ESAIM: Control, Optimisation and Calculus of Variations July 1999, Vol. 4, p. 377–403
URL: http://www.emath.fr/cocv/

SUB-RIEMANNIAN METRICS: MINIMALITY OF ABNORMAL GEODESICS VERSUS SUBANALYTICITY

Andrei A. Agrachev[1,2] and Andrei V. Sarychev[3]

Abstract. We study sub-Riemannian (Carnot-Caratheodory) metrics defined by noninvolutive distributions on real-analytic Riemannian manifolds. We establish a connection between regularity properties of these metrics and the lack of length minimizing abnormal geodesics. Utilizing the results of the previous study of abnormal length minimizers accomplished by the authors in [Annales IHP. *Analyse nonlinéaire* **13**, p. 635-690] we describe in this paper two classes of the germs of distributions (called 2-generating and medium fat) such that the corresponding sub-Riemannian metrics are subanalytic. To characterize these classes of distributions we determine the dimensions of the manifolds on which generic germs of distributions of given rank are respectively 2-generating or medium fat.

Résumé. On étudie des métriques sous-Riemanniennes (des Carnot-Carathéodory) définies par les distributions non involutives sur les variétés Riemanniennes analytiques réelles. On établie la connexion entre les propriétés de la régularité de ces métriques et l'absence des géodésiques anormales de longueur minimale. En utilisant les résultats des études précédentes sur les minimiseurs anormaux accomplies par les auteurs dans [Annales IHP. *Analyse nonlinéaire* **13**, p. 635-690], on décrit dans cet article, pour certains types des germes, des distributions (appelées 2-générées et d'une croissance moyenne) telles que les métriques sous-Riemanniennes correspondantes sont sous-analytiques. Pour caractériser la classe de ces types des distributions, on détermine les dimensions des variétés sur lesquelles les germes génériques des distributions du rang donné sont 2-générées ou d'une croissance moyenne.

AMS Subject Classification. 58A30, 93B29.

Received October 23, 1998. Revised May 31, 1999.

Introduction

We study local properties of sub-Riemannian metrics on real-analytic manifolds.

Recall that a *sub-Riemannian structure* on a connected n-dimensional Riemannian manifold M is defined by means of a *distribution* $\mathcal{D}$ on M. Recall that distribution is a subbundle of TM. Below we assume $\mathcal{D}$ to be *bracket generating.* This means that for each point $q \in M$ there exists integer $\kappa_q \geq 1$ such that the values of the iterated Lie brackets of orders $\leq \kappa_q$ of the vector fields subject to $\mathcal{D}$ span the tangent space T_qM. The minimal integer $\kappa_q \geq 1$ is called *degree of nonholonomy* of the distribution $\mathcal{D}$ at the point q.

Keywords and phrases: Sub-Riemannian metrics, subanalitycity, abnormal length minimizers.

[1] Steklov Mathematical Institute, Russian Academy of Sciences, Gubkina ul. 8, 117966 Moscow, Russia; e-mail: andrei@agrachev.mian.su

[2] SISSA, via Beirut 2-4, 34014 Trieste, Italy; e-mail: agrachev@sissa.it

[3] Department of Mathematics, University of Aveiro 3810-193 Aveiro, Portugal; e-mail: ansar@mat.ua.pt

Admissible or *horizontal paths* are Lipschitzian curves $t \mapsto q(t)$ which are tangent to the distribution at (almost) each of their points: $\dot{q}(t) \in \mathcal{D}_{q(t)}$. As far as $\mathcal{D}$ is bracket generating then according to Rashevsky-Chow theorem (see Sect. 1.2) any two points $q^0, q^1 \in M$ can be connected by a Lipschitzian admissible path $q(t)$.

According to classical existence results of optimal control theory (Filippov theorem; see [15]) for any pair of points from a sufficiently small neighborhood of q^0 there exists (possibly non unique) path of minimal length in the set of admissible paths connecting these two points. Global existence can be asserted under standard assumption of completeness.

Attributing to each couple of points the infimum of the lengths of admissible paths we obtain so-called *sub-Riemannian* or *Carnot-Caratheodory metric* on the manifold M. The properties of the sub-Riemannian metrics and of corresponding length minimizers (*sub-Riemannian geodesics*) were intensively studied during the last decade not only for their own sake but also due to numerous applications to physics.

There are many similarities between Riemannian and sub-Riemannian metrics, and also some major differences: i) the "exponential maps" in the sub-Riemannian case have singularity at the origin q^0, since all the geodesics are necessarily tangent to a proper subspace $\mathcal{D}_{q^0} \subset T_{q^0}M$; ii) the conjugate locus of a point $q^0 \in M$ in the sub-Riemannian case may contain q^0 in its closure, and due to it, iii) arbitrarily short geodesics paths may fail to be length-minimizers, iv) small sub-Riemannian spheres may fail to be smooth, on the contrary to the case of small Riemannian spheres which are C^ω-diffeomorphic to Euclidean spheres.

In the sub-Riemannian case the set of admissible paths, connecting two given points, may fail to be a Banach manifold (on the contrary to the Riemannian case). It may have singularities which correspond to an interesting phenomenon discovered some years ago – an occurrence of so-called *abnormal sub-Riemannian length-minimizers* or *length minimizing abnormal sub-Riemannian geodesics*. The first example has been constructed by Montgomery [29], further examples, theoretical study of this phenomenon and bibliographic references can be found in [8–10, 27, 30, 35].

The abnormal sub-Riemannian geodesics do not depend on Riemannian metric and are determined by the distribution $\mathcal{D}$. As it has been just said they are singular "points" in the set of admissible paths on the manifold connecting two given points. They may exhibit some exotic properties, like for example be *rigid*, *i.e.* isolated in $W_{1,\infty}$-metric of Lipschitzian paths, in the above mentioned set. For an abnormal sub-Riemannian geodesic the adjoint covector in the Hamiltonian form of geodesic equation is orthogonal to the distribution $\mathcal{D}$ along the geodesic path.

It is clear that if one studies properties of sub-Riemannian metrics then some pathologies arise exactly when abnormal sub-Riemannian minimizers appear. On the contrary if there are no abnormal length-minimizers one can expect "more regularity" for the corresponding metrics.

To formulate more rigorously what we mean by regularity we have to recall that a sub-Riemannian ρ-sphere – the set of points located at (Carnot-Caratheodory) distance $\rho > 0$ from the given point q^0 – is a subset of the image of the *sub-Riemannian exponential map*. (The whole image – the set of end-points of sub-Riemannian geodesics of given length ρ – is called *wave front*.) This map is defined on a subset of T^*M and is real analytic. Therefore the "right" category of sets to which the sub-Riemannian spheres may belong is the one which includes images of analytic and semianalytic sets under real analytic maps.

This category is the one of *subanalytic sets* (see [16, 22, 24]), which are exactly the images of semianalytic sets under *proper* real analytic maps. The subanalytic sets form a wider class than the semianalytic ones, but still share with them an important regularity property – they admit *stratifications* [22, 24].

The subanalytic sets have been utilized in control theory since the end of 70's in order to construct optimal synthesis and to study the structure of reachable sets of control systems [12,34]. It has been established that even this large category is not enough to describe all control-theoretic objects. Thus Sussmann and Lojasiewicz Jr. constructed [28] examples of control systems with non subanalytic reachable sets.

The sub-Riemannian spheres *a priori* could be expected to exhibit more regularity as far as they are parts of reachable sets of *symmetric control systems*. They indeed are more regular when abnormal length-minimizers are lacking.

A good example of this kind is provided by a class of *fat* distributions. A distribution is called fat [13] at a point $q \in M$ if for any vector field X with $X(q) \in \mathcal{D}_q,\ X(q) \neq 0$, there holds

$$[X, \mathcal{D}](q) = \mathcal{D}(q) + \mathrm{span}\{[X, Y](q) :\ Y \in \mathcal{D}\} = T_q M.$$

It is easy to see that for a distribution which is fat at the point $q^0 \in M$, there are no abnormal sub-Riemannian geodesic paths starting at q^0. Sussmann [36] and Zhong Ge [18] studying the properties of sub-Riemannian fat metrics have established their subanalyticity.

On the other hand it is known that occurrence of abnormal length-minimizers may destroy the regularity. For example sub-Riemannian metric arising from *Martinet distribution* – first example in which abnormal length-minimizers have been discovered – fails to be subanalytic, as has been established by Agrachev *et al.* [4].

In the present paper we make the next step studying some classes of sub-Riemannian structures which do possess abnormal geodesics, but these latter are not length-minimizing. For some of these classes we manage to prove that the corresponding sub-Riemannian metrics are subanalytic.

Before introducing these classes of distributions let us recall that earlier in [9] we derived *necessary* conditions for rigidity and weak ($= W_{1,\infty}$-local) minimality of abnormal sub-Riemannian geodesics. These conditions are counterparts of optimality conditions for singular extremals known in optimal control theory since 60's. One of them called *Goh condition*, appeared in [9] as necessary condition for rigidity of abnormal sub-Riemannian geodesics (see [9], Prop. 4.3). For the sub-Riemannian length-minimization problems the Goh condition amounts to existence of a Hamiltonian multiplier which is orthogonal to the Lie square $\mathcal{D}^2(q) = (\mathcal{D} + [\mathcal{D}, \mathcal{D}])\,(q)$ of the distribution along the abnormal geodesic. In [9] it has been also stated (Prop. 4.2) that the Goh condition is necessary for weak minimality whenever abnormal geodesic path is *strictly abnormal*, *i.e.* when *all* possible Hamiltonian multipliers corresponding to this path annihilate (or are orthogonal to) the distribution $\mathcal{D}$. In the present paper we will obtain this necessary condition on the basis of a more general result which relates Morse index of a critical point of a smooth mapping to its local openness (Th. 3.4).

There are some classes of distributions for which abnormal geodesics do exist but can not satisfy the Goh condition. For example, it happens when the Lie square $\mathcal{D}^2(q) = (\mathcal{D} + [\mathcal{D}, \mathcal{D}])(q)$ of the distribution $\mathcal{D}$ at the point $q \in M$ coincides with the tangent space T_qM. In this case there may exist strictly abnormal geodesic paths starting at q but due to the Goh necessary condition no one of them is weak length minimizer. We shall call (following Sussmann) these distributions *2-generating* at the point $q \in M$. Obviously all fat distributions are 2-generating.

A larger class is formed by distributions $\mathcal{D}$, which we call *medium fat*. They satisfy the condition that for any vector field $X \in \mathcal{D},\ X(q) \neq 0$:

$$[X, \mathcal{D}^2](q) = \mathcal{D}^2(q) + \mathrm{span}\{[X, Y](q) :\ Y \in \mathcal{D}^2\} = T_q M.$$

The class of medium fat distributions is bigger than the one of 2-generating distributions. It is also bigger than the class of distributions with fat Lie square. Also for medium fat distributions we prove that there are no strictly abnormal length-minimizers (Th. 3.8).

It turns out that the same conditions which determine weak length minimality or non minimality of strictly abnormal geodesic paths also have crucial influence on the properties of the sub-Riemannian metrics.

The main results obtained in this paper include the following statements:

- if $\mathcal{D}$ is a germ at q^0 of a 2-generating distribution then the germ of corresponding sub-Riemannian (Carnot-Caratheodory) metric $\rho_{\mathcal{D}}(q, q')$ is subanalytic (Th. 4.2)[1];
- if $\mathcal{D}$ is a germ at q^0 of a medium fat distribution then the germ of corresponding sub-Riemannian (Carnot-Caratheodory) metric $\rho_{\mathcal{D}}(q, q')$ is subanalytic outside the diagonal $\{q = q'\}$ (Th. 4.1).

