delta_\varepsilon x, \delta - \varepsilon y]$$

is a Lie bracket (bilinear, antisymmetric and it satisfies the Jacobi identity). Therefore at the limit $\varepsilon \to 0$ we get a Lie bracket. Moreover, it is straightforward to see from the definition of $[x,y]_n$ that δ_ε is an algebra isomorphism. We conclude that $(\mathfrak{g}, [\cdot, \cdot]_n)$ is the Lie algebra of a Carnot group with dilatations δ_ε. $\blacksquare$

Proposition 4.3 *Let $\{X_1, ..., X_{dim\ \mathfrak{g}}\}$ be a basis of $\mathfrak{g}$ constructed from a basis $\{X_1, ..., X_p\}$ of D. Set*

$$W_j = span\ \{X_i : l(X_i) = j\}$$

Then the spaces W_j provides a direct sum decomposition (4.1.2). Moreover the identification $\hat{X}_i = X_i$ gives a Lie algebra isomorphism between $(\mathfrak{g}, [\cdot, \cdot]_n)$ and $\mathfrak{n}(\mathfrak{g}, D)$.

Proof. In the basis $\{X_1, ..., X_{dim\ \mathfrak{g}}\}$ the Lie bracket on $\mathfrak{g}$ looks like this:

$$[X_i, X_j] = \sum C_{ijk} X_k$$

where $c_{ijk} = 0$ if $l(X_i) + l(X_j) < l(X_k)$. From here the first part of the proposition is straightforward. The expression of the Lie bracket generated by the decomposition (4.1.2) is obtained from (4.1.4)). We have

$$[X_i, X_j]_n = \sum \lambda_{ijk} C_{ijk} X_k$$

where $\lambda_{ikj} = 1$ if $l(X_i) + l(X_j) = l(X_k)$ and 0 otherwise. The Lie algebra isomorphism follows from the expression of the Lie bracket on $\mathfrak{n}(\mathfrak{g}, D)$:

$$[\hat{X}_i, \hat{X}_j] \ = \ [X_i, X_j] \ + \ V^{l(X_i) + l(X_j) - 1}$$

■

In conclusion the expression of the nilpotent Lie bracket depends on the choice of basis B trough the transport of the dilatations group from $\mathfrak{n}(\mathfrak{g}, D)$ to $\mathfrak{g}$.

Let $N(G, D)$ be the simply connected Lie group with Lie algebra $\mathfrak{n}(\mathfrak{g}, D)$. As previously, we identify $N(G, D)$ with $\mathfrak{n}(\mathfrak{g}, D)$ by the exponential map.

4.2 Commutative smoothness for uniform groups

This section is just a sketch and certainly needs to be improved. It is inspired from Bellaïche [3], last section. I think that by improving the notions explained here and by making axioms from the conclusions of lemma 4.13 and Ball-Box theorem 4.19, then it is possible to avoid the use of the notion of Lie group. For example, one could use consequences of proposition 4.7 in order to construct a general manifold structure, noncommutative derivative, moment map, as in sections 4.3, 4.5, 4.7.

We start with the following setting: G is a topological group endowed with an uniformity such that the operation is uniformly continuous. More specifically, we introduce first the double of G, as the group $G^{(2)} \ = \ G \times G$ with operation

$$(x, u)(y, v) \ = \ (xy, y^{-1}uyv)$$

The operation on the group G, seen as the function

$$op : G^{(2)} \to G \ , \quad op(x, y) \ = \ xy$$

is a group morphism. Also the inclusions:

$$i' : G \to G^{(2)} \ , \quad i'(x) \ = \ (x, e)$$

$$i'' : G \to G^{(2)} \ , \quad i''(x) \ = \ (x, x^{-1})$$

are group morphisms.

Definition 4.4 *G is an uniform group if we have two uniformity structures, on G and G^2, such that op, i', i'' are uniformly continuous.*

A local action of a uniform group G on a uniform pointed space (X, x_0) is a function $\phi \in W \in \mathcal{V}(e) \mapsto \hat{\phi} : U_\phi \in \mathcal{V}(x_0) \to V_\phi \in \mathcal{V}(x_0)$ such that the map $(\phi, x) \mapsto \hat{\phi}(x)$ is uniformly continuous from $G \times X$ (with product uniformity) to to X. Moreover, the action has the property: for any $\phi, \psi \in G$ there is $D \in \mathcal{V}(x_0)$ such that for any $x \in D$ $\phi \hat{\psi}^{-1}(x)$ and $\hat{\phi}(\hat{\psi}^{-1}(x))$ make sense and $\phi \hat{\psi}^{-1}(x) = \hat{\phi}(\hat{\psi}^{-1}(x))$.

Finally, a local group is an uniform space G with an operation defined in a neighbourhood of $(e, e) \subset G \times G$ which satisfies the uniform group axioms locally.

Remark that a local group acts locally at left (and also by conjugation) on itself.

Definition 4.5 *A conical local uniform group N is a local group with a local action of $(0, +\infty)$ by morphisms δ_ε such that $\lim_{\varepsilon \to 0} \delta_\varepsilon x = e$ for any x in a neighbourhood of the neutral element e.*

We shall make the following hypotheses on the local uniform group G: there is a local action of $(0, +\infty)$ (denoted by δ), on (G, e) such that

H0. $\delta_\varepsilon(e) = e$ and $\lim_{\varepsilon \to 0} \delta_\varepsilon x = e$ uniformly with respect to x.

H1. the limit

$$\beta(x, y) = \lim_{\varepsilon \to 0} \delta_\varepsilon^{-1} \left((\delta_\varepsilon x)(\delta_\varepsilon y) \right)$$

is well defined in a neighbourhood of e and the limit is uniform.

H2. the limit

$$\hat{x}^{-1} = \lim_{\varepsilon \to 0} \delta_\varepsilon^{-1} \left((\delta_\varepsilon x)^{-1} \right)$$

is well defined in a neighbourhood of e and the limit is uniform.

Proposition 4.6 *Under the hypotheses H0, H1, H2 (G, β) is a conical local uniform group.*

Proof.   All the uniformity assumptions permit to change at will the order of taking limits. We shall not insist on this further and we shall concentrate on the algebraic aspects.

We have to prove the associativity, existence of neutral element, existence of inverse and the property of being conical. The proof is straightforward. For the associativity $\beta(x, \beta(y, z)) = \beta(\beta(x, y), z)$ we compute:

$$\beta(x, \beta(y, z)) = \lim_{\varepsilon \to 0, \eta \to 0} \delta_\varepsilon^{-1} \left\{ (\delta_\varepsilon x) \delta_{\varepsilon/\eta} \left((\delta_\eta y)(\delta_\eta z) \right) \right\}$$

We take $\varepsilon = \eta$ and we get

$$= \beta(x, \beta(y, z)) = \lim_{\varepsilon \to 0} \left\{ (\delta_\varepsilon x)(\delta_\varepsilon y)(\delta_\varepsilon z) \right\}$$

In the same way:

$$\beta(\beta(x, y), z) = \lim_{\varepsilon \to 0, \eta \to 0} \delta_\varepsilon^{-1} \left\{ (\delta_{\varepsilon/\eta} x) \left((\delta_\eta x)(\delta_\eta y) \right) (\delta_\varepsilon z) \right\}$$

and again taking $\varepsilon = \eta$ we obtain

$$\beta(\beta(x, y), z) = \lim_{\varepsilon \to 0} \left\{ (\delta_\varepsilon x)(\delta_\varepsilon y)(\delta_\varepsilon z) \right\}$$

The neutral element is e, from H0 (first part): $\beta(x, e) = \beta(e, x) = x$. The inverse of x is $\hat{x}^{-1}$, by a similar argument:

$$\beta(x, \hat{x}^{-1}) = \lim_{\varepsilon \to 0, \eta \to 0} \delta_\varepsilon^{-1} \left\{ (\delta_\varepsilon x) \left(\delta_{\varepsilon/\eta}(\delta_\eta x)^{-1} \right) \right\}$$

and taking $\varepsilon = \eta$ we obtain

$$\beta(x, \hat{x}^{-1}) \;=\; \lim_{\varepsilon \to 0} \delta_\varepsilon^{-1}\left((\delta_\varepsilon x)(\delta_\varepsilon x)^{-1}\right) \;=\; \lim_{\varepsilon \to 0} \delta_\varepsilon^{-1}(e) \;=\; e$$

Finally, β has the property:

$$\beta(\delta_\eta x, \delta_\eta y) \;=\; \delta_\eta \beta(x, y)$$

which comes from the definition of β and commutativity of multiplication in $(0, +\infty)$. This proves that (G, β) is conical. ∎

We arrive at a natural realization of the tangent space to the neutral element. Let us denote by $[f, g] \;=\; f \circ g \circ f^{-1} \circ g^{-1}$ the commutator of two transformations. For the group we shall denote by $L_x^G y \;=\; xy$ the left translation and by $L_x^N y \;=\; \beta(x, y)$. The preceding proposition tells us that (G, β) acts locally by left translations on G.

Proposition 4.7 *We have the equality*

$$\lim_{\lambda \leq \varepsilon \to 0} \delta_\lambda^{-1}[\delta_\varepsilon, L_{(\delta_\lambda x)^{-1}}^G]\delta_\lambda \;=\; L_x^N \tag{4.2.5}$$

and the limit is uniform with respect to x.

