Strong Minimality of Abnormal Geodesics for 2-Distributions

Andrei A. Agrachev* Andrei V. Sarychev†

Abstract

We investigate the local length minimality (by the $W_{1,1}$ or H_1-topology) of abnormal sub-Riemannian geodesics for rank 2 distributions. In particular, we demonstrate that this kind of local minimality is equivalent to the rigidity for generic abnormal geodesics, and introduce an appropriate *Jacobi equation* in order to compute *conjugate points*. As a corollary, we obtain a recent result of Sussmann and Liu about the global length minimality of short pieces of the abnormal geodesics.

1 Introduction

In this paper we study abnormal sub-Riemannian geodesics. Let us recall that a *sub-Riemannian structure* on a Riemannian manifold M is defined by a *bracket generating* (or a *possessing full Lie rank*) *distribution* $\mathcal{D}$ on M. A locally Lipschitzian path $q(\tau)$ ($\tau \in [0,T]$) is *admissible* if its tangents lie in $\mathcal{D}$ for almost all $\tau \in [0,T]$. Given two points q^0 and q^1 one can set out the problem of finding minimal (i.e. length-minimizing) admissible path connecting q^0 with q^1.

An essential distinction of this setting from the classical Riemannian case is that the space of *all* locally Lipschitzian paths connecting q^0 with q^1 has a structure of Banach manifold with minimal paths being critical points of the length functional, or *Riemannian geodesics* on the manifold M, whereas the space of admissible paths is not, in general, a manifold and may have singularities. These singularities correspond to the so-called *abnormal geodesics*. In fact these abnormal geodesics do not depend on the Riemannian structure and are determined by the distribution $\mathcal{D}$.

The term 'abnormal' comes from the calculus of variations since the problem of finding minimal admissible paths can be reformulated as the Lagrange problem of the calculus of variations. The *Euler-Lagrange equation* for the Lagrange problem is called a *geodesic equation*; its solutions are *extremals* of the Lagrange problem or *sub-Riemannian geodesics.* In particular, *abnormal extremals* with a vanishing Lagrange multiplier for the (length) functional are abnormal geodesics.

For a long time abnormal sub-Riemannian geodesics were not treated by geometers as proper candidates for minimizers until Montgomery gave in [15] an example of a minimal admissible path which does not correspond to any normal sub-Riemannian geodesic.

*V.A.Steklov Mathematical Institute, Russian Academy of Sciences, ul. Vavilova, 42, 117966, Moscow, Russia; E-mail: andrei@agrachev.mian.su Partially supported by Russian Fund for Fundamental Research under grant No. 93-011-1728, and by International Science Foundation under grant MSD000

†Department of Mathematics, University of Aveiro, 3800, Aveiro, Portugal; E-mail: ansar@mat.ua.pt On leave from the Institute for Control Sciences, Russian Academy of Sciences, Moscow, Russia

Later another example was constructed by Kupka ([13]), and Sussmann established in [19] the minimality of short abnormal geodesic subarcs for generic 2-distributions in R^4. Later Sussmann and Liu generalized the last result to the 2-distributions in R^n ([20]).

Another approach to the investigation of *weak* (i.e. $W_{1,\infty}$-local) [1] minimality of abnormal extremals of the Lagrange problem and the abnormal sub-Riemannian geodesics was suggested by the authors in [5, 6]. It is a kind of Legendre-Jacobi-Morse-type theory of a second variation for abnormal extremals of the Lagrange problem and the abnormal geodesics and, therefore, deals with geodesics of an arbitrary length. Among the results established in [6] are second order Jacobi-type conditions of weak minimality for abnormal geodesics, which turned out to be also conditions of *rigidity* for the corresponding abnormal geodesic paths. Recall that rigidity means that an admissible path is isolated (up to a reparametrization) in $W_{1,\infty}-$topology of in the set of all admissible paths connecting the given points $q^0, q^1 \in M$, see [7]. As it was demonstrated in [6], the rigidity conditions follow from a general necessary/sufficient conditions for critical points of a smooth mapping to be isolated at the corresponding critical level. Developing in [6] the Jacobi-Morse-type approach to abnormal geodesics, the authors introduced the notions of *Morse index* and *nullity* and derived explicit formulas for these invariants. This made it possible to establish local rigidity of an abnormal geodesic meeting th Strong Generalized Legendre Condition.

In this paper we are going to establish sufficient conditions for $W_{1,1}$-*local minimality* of abnormal sub-Riemannian geodesic paths. We call it *strong* minimality, although it differs from the traditional definition of strong minimality in the calculus of variations, which is C^0-local minimality.

It turns out that unlike a weak minimality a strong minimality does not, in general, result from positive definiteness of the second variation unless the distribution $\mathcal{D}$ has rank 2. We choose to limit our consideration to the scope of second order conditions and, therefore, deal with abnormal geodesics of rank 2 distributions.

The paper is organized in the following way. Section 2 contains preliminary material. In Section 3 we reduce the problem of strong minimality of admissible paths to time optimal control problem, present the Hamilton-Pontryagin form of the geodesic equation and define normal and abnormal sub-Riemannian geodesics. In Section 4 we define the *first* and the *second variations* along an abnormal geodesic, introduce the *Generalized Legendre Conditions,* the *Jacobi equation* and *conjugate points.* Section 5 contains all substantial results of the paper. Thus Theorem 3 provides sufficient strong (= $W_{1,1}$-local) minimality conditions for abnormal geodesic paths. Since sufficiently short subarcs of a strongly minimal path are automatically *globally* minimal, we can establish (Corollary 4) with the aid of the previous theorem, the global minimality of short geodesic arcs which satisfy the Strong Generalized Legendre Condition. It was already mentioned that the weak (= $W_{1,\infty}$-local) minimality was often realized in the form of rigidity or isolation (up to a reparametrization) of an abnormal geodesic path in $W_{1,\infty}$-topology. On the contrary, as follows from the proof of the Rashevsky-Chow theorem (Theorem 2.1), an admissible path of a distribution of full Lie rank is never isolated in $W_{1,1}$-topology in the

[1] By $W_{1,k}[0,T]$, $k = 1, 2, \ldots \infty$ we denote the spaces of absolutely continuous (vector-) functions on $[0,T]$ $(T < \infty)$ whose derivatives belong to $L_k[0,T]$. They become Banach spaces when provided with the norms: $\|w(\cdot)\|_{1,k} = (|w(0)|^2 + \|\dot{w}(\cdot)\|^2_{L_k})^{1/2}$. In particular, $W_{1,1}[0,T]$ is the space of absolutely continuous functions, $W_{1,2}$ is Sobolev space $H_1[0,T]$.

space of admissible paths with given end–points. Nevertheless, for a generic abnormal geodesic of rank 2 distribution the intervals of its rigidity and strong minimality coincide (Theorem 5), i.e., the property of strong minimality does not depend on the Riemannian structure. It can be explained by the that for a generic geodesic strong minimality is equivalent to *strong constrained rigidity.* Theorem 1 that gives sufficient conditions for strong constrained rigidity is proved in Section 7, and Section 6 contains a reduced form of the Jacobi equation for abnormal geodesics satisfying some regularity conditions and also some examples of strongly minimal abnormal geodesic paths.

We are grateful to F. Silva Leite who suggested some improvements of the text and especially to M. Zhitomirskii for the detailed reviewing of the manuscript and a number of helpful remarks and advice.

2 Preliminaries

Below we use notation and technical tools of chronological calculus developed by Agrachev and Gamkrelidze (see [3, 4]).

We identify C^∞ *diffeomorphisms* $P : M \to M$ with automorphisms of the algebra $C^\infty(M)$ of smooth functions on M: $\phi(\cdot) \mapsto P\phi = \phi(P(\cdot))$. The image of point $q \in M$ under the diffeomorphism P will be denoted by $q \circ P$. *Vector fields* on M are first-order differential operators on M or arbitrary derivations of the $C^\infty(M)$ algebra, i.e., R-linear mappings $X : C^\infty(M) \longrightarrow C^\infty(M)$, that obey the Leibnitz rule: $X(\alpha\beta) = (X\alpha)\beta + \alpha(X\beta)$. The value $X(q)$ of the vector field X at the point $q \in M$ lies in the space $\mathcal{T}_q M$ tangent to the manifold M at the point q. We denote by $[X^1, X^2]$ *the Lie bracket* or commutator $X^1 \circ X^2 - X^2 \circ X^1$ of the vector fields X^1, X^2. It is again a first order differential operator which can be presented in local coordinates on M as

$$[X^1, X^2] = [\sum_{i=1}^n X_i^1 \partial/\partial x_i, \sum_{i=1}^n X_i^2 \partial/\partial x_i] = \sum_{i=1}^n (\partial X_i^2/\partial x X^1 - \partial X_i^1/\partial x X^2)\partial/\partial x_i.$$

This operation introduces, in the space of vector fields, a structure of Lie algebra, which is denoted by Vect M. For $X \in \mathrm{Vect} M$ the notation ad X is used for the inner derivation of Vect M: $(\mathrm{ad} X)X' = [X, X'], \forall X' \in \mathrm{Vect} M$.

For a diffeomorphism P we use the notation $\mathrm{Ad}\, P$ for the following inner automorphism of the Lie algebra Vect M: $\mathrm{Ad}\, P X = P \circ X \circ P^{-1} = P_*^{-1} X$. The last notation is used for the result of translation of the vector field X by means of the differential P_*^{-1} of the diffeomorphism P^{-1}.

A flow on M is an absolutely continuous with respect to $\tau \in R$ curve $\tau \mapsto P_\tau$ in the group of diffeomorphisms Diff M subject to the condition $P_0 = I$ (where I is the identical diffeomorphism). We assume all time-dependent vector fields X_τ to be locally integrable (see [3]) with respect to τ. The time-dependent vector field X_τ defines the ordinary differential equation $\dot{q} = X_\tau(q(\tau)), q(0) = q^0$ on the manifold M; if any solutions of this differential equation exist for all $q^0 \in M, \tau \in R$, then the vector field X_τ is said to be complete and defines a flow on M, being the unique solution of the (operator) differential equation

$$dP_\tau/d\tau = P_\tau \circ X_\tau, P_0 = I. \tag{2.1}$$

We denote this solution by $P_t = \overrightarrow{\exp} \int_0^t X_\tau d\tau$, and call it (see [3, 4]) a *right chronological exponential* of X_τ. If the vector field $X_\tau \equiv X$ is time-independent, then the corresponding flow is denoted by $P_t = e^{tX}$.

We also introduce *a Volterra expansion* (or Volterra series) for the chronological exponential. It is (see [3, 4])

$$\overrightarrow{\exp} \int_0^t X_\tau d\tau \asymp I + \sum_{i=1}^{\infty} \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \dots \int_0^{\tau_{i-1}} d\tau_i (X_{\tau_i} \circ \dots \circ X_{\tau_1}) \asymp$$
$$\asymp I + \int_0^t X_{\tau_1} d\tau_1 + \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 (X_{\tau_2} \circ X_{\tau_1}) + \cdots . \tag{2.2}$$

For time-independent X we obtain

$$e^{tX} \asymp \sum_{i=0}^{\infty} (t^k/k!) \underbrace{X \circ \dots \circ X}_{k} \asymp I + tX + (t^2/2) X \circ X + \cdots \tag{2.3}$$

One more formula of chronological calculus will be intensively used. It is a "generalized variational formula" for a chronological exponential $\overrightarrow{\exp} \int_0^t (\hat{X}_\tau + X_\tau) d\tau$ of a "perturbed" vector field $\hat{X}_\tau + X_\tau$. We give two (left and right) variants of the formula (see [3, 4] for their drawing):

$$\overrightarrow{\exp} \int_0^t (\hat{X}_\tau + X_\tau) d\tau = \overrightarrow{\exp} \int_0^t \hat{X}_\tau d\tau \circ \overrightarrow{\exp} \int_0^t \mathrm{Ad}(\overrightarrow{\exp} \int_t^\tau \hat{X}_\theta d\theta) X_\tau d\tau =$$
$$= \overrightarrow{\exp} \int_0^t \mathrm{Ad}(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta) X_\tau d\tau \circ \overrightarrow{\exp} \int_0^t \hat{X}_\tau d\tau. \tag{2.4}$$

By applying the operator $\mathrm{Ad}(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta)$ to a vector field Y and differentiating $\mathrm{Ad}(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta) Y = (\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta) \circ Y \circ (\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta)^{-1}$ with respect to τ, we come to the relation (see [3, 4])

$$\frac{d}{d\tau} \mathrm{Ad}(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta Y) = \mathrm{Ad}(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta) \,\mathrm{ad}\, \hat{X}_\tau Y,$$

which is of the same form as (2.1). Therefore $\mathrm{Ad}(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta)$ can be presented (at least formally) as an operator chronological exponential $\overrightarrow{\exp} \int_0^t \mathrm{ad}\, \hat{X}_\theta d\theta$ which for the time-independent vector field $\hat{X}_\tau \equiv \hat{X}$ can be written as $e^{t \,\mathrm{ad}\, \hat{X}}$.

