THE PROBLEM OF GEODESICS, INTRINSIC DERIVATION AND THE USE OF CONTROL THEORY IN SINGULAR SUB-RIEMANNIAN GEOMETRY

Fernand PELLETIER Liane VALÈRE BOUCHE

LAMA
Université de Savoie, Campus Scientifique
F-73376 Le Bourget du Lac Cedex (France)
pelletier@univ-savoie.fr valere@univ-savoie.fr

Abstract. We try to convince geometers that it is worth using Control Theory in the framework of sub-Riemannian structures, not only to get necessary conditions for length-minimizing curves, but also, from the very beginning, to give a description of sub-Riemannian structures by means of a global control vector bundle. This method is particularly efficient in characterizing admissible metrics with rank singularities. Some examples are developed.

Résumé. Notre but est d'essayer de convaincre les géomètres que cela vaut la peine d'appliquer les méthodes de la Théorie du Contrôle dans le contexte de structures sous-riemanniennes, non seulement pour obtenir des conditions nécessaires concernant les courbes minimisant la longueur, mais aussi, dès l'origine de la théorie, afin de définir globalement les structures sous-riemanniennes par des fibrés vectoriels dits de contrôle. Cette méthode est particulièrement efficace dans la caractérisation des métriques admissibles présentant des singularités de rang ; nous donnons des exemples.

M.S.C. Subject Classification Index (1991) : 49F22, 53B99, 53C22.

Acknowledgements. L. Valère wants to thank Marcel Berger and the members of his "seminar" for having welcome J. Houillot and herself, first as occasional participants in Paris VII in the seventies and then as contributors when writing the book about Einstein manifolds and Calabi's conjecture. Then, she would like to thank also the organizers of this present "Table Ronde en l'Honneur de Marcel Berger" for the nice work they have done, and the results of it.

TABLE OF CONTENTS

INTRODUCTION

1. Description of main results.

The main motivation for my talking here is to convince geometers that the Control Theory framework is providing a better understanding and an adapted tool in sub-Riemannian geometry. Our original presentation permits to associate to a singular plane distribution a family of natural sub-Riemannian metrics, with respect to which the regular case results are extendible to the singular one (section 4).

Another motivation is to give a really intrinsic definition in this context of a sub-Riemannian derivation (generalization of [S], section 8).

And the last motivation is to give an alternate proof that the abnormal horizontal helix in the Montgomery-Kupka example is length minimizing (section 9, [V], [V-P]). This method allows, as we know now, a generalization to any sub-Riemannian metric on a "generic" two distribution in $\mathbb{R}^3$.

Though looking far from the main concerns of Marcel Berger, the subject of this lecture has something to do with what has been a good deal of his own work ; namely, one of his successes has been the interpretation, in terms of Riemannian geometric invariants, of the asymptotic development of the heat kernel of the Laplace operator. In a parallel direction, G. Ben Arous [B-A], R. Léandre [L], G. Besson, (see also [A], [Bi], [G]) working on the asymptotic expansion of the Green kernel in the theory of hypo-elliptic operators, have pointed out the essential link between this expansion and the distance and geodesic notions in an associated regular or non regular plane distribution endowed with a Carnot-Carathéodory metric. The alternate name for such a framework is "sub-Riemannian geometry".

Anyway, geometers should be interested in sub-Riemannian structures for themselves, as did R.W. Brockett, R.S. Strichartz, C. Bär, U. Hamenstädt and also M. Gromov, P. Pansu, J. Mitchell, because they are nice particular examples of non integrable distributions on manifolds, besides the expansion of the Green kernel of hypo-elliptic operators.

One way of describing a regular or singular sub-Riemannian manifold M is providing M with a locally free, finite, constant rank p, bracket generating submodule $\mathcal{E}$ of the module of vector fields $\chi(M)$. An absolutely continuous (a.c.) curve is called **horizontal** if its velocity vector lies a.e. in $\mathcal{E}$.

Chow's theorem [C], using the bracket generating condition, says that the space of horizontal piecewise C^1-curves joining two fixed points x_0 and x_1 is not empty.

The two main problems are then,

(i) among the a.c. horizontal curves joining x_0 and x_1, does there exist some length minimizing curve ?

(ii) if yes, how to characterize these curves ?

Now, provided the Riemannian manifold $(M, \mathbf{g})$ is complete, it is well-known, in the regular case, that the minimum exists and that standard variational methods of Riemannian Geometry do not solve the sub-Riemannian minimization problem. In contrast to the Riemannian case, where the energy minimizing curves are characterized as solution of a differential system (**G**), here, both notions can be generalized but they are no longer equivalent [S]. The Maximum Principle of Control Theory was already known as a very good tool giving account of "abnormal" geodesics, i.e., curves minimizing the energy between two given points but not verifying the differential "geodesic" equation (**G**), generalizing the Riemannian geodesic equation obtained by a classical variational principle (this was already realized in the regular case, see [Br], [S], see also [Gr], [Mi]).

Here, we are using Control Theory from the very beginning of the definition of **singular,** i.e., not constant rank, plane-distribution. This last setting out is original and allows plenty of sub-Riemannian metrics on a given plane distribution. The main result is showing the link between metric and distribution in the neighbourhood of singularities through the Control space ; in the regular case, any sub-Riemannian metric can be seen as the restriction to the plane distribution of some (actually infinetely many) Riemannian metrics on M, whereas **in the singular case, given any sub-Riemannian metric, there exists no Riemannian metric on M, such that its restriction to the plane distribution could be the given one**.

In section 2, we give an account of what is known about regular sub-Riemannian manifolds M (the plane distribution is then of constant rank).

In section 3, we do our best to give a quick survey of the main ideas explaining how the maximum principle works, following the inventors of the theory, see [P].

In section 4, we use, from the beginning, ideas of Optimal Control Theory and describe the framework of a singular sub-Riemannian geometry, where the "horizontal" singular distribution is generated by a module of vector fields, locally free of finite rank p (definition (4-6)) ; possible metrics on such a plane distribution have to be chosen carefully, otherwise **the distance between two given distinct points of the singular set in M could be zero or never be achieved by any horizontal curve**, as illustrated by means of the very simple Example (4-1).

In section 5, we merely prove that, even in this context, looking for a horizontal length minimizing curve among horizontal a.c. curves $\gamma : I \longrightarrow M$ joining two fixed points x_0 and x_1, is equivalent to looking for a horizontal energy minimizing curve between x_0 and x_1. The first one is defined up to a.c. reparametrizations. One of these provides the curve with a velocity vector of constant norm and is then energy minimizing.

In section 6, we prove, applying Bellaïche's method to this context [Bel], that between two distinct points, within a compact cell K, the minimum of energy is finite and is actually achieved on some curve.

In section 7, we use the Maximum Principle, knowing that the minimum of energy is achieved on some curve to display necessary conditions in the form of differential equations or conditions involving derivatives which are to be defined carefully in this case. The result is that there exist three kinds of minimizing curves, either normal (N) or strictly abnormal (SAN), or both (NAN), exactly as in regular sub-Riemannian geometry. Conversely, a curve satisfying the (N) or (NAN) condition are locally energy minimizing curves, but as far as we know, there does not exist criteria to tell when a (SAN)-curve is locally length minimizing or not. Actually, we have now (1993) examples of a non length-minimizing (SAN)-curve for some codimension one distributions in $\mathbb{R}^{2p}$ (see [P-V-2]). Since the end of 1993 we know also that, in dimension 3, the Montgomery example is a generic local model : the abnormal horizontal curves drawn on the singular surface are (NAN) or (SAN), always C^1-rigid, and locally minimizing, whatever the sub-Riemannian metric [V-P]. Finally, we illustrate the method, in the singular case, by resuming the Example (4-1) and constructing some special normal geodesics.

In section 8, we produce an intrinsic derivation $D_\xi \eta$ defined on the cotangent fiber bundle with values in the tangent bundle. It is an extension to the whole $(T^*M)^2$ of the projected $(\nabla_{\mathrm{sym}})_\xi \xi$ initiated by C. Bär [B], (its "derivation" was defined only on the diagonal of $(T^*M)^2$, ours verifies $D_\xi \xi = g \circ (\nabla_{\mathrm{sym}})_\xi \xi$). Our intrinsic derivation allows a new way of writing the equations of normal geodesics (N) in the regular or singular context as well.

In section 9, we go back to the regular case and give a new proof of the fact that the example exhibited by R. Montgomery and simplified by I. Kupka (see also [Mo], [K], [L-S]) of an abnormal (SAN)-extremal of the maximum principle is actually a globally minimizing curve between two of its not too far away points, and is C^1-rigid, i.e., isolated with respect to the C^1-topology, though evidently not isolated with respect to the H^1-topology.

2. REGULAR SUB-RIEMANNIAN STRUCTURES

In this section we merely sum up what is already known about geodesics in sub-Riemannian geometry. Let us call *sub-Riemannian* manifold (M, E, G) an n-dimensional manifold M, with TM its tangent bundle, T^*M its cotangent bundle, provided with a C^∞ p-plane distribution $(p \leq n)$ of vectors termed as "horizontal" vectors $(E_x \subset T_xM)$, verifying the so-called Hörmander condition, that all the iterated Lie derivatives of local horizontal vector fields by local horizontal vector fields above a point x of M generate T_xM. Let X_x be an element of the fiber E_x, and X any local horizontal vector field extending X_x ; then, let us denote by $E_1(X)_x = E_x$, $E_2(X)_x = E_x + [X, E]_x$, $E_k(X)_x = E_x + [X, E_{k-1}(X)]_x$, and $p_k(X)_x$ the dimension of $E_k(X)_x$. The non-decreasing sequence

$$\Big(p_1(X)_x, \ldots, p_k(X)_x, \ldots\Big)$$

is such that for $k \geq 2$, $p_k(X) : M \longrightarrow \mathbb{N}$ is lower semi-continuous. The vector space $E_2(X)_x$ does not depend on the choice of the locally extending fields X in E, but only on the distribution E and the value of X_x above x. Let us denote now by $(E_1)_x = E_x, \quad (E_k)_x = E_x + \sum_{X \in E_x} [X, E_{k-1}]_x$, and $p_k(x)$ the dimension of $(E_k)_x$. The Hörmander condition merely means that

$$\forall x \in M,\ \exists r_0(x)\ /\ p_{r_0-1}(x) < n, \quad \text{and} \quad \forall r \geq r_0(x), p_r(x) = n\ .$$

The map $r_0 : M \longrightarrow \mathbb{N}$ is upper semi-continuous.

Further, every E_x is provided with a positive definite quadratic form, G_x depending smoothly on x. To the quadratic form G is canonically associated a linear fiber bundle morphism $g : T^*M \longrightarrow TM$, above the identity, and related to G by

$$G(X,Y)_x =< \xi, Y >_x =< \eta, X >_x =< \xi, g\eta >_x =< \eta, g\xi >_x\ ,$$

where X_x and Y_x are two horizontal vectors above x, $\xi_x \in g^{-1}(X_x)$ and $\eta_x \in g^{-1}(Y_x)$, are one of their respective inverse image by g_x, and $< \ , \ >_x$ is the duality product above x.

Let $\gamma : [a,b] \longrightarrow M, \quad [a,b] \subset \mathbb{R}$ be any continuous piecewise C^1 curve ; the curve γ is called "horizontal" if its tangent vector $\dot\gamma(t)$ at almost every point $t,\ t \in [a,b]$, is in $E_{\gamma(t)}$. As a matter of fact, a well known theorem due to W. L. Chow [Ch] says that any two points of M can be joined by a continuous piecewise C^1 horizontal curve, provided that the Hörmander condition is fulfilled. Then, it has been proved that the definition of G permits to define a distance on M, called the Carnot-Carathéodory distance.

Let x_0 and x_1 be any two points in M, let $\mathcal{C}_{x_0,x_1}$ be the set of continuous piecewise C^1 horizontal curves γ such that $\gamma(a) = x_0\ ,\ \gamma(b) = x_1$, we get the definition of the G-length for such a curve γ as

$$l_G(\gamma) = \int_a^b \sqrt{G(\dot\gamma, \dot\gamma)}\ dt\ .$$

Then, it is known that

$$d_G(x_0, x_1) = \inf\Big\{ l_G(\gamma)\ /\ \gamma \in \mathcal{C}_{x_0,x_1} \Big\}$$

exists and is achieved on some horizontal curve γ ([S-1], [H]).

R. Strichartz showed also that, for some of these locally length minimizing curves, one of the lifts ξ in $g^{-1}(\dot\gamma)$ of their tangent vector verifies a differential equation (**G**) which is a generalization of the classical one in the Riemannian case ($p = n$), namely, in a coordinate chart

$$(\mathbf{G}) \qquad \left(\dot\xi_\alpha + \frac{1}{2}\frac{\partial g^{\lambda\mu}}{\partial x^\alpha}\xi_\lambda\xi_\mu\right)_{x(t)} = 0\ .$$

In the case of the two steps strong generating Hörmander condition (i.e., $\forall X,\ p_2(X) = n$), it is easy to prove that p is even and all local length minimizing curves verify (**G**), and reciprocally. In other cases there exist examples of curves which are length minimizing, but do not verify (**G**), see [Mo], and section 10 below. R.S. Strichartz pointed out this difficulty and showed, as already did R.W. Brockett, that the constraint for curves being horizontal could be translated in terms of commands in the framework of Control Theory, and that this kind of curves is known and called "abnormal" extremals in Control Theory. These abnormal locally minimizing curves which are not solution of the classical Euler-Lagrange equations had been already detected by C. Carathéodory [C], Mayer [Ma], and R. Hermann [He-1], [He-2].

3. OPTIMAL CONTROL FRAMEWORK

In order to formulate the basic problem of Optimal Control, which we shall have to solve in the sections following this one, we recall the definitions and results of the theory which will be of some use for us. One first needs the definition of a system S, which will be given by the following data :

- a differential equation

$$(\mathbf{C}) \qquad \dot x = f(x,u)\ ;$$

- x belongs to a phase space M, which is an open subset of a Euclidean space $\mathbb{R}^n$ or the closure of an open subset of $\mathbb{R}^n$;
- u belongs to a control space domain $\mathbf{U}$, bounded closure of an open subset of a Euclidean space $\mathbb{R}^p$;
- the map $f : M \times \mathbf{U} \longrightarrow \mathbb{R}^n$ is a smooth field over M, C^k $(k \geq 1,)$ C^∞ or C^ω .

Let $I = [a, b]$ be any closed interval in $\mathbb{R}$; we shall denote by $\mathcal{M}(I; \mathbf{U})$ the set of measurable curves : $\{ \tilde{u} : [a, b] \longrightarrow \mathbf{U} \}$. So, as soon as an initial point $x(a) = x_0$ is chosen, to any such control curve $\tilde{u}$ are associated a uniquely determined maximum real value $\tilde{t}_1 \in [a, b]$, depending smoothly on x_0 and $\tilde{u}$, $\tilde{t}_1(x_0, \tilde{u})$, and a unique absolutely continuous curve, integral of ($\mathbf{C}$),

$$\tilde{x} : [a, \tilde{t}_1] \longrightarrow M \ ,$$

called a $\mathbf{C}$-path. We then give the following

3.1. Definition. — *Let $\tilde{u} \in \mathcal{M}(I; \mathbf{U})$; let x_0 be any point in M and $\tilde{t}_1,\ a \leq \tilde{t}_1 \leq b$, be the maximum real values such that*

$$\forall t \in [a, \tilde{t}_1], \qquad \tilde{x}(t) = x_0 + \int_a^t f(\tilde{x}(t), \tilde{u}(t))\, dt \quad \textit{exists} \ ,$$

the pair of functions

$$(\tilde{x}, \tilde{u}) : [a, \tilde{t}_1] \longrightarrow M \times \mathbf{U}$$

is called "trajectory of the controlled system S" •

Let us denote by $\mathcal{T}_{x_0}$ the set of trajectories such that $x(a) = x_0$.

From now on we shall often use the terms "almost everywhere", or "for almost every t", this will be equivalent to saying "for every regular value of t", with respect to the control maps, thanks to the hypothesis of measurability. Let us then define what it is.

3.2. Definition. — *A real value θ, $\theta \in [a, b]$ is called regular with respect to the admissible control $\tilde{u}$, if for any neighbourhood $U \subset \mathbf{U}$ of $\tilde{u}(\theta)$,*

$$\lim_{\mu(I) \to 0} \frac{\mu(\tilde{u}^{-1}(U) \cap I)}{\mu(I)} = 1 \ ,$$

where μ is the Borel measure •

In the following, the first problem will be to study the effects of variations of controls onto the paths in M, and to manage to get any possible path $\tilde{x}^*$ close to an original one $\tilde{x}$, through a class of variations in $\mathcal{M}(I;\mathbf{U})$, with nice properties. The class used by L. Pontryagin and coll. is the class of Mac Shane variations, i.e., the admissible controls different from the original one $\tilde{u}$, only on a finite number of small intervals, but such that $\tilde{u}^* - \tilde{u}$ is an arbitrary constant on each of these intervals. So, let $(\tilde{x}, \tilde{u})$ be a trajectory in $\mathcal{T}_{x_0}$, and let us consider Mac Shane variations, $\tilde{u}^* : [0, t+\delta t] \longrightarrow \mathbf{U}$ of $\tilde{u} : [0,t] \longrightarrow \mathbf{U}$, where δt is any real number (see [P] for more precisions) ; then, it can be proved that, in $T_{\tilde{x}(t)} = \mathbb{R}^n$, the set of vectors

$$K(t) = \left\{ \ \tilde{x}^*(t+\delta t) - \tilde{x}(t) \ / \ (\tilde{x}^*(t), \tilde{u}^*(t)) \in \mathcal{T}_{x_0} \ \right\}$$

describes a cone. Reciprocally, for any X in $K(t)$, there exists a real number $\varepsilon > 0$ and a conic ε-neighbourhood of X,

$$K_\varepsilon(X) = \left\{ \ \varepsilon X + \varepsilon Y \ / \ Y \in X^\perp, ||Y|| = 1 \ \right\}$$

such that any point inside the ε-cone $K_\varepsilon(X)$ is the end point $\tilde{x}^*(t+\varepsilon\delta t)$ of a "pushed" path through a Mac Shane variation $\tilde{u}^*$.

The previous tools and notions take place in $\mathbb{R}^n$ but can easily be interpreted in an n-dimensional Riemannian manifold $(M, \mathbf{g})$, by means of the exponential map and the theory of differential equations, or simpler, by means of the Nash isometric imbedding theorem.