Proof.   We shall use the equality:

$$\lim_{\varepsilon \to 0} \delta_\varepsilon^{-1}\left(\beta((\delta_\varepsilon x), \delta_\varepsilon y)\right) \;=\; \lim_{\varepsilon \to 0} \delta_\varepsilon\left(\beta(\hat{x}^{-1}, \delta_\varepsilon^{-1}(\beta(x, y)))\right) \tag{4.2.6}$$

which is a consequence of H0 and the previous proposition. The left hand side of this equality equals $\beta(x, y)$. We shall look now at the right hand side (RHS) of the equality. The uniformity assumptions and some computations lead us to the following equality, under the constraint $0 < \max\{\mu, \lambda, \eta\} < \varepsilon$:

$$RHS \;=\; \lim_{(\varepsilon, \mu, \lambda, \eta) \to 0} \delta_{\varepsilon/\mu}\left\{\delta_{\mu/\lambda}\left((\delta_\lambda x)^{-1}\right)\delta_{\mu/(\varepsilon\eta)}\left((\delta_\eta x)(\delta_\eta y)\right)\right\}$$

The mentioned constraint forces all expressions to make sense. Take now $\mu = \lambda = \eta < \varepsilon$ and get

$$RHS \;=\; \lim_{\varepsilon, \lambda \to 0} \delta_\lambda^{-1}\delta_\varepsilon\left(\left((\delta_\lambda x)^{-1}\right)\delta_\varepsilon^{-1}\left((\delta_\lambda x)(\delta_\lambda y)\right)\right)$$

as desired. ∎

An easy corollary is that

$$\lim_{\lambda \to 0}[L_{(\delta_\lambda x)^{-1}}^G, \delta_\lambda^{-1}] \;=\; L_x^N$$

Definition 4.8 *The group VT_eG formed by all transformations L_x^N is called the virtual tangent space at e to G.*

The virtual tangent space VT_xG at $x \in G$ to G is obtained by translating the group operation and the dilatations from e to x. This means: define a new operation on G by

$$y \overset{x}{\cdot} z = yx^{-1}z$$

The group G with this operation is isomorphic to G with old operation and the left translation $L_x^G y = xy$ is the isomorphism. The neutral element is x. Introduce also the dilatations based at x by

$$\delta_\varepsilon^x y = x\delta_\varepsilon(x^{-1}y)$$

Then $G^x = (G, \overset{x}{\cdot})$ with the group of dilatations δ_ε^x satisfy the axioms Ho, H1, H2. Define then the virtual tangent space VT_xG to be: $VT_xG = VT_xG^x$. A short computation using proposition 4.7 shows that

$$VT_xG = \left\{ L_y^{N,x} = L_x L_{x^{-1}y}^N L_x : y \in U_x \in \mathcal{V}(X) \right\}$$

where

$$L_y^{N,x} = \lim_{\lambda \leq \varepsilon \to 0} \delta_\lambda^{-1,x} [\delta_\varepsilon^x, L_{(\delta_\lambda x)^x, -1}^G] \delta_\lambda^x$$

We shall introduce the notion of commutative smoothness, which contains a derivative resembling with Pansu derivative. Independently, the author introduced a general "topological" derivative in [4] (there are a lot of typographical errors in this reference, but the ideas behind are quite nice in my opinion). Lack of knowledge stopped the development of this notion until recently, when I have learned about Carnot groups and Pansu derivative. As it will be seen further, interesting information will be found when noncommutative smoothness is considered.

Definition 4.9 *A function $f : G_1 \to G_2$ is commutative smooth, where G_1, G_2 are two groups satisfying H0, H1, H2, if the application*

$$(x,u) \in G_1^{(2)} \mapsto (f(x), Df(x)u) \in G_2^{(2)}$$

exists and it is continuous, where

$$Df(x)u = \lim_{\varepsilon \to 0} \delta_\varepsilon^{-1} \left(f(x)^{-1} f(x\delta_\varepsilon u) \right)$$

and the convergence is uniform with respect to (x,u).

For example the left translations L_x are commutative smooth and the derivative equals identity. If we want to see how the derivative moves the virtual tangent spaces we have to give a definition.

Definition 4.10 *Le $f : G \to G$ be a commutative smooth function. The virtual tangent to f is defined by:*

$$VTf(x) : VT_xG \to VT_{f(x)}G , \quad VTf(x)L_y^{N,x} = L_{f(x)} L_{Df(x)y}^N L_{f(x)}^{-1}$$

With this definition L_x is commutative smooth and it's virtual tangent in any point y is a group morphism from VT_yG to $VT_{xy}G$. More generally, reasoning as in proposition 2.14, we get:

Proposition 4.11 *The virtual tangent is morphism of conical groups.*

Nevertheless the right translations are not commutative smooth. This failure pushes us to consider a notion of noncommutative derivative. We shall meet the same failure soon, when looking to the manifold structure, in the sense of definition 2.20, of the group G.

Now that we have a model for the tangent space to e at G, we can show that the operation is commutative smooth.

Proposition 4.12 *Let G satisfy H0, H1, H2 and $\delta_\varepsilon^{(2)} : G^{(2)} \to G^{(2)}$ be defined by*

$$\delta_\varepsilon^{(2)}(x, u) \;=\; (\delta_\varepsilon x, \delta_\varepsilon y)$$

Then $G^{(2)}$ satisfies H0, H1, H2, op is commutative smooth and we have the relation:

$$D \; op \; (x, u)(y, v) \;=\; \beta(y, v)$$

Proof.   It is sufficient to use the morphism property of the operation. Indeed, the right hand side of the relation to be proven is

$$RHS \;=\; \lim_{\varepsilon \to 0} \delta_\varepsilon^{-1} \left(op(x, u)^{-1} op(x, u) op \left(\delta_\varepsilon^{(2)}(y, v) \right) \right) \;=$$

$$=\; \lim_{\varepsilon \to 0} \delta_\varepsilon^{-1} \left(op(\delta_\varepsilon^{(2)}(y, v)) \right) \;=\; \beta(y, v)$$

The rest is trivial.   ∎

4.3   Manifold structure

The notion of virtual tangent space is not based on the use of distances, but on the use of dilatations. In fact, any manifold has a tangent space to any of its points, not only the Riemannian manifolds. We shall prove in this section that VT_eG is isomorphic to the nilpotentisation $N(G, D)$.

Nevertheless G does not have the structure of a $N(G, D)$ $\mathcal{C}^1$ manifold, in the sense of definition 2.20.

We start from Euclidean norm on D and we choose an orthonormal basis of D. We can then extend the Euclidean norm to $\mathfrak{g}$ by stating that the basis of $\mathfrak{g}$ constructed, as explained, from the basis on D, is orthonormal. By left translating the Euclidean norm on $\mathfrak{g}$ we endow G with a structure of Riemannian manifold. The induced Riemannian distance d_R will give an uniform structure on G. This distance is left invariant:

$$d_R(xy, xz) \;=\; d_R(y, z)$$

for any $x, y, z \in G$.

Any left invariant distance d is uniquely determined if we set $d(x) = d(e, x)$.

The following lemma is important (compare with lemma 2.3).

Lemma 4.13 *Let $X_1, ..., X_p$ be a basis of D. Then there are $U \subset G$ and $V \subset N(G, D)$, open neighbourhoods of the neutral elements e_G, e_N respectively, and a surjective function $g : \{1, ..., M\} \rightarrow \{1, ..., p\}$ such that any $x \in U$, $y \in V$ can be written as*

$$x = \prod_{i=1}^{M} \exp_G(t_i X_{g(i)}) \ , \ y = \prod_{i=1}^{M} \exp_N(\tau_i X_{g(i)}) \tag{4.3.7}$$

Proof. We shall make the proof for G; the proof for $N(G, D)$ will follow from the identifications explained before.

Denote by n the dimension of $\mathfrak{g}$. We start the proof with the remark that the function

$$(t_1, ..., t_n) \ \mapsto \ \prod_{i=1}^{n} \exp_G(t_i X_i) \tag{4.3.8}$$

is invertible in a neighbourhood of $0 \in R^n$, where the X_i are elements of a basis B constructed as before. Remember that each $X_i \in B$ is a multi-bracket of elements from the basis of D. If we replace a bracket $\exp_G(t[x, y])$ in the expression (4.3.8) by $exp_G(t_1 x) \, exp_G(t_2 y) \, exp_G(t_3 x) \, exp_G(t_4 y)$ and we replace t by $(t_1, ..., t_4)$ then the image of a neighbourhood of 0 by the obtained function still covers a neighbourhood of the neutral element. We repeat this procedure a finite number of times and the thesis is proven. ■

As a corollary we obtain the Chow theorem for our particular example.

Theorem 4.14 *Any two points $x, y \in G$ can be joined by a horizontal curve.*

Let d_G be the Carnot-Carathéodory distance induced by the distribution D and the metric. This distance is also left invariant. We obviously have $d_R \leq d_G$. We want to show that d_R and d_G induce the same uniformity on G.

Let us introduce another left invariant distance on G

$$d_G^1(x) \ = \ \inf \left\{ \sum |t_i| \ : \ x \ = \ \prod \exp_G(t_i Y_i) \ , \ Y_i \in D \right\}$$

and the auxiliary functions :

$$\Delta_G^1(x) \ = \ \inf \left\{ \sum_{i=1}^{M} |t_i| \ : \ x \ = \ \prod \exp_G(t_i X_{g(i)}) \right\}$$

$$\Delta_G^\infty(x) \ = \ \inf \left\{ \max |t_i| \ : \ x \ = \ \prod \exp_G(t_i X_{g(i)}) \right\}$$

From theorems 1.5 and 1.3 we see that $d_G^1 = d_G$. Indeed, it is straightforward that $d_G \leq d_G^1$. On the other part $d_G^1(x)$ is less equal than the variation of any Lipschitz curve joining e_G with x. Therefore we have equality.

