Sub-Riemannian Geometry

Frédéric Jean*

ENSTA, Paris, France

November 25, 2003

Preliminary (incomplete) version†

Lectures given at the
Trimester on Dynamical and Control Systems
Trieste, September - December 2003

*fjean@ensta.fr

†Comments and corrections are welcome

Contents

Part I. Sub-Riemannian manifolds

1 Sub-Riemannian structures

1.1 Definition

Let M be a smooth n-dimensional manifold.

Definition. A *sub-Riemannian structure* on M is a pair (D, g) where D is a distribution and g is a Riemannian metric on D.

A *sub-Riemannian manifold* (M, D, g) is a smooth manifold M equipped with a sub-Riemannian structure (D, g).

Recall that a *distribution* D of rank m ($m \leq n$) is a family of m-dimensional linear subspaces $D_q \subset T_q M$ depending smoothly on $q \in M$. A *Riemannian metric* on D is a smooth function $g : D \to \mathbb{R}$ which restrictions g_q to D_q are positive definite quadratic forms.

Let (M, D, g) be a sub-Riemannian manifold. A *horizontal curve* $\gamma : I \subset \mathbb{R} \to M$ is an absolutely continuous curve such that $\dot\gamma(t) \in D_{\gamma(t)}$ for almost every $t \in I$. We define the *length* of a horizontal curve, as in Riemannian geometry, by:

$$\text{length}(\gamma) = \int_I \sqrt{g_{\gamma(t)}\big(\dot\gamma(t)\big)}dt.$$

Definition. The *sub-Riemannian distance* on (M, D, g) is defined by

$$d(p, q) = \inf \left\{ \text{length}(\gamma) : \begin{array}{l} \gamma \text{ horizontal curve} \\ \gamma \text{ joins } p \text{ to } q \end{array} \right\}$$

We use the convention $\inf \emptyset = +\infty$. Thus, if p and q can not be joined by a horizontal curve, $d(p, q) = +\infty$.

Local formulation Locally, on some open subset U, there exist vector fields $X_1, \dots, X_m$ which values at each point $q \in U$ form an orthonormal basis of D_q for the quadratic form g_q. We call $(X_1, \dots, X_m)$ a *local orthonormal frame* of the sub-Riemannian structure (D, g).

The horizontal curves in U coincides then with the solutions in U of the control system

$$\dot{q} = \sum_{i=1}^{m} u_i(t) X_i(q),$$

with L_1 controls $u(\cdot)$. Such control systems are called *nonholonomic control systems.*

As a trajectory, the length of a horizontal curve $q(t)$, $t \in I$, is given by

$$\text{length}\big(q(\cdot)\big) = \int_I \sqrt{u_1^2(t) + \cdots + u_m^2(t)}dt.$$

Thus, locally, to give a sub-Riemannian structure is equivalent to the given of a nonholonomic control system of constant rank (*i.e.* the rank of $X_1(q), \dots, X_m(q)$ is constant).

1.2 Generalized sub-Riemannian structures

The definition of sub-Riemannian structure does not take account of important cases like the following ones:

Riemannian metric with singularities: consider $\mathbb{R}^2 \backslash \{x = 0\}$ equipped with the Riemannian metric $g = dx^2 + \frac{1}{x^2} dy^2$ (known as Grušin plane). An orthonormal frame is $X_1 = \partial_x$, $X_2 = x\partial_y$, which is defined everywhere on $\mathbb{R}^2$ but does not generate a distribution on $\{x = 0\}$. So it does not define a sub-Riemannian structure.

General nonholonomic control systems: take a nonholonomic control system

$$\dot{q} = \sum_{i=1}^{m} u_i(t) X_i(q), \quad q \in M. \tag{1}$$

In general, the rank of $(X_1, \dots, X_m)$ is not constantly equal to m (it can even be impossible for global topological reasons, for instance in S^2).

In both cases, the problem is characterized by the given of a family of vector fields $(X_1, \dots, X_m)$ playing the role of an orthonormal frame. Let us show how we can associate a distance to such a family of vector fields.

We define first a metric. For $q \in M$ and $v \in T_qM$, set

$$g(q,v) = \inf\{u_1^2 + \cdots + u_m^2 \ : \ v = \sum_{i=1}^m u_i X_i(q)\}.$$

Clearly, $g : TM \to \mathbb{R}$ is a smooth function and

- $g(q,v) = +\infty$ if $v \notin F_q = \operatorname{span}\{X_1(q), \dots, X_m(q)\}$,
- $g|_{F_q}$ is a positive definite quadratic form.

The length of an absolutely continuous curve $\gamma(t)$, $t \in I$, is

$$\operatorname{length}(\gamma) = \int_I \sqrt{g\big(\gamma(t), \dot{\gamma}(t)\big)}dt.$$

The distance associated to $X_1, \dots, X_m$ is defined by

$$d(p,q) = \inf \operatorname{length}(\gamma)$$

where the infimum is taken over all absolutely continuous curves joining p to q.

Observe that only trajectories of the control system (1) can have finite length. The distance could then have been defined as the infimum over all the trajectories joining p to q.

This is not a sub-Riemannian distance with our definition. We give then a generalization of sub-Riemannian structures that will include this notion of distance.

Definition. A *generalized sub-Riemannian structure* on M is a triple (E, σ, g) where

- E is a vector bundle over M;
- $\sigma : E \to TM$ is a morphism of vector bundles;
- g is a Riemannian metric on E.

To a generalized sub-Riemannian structure is associated a metric defined, for $q \in M$ and $v \in T_qM$, by

$$g(q,v) = \inf\{g_q(u) \ : \ u \in E_q, \ \sigma(u) = v\}.$$

The length of absolutely continuous curves and the generalized sub-Riemannian distance are defined as above, in the same way than in Riemannian geometry.

Examples.

- Take $E = M \times \mathbb{R}^m$, $\sigma : E \to TM$, $\sigma(q,u) = \sum_{i=1}^m u_i X_i(q)$ and g the Euclidean metric on $\mathbb{R}^m$. The resulting generalized sub-Riemannian distance is the distance associated to the family $X_1, \dots, X_m$.
- Take $E = D$, $\sigma : D \hookrightarrow TM$ the inclusion, and g a Riemannian metric on D. We recover the sub-Riemannian distance associated to the sub-Riemannian structure (D, g).

A generalized sub-Riemannian structure can always be defined locally by a single finite family $X_1, \dots, X_m$ of vector fields.

In this lecture we will work in the framework of "classic" sub-Riemannian structures, which is simpler. However the results of Part II (sections 3 to 5), and some of Part III, can be easily extended for generalized sub-Riemannian structure.

2 Topological structure

Let (M, D, g) be a sub-Riemannian manifold.

Denote by D^1 the set of smooth sections of D, that is

$$D^1 = \{X \in VF(M) \ : \ X(q) \in D_q \ \forall q \in M\}.$$

In this lecture, we will identify systematically a distribution with the set of its smooth sections, in particular D with D^1. For $s \geq 1$, set $D^{s+1} = D^s + [D^1, D^s]$, where $[D^1, D^s] = \text{span}\{[X, Y] \ : \ X \in D^1, \ Y \in D^s\}$, and $\text{Lie}(D) = \bigcup_{s \geq 1} D^s$.

Locally, a local frame $X_1, \dots, X_m$ of D generates D^1. Denote by $I = i_1 \cdots i_k$ a multi-index of $\{1, \dots, m\}$, its length by $|I| = k$, and

$$X_I = [X_{i_1}, [\dots, [X_{i_{k-1}}, X_{i_k}] \dots].$$

With these notations, locally, $D^s = \text{span}\{X_I \ : \ |I| \leq s\}$.

Definition. The distribution D is *bracket generating* if, $\forall q \in M$, $\text{Lie}(D)_q = T_q M$.

Equivalently, for any $q \in M$, there exists an integer $r = r(q)$ such that $\dim D_q^r = n$.

This property is also known as Chow's Condition or Lie Algebra rank condition (LARC) in control theory, and as Hörmander condition in the context of PDE.

Theorem 1 (Chow's theorem). *If M is connected and if D is bracket-generating, then any two points of M can be joined by a horizontal curve, and so $d < \infty$.*

Proof. Let $p \in M$. Denote by $\mathcal{A}_p$ the set of points joined to p by a horizontal curve. We just have to prove that $\mathcal{A}_p$ is a neighborhood of p. Indeed, if it is this case, then the equivalence classes of the equivalence relation $d(p,q) < \infty$ are open sets. Since M is connected, there is only one such class.

Let $X_1, \ldots, X_m$ be a frame of D on a small neighborhood $U \subset M$ of p. We fix local coordinates and we identify U with a neighborhood of 0 in $\mathbb{R}^n$.

Let ϕ_t^i be the flow of the vector field X_i. The curves $t \mapsto \phi_t^i(q)$ are horizontal curves and we have

$$\phi_t^i = \mathrm{id} + tX_i + o(t).$$

Define then ϕ_t^{ij} as the commutator of flows, that is

$$\phi_t^{ij} = [\phi_t^i, \phi_t^j] = \phi_{-t}^j \circ \phi_{-t}^i \circ \phi_t^j \circ \phi_t^i.$$

It is a well-known fact that $\phi_t^{ij} = \mathrm{id} + t^2 X_{ij} + o(t^2)$.

For a multi-index $I = iJ$, we define by induction the local diffeomorphisms $\phi_t^I = [\phi_t^i, \phi_t^J]$. We have

$$\phi_t^I = \mathrm{id} + t^{|I|} X_I + o(t^{|I|}),$$

and $\phi_t^I(q)$ is the endpoint of a horizontal curve.

To obtain a diffeomorphism which derivative with respect to the time is exactly X_I, we set

$$\psi_t^I = \begin{cases} \phi^I_{t^{1/|I|}} & \text{if } t \geq 0, \\ \phi^I_{-|t|^{1/|I|}} & \text{if } t < 0 \text{ and } |I| \text{ is odd}, \\ [\phi^J_{|t|^{1/|I|}}, \phi^i_{|t|^{1/|I|}}] & \text{if } t < 0 \text{ and } |I| \text{ is even}. \end{cases}$$

We have

$$\psi_t^I = \mathrm{id} + tX_I + o(t), \tag{2}$$

and $\psi_t^I(q)$ is the endpoint of a horizontal curve.

Let us choose now commutators $X_{I_1}, \ldots, X_{I_n}$ which values at p span T_pM. It is possible since D is bracket-generating. Introduce the map φ defined on a small neighborhood Ω of 0 in $\mathbb{R}^n$ by

$$\varphi(t_1, \ldots, t_n) = \psi_{t_n}^{I_n} \circ \cdots \circ \psi_{t_1}^{I_1}(p) \in M.$$

Due to (2), this map is C^1 near 0 and its derivative at 0 is invertible. This implies that φ is a local C^1-diffeomorphism and so that $\phi(\Omega)$ contains a neighborhood of p.

Now, $\varphi(t)$ results from a concatenation of horizontal curves, the first one starting at p. Each point in $\phi(\Omega)$ is then the endpoint of a horizontal curve starting at p. Therefore $\phi(\Omega) \subset \mathcal{A}_p$ and so $\mathcal{A}_p$ is a neighborhood of p. ■

Remark. Chow's theorem appears also as a consequence of the Orbit Theorem (Sussmann, Stefan): each set $\mathcal{A}_p$ is a connected immersed submanifold of M and, at each point $q \in \mathcal{A}_p$, $\mathrm{Lie}(D)_q \subset T_q\mathcal{A}_p$. Moreover, when $\mathrm{Lie}(D)$ has constant rank on M, both spaces are equal: $\mathrm{Lie}(D)_q = T_q\mathcal{A}_p$.