According to this new notation, the generalized variational relation (2.4) can be represented as

$$\overrightarrow{\exp} \int_0^t (\hat{X}_\tau + X_\tau) d\tau = \overrightarrow{\exp} \int_0^t \hat{X}_\tau d\tau \circ \overrightarrow{\exp} \int_0^t (\overrightarrow{\exp} \int_t^\tau \mathrm{ad}\, \hat{X}_\theta d\theta) X_\tau d\tau =$$
$$= \overrightarrow{\exp} \int_0^t (\overrightarrow{\exp} \int_0^\tau \mathrm{ad}\, \hat{X}_\theta d\theta) X_\tau d\tau \circ \overrightarrow{\exp} \int_0^t \hat{X}_\tau d\tau. \tag{2.5}$$

The exponentials $\overrightarrow{\exp} \int_0^t \mathrm{ad}\, \hat{X}_\theta d\theta$ and $e^{t \,\mathrm{ad}\, \hat{X}}$ also admit the Volterra expansions

$$\overrightarrow{\exp} \int_0^t \mathrm{ad}\, X_\tau d\tau \asymp I + \sum_{i=1}^{\infty} \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \dots \int_0^{\tau_{i-1}} d\tau_i (\mathrm{ad}\, X_{\tau_i} \circ \cdots \mathrm{ad}\, X_{\tau_1}) \asymp$$
$$\asymp I + \int_0^t \mathrm{ad}\, X_{\tau_1} d\tau_1 + \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 (\mathrm{ad}\, X_{\tau_2} \circ \mathrm{ad}\, X_{\tau_1}) + \cdots \tag{2.6}$$

and

$$e^{t\,\mathrm{ad}\,X} \asymp I + t\,\mathrm{ad}\,X + (t^2/2)\,\mathrm{ad}\,X \circ \mathrm{ad}\,X + \cdots. \tag{2.7}$$

We call a *r-distribution* $\mathcal{D}$ on M the space of smooth sections of a sub-bundle of the tangent bundle $\mathcal{T}M$; $\dim \mathcal{D}_q \equiv r$ is constant for all $q \in M$. A generalization of the concept of distribution is *a differential system* or *a distribution with singularities*, which is a space of sections of a sub-bundle with nonconstant $\dim \mathcal{D}_q$. In other words, the differential systems $\mathcal{D}$ are $C^\infty(M)$-submodules of $\mathrm{Vect}\, M$, while the distributions correspond to projective $C^\infty(M)$-submodules. Locally one can treat the germ of a distribution as a free module.

If $\mathcal{D}$ is a differential system, then, taking theC^∞-modules generated by the Lie brackets of order $\leq k$, $k = 1, \dots,$ of the vector fields subject to $\mathcal{D}$, we obtain an expanding sequence of differential systems:

$$\mathcal{D} \subseteq \mathcal{D}^2 = [\mathcal{D}, \mathcal{D}] \cdots \subseteq \mathcal{D}^k = [\mathcal{D}, \mathcal{D}^{k-1}] \subseteq \cdots.$$

For any $q \in M$ the sequence of subspaces

$$\mathcal{D}_q \subseteq \cdots \mathcal{D}_q^k \subseteq \mathcal{T}_q M$$

is called *a flag of the differential system* $\mathcal{D}$ at the point $q \in M$, and the sequence $n_1(q) \leq \cdots n_k(q) \leq \cdots$, where $n_i(q) = \dim \mathcal{D}_q^i$, is called *a growth vector of the differential system* $\mathcal{D}$ *at the point* q. A differential system is *bracket generating* or *having full Lie rank* at a point $q \in M$ if $\mathcal{D}_q^{\bar{k}} = \mathcal{T}_q M$ for a certain $\bar{k}$. A differential system is *bracket generating* or or *having full Lie rank* if $\mathcal{D}_q^{k_q} = \mathcal{T}_q M$ for a certain k_q and all $q \in M$.

A fundamental property of the full-Lie-rank differential systems is established by the following theorem.

Theorem 2.1 (Rashevsky-Chow Theorem; see [14]) *If a differential system has a full Lie rank on the manifold M, then any two points of M can be connected by an admissible path.* $\square$

If $\mathcal{D}$ is a distribution ($n_1(q) \equiv \mathrm{const}$), then $\mathcal{D}^k$ may still have singularities since the growth vector of a distribution changes, in general, with q. A distribution is called *regular* if its growth vector is constant for all q.

Below we use some standard (see, for example, [9]) concepts of symplectic geometry. A *symplectic structure* in an even-dimensional linear space Σ is defined by a nondegenerate bilinear skew-symmetric 2-form $\sigma(\cdot,\cdot)$. Two vectors $\xi_1, \xi_2 \in \Sigma$ are *skew-orthogonal*, written $\xi_1 \flat \xi_2$, if $\sigma(\xi_1, \xi_2) = 0$. If N is a subspace of Σ, then $N^\flat$ is its skeworthogonal complement: $N^\flat = \{\xi \in \Sigma \mid \sigma(\xi, \nu) = 0, \forall \nu \in N\}$. Obviously, $\dim N + \dim N^\flat = \dim \Sigma$. A subspace $\Gamma \subseteq \Sigma$ is *isotropic* when $\Gamma \subseteq \Gamma^\flat$, and *coisotropic* when $\Gamma \supseteq \Gamma^\flat$. A subspace $\Lambda \subset \Sigma$ is *Lagrangian* if $\Lambda^\flat = \Lambda$. If Λ is a Lagrangian subspace and Γ is isotropic, then it is easy to prove that $(\Lambda \cap \Gamma^\flat) + \Gamma = (\Lambda + \Gamma) \cap \Gamma^\flat$ is a Lagrangian subspace. We denote it by Λ^Γ.

3 Minimal Paths, Geodesics, Abnormal Geodesics

Recall that the problem we started with is: given a rank 2 distribution $\mathcal{D}$ on a Riemannian manifold M, establish whether a given admissible nonselfintersecting path $t \mapsto \hat{q}(t)$, $t \in [0, T]$, connecting points $q^0 = \hat{q}(0), q^1 = \hat{q}(T) \in M$, is a $W_{1,1}$-local length-minimizer in the set of all admissible paths connecting q^0 and q^1. In order to pose the problem properly, we have to define the $W_{1,1}$-neighborhood of the given Lipschitzian path $t \mapsto \hat{q}(t),\ t \in [0,T]$.

Let us consider the graph $(t, \hat{q}(t)) : [0,T] \to [0,T] \times M$ of this path. In the sufficiently small neighborhood Ω of this graph in $R \times M$ we can choose a basis $B_{t,q} : T_qM \to R^n$ of T_qM continuously depending on $(t,q) \in W$. Then any Lipschitzian path $q(\cdot)$ on M parametrized by $[0,T]$ corresponds to a R^n-valued vector function $t \mapsto B_{t,q(t)}\dot{q}(t)$ defined almost everywhere on $[0,T]$. We shall identify $\dot{q}(t)$ with $B_{t,q(t)}\dot{q}(t)$. Assuming that the distance ρ between two points of M is defined by the Riemannian metric, we can define a $W_{1,1}$-norm locally in a small C^0-neighborhood of $\hat{q}(\cdot)$:

$$\|q^1(\cdot) - q^2(\cdot)\|_{1,1} = \rho(q^1(0), q^2(0)) + \int_0^T |\dot{q}^1(t) - \dot{q}^2(t)| dt.$$

Definition 3.1 *A nonselfintersecting admissible path $t \mapsto \hat{q}(t),\ t \in [0,T]$ of the distribution $\mathcal{D}$ with the end-points q^0 and q^1 is $W_{1,1}-$local minimizer if, for some neighborhood of $\hat{q}(\cdot)|_{[0,T']}$ in $W_{1,1}[0,T']$, the points q^0 and q^1 cannot be connected by a shorter admissible path $t \mapsto \bar{q}(t),\ t \in [0,T]$ belonging to this neighborhood.* □

The problem of finding a minimal admissible path for a 2-dimensional distribution can be represented as the following *time-optimal control problem*:

$$T \longrightarrow \min, \tag{3.1}$$

$$\dot{q} = g^1(q)u_1(\tau) + g^2 u_2(\tau),\ q(0) = q^0,\ u_1^2 + u_2^2 \le 1, \tag{3.2}$$

$$q(T) = q^1, \tag{3.3}$$

where $g^1(q), g^2(q)$ are smooth vector fields, which form a basis of the distribution $\mathcal{D}$ in a small neighborhood of the given nonselfintersecting admissible path $\hat{q}(\cdot)$ on M. We denote $G(q) = (g^1(q), g^2(q))$. The admissible controls $u(\tau) = (u_1(\tau), u_2(\tau))$ are measurable functions with the values in the unit ball $B \subset R^2$; the set of admissible controls is denoted by $\mathcal{U} : \mathcal{U} \subset L_\infty$.

The following proposition establishes the equivalence of the optimal control problem (3.1-3.3) with the one of finding $W_{1,1}$-locally minimal admissible path.

Lemma 3.1 (Reduction Lemma) *An admissible path parametrized by the length of arc $\tau \mapsto \hat{q}(\tau),\ 0 \le \tau \le T$ is a $W_{1,1}$-local minimizer if and only if the corresponding control $\hat{u}(\cdot)$ is an L_1-local minimizer for the time-optimal problem (3.1)-(3.3). The corresponding minimal time T is the length of the minimal admissible path.* □

Proof. a) We start with establishing the following inequality for the Euclidean norm $|\cdot|$ in R^n :

$$v, w \in R^n, |w| = 1, v \neq 0 \Rightarrow |w - \frac{v}{|v|}| \le 2|w - v|. \tag{3.4}$$

Indeed, by arranging the terms in the equivalent inequality

$$(w - \frac{v}{|v|}) \cdot (w - \frac{v}{|v|}) \le 4(w - v) \cdot (w - v)$$

and dividing it by 2, we transform it into

$$1 - 4w \cdot v + \frac{w \cdot v}{|v|} + 2|v|^2 \ge 0,$$

or, if θ is the cosine of the angle between v and w,

$$1 - 4|v|\cos\theta + \cos\theta + 2|v|^2 \geq 0.$$

When $cos\theta \leq 0$, the last inequality follows from the obvious inequalities $1+\cos\theta \geq 0$ and $-4|v|\cos\theta + 2|v|^2 \geq 0$. If, on the contrary, $\cos\theta \geq 0$, then $1+\cos\theta \geq 2\cos^2\theta$ and, therefore

$$1 + \cos\theta + 2|v|^2 \geq 2(\cos^2\theta + |v|^2) \geq 4|v|\cos\theta$$

by virtue of the arithmetic-geometric mean inequality.

Before proving the equivalence of time-optimality of $\hat{u}(\cdot)$ and strong minimality of $\hat{q}(\cdot)$, we fix a monotonic sequence $\{\varepsilon_k\} > 0$, $\lim_{k\to\infty}\varepsilon_k = 0$. Note that when the controls $u^k(\cdot)$ tend to $\hat{u}(\cdot)$ in the L_1-norm, then the corresponding trajectories of the systems $\dot{y}^k = G(y^k)u^k(t)$, $y^k(0) = q^0$ tend to $\hat{q}(\cdot)$ in the $W_{1,1}$-norm and hence in the C^0-norm as well.

(b)Suppose that the control $\hat{u}(\cdot)$ producing the admissible path $\hat{q}(\cdot)$ is not L_1-locally optimal for Problem (3.1)-(3.3). Then there exists a sequence of admissible (with their values in the unit ball of R^2) controls $u^k(\cdot)$ belonging to the ε_k-neighborhoods of $\hat{u}(\cdot)$ in L_1^2 and steering system (3.2) from q^0 to q^1 in the time $T_k < T$ along the paths $y^k(\cdot)$. Obviously, $length(y^k(\cdot)) < length(\hat{q}(\cdot)) = T$. Without loss of generality, we can assume that T_k converges. If $\lim_{k\to\infty} T_k = T' < T$ then, since $\lim_{k\to\infty}\|u^k - u|_{[0,T_k]}\|_{L_1} = 0$, we find that $q^1 = \lim_{k\to\infty} y^k(T_k) = \hat{q}(T')$ and, therefore, $\hat{q}(\cdot)$ must be selfintersecting. If $\lim_{k\to\infty} T_k = T$, then, choosing k such that $\|u^k - u|_{[0,T_k]}\|_{L_1} \leq \varepsilon/2$, $T - T_k \leq \varepsilon/2$ and defining $y^k(t) = q^1$ for $t \in [T_k, T]$, we obtain shorter admissible path between q^0 and q^1 which belongs to the ε-neighborhood of $\hat{q}(\cdot)$ in the $W_{1,1}$-metric.

