Uniform Estimation of Sub-Riemannian Balls

Frédéric JEAN
Laboratoire de Mathématiques Appliquées
École Nationale Supérieure de Techniques Avancées
32, bd Victor, 75739 Paris, France.
Email : fjean@ensta.fr
Tel: (33+1) 45 52 48 90, Fax: (33+1) 45 52 52 82.

July 3, 2001

Abstract

A fundamental result of sub-Riemannian geometry, the Ball-Box theorem, states that, in privileged coordinates, small sub-Riemannian balls look like boxes $[-\varepsilon^{w_1}, \varepsilon^{w_1}] \times \cdots \times [-\varepsilon^{w_n}, \varepsilon^{w_n}]$. This description is not uniform in general. Thus it does allow neither to compute Hausdorff measures and dimensions nor to prove the convergence of certain motion planning algorithms.

In this paper we present a description of the shape of small sub-Riemannian balls depending uniformly on their center and their radius. This result is a generalization of the Ball-Box Theorem. The proof is based on the one hand on a lifting method, which replaces the sub-Riemannian manifold by an extended equiregular one (where the Ball-Box Theorem is uniform); on the other hand on an estimate of sets defined by families of vector fields, which allows to project the balls in suitable coordinates.

Key words: uniform distance estimate, lifting method, Ball-Box theorem, sub-Riemannian geometry.

[1]1991 Mathematics Subject Classification. 53C99, 93B29
[2]2000 Mathematics Subject Classification. 53C17

1 Introduction

In sub-Riemannian geometry the shape of small balls is describe by the so-called Ball-Box Theorem (Bellaïche [4], Gromov [9]): for ε smaller than a given ε_p, the sub-Riemannian ball centered at p of radius ε looks like a box $[-\varepsilon^{w_1}, \varepsilon^{w_1}] \times \cdots \times [-\varepsilon^{w_n}, \varepsilon^{w_n}]$ in suitable coordinates. However this description is not uniform in general: the integers w_i can be different from one point to another and the positive number ε_p can depend discontinuously on p (in particular ε_p can be infinitely small on a compact set, though everywhere non-zero).

This reduces the scope of applications of the Ball-Box Theorem. Indeed the uniformity of the description of the small balls is crucial for a lot of computations. It allows first to count the number of balls of a given radius needed to cover a compact set. Such counting give the Hausdorff dimension and estimates of the Hausdorff measure and the entropy (Jean [11, 12]). It can also be used to show the convergence of certain motion planning algorithms taking into account singularities (Laumond, Oriolo and Venditelli [25, 13]).

Notice however that the Ball-Box Theorem is uniform near regular points, around which the integers w_i are constant. Results on Hausdorff measure and dimension exist then for equiregular manifolds (Mitchell [16], Pansu [19]). This uniformity is also responsible for the group structure of the tangent cone at regular points (Bellaïche [4], Margulis and Mostow [15]).

In this paper we present a description of the shape of small sub-Riemannian balls depending uniformly on their center and their radius. This result, Theorem 2, is a generalization of the Ball-Box Theorem. The idea is to lift the sub-Riemannian manifold into an equiregular one, where the Ball-Box Theorem is uniform, and then to project the balls in suitable coordinates. The lifting part is based on works of Bellaïche [4], Rothschild and Stein [22] and Goodman [8]. The choice of coordinates for the projection is inspired by results of Nagel, Stein and Wainger [18] on certain sets defined by families of vector fields.

The outline of this paper is as follows. Section 2 contains some recalls of sub-Riemannian geometry, in particular the definitions of regular and singular points and a precise statement of the Ball-Box Theorem.

The main result of this paper, Theorem 2, is established in Section 3. We express first the Ball-Box Theorem in a continuous form near regular points. This formulation is related to the notion of minimal basis of the tangent space. The statement of our theorem requires a generalization of this notion. We see in fact that we must choose the basis of the tangent space in function of the radius of the balls.

The proof of Theorem 2 is given in Section 4. We first construct an extended sub-Riemannian manifold. This manifold is equiregular and its balls are projected onto balls of the original manifold. The principle of the construction is to consider the tangent space as the quotient of a free nilpotent Lie group by a sub-group. The extended manifold is then the sub-Riemannian manifold which nilpotent approximation generates the corresponding free nilpotent Lie algebra.

In a second step we project the balls of the lifted manifold onto the original one. The problem here is to choose a "good" coordinates system, in which the projected balls look like boxes. This amounts to select a basis of the tangent space among a family of commutators.

In this last part we need to estimate and compare all the possible determinants by taking into account the length of the commutators and the radius of the ball. These results, grouped together in Lemma 7, are quite technical. Moreover they will be used in a more general form in a forthcoming paper. We therefore postpone their proof in Section 5.

2 Sub-Riemannian manifolds

We recall here some definitions and basic results of sub-Riemannian geometry. More general presentations can be found in Bellaïche [4] (the main reference for this section) or in Kupka [14].

Sub-Riemannian distance. Let M be a real analytic n-dimensional manifold and $X_1, \dots, X_m$ analytic vector fields on M. We define a *sub-Riemannian metric g on M* by setting, for each $q \in M$ and $v \in T_qM$,

$$g_q(v) = \inf \left\{ u_1^2 + \cdots + u_m^2 \;\middle|\; u_1 X_1(q) + \cdots + u_m X_m(q) = v \right\} .$$

Notice that the infimum of the empty set is defined as $+\infty$. The value $g_q(v)$ is then infinite when v does not belong to the linear space spanned by the $X_i(q)$.

The *length* of an absolutely continuous path $c(t)$ $(0 \leq t \leq \tau)$ is defined as

$$\mathrm{length}(c) = \int_0^\tau \sqrt{g_{c(t)}(\dot{c}(t))}dt .$$

The *sub-Riemannian distance* is $d(p,q) = \inf \mathrm{length}(c)$, where the infimum is taken on all the absolutely continuous paths joining p to q.

Remark. Only *horizontal paths* can have a finite length, that is the paths tangent almost everywhere to the distribution spanned by the X_i. The sub-Riemannian distance is then also equal to the infimum of the length taken on all the horizontal paths.

The manifold M endowed with the distance d, denoted (M, d), is called the *sub-Riemannian manifold* attached to $X_1, \dots, X_m$.

The sub-Riemannian distance between any two points of M is finite if M is connected and the following condition holds (Chow [6], Rashevsky [20]):

(Chow's Condition) *The vector fields $X_1, \dots, X_m$ and their iterated brackets $[X_i, X_j]$, $[[X_i, X_j], X_k]$, etc. span the tangent space T_pM at every point p of M.*

Notice that the condition is also necessary here, since $X_1, \dots, X_m$ are analytic vector fields (Nagano [17], Sussmann [24]).

When Chow's Condition is satisfied, the sub-Riemannian distance d is continuous. The topology defined by d is then the original topology of M. In the sequel we always assume that Chow's Condition is satisfied.

Singular points. Let $\mathcal{L}^1 = \mathcal{L}^1(X_1, \dots, X_m)$ be the set of linear combinations, with real coefficients, of the vector fields $X_1, \dots, X_m$. We define recursively $\mathcal{L}^s = \mathcal{L}^s(X_1, \dots, X_m)$ by setting, for $s > 1$,

$$\mathcal{L}^s = \mathcal{L}^{s-1} + [\mathcal{L}^{s-1}, \mathcal{L}^1]. \tag{1}$$

Due to Jacobi identity $\mathcal{L}^s$ is the set of linear combinations of all commutators of $X_1, \dots, X_m$ with a length $\leq s$. The union $\mathcal{L}$ of all $\mathcal{L}^s$ is a Lie sub-algebra of the Lie algebra of vector fields on M. It is generated by the commutators $[[X_{i_1}, X_{i_2}], \dots, X_{i_k}]$. Such a commutator is denoted $[X_I]$, where I is the multi-index $I = (i_1, \dots, i_k)$ and its length is $|I| = k$.

For $p \in M$, let $L^s(p)$ be the subspace of T_pM which consists of the values $X(p)$ taken, at the point p, by the vector fields X belonging to $\mathcal{L}^s$. By Chow's Condition, at each point $p \in M$ there is a smallest integer $r = r(p)$ such that $L^{r(p)}(p) = T_pM$ (and so $\dim L^{r(p)}(p) = n$). This integer is called the *degree of nonholonomy* at p.

We say that p is a *regular point* if the sequence

$$1 \leq \dim L^1(p) \leq \cdots \leq \dim L^s(p) \leq \cdots \leq \dim L^{r(p)}(p) = n$$

remains constant in a neighborhood of p. Otherwise we say that p is a *singular point*.

Privileged coordinates. Set $n_s = \dim L^s(p)$ for $s = 1, \ldots, r$, $n_0 = 0$ and define another sequence $w_1 \leq \cdots \leq w_n$ by setting $w_j = s$ if $n_{s-1} < j \leq n_s$.

Call $X_1 f, \ldots, X_m f$ the nonholonomic partial derivatives of order 1 of f. Call further $X_i X_j f$, $X_i X_j X_k f$,... the nonholonomic derivatives of order 2, 3,... of f. If the nonholonomic derivatives of order $\leq s-1$ of f vanish at p, we say that f is *of order $\geq s$ at p*. A function f is *of order s at p* if it is of order $\geq s$ but not of order $\geq s+1$.

We say that local coordinates $(q_1, \ldots, q_n)$ centered at p are *privileged coordinates at p* if the order of q_j at p is equal to w_j, for $j = 1, \ldots, n$. Privileged coordinates allow to estimate the sub-Riemannian distance.

Ball-Box Theorem (Bellaïche [4], Gromov [9]). *The following estimate holds if and only if $(q_1, \ldots, q_n)$ form a system of privileged coordinates at p:*

there exist constants c_p, C_p and $\varepsilon_p > 0$ such that, for q in M with $d(p,q) < \varepsilon_p$,

$$c_p\big(|q_1|^{1/w_1} + \cdots + |q_n|^{1/w_n}\big) \leq d\left(0, (q_1, \ldots, q_n)\right) \leq C_p\big(|q_1|^{1/w_1} + \cdots + |q_n|^{1/w_n}\big).$$

We can then describe locally the shape of sub-Riemannian balls. For $\varepsilon < \varepsilon_p$, the ball $B(p, \varepsilon)$ looks like a box $[-\varepsilon^{w_1}, \varepsilon^{w_1}] \times \cdots \times [-\varepsilon^{w_n}, \varepsilon^{w_n}]$ in privileged coordinates. This description is not uniform. In fact the constants c_p, C_p and ε_p depend on p and they may present discontinuities. In particular, if p_N is a sequence of regular points converging to a singular point p_∞, ε_{p_N} tends to 0, though ε_{p_∞} is non-zero.

3 Shape of sub-Riemannian balls

Let (M, d) be the sub-Riemannian manifold attached to a system $X_1, \ldots, X_m$. In the following we need the notion of determinant of n vectors on TM (recall that n is the dimension of M). We then assume that M is an orientable manifold with a given volume form and denote by det the determinant n-form with respect to this volume form (see for instance [1]).

3.1 Minimal basis

Near regular points, the estimate of Ball-Box Theorem can be stated in a uniform way. It requires the construction of a system of privileged coordinates varying continuously with their base point. We will see that the basis of the tangent space involved in this construction are well determined for “free” distributions.

Let $p \in M$. We call *minimal basis at* p a family of commutators $([X_{I_1}], \dots, [X_{I_n}])$ which values at p form a basis of T_pM and such that the total length $\sum_{i=1}^n |I_i|$ equals $\sum_{i=1}^n w_i$. It implies that, up to a permutation of indices, each $|I_i|$ equals w_i.