The proofs of these statements are based on a study of second variation along normal geodesics and on a *compactness* result for the set of length minimizers (of given length) for the medium fat and 2-generating

[1] This result has been independently obtained by Jacquet in his PhD thesis, Savoie, 1998.

distributions. Obviously the second statement refers to a bigger class of distributions, though there remains a problem of studying the metric at the diagonal.

The following statements describe how large are these two classes of distributions:

- generic germ of rank r distribution on n-dimensional manifold with $r \leq n \leq \frac{r(r+1)}{2}$ is 2-generating (Prop. 5.1);
- generic germ of rank r distribution on n-dimensional manifold with $r \leq n \leq r + (r-1)^2$ is medium fat (Th. 5.2).

The following tables illustrate these results representing information about (low) dimensions for which the properties of being 2-generating or medium fat are generic (evidently one assumes $\dim M \geq \dim \mathcal{D}$).

2-generating distributions

$\dim \mathcal{D}$	3	4	5	6	$\cdots$	r
$\dim M$	≤ 6	≤ 10	≤ 15	≤ 21	$\cdots$	$\leq r(r+1)/2$

Medium fat distributions

$\dim \mathcal{D}$	3	4	5	6	$\cdots$	r
$\dim M$	≤ 7	≤ 13	≤ 21	≤ 31	$\cdots$	$\leq r + (r-1)^2$

The present paper has the following structure. Section 1 contains some basic definitions concerning analytic, semianalytic and subanalytic sets, distributions on manifolds and formulae of chronological calculus for flows. These formulae will be needed for computation of first and second variations along a geodesic. In Section 2 we introduce the sub-Riemannian length-minimization problem and formulate the corresponding first-order necessary minimality condition – Hamiltonian form of geodesic equation. We also introduce normal and abnormal sub-Riemannian geodesics. In the Section 2.2 we introduce first and second variations for the length-minimization problem. In Section 3 we formulate some necessary conditions of weak minimality for sub-Riemannian geodesics in terms of their Morse index (Th. 3.3). These conditions are derived from a general result relating Morse index of a critical point of smooth mapping to stable local openness of this map at this point (Th. 3.4). The Goh condition appears then as necessary condition for finiteness of Morse index and therefore for weak minimality of strictly abnormal geodesic paths (Prop. 3.6 and Th. 3.7). Theorem 3.8 affirms absence of abnormal length-minimizers for medium fat and 2-generating distributions. Section 4 contains main results of the paper on subanalyticity: we prove subanalyticity of metrics for medium fat sub-Riemannian structures (Ths. 4.1 and 4.2). In Section 5 we evaluate the broadness of the classes of 2-generating and medium fat distributions (Prop. 5.1 and Th. 5.2). Section 6 contains major proofs.

1. Preliminaries

In what follows the manifold M is supposed to be connected.

1.1. Analytic, semianalytic and subanalytic sets

Recall that a subset S of a real analytic manifold M is called *analytic* if $\forall p \in S$ there exists a neighborhood U of p in M such that $U \cap S$ is defined by a finite system of equations $\varphi_i(x) = 0, i = 1, \ldots, m$, where $\varphi_i : M \to \mathbb{R}$ are real analytic functions.

A subset S of a real analytic manifold M is called *semianalytic* if $\forall p \in S$ there exists a neighborhood U of p in M such that $U \cap S$ is defined by a finite system of inequalities $\varphi_i(x) \geq 0, i = 1, \ldots, m, \ \chi_i(x) > 0, i = 1, \ldots, s$, where $\varphi_i : M \to \mathbb{R}, \chi_j : M \to \mathbb{R}$, are real analytic functions.

An extension of the class of semianalytic sets has been introduced [16,22,24] in order to recover an important property of "closeness under projection" which is valid for semialgebraic sets (Tarski-Seidenberg theorem) but may fail for the semianalytic ones. The images of semianalytic sets under analytic maps do not in general belong to the semianalytic category.

A subset S of a real analytic manifold M is called *subanalytic* if there exists a semianalytic subset T of a real analytic manifold N and a real analytic map $f : N \to M$, which is proper on clos$T$, and such that $f(T) = S$. Recall that a map f is *proper* if the inverse images $f^{-1}(K)$ of the compact sets K are compact.

Proper analytic maps map subanalytic sets into subanalytic ones. The class of subanalytic sets is also closed under the operations of taking finite unions, finite intersections, complements and under inverse images of analytic maps.

The subanalytic sets though being much more general than the semianalytic ones share with them an important regularity property – they are *stratifiable.*

Stratum is an analytic connected embedded submanifold of M. *Whitney stratification* $\mathcal{S}$ in M is a set of pairwise disjoint strata, such that:

i) $\mathcal{S}$ is locally finite, *i.e.* for any compact $K \subset M$ only a finite number of elements of $\mathcal{S}$ have nonempty intersections with K;
ii) for any $S \in \mathcal{S}$ the *frontier* $\mathrm{Fr}S = \mathrm{clos}S \setminus S$ is a union of elements of $\mathcal{S}$ of dimension strictly less than $\dim S$;
iii) for any pair of strata (S_α, S_β) such that $S_\alpha \subset \mathrm{clos}S_\beta$, the following *Whitney conditions* hold: suppose that sequences $x_i \in S_\beta$ and $y_i \in S_\alpha$ converge to one and the same point $y \in S_\alpha$, and (in some local coordinates on M) the secants $l_i = \overline{x_i y_i}$ converge to a straight line l and that the tangent spaces $T_{x_i}S_\beta$ converge to a subspace $\mathcal{T}$ then iv) $T_y S_\alpha \subset \mathcal{T}$ and v) $l \subset \mathcal{T}$.

The Whitney stratifications enjoy some nice properties; in particular, they admit triangulations.

$\mathcal{S}$ is a *stratification of* A, if $\cup_{S\in\mathcal{S}} = A$.

The following fact is fundamental for the subanalytic subsets: *any closed subanalytic subset admits Whitney stratification*; see [21] for details and references.

A map $f : M \to N$ between two C^ω-manifolds is called *subanalytic* if its graph is a subanalytic subset of $M \times N$.

1.2. Distributions and sub-Riemannian metrics

A *distribution* $\mathcal{D}$ on M is a subbundle of tangent bundle $\mathcal{T}M$. A vector field X is subject to $\mathcal{D}$ if $X(q) \in \mathcal{D}_q \subset \mathcal{T}_q M$ for every $q \in M$.

Generalizations of distributions are *differential systems* or *distributions with singularities*[2] for which $\dim \mathcal{D}_q$ may vary with $q \in M$. A differential system can be viewed as $\mathrm{C}^\infty(M)$-submodule of Vect$M$; then distributions correspond to projective $\mathrm{C}^\infty-$modules. Locally one may treat germ of distribution as a free module.

If $\mathcal{D}$ is a germ of differential system, then taking $\mathrm{C}^\infty-$modules generated by the Lie brackets of order $\leq k,\ k = 1, \dots,$ of the vector fields, subject to $\mathcal{D}$, one obtains 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

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

is called *flag of the differential system* $\mathcal{D}$ at the point $q \in M$, while the sequence $n_1(q) \leq \cdots \leq n_k(q) \leq \cdots$, where $n_i(q) = \dim \mathcal{D}_q^i$, is called *vector of growth of the differential system* $\mathcal{D}$ *at the point* q. Differential system is called *completely nonholonomic* or *having complete Lie rank* or *bracket generating* at a point $q \in M$ if $\mathcal{D}_q^\kappa = \mathcal{T}_q M$ for some κ. This κ is called *degree of nonholonomy of the distribution* at q.

A locally Lipschitzian path $\tau \mapsto q(\tau)$ on M is called *admissible* if its tangents belong to $\mathcal{D}$ for almost all $\tau \in [0, T]$.

A fundamental property of the bracket generating differential systems is established by the following theorem.

[2]Not to be mixed with the differential systems determined by the differential forms; those have different kind of singularities.

Theorem 1.1 (Rashevsky-Chow theorem; see [14, 31]). *If a differential system is bracket generating on the manifold M, then any two points of M can be connected by an admissible path.*

If one deals with distributions ($\dim \mathcal{D}_q = \text{const}$) then from existence results of optimal control theory (Filippov theorem) it follows (see [15]) that, locally, for any pair of points from a sufficiently small neighborhood of q^0 there exists (possibly non unique) path of minimal length in the set of admissible paths connecting these two points. Attributing to each couple of points the infimum of the lengths of admissible paths we obtain so-called *sub-Riemannian* or *Carnot-Caratheodory metric* on the manifold M.

One can define in an obvious way a sub-Riemannian ρ-sphere and sub-Riemannian ρ-ball centered at $q^0 \in M$ as a set of points located at sub-Riemannian distance ρ (respectively $\leq \rho$) from q^0.

The following parallelepiped-like (see [23, 37]) lower, or interior, estimate is valid for the sub-Riemannian balls. Let $\kappa > 1$ be the degree of nonholonomy of $\mathcal{D}$ at q^0 and $n_1(q^0) \leq \cdots \leq n_\kappa(q^0)$ its vector of growth at the same point. Let us define for $i = 1, \ldots, n,\ w_i = j$, if $n_{j-1} < i \leq n_j$. Consider the parallelepiped Π_ρ in $\mathbb{R}^n :\ \Pi_{c,\rho} = \{(x_1, \ldots, x_n)|\ |x_i| \leq c\rho^{w_i}\}$. The following statement is a corollary of the results of [7, 19, 23]: if B_ρ is sub-Riemannian ρ-ball centered at q^0, then there exists local coordinate system and constants $\rho_0 > 0, c, C > 0$ such that for all $\rho < \rho_0$: $\Pi_{c,\rho} \subset B_\rho \subset \Pi_{C,\rho}$. An easy corollary of this estimate is the fact that for some constant $c' > 0$ and for all $\rho < \rho_0$ the sub-Riemannian balls B_ρ contain the Riemannian balls of the radius $c'\rho^\kappa$ – Riemannian distance from q^0 a a sub-Riemannian ρ-sphere is estimated from below by $c'\rho^\kappa$.

Let us note that even if $\mathcal{D}$ is a distribution ($n_1(q) \equiv \text{const}$), then still $\mathcal{D}^k$ may have non constant dimensions – the growth vector of a distribution in general changes with q. A germ of distribution is called *regular* if its growth vector is constant. For regular sub-Riemannian metrics associated with regular distributions one can derive from the above mentioned estimates the following formula for the *Hausdorff dimension* of the sub-Riemannian manifold:

$$\dim_H M = n_1 + \sum_{i=2}^{\kappa} i(n_i - n_{i-1}).$$

Let us introduce some notions related to the growth vector of a germ of a distribution.

A distribution is called *fat* [13], if $\forall X \in \mathcal{D}$:

$$X(q) \neq 0 \Rightarrow [X, \mathcal{D}](q) = \mathcal{D}(q) + \text{span}\{[X, Y](q) :\ Y \in \mathcal{D}\} = T_q M.$$

A germ of distribution is called *2-generating* at a point $q \in M$ if its degree of nonholonomy $\kappa \leq 2$, *i.e.* $\mathcal{D}_q^2 = \mathcal{T}_q M$.

We call distribution *medium fat* at q, if $\forall X \in \mathcal{D}$:

$$[X, \mathcal{D}^2](q) = \mathcal{D}^2(q) + \text{span}\{[X, Y](q) :\ Y \in \mathcal{D}^2\} = T_q M.$$

The degree of nonholonomy of a medium fat distributions is ≤ 3.