Now, in $\mathbb{R}^n$, let t and t', $t < t'$ be two regular values ; then, the differential equation

(3-3)
$$\frac{dX^\alpha}{dt} = \frac{\partial f^\alpha}{\partial x^\beta} X^\beta$$

permits to define a translation of $T_{\tilde{x}(t)}M$ to $T_{\tilde{x}(t')}M$, called $A_{t't}$, which translates $K(t)$ to $K(t')$. Then, we call "limit cone $K(t_1)$" the limit of the following set

$$K(t_1) = \sum_{\text{regular } t's} A_{t_1 t} K(t) \ .$$

Now, if x_0 and x_1 are two given distincts points of M, we denote by

$$\mathcal{T}_{x_0,x_1} = \Big\{ \ (\tilde{x},\tilde{u}) \in \mathcal{T}_{x_0} \quad / \quad \exists \tilde{t}_1(x_0,\tilde{u}) \in [a,b], \quad \exists \tilde{u} : [a,b[\longrightarrow \mathbf{U} \ ; \ \tilde{x}(\tilde{t}_1) = x_1 \ \Big\}$$

and by $\mathbf{U}_{x_0,x_1}$, the projection of $\mathcal{T}_{x_0,x_1}$ on $\mathbb{R}^p$.

Besides this, we define a positive density cost function along a trajectory $(\tilde{x},\tilde{u})$, $f^0\big(\tilde{x}(t),\tilde{u}(t)\big)$, and a positive functional, called the "cost" of the system (S)

$$\tilde{y}^0\big(\tilde{x}(t),\tilde{u}(t)\big) = \int_a^{\tilde{t}_1} f^0\big(\tilde{x}(t),\tilde{u}(t)\big)\ dt\ .$$

The last definition implies that y^0 is the solution of the differential equation

$$(\mathbf{C_0}) \qquad \dot{y}^0 = f^0(x,u)\ .$$

Now, let y denote the points of $\mathbb{R} \times M$,

$$\Big\{ y = (y^0, x) \Big\}\ ,$$

where y^0 is the cost functional, the value of which being considered as a new independent coordinate. Now we can transform the definition (3-1) into (3-4).

3.4. Definition. — *Let $[a,b]$ be any closed interval in $\mathbb{R}$, $\tilde{u} : [a,b] \longrightarrow \mathbf{U}$, a measurable map, let x_0 be any point in M. Let $\tilde{t}_1,\ a \le \tilde{t}_1 \le b$, be the maximum real value such that*

$$\forall t \in [a,\tilde{t}_1], \qquad \tilde{y}(t) = (0,x_0) + \int_a^t f(\tilde{x}(t),\tilde{u}(t))\ dt \quad \textit{exists}\ ;$$

the pair of functions

$$(\tilde{y},\tilde{u}) : [a,\tilde{t}_1] \longrightarrow (\mathbb{R} \times M) \times \mathbf{U}$$

is called "Trajectory" (with a capital T) of the controlled system S, with cost density function f^0 •

Of course, as soon as an initial point $\tilde{x}(a) = x_0 \in M$ is chosen, to any control $\tilde{u}(t) : [a,b] \longrightarrow \mathbf{U}$ is associated a uniquely determined Trajectory

$$(\tilde{y}(t),\tilde{u}(t)) : [a,\tilde{t}_1] \longrightarrow (\mathbb{R} \times M) \times \mathbf{U}$$

because of the differential equation $(\mathbf{C}, \mathbf{C_0})$.

Let us call **accessible set from** x_0 the set of all $\tilde{y}(t)$ that we just defined, for any t and through any measurable map $\tilde{u} : [a,b] \longrightarrow \mathbf{U}$. The problem of Optimal Control is then :

3.5. Problem. — *Let x_0, x_1 be two given distincts points of M, find at least one control curve $\bar{u} : [a,b] \longrightarrow \mathbf{U}$, such that*

$$(\bar{x}, \bar{u}) \in \mathcal{T}_{x_0,x_1} \ ,$$

and

$$\bar{y}^0(\bar{t}_1) = \inf \left\{ \tilde{y}^0(\tilde{t}_1) \quad / \quad (\tilde{x}, \tilde{u}) \in \mathcal{T}_{x_0,x_1} \right\} \bullet$$

3.6. Notation and Definition. — *A Trajectory as just defined $(\bar{y}, \bar{u})$ is called an "optimal Trajectory", $\bar{x}$, $\bar{u}$, $\bar{y}^0$ are respectively called "optimal path" from x_0 to x_1, "Optimal Control", and "optimal cost"* •

In many technical problems of Optimal Control, **U** is a polyhedron, and the Optimal Control because of the Maximum Principle (see Theorem (3-11) below) jumps from a vertex to another one ; this is why the class of functions $\tilde{u}$ must contain at least piecewise C^0 ones ; it is even possible to deal with measurable functions. The various controls in action are not necessarily in the neighbourhood of one of them ; this is the reason why the proof of the Maximum Principle is not simple, but at the same time, more powerful than the classical Lagrange calculus of variations (which becomes a particular case of Optimal Control theory), as was pointed out by L. Pontryagin himself ([P] chap. 5).

The idea is the following. When the controlled system is not linear, the set of accessible points $\tilde{y}(t)$ obtained as points of the integral curves of $(\mathbf{C})$ through all controls in $\mathbf{U}$, is non-convex and infinite. The Trajectory $(\bar{y}(t), \bar{u}(t))$ is optimal if and only if the zero component $\bar{y}^0(t_1) = \int_a^{t_1} f^0(\bar{x}(t), \bar{u}(t))\, dt$ is minimum compared to the other $\tilde{y}^0(t_1)'$s, and then the end point of the optimal Trajectory lies on the boundary of the accessible set in $\mathbb{R}^{n+1}$. Moreover, if one of the Trajectories which goes from

$(0, x_0)$ to (y^0, x_1) is optimal, the only controls $\tilde{u} \in \mathbf{U}$, which will be chosen in order to be compared with the Optimal Control $\bar{u}$ are Mac Shane variations.

For almost every $t,\ t \in [a, \tilde{t}_1]$, the set of accessible points from $\{y(a)\}$ by trajectories generated by means of Mac Shane perturbations on controls is the cone $K(t)$, and it can be proved ([P] Lemma 4, p. 90) that the half-line with $\bar{y}(t)$ as origin and oriented towards negative y^0's has an empty intersection with the interior of the cone $K(t)$. Then, there exists at least a supporting hyperplane $P_{\bar{y}(t)}$ passing through $\bar{y}(t)$, and a perpendicular vector $\bar{\lambda}(t)$ to $P_{\bar{y}(t)}$, which can be seen better as a non-zero element of $T^*_{\bar{y}(t)}(M \times \mathbb{R}) = \mathbb{R}^{n+1}$ with $P_{\bar{y}(t)}$ as its kernel. Thus, the half line $[\bar{y}^0(t), -\infty[$, oriented towards negative y^0's, is either outside the cone, or at most lies on its boundary. The 1-form $\bar{\lambda}(t)$ is determined up to a multiplicative factor, usually it is chosen in order to make the function $< \bar{\lambda}(t), X >$ negative for any X inside the cone $K(t)$, and such that

$$\mathcal{H}\big(\bar{y}(t), \bar{u}(t), \bar{\lambda}(t)\big) =< \bar{\lambda}(t), f\big(\bar{y}(t), \bar{u}(t)\big) >= 0 \ ;$$

then, intuitively,

$$\mathcal{H}\big(\bar{y}(t), v, \bar{\lambda}(t)\big) =< \bar{\lambda}(t), f\big(\bar{y}(t), v\big) >\leq 0 \ ,$$

for any control v in $\mathbf{U}$. Furthermore, when $u = \bar{u}(t)$, the 1-form $\bar{\lambda}(t)$ is also proved to satisfy the following adjoint equation of the translation (3-3)

(3-7) $$\dot{\lambda}_\alpha = -\frac{\partial \mathcal{H}}{\partial x^\alpha} \ .$$

These properties are proved to be realized for almost every t and also necessarily for t_1, thanks to the limit cone $K(t_1)$. This, intuitively, leads to the contention of the Maximum Principle.

Let $\lambda_0, \lambda_1, \ldots, \lambda_n$ be introduced as auxiliary functions, namely the $(n+1)$ components of a 1-form over $\mathbb{R} \times M$, $\lambda : [a, b] \longrightarrow \mathbb{R} \times M$, supposed to be solutions of the differential equation (3-7) for almost all $t,\ t \in [a, b]$. Again, as soon as $x(a)$ and an admissible $\tilde{u}$ are chosen, the Trajectory $(\tilde{y}, \tilde{u})$ is completely determined and then $\tilde{\lambda} : [a, b] \longrightarrow \mathbb{R}^{n+1}$, up to a positive multiplicative factor, as well. The solution $\tilde{\lambda}$ of the linear equation (3-7) is also absolutely continuous with measurable derivatives. Let us denote by $\mathcal{T}^*_{x_0,x_1}$, and call *lifted Trajectories* the corresponding triplets

$$\mathcal{T}^*_{x_0,x_1} = \left\{ \ (\tilde{y},\tilde{u},\tilde{\lambda}) \ / \ (\tilde{y},\tilde{u}) \in \mathcal{T}_{x_0,x_1} \ \right\} .$$

Now it is possible to give the following

3.8. Definition. — *Let us call Hamiltonian of Control Theory, the C^∞-function $\mathcal{H} : (\mathbb{R} \times M) \times \mathbf{U} \times \mathbb{R}^{n+1} \longrightarrow \mathbb{R}$ such that*

$$\mathcal{H}(y,u,\lambda) =< \lambda, f(y,u) > \ \bullet$$

Then, the differential equations (**C**) and (3-7) can be reformulated as

$$\text{(3-}\mathcal{H}\text{)} \qquad \begin{cases} (3-\mathcal{H}-1) & \dot{y}^\alpha & = & \dfrac{\partial \mathcal{H}}{\partial \lambda_\alpha} \\ (3-\mathcal{H}-2) & \dot{\lambda}_\alpha & = & -\dfrac{\partial \mathcal{H}}{\partial y^\alpha} \end{cases}$$

with $\alpha = 0, 1, \ldots, n$.

3.9. Notation. — *Let us denote by $\mathcal{T}^*_{\mathcal{H}}$ the set of lifted Trajectories satisfying (3-$\mathcal{H}$)* $\bullet$

3.10. Remark. — *The map $\mathcal{H}$ does not depend on y^0, so that the zero coordinate equation implies immediately the following result : along any lifted Trajectory,*

$$\dot{\lambda}_0 = -\frac{\partial \mathcal{H}}{\partial y^0} = 0 \ ;$$

then, $\tilde{\lambda}_0$ remains constant all along a lifted Trajectory, and furthermore constant and non-positive all along a lifted optimal Trajectory, because of the chosen sign of $\bar{\lambda}$ $\bullet$

3.11. Maximum Principle. — *Let $\bar{u} : [a,b] \longrightarrow \mathbf{U}$ be a measurable control, such that the associated lifted Trajectory $(\bar{y},\bar{u},\bar{\lambda})$ lies in $\mathcal{T}^*_{x_0,x_1}$. Then, if $(\bar{y},\bar{u},\bar{\lambda})$ is optimal on $[a,\bar{t}_1] \subset [a,b]$,*

1°) *$(\bar{y},\bar{u},\bar{\lambda})$ lies in $\mathcal{T}^*_{\mathcal{H}}$,* $\quad \bar{\lambda} \neq 0$;

2°) *there exists a real non-negative constant* B*, such that, for almost every* t*,* $t \in [a, t_1]$

$$(3\text{-}\mathcal{M}) \qquad \begin{cases} (i) & \mathcal{H}\big(\bar{y}(t), \bar{u}(t), \bar{\lambda}(t)\big) = \sup_{v \in \mathbf{U}} \mathcal{H}\big(\bar{y}(t), v, \bar{\lambda}(t)\big) = \mathcal{M}\big(\bar{y}(t), \bar{\lambda}(t)\big) \\ (ii) & \bar{\lambda}_0(t) = -B \leq 0 \quad , \quad \mathcal{M}\big(\bar{y}(t), \bar{\lambda}(t)\big) = 0 \quad \bullet \end{cases}$$

3.12. Remark. — *In case it would be specified that* t_1 *is fixed and equal to* b*, the maximum principle is unchanged except the very last conclusion : there exists a real non negative constant* B*, such that, for almost every* $t, t \in [a, b]$*,*

$$(3\text{-}\mathcal{M}\text{-b}) \qquad \begin{cases} (ii) \quad \bar{\lambda}_0(t) = -B \leq 0 \quad , \quad \mathcal{M}\big(\bar{y}(t), \bar{\lambda}(t)\big) \textit{ is constant } \bullet \end{cases}$$

4. THE SINGULAR CASE : AN EXAMPLE

In this section, we show how the formalism of Control Theory has to be used from the very beginning of the theory of singular sub-Riemannian structures in order to give a meaning to the quadratic form G. The new formalism leads us to claim that, in the neighbourhood of singular points,

(i) the singular sub-Riemannian metric has to be chosen carefully ;

(ii) it is impossible to extend the metric G_x, defined on $\mathcal{E}_x$, to any Riemannian metric $\tilde{G}_x$, defined on T_xM (Theorem (4-8)).

4.1. Example. — To point out the difficulties which could occur in the singular case with respect to the sub-Riemannian metric, if not chosen carefully, we will have a look at the very simple following example. On $M = \mathbb{R}^2$, let us consider the module $\mathcal{E}$ generated by

$$\text{(4-1)} \qquad \begin{cases} \varepsilon_1 &= \dfrac{\partial}{\partial x} \\ \varepsilon_2 &= x\dfrac{\partial}{\partial y} \; . \end{cases}$$

Let us suppose that each fiber $\mathcal{E}_{(x,y)}$ of $\mathcal{E}$ is provided with a scalar product $G_{(x,y)}$ such that, for any C^∞-vector fields X and Y in $\mathcal{E}$, the map $G_{(\cdot,\cdot)}(X,Y) : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is C^∞. Then, whatever the curve $\gamma : [a,b] \longrightarrow \mathbb{R}^2$, horizontal, i.e., $\dot\gamma(t) \in \mathcal{E}_{\gamma(t)}$ a.e., and of class H^1 (see section 6), we can define its energy

$$E_G(\gamma;[a,b];t) = \frac{1}{2}\int_a^b G_{\gamma(t)}\left(\dot\gamma(t),\dot\gamma(t)\right)dt$$

and its length

$$l_G(\gamma) = \int_a^b \sqrt{G_{\gamma(t)}\left(\dot\gamma(t),\dot\gamma(t)\right)}\,dt \; .$$

The generating Hörmander condition is verified ; then, any two points in $\mathbb{R}^2$ can be joined by a horizontal H^1-curve, and it is then possible to define a map with non-negative values

$$\delta_G\big((x_0,y_0),(x_1,y_1)\big) \;=\; \inf \; \Big\{ l(\gamma) \quad / \quad \gamma : [0,1] \longrightarrow \mathbb{R}^2, \quad \gamma \in H^1 \\ \gamma(0) = (x_0,y_0), \quad \gamma(1) = (x_1,y_1) \Big\} \; .$$

The question is "what are the sufficient conditions on G in order to make δ a distance ?" Let us develop two distinct simple examples.

(4-1-i) - If G is the induced metric on $\mathcal{E}_{(x,y)}$ by the canonical metric of $M = \mathbb{R}^2$. Let us consider the broken lines $\gamma_n : [0, \frac{n+2}{n}] \longrightarrow \mathbb{R}^2$, such that A $= \gamma_n(0) = (0,0)$, $\gamma_n(\frac{1}{n}) = (\frac{1}{n},0)$, $\gamma_n(\frac{n+1}{n}) = (\frac{1}{n},1)$, B $= \gamma_n(\frac{n+2}{n}) = (0,1)$. These curves, γ_n, are horizontal and their length is $\frac{n+2}{n}$, thus $\delta_g(A,B) \leq 1$. But if $\bar\gamma$ is any horizontal path joining A to B, there exists n such that

$$l_G(\gamma_n) < l_G(\bar\gamma) \; .$$

Thus $\delta_G(\mathrm{A},\mathrm{B}) = 1$, and **there does not exist any horizontal length minimizing path joining A to B**.

(4-1-ii) - If G is the metric induced by the quadratic form

$$ds^2 = dx^2 + x^2 dy^2 \ ,$$

let us consider the same sequence of paths joining A to B. Then, here

$$l_G(\gamma_n) = \frac{\sqrt{3}}{n} \quad \text{and} \quad \inf l_G(\gamma_n) = 0 \ ,$$

implies that

$$\delta_G(A,B) = 0 \ .$$

These examples show that G cannot be chosen without caution, in some sense it has to be bounded from below. This result is justifying the way we shall define G in this section.

We shall give now a **particular notion of a singular sub-Riemannian manifold.** Let $(M, \mathbf{g}, \mathcal{E}, g)$ be an n-dimensional, paracompact C^∞-manifold M, $\mathbf{g}$ its Riemannian metric, TM its tangent bundle, T^*M its cotangent bundle, $\mathcal{E}$ a rank p, locally free C^∞-module of vector fields, $(p \leq n)$. Similarly to the regular case, let us call "horizontal" the vector fields in $\mathcal{E}$ ($\mathcal{E}_x \subset T_xM \quad x \in M$), the dimension of $\mathcal{E}_x$, $p(x)$ is a lower semi-continuous function of maximum value p. Furthermore, $g_x : T_x^*M \longrightarrow T_xM$, is a C^∞-field of linear maps, positive and symmetric in the sense that for all $X, Y \in T_xM$ and for all $\xi_x \in g_x^{-1}(X_x)$, for all $\eta_x \in g_x^{-1}(Y_x)$,

$$< g_x\xi_x, Y_x > = < g_x\eta_x, X_x >, \qquad \text{and} \quad < g_x\xi_x, X_x > \geq 0 \ ,$$

and such that

$$\operatorname{Im} g_x = \mathcal{E}_x \ .$$

The module $\mathcal{E}$ is verifying the so-called Hörmander condition, i.e., all the iterated Lie derivatives of local horizontal vector fields by local horizontal vector fields, above a point x of M, generate T_xM.

