L. Rifford · E. Trélat

On the stabilization problem for nonholonomic distributions

Received October 9, 2006 and in revised form May 9, 2007

Abstract. Let $M$ be a smooth connected complete manifold of dimension $n$, and $\Delta$ be a smooth nonholonomic distribution of rank $m \leq n$ on $M$. We prove that if there exists a smooth Riemannian metric on $\Delta$ for which no nontrivial singular path is minimizing, then there exists a smooth repulsive stabilizing section of $\Delta$ on $M$. Moreover, in dimension three, the assumption of the absence of singular minimizing horizontal paths can be dropped in the Martinet case. The proofs are based on the study, using specific results of nonsmooth analysis, of an optimal control problem of Bolza type, for which we prove that the corresponding value function is semiconcave and is a viscosity solution of a Hamilton–Jacobi equation, and we establish fine properties of optimal trajectories.

Keywords. Nonholonomic distributions, stabilization, SRS feedback, minimizing singular path, Martinet case, nonsmooth analysis

1. Introduction

Throughout this paper, $M$ denotes a smooth connected manifold of dimension $n$.

1.1. Stabilization of nonholonomic distributions

Let $\Delta$ be a smooth distribution of rank $m \leq n$ on $M$, that is, a rank $m$ subbundle of the tangent bundle $TM$ of $M$. This means that, for every $x \in M$, there exist a neighborhood $V_x$ of $x$ in $M$ and an $m$-tuple $(f_1^x, \ldots, f_m^x)$ of smooth vector fields on $V_x$, linearly independent on $V_x$, such that

$$\Delta(y) = \text{span}\{f_1^x(y), \ldots, f_m^x(y)\}, \quad \forall y \in V_x.$$ 

One says that the $m$-tuple $(f_1^x, \ldots, f_m^x)$ locally represents the distribution $\Delta$. The distribution $\Delta$ is said to be nonholonomic (also called totally nonholonomic e.g. in [3]) if,
for every $x \in M$, there is an $m$-tuple $(f_1^x, \ldots, f_m^x)$ of smooth vector fields on $V_x$ which locally represents the distribution and

$$\text{Lie}[f_1^x, \ldots, f_m^x](y) = T_yM, \quad \forall y \in V_x,$$

that is, the Lie algebra spanned by $f_1^x, \ldots, f_m^x$ is equal to the whole tangent space $T_yM$ at every point $y \in V_x$. This Lie algebra property is often called Hörmander’s condition.

A horizontal path joining $x_0$ to $x_1$ is an absolutely continuous curve $\gamma(\cdot) : [0, 1] \to M$ such that $\gamma(0) = x_0$, $\gamma(1) = x_1$, and $\dot{\gamma}(t) \in \Delta(\gamma(t))$ for almost every $t \in [0, 1]$. According to the classical Chow–Rashevsky Theorem (see [9, 19, 33, 36]), since the distribution is nonholonomic on $M$, any two points of $M$ can be joined by a horizontal path.

Let $\Delta$ be a nonholonomic distribution and $\tilde{x} \in M$ be fixed. We recall that, for a smooth vector field $X$ on $M$, the dynamical system $\dot{x} = X(x)$ is said to be globally asymptotically stable at the point $\tilde{x}$ if the following two properties are satisfied:

- Lyapunov stability: for every neighborhood $V$ of $\tilde{x}$, there exists a neighborhood $W$ of $\tilde{x}$ such that, for every $x \in W$, the solution of $\dot{x}(t) = X(x(t))$, $x(0) = x$, satisfies $x(t) \in V$ for every $t \geq 0$.
- Attractivity: for every $x \in M$, the solution of $\dot{x}(t) = X(x(t))$, $x(0) = x$, tends to $\tilde{x}$ as $t$ tends to $\infty$.

The stabilization problem for nonholonomic distributions consists in finding, if possible, a smooth stabilizing section $X$ of $\Delta$, that is, a smooth vector field $X$ on $M$ satisfying $X(x) \in \Delta(x)$ for every $x \in M$, such that the dynamical system $\dot{x} = X(x)$ is globally asymptotically stable at $\tilde{x}$.

There exist two main obstructions for a distribution to admit a stabilizing section. The first one is of global nature: it is well known that, if the manifold $M$ admits a smooth Lyapunov function, i.e., a Morse function having only one (possibly degenerate) critical point in $M$. Consequently, $M$ must be homeomorphic to the Euclidean space $\mathbb{R}^n$ (we refer the reader to [39] for further details). The second one is of local nature: due to Brockett’s condition (see [13] Theorem 1(iii); see also [23] [44]), the distribution $\Delta$ cannot admit a smooth stabilizing section whenever $m < n$.

The absence of smooth stabilizing sections motivates one to define a new kind of stabilizing section. The first author has recently introduced the notion of smooth repulsive stabilizing feedback for control systems (see Section 1.4), whose definition can be easily translated in terms of stabilizing section.

Let $\tilde{x} \in M$ be fixed. Let $S$ be a closed subset of $M$ and $X$ be a vector field on $M$. The dynamical system $\dot{x} = X(x)$ is said to be smooth repulsive globally asymptotically stable at $\tilde{x}$ with respect to $S$ (denoted $\text{SRS}_{\tilde{x},S}$ for short) if the following properties are satisfied:

1. The vector field $X$ is locally bounded on $M$ and smooth on $M \setminus S$.

1 If one locally represents the distribution $\Delta$ by an $m$-tuple $(f_1, \ldots, f_m)$ of smooth vector fields, then the existence of a local stabilizing section for $\Delta$ is equivalent to the existence of a stabilizing feedback for the associated control system $\dot{x} = \sum_{i=1}^m u_i f_i(x)$. There is a large literature on alternative types of stabilizing feedbacks for control systems (see Section 1.4).
(ii) The dynamical system $\dot{x} = X(x)$ is globally asymptotically stable at $\bar{x}$ in the sense of Carathéodory, that is, for every $x \in M$, there exists a solution of
\begin{equation}
\dot{x}(t) = X(x(t)) \quad \text{for almost every } t \in [0, \infty), \quad x(0) = x,
\end{equation}
and, for every $x \in M$, every solution of $\dot{x} = X(x)$ on $[0, \infty)$ tends to $\bar{x}$ as $t$ tends to $\infty$. Moreover, for every neighborhood $\mathcal{V}$ of $\bar{x}$, there exists a neighborhood $\mathcal{W}$ of $\bar{x}$ such that, for $x \in \mathcal{W}$, the solutions of $\dot{x} = X(x)$ satisfy $x(t) \in \mathcal{V}$ for every $t \geq 0$.

(iii) For every $x \in M$, the solutions of $\dot{x} = X(x)$ satisfy $x(t) \notin S$ for every $t > 0$.

In view of what happens whenever $\Delta = TM$, and having in mind the above obstructions for the stabilization problem, a natural question is whether, given a smooth nonholonomic distribution $\Delta$, there exists a section $X$ of $\Delta$ on $M$ and a closed nonempty subset $S$ of $M$ such that $X$ is SRS$_{x,S}$. In this paper, we provide a positive answer in a large number of situations. To state our main results, we need to endow the distribution $\Delta$ with a Riemannian metric, thus encountering the framework of sub-Riemannian geometry, and we require the concept of a singular path, recalled next.

1.2. Sub-Riemannian geometry

For $x_0 \in M$, let $\Omega_\Delta(x_0)$ denote the set of horizontal paths $\gamma(\cdot) : [0, 1] \to M$ such that $\gamma(0) = x_0$. The set $\Omega_\Delta(x_0)$, endowed with the $W^{1,1}$-topology, has a Banach manifold structure.$^2$ For $x_0, x_1 \in M$, denote by $\Omega_\Delta(x_0, x_1)$ the set of horizontal paths $\gamma(\cdot) : [0, 1] \to M$ such that $\gamma(0) = x_0$ and $\gamma(1) = x_1$. Note that $\Omega_\Delta(x_0, x_1) = E^{-1}_{\bar{x}}(x_1)$, where the end-point mapping $E_{x_0} : \Omega_\Delta(x_0) \to M$ is the smooth mapping defined by $E_{x_0}(\gamma(\cdot)) := \gamma(1)$. A path $\gamma(\cdot)$ is said to be singular if it is horizontal and if it is a critical point of the end-point mapping $E_{x_0}$.

The set $\Omega_\Delta(x_0, x_1)$ is a Banach submanifold of $\Omega_\Delta(x_0)$ of codimension $n$ in a neighborhood of a nonsingular path, but may fail to be a manifold in a neighborhood of a singular path. Singular paths play a crucial role in the calculus of variations with nonholonomic constraints (see [17] for details and for properties of such curves).

Let $T^*M$ denote the cotangent bundle of $M$, $\pi : T^*M \to M$ the canonical projection, and $\omega$ the canonical symplectic form on $T^*M$. Let $\Delta^\perp$ denote the annihilator of $\Delta$ in $T^*M$ minus its zero section. Define $\bar{\sigma}$ as the restriction of $\sigma$ to $\Delta^\perp$. An absolutely continuous curve $\psi(\cdot) : [0, 1] \to \Delta^\perp$ such that $\psi(t) \in \ker \bar{\sigma}(\psi(t))$ for almost every $t \in [0, 1]$ is called an abnormal extremal of $\Delta$. It is well known that a path $\gamma(\cdot) : [0, 1] \to M$ is singular if and only if it is the projection of an abnormal extremal $\psi(\cdot)$ of $\Delta$ (see [29] or [17]). The curve $\psi(\cdot)$ is said to be an abnormal extremal lift of $\gamma(\cdot)$.

Let $g$ be a smooth Riemannian metric defined on the distribution $\Delta$. The triple $(M, \Delta, g)$ is called a sub-Riemannian manifold. The length of a path $\gamma(\cdot) \in \Omega_\Delta(x_0)$ is defined by
\begin{equation}
\text{length}_g(\gamma(\cdot)) := \int_0^1 \sqrt{g_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))} \, dt.
\end{equation}

$^2$ This is a straightforward adaptation of results of Bismut [10] (see also [33]).
The sub-Riemannian distance $d_{SR}(x_0, x_1)$ between two points $x_0$, $x_1$ of $M$ is the infimum over the lengths (for the metric $g$) of the horizontal paths joining $x_0$ and $x_1$. According to the Chow–Rashevsky Theorem (see [9, 19, 33, 36]), since the distribution is nonholonomic on $M$, the sub-Riemannian distance is well defined and continuous on $M \times M$. Moreover, if the manifold $M$ is a complete metric space for the sub-Riemannian distance $d_{SR}$, then, since $M$ is connected, for every pair $(x_0, x_1)$ of points of $M$ there exists a horizontal path $\gamma(\cdot)$ joining $x_0$ to $x_1$ such that

$$d_{SR}(x_0, x_1) = \text{length}_g(\gamma(\cdot)).$$

Such a horizontal path is said to be minimizing.

Define the Hamiltonian $H : T^*M \to \mathbb{R}$ as follows. For every $x \in M$, the restriction of $H$ to the fiber $T^*_x M$ is given by the nonnegative quadratic form

$$p \mapsto \frac{1}{2} \max \left\{ \frac{p(v)^2}{g_s(v,v)} \mid v \in \Delta(x) \setminus \{0\} \right\},$$

(3)

Let $\tilde{H}$ denote the Hamiltonian vector field on $T^*M$ associated to $H$, that is, $\omega_{\tilde{H}} = -dH$.

A normal extremal is an integral curve of $\tilde{H}$ defined on $[0, 1]$, i.e., a curve $\psi(\cdot) : [0, 1] \to T^*M$ such that $\dot{\psi}(t) = \tilde{H}(\psi(t))$ for $t \in [0, 1]$. Note that the projection of a normal extremal is a horizontal path. The exponential mapping $\exp_{x_0}$ is defined on $T^*_x M$ by $\exp_{x_0}(p_0) := \pi(\psi(1))$, where $\psi(\cdot)$ is the normal extremal so that $\psi(0) = (x_0, p_0)$ in local coordinates. Note that $H(\psi(t))$ is constant along a normal extremal $\psi(\cdot)$, and the length of the path $\pi(\psi(\cdot))$ is equal to $(2H(\psi(0)))^{1/2}$.

According to the Pontryagin maximum principle (see [35]), a necessary condition for a horizontal path to be minimizing is to be the projection either of a normal extremal or of an abnormal extremal. In particular, singular paths satisfy this condition. However, a singular path may also be the projection of a normal extremal. A singular path is said to be strictly abnormal if it is not the projection of a normal extremal.

A point $x \in \exp_{x_0}(T^*_x M)$ is called conjugate to $x_0$ if it is a critical value of the mapping $\exp_{x_0}$. The conjugate locus, denoted by $C(x_0)$, is defined as the set of all points conjugate to $x_0$. Note that the Sard Theorem applied to the mapping $\exp_{x_0}$ implies that the conjugate locus $C(x_0)$ has Lebesgue measure zero in $M$.

**Remark 1.1.** It has been established in [33] that the image of the exponential mapping $\exp_{x_0}$ is dense in $M$, and is of full Lebesgue measure for corank one distributions.

**Remark 1.2.** Let $x \in \exp_{x_0}(T^*_x M)$, let $p_0 \in T^*_x M$ be such that $x = \exp_{x_0}(p_0)$, and let $\psi(\cdot)$ denote the normal extremal so that $\psi(0) = (x_0, p_0)$ in local coordinates. If $x$ is not conjugate to $x_0$, then the path $x(\cdot) := \pi(\psi(\cdot))$ admits a unique normal extremal lift. Indeed, if it had two distinct normal extremal lifts $\psi_1(\cdot)$ and $\psi_2(\cdot)$, then the extremal $\psi_1(\cdot) - \psi_2(\cdot)$ would be an abnormal extremal lift of the path $x(\cdot)$. Hence, the path $x(\cdot)$ is singular, and not strictly abnormal, and thus, in particular, the point $x$ is conjugate to $x_0$. This is a contradiction.

Note that, since the distribution $\Delta$ is nonholonomic on $M$, the topology defined by the sub-Riemannian distance $d_{SR}$ coincides with the original topology of $M$ (see [9, 33]).
We also recall the notion of a cut point, required in this article. Let \( x_0 \in M \); a point \( x \in M \) is not a cut point with respect to \( x_0 \) if there exists a minimizing path joining \( x_0 \) to \( x \), which is the strict restriction of a minimizing path starting from \( x_0 \). In other words, a cut point is a point at which a minimizing path ceases to be optimal. The cut locus of \( x_0 \), denoted by \( L(x_0) \), is defined as the set of all cut points with respect to \( x_0 \). The following result comes from [45]. In Section 2.2.3 we provide a new (and self-contained) proof of this result, using techniques of nonsmooth analysis.