To a family $\underline{I} = ([X_{I_1}], \dots, [X_{I_n}])$ we associate the application

$$\phi_{\underline{I}}^p \text{ (or } \phi_{\underline{I}}) : (u_1, \dots, u_n) \mapsto p \exp(u_n[X_{I_n}]) \cdots \exp(u_1[X_{I_1}]).$$

When $\underline{I}$ is a minimal basis at p, $\phi_{\underline{I}}^p$ is a diffeomorphism from a neighborhood of $0 \in \mathbb{R}^n$ into a neighborhood of p in M. It defines local coordinates, called *canonical coordinates of the second kind*, which are privileged at p (Hermes [10]).

Remark. Here and in the sequel we note on the right the action of diffeomorphisms: $p \exp(tX)$ results from the action of $\exp(tX)$ on point p. This notation is consistent with the notation for Lie group (used below in the proof of Lemma 4): diffeomorphisms which come from flows of left-invariant vector fields are defined by right multiplication. It means also that the Lie bracket is defined as $[X, Y] = L_X Y = -L_Y X$ (L_X is the Lie derivative in the direction X).

Let us consider now a regular point p_0 in M and a neighborhood U of p_0 containing only regular points. Let $([X_{I_1}], \dots, [X_{I_n}])$ be a minimal basis at p_0. Reducing if necessary U, we can assume that it is a minimal basis at every point $p \in U$. Thus $\phi_{\underline{I}}^p$ is a local diffeomorphism for all point $p \in U$. This application defines at every point $p \in U$ a system of privileged coordinates varying continuously with p.

In this case the constants c_p, C_p and ε_p given in Ball-Box Theorem depend continuously on p (Bellaïche [4, p. 75]; see also Mitchell [16], Margulis and Mostow [15]). Hence we have a continuous version of Ball-Box Theorem near regular points.

Lemma 1. *Let $p_0 \in M$ be a regular point and $([X_{I_1}], \dots, [X_{I_n}])$ a minimal basis at p_0. There exist an open neighborhood U of p_0 and functions $c_1(p)$, $c_2(p)$ and $\varepsilon_1(p) > 0$ continuous on U such that, for all $p \in U$ and $\varepsilon < \varepsilon_1(p)$,*

$$B_{\underline{I}}(p, c_1(p)\varepsilon) \subset B(p, \varepsilon) \subset B_{\underline{I}}(p, c_2(p)\varepsilon)$$

where $B_{\underline{I}}(p, \varepsilon) = \left\{ p \exp(u_n[X_{I_n}]) \cdots \exp(u_1[X_{I_1}]),\ |u_i| < \varepsilon^{|I_i|},\ 1 \leq i \leq n \right\}$.

In the proof of Theorem 2 below (§ 4), we will apply this lemma to a particular class of system for which the minimal basis are easy to determine. Let us introduce some definitions.

We say that the Lie algebra $\mathfrak{g}$ is *nilpotent of step* r if $\mathfrak{g}^{(r+1)} = 0$, where $\mathfrak{g}^{(s)}$ is defined inductively $\mathfrak{g}^{(1)} = \mathfrak{g}$, $\mathfrak{g}^{(s+1)} = [\mathfrak{g}^{(s)}, \mathfrak{g}]$. In other words, denoting $A_1, \dots, A_m$

the generators of $\mathfrak{g} = \mathcal{L}(A_1, \dots, A_m)$, the nilpotency condition reads as $\mathcal{L}^{r+1} = \mathcal{L}^r$ ($\mathcal{L}^r$ is defined by (1)). Let $\mathcal{F}$ be the free Lie algebra on m generators ($\mathcal{F}$ is infinite-dimensional). Then $\mathfrak{n}_{m,r} = \mathcal{F}/\mathcal{F}^{(r+1)}$ is clearly nilpotent and is called *the free nilpotent Lie algebra of step r on m generators*. We denote by $\widetilde{n}(m,r)$ its dimension.

The vector fields $X_1, \dots, X_m$ on M are said *free up to order r at $p \in M$* if the vector space $L^r(X_1, \dots, X_m)(p)$ has the same dimension as $\mathfrak{n}_{m,r}$. Notice that it is the greatest possible dimension for this vector space (among all distributions of m vector fields and all points of M).

Now, consider a manifold $\widetilde{M}$ of dimension $\widetilde{n}(m,r)$ and vector fields $\xi_1, \dots, \xi_m$ on $\widetilde{M}$ free up to order r at $\widetilde{p} \in \widetilde{M}$. Since $\dim L^r(\xi_1, \dots, \xi_m)(\widetilde{p}) = \dim \widetilde{M}$, $\widetilde{p}$ is a regular point for the distribution $\xi_1, \dots, \xi_m$ and r is the degree of nonholonomy at $\widetilde{p}$.

The other important property of this distribution is that the minimal basis at $\widetilde{p}$ are determined only by r and m. In fact, a family $([\xi_{J_1}], \dots, [\xi_{J_{\widetilde{n}(m,r)}}])$ is a minimal basis at $\widetilde{p}$ if and only if the set of multi-indices $\mathcal{J} = \{J_1, \dots, J_{\widetilde{n}(m,r)}\}$ corresponds to a basis $[A_{J_1}], \dots, [A_{J_{\widetilde{n}(m,r)}}]$ of $\mathfrak{n}_{m,r}$ ($A_1, \dots, A_m$ denote the generators of $\mathfrak{n}_{m,r}$).

3.2 Associated basis

The notion of minimal basis is related to the one of privileged coordinates and so to the Ball-Box Theorem. We will define now a larger class of basis suited for uniform estimates.

From now on, we restrict ourselves to a compact subset Ω of M. We denote by r the maximum of the degree of nonholonomy on Ω.

Let $p \in \Omega$ and $\varepsilon > 0$. We consider the families of vector fields $([X_{I_1}], \dots, [X_{I_n}])$ such that each bracket $[X_{I_j}]$ is of length $|I_j| \leq r$. On the (finite) set of these families, we have a function

$$\left| \det \Big([X_{I_1}]\varepsilon^{|I_1|}, \dots, [X_{I_n}]\varepsilon^{|I_n|} \Big)(p) \right|. \tag{2}$$

We say that the family $\underline{I} = ([X_{I_1}], \dots, [X_{I_n}])$ is *associated with* (p, ε) on Ω if it achieves the maximum of this function. In particular the value at p of a family associated with (p, ε) forms a basis of T_pM.

We will need also a slightly different notion when the commutators are chosen in a restricted set. Let $\mathcal{J}$ be a set of multi-indices. We say that the family $\underline{I} = ([X_{I_1}], \dots, [X_{I_n}])$ is *$\mathcal{J}$-associated with* (p, ε) if each I_i belongs to $\mathcal{J}$ and if $\underline{I}$ achieves the maximum of the function (2) taken on the families with multi-indices in $\mathcal{J}$.

Let us explain how the associated families vary in function of ε. Fix $p \in \Omega$. For ε small enough, every family associated with (p, ε) is of minimal total length, that is

$$\sum_{i=1}^{n} |I_i| = \sum_{s \geq 1} s\big(\dim L^s(p) - \dim L^{s-1}(p)\big) = \sum_{i=1}^{n} w_i = D(p).$$

This implies that the family is a minimal basis at p.

Set $\sum_{i=1}^{n} w_i = D(p)$. We obtain a sequence $0 < \varepsilon_{D(p)}(p) \leq \varepsilon_{D(p)+1}(p) \leq \cdots \leq \varepsilon_N(p) \leq 1$ such that, for ε in $]\varepsilon_{k-1}(p), \varepsilon_k(p)[$, every family associated with (p, ε) has a total length $\sum_i |I_i|$ equal to k. Setting $\varepsilon_k(p) = 0$ for $k < D(p)$ we see that each $\varepsilon_k(p)$ is a continuous function of p on Ω. However the function $\varepsilon_{D(p)}(p)$ is not continuous everywhere: if p_0 is a singular point and p a regular point near p_0, we have $D(p_0) > D(p)$, so $\varepsilon_{D(p)}(p)$ tends to 0 as p tends to p_0 when $\varepsilon_{D(p_0)}(p_0)$ is nonzero.

Let us illustrate these notions with an example. Consider the Martinet distribution defined in $\mathbb{R}^3$ by the vector fields

$$X_1 = \frac{\partial}{\partial x}, \quad X_2 = \frac{\partial}{\partial y} + \frac{x^2}{2}\frac{\partial}{\partial z}.$$

The only non zero commutators are

$$X_{12} = [X_1, X_2] = x\frac{\partial}{\partial z}, \qquad X_{121} = [[X_1, X_2], X_1] = -\frac{\partial}{\partial z}.$$

Every point in the plane $\{x = 0\}$ is singular. The degree of nonholonomy is 2 at regular points and 3 at singular points. Two families of vector fields have a non identically zero determinant, $\underline{I} = (X_1, X_2, X_{12})$ and $\underline{J} = (X_1, X_2, X_{121})$. We have

$$|\det(\varepsilon X_1, \varepsilon X_2, \varepsilon^2 X_{12})(x, y, z)| = |x|\varepsilon^4, \qquad |\det(\varepsilon X_1, \varepsilon X_2, \varepsilon^3 X_{121})(x, y, z)| = \varepsilon^5.$$

Thus the families associated with $\big((x, y, z), \varepsilon\big)$ are:

$$\underline{J} \text{ if } |x| < \varepsilon, \qquad \underline{J} \text{ and } \underline{I} \text{ if } |x| = \varepsilon, \qquad \underline{I} \text{ if } |x| > \varepsilon.$$

At a singular point $p_0 = (0, y, z)$, the minimal basis is $\underline{J}$ and we have $D(p_0) = 5$ and $\varepsilon_5(p_0) = 1$. Actually $\varepsilon_5(p_0)$ should be equal to $+\infty$ but we are interested in small values of ε, so we have required $\varepsilon_k(p_0) \leq 1$. At a regular point $p = (x, y, z)$, $x \neq 0$, near p_0, the minimal basis is $\underline{I}$ and $D(p) = 4$. We have then $\varepsilon_4(p) = |x|$ and $\varepsilon_5(p) = 1$. We see that $\varepsilon_{D(p)}(p) = \varepsilon_4(p)$ tends to 0 as $|x| \to 0$, when $\varepsilon_{D(p_0)}(p_0) = \varepsilon_5(p_0) = 1$.

3.3 Uniform estimate of sub-Riemannian balls

For $p \in M$ and $\varepsilon > 0$, $B(p,\varepsilon)$ (or $B^d(p,\varepsilon)$) denotes the open ball centered at p of radius ε for the sub-Riemannian distance d. As in Lemma 1, we set also, for a family $\underline{I} = ([X_{I_1}], \dots, [X_{I_n}])$ of vector fields,

$$B_{\underline{I}}(p,\varepsilon) = \left\{ p \exp(u_n[X_{I_n}]) \cdots \exp(u_1[X_{I_1}]),\ |u_i| < \varepsilon^{|I_i|},\ 1 \leq i \leq n \right\}.$$

Theorem 2. *Let $\Omega \subset M$ be a compact set. There exist a constant $\delta_0 > 0$ and functions $k(\delta)$, $K(\delta)$, $0 < k(\delta) < K(\delta)$, with $\lim_{\delta \to 0} K(\delta) = 0$, such that:*
for every $p \in \Omega$, $\varepsilon < 1$, $\delta < \delta_0$ and every family $\underline{I}$ associated with (p,ε) on Ω,

$$B_{\underline{I}}(p, k(\delta)\varepsilon) \subset B(p, \delta\varepsilon) \subset B_{\underline{I}}(p, K(\delta)\varepsilon). \tag{3}$$

It is worth writing the more simple statement where $\delta = \delta_0$ is fixed.