Now, let X_x be the value of a horizontal vector field above x, and X any local horizontal vector field extending X_x ; let us use the same notations as in the regular case $\mathcal{E}_1(X)_x = \mathcal{E}_x$, $\mathcal{E}_2(X)_x = \mathcal{E}_x + [X, \mathcal{E}]_x$, $\mathcal{E}_k(X)_x = \mathcal{E}_x + [X, \mathcal{E}_{k-1}(X)]_x$, and $p_k(X)_x$ the dimension of $\mathcal{E}_k(X)_x$. The vector space $\mathcal{E}_2(X)_x$ does not depend on the choice of the locally extending fields X in $\mathcal{E}$, but only on the module $\mathcal{E}$ and on the value of X_x at x ; more generally, the vector space $\mathcal{E}_k(X)_x$ depends only on the $(k-2)$-jet of the field X in $\mathcal{E}$. The non-decreasing sequence

$$\Big(p_1(X)_x, \ldots, p_k(X)_x, \ldots\Big)$$

is such that, for any $k \geq 1$, $p_k(X) : M \longrightarrow \mathbb{N}$ is lower semi-continuous. Let us denote now by

$$(\mathcal{E}_1)_x = \mathcal{E}_x,\ (\mathcal{E}_k)_x = \mathcal{E}_x + \sum_{X \in \mathcal{E}_x} [X, \mathcal{E}_{k-1}(X)]_x$$

and, as before, in the regular case, by $p_k(x)$ the dimension of $(\mathcal{E}_k)_x$, and the lower semi-continuous non-decreasing sequence, by

$$\Big(p_1(x), \ldots, p_k(x), \ldots\Big) ,$$

called **growth vector at** x of the module $\mathcal{E}$. The Hörmander condition merely means that

$$\forall x \in M,\ \exists r_0(x) \in \mathbb{N} \quad / \quad p_{r_0 - 1}(x) < n, \quad \forall r \geq r_0(x), \quad p_r(x) = n\ .$$

The map $r_0 : M \longrightarrow \mathbb{N}$ is upper semi-continuous. If $p(x)$ were a constant p, the structure would be regular as the one described in section 2.

The following two propositions will help us to use Control formalism in our own definition of singular sub-Riemannian geometry (see [Os] pp. 122–123).

4.2. Proposition. — *For any smooth manifold M and any integer $p \geq 0$ there is a one-to-one correspondence between smooth real vector bundles* $\mathbf{U}$ *of rank p over M and isomorphism classes of locally free $C^\infty(M)$-modules $\mathcal{E}$ of rank p* •

4.3. Proposition. — *For any smooth manifold M, let $\mathcal{E}$ be a locally free $C^\infty(M)$-module of fixed rank $p > 0$, and let $\mathcal{E}^*$ be the dual. Then $\mathcal{E}$ and $\mathcal{E}^*$ are modules of* *smooth sections of smooth coordinate bundles representing the same smooth p-plane bundle over M, say* $\mathbf{U}$ •

In order to translate the constraint of being horizontal for vector fields, we shall consider the real vector space $\mathbb{R}^p$ in which lie the controls, as the model Euclidean space for the fiber space of rank p, $\mathbf{U}$, the one described in proposition (4-3).

Let us denote h any Riemannian metric on the vector bundle $\mathbf{U}$, and $h^\sharp : \mathbf{U}^* \longrightarrow \mathbf{U}$, the canonical isomorphism associated to h. One gets the following diagram

(4-4)
$$\begin{array}{ccc} \mathbf{U}^* & \xleftarrow{\;H^*\;} & T^*M \\ {\scriptstyle h^\sharp}\downarrow\wr & & \downarrow{\scriptstyle g} \\ \mathbf{U} & \xrightarrow{\;H\;} & TM \\ {\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle P} \\ M & \xrightarrow{\;id\;} & M \end{array}$$

where H is a singular vector fiber bundle homomorphism above the identity, and $\mathcal{E}$ is the pushforward by H of the space of sections of $\mathbf{U}$. Let P be the natural projection $P : TM \longrightarrow M$.

4.5. Notation. — *Let us denote by* $H(x) \cdot s(x)$, *or* $H_x \cdot s(x)$, *the image through* H *of a local section* s *of* $\mathbf{U}$, *above the point* $x \in M$ •

The existence of H is guaranteed thanks to proposition (4-3), evidently the rank of the linear operator $H(x)$ is $p(x)$.

Then it is natural to choose as quadratic form G_x on $\mathcal{E}_x$ the one corresponding to the vector bundle morphism $g = H \circ h^\sharp \circ H^*$, making the diagram (4-4) commutative. So, G_x is completely determined above each point x, and we get the following

4.6. Definition and notation. — *Let us denote by* $(M, \mathcal{E}, g)$, *and call "sub-Riemannian manifold", an* n*-manifold* M, *provided with*

(i) a locally free rank p, $p < n$, *submodule of the module of vector fields on* M, *denoted by* $\mathcal{E}$, *which can be seen as the pushforward by some* C^∞ *fiber bundle*

homomorphism $H : \mathbf{U} \longrightarrow TM$ *of the space of sections of some rank* p *fiber space on* M : $\mathbf{U}$,

(ii) *a linear fiber bundle homomorphism* $g : T^*M \longrightarrow TM$, *such that* $g = H \circ h^\sharp \circ H^*$, *where* h *is any fiber metric on* $\mathbf{U}$, $h^\sharp : \mathbf{U}^* \longrightarrow \mathbf{U}$ *is the associated canonical fiber isomorphism between the dual fiber space of* $\mathbf{U}^*$ *and* $\mathbf{U}$.

A manifold provided with such a structure $(M, \mathcal{E}, g)$ *will be called "regular" if* $\mathrm{Im}\ g$ *is a subbundle of* TM, *of (constant) rank* p, **singular,** *if* $\mathrm{Im}\ g$ *is not of constant rank* •

Actually, the definition of $(M, \mathcal{E}, g)$ is stable with respect to the fiber bundle **isometries** $\varphi : \mathbf{U}' \longrightarrow \mathbf{U}$ for

$$g' = H' \circ h'^\sharp \circ H'^* = H \circ \varphi \circ h'^\sharp \circ \varphi^* \circ H^* = H \circ h^\sharp \circ H^* = g \ .$$

Now, because of the regularity of $h_x^\sharp$, we get the following

4.7. Proposition. — *For any* $x \in M$, *there is a one-to-one correspondence between horizontal vectors* X_x *in* $\mathrm{Im}\ H_x$ *and "control vectors"* s_x *such that*

$$s_x \in (\mathrm{Ker}\ H_x)^{\perp_h} \subset \mathbf{U}_x \ .$$

Furthermore, for any $\xi_x \in g_x^{-1}(X_x)$, $h_x^\sharp \circ H_x^*(\xi_x) = s_x$, *and it is possible to define a unique quadratic form* G_x *on* $\mathrm{Im}\ H_x$ *by setting*

$$G_x(X_x, X_x) = h_x\big(s(x), s(x)\big) = \inf\Big\{\ h_x\big(\sigma(x), \sigma(x)\big)\ /\ H_x \cdot \sigma(x) = X_x\ \Big\}\ .$$

Then, G_x *is a positive non-degenerate quadratic form on* $\mathrm{Im}\ H_x$, *and, for any two horizontal vectors* X_x *and* Y_x *of* $\mathcal{E}_x$, $\xi_x \in g_x^{-1}(X_x)$ *and* $\eta_x \in g_x^{-1}(Y_x)$, $s_1(x) = h_x^\sharp \circ H_x^* \cdot \xi_x$ *and* $s_2(x) = h_x^\sharp \circ H_x^* \cdot \eta_x$.

$$G_x(X_x, Y_x) =< \xi_x, Y_x >=< \eta_x, X_x >=< \xi_x, g_x\eta_x >$$

$$=< \eta_x, g_x\xi_x >= h_x\big(s_1(x), s_2(x)\big) \qquad \bullet$$

Let $\mathcal{N}$ be the annihilator of $\mathcal{E}$ in T^*M ; then, at each point x,

$$\mathrm{Ker}\, g_x \supset \mathrm{Ker}\, H_x^* \supset \mathcal{N}_x \ ,$$

they are equal if and only if $p(x) = p = \text{const.}$

Proof. Let ξ_x, η_x, be any two 1-forms of T_x^*M, h_x^{-1} be the quadratic form induced by h_x on $\mathbf{U}_x^*$; then,

$$< \xi_x \,,\, g_x\eta_x >=< \xi_x \,,\, (H_x \circ h_x^\sharp \circ H_x^*) \cdot \eta_x >=< H_x^* \cdot \xi_x \,,\, (h_x^\sharp \circ H_x^*) \cdot \eta_x >$$

$$= h_x^{-1}(H_x^* \cdot \xi_x \,,\, H_x^* \cdot \eta_x) \ .$$

Then, $< \xi_x \,,\, g_x\eta_x >$ is symmetric, and $< \xi_x \,,\, g_x\xi_x >$ is zero if and only if $\xi_x \in \mathrm{Ker}\, H_x^*$; we also get

$$\mathrm{Ker}\, g_x = \mathrm{Ker}\, H_x^* \ .$$

Furthermore, it is a well known result of linear algebra and the theory of quadratic forms that

$$\mathrm{Im}\, \left(h^\sharp \circ H^*\right)_x = (\mathrm{Ker}\, H_x)^{\perp_h} \ ,$$

and, for all σ_x in $\mathbf{U}_x$, there exists $s_x \in (\mathrm{Ker}\, H_x)^{\perp_h}$, such that

$$\sigma_x = s_x + \tau_x \text{ with } \tau_x \in \mathrm{Ker}\ H_x \ ,$$

and then,

$$h_x(\sigma_x, \sigma_x) = h_x(s_x, s_x) \ + \ h_x(\tau_x, \tau_x) \geq h_x(s_x, s_x) \ . \qquad \square$$

Our next remark will make obvious the essential difference between the singular case and the regular one. Let x be a point such that $\mathrm{Ker}\ H(x) \neq \{0\}$. Let V_x be a coordinate open cell of M, $\mathbf{g}$-neighbourhood for x, trivializing both the vector bundle TM and $\mathbf{U}$. As $\mathcal{E}$ is locally free, there exists a sequence $(x_j) \in V_x$ converging to x with respect to the topology induced by the metric $\mathbf{g}$, such that $H(x_j)$ is of maximal rank p. Thus, there exists a control u in $\mathrm{Ker}\,(H(x))$, such that $h_x(u, u) = 1$, and a sequence of controls (u_j), $u_j \in \left(\mathrm{Ker}\ H_{x_j}\right)^{\perp_h} \subset \mathbf{U}_{x_j} = (\mathbb{R}^p, h_{x_j})$, such that $h_{x_j}(u_j, u_j) = 1$, converging in the sense of the product $(\mathbf{g} \times h)$-topology to u. Then, to the sequence

$(x_j, u_j) \in \mathbf{U}$ is associated through H, a sequence of horizontal vectors $\big(x_j, H(x_j)\cdot u_j\big)$ converging necessarily to $(x,0)$ in TM, with respect to the regular metric $\mathbf{g}$, because of the smoothness of H, but such that

$$\forall j, \quad G\big(H_{x_j}\cdot u_j\,,\, H_{x_j}\cdot u_j\big) = h(u_j,u_j) = 1\ ,$$

because of the definition of G, (4-7), though of course

$$\lim\ \mathbf{g}\big(H_{x_j}\cdot u_j\ ,\ H_{x_j}\cdot u_j\big) = 0\ .$$

Thus, we get the following

4.8. Theorem. — *In singular sub-Riemannian geometry, if* Ker $H(x) \neq \{0\}$, *for some x, in any $\mathbf{g}$-neighbourhood of x, there exists a sequence of points (x_j), $\mathbf{g}$-converging to x, and a sequence of non-zero controls $\big(u_j \in \pi^{-1}(x_j)\big)$ such that*

$$\lim \frac{\mathbf{g}_{x_j}\big(H_{x_j}\cdot u_j\ ,\ H_{x_j}\cdot u_j\big)}{G_{x_j}\big(H_{x_j}\cdot u_j\ ,\ H_{x_j}\cdot u_j\big)} = 0\ .$$

So, it is impossible to extend the metric G_x, defined on $\mathcal{E}_x$, to any Riemannian metric $\tilde{G}_x$, defined on T_xM •

Actually, let $\mathbf{g}$, G be given, let K be a compact cell of M, and denote by Σ the set of singular points of H inside K and

$$U_{\mathbf{g}}M = \Big\{\ X \in TM\ /\ \mathbf{g}(X,X) = 1\ \Big\}\ .$$

Then, there exists $\delta > 0$ and a horizontal thickening δ-strip of Σ with respect to $\mathbf{g}$, namely

$$\mathrm{HStrip}_\delta\Sigma = \Big\{\ (\exp_{\mathbf{g}})_x\ tX\ /\ 0 \le t \le \delta\ ,\ x \in \Sigma\ ,\ X \in U_{\mathbf{g}}\Sigma\ \cap\ \mathcal{E}\ \Big\}$$

such that, for any horizontal vector field such that $\mathbf{g}(X,X) = 1$, inside $\mathrm{HStrip}_\delta\Sigma$, $G(X,X) \geq 1$, and, outside $\mathrm{HStrip}_\delta\Sigma$, there exist positive constants A and B such that $A < G(X,X) < B$. Thus we get the following theorem

4.9. Theorem. — *Let K be a compact cell in M and $U_{\mathbf{g}}K$ be the unitary fiber bundle with respect to the metric* $\mathbf{g}$*. Then, there exist two strictly positive constants, a and A, such that*

$$\forall X \in U_{\mathbf{g}}K, \qquad a < G(X,X) < A \ ,$$

if and only if all points in K are regular •

We get also the following

4.10. Corollary. — *For any positive numbers δ and ε, it is possible to choose a Riemannian metric* $\mathbf{g}$ *on M such that, for any horizontal vector field $Y \neq 0$,*

$$G(Y,Y) > \mathbf{g}(Y,Y) \quad \textit{inside} \quad \mathrm{HStrip}_{\delta}\Sigma \ ,$$

and

$$G(Y,Y) = \mathbf{g}(Y,Y) \quad \textit{outside} \quad \mathrm{HStrip}_{\delta+\varepsilon}\Sigma \, \bullet$$

4.11. Definition. — *From now on, we suppose that* $\mathbf{g}$ *is chosen in order to have everywhere in K*

$$\forall Y \in \mathcal{E}, \quad Y \neq 0, \quad G(Y,Y) \geq \mathbf{g}(Y,Y) \, \bullet$$

4.12. Example (4-1) revisited. — We go back to Example (4-1), and now, we shall use one of these metrics described in this section, with necessarily $\mathbf{U} = TM = \mathbb{R}^4$, choosing as h, the canonical metric on each $\mathbf{U}_{(x,y)} = \mathbb{R}^2$. Then, the matrices of g and H, in the frames $\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\}$, and $\{dx, dy\}$ are such that

(4-1-iii)
$$g = H \circ h^{\sharp} \circ H^* = \begin{pmatrix} 1 & 0 \\ 0 & x^2 \end{pmatrix} .$$

Let us choose $\mathbf{g}$ as the canonical metric, then

$$\lim_{x \to 0} \mathbf{g}(\varepsilon_2, \varepsilon_2)/G(\varepsilon_2, \varepsilon_2) = 0 \ .$$

The horizontal curve $\gamma : [0,1] \longrightarrow \mathbb{R}^2$, given by $\Big\{\gamma(t) = (1-t, 1-t)/\ t \in [0,1[\Big\}$, has $\mathbf{g}$-length $\sqrt{2}$, whereas its G-length is infinite. A horizontal curve going to the y-axis has finite G-length if and only if the vertical component of the velocity goes to zero faster than x on the curve.

As a matter of fact, the horizontal curves $\Big\{\gamma(t) = \big(1-t, (1-t)^\alpha\big)/\ t \in [0,1[\Big\}$ have finite G-length as soon as $\alpha \geq 1$, as soon as they arrive to the origin perpendicularly to the y-axis, in the horizontal direction, otherwise naturally the vertical part of the tangent vector tends to a non-horizontal vector of infinite G-norm.

We shall resume Example (4-1) in section 7, illustrating the construction of geodesics by the application of the Maximum Principle.

5. HORIZONTAL CURVES, LENGTH AND ENERGY

In this section we will show that, to any horizontal path $\gamma : I \longrightarrow M$, where I is an interval $[a,b] \in \mathbb{R}$, can be associated a unique control $s : I \longrightarrow \gamma^*\mathbf{C}$ and a unique 1-form $\xi : I \longrightarrow \gamma^*(T^*M)$ with nice properties. Then, among the horizontal curves joining two given points of M, as in Riemannian geometry, seeking the minimum of G-length is equivalent to seeking the minimum of the G-energy. Let us consider $\sigma : I \longrightarrow \mathbf{C}$, a measurable map, such that $\pi \circ \sigma(t) = \gamma(t)$ is an absolutely continuous curve in M, i.e., necessarily $\pi \circ \sigma$ is an injective map. Then, $(H \circ \pi \circ \sigma) \cdot \sigma$ is a section of TM above the curve, using the following notations

$$H \circ \sigma(t) = H\Big(\pi \circ \sigma(t)\Big) \cdot \sigma(t)\ ,$$

where $\cdot$ is the matrix multiplication.

5.1. Definition. — *A curve, $\gamma : [a,b] \longrightarrow M$, is called "horizontal", if there exists above $\gamma(t)$ a measurable section of $\mathbf{C}$, $\sigma(t)$, such that*

$$\forall t \in [a,b], \quad \gamma(t) = \pi \circ \sigma(t) = \pi \circ \sigma(a) + \int_a^t (H \circ \pi \circ \sigma) \cdot \sigma\ (t)\ dt\ ,$$

or equivalently

$$(5\text{-}2) \qquad \begin{cases} \dot{\gamma} &= (H \circ \gamma) \cdot \sigma \quad \text{a.e.} \\ \gamma &= \pi \circ \sigma \quad \bullet \end{cases}$$

Thus, the curve γ is absolutely continuous, its tangent vector exists a.e., and when it exists, it belongs to Im (H).