Obviously this latter property is weaker than fatness and even weaker than the fatness of the Lie square $\mathcal{D}^2$. The class of fat distributions is contained in the class of 2-generating distributions which is contained in the class of distributions with fat Lie square which is contained in the class of medium fat distributions.

1.3. Exponential representation of flows: Some formulae

Here we introduce some notation and formulae of *chronological calculus* which will be utilized in the next section to obtain expressions for the first and the second variations along a geodesic. The chronological calculus has been developed in [6], one can find in [7] exposition of some part of it which is sufficient for our objectives.

A *flow* on M is an absolutely continuous (with respect to $\tau \in \mathbb{R}$) curve $\tau \longrightarrow P_\tau$ in the group of diffeomorphisms Diff M, satisfying the condition $P_0 = I$ – the identity diffeomorphism. We assume all time-dependent vector fields X_τ to be locally integrable with respect to τ, below they turn out to be real analytic. A time-dependent vector field X_τ defines an ordinary differential equation $\dot{q} = X_\tau(q(\tau)), q(0) = q^0$ on the manifold M. If the solutions of this differential equation exist for all $q^0 \in M, \tau \in R$, then the vector field X_τ is called *complete* and defines a flow on M, which is the unique solution of the (operator) differential equation:

$$dP_\tau/d\tau = P_\tau \circ X_\tau, P_0 = I. \tag{1.1}$$

This solution will be denoted by $P_t = \overrightarrow{\exp} \int_0^t X_\tau d\tau$, and is called (see [6] or [7]) 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 $P_t = e^{tX}$.

The chronological exponential admits Volterra expansion (see [6, 7]); we will need only its terms of zero-, first- and second-order:

$$\overrightarrow{\exp} \int_0^t X_\tau d\tau \asymp I + \int_0^t X_\tau d\tau + \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 (X_{\tau_2} \circ X_{\tau_1}) + \cdots \tag{1.2}$$

One more tool from the chronological calculus is a "generalized variational formula" (see [6, 7] for its drawing). In this formula we use the notation $\operatorname{Ad} P$ (where P is a diffeomorphism) for the following inner automorphism of the Lie algebra of vector fields on M: $\operatorname{Ad} PX = P \circ X \circ P^{-1} = P_*^{-1} X$. The last notation stands for the result of translation of the vector field X by the differential of the diffeomorphism P^{-1}.

The generalized variational formula can be written as

$$\overrightarrow{\exp} \int_0^t (\hat{X}_\tau + X_\tau) d\tau = \overrightarrow{\exp} \int_0^t \operatorname{Ad} \left(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta \right) X_\tau d\tau \circ \overrightarrow{\exp} \int_0^t \hat{X}_\tau d\tau. \tag{1.3}$$

Applying the operator $\operatorname{Ad} \left(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta \right)$ to a vector field Y and differentiating

$$\operatorname{Ad} \left(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta \right) Y = \left(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta \right) \circ Y \circ \left(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta \right)^{-1}$$

with respect to τ one comes to the equality (see [6, 7]):

$$\frac{d}{d\tau} \operatorname{Ad} \left(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta \right) Y = \operatorname{Ad} \left(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta \right) \circ \operatorname{ad} \hat{X}_\tau Y, \tag{1.4}$$

which is of the same form as (1.1). Therefore $\operatorname{Ad} \left(\overrightarrow{\exp} \int_0^\tau \hat{X}_\theta d\theta \right)$ can be represented as an operator chronological exponential $\overrightarrow{\exp} \int_0^t \operatorname{ad} \hat{X}_\theta d\theta$. For a time-independent vector field $\hat{X}_\tau \equiv \hat{X}$ it can be written as $e^{t \operatorname{ad} \hat{X}}$. These exponentials also admit Volterra expansions:

$$\overrightarrow{\exp} \int_0^t \operatorname{ad} X_\tau d\tau \asymp I + \int_0^t \operatorname{ad} X_\tau d\tau + \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 (\operatorname{ad} X_{\tau_2} \circ \operatorname{ad} X_{\tau_1}) + \cdots \tag{1.5}$$

In this new notation the generalized variational formula (1.3) can be represented as:

$$\overrightarrow{\exp} \int_0^t (\hat{X}_\tau + X_\tau) d\tau = \overrightarrow{\exp} \int_0^t \left(\overrightarrow{\exp} \int_0^\tau \operatorname{ad} \hat{X}_\theta d\theta \right) X_\tau d\tau \circ \overrightarrow{\exp} \int_0^t \hat{X}_\tau d\tau. \tag{1.6}$$

2. Sub-Riemannian problem of length minimization

2.1. **Normal and abnormal sub-Riemannian geodesics. Weak minimality and rigidity**

In this subsection we set a problem of sub-Riemannian length minimization (minimization of the Carnot-Caratheodory length), define sub-Riemannian geodesics as extremals of this minimization problem, introduce normal and abnormal geodesics and define rigidity.

From now on we deal with a real-analytic *sub-Riemannian structure* on a connected Riemannian manifold M. This structure is defined by means of a bracket generating distribution $\mathcal{D}$ on M. Given two points q^0 and q^1 we set a problem of finding admissible path of minimal length connecting q^0 with q^1. According to Rashevsky-Chow theorem the set of admissible paths connecting these two points is nonempty.

The problem we set is a generalization of the classical Riemannian length minimization problem, but in fact there are essential differences. Thus the space of *all* locally Lipschitzian paths, which connect q^0 and q^1, has natural structure of Banach manifold. Critical points of the length functional on this manifold are Riemannian geodesics and all paths of minimal length are among them. On the contrary the space of *admissible* paths, which connect q^0 and q^1, is not, in general, a manifold; it may have singularities. These singularities correspond to *abnormal* sub-Riemannian geodesics, which do not depend on Riemannian structure on M and are determined by the distribution $\mathcal{D}$.

The term "abnormal" came from optimization theory, since the problem of finding minimal admissible path can be formulated as a Lagrange problem of the Calculus of Variations. The *extremals* of this latter problem are called *sub-Riemannian geodesics* and, in particular, *abnormal extremals,* with vanishing Lagrange multiplier for the (length) functional, are called *abnormal sub-Riemannian geodesics*.

In spite of a lot of activity tended to elimination of abnormal sub-Riemannian geodesics from the candidates to minimizers, they can be minimal. Preprint [29] of Montgomery contains first example of strictly abnormal sub-Riemannian length-minimizer. Other examples and the corresponding theory of abnormal minimizers can be found in [8–10, 27, 30, 35].

The problem of finding minimal admissible path locally in a neighborhood of q^0 can be represented as the following Lagrange problem of the Calculus of Variations:

$$\begin{aligned} \ell_1(u(\cdot)) &= \int_0^1 \langle u(\tau), u(\tau)\rangle^{1/2} d\tau \longrightarrow \min, \\ \dot{q} &= G(q)u,\ q(0) = q^0,\ u \in \mathbb{R}^r, \qquad (2.1) \\ q(1) &= q^1. \qquad (2.2) \end{aligned}$$

Here $\langle \cdot, \cdot \rangle$ stays for the inner product in the tangent spaces T_qM; the "control parameter" u belongs to $\mathbb{R}^r$; the "controls" $u(\tau)$ are integrable; $G(q) = (g^1(q), \ldots, g^r(q))$ is a $r-$tuple of smooth vector fields, which form an orthonormal basis of the distribution $\mathcal{D}$.

It is easy to prove that the length-minimization problem is equivalent to the minimization of the functional

$$\ell(u(\cdot)) = \int_0^1 \langle u(\tau), u(\tau)\rangle d\tau \longrightarrow \min. \qquad (2.3)$$

Indeed by virtue of Cauchy-Schwartz inequality

$$\ell_1(u(\cdot)) = \int_0^1 \langle u(\tau), u(\tau)\rangle^{1/2} d\tau \leq \left(\int_0^1 \langle u(\tau), u(\tau)\rangle d\tau \right)^{1/2} = (\ell(u(\cdot)))^{1/2},$$

and the equality holds if and only if $\|u(\tau)\| \equiv$ const. As far as the functional of length ℓ_1 does *not* depend on a parametrization of a curve, we may always assume that $\|u(\tau)\| \equiv$ const for a chosen parametrization. This implies that the minimizers of the functional ℓ are the same as the minimizers of ℓ_1. At the same time we have established that minimizers of the problem (2.1–2.3) have constant magnitude. This magnitude is equal to the corresponding value of the functional ℓ_1 or, the same, to the sub-Riemannian length of $\hat{q}(\tau),\ \tau \in [0, 1]$.

To define $W_{1,\infty}$-local or *weak* minimality let us introduce first $W_{1,\infty}$-metric in the space of Lipschitzian paths $t \mapsto \hat{q}(t),\ (t \in [0, T])$ with fixed terminal points.

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 $\mathbb{R} \times M$ we can choose a basis $B_{t,q} : T_qM \to \mathbb{R}^n$ 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 $\mathbb{R}^n$-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 $B_{t,q(t)}\dot{q}(t)$. Then we can define a $W_{1,\infty}$-metric or norm in a small C^0-neighborhood of $\hat{q}(\cdot)$ in the space of paths with fixed terminal points by putting:

$$\|q^1(\cdot) - q^2(\cdot)\|_{1,\infty} = \text{vraisup}_{t\in[0,T]}\|\dot{q}^1(t) - \dot{q}^2(t)\|.$$

Definition 2.1. An 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 $W_{1,\infty}$-local or weak minimizer if, for some neighborhood of $\hat{q}(\cdot)|_{[0,T]}$ in $W_{1,\infty}[0,T]$, the points q^0 and q^1 can not be connected by a shorter admissible path $t \mapsto \bar{q}(t),\ t \in [0,T]$ belonging to this neighborhood.

Now we formulate first-order necessary weak minimality condition for the problem (2.1–2.3) – the *Pontryagin Maximum Principle* for the Lagrange problem of the Calculus of Variations.

Theorem 2.2 (Pontryagin Maximum Principle)**.** *If $\hat{u}(\cdot)$ is weak minimizer for the problem (2.1–2.3), i.e. corresponding trajectory $\hat{q}(\tau)$ ($\tau \in [0,1]$) of (2.1), satisfying the boundary conditions $\hat{q}(0) = q^0,\ \hat{q}(1) = q^1$, is $W^1_\infty-$locally minimal admissible path, then there exists a nonzero pair $(\hat{\psi}_0, \hat{\psi}(\cdot))$, where $\hat{\psi}_0$ is a nonpositive constant and $\hat{\psi}(\tau)$ is an absolutely continuous covector-function on $[0,1]$, such that $\hat{\psi}(\tau) \in \mathcal{T}^*_{\hat{q}(\tau)}M$ and the quadruple $(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}_0, \hat{\psi}(\cdot))$:*
1) satisfies Hamiltonian system associated with a Hamiltonian

$$H(u, q, \psi_0, \psi) = \hat{\psi}_0\langle u, u\rangle + \psi \cdot G(q)u; \tag{2.4}$$

in local coordinates on M this system has a form

$$\dot{q} = \partial H/\partial\psi, \tag{2.5}$$

$$\dot{\psi} = -\partial H/\partial q, \tag{2.6}$$

2) meets stationarity condition for almost all $\tau \in [0,1]$:

$$\left.\frac{\partial H}{\partial u}\right|_{(\hat{u}(\tau),\hat{q}(\tau),\hat{\psi}_0,\hat{\psi}(\tau))} = 2\hat{\psi}_0\hat{u}(\tau) + \hat{\psi}(\tau)G(\hat{q}(\tau)) = 0, \tag{2.7}$$

and for almost all $\tau \in [0,1]$

$$H(\hat{u}(\tau), \hat{q}(\tau), \hat{\psi}_0, \hat{\psi}(\tau)) = \hat{\psi}_0\langle\hat{u}(\tau), \hat{u}(\tau)\rangle + \hat{\psi}(\tau) \cdot G(\hat{q}(\tau))\hat{u}(\tau) = h_0 = \textit{const}. \tag{2.8}$$

Definition 2.3. An extremal of the Lagrange problem (2.1–2.3), *i.e.* a quadruple $\left(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}_0, \hat{\psi}(\cdot)\right)$ meeting the conditions of the Theorem 2.2 is called sub-Riemannian geodesic. Sub-Riemannian geodesic is called normal, if $\hat{\psi}_0 < 0$, and abnormal, if $\hat{\psi}_0 = 0$. The corresponding triple $(\hat{u}(\cdot), \hat{q}(\cdot))$ is called sub-Riemannian geodesic path.