Thus, when $\mathrm{Lie}(D)$ has constant rank, the bracket generating hypothesis is not restrictive: it is indeed satisfied on each $\mathcal{A}_p$ by the restricted distribution $D|_{\mathcal{A}_p}$.

This proof of Chow's theorem gives a little bit more than accessibility. Assume first that, in our construction, the vector fields $X_1, \dots, X_m$ are an orthonormal frame of (D, g).

For ε small enough, any $\phi_t^i(q)$, $0 \leq t \leq \varepsilon$ is a horizontal curve of length ε. Thus $\varphi(t_1, \dots, t_n)$ is the endpoint of a horizontal curve of length less than $N\big(|t_1|^{1/|I_1|} + \dots + |t_n|^{1/|I_n|}\big)$, where N counts the maximal number of concatenations involved in the $\psi_t^{I_i}$'s. This gives an upper bound for the distance:

$$d\big(p, \phi(t)\big) \leq N\big(|t_1|^{1/|I_1|} + \dots + |t_n|^{1/|I_n|}\big). \tag{3}$$

This kind of estimates of the distance in function of local coordinates plays an important role in sub-Riemannian geometry. However here $(t_1, \dots, t_n)$ are not local coordinates: φ is only a C^1-diffeomorphism, not a smooth diffeomorphism.

Let's try to replace $(t_1, \dots, t_n)$ by local coordinates. Choose local coordinates $(y_1, \dots, y_n)$ centered at p such that $\frac{\partial}{\partial y_i}|_p = X_{I_i}(p)$. The map $\varphi^y = y \circ \varphi$ is a C^1-diffeomorphism between neighborhoods of 0 in $\mathbb{R}^n$, and its differential at 0 is $d\varphi_0^y = \mathrm{Id}_{\mathbb{R}^n}$.

Denoting by $\|\cdot\|_{\mathbb{R}^n}$ the Euclidean norm on $\mathbb{R}^n$, we obtain, for $\|t\|_{\mathbb{R}^n}$ small enough, $y_i(t) = t_i + o(\|t\|_{\mathbb{R}^n})$. The inequality (3) becomes

$$d(p, q^y) \leq N' \|y\|_{\mathbb{R}^n}^{1/r},$$

where q^y denotes the point of coordinates y and $r = \max_i |I_i|$. This inequality allows to compare d to a Riemannian distance.

Let g^R be a Riemannian metric compatible with g, that is $g^R|_D = g$, and d_R the associate Riemannian distance. By construction $d_R(p,q) \leq d(p,q)$. Moreover, near p, $d_R(p, q^y) \geq Cst \|y\|_{\mathbb{R}^n}$. We obtain like that a first estimate of the sub-Riemannian distance.

Theorem 2. *Assume D is a bracket generating distribution. For any Riemannian metric g^R compatible with g, we have, for q close enough to p,*

$$d_R(p,q) \leq d(p,q) \leq Cst\, d_R(p,q)^{1/r}$$

where $D_p^r = T_pM$.

As a consequence, the sub-Riemannian distance is continuous.

Corollary 3. *If the distribution D is bracket generating, then the topology of the metric space (M, d) is the original topology of M.*

Remark. The converse of Chow's theorem is false in general. Consider for instance the sub-Riemannian structure in $\mathbb{R}^3$ defined by the orthonormal frame $X_1 = \partial_x$, $X_2 = \partial_y + f(x)\partial_z$ where $f(x) = e^{-1/x^2}$ for positive x and $f(x) = 0$ otherwise.

The associated sub-Riemannian distance is finite whence the distribution is not bracket generating. Moreover the topology of the metric space is different from the canonical topology of $\mathbb{R}^3$.

However, for an analytic sub-Riemannian manifold (that is when M, D, g are in the analytic category), the bracket generating property is equivalent to the finiteness of the distance.

Part II. Infinitesimal point of view

3 Nilpotent approximations

In all this section, we consider a sub-Riemannian manifold (M, D, g), with D bracket generating, and the associated sub-Riemannian distance d. We fix a point $p \in M$ and a local orthonormal frame $X_1, \dots, X_m$ of the sub-Riemannian structure (D, g) in a neighborhood of p.

3.1 Nonholonomic orders

Let f be a smooth function defined near p.

Definition. We call *nonholonomic derivatives of f of order 1* the Lie derivatives $X_1 f, \dots, X_m f$. We call further $X_i(X_j f)$, $X_i(X_j(X_k f))$,... the *nonholonomic derivatives of f of order 2, 3,...*. The nonholonomic derivative of f of order 0 is $f(p)$.

When $M = \mathbb{R}^n$, $m = n$, and $X_i = \partial_{x_i}$, nonholonomic derivatives are just the usual partial derivatives. In fact, the nonholonomic derivatives will play in sub-Riemannian geometry a role analogous to that of ∂_{x_i} in $\mathbb{R}^n$.

Proposition 4. *Let $s \geq 0$ be an integer. For a smooth function f defined near p, the following conditions are equivalent:*

(i) $f(q) = O\big(d(p,q)^s\big)$,

(ii) all nonholonomic derivatives of f of order smaller than s vanish at p.

Proof.

(i)⇒(ii) Note first that a nonholonomic derivative of f of order k can be written as

$$(X_{i_1} \dots X_{i_k} f)(p) = \frac{\partial^k}{\partial t_1 \cdots \partial t_k} f\big(e^{t_k X_{i_k}} \circ \dots \circ e^{t_1 X_{i_1}}(p)\big)\Big|_{t=0}.$$

The point $q = e^{t_k X_{i_k}} \circ \cdots \circ e^{t_1 X_{i_1}}(p)$ is the endpoint of a horizontal curve of length $|t_1|+\cdots+|t_n|$. Therefore, $d(p,q) \leq |t_1|+\cdots+|t_n|$ and, since f satisfies (i), $f(q) = O\big((|t_1|+\cdots+|t_n|)^s\big)$. This implies that, for $k < s$,

$$(X_{i_1} \dots X_{i_k} f)(p) = \frac{\partial^k}{\partial t_1 \cdots \partial t_k} f(q)\Big|_{t=0} = 0.$$

(ii)⇒(i) The proof goes by induction on s. For $s = 0$, there is nothing to prove.

Assume that (ii)⇒(i) for a given $s \geq 0$ (induction hypothesis) and take a function f such that all its nonholonomic derivatives of order $< s+1$ vanish at p.

Observe that, for $i = 1, \dots, m$, all the nonholonomic derivatives of $X_i f$ of order $< s$ vanish at p. Indeed, $X_{i_1} \dots X_{i_k}(X_i f) = X_{i_1} \dots X_{i_k} X_i f$. Applying the induction hypothesis to $X_i f$ yields $X_i f(q) = O\big(d(p,q)^s\big)$. In other terms, there exist positive constants $C_1, \dots, C_m$ such that, for q close enough to p,

$$X_i f(q) \leq C_i d(p,q)^s.$$

Fix now a point q near p. Let $\gamma(\cdot)$ be a minimizing curve joining p to q. We will see in Part III, § **??**, that such a curve exists and that we can assume it has velocity one. This means that γ satisfies

$$\dot\gamma(t) = \sum_{i=1}^m u_i(t) X_i\big(\gamma(t)\big) \quad \text{for a.e. } t \in [0,T], \qquad \gamma(0) = p,\ \gamma(T) = q,$$

with $\sum_i u_i^2(t) = 1$ a.e. and $d(p,q) = \text{length}(\gamma) = T$. Actually every sub-arc of γ is also clearly minimizing, so $d\big(p,\gamma(t)\big) = t$ for any $t \in [0,T]$.

To estimate $f\big(\gamma(T)\big)$, we compute the derivative of $f\big(\gamma(t)\big)$ with respect to t:

$$\begin{aligned} \frac{d}{dt} f\big(\gamma(t)\big) &= \sum_{i=1}^m u_i(t) X_i f\big(\gamma(t)\big), \\ \Rightarrow \left|\frac{d}{dt} f\big(\gamma(t)\big)\right| &\leq \sum_{i=1}^m |u_i(t)| C_i d\big(p,\gamma(t)\big)^s \leq C t^s, \end{aligned}$$

where $C = C_1 + \cdots + C_m$. Integrate this inequality between 0 and t:

$$\big|f\big(\gamma(t)\big)\big| \leq |f(p)| + \frac{C}{s+1} t^{s+1}.$$

We have $f(p) = 0$ since it is the nonholonomic derivative of f of order 0. Finally, at $t = T = d(p,q)$, we obtain

$$|f(q)| \leq \frac{C}{s+1} T^{s+1}.$$

$\blacksquare$

Definition. If Property (i), or (ii), holds, we say that f is *of (nonholonomic) order* $\geq s$ *at* p. If f is of order $\geq s$ but not of order $\geq s+1$, we say that f is *of order* s *at* p and we denote it as $\mathrm{ord}_p(f) = s$.

In other terms,

$$\mathrm{ord}_p(f) = \min\big\{s \in \mathbb{N} \ : \ \exists\, i_1, \dots, i_s \in \{1, \dots, m\} \ \text{ s.t. } \ (X_{i_1} \dots X_{i_s} f)(p) \neq 0\big\}.$$

Remark. The order does not depend on the chosen frame $X_1, \dots, X_m$. Actually, the order is a nonholonomic notion rather than a sub-Riemannian one: it depends only on the distribution D and not on the Riemannian metric g (exercise: prove that the orders defined by two different frames of D are the same).

Any smooth function defined near p is of order ≥ 0. Functions vanishing at p are of order ≥ 1. Moreover, we have the classical properties of orders:

$$\begin{aligned}
\mathrm{ord}_p(fg) &\geq \mathrm{ord}_p(f) + \mathrm{ord}_p(g),\\
\mathrm{ord}_p(\lambda f) &= \mathrm{ord}_p(f) \quad \text{for } \lambda \in \mathbb{R}^*,\\
\mathrm{ord}_p(f+g) &\geq \min\big(\mathrm{ord}_p(f), \mathrm{ord}_p(g)\big).
\end{aligned}$$

Notice that the first inequality is actually an equality. However the proof of this fact requires an additional result (see Proposition 5).

We define in the same way order for vector fields.

Definition. Let $\sigma \in \mathbb{Z}$. We say that a vector field X is *of order* $\geq \sigma$ *at* p if, for any smooth function f defined near p, Xf is of order $\geq \sigma + \mathrm{ord}_p(f)$. If X is of order $\geq \sigma$ but not of order $\geq \sigma + 1$ at p, we say that X is *of order* σ *at* p, and we denote it as $\mathrm{ord}_p(X) = \sigma$.

In other terms, we have $\mathrm{ord}_p(0) = +\infty$ and, for $X \neq 0$,

$$\mathrm{ord}_p(X) = \max\big\{\sigma \in \mathbb{Z} \ : \ \forall f \in C^\infty(p),\ \mathrm{ord}_p(Xf) \geq \sigma + \mathrm{ord}_p(f)\big\}.$$

The order of a differential operator is defined in exactly the same way.

The order of a function coincides with its order as a differential operator acting by multiplication. We have then the following properties:

$$\begin{aligned}
\mathrm{ord}_p([X,Y]) &\geq \mathrm{ord}_p(X) + \mathrm{ord}_p(Y),\\
\mathrm{ord}_p(fX) &\geq \mathrm{ord}_p(f) + \mathrm{ord}_p(X) \quad \text{for } f \in C^\infty(p),\\
\mathrm{ord}_p(X+Y) &\geq \min\big(\mathrm{ord}_p(X), \mathrm{ord}_p(Y)\big).
\end{aligned}$$

As already noticed for functions, the second inequality is in fact an equality. This is not the case for the first inequality (take for instance commuting nonzero vector fields).