**Lemma 1.1.** Let \( M \) be a smooth closed connected manifold of dimension \( n \), and \( \Delta \) be a smooth nonholonomic distribution of rank \( m \leq n \) on \( M \). Let \( g \) be a metric on \( \Delta \) for which no nontrivial singular path is minimizing, and let \( x_0 \in M \). Then

\[
C_{\text{min}}(x_0) \subseteq L(x_0),
\]

where \( C_{\text{min}}(x_0) \) denotes the set of points \( x \in M \setminus \{x_0\} \) such that there exists a critical point \( p_0 \in T^*_x M \) of the mapping \( \exp_{x_0} \), and such that the projection of the normal extremal \( \psi(\cdot) \), satisfying \( \psi(0) = (x_0, p_0) \) in local coordinates, is minimizing between \( x_0 \) and \( x \).

In other words, under the assumptions of the lemma, every (nonsingular) minimizing trajectory ceases to be minimizing beyond its first conjugate point.

### 1.3. The main results

**Theorem 1.** Let \( M \) be a smooth connected manifold of dimension \( n \), and \( \Delta \) be a smooth nonholonomic distribution of rank \( m \leq n \) on \( M \). Let \( \bar{x} \in M \). Assume that there exists a smooth Riemannian metric \( g \) on \( \Delta \) for which \( M \) is complete and no nontrivial singular path is minimizing. Then there exist a section \( X \) of \( \Delta \) on \( M \), and a closed nonempty subset \( S \) of \( M \), of Hausdorff dimension smaller than or equal to \( n - 1 \), such that \( X \) is SRS \( \bar{x}, S \).

**Remark 1.3.** If the manifold \( M \), the distribution \( \Delta \), and the metric \( g \) are moreover real-analytic, then the set \( S \) of the theorem can be chosen to be a subanalytic subset of \( M \setminus \{\bar{x}\} \), of codimension greater than or equal to one (see [27, 28] for the definition of a subanalytic set). Note that, in this case, since \( S \) is subanalytic (in \( M \setminus \{\bar{x}\} \)), it is a stratified (in the sense of Whitney) submanifold of \( M \setminus \{\bar{x}\} \).

**Remark 1.4.** If \( m = n \), then obviously there exists no singular path (it is the Riemannian situation).

**Remark 1.5.** The distribution \( \Delta \) is called fat (see [33]) at a point \( x \in M \) if, for every vector field \( X \) on \( M \) such that \( X(x) \in \Delta(x) \setminus \{0\} \),

\[
T_x M = \Delta(x) + \text{span}([X, f_i](x) \mid 1 \leq i \leq m),
\]

where \((f_1, \ldots, f_m)\) is an \( m \)-tuple of vector fields locally representing the distribution \( \Delta \). With the same notations, it is called medium-fat at \( x \) (see [4]) if

\[
T_x M = \Delta(x) + \text{span}([f_i, f_j](x) \mid 1 \leq i, j \leq m) + \text{span}([X, [f_i, f_j]](x) \mid 1 \leq i, j \leq m).
\]

If \( \Delta \) is fat at every point of \( M \), then there exists no nontrivial singular path (see [33]). On the other hand, for a generic smooth Riemannian metric \( g \) on \( M \), every nontrivial
singular path must be strictly abnormal (see [18]); it follows from [4] Theorem 3.8 that, if \( \Delta \) is medium-fat at every point of \( M \), then, for generic metrics, there exists no nontrivial minimizing singular path. Note that, if \( n \leq m(m - 1) + 1 \), then the germ of an \( m \)-tuple \((f_1, \ldots, f_m)\) of vector fields is generically (in the Whitney \( C^\infty \) topology) medium-fat (see [4]).

**Remark 1.6.** Let \( m \geq 3 \) be a positive integer, and \( G_m \) be the set of pairs \((\Delta, g)\), where \( \Delta \) is a rank \( m \) distribution on \( M \) and \( g \) is a Riemannian metric on \( \Delta \), endowed with the Whitney \( C^\infty \) topology. There exists an open dense subset \( W_m \) of \( G_m \) such that no element of \( W_m \) admits nontrivial minimizing singular paths (see [16, 17]). This means that, for \( m \geq 3 \), generically, the main assumption of Theorem 1 is satisfied.

In the next result, we are able to remove, in the compact and orientable three-dimensional case, the assumption on the absence of singular minimizing paths. Assume from now on that \( M \) is a smooth closed manifold of dimension 3 which is orientable, and denote by \( \Omega \) an orientation form on \( M \). Any nonvanishing one-form \( \alpha \) generates a smooth rank two distribution \( \Delta := \ker \alpha \). Assume that \( \Delta \) is nonholonomic on \( M \). There exists a unique smooth function \( f \) on \( M \) such that \( \alpha \wedge d\alpha = f \Omega \) on \( M \). Since \( \Delta \) is nonholonomic, the set \( \{ f \neq 0 \} \) is open and dense in \( M \). The singular set \( \Sigma_\Delta \) of \( \Delta \) is defined by

\[
\Sigma_\Delta := \{ x \in M \mid f(x) = 0 \}.
\]

Note that, if \( M \) and \( \alpha \) are analytic, then the singular set is an analytic subset of \( M \). The set \( \Sigma_\Delta \) is said to be a **Martinet surface** if, for every \( x \in \Sigma_\Delta \), \( df(x) \neq 0 \), so that \( \Sigma_\Delta \) is a smooth orientable hypersurface on \( M \). We define a **Martinet distribution** to be any nonholonomic distribution \( \Delta \) associated with a nonvanishing one-form as above such that \( \Sigma_\Delta \) is a Martinet surface. In fact, it follows from the generic classification of rank two distributions on a three-dimensional manifold (see [48], see also [11]) that, for every \( x \) in \( \Sigma_\Delta \), the distribution \( \Delta \) is, in a neighborhood of \( x \), isomorphic to \( \ker \alpha \), where the one-form \( \alpha \) is defined by \( \alpha := dx_3 - x_2^2dx_1 \) in local coordinates \((x_1, x_2, x_3)\). In this neighborhood, the Martinet surface \( \Sigma_\Delta \) coincides with the surface \( x_2 = 0 \), and the singular paths are the integral curves of the vector field \( \partial / \partial x_1 \) restricted to \( x_2 = 0 \). This situation corresponds to the so-called **Martinet case**, and these singular paths are minimizing in the context of sub-Riemannian geometry, for every smooth metric \( g \) on \( \Delta \) (see [2, 11, 32]).

**Theorem 2.** Let \( M \) be a smooth connected orientable compact Riemannian manifold of dimension three, and \( \Delta \) be a Martinet distribution on \( M \). Let \( \bar{x} \in M \). Then there exist a section \( X \) of \( \Delta \) on \( M \), and a closed nonempty subset \( S \) of \( M \), of Hausdorff dimension smaller than or equal to two, such that \( X \) is SRS\( \bar{x}, S \).

**Remark 1.7.** The compactness assumption on the manifold \( M \) can actually be dropped (see Remark 2.5). It is set to avoid technical difficulties in the proof.

### 1.4. Stabilization of nonholonomic control systems

We begin this section with a remark on the local formulation of Theorem 1. Let \( U \) be an open neighborhood of \( \bar{x} \) in \( M \) such that \( \Delta|_U \) is spanned by an \( m \)-tuple \((f_1, \ldots, f_m)\)
of smooth vector fields on $U$ which are everywhere linearly independent on $U$. Every horizontal path $x(\cdot) \in \Omega(\bar{x})$ contained in $U$ satisfies

$$\dot{q}(t) = \sum_{i=1}^{m} u_i(t) f_i(q(t)) \quad \text{for a.e. } t \in [0, 1],$$

(4)

where $u_i \in L^1([0, 1], \mathbb{R})$ for $i = 1, \ldots, m$. The function $u(\cdot) = (u_1(\cdot), \ldots, u_m(\cdot))$ is called the control associated to $x(\cdot)$, and the system (4) is a control system. Hence, Theorem 1 translated in local coordinates, yields a stabilization result for control systems of the form (4).

There are however slight differences between the geometric formulation adopted in Theorem 1 and the corresponding result for control systems. Indeed, when considering control systems of the form (4), the vector fields $f_1, \ldots, f_m$ need not be everywhere linearly independent. Moreover, a rank $m$ distribution $\Delta$ on the manifold $M$ is not necessarily globally represented by an $m$-tuple of linearly independent vector fields (for example, consider a rank two distribution on the two-dimensional sphere of $\mathbb{R}^3$).

For these reasons, we derive hereafter a stabilization result, similar to Theorem 1, valid for control systems of the form (4), and of independent interest.

Consider on the manifold $M$ the control system

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

(5)

where $f_1, \ldots, f_m$ are smooth vector fields on $M$ (not necessarily linearly independent), and the control $u = (u_1, \ldots, u_m)$ takes values in $\mathbb{R}^m$.

The system (5) is said to be (totally) nonholonomic if the $m$-tuple $(f_1, \ldots, f_m)$ satisfies Hörmander’s condition everywhere on $M$. According to the Chow–Rashevsky Theorem, any two points of $M$ can be joined by a trajectory of (5).

Let $\bar{x} \in M$ be fixed. The stabilization problem consists in finding a feedback control function $k = (k_1, \ldots, k_m) : M \to \mathbb{R}^m$ such that the closed-loop system

$$\dot{x} = \sum_{i=1}^{m} k_i(x) f_i(x)$$

(6)

is globally asymptotically stable at $\bar{x}$. It follows from the discussion above, and in particular from Brockett’s condition, that smooth or even continuous stabilizing feedbacks do not exist in general. This fact has generated a wide-ranging research with a view to deriving adapted notions for stabilization issues, such as discontinuous piecewise analytic feedbacks (see [46]), discontinuous sampling feedbacks (see [2], [57]), continuous time varying control laws (see [24]), patchy feedbacks (see [6]), and almost globally asymptotically stabilizing feedbacks (see [38]), enjoying various properties. The notion of smooth repulsive stabilizing feedback (see [39]–[41]), whose definition is recalled below, is considered in the present article.

Let $\bar{x} \in M$ be fixed. Let $S$ be a closed subset of $M$ and $k = (k_1, \ldots, k_m) : M \to \mathbb{R}^m$. The feedback $k$ is said to be smooth repulsive globally asymptotically stable at $\bar{x}$ with respect to $S$ (denoted SRS$_{\bar{x},S}$) if the following properties are satisfied:
(i) The mapping \( k \) is locally bounded on \( M \) and smooth on \( M \setminus S \).

(ii) The dynamical system (5) is globally asymptotically stable at \( \bar{x} \) in the sense of Carathéodory.

(iii) For every \( x \in M \), the Carathéodory solutions of (6) satisfy \( x(t) \not\in S \) for every \( t > 0 \).

We next associate to the control system (5) an optimal control problem. For \( x_0 \in M \) and \( T > 0 \), a control \( u \in L^\infty([0,T], \mathbb{R}^m) \) is called admissible if the solution \( x(\cdot) \) of (5) associated to \( u \) and starting at \( x_0 \) is well defined on \([0,T]\). On the set \( U_{x_0, T} \) of admissible controls, and with the previous notations, define the end-point mapping by \( E_{x_0, T}(u) := x(T) \). It is classical that \( U_{x_0, T} \) is an open subset of \( L^\infty([0,T], \mathbb{R}^m) \) and that \( E_{x_0, T} : U_{x_0, T} \to M \) is a smooth map.

A control \( u \in U_{x_0, T} \) is said to be singular if \( u \) is a critical point of the end-point mapping \( E_{x_0, T} \); in this case the corresponding trajectory \( x(\cdot) \) is said to be singular.

Let \( x_0 \) and \( x_1 \) be two points of \( M \), and \( T > 0 \). Consider the optimal control problem of determining, among all the trajectories of (5) steering \( x_0 \) to \( x_1 \), a trajectory minimizing the cost

\[
C_U(T, u) = \int_0^T u(t)^T U(x(t))u(t) \, dt, \tag{7}
\]

where \( U \) takes values in the set \( S^+_m \) of symmetric positive definite \( m \times m \) matrices.

**Theorem 3.** Assume that there exists a smooth function \( U : M \to S^+_m \) such that no nontrivial singular trajectory of the control system (5) minimizes the cost (7) between its extremities. Then there exist a mapping \( k : M \to \mathbb{R}^m \), and a closed nonempty subset \( S \) of \( M \), of Hausdorff dimension smaller than or equal to \( n - 1 \), such that \( k \) is an SRS\(_{\bar{x}, S}\) feedback.

**Remark 1.8.** The same remarks as those following Theorem 1 are valid. In particular, it is proved in [18] that, for a fixed smooth function \( U : M \to S^+_m \), if \( m \geq 3 \), then there exists an open and dense subset \( O_m \) of the set of \( m \)-tuples of smooth vector fields on \( M \) such that the optimal control problem (5)–(7) defined with an \( m \)-tuple of \( O_m \) does not admit nontrivial minimizing singular trajectories.

### 2. Proof of the main results

This section is organized as follows. In Section 2.1 we recall some tools of nonsmooth analysis that are required to prove our main results. Section 2.2 is devoted to the proof of Theorem 1. We first define a Bolza problem, equivalent to the sub-Riemannian problem, for which we derive some fine properties of the value function and of optimal trajectories. In particular, we prove that the value function is smooth outside a singular set which is defined using a specific notion of subdifferential. Theorem 1 is then derived in Section 2.2.4. Theorem 2 is proved in Section 2.3. The proof of Theorem 3 is similar to the one of Theorem 1 and therefore skipped.

\(^4\) Note that, in what follows, the value of \( T \) is not important. It can be assumed for instance that \( T = 1 \).
2.1. Preliminaries: some tools of nonsmooth analysis

Let $M$ be a smooth manifold of dimension $n$.

2.1.1. Viscosity subsolutions, supersolutions and solutions. For an introduction to viscosity solutions of Hamilton–Jacobi equations, we refer the reader to [7, 8, 25, 31]. Assume that $F : T^*M \times \mathbb{R} \to \mathbb{R}$ is a continuous function on $M$. A function $u : U \to \mathbb{R}$, continuous on the open set $U \subset M$, is a viscosity subsolution (resp., supersolution) of $F(x, du(x), u(x)) = 0$, if, for every $C^1$ function $\phi : U \to \mathbb{R}$ (resp., $\psi : U \to \mathbb{R}$) satisfying $\phi \geq u$ (resp., $\psi \leq u$), and every point $x_0 \in U$ satisfying $\phi(x_0) = u(x_0)$ (resp., $\psi(x_0) = u(x_0)$), we have $F(x_0, d\phi(x_0), u(x_0)) \leq 0$ (resp., $F(x_0, d\psi(x_0), u(x_0)) \leq 0$). A function is a viscosity solution of (8) if it is both a viscosity subsolution and a viscosity supersolution of (8).