The functions Δ_G^1, Δ_G^∞ don't induce left invariant distances. Nevertheless they are useful, because of their equivalence:

$$\Delta_G^\infty(x) \ \leq \ \Delta_G^1(x) \ \leq \ M \ \Delta_G^\infty(x) \tag{4.3.9}$$

for any $x \in G$. This is a consequence of the lemma 4.13.

We have therefore the chain of inequalities:

$$d_R \ \leq \ d_G \ \leq \ \Delta_G^1 \ \leq \ M\Delta_G^\infty$$

But from the proof of lemma 4.13 we see that Δ_G^∞ is uniformly continuous. This proves the equivalence of the uniformities.

Because $\exp_G$ does not deform much the d_R distances near e, we see that the group G with the dilatations

$$\tilde{\delta}_\varepsilon(\exp_G x) \ = \ \exp_G(\delta_\varepsilon x)$$

satisfies H0, H1, H2.

The same conclusion is true for the local uniform group (with the uniformity induced by the Euclidean distance) $\mathfrak{g}$ with the operation:

$$XY \ = \ \log_G\left(\exp_G(X)\exp_G(Y)\right)$$

for any X, Y in a neighbourhood of $0 \in \mathfrak{G}$. Here the dilatations are δ_ε. We shall denote this group by $\log G$ These two groups are isomorphic as local uniform groups by the map $\exp_G$. Dilatations commute with the isomorphism. They have therefore isomorphic (by $\exp_G$) virtual tangent spaces.

Theorem 4.15 *The virtual tangent space VT_eG is isomorphic to $N(G,D)$. More precisely $N(G,D)$ is equal (as local group) to the virtual tangent space to $\log G$:*

$$N(G,D) \ = \ VT_0 \log G$$

Proof. The product XY in $\log G$ is given by Baker-Campbell-Hausdorff formula

$$X \overset{\mathfrak{g}}{\cdot} Y \ = \ X + Y + \frac{1}{2}[X,Y] + ...$$

Use proposition 4.2 to compute $\beta(X,Y)$ and show that $\beta(X,Y)$ equals the nilpotent multiplication. ∎

Let $\{X_1, ..., X_{dim\ \mathfrak{g}}\}$ be a basis of $\mathfrak{g}$ constructed from a basis $\{X_1, ..., X_p\}$, as explained before. Denote by

$$b : \mathfrak{g} \to \mathfrak{n}(\mathfrak{g}, D)$$

the identification of the vectorspace $\mathfrak{g}$ with $\mathfrak{n}(\mathfrak{g}, D)$ using the basis $\{X_1, ..., X_{dim\ \mathfrak{g}}\}$. The exponential and the logarithm of the groups G and $N(G,D)$ will be denoted by $\exp_G$, $\log_G$ and $\exp_N$, $\log_N$ respectively.

We shall consider the following $N(G, D)$ atlas for G:

$$\mathcal{A} = \{\psi_x : U_x \subset G \to N(D, G) : x \in G ,$$

$$\psi_x(y) = \exp_N \circ b \circ \log_G(x^{-1}y)\}$$

We shall study the differentiability properties of the transition functions. We want to work in $\mathfrak{g}$. For this we have to transport (in a neighbourhood $0 \in \mathfrak{g}$) the interesting operations, namely:

$$X \overset{\mathfrak{g}}{\cdot} Y = \log_G (\exp_G(X) \exp_G(Y))$$

$$X \overset{\mathfrak{n}}{\cdot} Y = \log_N (\exp_N(X) \exp_N(Y))$$

and to use instead of the chart $\psi_x : U_x \subset G \to N(D, G)$, the chart $\phi_X : U_x \subset G \to \mathfrak{g}$, where $x = \exp_G X$ and $\phi_X(y) = (-X) \overset{\mathfrak{g}}{\cdot} \log_G(y)$.

The transition function from ϕ_X to ϕ_0 is then the left translation by X, with respect to the operation $\overset{\mathfrak{g}}{\cdot}$. We denote this translation by $L_X^{\mathfrak{g}}$. We want to know if this function is Pansu derivable from $(\mathfrak{g}, \overset{\mathfrak{n}}{\cdot})$ to itself. See the difference: the function is certainly derivable (or commutative smooth) in the sense of definition 4.9, for the operation $\overset{\mathfrak{g}}{\cdot}$, but what about the derivability with respect to the operation $\overset{\mathfrak{n}}{\cdot}$? In order to answer we shall use the following trick: a C^∞ function $f : N \to N$ is Pansu derivable if and only if it preserves the horizontal distribution and the derivative in each point is a morphism from N to N.

The (classical) derivative of $L_X^{\mathfrak{g}}$ moves the distribution $D^n(Y) = DL_Y^{\mathfrak{n}}(e)$ into $DL_X^{\mathfrak{g}} D^n(Y) \subset T_{X \overset{\mathfrak{g}}{\cdot} Y} \mathfrak{g}$. The horizontal distribution in $X \overset{\mathfrak{g}}{\cdot} Y$, corresponding to the group operation $\overset{\mathfrak{n}}{\cdot}$, is $D^n(X \overset{\mathfrak{g}}{\cdot} Y)$. The difference between these two distributions is measured by one of the linear transformations:

$$A_{X,Y} : T_{x \overset{\mathfrak{g}}{\cdot} Y} \mathfrak{g} \to T_{x \overset{\mathfrak{g}}{\cdot} Y} \mathfrak{g}$$

$$A_{X,Y} = DL_{X \overset{\mathfrak{g}}{\cdot} Y}^{\mathfrak{g}}(0) \left(DL_Y^{\mathfrak{g}}(0)\right)^{-1} DL_Y^{\mathfrak{n}}(0) \left(DL_{X \overset{\mathfrak{g}}{\cdot} Y}^{\mathfrak{n}}(0)\right)^{-1} \qquad (4.3.10)$$

$$A^{X,Y} : \mathfrak{g} \to \mathfrak{g}$$

$$A^{X,Y} = \left(DL_{X \overset{\mathfrak{g}}{\cdot} Y}^{\mathfrak{n}}(0)\right)^{-1} DL_{X \overset{\mathfrak{g}}{\cdot} Y}^{\mathfrak{g}}(0) \left(DL_Y^{\mathfrak{g}}(0)\right)^{-1} DL_Y^{\mathfrak{n}}(0) \qquad (4.3.11)$$

Let then $\mathfrak{J}(G, D)$ be the Lie group generated by these transformations

$$\mathfrak{J}(G, D) = \langle\{A^{X,Y} : X, Y \in \mathfrak{g}\}\rangle$$

It is then straightforward to see that the algebra $\mathfrak{j}(G, D)$ of this group contains the algebra generated by all the linear transformation with the form

$$a_x = ad_x^G - ad_x^N \qquad (4.3.12)$$

The necessary and sufficient condition for the group $\mathcal{J}(G, D)$ to be included in the group $End((\mathfrak{g}, \overset{\mathfrak{n}}{\cdot}), D)$ is that all the elements a_X, to be in the algebra of the mentioned linear group. But this is equivalent with one of the (equivalent) conditions:

(i) Ad_x^G is $[\cdot,\cdot]_N$ morphism, for any x in a neighbourhood of the identity,

(ii) for any $X,U,V \in \mathfrak{g}$ we have the identity:

$$[[X,U]_G,V]_N \; + \; [U,[X,V]_G]_N \;\; = \;\; [X,[U,V]_N]_G \qquad\qquad (4.3.13)$$

We collect what we have found in the following theorem.

Theorem 4.16 *Set Ad^G to be the adjoint representation of G and Ad^N the adjoint representation of $N(G,D)$, seen as group of linear transformations on $\mathfrak{g}$ (via the identification function b). The following are equivalent:*

(a) $\mathcal{J}(G,D) \;\subset\; End(\mathfrak{g}, \overset{\text{n}}{\cdot}),$

(b) $Ad^G \subset\; End(\mathfrak{g}, \overset{\text{n}}{\cdot}),$

(c) the relation (4.3.13) is true.

If (4.3.13) holds then $Ad^G Ad^N$ is a group and its Lie algebra is the adjoint representation of the algebra $\mathfrak{g} \oplus \mathfrak{g}$ with the bracket:

$$[(X,U),(Y,V)] \;\; = \;\; ([X,Y]_G, [X,V]_G + [U,Y]_G + [V,U]_N) \qquad\qquad (4.3.14)$$

Proof.   The equivalence of (a), (b), (c) has been explained. Suppose now that (4.3.13) is true. It is then straightforward to show that the space of all elements

$$a_{(X,Y)} \;\; = \;\; ad_X^G \;-\; ad_Y^N$$

forms a Lie algebra with the linear commutator as bracket. Moreover, we have:

$$[a_{(X,U)}, a_{(Y,V)}] \;\; = \;\; a_{[(X,U),(Y,V)]}$$

This shows that $Ad^G Ad^N$ is a group and the property of its algebra. In order to finish the proof remark that $a_{(X,X)} \;=\; a_X$, defined at (4.3.12).   ∎

Corollary 4.17 *The atlas $\mathcal{A}$ gives to G a $\mathcal{C}^1$ $N(D,G)$ manifold structure if and only if (4.3.13) is true and for any $\varepsilon > 0$, $X,Y \in \mathfrak{g}$*

$$\delta_\varepsilon^{-1}[X,\delta_\varepsilon Y]_N \;-\; \delta_\varepsilon^{-1}[X,\delta_\varepsilon Y]_G \;\; = \;\; [X,Y]_N \;-\; [X,Y]_G \qquad\qquad (4.3.15)$$

Proof. The atlas $\mathcal{A}$ gives to G a $\mathcal{C}^1$ $N(D, G)$ manifold structure if and only if $\mathcal{J}(G, D) \subset HL(\mathfrak{g}, \overset{n}{\cdot})$ and any of its elements commute with dilatations δ_ε. This is equivalent with (4.3.13) and (4.3.15).   ■

Remark 4.18 *A Lie group is a manifold endowed with a smooth operation. In what sense is then G a (sub-Riemannian) Lie group? We already have problems to assign an atlas with smooth transition functions to G. The real meaning of the corollary 4.17 is that even if we succeed to give to G an atlas with smooth transition functions then the left translations and the exponential map will not be smooth. This is a problem. If we renounce to give G a manifold structure, at least the operation is smooth in the sense of proposition 4.12.*

Returning to the problem of the atlas, if $N(G, D)$ is a Heisenberg group, or equivalently the distribution has codimension one, then the commentary of Bellaïche [3], page 73, shows that such an atlas exists, as a consequence of Darboux theorem on normal forms for contact differential forms. But such an atlas is not compatible with the operation.