As a consequence, $X_1, \dots, X_m$ are of order ≥ -1, $[X_i, X_j]$ of order ≥ -2, and X_I of order $\geq -|I|$.

Remark. When $M = \mathbb{R}^n$, $m = n$, and $X_i = \partial_{x_i}$, nonholonomic orders are the usual vanishing orders from analysis. At $x = 0$, it is the smallest degree of the monomials which appears with a nonzero coefficient in the Taylor series

$$f(x) \sim \sum c_\alpha x_1^{\alpha_1} \dots x_n^{\alpha_n}$$

of f at 0.

On the other hand the order of a constant differential operator is the negative of its usual order. For instance ∂_{x_i} is of nonholonomic order -1.

We are able now to precise the meaning of approximation of a family $X_1, \dots, X_m$ of vector fields.

Definition. A system of vector fields $\widehat{X}_1, \dots, \widehat{X}_m$ defined near p is called a *first-order approximation of* $X_1, \dots, X_m$ *at* p if the vector fields $X_i - \widehat{X}_i$, $i = 1, \dots, m$, are of order ≥ 0 at p.

In particular the order at p defined by the vector fields $\widehat{X}_1, \dots, \widehat{X}_m$ coincides with the one defined by $X_1, \dots, X_m$.

To go further in the characterization of orders and approximations, we need suitable systems of coordinates.

3.2 Privileged coordinates

We have introduced in § 2 the sets of vector fields D^s, defined locally as $D^s = \text{span}\{X_I \ : \ |I| \leq s\}$. Since D is a bracket generating distribution, the values of these sets at p form a flag of subspaces of T_pM

$$D_p^1 \subset D_p^2 \subset \cdots \subset D_p^r = T_pM. \tag{4}$$

Here $r = r(p)$ is the smaller integer such that $D_p^r = T_pM$. This integer is called the *degree of nonholonomy at p*.

Set $n_i(p) = \dim D_p^i$. The integer list $(n_1(p), \dots, n_r(p))$ is called the *growth vector at p*. The first integer $n_1(p) = m$ is the rank of the distribution and the last one $n_r(p) = n$ is the dimension of the manifold M.

By definition, D^1 is a distribution (recall that we identify D with D^1). For $s > 1$, the set D^s is a distribution if and only if $n_s(q)$ is constant on M. We then distinguish two kind of points.

Definition. The point p is a *regular point* if the growth vector is constant in a neighborhood of p. Otherwise, p is a *singular point*.

Thus, near a regular point, all sets D^s are locally distributions.

Some properties of growth vector and degree of nonholonomy

- The regular points form an open and dense subset of M.
- At a regular point, the growth vector is a strictly increasing sequence: $n_1(p) < \cdots < n_r(p)$. Indeed, if $n_s(q) = n_{s+1}(q)$ in a neighborhood of p, then D^s is locally an involutive distribution and so $s = r$.

  As a consequence, at a regular point, $r(p) \leq n - m + 1$.
- In a connected analytic sub-Riemannian manifold, the growth vector – and so the degree of nonholonomy – is the same at all regular points.
- The degree of nonholonomy is an upper continuous function, that is $r(q) \leq r(p)$ for q near p. In general it can be unbounded on M. Thus the finiteness of a sub-Riemannian distance is a non decidable problem: the computation of an infinite number of brackets may be needed to decide if the distribution is bracket generating.

  In the case (important in the sequel) of a distribution in $\mathbb{R}^n$ generated by polynomial vector fields, the degree of nonholonomy can be bounded by a universal function of the degree d of the polynomials:
$$r(x) \leq 2^{3n^2} n^{2n} d^{2n}.$$

The structure of the flag (4) may also be described by the sequence of integers $w_1 \leq \cdots \leq w_n$ where $w_j = s$ if $n_{s-1}(p) < j \leq n_s(p)$ (we set $n_0 = 0$), that is
$$w_1 = \cdots = w_{n_1} = 1,\ w_{n_1+1} = \cdots = w_{n_2} = 2,\ \ldots,\ w_{n_{r-1}+1} = \cdots = w_{n_r} = r.$$
The integers $w_i = w_i(p)$, $i = 1, \ldots, n$, are called the *weights at* p.

The meaning of this sequence is best understood in terms of basis of T_pM. Choose first vector fields $Y_1, \ldots, Y_{n_1}$ in D^1 which values at p form a basis of D^1_p (for instance $Y_i = X_i$). Choose then vector fields $Y_{n_1+1}, \ldots, Y_{n_2}$ in D^2 such that the values $Y_1(p), \ldots, Y_{n_2}(p)$ form a basis of D^2_p (for instance $Y_i = X_{I_i}$, with I_i of length 2). For each s, choose $Y_{n_{s-1}+1}, \ldots, Y_{n_s}$ in D^s such that $Y_1(p), \ldots, Y_{n_s}(p)$ form a basis of D^s_p (for instance $Y_i = X_{I_i}$, with $|I_i| = s$). We obtain in this way a sequence of vector fields $Y_1, \ldots, Y_n$ such that
$$\begin{cases} Y_1(p), \ldots, Y_n(p) \text{ is a basis of } T_pM, \\ Y_i \in D^{w_i},\ i = 1, \ldots, n. \end{cases} \tag{5}$$

A sequence of vector fields satisfying (5) is called an *adapted frame at p*. The word "adapted" means "adapted to the flag (4)", since the values at p of an adapted frame contain a basis $Y_1(p), \dots, Y_{n_s}(p)$ of each subspace D_p^s of the flag. The values of $Y_1, \dots, Y_n$ at q near p form also a basis of T_qM. However, this basis may be not adapted to the flag (4) at q if p is singular.

Let us relate now the weights to the orders. Write first the tangent space as a direct sum

$$T_pM = D_p^1 \oplus D_p^2/D_p^1 \oplus \cdots \oplus D_p^r/D_p^{r-1},$$

where D_p^s/D_p^{s-1} denotes a supplementary of D_p^{s-1} in D_p^s, and take a local system of coordinates $(y_1, \dots, y_n)$. The dimension of each D_p^s/D_p^{s-1} is equal to $n_s - n_{s-1}$, so we can assume that, up to a reordering, $dy_j(D_p^s/D_p^{s-1}) \neq 0$ for $n_{s-1} < j \leq n_s$.

Thus, for $0 < j \leq n_1$, we have $dy_j(D_p^1) \neq 0$. There exists then X_i such that $dy_j(X_i(p)) \neq 0$. Since $dy_j(X_i) = X_i y_j$ is a first-order nonholonomic derivative of y_j, we have $\mathrm{ord}_p(y_j) \leq 1 = w_j$.

In the same way, for $n_{s-1} < j \leq n_s$, there exists a multi-index I of length s such that $dy_j(X_I(p)) = (X_I y_j)(p) \neq 0$, and so $\mathrm{ord}_p(y_j) \leq w_j$.

To sum up, for any system of local coordinates $(y_1, \dots, y_n)$, we have, up to a reordering, $\mathrm{ord}_p(y_j) \leq w_j$ (or, without reordering, $\sum_{i=1}^n \mathrm{ord}_p(y_i) = \sum_{i=1}^n w_i$). We distinguish the coordinates with the maximal possible order.

Definition. A *system of privileged coordinates at p* is a system of local coordinates $(z_1, \dots, z_n)$ such that $\mathrm{ord}_p(z_j) = w_j$ for $j = 1, \dots, n$.

Remark. Kupka [Kup96] defines privileged functions by ***

Privileged coordinates are an essential tool to compute nonholonomic orders, characterize first-order approximations, and estimate the distance. We first show how to compute orders.

Notice that privileged coordinates $(z_1, \dots, z_n)$ satisfy

$$dz_i(D_p^{w_i}) \neq 0, \quad dz_i(D_p^{w_i - 1}) = 0, \quad i = 1, \dots, n, \tag{6}$$

or, equivalently, $\partial_{z_i}|_p \in D_p^{w_i}$ but $\notin D_p^{w_i-1}$. Local coordinates satisfying (6) are called *linearly adapted coordinates* ("adapted" because the differentials at p of the coordinates form a basis of T_p^*M dual to the values of an adapted frame).

Thus privileged coordinates are always linearly adapted coordinates. Notice that the converse is false.

Example. Take $X_1 = \partial_x$, $X_2 = \partial_y + (x^2 + y)\partial_z$ in $\mathbb{R}^3$. The weights at 0 are $(1, 1, 3)$ and (x, y, z) are adapted at 0. But they are not privileged: $(X_2^2 z)(0) = 1$.

Consider now a system of privileged coordinates $(z_1,\ldots,z_n)$. Given a sequence of integers $\alpha=(\alpha_1,\ldots,\alpha_n)$, we define the weighted degree of the monomial $z^\alpha = z_1^{\alpha_1}\cdots z_n^{\alpha_n}$ as $w(\alpha)=w_1\alpha_1+\cdots+w_n\alpha_n$ and the weighted degree of the monomial vector field $z^\alpha\partial_{z_j}$ as $w(\alpha)-w_j$. The weighted degrees allow to compute the orders of functions and vector fields in a purely algebraic way.

Proposition 5. *For a smooth function f with a Taylor expansion*

$$f(z)\sim\sum_\alpha c_\alpha z^\alpha,$$

the order of f is the least weighted degree of a monomial appearing in the Taylor series with a nonzero coefficient.

For a vector field X with a Taylor expansion

$$X(z)\sim\sum_{\alpha,j} a_{\alpha,j} z^\alpha \partial_{z_j},$$

the order of X is the least weighted degree of a monomial vector field appearing in the Taylor series with a nonzero coefficient.

Proof. For $i=1,\ldots,n$, we have $\partial_{z_i}|_p\in D_p^{w_i}$. We can then find a frame $Y_1,\ldots,Y_n$ adapted at p such that $Y_1(p)=\partial_{z_1}|_p$, …, $Y_n(p)=\partial_{z_n}|_p$. For each i, Y_i of order $\geq -w_i$ at p (since it belongs to D^{w_i}). Moreover $(Y_iz_i)(p)=1$ and $\mathrm{ord}_p(z_i)=w_i$. Thus $\mathrm{ord}_p(Y_i)=-w_i$.

Take a sequence of integers $\alpha=(\alpha_1,\ldots,\alpha_n)$. The monomial z^α is of order $\geq w(\alpha)$ at p and the differential operator $Y^\alpha=Y_1^{\alpha_1}\cdots Y_n^{\alpha_n}$ is of order $\geq -w(\alpha)$. Noticing that $(Y_iz_j)(p)=0$ if $j\neq i$, we prove easily that $(Y^\alpha z^\alpha)(p)=\frac{1}{\alpha_1!\ldots\alpha_n!}\neq 0$, which implies $\mathrm{ord}_p(z^\alpha)=w(\alpha)$.

In the same way, we obtain that, if z^α, z^β are two different monomials and λ, μ two nonzero real numbers, then $\mathrm{ord}_p(\lambda z^\alpha+\mu z^\beta)=\min\big(w(\alpha),w(\beta)\big)$. Thus the order of a series is the least weighted degree of monomials actually appearing in it. This shows the result on order of functions.

As a consequence, for any smooth function f, the order at p of $\partial_{z_i}f$ is $\geq \mathrm{ord}_p(f)-w_i$. Since moreover $\partial_{z_i}z_i=1$, we obtain $\mathrm{ord}_p(\partial_{z_i})=-w_i$ and the result on order of vector fields. ■

We see that, when using privileged coordinates, the notion of nonholonomic order amounts to the usual notion of vanishing order at some point, only assigning weights to the variables.