Corollary 3. *Let $\Omega \subset M$ be a compact set. There exist constants c, C and $\delta_0 > 0$ such that, for every $p \in \Omega$, $\varepsilon < \delta_0$ and every family $\underline{I}$ associated with $(p, \varepsilon/\delta_0)$ on Ω,*

$$B_{\underline{I}}(p, c\varepsilon) \subset B(p,\varepsilon) \subset B_{\underline{I}}(p, C\varepsilon). \tag{4}$$

These results extend Ball-Box Theorem (page 5). Indeed, with the notations of § 3.2, for $\varepsilon < \delta_0 \varepsilon_{D(p)}(p)$, a family $\underline{I}$ associated with $(p, \varepsilon/\delta_0)$ is a minimal basis at p. It defines then canonical coordinates $(u_1, \dots, u_n)$ of the second kind on $B_{\underline{I}}(p,\varepsilon)$, which are privileged at p. For these values of ε, (4) is equivalent to the estimate of Ball-Box Theorem.

Hence the estimate of Ball-Box Theorem holds for balls of radius less than $\varepsilon(p) = \varepsilon_{D(p)}(p)$. We have seen in particular that $\varepsilon(p)$ can have no minimum on Ω, though (3) holds for radius less than δ_0, independent of p. Theorem 2 gives then a uniform estimate of sub-Riemannian balls.

In estimate (3) the family $\underline{I}$ used to construct $B_{\underline{I}}$ and the coordinates u_i depends on ε. That's why we have introduced δ: it allows to compare balls $B(p, \delta\varepsilon)$ of different radius in the same coordinates u_i (that is with the same $\underline{I}$).

4 Proof of Theorem 2

To show Theorem 2 we first construct locally an extended manifold $(\widetilde{M}, \widetilde{d})$ for which we are able to estimate the balls. We project then the estimates onto M and use the compactness to make the result global on Ω. Finally we show that the projected estimates look like boxes in suitable coordinates.

4.1 Lifting of the manifold

We can reduce locally the study to a sub-Riemannian manifold with no singular points thanks to the following lemma.

Lemma 4. *Let $p \in M$ and let r be an integer greater than, or equal to, the degree of nonholonomy r_p at p. Denote by $\widetilde{n} = \widetilde{n}(m, r)$ the dimension of the free nilpotent Lie algebra of step r on m generators and by $\widetilde{M}$ the manifold $\widetilde{M} = M \times \mathbb{R}^{\widetilde{n}-n}$. Then there exist a neighborhood $\widetilde{U} \subset \widetilde{M}$ of $(p, 0)$; a neighborhood $U \subset M$ of p, $U \times \{0\} \subset \widetilde{U}$; coordinates (y, z) on $\widetilde{U}$; and vector fields on $\widetilde{U}$*

$$\xi_i(y, z) = X_i(y) + \sum_{j=n+1}^{\widetilde{n}} b_{ij}(y, z) \frac{\partial}{\partial z_j} \tag{5}$$

such that,

- *the vector fields $\xi_1, \dots, \xi_m$ are free up to order r at every point in $\widetilde{U}$;*
- *every $\widetilde{q} \in \widetilde{U}$ is regular for the distribution $\xi_1, \dots, \xi_m$ and $L^r(\xi_1, \dots, \xi_m)(\widetilde{q}) = T_{\widetilde{q}}\widetilde{M}$;*
- *denoting π the canonical projection from $\widetilde{M}$ onto M and $\widetilde{d}$ the sub-Riemannian distance on $\widetilde{M}$ attached to $(\xi_1, \dots, \xi_m)$, we have, for $q \in U$ and $\varepsilon > 0$ such that $B^{\widetilde{d}}((q, 0), \varepsilon) \subset \widetilde{U}$,*

$$B(q, \varepsilon) = \pi\big(B^{\widetilde{d}}((q, 0), \varepsilon)\big). \tag{6}$$

Remark. As noticed in 3.1, the second point is a direct consequence of the first one.

Most of this result is originally due to Bellaïche [4, p. 49]. It is based on the lifting method introduced by Rothschild and Stein [22]. There is however a mistake in Bellaïche's proof. Our proof is then inspired of his own and include the necessary correction.

Proof. Fix $p \in M$ and consider a system of privileged coordinates $(q_1, \dots, q_n)$ at p. With these coordinates we build the nilpotent system $\widehat{X}_1, \dots, \widehat{X}_m$ approximating $X_1, \dots, X_m$ at p. Various nilpotent approximations have been used since the works of Rothschild and Stein [22] and Goodman [8] (see for instance [2, 3, 10, 21, 23]). Let us recall here the construction presented by Bellaïche [4].

Let $(w_1, \dots, w_n)$ be the orders of $(q_1, \dots, q_n)$ at p. We say that a polynomial is homogeneous of weighted degree s if it is a linear combination of monomials $q_1^{l_1} \cdots q_n^{l_n}$, with $w_1 l_1 + \cdots + w_n l_n = s$. Its order at 0 is equal to its weighted degree. Each X_i may be rewritten near p as $X_i(q) = \sum_{j=1}^n (f_{ij}(q_1, \ldots, q_{j-1}) + g_{ij}(q))\partial/\partial q_j$, where f_{ij} is an homogeneous polynomial of weighted degree $w_j - 1$ and g_{ij} a function of order greater than $w_j - 1$ at 0. We define $\widehat{X}_i$ as $\sum_{j=1}^n f_{ij}(q_1, \ldots, q_{j-1})\partial/\partial q_j$.

Thus each $\widehat{X}_i$ is a vector field on $\mathbb{R}^n$. The Lie algebra $\mathcal{L}(\widehat{X}_1, \ldots, \widehat{X}_m)$ is nilpotent of step r_p. Moreover, for $s \geq 1$, $L^s(\widehat{X}_1, \ldots, \widehat{X}_m)(0)$ and $L^s(p)$ have the same dimension.

We are going to prove the lemma in two steps. We first construct vector fields $\widehat{\xi}_i$ on $\mathbb{R}^{\tilde{n}}$ satisfying (5) with $\widehat{X}_i$ instead of X_i. Then we give the vector fields ξ_i and prove (6).

Let $\mathfrak{n}_{m,r}$ be the free nilpotent Lie algebra of step r on m generators and $A_1, \ldots, A_m$ its generators. The corresponding simply connected Lie group is $N = \exp(\mathfrak{n}_{m,r})$.

Since $r \geq r_p$, the correspondence $\varpi(A_i) = \widehat{X}_i$, $1 \leq i \leq m$, extends to a surjective Lie algebra homomorphism from $\mathfrak{n}_{m,r}$ onto $\mathcal{L}(\widehat{X}_1, \ldots, \widehat{X}_m)$: for $A = \sum_j c_j [A_{I_j}]$ in $\mathfrak{n}_{m,r}$, $\varpi(A) = \sum_j c_j [\widehat{X}_{I_j}]$. In the same way, every element $g = \exp A$ in N corresponds to a diffeomorphism $\exp(\varpi(A))$ of $\mathbb{R}^n$. We can then define a right action of N on $\mathbb{R}^n$:

$$\theta : (q, g = \exp A) \mapsto q \exp(\varpi(A)),$$

which is transitive (because $\widehat{X}_1, \ldots, \widehat{X}_m$ satisfy Chow's condition).

Denote by H the isotropy subgroup of 0, that is $H = \{g \in N,\ \theta(0, g) = 0\}$. Assigning $\theta(0, g)$ to g defines a map from N to $\mathbb{R}^n$ mapping the identity of N to 0. This map gives rise to a diffeomorphism $\phi : H \backslash N \to \mathbb{R}^n$.

Moreover H is connected and simply connected (because H is invariant by a one-parameter group of automorphisms of N, see Bellaïche [4, p. 49] for details). This implies that $H = \exp \mathfrak{h}$ where the Lie sub-algebra $\mathfrak{h} \subset \mathfrak{n}_{m,r}$ is the set of elements A in $\mathfrak{n}_{m,r}$ such that $\varpi(A)(0) = 0$, that is

$$\mathfrak{h} = \left\{ \sum_j c_j [A_{I_j}] \text{ such that } \sum_j c_j [\widehat{X}_{I_j}](0) = 0 \right\}.$$

Let us construct now coordinates on N, $H \backslash N$ and $\mathbb{R}^n$. Consider $(B_1, \ldots, B_n)$ in $\mathfrak{n}_{m,r}$ which span a supplement of $\mathfrak{h}$ in $\mathfrak{n}_{m,r}$ and complete it with a basis $(D_{n+1}, \ldots, D_{\tilde{n}})$ of $\mathfrak{h}$. Set $\widehat{Y}_i = \varpi(B_i)$, $i = 1, \ldots, n$.

We define coordinates $y \in \mathbb{R}^n$ on $H \backslash N$ and $\mathbb{R}^n$ by the diffeomorphisms:

$$y \quad \longmapsto \quad H \exp\Big(\sum_{i=1}^n y_i B_i\Big) \in H \backslash N \quad \stackrel{\phi}{\longleftrightarrow} \quad q = 0 \exp\Big(\sum_{i=1}^n y_i \widehat{Y}_i\Big) \in \mathbb{R}^n,$$

and coordinates $(y,z) \in \mathbb{R}^{\widetilde{n}}$ on N by

$$(y,z) \quad \longmapsto \quad g = \exp\big(\sum_{j=n+1}^{\widetilde{n}} z_j D_j\big)\exp\big(\sum_{i=1}^{n} y_i B_i\big). \tag{7}$$

Let $\widehat{\xi_i}$ (resp. $\widehat{\xi_i^h}$) be the element A_i of $\mathfrak{n}_{m,r}$ viewed as a left invariant vector field on N (resp. on $H\backslash N$), i.e. $\widehat{\xi_i}$ (resp. $\widehat{\xi_i^h}$) is the infinitesimal generator of $\exp(tA_i)$ on N (resp. on $H\backslash N$). Identifying coordinates y on $H\backslash N$ with coordinates y on $\mathbb{R}^n$ by ϕ, we identify $\widehat{\xi_i^h}(y)$ to $\widehat{X}_i(y)$.

Claim. *There exist smooth functions b_{ij} on $\mathbb{R}^{\widetilde{n}}$ such that*

$$\widehat{\xi_i}(y,z) = \widehat{X}_i(y) + \sum_{j=n+1}^{n'} b_{ij}(y,z)\frac{\partial}{\partial z_j}.$$

Proof of the claim. We only have to prove that $\widehat{\xi_i}$ and $\widehat{X}_i$ have the same coordinate on $\partial/\partial y_k$, for $k = 1,\ldots,n$.

The vector fields $\widehat{\xi_i}$ and $\widehat{\xi_i^h}$ are defined as

$$\widehat{\xi_i}(g) = \frac{d}{dt}\big[g\exp(tA_i)\big]\big|_{t=0}, \qquad \widehat{\xi_i^h}(Hg) = \frac{d}{dt}\big[Hg\exp(tA_i)\big]\big|_{t=0}.$$

In coordinates (y,z), we have

$$\begin{aligned} g\exp(tA_i) &= \exp\big(\sum_{j=n+1}^{\widetilde{n}} z_j D_j\big)\exp\big(\sum_{k=1}^{n} y_k B_k\big)\exp(tA_i), \\ Hg\exp(tA_i) &= H\exp\big(\sum_{k=1}^{n} y_k B_k\big)\exp(tA_i). \end{aligned}$$

The product $\exp\big(\sum_{k=1}^{n} y_k B_k\big)\exp(tA_i)$ is an element of N. Hence it can be written as $h(y,t)g(y,t)$ with $h(y,t) \in H$ and

$$g(y,t) = \exp\big(\sum_{k=1}^{n} c_k(y,t)B_k\big).$$

Since H is a subgroup of N, $\exp(\sum_{j=n+1}^{\widetilde{n}} z_j D_j)h(y,t)$ belongs to H and we have

$$\begin{aligned} g\exp(tA_i) &= \exp\big(\sum_{j=n+1}^{\widetilde{n}} d_j(y,z,t)D_j\big)\exp\big(\sum_{k=1}^{n} c_k(y,t)B_k\big), \\ Hg\exp(tA_i) &= H\exp\big(\sum_{k=1}^{n} c_k(y,t)B_k\big). \end{aligned}$$

The coordinate of $\widehat{\xi}_i$ on $\partial/\partial y_k$ is $b_k(y) = \frac{dc_k}{dt}(y,0)$ and is equal to the one of $\widehat{\xi}_i^h$. Since we have identified $\widehat{\xi}_i^h(y)$ to $\widehat{X}_i(y)$, it proves the claim.