(c) If now $\hat{q}(\cdot)$ is not a $W_{1,1}$-locally minimal path, then there exist admissible paths $y^k(\cdot)$, parametrized by $[0,T]$, which are ε_k-close to $\hat{q}(\cdot)$ in the $W_{1,1}$-metric and have $thelength(y^k(\cdot)) < length(\hat{q}(\cdot) = T$. Obviously, $\lim_{k\to\infty} l_k = T$. The relations $\dot{y}^k(t) = G(y^k(t))u^k(t)$, $t \in [0,T]$ unequely define $u^k(\cdot)$, $k = 1, \ldots$ The controls $u^k(\cdot)$ may have values outside of the unit ball in R^2. Parametrizing each $y^k(\cdot)$ by the length of ar,c we represent them as trajectories of the differential equations

$$\dot{y}^k = G(q)u^k(t_k(\tau))/\|u^k(t_k(\tau))\|, \ \tau[0, l_k],$$

where $t_k(\tau)$ is the inverse to the function $\tau_k(t) = \int_0^t \|u^k(s)\|ds$. Then, by virtue of (3.4)

$$\int_0^{l_k} \left| \frac{u^k(t_k(\tau))}{|u^k(t_k(\tau))|} - u^k(\tau) \right| d\tau \leq 2\int_0^{l_k} |u^k(t_k(\tau)) - u^k(\tau)|d\tau \leq$$
$$\leq 2\left(\int_0^{l_k} |u^k(t_k(\tau)) - u^k(\tau)|d\tau + \int_0^{l_k} |u^k(\tau) - u(\tau)|d\tau \right).$$

The second term on the right-hand side obviously tends to zero as $k \to \infty$ and, since $\{t_k(\tau)\}$ converges uniformly to $t(\tau) \equiv \tau$ on any subinterval $[0,T'] \subset [0,T]$, the first term tends to zero as well. □

A first-order necessary minimality condition for the time-optimal control problem (3.1)-(3.3) is provided by the *Pontryagin Maximum Principle.* We assume the local coordinates $(t,q) = (t, q_1, \ldots, q_n)$ to be defined in some neighborhood of the curve $(t, \hat{q}(t))$, $t \in [0,T]$.

Theorem 3.2 (Pontryagin Maximum Principle; [16]) *If $\hat{u}(\cdot)$ is a weak (= L_∞-local) minimizer for Problem (3.1)-(3.3), then there exists a nonvanishing absolutely continuous covector-function $\hat{\psi}(\cdot)$) on $[0,T]$ such that $\hat{\psi}(\tau) \in \mathcal{T}^*_{\hat{q}(\tau)}M$ and, in the local coordinates (t,q), the 4-tuple $(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}(\cdot), T)$*

(1) satisfies the Hamiltonian system

$$\dot{q} = \frac{\partial H}{\partial \psi},\ q(0) = q^0,\ q(T) = q^1, \tag{3.5}$$

$$\dot{\psi} = -\frac{\partial H}{\partial q}, \tag{3.6}$$

with the Hamiltonian

$$H(u, q, \psi) = \psi \cdot G(q)u; \tag{3.7}$$

(2) obeys the maximality condition a.e. on $[0,T]$:

$$0 \leq \text{const} = H(\hat{u}(\tau), \hat{q}(\tau), \hat{\psi}(\tau)) = \max\{H(u, \hat{q}(\tau), \hat{\psi}(\tau)) \mid u \in R^2, \|u\| \leq 1\}. \ \Box \tag{3.8}$$

Definition 3.2 *The 4-tuple $(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}(\cdot), T)$ subject to the conditions of the Pontryagin Maximum Principle is an extremal of the optimal control problem (3.1)-(3.3) or a sub-Riemannian geodesic. A sub-Riemannian geodesic is normal if $H > 0$ and abnormal, if $H = 0$. The corresponding triple $(\hat{u}(\cdot), \hat{q}(\cdot), T)$ is called sub-Riemannian geodesic path.* $\Box$

Remark. Obviously for any normal or abnormal sub-Riemannian geodesic $(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}(\cdot), T)$ its restriction $(\hat{u}(\cdot)|_{[0,t]}, \hat{q}(\cdot)|_{[0,t]}, \hat{\psi}(\cdot)|_{[0,t]}, t)$ to the subinterval $[0,t] \subset [0,T]$ is also a normal or abnormal sub-Riemannian geodesic correspondingly. $\Box$

Remark. Several geodesics with different $\hat{\psi}(\cdot)$ may correspond to the geodesic path $(\hat{u}(\cdot), \hat{q}(\cdot), T)$. $\Box$

Definition 3.3 *The geodesic path $(\hat{u}(\cdot), \hat{q}(\cdot), T)$ is a corank k abnormal geodesic path if the space of $\hat{\psi}(\cdot)$, which, together with $(\hat{u}(\cdot), \hat{q}(\cdot), T)$, satisfies Theorem 3.2 with $H \equiv 0$, is k-dimensional.* $\Box$

When a geodesic is abnormal, i.e., $H = 0$ in the Pontryagin Maximum Principle, then the maximality condition (3.8) becomes $\max\{\hat{\psi}(\tau)G(\hat{q}(\tau))u \mid u \in R^2, \|u\| \leq 1\} = 0$, which is equivalent to

$$H(u, \hat{q}(\tau), \hat{\psi}(\tau)) = \hat{\psi}(\tau)G(\hat{q}(\tau))u \equiv 0, \forall u \in R^2. \tag{3.9}$$

This means that $\hat{\psi}(\tau)$ is orthogonal to the distribution $\mathcal{D}$ at $\hat{q}(\tau)$. Therefore we have one more equivalent definition of abnormality.

Definition 3.4 *The geodesic $(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}(\cdot), T)$ is abnormal if $\hat{\psi}(\tau) \perp \mathcal{D}(\hat{q}(\tau))$ at every point $\hat{q}(\tau)$.* $\Box$

We can see that the definition of an abnormal geodesic does not depend on a Riemannian structure but is only defined by the distribution $\mathcal{D}$.

Abnormal admissible geodesic paths of a distribution often exhibit a phenomenon called *rigidity*.

Definition 3.5 *The admissible path $q(\cdot)$ of the distribution $\mathcal{D}$ with end-points q^0 and q^1 is rigid if it is isolated (up to a reparametrization) in the topology of $W_{1,\infty}$ in the set of all admissible paths connecting q^0 with q^1.* $\square$

Remark. Paper [6] contains detailed consideration of the rigidity phenomenon for distributions and differential systems. $\square$

In [6] some necessary weak minimality conditions for abnormal geodesics are established. First, differentiating identity (3.9) with respect to τ, we derive

$$\hat{\psi}(\tau) \cdot [G\hat{u}(\tau), Gw](\hat{q}(\tau)) = 0, \ \forall w \in R^2, \text{ a.e. on } [0,T],$$

for almost all $\tau \in [0,T]$, which means that for all $\tau \in [0,T]$ we have

$$\hat{\psi}(\tau) \cdot [Gv, Gw](\hat{q}(\tau)) = 0 \ \forall v, w \in R^2, \tag{3.10}$$

i.e., at every point $\hat{q}(\tau)$ of the abnormal geodesic $(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}(\cdot), T)$ the covector $\hat{\psi}(\tau)$ annihilates the distribution $[\mathcal{D}, \mathcal{D}](\hat{q}(\tau))$, spanned by the vector fields f, g from $\mathcal{D}$ and their Lie bracket $[f, g]$.

The following *Generalized Legendre Condition* (see [11, 2, 12]) [2] is necessary for the weak and therefore also strong minimality of an abnormal geodesic path: for some $\hat{\psi}(\cdot)$ satisfying Pontryagin Maximum Principle (Theorem 3.2) we have

$$\frac{\partial}{\partial u} \frac{d^2}{d\tau^2} \frac{\partial H}{\partial u} \Big|_{\hat{u}(\tau)} (v, v) = \gamma_\tau(v, v) = \hat{\psi}(\tau) \cdot [Gv, [G\hat{u}(\tau), Gv]](\hat{q}(\tau)) \geq 0. \tag{3.11}$$

for all $\tau \in [0,T]$, and $v \perp \hat{u}(\tau)$.

In order to set out Jacobi-type minimality conditions, we introduce *Strong Generalized Legendre Condition*: for some $\beta > 0$ and for all $\tau \in [0,T]$ and $v \perp \hat{u}(\tau)$: we have

$$\gamma_\tau(v, v) = \hat{\psi}(\tau) \cdot [Gv[G\hat{u}(\tau), Gv]](\hat{q}(\tau)) \geq \beta \|v\|^2. \tag{3.12}$$

The last condition also implies the smoothness of the geodesic.

Theorem 3.3 **(Smoothness of Abnormal Geodesics; see [6, Theorem 4.4])** *If the Strong Generalized Legendre Condition (3.12) holds along the abnormal geodesic $(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}(\cdot), T)$, then $\hat{u}(\tau), \hat{q}(\cdot), \hat{\psi}(\cdot)$ are smooth on* $[0,T]$. $\square$

Note that (3.10) and (3.12) imply $\hat{\psi}(\tau) \in \mathcal{D}^2_{\hat{q}(\tau)} \setminus \mathcal{D}^3_{\hat{q}(\tau)}$, $\forall \tau \in [0,T]$. On the other hand, differentiating (3.10) w.r.t. τ, we derive

$$\hat{\psi}(\tau) \cdot [G\hat{u}(\tau), [Gv, Gw]](\hat{q}(\tau)) = 0 \ \forall v, w \in R^2.$$

Therefore, if $\hat{\psi}(\tau) \in \mathcal{D}^2_{\hat{q}(\tau)} \setminus \mathcal{D}^3_{\hat{q}(\tau)}$, $\forall \tau \in [0,T]$ for the smooth abnormal extremal $(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}(\cdot))$, then the Strong Generalized Legendre Condition holds either for $(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}(\cdot))$ or for the abnormal extremal $(\hat{u}(\cdot), \hat{q}(\cdot), -\hat{\psi}(\cdot))$.

The abnormal extremals which are subject to the conditions $\hat{\psi}(\tau) \in \mathcal{D}^2_{\hat{q}(\tau)} \setminus \mathcal{D}^3_{\hat{q}(\tau)}$, $\forall \tau \in [0,T]$ are called *regular abnormal biextremals* in [20]. Therefore, the class of abnormal

[2] also called Generalized Legendre-Klebsch Condition, or Kelley Condition

extremals which satisfies condition (3.10) together with the Strong Generalized Legendre Condition coincides with the class of "regular abnormal biextremals" introduced in [20].

Below we assume the Strong Generalized Legendre Condition to hold and our abnormal geodesic $(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}(\cdot), T)$ to be smooth. Then we can choose smooth vector fields f, g spanning the distribution $\mathcal{D}$ in some neighborhood of the curve $\hat{q}(\tau),\ 0 \leq \tau \leq T$ such that the abnormal geodesic path $\hat{q}(\cdot) = q^0 \circ e^{tf}$ starts at the q^0 trajectory of the vector field f (other trajectories of f need not be geodesics). It is more convenient from the technical point of view to introduce new notations for system (3.2), namely, we shall consider the admissible paths of the distribution $\mathcal{D}$ starting at q^0 as trajectories of the affine control system

$$\dot{q} = f(q)(1 + \dot{v}(\tau)) + g(q)\dot{w}(\tau),\ q(0) = q^0, \tag{3.13}$$

where $u(\tau) = (v(\tau), w(\tau))$ is a vector function treated as control; $u(0) = 0$. What we have to do is to find out when the reference control $\hat{u} \equiv 0$ shifting the affine control system (3.13) from q^0 to q^1 is $W_{1,1}$-locally time optimal among the controls subject to the constraint

$$(1 + \dot{v})^2 + \dot{w}^2 \leq 1. \tag{3.14}$$

4 The First and Second Variations along Abnormal Geodesics

Let us introduce a family of *input/state mappings* F_t for the control system (3.13). It is defined in a neighborhood of the origin of $W^2_{1,\infty}[0,T]$ (we ignore for the timebeing the restrictions imposed on the control): F_t maps the input $u(\cdot)$ into the point $q(t) \in M$ of the trajectory $q(\cdot)$ of the differential equation $\dot{q} = f(q)(1 + \dot{v}(\tau)) + g(q)\dot{w}(\tau),\ q(0) = q^0$. Obviously, $F_t(0) = \hat{q}(t)$ and $F_T(0) = q^1$.

In order to study $F_t(u(\cdot))$, we represent it as a chronological exponential:

$$F_t(u(\cdot)) = q^0 \circ \overrightarrow{\exp} \int_0^t f(1 + \dot{v}(\tau)) + g\dot{w}(\tau)d\tau.$$

Rearranging the vector field $f(1+\dot{v}) + g\dot{w}$ as the sum $f + (f\dot{v} + g\dot{w})$ and applying the first of the generalized variational formulas (2.5), we transform the exponential into

$$F_t(u(\cdot)) = q^0 \circ e^{tf} \circ \overrightarrow{\exp} \int_0^t (f\dot{v}(\tau) + Y_{t,\tau}\dot{w}(\tau))d\tau, \tag{4.1}$$

where

$$Y_{t,\tau} = e^{(\tau - t)\operatorname{ad} f} g.$$

Taking into account $q^0 \circ e^{tf} = \hat{q}(t)$ and applying the second of the relations (2.5) to the chronological exponential in (4.1) , we obtain

$$F_t(u(\cdot)) = \hat{q}(t) \circ \overrightarrow{\exp} \int_0^t e^{\operatorname{ad} f v(\tau)} Y_{t,\tau}\dot{w}(\tau)d\tau \circ e^{fv(t)}. \tag{4.2}$$

Let us denote

$$Y_\tau = Y_{T,\tau} = e^{(\tau - T)\operatorname{ad} f} g,\ F_T = F; \tag{4.3}$$

obviously, $Y_T = g$.

The first and second variations along the chosen abnormal extremal are, correspondingly, first and second differentials of the input/state mapping F at the origin (see [6]). Using the Volterra expansions (2.2)-(2.7), we derive (see [6] for details) the expression

$$F'|_0 u(\cdot) = f(q^1)v(T) + \int_0^T Y_\tau(q^1)\dot{w}(\tau)d\tau =$$

$$f(q^1)v(T) + g(q^1)w(T) - \int_0^T \dot{Y}_\tau(q^1)w(\tau)d\tau, \tag{4.4}$$

where $\dot{Y}_\tau = [f, Y_\tau]$ by virtue of (4.3) , for the first differential.