Homogeneity Fix a system of privileged coordinates $(z_1, \dots, z_n)$. A notion of homogeneity is naturally associated to this system. Define first the one-parameter group of dilations

$$\delta_t : (z_1, \dots, z_n) \mapsto (t^{w_1} z_1, \dots, t^{w_n} z_n), \qquad t \in \mathbb{R}.$$

The dilation δ_t acts on functions and vector fields by pull-back: $\delta_t^* f = f \circ \delta_t$ and $\delta_t^* X$, so that $(\delta_t^* X)(\delta_t^* f) = \delta_t^*(Xf)$.

Definition. A function f is *homogeneous of degree* s if $\delta_t^* f = t^s f$. A vector field X is *homogeneous of degree* σ if $\delta_t^* X = t^\sigma X$.

For a smooth function (resp. for a vector field), this is the same as being a finite sum of monomials (resp. monomial vector fields) of weighted degree s.

We introduce also the function $\|z\| = |z_1|^{1/w_1} + \cdots + |z_n|^{1/w_n}$, which is homogeneous of degree 1. This function, called *pseudo-norm*, will be of great use in the sequel. Observe already that, by definition of orders, $z_i = O\big(d(p, q^z)^{w_i}\big)$ and so

$$\|z\| = O\big(d(p, q^z)\big)$$

where q^z is the point of coordinates z.

Examples of privileged coordinates Of course all the results above on algebraic computation of orders hold only if privileged coordinates do exist. Two types of privileged coordinates are commonly used in the literature.

a. Exponential coordinates. Choose an adapted frame $Y_1, \dots, Y_n$ at p. The inverse of the local diffeomorphism

$$(z_1, \dots, z_n) \mapsto e^{z_1 Y_1 + \cdots + z_n Y_n}(p)$$

defines a system of local privileged coordinates at p, called *canonical coordinates of the first kind*. These coordinates are mainly used in the context of hypoelliptic operator and for nilpotent Lie groups with right (or left) invariant sub-Riemannian structure. The fact that these coordinates are privileged is proved – in different terms – in [RS76].

The inverse of the local diffeomorphism

$$(z_1, \dots, z_n) \mapsto e^{z_n Y_n} \circ \cdots \circ e^{z_1 Y_1}(p)$$

also defines privileged coordinates at p, called *canonical coordinates of the second kind*. They are easier to work with than the one of the first kind. For instance, in these coordinates, the vector field Y_n read as ∂_{z_n}. One can also exchange the order of the flows in the definition to obtain any of the Y_i as ∂_{z_i}. The fact that these coordinates are privileged is proved in [Her91] (see also [Mon02]).

Exercise 1. Prove that the diffeomorphism

$$(z_1, \dots, z_n) \mapsto e^{z_n Y_n + \cdots + z_{s+1} Y_{s+1}} \circ e^{z_s Y_s} \cdots \circ e^{z_1 Y_1}(p)$$

induces privileged coordinates. Show in fact that any "mix" between first and second kind canonical coordinates are privileged coordinates.

b. Algebraic algorithm. There exist also effective constructions of privileged coordinates (the construction of exponential coordinates is not effective in general since it requires to integrate flows). We present here Bellaïche's algorithm (see also [Ste86] and [AS87]).

1. Choose an adapted frame $Y_1, \dots, Y_n$ at p.
2. Choose coordinates $(y_1, \dots, y_n)$ centered at p such that $\partial_{y_i}|_p = Y_i(p)$.
3. Build privileged coordinates $z_1, \dots, z_n$ by the iterative formula: for $j = 1, \dots, n$,
$$z_j = y_j + \sum_{k=2}^{w_j - 1} h_k(y_1, \dots, y_{j-1}),$$
where, for $k = 2, \dots, w_j - 1$,
$$h_k(y_1, \dots, y_{j-1}) = - \sum_{\substack{|\alpha| = k \\ w(\alpha) < w_j}} Y_1^{\alpha_1} \dots Y_{j-1}^{\alpha_{j-1}} \Big(y_j + \sum_{q=2}^{k-1} h_q(y) \Big)(p) \, \frac{y_1^{\alpha_1}}{\alpha_1 !} \cdots \frac{y_{j-1}^{\alpha_{j-1}}}{\alpha_{j-1}!}$$
where $|\alpha| = \alpha_1 + \cdots + \alpha_n$.

The coordinates $y_1, \dots, y_n$ are linearly adapted coordinates. Starting from any system of coordinates, they can be obtained by an affine change of coordinates. Privileged coordinates are obtained from linearly adapted coordinates by expressions of the form

$$\begin{aligned} z_1 &= y_1 \\ z_2 &= y_2 + \mathrm{pol}(y_1) \\ &\vdots \\ z_n &= y_n + \mathrm{pol}(y_1, \dots, y_{n-1}) \end{aligned}$$

where pol are polynomials without constant or linear terms.

To prove that these coordinates are actually privileged, the key result is the following lemma.

Lemma 6. *A function f is of order $\geq s$ at p if and only if*

$$(Y_1^{\alpha_1} \cdots Y_n^{\alpha_n} f)(p) = 0$$

for all α such that $w(\alpha) < s$.

Remark. This lemma seems to be an easy consequence of Proposition 5 and of its proof. However, in the latter proposition, existence of privileged coordinates is assumed, whence here the aim is to prove this existence.

Roughly speaking, the idea to obtain z_j from y_j is the following. For each α with $w(\alpha) < w_j$ (and so $\alpha_j = \cdots = \alpha_n$), compute $(Y_1^{\alpha_1} \cdots Y_{j-1}^{\alpha_{j-1}} y_j)(p)$. If it is nonzero, then replace y_j by $y_j - (Y_1^{\alpha_1} \cdots Y_{j-1}^{\alpha_{j-1}} y_j)(p)\ \frac{y_1^{\alpha_1}}{\alpha_1!} \cdots \frac{y_{j-1}^{\alpha_{j-1}}}{\alpha_{j-1}!}$. With the new value of y_i, $(Y_1^{\alpha_1} \cdots Y_{j-1}^{\alpha_{j-1}} y_j)(p) = 0$.

3.3 Nilpotent approximation

Fix a system of privileged coordinates at p. We already know that each vector field X_i is of order ≥ -1. Moreover, for at least one coordinate z_j among $z_1, \ldots, z_m$, the derivative $(X_i z_j)(p)$ is nonzero (since $dz_j(D_p^1) \neq 0$). This implies that all X_i's are of order -1.

In z coordinates, X_i has a Taylor expansion

$$X_i(z) \sim \sum_{\alpha,j} a_{\alpha,j} z^\alpha \partial_{z_j},$$

where $w(\alpha) \geq w_j - 1$ if $a_{\alpha,j} \neq 0$. Grouping together the monomial vector fields of same weighted degree, we express X_i as a series

$$X_i = X_i^{(-1)} + X_i^{(0)} + X_i^{(1)} + \cdots$$

where $X_i^{(s)}$ is a homogeneous vector field of degree s.

Set $\widehat{X}_i = X_i^{(-1)}$, $i = 1, \ldots, m$. By construction, $\widehat{X}_1, \ldots, \widehat{X}_m$ is a first-order approximation of $X_1, \ldots, X_m$.

Proposition 7. *The vector fields $\widehat{X}_1, \ldots, \widehat{X}_m$ generate a nilpotent Lie algebra of step $r = w_n$.*

Proof. Note first that any homogeneous vector field of degree smaller than $-w_n$ is zero (clear in privileged coordinates). Secondly, if X and Y are homogeneous of degree respectively k and l, then the bracket $[X, Y]$ is homogeneous of degree $k+l$ because $\delta_t^*[X, Y] = [\delta_t^* X, \delta_t^* Y] = t^{k+l}[X, Y]$.

Hence, every bracket $\widehat{X}_I$ of the vector fields $\widehat{X}_1, \ldots, \widehat{X}_m$ is homogeneous of degree $-|I|$ and is zero if $|I| > w_n$. ∎

Definition. The system of vector fields $(\widehat{X}_1, \dots, \widehat{X}_m)$ is the *(homogeneous) nilpotent approximation* of $(X_1, \dots, X_m)$ at p associated to the coordinates z.

This homogeneous nilpotent approximation is not intrinsic, it depends on the chosen system of privileged coordinates. However, if $\widehat{X}_1, \dots, \widehat{X}_m$ and $\widehat{X}'_1, \dots, \widehat{X}'_m$ are the nilpotent approximations associated to two different systems of coordinates, then the Lie algebras $\mathrm{Lie}(\widehat{X}_1, \dots, \widehat{X}_m)$ and $\mathrm{Lie}(\widehat{X}'_1, \dots, \widehat{X}'_m)$ are isomorphic. If moreover p is a regular point, then $\mathrm{Lie}(\widehat{X}_1, \dots, \widehat{X}_m)$ is isomorphic to the graded nilpotent Lie algebra

$$\mathrm{Gr}(D)_p = D_p \oplus (D^2/D^1)_p \oplus \cdots \oplus (D^{r-1}/D^r)_p.$$

Notice finally that an intrinsic definition of nilpotent approximations has been recently proposed by Agrachev and Marigo [**?**].

Remark. The nilpotent approximation denotes in fact two different objects. Each $\widehat{X}_i$ can be seen as a vector field on $\mathbb{R}^n$ or as the representation in z coordinates of the vector field $z^*\widehat{X}_i$ defined on a neighborhood of p in M. This is not a problem since the nilpotent approximation is associated to a given system of privileged coordinates.

The nilpotent approximation has a particular form in privileged coordinates. Indeed, write $\widehat{X}_i = \sum_{j=1}^n f_{ij}(z)\partial_{z_j}$. Since $\widehat{X}_i$ is homogeneous of degree -1 and ∂_{z_j} of degree $-w_j$, f_{ij} is homogeneous of degree $w_j - 1$ and can not involve variables of weight greater than $w_j - 1$. Thus

$$\widehat{X}_i(z) = \sum_{j=1}^n f_{ij}(z_1, \dots, z_{n_{w_j}-1})\partial_{z_j},$$

where f_{ij} is a homogeneous polynomial of weighted degree $w_j - 1$, and

$$X_i(z) = \sum_{j=1}^n \Big(f_{ij}(z_1, \dots, z_{n_{w_j}-1}) + O\big(\|z\|^{w_j}\big) \Big)\, \partial_{z_j}.$$

The nonholonomic control system $\dot{z} = \sum_{i=1}^m u_i \widehat{X}_i(z)$ associated to the nilpotent approximation is then polynomial and in a triangular form:

$$\dot{z}_j = \sum_{i=1}^m u_i f_{ij}(z_1, \dots, z_{n_{w_j}-1}).$$

Such a form is "easy" to integrate: given the input function $(u_1(t), \dots, u_m(t))$, it is possible to compute the coordinates z_j one after the other, only by computing primitives.

As vector fields on $\mathbb{R}^n$, $\widehat{X}_1, \dots, \widehat{X}_m$ generates an homogeneous sub-Riemannian structure.

Lemma 8.

(i) $\widehat{D} = \mathrm{span}\{\widehat{X}_1, \dots, \widehat{X}_m\}$ *is a bracket generating distribution on* $\mathbb{R}^n$*.*

Denote by $(\widehat{D}, \widehat{g})$ *the sub-Riemannian structure on* $\mathbb{R}^n$ *having* $\widehat{X}_1, \dots, \widehat{X}_m$ *as global orthonormal frame, and by* $\widehat{d}$ *the associated sub-Riemannian distance on* $\mathbb{R}^n$*.*

(ii) *The distance* $\widehat{d}$ *is homogeneous of degree 1, that is*

$$\widehat{d}(\delta_t x, \delta_t y) = t\widehat{d}(x, y)$$

(iii) *There exists a constant* $C > 0$ *such that, for all* $z \in \mathbb{R}^n$*,*

$$\frac{1}{C}\|z\| \leq \widehat{d}(0, z) \leq C\|z\|.$$

Proof.