2.1.2. Generalized differentials. Let $u : U \to \mathbb{R}$ be a continuous function on an open set $U \subset M$. The viscosity subdifferential of $u$ at $x \in U$ is the subset of $T^*_x M$ defined by

$$D^- u(x) := \{d\psi(x) | \psi \in C^1(U) \text{ and } f - \psi \text{ attains a global minimum at } x\}.$$ 

Similarly, the viscosity superdifferential of $u$ at $x$ is the subset of $T^*_x M$ defined by

$$D^+ u(x) := \{d\phi(x) | \phi \in C^1(U) \text{ and } f - \phi \text{ attains a global maximum at } x\}.$$ 

Notice that $u$ is a viscosity subsolution (resp., supersolution) of (8) if and only if, for every $x \in U$ and every $\xi \in D^+ u(x)$ (resp., $\xi \in D^- u(x)$), one has $F(x, \xi, u(x)) \leq 0$ (resp., $F(x, \xi, u(x)) \geq 0$).

The limiting subdifferential of $u$ at $x \in U$ is the subset of $T^*_x M$ defined by

$$\partial_L u(x) := \{ \lim_{k \to \infty} \xi_k | \xi_k \in D^- u(x_k), x_k \to x \}.$$ 

By construction, the graph of the limiting subdifferential is closed in $T^* M$. Moreover, the function $u$ is locally Lipschitzian on its domain if and only if the limiting subdifferential of $u$ at any point is nonempty and its graph is locally bounded (see [22, 42]).

Let $u : U \to \mathbb{R}$ be a locally Lipschitzian function. The Clarke’s generalized gradient of $u$ at the point $x \in U$ is the subset of $T^*_x M$ defined by

$$\partial u(x) := \text{co}(\partial_L u(x)),$$

that is, the convex hull of the limiting differential of $u$ at $x$. Notice that, for every $x \in U$,

$$D^- u(x) \subset \partial_L u(x) \subset \partial u(x) \quad \text{and} \quad D^+ u(x) \subset \partial u(x).$$

2.1.3. Locally semiconcave functions. For an introduction to semiconcavity, we refer the reader to [15]. A function $u : U \to \mathbb{R}$, defined on the open set $U \subset M$, is locally
semiconcave on \( U \) if for every \( x \in U \), there exist a neighborhood \( U_x \) of \( x \) and a smooth diffeomorphism \( \varphi_x : U_x \to \varphi_x(U_x) \subset \mathbb{R}^n \) such that \( f \circ \varphi_x^{-1} \) is locally semiconcave on the open subset \( U_x = \varphi_x(U_x) \subset \mathbb{R}^n \). For completeness, we recall that the function \( u : U \to \mathbb{R} \), defined on the open set \( U \subset \mathbb{R}^n \), is locally semiconcave on \( U \) if for every \( x \in U \) there exist \( C, \delta > 0 \) such that

\[
\mu u(y) + (1 - \mu)u(x) - u(\mu x + (1 - \mu)y) \leq \mu(1 - \mu)C|x - y|^2
\]

for all \( x, y \in \bar{x} + \delta B \) (where \( B \) denotes the open unit ball in \( \mathbb{R}^n \)) and every \( \mu \in [0, 1] \). This is equivalent to saying that \( u \) can be written locally as

\[
u(x) = (u(x) - C|x|^2) + C|x|^2, \quad \forall x \in \bar{x} + \delta B,
\]

that is, as the sum of a concave function and a smooth function. Note that every semiconcave function is locally Lipschitzian on its domain, and thus, by Rademacher’s Theorem, is differentiable almost everywhere on its domain. The following result will be useful in the proof of our theorems.

**Lemma 2.1.** Let \( u : U \to \mathbb{R} \) be a function defined on an open set \( U \subset \mathbb{R}^n \). If, for every \( \bar{x} \in U \), there exist a neighborhood \( V \) of \( \bar{x} \) and a positive real number \( \sigma \) such that, for every \( x \in V \), there exists \( p_x \in \mathbb{R}^n \) such that

\[
u(y) \leq u(x) + \langle p_x, y - x \rangle + \sigma|y - x|^2
\]

for every \( y \in V \), then the function \( u \) is locally semiconcave on \( U \).

**Proof.** Without loss of generality, assume that \( V \) is an open ball \( B \). Let \( x, y \in B \) and \( \mu \in [0, 1] \). The point \( \bar{x} := \mu x + (1 - \mu)y \) belongs to \( B \) by convexity. By assumption, there exists \( \bar{p} \in \mathbb{R}^n \) such that

\[
u(z) \leq u(\bar{x}) + \langle \bar{p}, z - \bar{x} \rangle + \sigma|z - \bar{x}|^2, \quad \forall z \in B.
\]

Hence,

\[
\mu u(y) + (1 - \mu)u(x) \leq u(\bar{x}) + \mu \sigma|x - \bar{x}|^2 + (1 - \mu)\sigma|y - \bar{x}|^2
\]

\[
\leq u(\bar{x}) + (\mu(1 - \mu)^2\sigma + (1 - \mu)\mu^2\sigma)|x - y|^2
\]

\[
\leq u(\bar{x}) + 2\mu(1 - \mu)\sigma|x - y|^2,
\]

and the conclusion follows. \( \square \)

The converse result can be stated as follows.

**Proposition 4.** Let \( U \) be an open and convex subset of \( \mathbb{R}^n \) and \( u : U \to \mathbb{R} \) be a function which is \( C \)-semiconcave on \( U \), that is,

\[
\mu u(y) + (1 - \mu)u(x) - u(\mu x + (1 - \mu)y) \leq \mu(1 - \mu)C|x - y|^2
\]

for all \( x, y \in U \). Then, for every \( x \in U \) and every \( p \in D^+u(x) \), we have

\[
u(y) \leq u(x) + \langle p, y - x \rangle + \frac{C}{2}|y - x|^2, \quad \forall y \in \Omega,
\]

In particular, \( D^+u(x) = \partial u(x) \) for every \( x \in U \).
Remark 2.1. As a consequence (see [15, 42]), if $u : U \to \mathbb{R}$ is locally semiconcave on an open set $U \subset M$, then, for every $x \in U$,

$$\partial_L u(x) = \{ \lim_{k \to \infty} du(x_k) | x_k \in D_u, x_k \to x \},$$

where $D_u$ denotes the set of points of $U$ at which $u$ is differentiable.

The following result is useful in obtaining several characterizations of the singular set of a given locally semiconcave function. We refer the reader to [15, 42] for its proof.

**Proposition 5.** Let $U$ be an open subset of $M$ and $u : U \to \mathbb{R}$ be a function which is locally semiconcave on $U$. Then, for every $x \in U$, $u$ is differentiable at $x$ if and only if $\partial u(x)$ is a singleton.

The next result will turn out to be useful (see [15, Corollary 3.3.8]).

**Proposition 6.** Let $u : U \to \mathbb{R}$ be a function defined on an open set $U \subset M$. If both $u$ and $-u$ are locally semiconcave on $U$, then $u$ is of class $C^{1,1}_{\text{loc}}$ on $U$.

2.1.4. Singular sets of semiconcave functions. Let $u : U \to \mathbb{R}$ be a function which is locally semiconcave on the open set $U \subset M$. We recall that since such a function is locally Lipschitzian on $U$, its limiting subdifferential is always nonempty on $U$. We define the singular set of $u$ as

$$\Sigma(u) := \{ x \in U | \partial_L u(x) \text{ is not a singleton} \}.$$  

Alberti, Ambrosio and Cannarsa [5] proved the following result.

**Theorem 7.** Let $U$ be an open subset of $M$. The singular set of a locally semiconcave function $u : U \to \mathbb{R}$ is of Hausdorff dimension smaller than or equal to $n - 1$.

The following lemma, proved in the Appendix (Section 3.1), will be useful for the proof of Theorems 1 and 2.

**Lemma 2.2.** Let $u : U \to \mathbb{R}$ be a locally semiconcave function on an open subset $U \subset M$ and $\gamma : [a, b] \to U$ be a locally Lipschitzian curve. Then, for every measurable map $p : [a, b] \to T^*M$ satisfying

$$p(t) \in D^+ u(\gamma(t)) \quad \text{for a.e. } t \in [a, b],$$

we have

$$\frac{d}{dt}(u(\gamma(t))) = p(t)(\dot{\gamma}(t)) \quad \text{for a.e. } t \in [a, b].$$

\[5\] In fact, this result has been strengthened later as follows: the singular set of a locally semiconcave function is countably $n - 1$-rectifiable, i.e., contained in a countable union of locally Lipschitzian hypersurfaces of $M$ (see [15, 42]).
2.2. Proof of Theorem 1

From now on, assume that the assumptions of Theorem 1 hold. In particular, assume that there exists no nontrivial singular minimizing path for the metric \( g \).

2.2.1. An equivalent optimal control problem. Define the running cost \( L_g \) by

\[
L_g(x, v) := g_x(v, v)
\]

for \( x \in M \) and \( v \in \Delta(x) \), and define the functional \( J_g : \Omega_\Delta(\bar{x}) \to \mathbb{R}^+ \) by

\[
J_g(\gamma) := \int_0^1 L_g(\gamma(t), \dot{\gamma}(t)) \, dt.
\]

The Bolza optimization problem under consideration, denoted by \((BP)_g,1\), consists in minimizing the functional \( J_g \), called energy, over all horizontal paths \( \gamma \) joining \( \bar{x} \) to \( x \in M \). Since \( M \) is connected and complete, and since the running cost \( L_g \) is coercive in every fiber, for every \( x \in M \) there exists a horizontal path \( \gamma \in \Omega_\Delta(\bar{x}, x) \) minimizing the energy \( J_g \). The value function associated to the Bolza problem \((BP)_g,1\) is defined by

\[
V_{g,1}(x) := \inf \{ J_g(\gamma) \mid \gamma \in \Omega_\Delta(\bar{x}, x) \}
\]

for every \( x \in M \).

Note that the length of a horizontal path \( \gamma \), defined by \((2)\), does not depend on its parametrization. Hence, up to reparametrizing, one can assume that the horizontal paths are parametrized by arc-length, i.e., \( g_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t)) = 1 \). In this case, the length minimizing problem is equivalent to the minimal time problem. Moreover, if all paths are defined on the same interval, then length and energy minimization problems are equivalent, and the value function \( V_{g,1} \) satisfies

\[
V_{g,1}(x) = d_{SR}(\bar{x}, x)^2.
\]

(13)

In other terms, the problem of minimizing the length between two points \( \bar{x} \) and \( x \), for the sub-Riemannian manifold \((M, \Delta, g)\), is equivalent to the Bolza problem \((BP)_g,1\).

We next provide another equivalent formulation of this optimization problem, in terms of optimal control theory, which will be useful in the proofs of Theorems 1 and 2. Let \( x \in M \), and let \( \gamma \) be a minimizing horizontal path joining \( \bar{x} \) to \( x \). Since \( \gamma \) is necessarily non-self-intersecting, there exists a tubular neighborhood \( \mathcal{V} \) of the path \( \gamma \) in \( M \), and \( m \) smooth vector fields \( f_1, \ldots, f_m \) on \( \mathcal{V} \), such that

\[
\Delta(x) = \text{span} \{ f_i(x) \mid i = 1, \ldots, m \}
\]

for every \( x \in \mathcal{V} \). Then every horizontal path \( x(\cdot) \) contained in \( \mathcal{V} \) is a solution of the control system

\[
\dot{x}(t) = \sum_{i=1}^m u_i(t) f_i(x(t)),
\]
where \( u(\cdot) = (u_1(\cdot), \ldots, u_m(\cdot)) \in L^\infty([0, 1], \mathbb{R}^m) \) is called the control. Without loss of generality, we assume that the \( m \)-tuple \((f_1, \ldots, f_m)\) of vector fields is orthonormal for the metric \( g \). In these conditions, the energy of the path \( x(\cdot) \) is

\[
J_g(x(\cdot)) = \int_0^1 \sum_{i=1}^m u_i(t)^2 \, dt.
\]

Since the optimal control problem does not admit any nontrivial singular minimizing path, it follows from the Pontryagin maximum principle (see [35]) that every minimizing path \( \gamma \) is the projection of a normal extremal \( \psi(\cdot) = (\gamma(\cdot), p(\cdot)) \), associated with the control \( u(\cdot) = (u_1(\cdot), \ldots, u_m(\cdot)) \), where

\[
u_i(t) = \langle p(t), f_i(\gamma(t)) \rangle, \quad i = 1, \ldots, m.
\]

### 2.2.2. Properties of the value function \( V_{g,\Delta} \)

Consider the Hamiltonian function \( H_{g,\Delta} : T^*M \to \mathbb{R} \) defined by

\[
H_{g,\Delta}(x, p) := \max_{v \in \Delta(x)} \left( p(v) - \frac{1}{2} g_x(v, v) \right).
\]

Note that this Hamiltonian coincides with the Hamiltonian \( H \) defined by [3] (as can be seen in local coordinates).

**Proposition 8.** If the distribution \( \Delta \) is nonholonomic on \( M \), then the value function \( V_{g,\Delta} : M \to \mathbb{R} \) is continuous on \( M \) and is a viscosity solution of the Hamilton–Jacobi equation

\[
-\frac{1}{2} V_{g,\Delta}(x) + H_{g,\Delta}(x, \frac{1}{2} dV_{g,\Delta}(x)) = 0, \quad \forall x \in M \setminus \{\bar{x}\}.
\]

Note that this proposition still holds if there exist some minimizing singular paths.