If 0 is a critical point of F, i.e., $\operatorname{Im} F'|_0 \neq \mathcal{T}_{q^1}M$, then there exists a nonzero covector $\hat{\psi}_T \in \mathcal{T}^*_{q^1}M$ that annihilates $\operatorname{Im} F'|_0$. This implies

$$\hat{\psi}_T \cdot f(q^1) = \hat{\psi}_T \cdot g(q^1) = \hat{\psi}_T \cdot \dot{Y}_t(q^1) = 0, \text{a.e. on } [0,T]. \tag{4.5}$$

The critical character of $0 \in W^2_{1,\infty}[0,T]$ for F is another formulation of $\hat{u} \equiv 0$ which is an abnormal extremal control or $\hat{q}(\cdot)$ which is an abnormal geodesic path. Indeed, we can easily establish the equivalence of conditions (4.5) and (3.9). If we take the covector $\hat{\psi}_T$ from (4.5) and choose the solution $\hat{\psi}(\cdot)$ of the adjoint equation (3.6) with the end-point value $\hat{\psi}(T) = \hat{\psi}_T$, then $\hat{\psi}(t)$ annihilates the distribution $\mathcal{D}$ at any point $\hat{q}(t)$ of the abnormal geodesic path.

Definition 4.1 *The first differential $F'|_0 : W^2_{1,\infty}[0,T] \to \mathcal{T}_{q^1}M$, defined by relation (4.4) is called the first variation along the abnormal geodesic path $\hat{q}(t) = q^0 \circ e^{tf},\ t \in [0,T]$.* □

Let us introduce the second variation along the abnormal geodesic $(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}(\cdot), T)$. This is a Hessian, or a quadratic differential of F at $0 \in W^2_{1,\infty}[0,T]$. It is a quadratic form whose domain is the kernel of the first variation. Choosing a smooth function $\chi : M \longrightarrow R$, such that $d\chi|_{q^1} = \hat{\psi}_T$, we consider the function $\phi(u(\cdot)) = \chi(F(u(\cdot)))$. Since $\hat{\psi}_T$ annihilates $\operatorname{Im}F'|_0$, it follows that 0 is a critical point for this function.

Let us compute the quadratic term of the Taylor expansion for $\phi(u(\cdot))$ at 0. Using the Volterra expansion (2.2) for the chronolological and ordinary exponent in (4.2), we derive

$$\begin{aligned}\phi''|_0(u(\cdot)) = ((\textstyle\int_0^T \int_0^\tau Y_\xi \dot{w}(\xi)d\xi \circ Y_\tau \dot{w}(\tau)d\tau +\\ +\frac{1}{2} fv(T) \circ fv(T) + \textstyle\int_0^T Y_\tau \dot{w}(\tau)d\tau \circ fv(T))\chi)(q^1).\end{aligned} \tag{4.6}$$

When carrying out the computation we took into account the relations

$$([f, Y_\tau] \cdot \chi)(q^1) = \hat{\psi}_T[f, Y_\tau](q^1) = \hat{\psi}_T \dot{Y}_\tau(q^1) \equiv 0. \tag{4.7}$$

If we restrict the quadratic form (4.6) to the kernel of $F'|_{(T,\hat{u}(\cdot))}$, we can subtract from (4.6) the vanishing value of

$$\frac{1}{2}((fv(T) + \int_0^T Y_\tau \dot{w}(\tau)d\tau) \circ (fv(T) + \int_0^T Y_\tau \dot{w}(\tau)d\tau)\chi)(q^1),$$

and, arranging the terms and taking (4.7) into account, transform (4.6) into

$$\frac{1}{2}((\int_0^T [\int_0^\tau Y_\xi \dot{w}(\xi)d\xi, Y_\tau \dot{w}(\tau)]d\tau \circ \chi)(q^1).$$

The last expression does not depend on the choice of χ but only on $\hat{\psi}_T = d\chi|_{q^1}$ and, therefore, we can write it as

$$F''|_0(u(\cdot)) = \hat{\psi}_T \int_0^T [\int_0^t Y_\tau \dot{w}(\tau) d\tau, Y_t \dot{w}(t)](q^1) dt, \tag{4.8}$$

where $u(\cdot) = (v(\cdot), w(\cdot))$ satisfies the equality

$$f(q^1)v(T) + g(q^1)w(T) - \int_0^T \dot{Y}_\tau(q^1)w(\tau)d\tau = 0. \tag{4.9}$$

To get rid of the derivatives of $w(\cdot)$, we twice integrate (4.8) by parts transforming it into a quadratic form in $w(\cdot)$:

$$F''|_0(u(\cdot)) = \int_0^T \hat{\psi}_T[\dot{Y}_t, Y_t](q^1)w^2(t)dt +$$
$$\int_0^T \hat{\psi}_T[gw(T) + \int_0^t \dot{Y}_\tau w(\tau)d\tau, \dot{Y}_t w(t)](q^1)dt. \tag{4.10}$$

Definition 4.2 *The quadratic form (4.10)-(4.9) is called the second variation along the abnormal geodesic* $(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}(\cdot), T)$. □

Note that the quadratic form $\hat{\psi}_T[\dot{Y}_t, Y_t](q^1)w^2$, which appears in relation (4.10), coincides with the form $\gamma_t w^2$ appearing in the Generalized Legendre Conditions (3.11)-(3.12). Recall that we assumed the Strong Generalized Legendre Condition to be fulfilled, i.e., $\gamma_t \geq \kappa > 0, \forall t \in [0,T]$.

Next we introduce (as in [6]) a symplectic representation of the second variation (4.10)–(4.9) along the abnormal geodesic. Let us set

$$W = \mathrm{span}\{f(q^1) \cup g(q^1) \cup \{\dot{Y}_\tau(q^1)|\tau \in [0,T]\} \subset \mathcal{T}_{q^1}M\}. \tag{4.11}$$

Obviously, W coincides with the image $\mathrm{Im}\, F'|_0$ of the first variation (4.4) and $\hat{\psi}_T$ annihilates W by virtue of (4.5).

Let us take the space $\mathcal{E}_W$ of the vector fields, whose values at q^1 lie in W, and consider the skew-symmetric bilinear form on $\mathcal{E}_W$:

$$\hat{\psi}_T \cdot [X, X'](q^1), \ \forall X, X' \in \mathcal{E}_W. \tag{4.12}$$

This form has a kernel of finite codimension in $\mathcal{E}_W$; it is defined by the relations

$$X(q^1) = 0; \hat{\psi}_T \cdot (\partial X/\partial \xi)(q^1) = 0, \forall \xi \in W.$$

Taking the quotient of $\mathcal{E}_W$ with respect to this kernel, we obtain a (induced from (4.12)) *nondegenerate* skew-symmetric bilinear form $\sigma(\cdot,\cdot)$ on the finite-dimensional quotient space Σ, which defines the symplectic structure on Σ. A direct calculation gives $\dim \Sigma = 2 \dim W = 2(n-k)$. We denote by $\underline{X}$ the image of the $X \in \mathcal{E}_W$ under the canonical projection $\mathcal{E}_W \longrightarrow \Sigma$.

Choosing local coordinates $(x_1, \ldots x_n) : \mathcal{O} \longrightarrow R^n$ in some neghborhood $\mathcal{O}$ of q^1 in M such that (1) $x_i(q^1) = 0$, $(i = 1, \ldots n)$, (2) the subspace W is defined by the equalities

$x_1 = \dots = x_k = 0$ and (3) $\hat{\psi}_T = (\psi_1, \dots, \psi_k, 0, \dots 0)$, we can represent the canonical projection $X \mapsto \underline{X}$ as

$$X = \textstyle\sum_{i=1}^n X_i(x)\partial/\partial x_i \mapsto \underline{X} = \\ (X_{k+1}(0), \dots X_n(0), \partial(\textstyle\sum_{i=1}^k \psi_i X_i)/\partial x_{k+1}|_0, \dots \partial(\textstyle\sum_{i=1}^k \psi_i X_i)/\partial x_n|_0). \tag{4.13}$$

The symplectic form $\sigma(\underline{X}, \underline{Y})$ can now be represented as

$$\sigma(\underline{X}, \underline{Y}) = \sum_{j=k+1}^n (X_j(0)\partial(\sum_{i=1}^k \psi_i Y_i)/\partial x_j|_0 - Y_j(0)\partial(\sum_{i=1}^k \psi_i X_i)/\partial x_j|_0).$$

Let us denote by Π the image under the canonical projection of the space of the vector fields which vanish at q^1. Since the Lie bracket for two vector fields vanishing at q^1 also vanishes at q^1, Π is a Lagrangian subspace (see Section 2).

Using the notation introduced above we can represent the second variation (4.10)-(4.9) as the quadratic form

$$2F''|_0(u(\cdot)) = \int_0^T \gamma_\tau w^2(\tau)d\tau + \int_0^T \sigma(\underline{g}w(T) + \int_0^t \underline{\dot{Y}}_\tau w(\tau)d\tau, \underline{\dot{Y}}_t w(t))dt \tag{4.14}$$

with the domain

$$\{u(\cdot) = (v(\cdot), w(\cdot)) | \underline{f}v(T) + \underline{g}w(T) - \int_0^T \underline{\dot{Y}}_\tau w(\tau)d\tau \in \Pi\}. \tag{4.15}$$

Let us extend the domain of the second variation by considering not only the absolutely continuous but also arbitrary $w(\cdot) \in L_2[0, T]$ such that

$$\int_0^T \underline{\dot{Y}}_\tau w(\tau)d\tau \in \Pi + \mathrm{span}\{\underline{f}, \underline{g}\}, \tag{4.16}$$

or, equivalently,

$$\underline{f}v_T + \underline{g}w_T - \int_0^T \underline{\dot{Y}}_\tau w(\tau)d\tau \in \Pi,$$

where v_T, w_T are uniquely defined linear continuous (by L_1-norm) functionals in $w(\cdot)$: $v_T = \alpha(w(\cdot))$, $w_T = \beta(w(\cdot))$. The quadratic form (4.14) becomes

$$2F''|_0(u(\cdot)) = \int_0^T \gamma_\tau w^2(\tau)d\tau + \int_0^T \sigma(\underline{g}w_T + \int_0^t \underline{\dot{Y}}_\tau w(\tau)d\tau, \underline{\dot{Y}}_t w(t))dt \tag{4.17}$$

with $w_T = \beta(w(\cdot))$. Relation (4.16) can be represented as a system of relations:

$$\int_0^T \sigma(\nu, \underline{\dot{Y}}_\tau)w(\tau)d\tau = 0, \ \forall \nu \in \Pi \cap \{\underline{f}, \underline{g}\}^\flat. \tag{4.18}$$

Since $\sigma(\underline{g}, \underline{f}) = \sigma(\underline{Y}_t, \underline{f}) \equiv 0$, i.e., all the vectors that appear in (4.17)-(4.15) are skew-orthogonal to $\underline{f}$, we can make a reduction taking, instead of the symplectic space Σ, the quotient of its subspace $\underline{f}^\flat$ by $\underline{f}$; we denote it by Σ^f. From now on we consider

$\Pi^f = \Pi \cap \underline{f}^\flat + R\underline{f}$ instead of Π and preserve the same notations for the vectors $\underline{g}$ and $\dot{\underline{Y}}_t$ treated further modulo $R\underline{f}$. Relation (4.18) becomes

$$\int_0^T \sigma(\nu, \dot{\underline{Y}}_\tau) w(\tau) d\tau = 0, \ \forall \nu \in \Pi^f \cap \underline{g}^\flat. \tag{4.19}$$

Recall that the Legendre-Jacobi necessary/sufficient minimality conditions for normal extremals in the Calculus of Variations amount to the nonnegativeness/positive definiteness of the second variation. The positive definiteness is also essential for the minimality of abnormal geodesics.

Let us start with an observation that the Strong Generalized Legendre Condition (3.12) provides positive definiteness of the second variation for sufficiently small $T > 0$. Indeed, since $|w_T| = O(\|w(\cdot)\|_{L_1}) = O(\sqrt{T}\|w\|_{L_2})$, the term

$$\int_0^T \sigma(\underline{g} w_T + \int_0^t \dot{\underline{Y}}_\tau w(\tau) d\tau, \dot{\underline{Y}}_t w(t)) dt$$

of (4.17) admits the upper estimate $c(\int_0^T |w(t)| dt)^2 \leq cT \int_0^T |w(t)|^2 dt$. Since $\gamma_t \geq \kappa > 0$ on $[0, T]$ it follows that

$$2F''|_0(u(\cdot)) \geq (\kappa - cT) \int_0^T |w(t)|^2 dt \tag{4.20}$$

for some constant $\kappa > 0$ and therefore is positive definite for small $T > 0$.

Following the Jacobi approach, we introduce the notion of *conjugate points for abnormal geodesics.*

Definition 4.3 *The conjugate points of an abnormal geodesic are the time instants T for which the (depending on T) quadratic form (4.17) with the domain determined by (4.15) has a nontrivial kernel; the dimension of this kernel is the multiplicity of the conjugate point.* $\square$

It follows from the aforesaid that under the Strong Generalized Legendre Condition (3.12) the conjugate points of abnormal geodesic are isolated from $0 \in [0, T]$.