(i) It results clearly from the proof of Proposition 7 that $X_I = \widehat{X}_I +$ terms of order $> -|I|$. It implies

$$\widehat{X}_I(0) = X_I(p) \bmod \mathrm{span}\{\partial_{z_j}\big|_p \,:\, w_j < |I|\} = X_I(p) \bmod D_p^{|I|-1}.$$

Then the nilpotent approximation has the same growth vector at 0 than D at p. In particular,

$$\mathrm{rk}(\widehat{X}_1, \dots, \widehat{X}_m)(0) = m \quad \text{and} \quad \mathrm{rk}(\widehat{X}_{I_1}, \dots, \widehat{X}_{I_n})(0) = n$$

if $(X_{I_1}, \dots, X_{I_n})$ is an adapted frame at p. By continuity, both properties holds near 0.

Consider now a nonzero minor $\det\big(f_{ij}(z)\big)$, $1 \leq i \leq m$, $j = i_1, \dots, i_m$, in the matrix of the components of $\widehat{X}_1, \dots, \widehat{X}_m$. It is an homogeneous polynomial of degree $w_{i_1} + \cdots + w_{i_m} - m$. Since it is nonzero near 0, it is nonzero everywhere, and then $\mathrm{rk}(\widehat{X}_1, \dots, \widehat{X}_m) = m$ on $\mathbb{R}^n$. The same reasoning holds for $\mathrm{rk}(\widehat{X}_{I_1}, \dots, \widehat{X}_{I_n})$. This shows (i).

(ii) Observe that, if $\widehat{\gamma}$ is a horizontal curve of $\widehat{D}$, that is

$$\dot{\widehat{\gamma}}(t) = \sum_{i=1}^m u_i \widehat{X}_i\big(\widehat{\gamma}(t)\big), \quad t \in [0, T],$$

then $\delta_\lambda \widehat{\gamma}$ satisfies

$$\frac{d}{dt}\delta_\lambda \widehat{\gamma}(t) = \sum_{i=1}^m \lambda u_i \widehat{X}_i\big(\delta_\lambda \widehat{\gamma}(t)\big), \quad t \in [0, T].$$

Thus $\delta_\lambda\widehat{\gamma}$ is a horizontal curve of $\widehat{D}$ of length $\lambda \times \text{length}(\widehat{\gamma})$ with extremities $(\delta_\lambda\widehat{\gamma})(0) = \delta_\lambda(\widehat{\gamma}(0))$ and $(\delta_\lambda\widehat{\gamma})(T) = \delta_\lambda(\widehat{\gamma}(T))$. This shows the homogeneity of $\widehat{d}$.

(iii) Choose $C > 0$ such that, on the compact set $\{\|z\| = 1\}$, $1/C \leq \widehat{d}(0, z) \leq C$. Both functions $\widehat{d}(0, z)$ and $\|z\|$ being homogeneous of degree 1, the inequality follows.

■

4 Distance estimates

4.1 Ball-Box theorem

Privileged coordinates give estimates of the sub-Riemannian distance.

Theorem 9. *The following theorem holds if and only if $z_1, \ldots, z_n$ are privileged coordinates at p:*

there exist constants C_p and $\varepsilon_p > 0$ such that, if $d(p, q^z) < \varepsilon_p$, then

$$\frac{1}{C_p}\|z\| \leq d(p, q^z) \leq C_p\|z\| \tag{7}$$

(as usual, q^z denotes the point near p with coordinates z).

Corollary 10 (Ball-Box Theorem). *Expressed in a given system of privileged coordinates, the sub-Riemannian balls $B(p, \varepsilon)$ satisfy, for $\varepsilon < \varepsilon_p$,*

$$\text{Box}\big(\frac{1}{C_p}\varepsilon\big) \subset B(p, \varepsilon) \subset \text{Box}\big(C_p\varepsilon\big),$$

where $\text{Box}(\varepsilon) = [-\varepsilon^{w_1}, \varepsilon^{w_1}] \times \cdots \times [-\varepsilon^{w_n}, \varepsilon^{w_n}]$.

Remark. The constants C_p and ε_p depend on the base point p. Around a regular point p_0, it is possible to construct systems of privileged coordinates depending continuously on the base point p. In this case, the corresponding constants C_p and ε_p depend continuously on p. This is not true at a singular point. In particular, if p_0 is singular, the estimate (7) does not hold uniformly near p_0: we can not choose the constants C_p and ε_p independently on p near p_0 (see § 5.2 for uniform versions of Ball-Box Theorem).

Ball-Box Theorem is stated in different papers, often under the hypothesis that the point p is regular. As far as I know, two valid proofs exist: in [NSW85] and in [Bel96]. The result also appears without proof in [Gro96] and in [Ger84][Ger84], and with erroneous proofs in [Mit85] and in [Mon02].

We present here a proof adapted from the one of Bellaïche (our is much simpler because Bellaïche actually proves a more general result, namely (10)). Basically, the idea is to compare the distances d and $\widehat{d}$. The main step is Lemma 11 below, which is essential to explain the role of nilpotent approximations in control theory.

Observe first that, by the definition of order, a system of coordinates z is privileged if and only if $d(p, q^z) \geq Cst\,\|z\|$. The only thing to prove is that, if z are privileged coordinates, then $d(p, q^z) \leq Cst\,\|z\|$.

From now on, we work locally. We fix then a point p, a local orthonormal frame $X_1, \dots, X_m$ of the sub-Riemannian structure, and a system of privileged coordinates at p. Through these coordinates we identify a neighborhood of p in M with a neighborhood of 0 in $\mathbb{R}^n$. Finally, we denote by $\widehat{X}_1, \dots, \widehat{X}_m$ the homogeneous nilpotent approximation of $X_1, \dots, X_m$ at p (associated to the given privileged coordinates) and by $\widehat{d}$ the induced sub-Riemannian distance on $\mathbb{R}^n$. Recall also that $r = w_n$ denotes the degree of nonholonomy.

Lemma 11. *There exist constants C and $\varepsilon > 0$ such that, for any $x_0 \in \mathbb{R}^n$ and any $t \in \mathbb{R}^+$ with $\tau = \max(\|x_0\|, t) < \varepsilon$, we have*

$$\|x(t) - \widehat{x}(t)\| \leq C\tau t^{1/r},$$

*where $x(\cdot)$ and $\widehat{x}(\cdot)$ are trajectories of the control systems associated respectively to $X_1, \dots, X_m$ and $\widehat{X}_1, \dots, \widehat{X}_m$, starting at the same point x_0, defined by the same control function $u(\cdot)$, and with velocity one (*i.e. *$\sum_i u_i^2 \equiv 1$).*

Proof. The first step is to prove that there exists a constant such that $\|x(t)\|$ and $\|\widehat{x}(t)\| \leq Cst\,\tau$ for small enough τ. Let us do it for $x(t)$, the proof is exactly the same for $\widehat{x}(t)$.

The equation of a trajectory of the control system associated to $X_1, \dots, X_m$ is

$$\dot{x}_j = \sum_{i=1}^{m} u_i \big(f_{ij}(x) + r_{ij}(x)\big), \quad j = 1, \dots, n,$$

where $f_{ij}(x) + r_{ij}(x)$ is of order $\leq w_j - 1$ at 0. There exist then a constant such that, when $\|x\|$ is small enough, $|f_{ij}(x) + r_{ij}(x)| \leq Cst\,\|x\|^{w_j - 1}$ for any $j = 1, \dots, n$ and any $i = 1, \dots, m$. Note that, along a trajectory starting at x_0, $\|x\|$ is small when τ is. If moreover the trajectory has velocity one, we obtain:

$$|\dot{x}_j| \leq Cst\,\|x\|^{w_j - 1}. \tag{8}$$

To integrate this inequality, choose an integer N such that all N/w_j are even integers and set $|||x||| = \big(\sum_{i=1}^n x_i^{N/w_i}\big)^{1/N}$. The function $|||x|||$ is equivalent to $\|x\|$ in the norm sense and is differentiable out of the origin. Inequality (8) implies $\frac{d}{dt}|||x||| \leq Cst$, and then, by integration,

$$|||x(t)||| \leq Cst \ \times t + |||x(0)||| \leq Cst \ \times \tau.$$

The pseudo-norms $|||x|||$ and $\|x\|$ being equivalent, we obtain, for a trajectory starting at x_0, $\|x(t)\| \leq Cst \ \times \tau$ when τ is small enough.

The second step is to prove $|x_j(t) - \widehat{x}_j(t)| \leq Cst\, \tau^{w_j} t$. The function $x_j - \widehat{x}_j$ satisfies the differential equation

$$\begin{aligned}\dot{x}_j - \dot{\widehat{x}}_j &= \sum_{i=1}^m u_i\big(f_{ij}(x) - f_{ij}(\widehat{x}) + r_{ij}(x)\big),\\ &= \sum_{i=1}^m u_i\big(\sum_{\{k\,:\, w_k < w_j\}} (x_k - \widehat{x}_k) Q_{ijk}(x, \widehat{x}) + r_{ij}(x)\big),\end{aligned}$$

where $Q_{ijk}(x, \widehat{x})$ is a homogeneous polynomial of weighted degree $w_j - w_k - 1$. For $\|x\|$ and $\|\widehat{x}\|$ small enough, we have

$$|r_{ij}(x)| \leq Cst\, \|x\|^{w_j} \quad \text{and} \quad |Q_{ijk}(x, \widehat{x})| \leq Cst\, (\|x\| + \|\widehat{x}\|)^{w_j - w_k - 1}.$$

By using the inequalities of the first step, we obtain finally, for τ small enough,

$$|\dot{x}_j(t) - \dot{\widehat{x}}_j(t)| \leq Cst \sum_{\{k\,:\, w_k < w_j\}} |x_k(t) - \widehat{x}_k(t)| \tau^{w_j - w_k - 1} + Cst\, \tau^{w_j}. \tag{9}$$

This system of inequalities has a triangular form. We can then integrate it by induction. For $w_j = 1$, the inequality is $|\dot{x}_j(t) - \dot{\widehat{x}}_j(t)| \leq Cst\, \tau$, and so $|x_j(t) - \widehat{x}_j(t)| \leq Cst\, \tau t$. By induction, let $j > n_1$ and assume $|x_k(t) - \widehat{x}_k(t)| \leq Cst\, \tau^{w_k} t$ for $k < j$. Inequality (9) implies

$$|\dot{x}_j(t) - \dot{\widehat{x}}_j(t)| \leq Cst\, \tau^{w_j - 1} t + Cst\, \tau^{w_j} \leq Cst\, \tau^{w_j},$$

and so $|x_j(t) - \widehat{x}_j(t)| \leq Cst\, \tau^{w_j} t$.

Finally,

$$\|x(t) - \widehat{x}(t)\| \leq Cst\, \tau (t^{1/w_1} + \cdots + t^{1/w_n}) \leq Cst\, \tau t^{1/r},$$

which completes the proof of the lemma. ■

Proof (of Theorem 9). We will show that, for $\|x^0\|$ small enough,

$$d(0,x^0) \leq 2\widehat{d}(0,x^0),$$

and so $d(0,x^0) \leq Cst\,\|x^0\|$ by Lemma 8. As noticed earlier, this proves Theorem 9.