We shall see now how it can be possible that the operation is commutative smooth but the right translations are not.

We shall look to the double $G^{(2)}$. This group is isomorphic with $G^2 = G \times G$ with componentwise multiplication by the isomorphism

$$F : G^{(2)} \to G^2 , \quad F(x, u) = (x, xu)$$

The Lie algebra of $G^{(2)}$ is easily then $\mathfrak{g}^{(2)} = \mathfrak{g} \times \mathfrak{g}$, with Lie bracket

$$[(x, u), (y, v)]_{(2)} = ([x, u], [x + u, y + v] - [x, y])$$

Take in $\mathfrak{g}^{(2)}$ the distribution $D^{(2)} = D \times D$. This distribution generates $\mathfrak{g}^{(2)}$, and the dilatations are

$$\delta_\varepsilon^{(2)}(x, u) = (\delta_\varepsilon x, \delta_\varepsilon u)$$

We translate the distribution at left all over $G^{(2)}$.

We know that the operation is commutative smooth. We shall check if the classical derivative of the operation transports the distribution $D^{(2)}$ in the distribution D. Straightforward computation shows that

$$D \; op(e, e) D^{(2)} = D$$

where D here means classical derivative. Indeed, $D \; op(e, e)(X, Y) = X + Y$ and D is a vector space, therefore $(X, Y) \in D^{(2)}$ implies $X + Y \in D$.

For general $(x, u) \in G^{(2)}$ this is no longer true. Indeed, we have:

$$D \; op(x, u)(Y, V) = DL_{xu}^G(0) \left\{ Y + V + (Ad_{xu}^G - Ad_u^G)Y \right\}$$

therefore a sufficient condition for $D \; op(x, u) D^{(2)} = D$ is that $x = e$.

In conclusion the fact that the operation is commutative smooth does not imply that so are the right translations.

We can hope that by a slight modification in the definition of smoothness we shall be able to give to G a manifold structure and simultaneously to have smooth right translations. For the manifold structure we have to "smooth" the group generated by the elements with the form

$$\left(Ad_X^G\right)^{-1} Ad_X^N$$

To solve the problem of the right translations we need that the group Ad^G be "smoothed".

Before doing this we shall see that the pure metric point of view does not feel these problems but in such a precise way.

4.4 Metric tangent cone

We shall prove the result of Mitchell [23] theorem 1, that G admits in any point a metric tangent cone, which is isomorphic with the nilpotentisation $N(G, D)$. Mitchell theorem is true for regular sub-Riemannian manifolds. The proof that we give here is based on the lemma 4.13 and Gromov [13] section 1.2.

Because left translations are isometries, it is sufficient to prove that G admits a metric tangent cone in identity and that the tangent cone is isometric with $N(G, D)$.

For this we transport all in the algebra $\mathfrak{g}$, endowed with two brackets $[\cdot, \cdot]_G$, $[\cdot, \cdot]_N$ and with two operations $\overset{\mathfrak{g}}{\cdot}$ and $\overset{\mathfrak{n}}{\cdot}$. Denote by d_G, d_N the Carnot-Carathéodory distances corresponding to the $\overset{\mathfrak{g}}{\cdot}$, respectively $\overset{\mathfrak{n}}{\cdot}$ left invariant distributions on (a neighbourhood of 0 in) $\mathfrak{g}$. l_G, l_N are the corresponding length functionals. We shall denote by $B_G(x, R)$, $B_N(x, R)$ the balls centered in x with radius R with respect to d_G, d_N.

We can refine lemma 4.13 in order to obtain the Ball-Box theorem in this more general situation.

Theorem 4.19 *(Ball-Box Theorem) Denote by*

$$Box_G^1(\varepsilon) = \left\{ x \in G : \Delta_G^1(x) < \varepsilon \right\}$$

$$Box_N^1(\varepsilon) = \left\{ x \in N : \Delta_N^1(x) < \varepsilon \right\}$$

For small $\varepsilon > 0$ there is a constant $C > 1$ such that

$$\exp_G\left(Box_N^1(\varepsilon)\right) \subset Box_G^1(C\varepsilon) \subset \exp_G\left(Box_N^1(C^2\varepsilon)\right)$$

Proof. Reconsider the proof of lemma 4.13. This time, instead of the trick of replacing commutators $\exp_G[X, Y](t)$ with four letters words

$$\exp_G X(t_1) \exp_G Y(t_2) \exp_G X(t_3) \exp_G Y(t_4)$$

we shall use a smarter replacement (which, important! , works equally for the nilpotentisation $N(G, D)$).

We start with a basis $\{Y_1, ..., Y_n\}$ of the algebra $\mathfrak{g}$, constructed from multibrackets of elements $\{X_1, ..., X_p\}$ which form a basis for the distribution D. Introduce an Euclidean distance by declaring the basis of $\mathfrak{g}$ orthonormal. Set

$$[X(t), Y(t)]^\circ \ = \ (tX) \overset{\mathfrak{g}}{\cdot} (tY) \overset{\mathfrak{g}}{\cdot} (-tX) \overset{\mathfrak{g}}{\cdot} (-tY)$$

It is known that for any set of vectors $Z_1, ..., Z_q \in \mathfrak{g}$, if we denote by $\alpha^\circ(Z_1(t), ..., Z_q(t))$ a $[\cdot, \cdot]^\circ$ multibracket and by $\alpha(Z_1, ..., Z_q)$ the same constructed $[\cdot, \cdot]$ multibracket, then we have

$$\alpha^\circ(Z_1(t), ..., Z_q(t)) \ = \ t^q \alpha(Z_1, ..., Z_q) \ + \ o(t^q)$$

with respect to the Euclidean norm.

Remark as previously that the function

$$(t_1, ..., t_n) \ \mapsto \ \prod_{i=1}^{n} (t_i Y_i) \tag{4.4.16}$$

is invertible in a neighbourhood of $0 \in R^n$. Each X_i from the basis of $\mathfrak{g}$ can be written as a multibracket

$$X_i \ = \ \alpha_i(X_{i_1}, ..., X_{j^i})$$

which has the length $l_i = j^i - 1$. If l_i is odd then replace $(t_i X_i)$ by

$$\alpha^\circ(X_{i_1}(t^{1/l_i}), ..., X_{j^i}(t^{1/l_i}))$$

If l_i is even then the multibracket α_i can be rewritten as

$$\alpha_i(X_{i_1}, ..., X_{j^i}) \ = \ \beta_i(X_{i_1}, ..., [X_{i_k}, X_{i_{k+1}}], ..., X_{j^i})$$

Replace then $(t_i X_i)$ by

$$\begin{cases} \beta_i^\circ(X_{i_1}(|\,t\,|^{1/l_i}), ..., [X_{i_k}(|\,t\,|^{1/l_i}), X_{i_{k+1}}(|\,t\,|^{1/l_i})]^\circ, ..., X_{j^i}(|\,t\,|^{1/l_i})) & \text{if } t \geq 0 \\ \beta_i^\circ(X_{i_1}(|\,t\,|^{1/l_i}), ..., [X_{i_{k+1}}(|\,t\,|^{1/l_i}), X_{i_k}(|\,t\,|^{1/l_i})]^\circ, ..., X_{j^i}(|\,t\,|^{1/l_i})) & \text{if } t \leq 0 \end{cases}$$

After this replacements in the expression (4.4.16) one obtains a function E_G which is still invertible in a neighbourhood of 0. We obtain a function E_N with the same algebraic expression as E_G, but with $[\cdot, \cdot]_N^\circ$ brackets instead of $[\cdot, \cdot]_G^\circ$ ones. Use these functions to (obviously) end the proof of the theorem. ∎

Theorem 4.20 *The Gromov-Hausdorff limit of pointed metric spaces $(\mathfrak{g}, 0, \lambda d_G)$ as $\lambda \to \infty$ exists and equals $(\mathfrak{g}, d_N)$.*

Proof. We shall use the proposition 1.24. For this we shall construct ε isometries between $Box_N^1(1)$ and $Box_G^1(C\varepsilon)$. These are provided by the function $E_G \circ E_N^{-1} \circ \delta_\varepsilon$.