Remark 2.4. It follows immediately from the relations (2.7–2.8) of this theorem, that $\hat{\psi}_0\langle\hat{u}(\tau), \hat{u}(\tau)\rangle = \text{const}$.

Remark 2.5. Obviously any restriction $\left(\hat{u}(\cdot)|_{[0,t]}, \hat{q}(\cdot)|_{[0,t]}, \hat{\psi}_0, \hat{\psi}(\cdot)|_{[0,t]}, t\right)$ of normal or abnormal sub-Riemannian geodesic $\left(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}_0, \hat{\psi}(\cdot)\right)$ onto a subinterval $[0,t] \subset [0,1]$ satisfies the equations (2.4–2.8) on $[0,t]$ and is also called normal or abnormal sub-Riemannian geodesic correspondingly.

Remark 2.6. A geodesic path $(\hat{u}(\cdot), \hat{q}(\cdot))$ may correspond to different geodesics with different $\hat{\psi}_0, \hat{\psi}(\cdot)$.

In normal case one may assume $\hat{\psi}_0 = -1/2$. Then (2.7) becomes

$$\hat{u}(\tau) = \hat{\psi}(\tau)G(\hat{q}(\tau)). \tag{2.9}$$

Definition 2.7. A corank of a geodesic path $(\hat{u}(\cdot), \hat{q}(\cdot))$ is the dimension of the space of pairs $(\hat{\psi}_0, \hat{\psi}(\cdot))$, which together with $(\hat{u}(\cdot), \hat{q}(\cdot))$ satisfy Theorem 2.2.

If a geodesic is abnormal, then the length functional ℓ does not appear in the minimality conditions, provided by the Theorem 2.2. No surprise that corresponding geodesic paths have not too much to do with the sub-Riemannian metric and the minimality of length. They rather correspond to some extremality of the distribution. It turns out that the abnormal geodesics "often" exhibit a phenomenon called in [38] *rigidity*.

Definition 2.8. An admissible path $q(\cdot)$ of the vector distribution $\mathcal{D}$ with end-points q^0 and q^1 is called rigid if it is isolated up to a reparametrization in the metric of $W_{1,\infty}$ in the set $\mathcal{P}_{q^0}^{q^1}$ of all admissible paths, which connect q^0 and q^1.

Rigid admissible paths are formally weakly minimal. Theorem 2.2 provides necessary conditions of weak minimality. When it is applied to rigid paths then in addition one should take $\hat{\psi}_0 = 0$. This implies that if an admissible path $(\hat{u}(\cdot), \hat{q}(\cdot))$ is rigid on $[0, 1]$, then $(\hat{u}(\cdot), \hat{q}(\cdot))$ is an abnormal geodesic path.

We dealt with the rigidity in [8,9]; the readers may consult that paper for detailed study of the phenomenon.

To finish with the first-order condition given by the Theorem 2.2 let us note that in the abnormal case the Hamiltonian (2.4) degenerates into an abnormal Hamiltonian

$$H = \psi \cdot G(q)u. \tag{2.10}$$

The stationarity condition (2.7) for an abnormal geodesic amounts to the orthogonality of $\hat{\psi}(\tau)$ to the distribution $\mathcal{D}$ at every point $\hat{q}(\tau)$:

$$\hat{\psi}(\tau) \cdot G(\hat{q}(\tau)) = 0, \ \forall \tau \in [0, 1]. \tag{2.11}$$

2.2. Second variation along normal and abnormal geodesics

In the previous subsection we have reduced the problem of finding minimal admissible path between given points q^0 and q^1, to the Lagrange problem (2.1–2.3). We have formulated first-order necessary condition of weak minimality and defined sub-Riemannian geodesics. The solutions of the length minimization problem should be sought among geodesic paths. In this Section we are going to introduce first and second variation along a geodesic $\left(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}_0, \hat{\psi}(\cdot)\right)$.

Let us introduce an *input/state map* $\Phi_t : \mathbb{R} \times L^r_\infty[0, 1] \longrightarrow M$, which maps an input $u(\cdot)$ into the point $q(t)$ of the trajectory $q(\cdot)$ of the system $\dot{q} = G(q)u(\tau),\ q(0) = q^0$. Obviously, $\Phi_t(\hat{u}(\cdot)) = \hat{q}(t)$ and $\Phi_1(\hat{u}(\cdot)) = \hat{q}(1) = q^1$.

Let $\hat{P}_t$ be the flow generated by the reference vector field $\hat{f}_t(q) = G(q)\hat{u}(t)$. Let us define a "pull-backed" input/state map $F_t(u(\cdot)) = \hat{P}_t^{-1}(\Phi_t(u(\cdot))$ and put $F(u(\cdot)) = F_1(u(\cdot))$. Obviously $F(\hat{u}(\cdot)) = q^0$.

We put

$$\ell(u(\cdot)) = \int_0^1 \langle u(\tau), u(\tau)\rangle d\tau; \ \ell : L^r_\infty[0, 1] \to \mathbb{R}.$$

A well known fact is that for $\hat{u}(\cdot) \in L^r_\infty$ to be a weak minimizer for the Lagrange problem (2.1–2.3) it must be a critical point of the map (ℓ, F). Indeed otherwise the Implicit Function Theorem implies existence (locally, in a L_∞-neighborhood of $\hat{u}(\cdot)$) of solutions of the system of equations

$$\ell(u(\cdot)) = \ell(\hat{u}(\cdot)) - \epsilon, \ F(u(\cdot)) = q^0,$$

for any sufficiently small $\epsilon > 0$. This means that q^0 and q^1 can be connected by an admissible path of length $\ell(\hat{u}(\cdot)) - \epsilon < \ell(\hat{u}(\cdot))$.

If $\hat{u}(\cdot)$ is a critical point for the map (ℓ, F), *i.e.* the differential $(\ell', F')|_{\hat{u}(\cdot)} : L^r_\infty \to \mathbb{R} \times \mathcal{T}_{q^0}M$ is not onto, then there exists a pair $(\hat{\psi}_0, \hat{\psi}_T) \in \mathbb{R} \times \mathcal{T}^*_{q^0}M$, which annihilates the image of $(\ell', F')|_{(\hat{u}(\cdot))}$:

$$\hat{\psi}_0 \ell' + \hat{\psi}_T F' \equiv 0. \tag{2.12}$$

This equality is equivalent to the statement of the Theorem 2.2 with $\hat{\psi}_T$ being the initial value $\hat{\psi}(0)$ for the solution of the adjoint equation (2.6). If $\hat{\psi}_0 = 0$, then the functional ℓ appears neither in (2.12) nor in the Theorem 2.2. In this case $\hat{u}(\cdot)$ is a critical point of the map F, or equivalently is an abnormal geodesic control corresponding to the geodesic $(\hat{u}(\cdot), \hat{q}(\cdot), 0, \hat{\psi}(\cdot))$.

To study the geodesics – the critical points of (ℓ, F) – we have to invoke initial terms of Taylor expansion for F. Let us present $F(u(\cdot))$ as a composition of chronological exponentials (see Sect. 2 for the notation):

$$\overrightarrow{\exp} \int_0^1 Gu(\tau)d\tau \circ \left(\overrightarrow{\exp} \int_0^1 G\hat{u}(\tau)d\tau \right)^{-1},$$

acting on the point q^0. Putting $u(\tau) = \hat{u}(\tau) + v(\tau)$ and applying the variational formula (1.3) to the first of the chronological exponentials we obtain

$$\overrightarrow{\exp} \int_0^1 G\left(\hat{u}(\tau) + v(\tau)\right) d\tau = \overrightarrow{\exp} \int_0^1 Y_\tau v(\tau) d\tau \circ \overrightarrow{\exp} \int_0^1 G\hat{u}(\tau)d\tau,$$

where

$$Y_\tau v = \mathrm{Ad}\, \overrightarrow{\exp} \int_0^\tau G\hat{u}(\xi)d\xi Gv.$$

Therefore

$$F(u(\cdot)) = \overrightarrow{\exp} \int_0^1 Y_\tau v(\tau)d\tau(q^0). \tag{2.13}$$

Invoking the Volterra expansion of the chronological exponential (2.13) we obtain the following expression for the first differential of F at $\hat{u}(\cdot)$:

$$F'|_{\hat{u}(\cdot)}(u(\cdot)) = \int_0^1 Y_\tau(q^0)u(\tau)d\tau. \tag{2.14}$$

If $\hat{u}(\cdot)$ is a critical point of (ℓ, F), then $\mathrm{Im}(\ell', F')|_{\hat{u}(\cdot)} \neq \mathbb{R} \times \mathcal{T}_{q^0}M$, and there exists a nonzero covector $(\hat{\psi}_0, \hat{\psi}_T) \in \mathcal{T}^*_{q^0}M$, which annihilates $\mathrm{Im}(\ell', F')|_{\hat{u}(\cdot)}$:

$$\int_0^1 \left(2\hat{\psi}_0 \hat{u}(\tau) + \hat{\psi}_T \cdot Y_\tau(q^0)\right) u(\tau)d\tau = 0,$$

for all $u(\cdot) \in L^r_\infty[0, 1]$. The latter equality implies:

$$2\hat{\psi}_0 \hat{u}(\tau) + \hat{\psi}_T \cdot Y_\tau(q^0) = 0, \text{ for almost all } \tau \in [0, T]. \tag{2.15}$$

This condition is equivalent to the condition (2.7) of the Theorem 2.2. Namely if one chooses the solution $\hat{\psi}(\cdot)$ of the adjoint equation (2.6) with the initial value $\hat{\psi}(0) = \hat{\psi}_T$, then $\hat{\psi}(\cdot)$ satisfies the stationarity condition (2.7). One defines *corank of the geodesic path* $(\hat{u}(\cdot), \hat{q}(\cdot))$ as the corank of $(\ell', F')|_{\hat{u}(\cdot)}$.

Definition 2.9. The first differential $(\ell', F')|_{\hat{u}(\cdot)} : L^r_\infty \to \mathcal{T}_{q^0}M$, at a critical "point" $\hat{u}(\cdot)$ is called first variation along the geodesic path $(\hat{u}(\cdot), \hat{q}(\cdot))$. It is computed according to the formula (2.14).

In what follows we assume either $\hat{\psi}_0 = -1/2$ or $\hat{\psi}_0 = 0$.