Now, we want to prove the following theorem

5.3. Theorem. — *Let γ be a horizontal curve defined as above. Then, above γ, there exists a unique control $s : I \longrightarrow \gamma^*\mathbf{C}$, and a unique 1-form $\xi : I \longrightarrow \gamma^*(T^*M)$, modulo $\mathbf{g}$, and modulo a set of t's of measure zero such that*

$$\begin{aligned} s(t) &\in \Big(\text{Ker } H_{\gamma(t)}\Big)^{\perp_h} \quad a.e. \ , \\ \xi(t) &\in \Big(\text{Ker } (H^*_{\gamma(t)}\Big)^{\perp_{\mathbf{g}}} \quad a.e. \quad \bullet \end{aligned}$$

Proof. In order to do this, let us consider a covering of I by means of sets A_k , $0 \leq k \leq p$, where

$$A_k = \Big\{ \ t \in I \ / \ \dim \text{Im } H_{\gamma(t)} = k \ \Big\} \ .$$

Let us call $p(t)$ the rank of $H_{\gamma(t)}$. The function $p : I \longrightarrow \mathbb{N}$ is well defined for any t and is lower semi-continuous ; then,

$$\bigcup_{l>k} A_l \ = \ p^{-1} \,]k, +\infty[$$

is open in I, and is, then, a union of open intervals, then it is measurable. For any $k \in \mathbb{N}$, the set A_k is given by

$$A_k \ = \ p^{-1} \,]k-1, +\infty[\quad \backslash \quad p^{-1} \,]k, +\infty[\ .$$

The set A_k is then measurable as difference of two measurable sets. Further it is the disjoint union of semi-open intervals and single points. Let us call them I_{k,μ_k} / $\mu_k \in \mathcal{M}_k$. Then,

$$I = \bigcup_{0 \le k \le p} \bigcup_{\mu_k \in \mathcal{M}} I_{k,\mu_k} \qquad \text{where} \quad \mathcal{M} = \bigcup_{0 \le k \le p} \mathcal{M}_k \ .$$

The set $\mathcal{M}$ could be any huge index set, and the union is a disjoint union. The measure of the subset of I, union of I_{k,μ_k} which are single points, is not necessarily zero.

The curve γ is supposed to be absolutely continuous. Then, the set of points where its tangent vector is not well defined is of measure zero. The set of points which are boundaries of I_{k,μ_k} on the whole of I could be of measure not zero, then necessarily in this case, the curve γ goes through the "main part" of the set of boundary points with a well defined velocity vector. We have just seen that the total measure of the set of those points, with no velocity vector, either because of the curve itself at a regular point of H or because of the singularities of H is necessarily zero. There is an illustration of some of these situations associated to the singularities of H in the example of the section 8.

Above each I_{k,μ_k} the fiber spaces $\gamma^*(\mathbf{C})$ and $\gamma^*(T^*M)$ are trivial fiber spaces. Then there exists a trivialization such that

$$\begin{array}{rclcl} \gamma^*(\mathbf{C})/I_{k,\mu_k} & = & I_{k,\mu_k} \times \left(\mathrm{Ker}\ H\right)^{\perp_h} \times \mathrm{Ker}\ H & = & I_{k,\mu_k} \times \mathbb{R}^k \times \mathbb{R}^{p-k} \ , \\ \gamma^*(T^*M)/I_{k,\alpha_k} & = & I_{k,\mu_k} \times \left(\mathrm{Ker}\ H^*\right)^{\perp_{\mathbf{g}}} \times \mathrm{Ker}\ H^* & = & I_{k,\mu_k} \times \mathbb{R}^k \times \mathbb{R}^{n-k} \ . \end{array}$$

The trivialization fiber frames above each I_{k,μ_k} can be made orthonormal with respect to h in $(\mathbf{C})$, and with respect to $\mathbf{g}^{-1}$ in $\gamma^*(T^*M)$, where $\mathbf{g}^{-1}$ is the non-degenerate positive quadratic form induced by $\mathbf{g}$ on T^*M. This makes the matrix of the restriction of $h^\sharp \circ H^*$ to I_{k,μ_k} from $\mathbb{R}^k$ to $\mathbb{R}^k$ diagonal and non-degenerate.

Then, if we come back to σ, the control given in the definition of the horizontal curve γ, on each I_{k,μ_k}, the restriction of σ, in the previous trivialization, can be written

$$\sigma_{k,\mu_k}(t) = \left(\ \gamma_{k,\mu_k}(t), s_{k,\mu_k}(t), u_{k,\mu_k}(t) \ \right) \ .$$

Then, on each I_{k,μ_k} ,

$$\begin{array}{rcl} \pi \circ s \ (t) & = & \pi \circ \sigma \ (t) \\ H \circ s \ (t) & = & H \circ \sigma \ (t) \ . \end{array}$$

And

$$\gamma(t) = \gamma(a) + \sum \int_{I_{k,\mu_k} \cap [0,t]} \big((H \circ \gamma) \cdot s_{k,\mu_k}\big)\ (t)\ dt$$

and, as a consequence, the section s of $\gamma^* \mathbf{C}$, defined by these restrictions on each I_{k,μ_k}, is a measurable section. The associated 1-form, ξ, in $\big(\operatorname{Ker} H^*\big)^{\perp_{\mathbf{g}}}$, is defined modulo $\mathbf{g}$. □

Furthermore, for every t in I, we get

$$h_{\gamma(t)}(s(t), s(t)) = \inf\Big\{\ h_{\gamma(t)}\big(\sigma(t), \sigma(t)\big) \quad / \quad \gamma = \pi \circ \sigma\ \Big\}$$

because of the definition of G, through the proposition (4-7). It is possible to define formally the following positive functionals on horizontal curves, but in the singular case they are not necessarily finite, even if the $\mathbf{g}$-distance of the two points is finite (Example (4-1-iii)).

5.4. Definition. — *Let us denote by*

$$l_G(\gamma) = \int_a^b \sqrt{G(\dot\gamma, \dot\gamma)_t}\ dt = \int_a^b \sqrt{h_{\gamma(t)}\big(s(t), s(t)\big)}\ dt$$

and call it, when it exists, the G-length of the curve γ.

Let us denote by

$$E_G(\gamma; t) = \frac{1}{2}\int_a^b G(\dot\gamma, \dot\gamma)\ dt = \frac{1}{2}\int_a^b h_{\gamma(t)}\big(s(t), s(t)\big)\ dt$$

and call it the G-energy of the t-parametrized curve γ •

But, we know that

$$l_G(\gamma) = \int_a^b \sqrt{G(\dot\gamma, \dot\gamma)_t}\ dt = \int_a^b \sqrt{h_{\gamma(t)}\big(s(t), s(t)\big)}\ dt = \int_a^b \sqrt{< \xi(t), \dot\gamma(t) >}\ dt\ .$$

Similarly,

$$E_G(\gamma; t) = \frac{1}{2}\int_a^b G(\dot\gamma, \dot\gamma)\ dt = \frac{1}{2}\int_a^b h_{\gamma(t)}\big(s(t), s(t)\big)\ dt = \frac{1}{2}\int_a^b < \xi(t), \dot\gamma(t) >\ dt$$

where s, ξ are the ones just defined in Theorem (5-3). In both cases, the last integral does not depend on the chosen metric $\mathbf{g}$, and, as soon as one of these three integrals exist, necessarily, the other two do too.

Moreover, it is well known that, because of the Schwarz inequality,

$$\Big(\int_a^b |\dot\gamma(t)|_G \, dt\Big)^2 \leq (b-a) \int_a^b |\dot\gamma(t)|_G^2 \, dt$$

a horizontal parametrized, absolutely continuous curve which realizes the minimum for the energy, if it ever exists, has necessarily its parameter proportional to the G-arc length such that

$$|\dot\gamma(t)|_G = l_G(\gamma)/(b-a) = \text{constant}$$

(see [L-S] for the existence of such reparametrizing of horizontal curves). It is then also a minimum for the length. Conversely, if a horizontal parametrized a.c. curve realizes the minimum of the length, among the reparametrized curves defined on the same interval $[a,b]$, the one with its parameter proportional to the G-arc length realizes the minimum for the energy. Then, when looking for the minimum of length, among curves which are defined on a given interval $[a,b]$, we are led to look for energy minimizing curves among those horizontal a.c. curves which are defined on the same interval $[a,b]$.

6. DISTANCE AND ENERGY

In this section, we suppose the Riemannian manifold $(M, \mathbf{g})$ (definition (4-11)) connected and complete as before. Let x_0 and x_1 be any two points in M, and we want to prove that the sub-Riemannian distance is achieved on some horizontal absolutely continuous curve, even in the singular case. The proof will be adaptated from the proof used by A. Bellaïche [Bel].

Thanks to the Hörmander condition and the theorem of Chow [Ch] the two distinct points x_0 and x_1 can be joined, at least, by one horizontal piecewise C^1-curve $\tilde\gamma : I = [a,b] \longrightarrow M$, with G-energy $E_G(\tilde\gamma) = \mathrm{A} > 0$. Let us consider the set of parametrized absolutely continuous curves

$$\mathcal{C}_{x_0,x_1}(I;\mathrm{A}) = \{\ \gamma : I \longrightarrow M \quad / \quad \gamma(a) = x_0\,,\ \gamma(b) = x_1\,, \quad E_{\mathbf{g}}(\gamma) \leq \mathrm{A}\ \}$$

provided with the topology of uniform convergence associated to $\mathbf{g}$ such that

$$\forall X \in \mathcal{E}, \quad \mathbf{g}(X,X) \leq G(X,X) \Longrightarrow \forall \gamma \in \mathcal{C}_{x_0,x_1}(I;\mathrm{A}), \quad E_{\mathbf{g}}(\gamma) \leq E_G(\gamma)\ .$$

Thanks (4-7), it is still possible to define the energy of these horizontal curves.

Let us denote by

$$\mathcal{H}_{x_0,x_1}(I;\mathrm{A}) = \{\ \gamma : I \longrightarrow M \quad / \quad \gamma(a) = x_0\,,\ \gamma(b) = x_1\,,\ \dot{\gamma} \in \mathcal{E}\,,\ E_G(\gamma) \leq \mathrm{A}\ \}\ .$$

Clearly, $\mathcal{H}_{x_0,x_1}(I;\mathrm{A}) \subset \mathcal{C}_{x_0,x_1}(I;A)$, because of (4-11).

The set $K = \bigcup_{E_{\mathbf{g}}(\gamma)\leq A} \operatorname{Im}\gamma$ is a compact subset of M with respect to the metric $\mathbf{g}$. Moreover, x_0, x_1, and the images of the curves in $\mathcal{H}_{x_0,x_1}(I;\mathrm{A})$ lie in K.

The set $\mathcal{H}_{x_0,x_1}(I;\mathrm{A})$ can be provided with the H^1 topology, i.e., the topology induced by the H^1-distance, defined as follows

$$d^2_{H^1}(\gamma_1,\gamma_2) = d^2_{\mathbf{g}}\big(\gamma_1(a),\gamma_2(a)\big) + \int_a^b h\left(s_1(t)-s_2(t), s_1(t)-s_2(t)\right)\,dt\ ,$$

where γ_1 and γ_2 are two curves of $\mathcal{H}_{x_0,x_1}(I;\mathrm{A})$, s_1 and s_2 are the associated unique sections of Theorem (5-3), and the integral is computed on the union of disjoint intervals $I_{k_1,\mu_{k_1}} \cap I_{k_2,\mu_{k_2}}$. It is worth paying attention to the fact that C^1 horizontal curves are dense in $\mathcal{H}_{x_0,x_1}(I;K;\mathrm{A})$ with respect to the H^1-topology, (Süssman, private communication).

We shall prove the following

6.1. Theorem. — *Let $(M,\mathcal{E},g)$ be a singular sub-Riemannian manifold of class at least C^1, complete with respect to some Riemannian metric. Let x_0, x_1, I, A be defined as above, then, among the curves of $\mathcal{H}_{x_0,x_1}(I;\mathrm{A})$, there exists at least one horizontal curve γ, such that the infimum of the energy is achieved on γ* •

As we have seen in section 5, the infimum of length is also achieved on γ.

6.2. Definition. — *Let $(M, \mathcal{E}, g)$ be a singular sub-Riemannian manifold of class at least C^1. Let $x_0, x_1, I, K, \mathrm{A}$ be defined as above, we shall call "horizontal distance" between x_0 and x_1, and denote $d_G(x_0, x_1)$,*

$$d_G(x_0, x_1) = \inf\Big\{ \sqrt{2\ (b-a)\ E_G(\gamma; t)}\ /\ \gamma \in \mathcal{H}_{x_0,x_1}(I; K; \mathrm{A})\Big\} \bullet$$

It is well known that the infimum of a lower semi-continuous function on a compact set is achieved. So Theorem (6-1) will follow from the following lemma

6.3. Lemma. — *Inside the functional space $\mathcal{C}(I; M)$ provided with the topology of uniform convergence*

$$(i) \quad \mathcal{H}_{x_0,x_1}(I; \mathrm{A}) \textit{ is compact in } \mathcal{C}_{x_0,x_1}(I; \mathrm{A})\ ;$$
$$(ii) \quad E_G \textit{ is lower semi-continuous on } \mathcal{H}_{x_0,x_1}(I; \mathrm{A}) \bullet$$

Proof (of Lemma 6-3). We have to first prove that for all t, $t \in [a, b]$, $\mathcal{H}_{x_0,x_1}(I; \mathrm{A})(t)$ is compact in K, and second that $\mathcal{H}_{x_0,x_1}(I; \mathrm{A})$ is equicontinuous. Then, Ascoli's theorem implies (i). On the way, it will be necessary to prove (ii).

Consider a sequence $(\gamma_j = \pi \circ s_j)$ in $\mathcal{H}_{x_0,x_1}(I; \mathrm{A})$ converging uniformly to a continuous curve γ in $\mathcal{C}_{x_0,x_1}(I; \mathrm{A})$, with respect to the metric $\mathbf{g}$. For every j, for every $t \in [a, b]$, we have

$$\gamma_j(t) = \gamma_j(a) + \int_a^t H(\gamma_j(\tau)) \cdot u_j(\tau) d\tau\ .$$

We already know that Im (γ) lies in K which is $\mathbf{g}$-compact by definition.

The sequence $E_G(\gamma_j)$ of strictly positive real numbers is bounded by A, and it is possible to extract from (γ_j) a subsequence (indexed by the same letter) such that $E_G(\gamma_j)$ converges to $\liminf E_G(\gamma_j) = E_0 \leq A$.

Let us create a finite subdivision

$$a = t_0 < t_1 < \ldots < t_k < \ldots < t_m = b \quad , \quad I_k = [t_{k-1}, t_k]$$

such that, for any j, $\gamma_j([t_{k-1}, t_k]) \subset \bar{V}_k \subset U_k$, where U_k's are TM and $\mathbf{C}$ trivializing coordinate open sets, and $\bar{V}_k$ is a compact cell. Now, because of the trivialization of

C above U_k, to every γ_j/I_k is bijectively associated a control vector $s_j/I_k : I_k \longrightarrow \mathbb{R}^p$ such that $s_j(t) \in (\operatorname{Ker} H_{\gamma_j(t)})^{\perp_h}$, and

$$\gamma_j(t) = \gamma_j(t_{k-1}) + \int_{t_{k-1}}^{t} (H \circ \gamma_j) \cdot s_j \; d\tau \; .$$

Then,

$$E_G(\gamma_j) = \sum_{k=1}^{m} \frac{1}{2} \int_{t_{k-1}}^{t_k} h_{\pi \circ s_j(\tau)}(s_j(\tau), s_j(\tau)) d\tau \leq \ \mathrm{A} \; .$$

The spaces

$$\mathcal{W}_k = \Big\{ \; u : I_k \longrightarrow \mathbb{R}^p \; \Big\}, \; \forall k \; , \; 0 \leq k \leq m$$

provided with the L^2 norm

$$\int_{t_{k-1}}^{t_k} h(u(t), u(t)) dt$$

are Hilbert spaces.

In the Hilbert space $\mathcal{W}_k$, the closed ball

$$\mathcal{B}(0, 2\mathrm{A})$$

is a weakly compact subset of $\mathcal{W}_k$. Thus, there exists a control function $v_k : I_k \longrightarrow \mathbb{R}^p$ and a subsequence (u_{j_k}) of $(u_{j,k}) = (s_j/I_k)$ such that v_k is the weak limit of (u_{j_k}).

As

$$\forall t \in I_k, \qquad \gamma(t) = \lim_{j \to +\infty} \gamma_j(t) \; ,$$

$$\gamma(t) = \lim_{j \to +\infty} \Big(\; \gamma_j(t_{k-1}) + \int_{t_{k-1}}^{t} (H \circ \gamma_j) \cdot u_{j,k} \; dt \; \Big) \; ,$$

or,
$$\gamma(t) = \gamma(t_{k-1}) + \lim_{j \to +\infty} \Big(\int_{t_{k-1}}^{t} (H \circ \gamma_j) \cdot u_{j,k} \; dt \; \Big) \; .$$

But, whatever $w : I_k \longrightarrow \mathbb{R}^p$, and $t \in I_k$, there exist strictly positive numbers B, C such that

$$\Big| \int_{t_{k-1}}^{t} (H \circ \gamma_j - H \circ \gamma) \cdot w \; d\tau \; \Big|_{\mathbf{g}}^2 \leq \; B \; \Big| \int_{t_{k-1}}^{t} (H \circ \gamma_j - H \circ \gamma) \cdot w \; d\tau \; \Big|_{(eucl)}^2$$

$$\leq \; B\ (t_k - t_{k-1}) \int_{t_{k-1}}^{t} |\, (H\circ\gamma_j - H\circ\gamma)\cdot w\,|^2_{(eucl)}\ d\tau$$

$$\leq \; B\ C\ (t_k - t_{k-1}) \int_{t_{k-1}}^{t} |\, (H\circ\gamma_j - H\circ\gamma)\cdot w\,|^2_{\mathbf{g}}\ d\tau\ ,$$

thanks to Schwarz inequality and the equivalence above $\bar{V}_k$ between $\mathbf{g}$ and the Euclidean metric on fibers of $TM/U = \mathbb{R}^n$.