The trick consists in the definition of the nets. We shall exemplify the construction for the case of a 3 dimensional algebra $\mathfrak{g}$. The basis of $\mathfrak{g}$ is $X_1, X_2, X_3 \ = \ [X_1, X_2]_G$. Divide the interval $[0, X_1]$ into P equal parts, same for the interval $[0, X_2]$. The interval $[0, X_3]$ though will be divided into P^2 intervals. The net so obtained, seen in $N \ \equiv \ \mathfrak{g}$,

turn $E_G \circ E_N^{-1} \circ \delta_\varepsilon$ into a ε isometry between $Box_N^1(1)$ and $Box_G^1(C\varepsilon)$. To check this is mostly a matter of smart (but still heavy) notations. ∎

The proof of this theorem is basically a refinement of Proposition 3.15, Gromov [13], pages 85–86, mentioned in the introduction of these notes. Close examination shows that the theorem is a consequence of the following facts:

(a) the identity map $id : (\mathfrak{g}, d_G) \to (\mathfrak{g}, d_N)$ has finite dilatation in 0 (equivalently $\exp_G : N(D, G) \to G$ has finite dilatation in 0),

(b) the change from the G-invariant distribution induced by D to the N invariant distribution induced by the same D is classically smooth.

It is therefore natural that the result does not feel the non-derivability of id map in $x \neq 0$.

4.5 Noncommutative smoothness in Carnot groups

Let N be a Carnot group and

$$N = V_1 + \dots + V_m \tag{4.5.17}$$

the graduation of its algebra. Let $End(N)$ be the group of endomorphisms of N.

Definition 4.21 *The group of vertical endomorphisms of N, noted by $VL(N)$, is the subgroup of $End(N)$ of all N endomorphisms A which admit the form $A = (A_{i,j})$ with respect to the direct sum decomposition of N (4.5.17), such that*

$$A_{ii} = id_{V_i} \ , \ A_{ij} = 0 \ \ \forall i < j \in \{1, ..., m\}$$

Proposition 4.22 *(a) The elements of the linear group $HL(N)$ have diagonal form with respect to the direct sum decomposition (4.5.17).*

(b) Any $A \in End(N)$ admits the unique decomposition

$$A = A^v A^h \ , \ A^v \in VL(N) \ , \ A^h \in HL(N)$$

(c) $Ad(N)$ is a normal subgroup of $VL(N)$. The quotient group is Carnot.

Proof. (b) Proof by induction. Just write what the algebra morphism condition for $A \in End(N)$ means and get the inferior diagonal form of A. Such a matrix (with elements linear transformations) decomposes uniquely in a product of a vertical morphism A^v and a diagonal morphism A^h. Check that the latter commutes with dilatations.

(a) We know from (b) that any element $A \in HL(N)$ has inferior diagonal form. Commutation with dilatations forces A to have diagonal form.

(c) $Ad(N)$ is normal in $End(N)$ and it is a subgroup of $VL(N)$. It is therefore normal in $VL(N)$. The quotient group is nilpotent, because $VL(N)$ is nilpotent. It is Carnot because it has dilatations. Consider the map

$$\delta^\varepsilon \; : End(N) \to End(N) \; , \;\; \delta^\varepsilon A \;=\; \delta_\varepsilon A \delta_\varepsilon^{-1}$$

From (b) we get that δ^ε is an endomorphism of $VL(N)$. From the fact that $\delta_\varepsilon \in End(N)$ we get that

$$\delta^\varepsilon Ad(N) \;=\; Ad(N)$$

Therefore δ^ε factorizes to an endomorphism of the quotient $VL(N)/Ad(N)$. The rest is trivial. ∎
We shall call the quotient

$$R(N) \;=\; VL(N)/Ad(N)$$

the rest of N.

The definition of noncommutative derivative follows. For any function $f : N \to N$ and $x \in N$ set

$$f_x : N \to N \; , \quad f_x(y) \;=\; f(x)^{-1} f(xy)$$

f is Pansu derivable in x if and only if f_x is Pansu derivable in 0; moreover the Pansu derivative of f in x equals the Pansu derivative of f_x in 0.

From the proof of Pansu's Rademacher theorem one can extract the following information:

Proposition 4.23 *if f is Pansu derivable and the convergence*

$$\lim_{\varepsilon \to 0} \delta_\varepsilon^{-1} f_x(x \delta_\varepsilon y)$$

is uniform with respect to x then f is classically derivable in x along the horizontal directions and we have the connection between Pansu derivative and classical derivative:

$$Pf(x) \;=\; \nabla f(x) \;=\; \left(DL_{f(x)} \right)^{-1}(0) Df(x) DL_x(0)$$

Conversely if f is classically derivable along the distribution, such that the quantity

$$\nabla f(x)_{|D} \;=\; \left(DL_{f(x)} \right)^{-1}(0) Df(x) DL_x(0)_{|D}$$

can be extended to an algebra morphism of N, an such that the convergence

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left(f(x \delta_\varepsilon y) - f(x) \right)$$

(with $y \in D$) is uniform with respect to x then f is Pansu derivable.

Definition 4.24 *A function $f : N \to N$ is noncommutative smooth if for any $B \in VL(N)$ there is a continuous function $A_B : N \to VL(N)$ such that:*

(a) for any $x \in N$ the map

$$y \;\mapsto\; \tilde{f}^B_x(y) \;=\; A(x)^{-1} f_x(By)$$

is Pansu derivable in 0,

(b) the map

$$(x, y) \in N^2 \mapsto (f(x), D\tilde{f}^B_x(0)y) \in N^2$$

is continuous,

(c) the convergence

$$Df_x(0)y \;=\; \lim_{\varepsilon \to 0} \delta_\varepsilon^{-1} f_x(x\delta_\varepsilon y)$$

is uniform with respect to x.

The derivative of f in x is by definition

$$Df(x) \;=\; A_I(x) D\tilde{f}^I_x(x)$$

Note that if f is invertible then we have the decomposition:

$$Df(x)^v \;=\; A_I(x) \;, \quad Df(x)^h \;=\; D\tilde{f}^I_x(x)$$

Moreover, if f is invertible and classically derivable then the conclusion of proposition 4.23 is still true, namely $Df(x) \;=\; \nabla f(x)$, where the "gradient" $\nabla f(x)$ is given by

$$\left(DL_{f(x)}\right)^{-1}(0) Df(x) DL_x(0)$$

The derivative is called noncommutative because it does not commute with dilatations.

Proposition 4.25 *If $f : N \to N$ is commutative smooth then it is also noncommutative smooth.*

If $f, g : N \to N$ are noncommutative smooth then so it is $f \circ g$.

Proof.   Because the noncommutative derivative splits the function in a horizontal and a vertical part, it is sufficient to prove that if f is commutative smooth and g is noncommutative smooth such that $D^h g(x) \;=\; id$ then $f \circ g$ is noncommutative smooth. This is contained in the definiton of noncommutative smoothness.   ∎

The chain formula is true for noncommutative derivative. But obviously not for the vertical and the horizontal parts.

With this definition of smoothness the right translations in a Carnot group become smooth. It is easy to show that

$$D\,R_x \;=\; Ad_x^{-1} \;, \quad D^v R_x \;=\; Ad_x^{-1} \;, \quad D^h R_x \;=\; id$$

This smoothness definition seems to be a good one for Carnot groups. It is not good enough though, as the following example shows.

4.6   Moment maps and the Heisenberg group

Let $G = T^k \oplus N$ be the direct sum of the k dimensional torus T^k with the Carnot group N. The Lie algebra of G is $\mathfrak{g} = R^k \oplus N$ (recall that we use the same notation for the Carnot group N and its Lie algebra).

The algebra $\mathfrak{g}$ is Carnot, with respect to the dilatations

$$\delta_\varepsilon(T, X) = (\varepsilon T, \delta_\varepsilon X)$$

The distribution is generated by $R^k \oplus V^1$ and the nilpotentisation of G with respect to this distribution is $N' = R^k \oplus H$.

The corollary 4.17 tells that G admits a N' manifold structure. As the construction of the noncommutative derivative is local in nature, we could reproduce it for N' manifolds. In this case that simply means: in reasonings of local nature one can work on N' instead of G.

Take now $N = H(n)$ and a toric Hamiltonian action

$$(s_1, ..., s_k) \in T^k \; \mapsto \; \phi_{(s_1,...,s_k)} \in Sympl(R^{2n})$$

For any $v \in R^k = Lie \, T^k$ the infinitesimal generator is

$$\xi_v(x) = \frac{d}{dt}_{|t=0} \phi_{\exp tv}(x)$$

Let H_v be the Hamiltonian of $t \mapsto \phi_{\exp tv}$. The moment map associated to the Hamiltonian toric action is then

$$j_\phi : R^{2n} \; \to \; \left(R^k\right)^* \;, \quad \langle j(x), v \rangle = H_v(x)$$

We shall associate to the toric action a function on $T^K \times H(n)$.

$$\Phi : G = T^k \times H(n) \to G \;, \quad \Phi(s, x, \bar{x}) = (s, \tilde{\phi}_s(x, \bar{x})) \tag{4.6.18}$$

where $\tilde{\phi}_s$ is the lift of $(\phi_s, 0)$ (see theorem 3.3) and it has the expression:

$$(\phi_s(x), \bar{x} + F_s(x))$$

Let us compute the "gradient"

$$\nabla \Phi(s, x\bar{x}) = \left(DL_{\Phi(s,x,\bar{x})}(0)\right)^{-1} D\Phi(s, x, \bar{x}) \, DL_{(s,x,\bar{x})}(0)$$

In matrix form with respect to the decomposition of the algebra

$$R^k \oplus H(n) = R^k \oplus R^{2n} \oplus R$$

we obtain the following expression:

$$\nabla \Phi(s, x, \bar{x}) = \begin{pmatrix} 1 & 0 & 0 \\ \frac{\partial \phi_s}{\partial s} & D\phi_s & 0 \\ a(s, x) & 0 & 1 \end{pmatrix}$$

where $a(s, x)$ is the function:

$$a(s, x) = \frac{\partial F}{\partial s} - \frac{1}{2}\omega(\phi, \frac{\partial \phi}{\partial s})$$

We can improve the expression of the function a if we use the fact that the toric action is Hamiltonian. Indeed, by direct computation one can show that

$$\frac{\partial a}{\partial x}(s, x) = -\frac{\partial}{\partial x}(j_\phi(\phi_s(x)))$$

This is equivalent with

$$a(s, x) = \lambda(s) - j_\phi(\phi_s(x))$$

Easy computation shows that $End(R^k \oplus H(n))$ is made by all matrices with the form:

$$\begin{pmatrix} A & 0 & 0 \\ 0 & B & 0 \\ a & b & c \end{pmatrix} , \ A \in GL(R^k) , \ \begin{pmatrix} B & 0 \\ b & c \end{pmatrix} \in HL(H(n)) , \ a \in \left(R^k\right)^* = \left(Lie \ T^k\right)^*$$

It follows that $\nabla\Phi \notin End(R^k \oplus H(n))$ and we already saw this in a particular form in proposition 3.6. Therefore Φ is not even noncommutative smooth in the sense of definition 4.24.