Now we introduce the *second variation* along a geodesic $\left(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}(\cdot)\right)$. It is the Hessian, or quadratic differential of (ℓ, F), at the critical point $\hat{u}(\cdot) \in L^r_\infty$ (see [11]). Choosing a function $\chi : M \longrightarrow \mathbb{R}$, such that $d\chi|_{q^0} = \hat{\psi}_T$, let us consider a function $\phi(u(\cdot)) = \hat{\psi}_0 \ell(u(\cdot)) + \chi(F(u(\cdot)))$. Since the pair $(\hat{\psi}_0, \hat{\psi}_T)$ annihilates $\mathrm{Im}(\ell', F')|_{\hat{u}(\cdot)}$, then $\hat{u}(\cdot)$ is a critical point for this function.

Let us compute the quadratic term of Taylor expansion for $\Phi(u(\cdot))$ at $\hat{u}(\cdot)$. Appealing to the Volterra expansion (1.2) for the right chronological exponential (2.13), we obtain

$$\phi''|_{\hat{u}(\cdot)}(u(\cdot)) = \hat{\psi}_0 \int_0^1 \langle u(\tau), u(\tau)\rangle d\tau + \left(\int_0^1 Y_\tau u(\tau) \circ \int_0^\tau Y_\xi u(\xi) d\xi d\tau \chi)(q^0) \right). \tag{2.16}$$

When restricting the quadratic form (2.16) to the kernel of $(\ell', F')|_{\hat{u}(\cdot)}$, determined by the equalities

$$\int_0^1 \hat{u}(\tau) u(\tau) d\tau = 0, \ \int_0^1 Y_\tau(q^0) u(\tau) d\tau = 0, \tag{2.17}$$

we are able to subtract from (2.16) a vanishing value of

$$\frac{1}{2} \left(\int_0^1 Y_\tau u(\tau) d\tau \right) \circ \left(\left(\int_0^1 Y_\tau u(\tau) d\tau \right) \chi \right) (q^0),$$

and transform (2.16) into

$$\hat{\psi}_0 \int_0^1 \langle u(\tau), u(\tau)\rangle d\tau + \frac{1}{2} \left(\left(\int_0^1 \left[\int_0^\tau Y_\xi u(\xi) d\xi, Y_\tau u(\tau) \right] d\tau \right) \chi \right) (q^0).$$

The last expression does not depend on a choice of χ but only on $\hat{\psi}_T = d\chi|_{q^0}$ and therefore we may introduce the following definition.

Definition 2.10. The second variation along the geodesic $\left(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}(\cdot)\right)$ is the quadratic form

$$Q[\hat{\psi}_0, \hat{\psi}_T](u(\cdot)) = \hat{\psi}_0 \int_0^1 \langle u(\tau), u(\tau)\rangle d\tau + \hat{\psi}_T \cdot \int_0^1 \left[\int_0^\tau Y_\xi u(\xi) d\xi, Y_\tau u(\tau) \right] (q^0) d\tau, \tag{2.18}$$

whose domain is subspace of L^r_∞ defined by the conditions (2.17), and either $\hat{\psi}_0 = -1/2$ (normal case) or $\hat{\psi}_0 = 0$ (abnormal case).

3. Morse index and necessary conditions for weak minimality

In this section we define an important invariant of a geodesic – *Morse index*. Then we establish its interrelation with weak minimality of geodesics and provide necessary condition for finiteness of index for abnormal geodesics.

Definition 3.1. Morse index of the geodesic $\left(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}_0, \hat{\psi}(\cdot)\right)$ is positive index of the quadratic form (2.18–2.17), *i.e.* maximal among the dimensions of the subspaces in its domain, on which the quadratic form is positive definite.

Definition 3.2. Morse index of the geodesic path $(\hat{u}(\cdot),\hat{q}(\cdot))$ is minimum of Morse indices of the geodesics $\left(\hat{u}(\cdot),\hat{q}(\cdot),\hat{\psi}_0,\hat{\psi}(\cdot)\right)$, which correspond to this geodesic path, or minimum of the positive indices of quadratic forms $Q|_{\hat{u}(\cdot)}[\hat{\psi}_0,\hat{\psi}_T]$ for all possible $(\hat{\psi}_0,\hat{\psi}_T)\perp \operatorname{Im} F'|_{\hat{u}(\cdot)}$.

We now formulate second-order necessary weak minimality condition for corank k geodesics paths.

Theorem 3.3 (Morse Index and Weak Minimality of Geodesic Paths)**.** *Morse index of a weakly minimal corank m geodesic path $(\hat{u}(\cdot),\hat{q}(\cdot))$ must be $\leq m-1$.*

This theorem is a corollary of a general result on stable local openness of a smooth map – Theorem 3.4, which provides sufficient condition for the stable local openness at a critical point of the map (stable local openness of a map at its regular point is corollary of the Implicit Function Theorem).

Let a map $G: X\to\mathbb{R}^k$ be twice Frechet differentiable at a point $\hat{x}$ which is critical point of G. Let us introduce the Hessian of G at the point $\hat{x}$, which is the quadratic map $G''(\hat{x}):\ker G'|_{\hat{x}}\to \operatorname{coker} G'|_{\hat{x}}$ (see [11]). One can represent it as a bundle of quadratic forms

$$\lambda\mapsto \lambda G''(\hat{x})(\xi,\xi),\ \xi\in\ker G'|_{\hat{x}},\ \lambda\perp \operatorname{Im} G'|_{\hat{x}}.$$

We define *index* and *nullity* of $(\hat{x},\lambda)$ as positive index and dimension of the kernel of the quadratic form $\lambda G''(\hat{x})(\xi,\xi)$. We define *index of critical point $\hat{x}$* as

$$\min\{\operatorname{ind}\lambda G''|\lambda\perp \operatorname{Im} G'|_{\hat{x}}\}\cdot$$

Theorem 3.4 (Morse Index and Stable Local Openness of a Smooth Map)**.** *Let X be a Banach space and $\hat{x}\in X$ be a corank m critical point for the map $G: X\to\mathbb{R}^k$, which is twice Frechet differentiable at $\hat{x}$. If the Morse index of $\hat{x}$ is $\geq m$, then the map G is locally open in a stable way i.e. for any neighborhood U of $\hat{x}\in X$ there exists a C^0-neighborhood $\mathcal{V}$ of G such that for any $\bar{G}\in\mathcal{V}$: $G(\hat{x})\in int\bar{G}(U)$.*

Some version of this result has been proved in [3]. As far as this source is not very accessible we provide in Section 6.1 a sketch of the proof.

Now basing on this result we provide

Proof of the Theorem 3.3. Let us consider the Banach space $X=L^r_\infty[0,1]$ and the map $G=(\ell,F)$: $L_\infty\longrightarrow\mathbb{R}\times M$, where ℓ,F are as defined above. If $\hat{u}(\cdot)$ is a geodesic control, *i.e.* a critical point of G of corank m, and the Morse index of this geodesic control is $\geq m$, then, according to the Theorem 3.4, $G(\hat{u}(\cdot))\in \operatorname{int} G(U)$ for any neighborhood U of $\hat{u}(\cdot)$ in L_∞. Therefore there exists $\bar{u}(\cdot)\in U$ such that $\ell(\bar{u}(\cdot))<\ell(\hat{u}(\cdot))$, $F(\bar{u}(\cdot))=F(\hat{u}(\cdot))=q^1$. *i.e.* the path generated by the control $\bar{u}(\cdot)$ connects q^0 with q^1 and is *strictly shorter* then the reference path. □

From the Theorem 3.3 it follows, that *finiteness of Morse index is necessary for weak minimality* of a geodesic path. There is a classical result which claims that *Morse index of normal geodesics is finite.* We are going to introduce a necessary condition for the finiteness of index for abnormal geodesics.

Definition 3.5. An abnormal geodesic satisfies Goh condition if for all $\tau\in[0,1]$

$$\hat{\psi}(\tau)\cdot[Gv,Gw](\hat{q}(\tau))=0,\ \forall v,w\in\mathbb{R}^r, \tag{3.1}$$

i.e. $\hat{\psi}(\tau)$ annihilates the Lie square $\mathcal{D}^2(\hat{q}(\tau))$ of the distribution $\mathcal{D}$ at every point of the geodesic path $\hat{q}(\cdot)$.

For singular extremals of optimal control problems this condition has been introduced by Goh in [20]. One can find in [9] the proof of the following result for abnormal geodesics corresponding to arbitrary measurable bounded geodesic control $\hat{u}(\cdot)$.

Proposition 3.6 (Goh Condition and Finiteness of Morse Index for Abnormal Geodesics)**.** *For index of an abnormal geodesic to be finite the Goh condition must hold along it.*

Now we are able to formulate a condition of weak minimality for *strictly* abnormal geodesic paths, *i.e.* such that all the corresponding geodesics are abnormal.

Theorem 3.7 (Necessity of Goh Condition for Abnormal Weak Minimality). *Let* $(\hat{u}(\cdot), \hat{q}(\cdot))$ *be strictly abnormal geodesic path. If this path is weak length minimizer then the Goh condition (3.1) must hold for some of the corresponding (abnormal) geodesics, i.e. for some choice of* $\hat{\psi}(\cdot)$.

Proof of the Theorem 3.7. It has been established in ([9] Appendix 1) (see also proof of the Lem. 6.7 below) that if Goh condition is not satisfied along an abnormal geodesic than Morse index of this geodesic is infinite. If all geodesics corresponding to the weakly minimal geodesic path $(\hat{u}(\cdot), \hat{q}(\cdot))$ are abnormal, then according to the Theorem 3.3 Morse index of at least one of them must be finite and therefore the Goh condition (3.1) must hold along this abnormal geodesic. □

An immediate corollary of this theorem is that *if a germ of distribution is* 2*-generating, i.e.* $\mathcal{D}^2(q^0) = T_{q^0}M$, *then there are no weakly minimal strictly abnormal geodesic paths starting at* q^0.

It turns out that the same is true for a bigger class of *medium fat* distributions.

Theorem 3.8 (Lack of Abnormal Weak Minimizers for Medium Fat Distributions). *If a germ of distribution is medium fat, i.e.*

$$\forall X \in \mathcal{D},\ X(q^0) \neq 0 :\ [X, \mathcal{D}^2](q^0) + \mathcal{D}^2(q^0) = T_{q^0}M,$$

then there are no weakly minimal strictly abnormal geodesic paths starting at q^0.

The proof of this result is provided in the Section 6.3.

4. Subanalyticity of sub-Riemannian metrics for medium fat and 2-generating distributions

It turns out that the same conditions which rule out existence of abnormal weak length minimizers for some classes of distributions guarantee subanalyticity of germs of sub-Riemannian metrics for these classes of distributions.

Theorem 4.1 (Subanalyticity of sub-Riemannian Metrics for Medium Fat Distributions). *Let* $\mathcal{D}$ *be a germ at* $q^0 \in M$ *of distribution which is medium fat at this point. Then small sub-Riemannian* ρ*-spheres* $\{q \in M|\ \rho_{\mathcal{D}}(q^0, q) = \rho > 0\}$ *are subanalytic subsets of* M *for all sufficiently small* $\rho > 0$. *Moreover the sub-Riemannian metric* $\rho_{\mathcal{D}}(q, q')$ *is subanalytic in a small neighborhood of* (q^0, q^0) *in* $M \times M$ *outside the diagonal* $\{(q, q)\}$.

One can say more if the distribution $\mathcal{D}$ is 2-generating.

Theorem 4.2 (Subanalyticity of sub-Riemannian Metrics for 2-generating Distributions). *Let* $\mathcal{D}$ *be a germ at* $q^0 \in M$ *of a distribution which is 2-generating at this point. Then the sub-Riemannian metric* $\rho_{\mathcal{D}}(q, q')$ *is subanalytic in a small neighborhood of* (q^0, q^0) *in* $M \times M$.

The crucial fact for establishing the subanalyticity is the one that the length-minimizing geodesics $\left(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}_0, \hat{\psi}(\cdot)\right)$ can be parametrized by a *compact* set.