Further, we can consider H as a section of the local trivial fiber bundle $U \times \mathbb{R}^p \otimes \mathbb{R}^n$ provided with the fiber metric $h^{-1}\otimes\mathbf{g}$, smooth on $\bar{V}_k$, and $\lim \gamma_j = \gamma$ uniformly with respect to the $\mathbf{g}$ topology. Then,

$$\lim_{j_k\to\infty} \int_{t_{k-1}}^{t} |\, (H\circ\gamma_j - H\circ\gamma)\cdot w\,|^2_{\mathbf{g}}\ d\tau = 0\ .$$

Moreover,

$$\begin{aligned}\gamma(t) &= \gamma(t_{k-1}) + \lim_{j\to+\infty}\Big(\int_{t_{k-1}}^{t_k} \mathbf{1}_{[t_{k-1},t]}\big((H\circ\gamma)\cdot s_j\big)\ (\tau)\ d\tau\ \Big)\ ,\\ &= \gamma(t_{k-1}) + \int_{t_{k-1}}^{t_k} \mathbf{1}_{[t_{k-1},t]}\big((H\circ\gamma)\cdot v_k\big)\ d\tau\ .\end{aligned}$$

So we get

$$\text{(6-4)} \qquad \forall t\in I_k\ , \qquad \gamma(t) = \gamma(t_{k-1}) + \int_{t_{k-1}}^{t} \big((H\circ\gamma)\cdot v_k\big)\ (\tau)\ d\tau\ .$$

Now, we begin with the sub-sequence (u_{j_1}), such that the restriction to I_1 converges weakly to the measurable function $v_1 : I \longrightarrow \mathbb{R}^p$. Then on I_1,

$$\forall t\in I_1\ , \quad \gamma(t) = \gamma(a)\ +\ \int_a^t \big((H\circ\gamma)\cdot v_1\big)\ (\tau)\ d\tau = \lim(\gamma_{j_1}(t))\ ,$$

because of the weak convergence. Now, starting from the previous global sequence $(\gamma_{j_1}(t))$, we extract from the associated (u_{j_1}/I_2) a new sub-sequence (u_{j_2}), weakly converging on I_2 to v_2, such that

$$\forall t\in I_2, \quad \gamma(t) = \gamma(t_1)\ +\ \int_{t_1}^{t} \big((H\circ\gamma)\cdot v_2\big)\ (\tau)\ d\tau\ .$$

Finally, collecting the results step by step, from 1 to m, the last extracted subsequences $(\gamma_{j_m}(t))$ and the associated (u_{j_m}) are well defined on the whole interval I, and (u_{j_m}) admits as weak limit the measurable map v, such that $v/I_k = v_k,\ 1 \leq k \leq m$. One gets, globally,

$$\gamma(t) = \gamma(0) \ + \ \int_0^t \big((H \circ \gamma) \cdot v\big)\ (\tau)\ d\tau\ . \tag{6-5}$$

This relation means that the limit curve γ is absolutely continuous and horizontal for, since H, though singular, is a fiber bundle homomorphism, $v_k(t)$ belongs to $\pi^{-1}(\gamma(t))/I_k$ a.e. and there exists a measurable section of $\mathbf{C}$ locally described by v_k such that

$$\frac{d\gamma}{dt}(t) = (H \circ \gamma) \cdot v(t)$$

and $\gamma(t) = P \circ \dfrac{d\gamma}{dt}(t) = \pi \circ v(t)$, where P is the canonical projection $TM \longrightarrow M$. Furthermore, to prove the relative compactness of $\mathcal{H}_{x_0,x_1}(I;\mathrm{A})$ with regard to $\mathcal{C}_{x_0,x_1}$ $(I;\mathrm{A})$, it remains to prove $E_G(\gamma) < \mathrm{A}$, but the lower semi-continuity will imply

$$E_G(\gamma) \leq \liminf E_G(\gamma_j) \leq \mathrm{A}\ ,$$

and the proof is over.

To prove that E_G is lower semi-continuous, let us remark that for any w_k, and any $w_k^* \in \mathbb{R}^p$, and (γ_j) being the extracted sub-sequence of the last step

$$\lim_{j\to+\infty} \big(h_{\gamma_j(t)}(w_k, w_k^*) - h_{\gamma(t)}(w_k, w_k^*)\big) = 0\ ,$$

but

$$\lim_{j\to+\infty} \big(h_{\gamma_j}(t)(s_j - v, s_j - v)\big) \geq 0$$

implies, in restriction to I_k,

$$\lim_{j\to+\infty} \Big(\ E_G(\gamma_j) - \int_a^b h_{\gamma_j(t)}(s_j, v)\ dt + \frac{1}{2}\int_a^b h_{\gamma_j}(t)(v,v)\ dt\ \Big) \geq 0\ .$$

Then, for the restrictions to $I_k' s$,

$$\liminf\ E_G(\gamma_j) + E_G(\gamma) \geq \lim_{j\to+\infty} \Big(\ \int_a^b h_{\gamma(t)}(s_j, v)\ dt\ \Big)\ ,$$

but v is the weak limit of s_j along I_k, and

$$\lim_{j\to+\infty}\left(\int_a^b h_{\gamma(t)}(s_j, v)\ dt \right) = 2\ E_G(\gamma)\ ;$$

so

$$\liminf\ E_G(\gamma_j) \geq E_G(\gamma)\ .$$

Then, $E_G(\gamma) \leq A$ and $\mathcal{H}_{x_0,x_1}(I;\mathrm{A})$ is closed. Now

$$d^2_{\mathbf{g}}(\gamma_j(t), \gamma_j(t')) \leq \left(\int_t^{t'} \sqrt{\mathbf{g}_{\gamma_j(\tau)}(\dot\gamma_j, \dot\gamma_j)}\ d\tau \right)^2$$

$$\leq |t-t'| \int_t^{t'} \mathbf{g}_{\gamma_j(\tau)}(\dot\gamma_j, \dot\gamma_j)\, d\tau$$

and, because of (4-10) and (4-11), the relations

$$\begin{aligned} d^2_{\mathbf{g}}\left(\gamma_j(t), \gamma_j(t')\right) &\leq |t-t'| \int_t^{t'} G_{\gamma_j(\tau)}(\dot\gamma_j, \dot\gamma_j)\ d\tau \\ &\leq 2\,|t-t'|\, E_G(\gamma_j; t)\ \leq 2\,|t-t'|\,\mathrm{A} \end{aligned}$$

imply the equicontinuity of $\mathcal{H}_{x_0,x_1}(I;\mathrm{A})$ in $\mathcal{C}_{x_0,x_1}(I;\mathrm{A})$. □

6.6. Remark. — *The G-length minimizing curve between x_0 and x_1 does not need to be even piecewise C^1, it could a priori be only absolutely continuous.*

7. MAXIMUM PRINCIPLE AND HORIZONTAL GEODESICS

From now on, we will term geodesic a G-energy minimizing curve called either normal or abnormal, or both, according to the different cases pointed out by the Maximum Principle. In this section, we show how the Maximum Principle is providing necessary conditions involving a lift of the velocity vector in $g^{-1}(\dot\gamma)$ of a G-energy minimizing curve $\bar{\gamma}$, even in the singular case, the abnormal geodesics appear as a limit case, when λ_0 goes to zero (7-G-1), (7-G-2).

Let, as above, $(M, \mathcal{E}, g) = (M, \tilde{\mathbf{U}}, \tilde{H}, \tilde{h})$ (see (4-6)) be a singular sub-Riemannian manifold,with H of class C^k, $(1 \leq k)$, K a compact subset of M, x_0 and x_1 be any two points in K, $I = [0, 1]$.

We know that the length of a curve does not depend on the bijective absolutely continuous changes of parameter, and that the minimum of length is particularly achieved on a parametrized curve which realizes as well the minimum with respect to the energy with a velocity vector of constant G-norm, c almost everywhere. As, in this section, the interval of definition of the curves is chosen to be $[0, 1]$, the velocity vector constant G-norm is

$$|\dot{\bar{\gamma}}(t)|_G = l_G(\bar{\gamma}) = c\,, \quad \text{a.e. .}$$

Under these conditions, any curve in $\mathcal{H}_{x_0x_1}(I; K; A)$, whatever its parametrization, has its energy larger than

$$E_G(\bar{\gamma}) = \frac{1}{2}\int_0^1 l_G(\bar{\gamma})^2\,dt = \frac{1}{2}\,l_G(\bar{\gamma})^2\ .$$

Let $\tilde{x}$ be any curve in $\mathcal{H}_{x_0x_1}(I; K; A)$ such that its image lies in K and its energy $E(\tilde{x}; I; t) \geq \frac{1}{2}\,l_G(\tilde{x})^2$ is finite. Then, if we choose any positive constant A larger than $E(\tilde{x}; I; t)$, we know that there exists at least one horizontal curve $\bar{\gamma} \in \mathcal{H}_{x_0x_1}(I; K; A)$ with $\frac{1}{2}\,l_G(\bar{\gamma})^2 < A$, such that the minimum of G-energy is achieved on this curve between x_0 and x_1 (see section 6). We will look via the Maximum Principle for the necessary conditions verified by such a curve $\bar{\gamma}$. In order to do so, we want first to specify the domain of controls in $\mathbb{R}^p$.

Evidently, if a trajectory is optimal between x_0 and x_1, it will be optimal between x_0 and $\bar{\gamma}(t)$, for any t, $0 \leq t \leq 1$. For if it were not, there would exist a new a.c. horizontal minimizing curve $\gamma_1 \in \mathcal{H}_{x_0\bar{\gamma}(t)}([0, t]; K; A)$ between x_0 and $\bar{\gamma}(t)$, strictly shorter than the previous one $\bar{\gamma}/[0, t]$, with $|\dot{\gamma}_1(\tau)|_G < l_G(\bar{\gamma})$ and the curve

$$\gamma \in \mathcal{H}_{x_0x_1}([0, 1]; K; A)\ ,$$

such that $\gamma/[0,t] = \gamma_1/[0,t]$, $\gamma/[t,1] = \bar{\gamma}/[t,1]$ would be strictly G-shorter than $\bar{\gamma}$, with velocity norm

$$t\ |\dot{\gamma}_1|_G \ + \ (1-t)\ l_G(\bar{\gamma}) < l_G(\bar{\gamma})\ .$$

So the problem of seeking necessary conditions for a curve being an energy minimizer has become a local problem around a regular value t (definition (3-2)). In order to do so, we will choose a domain of controls $\mathbf{W} \in \mathbf{U}$ (see (4-6)), the bounded and close tubular neighbourhood of the null section in $\mathbf{C}$, so that

$$\mathbf{W}_x \ = \left\{ \ s(x) \in \mathbf{U}_x \ / \ l_G^2(\bar{\gamma}) \leq h(s,s)_x \leq 2A \ \right\} .$$

The cost density function that we consider is then the energy density. A minimum with respect to the energy is exactly a minimum with respect to the length, among the curves γ parametrized a.e. by $t = l_G(\gamma)\ \sigma$, where σ is the G-arc length.

Let us create a finite subdivision as in section 6.

(7-1) $$0 = t_0 < t_1 < \ldots < t_l < \ldots < t_m = 1 \quad , \quad I_l = [t_{l-1}, t_l]$$

such that, for any l, $\bar{x}([t_{l-1}, t_l]) \subset \bar{W}_l \subset V_l$, where V_l is a TM and $\mathbf{U}$ (see (4-6)) trivializing coordinate open set, and $\bar{W}_l$ is a compact cell. Let $\{e_i \ / \ 1 \leq i \leq p\}$ be an h-orthonormal frame on $\mathbb{R}^p \simeq \mathbf{U}\,/\,V_l$. Further, let (x^α) be local coordinates on V_l. The greek indices will be running from now on, between 1 and n, the latin ones will be running between 1 and p. We will use the Einstein sommation convention on the greek indices only. Above this trivialized open set V_l, H becomes identified with the $(n \times p)$-matrix $H = (H_i^\alpha)$, and, if γ is a horizontal path, in $\mathcal{H}_{x_0 x_1}(I;K;A)$, there exists a unique control $s(t)$ (Theorem (5-3)) such that

$$\big((H \circ \gamma) \cdot s\big)\ (t) = \dot{\gamma}(t), \qquad s(t) \subset \big(\mathrm{Ker}\, H_{\gamma(t)}\big)^{\perp_h} \cap \mathbf{U}\ .$$

And then we get

$$E_G(\gamma; I_l; t) = \frac{1}{2} \int_{t_{l-1}}^{t_l} \sum_{i,j=1}^{p} \big(h_{\gamma(t)}\big)_{ij} s^i(t) s^j(t)\, dt\ .$$

Now, we are in a position to write for any t the "Hamiltonian function" of the Maximum Principle (see definition (3-8))

$$\mathcal{H} : (\mathbb{R} \times V_l) \times \mathbf{W} \times \mathbb{R}^{n+1} \longrightarrow \mathbb{R} \ ,$$

such that, *with an opposite* λ_0 *with respect to the paragraph 3,*

$$\mathcal{H}(y,u,\lambda) \ = < \lambda, f(y,u) > = \sum_{\alpha=1}^{n} \lambda_\alpha f^\alpha(x,u) \ - \ \lambda_0 f^0(x,u)$$

with

(7-2) $$f^\alpha = \sum_{i=1}^{p} (H_x)_i^\alpha \, u^i, \quad \text{and} \quad f^0 = \frac{1}{2} \sum_{i,j=1}^{p} (h_x)_{ij} \, u^i u^j .$$

Let, as above, the curve $\bar\gamma$ be G-length minimizing. Then, it satisfies the Maximum Principle on V_l. Thus, there exists one 1-form $\bar\lambda : I_l \longrightarrow T^*M$ such that, $x(t) = \bar\gamma(t)$ and $\bar\lambda(t)$ are solutions a.e. of the Hamiltonian system (see section 3) :

($\mathcal{H}$) $$\begin{cases} (\mathcal{H}-1) & \dot{x}^\alpha \ = \ \dfrac{\partial \mathcal{H}}{\partial \lambda_\alpha} \\ (\mathcal{H}-2) & \dot{\lambda}_\alpha \ = \ -\dfrac{\partial \mathcal{H}}{\partial x^\alpha} \ , \end{cases}$$

and verify

$$\mathcal{H}\big(\ \bar\gamma(t), \bar s(t), \bar\lambda(t)\big) \ = \ \sup_{u \in \mathbf{U}} \mathcal{H}\big(\ \bar\gamma(t), \ u \, , \bar\lambda(t)\big) \ ,$$

with

$$\mathcal{H}\big(\ \bar\gamma(t), \ u \, , \bar\lambda(t)\big) \ = \ \sum_{i=1}^{p} \bar\lambda_\alpha(t) \big(H_{\bar\gamma(t)}\big)_i^\alpha \, u^i \ - \ \frac{1}{2} \bar\lambda_0 \sum_{i,j=1}^{p} \big(h_{\bar\gamma(t)}\big)_{ij} \, u^i u^j \ .$$

Then, here

$$\frac{\partial \mathcal{H}}{\partial u^i}\big(\ \bar\gamma(t) \, , \ u \, , \ \bar\lambda(t)\big) = 0, \qquad 1 \le i \le p \ .$$

Thus, because of (7-2),

(7-M-P) $$\bar\lambda_\alpha \big(H_{\bar\gamma(t)}\big)_i^\alpha = \bar\lambda_0 \big(h_{\bar\gamma(t)}\big)_{ij} \bar s^j \ .$$

We have then to distinguish two cases :

$$\bar{\lambda}_0 \neq 0, \qquad \text{and } \bar{\lambda}_0 = 0 \ .$$

I - Case $\bar{\lambda}_0 \neq 0$.

We get, if $\bar{\lambda}_0 \neq 0$ on V_l,

$$\mathcal{H} = \frac{1}{2\bar{\lambda}_0} \sum_{i,j=1}^{p} h^{ij} H_i^\alpha H_j^\beta \bar{\lambda}_\alpha \bar{\lambda}_\beta = \frac{1}{2\bar{\lambda}_0} g^{\alpha\beta} \bar{\lambda}_\alpha \bar{\lambda}_\beta \ .$$

Hamilton equations $(\mathcal{H}-1,2)$ in section 3 become almost everywhere (denoting $\bar{\gamma}(t) = x(t)$, and $\bar{\lambda}(t) = \lambda(t)$)

$$(7\text{-}\mathcal{H}) \qquad \begin{cases} (7-\mathcal{H}-1) & \dot{x}^\alpha = \dfrac{\partial \mathcal{H}}{\partial \lambda_\alpha} = g^{\alpha\beta}\lambda_\beta/\lambda_0 \\ (7-\mathcal{H}-2) & \dot{\lambda}_\alpha = -\dfrac{\partial \mathcal{H}}{\partial x^\alpha} = -\dfrac{\partial g^{\rho\sigma}}{\partial x^\alpha}\lambda_\rho\lambda_\sigma/\lambda_0 \ . \end{cases}$$

For $\lambda_0 = 1$, the previous system of differential equations is well known as the Hamiltonian system associated to the function $\tilde{g} : T^*M \,/\, V_l \longrightarrow \mathbb{R}$ such that $\tilde{g}(\xi) =< \xi, g\xi >= G(\dot{x}, \dot{x})$. Then, above V_1, there exists a solution of class C^k of the system (7-$\mathcal{H}$) with the same initial conditions as $(\bar{\gamma}, \bar{\lambda})$, and as the derivatives of this smooth solution are a.e. equal to the derivatives of $(\bar{\gamma}, \bar{\lambda})$, they are the same everywhere ; thus, $(\bar{\gamma}, \bar{\lambda})$ is of class C^k, on V_1. The same argument works on V_2 with initial conditions $(\bar{\gamma}(t_1), \bar{\lambda}(t_1))$. Furthermore, as $V_1 \cap V_2 \neq 0$, by a connexity argument, $(\bar{\gamma}, \bar{\lambda})$ is also an integral curve of the Hamiltonian system associated to $\tilde{g}$, of class C^k, on V_2 and so on, step by step, until $t_m = 1$ in V_m. Then, $(\bar{\gamma}, \bar{\lambda}) \,/\, V_l$ satisfies also the Maximum Principle with the same $\bar{\lambda}_0 \neq 0$ on any V_l. Thus, $\bar{\lambda}$ being determined up to a multiplicative constant, we could suppose $\bar{\lambda}_0 = 1$ all along $\bar{\gamma}$, and $\bar{\lambda}$ is of class C^k.

Now, if $\xi \in g^{-1}(\dot{\bar{\gamma}})$ is any lift in T^*M of the velocity vector field along $\bar{\gamma}$, (such a lift exists after the results of section 5), there exists a field of 1-forms ν along $\bar{\gamma}$, such that $\nu(t)$ belongs to Ker $g_{\bar{\gamma}(t)}$ = Ker $H^*_{\bar{\gamma}(t)}$ and satisfies the following relation

$$\bar{\lambda}_\alpha = \bar{\lambda}_0 \xi_\alpha + \nu_\alpha \ .$$

Furthermore, the condition that $\mathcal{H}$ is a constant along $\bar{\gamma}$ implies that

$$g^{\alpha\beta}\xi_\alpha\xi_\beta = 2\bar{\lambda}_0 C = G(\dot{\bar{\gamma}}, \dot{\bar{\gamma}}) \ .$$

Anyhow, whatever the field η of absolutely continuous 1-forms along any absolutely continuous curve γ, it is possible to give a definition of the "Lie" derivative of η along the curve γ.