Define now a lifted manifold $\widetilde{M} = M \times \mathbb{R}^{\tilde{n}-n}$. On a neighborhood $\widetilde{U}$ of $(p,0)$ in $\widetilde{M}$, we construct vector fields

$$\xi_i(y,z) = X_i(y) + \sum_j b_{ij}(y,z)\frac{\partial}{\partial z_j}$$

with the same b_{ij} as in the expression of $\widehat{\xi}_i$. Remark that, by means of the coordinates $(q_1,\dots,q_n)$, we have identified a neighborhood of p in M with a neighborhood of 0 in $\mathbb{R}^n$. The coordinates y can then be considered as coordinates on M near p.

By construction $\widehat{\xi}_1,\dots,\widehat{\xi}_m$ is the nilpotent approximation of $\xi_1,\dots,\xi_m$ at $(p,0)$. Therefore $\xi_1,\dots,\xi_m$ are free up to order r at $(p,0)$. Reducing if necessary $\widetilde{U}$ we can assume that this property holds on the whole of $\widetilde{U}$. It implies in particular that every point $\widetilde{q} \in \widetilde{U}$ is regular for the sub-Riemannian structure defined on $\widetilde{M}$ by $\xi_1,\dots,\xi_m$ and that $L^r(\xi_1,\dots,\xi_m)(\widetilde{q}) = T_{\widetilde{q}}\widetilde{M}$. This shows the first two points of Lemma 4.

Denote by π the canonical projection from $\widetilde{M}$ onto M and by $\widetilde{d}$ the sub-Riemannian distance on $\widetilde{M}$. Consider an open set $U \ni p$ in M such that $U \times \{0\} \subset \widetilde{U}$. Let $q \in U$ and $\varepsilon > 0$ be such that $B^{\widetilde{d}}((q,0),\varepsilon) \subset \widetilde{U}$. The ball $B(q,\varepsilon)$ (resp. $B^{\widetilde{d}}((q,0),\varepsilon)$) is the set of points in M (resp. $\widetilde{M}$) that can be joined to q (resp. $(q,0)$) by an horizontal path of the X_is (resp. ξ_is) of length smaller than ε.

Now, due to the form of ξ_i, every horizontal path of the X_is going through q and with a length smaller than ε is the projection of an horizontal path of the ξ_is going through $(q,0)$ with the same length, and vice-versa. We have then

$$B(q,\varepsilon) = \pi(B^{\widetilde{d}}((q,0),\varepsilon)).$$

This ends the proof of Lemma 4.

Remark. In [4, p. 49], Bellaïche constructed vector fields $\widehat{\xi}_1,\dots,\widehat{\xi}_m$ in the same way, but instead of N, he uses the group generated by the diffeomorphisms $\exp t\widehat{X}_i$ acting on T_pM. The problem with this construction is that $(p,0)$ can be not regular for the distribution $\xi_1,\dots,\xi_m$. Indeed, a point can be singular for a system and regular for the nilpotent approximation taken at this point (for instance $0 \in \mathbb{R}^3$ for the system $X_1 = \frac{\partial}{\partial x}$, $X_2 = \frac{\partial}{\partial y} + x\frac{\partial}{\partial z}$ and $X_3 = x^{100}\frac{\partial}{\partial z}$).

4.2 Projection onto the original manifold

Let us apply Lemma 4 to a point $p \in \Omega$ and to the maximum r of the degree of nonholonomy on Ω. We thus obtain a neighborhood $\widetilde{U} \subset \widetilde{M}$ of $(p, 0)$ and a distribution $\xi_1, \ldots, \xi_m$ on $\widetilde{U}$.

Choose a set of multi-indices $\mathcal{J} = \{J_1, \ldots, J_{\widetilde{n}}\}$ corresponding to a basis of $\mathfrak{n}_{m,r}$ (see § 3.1). The family $[\xi_{J_1}], \ldots, [\xi_{J_{\widetilde{n}}}]$ is a minimal basis of $T_{(p,0)}\widetilde{M}$ at $(p, 0)$. Every point in $\widetilde{U}$ is regular. Reducing $\widetilde{U}$ if necessary, we can assume that $[\xi_{J_1}], \ldots, [\xi_{J_{\widetilde{n}}}]$ is a minimal basis of $T_{\widetilde{q}}\widetilde{M}$ at every point $\widetilde{q} \in \widetilde{U}$.

We are in a position to apply Lemma 1: there exist functions $C_1(\widetilde{q})$, $C_2(\widetilde{q})$ and $\varepsilon_1(\widetilde{q}) > 0$ continuous on $\widetilde{U}$ such that, for every $\widetilde{q} \in \widetilde{U}$, if $\varepsilon < \varepsilon_1(\widetilde{q})$, then

$$\widetilde{B}_{\mathcal{J}}(\widetilde{q}, C_1(\widetilde{q})\varepsilon) \subset B^{\widetilde{d}}(\widetilde{q}, \varepsilon) \subset \widetilde{B}_{\mathcal{J}}(\widetilde{q}, C_2(\widetilde{q})\varepsilon) \tag{8}$$

where $\widetilde{B}_{\mathcal{J}}(\widetilde{q}, \varepsilon) = \left\{\widetilde{q}\exp(x_{\widetilde{n}}[\xi_{J_{\widetilde{n}}}]) \cdots \exp(x_1[\xi_{J_1}]),\ |x_i| < \varepsilon^{|J_i|},\ 1 \leq i \leq \widetilde{n}\right\}$.

There exist only a finite number of sets $\mathcal{J}$ corresponding to a basis of $\mathfrak{n}_{m,r}$. We can then assume that (8) holds on $\widetilde{U}$ with the same functions $C_1(\widetilde{q})$, $C_2(\widetilde{q})$ and $\varepsilon_1(\widetilde{q})$ for every $\mathcal{J}$.

Moreover $\widetilde{n}$ and the sets $\mathcal{J}$ corresponding to a basis of $\mathfrak{n}_{m,r}$ depend only on r and m. They are thus the same at every point p in Ω.

Projection of $\widetilde{M}$ onto M. To come back to the original manifold M, we have to project the sets $\widetilde{B}_{\mathcal{J}}(\widetilde{q}, \varepsilon)$. Using (5), we write a commutator $[\xi_J]$ as

$$[\xi_J](y, z) = [X_J](y) + \sum_j b_{Jj}(y, z)\frac{\partial}{\partial z_j}. \tag{9}$$

Now, if $\widetilde{q}' = \widetilde{q}\exp(x_{\widetilde{n}}[\xi_{J_{\widetilde{n}}}]) \cdots \exp(x_1[\xi_{J_1}])$ belongs to $\widetilde{B}_{\mathcal{J}}(\widetilde{q}, \varepsilon)$ and if $\widetilde{q} = (q, 0)$, then the projection of $\widetilde{q}'$ onto M is

$$\pi(\widetilde{q}') = q\exp(x_{\widetilde{n}}[X_{J_{\widetilde{n}}}]) \cdots \exp(x_1[X_{J_1}]).$$

Denoting by $B_{\mathcal{J}}(q, \varepsilon)$ the projection of the set $\widetilde{B}_{\mathcal{J}}((q, 0), \varepsilon)$, we have

$$B_{\mathcal{J}}(q, \varepsilon) = \left\{q\exp(x_{\widetilde{n}}[X_{J_{\widetilde{n}}}]) \cdots \exp(x_1[X_{J_1}]),\ |x_i| < \varepsilon^{|J_i|},\ 1 \leq i \leq \widetilde{n}\right\}.$$

As $U \times \{0\} \subset \widetilde{U}$, it follows from (6) and (8) that, for $q \in U$ and $\varepsilon < \varepsilon_1(q, 0)$,

$$B_{\mathcal{J}}(q, C_1(q)\varepsilon) \subset B(q, \varepsilon) \subset B_{\mathcal{J}}(q, C_2(q)\varepsilon). \tag{10}$$

For every p in Ω, the construction above gives a neighborhood U_p and continuous functions $C_1(q)$, $C_2(q)$ and $\varepsilon_1(q)$ on U_p such that inclusions (10) hold for every set $\mathcal{J}$ corresponding to a basis of $\mathfrak{n}_{m,r}$. Since Ω is compact, there exists a finite covering of Ω with compact sets $\Omega_l \subset U_{p_l}$ and constants $C_1(\Omega_l)$, $C_2(\Omega_l)$, $\varepsilon_1(\Omega_l) > 0$ such that (10) with $C_1(\Omega_l)$ and $C_2(\Omega_l)$ holds on each Ω_l for $\varepsilon < \varepsilon_1(\Omega_l)$.

Set $C_1 = \min_l C_1(\Omega_l)$, $C_2 = \max_l C_2(\Omega_l)$ and $\varepsilon_1 = \min_l \varepsilon_1(\Omega_l)$. We have obtained constants C_1, C_2 and $\varepsilon_1 > 0$ such that, for every $q \in \Omega$, $\varepsilon < \varepsilon_1$ and every set $\mathcal{J}$ corresponding to a basis of $\mathfrak{n}_{m,r}$,

$$B_{\mathcal{J}}(q, C_1\varepsilon) \subset B(q,\varepsilon) \subset B_{\mathcal{J}}(q, C_2\varepsilon). \tag{11}$$

4.3 Choice of coordinates

Consider now one of the sets $\mathcal{J}$ given above.

Lemma 5. *There exist a constant* $0 < \delta_0 < 1$ *and a function* $K'(\delta)$, $\lim_{\delta\to 0} K'(\delta) = 0$, *such that, if* $p \in \Omega$, $\varepsilon < 1$ *and if* $\underline{I}$ *is a family* $\mathcal{J}$*-associated with* (p,ε)*, then, for all* $\delta < \delta_0$,

$$B_{\underline{I}}(p, \delta\varepsilon) \subset B_{\mathcal{J}}(p, \delta\varepsilon) \subset B_{\underline{I}}(p, K'(\delta)\varepsilon).$$

This lemma allows to achieve the proof of Theorem 2. Indeed, since there is only a finite number of $\mathcal{J}$, we can assume that δ_0 and K' are the same for all these sets. But, for at least one set $\mathcal{J}$, a family $\mathcal{J}$-associated with (p,ε) is also associated with (p,ε) (associated and $\mathcal{J}$-associated families are defined in § 3.2). By using (11) we then obtain the estimate (3) on the whole Ω, that is Theorem 2.

Proof (of Lemma 5). One of the inclusions is obvious: since $\underline{I}$ is a $\mathcal{J}$-family, for any radius R we have

$$B_{\underline{I}}(p, R) \subset B_{\mathcal{J}}(p, R),$$

whenever these sets are defined.

To obtain the reverse inclusion, we need some intermediate result. Let us first introduce some notations. For $\varepsilon > 0$, we denote by $\mathrm{Box}(\varepsilon)$ the set $\{|u_i| < \varepsilon^{|I_i|},\ 1 \leq i \leq n\}$ in $\mathbb{R}^n$. Given $p \in M$ and ε small enough, $B_{\underline{I}}(p,\varepsilon)$ is then the image of $\mathrm{Box}(\varepsilon)$ by the application $\phi_{\underline{I}}(u) = p\exp(u_n[X_{I_n}]) \cdots \exp(u_1[X_{I_1}])$.