In order to derive a condition for conjugate points, let us first note that the kernel of the quadratic form consists of all $(v^0(\cdot), w^0(\cdot))$ such that $w^0(\cdot)$ satisfy (4.19) and

$$\int_0^T (\gamma_t w^0(t) + \sigma(\int_0^t g w_T^0 + \dot{\underline{Y}}_\tau w^0(\tau) d\tau, \dot{\underline{Y}}_t)) w(t) dt = 0$$

for any $w(\cdot)$ satisfying (4.19). This means that

$$\gamma_t w^0(t) + \sigma(\int_0^t g w_T^0 + \dot{\underline{Y}}_\tau w^0(\tau) d\tau, \dot{\underline{Y}}_t) = \sigma(-\nu, \dot{\underline{Y}}_\tau) \text{ for some } \nu \in \Pi^f \cap \underline{g}^\flat. \tag{4.21}$$

If we set $x(t) = \int_0^t \dot{\underline{Y}}_\tau w^0(\tau) d\tau + g w_T^0 + \nu$, then the integral equation (4.21) is equivalent to the equation *Jacobi differential equation*

$$\dot{x} = \gamma_t^{-1} \sigma(\dot{\underline{Y}}_t, x) \dot{\underline{Y}}_t, \ x(0) \in \Pi^{f,g}, \tag{4.22}$$

where $\Pi^{f,g} = \Pi^f \cap \underline{g}^\flat + \text{span}\{\underline{g}\}$. Condition (4.15) turns into the inclusion

$$x(T) \in \Pi^f. \tag{4.23}$$

The boundary value problem (4.22)-(4.23) has a nontrivial solution if and only if the second variation is degenerate. Therefore, Definition 4.3 is equivalent to the following definition.

Definition 4.4 *The time momemt $T > 0$ is a conjugate point for an abnormal geodesic if there exists a nontrivial solution of the boundary value problem (4.22)-(4.23) on* $[0, T]$. $\Box$

The evolution of the second variation with the growth of T was studied in [10] (see also [17]). It was shown that, provided the Strong Generalized Legendre Condition (3.12) was fullfilled, the minimal eigenvalue of the self-adjoint operator, which corresponded to the quadratic form (4.17)-(4.16), was a continuous nonincreasing function of T. Since the quadratic form is positive definite for sufficiently small $T > 0$, the conjugate points are isolated from 0. If a conjugate point appears, then the second variation *cannot* again become positive definite with a further growth of T. This means that the second variation possesses a lower estimate of the kind of(4.20) on $[0, T]$ if and only if there are no conjugate points on $[0, T]$.

In Sec.6 we shall derive, essentially following [6], another, reduced, representation of the Jacobi equation (4.22) for the abnormal extremals of the 2-dimensional distributions which satisfy some additional regularity conditions.

Now we are ready to formulate the main results of the paper.

5 Main Results

In the Section 3 we have reduced the problem of finding a minimal admissible path for the distribution $\mathcal{D}$ to the time optimal control problem for the 2-input affine control system (3.13) with the inputs (v, w) subjected to the constraint: $(1 + \dot{v})^2 + \dot{w}^2 \leq 1$. Note that the reference control $\hat{u} = (v, w) \equiv 0$, which we investigate, is an extremal control for the affine system (3.13); its values lie *on the boundary of the constraints.*

The authors established in the[6] rigidity conditions for extremals of affine control systems with *unconstrained* controls or with the reference control lying in the relative interior of the constraints. Recall that rigidity means that $\hat{u}(\cdot)$ is isolated by L_∞-metric (up to a reparametrization) in the set of controls which steer the affine system from the given initial point to the given end-point. It was already mentioned that when $\dim \mathrm{Lie}\{f, g\} = \dim M$ at every point of $\hat{q}(\cdot)$, then, by virtue of Rashevsky-Chow Theorem, $\hat{u}(\cdot)$ is *never* isolated by the L_1-metric (strongly isolated). It turns out that for affine systems with *constrained* controls such an isolation may occur. We call it a *strong constrained rigidity.* The following theorem provides sufficient condition for the strong constrained rigidity.

Theorem 1 (Strong Constrained Rigidity for Affine Control Systems) *Let an abnormal geodesic $(\hat{q}(t), \hat{\psi}(t))$ of a 2-dimensional distribution $\mathcal{D} = \mathrm{span}\{f, g\}$ satisfy the Strong Generalized Legendre Condition (3.12) and $\hat{q}(t)$ be the trajectory of the vector field f starting at q^0. Suppose that the set $U \subset R^2$ is convex and bounded and*

$$(0, 0) \notin \mathrm{int}U, \ \mathrm{int}U \cap (R \times \{0\}) \neq \emptyset.$$

If there are no conjugate points of the abnormal geodesic on $[0, t]$, then, for some $\varepsilon > 0$, we have

$$q^0 \circ \overrightarrow{\exp} \int_0^t f(1 + \dot{v}(\tau)) + g\dot{w}(\tau) d\tau \neq \hat{q}(t) \tag{5.1}$$

provided that $(\dot{v}(\tau), \dot{w}(\tau)) \in U$, for $\tau \in [0, t]$ and

$$0 < \int_0^t |(\dot{v}(\tau), \dot{w}(\tau))| d\tau < \varepsilon, \tag{5.2}$$

where $|(\dot{v},\dot{w})| = (\dot{v}^2+\dot{w}^2)^{1/2}$. $\Box$

Remark. Conditions (5.1)-(5.2) mean that for sufficiently small $\varepsilon > 0$ there is no *admissible* control from the ε-neighborhood of the origin in L_1 which steers in time t the affine system $\dot{q} = f(q)(1+\dot{v}(\tau))+g(q)\dot{w}(\tau)$ from q^0 to $\hat{q}(t)$, i.e., the trajectory $\hat{q}(\cdot)$ exhibits strong constrained rigidity. $\Box$

We shall prove of the theorem in the next section and shall now derive an important corollary.

Corollary 2 *Suppose that the abnormal geodesic* $(\hat{q}(t),\hat{\psi}(t))$, *where* $\hat{q}(t) = q^0 \circ e^{tf}$, *satisfies the Strong Generalized Legendre Condition (3.12). Then for the arbitrary norm* $\|\cdot\|$ *in* R^2 *there exists* $\bar{t} > 0$ *(depending on the norm) such that* $\forall t \in [0,\bar{t}]$, $\forall s > 0$ *the relations*

$$q^0 \circ \overrightarrow{\exp} \int_0^s f\dot{v}^\star(\tau) + g\dot{w}(\tau)d\tau = \hat{q}(t), \tag{5.3}$$

$$\int_0^s \|(\dot{v}^\star(\tau),\dot{w}(\tau)\|d\tau \le \|(1,0)\|t \tag{5.4}$$

can only hold if $\dot{v}^\star(\tau) \equiv 1,\ \dot{w}(\tau) \equiv 0,$ *on* $[0,s]$. $\Box$

Proof. According to the aforesaid there are no conjugate points on $[0,\bar{t}]$ for sufficiently small $\bar{t} > 0$. Without loss of generality, we can assume that $\|(1,0)\| = 1$. We can also rescale the time variable in the chronological exponential of (5.3) and in the integral of (5.4) in such a way that $\|(\dot{v}^\star(t),\dot{w}(t)\| \equiv 1$, there by transforming inequality (5.4) into $s \le t$ for the rescaled s.

Let us consider the convex bounded set $U = \{u \in R^2 | \|(1,0)+u\| \le 1\}$ and assume that $|u| \le c\|u\|, \forall u \in R^2$. If we extend $(\dot{v}^\star(t),\dot{w}(t))$ by means of zero from $[0,s]$ onto $[0,t]$ and set $\dot{v}(\tau) = \dot{v}^\star(\tau) - 1$, then

$$\int_0^t |(\dot{v}(\tau),\dot{w}(\tau))|d\tau = \int_0^t |(\dot{v}^\star(\tau),\dot{w}(\tau)) - (1,0)|d\tau \le$$
$$\le \int_0^s |(\dot{v}^\star(\tau),\dot{w}(\tau))|d\tau + t \le cs + t \le (c+1)t.$$

Choosing a positive $\bar{t} < \frac{\varepsilon}{c+1}$, we satisfy the upper inequality of (5.2) and use (5.1) to prove the Corollary. $\Box$

The previous results do not depend on any Riemannian structure on M but only depend on the distribution $\mathcal{D}$. Suppose now that M possesses a Riemannian structure. The following two theorems are direct corollaries of the results that we formulated above.

Theorem 3 (The Sufficient Condition of Strong Minimality for Abnormal Geodesics) *If an abnormal sub-Riemannian geodesic satisfies the Strong Generalized Legendre Condition (3.12), there are no conjugate points in its domain* $[0,T]$, *and the corresponding geodesic path connecting the points* q^0 *and* q^1 *is nonselfintersecting, then it is strongly (*$= W_{1,1}$*-locally) minimal.* $\Box$

Proof. We can assume that the abnormal geodesic path $\hat{q}(\tau)$, $0 \le \tau \le T$, is a trajectory of the vector field f and f,g form an *orthonormal basis* of the distribution $\mathcal{D}$

in its neighborhood. Then the length of the geodesic path is T. Considering the convex bounded set $U = \{u \in R^2 |\ |u + (1,0)| \leq 1\}$, we can apply the Theorem 1 and infer that there is not a single control $u(\cdot) = (u_1(\cdot), u_2(\cdot))$ which is ε-close by L_1-metric to $\hat{u}(\cdot) \equiv (1,0)$ and different from $\hat{u}(\cdot)$ with the values $|u(\tau)| \leq 1$ and which can steer the control system $\dot{q} = fu_1 + gu_2$ from q^0 to q^1 in time $\leq T$. Now the strong ($W_{1,1}$-local) minimality follows from the Reduction Lemma 3.1. □

Remark. Note that we derive the strong minimality from the strong constrained rigidity and since the sufficient condition for the Strong Constrained Rigidity established in Theorem 1 is valid for the arbitrary convex bounded set U, the abnormal geodesic path satisfying the conditions of Theorem 3 is $W_{1,1}$-locally minimal for any choice of the Riemannian metric. □

From Corollary 2 we can find the sufficient condition for the global minimality of the abnormal geodesic arcs established in [20, Theorem 6].

Corollary 4 (The Sufficient Condition for The Global Minimality of Abnormal Geodesic Subarcs). *If the abnormal sub-Riemannian geodesic $(\hat{q}(\cdot), \hat{\psi}(\cdot))$ satisfies the Strong Generalized Legendre Condition (3.12) and the corresponding geodesic path is nonselfintersecting, then there exists $\bar{t} > 0$ such that $\forall t \in (0, \bar{t})$ the restriction $\hat{q}(\cdot)|_{[0,t]}$ is globally minimal among the admissible paths connecting the point q^0 with the point $\hat{q}(t)$.* □

Proof. We can again assume that $\hat{q}(\cdot)$ is the trajectory of the vector field f or, all the same, of the control system $\dot{q} = f(q)\dot{v}^\star + g(q)\dot{w}$ driven by the control $\hat{u} = (\dot{v}, \dot{w}) \equiv (1,0)$. Taking $\|u\| = |u|,\ |(1,0)| = 1$, we derive from Corollary 2 the existence of $\bar{t}$ such that for any $t \in (0, \bar{t}]$ the system *cannot* be steered in time $\leq t$ from q^0 to $\hat{q}(t)$ by any other control than $\hat{u}(\cdot)$ with the values in the unit ball $|u| \leq 1$. This means that the restriction $\hat{q}(\cdot)|_{[0,t]}$ is a strictly minimal path connecting q^0 with $\hat{q}(t)$. □

Remark. In contrast to the previous remark, here the interval $[0, \bar{t}]$ depends on the choice ofthe Riemannian structure. □

The following theorem is a corollary of Theorem 1 and of the rigidity conditions for abnormal geodesic paths derived in [6] . Here rigidity is understood in accordance with Definition 3.5.

Theorem 5 (Minimality and Rigidity) *Suppose that the Strong Generalized Legendre Condition holds for the abnormal geodesic path $\hat{q}(t),\ t \in [0,T]$, and its restrictions $\hat{q}(\cdot)|_{[t_1,t_2]}$ on any nontrivial interval $[t_1, t_2]$ have corank 1. Then:*

(1) if $\hat{q}(\cdot)|_{[0,T]}$ is rigid then $\forall t \in [0,T)$ the restriction $\hat{q}(\cdot)|_{[0,t]}$ is a strict $W_{1,1}$-local length-minimizer in the set of the admissible paths connecting q^0 with $\hat{q}(t)$;

(2) if $\hat{q}(\cdot)|_{[0,T]}$ is not a normal geodesic path and is a $W_{1,\infty}$-local length-minimizer in the set of the admissible paths connecting q^0 with $\hat{q}(T)$, then $\hat{q}(\cdot)|_{[0,t]}$ is rigid for any $t \in (0,T)$. □

Proof. 1) If the corank 1 geodesic path $\hat{q}(\cdot)|_{[0,T]}$ is rigid, then, as it was established in [6, Theorem 4.1], the second variation along the geodesic must be nonnegative. Since a restriction of the abnormal geodesic path $\hat{q}(\cdot)$ on any nontrivial interval $[t_1, t_2]$ has corank 1, i.e., the so-called *strong regularity condition* (see [17]) holds, then (see [10, 17]) the absence of conjugate points on $(0,T)$ is necessary for the nonnegativeness and we can apply Theorem 3 to establish the strict $W_{1,1}$-local minimality of the restriction $\hat{q}(\cdot)|_{[0,t]}$ on any $t \in (0, T)$.