7.3. Definition. — *Let η be a field of absolutely continuous 1-forms along a horizontal absolutely continuous curve γ, such that*

$$\dot\gamma(t) = \big(H \circ \gamma(t)\big) \cdot s(t), \quad s(t) \in (\mathrm{Ker}\, H_{\gamma(t)})^{\perp_h} \ .$$

Then, let us call "Lie derivative" of η along the curve γ, the 1-form above $\gamma(t)$ given by

$$< \mathcal{L}_{\dot\gamma(t)}\eta_{\gamma(t)} \ , \ Z_{\gamma(t)} > = \frac{d}{dt} < \eta, Z >_{\gamma(t)} \ - \sum_{i=1}^{p} < \eta, \ [\dot\gamma(t) \ , \ Z] \ >_{\gamma(t)} \ ,$$

for any absolutely continuous vector field Z along γ •

It is easy to verify that this definition does not depend on the choice of fields $\tilde\eta$ extending η, and $\tilde Z$ extending Z.

The previous necessary conditions can now be written in the following way.

7.4. Proposition. — *Let a G-energy minimizing curve $\bar\gamma : I \longrightarrow M$ and a subdivision of I as the one defined by (7-1). If there exists an integer $l_0 \in \{1, 2, \ldots, m\}$, such that in the chart V_{l_0}, $(\bar\gamma, \bar s, \bar\lambda)$ is an extremal lifted Trajectory of the Maximum Principle, with a constant non-zero $\bar\lambda_0$, then, it is again true in the chart V_l for any integer l, $l \in \{1, 2, \ldots, m\}$, with the same $\bar\lambda_0$. This situation occurs if and only if there exists a lift $\xi \in g^{-1}(\dot{\bar\gamma})$ and a 1-form ν in* Ker $H^*_{\bar\gamma(t)}$, *such that*

(7-G-1)
$$\begin{cases} G(\dot{\bar\gamma}, \dot{\bar\gamma}) & = \ 2\ \bar\lambda_0\ C \ = \ < \xi, \dot{\bar\gamma} > \\ \mathcal{L}_{\dot x}\ (\xi + \dfrac{\nu}{\bar\lambda_0}) & = \ 0 \end{cases} ,$$

with

(7-G-2)
$$\begin{cases} \dot{\bar\gamma} & = \ g\xi \\ \bar\lambda & = \ \bar\lambda_0 \xi + \nu \end{cases} .$$

And $\bar\gamma$ is C^k •

This last equation shows then how it can happen that $\bar{\lambda}$ lies in Ker H^* when $\bar{\lambda}_0$ goes to zero. By the way we get also the following

7.5. Proposition. — *The path $\bar{\gamma}$ is the projection of the Hamiltonian integral $(\bar{\gamma}, \bar{\lambda})$ of the Hamiltonian system of $\tilde{g}$ similarly as the solution of the classical Riemannian variational problem* •

II - Case $\bar{\lambda}_0 = 0$.

The first remark to do is that, as $\bar{\gamma}$ is absolutely continuous, if $\bar{\lambda}_0$ was not zero on some measurable set of positive measure, it would be not zero on the whole interval I, as we have seen in (7-I), so it must be zero, all along $\bar{\gamma}$. We recall that the restriction above V_l of the control "Hamiltonian" of a lifted optimal Trajectory (definition (3-8)) can be written

$$\mathcal{H}_t = \bar{\lambda}_\alpha(t) \sum_{i=1}^{p} \big(H_{\bar{\gamma}(t)}\big)_i^\alpha \; \bar{s}^i(t) - \frac{1}{2}\bar{\lambda}_0 \sum_{i=1}^{p} \mu_i\big(\bar{\gamma}(t)\big)(\bar{s}^i)^2 = \bar{\lambda}_\alpha(t) \sum_{i=1}^{p} \big(H_{\bar{\gamma}(t)}\big)_i^\alpha \, \bar{u}^i(t) \ ,$$

if $\bar{\lambda}_0 = 0$, and $\bar{u}^i(t)$ can be any possible Trajectory control such that $\bar{u}^i(t) - \bar{s}^i(t) \in$ Ker $H_{\bar{\gamma}(t)}$, as the energy is no more involved in the equation, $\bar{u}^i(t)$ satisfies the Maximum Principle as well.

With $\bar{\lambda}_0 = 0$, the control Hamiltonian function

$$\mathcal{H} : (\mathbb{R} \times V_l) \times \mathbf{W} \times \mathbb{R}^{n+1} \longrightarrow \mathbb{R} \ ,$$

becomes linear with respect to v ; it can be constant and maximum with respect to the controls v along the lifted minimizing Trajectory $(\bar{\gamma}, \bar{u}, \bar{\lambda})$ if and only if it is equal to zero for almost every t, thus for any t. This is the same as saying that, above V_l,

$$\bar{\lambda}_\alpha(t) \in \text{Ker } g_{\bar{\gamma}(t)} = \text{Ker } H^*_{\bar{\gamma}(t)} \ .$$

Then,

$$\mathcal{H}(t) \ = \ \sum_{i=1}^{p} \bar{\lambda}_\alpha(t) \left(H_{\bar{\gamma}(t)}\right)_i^\alpha v^i \ = \ 0 \qquad \forall v \in \mathbf{U}_{\bar{\gamma}(t)} \subset \mathbb{R}^p \ ,$$

in other words

$$< \bar{\lambda}(t) \, , \mathcal{E}_{\bar{\gamma}(t)} >= \text{Constant} = 0$$

all along the minimizing curve, and its derivatives of any order with respect to t are zero everywhere, whatever the control v.

So let us set from now on in this section (7-II), $\bar\lambda := \bar\nu$.

Hamilton's equations are still valid a.e.

$$(\mathcal{H}) \qquad \begin{cases} (II-\mathcal{H}-1) & \dot{\bar\gamma}^\alpha = \dfrac{\partial \mathcal{H}}{\partial \lambda_\alpha} = H_i^\alpha \bar u^i \\ (II-\mathcal{H}-2) & \dot{\bar\nu}_\alpha = -\dfrac{\partial \mathcal{H}}{\partial x^\alpha} = -\bar\nu_\beta \dfrac{\partial H_i^\beta}{\partial x^\alpha}\bar u^i \ , \end{cases}$$

and

$$\text{(II-M-P)} \qquad \forall i,\ 1\le i\le p, \quad \forall k,\ k\in \mathbb{N} \quad \begin{cases} \bar\nu_\alpha(t)\left(H_{\bar\gamma(t)}\right)_i^\alpha = 0 \\ \dfrac{d^k}{dt^k}\left(\bar\nu_\alpha(t)\left(H_{\bar\gamma(t)}\right)_i^\alpha\right) = 0 \ . \end{cases}$$

As $\bar\nu$ is solution of the differential equation (II-$\mathcal{H}$-2), it is a.c., and we can use the definition (7-3) to write the "Lie derivative" of $\bar\nu_{\bar\gamma(t)}$ with respect to $\dot{\bar\gamma}(t)$ for almost every t, locally, as

$$\mathcal{L}_{\dot{\bar\gamma}(t)}\bar\nu(t) = \dot{\bar\nu}_\alpha(t) + \bar\nu_\beta \sum_{i=1}^{p} \frac{\partial H_i^\beta}{\partial x^\alpha}\bar u^i = 0 \quad \text{a.e. .}$$

Now, we are able to prove the following

7.6. Lemma. — *Let $\bar\nu(t)$ be the a.c. lift $\bar\lambda(t)$ in T^*M of an optimal lifted Trajectory $(\bar\gamma,\bar s,\bar\lambda)$ of the Maximum Principle with $\bar\lambda_0 = 0$. Let t_0 be a regular value of t. Then, for any vector field Z, along $\bar\gamma$, such that $<\bar\nu, Z>_{\bar\gamma(t)} = 0$,*

$$< \nu\, ,\, [\dot{\bar\gamma}(t), Z] >_{\bar\gamma(t)}$$

is tensorial in Z, above $\bar\gamma(t)$, and is zero everywhere •

Proof. The lemma is obtained by a straightforward calculation, assuming that along the curve, for almost every t, $\mathcal{L}_{\dot{\bar\gamma}}\bar\nu = 0$ and $\dfrac{d}{dt} <\bar\nu, Z> = 0$. Furthermore, if we define $<\bar\nu, [\dot{\bar\gamma}(t), Z]>$ as

$$<\bar\nu, [\dot{\bar\gamma}(t), Z]> = <\mathcal{L}_{\dot{\bar\gamma}}\bar\nu, Z> - \frac{d}{dt} <\bar\nu, Z> \ ,$$

the result is tensorial in $\dot{\bar{\gamma}}(t)$ and in Z, because of the hypothesis that Z_t belongs to $\operatorname{Ker}\bar{\nu}_t$ for any t. Under these conditions $< \bar{\nu}, [\dot{\bar{\gamma}}(t), Z] >$ is well defined, even if Z_t and $\dot{\bar{\gamma}}(t)$ are not continuous.

If we recall that $\mathcal{E}_1(\dot{\bar{\gamma}}) = \mathcal{E}$, and $\mathcal{E}_{j+1}(\dot{\bar{\gamma}}) = \mathcal{E} + [\dot{\bar{\gamma}}, \mathcal{E}_j(\dot{\bar{\gamma}})]$, as in section 4. This leads to the following

7.7. Theorem. — *If a singular or regular sub-Riemannian manifold admits an optimal lifted Trajectory* $(\bar{\gamma}, \bar{s}, \bar{\lambda})$ *satisfying the Maximum Principle with* $\bar{\lambda}_0 = 0$, *on some set of positive measure, then,* $\bar{\lambda}_0$ *is zero all along* $\bar{\gamma}$, *the 1-form* $\bar{\lambda} := \bar{\nu}\,(\bar{\nu} \neq 0)$ *is absolutely continuous and such that, for all regular* t,

$$\forall j \in \mathbb{N} \quad < \nu \ , \ \mathcal{E}_j(\dot{\bar{\gamma}}) >_{\bar{\gamma}(t)} = 0 \ .$$

Then, necessarily, along this curve, $\bar{\gamma} : [0,1] \longrightarrow M$,

$$\bigcup_{j=1}^{\infty} \mathcal{E}_j(\dot{\bar{\gamma}})_{\bar{\gamma}(t)} \subsetneq T_{\bar{\gamma}(t)}M$$

and

$$\forall j \in \mathbb{N} \ , \qquad < \mathcal{L}_{\dot{\bar{x}}}\tilde{\nu}, \mathcal{E}_j(\dot{\bar{x}}) >= 0, \quad a.e. \quad \bullet$$

That is the reason why it never happens that $\bar{\lambda}_0 = 0$ when the strong Hörmander generating condition is verified, i.e., for every horizontal vector X_x, $\big(\mathcal{E}_2(X)\big)_x = T_xM$.

R. Hermann found already this condition in a different context (see [He]).

Proof. Because of the Maximum Principle, $< \bar{\nu}\ ,\ \mathcal{E} >_{\bar{\gamma}(t)} = < \bar{\nu}\ , \mathcal{E}_1(\dot{\bar{\gamma}}) >_{\bar{\gamma}(t)} = 0$ all along the curve. Then, we have to apply lemma (7-6) to any $Z \in \bar{\gamma}^*\mathcal{E}$ and recall that $\mathcal{E}_2(\dot{\bar{\gamma}}) = \mathcal{E} + [\dot{\bar{\gamma}}, \mathcal{E}]$ in $\bar{\gamma}^*(TM)_t$ is tensorial in $\dot{\bar{\gamma}}(t)$, and depends only on its value at the point $\bar{\gamma}(t)$. Thus, $< \bar{\nu}\ ,\ \mathcal{E}_2(\dot{\bar{\gamma}}) >_{\bar{\gamma}(t)} = 0$, for every t. Then, we have to apply again lemma (7-6) to any vector Z in $\bar{\gamma}^*\mathcal{E}_2(\dot{\bar{\gamma}})$, at t. Thus, $< \bar{\nu}\ ,\ \mathcal{E}_3(\dot{\bar{\gamma}}) >_{\bar{\gamma}(t)} = 0$, for every t, and so on, step by step.

III - Statement of the results.

Let us give a more precise definition.

7.8. Definition. — *Let* $\tilde{x}$ *be a curve in* $\mathcal{H}_{x_0x_1}(I;K;A)$. *Let* $\tilde{\xi} : I \longrightarrow T^*M$ *be any "lift" 1-form such that* $\tilde{\xi}(t) \in g^{-1}\big(\dot{\tilde{x}}(t)\big)$, $\tilde{x}$ *is called a*

1°) *normal extremal satisfying the Maximum Principle if there exists a 1-form* $\tilde{\nu} : t \longrightarrow \mathrm{Ker}\ H^*_{\tilde{x}(t)}$ *such that*

$$\text{(N)} \qquad \begin{cases} \mathcal{L}_{\dot{\tilde{x}}}(\tilde{\xi}+\tilde{\nu}) & = & 0 \\ d < \tilde{\xi}, \dot{\tilde{x}} > & = & 0\ ; \end{cases}$$

2°) *strictly abnormal extremal satisfying the Maximum Principle if*

$$\text{(SAN)} \qquad \begin{cases} \forall \tilde{\mu} : I \longrightarrow \mathrm{Ker}\ H^* & \mathcal{L}_{\dot{\tilde{x}}}(\tilde{\xi}+\tilde{\mu}) & \neq & 0 \\ \forall k \in \mathbb{N} & p_k(\dot{\tilde{x}}) & < & n\ ; \end{cases}$$

3°) *non-strictly abnormal extremal satisfying the Maximum Principle if*

$$\text{(NAN)} \qquad \begin{cases} \qquad (N) & is\ verified \\ \forall k \in \mathbb{N} & p_k(\dot{\tilde{x}}) < n \end{cases} \quad \bullet$$

Finally, we get the following

7.9. Theorem. — *Let* $(M, \mathcal{E}, g)$ *be a regular or singular sub-Riemannian manifold, let* $\mathcal{H}_{x_0x_1}(I;K;A)$ *be defined as above, with* H *of class* C^k. *Then, a* G*-length minimizing curve* $\tilde{x} \in \mathcal{H}_{x_0x_1}(I;K;A)$ *is an extremal satisfying the Maximum Principle of one of the three kinds. Moreover, if it is an* (N)*- or* (NAN)*-extremal, then it is of class* C^k *and* G*-length minimizing. In this case, the extremal is the projection of the sub-Riemannian Hamiltonian trajectory of* $\tilde{g}$. *Furthermore, if it is an* (SAN)*-extremal, then there exists an a.c. curve* $\tilde{\nu} : t \longrightarrow \mathrm{Ker}\ g^*_{\tilde{x}(t)}$, *such that*

$$\forall j \in \mathbb{N}\ , \quad < \mathcal{L}_{\dot{\tilde{x}}}\tilde{\nu}, \mathcal{E}_j(\dot{\tilde{x}}) >= 0\ , \quad a.e. \quad \bullet$$

To our knowledge there does not exist, up to now, a simple criteria telling in which cases a strictly abnormal (SAN)-extremal is locally G-length minimizing.

R. Montgomery [Mo] has exhibited a regular three-dimensional sub-Riemannian manifold $(p = 2) M$ verifying the strong two steps Hörmander generating condition $p_2(x) = 3$ everywhere but on a cylinder C, where $p_2/C = 2 = p < 3$, and $p_3/C = 3 = n$. The manifold M is provided with a helicoidal vector field, along C, verifying (SAN). It has not been so easy to prove that the integral curves of this vector field are even locally length minimizing. I. Kupka ([K]) has shown that, up to a certain distance of the initial point, there exists no cut-point. We chose another method and we develop this last proof in section 10. See also [L-S] for a local proof.

7.10. Example (4-1-iii) continued. — Now, we shall take up again Example (4-1) to illustrate the previous method in the singular case. We shall use the metric described in section 4, with $\mathbf{U} = TM = \mathbb{R}^4$, and h the canonical metric on each $\mathbf{U}_{(x,y)} = \mathbb{R}^2$. Then the matrices of g and H, in the frames $\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\}$ and $\{dx, dy\}$, are such that

$$g = H \circ h^\sharp \circ H^* = \begin{pmatrix} 1 & 0 \\ 0 & x^2 \end{pmatrix} .$$

The control fiber bundle is the trivial bundle $M \times \mathbb{R}^2$. Let its canonical basis $\{e_1, e_2\}$ be an h-orthonormal moving frame, such that $H(e_1) = \frac{\partial}{\partial x}$ and $H(e_2) = x \frac{\partial}{\partial y}$, so that, applying Control Theory, we get

$$\mathcal{H} = \lambda_1\ s_1 + \lambda_2\ x\ s_2 - \frac{1}{2}\ (s_1^2 + s_2^2)\ .$$

Here, we take $\lambda_0 = 1$, because the strong two steps generating Hörmander condition is satisfied, and no abnormal geodesic can appear. The maximum principle implies

$$\begin{cases} \dfrac{\partial \mathcal{H}}{\partial s_1} = \lambda_1 - s_1 = 0 \\ \dfrac{\partial \mathcal{H}}{\partial s_2} = \lambda_2\ x - s_2 = 0\ . \end{cases}$$

Then, $\mathcal{H} = \frac{1}{2}\ (\lambda_1^2 + \lambda_2^2\ x^2)$, and both Hamilton's equations $(\mathcal{H} - 1, 2)$ imply

$$\begin{cases} \dot{x} = \lambda_1 \quad , \quad \dot{y} = \lambda_2\ x^2, \\ \dot{\lambda}_1 = -\lambda_2^2\ x \quad , \quad \dot{\lambda}_2 = 0\ . \end{cases}$$

Writing the first integral $\mathcal{H} = B =$ constant, we get

$$\dot{x}^2 + a\dot{y} = 2\ B\ .$$

A very easy computation gives the "horizontal" geodesics joining A $= (0,0)$ to B $= (0,1)$. Namely, with $k \in \mathbb{Z}$,

$$\begin{cases} x &= \sqrt{\dfrac{2}{|k\pi|}}\ sin\ k\pi\ t \\ y &= t - \dfrac{sin\ 2k\pi\ t}{2k\pi} \end{cases} .$$

Their G-length is equal to

$$l_G(\gamma) = \int_0^1 \sqrt{\dot{x}^2 + \frac{\dot{y}^2}{x^2}}\ dt\ =\ 2|k|\pi\ ,$$

the shortest are both obtained for $k = \pm 1$. The result is that the G-distance between $(0,0)$ and $(0,1)$ is 2π. Furthermore, it is obvious that any point on the y-axis can be joined to A by two minimizing curves. Then, all points $(0,y)$ are cut points for the origin, whatever y.