Proposition 6. *There exist constants* $0 < \delta_1 < 1$ *and* $C > 0$ *such that, if* $p \in \Omega$, $\varepsilon \leq 1$ *and if* $\underline{I}$ *is a family* $\mathcal{J}$*-associated with* (p,ε)*, then the following properties are satisfied.*

(i) $\phi_{\underline{I}}$ *is a local diffeomorphism in a neighborhood of every point of* $\mathrm{Box}(\delta_1\varepsilon)$.

(ii) *At every* $q \in B_{\underline{I}}(p,\delta_1\varepsilon)$*, the values of* $[X_{I_1}],\dots,[X_{I_n}]$ *form a basis of* T_qM *and for every* $J \in \mathcal{J}$
$$\left|X_J^i(q)\right| \le C\varepsilon^{|I_i|-|J|}$$
where X_J^i *is defined by* $[X_J](q) = \sum_{i=1}^n X_J^i(q)\,[X_{I_i}](q)$.

(iii) *Denoting by* $(\psi_1,\dots,\psi_n)$ *the local inverse application of* $\phi_{\underline{I}}$*, we have, for every* $q \in B_{\underline{I}}(p,\delta_1\varepsilon)$*,*
$$\left|[X_{I_j}]\cdot\psi_i(q)\right| \le 2\varepsilon^{|I_i|-|I_j|}.$$

The proof of this proposition is long and technical. It is postponed to § 5.

Now, set $K'(\delta) = 2(4C\widetilde{n}n\delta)^{1/r}$ and $\delta_0 = K'^{-1}(\delta_1)$. Consider $p\in\Omega$, $\varepsilon\le 1$, $\underline{I}$ a family $\mathcal{J}$-associated with (p,ε) and $\delta\le\delta_0$. We will show that, with this choice of $K'(\delta)$, $B_{\mathcal{J}}(p,\delta\varepsilon)$ is included in $B_{\underline{I}}\big(p,K'(\delta)\varepsilon\big)$.

Let q be a point in $B_{\mathcal{J}}(p,\delta\varepsilon)$. By definition $q = p\exp(x_{\widetilde{n}}[X_{J_{\widetilde{n}}}])\cdots\exp(x_1[X_{J_1}])$ with $|x_i| < (\delta\varepsilon)^{|J_i|}$. We define an absolutely continuous path $\gamma:[0,\widetilde{n}]\to\Omega$ by $\gamma(0)=p$ and
$$\dot\gamma(t) = x_i[X_{J_i}]\big(\gamma(t)\big) \quad\text{when}\quad t\in[\widetilde{n}-i,\widetilde{n}-i+1[.$$
In other words we have
$$\gamma(t) = p\exp(x_{\widetilde{n}}[X_{J_{\widetilde{n}}}])\cdots\exp((t-\widetilde{n}+i)x_i[X_{J_i}])\text{ when }t\in[\widetilde{n}-i,\widetilde{n}-i+1[.$$
This path γ verifies $\gamma(\widetilde{n}) = q$ and we can assume that it is one-to-one (changing if necessary $(x_1,\dots,x_{\widetilde{n}})$).

From Proposition 6, $\phi_{\underline{I}}$ is a local diffeomorphism on $\mathrm{Box}(\delta_1\varepsilon)$. Therefore, for $t_0\in[0,\widetilde{n}]$ small enough, there exists a unique absolutely continuous application $\theta: [0,t_0]\to\mathbb{R}^n$ such that, if $t\in[0,t_0]$, $\theta(t)\in\overline{\mathrm{Box}}\big(\frac{1}{2}K'(\delta)\varepsilon\big)$ and $\phi_{\underline{I}}(\theta(t)) = \gamma(t)$. We take the closure $\overline{\mathrm{Box}}$ of Box to have a compact set. This set is included in $\mathrm{Box}(\delta_1\varepsilon)$ since we have chosen $K'(\delta_0)=\delta_1$.

Let T be the least upper bound of t_0 such that θ does exist. We want to show that $T=\widetilde{n}$, that is $\gamma(\widetilde{n}) = q$ belongs to $B_{\underline{I}}\big(p,K'(\delta)\varepsilon\big)$.

Notice first that γ and θ are one-to-one on $[0,T]$ and that $\phi_{\underline{I}}$ is a local diffeomorphism on $\mathrm{Box}(\delta_1\varepsilon)$. Therefore $\phi_{\underline{I}}$ is a global diffeomorphism on a neighborhood of $\theta([0,T])$ and its inverse application $\phi_{\underline{I}}^{-1} = (\psi_1,\dots,\psi_n)$ is well defined on this neighborhood.

Suppose $T < \widetilde{n}$. This implies that $\theta(T)$ belongs to the boundary of $\mathrm{Box}(\frac{1}{2}K'(\delta)\varepsilon)$. Hence there exists j such that $|\psi_j(\gamma(T))| = \left(\frac{1}{2}K'(\delta)\varepsilon\right)^{|I_j|}$.

On the other hand we can compute $\psi_j(\gamma(T))$ as an integral.

$$\psi_j(\gamma(T)) = \int_0^T \frac{d}{dt}\psi_j(\gamma(t))dt = \int_0^T \dot{\gamma}(t)\cdot\psi_j(\gamma(t))dt. \tag{12}$$

But for each t there is one i such that

$$\dot{\gamma}(t) = x_i[X_{J_i}](\gamma(t)) = \sum_{l=1}^{n} x_i X_{J_i}^l(\gamma(t))[X_{I_l}](\gamma(t)).$$

Thanks to Proposition 6, we can bound $X_{J_i}^l$ and $[X_{I_l}]\cdot\psi_j$:

$$\left|\dot{\gamma}(t)\cdot\psi_j(\gamma(t))\right| \leq \sum_{l=1}^{n} 2C\,|x_i|\,\varepsilon^{|I_l|-|J_i|}\,\varepsilon^{|I_j|-|I_l|}.$$

Since $|x_i| \leq (\delta\varepsilon)^{|J_i|}$, we obtain $|\dot{\gamma}(t)\cdot\psi_j(\gamma(t))| \leq 2Cn\delta^{|J_i|}\varepsilon^{|I_j|}$. This yields to

$$\left|\psi_j(\gamma(T))\right| \leq 2Cn\widetilde{n}\delta\varepsilon^{|I_j|} = \frac{1}{2}\left(\frac{1}{2}K'(\delta)\right)^r\varepsilon^{|I_j|}.$$

Since $|I_j| \leq r$, we have proven $\left|\psi_j(\gamma(T))\right| < \left(\frac{1}{2}K'(\delta)\varepsilon\right)^{|I_j|}$. This is in contradiction with $\psi_j(\gamma(T)) = \left(\frac{1}{2}K'(\delta)\varepsilon\right)^{|I_j|}$.

We have then $T = \widetilde{n}$, which ends the proof.

5 Estimations of some determinants

This part is devoted to the proof of Proposition 6, the technical result used above when proving Lemma 5. This proposition, as well as certain consequences of its proof, have their own interest: in particular they will be used in a forthcoming paper (see also [11]). We group then together these results in Lemma 7, that we prove in this section. Proposition 6 appears as a corollary.

This lemma comes essentially from a result of Nagel, Stein and Wainger [18]. Our proof follows the same scheme but provides some simplifications.

5.1 Notations

Let us introduce first some notations (and recall some defined previously). For a family $\underline{I}$ of n brackets $([X_{I_1}],\dots,[X_{I_n}])$, we denote by $D(\underline{I}) = |I_1|+\cdots+|I_n|$ the total length and by $\Delta_{\underline{I}}(q) = \det([X_{I_1}],\dots,[X_{I_n}])(q)$ the determinant.

Given $p \in M$, the application $\phi_{\underline{I}}$ from an open neighborhood of $0 \in \mathbb{R}^n$ to M is defined by

$$\phi_{\underline{I}}(u) = p \exp(u_n[X_{I_n}]) \cdots \exp(u_1[X_{I_1}]).$$

For $\varepsilon > 0$, $\mathrm{Box}(\varepsilon)$ denotes the set $\{|u_i| < \varepsilon^{|I_i|},\ 1 \leq i \leq n\}$ in $\mathbb{R}^n$ and its image is $B_{\underline{I}}(p,\varepsilon) = \phi_{\underline{I}}(\mathrm{Box}(\varepsilon))$.

Consider a set of multi-indices $\mathcal{J} = \{J_1, \dots, J_{\tilde{n}}\}$ corresponding to a basis of $\mathfrak{n}_{m,r}$ (see § 3.1). A family $\underline{I}$ is said $\mathcal{J}$-associated with (p,ε) if each I_i belongs to $\mathcal{J}$ and

$$|\Delta_{\underline{I}}(p)|\varepsilon^{D(\underline{I})} = \max\{|\Delta_{\underline{K}}(p)|\varepsilon^{D(\underline{K})},\ \underline{K} \text{ s. t. each } K_i \in \mathcal{J}\}.$$

Finally remark that every bracket $[X_I]$ can be written on Ω as

$$[X_I](q) = \sum_{|J_k| \leq |I|} \mu_{Ik}(q)\, [X_{J_k}](q), \tag{13}$$

where each μ_{Ik}, $k = 1, \dots, \tilde{n}$, is a C^∞ function on Ω.

Lemma 7. *Let $\Omega \subset M$ be a compact set, r the maximum of the degree of nonholonomy on Ω and $\mathcal{J}$ a set of multi-indices corresponding to a basis of $\mathfrak{n}_{m,r}$. There exist constants $0 < \delta_1 < 1$ and $C > 0$ such that, if $p \in \Omega$, $\varepsilon \leq 1$ and if $\underline{I}$ is a family $\mathcal{J}$-associated with (p,ε), then the following properties are satisfied.*

(i) *$\phi_{\underline{I}}$ is a local diffeomorphism in a neighborhood of every point of* $\mathrm{Box}(\delta_1\varepsilon)$.

(ii) *For every $q \in B_{\underline{I}}(p, \delta_1\varepsilon)$,*

$$\frac{1}{2}\big|\Delta_{\underline{I}}(p)\big| \leq \big|\Delta_{\underline{I}}(q)\big| \leq 2\big|\Delta_{\underline{I}}(p)\big|,$$
$$|\Delta_{\underline{I}}(q)|\varepsilon^{D(\underline{I})} \geq \frac{1}{C}\max\{|\Delta_{\underline{K}}(q)|\varepsilon^{D(\underline{K})},\ \underline{K} \text{ s. t. each } K_i \in \mathcal{J}\},$$

and for every $J \in \mathcal{J}$

$$\big|X_J^i(q)\big| \leq C\varepsilon^{|I_i|-|J|}$$

where X_J^i is defined by $[X_J](q) = \sum_{i=1}^n X_J^i(q)\,[X_{I_i}](q)$.

(iii) *Denoting by $(\psi_1, \dots, \psi_n)$ the local inverse application of $\phi_{\underline{I}}$, we have, for every $q \in B_{\underline{I}}(p, \delta_1\varepsilon)$,*

$$\begin{aligned} \big|[X_{I_j}] \cdot \psi_i(q)\big| &\leq 2\varepsilon^{|I_i|-|I_j|} \\ \big|[X_{I_i}] \cdot \psi_i(q)\big| &\geq \frac{1}{2} \end{aligned}$$

Moreover, given $\tau > 0$, there exists $\delta(\tau) \in]0, \delta_1]$ independent of p, ε and $\underline{I}$ such that, if $q \in B_{\underline{I}}(p, \delta(\tau)\varepsilon)$, then

$$\left|[X_{I_j}] \cdot \psi_i(q)\right| \leq \tau \varepsilon^{|I_i|-|I_j|} \qquad \text{if } j \neq i.$$

Remark. Proposition 6 is a direct consequence of this lemma.