Suppose that the toric action is by diffeomorphisms with compact support. Then $\lambda(s)$ can be taken equal to 0 hence $a(x, s) = -j_\phi(x)$.

Proposition 4.26 *The map Φ transforms the distribution*

$$D^N(s, x, \bar{x}) = \left\{(S, Y, \bar{Y}) \ : \ -\frac{1}{2}\omega(x, Y) + \bar{Y} = 0\right\}$$

in the distribution

$$D^\phi(s, x, \bar{x}) = \left\{(S, Y, \bar{Y}) \ : \ \langle j_\phi(x), S\rangle - \frac{1}{2}\omega(x, Y) + \bar{Y} = 0\right\}$$

Proof. By direct computation. ■

4.7 General noncommutative smoothness

The introduction of noncommutative derivative makes right translations in Carnot groups "smooth". We shall apply the same strategy for having a manifold structure compatible with the group operation, in the case of a Lie group G with a left invariant distribution D.

The transition functions of the chosen atlas have the form $\left(L_X^\mathfrak{g}\right)^{-1}$, with $X \in \mathfrak{g}$. These transition functions forms a group. Note that a natural generalization is the case where the transition functions forms a groupoid. This generalisation is not meaningless

since it leads to the introduction of intrinsic SR manifolds, if we identify a manifold with the groupoid of transition functions of an atlas.

We want this family of transition functions to be "smooth". We shall proceed then as in the motivating example.

The group G is identified with the group of left translations L^G acting on the Carnot group $(\mathfrak{g}, \overset{n}{\cdot})$. These transformation are not smooth, they preserve another distribution than the N left invariant distribution on G.

We shall give the following "upgraded" variant of the definition 4.24, concerning the noncommutative derivative.

We shall denote by N the set $\mathfrak{g}$ endowed with the nilpotent multiplication $\overset{n}{\cdot}$. The same notation will be used for the Lie algebra $(\mathfrak{g}, [\cdot, \cdot]_N)$. We fix on N an Euclidean norm, given, for example, by declaring orthonormal a basis constructed from multi-brackets from a basis of D. The Euclidean norm will be denoted by $\| \cdot \|$.

We introduce the group of linear transformations $\Sigma(G, D)$ generated by all transformations $DL^g_X(0)$, $DR^g_X(0)$, $HL(N)$. In particular $\delta_\varepsilon \in \Sigma(G, D)$ for any $\varepsilon > 0$. From the formula:

$$DL^N_X(0) \;=\; \lim_{\varepsilon \to 0} \delta_\varepsilon^{-1} \circ DL^g_{\delta_\varepsilon X}(0) \circ \delta_\varepsilon$$

we get that $DL^N_X(0) \in \Sigma(G, D)$. In the same way we obtain $DR^N_X(0) \in \Sigma(G, D)$.

For any $F \in \Sigma(G, D)$ we define the group $N(F)$ to be the Carnot group N with the bracket:

$$[X, Y]^F_N \;=\; F[F^{-1}X, F^{-1}Y]_N$$

and with the dilatations

$$\delta_\varepsilon^F \;=\; F\delta_\varepsilon F^{-1}$$

This group is endowed with the Carnot-Carathéodory distance generated by the Euclidean norm and the distribution $D^N(F) \;=\; FD$.

We can see the class $\{N(F) : F \in \Sigma(G, D)\}$ in two ways, as an orbit. First, it is the orbit

$$\bar{\Sigma}(G, D) \;=\; \left\{ (F\delta_\varepsilon F^{-1}, F \, ad^N_{F^{-1}X} F^{-1}) \; : \; F \in \Sigma(G, D) \right\}$$

in the space $GL(N) \times gl(N, gl(N))$, under the action of $\Sigma(G, D)$

$$F.(A, B) \;=\; (FAF^{-1}, F(B)) \;, \quad F(B)(Z) \;=\; FB(F^{-1}Z)F^{-1}$$

Second, it is the right coset $\Sigma(G, D)/HL(N)$. Indeed, if $N(F) \;=\; N(P)$ then $F^{-1}P \in HL(N)$.

Definition 4.27 *A function $f : N \to N$ is (G, D) noncommutative derivable in $x \in N$ if for any $F \in \Sigma(G, D)$ there exists $P \in \Sigma(G, D)$ such that $f : N(F) \to N(P)$ is Pansu derivable in x. The derivative of f in x is by definition:*

$$Df(x)(F, P) \;=\; P^{-1} \circ D_{Pansu}f(x) \circ F$$

The function f is noncommutative smooth if

(a) the map

$$(x, y) \in N^2 \mapsto (f(x), Df(x)y) \in N^2$$

is continuous,

(c) the convergence

$$D_{Pansu}f(x)(y \ = \ \lim_{\varepsilon \to 0} \delta_\varepsilon^{-1,P} \left(P(P^{-1}(f(x))^{-1} \ P^{-1} f(x\delta_\varepsilon^F y))) \right)$$

is uniform with respect to x.

If f is invertible then following map is well defined:

$$\Lambda(f, x) : \bar{\Sigma}(G, D) \to \bar{\Sigma}(G, D)$$

$\Lambda(f, x)(F \ HL(N)) \ = \ P \ HL(N)$, where $f : N(F) \to N(P)$ is Pansu derivable in x. In this case the noncommutative derivative can be written as $Df(x, N(F))$.

The definition can be adapted for functions $f : N \to R$, or $f : R \to N$.

Definition 4.28 *A function $f : N \to R$ is noncommutative derivable in x if for any $F \in \Sigma(G, D)$ the function $f : N(F) \to R$ is Pansu derivable in x. In this case*

$$Df(x, N(F)) \ = \ D_{Pansu}f(x) \circ F$$

A function $f : R \to N$ is noncommutative derivable in $x \in R$ if there exists $P \in \Sigma(G, D)$ such that $f : R \to N(P)$ is Pansu derivable in $x \in R$. In this case

$$Df(x) \ = \ P^{-1} \circ D_{Pansu}f(x)$$

The definition of noncommutative smoothness adapts obviously.

We might risk to have no noncommutative derivable functions from N to R. On the contrary we shall have a lot of noncommutative derivable functions from R to N.

What is the connection with the previous notion of noncommutative smoothness? The noncommutative smoothness in the sense of definition 4.24 corresponds to the definition 4.27 if we replace $\Sigma(G, D)$ with $End(N)$.

The following proposition has now a straightforward proof.

Proposition 4.29 *The transition functions of the $N(G, D)$ atlas $\mathcal{A}$ are $\Sigma(G, D)$ smooth.*

Proof.   Indeed, it is sufficient to remark that $\mathcal{J}(G, D) \subset \Sigma(G, D)$.   ∎

4.8   Margulis & Mostow tangent bundle

In this section we shall apply Margulis & Mostow [21] construction of the tangent
bundle to a SR manifold for the case of a group with left invariant distribution. It will
turn that the tangent bundle does not have a group structure, due to the fact that, as
previously, the non-smoothness of the right translations is not studied.

   The main point in the construction of a tangent bundle is to have a functorial
definition of the tangent space. This is achieved by Margulis & Mostow [21] in a very
natural way. One of the geometrical definitions of a tangent vector v at a point x, to
a manifold M, is the following one: identify v with the class of smooth curves which
pass through x and have tangent v. If the manifold M is endowed with a distance then
one can define the equivalence relation based in x by: $c_1 \equiv_x c_2$ if $c_1(0) = c_2(0) = x$ and
the distance between $c_1(t)$ and $c_2(t)$ is of order t^2 for small t. The set of equivalence
classes is the tangent space at x. One has to put then some structure on the tangent
space (as, for example, the nilpotent multiplication).

   To put is practice this idea is not so easy though. This is achieved by the following
sequence of definitions and theorems. For commodity we shall explain this construction
in the case $M = G$ connected Lie group, endowed with a left invariant distribution D.
The general case is the one of a regular sub-Riemannian manifold. We shall denote by
d_G the CC distance on G and we identify G with $\mathfrak{g}$, as previously. The CC distance
induced by the distribution D^N, generated by left translations of G using nilpotent
multiplication $\overset{\text{n}}{\cdot}$, will be denoted by d_N.

Definition 4.30 *A C^∞ curve in G with $x = c(0)$ is called rectifiable at $t = 0$ if*
$d_G(x, c(t)) \leq Ct$ as $t \to 0$.