(2) If a corank 1 abnormal geodesic path $\hat{q}(\cdot)|_{[0,T]}$ is a $W_{1,\infty}$-locally minimal, then, as it was established in [6, Theorem 4.2] the second variation along the geodesic must be nonnegative. Together with the strong regularity condition, it implies the absence of conjugate points on $(0, T)$ and positive definiteness of the second variation along any restriction $\hat{q}(\cdot)|_{[0,t]}$ with $t \in (0, T)$. According to [6, Theorem 4.8], all these restrictions are rigid. □

6 A Reduced Form of the Jacobi Equation for Regular Distributions. Examples

In this section we introduce (essentially following [6]) another form of the Jacobi equation for the 2-dimensional distributions which satisfy some regularity condition and also provide some examples of strongly minimal abnormal geodesics for 2-*dimensional* smooth distributions on 3- and 4-dimensional manifolds (compare with [15, 7, 20]).

Let us consider a 2-dimensional distribution $\mathcal{D}$ on the $(n+2)-$dimensional manifold M; let the vector fields $f, g \in \mathrm{Vect}M$ span $\mathcal{D}$. Suppose that:

(i) the vector fields

$$f, g, [f, g], \ldots (\mathrm{ad}f)^{n-1}g$$

are linearly independent at every point of the domain that we consider;

(ii) $(\mathrm{ad}f)^n g$ can be presented as a linear combination with C^∞-coefficients of these $n+1$ vector fields:

$$(\mathrm{ad}f)^n g = \beta f + \sum_{i=0}^{n-1} \alpha^i (\mathrm{ad}f)^i g \ (\beta, \alpha^i \in C^\infty(M)). \tag{6.1}$$

Then the trajectories of the vector field f are corank 1 abnormal geodesics for the distribution $\mathcal{D}$.

Let us consider the distribution (free $C^\infty(M)$-module of vector fields)

$$V = \mathrm{span}\{f, g, [f, g], \ldots (\mathrm{ad}f)^{n-1}g\}$$

and assume that:

(iii) in the domain being considered we have

$$[[f, g]g]](q) \notin V(q).$$

Let ψ be a 1-form defined in the domain by the conditions

$$\psi \perp V, \ \psi[[f, g]g] = 1.$$

We shall derive the Jacobi equation for the abnormal geodesic, which corresponds to the vector field f. We denote by $\hat{q}(\cdot) = q^0 \circ e^{tf}$ the trajectory of f; starting at $q^0 = \hat{q}(0)$ $\hat{q}(T) = q^1$. Following the approach Section 4 we shall consider the skew-symmetric bilinear form $(v_1, v_2) \mapsto \psi \cdot [v_1, v_2](q^1), \ v_1, v_2 \in V$. Taking the quotient of V with respect to the kernel of this form, we obtain a $2(n+1)$-dimensional symplectic space Σ'. We reduce the symplectic space by taking the $(2n+1)$-dimensional skew-orthogonal complement of the canonical projection $\underline{f}$ of the vector field f onto Σ' and then, by taking its quotient with respect to span$\{\underline{f}\}$. The result is denoted by Σ^f; it is a $2n$-dimensional symplectic space with a skew-scalar product denoted by σ. We again denote by $\underline{Y}$ the image of the vector field $Y \in V$ under the canonical projection $V \to \Sigma^f$.

We are going to introduce special coordinates in Σ^f and derive one more representation of the Jacobi equation (4.22).

Let us set

$$Y_t = e^{(t-T)\mathrm{ad}f}g,\ Y_t^i = e^{(t-T)\mathrm{ad}f}(\mathrm{ad}f)^i g = \partial^i Y_t/\partial t^i,$$
$$\gamma_t^i = \hat{\psi}\cdot[Y_t^1, Y_t^i](q^1) = \hat{\psi}(t)[[f,g],(\mathrm{ad}f)^i g](\hat{q}(t))$$

for $i \geq 0$.

Returning to relation (6.1), we set $\alpha_t^i = \alpha^i(\hat{q}(t))$ $(i = 0,\dots n-1)$, $\beta_t = \beta(\hat{q}(t))$, and derive

$$Y_t^n(q^1) = \beta_t f(q^1) + \sum_{i=0}^{n-1}\alpha_t^i Y_t^i(q^1)$$

from (6.1)

Lemma 6.1

$$\underline{Y}_t^n = \sum_{i=0}^{n-1}\alpha_t^i \underline{Y}_t^i.\ \Box \tag{6.2}$$

Proof. Chosing coordinates in Σ^f as in (4.13) (with $k=1$) we only need to establish that

$$\partial(\psi\cdot Y_t^n)/\partial x|_{q^1} = \sum_{i=0}^{n-1}\alpha_t^i\partial(\psi\cdot Y_t^i)/\partial x|_{q^1} + \beta_t\partial(\psi\cdot f)/\partial x|_{q^1}$$

for the local coordinates $x = (x_1,\dots x_n)$ in the neighborhood of $q^1 \in M$. But this follows directly from (6.1) and the identities $(\psi Y_t^i)|_{q^1} \equiv 0,\ i = 0,\dots n-1.$ $\Box$

Let Π be the image under the canonical projection of the vector fields Y, which satisfy the condition $\psi\cdot[f,Y](q^1) = 0$ and vanish at q^1; Π is a Lagrangian plane in Σ^f. It follows from (6.1)-(6.2) that $\Sigma^f = \Pi \oplus \mathrm{span}\{\underline{Y}_t, t\in R\}$ and for any $\tau \in R$ the vectors $\underline{Y}_\tau, \underline{Y}_\tau^1 \dots \underline{Y}_\tau^{n-1}$ form a basis of the subspace span$\{\underline{Y}_t, t \in R\} = \Delta$. It should be emphasize, that the subspace Δ *is not* Lagrangian and σ defines a nondegenerate coupling between Π and Δ.

Representing $x \in \Sigma^f$ as $x = z + \xi$, where $z \in \Delta,\ \xi \in \Pi$, we can write the Jacobi equation (see (4.22)) in these coordinates as

$$\gamma_t^0(\dot{z}+\dot{\xi}) = \sigma(\underline{Y}_t^1, z+\xi)\underline{Y}_t^1$$

or

$$\gamma_t^0\dot{z} = \sigma(\underline{Y}_t^1, z)\underline{Y}_t^1 + \sigma(\underline{Y}_t^1,\xi)\underline{Y}_t^1,\ \dot{\xi} = 0. \tag{6.3}$$

Obviously, one of the solutions of this equation is $z_t \equiv \underline{Y}_t, \xi_t = 0$.

The point $\bar{t}$ is *a conjugate point of multiplicity* $k > 0$ *for the abnormal geodesic* $\hat{q}(t) = q^0 \circ e^{tf}$, if, for Eq. (6.3), the space of solutions which satisfy the boundary conditions

$$z_0 = 0,\ z_{\bar{t}} \parallel \underline{Y}_{\bar{t}},\ \sigma(\underline{g},\xi_0) = 0, \tag{6.4}$$

is k-dimensional.

Let us set $\zeta_t = \sigma(\underline{Y}_t, \xi_0)$ and present z_t in the form $z_t = \sum_{i=0}^{n-1} z_t^i \underline{Y}_t^i$. Then the equation (6.3) can be transformed into the following system

$$
\begin{gathered}
\gamma_t^0(\dot{z}_t^1 + \alpha_t^1 z_t^{n-1}) = \sum_{j=2}^{n-1} \gamma_t^j z_t^j + \dot{\zeta}_t, \ z_0^1 = 0, \\
\dot{z}_t^j + \alpha_t^j z_t^{n-1} = -z_t^{j-1}, \ z_0^j = 0, \ j = 2, \dots n-1, \\
\zeta^{(n)} = \sum_{i=0}^{n-1} \alpha_t^i \zeta_t^{(i)}, \ \zeta_0 = 0
\end{gathered}
\tag{6.5}
$$

(the equation for z^0, which does not appear neither in (6.4) nor in (6.5), is omitted)

The multiplicity of the conjugate point is equal to the dimension of the space of the solutions of system (6.5) which satisfy the conditions

$$
z_{\bar{t}}^i = 0, \ i = 1, \dots n-1. \tag{6.6}
$$

Summarizing the aforesaid, we can formulate the following theorem (compare with [6, Theorem 7.1]).

Theorem 6.1 *Suppose that conditions (i),(ii) and (iii) hold for the trajectory* $\hat{q}(t) = q^0 \circ e^{tf}$ *of the 2-dimensional distribution on an* $(n+2)$*-dimensional manifold starting at* q^0 *. Then:*

(1) $\hat{q}(t)$, $t \in [0, T]$, *is corank 1 abnormal geodesic path of the distribution;*

(2) *it has a finite number (which can be zero) of conjugate points* $\bar{t}_i$ *and the multiplicity of the conjugate point* $\bar{t}$ *is equal to the dimension of the space of solutions of system (6.5) which satisfy the boundary conditions (6.6);*

(3) *for the abnormal geodesic path to be strongly minimal, it is necessary (corr. sufficient) that* $(0, T)$ *(corr.* $(0, T]$*) does not contain conjugate points.* □

Proof. Statement (1) was established at the beginning of the section, and the finiteness of the set of conjugate points follows from the *strong regularity* (see [17]) of the abnormal geodesic path $\hat{q}(\cdot)$. Finally, for a strongly regular abnormal extremal (see the proof of Theorem 5) the absence of conjugate points on $[0, T)$ is necessary for the rigidity of every restriction of the corresponding geodesic path on $[0, t] \subset [0, T]$ and, therefore, by virtue of Theorem 5, is necessary for strong minimality. The sufficiency of the absence of conjugate points on $[0, T]$ for the strong minimality follows from Theorem 3. □

Example 6.1

Let us consider in greater detail the 2-dimensional distributions on 4-dimensional manifolds. Here the vector field f, which satisfies the condition (6.1), exists and is unique for any 2-dimensional distribution $\mathcal{D}$ of maximal growth; such distributions define the so-called *Engel structure* on 4-dimensional manifolds (readers can find in [8] a detailed survey of various problems connected with these structures). For $n = 2$ system (6.5) takes the form

$$
\dot{z}^1 = -\alpha_t^1 z^1 + \frac{\dot{\zeta}}{\gamma_t^0}, \ z_0^1 = 0,
$$

$$\ddot{\zeta} = \alpha_t^0 \zeta + \alpha_t^1 \dot{\zeta}, \ \zeta_0 = 0, \dot{\zeta}_0 = 1. \tag{6.7}$$

In addition,

$$\dot{\gamma}_t^0 = d(\psi[Y_t^1, Y_t])/dt = \psi[Y_t^2, Y_t] = \alpha_t^1 \gamma_t^0, \ \gamma_0^0 = 1.$$

Hence $\gamma_t^0 = e^{\int_0^t \alpha_\tau^1 d\tau}$, and therefore

$$z_t^1 = \int_0^t \frac{1}{\gamma_\tau^0} e^{-\int_\tau^t \alpha_\theta^1 d\theta} \dot{\zeta}_\tau d\tau = \zeta_t e^{-\int_0^t \alpha_\tau^1 d\tau}.$$

We can see that $\bar{t}$ is a conjugate point if and only if $\zeta_{\bar{t}} = 0$. The multiplicity of any conjugate point is 1.

Therefore, for the abnormal geodesic path to be strongly minimal it is necessary (corr. sufficient) that there are no zeros of the solution $\zeta(\cdot)$ of (6.7) on $(0, T)$ (corr. on $(0, T]$). □

Example 6.2.

Suppose now that $\mathcal{D}$ is a 2-dimensional distribution on a 3-dimensional Riemannian manifold M. Let us define

$$N = \{q \in M | \mathcal{D}_q = \mathcal{D}_q^2 \neq \mathcal{D}_q^3\}.$$

Then N is either empty or a smooth 2-dimensional submanifold of M.

In order to prove it, let us take any point of N and choose a basis $\{f, g\}$ of $\mathcal{D}$ in the neighborhood W of this point. We assume that $\mathcal{D}_q^2 \neq \mathcal{D}_q^3$ in W and choose local coordinates $q_i, \ i = 1, 2, 3$, in W. Then $N \cap W$ is determined by the equation $\Phi(q) = (f \wedge g \wedge [f, g])(q) = 0$ and we only have to establish that 0 is a regular value of Φ.

Since $\mathcal{D}_q^2 \neq \mathcal{D}_q^3$, it follows that there exists a vector field X subjected to $\mathcal{D}$ such that $[X, \mathcal{D}^2]_q \not\subseteq \mathcal{D}_q^2$; without loss of generality we can assume that $X = f$ and, changing the coordinates in W, transform f into a constant vector field. Then, obviously, $(f \circ \Phi)(q) \neq 0$ in W, i.e., every $q \in W$ is a regular point of Φ. At the same time we have proved that $f(q)$ is transversal to N at every point $q \in N$ and, hence, $\mathcal{D}_q|_N$ are transversal to $T_q N$. Actually $\mathcal{D}$ has the well-known Martinet canonical form in the neighborhood of $q \in N$.