**Proof.** The continuity of \( V_{g,\Delta} \) follows from the continuity of the sub-Riemannian distance, associated to the metric \( g \), on \( M \times M \). Notice that, since the running cost \( L_g \) is coercive in the fibers, and since \( M \) is connected, for every \( x \in M \setminus \{\bar{x}\} \) there exists a horizontal path \( \gamma(\cdot) \in \Omega_{\Delta}(\bar{x}, x) \) such that

\[
V_{g,\Delta}(x) = J_g(\gamma(\cdot)) = \int_0^1 L_g(\gamma(s), \dot{\gamma}(s)) \, ds.
\]

Let us prove that \( V_{g,\Delta} \) is a viscosity solution of [15] on \( M \setminus \{\bar{x}\} \). Let \( x \in M \setminus \{\bar{x}\} \), and let \( \gamma : [0, 1] \to M \) be a horizontal path joining \( \bar{x} \) to \( x \). For \( t \in (0, 1) \), there exists \( \tilde{\gamma} \in \Omega_{\Delta}(\bar{x}, \gamma(t)) \) such that

\[
V_{g,\Delta}(\gamma(t)) = \int_0^1 L_g(\tilde{\gamma}(s), \dot{\tilde{\gamma}}(s)) \, ds = \frac{1}{t} \int_0^t L_g\left( \tilde{\gamma}\left( \frac{s}{t} \right), \dot{\tilde{\gamma}}\left( \frac{s}{t} \right) \right) \, ds.
\]
Define $\gamma_1 \in \Omega_\Delta(\bar{x}, x)$ by
\[
\gamma_1(s) := \begin{cases} 
\tilde{\gamma}(s/t) & \text{if } s \in [0, t], \\
\gamma(s) & \text{if } s \in [t, 1].
\end{cases}
\]

Then
\[
V_{g,\Delta}(x) \leq \int_0^1 L_g(\gamma_1(s), \dot{\gamma}_1(s)) \, ds \leq \int_0^t L_g\left(\tilde{\gamma}\left(\frac{s}{t}\right), \frac{1}{t}\tilde{\gamma}'\left(\frac{s}{t}\right)\right) \, ds + \int_t^1 L_g(\gamma(s), \dot{\gamma}(s)) \, ds \\
\leq \frac{1}{t} V_{g,\Delta}(\gamma(t)) + \int_t^1 L_g(\gamma(s), \dot{\gamma}(s)) \, ds.
\]

If $\phi : M \to \mathbb{R}$ is a $C^1$ function satisfying $\phi \geq V_{g,\Delta}$ and $\phi(x) = V_{g,\Delta}(x)$, then
\[
\phi(x) = V_{g,\Delta}(x) \leq \frac{1}{t} V_{g,\Delta}(\gamma(t)) + \int_t^1 L_g(\gamma(s), \dot{\gamma}(s)) \, ds \\
\leq \frac{1}{t} \phi(\gamma(t)) + \int_t^1 L_g(\gamma(s), \dot{\gamma}(s)) \, ds.
\]

Letting $t$ tend to 1, and considering all $C^1$ horizontal paths joining $\bar{x}$ to $x$, we infer that, for every $v \in \Delta(x)$,
\[
d\phi(x)(v) \leq \phi(x) + L_g(x, v).
\]

On the other hand, consider some path $\gamma \in \Omega_\Delta(\bar{x}, x)$ satisfying $V_{g,\Delta}(x) = J_g(\gamma)$. For every $t \in (0, 1)$, up to a change of variable, this path is necessary minimizing between $\bar{x}$ and $\gamma(t)$. Therefore, for every $t \in (0, 1)$,
\[
V_{g,\Delta}(x) = \frac{1}{t} V_{g,\Delta}(\gamma(t)) + \int_t^1 L_g(\gamma(s), \dot{\gamma}(s)) \, ds.
\]

If $\psi : U \to \mathbb{R}$ is a $C^1$ function satisfying $\psi \leq V_{g,\Delta}$ and $\psi(x) = V_{g,\Delta}(x)$, then
\[
\psi(x) \geq \frac{1}{t} \psi(\gamma(t)) + \int_t^1 L_g(\gamma(s), \dot{\gamma}(s)) \, ds.
\]

As previously, passing to the limit yields the existence of $v \in \Delta(x)$ such that
\[
d\psi(x)(v) \geq \psi(x) + L_g(x, v).
\]

The conclusion follows. \qed

**Remark 2.2.** Notice that, since $V_{g,\Delta}$ is a viscosity solution of (15) on $M \setminus \{\bar{x}\}$, we have, for every horizontal path $\gamma : [a, b] \to M \setminus \{\bar{x}\}$ (with $a < b$),
\[
V_{g,\Delta}(\gamma(b)) - V_{g,\Delta}(\gamma(a)) \leq \int_a^b V_{g,\Delta}(\gamma(s)) \, ds + \int_a^b L_g(\gamma(s), \dot{\gamma}(s)) \, ds.
\]
Remark 2.3. We also notice that, since $V_{g,\Delta}$ is a viscosity solution of (15) on $M \setminus \{\bar{x}\}$, we have

$$-\frac{1}{2} V_{g,\Delta}(x) + H_{\Delta}(x, \frac{1}{2} \xi) = 0, \quad \forall x \in M \setminus \{\bar{x}\}, \forall \xi \in \partial L V_{g,\Delta}(x). \quad (16)$$

Finally, we have the following result.

Proposition 9. If the distribution $\Delta$ is nonholonomic on $M$, then the value function $V_{g,\Delta}$ is continuous on $M$, and locally semiconcave on $M \setminus \{\bar{x}\}$.

We just sketch the proof of Proposition 9; we refer the reader to [14, 42] for further details.

Proof. Recall that since $M$ is connected and complete, for every $x \in M \setminus \{\bar{x}\}$ there exists a horizontal path $\gamma \in \Omega_{\Delta}(\bar{x}, x)$ such that $V_{g,\Delta}(x) = J_g(\gamma) = \int_0^1 L_g(\gamma(s), \dot{\gamma}(s)) \, ds.$

By assumption, this minimizing path $\gamma$ is necessarily nonsingular, and thus it is the projection of a normal extremal. It is well known (see [1, 47]) that, for every $x \in M \setminus \{\bar{x}\}$, there exists a neighborhood $V$ of $x$ in $M \setminus \{\bar{x}\}$ such that the set of cotangent vectors $p_0 \in T^* \bar{x}M$ for which $\exp_{\bar{x}}(p_0) \in V$ and the projection of the corresponding normal extremal minimizes the length between $\bar{x}$ and $\exp_{\bar{x}}(p_0)$, is compact in $T^* \bar{x}M$. On the other hand, we know from [43, Proposition 4, p. 153] that if $\xi \in \partial L V_{g,\Delta}(x)$, then there exists a normal extremal $\psi : [0, 1] \to T^* M$ whose projection is minimizing between $\bar{x}$ and $x$ and such that $\psi(1) = (x, \frac{1}{2} \xi)$. This proves that the function $V_{g,\Delta}$ is locally Lipschitzian on $M \setminus \{\bar{x}\}$.

Let $x \in M \setminus \{\bar{x}\}$, and let $\bar{\gamma}$ be a minimizing horizontal path joining $\bar{x}$ to $x$. By assumption, this path is nonsingular, and thus it is not a critical point of the end-point mapping $E_{\bar{x}}$. Hence, there exists a submanifold $N$ of $\Omega_{\Delta}(\bar{x}, x)$, of dimension $n$, such that the mapping

$$E : N \to M, \quad \gamma(\cdot) \mapsto E_{\bar{x}}(\gamma(\cdot)) = \gamma(1),$$

is a local diffeomorphism from a neighborhood of $\bar{\gamma}(\cdot)$ in $N$ into a neighborhood $\mathcal{W}$ of $x = \bar{\gamma}(1)$. We infer that, for every $y \in \mathcal{W}$,

$$V_{g,\Delta}(y) \leq J_g(E^{-1}(y)).$$

Since $J_g$ is smooth on the submanifold $N$, up to diffeomorphism, one can put a parabola over the graph of $J_g$ on $N$, and thus over the graph of the function $V_{g,\Delta}$ at every $x \in M \setminus \{\bar{x}\}$. The second-order term of this parabola depends on the minimizing controls which are associated to the points $x$. Using the compactness of the minimizers that we recalled above, we deduce that the function $V_{g,\Delta}$ is locally semiconcave on $M \setminus \{\bar{x}\}$. □

In the following, the singular set of $V_{g,\Delta}$, denoted $\Sigma(V_{g,\Delta})$, is

$$\Sigma(V_{g,\Delta}) := \{x \in M \setminus \{\bar{x}\} \mid \partial_t V_{g,\Delta}(x) \text{ is not a singleton}\}.$$ 

Recall that since the function $V_{g,\Delta}$ is locally semiconcave on $M \setminus \{\bar{x}\}$, its limiting subdifferential is nonempty at any point of $M \setminus \{\bar{x}\}$ (see [15]).
2.2.3. Properties of optimal trajectories of \((BP)_{g,\Delta}\). We stress that, due to the assumption of the absence of singular minimizing paths, every minimizing curve of the Bolza problem \((BP)_{g,\Delta}\) is the projection of a normal extremal, i.e., an integral curve of the Hamiltonian vector field \(\vec{H}\) defined by \([3]\), associated with \(H\). In particular, every minimizing curve of \((BP)_{g,\Delta}\) is smooth on \([0, 1]\).

**Lemma 2.3.** For every \(x \in M \setminus \{\bar{x}\}\) and every \(\xi \in \partial_L V_{g,\Delta}(x)\), there exists a unique normal extremal \(\psi(\cdot) : [0, 1] \to T^*M\) whose projection \(\gamma(\cdot) : [0, 1] \to M\) is minimizing between \(\bar{x}\) and \(x\), and such that \(\psi(1) = (x, \frac{1}{2}\xi)\) in local coordinates. In addition, \(\psi(\cdot)\) is the unique (up to a scalar factor) normal extremal lift of \(\gamma(\cdot)\).

**Proof.** The first part of the statement is a consequence of \([43]\), Proposition 4, p. 153. Uniqueness follows from the Cauchy–Lipschitz Theorem. Uniqueness (up to a scalar factor) of the normal extremal lift of \(\gamma(\cdot)\) is a consequence of the absence of singular minimizing paths (see \([43]\), Remark 8, p. 149).

**Lemma 2.4.** Let \(x \in M \setminus \{\bar{x}\}\) and \(\gamma(\cdot) : [0, 1] \to M\) be a minimizing curve of \((BP)_{g,\Delta}\) such that \(\gamma(1) = x\). Then, for every \(t \in (0, 1)\), the curve \(\tilde{\gamma}^t(\cdot) : [0, 1] \to M\) defined by \(\tilde{\gamma}^t(s) := \gamma(st)\) for \(s \in [0, 1]\) is the unique minimizing curve of \((BP)_{g,\Delta}\) steering \(\bar{x}\) to \(\gamma(t)\). Moreover, \(\tilde{\gamma}^t(\cdot)\) is the projection of the normal extremal \((\tilde{\gamma}^t(\cdot), \tilde{p}^t(\cdot))\) in local coordinates, by \(\tilde{p}^t(s) = tp(st)\) for every \(s \in [0, 1]\).

**Proof.** We argue by contradiction. If there is another horizontal curve \(\gamma_2(\cdot) : [0, 1] \to M\) which minimizes the sub-Riemannian distance between \(\bar{x}\) and \(\gamma(t)\), then there exists a nontrivial minimizing path \(x(\cdot)\) joining the points \(\gamma(t)\) and \(\gamma(1) = x\), and having two distinct normal extremal lifts \(\psi_1(\cdot)\) and \(\psi_2(\cdot)\). Then the extremal \(\psi_1(\cdot) - \psi_2(\cdot)\) is an abnormal extremal lift of the path \(x(\cdot)\). Hence, the path \(x(\cdot)\) is singular and minimizing, and this contradicts our assumption.

We next prove that the adjoint vector associated to \(\tilde{\gamma}^t(\cdot)\) is given by \(\tilde{p}^t(s) = tp(st)\) for \(s \in [0, 1]\). In local coordinates, by the expression \([14]\) of normal controls, \(\gamma(\cdot)\) is a solution of the system

\[
\dot{\gamma}(t) = \sum_{i=1}^{n} \langle p(t), f_i(\gamma(t)) \rangle f_i(\gamma(t)) \quad \text{for a.e. } t \in [0, 1].
\]

Hence, \(\tilde{\gamma}^t(\cdot)\) is a solution of

\[
\frac{d}{ds} \tilde{\gamma}^t(s) = t \sum_{i=1}^{n} \langle p(st), f_i(\tilde{\gamma}^t(s)) \rangle f_i(\tilde{\gamma}^t(s)) \quad \text{for a.e. } s \in [0, 1].
\]

The conclusion follows.

**Lemma 2.5.** Any normal extremal \(\psi(\cdot) : [0, 1] \to T^*M\) whose projection is minimizing between \(\bar{x}\) and \(x \in M \setminus \{\bar{x}\}\) satisfies \(\xi \in \partial_L V_{g,\Delta}(x)\), where \(\psi(1) = (x, \frac{1}{2}\xi)\) in local coordinates.
Proof. Let \( \psi(\cdot) : [0, 1] \rightarrow T^*M \) be a normal extremal whose projection \( \gamma(\cdot) \) is minimizing between \( \bar{x} \) and \( x \in M \setminus \{\bar{x}\} \). Since \( V_{\bar{g}, \Delta} \) is locally semiconcave on \( M \setminus \{\bar{x}\} \), its limiting subdifferential is always nonempty on \( M \setminus \{\bar{x}\} \). We infer from Lemmas 2.3 and 2.4 that, for every \( t \in (0, 1) \), we have \( \partial_{\bar{g}}V_{\bar{g}, \Delta}(\gamma(t)) = \{\xi(t)\} \), where \( \psi(\cdot) = (s(t), \frac{1}{2}s(t)) \xi(t) \) in local coordinates. Consider a sequence \( (t_k) \) of real numbers converging to 1. Then, on the one hand, the sequence \( (\psi(t_k)) \) converges to \( \psi(1) \), and on the other hand, by construction of the limiting subdifferential, \( \xi = \xi(1) \in \partial_{\bar{g}}V_{\bar{g}, \Delta}(x) \).

\[ \square \]

Lemma 2.6. The following inclusion holds:

\[ \Sigma(V_{\bar{g}, \Delta}) \setminus \Sigma(V_{\bar{g}, \Delta}) \subset C_{\min}(\bar{x}) \cup \{\bar{x}\}. \]

In particular, the set \( \Sigma(V_{\bar{g}, \Delta}) \) is of Hausdorff dimension smaller than or equal to \( n - 1 \).