Fix $x^0 \in \mathbb{R}^n$, $\|x^0\| < \varepsilon$. Let $\widehat{x}_0(t)$, $t \in [0,T_0]$, be a minimizing curve for $\widehat{d}$, having velocity one, and joining x^0 to 0. Let $x_0(t)$, $t \in [0,T_0]$, be the trajectory of the control system associated to $X_1,\dots,X_m$ starting at x^0 and defined by the same control function than $\widehat{x}_0(t)$. Set $x^1 = x_0(T_0)$.

Thus $T_0 = \widehat{d}(0,x^0) = \text{length}\big(x_0(\cdot)\big)$. Moreover, by Lemma 11,

$$\|x^1\| = \|x_0(T_0) - \widehat{x}_0(T_0)\| \leq C\tau T_0^{1/r},$$

where $\tau = \max(\|x^0\|, T_0)$. By Lemma 8, $T_0 = \widehat{d}(0,x^0)$ satisfies $T_0 \geq \|x^0\|/C'$, so $\tau \leq C'T_0$, and

$$\widehat{d}(0,x^1) \leq C'\|x^1\| \leq C''\widehat{d}(0,x^0)^{1+1/r},$$

with $C'' = C'^2C$.

Choose now $\widehat{x}_1(t)$, $t \in [0,T_1]$, a minimizing curve for $\widehat{d}$, having velocity one, and joining x^1 to 0. Let $x_1(t)$, $t \in [0,T_1]$, be the trajectory of the control system associated to $X_1,\dots,X_m$ starting at x^1 and defined by the same control function than $\widehat{x}_1(t)$. Set $x^2 = x_1(T_1)$. As previously, we have $\text{length}\big(x_1(\cdot)\big) = \widehat{d}(0,x^1)$ and $\widehat{d}(0,x^2) \leq C''\widehat{d}(0,x^1)^{1+1/r}$.

Continuing in this way, we construct a sequence $x^0, x^1, x^2, \dots$ of points such that $\widehat{d}(0,x^{k+1}) \leq C''\widehat{d}(0,x^k)^{1+1/r}$, and horizontal curves $x_k(\cdot)$ joining x^k to x^{k+1} of length equal to $\widehat{d}(0,x^k)$.

By taking $\|x^0\|$ small enough, we can assume $C''\widehat{d}(0,x^0)^{1/r} \leq 1/2$. We have then $\widehat{d}(0,x^1) \leq \widehat{d}(0,x^0)/2, \dots, \widehat{d}(0,x^k) \leq \widehat{d}(0,x^0)/2^k, \dots$ As a consequence, x^k tends to 0 as $k \to +\infty$, and putting end to end the curves $x_k(\cdot)$, we obtain a horizontal curve joining x^0 to 0 of length $\widehat{d}(0,x^0) + \widehat{d}(0,x^1) + \cdots \leq 2\widehat{d}(0,x^0)$. This implies $d(0,x^0) \leq 2\widehat{d}(0,x^0)$, and so the theorem. ■

4.2 Approximate motion planning

Given a control system (Σ), the *motion planning problem* is to steer (Σ) from an initial point to a final point. For nonholonomic control systems, the exact problem is in general unsolvable. However methods exist for particular class of system, in particular for nilpotent (or nilpotentizable) systems. It is then of interest to devise approximate motion planning techniques based on nilpotent approximations. These techniques are Newton type methods, the nilpotent approximation playing the role of the usual linearization.

Precisely, consider a nonholonomic control system

$$(\Sigma) \ : \qquad \dot{x} = \sum_{i=1}^{m} u_i X_i(x), \quad x \in \mathbb{R}^n,$$

and initial and final points a and b in $\mathbb{R}^n$. Denote by $\widehat{X}_1, \dots, \widehat{X}_m$ a nilpotent approximation of $X_1, \dots, X_m$ at b. The k-step of an approximate motion planning algorithm take the following form (x^k denotes the state of the system):

1. compute a control law $u(t)$, $t \in [0, T]$, steering the control system associated to $\widehat{X}_1, \dots, \widehat{X}_m$ from x^k to b;
2. compute the trajectory $x(\cdot)$ of (Σ) with control law $u(\cdot)$ starting from x^k;
3. set $x^{k+1} = x(T)$.

The question is the following: is this algorithm convergent? or, at least, locally convergent? The answer to the latter question will be positive, but we need an extra hypothesis on the control law given in point 2 of the algorithm, namely,

(H) there exists a constant K such that, if x^k and b are close enough, then

$$\int_0^T \sqrt{\sum u_i^2(t)} dt \leq K \widehat{d}(b, x_k).$$

Note that a control corresponding to a minimizing curve for $\widehat{d}$ satisfies this condition. Other standards methods using Lie groups (like the one in [LS91]) or based on the triangular form of the homogeneous nilpotent approximation satisfy also this hypothesis.

We can also assume without restriction that the control is normalized: $\sum u_i^2(t) \equiv 1$.

The local convergence is then proved in exactly the same way than Theorem 9: from Lemmas 11 and 8, we have $\widehat{d}(b, x^{k+1}) \leq C'' T^{1+1/r}$, and using hypothesis (H), we obtain

$$\widehat{d}(b, x^{k+1}) \leq C'' K^{1+1/r} \widehat{d}(b, x_k)^{1+1/r}.$$

If a is close enough to b, we have, at each step of the algorithm, $\widehat{d}(b, x^{k+1}) \leq \widehat{d}(b, x_k)/2$, which proves the local convergence of the algorithm, that is:
for each point $b \in M$, there exists a constant $\varepsilon_b > 0$ such that, if $d(a, b) < \varepsilon_b$, then the approximate motion planning algorithm steering the system from a to b converges.

To obtain a globally convergent algorithm, a natural idea is to iterate the locally convergent one. This requires the construction of a finite sequence of intermediate goals $b_0 = a, b_1, \ldots, b_N = b$ such that $d(b_{i-1}, b_i) < \varepsilon_{b_i}$. However the constant ε_b depends on b and, as already noticed for Theorem 9, it is not possible to have a uniform nonzero constant near singular points. Thus this method gives a globally convergent algorithm only when every point is regular.

4.3 Hausdorff dimension

***** to be written ******

5 Tangent structure

5.1 Metric tangent space

A notion of tangent space can be defined for a general metric space. Indeed, in describing the tangent space to a manifold, we usually imagine looking at smaller and smaller neighborhoods of a given point, the manifold being fixed. Equivalently, we can imagine looking at a fixed neighborhood, but expanding the manifold. As noticed by Gromov, this idea can be used for metric spaces.

If X is a metric space with distance d, we define $\lambda\mathsf{X}$, for $\lambda > 0$, as the metric space with same underlying set than X and distance λd. A *pointed metric space* (X, x) is a metric space with a distinguished point x.

Definition. The *metric tangent space* $C_x\mathsf{X}$ to the metric space X at x is defined as

$$C_x\mathsf{X} = \lim_{\lambda \to +\infty} (\lambda\mathsf{X}, x)$$

provided the limit exists. When it exists, the metric tangent space is a pointed metric space.

Of course, for this definition to make sense, we have to define the limit of pointed metric spaces.

Let us first define the Gromov-Hausdorff distance between metric spaces. Recall that, in a metric space X, the Hausdorff distance $\text{H-}\dim(A, B)$ between two subsets A and B of X is the infimum of ρ such that any point of A is within a distance ρ of B and any point of B is within a distance ρ of A. The Gromov-Hausdorff distance $\text{GH-}\dim(X, Y)$ between two metric spaces X and Y is the infimum of Hausdorff distances $\text{H-}\dim(i(\mathsf{X}), j(\mathsf{Y}))$ over all metric spaces Z and all isometric embeddings $i : \mathsf{X} \to \mathsf{Z}$, $j : \mathsf{Y} \to \mathsf{Z}$.

Thanks to Gromov-Hausdorff distance, one can define the notion of limit of a sequence of pointed metric spaces: (X_n, x_n) converge to (X, x) if

$$\text{GH-}\dim\left(B^{\mathsf{X}_n}(x_n, r), B^{\mathsf{X}}(x, r)\right) \to 0 \quad \text{as } n \to +\infty$$

for any positive r. Note that the limit of a sequence of metric spaces is unique provided the closed balls around the distinguished point are compact.

Thus the metric tangent space can also be characterized as follows: denoting by 0 the distinguished point of $C_x\mathsf{X}$, we have, for any $r > 0$ ($r = 1$ suffices),

$$\text{GH-}\dim\big(\frac{1}{\varepsilon}B^{\mathsf{X}}(x, r\varepsilon), B^{C_x\mathsf{X}}(0, r)\big) \to 0 \quad \text{as } \varepsilon \to 0.$$

For a Riemannian manifold (M, g), the metric tangent space at a point p exists and is the Euclidean space (T_pM, g_p), that is, its standard tangent space endowed with the scalar product defined by g_p.

For a sub-Riemannian manifold, the metric tangent space is given by the nilpotent approximation.

Theorem 12. *A sub-Riemannian manifold (M, d) admits a metric tangent space $(C_pM, 0)$ at every point $p \in M$. The space C_pM is a sub-Riemannian manifold isometric to $(\mathbb{R}^n, \widehat{d})$, where $\widehat{d}$ is the sub-Riemannian distance associated to a homogeneous nilpotent approximation at p.*

This theorem is a consequence of a strong version of Theorem 9 established by Bellaïche: there exists $C > 0$ such that, for q and q' in a neighborhood of p,

$$|d(q, q') - \widehat{d}(q, q')| \leq C\widehat{d}(p, q)d(q, q')^{1/r}. \tag{10}$$

Intrinsic characterization (*i.e.* up to a unique isometry) of the metric tangent space can be found in [MM00] and [FJ03].

What is the algebraic structure of C_pM? Of course it is not a linear space in general: for instance, $\widehat{d}$ is homogeneous of degree 1 but with respect to dilations δ_t and not to the usual Euclidean dilations. We will see that C_pM has a natural structure of group, or at least of quotient of groups.

Denote by G_p the group generated by the diffeomorphisms $\exp(t\widehat{X}_i)$ acting on the left on $\mathbb{R}^n$. Since $\mathfrak{g}_p = \text{Lie}(\widehat{X}_1, \dots, \widehat{X}_m)$ is a nilpotent Lie algebra, $G_p = \exp(\mathfrak{g}_p)$ is a simply connected group, having $\mathfrak{g}_p$ as its Lie algebra.

This Lie algebra $\mathfrak{g}_p$ splits into homogeneous components

$$\mathfrak{g}_p = \mathfrak{g}^1 \oplus \cdots \oplus \mathfrak{g}^r,$$

where $\mathfrak{g}^s$ is the set of homogeneous vector fields of degree $-s$, and so is a graded Lie algebra. The first component $\mathfrak{g}^1 = \text{span}\langle\widehat{X}_1, \dots, \widehat{X}_m\rangle$ generates $\mathfrak{g}_p$ as a Lie algebra. All these properties imply that G_p is a Carnot groups.

Definition. A *Carnot group* is a simply connected Lie group whose Lie algebra is graded, nilpotent, and generated by its first component.