The intersections $\mathcal{D}_q \cap T_q N$ define the 1-dimensional distribution on N and the equality $(f \wedge g \wedge [f, g])(q) = 0$ implies that the integral curves of this distribution are corank 1 abnormal geodesics in M. Note that they do not satisfy condition (i) of regularity for 2-distributions formulated at the beginning of this section.

To establish their strong minimality, we shall apply Theorem 3 and the Jacobi equation (4.22). In this case the symplectic space Σ^f is 2-dimensional and, therefore, any solution of (4.22) that satisfies the boundary condition (4.23) must vanish identically. Hence there are no conjugate points and every subarc of these abnormal geodesics is a strong minimizer. □

7 Proof of Theorem 1

The proof is based on the following result which is a modification of [6, Theorem 9.5].

Theorem 7.1 (Isolated Points at Critical Levels of Smooth Mappings: A Sufficient Condition) *Let $\mathcal{U}$ be a closed convex subset of the normed space X, which is densely embedded into a separable Hilbert space $H: \ X \hookrightarrow H$. Suppose that the mapping $F : X \to$*

R^m is twice Frechet differentiable at $\hat{x} \in \mathcal{U}$ and $\hat{x}$ is a critical point of F, i.e., $\lambda F'(\hat{x}) = 0$ for some $\lambda \in R^{m^*} \setminus 0$. We denote by $\mathcal{K}_{\hat{x}}\mathcal{U}$ a cone tangent to $\mathcal{U}$ at $\hat{x}$. Assume that:

$$(1)\ \|F(\hat{x}+x) - F(\hat{x}) - F'(\hat{x})x\| = o(1)\|x\|_H,\ \textit{as}\ \|x\|_X \to 0; \tag{7.1}$$

$$(2)\ \|\lambda F'(x)z\| = O(1)\|x\|_H, \textit{as}\ \|x\|_X \to 0; \textit{for an arbitrary}\ z \in X; \tag{7.2}$$

(3) *the quadratic form $\lambda F''(\hat{x})(\xi,\xi)$ admits a continuous extension from* $\ker F'(\hat{x})$ *to the completion of* $\ker F'(\hat{x})$ *in H and is H-positive definite on* $\ker F'(\hat{x}) \cap \mathcal{K}_{\hat{x}}\mathcal{U}$, *i.e.,*

$$\lambda F''(\hat{x})(\xi,\xi) \geq 2\gamma\|\xi\|_H^2,\quad \forall \xi \in (\ker F'(\hat{x}) \cap \mathcal{K}_{\hat{x}}\mathcal{U}); \tag{7.3}$$

$$4)\ \|\lambda(F(\hat{x}+\xi) - F(\hat{x})) - \frac{1}{2}\lambda F''(\hat{x})(\xi,\xi))\| = o(1)\|\xi\|_H^2,\ \textit{as}\ \|\xi\|_X \to 0, \tag{7.4}$$

for some $\gamma > 0$. Then $\hat{x}$ is an isolated point in X of the level set $F^{-1}(F(\hat{x})) \cap \mathcal{U}$. □

Remark. The cone $\mathcal{K}_{\hat{x}}\mathcal{U}$ tangent to the convex set $\mathcal{U}$ at $\hat{x}$ is a conic hull of $\mathcal{U} - \hat{x}$; obviously, $\mathcal{U} \subset \hat{x} + \mathcal{K}_{\hat{x}}\mathcal{U}$. □

Proof of Theorem 7.1. Without loss of generality we cany assume that $F(\hat{x}) = 0$ and $\hat{x}$ is the origin of X. We are going to establish that $\|F(x)\| \geq \mu\|x\|_H^2$ for some $\mu > 0$ and all x from some small neighborhood of the origin of X.

Let us take for Z a finite-dimensional complement of $\ker F'(0)$ in X; $F'(0)$ isomorphically maps Z onto the image $F'(0)X$ and

$$\|F'(0)z\| \geq c\|z\|\ \forall z \in Z \text{ for some } c > 0. \tag{7.5}$$

We defining $N = \{y \in R^m | \lambda \cdot y = 0\}$ and choose a vector $\eta \in R^m$ such that $\lambda \cdot \eta = 1$. Then $R^m = R\eta + N$ and $\mathrm{Im}F'(0) \subseteq N$.

If $x = z + \xi$, then, using the Hadamard lemma, we can present $F(x)$ as

$$F(x) = \Phi(\xi) + F'(0)z + A(x)z.$$

Here $A(x)z = \int_0^1 (F'(\xi + tz) - F'(0))z dt$. By virtue of (7.1),

$$\|\Phi(\xi) + A(x)z\| = o(1)(\|\xi\|_H + \|z\|),\ \text{as}\|x\|_X \to 0,$$

and by virtue of (7.2),

$$|\lambda A(x)z| = O(1)\|x\|_H\|z\|,\ \forall z \in Z,\ \text{as}\ \|x\|_X \to 0.$$

Let us consider the projections of $F(x)$ onto the vector η and the subspace N; they are $\lambda \cdot (\Phi(\xi) + A(x)z)\eta$ and $R(x) = F'(0)z + \Phi_N(\xi) + A_N(x)z$ correspondingly.

Fixing arbitrarily small $\epsilon > 0$ we may choose a small neighborhood V of X such that for $x \in V$ and a certain positive k we have

$$\|\Phi_N(\xi) + A_N(x)z\| \leq \epsilon\|x\|_H,\ |\lambda \cdot (\Phi(\xi) - \frac{1}{2}F''(0)(\xi,\xi) + A(x)z)| \leq k\|x\|_H\|z\| + \epsilon\|\xi\|_H^2.$$

It follows from (7.5) that

$$\|R(x)\| \geq \max(0, (c-\epsilon)\|z\| - \epsilon\|\xi\|_H),\ \forall x \in V,$$

$$|\lambda(\Phi(\xi)+A(x)z)| \geq \max(0,(\gamma-\epsilon)\|\xi\|_H^2 - k\|x\|_H\|z\|).$$

Settin $c=c-\epsilon,\ \gamma=\gamma-\epsilon$, we obtain

$$\|F(x)\| \geq a(\max(0,c\|z\|-\epsilon\|\xi\|_H)) + \max(0,\gamma\|\xi\|_H^2 - k\|x\|_H\|z\|)$$

for a certain $a>0$. Without loss of generality, we can assume that $k(1+4\epsilon/c)4\epsilon/c \leq \gamma/2$.

Now, if $c\|z\| \geq 4\epsilon\|\xi\|_H$, then

$$\|F(x)\| \geq a(\frac{c}{2}\|z\| + \epsilon\|\xi\|_H) \geq a\alpha(\epsilon,c)\|x\|_H^2$$

with $\alpha(\epsilon,c)>0$.

Otherwise, if $c\|z\| \leq 4\epsilon\|\xi\|_H$, then

$$\|F(x)\| \geq a(\gamma\|\xi\|_H^2 - k\|x\|_H\|z\|) \geq a(\gamma\|\xi\|_H^2 - \|\xi\|_H^2 k(1+4\epsilon/c)4\epsilon/c) \geq$$
$$\geq (a\gamma/2)\|\xi\|_H^2 \geq a\beta(\gamma,c,\epsilon)\|x\|_H^2$$

with $\beta(\gamma,c,\epsilon)>0$. □

Theorem 1 follows if we apply Theorem 7.1 to the end-point mapping

$$\begin{aligned} &F(v(\cdot),w(\cdot)) = \\ &= q^0\circ \overrightarrow{\exp}\int_0^T f(1+\dot v(\tau)) + g\dot w(\tau)d\tau = q^1\circ \overrightarrow{\exp}\int_0^T f\dot v(\tau) + Y_\tau \dot w(\tau)d\tau. \end{aligned} \tag{7.6}$$

Since our consideration is local, we fix local coordinates in the neighborhood of $q^1\in M$ and treat the input/state mapping F as a mapping into R^n.

We denote by $\mathcal{U}$ the set of Lipschitzian vector functions $u(\cdot)=(v(\cdot),w(\cdot))$ such that $u(0)=(v(0),w(0))=0$, and $(\dot v(\tau),\dot w(\tau))$ belongs to the set U from the statement of Theorem 1. Since U is convex, bounded and closed,it follows that $\mathcal{U}$ is a convex bounded closed subset of both $W^2_{1,\infty}[0,T]$ and $W^2_{1,1}[0,T]$; we choose the normed space X which is $W^2_{1,\infty}[0,T]$ equipped with the norm of $W^2_{1,1}[0,T]$. Introducing in X the scalar product $\langle u_1(t),u_2(t)\rangle = u_1(T)u_2(T)+\int_0^T u_1(\tau)u_2(\tau)d\tau$ and taking the completion of $W^2_{1,1}[0,T]$ with respect to the corresponding norm denoted by $\|\cdot\|_2$, we obtain a Hilbert space H, which can be identified with the Sobolev space $H^2_{-1}[0,T]$. Obviously isolation of $\hat u(\cdot)$ in $F^{-1}(q^1)\cap\mathcal{U}$ with respect to the metric of $W^2_{1,1}[0,T]$ is equivalent to the strong constrained rigidity, and, therefore, all that we need is to check whether the input/state mapping F satisfies the assumptions of Theorem 7.1.

First note that the input/state mapping F is smooth in $W_{1,1}[0,T]$ (see [3, 4]), and the abnormal extremal control $\hat u\equiv 0$, is, by definition, a critical point of F; $\lambda=\hat\psi_T$ annihilates $\operatorname{Im} F'(0)$. Then $\lambda F''(0)$ coincides with the second variation (4.17)-(4.15) along the abnormal extremal. Due to the absence of conjugate points on $[0,T]$ and the Strong Generalized Legendre Condition (3.12) the second variation (4.17) is positive definite and, hence, the condition (7.3) is fulfilled.

We have to verify estimates (7.1), (7.2) and (7.4) for the mapping (7.6). Since U is convex, we can always transform the basis f,g of $\mathcal{D}$ into $f,g+af$ in such a way that after the corresponding transformation of R^2 the set U will lie either in the left or in the right half-plane. Since the two cases are similar we choose the first one, i.e., from now on

$$(\dot v,\dot w)\in U \Rightarrow \dot v\leq 0.$$

We also assume that $(\dot v,\dot w)\in U \Rightarrow |\dot v|\le a$ and without loss of generality can take $t=T$ in the formulation of Theorem 1 since we can reduce the case of arbitrary $t\le T$ to this one by extending $(\dot v(\cdot),\dot w(\cdot))$ by means of zero from $[0,t]$ to $[0,T]$. Recall that $Y_\tau = e^{(\tau-T)\,\mathrm{ad}\, f}g$ and $\hat q(T)=q^0\circ e^{Tf}=q^1$.

Introducing the notation

$$\bar v(t)=v(t)-v_T=v(t)-\int_0^T \dot v(t)dt,\ w_T=\int_0^T \dot w(t)dt,$$

we apply successively the two variants of the generalized variational formula (2.5) to the chronological exponential $q^0\circ \overrightarrow{\exp}\int_0^T f(1+\dot v(t))+g\dot w(t)dt$ and obtain

$$\begin{aligned}
&q^0\circ \overrightarrow{\exp}\int_0^T \underbrace{f(1+\dot v(t))}_{perturbation}+g\dot w(t)dt=q^0\circ \overrightarrow{\exp}\int_0^T (1+\dot v(t))e^{w(t)\,\mathrm{ad}\, g}f dt\circ e^{w_T g}=\\
&=q^0\circ \overrightarrow{\exp}\int_0^T \underbrace{(1+\dot v(t))(e^{w(t)\,\mathrm{ad}\, g}f-f)}_{perturbation}+(1+\dot v(t))f dt\circ e^{w_T g}=\\
&=q^0\circ e^{Tf}\circ e^{v_T f}\circ \overrightarrow{\exp}\int_0^T (1+\dot v(t))(e^{\bar v(t)\,\mathrm{ad}\, f}e^{(t-T)\,\mathrm{ad}\, f}e^{w(t)\,\mathrm{ad}\, g}f-f)dt\circ e^{w_T g}=\\
&q^1\circ e^{v_T f}\circ \overrightarrow{\exp}\int_0^T (1+\dot v(t))(e^{\bar v(t)\,\mathrm{ad}\, f}e^{w(t)\,\mathrm{ad}\, Y_t}f-f)dt\circ e^{w_T g}=q^1\circ e^{v_T f}\circ\\
&\circ \overrightarrow{\exp}\int_0^T (1+\dot v(t))(w(t)[Y_t,f]+(w^2(t)/2)[Y_t,[Y_t,f]]+\bar v(t)w(t)[f,[Y_t,f]])dt\\
&\circ e^{w_T g}+o(1)(|v_T|^2+\|w(\cdot)\|_2^2),\ \text{ as } \|(\dot v(\cdot),\dot w(\cdot))\|_{L_1}\to 0
\end{aligned}$$

(recall that $|v_T|=\int_0^T|\dot v(\tau)|d\tau$, since $v(0)=0,$ and $\dot v(\tau)\le 0$). Choosing local coordinates in the neighborhood of $q^1\in M$ and using the Volterra expansions for the ordinary and chronological exponentials, we derive, from the last formula,

$$\begin{aligned}
&q^0\circ \overrightarrow{\exp}\int_0^T f(1+\dot v(t))+g\dot w(t)dt-q^1=\underbrace{(fv_T+gw_T+\int_0^T[Y_t,f]w(t)dt)(q^1)+}_{1}\\
&+\underbrace{(fv_T\circ\int_0^T[Y_t,f]w(t)dt+\int_0^T[Y_t,f]w(t)dt\circ gw_T+\int_0^T[Y_t,[Y_t,f]]\frac{w^2(t)}{2}dt+}_{2}\\
&\underbrace{+\frac12(v_T)^2(f\circ f)++\frac12(w_T)^2(g\circ g))+\int_0^T[f,[Y_t,f]]\bar v(t)w(t)dt+}_{2}\\
&\underbrace{+\int_0^T[Y_t,f]\dot v(t)w(t)dt+\int_0^T\int_0^t[Y_\tau,f]w(\tau)d\tau\circ[Y_t,f]w(t)dt)(q^1)+}_{2}\\
&+\underbrace{\int_0^T[Y_t,[Y_t,f]](q^1)\dot v(t)\frac{w^2(t)}{2}dt}_{3}+\underbrace{\int_0^T[f,[Y_t,f]](q^1)\dot v(t)\bar v(t)w(t)dt+}_{4}\\
&+o(1)(|v_T|^2+\|w(\cdot)\|_2^2)
\end{aligned}\tag{7.7}$$

(when establishing the estimate for the other term, we have used the estimate

$$\forall t \in [0,T]: \int_0^t |\dot v(\tau) w(\tau)| d\tau \le \|\dot v\|_{L_1}^{1/2} a^{1/2} \|w\|_{L_2}$$

which is valid when $|\dot v(\tau)| \le a$ on $[0,T]$). To establish estimate (7.1) for mapping (7.6), we must only note that, if the terms marked by "1" in relation (7.7) vanish, then the remainder is $O((\|\dot v\|_{L_1}^{\frac{1}{2}} + \|w\|_2)\|w\|_2)$ as $\|u(\cdot)\|_{L_1} \to 0$.