Proof. Let \( x \in \Sigma(V_{\bar{g}, \Delta}) \setminus \Sigma(V_{\bar{g}, \Delta}) \) be such that \( x \neq \bar{x} \). By definition, the set \( \partial_{\bar{g}}V_{\bar{g}, \Delta}(x) \) is a singleton. Hence by Lemmas 2.3 and 2.4 there is a unique minimizing path \( \gamma(\cdot) \in \Omega(\bar{x}, x) \) and a unique normal extremal \( \psi(\cdot) : [0, 1] \rightarrow T^*M \) such that \( \gamma(\cdot) = \pi(\psi(\cdot)) \); moreover, \( \partial_{\bar{g}}V_{\bar{g}, \Delta}(x) = \{\xi\} \), where \( \psi(1) = (x, \frac{1}{2} \bar{\xi}) \) in local coordinates. We argue by contradiction; if \( x \notin C_{\min}(\bar{x}) \), then the exponential mapping \( \exp_{\bar{g}} \) is not singular at \( p_0 \), where \( \psi(0) = (\bar{x}, p_0) \) in local coordinates. Furthermore, since \( x \in \Sigma(V_{\bar{g}, \Delta}) \), there is a sequence \( (x_k) \) of points in \( \Sigma(V_{\bar{g}, \Delta}) \) which converges to \( x \). For every \( k \), the set \( \partial_{\bar{g}}V_{\bar{g}, \Delta}(x_k) \) admits at least two elements. Hence for every \( k \), there are two distinct normal extremals \( \psi^1_k(\cdot), \psi^2_k(\cdot) : [0, 1] \rightarrow T^*M \) whose projections \( \gamma_k^1(\cdot), \gamma_k^2(\cdot) \) are minimizing between \( \bar{x} \) and \( x_k \). Since the limiting subdifferential of \( V_{\bar{g}, \Delta} \) is a singleton, the sequences \( (\psi^1_k(1)), (\psi^2_k(1)) \) necessarily converge to \( \psi(1) \). Moreover, by regularity of the Hamiltonian flow, the sequences \( (\psi^1_k(0)), (\psi^2_k(0)) \) converge to \( \psi(0) \). But the exponential mapping \( \exp_{\bar{g}} \) must be a local diffeomorphism from a neighborhood of \( p_0 \) into a neighborhood of \( \pi(\psi(1)) \). This is a contradiction.

The second part of the lemma follows from the fact that the singular set \( \Sigma(V_{\bar{g}, \Delta}) \) is of Hausdorff dimension \( \leq n - 1 \) (see Theorem 7), and of the fact that \( C_{\min}(\bar{x}) \) is contained in \( C(\bar{x}) \) which is of Hausdorff dimension \( \leq n - 1 \) by Theorem 3.4.3). \[ \square \]

Lemma 2.7. The function \( V_{\bar{g}, \Delta} \) is of class \( C^1 \) on the open set \( M \setminus (\Sigma(V_{\bar{g}, \Delta}) \cup \{\bar{x}\}) \).

Proof. The set \( \partial_{\bar{g}}V_{\bar{g}, \Delta}(x) \) is a singleton for every \( x \) in the set \( M \setminus (\Sigma(V_{\bar{g}, \Delta}) \cup \{\bar{x}\}) \) which is open in \( M \). From Remark 2.1 and the fact that \( u \) is differentiable at some \( x \in M \setminus \{\bar{x}\} \) if and only if \( x \notin \Sigma(u) \), we infer that \( V_{\bar{g}, \Delta} \) is of class \( C^1 \) on \( M \setminus (\Sigma(V_{\bar{g}, \Delta}) \cup \{\bar{x}\}) \). \[ \square \]

Lemma 2.8. Let \( x \in M \setminus \{\bar{x}\} \) and \( \bar{\gamma}(\cdot) : [0, 1] \rightarrow M \) be a minimizing curve of \((BP)_{\bar{g}, \Delta}\) such that \( \bar{\gamma}(1) = x \). Let \( U_x \) be an open neighborhood of \( x \) and \( \varphi_x : U_x \rightarrow \varphi_x(U_x) \subset \mathbb{R}^n \) be a smooth diffeomorphism such that \( V := V_{\bar{g}, \Delta} \circ \varphi_x^{-1} \) is locally semiconcave on the open subset \( U := \varphi_x(U_x) \subset \mathbb{R}^n \). Let \( t \in (0, 1) \) be such that \( \bar{\gamma}(s) \in U_x \) for every \( s \in [t, 1] \). Then there exist a neighborhood \( W_t \) of \( \bar{\gamma}(t) \) and \( \sigma(t) > 0 \) such that

\[ V(y) \geq V(\varphi_x(\bar{\gamma}(t)))+dV(\varphi_x(\bar{\gamma}(t)))(y-\varphi_x(\bar{\gamma}(t)))-\sigma(t)|y-\varphi_x(\bar{\gamma}(t))|^2, \quad \forall y \in W_t. \]
Proof. Without loss of generality, we assume that \( M = \mathbb{R}^n \), that \( \varphi_x \) is the identity, and that the closure of \( U_x \) is a compact subset of \( M \setminus \{ \bar{x} \} \). Set \( x_s := \tilde{\gamma}(s) \) for every \( s \in [t, 1] \).
Since \( V = V_{g, \Delta} \) is locally semiconcave on \( M \setminus \{ \bar{x} \} \), there exists \( \sigma \in \mathbb{R} \) such that
\[
V(y) \leq V(x_s) + dV(x_s)(y - x_s) + \sigma|y - x_s|^2, \quad \forall y \in U, \forall s \in [t, 1].
\] (18)
The horizontal path \( \tilde{\gamma}(\cdot) : [0, 1 - t] \to M \), defined by
\[
\tilde{\gamma}(s) := \gamma(1 - s), \quad \forall s \in [0, 1 - t],
\]
is minimizing between \( x \) and \( \gamma(t) \). Hence, by assumption, it is nonsingular, and thus it is not a critical point of the end-point mapping
\[
E_t : \Omega_{\Delta}(x) \to M, \quad \gamma(\cdot) \mapsto E(\gamma(\cdot)).
\]
Therefore, there exists a submanifold \( N \) of \( \Omega_{\Delta}(x) \) of dimension \( n \) such that the mapping
\[
E_t : N \to M, \quad \gamma(\cdot) \mapsto E(\gamma(\cdot)),
\]
is a local diffeomorphism from a neighborhood of \( \tilde{\gamma}(\cdot) \) in \( N \) into a neighborhood \( W_t \) of \( \tilde{\gamma}(1 - t) = x_t \). From Remark 2.2, we infer that, for every \( y \in W_t \),
\[
V(y) \geq V(x_t) + \int_t^1 V(\gamma_s) \, ds - \int_t^1 L_g(\gamma_s, \dot{\gamma}_s) \, ds.
\] (19)
where \( \gamma_s(\cdot) : [t, 1] \to M \) is defined by
\[
\gamma_s := E_t^{-1}(y)(1 - s), \quad \forall s \in [t, 1].
\]
By (18), we have
\[
- \int_t^1 V(\gamma_s) \, ds \geq - \int_t^1 (V(x_s) + dV(x_s)(\gamma_s - x_s) + \sigma|\gamma_s - x_s|^2) \, ds.
\] (20)
Moreover, since \( \tilde{\gamma}(\cdot) \) is minimising between \( \bar{x} \) and \( x \),
\[
V(x) = V(x_t) + \int_t^1 V(x_s) \, ds + \int_t^1 L_g(x_s, \dot{\gamma}(s)) \, ds.
\]
Hence, from (18)–(20), we deduce that, for every \( y \in W_t \),
\[
V(y) \geq V(x_t) + \Phi(y),
\]
where
\[
\Phi(y) := \int_t^1 (L_g(x_s, \dot{\gamma}(s)) - L_g(\gamma_s, \dot{\gamma}_s)) \, ds
- \int_t^1 (dV(x_s)(\gamma_s - x_s) + \sigma|\gamma_s - x_s|^2) \, ds.
\]
Since the mapping \( \Phi_t : W \to \mathbb{R} \) is smooth and since \( \Phi_t(x_t) = 0 \), a parabola can be put under the graph of \( V \) at \( x_t \). This proves (17). \( \Box \)
Lemma 2.9. The following inclusion holds:

\[ C_{\min}(\bar{x}) \subset \Sigma(V_{g,\Delta}). \]

Proof. Let \( x \in C_{\min}(\bar{x}) \); note that, by definition of \( C_{\min}(\bar{x}) \), one has \( x \neq \bar{x} \). We argue by contradiction. If \( x \) does not belong to \( \Sigma(V_{g,\Delta}) \), then \( V_{g,\Delta} \) is \( C^1 \) in a neighborhood of \( x \). This means that there exist a neighborhood \( V \) of \( x \) and \( t \in (0,1) \) such that for every \( y \in V \), there is a minimizing curve of \((\text{BP})_{g,\Delta}\) such that \( \bar{y}(t) = y \). From the previous lemma and by compactness of the minimizers, we deduce that the function \( -V_{g,\Delta} \) is locally semiconcave on \( V \). Hence by Proposition 6, \( V_{g,\Delta} \) is \( C^1_{\text{loc}} \) in \( V \). Define

\[ \Psi : V \to T^* \bar{x} M, \quad y \mapsto \psi(0), \]

where \( \psi(\cdot) : [0,1] \to T^* M \) is the normal extremal satisfying \( \psi(1) = (y, \frac{1}{2} dV_{g,\Delta}(y)) \). This mapping is locally Lipschitz on \( V \). Moreover, by construction, \( \Psi \) is the inverse of the exponential mapping. This proves that \( p_0 := \Psi(x) \) is not a conjugate point. We obtain a contradiction. \( \square \)

Lemma 2.10. Let \( p_0 \in T^* \bar{x} M \) be such that \( H(\bar{x}, p_0) \neq 0 \). There exist a neighborhood \( W \) of \( p_0 \) in \( T^* \bar{x} M \) and \( \epsilon > 0 \) such that every normal extremal with \( \psi(0) = (\bar{x}, p) \) (in local coordinates) belonging to \( W \) is minimizing on the interval \([0, \epsilon]\).

The proof of Lemma 2.10 is postponed to the Appendix (Section 3.2).

We are now ready to provide a proof for Lemma 1.1.

Proof of Lemma 1.1. For the sake of simplicity, we assume that \( M = \mathbb{R}^n \), endowed with the Euclidean metric. We have to prove that \( C_{\min}(\bar{x}) \subset L(\bar{x}) \). Let \( y \in C_{\min}(\bar{x}) \). We argue by contradiction. Suppose that \( y \) does not belong to \( L(\bar{x}) \). This means that there exists a minimizing curve \( \gamma(\cdot) \) of \((\text{BP})_{g,\Delta}\) and \( t_0 \in (0,1) \) such that \( \gamma(t_0) = y \). Set \( x := \gamma(1) \), and let \( \bar{t} \) be the minimum of times \( t \in (0,1) \) such that \( \gamma(t) \notin \Sigma(V_{g,\Delta}) \). We claim that \( \bar{t} \in (0,t_0) \). As a matter of fact, we know by Lemma 2.9 that \( \gamma(t_0) = y \in \Sigma(V_{g,\Delta}) \).

Moreover, from Lemma 2.10 and the absence of (nontrivial) singular minimizing paths, the mapping

\[ W \to M, \quad p \mapsto \pi(\psi(\epsilon)), \]

where \( \psi(0) = (\bar{x}, p) \), is injective. Hence from the Invariance of Domain Theorem\(^6\) this mapping is open. This means that \( V = V_{g,\Delta} \) is necessarily of class \( C^1 \) on a neighborhood of each \( \gamma(s) \) with \( s \in (0, \epsilon] \). We conclude that \( \bar{t} \in (0,t_0) \).

Set \( \bar{x} := \gamma(\bar{t}) \) and \( x_s := \gamma(s) \) for every \( s \in [0,1] \). By local semiconcavity of \( V \) (see Proposition 5), there exist a neighborhood \( V \) of \( \bar{x} \) in \( M \setminus \{\bar{x}\} \) and \( \sigma \in \mathbb{R} \) such that

\[ V(z') \leq V(z) + \langle dV(z), z' - z \rangle + \sigma |z' - z|^2, \quad \forall z, z' \in V. \quad (21) \]

\(^6\) The Invariance of Domain Theorem states that, for a topological manifold \( N \), if \( f : N \to N \) is continuous and injective, then it is open. We refer the reader to the book\(^{12}\) for a proof of that result.
Let $\tilde{p} \in T^*_xM$ be such that $\tilde{x} = \exp_{\tilde{p}}(\tilde{p})$. Since $V$ is of class $C^1$ in a neighborhood of the curve $s \in (0, \tilde{t}) \mapsto \gamma(s)$, there exists a neighborhood $\mathcal{W}'$ of $\tilde{p}$ in $T^*_xM$ such that

$$\forall p \in \mathcal{W}', \quad H(\tilde{x}, p) = H(\tilde{x}, \tilde{p}) \Rightarrow V(\exp_{\tilde{x}}(p)) = V(\tilde{x}).$$

Thus, by (21), for every $p \in \mathcal{W}'$ satisfying $H(\tilde{x}, p) = H(\tilde{x}, \tilde{p})$ we have

$$\langle dV(\tilde{x}, \exp_{\tilde{x}}(p) - \tilde{x}) \rangle \geq -\sigma |\exp_{\tilde{x}}(p) - \tilde{x}|^2. \quad (22)$$

Furthermore, from Lemma 2.5, there exist a neighborhood $\mathcal{V}$ of $\tilde{x}$ and $\sigma' > 0$ such that

$$V(z) \geq V(\tilde{x}) + \langle dV(\tilde{x}), z - \tilde{x} \rangle - \sigma'|z - \tilde{x}|^2, \quad \forall z \in \mathcal{V}. \quad (23)$$

Without loss of generality, assume that $\mathcal{V}' = \mathcal{V}$ and $\sigma' = \sigma$. For every $p \in \mathcal{W}'$, set $x(p) := \exp_{\tilde{x}}(p)$. By (22) and (23), we deduce that for every $p \in \mathcal{W}'$ satisfying $H(\tilde{x}, p) = H(\tilde{x}, \tilde{p})$ and for every $z \in \mathcal{V}$, we have

$$V(z) \geq V(\tilde{x}) + \langle dV(\tilde{x}), z - x(p) \rangle + \langle dV(\tilde{x}), x(p) - \tilde{x} \rangle - \sigma|z - \tilde{x}|^2$$

$$\geq V(\tilde{x}) + \langle dV(\tilde{x}), z - x(p) \rangle - \sigma|x(p) - \tilde{x}|^2 - \sigma|z - \tilde{x}|^2.$$