This lemma contains only estimates of determinants and of ratio of determinants. Its proof is organized as follow. Preliminaries computations in § 5.2 give Taylor expansion of some functions in exponential coordinates. We then estimate the determinants $\Delta_{\underline{I}}$ in § 5.3. With these estimates we show in § 5.4 that, if $\underline{I}$ is $\mathcal{J}$-associated with (p, ε), then the Jacobian of $\phi_{\underline{I}}$ is nonzero on a set $B_{\underline{I}}(p, k\varepsilon)$. Finally we use all these results to prove Lemma 7.

5.2 Preliminary results

Consider a point $p \in M$, a family of brackets $\underline{I} = ([X_{I_1}], \ldots, [X_{I_n}])$ and the application $\phi_{\underline{I}}(u) = p \exp(u_n[X_{I_n}]) \cdots \exp(u_1[X_{I_1}])$.

Proposition 8. *The Jacobian of $\phi_{\underline{I}}$ is*

$$J\phi_{\underline{I}}(u) = \det\left(\frac{\partial \phi_{\underline{I}}}{\partial u_1}(u), \ldots, \frac{\partial \phi_{\underline{I}}}{\partial u_n}(u)\right)$$

with, for $i = 1, \ldots, n$,

$$\frac{\partial \phi_{\underline{I}}}{\partial u_i}(u) = \sum_{\alpha_1, \ldots, \alpha_{i-1} \geq 0} \frac{u_1^{\alpha_1}}{\alpha_1 !} \cdots \frac{u_{i-1}^{\alpha_{i-1}}}{\alpha_{i-1}!} \big[[X_{I_i}] \; \mathrm{ad}^{\alpha_{i-1}}[X_{I_{i-1}}] \cdots \mathrm{ad}^{\alpha_1}[X_{I_1}]\big](\phi_{\underline{I}}(u)). \qquad (14)$$

Proof. Each partial derivative of $\phi_{\underline{I}}$ is written as

$$\frac{\partial \phi_{\underline{I}}}{\partial u_i}(u) = \frac{d}{dt}\Big(p \exp\big(u_n[X_{I_n}]\big) \cdots \exp\big((u_i + t)[X_{I_i}]\big) \cdots \exp\big(u_1[X_{I_1}]\big)\Big)\Big|_{t=0}. \qquad (15)$$

Let us recall first a classical result on Lie algebras. Set $[Y \,\mathrm{ad}\, Z] = [Y, Z]$ and $[Y \exp(\mathrm{ad}\, Z)] = \sum_{i \geq 0} \frac{1}{i!}[Y \mathrm{ad}^i Z]$. We have (see for instance Bourbaki [5])

$$\exp(Y)\exp(Z) = \exp(Z)\exp([Y \exp(\mathrm{ad}\, Z)]).$$

Using this relation in 15, we "push" the term depending on t to the right of the exponential product, that is

$$\begin{aligned}\frac{\partial \phi_{\underline{I}}}{\partial u_i}(u) &= \frac{d}{dt}\Big(\phi_{\underline{I}}(u) \exp\big(t \,\big[\big[[X_{I_i}] \exp(\mathrm{ad}\, u_{i-1}[X_{I_{i-1}}])\big] \cdots \exp(\mathrm{ad}\, u_1[X_{I_1}])\big]\big)\Big)\Big|_{t=0} \\ &= \big[\big[[X_{I_i}] \exp(\mathrm{ad}\, u_{i-1}[X_{I_{i-1}}])\big] \cdots \exp(\mathrm{ad}\, u_1[X_{I_1}])\big](\phi_{\underline{I}}(u)).\end{aligned}$$

By using the definition of $[Y \exp(\mathrm{ad}\, Z)]$ we obtain the required expression.

Proposition 9. *If f is a C^∞ real-valued function defined on a neighborhood $U \subset M$ of p, then the Taylor expansion at 0 of $F = f \circ \phi_{\underline{I}}$, is*

$$f(p) + \sum_{\substack{1 \le s \le |\underline{\alpha}| \\ \sum_j |K_j| = \langle \underline{\alpha}, \underline{I} \rangle}} \lambda_{\underline{\alpha}} \, u^{\underline{\alpha}} \, [X_{K_1}] \cdots [X_{K_s}] \cdot f(p)$$

where the $\lambda_{\underline{\alpha}}$ are constants and $\langle \underline{\alpha}, \underline{I} \rangle = \alpha_1 |I_1| + \cdots + \alpha_n |I_n|$.

Proof. The Taylor expansion of $F(u_1, \ldots, u_n)$ at 0 is

$$\sum_{|\underline{\alpha}|=0}^{\infty} \frac{u^{\underline{\alpha}}}{\underline{\alpha}!} \frac{\partial^{|\underline{\alpha}|} F}{\partial u^{\underline{\alpha}}}(0).$$

Here we use classical notations for multi-indices: $\underline{\alpha} = (\alpha_1, \ldots, \alpha_n)$, $|\underline{\alpha}| = \alpha_1 + \cdots + \alpha_n$, $\underline{\alpha}! = \alpha_1! \ldots \alpha_n!$ and $u^{\underline{\alpha}} = u_1^{\alpha_1} \ldots u_n^{\alpha_n}$.

Let us compute the partial derivatives of F at $u = (u_1, \ldots, u_n)$. For $|\underline{\alpha}| = 1$, we have

$$\frac{\partial F}{\partial u_i}(u) = T_{\phi_{\underline{I}}(u)} f\Big(\frac{\partial \phi_{\underline{I}}}{\partial u_i}\Big) = \frac{\partial \phi_{\underline{I}}}{\partial u_i} \cdot f(\phi_{\underline{I}}(u)).$$

By using (14) we obtain

$$\frac{\partial F}{\partial u_i}(u) = \sum_{\nu_1, \ldots, \nu_{i-1} \ge 0} \frac{u_1^{\nu_1}}{\nu_1!} \cdots \frac{u_{i-1}^{\nu_{i-1}}}{\nu_{i-1}!} \, [[X_{I_i}] \, \mathrm{ad}^{\nu_{i-1}}[X_{I_{i-1}}] \cdots \mathrm{ad}^{\nu_1}[X_{I_1}]] \cdot f(\phi_{\underline{I}}(u)).$$

We introduce the notations $\underline{\nu} = (\nu_1, \ldots, \nu_n)$ and $\langle \underline{\nu}, \underline{I} \rangle = \sum_{i=1}^n \nu_i |I_i|$. Hence $\frac{\partial F}{\partial u_i}(u)$ is a linear combination of terms $u^{\underline{\nu}}[X_K] \cdot f(\phi_{\underline{I}}(u))$, with $|K| \ge |I_i|$ and $\langle \underline{\nu}, \underline{I} \rangle = |K| - |I_i|$.

To compute order 2 partial derivatives of F we differentiate each one of these terms with respect to u_j, $j = 1, \ldots, n$:

$$\frac{\partial}{\partial u_j}\Big(u^{\underline{\nu}}[X_K] \cdot f(\phi_{\underline{I}}(u))\Big) = \Big(\frac{\partial u^{\underline{\nu}}}{\partial u_j}\Big) [X_K] \cdot f(\phi_{\underline{I}}(u)) + u^{\underline{\nu}} \frac{\partial \phi_{\underline{I}}}{\partial u_j} \cdot [X_K] \cdot f(\phi_{\underline{I}}(u)).$$

If $\langle \underline{\nu}, \underline{I} \rangle < |I_j|$, then the first term is zero. Otherwise it is written $u^{\underline{\nu}'}[X_K] \cdot f(\phi_{\underline{I}}(u))$, with $|K| \ge |I_i| + |I_j|$ and $\langle \underline{\nu}', \underline{I} \rangle = |K| - |I_i| - |I_j|$. Using again (14) we see that the second term is a linear combination of $u^{\underline{\nu}''}[X_{K_1}] \cdot [X_{K_2}] \cdot f(\phi_{\underline{I}}(u))$, with $|K_1| + |K_2| \ge |I_i| + |I_j|$ and $\langle \underline{\nu}'', \underline{I} \rangle = |K_1| + |K_2| - |I_i| - |I_j|$.

An easy induction on $|\underline{\alpha}|$ shows that $\frac{\partial^{|\underline{\alpha}|} F}{\partial u^{\underline{\alpha}}}(u)$ is a linear combination of terms in the form $u^{\underline{\nu}}[X_{K_1}] \cdots [X_{K_s}] \cdot f(\phi_{\underline{I}}(u))$, with $s \le |\underline{\alpha}|$ and $\langle \underline{\nu}, \underline{I} \rangle = \sum_j |K_j| - \sum_i |I_{\alpha_i}| \ge 0$. Taking the values at 0 we obtain

$$\frac{\partial^{|\underline{\alpha}|} F}{\partial u^{\underline{\alpha}}}(0) = \sum_{\sum_j |K_j| = \sum_i |I_{\alpha_i}|} \lambda_{\underline{\alpha}} \, [X_{K_1}] \cdots [X_{K_s}] \cdot f(p)$$

and then the Taylor expansion of F at 0.

5.3 Determinant estimates

Let $p \in \Omega$, $\varepsilon > 0$ and let $\underline{I}$ be $\mathcal{J}$-associated with (p, ε). We are going to estimate quantities arising when computing the Jacobian of $\phi_{\underline{I}}$ on $B_{\underline{I}}(p, \varepsilon)$. According to Proposition 8, this Jacobian depends on determinants of the form $\Delta_{\underline{K}}$. In a first step we bound these determinants and their derivatives at p. We use then a Taylor expansion to give bounds at points near p.

Proposition 10. *Let N_0 be an integer. There exists a constant $C'_{N_0} > 0$ such that, for all $p \in \Omega$, $\varepsilon \in]0, 1[$ and $\underline{I}$ a family $\mathcal{J}$-associated with (p, ε), we have: if $[X_{K_1}], \ldots, [X_{K_s}]$ satisfy $\sum_i |K_i| \leq N_0$, and $\underline{J}$ is a family of n brackets of total length $D(\underline{J}) \leq N_0$, then*

$$\left|[X_{K_1}] \cdots [X_{K_s}] \cdot \Delta_{\underline{J}}(p)\right| \leq C'_{N_0} \left|\Delta_{\underline{I}}(p)\right| \varepsilon^{D(\underline{I}) - D(\underline{J}) - \sum_i |K_i|}.$$

Proof. Notice that, for a given N, there is a finite number of multi-indices K such that $|K| \leq N$. Therefore, for such multi-indices, we can find a constant greater than $|\mu_{Kk}(p)|$, for every $p \in \Omega$ and $1 \leq k \leq \widetilde{n}$ ($\mu_{Kk}(p)$ comes from (13)).

Let us consider a family $\underline{J}$ of n brackets, with $D(\underline{J}) \leq N$. Property (13) implies that, for every $p \in \Omega$,

$$|\Delta_{\underline{J}}(p)| \leq Cst \max\{|\Delta_{\underline{K}}(p)|,\ D(\underline{K}) \leq D(\underline{J}),\ \text{each } K_i \in \mathcal{J}\}.$$

Let $\underline{I}$ be a family $\mathcal{J}$-associated with (p, ε). Its definition and the fact that $\varepsilon \leq 1$ yield to

$$|\Delta_{\underline{J}}(p)| \leq Cst\, |\Delta_{\underline{I}}(p)| \varepsilon^{D(\underline{I}) - D(\underline{J})}. \tag{16}$$

We need a classic formula (see for instance [7, Ex. p. 93]): if $V, Y_1, \ldots, Y_n$ are vector fields on M, then

$$V \cdot \det(Y_1, \ldots, Y_n) = \sum_{i=1}^{n} \det(Y_1, \ldots, [V, Y_i], \ldots, Y_n) + \\ + \operatorname{div}(V)\, \det(Y_1, \ldots, Y_n). \tag{17}$$

Recall that div is defined with respect to the chosen volume form on M (see [1, p. 130]).