8. NORMAL GEODESICS AND G-DERIVATION

In the framework of regular or singular sub-Riemannian geometry it is possible to define an intrinsic derivative generalizing the Levi-Civita connection of the Riemannian geometry. It will take the shape of an intrinsic bilinear form on T^*M, with values in TM, the restriction of which to the diagonal of $T^*M \times T^*M$ is merely the projection of the ∇_{sym} of C. Bär [B]. This connection will allow us to introduce the idea of G-parallel translation along a horizontal curve, without enlarging G to the whole of TM, for this extension is impossible in the singular case, as we have seen (4-8).

Namely, if $\chi(M)$ (resp. $\wedge M$) are spaces of local sections of TM (of T^*M resp.), and $g = H \circ h^\sharp \circ H^*$, as above.

8.1. Definition. — *Let us call G-derivative and denote by*

$$D : (\wedge M)^2 \longrightarrow \chi(M) \ ,$$

the map such that, for all $\alpha, \beta, \gamma \ \in \ \wedge M$,

$$\begin{aligned} < \gamma, D_\alpha \beta >= \ & \tfrac{1}{2} \left\{ g\alpha < \beta, g\gamma > \ + \ g\beta < \gamma, g\alpha > \ - \ g\gamma < \alpha, g\beta > \right. \\ & \left. - < \alpha, [g\beta, g\gamma] > \ + \ < \beta, [g\gamma, g\alpha] > \ + \ < \gamma, [g\alpha, g\beta] > \right\} \bullet \end{aligned}$$

The following proposition (8-2) (i) to (iv) implies that D is actually an actual global derivation, (v) implies that the D-connection is a generalization of a symmetric Levi-Civita connection. The remaining results and definitions of (8-2), (8-3), and (8-4) constitute a practical formulary about D.

8.2. Proposition. — *Let f be any function of class C^1 on M, then, for all $\alpha, \ \beta, \ \gamma \in \wedge M$ and for all $\mu, \ \nu \in$ Ker H^**

$$\begin{array}{rl} (i) & D_\alpha \ \beta \textit{ is } \mathbb{R}\textit{-linear with respect to } \alpha \textit{ and } \beta \\ (ii) & D_{(f\alpha)} \ \beta \ = \ f(D_\alpha \ \beta) \\ (iii) & D_\alpha \ (f\beta) = \ f(D_\alpha \ \beta) \ + \ ((g\alpha)f).g\beta \\ (iv) & (g\alpha) < \beta, g\gamma > = < \beta, D_\alpha \ \gamma > \ + \ < \gamma, D_\alpha \ \beta > \\ (v) & D_\alpha \ \beta \ - \ D_\beta \ \alpha \ = \ [\ g\alpha, g\beta \ ] \\ & \textit{particularly } D_\alpha \nu \ = \ D_\nu \alpha \ = \ \frac{1}{2}(D_\alpha \nu + D_\nu \alpha) \\ (vi) & < \nu, D_\alpha \ \beta \ + \ D_\beta \ \alpha > = \ 0 \\ & \textit{particularly } \ D_\alpha \ \alpha \in \ \mathcal{E} \textit{ and } D_\alpha \nu \ = \ D_\nu \alpha \in \ \mathcal{E} \\ (vii) & D_\mu \ \nu \ = \ 0 \\ (viii) & < \nu, D_\alpha \ \beta > = < \alpha, D_\nu \ \beta > = \ - < \beta, D_\nu \ \alpha > \\ & = \ \frac{1}{2} < \nu, [g\alpha, g\beta] > \ \bullet \end{array}$$

This G-derivation allows a very nice intrinsic formalism to translate the R.S. Strichartz map Γ ([S] p. 227), namely, let us define the map

$$F : \text{Ker } H^* \times \mathcal{E} \longrightarrow \mathcal{E} \quad \text{and} \quad < \eta, F(\nu, X) >= 2 \ \Gamma(\xi, \nu)\eta \ .$$

8.3. Definition. — *For any ξ and η in $\wedge M$, and any ν in* Ker H^*, *let us consider $X = g\xi,\ Y = g\eta$ in $\chi(M)$. Then,*

$$< \eta, F(\nu, X) > = < \nu, [X, Y] > \bullet$$

These scalar quantities give an idea of how the first order brackets leave $\mathcal{E}$.

8.4. Proposition. — *Using the preceding notations*

$$\begin{array}{rcccl} < \eta, F(\nu\ , X) > & = & < \nu, (D_\xi \eta - D_\eta \xi\ ) > & = & 2 < \nu, D_\xi \eta > \\ & = & 2 < \xi, D_\eta \nu > & = & 2 < \xi, D_\nu \eta > \\ & = & -2 < \eta, D_\xi \nu > & = & -2 < \eta, D_\nu \xi > \\ & = & < \nu, [X, Y] > & = & -d\nu(X, Y)\ . \end{array}$$

Furthermore, $\quad D_\xi \nu \ = \ D_\nu \xi \ = \ -\frac{1}{2} F(\nu\ ,\ X) \bullet$

Then, we get the following characterization for the normal geodesic flow.

8.5. Theorem. — *A vector field X of $\chi(M)$ is the vector field of a "normal geodesic" flow if and only if there exists at least one 1-form $\xi \in g^{-1}(X)$, and one 1-form $\nu \in$ Ker H^* such that*

$$\text{(8-}\mathbf{G}\text{-1)} \qquad \left\{ \begin{array}{rcccc} D_\xi \xi & = & F(\nu, X) & = & -2 D_\nu \xi \\ \mathcal{L}_X(\xi) & = & -\mathcal{L}_X(\nu) & & \bullet \end{array} \right.$$

As $\xi + \nu$ is still a "lift" of X, the Theorem (8-5) is equivalent to the following

8.6. Theorem. — *A vector field X of $\chi(M)$ is the vector field of a "normal geodesic" flow if and only if there exists at least one 1-form $\xi \in g^{-1}(X)$, such that*

$$\text{(8-}\mathbf{G}\text{-2)} \qquad \left\{ \begin{array}{rcl} D_\xi \xi & = & 0 \\ \mathcal{L}_X(\xi) & = & 0 \quad \bullet \end{array} \right.$$

The first equation $D_\xi \xi = 0$, here, is stronger than the previous first integral in (7-**G**-1), $< \xi, X >=$ Constant.

9. THE ABNORMAL GEODESIC OF MONTGOMERY-KUPKA

The purpose in this section is to set a new proof of the length minimizing property of the Montgomery-Kupka abnormal extremal, using the measurable properties of a.c. curves, contrary to the nice simple simultaneous proof due to Liu and Süssmann [L-S]. Actually, we now know that our method leads to a generalization to any "generic" 2-distribution in $\mathbb{R}^3$ with a growth vector (2,2,3) on some hypersurface, whatever the metric [P-V-2]. Let M be the manifold $M = \mathbb{R}^3 \setminus (0,0,\mathbb{R})$ provided with the following regular sub-Riemannian structure. Using systematically cylinder coordinates, let us consider in $\mathcal{E}$ the convenient moving frame denoted by (e)

$$(e) \qquad \begin{cases} e_1 = \dfrac{\partial}{\partial r} & \dots \quad \theta^1 = dr \\ e_2 = \dfrac{1}{r}\Big(\dfrac{\partial}{\partial \theta} + A(r)\dfrac{\partial}{\partial z}\Big) & \dots \quad \theta^2 = r d\theta \\ e_3 = \dfrac{\partial}{\partial z} & \dots \quad \theta^3 = dz - A(r)d\theta \end{cases}$$

where $e_i = g\theta^i$, $i = 1,2,3$, and the $A(r)$ simplified by I. Kupka is given by

$$A(r) = 1 - (1-r)^2 \ , \ A'(r) = 2(1-r) \ , \ A''(r) = -2 \ .$$

The horizontal planes generated by e_1 and e_2, above the points x of M, are denoted by E_x and generate the fiber space E.

The Montgomery positive non-degenerate quadratic form on the horizontal planes G is then well defined by the following matrix

$$g = H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} ,$$

where $\mathbb{R}\theta^3 = \mathbb{R}\ (dz - A(r)d\theta)$ is Ker $H^* =$ Ker g.

All over M except on the cylinder C ($C = \{ r = 1\}$), the plane distribution satisfies the strong generating Hörmander condition, i.e.,

$$\forall X \in E, \quad \forall x \in M \backslash C, \quad (E + [X, E])_x = T_x M$$

for we have

$$[e_1, e_2] = -\frac{1}{r} e_2 + \frac{A'(r)}{r} e_3 \ .$$

Nevertheless, everywhere

$$[e_1, [e_1, e_2]] = \frac{2}{r^2} e_2 - \frac{2}{r} e_3 - 2\frac{A'(r)}{r^2} e_3 \ ,$$

which becomes $2(e_2 - e_3)$ on the cylinder C. Thus, the Hörmander condition is satisfied everywhere.

But, on the cylinder C, $[e_2, [e_2, [e_2, e_1]]] = 0 = [e_2, [e_2, e_1]] = 0 = [e_2, e_1]$, and so on. Thus, $e_2 \,/\, C$ satisfies condition (8-*SAN*) and C is an abnormal extremal of the Maximum Principle. We do not yet know if it satisfies (8-*NAN*), i.e., the condition for a normal automatically length minimizing geodesic.

We shall prove the following

9.1. Theorem. — *Let* $\mathbf{H}_{x_0, x_1} = \Big\{ (1, t, t) \in \mathbb{R}^3 \ / \ x_0 = (1, 0, 0) \ , x_1 = (1, \theta_1, \theta_1) \ ,$ $0 < \theta_1 < 2 \Big\}$ *be the Montgomery-Kupka helix. Then, the length of any horizontal a.c. curve joining* x_0 *to* x_1 *is bigger than the length of* $\mathbf{H}_{x_0,x_1}$*, i.e.,* θ_1 •

Let us write $F : \mathrm{Ker}\ g \times E \longrightarrow E$, such that

$$F(\theta^3, e_i) = \sum_{j \neq i} < \theta^3, [e_i, e_j] > e_j \ ;$$

here $F(\theta^3, e_1) =< \theta^3, [e_1, e_2] > e_2 = \frac{A'(r)}{r} e_2$, and $F(\theta^3, e_2) =< \theta^3, [e_2, e_1] > e_1 = -\frac{A'(r)}{r} e_1$. Now, using the G-derivative just defined in section 8, we see that the flow e_2 is a geodesic flow on $M \backslash C$, because the characterizing conditions (8-**G**-1), equations of normal geodesics, are verified. These conditions are equivalent to

$$\text{(8-}\mathbf{G}\text{-1)} \qquad \exists \xi \in \wedge M \quad \text{such that} \quad \left\{ \begin{array}{rcl} g\xi & = & X \\ D_\xi \xi & = & 0 \\ \mathcal{L}_X(\xi) & = & 0 \,. \end{array} \right.$$

The integral flow of the vector field e_2 is a normal geodesic flow on $M\backslash C$, because the Hamilton-Jacobi equations for the energy minimizing geodesics are satisfied ([S-1]) and the intrinsic characterizing conditions (8-**G**-1) are verified.

A - H does not verify the normal geodesic equation.

As a matter of fact, between any two points on an integral curve of the field e_2 outside C, we have

$$D_{\theta^2}\theta^2 = -\frac{1}{r}e_1 = \frac{1}{A'(r)}F(\theta^3, e_2) = -2\, D_{\theta^2}\frac{\theta^3}{A'(r)} \ ;$$

then,

$$D_{(\theta^2+\nu)}\,(\theta^2+\nu) = 0, \quad \text{with} \quad \nu = \frac{\theta^3}{A'(r)}$$

and

$$\mathcal{L}_{e_2}(\theta^2) = i_{e_2}\frac{\theta^1\wedge\theta^2}{r} = -\frac{\theta^1}{r} \ ,$$

$$\mathcal{L}_{e_2}(\frac{\theta^3}{A'(r)}) = i_{e_2}(\frac{d\theta^3}{A'(r)}) = \frac{\theta^1}{r} \ .$$

The helix $\mathbf{H} = \Big\{(1,t,t)\ /\ t\in \mathbb{R}\Big\}$ is the integral curve of the vector field e_2, restricted to the cylinder C, and

9.2. Remark. — *On* C, $\nu = \dfrac{\theta^3}{A'(r)}$, *is not defined. This is the reason why the helix* **H**, *integral curve of* e_2 *on* C, *cannot verify (8-*__G__*-1), and then cannot be considered as a normal geodesic* •

But $\nu = \dfrac{\theta^3}{A'(r)}$ gives an idea of how the geodesic conditions (**G**) could perhaps be extended to infinity.

Now, let x_0 be the point such that $r=1, \theta=0, z=0$, and x_1 be the point such that $r=1, \theta=\theta_1>0, z=\theta_1$. These two points lie in the helix **H**.

I. Kupka [K] proved that in a tubular neighbourhood of the helix, and for $\theta_1 < \sqrt{2}$, there could not exist any normal geodesic joining x_0 to x_1, for the normal geodesic local equation is not integrable taking account of the end points condition.

So, knowing that the distance is achieved on some curve among the extremal curves deduced from the Maximum Principle, we can conclude that, necessarily, the abnormal arc of helix $\mathbf{H}_{x_0,x_1}$ is G-length minimizing and has no intersection with the "cut-locus" of x_0.

Here, what we are showing is that this abnormal helix is globally minimizing among a.c. curves joining x_0 and x_1, as soon as

$$d_G(x_0, x_1) = \theta_1 < 2 ,$$

the proof is constructive and very simple.

B - The set of horizontal curves joining any two points of H, x_0 **and** x_1.

Let x_0 be the point such that

$$\begin{cases} r &= 1 \\ \theta &= 0 \\ z &= 0 . \end{cases}$$

Let x_1 be the point such that

$$\begin{cases} r &= 1 \\ \theta &= \theta_1, \quad 0 < \theta_1 < 2\pi \\ z &= \theta_1 . \end{cases}$$

These two points are joined by the helix **H**

$$(\mathbf{H}) \qquad \begin{cases} r &= 1 \\ z &= \theta, \quad 0 < \theta < \theta_1 . \end{cases}$$

9.3. Remark. — *If $r(t) \geq 2$, for some t in $]0, t_1[$, any horizontal curve joining x_0 and x_1 through $(r(t), \theta(t), z(t))$ has length larger than 2 whatever $z(t)$.*

Define

$$\forall t \in [0, t_1] , \qquad \rho(t) = 1 - r(t) \qquad 0 < r(t) < 2 .$$

The set of all such horizontal absolutely continuous curves joining x_0 to x_1 can be described by the following conditions, with $|\rho| < 1$

$$x : [0, t_1] \longrightarrow \big(\rho(t), \theta(t), z(t)\big)$$

and

$$z(t) = \int_0^t A\big(r(t)\big) \frac{d\theta}{dt} dt ,$$

and the end point conditions give

$$\begin{array}{lll} \text{(9-I)} & \theta(0)=0 & \theta(t_1)=\theta_1+2k\pi \quad (k\in\mathbb{Z})\ , \\ \text{(9-II)} & z(0)=0 & z(t_1)=\theta_1\ , \\ \text{(9-III)} & \rho(0)=0 & \rho(t_1)=0\ . \end{array}$$

9.4. Remark. — *If a curve* $\gamma : [0,t_1] \longrightarrow \big(\rho(t),\theta(t),z(t)\big)$ *is absolutely continuous on* $[0,1]$*, then, the subsets of* $[0,t_1]$ *where* $\frac{d\theta}{dt}$*, (or* $\frac{d\rho}{dt}$*) is either zero or negative, or positive, are measurable. Furthermore,* $|\dot\theta|dt$ *is a measure density on* $[0,t_1]$ •

Let Σ_0 be the subset of $[0,t_1]$ where $\frac{d\theta}{dt}=0$. Let Σ_- be the subset of $[0,t_1]$ where $\frac{d\theta}{dt}<0$, and, $\sigma_-=\int_{\Sigma_-}|\dot\theta|dt$. Let Σ_+ be the subset of $[0,t_1]$ where $\frac{d\theta}{dt}>0$ and $\sigma_+=\int_{\Sigma_+}\dot\theta dt$. The first endpoint condition becomes

$$\text{(9-I)} \qquad \sigma_+ \;=\; \sigma_- \;+\; \theta_1 \;+\; 2k\pi\ .$$

Let us denote again by (9-II) the second end point condition

$$\text{(9-II)} \qquad z(t_1) \;=\; \theta_1 \;=\; \int_0^{t_1} A\big(r(t)\big)\frac{d\theta}{dt}dt\ ,$$

which becomes

$$\text{(9-II)} \qquad z(t_1) \;=\; \theta_1 \;=\theta_1+2k\pi-\int_{\Sigma_-}\rho^2\frac{d\theta}{dt}dt-\int_{\Sigma_+}\rho^2\frac{d\theta}{dt}dt\ .$$

Thus, the horizontal absolutely continuous curves have the same endpoints as **H**, only if

$$\text{(9-II)} \qquad D^2=\int_{\Sigma_-}\rho^2|\dot\theta|dt=\int_{\Sigma_+}\rho^2|\dot\theta|dt-2k\pi\ .$$

Let us denote by

$$\bar\rho_-^2=\frac{1}{\sigma_-}\int_{\Sigma_-}\rho^2|\dot\theta|dt\ ,\ \bar\rho_+^2=\frac{1}{\sigma_+}\int_{\Sigma_+}\rho^2|\dot\theta|dt\ .$$

So, the second end point condition becomes

$$\text{(9-II)} \qquad \bar\rho_-^2\sigma_-=\bar\rho_+^2\sigma_+-2k\pi\ .$$

The third endpoint condition is

$$\int_0^{t_1} \dot{\rho}(t)dt = 0 \ . \tag{9-III}$$

9.5. Proposition. — *Any horizontal curve joining x_0 to x_1 with $0 < \theta_1 < 2\pi$ and k not zero in condition (9-I) has length larger than 2* •

Proof. The a.c. function $\theta(t)$ goes from 0 to $\theta_1 + 2k\pi$, and then goes through the value $\dfrac{\theta_1}{2} + k\pi$.

$$\begin{array}{llccccccc} \text{If} & k > 0, & 0 & < & \theta_1/2 + \pi & \leq & \theta_1/2 + k\pi & < & \theta_1 + 2k\pi \ ; \\ \text{if} & k < 0, & \theta_1 + 2k\pi & \leq & \theta_1 - 2\pi & < & \theta_1/2 - \pi & < & 0 \qquad . \end{array}$$

In both cases there exists a value of t, t_π such that

$$\theta(t_\pi) = \frac{\theta_1}{2} \pm \pi \text{ with } 0 \leq \rho(t_\pi) < 1 \ ;$$

the point $\gamma(t_\pi) = (\rho(t_\pi), \theta(t_\pi))$ is equidistant from $(1, 0)$ and $(1, \theta_1)$ over the origin, then, necessarily, a curve with $k \neq 0$ has length larger than twice the radius of the cylinder, i.e, 2.