The vector fields $\widehat{X}_1, \dots, \widehat{X}_m$ induce a right-invariant sub-Riemannian structure (Δ, γ) on G_p. Indeed, by identifying $\mathfrak{g}_p$ to the set of right-invariant vector fields on G, we identify $\mathfrak{g}^1 \subset \mathfrak{g}_p$ to a right-invariant distribution Δ on G. The inner-product on $\mathfrak{g}^1$ which has $\widehat{X}_1, \dots, \widehat{X}_m$ as an orthonormal basis is then identified to a right-invariant Riemannian metric γ on Δ. Equivalently, let $\widehat{\xi}_1, \dots, \widehat{\xi}_m$ be the elements $\widehat{X}_1, \dots, \widehat{X}_m$ of $\mathfrak{g}_p$ viewed as right-invariant vector fields on G_p, that is

$$\widehat{\xi}(g) = \frac{d}{dt}\left[\exp(t\widehat{X}_i)g\right]\big|_{t=0}.$$

Then $\widehat{\xi}_1, \dots, \widehat{\xi}_m$ is an orthonormal frame of the sub-Riemannian structure (Δ, γ) on G_p.

The action of G_p on $\mathbb{R}^n$ is transitive, since $\widehat{D}$ is bracket-generating: the orbit of 0 under the action of G_p is the whole $\mathbb{R}^n$. The mapping $\phi_p : G_p \to \mathbb{R}^n$, $\phi_p(g) = g(0)$, is then surjective.

Case p regular:

Proposition 13. *If p is a regular point, then* $\dim G_p = n$.

Proof. Let $X_{I_1}, \dots, X_{I_n}$ be an adapted frame at p. Due to the regularity of p, it is also an adapted frame near p, so any bracket X_J can be written as

$$X_J(z) = \sum_{\{i\,:\,|I_i| \leq |J|\}} a_i(z) X_{I_i}(z),$$

where each a_i is a function of order $\geq |I_i| - |J|$. By taking the homogeneous terms of degree $-|J|$ in this expression, we obtain

$$\widehat{X}_J(z) = \sum_{\{i\,:\,|I_i| = |J|\}} a_i(0) \widehat{X}_{I_i}(z),$$

and so $\widehat{X}_J \in \operatorname{span}\langle \widehat{X}_{I_1}, \dots, \widehat{X}_{I_n} \rangle$. Thus $\widehat{X}_{I_1}, \dots, \widehat{X}_{I_n}$ is a basis of $\mathfrak{g}_p$, and so $\dim G_p = n$. ∎

If p is regular, the mapping ϕ_p is a diffeomorphism. Moreover $\phi_{p*}\widehat{\xi}_i = \widehat{X}_i$, that is ϕ_p maps the sub-Riemannian structure (Δ, γ) to $(\widehat{D}, \widehat{g})$.

Lemma 14. *When p is a regular point, the metric tangent space C_pM and $(\mathbb{R}^n, \widehat{D}, \widehat{g})$ are isometric to the Carnot group G_p endowed with the right-invariant sub-Riemannian structure (Δ, γ).*

Carnot groups are to sub-Riemannian geometry as Euclidean spaces are to Riemannian geometry: the internal operation – addition – is replaced by the law group and the external operation – product by a real number – by the dilations. Indeed, recall that the dilations δ_t act on $\mathfrak{g}_p$ as multiplication t^s on $\mathfrak{g}^s$; it extends to G_p by the exponential mapping. Notice that, when G_p is Abelian (*i.e.* commutative) it has a linear structure and the metric (Δ, γ) is a Euclidean metric.

<u>General case:</u>

Without hypothesis on p, G_p can be of dimension greater than n and the map ϕ_p not injective. Denoting by $H_p = \{g \in G_p \ : \ g(0) = 0\}$ the isotropy subgroup of 0, ϕ_p induces a diffeomorphism

$$\psi_p : G_p/H_p \to \mathbb{R}^n.$$

Beware: G_p/H_p is in general only a coset space, not a quotient group.

Observe that H_p is invariant under dilations, since $\delta_t g(\delta_t x) = \delta_t(g(x))$. Hence H_p is connected and simply connected, and so $H_p = \exp(\mathfrak{h}_p)$, where $\mathfrak{h}_p$ is the Lie sub-algebra of $\mathfrak{g}_p$ containing the vector fields vanishing at 0:

$$\mathfrak{h}_p = \{Z \in \mathfrak{g}_p \ : \ Z(0) = 0\}.$$

As $\mathfrak{g}_p$, $\mathfrak{h}_p$ is invariant under dilations and splits into homogeneous components.

Now, the elements $\widehat{X}_1, \dots, \widehat{X}_m$ of $\mathfrak{g}_p$ acts on the left on $G_p/H_p = \{gH_p \ : \ g \in G_p\}$ under the denomination $\overline{\xi}_1, \dots, \overline{\xi}_m$:

$$\overline{\xi}_i(g) = \frac{d}{dt}\big[\exp(t\widehat{X}_i)gH_p\big]\big|_{t=0}.$$

These vector fields form the orthonormal basis of a sub-Riemannian structure $(\overline{\Delta}, \overline{\gamma})$ on G_p/H_p. We clearly have $\psi_{p*}\overline{\xi}_i = \widehat{X}_i$, so ψ_p maps $(\overline{\Delta}, \overline{\gamma})$ to $(\widehat{D}, \widehat{g})$.

Theorem 15. *The metric tangent space C_pM and $(\mathbb{R}^n, \widehat{D}, \widehat{g})$ are isometric to the homogeneous space G_p/H_p endowed with the sub-Riemannian structure $(\overline{\Delta}, \overline{\gamma})$.*

5.2 Desingularization and uniform distance estimate

To get rid of singular points, the usual way is to consider a singularity as the projection of a regular object. The algebraic structure of the metric tangent space yields a good way of lifting and projecting sub-Riemannian manifolds. We start with nilpotent approximations.

We keep the notations and definitions of the preceding section. At a singular point p of the sub-Riemannian manifold, we have the following diagram between sub-Riemannian manifolds:

$$\begin{array}{ccc} (G_p, \Delta, \gamma) & & \\ \pi \downarrow & \phi_p \searrow & \\ (G_p/H_p, \overline{\Delta}, \overline{\gamma}) & \overset{\psi_p}{\underset{\sim}{\longrightarrow}} & (\mathbb{R}^n, \widehat{D}, \widehat{g}) \end{array}$$

Since (Δ, γ) is a right-invariant structure on G_p, every point in the sub-Riemannian manifold (G_p, Δ, γ) is regular. We say that (G_p, Δ, γ) is an *equiregular manifold.* Thus $(\mathbb{R}^n, \widehat{D}, \widehat{g})$ is the projection of an equiregular manifold.

Recall that $\widehat{\xi}_1, \dots, \widehat{\xi}_m$ (resp. $\overline{\xi}_1, \dots, \overline{\xi}_m$) is the orthonormal frame of (Δ, γ) (resp. $(\overline{\Delta}, \overline{\gamma})$) which is mapped to $\widehat{X}_1, \dots, \widehat{X}_m$ by ϕ_p (resp. ψ_p). Working in a system of coordinates, we identify G_p/H_p with $\mathbb{R}^n$ and $\overline{\xi}_i$ with $\widehat{X}_i$. These coordinates x on $\mathbb{R}^n \simeq G_p/H_p$ induce coordinates $(x, z) \in \mathbb{R}^N$ on G_p and we have:

$$\widehat{\xi}_i(x, z) = \widehat{X}_i(x) + \sum_{j=n+1}^{N} b_{ij}(x, z)\partial_{z_j}. \tag{11}$$

Consider a horizontal curve $(x(t), z(t))$ in G_p associated to the control $u(t)$, that is

$$(\dot{x}(t), \dot{z}(t)) = \sum_{i=1}^{m} u_i(t)\widehat{\xi}_i(x, z).$$

Then $x(t)$ is a horizontal curve in $\mathbb{R}^n$ with the same length: $\text{length}\big(x(\cdot)\big) = \text{length}\big((x, z)(\cdot)\big) = \int \sqrt{\sum u_i^2(t)}dt$. It implies that $\widehat{d}$ can be obtained from the sub-Riemannian distance d_{G_p} in G_p by

$$\widehat{d}(q_1, q_2) = \inf_{\widetilde{q}_2 \in q_2 H_p} d_{G_p}(\widetilde{q}_1, \widetilde{q}_2), \qquad \text{for any } \widetilde{q}_1 \in q_1 H_p,$$

or, equivalently, $B^{\widehat{d}}(q_1, \varepsilon) = \phi_p\big(B^{d_{G_p}}(\widetilde{q}_1, \varepsilon)\big)$.

We will use this idea to desingularize the original sub-Riemannian manifold. Choose for x privileged coordinates at p, so that

$$X_i(x) = \widehat{X}_i(x) + R_i(x) \qquad \text{with } \text{ord}_p R_i \geq 0.$$

Set $\widetilde{M} = M \times \mathbb{R}^{N-n}$, and in local coordinates (x, z) on $\widetilde{M}$, define vector fields on a neighborhood of $(p, 0)$ as

$$\xi(x, z) = X_i(x) + \sum_{j=n+1}^{N} b_{ij}(x, z)\partial_{z_j},$$

with the same functions b_{ij} than in (11).

We define in this way a sub-Riemannian structure on an open set $\widetilde{U} \subset \widetilde{M}$ which nilpotent approximation at $(p, 0)$ is, by construction, given by $(\widehat{\xi}_1, \ldots, \widehat{\xi}_m)$. Unfortunately, $(p, 0)$ can be itself a singular point. Indeed, a point can be singular for a sub-Riemannian structure and regular for the nilpotent approximation taken at this point.

Example. Take the distribution on $\mathbb{R}^5$ generated by $X_1 = \partial_{x_1}$, $X_2 = \partial_{x_2} + x_1\partial_{x_3} + x_1^2\partial_{x_4}$ and $X_3 = \partial_{x_5} + x_1^{100}\partial_{x_4}$. The origin 0 is a singular point. However the nilpotent approximation at 0 is $\widehat{X}_1 = X_1$, $\widehat{X}_2 = X_2$, $\widehat{X}_3 = \partial_{x_5}$, for which 0 is not singular.

To avoid this difficulty, we take a group bigger than G_p, namely the free nilpotent group N_r of step r with m generators. It is a Carnot group and its Lie algebra $\mathfrak{n}_r$ is the free nilpotent Lie algebra of step r with m generators. Given generators $\alpha_1, \ldots, \alpha_m$ of $\mathfrak{n}_r$ define on N_r a right-invariant sub-Riemannian structure (Δ_N, γ_N).

The group N_r can be thought as a group of diffeomorphisms and so define a left action on $\mathbb{R}^n$. Denoting by J the isotropy subgroup of 0 for this action, we obtain that $(\mathbb{R}^n, \widehat{X}_1, \ldots, \widehat{X}_m)$ is isometric to N_r/J endowed with the restrictions of (Δ_N, γ_N).

By the same reasoning as above, we are able to lift locally the sub-Riemannian structure (D, g) on M to a sub-Riemannian structure in $M \times \mathbb{R}^{\widetilde{n}-n}$, $\widetilde{n} = \dim N_r$, having for nilpotent approximation at $(p, 0)$ the orthonormal basis of (Δ_N, γ_N) defined by $\alpha_1, \ldots, \alpha_m$. And, because N_r is free up to step r, this implies that $(p, 0)$ is a regular point for that structure in $M \times \mathbb{R}^{\widetilde{n}-n}$. We obtain in this way a result of desingularization.