Let us recall that according to the Pontryagin Maximum Principle (Th. 2.2) any normal (with $\hat{\psi}_0 = -1/2$) geodesic control $\hat{u}(\cdot)$ is computed as $\hat{u}(t) = \hat{\psi}(t)G(\hat{q}(t))$. Assume that all geodesics are parametrized by the interval $[0, 1]$. Then the constant value h of the Hamiltonian (2.4) along this geodesic equals $\|\hat{\psi}(t)G(\hat{q}(t))\|^2/2$. Hence $\hat{u}(\cdot)$ has constant magnitude $\|\hat{u}(t)\| \equiv \sqrt{2h}$ on $[0, 1]$ and the length of the geodesic path equals $\rho = \sqrt{2h}$ as well.

Obviously each geodesic is determined by the initial value $\hat{\psi}_T = \hat{\psi}(0)$ of $\hat{\psi}(\cdot)$. As far as for a geodesic of fixed length ρ one has $\|\hat{\psi}(t)G(\hat{q}(t))\| \equiv \rho$, then we may think that all initial values $\hat{\psi}_T$ have horizontal component of norm ρ. Thus the space of $\hat{\psi}_T$ can be identified with the cylinder $S^{r-1} \times \mathbb{R}^{n-r}$. Our goal is proving that *the length-minimizing geodesics are parametrized by $\hat{\psi}_T$ from a compact part K of the cylinder.*

We shall also need the asymptotic dependence of this compact part on the value of ρ, as $\rho \to +0$.

Theorem 4.3 (Compactness of the Set of Minimal Geodesics for 2-Generating and Medium Fat Distributions)**.** *Assume that the geodesics are parametrized by the interval $[0,1]$. Then for some $\rho_0 > 0$:*

i) *if $\mathcal{D}$ is a germ at q^0 of medium fat distribution then for any q from the sub-Riemannian ρ_0-neighborhood of q^0 the set of minimal geodesics, starting at q, of given length $\rho < \rho_0$ is compact (parametrized by a compact set of $\hat{\psi}_T$, the same for all q);*
ii) *if this germ is 2-generating then there exists $A > 0$ such that for all the length-minimizers of length $\rho < \rho_0$ there holds $\|\hat{\psi}_T\| \le A$;*
iii) *if the germ is medium fat then for any $\alpha > 3$ there exists $\bar{A} > 0$ such that for all the length-minimizers of positive length $\rho \le \rho_0$ there holds $\|\hat{\psi}_T\| \le \bar{A}\rho^{-\alpha-2}$.*

The subanalyticity claimed in the Theorems 4.1 and 4.2 follows from this latter theorem by a standard reasoning. We shall prove subanalyticity of the square of the sub-Riemannian length $\rho_{\mathcal{D}}^2$, which immediately implies the subanalyticity of the sub-Riemannian length $\rho_{\mathcal{D}}$.

Proof of the Theorems 4.1 and 4.2. Let us consider the sub-Riemannian exponential map $E(q, \hat{\psi}_T)$ which maps the pair $(q, \hat{\psi}_T)$ into the final point $\hat{q}(1)$ of the corresponding geodesic path; this map is obviously analytic. For given geodesic path defined by $(q, \hat{\psi}_T)$ the square of its length equals to $2h(q, \hat{\psi}_T)$ – the doubled value of the (analytic) Hamiltonian along this geodesic. Then we may calculate $\rho_{\mathcal{D}}^2(q, q')$ as

$$\rho_{\mathcal{D}}^2(q, q') = \min\{2h(q, \hat{\psi}_T) : \; E(q, \hat{\psi}_T) = q'\}\cdot$$

According to the Theorem 4.3 one can assume in the 2-generating case that $\hat{\psi}_T \in T^*_{q^0}M$ satisfies $\|\hat{\psi}_T\| \le A$ provided that $\rho_{\mathcal{D}}^2(q, q') \le \rho_0^2$. In the medium fat case if $0 < \delta^2 \le \rho_{\mathcal{D}}^2(q, q') \le \rho_0^2$, then $\|\hat{\psi}_T\| \le \bar{A}\delta^{-\alpha-2}$ for some $\alpha > 3$. In both cases $\hat{\psi}_T \in K,\ K$ – compact.

Consider the set of triples $(q, q', \hat{\psi}_T)$ from $M \times M \times K$ determined by the relation $E(q, \hat{\psi}_T) = q'$. This set is obviously subanalytic, and the image of this set under the analytic map $(q, q', \hat{\psi}_T) \mapsto \Big(q, q', \hat{\psi}_T, 2h(q, \hat{\psi}_T)\Big)$ is subanalytic as well. The image of the latter set under the projection $(q, q', \hat{\psi}_T, c) \mapsto (q, q', c)$ is subanalytic. Here is where the compactness of K plays crucial role[3]. This image is the epigraph of the function $\rho_{\mathcal{D}}^2(q, q')$. Therefore both this function and the sub-Riemannian metric are subanalytic. □

5. Generic 2-generating and medium fat distributions

In the previous section we established subanalyticity of sub-Riemannian metrics for the classes of 2-generating and medium fat distributions. These classes are substantially larger than the class of fat distributions. Here we discuss in what dimensions the properties of medium fatness or of being 2-generated are generic.

Let us start with the class of fat distributions for which subanalyticity of sub-Riemannian metrics has been established earlier by Sussmann [36] and Zhong Ge [18]. This class is relatively poor; if we restrict ourselves to generic distributions then it basically involves germs of corank 1 distributions on odd-dimensional manifolds.

Indeed if we choose a basis $X^1, \ldots, X^{2k}$ of the $2k$-distribution $\mathcal{D}$ on a $(2k+1)$-dimensional manifold, and a (unique up to multiplication by a scalar) covector $\psi \in T_q^*M$ which annihilates $\mathcal{D}_q$, then for generic $\mathcal{D}$ the *even-dimensional skew-symmetric matrix* $C = (c_{ij}^{\psi}) = \langle \psi, [X^j, X^i](q) \rangle$, $i, j = 1, \ldots, 2k$ is nonsingular. Hence for $X = \sum_{i=1}^{2k} X^i v_i$ the condition

$$[X, \mathcal{D}](q) \subset \mathcal{D}(q) \Leftrightarrow \langle \psi, [X, \mathcal{D}](q) \rangle = 0,$$

[3]Indeed let S be a subanalytic subset of $X \times Y$ (X, Y – C^ω-manifolds). Let $\pi : X \times Y \to X$ be the projection onto X. If for any compact $C \subset X$ there exists compact $C' \subset Y$ such that $\forall x \in C : \; \pi^{-1}(x) \subseteq C'$, then $\pi(S)$ is subanalytic.

is equivalent to the equality $Cv = 0$ where C is defined above and $v = (v_1, \dots, v_{2k})$. This equality is only satisfied when $v = 0$.

If one considers a corank 1 distribution $\mathcal{D}$ on odd-dimensional manifold, then repeating the same argument one comes to an odd-dimensional skew-symmetric matrix C, which is always singular and therefore always has a nontrivial vector $v = (v_1, \dots, v_n)$ in its kernel. Then $\langle \psi, [\sum_{i=1}^n X^i v_i, \mathcal{D}](q) \rangle = 0$ and hence the distribution $\mathcal{D}$ is *not* fat.

It is easy to see that in a class of distributions of given corank > 1 there exists open (in C^1-metric) subset consisting of nonfat distributions. Indeed let us take a pair of linearly-independent covectors ψ^1, ψ^2 from the annihilator $\mathcal{D}^\perp$ of a distribution and construct the matrices C^{ψ^1}, C^{ψ^2} as above. If there is a stable intersection at the "point" $C_0 = \alpha_0 C^{\psi^1} + (1-\alpha_0) C^{\psi^2}$ of the pencil $\alpha C^{\psi^1} + (1-\alpha) C^{\psi^2}$ $(\alpha \in R)$ with the cone $\{C : | \det C = 0\}$, then taking any vector v from the kernel of C_0, we construct, as above, the vector field $X = \sum_{i=1}^{2k} X^i v_i$ such that $(\mathcal{D} + [X, \mathcal{D}])(q)$ is annihilated by the covector $\alpha_0 \psi^1 + (1-\alpha_0)\psi^2$.

Still in some classes of germs of corank $c > 1$ distributions there are open subsets consisting of fat distributions. The corresponding pairs (c, n) (with n being a dimension of manifold) have been characterized by Rayner ([32]; see also [30]).

The class of 2-generating distributions is essentially bigger.

It is known (see [37], and Sect. 1.2) that for a generic germ of r-distribution on n-dimensional manifold the initial segment of the vector of growth $n_1 \le n_2 \le \cdots$ is:

$$n_1 = r, n_2 = \min\{n, r(r+1)/2\}, \dots$$

We are going to prove this result for the manifolds of dimension $n \le r(r+1)/2$.

Let us take any germ $\mathcal{D}$ of distribution at $q \in M$ and introduce an adapted coordinate system $(x, y) : \mathcal{O} \to \mathbb{R}^n$ such that:

i) $(x, y)(q) = 0_{R^n}$;
ii) $(x, y)_*|_q f_i(q)$ $(i = 1, \dots, r)$ coincide with the first r elements $e_i, i = 1, \dots, r$, of the standard basis of $\mathbb{R}^n$;
iii) the subspace $\Delta_q^2 = \text{span}\{\mathcal{D}_q, [f^i, f^j](q), |i < j\}$, is mapped isomorphically by the differential $x_*|_q$ onto a subspace of $\mathbb{R}^n$, while $y_*|_q \Delta_q^2 = 0$.

Denote by $\bar{\Gamma}$ a minimal subset of $\Gamma = \{(ij)|i < j\}$ such that Δ_q^2 is spanned by $\mathcal{D}_q$ and $\{[f^i, f^j](q), | \ (ij) \in \bar{\Gamma}\}$. Let $y = (y_1, \dots, y_s)$; $s = \text{codim } \Delta_q^2$. Obviously

$$s = \text{corank}\, \mathcal{D}_q - \sharp\bar{\Gamma} \le r(r-1)/2 - \sharp\bar{\Gamma} = \sharp(\Gamma \setminus \bar{\Gamma}).$$

Let us choose an arbitrary injection $\sigma : \{1, \dots, s\} \to \Gamma \setminus \bar{\Gamma}$. Denote by π the projection $(ij) \mapsto i$ defined on Γ and introduce vector fields

$$\tilde{f}_j = f_j + \sum_{\alpha=1}^{s} \varepsilon_\alpha x_{\pi(\sigma(\alpha))} \partial/\partial y_\alpha, \ (\varepsilon_\alpha \in \mathbb{R}, \ \alpha = 1, \dots, s).$$

For these vector field we establish by direct computation that

$$[\tilde{f}_i, \tilde{f}_j](q) = \begin{cases} [f_i, f_j](q), \ \text{if } (ij) \notin \sigma(\{1, \dots, s\}), \text{ in particular if } (ij) \in \bar{\Gamma}, \\ [f_i, f_j](q) + \varepsilon_{\sigma^{-1}(ij)} (\partial/\partial y_{\sigma^{-1}(ij)})|_q \text{ otherwise.} \end{cases}$$

Thus we constructed a distribution $\tilde{\mathcal{D}}$ generated by $\tilde{f}^1, \dots, \tilde{f}^r$ which is 2-generating and is arbitrarily close to $\mathcal{D}$ in C^∞ topology if $|\varepsilon_\alpha|, \ \alpha = 1, \dots, s$ are sufficiently small.