In conclusion, by (21), for every $p \in \mathcal{W}'$ satisfying $H(\tilde{x}, p) = H(\tilde{x}, \tilde{p})$ and every $z \in \mathcal{V}$, we have

$$\langle dV(\tilde{x}), z - x(p) \rangle - \sigma|x(p) - \tilde{x}|^2 - \sigma|z - \tilde{x}|^2 \leq \langle dV(x(p)), z - x(p) \rangle + \sigma|z - x(p)|^2.$$

Hence, for every $p \in \mathcal{W}'$ satisfying $H(\tilde{x}, p) = H(\tilde{x}, \tilde{p})$ and every $z \in \mathcal{V}$,

$$2\sigma|z - x(p)|^2 + \langle dV(x(p)) - dV(\tilde{x}), z - x(p) \rangle$$

$$+ 2\sigma|x(p) - \tilde{x}, z - x(p) \rangle + 2\sigma|x(p) - \tilde{x}|^2 \geq 0. \quad (24)$$

Now, since $\tilde{x} = \gamma(\tilde{t})$ belongs to $\Sigma(V)$, we know from Lemma 2.6 that the exponential mapping is singular at $\tilde{p}$. Define the mapping $\Phi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^n$ by

$$\forall (z, p) \in \mathbb{R}^n \times \mathbb{R}^n, \quad \Phi(x, p) := \psi(x(1), p(1)),$$

where $(x(\cdot), p(\cdot)) : [0, 1] \rightarrow T_x^*M$ is the normal extremal satisfying $(x(0), p(0)) = (x, p)$. Since $\Phi$ is a flow, its differential is always invertible. Hence there exist $P \in \mathbb{R}^n$ and $Q \in \mathbb{R}^n \setminus \{0\}$ such that

$$D\Phi(\tilde{x}, \tilde{p}) \cdot (0, P) = (0, Q).$$

This means that there exist two continuous functions $\epsilon_1, \epsilon_2 : \mathbb{R} \rightarrow \mathbb{R}^n$ and a mapping $\lambda \mapsto p(\lambda) \in \mathcal{W}'$ such that, for every $\lambda$ sufficiently small, the following properties are satisfied:

(i) $H(\tilde{x}, p(\lambda)) = H(\tilde{x}, \tilde{p})$;
(ii) $V(x_\lambda) = V(\tilde{x})$ where $x_{\lambda} := x(p(\lambda))$;
(iii) $x_\lambda = \tilde{x} + \lambda^2 \epsilon_1(\lambda)$;
(iv) $dV(x_\lambda) = dV(\tilde{x}) + \lambda Q + \lambda^2 \epsilon_2(\lambda)$. 


From (24), we deduce that, for every $z \in V$,

$$2\sigma |z - x_\lambda|^2 + \lambda \langle Q, z - x_\lambda \rangle + \lambda^2 (e_2(\lambda), z - x_\lambda) + 2\sigma \lambda^2 (e_1(\lambda), z - x_\lambda) + 2\sigma \lambda^4 |e_1(\lambda)|^2 \geq 0.$$ 

We can apply this inequality for every $\alpha$ sufficiently small with $z = x_\lambda - \alpha Q$. This yields

$$2\sigma \alpha^2 |Q|^2 - \lambda \alpha |Q|^2 + \lambda^2 \alpha (e_2(\lambda), -Q) + 2\sigma \lambda^2 \alpha (e_1(\lambda), -Q) + 2\sigma \lambda^4 |e_1(\lambda)|^2 \geq 0$$

for all $\lambda, \alpha$ sufficiently small. Taking $\alpha := \lambda \sqrt{\lambda}$, we find a contradiction. □

**Lemma 2.11.** We have $\Sigma(V_{g,1}) = L(\bar{x}) \cup \{\bar{x}\}$.

In particular, the cut locus is closed in $M \setminus \{\bar{x}\}$, and is of Hausdorff dimension smaller than or equal to $n - 1$.

**Proof.** From Lemma 2.3, any point of $\Sigma(V_{g,1})$ is joined to $\bar{x}$ by several minimizing curves. Hence, from Lemma 2.4, any such point belongs to the cut locus $L(\bar{x})$. From Lemmas 2.6 and 1.1, we deduce that $\Sigma(V_{g,1}) \subset L(\bar{x}) \cup \{\bar{x}\}$.

If $x \in M \setminus \{\bar{x}\}$ does not belong to $\Sigma(V_{g,1})$, then, from Lemma 2.7, the function $V_{g,1}$ is of class $C^1$ in a neighborhood $U$ of $x$. Then the continuous mapping

$$F : U \to T^*M, \quad x \mapsto F(x) = -\tilde{H}(x, \frac{1}{2}dV_{g,1}(x)),$$

is such that $F(x) = (\tilde{x}, p_0)$ with $\exp_{\tilde{x}}(p_0) = x$. This means that the exponential mapping $\exp_{\tilde{x}}$ is a homeomorphism from $F(U)$ into $U$, with inverse mapping $F$. In particular, it follows that $x \notin L(\bar{x})$. The fact that $\bar{x} \in \Sigma(V_{g,1})$ results from [1, Theorem 1]. □

**Remark 2.4.** Lemma 2.11 asserts that the cut locus $L(\bar{x})$ has Hausdorff dimension $n - 1$.

Recently, proving a Lipschitz regularity property of the distance function to the cut locus, Li and Nirenberg showed in [30] that the $(n - 1)$-dimensional Hausdorff measure of the cut locus in the Riemannian framework is finite. It would be interesting to study the regularity of the distance function to the cut locus to obtain such a result in the sub-Riemannian case.

**Lemma 2.12.** The function $V_{g,1}$ is of class $C^\infty$ on the open set $M \setminus \Sigma(V_{g,1})$. Moreover, if $\gamma : [0, 1] \to M$ is a minimizing curve for $(BP)_{g,1}$, then $\gamma(t) \notin \Sigma(V_{g,1})$ for every $t \in (0, 1)$.

**Proof.** Let $\gamma : [0, 1] \to M$ be a minimizing curve for $(BP)_{g,1}$. It follows from Lemmas 1.1 and 2.11 that $\gamma(t) \notin \Sigma(V_{g,1})$ for every $t \in (0, 1)$.

Let $x \in M \setminus \Sigma(V_{g,1})$, and let $\gamma(\cdot)$ be a minimizing horizontal path joining $\bar{x}$ to $x$. By assumption, $\gamma(\cdot)$ is necessarily nonsingular, and admits a unique normal extremal lift $\psi(\cdot) : [0, 1] \to T^*M$. From Lemmas 1.1 and 2.11, the point $x$ is not conjugate to $\bar{x}$, and hence the exponential mapping $\exp_{\bar{x}}$ is a (smooth) local diffeomorphism from
a neighborhood of \( p_0 \) into a neighborhood of \( x \), where \( \psi(0) = (\bar{x}, p_0) \) in local coordinates. As recalled in the first section, the length of the path \( \gamma(\cdot) = \pi(\psi(\cdot)) \) is equal to \((2H(\psi(0)))^{1/2}\). Since \( \gamma(\cdot) \) is minimizing, it is also equal to \( d_{SR}(\bar{x}, x) \). Then, using local coordinates, and from (13),

\[
V_{x, \Delta}(x) = 2H(\bar{x}, (\exp_p)^{-1}(x))
\]

in a neighborhood of \( x \) (see also [43] Corollary 1, p. 157). It follows that \( V_{x, \Delta} \) is of class \( C^\infty \) at the point \( x \).

\[\Box\]

2.2.4. Conclusion: proof of Theorem 1

Define \( S := \Sigma(V_{x, \Delta}) \). From Lemma 2.11 we have \( S = \mathcal{L}(\bar{x}) \cup [\bar{x}] \). We next define a section \( X \) of \( \Delta \) that is smooth outside \( S \). To this end, it is convenient to consider local coordinates, and to express the problem in terms of optimal control. Let \( x \in M \setminus S \). In a neighborhood \( U \) of \( x \), one has, in local coordinates,

\[\Delta = \text{span}\{f_1, \ldots, f_m\},\]

where \( (f_1, \ldots, f_m) \) is an \( m \)-tuple of smooth vector fields which is orthonormal for the metric \( g \). We proceed as in [3].

Fix \( x \in M \setminus \bar{x} \) (of course, we set \( X(x) := 0 \) if \( x = \bar{x} \)), and pick some \( \xi \in \partial L V_{x, \Delta}(x) \). Note that, since \( V_{x, \Delta} \) is smooth outside the set \( S \), one has \( \xi = dV_{x, \Delta}(x) \) whenever \( x \in M \setminus S \). Define the control \( \tilde{u}_i(x) = (\tilde{u}_1(x), \ldots, \tilde{u}_m(x)) \) by

\[\tilde{u}_i(x) := \frac{1}{2} \xi(f_i(x)), \quad \forall i = 1, \ldots, m. \tag{25}\]

For \( x \in M \setminus \bar{S}, \tilde{u}_i(x) = \frac{1}{2}(dV_{x, \Delta}(x), f_i(x)) \) is the closed-loop form of the optimal control [14]. For \( x \in S \), the expression of \( \tilde{u}_i(x) \) depends on the choice of \( \xi \in \partial L V_{x, \Delta}(x) \). Define

\[X(x) := -\sum_{i=1}^m \tilde{u}_i(x) f_i(x). \tag{26}\]

Geometrically, \( X(x) \) coincides with the projection of \(-\frac{1}{2} \xi \) onto \( \Delta(x) \). At the point \( \bar{x} \), we set \( X(\bar{x}) = 0 \). This defines a vector field \( X \) on \( M \) which is smooth on \( M \setminus \bar{S} \), but may be totally discontinuous on \( \bar{S} \).

We next prove that \( X \) is SRS\( S_x \). Property (i) is obviously satisfied, but properties (ii) and (iii) are not so direct to derive.

We first prove that every minimizing trajectory yields a Carathéodory solution of \( \bar{x} = X(x) \). Let \( x \in M \setminus \bar{x} \) be fixed and \( \gamma(\cdot) : [0, 1] \to M \) be a minimizing curve of the Bolza problem (BP)\( _{x, \Delta} \) between \( \bar{x} \) and \( x \). It follows from the Pontryagin maximum principle that \( \gamma \) is the projection of a normal extremal expressed in local coordinates by \( \psi(\cdot) = (\gamma(\cdot), p(\cdot)) \). Let \( t \in (0, 1) \); from Lemma 2.4, the curve \( \gamma'(\cdot) : [0, 1] \to M \) defined by \( \gamma'(s) := \gamma(s) \) for \( s \in [0, 1] \) is the unique minimizing curve of (BP)\( _{\bar{x}, \Delta} \) steering \( \bar{x} \) to \( \gamma(t) \). Moreover, from Lemma 2.4 it is the projection of the normal extremal \( \bar{\psi}'(\cdot) = (\bar{\psi}'(\cdot), \bar{p}'(\cdot)) \), where \( \bar{p}'(\cdot) \) is defined by \( \bar{p}'(s) = tp(st) \) for every \( s \in [0, 1] \). It then follows from Lemmas 2.5 and 2.12 that, along the curve \( \gamma(\cdot) \),

\[dV_{x, \Delta}(\gamma(t)) = 2tp(t), \quad \forall t \in (0, 1).\]
Therefore, \( \gamma(\cdot) \) is a solution of
\[
\dot{\gamma}(t) = \frac{1}{2t} \sum_{i=1}^{m} (dV_{g,\Delta}(\gamma(t))(f_i(\gamma(t)))) f_i(\gamma(t)), \text{ a.e. on } (0, 1),
\]
in local coordinates along \( \gamma(\cdot) \). This implies that the curve \( x(\cdot): [0, \infty) \to M \) defined by
\[
x(t) := \gamma(e^{-t}), \quad \forall t \in (0, \infty),
\]
is a Carathéodory solution of \( \dot{x} = X(x) \) such that \( x(0) = \gamma(1) = x \).

We next prove that any Carathéodory solution of \( \dot{x} = X(x) \), \( x(0) = x \), tends to \( \bar{x} \) as \( t \) tends to \( +\infty \). Having in mind the minimizing properties (by construction) of the vector field \( X \), it suffices actually to prove the following lemma.

**Lemma 2.13.** Let \( x(\cdot) \) be any Carathéodory solution of \( \dot{x} = X(x) \). Then there does not exist a nontrivial interval \([a, b]\) such that \( x(t) \in \mathcal{S} \) for every \( t \in [a, b] \).

**Proof.** The proof goes by contradiction. Suppose that there exist \( \epsilon > 0 \) and a curve \( x(\cdot): [0, \epsilon] \to M \) such that
\[
\dot{x}(t) = X(x(t)) \quad \text{for almost every } t \in [0, \epsilon],
\]
and
\[
x(t) \in \mathcal{S}, \quad \forall t \in [0, \epsilon].
\]
In local coordinates in a neighborhood of \( x(0) = x \), one has
\[
\dot{x}(t) = X(x(t)) = -\frac{1}{2} \sum_{i=1}^{m} \xi_t(f_i(x(t))) f_i(x(t)) \quad \text{for almost every } t \in [0, \epsilon],
\]
where \( \xi_t \in \partial_x V_{g,\Delta}(x(t)) \) for almost every \( t \in [0, \epsilon] \). At this stage, we need to use Lemma 2.2 whose proof is provided in the Appendix (Section 3.1). According to this lemma, using (25) and the Hamilton-Jacobi equation (16) satisfied by \( V_{g,\Delta} \), we deduce that, for almost every \( t \in [0, \epsilon] \),
\[
\frac{d}{dt}(V_{g,\Delta}(x(t))) = \xi_t(\dot{x}(t)) = -\frac{1}{2} \sum_{i=1}^{m} (\xi_t(f_i(x(t))))^2
\]
\[
= -H_{g,\Delta}(x(t), \xi_t) = -2V_{g,\Delta}(x(t)), \quad (27)
\]
since the Hamiltonian function \( H_{g,\Delta}(x, p) \) is quadratic in \( p \). Therefore,
\[
V_{g,\Delta}(x(t)) = V_{g,\Delta}(x)e^{-2t}, \quad \forall t \in [0, \epsilon].
\]
Let \( \gamma(\cdot): [0, 1] \to M \) be a minimizing curve of the Bolza problem (BP)\(_{g,\Delta} \) between \( \bar{x} \) and \( x(\epsilon) \). Define the horizontal path \( \tilde{\gamma}(\cdot): [0, 1] \to M \) by
\[
\tilde{\gamma}(t) = \begin{cases} x(-\ln t) & \text{if } e^{-\epsilon} \leq t \leq 1, \\ \gamma(e^{t}) & \text{if } 0 \leq t \leq e^{-\epsilon}. \end{cases}
\]
The cost of \( \tilde{\gamma}(\cdot) \) is

\[
J_g(\tilde{\gamma}(\cdot)) = \int_0^{\epsilon^{-1}} L_g(\tilde{\gamma}(t), \dot{\tilde{\gamma}}(t)) \, dt + \int_{\epsilon^{-1}}^1 L_g(\tilde{\gamma}(t), \dot{\tilde{\gamma}}(t)) \, dt
\]

\[
= \int_0^{\epsilon^{-1}} L_g(\gamma(\epsilon t), \dot{\gamma}(\epsilon t)) \, dt + \int_{\epsilon^{-1}}^1 L_g(\tilde{\gamma}(t), \dot{\tilde{\gamma}}(t)) \, dt
\]

\[
= e^\epsilon V_{g,\Delta}(x(e) \times x(\frac{1}{t})) + \int_{\epsilon^{-1}}^1 \frac{1}{t} \sum_{i=1}^m \tilde{a}_i(x(-\ln t))^2 \, dt
\]

\[
= e^\epsilon V_{g,\Delta}(x(e)) + \int_0^1 e^\epsilon \sum_{i=1}^m \tilde{a}_i(x(s))^2 \, ds.
\]