To find estimate (7.4) for mapping (7.6), let us first note that in the last relation the underbraced terms marked by "1" correspond to the first variation (first differential of F) whereas those marked by "2" correspond to the second differential; all the other forms of the second order are the terms which we have to estimate. The term marked by "4" is annihilated by $\hat\psi_T = \lambda$ and, hence, does not perturb estimate (7.4). The main obstacle is the "principal" third-order term marked by "3" , which, in general, *does not* admit estimate (7.4) (see [18]); in this particular case we achieve the result by using essentially the sign definiteness of $\dot v(\cdot)$.

Indeed, integrating the principal term by parts, we derive

$$\int_0^T [Y_t,[Y_t,f]](q^1)\frac{w^2(t)}{2}\underbrace{\dot v(t)dt} = [Y_T,[Y_T,f]](q^1)v_T\frac{w_T^2}{2} -$$
$$- \int_0^T v(t)\frac{w^2(t)}{2}\frac{d}{dt}[Y_t,[Y_t,f]](q^1)dt - \int_0^T [Y_t,[Y_t,f]](q^1)v(t)w(t)\dot w(t)dt.$$

The first two terms on the right-hand side admit an estimate $O(1)\|\dot v(\cdot)\|_{L_1}\|w(\cdot)\|_2^2$, and the last integral can be estimated from above as

$$|\int_0^T [Y_t,[Y_t,f]](q^1)v(t)w(t)\dot w(t)| \le \sup_{0\le t\le T}(\|[Y_t,[Y_t,f]](q^1)\||v(t)|)\int_0^T |w(t)||\dot w(t)|dt.$$

Since $v(t)$ is *a monotonically decreasing function* and $v(0)=0$, it follows that $\max_{0\le t\le T}|v(t)| = |v_T|$. Denoting $b = \sup_{0\le t\le T}\|[Y_t,[Y_t,f]](q^1)\|$ and applying the Cauchy-Schwartz inequality to the last integral, we find the upper estimate $b|v_T|\|w\|_2\|\dot w\|_{L_2}$ for the principal term, which is $o(1)(|v_T|^2 + \|w\|_2^2)$ as $\|\dot w\|_{L_1} \to 0$. Therefore we come to estimate (7.4) for mapping (7.6).

To find estimate (7.2), we have to compute the first differential of the input/state mapping (7.6) at the point $u^0(\cdot) = (v^0(\cdot), w^0(\cdot))$. Substituting $u^0(\cdot) + u(\cdot)$ for $u(\cdot)$ in relation (4.1), we obtain

$$F(u^0(\cdot)+u(\cdot)) = q^1 \circ \overrightarrow{\exp}\int_0^t (f(\dot v^0(\tau) + \dot v(\tau)) + Y_\tau(\dot w^0(\tau) + \dot w(\tau))d\tau.$$

Setting

$$X_t^0 = f\dot v^0(t) + Y_\tau \dot w^0(t),\ P_t^0 = \overrightarrow{\exp}\int_0^t X_\tau^0 d\tau$$

and applying the generalized variational formula (2.5), we obtain

$$F(u^0(\cdot)+u(\cdot)) = q^1 \circ P_T^0 \circ \overrightarrow{\exp}\int_0^T (X_t^v \dot v(t) + X_t^w \dot w(t))dt,$$

where

$$X_t^v = \mathrm{Ad} P_t^0 f,\ X_t^w = \mathrm{Ad} P_t^0 Y_t.$$

The first differential of F at $u^0(\cdot)$ is given by the relation

$$F'|_{u^0} = q^1 \circ P_T^0 \circ \int_0^T (X_t^v \dot{v}(t) + X_t^w \dot{w}(t))dt$$

and, in particular,

$$F'|_0 = q^1 \circ \int_{00}^T (f\dot{v}(t) + Y_t \dot{w}(t))dt.$$

We know that $\hat{\psi}_T F'|_0 = 0$ and therefore $\hat{\psi}_T F'|_{u^0} = \hat{\psi}_T F'|_{u^0} - \hat{\psi}_T F'|_0$.

We obtain

$$(F'|_{u^0} - F'|_0)u(\cdot) = q^1 \circ (P_T^0 \circ \int_0^T (X_t^v \dot{v}(t) + X_t^w \dot{w}(t))dt - \int_0^T (f\dot{v}(t) + Y_t \dot{w}(t))dt) =$$

$$q^1 \circ (P_T^0 \circ \int_0^T ((X_t^v - f)\dot{v}(t) + (X_t^w - Y_t)\dot{w}(t))dt + (P_T^0 - I) \circ \int_0^T (f\dot{v}(t) + Y_t \dot{w}(t))dt. \quad (7.8)$$

To find (7.2) we fix $(\dot{v}(t), \dot{w}(t)) \in \mathcal{U} \subset L_\infty$.

Let us estimate $(X_t^w - Y_t),\ (X_t^v - f)$. By definition,

$$|(X_t^w - Y_t)(q)| = |(\mathrm{Ad}\, P_t^0 - I)Y_t(q)| = |\int_0^t \mathrm{Ad}\, P_\tau^0 \,\mathrm{ad}(f\dot{v}^0(\tau) + Y_\tau \dot{w}^0(\tau))d\tau Y_t(q)| =$$

$$= |\,\mathrm{Ad}\, P_t^0 \circ \mathrm{ad}(fv^0(t) + Y_t w^0(t))Y_t(q) - \int_0^t (\mathrm{Ad}\, P_\tau^0 \dot{Y}_\tau w^0(\tau) +$$

$$+ \mathrm{Ad}\, P_\tau^0 \circ \mathrm{ad}\, X_\tau^0 (fv^0(\tau) + Y_\tau w^0(\tau)))d\tau Y_t(q)| \le C(|(v^0(t), w^0(t))| + \|(v^0(\cdot), w^0(\cdot))\|_{L_1}).$$

When deriving the last inequality, we have used the fact that for any $k \ge 0$ and a compact $K \subset R^n$ the diffeomorphisms P_t^0 and the vector fields X_t^0 and their derivatives in q of order $\le k$ are bounded on $[0,T] \times K$ by a constant depending on k, K and independent of $u^0(\cdot) \in \mathcal{U}$ (see [3]). Obviously, $(X_t^v - f)$ admits a similar estimate and, recalling that the values of$\dot{v}(t), \dot{w}(t)$ are bounded by the constant a, we infer that the first term on the right-hand side of (7.8) admits the estimate

$$O(1)\|(v^0(\cdot), w^0(\cdot))\|_{L_1} = O(1)\|(v^0(\cdot), w^0(\cdot))\|_2. \quad (7.9)$$

To estimate the second term we compute

$$|q^1 \circ P_T^0 - q^1| = \int_0^T q^1 \circ P_t^0 \circ (f\dot{v}^0(t) + Y_t \dot{w}^0(t))dt =$$

$$= q^1 \circ (P_T^0 \circ (fv^0(T) + Y_T w^0(T)) - \int_0^T (P_t^0 \circ X_t^0 \circ (fv^0(t) + Y_t w^0(t)) + P_t^0 \circ \dot{Y}_t w^0(t))dt =$$

$$= O(1)(|(v^0(T), w^0(T))| + \|(v^0(\cdot), w^0(\cdot))\|_{L_2}) = O(1)\|(v^0(\cdot), w^0(\cdot))\|_2$$

and, using again the boundedness of $\dot{v}(t), \dot{w}(t), \forall t$ we find estimate (7.9) for the second term of (7.8) and complete the proof of Theorem 1. □

References

[1] A.A. Agrachev, *Quadratic Mappings in Geometric Control Theory,* in: Itogi Nauki i Tekhniki; Problemy Geometrii, VINITI, Acad. Nauk SSSR, Moscow, Vol.20,pp.11-205,1988. English transl. in J.Soviet Math.,Vol.51,pp.2667-2734, 1990.

[2] A.A.Agrachev and R.V.Gamkrelidze, *Second-Order Optimality Condition for the Time-Optimal Problem,* Matem Sbornik, Vol.100,pp.610-643,1976. English transl. in: Math. USSR Sbornik,Vol.29,pp.547-576,1976.

[3] A.A.Agrachev and R.V.Gamkrelidze, *Exponential Representation of Flows and Chronological Calculus,* Matem. Sbornik, Vol.107, pp.467-532,1978. English transl. in: Math. USSR Sbornik, Vol.35, pp.727-785,1979.

[4] A.A. Agrachev, R.V. Gamkrelidze and A.V. Sarychev, *Local Invariants of Smooth Control Systems,* Acta Applicandae Mathematicae, Vol.14, pp.191-237, 1989.

[5] A.A.Agrachev and A.V. Sarychev, *On Abnormal Extremals for Lagrange Variational Problems,* to appear in J.Mathematical Systems, Estimation and Control.

[6] A.A.Agrachev and A.V. Sarychev, *Abnormal sub-Riemannian Geodesics: Morse Index and Rigidity,* submitted to 'Annales de l'Institut Henri Poincaré - Analyse non linéaire'.

[7] R. Bryant and L. Hsu, *Rigidity of Integral Curves of Rank 2 Distributions*, Inventiones Mathematicae, V. 114, pp. 435-461, 1993.

[8] V. Gershkovich, *Engel Structures on Four Dimensional Manifolds,* University of Melbourne, Department of Mathematics, Preprint Series No. 10, 1992.

[9] V.Guillemin and S.Sternberg, *Geometric Asymptotics,* Amer.Math.Soc., Providence, Rhode Island,1977.

[10] M.Hestenes, *Application of the Theory of Quadratic Forms in Hilbert Space to the Calculus of Variations,* Pacific J. Math, v.1, pp.525-582,1951.

[11] H.J.Kelley, R.Kopp and H.G. Moyer, *Singular Extremals,* in: G.Leitman Ed., Topics in Optimization, Academic Press, New York,N.Y.,pp.63-101,1967.

[12] A.J.Krener, *The High-Order Maximum Principle and its Applications to Singular Extremals,* SIAM J. on Control and Optimiz.,Vol.15,pp.256-293,1977.

[13] I.Kupka, *Abnormal Extremals,* preprint, 1992.

[14] C. Lobry, *Dynamical Polysystems and Control Theory,* in D.Q. Mayne and R.W. Brockett Eds., Geometric Methods in Systems Theory, Reidel, Dordrecht-Boston, pp.1-42,1973.

[15] R.Montgomery, *Geodesics, Which Do Not Satisfy Geodesic Equations*, Preprint, 1991.

[16] L.S.Pontryagin, V.G.Boltyanskii, R.V.Gamkrelidze and E.F.Mischenko, *The Mathematical Theory of Optimal Processes,* Pergamon Press, Oxford, 1964.

[17] A.V.Sarychev, *The index of the second variation of the control system*, Matemat. Sbornik, Vol.113,pp.464-486,1980. English transl. in: Math.USSR Sbornik, Vol.41,pp.383-401,1982.

[18] A.V.Sarychev, *On Legendre-Jacobi-Morse-type Theory of Second Variation for Optimal Control Problems*, Schwerpunktprogramm 'Anwendungsbezogene Optimierung und Steuerung' der Deutschen Forschungsgemeinschaft, Report No.382, Würzburg,1992.

[19] H.J.Sussmann, *A Cornucopia of Abnormal SubRiemannian Minimizers. Part I: the Four-Dimensional Case*, IMA Preprint Series #1073,1992.

[20] H.J.Sussmann, Wensheng Liu, *Shortest paths for sub-Riemannian metrics on rank-2 distributions,* Report SYCON-93-08, Rutgers Center for Systems and Control.