Because of (9-3) and (9-4) **from now on we will be interested only in curves γ such that**

$$0 < \theta_1 \leq 2, \quad \forall t \in [0, t_1], \quad -1 < \rho(t) < 1, \quad k = 0 \ .$$

So, the second endpoint condition becomes

$$\bar{\rho}_-^2 \sigma_- = \bar{\rho}_+^2 \sigma_+ \ . \tag{9-II}$$

9.6. Proposition. — *Let $x_0 = (1, 0, 0)$, $x_1 = (1, \theta_1, \theta_1)$ with $\theta_1 < 2$. If there is no return in the $x0y$ plane, i.e., if $\dot{\theta} \geq 0$, a.e., then, the only ways to go from x_0 to x_1 are either* **H** *or* **H** *with radial horizontal "there and back" segments starting from points of a subset of* **H** *of measure zero. The lengths of these last curves, say of "T_1-type" are larger than the length of* **H**, *namely θ_1* •

Proof. The condition $\sigma_- = 0$ implies

$$D^2 = \int_{\Sigma_-} \rho^2(t)\ |\dot\theta(t)|\ dt = 0\ .$$

Then, the condition

$$\text{(9-II)} \qquad D^2 = \int_{\Sigma_-} \rho^2(t)\ |\dot\theta(t)|(t)\ dt = \int_{\Sigma_+} \rho^2(t)\ \dot\theta(t)\ dt$$

implies

$$\int_{\Sigma_+} \rho^2\ \dot\theta(t)\ dt = 0\ ,$$

then $\rho_/\Sigma_+ = 0$ a.e., and

$$\text{(9-III)} \quad \Longrightarrow \quad \int_{\Sigma_0} \dot\rho(t)\ dt\ =\ 0\ .$$

The result follows.

So, if the curve is not of the previous type, there must exist some set of positive measure on which $\dot\theta < 0$, and then, either $\rho/\Sigma_-^2 = 0$ a.e., and (9-II) implies $\rho_+^2 = 0$ a.e., or $\bar\rho_-^2$ is necessarily > 0 along a curve candidate to be shorter than **H**, and then there is some positive measure subset of Σ_- where $\dot\rho$ is necessarily > 0. In the first case, the curves said of "T_2-type" coincide geometrically with **H**, but they cover some positive measure subset of it more than once and have then, length larger than **H** ; in the second case, the necessary loops in the $x0y$ plane of these curves imply that **H** is C^1**-rigid**, for $\dot\theta(t) = 1$ on $\mathcal{H}$, and is necessarily ≤ 0 on some positive measure measurable subset of T_1-type and T_2-type curves.

C - Horizontal curves joining x_0 to x_1, such that $\theta_1 < 2$ and $0 < r$, if they ever exist, have length greater than H.

From now on, let us examine only the cases where

$$\text{(9-7)} \qquad D^2 = \int_{\Sigma_-} \rho^2(t)\ \dot\theta(t)\ dt > 0\ .$$

Condition (9-7) implies that ρ^2 must be positive on some set of positive measure.

9.8. Lemma. — *For the comparison of lengths between* **H** *and the curves* γ_{x_0,x_1}, *with* $r < 2$, *it is sufficient to study curves* γ_{x_0,x_1} *which lie inside the cylinder* C •

Proof. Let $\gamma : [0, t_1] \longrightarrow \mathbb{R}^3$ be a horizontal curve $\gamma(t) = \big(\rho_1(t), \theta(t), z(t)\big)$ joining $x_0 = (0,0,0)$ to $x_1 = (0, \theta_1, \theta_1)$ with $\theta_1 < 2$. For any absolutely continuous function $\rho_1(t)$, if there exists a subset A of $[0, t_1]$ where

$$-1 < \rho_1(t) < 0 \Longleftrightarrow\ 1 < r_1(t) = 1 - \rho_1(t) < 2\ ,$$

the subset A is necessarily measurable, and we can change ρ_1 to ρ_2 on A such that

$$0 < -\rho_1(t) = \rho_2(t) \Longleftrightarrow\ 0 < r_2(t) = 1 - \rho_2(t) < 1\ , \quad t \in A\ ,$$

and the vertical defect integral is unchanged

$$\int_A \rho_1^2(t)\dot\theta dt\ =\ \int_A \rho_2^2(t)\dot\theta dt$$

the endpoint condition (9-II) is yet fulfilled and

$$l(\gamma_1)/A = \int_A \sqrt{(1+|\rho_1|)^2\dot\theta^2 + \dot\rho_1^2}\ dt\ >\ l(\gamma_2)/A = \int_A \sqrt{(1-|\rho_2|)^2)\dot\theta^2 + \dot\rho_2^2}\ dt\ ,$$

the new curve, inside the cylinder C is shorter than the one outside.

From now on, without lack of generality, we can suppose that the curve γ is parametrized by arc length, and $0 \le \rho(s) < 1$. We also assume $\sigma_- > 0$.

9.9. Lemma. — *Let* R_+ *(resp.* R_-*) be the subset of* $[0, l(\gamma)]$ *where* $\dot\rho > 0$ *(resp.* $\dot\rho < 0$*). Then, the end point condition (9-III) implies*

$$\text{(9-IV)} \qquad \int_{R_+} \dot\rho\ ds\ =\ \int_{R_-} |\dot\rho|\ ds\ \ge\ \sup|\rho|\ \bullet$$

Proof. We have

$$\text{(9-III)} \qquad \rho(0) = 0, \qquad \int_0^{l(\gamma)} \dot\rho\ ds\ = 0\ .$$

The function $\rho(s)$ is absolutely continuous. Thus, there exists a value s_M, $s_M \in]0, l(\gamma)[$ such that

$$\sup|\rho| = |\rho(s_M)| = |\int_0^{s_M} \dot\rho \; ds \; | \; .$$

Then, (9-III) becomes

$$\text{(9-III)} \qquad \int_{R_+} \dot\rho \; ds \; = \int_{R_-} |\dot\rho| \; ds \; .$$

If $\sup|\rho| = \sup \rho > 0$,

$$\sup\rho = \int_{[0,s_M]\cap(R_+\cup R_-)} \dot\rho \; ds \; = \int_{[0,s_M]\cap R_+} \dot\rho \; ds \; - \int_{[0,s_M]\cap R_-} |\dot\rho| \; ds \; ,$$

$$\int_{R_+} \dot\rho \; ds \; \geq \int_{R_+\cap[0,s_M]} \dot\rho \; ds \; = \int_0^{s_M} \dot\rho \; ds \; + \int_{[0,s_M]\cap R_-} |\dot\rho| \; ds$$

$$= \; \sup\rho \; + \int_{[0,s_M]\cap R_-} |\dot\rho| \; ds \; \geq \; \sup\rho \; > 0 \; .$$

If $\sup|\rho| = \sup(-\rho) = -\inf\rho > 0$. This case, of no use here, will be useful in the generalization [V-P],

$$\inf\rho = \int_{[0,s_M]\cap(R_+\cup R_-)} \dot\rho \; ds \; = \int_{[0,s_M]\cap R_+} \dot\rho \; ds \; - \int_{[0,s_M]\cap R_-} |\dot\rho| \; ds \; < 0 \; ,$$

$$\int_{R_-} |\dot\rho| \; ds \; \geq \int_{R_-\cap[0,s_M]} |\dot\rho| \; ds \; = \; -\int_0^{s_M} \dot\rho \; ds \; + \int_{[0,s_M]\cap R_+} \dot\rho \; ds \; > 0$$

$$= \; \sup|\rho| \; + \int_{[0,s_M]\cap R_+} \dot\rho \; ds \; \geq \; \sup|\rho| \; > 0 \; . \qquad \square$$

Let us now write the length of $\gamma_{x_0x_1}$

$$l(\gamma) = 2E(\gamma) = \int_0^{l(\gamma)} \left((1-\rho)^2\dot\theta^2 + \dot\rho^2\right) \; ds$$

$$= \int_{\Sigma_+\cup\Sigma_-} (1-\rho)^2\dot\theta^2 \; ds + \int_{R_+\cup R_-} \dot\rho^2 \; ds \; ,$$

using Schwarz inequality

$$\geq \frac{1}{l(\gamma)} \left(\int_{R_+\cup R_-} |\dot\rho| \; ds \right)^2 \; + \; \frac{1}{l(\gamma)} \left(\int_{\Sigma_+\cup\Sigma_-} (1-\rho)|\dot\theta| \; ds \right)^2 ,$$

using the end point conditions (9-I) and lemma (9-9) (end point conditions (9-IV)),

$$l(\gamma)^2 \geq 4 \sup \rho^2 \ + \ \left(\theta_1 + 2\sigma_- - \int_{\Sigma_+ \cup \Sigma_-} \rho |\dot{\theta}| \ ds \right)^2 .$$

We recall that

$$0 < |\rho| < 1 .$$

Now, using Schwarz inequality again with respect to the measure density $|\dot{\theta}| \ ds$,

$$l(\gamma)^2 \ \geq \ 4 \sup \rho^2 \ + \ \left(\ \theta_1 + 2\sigma_- - \sqrt{\theta_1 + \sigma_-} \ D \ - \ \sqrt{\sigma_-} \ D \ \right)^2 ,$$

for, because of the first and second end point conditions, with $k = 0$,

$$\text{(9-I)} \quad \sigma_+ = \theta_1 + \sigma_- \quad \text{and} \quad \text{(9-II)} \quad \bar{\rho}_-^2 \sigma_- = \bar{\rho}_+^2 \sigma_+ = D^2 .$$

Thus,

$$l(\gamma)^2 \ \geq \ 4 \sup \rho^2 \ + \ \left(\ \theta_1 + 2\sigma_- \ - \ \sqrt{\theta_1 + \sigma_-} \ \sqrt{\sigma_-} \ \ (\bar{\rho}_+ + \bar{\rho}_-) \ \right)^2 ,$$

but

$$\theta_1 + 2\sigma_- \ - \ \sqrt{\theta_1 + \sigma_-} \ \sqrt{\sigma_-} \ \ (\bar{\rho}_+ + \bar{\rho}_-) \ > 0$$

for

$$\theta_1 + 2\sigma_- \ - \ \sqrt{\theta_1 + \sigma_-} \ \sqrt{\sigma_-} \ \ (\bar{\rho}_+ + \bar{\rho}_-) \ \geq \ \theta_1 + 2\sigma_- - 2\sqrt{\theta_1 + \sigma_-} \ \sqrt{\sigma_-} \ \sup |\rho|$$

$$= \sqrt{\theta_1 + \sigma_-}^2 + \sqrt{\sigma_-}^2 - 2\sqrt{\theta_1 + \sigma_-} \ \sqrt{\sigma_-} \ \sup |\rho|$$

$$= \left(\sqrt{\theta_1 + \sigma_-} - \sqrt{\sigma_-}\right)^2 + 2\sqrt{\theta_1 + \sigma_-} \ \sqrt{\sigma_-} \ \left(1 - \sup |\rho|\right) > 0 \ .$$

Then,

$$l(\gamma)^2 \ \geq \ 4 \sup \rho^2 \ + \ \left(\ \theta_1 + 2\sigma_- \ - \ 2\sqrt{\theta_1 + \sigma_-} \ \sqrt{\sigma_-} \ \ \sup |\rho| \ \right)^2 ,$$

and

$$l(\gamma)^2 \ \geq \ \theta_1^2 + \ 4\sigma_- \ (\theta_1 + \sigma_-)$$

$$+ \ 4 \sup \rho^2 \ \left(1 + (\theta_1 + \sigma_-) \ \sigma_-\right) \ - \ 4\sqrt{\theta_1 + \sigma_-} \ \sqrt{\sigma_-} \ (\theta_1 + 2\sigma_-) \ \sup |\rho| \ .$$

Finally, the horizontal curve γ_{x_0,x_1} is longer than **H** as soon as

$$(1 + (\theta_1 + \sigma_-)\sigma_-) \sup\rho^2 \; - \sqrt{\theta_1 + \sigma_-}\,\sqrt{\sigma_-}\,(\theta_1 + 2\sigma_-)\; \sup\rho \; + \; \sigma_-(\theta_1 + \sigma_-) > 0 \; .$$

It is easy to see that the polynomial in $\sup\rho$ has a strictly positive minimum as soon as

$$\theta_1^2 < 4 \; .$$

Then, for $\theta_1 < 2$, if there exist horizontal a.c. curves other than those of T_1-type and T_2-type, joining x_0 to x_1, and **for any** ρ their lengths are greater than the length of **H** as well as the length of those of T_1-type and T_2-type. Furthermore, $\sigma_- \neq 0$ implies that **H is C^1-rigid**, as we have already seen. So we have proved Theorem (9-1).

Furthermore, on the way of the proof of this global result we showed that, as any curve is either of T_1-type or $\sigma_- \neq 0$, the only C^1-curves in a C^1-neighbourhood of **H** are reparametrizations of **H**. This is the actual definition of C^1-rigidity. We saw also the way of constructing horizontal C^1-curves γ_{x_0,x_1} close to **H** with respect to the topology of the uniform convergence, even in the sense of the H^1-topology. Thus, **H** is not H^1-rigid. We now know that in dimensions greater than 3 there are examples of codimension 1 distributions with horizontal **non-minimizing abnormal**, non C^1-rigid C^1-curves [P-V-2], and, in dimension 3, the Montgomery example is a generic local model for the 2-plane distributions with growth vector (2,3) on a dense subset of M, and (2,2,3) on a local hypersurface, whatever the sub-Riemannian metric [P-V-2].

BIBLIOGRAPHY

[A] R. Azencott, *Géodésiques et diffusion en temps petit*, Séminaire de probabilités, Université de Paris VII, Astérisque, S.M.F. **84-85** (1981).

[B] C. Bär, *Carnot-Caratheodory-Metriken*, Diplomarbeit, Bonn (1988).

[B-A] G. Ben Arous, *Noyau de la chaleur hypoelliptique et géométrie sous-riemannienne*, (preprint Univ. Orsay).

[Bel] A. Bellaïche, *Géodésiques et diffusion en temps petit*, Séminaire de probabilités, Université de Paris VII, Astérisque, S.M.F. **84-85** (1981).

[Bi] J.M. Bismut, *Large deviations and the Malliavin calculus*, Progress in Math., Birkhäuser, Basel **45** (1984).

[Br] R.W. Brockett, *Control theory and singular Riemannian geometry*, New Directions in Appl. Math. (P.J. Hilton and G.S. Young, eds.), Springer, Berlin (1981), 11–27.

[C] Carnot-Carathéodory, *Untersuchungen über die Grundlagen der Thermodynamik*, Math. Ann. (1909), 355–386.

[Ch] W.L. Chow, *Über Systeme von Linearen Partiellen Differentialgleichungen erster Ordnung*, Math. Ann. **117** (1939), 98–105.

[G] B. Gaveau, *Systèmes dynamiques associés à certains opérateurs hypoelliptiques*, Bull. Sci. Math. **102** (1978), 203–229.

[Gr] M. Gromov, J. Lafontaine, P. Pansu, *Structures métriques pour les variétés riemanniennes*, Paris, Cedic (1981).

[H] U. Hamenstädt, *Zur Theorie von Carnot-Carathéodory Metriken und ihren Anwendungen*, Bonner Math. Schriften, Univ. Bonn **180** (1987).

[He-1] R. Hermann, *On The Accessibility Problem in Control Theory*, Lassalle and Lefschetz (1962).

[He-2] R. Hermann, *Geodesics of Singular Riemannian metrics*, Bull. Amer. Math. **79** 4(1973), 780–782.

[K] I. Kupka, *Abnormal extremals*, preprint (1992).

[L] R. Léandre, *Développements asymptotiques de la densité de diffusions dégénérées* to appear.,

[L-S] W. Liu, H.J. Süssmann, *Abnormal sub-Riemannian Minimizers*, preprint (1992), to appear in Trans. Amer. Math. Soc. (1992).

[Ma] A. Mayer, *Begründung der Lagrangenschen Multiplicatorenmethode in der Variationsrechnung*, Math. Ann., XXVI **74** (1886).

[Mi] J. Mitchell, *On Carnot-Carathéodory metrics*, J. Differential Geom. **21** (1985), 35–45.

[Mo] R. Montgomery, *Geodesics Which Do Not Satisfy the Geodesic Equations*, preprint (1991), to appear in SIAM J. Control Optimal.

[Os] H. Osborn, *Vector bundles, Vol 1, Foundations and Stiefel-Whitney classes*, Academic Press (1982).

[P] L. Pontriaguine, V. Boltianski, R. Gamkrelidze, E. Michtchenko, *Théorie mathématique des processus optimaux*, éd. MIR, Moscou, trad. française (1974).

[P-V-1] F. Pelletier, L. Valère, *Le problème des géodésiques en géométrie sous-riemannienne singulière*, C. R. Acad. Sci. Paris **317** (1993), 71–76.

[P-V-2] F. Pelletier, L. Valère Bouche, *Abnormality of Trajectory in sub-Riemannian Structure*, in *Geometry in Nonlinear Control and Differential Inclusions*, Banach Center Publications **32** (1995) 301–317.

[S] R.S. Strichartz, *Sub-Riemannian geometry*, J. Differential Geom. **24** (1986), 221–263.

[S] R.S. Strichartz, *Corrections to "Sub-Riemannian Geometry"*, J. Differential Geom. **30** (1989), 595–596.