Using (25), (27), and (28), one has, for almost every \( s \in [0, \epsilon] \),

\[
\sum_{i=1}^m \tilde{a}_i(x(s))^2 = \sum_{i=1}^m \frac{1}{4} (\zeta_i(f_i(x(s))))^2 = V_{g,\Delta}(x(s)) = V_{g,\Delta}(x) e^{-2s},
\]

and, since \( V_{g,\Delta}(x(e)) = V_{g,\Delta}(x) e^{-2\epsilon} \), it follows that

\[
J_g(\tilde{\gamma}(\cdot)) = V_{g,\Delta}(x).
\]

Hence, \( \tilde{\gamma} \) is a minimizing curve of the Bolza problem \( (BP)_{g,\Delta} \) between \( \bar{x} \) and \( x \). From Lemma 2.12, it cannot stay on \( S \) for positive times. This yields a contradiction. \( \Box \)

It follows from this lemma, and from the construction of \( X \) using optimal controls, that any Carathéodory trajectory of \( \dot{x} = X(x), x(0) = x \), tends to \( \bar{x} \) as \( t \) tends to \( +\infty \). The property of Lyapunov stability is obvious to verify. Finally, the fact that the set \( S \) has Hausdorff dimension \( \leq n - 1 \) is a consequence of Lemma 2.6.

2.3. Proof of Theorem 2

Let \( g \) be a Riemannian metric on \( M \), and \( \bar{x} \) be fixed. Since \( \Delta \) is a smooth distribution of rank two on \( M \), for every \( x \in M \) there exists a neighborhood \( \mathcal{V}_x \) of \( x \) and two smooth vector fields \( f_1^x, f_2^x \) which represent \( \Delta \) in \( \mathcal{V}_x \), that is,

\[
\Delta(y) = \text{span}\{f_1^x(y), f_2^x(y)\}, \quad \forall y \in \mathcal{V}_x.
\]

Moreover, as recalled in the introduction, since \( \Delta \) is a Martinet distribution, for every \( x \in \Sigma_\Delta \), the two vector fields \( f_1^x, f_2^x \) can be chosen as

\[
f_1^x = \frac{\partial}{\partial x_1} + x_2^2 \frac{\partial}{\partial x_3} \quad \text{and} \quad f_2^x = \frac{\partial}{\partial x_2}
\]

in local coordinates. Recall that, in the neighborhood \( \mathcal{V}_x \), the Martinet surface \( \Sigma_\Delta \) coincides with the surface \( x_2 = 0 \), and the singular paths are the integral curves of the vector
field \( \partial / \partial x_1 \) restricted to \( x_2 = 0 \). For convenience, assume that the vector fields \( f_1^x, f_2^x \) are defined as well outside the neighborhood \( \mathcal{V}_x \). Thus, without loss of generality, for every \( x \in M \), we assume that the vector fields \( f_1^x, f_2^x \) are well defined, smooth on \( M \) and satisfy

\[
f_j^x(y) = f_j^x(y) = 0, \quad \forall y \in M \setminus \mathcal{W}_x,
\]

with \( \mathcal{V}_x \subset \mathcal{W}_x \), and

\[
\text{span}\{f_1^x(y), f_2^x(y)\} \subset \Delta(y), \quad \forall y \in M.
\]

By compactness of \( \Sigma_\Delta \), there is a finite collection \((x_i)_{i \in I}\) of points of \( \Sigma_\Delta \) such that

\[
\Sigma_\Delta \subset \bigcup_{i \in I} \mathcal{V}_{x_i}.
\]

Let \( \beta : M \to [0, \infty) \) be a smooth function such that

\[
\forall x \in M, \quad \beta(x) = 0 \Rightarrow x \in \Sigma_\Delta.
\]

For every \( i \in I \), define the smooth vector field \( g_i \), in local coordinates, by

\[
g_i(y) := \beta(y)f_1^{x_i}(y), \quad \forall y \in M.
\]

By compactness of \( M \), there is a finite collection \((y_j)_{j \in J}\) of points of \( M \) such that

\[
M \subset \bigcup_{i \in I} \mathcal{V}_{x_i} \cup \bigcup_{j \in J} \mathcal{V}_{y_j}.
\]

and

\[
\Sigma_\Delta \cap \bigcup_{j \in J} \mathcal{W}_{y_j} = \emptyset.
\]

By construction, we have

\[
\text{span}\{g_i(y), f_2^x(y), f_1^y(y), f_2^y(y) \mid i \in I, \ j \in J\} = \Delta(y), \quad \forall y \in M \setminus \Sigma_\Delta, \tag{30}
\]

and

\[
\text{span}\{g_i(y), f_2^x(y), f_1^y(y), f_2^y(y) \mid i \in I, \ j \in J\} \cap T_y \Sigma_\Delta = \{0\}, \quad \forall y \in \Sigma_\Delta. \tag{31}
\]

Indeed, for every \( y \in \Sigma_\Delta \),

\[
\text{span}\{g_i(y), f_2^x(y), f_1^y(y), f_2^y(y) \mid i \in I, \ j \in J\} = \text{span}\{f_2^x(y) \mid i \in I\}. \tag{32}
\]

It follows from (30) and (31) that any trajectory solving the control system

\[
\dot{x}(t) = \sum_{i \in I} (u_1^i(t)g_i(x(t)) + u_2^i(t)f_2^{x_i}(x(t)))
\]

\[
+ \sum_{j \in J} (v_1^j(t)f_1^{y_j}(x(t)) + v_2^j(t)f_2^{y_j}(x(t))), \tag{33}
\]

\[
\dot{y}(t) = \sum_{i \in I} (u_1^i(t)f_1^{y_i}(x(t)) + u_2^i(t)f_2^{y_i}(x(t)))
\]

\[
+ \sum_{j \in J} (v_1^j(t)f_1^{y_j}(x(t)) + v_2^j(t)f_2^{y_j}(x(t))),
\]

\[
\dot{z}(t) = \sum_{i \in I} (u_1^i(t)f_1^{z_i}(x(t)) + u_2^i(t)f_2^{z_i}(x(t)))
\]

\[
+ \sum_{j \in J} (v_1^j(t)f_1^{z_j}(x(t)) + v_2^j(t)f_2^{z_j}(x(t))),
\]

\[
\dot{w}(t) = \sum_{i \in I} (u_1^i(t)f_1^{w_i}(x(t)) + u_2^i(t)f_2^{w_i}(x(t)))
\]

\[
+ \sum_{j \in J} (v_1^j(t)f_1^{w_j}(x(t)) + v_2^j(t)f_2^{w_j}(x(t))].
\]
where $u(\cdot) = (u^1_1(\cdot), u^2_1(\cdot), \ldots, u^1_{|J|}(\cdot), u^2_{|J|}(\cdot), v^1_1(\cdot), v^2_1(\cdot), \ldots, v^1_{|J|}(\cdot), v^2_{|J|}(\cdot))$ belongs to the control set $\mathcal{U}$ defined by

$$\mathcal{U} := L^\infty([0, 1], \mathbb{R}^{2|J|})$$

is a horizontal path of $\Delta$. Note that, for every $u(\cdot) \in \mathcal{U}$, there exists a unique absolutely continuous curve $\gamma_{u(\cdot)} : [0, 1] \to M$ such that $\gamma_{u(\cdot)}(0) = \bar{x}$ and

$$\dot{\gamma}_{u(\cdot)}(t) = \sum_{i \in I} (u^i_1(t)g_i(\gamma_{u(\cdot)}(t)) + u^i_2(t)f_i^2(\gamma_{u(\cdot)}(t))) + \sum_{j \in J} (v^j_1(t)f_1^j(\gamma_{u(\cdot)}(t)) + v^j_2(t)f_2^j(\gamma_{u(\cdot)}(t)))$$

for almost every $t \in [0, 1]$. Moreover, it is clear by construction of the control system under consideration that, for every $x \in M$, there exists a control $u(\cdot) \in \mathcal{U}$ such that $\gamma_{u(\cdot)}(1) = x$. For every $u(\cdot) \in \mathcal{U}$, set

$$J(u(\cdot)) := \int_0^1 \left( \sum_{i \in I} (u^i_1(t)^2 + u^i_2(t)^2) + \sum_{j \in J} (v^j_1(t)^2 + v^j_2(t)^2) \right) dt.$$

Define the value function $W : M \to \mathbb{R}$ by

$$W(x) := \inf \{ J(u(\cdot)) \mid u(\cdot) \in \mathcal{U}, \gamma_{u(\cdot)}(0) = \bar{x}, \gamma_{u(\cdot)}(1) = x \}$$

for every $x \in M$. By coercivity of the cost function, it is easy to prove that, for every $x \in M \setminus \{\bar{x}\}$, there exists a control $u(\cdot) \in \mathcal{U}$ such that $\gamma_{u(\cdot)}(1) = x$ and $W(x) = J(u(\cdot))$ (i.e., a minimizing control). Moreover, by construction of the control system, more precisely, from (32), the trajectory $\gamma_{u(\cdot)}(\cdot)$ cannot stay on the Martinet surface on a nontrivial subinterval of $[0, 1]$. As a consequence, since any singular trajectory is contained in the Martinet surface, any nontrivial minimizing control is nonsingular. Using similar arguments to those in the proof of Theorem 1 it follows that the value function $W$ is a viscosity solution of a certain Hamilton–Jacobi equation, is continuous on $M$, and is locally semiconcave in $M \setminus \{\bar{x}\}$ (see [14]). Moreover, the optimal trajectories of the optimal control problem under consideration share the same properties as those of the Bolza problem (BP)$_{g, \Delta}$. The construction of a stabilizing feedback then follows the same lines as in Theorem 1.

**Remark 2.5.** For a noncompact manifold $M$, the above proof needs to be adapted by replacing a finite set of controls $(u_i)_{i \in I}$ and $(v_j)_{j \in J}$ with a locally finite set of controls.

### 3. Appendix

#### 3.1. Proof of Lemma 2.2

Without loss of generality, we assume that $M = \mathbb{R}^n$. Given $k \in \{1, \ldots, n\}$ and $\rho > 0$, denote by $\Sigma_k^\rho(u)$ the set of all $x \in U$ such that $D^+ u(x)$ contains a $k$-dimensional sphere of radius $\rho$, and define

$$\Sigma_k^\rho(u) := \{ x \in U \mid \dim(D^+ u(x)) = k \}.$$
By well known properties of convex sets, one has \( \Sigma_k^{\rho}(u) \subset \bigcup_{\rho > 0} \Sigma_k^{\rho}(u) \). Note that a point \( x \in \Sigma_k^{\rho}(u) \) does not necessarily belong to \( \Sigma_k(u) \), since \( D^+u(x) \) may be of dimension greater than \( k \). The following result is fundamental for the proof of Lemma 2.2 (we refer the reader to [15] for its proof).

**Lemma 3.1.** For every \( k \in \{1, \ldots, n\} \) and every \( \rho > 0 \), the set \( \Sigma_k^{\rho}(u) \) is closed and satisfies

\[
\text{Tan}(x, \Sigma_k^{\rho}(u)) \subset \{D^+u(x)\}^\perp, \quad \forall x \in \Sigma_k^{\rho}(u) \cap \Sigma_k(u) \]  

We return to the proof of Lemma 2.2. First, note that the map \( t \in [a, b] \mapsto u(\gamma(t)) \) is Lipschitzian. Hence, by Rademacher’s Theorem, it is differentiable almost everywhere on \([a, b]\). Moreover, by the chain rule for Clarke’s generalized gradients (see [22]), for every \( t \in [a, b] \) where \( \gamma \) is differentiable, there exists \( p \in \partial u(\gamma(t)) \) such that

\[
\frac{d}{dt}u(\gamma(t)) = \langle p, \dot{\gamma}(t) \rangle. \tag{34}
\]

For every \( k \in \{1, \ldots, n\} \) and any positive integer \( l \), set

\[
I_{k,l} := \{ t \in [a, b] \mid \gamma(t) \in (\Sigma^{\rho}_{k+1}(u) \cap \Sigma^k(u)) \setminus \Sigma^{\rho}_{k+1}(u) \}
\]

and

\[
J := [a, b] \setminus \bigcup_{k,l} I_{k,l}.
\]

Notice that, since \( u \) is locally semiconcave and \( \gamma \) is locally Lipschitzian, \( u \) is differentiable at almost every \( \gamma(t) \) with \( t \in J \). Thus, for every such \( t \), we have necessarily \( p(t) = \nabla u(\gamma(t)) \) and

\[
\frac{d}{dt}u(\gamma(t)) = \langle p(t), \dot{\gamma}(t) \rangle.
\]

It remains to prove that this equality holds for almost every \( t \) in \([a, b] \setminus J\). From the Lebesgue density theorem, there exists a sequence \( \{I'_{k,l}\} \) of measurable sets such that all sets \( I_{k,l} \setminus I'_{k,l} \) have Lebesgue measure zero and each point in any \( I'_{k,l} \) is a density point in that set. It is sufficient to prove the required equality on each \( I'_{k,l} \). Fix \( k, l \) and \( t \in I'_{k,l} \), and set \( x := \gamma(t) \). Since \( x \) is a density point in \( I'_{k,l} \), there exists a sequence \( \{t_i\} \) in \( I'_{k,l} \) converging to \( t \). Thus, the vector \( \dot{\gamma}(t) \) belongs to \( \text{Tan}(x, \Sigma_k^{\rho}(u)) \). Then, from Lemma 3.1, \( \dot{\gamma}(t) \) belongs to \( \{D^+u(x)\}^\perp \). By (34), we obtain the desired equality. This concludes the proof of Lemma 2.2.