Lemma 16. *Let p be a point in M, r the degree of nonholonomy at p, $\widetilde{n} = \widetilde{n}(m, r)$ the dimension of the free Lie algebra of step r with m generators, and $\widetilde{M}$ 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 with $U \times \{0\} \subset \widetilde{U}$; local coordinates (x, z) on $\widetilde{U}$; and smooth vector fields on $\widetilde{U}$:

$$\xi(x, z) = X_i(x) + \sum_{j=n+1}^{N} b_{ij}(x, z)\partial_{z_j},$$

such that:

- *the distribution generated by $\xi_1, \ldots, \xi_m$ is bracket-generating and has r for degree of nonholonomy everywhere (so its Lie algebra is free up to step r);*
- *every $\widetilde{q}$ in $\widetilde{U}$ is regular;*

- *denoting $\pi : \widetilde{M} \to M$ the canonical projection and $\widetilde{d}$ the sub-Riemannian distance defined by $\xi_1, \dots, \xi_m$ on $\widetilde{U}$, we have, for $q \in U$ and $\varepsilon > 0$ small enough,*
$$B(q, \varepsilon) = \pi\Big(B^{\widetilde{d}}\big((q, 0), \varepsilon\big)\Big),$$
or, equivalently,
$$d(q_1, q_2) = \inf_{\widetilde{q}_2 \in \pi^{-1}(q_2)} \widetilde{d}\big((q_1, 0), \widetilde{q}_2\big).$$

Remark. The lemma still holds if we take for r any integer greater than the degree of nonholonomy at p.

Thus any sub-Riemannian manifold is locally the projection of an equiregular sub-Riemannian manifold. This projection preserves the horizontal curves, the minimizing curves, the distance. We will see in Part III that it also preserves the extremals and the geodesics.

Application: uniform Ball-Box theorem

The key feature of the sub-Riemannian structure at regular points is the uniformity:

- uniformity of the flag (4);
- uniformity w.r.t. p of the convergence $(\lambda(M, d), p) \to C_pM$ (as explained by Bellaïche [Bel96, §8], this uniformity is responsible for the group structure of the metric tangent space);
- uniformity of the distance estimates (see Remark page 21).

This last property in particular is essential to compute Hausdorff dimension or to prove the global convergence of approximate motion planning algorithms. Recall what we mean by uniformity in this context: in a neighborhood of a regular point p_0, we can construct privileged coordinates depending continuously on the base point p and such that the distance estimate (7) holds with C_p and ε_p independent of p.

As already noticed, all these uniformity properties are lost at singular points. However, using the desingularization of the sub-Riemannian manifold, it is possible to give a uniform version of the distance estimates.

Let $\Omega \subset M$ be a compact set. We denote by $r_{\max}$ the maximum of the degree of nonholonomy on Ω. We assume that M is an oriented manifold, so the determinant n-form det is well-defined.

Given $q \in \Omega$ and $\varepsilon > 0$, we consider the families $\mathcal{X} = (X_{I_1}, \ldots, X_{I_n})$ of brackets of length $|I_i| \leq r_{\max}$. On the finite set of these families, we define a function

$$f_{q,\varepsilon}(\mathcal{X}) = \left|\det \left(X_{I_1}(q)\varepsilon^{|I_1|}, \ldots, X_{I_n}(q)\varepsilon^{|I_n|}\right)\right|.$$

We say that $\mathcal{X}$ is an *adapted frame at* (q, ε) if it achieves the maximum of $f_{q,\varepsilon}$.

The values at q of an adapted frame at (q, ε) clearly form a basis of T_qM. Moreover, q being fixed, the adapted frames at (q, ε) are adapted frame at q for ε small enough.

Theorem 17 (Uniform Ball-Box theorem). *There exist positive constants K and ε_0 such that, for all $q \in \Omega$ and $\varepsilon < \varepsilon_0$, if $\mathcal{X}$ is an adapted frame at (q, ε), then*

$$\mathrm{Box}_{\mathcal{X}}(q, \frac{1}{K}\varepsilon) \subset B(q, \varepsilon) \subset \mathrm{Box}_{\mathcal{X}}(q, K\varepsilon),$$

where $\mathrm{Box}_{\mathcal{X}}(q, \varepsilon) = \{e^{x_1 X_{I_1}} \circ \cdots \circ e^{x_n X_{I_n}}(q) \ : \ |x_i| \leq \varepsilon^{|I_i|}, \ 1 \leq i \leq n\}$.

Of course, q being fixed, this estimate is equivalent to the one of Ball-Box theorem for ε smaller than some $\varepsilon_1(q) > 0$. However $\varepsilon_1(q)$ can be infinitely close to 0 though the estimate here holds for ε smaller than ε_0, which is independent of q.

Part III. Minimizing curves

6 Optimal control problem

7 Singular curves

8 Extremals and geodesics

Bibliography

SR geometry in general A few surveys, [Str86], [VG94], [Kup96] (in french), [Mon02], and the collection [BR96] including the important contributions of Gromov [Gro96] and Bellaïche [Bel96].

Part II. Infinitesimal point of view On nilpotent approximations: [AS87], [BS90], [Her91] (control theory), [RS76], [Goo76] (analysis). Tangent cone, Hausorff dimensions: [Gro81], [Mit85], [MM00], [FJ03], [Jea03]. Lifting method, uniform Ball-Box theorem: [NSW85], [Jea01].

Part III. Minimizing curves On singular trajectories, rigidity: [BC03] (more general context than SR geometry), [BH93], [AS96], [CJT03]. Numerous papers on sub-analyticity of spheres and of the SR distance, on the optimal status of abnormal geodesics, for instance [LS95], [AS95], [Agr98], [AS99], [AG01], [ABCK97], [BC99], [BT01], [Tré01]. Three-dimensional SR metrics, exponential mappings: [Agr96], [BGG00], [AEAG98].

References

[ABCK97] A. Agrachev, B. Bonnard, M. Chyba, and I. Kupka. Sub-Riemannian sphere in Martinet flat case. *ESAIM: Control, Optimisation and Calculus of Variations*, 2:377–448, 1997.

[AEAG98] A. A. Agrachev, El-H. Ch. El-Alaoui, and J.-P. Gauthier. Sub-Riemannian metrics in R^3. *Proc. Canadian Math. Soc.*, 25:29–78, 1998.

[AG01] A. A. Agrachev and J.-P. Gauthier. On subanalyticity of Carnot-Carathéodory distances. *Ann. Inst. Henri Poincaré, Analyse Non Linéaire*, 18(3), 2001.

[Agr96] A. A. Agrachev. Exponential mappings for contact sub-Riemannian structures. *Journal on Dynamical and Control Systems*, 2(3):321–358, 1996.

[Agr98] A. A. Agrachev. Compactness for sub-Riemannian length minimizers and subanalyticity. *Rend. Semin. Mat. Torino*, 56, 1998.

[AS87] 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.

[AS95] A. A. Agrachev and A. V. Sarychev. Strong minimality of abnormal geodesics for 2-distributions. *Journal on Dynamical and Control Systems*, 1(2), 1995.

[AS96] A. A. Agrachev and A. V. Sarychev. Abnormal sub-Riemannian geodesics: Morse index and rigidity. *Ann. Inst. Henri Poincaré, Analyse Non Linéaire*, 13(6):635–690, 1996.

[AS99] A. A. Agrachev and A. V. Sarychev. Sub-Riemannian metrics: Minimality of abnormal geodesics versus subanalyticity. *ESAIM: Control, Optimisation and Calculus of Variations*, 4, 1999.

[BC99] B. Bonnard and M. Chyba. Exponential mapping, spheres and waves front in sub-Riemannian sphere: the Martinet case. *ESAIM: Control, Optimisation and Calculus of Variations*, 4, 1999.

[BC03] B. Bonnard and M. Chyba. *Singular Trajectories and Their Role in Control Theory.* Springer, 2003.

[Bel96] 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.

[BGG00] R. Beals, B. Gaveau, and P. C. Greiner. Hamilton-Jacobi theory and the heat kernel on Heisenberg groups. *Journal de Mathématiques Pures et Appliquées*, 79:633–689, 2000.

[BH93] R. L. Bryant and L. Hsu. Rigidity of integral curves of rank 2 distributions. *Inventiones Math.*, 114, 1993.

[BR96] A. Bellaïche and J.-J. Risler, editors. *Sub-Riemannian Geometry.* Progress in Mathematics. Birkhäuser, 1996.

[BS90] R. M. Bianchini and G. Stefani. Graded approximations and controllability along a trajectory. *SIAM J. on Control and Optimization*, 28(4):903–924, 1990.

[BT01] B. Bonnard and E. Trélat. On the role of abnormal minimizers in sub-Riemannian geometry. *Annales de la Faculté des Sciences de Toulouse, Série 6*, X(3):405–491, 2001.

[CJT03] Y. Chitour, F. Jean, and E. Trélat. Propriétés génériques des trajectoires singulières. *C. R. Acad. Sci. Paris, Serie I*, 337:49–52, 2003.

[FJ03] E. Falbel and F. Jean. Measures of transverse paths in sub-Riemannian geometry. *Journal d'Analyse Mathematique*, 2003. To appear.

[Ger84] V. Ya. Gershkovich. Two-sided estimates of metrics generated by absolutely non-holonomic distributions on Riemannian manifolds. *Sov. Math. Dokl.*, 30:506–510, 1984.

[Goo76] N. Goodman. *Nilpotent Lie groups*, volume 562 of *Springer Lecture Notes in Mathematics*. Springer Verlag, 1976.

[Gro81] M. Gromov. *Structures Métriques pour les Variétés Riemanniennes.* Cedic-Nathan, 1981.

[Gro96] M. Gromov. Carnot-Carathéodory spaces seen from within. In A. Bellaïche and J.-J. Risler, editors, *Sub-Riemannian Geometry*, Progress in Mathematics. Birkhäuser, 1996.

[Her91] H. Hermes. Nilpotent and high-order approximations of vector field systems. *SIAM Review*, 33(2):238–264, 1991.

[Jea01] F. Jean. Uniform estimation of sub-Riemannian balls. *Journal on Dynamical and Control Systems*, 7(4):473–500, 2001.

[Jea03] F. Jean. Entropy and complexity of a path in sub-Riemannian geometry. *ESAIM: Control, Optimisation and Calculus of Variations*, 9:485–506, 2003.

[Kup96] I. Kupka. Géométrie sous-riemannienne. In *Séminaire N. Bourbaki*, volume 817, June 1996.

[LS91] G. Lafferriere and H. Sussmann. Motion planning for controllable systems without drift. In *Proceedings of the 1991 IEEE International Conference on Robotics and Automation*, Sacramento, California, 1991.

[LS95] W. S. Liu and H. J. Sussmann. Shortest paths for sub-Riemannian metrics of rank two distributions. *Memoirs AMS*, 118(564), 1995.

[Mit85] J. Mitchell. On Carnot-Carathéodory metrics. *Journal of Differential Geom.*, 21:35–45, 1985.

[MM00] 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.

[Mon02] R. Montgomery. *A tour of sub-Riemannian geometries, their geodesics and applications*. Math. Surveys and Monographs. American Math. Soc., 2002.

[NSW85] A. Nagel, E. M. Stein, and S. Wainger. Metrics defined by vector fields. *Acta Math.*, 155:103–147, 1985.

[RS76] L. P. Rothschild and E. M. Stein. Hypoelliptic differential operators and nilpotent groups. *Acta Math.*, 137:247–320, 1976.

[Ste86] 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.

[Str86] R. S. Strichartz. Sub-Riemannian geometry. *J. Differential Geometry*, 24(2):221–263, 1986. Corrections in *J. Differential Geometry*, 30(2):595–596, 1989.

[Tré01] E. Trélat. Asymptotics of accessibility sets along an abnormal trajectory. *ESAIM: Control, Optimisation and Calculus of Variations*, 6:387–414, 2001.

[VG94] A. M. Vershik and V. Ya. Gershkovich. Nonholonomic dynamical systems, geometry of distributions and variational problems. In V. I Arnold and S. P. Novikov, editors, *Dynamical Systems VII*, volume 16 of *Encyclopaedia of Mathematical Sciences*. Springer, 1994.