Let $[X_K]$ be such that $|K| \leq N_0$ and $\underline{J}$ a family of total length $D(\underline{J}) \leq N_0$. According to (17),

$$[X_K] \cdot \Delta_{\underline{J}}(p) = \operatorname{div}([X_K])\, \Delta_{\underline{J}}(p) + \sum_i \Delta_{\underline{K}^i}(p),$$

where each $D(\underline{K}^i) = D(\underline{J}) + |K| \le 2N_0$. Applying (16) (with $N = 2N_0$) to each $\underline{K}^i$ and to $\underline{J}$, we obtain

$$|[X_K] \cdot \Delta_{\underline{J}}(p)| \le Cst\ |\Delta_{\underline{I}}(p)| \varepsilon^{D(\underline{I})-D(\underline{J})-|K|}.$$

A recursion on s allows to conclude.

Proposition 11. *Let N_0 be an integer. There exist $\delta_{N_0} \in]0,1[$ and a constant C_{N_0} such that, if $p \in \Omega$, $\varepsilon \in]0,1[$ and $\underline{I}$ is $\mathcal{J}$-associated with (p,ε), we have, for all $q \in B_{\underline{I}}(p, \delta_{N_0}\varepsilon)$,*

$$\left|\Delta_{\underline{I}}(q) - \Delta_{\underline{I}}(p)\right| \le \frac{1}{2}\left|\Delta_{\underline{I}}(p)\right|,$$

and for every family $\underline{J}$ of n brackets such that $D(\underline{J}) \le N_0$,

$$\left|\Delta_{\underline{J}}(q)\right| \le C_{N_0}\left|\Delta_{\underline{I}}(p)\right| \varepsilon^{D(\underline{I})-D(\underline{J})}.$$

Proof. The proof is based on Taylor expansions given in Proposition 9. Given a family $\underline{I}$ of n brackets, we define

$$F_p(u_1, \ldots, u_n) = \Delta_{\underline{I}}\Big(p \exp\big(u_n[X_{J_n}]\big) \cdots \exp\big(u_1[X_{J_1}]\big)\Big).$$

It is a smooth function on a neighborhood of 0 in $\mathbb{R}^n$ – say $\{(u_1, \ldots, u_n),\ \|u\| < \delta\}$ – and depends continuously on $p \in \Omega$. Hence, for a given N, there is a constant c_N such that, for $p \in \Omega$ and $\|u\| < \delta$,

$$\left| F_p(u) - \sum_{|\underline{\alpha}|=0}^{N} \frac{u^{\underline{\alpha}}}{\underline{\alpha}!} \frac{\partial^{|\underline{\alpha}|} F_p}{\partial u^{\underline{\alpha}}}(0) \right| \le c_N \|u\|^{N+1} \tag{18}$$

($\|u\|$ denotes here the Euclidean norm on $\mathbb{R}^n$).

The idea is to fix N such that the right hand side of (18) is less than $1/4\Delta_{\underline{I}}(p)$. We use then Proposition 9 to compute the sum in the left hand side and Proposition 10 to bound it by $1/4\Delta_{\underline{I}}(p)$.

Let $\underline{I}$ be a family $\mathcal{J}$-associated with (p,ε). We set $N = \max|D(\underline{J}) - D(\underline{K})|$ and $\eta = \inf_{q\in\Omega} \max|\Delta_{\underline{J}}(q)|$ (the maxima are taken on all families $\underline{J}$ and $\underline{K}$ with elements in $\mathcal{J}$). The real number η is nonzero since $[X_{J_1}](q), \ldots, [X_{J_{\tilde{n}}}](q)$ generate T_qM at every $q \in \Omega$. It follows from the definition of a $\mathcal{J}$-associated family that $\Delta_{\underline{I}}(p) \ge \eta\varepsilon^N$. So there is a constant $\delta > 0$ independent on p, ε and $\underline{I}$, such that, if $\|u\| \le \delta\varepsilon$, then

$$c_N \|u\|^{N+1} \le \frac{1}{4}\Delta_{\underline{I}}(p).$$

This inequality holds in particular when $|u_j| \leq (\delta\varepsilon)^{|I_j|}$ for each I_j, that is when the point $p \exp(u_n[X_{I_n}]) \cdots \exp(u_1[X_{I_1}])$ belongs to $B_{\underline{I}}(p, \delta\varepsilon)$.

Using this inequality in (18) we obtain, if $|u_j| \leq (\delta\varepsilon)^{|I_j|}$,

$$\left|F_p(u) - F_p(0)\right| \leq \frac{1}{4}\Delta_{\underline{I}}(p) + \left|\sum_{|\underline{\alpha}|=1}^{N} \frac{u^{\underline{\alpha}}}{\underline{\alpha}!} \frac{\partial^{|\underline{\alpha}|} F_p}{\partial u^{\underline{\alpha}}}(0)\right|.$$

But, according to Proposition 9, the above sum of partial derivatives equals

$$\sum_{\substack{1 \leq |\underline{\alpha}| \leq N \\ \sum_k |K_k| = \langle \underline{\alpha}, \underline{I} \rangle}} \lambda_{\underline{\alpha}}\, u^{\underline{\alpha}}\, [X_{K_1}] \cdots [X_{K_s}] \cdot \Delta_{\underline{I}}(p)$$

(recall that $\langle \underline{\alpha}, \underline{I} \rangle = \alpha_1|I_1| + \cdots + \alpha_n|I_n|$). Due to Proposition 10 we have, if $|u_j| \leq (\delta\varepsilon)^{|I_j|}$,

$$|u^{\underline{\alpha}}\, [X_{K_1}] \cdots [X_{K_s}] \cdot \Delta_{\underline{I}}(p)| \leq C'_{N_0} \delta^{\langle \underline{\alpha}, \underline{I} \rangle} |\Delta_{\underline{I}}(p)| \leq C'_{N_0} \delta |\Delta_{\underline{I}}(p)|.$$

To conclude we just have to choose δ_{N_0} such that the sum (on $\underline{\alpha}$ such that $1 \leq |\underline{\alpha}| \leq N$) of constants $\lambda_{\underline{\alpha}} C'_{N_0}$ is smaller than $1/(4\delta_{N_0})$.

We have shown the first inequality of the proposition. The second one can be proved in the same way.

Corollary 12. *Let N_1 be an integer and set $N_0 = N_1 + \sum_{i=1}^{\tilde{n}} |J_i|$. The constants δ_{N_0} and C_{N_0} are given by Proposition 11. If $p \in \Omega$, $\varepsilon \in]0,1[$ and if $\underline{I}$ is $\mathcal{J}$-associated with (p, ε), then, for every $q \in B_{\underline{I}}(p, \delta_{N_0}\varepsilon)$ and every $[X_J]$ of length $|J| \leq N_1$,*

$$\left|X_J^i(q)\right| \leq C_{N_0} \varepsilon^{|I_i| - |J|}$$

where X_J^i is defined by

$$[X_J](q) = \sum_{i=1}^{n} X_J^i(q)\, [X_{I_i}](q).$$

Proof. We only have to notice that

$$X_J^i(q) = \frac{\Delta_{\underline{K}}(q)}{\Delta_{\underline{I}}(q)}, \quad \text{where } \underline{K} = ([X_{I_1}], \ldots, [X_{I_{i-1}}], [X_J], [X_{I_{i+1}}], \ldots, [X_{I_n}]).$$

Proposition 11 gives then the result since $D(\underline{K}) = D(\underline{I}) - |I_i| + |J|$ is smaller than N_0.

5.4 Jacobian and inverse application

We first show that, if $\underline{I}$ is $\mathcal{J}$-associated with (p,ε), $\phi_{\underline{I}}$ is a local diffeomorphism. We give in a second step estimates for the inverse application $\phi_{\underline{I}}^{-1}$. We make use below of Formula (14) of Proposition 8 written in a slightly different form.

$$\frac{\partial \phi_{\underline{I}}}{\partial u_i}(u) = [X_{I_i}](\phi(u)) + \sum_{|\underline{\alpha}|\geq 1} \lambda'_{\underline{\alpha}}\, u^{\underline{\alpha}}\, [X_{J(\underline{\alpha})}](\phi(u)) \tag{19}$$

where each $J(\underline{\alpha})$ verifies $|J(\underline{\alpha})| = |I_i| + \langle \underline{\alpha}, \underline{I}\rangle$.

Proposition 13. *There exists a constant $\delta_2 > 0$ such that, if $p \in \Omega$, $\varepsilon \in\ ]0,1[$, $\underline{I}$ is $\mathcal{J}$-associated with (p,ε) and $\phi_{\underline{I}}$ is the application defined from $\phi_{\underline{I}}(0) = p$, then, for all $u \in \mathrm{Box}(\delta_2\varepsilon)$,*

$$\frac{1}{4}|\Delta_{\underline{I}}(p)| \leq |J\phi_{\underline{I}}(u)| \leq 2|\Delta_{\underline{I}}(p)|.$$

$\phi_{\underline{I}}$ is then a local diffeomorphism in a neighborhood of every point in $\mathrm{Box}(\delta_2\varepsilon)$.

Proof. According to Proposition 8 and Formula (19) the Jacobian of $\phi_{\underline{I}}$ at u is, setting $q = \phi_{\underline{I}}(u)$,

$$J\phi_{\underline{I}}(u) = \Delta_{\underline{I}}(q) + \sum_{|\underline{\alpha}|\geq 1} \lambda''_{\underline{\alpha}}\, u^{\underline{\alpha}} \Delta_{\underline{J}(\underline{\alpha})}(q),$$

with $D(\underline{J}(\underline{\alpha})) = D(\underline{I}) + \langle \underline{\alpha}, \underline{I}\rangle$.

Using the same reasoning as in the proof of Proposition 11, we find constants δ, N_2 and c_{N_2} independents of p, ε and $\underline{I}$ such that, for $u \in \mathrm{Box}(\delta\varepsilon)$,

$$\Big|J\phi_{\underline{I}}(u) - \Delta_{\underline{I}}(q) - \sum_{1\leq|\underline{\alpha}|<N_2} \lambda''_{\underline{\alpha}}\, u^{\underline{\alpha}} \Delta_{\underline{J}(\underline{\alpha})}(q)\Big| \leq c_{N_2}\|u\|^{N_2},$$

and $c_{N_2}\|u\|^{N_2} \leq 1/8|\Delta_{\underline{I}}(p)|$. Remind that $p = \phi_{\underline{I}}(0)$ and $J\phi_{\underline{I}}(0) = \Delta_{\underline{I}}(p)$. We write the above inequality as

$$\big|J\phi_{\underline{I}}(u) - \Delta_{\underline{I}}(p)\big| \leq \frac{1}{8}\big|\Delta_{\underline{I}}(p)\big| + \big|\Delta_{\underline{I}}(q) - \Delta_{\underline{I}}(p)\big| + \Big|\sum_{1\leq|\underline{\alpha}|<N_2} \lambda''_{\underline{\alpha}}\, u^{\underline{\alpha}}\, \Delta_{\underline{J}(\underline{\alpha})}(q)\Big|.$$

Set $N_0 = (N_2+1)\sum_{i=1}^{\widetilde{n}} |J_i|$. Reducing eventually δ we assume it is smaller than δ_{N_0} (δ_{N_0} is given by Proposition 11). Thus, if $u \in \mathrm{Box}(\delta\varepsilon)$, $q = \phi_{\underline{I}}(u)$ belongs to $B_{\underline{I}}(p, \delta_{N_0}\varepsilon)$ and we can apply the inequalities of Proposition 11: $|\Delta_{\underline{I}}(q) - \Delta_{\underline{I}}(p)| \leq$

$\frac{1}{2}|\Delta_{\underline{I}}(p)|$ and

$$\begin{aligned}
\Big|\sum_{1\le|\underline{\alpha}|<N_2} \lambda''_{\underline{\alpha}}\, u^{\underline{\alpha}}\, \Delta_{\underline{J}(\underline{\alpha})}(q)\Big| &\le \sum_{1\le|\underline{\alpha}|<N_2} |\lambda''_{\underline{\alpha}}|\, |u^{\underline{\alpha}}|\, C_{N_0}\, \varepsilon^{D(\underline{I})-D(\underline{J}(\underline{\alpha}))}\, \big|\Delta_{\underline{I}}(p)\big| \\
&\le \sum_{1\le|\underline{\alpha}|<N_2} |\lambda''_{\underline{\alpha}}|\, (\delta\varepsilon)^{\langle\underline{\alpha},\underline{I}\rangle}\, C_{N_0}\, \varepsilon^{\langle\underline{\alpha},\underline{I}\rangle}\, \big|\Delta_{\underline{I}}(p)\big| \\
&\le \big|\Delta_{\underline{I}}(p)\big|\, \delta \sum_{1\le|\underline{\alpha}|<N_2} |\lambda''_{\underline{\alpha}}|\, C_{N_0}.
\end{aligned}$$

Choose δ_2 such that $\delta_2 \sum_{1\le|\underline{\alpha}|<N_2} |\lambda''_{\underline{\alpha}}|\, C_{N_0}$ is smaller than $1/8$ (and $\delta_2 \le \delta_{N_0}$). We obtain $|J\phi_{\underline{I}}(u) - \Delta_{\underline{I}}(p)| \le 3/4|\Delta_{\underline{I}}(p)|$ and so the proposition.