   Two C^∞ curves c', c'' with $c'(0) = x = c''(0)$ are called equivalent at x if $t^{-1}d_G(c'(t), c''(t)) \to$
0 as $t \to 0$.

   The tangent cone to G as x, denoted by $C_x G$ is the set of equivalence classes of all
C^∞ paths c with $c(0) = x$, rectifiable at $t = 0$.

   Let $c : [-1, 1] \to G$ be a C^∞ rectifiable curve, $x = c(0)$ and

$$v = \lim_{t \to 0} \delta_t^{-1} \left(c(0)^{-1} \overset{\mathfrak{g}}{\cdot} c(t) \right) \tag{4.8.19}$$

The limit v exists because the curve is rectifiable.
   Introduce the curve $c_0(t) = x \exp_G(\delta_t v)$. Then

$$d(x, c_0(t)) = d(e, x^{-1}c(t)) < |v| t$$

as $t \to 0$ (by the Ball-Box theorem) The curve c is equivalent with c_0. Indeed, we have
(for $t > 0$):

$$\frac{1}{t}d_G(c(t), c_0(t)) = \frac{1}{t}d_G(c(t), x \overset{\mathfrak{g}}{\cdot} \delta_t v) = \frac{1}{t}d_G(\delta_t(v^{-1}) \overset{\mathfrak{g}}{\cdot} x^{-1} \overset{\mathfrak{g}}{\cdot} c(t), 0)$$

The latter expression is equivalent (by the Ball-Box Theorem) with

$$\frac{1}{t}d_N(\delta_t(v^{-1}) \overset{\mathfrak{g}}{\cdot} x^{-1} \overset{\mathfrak{g}}{\cdot} c(t), 0) = d_N(\delta_t^{-1} \left(\delta_t(v^{-1}) \overset{\mathfrak{g}}{\cdot} \delta_t \left(\delta_t^{-1} \left(x^{-1} \overset{\mathfrak{g}}{\cdot} c(t) \right) \right) \right))$$

The right hand side (RHS) converges to $d_N(v^{-1} \overset{\mathfrak{n}}{\cdot} v, 0)$, as $t \to 0$, as a consequence of the definition of v and theorem 4.15.

Therefore we can identify $C_x G$ with the set of curves $t \mapsto x \exp_G(\delta_t v)$, for all $v \in \mathfrak{g}$. Remark that the equivalence relation between curves c_1, c_2, such that $c_1(0) = c_2(0) = x$ can be redefined as:

$$\lim_{t \to 0} \delta_t^{-1} \left(c_2(t)^{-1} \overset{\mathfrak{g}}{\cdot} c_1(t) \right) = 0 \tag{4.8.20}$$

In order to define the multiplication Margulis & Mostow introduce the families of segments rectifiable at t.

Definition 4.31 *A family of segments rectifiable at $t = 0$ is a C^∞ map*

$$\mathcal{F} : U \to G$$

where U is an open neighbourhood of $G \times 0$ in $G \times R$ satisfying

(a) $\mathcal{F}(\cdot, 0) = id$

(b) the curve $t \mapsto \mathcal{F}(x,t)$ is rectifiable at $t = 0$ uniformly for all $x \in G$, that is for every compact K in G there is a constant C_K and a compact neighbourhood I of 0 such that $d_G(y, \mathcal{F}(y,t)) < C_K t$ for all $(y,t) \in K \times I$.

Two families of segments rectifiable at $t = 0$ are called equivalent if $t^{-1} d_G(\mathcal{F}_1(x,t), \mathcal{F}_2(x,t)) \to 0$ as $t \to 0$, uniformly on compact sets in the domain of definition.

Part (b) from the definition of a family of segments rectifiable can be restated as: there exists the limit

$$v(x) = \lim_{t \to 0} \delta_t^{-1} \left(x^{-1} \overset{\mathfrak{g}}{\cdot} \mathcal{F}(x,t) \right) \tag{4.8.21}$$

and the limit is uniform with respect to $x \in K$, K arbitrary compact set.

It follows then, as previously, that $\mathcal{F}$ is equivalent to $\mathcal{F}_0$, defined by:

$$\mathcal{F}_0(x,t) = x \overset{\mathfrak{g}}{\cdot} \delta_t v(x)$$

Also, the equivalence between families of segments rectifiable can be redefined as:

$$\lim_{t \to 0} \delta_t^{-1} \left(\mathcal{F}_2(x,t)^{-1} \overset{\mathfrak{g}}{\cdot} \mathcal{F}_1(x,t) \right) = 0 \tag{4.8.22}$$

uniformly with respect to $x \in K$, K arbitrary compact set.

Definition 4.32 *The product of two families $\mathcal{F}_1$, $\mathcal{F}_2$ of segments rectifiable at $t = 0$ is defined by*

$$(\mathcal{F}_1 \circ \mathcal{F}_2)(x,t) = \mathcal{F}_1(\mathcal{F}_2(x,t), t)$$

The product is well defined by Lemma 1.2 *op. cit.*. One of the main results is then the following theorem (5.5).

Theorem 4.33 *Let c_1, c_2 be C^∞ paths rectifiable at $t = 0$, such that $c_1(0) = x_0 = c_2(0)$. Let $\mathcal{F}_1$, $\mathcal{F}_2$ be two families of segments rectifiable at $t = 0$ with:*

$$\mathcal{F}_1(x_0, t) = c_1(t) \quad , \quad \mathcal{F}_2(x_0, t) = c_2(t)$$

Then the equivalence class of

$$t \mapsto \mathcal{F}_1 \circ \mathcal{F}_2(x_0, t)$$

depends only on the equivalence classes of c_1 and c_2. This defines the product of the elements of the tangent cone $C_{x_0} G$.

This theorem is the straightforward consequence of the following facts (5.1(5) and 5.2 in Margulis & Mostow [**?**]).

Remark 4.34 *Maybe I misunderstood the notations, but it seems to me that several times the authors claim that the exponential map which they construct is bi-Lipschitz (as in 5.1(4) and Corollary 4.5). This is false, as explained before. In Bellaïche [3], Theorem 7.32 and also at the beginning of section 7.6 we find that the exponential map is only $1/m$ Hölder continuous (where m is the step of the nilpotentization). However, (most of) the results of Margulis & Mostow hold true. It would be interesting to have a better written paper on the subject, even if the subject might seem trivial (which is not).*

We shall denote by $\mathcal{F} \approx \mathcal{F}'$ the equivalence relation of families of segments rectifiable; the equivalence relation of rectifiable curves based at x will be denoted by $c \overset{x}{\approx} c'$.

Lemma 4.35   (a) *Let $\mathcal{F}_1 \approx \mathcal{F}_2$ and $\mathcal{G}_1 \approx \mathcal{G}_2$. Then $\mathcal{F}_1 \circ \mathcal{G}_1 \approx \mathcal{F}_2 \circ \mathcal{G}_2$.*

   (b) *The map $\mathcal{F}$ mapsto$\mathcal{F}_0$ is constant on equivalence classes of families of segments rectifiable.*

Proof.   Let
$$\mathcal{F}_0(x, t) = x \overset{\mathfrak{g}}{\cdot} \delta_t w_1(x) \quad , \quad \mathcal{G}_0(x, t) = x \overset{\mathfrak{g}}{\cdot} \delta_t w_2(x)$$

For the point (a) it is sufficient to prove that

$$\mathcal{F} \circ \mathcal{G} \approx \mathcal{F}_0 \circ \mathcal{G}_0$$

This is true by the following chain of estimates.

$$\frac{1}{t} d_G(\mathcal{F} \circ \mathcal{G}(x, t), \mathcal{F}_0 \circ \mathcal{G}_0(x, t)) = \frac{1}{t} d_G(\delta_t w_1(\mathcal{G}_0(x, t))^{-1} \overset{\mathfrak{g}}{\cdot} \delta_t w_2(x)^{-1} \overset{\mathfrak{g}}{\cdot} x^{-1} \overset{\mathfrak{g}}{\cdot} \mathcal{F}(\mathcal{G}(x, t), t), 0)$$

The RHS of this equality behaves like

$$d_N(\delta_t^{-1} \left(\delta_t w_1(\mathcal{G}_0(x, t))^{-1} \overset{\mathfrak{g}}{\cdot} \delta_t w_2(x)^{-1} \overset{\mathfrak{g}}{\cdot} \delta_t \left(\delta_t^{-1} \left(x^{-1} \overset{\mathfrak{g}}{\cdot} \mathcal{G}(x, t) \right) \right) \overset{\mathfrak{g}}{\cdot} \delta_t \left(\delta_t^{-1} \left(\mathcal{G}(x, t)^{-1} \overset{\mathfrak{g}}{\cdot} \mathcal{F}(\mathcal{G}(x, t), t) \right) \right) \right), 0)$$

This quantity converges (uniformly with respect to $x \in K$, K an arbitrary compact) to

$$d_N(w_1(x)^{-1} \overset{\mathfrak{n}}{\cdot} w_2(x)^{-1} \overset{\mathfrak{n}}{\cdot} w_2(x) \overset{\mathfrak{n}}{\cdot} w_1(x), 0) \; = \; 0$$

The point (b) is easier: let $\mathcal{F} \approx \mathcal{G}$ and consider $\mathcal{F}_0$, $\mathcal{G}_0$, as above. We want to prove that $\mathcal{F}_0 \; = \; \mathcal{G}_0$, which is equivalent to $w_1 \; = \; w_2$.