Therefore we may formulate the following conclusion.

Theorem 5.1 (Generic 2-Generating Distributions)**.** *Generic germ of rank r distribution on n-dimensional manifold with $r \le n \le \frac{r(r+1)}{2}$ is 2-generating.*

The class of generic medium fat distributions is yet bigger.

Theorem 5.2 (Generic Medium Fat Distributions)**.** *Generic germ of rank r distribution on n-dimensional manifold with $r \leq n \leq r + (r-1)^2$ is medium fat.*

The proof of the latter result is more technically involved; it can be found in Section 6.4.

6. Proofs

6.1. Theorem 3.4 on stable local openness: Sketch of the proof

We assume without loss of generality, that $\hat{x}$ coincides with the origin of X. We denote by D and h, correspondingly, the differential and the Hessian of the map F at the origin. Suppose, that for any $\lambda \in (\operatorname{Im} D)^{\perp}$ index of the quadratic form λh on $\ker D$ is $\geq m$. We are going to prove that then F is locally open at $\hat{x}$ and this property is stable, *i.e.* persists under small perturbations of F in C^0-metric.

For any $w \in \mathbb{R}^k$ the equation $F(x) = w$ can be represented as a system of equations $f(y,z) = u,\ g(y,z) = v$, where $(y,z) = x, (u,v) = w$ are such splittings of x and w, that: i) u coordinatizes ImD; ii) v cordinatizes a complementary space to ImD; iii) z coordinatizes $\ker D$, and $\partial F/\partial z = 0$; iv) $\dim f = \dim y = \dim u = \operatorname{rank} D = k - m,\ \operatorname{rank} \partial F/\partial y|_0 = \operatorname{rank} \partial f/\partial y|_0 = \operatorname{rank} D$.

By virtue of Implicit Function Theorem the equation $f(y,z) = u$ can be resolved w.r.t. $y:\ y = y(z,u)$. The function $y(z,u)$ is continuous locally and $y(0,0) = 0$. Substituting $y(z,u)$ into the equation $g(y,z) = v$ we obtain an equation $g(y(z,u),z) = 0$. Obviously local openness of F at the origin is equivalent to local openness of the map $(u,z) \mapsto (u, g(y(z,u),z))$. To establish this local openness it is enough to prove that the map $\varphi(z) = g(y(z,0),z)$ covers the origin of $\mathbb{R}^k$ *substantially*, *i.e.* for any closed neighborhood U of the origin in X there exists $\delta > 0$ such that

$$\|\bar{\varphi} - \varphi\|_{C^0(U)} < \delta \Rightarrow 0 \in \operatorname{int} \bar{\varphi}(U).$$

We are dealing now with the map φ, whose differential vanishes at the origin: $\varphi'(0) = 0$. To avoid additional notation we will assume, instead of it, that $D = F'|_0 = 0$, and then $h = F''|_0$ is a quadratic map of X into $\mathbb{R}^m$. Again we assume, that for any nonzero $\lambda \in \mathbb{R}^{m^*}$ index of the quadratic form λh is $\geq m$.

First we get rid of infinite-dimensional space X, proving that under the conditions of the Theorem 3.4 there exists a finite-dimensional subspace $W \subset X$, such that for any nonzero $\lambda \in \mathbb{R}^{m^*}$ index of the quadratic form $\lambda h|_W$ is $\geq m$.

Indeed for any unit covector $\bar{\lambda} \in \mathbb{R}^{m^*}$, there exists a $m-$dimensional subspace $W_{\bar{\lambda}} \subset X$, such that the restriction $\bar{\lambda} h|_{W_{\bar{\lambda}}}$ is positive definite. For all λ's from some small neighborhood $\Omega_{\bar{\lambda}}$ of $\bar{\lambda}$ the quadratic forms $\lambda h|_{W_{\bar{\lambda}}}$ are also positive definite. Choosing a finite covering of the sphere $\|\lambda\| = 1$ by corresponding neighborhoods $\Omega_{\bar{\lambda}^1}, \ldots \Omega_{\bar{\lambda}^s}$ we may take $W = W_{\bar{\lambda}^1} + \cdots W_{\bar{\lambda}^s}$.

From now on we consider W in place of X or, all the same, assume $\dim X < \infty$.

The following statement enables us to study the quadratic map h instead of F.

Lemma 6.1. *If the cone $h^{-1}(0)$ contains a regular point of the quadratic map $h : X \to \mathbb{R}^m$, then F is locally open in a stable way.*

Proof. If $y \in h^{-1}(0)$ is a regular point of h, then there exists a m-dimensional subspace $Z \subset X$, such that $h|_{y+Z} : (y + Z) \to \mathbb{R}^m$ is local diffeomorphism at y. Since h is homogeneous, then the same holds for all points $\eta y,\ \eta \neq 0$.

Consider the map $\phi_\epsilon(z) = h(y + \epsilon z)$, where z belongs to the unit sphere $S^{m-1} \subset Z$. Obviously $h(\eta y + \eta\epsilon z) = \eta^2 \phi_\epsilon(z)$ and, for all sufficiently small $\epsilon > 0$, the topological degree of the map $\phi_\epsilon/\|\phi_\epsilon\| : S^{m-1} \to S^{m-1}$ is $+1$ or -1. Since the differentials of h at the points ηy are nondegenerate, then $\exists a > 0$, such that for small enough $\epsilon > 0$ and $\forall z \in S^{n-1}$: $\|h(\eta y + \eta\epsilon z)\| \geq a\eta^2\epsilon$. Also for any w with $\|w\| < a\eta^2\epsilon$ there exist ζ with $\|\zeta\| \leq 1$ such that $h(\eta y + \eta\epsilon\zeta) = w$. On the other side

$$\|F(\eta y + \eta\epsilon z) - h(\eta y + \eta\epsilon z)\| = o(\eta^2), \text{ as } \eta \to +0,$$

and therefore for some $\epsilon > 0$ and small enough $\eta > 0$ the topological degree of the map

$$z \longrightarrow F(\eta y + \eta\epsilon z)/\|F(\eta y + \eta\epsilon z)\|$$

is $+1$ or -1. Hence one can choose small enough $\eta > 0$ such that say for $\|w\| < (\epsilon/2)a\eta^2$ the equation $F(\eta y + \eta\epsilon z) = w$ has a solution $z_\eta(w)$ belonging to the unit ball $B \subset Z$. As far as we apply the topological degree argument the same will be true for any map $\bar{F}$ which is sufficiently close to F in C^0-metric. □

This Lemma allows us to deal with the quadratic map h instead of F. The conclusion of the Theorem 3.4 follows from the following

Lemma 6.2. *Let $P : X \to \mathbb{R}^m$ be quadratic map, such that $\mathrm{ind}\,\lambda P \geq m,\ \forall \lambda \in \mathbb{R}^{m^*} \setminus 0$. Then $P^{-1}(0)$ contains regular point of the map P.*

Proof of this Lemma can be found in ([9] Appendix 1, Lem. 9).

6.2. Proof of the Theorem 4.3 on compactness of the set of minimal geodesics of given length

Proof. Assume all length-minimizers, we consider, to be of positive length $\rho < 1$. It will be convenient for us to parametrize them by the length of arc. We assume that the vector fields $g^i(q),\ i = 1, \ldots, r$, form an orthonormal basis in T_qM for all q from a small neighborhood of q^0. Hence the corresponding minimizing controls have constant magnitude 1.

As we already know length-minimizing geodesic paths of medium fat sub-Riemannian structures are not strictly abnormal. (The proof of this fact can be found in the next subsection.) We choose a fixed normal geodesic path of length $\rho > 0$ and normal pair $(-1/2, \hat{\psi}(\cdot))$ of Hamiltonian multipliers and consider (multiplied by 2) second variation along the corresponding geodesic. This is a quadratic form (compare with (2.18–2.17)):

$$Q(u(\cdot)) = -\int_0^\rho \langle u(\tau), u(\tau)\rangle d\tau + \hat{\psi}_T \int_0^\rho \left[\int_0^\tau Y_\xi u(\xi)d\xi, Y_\tau u(\tau)\right](q^0)d\tau, \tag{6.1}$$

defined on the set of $u(\cdot)$'s which satisfy

$$\int_0^\rho \hat{u}(\tau)u(\tau)d\tau = 0,\ \int_0^\rho Y_\tau(q^0)u(\tau)d\tau = 0. \tag{6.2}$$

Recall that $\hat{\psi}_T$ coincides with the initial value $\hat{\psi}(0)$ of the adjoint covector function $\hat{\psi}(t)$ and

$$Y_\tau = (Y_\tau^1 \ldots Y_\tau^r),\ Y_\tau^i = \mathrm{Ad}\,\overrightarrow{\exp} \int_0^\tau G\hat{u}(\xi)d\xi g^i,\ i = 1, \ldots, r.$$

Due to real analyticity of any normal extremal control $\hat{u}(\cdot)$ the vector fields Y_τ^i are real analytic. For any small compact neighborhood K of q^0 their C^1-norms $\|Y_\tau^i\|_{1,K}$ have a uniform (with respect to $\hat{u}(\cdot)$ and $\tau \in [0, \rho]$) upper estimate B.

Our first goal is establishing that for given $\rho > 0$

$\exists B$ such that if $\|\hat{\psi}_T\| \geq B$, then the corresponding geodesic path fails to be a length-minimizer.

We shall also estimate the asymptotic dependence of B on ρ as $\rho \to +0$.

The proof is rather technical and before proceeding with it we shall provide its sketch in very general terms.

Let us first introduce the covector-functions $\psi^1(t), \psi^2(t)$ with the components:

$$\begin{aligned} \psi_i^1(t) &= \hat{\psi}(t)g^i(\hat{q}(t)), && i = 1, \ldots, r; \\ \psi_{ij}^2(t) &= \hat{\psi}(t)[g^i, g^j](\hat{q}(t)), && i, j = 1, \ldots, r. \end{aligned} \tag{6.3}$$

According to the Pontryagin Maximum Principle (Th. 2.2) $\hat{u}(t) = \hat{\psi}(t)G(\hat{q}(t)) = \psi^1(t)$ and hence $\|\psi^1(t)\| = \|\hat{u}(t)\| \equiv 1$. Also $\psi^2_{ij}(t) = \hat{\psi}_T[Y^i_t, Y^j_t](q^0)$.

Consider the second variation $Q(u(\cdot))$, defined by (6.1–6.2) for the $u(\cdot)$'s which vanish outside a small interval $[t_0, t_0+\Delta] \subset [0,\rho)$. Assume also for the moment that $\psi^2(t)$ "does not oscillate too fast", *e.g.* there exists an estimate $\mathrm{Var}\psi^2(\cdot) \le C\|\psi^2(\cdot)\|_{L_1}$. Then one may substitute the second addend of (6.1) by its principal term

$$\hat{\psi}_T \int_{t_0}^{t_0+\Delta} \left[Y_{t_0} \int_{t_0}^{\tau} u(\xi)d\xi, Y_{t_0} u(\tau) \right] d\tau. \tag{6.4}$$

which depends (linearly) only on $\psi^2(t_0)$.

Choosing sufficiently small $\delta \in (0,1)$ and taking $\Delta = \delta\rho$ in the 2-generating case and $\Delta = \delta\rho^\alpha$ $(\alpha > 3)$, in the medium fat case, we manage to prove that in both cases there exists a pair of linear subspaces L^+_N, L^-_N of arbitrarily large dimension N such that the quadratic form (6.4) is positive definite on one L^+_N and negative definite on L^-_N. On L^+_N it admits a lower estimate b_N and on L^-_N it admits an upper estimate $-b_N$. These estimates are $b_N = A_N\rho^{2\alpha+3}\|\hat{\psi}_T\|\|u(\cdot)\|^2_{L_\infty}$ in the medium fat case and $b_N = A_N\rho^2\|\hat{\psi}_T\|\|u(\cdot)\|^2_{L_\infty}$ in the 2-generating case.