For $u \in \mathrm{Box}(\delta_2\varepsilon)$, $\phi_{\underline{I}}$ has an inverse application on a neighborhood of $q = \phi_{\underline{I}}(u)$. Let us set $\phi_{\underline{I}}^{-1} = (\psi_1, \ldots, \psi_n)$ with $\psi_i(\phi_{\underline{I}}(u)) = u_i$. We give below estimations on derivatives of coordinates ψ_i.

Proposition 14. *For $\tau > 0$, there exists $\delta(\tau) \in]0, \delta_2]$ such that, for $p \in \Omega$, $\varepsilon \in]0,1[$ and $\underline{I}$ a family $\mathcal{J}$-associated with (p,ε), we have if $q \in B_{\underline{I}}(p, \delta(\tau)\varepsilon)$ and if $\phi_{\underline{I}}^{-1} = (\psi_1, \ldots, \psi_n)$ is locally the inverse of $\phi_{\underline{I}}$,*

$$\begin{aligned}
\big|[X_{I_i}]\cdot\psi_j(q)\big| &\le \tau\varepsilon^{|I_j|-|I_i|} \quad \text{if } j \ne i, \\
\frac{1}{2} \le \big|[X_{I_i}]\cdot\psi_i(q)\big| &\le 2.
\end{aligned}$$

Proof. For $q = \phi_{\underline{I}}(u)$ in $B_{\underline{I}}(p, \delta_2\varepsilon)$, we denote

$$T_q\phi_{\underline{I}}^{-1}([X_{I_i}]) = \sum_{j=1}^{n} a_{ij}(u) \frac{\partial}{\partial u_j}.$$

The component $a_{ij}(u)$ equals $[X_{I_i}]\cdot\psi_j(q)$. The formula can also be written

$$[X_{I_i}](q) = \sum_{j=1}^{n} a_{ij}(u) \frac{\partial\phi_{\underline{I}}}{\partial u_j}(u). \tag{20}$$

The proof goes as follows: let us write first the $\partial\phi_{\underline{I}}/\partial u_k$'s in the basis of the $[X_{I_l}](q)$'s thanks to (19), then inverse the coordinates matrix to obtain $a_{ij}(u)$.

We apply the reasoning of the proof of Proposition 11 to formula (19): we find constants $\delta \le \delta_2$, N_3 and c_{N_3} independent of p, ε and $\underline{I}$ such that, for $u \in \mathrm{Box}(\delta\varepsilon)$,

$$\Big\|\frac{\partial\phi_{\underline{I}}}{\partial u_k}(u) - [X_{I_k}](\phi(u)) - \sum_{1\le|\underline{\alpha}|<N_3} \lambda'_{\underline{\alpha}}\, u^{\underline{\alpha}}\, [X_{J(\underline{\alpha})}](\phi(u))\Big\| \le c_{N_3}\|u\|^{N_3} \tag{21}$$

with $|J(\underline{\alpha})| = |I_k| + \langle \underline{\alpha}, \underline{I} \rangle$. The l-th component of $[X_{J(\underline{\alpha})}](q)$ in the basis of the $[X_{I_i}](q)$'s is $X^l_{J(\underline{\alpha})}(q)$ and satisfies (see Corollary 12):

$$\left| X^l_{J(\underline{\alpha})}(q) \right| \leq C_{N_3'} \varepsilon^{|I_k| - |I_l| + \langle \underline{\alpha}, \underline{I} \rangle}$$

with $N_3' = (N_3 + 1) \sum_{i=1}^{\tilde{n}} |J_i|$.

Let $\kappa > 0$. According to the inequality just above and to (21), there exists $\delta(\kappa)$, $0 < \delta(\kappa) \leq \delta$, independent of p and ε, such that, if $q = \phi_{\underline{I}}(u)$ belongs to $B_{\underline{I}}(p, \delta\varepsilon)$,

$$\frac{\partial \phi_{\underline{I}}}{\partial u_k}(u) = [X_{I_k}](q) + \sum_{l=1}^{n} b_{kl}(u) [X_{I_l}](q)$$

with $|b_{kl}(u)| \leq \kappa \varepsilon^{|I_k| - |I_l|}$. We assume that κ is small enough for $\det \big(\mathbb{I} + (b_{kl}(u)) \big)$ to be nonzero ($\mathbb{I}$ is the identity $(n \times n)$-matrix). The inverse of this matrix is (a_{ij}) (defined by (20)), that is

$$a_{ij}(u) = (-1)^{i+j} \frac{\mathrm{minor}_{ij} \big(\mathbb{I} + \big(b_{kl}(u) \big) \big)}{\det \big(\mathbb{I} + \big(b_{kl}(u) \big) \big)}. \tag{22}$$

But we have $|b_{kl}(u)| \leq \kappa \varepsilon^{|I_k| - |I_l|}$. Therefore for a given τ we can choose κ (and so $\delta(\kappa) = \delta(\tau)$) such that, if $u \in \mathrm{Box}(\delta(\tau)\varepsilon)$,

$$\begin{aligned}
\frac{1}{\sqrt{2}} \leq \quad \left| \det \big(\mathbb{I} + \big(b_{kl}(u) \big) \big) \right| \quad &\leq \sqrt{2}, \\
\left| \mathrm{minor}_{ij} \big(\mathbb{I} + \big(b_{kl}(u) \big) \big) \right| \quad &\leq \frac{1}{\sqrt{2}} \tau \varepsilon^{|I_i| - |I_j|} \quad \text{if } i \neq j, \\
\frac{1}{\sqrt{2}} \leq \quad \left| \mathrm{minor}_{ii} \big(\mathbb{I} + \big(b_{kl}(u) \big) \big) \right| \quad &\leq \sqrt{2}.
\end{aligned}$$

Since $a_{ij}(u) = [X_{I_i}] \cdot \psi_j(\phi_{\underline{I}}(u))$, these inequalities and Equality (22) allow to conclude.

Proof of Lemma 7. We are now in a position to prove Lemma 7. Set first $N_0 = \sum_{i=1}^{\tilde{n}} |J_i|$. It results from Proposition 11 and Corollary 12 that property (ii) of Lemma 7 holds with $C = C_{N_0}$ provided $\delta_1 \leq \delta_{N_0}$.

On the other hand property (i) follows from Proposition 13 when $\delta_1 \leq \delta_2$ and property (iii) from Proposition 14 when $\delta_1 \leq \delta(2)$. Thus we end the proof by setting $\delta_1 = \min(\delta_{N_0}, \delta_2, \delta(2), 1)$.

Acknowledgements

I thank Jean-Jacques Risler for the advice and support he provided throughout the composition of this article.

References

[1] R. Abraham and J. E. Marsden. *Foundations of Mechanics*. Addison-Wesley, 1978.

[2] A. A. Agrachev, R. V. Gamkrelidze, and A. V. Sarychev. Local invariants of smooth control systems. *Acta Appl. Math.*, 14:191–237, 1989.

[3] A. A. Agrachev and A. V. Sarychev. Filtrations of a Lie algebra of vector fields and nilpotent approximations of control systems. *Dokl. Akad. Nauk SSSR*, 285:777–781, 1987. English transl.: *Soviet Math. Dokl.*, 36:104–108, 1988.

[4] A. Bellaïche. The tangent space in sub-Riemannian geometry. In A. Bellaïche and J.-J. Risler, editors, *Sub-Riemannian Geometry*, Progress in Mathematics. Birkhäuser, 1996.

[5] N. Bourbaki. *Groupes et Algèbres de Lie*. Hermann, Paris, 1972.

[6] W. L. Chow. Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung. *Math. Ann.*, 117:98–115, 1940.

[7] C. Godbillon. *Géométrie Différentielle et Mécanique Analytique*. Hermann, Paris, 1969.

[8] N. Goodman. *Nilpotent Lie groups*, volume 562 of *Springer Lecture Notes in Mathematics*. Springer Verlag, 1976.

[9] M. Gromov. Carnot-Carathéodory spaces seen from within. In A. Bellaïche and J.-J. Risler, editors, *Sub-Riemannian Geometry*, Progress in Mathematics. Birkhäuser, 1996.

[10] H. Hermes. Nilpotent and high-order approximations of vector field systems. *SIAM Review*, 33(2):238–264, 1991.

[11] F. Jean. Entropy and complexity of a path in sub-Riemannian geometry. Technical Report 331, ENSTA, July 1999.

[12] F. Jean. Complexity of nonholonomic motion planning. *International Journal of Control*, 74(8):776–782, 2001.

[13] F. Jean, J-P. Laumond, G. Oriolo, and M. Venditelli. Nonhomogeneous nilpotent approximations for systems with singularities. En préparation, 2001.

[14] I. Kupka. Géométrie sous-riemannienne. In *Séminaire N. Bourbaki*, volume 817, June 1996.

[15] G. A. Margulis and G. D. Mostow. Some remarks on the definition of tangent cones in a Carnot-Carathéodory space. *Journal d'Analyse Mathématique*, 80:299–317, 2000.

[16] J. Mitchell. On Carnot-Carathéodory metrics. *Journal of Differential Geom.*, 21:35–45, 1985.

[17] T. Nagano. Linear differential systems with singularities and an application to transitive Lie algebras. *J. Math. Soc. Japan*, 18:398–404, 1966.

[18] A. Nagel, E. M. Stein, and S. Wainger. Metrics defined by vector fields. *Acta Math.*, 155:103–147, 1985.

[19] P. Pansu. Une inégalité isopérimétrique sur le groupe d'Heisenberg. *C.R. Acad. Sci. Paris*, 295:127–130, 1982.

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

[21] C. Rockland. Intrinsic nilpotent approximations. *Acta Appl. Math.*, 8:213–270, 1987.

[22] L. P. Rothschild and E. M. Stein. Hypoelliptic differential operators and nilpotent groups. *Acta Math.*, 137:247–320, 1976.

[23] G. Stefani. On local controllability of a scalar-input system. In Lindquist Byrnes, editor, *Theory and Appl. of Nonlinear Control Syst.*, pages 167–179. North Holland, Amsterdam, 1986.

[24] H. J. Sussmann. An extension of theorem of Nagano on transitive Lie algebras. *Proc. Amer. Math. Soc.*, 45:349–356, 1974.

[25] M. Venditelli, G. Oriolo, and J-P. Laumond. Steering nonholonomic systems via nilpotent approximations: the general two-trailer system. In *IEEE Int. Conf. on Robotics and Automation*, pages 823–829, 1999.