---

\[7\] Here, \( \text{Tan}(x, \Sigma_k^{\rho}(u)) \) denotes the tangent set to \( \Sigma_k^{\rho}(u) \) at \( x \). Recall that, given a closed set \( S \subset \mathbb{R}^n \) and \( x \in S \), the tangent set to \( S \) at \( x \), denoted by \( \text{Tan}(x, S) \), is defined as the vector space spanned by the set

\[
T(x, S) := \lim_{t \to 0} \{ \frac{x_t - x}{t} \mid x_t \in S, x_t \to x, t_t \in \mathbb{R}^+, t_t \downarrow 0 \}.
\]

Recall also that if \( A \subset \mathbb{R}^n \), then \( A^\perp \) is the set of vectors \( v \in \mathbb{R}^n \) such that \( \langle v, p \rangle = \langle v, p' \rangle \) for any \( p, p' \in A \).
3.2. Proof of Lemma 2.10

The proof that we present here is taken from [42] (cf. [32, 34]). For simplicity, assume that \( M = \mathbb{R}^n \), endowed with the Euclidean metric. Since the property to be proved is local, we assume that there are \( m \) smooth vector fields \( f_1, \ldots, f_m \), orthonormal with respect to the Euclidean metric, such that

\[
\Delta(x) = \text{span}\{f_i(x) \mid i = 1, \ldots, m\}
\]

in a neighborhood \( V \) of \( \bar{x} \). With these notations, the associated Hamiltonian \( H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \) is

\[
H(x, p) := \max_{u \in \mathbb{R}^m} \left\{ \langle p, \sum_{i=1}^m u_i f_i(x) \rangle - \frac{1}{2} \sum_{i=1}^m u_i^2 \right\} = \frac{1}{2} \sum_{i=1}^m \langle p, f_i(x) \rangle^2
\]

for every \( (x, p) \in \mathbb{R}^n \times \mathbb{R}^n \).

Our aim is now to prove the following result: for every \( p_0 \in \mathbb{R}^n \) such that \( H(\bar{x}, p_0) \neq 0 \), there exist a neighborhood \( W \) of \( p_0 \) in \( \mathbb{R}^n \) and \( \epsilon > 0 \) such that every solution \( (x(\cdot), p(\cdot)) : [0, \epsilon] \to \mathbb{R}^n \times \mathbb{R}^n \) of the Hamiltonian system

\[
\begin{align*}
\dot{x}(t) &= \frac{\partial H}{\partial p}(x(t), p(t)) = \sum_{i=1}^m \langle p(t), f_i(x(t)) \rangle f_i(x(t)), \\
\dot{p}(t) &= -\frac{\partial H}{\partial x}(x(t), p(t)) = -\sum_{i=1}^m \langle p(t), f_i(x(t)) \rangle df_i(x(t))^* p(t),
\end{align*}
\]  

(35)

with \( x(0) = \bar{x} \) and \( p(0) \in W \), satisfies

\[
\int_0^\epsilon \sum_{i=1}^m \langle p(t), f_i(x(t)) \rangle^2 dt \leq \int_0^\epsilon \sum_{i=1}^m u_i(t)^2 dt
\]  

(36)

for every control \( u(\cdot) \in L^\infty([0, \epsilon], \mathbb{R}^m) \) such that the solution of

\[
\dot{y}(t) = \sum_{i=1}^m u_i(t) f_i(y(t)), \quad y(0) = \bar{x},
\]  

(37)

satisfies \( y(\epsilon) = x(\epsilon) \). Fix \( p_0 \in \mathbb{R}^n \setminus \{0\} \). We need the following lemma.

**Lemma 3.2.** There exist a neighborhood \( W \) of \( p_0 \) and \( \rho > 0 \) such that, for every \( p \in W \), there exists a function \( S : B(\bar{x}, \rho) \to \mathbb{R} \) of class \( C^1 \) which satisfies

\[
H(x, \nabla S(x)) = H(\bar{x}, p), \quad \forall x \in B(\bar{x}, \rho),
\]  

(38)

and such that if \( (x^p(\cdot), p^p(\cdot)) \) denotes the solution of (35) satisfying \( x^p(0) = \bar{x} \) and \( p^p(0) = p \), then

\[
\nabla S(x^p(t)) = p^p(t), \quad \forall t \in (-\rho, \rho).
\]  

(39)
Proof. We apply the method of characteristics. Let \( \Pi \) be the linear hyperplane such that \( \langle p_0, v \rangle = 0 \) for every \( v \in \Pi \). We first show how to construct \( \mathcal{S} \) locally as the solution of the Hamilton–Jacobi equation \( [38] \) which vanishes on \( \ddot{x} + \Pi \) and satisfies \( \nabla \mathcal{S}(\ddot{x}) = p_0 \).

Up to diminishing \( \mathcal{V} \), we assume that \( H(x, p_0) \neq 0 \) for every \( x \in \mathcal{V} \). For every \( x \in (\ddot{x} + \Pi) \cup \mathcal{V} \), set

\[
\tilde{\rho}(x) := \sqrt{\frac{H(\ddot{x}, p_0)}{H(x, p_0)}} p_0.
\]

Then \( H(x, \tilde{\rho}(x)) = H(\ddot{x}, p_0) \) and \( \tilde{\rho}(x) \perp \Pi \), for every \( x \in \mathcal{V} \). There exists \( \mu > 0 \) such that, for every \( x \in (\ddot{x} + \Pi) \cup \mathcal{V} \), the solution \( (x_i(x), p_{x_i}(x)) \) of \( [35] \) satisfying \( x_i(0) = x \) and \( p_{x_i}(0) = \ddot{x} \) is defined on the interval \( (-\mu, \mu) \). For every \( x \in (\ddot{x} + \Pi) \cup \mathcal{V} \) and every \( t \in (-\mu, \mu) \), set \( \theta(t, x) := x_i(t) \). The mapping \( (t, x) \mapsto \theta(t, x) \) is smooth. Moreover, \( \theta(0, x) = x \) for every \( x \in (\ddot{x} + \Pi) \cup \mathcal{V} \), and \( \theta(0, \ddot{x}) = \sum_{i=1}^{m} \langle \bar{\rho}(x), f_i(\ddot{x}) \rangle f_i(\ddot{x}) \) does not belong to \( \Pi \). Hence there exists \( \rho \in (0, \mu) \) with \( B(\ddot{x}, \rho) \subset \mathcal{V} \) such that \( \theta \) is smooth.

The mapping \( (t, x) \mapsto \theta(t, x) \) is smooth. Moreover, \( \theta(0, x) = x \) for every \( x \in (\ddot{x} + \Pi) \cup \mathcal{V} \), and \( \theta(0, \ddot{x}) = \sum_{i=1}^{m} \langle \bar{\rho}(x), f_i(\ddot{x}) \rangle f_i(\ddot{x}) \) does not belong to \( \Pi \). Hence there exists \( \rho \in (0, \mu) \) with \( B(\ddot{x}, \rho) \subset \mathcal{V} \) such that \( \theta \) is smooth.

Denote by \( \varphi = (\tau, \pi) \) the inverse function of \( \theta \), that is, the function such that \( (\theta \circ \varphi)(x) = (\tau(x), \pi(x)) = x \) for every \( x \in \mathcal{V} \). Define the two vector fields \( X \) and \( P \) by

\[
X(x) := \dot{\theta}(\tau(x), \pi(x)) \quad \text{and} \quad P(x) := p_{\pi(x)}(\tau(x)), \quad \forall x \in \mathcal{V}.
\]

Then

\[
X(\theta(t, x)) = \dot{\theta}(t, x) = \ddot{x}(t) = \sum_{i=1}^{m} \langle p_{x_i}(t), f_i(x_i(t)) \rangle f_i(x_i(t)) = \sum_{i=1}^{m} \langle P(\theta(t, x)), f_i(\theta(t, x)) \rangle f_i(\theta(t, x))
\]

and

\[
\sum_{i=1}^{m} (P(\theta(t, x)), f_i(x_i(t)))^2 = \sum_{i=1}^{m} (p_{x_i}(t), f_i(x_i(t)))^2 = 2H(x, \bar{\rho}(x)) = 2H(\ddot{x}, p_0)
\]

for every \( t \in (-\rho, \rho) \) and every \( x \in (\ddot{x} + \Pi) \cup B(\ddot{x}, \rho) \). For every \( x \in \mathcal{V} \), set \( \alpha_i(x) := \langle P(x), f_i(x) \rangle \). Hence,

\[
X(x) = \sum_{i=1}^{m} \alpha_i(x) f_i(x) \quad \text{and} \quad \sum_{i=1}^{m} \alpha_i(x)^2 = H(\ddot{x}, p_0)
\]

for every \( x \in \mathcal{V} \). Define the function \( \mathcal{S} : \mathcal{V} \to \mathbb{R} \) by

\[
\mathcal{S}(x) := 2H(\ddot{x}, p_0) \tau(x), \quad \forall x \in \mathcal{V}.
\]

We next prove that \( \nabla \mathcal{S}(x) = P(x) \) for every \( x \in \mathcal{V} \). For every \( t \in (-\rho, \rho) \), define \( W_t := \{ y \in \mathcal{V} \mid \tau(y) = t \} \). In fact, \( W_t \) coincides with the set of \( y \in \mathcal{V} \) such that \( \mathcal{S}(y) = 2H(\ddot{x}, p_0)t \). It is a smooth hypersurface which satisfies \( \nabla \mathcal{S}(y) \perp T_y W_t \) for every \( y \in W_t \). Fix \( y \in W_t \). There exists \( x \in (\ddot{x} + \Pi) \cup B(\ddot{x}, \rho) \) such that \( y = \theta(t, x) = x_i(t) \).
Let us first prove that $P(y) = p_x(t)$ is orthogonal to $T_yW_t$. To this end, without loss of generality we assume that $t > 0$. Let $w \in T_yW_t$. There exists $v \in \Pi$ such that $w = d_x\theta_t(x)v$. For every $s \in [0, t]$, set $z(s) := d_x\theta(s, x)v$. We have

$$\dot{z}(s) = \frac{d}{ds}d_x\theta(s, x)v = \frac{d}{dx}\dot{\theta}(t, x)v = \frac{d}{dx}X(\theta(t, x))v = dX(\theta(t, x))z(s).$$

Hence,

$$\frac{d}{ds} \langle z(s), p_x(s) \rangle = \langle \dot{z}(s), p_x(s) \rangle + \langle z(s), \dot{p}_x(s) \rangle = \langle dX(\theta(s, x))z(s), p_x(s) \rangle - \left\langle z(s), \sum_{i=1}^{m} (p_x(s), f_i(x(s)))df_i(x(s))^*p_x(s) \right\rangle.$$

Since $X(x) = \sum_{i=1}^{m} \alpha_i(x)f_i(x)$ and $\sum_{i=1}^{m} \alpha_i(x)^2 = H(\bar{x}, p_0)$ for every $x \in V'$,

$$dX(x(s))^*p_x(s) = \sum_{i=1}^{m} \alpha_i(x(s))df_i(x(s))^*p_x(s) + \sum_{i=1}^{m} (f_i(x(s)), p_x(s))\nabla\alpha_i(x(s))$$

$$= \sum_{i=1}^{m} \alpha_i(x(s))df_i(x(s))^*p_x(s) + \sum_{i=1}^{m} \alpha_i(x(s))\nabla\alpha_i(x(s))$$

$$= \sum_{i=1}^{m} \alpha_i(x(s))df_i(x(s))^*p_x(s).$$

We deduce that $\frac{d}{ds} \langle z(s), p_x(s) \rangle = 0$ for every $s \in [0, t]$. Hence,

$$\langle w, P(y) \rangle = \langle w, p_x(t) \rangle = \langle z(t), p_x(t) \rangle = \langle z(0), \tilde{p}(x) \rangle = 0.$$

This proves that $P(y)$ is orthogonal to $T_yW_t$, which implies that $P(y)$ and $\nabla S(y)$ are collinear. Furthermore, since $S(x(s)) = 2H(\bar{x}, p_0)s$ for every $s \in [0, t]$, one gets

$$\langle \nabla S(x(s)(t)), \dot{x}_s(t) \rangle = 2H(\bar{x}, p_0) = \langle p_x(t), \dot{x}_s(t) \rangle.$$

Since $\dot{x}_s(t) = X(y)$ does not belong to $T_yW_t$, we deduce that $\nabla S(x(s)(t)) = p_x(t)$. Consequently, we proved that $\nabla S(x) = P(x)$ for every $x \in V'$. □

Let us now conclude the proof of Lemma 2.10. Clearly, there exists $\epsilon > 0$ such that every solution $(x(\cdot), p(\cdot)) : [0, \epsilon] \to \mathbb{R}^n \times \mathbb{R}^n$ of (15) with $x(0) = \bar{x}$ and $p(0) \in W$ satisfies

$$x(t) \in B(\bar{x}, \rho), \quad \forall t \in [0, \epsilon].$$

Moreover, we have

$$S(x(\epsilon)) - S(\bar{x}) = 2\epsilon H(\bar{x}, p).$$
Let \( u(\cdot) \in L^\infty([0, \epsilon]; \mathbb{R}^m) \) be a control such that the solution \( y(\cdot) : [0, \epsilon] \rightarrow \mathcal{W} \) of (37) starting at \( \bar{x} \) satisfies \( y(\epsilon) = x(\epsilon) \). We have

\[
S(x(\epsilon)) - S(\bar{x}) = S(y(\epsilon)) - S(y(0)) = \int_{0}^{\epsilon} \frac{d}{dt}(S(y(t))) dt = \int_{0}^{\epsilon} \langle \nabla S(y(t)), \dot{y}(t) \rangle dt \\
\leq \int_{0}^{\epsilon} H(y(t), dS(y(t))) + \frac{1}{2} \sum_{i=1}^{m} u_i(t)^2 dt = \epsilon H(\bar{x}, p) + \int_{0}^{\epsilon} \sum_{i=1}^{m} u_i(t)^2 dt.
\]

The conclusion follows.

References


[34] Orro, P., Pelletier, F.: Differential properties of the distance function associated to a submanifold in sub-Riemannian geometry. Prépublication du LAMA (2001)