Because $\approx$ is an equivalence relation all we have to prove is that if $\mathcal{F}_0 \approx \mathcal{G}_0$ then $w_1 \; = \; w_2$. We have:

$$\frac{1}{t} d_G(\mathcal{F}_0(x,t), \mathcal{G}_0(x,t)) \; = \; \frac{1}{t} d_G(x \overset{\mathfrak{g}}{\cdot} \delta_t w_1(x), x \overset{\mathfrak{g}}{\cdot} \delta_t w_2(x))$$

We use the $\overset{\mathfrak{g}}{\cdot}$ left invariance of d_G and the Ball-Box theorem to deduce that the RHS behaves like

$$d_N(\delta_t^{-1}\left(\delta_t w_2(x)^{-1} \overset{\mathfrak{g}}{\cdot} \delta_t w_1(x)^{-1}\right), 0)$$

which converges to $d_N(w_1(x), w_2(x))$ as t goes to 0. The two families are equivalent, therefore the limit equals 0, which implies that $w_1(x) = w_2(x)$ for all x. ∎

We shall apply this theorem. Let $c_i(t) \; = \; x_0 \exp_G \delta_t v_i$, for $i = 1, 2$. It is easy to check that $\mathcal{F}_i(x, t) \; = \; x \exp_G(\delta_t v_i)$ are families of segments rectifiable at $t = 0$ which satisfy the hypothesis of the theorem. But then

$$(\mathcal{F}_1 \circ \mathcal{F}_2)(x, t) \; = \; x_0 \exp_G(\delta_t v_1) \exp_G(\delta_t v_2)$$

which is equivalent with

$$\mathcal{F} \exp_G\left(\delta_t(v_1 \overset{\mathfrak{n}}{\cdot} v_2)\right)$$

Therefore the tangent bundle defined by this procedure is the same as the virtual tangent bundle defined previously. It's definition is possible because the group operation has horizontal derivative in (e, e).

In terms of commutative derivative, theorem 10.5 Margulis & Mostow [20] (and restricting to bi-Lipschitz maps) becomes the Rademacher theorem. Indeed, in the case of a Lie group G endowed with a left invariant distribution D, with an associated Carnot-Carathéodory distance, the definition 4.9 of commutative derivative can be adapted in the following way:

Definition 4.36 *Let G_1, G_2 two groups endowed with left invariant distributions and d_1, d_2 tow associated Carnot-Carathéodory distances.*

A function $f : G_1 \rightarrow G_2$ is metrically commutative derivable in $x \in G_1$ if there is a $\varepsilon > 0$ such that the sequence

$$\Delta_\lambda f(x) u \; = \; \delta_\lambda^{-1}\left(f(x)^{-1} f(x \delta_\lambda u)\right)$$

converges uniformly with respect to $u \in B(x, \varepsilon)$. The derivative is

$$Df(x)u \; = \; \lim_{\varepsilon \to 0} \delta_\varepsilon^{-1}\left(f(x)^{-1}f(x\delta_\varepsilon u)\right)$$

and the virtual tangent is

$$VTf(x) : VT_x G_1 \rightarrow VT_{f(x)} G_2 \; , \quad VTf(x) L_y^{N_1, x} \; = \; L_{f(x)}^{G_2} L_{Df(x)y}^{N_2} L_{f(x)}^{-1, G_2}$$

Theorem 4.37 *Let $\phi : G \to G$ be a bi-Lipschitz map with respect to the Carnot-Carathéodory distance on G induced by the left invariant distribution D. Then for almost all $x \in G$ the map ϕ is commutative derivable at x and the virtual tangent to ϕ is an isomorphism of conical groups.*

We don't give the proof of this theorem. The reader might find interesting to adapt the proof of the mentioned Margulis & Mostow theorem in our particular case.

References

[1] D. Allcock, An isoperimetric inequality for the Heisenberg groups, *Geom. Funct. Anal.*, **8**, (1998), 219–233

[2] L. Ambrosio, Geometric measure theory and applications to the calculus of variations, (1999), chapters 1–4

[3] A. Bellaïche, The tangent space in sub-Riemannian geometry, in: *Sub-Riemannian Geometry*, A. Bellaïche, J.-J. Risler eds., Progress in Mathematics, **144**, Birchuser, (1996), 4 – 78

[4] M. Buliga, The topological substratum of the derivative, *Stud. Cerc. Mat. (Mathematical Reports)*, **45**, no. 6, (1993), 453–465

[5] D. Burago, Y. Burago, S. Ivanov, A Course in Metric Geometry, *Graduate Studies in Mathematics*, **33**, AMS Providence, Rhode Island, (2000)

[6] M. Cowling, A Dooley, A. Koranyi, F. Ricci, H-type groups and Iwasawa decompositions, *Adv. in Math.*, **87** (1991), 1–41

[7] Y. Eliashberg, L. Polterovich, Biinvariant metrics on the group of Hamiltonian diffeomorphisms, *International J. of Math.*, **4**, no. 5, (1993), 727–738

[8] G.B. Folland, E.M. Stein, Hardy spaces on homogeneous groups, *Mathematical notes*, **28**, Princeton - N.J.: Princeton University Press; Tokyo: University of Tokyo Press, (1982)

[9] B. Franchi, R. Serapioni, F. Serra Cassano, Rectifiability and Perimeter in the Heisenberg Group, *Math. Ann.*, **321**, (2001), 479–531

[10] N. Garofalo, D.M. Nhieu, Isoperimetric and Sobolev inequalitites for Carnot-Carathéodory spaces and the existence of minimal surfaces, *Comm. Pure Appl. Math.*, **49** (1996), no. 10, 1081–1144

[11] R. Goodman, Nilpotent Lie groups: structure and applications to analysis, *Lecture Notes in Math.* , **53** (1981)

[12] M. Gromov, Carnot-Caratheodory spaces seen from within, in: *Sub-Riemannian Geometry*, A. Bellaïche, J.-J. Risler eds., Progress in Mathematics, **144**, Birchuser, (1996), 79 – 323

[13] M. Gromov, Metric structures for Riemannian and non-Riemannian spaces, *Progress in Math.*, **152**, Birchäuser (1999), chapter 1.

[14] P. Hajlasz, P. Koskela, Sobolev met Poincaré, *Memoirs of the America Mathematical Society* (2000), **688**

[15] J. Hilgert, K.-H. Neeb, W. Plank, Symplectic convexity theorems and coadjoint orbits, *Compositio Mathematica*, **94**, 129–180, (1994)

[16] H. Hofer, E. Zehnder, Symplectic invariants and Hamiltonian dynamics, *Birkhuser Advanced Texts: Basler Lehrbcher*, Birkhuser Verlag, Basel, (1994)

[17] A. Koranyi, M. Reimann, Foundations for the theory of quasi-conformal mappings of the Heisenberg group, *Adv. in Math.*, **III:1** (1995), 1–87

[18] B. Kostant, On convexity, the Weyl group and the Iwasawa decomposition, *Ann. Scient. Ec. Norm. Sup.*, 4e série, **t. 6** (1973), 413–455

[19] V. Magnani, Differentiability and area formula on stratified Lie groups, preprint http:
www.cvgmt.it, (2000)

[20] G.A. Margulis, G.D. Mostow, The differential of a quasi-conformal mapping of a Carnot-Carathéodory space, *Geom. Funct. Analysis*, **8** (1995), 2, 402–433

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

[22] G.D. Mostow, Strong rigidity of locally symmetric spaces, *Ann. of Math. Studies*, **78** (1973), Princeton Univ. Press, Princeton

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

[24] R. Montgomery, Survey of singular geodesics, in: *Sub-Riemannian Geometry*, A. Bellaïche, J.-J. Risler eds., Progress in Mathematics, **144**, Birchuser, (1996), 325–339

[25] P. Pansu, Métriques de Carnot-Carathéodory at quasiisometries des espaces symetriques de rang un, *Ann. of Math.*, **129** (1989), 1–60

[26] P. Pansu, Croissance des boules et des géodésiques fermées dans les nilvariétés, *Ergod. Th. & Dynam. Sys.*, **3** (1983), 415–445

[27] P. Pansu, Une inégalité isopérimétrique sur le groupe d'Heisenberg, *C. R. Acad. Sci. Paris*, **295** (1982), 127–131

[28] L. Polterovich, The geometry of the group of symplectic diffeomorphisms, *Lectures in Mathematics ETH Zrich*, Birkhuser Verlag, Basel, (2001)

[29] H.M. Reimann, F. Ricci, The complexified Heisenberg group, *PRoc. on Analysis and Geometry*, Novosibirsk: Sobolev Institute Press, (2000), 465–480

[30] P. Sleewaegen, Application moment et theoreme de convexite de Kostant, PhD. thesis, Univ. Libre de Bruxelles, 1999

[31] H.J. Sussmann, A cornucopia of four-dimensional abnormal sub-Riemannian minimizers, in: *Sub-Riemannian Geometry*, A. Bellaïche, J.-J. Risler eds., Progress in Mathematics, **144**, Birchuser, (1996), 342–364

[32] N. Varopoulos, L. Saloff-Coste, Th. Coulhon, Analysis and Geometry on Groups, *Cambridge Univ. Press*, (1993)

[33] S.K. Vodop'yanov, $\mathcal{P}$-Differentiability on Carnot Groups in Different Topologies and Related Topics, *Proc. on Analysis ans Geometry*, Novosibirsk: Sobolev Institute Press, (2000), 603 – 670

[34] S.K. Vodop'yanov, A.D. Ukhlov, Approximately differentiable transformations and change of variables on nilpotent groups, *Sib. Math. J.*, **37** (1996), 1, 62–78