The first addend of (6.1) has order of smallness $O(\rho)$ as $\rho \to 0$. Therefore the second addend of (6.1) will dominate the first addend for sufficiently small $\rho > 0$ if $\|\hat{\psi}_T\| > C\rho^{-1}$ in the 2-generating case, or in the medium fat case $\|\hat{\psi}_T\| > C\rho^{-\alpha-3}$ with $C > 0$ (depending on N) sufficiently large. This will imply that the second variation has both positive and negative indices large enough (greater than corank!) *for all the geodesics with $\hat{\psi}_T$ satisfying the latter inequalities* and hence according to the Theorem 3.4 all these geodesics fail to be length minimizers.

Remark 6.3. When talking about indices we may ignore the conditions (6.2) , as far as their imposition merely diminishes the value of indices at most by $n+1$ $(n = \dim M)$.

The above reasoning is valid provided that $\psi^2(t)$ "does not oscillate too fast". This assumption holds (uniformly) for all geodesics in the 2-generating case. However in the medium fat case it may fail. Indeed differentiating $\psi^1(t), \psi^2(t)$ we obtain

$$\dot{\psi}^1_i(t) = \hat{\psi}(t)[G\hat{u}(t), g^i](\hat{q}(t)),\ \dot{\psi}^2_{ij}(t) = \hat{\psi}(t)[G\hat{u}(t), [g^i, g^j]](\hat{q}(t)), i, j = 1, \dots, r, \tag{6.5}$$

that shows that $\psi^2(t)$ can be "fast oscillating" if $\|\dot{\psi}^2\| \gg \|\psi^2\|$ and $\hat{u}(t)$ is fast oscillating. This may happen if the projection of $\hat{\psi}_T$ onto the directions orthogonal to $\mathcal{D}^2$ is large and if the derivative of $\hat{u}(t) = \psi^1(t)$ defined by the first of the equation (6.5) is large. A more thorough analysis, which will not be used in the further proof shows that there exists a family of geodesics with $\|\psi^2(0)\| \sim N$, $\|\hat{\psi}_T\| \sim N^2$ for which $\mathrm{Var}\psi^2(t)/\|\psi^2(\cdot)\|_{L_1} \sim N$, as $N \to \infty$. Fortunately for large N these geodesics can not attain the ρ-sphere ($\rho > 0$), because "too fast oscillating" geodesic controls can not drive the point from the origin far enough. So there is no need to consider them.

We start now the detailed proof of the Theorem 4.3 with the explanation of what is meant by fast oscillation and why too fast oscillating controls do not correspond to length minimizers. After it we proceed to the more technically involved part of arranging estimates for the second variation (Lem. 6.7).

To evaluate the rate of oscillation we introduce the following norm, which is useful for studying fast oscillating and relaxed controls.

Lemma 6.4. *The expression*

$$\|u(\cdot)\|_o = \sup_{t',t'' \in [0,\rho]} \left\| \int_{t'}^{t''} u(\tau)d\tau \right\|$$

defines a norm in the space of integrable functions (the corresponding topology is obviously weaker then the one of $L^r_1[0,\rho]$).

The subindex "o" stays for "oscillating". Indeed fast oscillating controls are small in this norm. For example $\|\sin Nt\|_o$ as well as $\|N\sin N^2t\|_o$ are $O(N^{-1})$ as $N\to\infty$. One can find in [17] examples of using this norm, *e.g.* a theorem on continuous, with respect to this norm, dependence of solutions of ODE on right-hand side. What follows is a much more particular result which proves that the fast oscillating controls "do not drive far". This result is local and we provide its formulation for a germ at $0_{\mathbb{R}^n}$ of distribution $\mathcal{D}=\operatorname{span}\{G(x)u \mid u\in\mathbb{R}^r\}$ in $\mathbb{R}^n$ in place of M.

Lemma 6.5 (compare with Chap. 4 of [17])**.** *For any sufficiently small $\rho_0>0$ there exists a constant $C>0$, such that for any $\rho<\rho_0$ and for all $u(\cdot)\in L^r_\infty[0,\rho]$ such that $\|u(\cdot)\|_{L_\infty}\leq 1$, the corresponding trajectories of the distribution satisfy*

$$\dot{x}(t)=G(x(t))u(t),\ x(0)=0\ \Rightarrow x(\rho)\leq C\|u(\cdot)\|_o. \tag{6.6}$$

It has been already mentioned in Section 1.2 that for a sub-Riemannian structure determined by a distribution of degree of nonholonomy κ the Riemannian distance from the center q^0 to the sub-Riemannian ρ-sphere admits (for $\rho<\rho_0$) a lower estimate $\underline{c}\rho^\kappa$ with $\underline{c}>0$ depending continuously on q^0.

Therefore coming back to the germ of distribution $\mathcal{D}$ at $q^0\in M$ of degree of nonholonomy κ we conclude with the following

Corollary 6.6. *For some $\rho_0>0$ there exists a constant $c>0$ such that $\forall\rho<\rho_0$ and for any control $u(\cdot)$ leading from q^0 to a point of sub-Riemannian ρ-sphere there holds $\|u(\cdot)\|_o\geq c\rho^\kappa$.*

Proof of the Lemma 6.5. Let $\|G(0)\|=L_0$ and $L\geq 0$ be a Lipschitz constant for the restriction of $G(x)$ on some compact neighborhood of 0. If $x(t)$ stays inside this compact neighborhood then transforming the differential equation (6.6) into an integral one:

$$x(t)=\int_0^t G(x(\tau))u(\tau)d\tau \tag{6.7}$$

we conclude

$$\|x(t)\|\leq\int_0^t\|G(x(\tau))u(\tau)\|d\tau\leq\int_0^t(L\|x(\tau)\|+L_0)d\tau,$$

and using the Gronwall inequality obtain

$$\|x(t)\|\leq L_0L^{-1}(e^{Lt}-1),\ \text{if}\ L>0,\ \text{and}\ \|x(t)\|\leq L_0t,\ \text{if}\ L=0.$$

Proceeding with the integration by parts in (6.7) we obtain

$$x(t)=G(x(t))\int_0^t u(\tau)d\tau-\int_0^t\left((G(x(\tau))u(\tau))\circ G\right)(x(\tau))\int_0^\tau u(\xi)d\xi d\tau.$$

As long as $\sup_\tau|\int_0^\tau u(\xi)d\xi|\leq\|u(\cdot)\|_o$ and $\|u(\cdot)\|_{L_\infty}\leq 1$, then $\|x(t)\|\leq C\|u(\cdot)\|_o$, where the constant C is calculated *via* $\|G(\cdot)\|_{C^1}$. ☐

The rest of our proof is based on the following technical result.

Lemma 6.7. *i) If the distribution $\mathcal{D}$ is 2-generating at a point $q^0\in M$, then there exist constants $\delta\in(0,1/2),A>0,\rho_0>0,C>0$ such that for all geodesics of length $\rho\leq\rho_0$:*

$$\|\hat{\psi}_T\|\geq A\Rightarrow\|\psi^2(t_0)\|\geq C\delta\|\hat{\psi}_T\|$$

at some point t_0 of the interval $[0,\rho]$;

ii) if the distribution $\mathcal{D}$ is medium fat at a point $q \in M$, then there exist constants $\delta \in (0, 1/2), A > 0, \rho_0 > 0$ such that for all geodesics $(\hat{u}(\cdot)|_{[0,\rho]}, \hat{q}(\cdot)|_{[0,\rho]}, \psi_0, \hat{\psi}(\cdot)|_{[0,\rho]})$, whose end-point $\hat{q}(\rho)$ belongs to the ρ-sphere ($0 < \rho < \rho_0$), there holds:

$$\|\hat{\psi}_T\| \geq A \Rightarrow \|\psi^2(t_0)\| \geq C\delta\rho^3\|\hat{\psi}_T\|, \tag{6.8}$$

at some point t_0 of the interval $[0, \rho]$.

Remark 6.8. In both statements the point t_0 may depend on geodesic, but the constants k, δ, ρ_0, C can be chosen in a uniform way.

Proof of the Lemma 6.7. Proof of i) is immediate. Indeed, if $\mathcal{D}$ is 2-generating, then the vector fields g^i, $[g^i, g^j]$, $i, j \in \{1, \dots, r\}$, span the space T_qM and therefore $\|\hat{\psi}(t)\| \leq \mu(\|\psi^1(t)\| + \|\psi^2(t)\|)$. Then

$$A \leq \|\hat{\psi}_T\| = \|\hat{\psi}(0)\| \leq \mu(\|\psi^1(0)\| + \|\psi^2(0)\|),$$

and

$$\|\psi^2(0)\| \geq \mu^{-1}\|\hat{\psi}_T\| - 1 \geq (\mu^{-1} - 1/A)\|\hat{\psi}_T\|.$$

Proof of ii). Recall that we deal with a geodesic of positive length $\rho \leq \rho_0$ (some conditions on ρ_0 will be imposed later). Introduce the covector-function $\psi^3(t)$:

$$\psi^3_{ijk}(t) = \hat{\psi}(t)[g^i, [g^j, g^k]](\hat{q}(t)), i, j, k = 1, \dots, r.$$

As long as the degree of nonholonomy of medium fat sub-Riemannian distributions is ≤ 3, then

$$\|\hat{\psi}(t)\| \leq \mu(\|\psi^1(t)\| + \|\psi^2(t)\| + \|\psi^3(t)\|).$$

Differentiating the functions $\psi^1(t), \psi^2(t), \psi^3(t)$ with respect to t we obtain

$$\dot{\psi}^3_{ijk}(t) = \hat{\psi}(t)[G\hat{u}(t), [g^i, [g^j, g^k]]](\hat{q}(t)), i, j, k = 1, \dots, r.$$

Using the notation $\psi(t)$ for the triple $(\psi^1(t), \psi^2(t), \psi^3(t))$ we conclude that for geodesics of length $\rho < \rho_0$ there exists a constant $B > 0$ depending only on the distribution $\mathcal{D}$ (and on ρ_0) such that

$$\|\dot{\psi}(t)\| \leq B\|\psi(t)\|.$$

Using the Gronwall inequality we obtain $\forall \bar{t} \geq \underline{t}$:

$$\|\psi(\bar{t})\| \leq \|\psi(\underline{t})\|e^{B(\bar{t}-\underline{t})}, \tag{6.9}$$

and also

$$\|\psi(\bar{t}) - \psi(\underline{t})\| \leq \|\psi(\underline{t})\| \int_{\underline{t}}^{\bar{t}} Be^{B(\tau-\underline{t})}d\tau = \|\psi(\underline{t})\|(e^{B(\bar{t}-\underline{t})} - 1), \ \|\psi(\bar{t})\| \geq \|\psi(\underline{t})\|(1 - (e^{B(\bar{t}-\underline{t})} - 1)). \tag{6.10}$$

If $\rho > 0$ is fixed and $t, t_0 \in [0, \rho]$, then we may derive from the inequalities (6.9–6.10) (increasing B if necessary) the estimate:

$$\|\psi(t)\| \geq \|\psi(t_0)\|(1 - B|t - t_0|).$$

In a similar way we derive $\forall \bar{t} \geq \underline{t}$:

$$\|\psi_2(\bar{t}) - \psi_2(\underline{t})\| \leq \|\psi(\bar{t})\| B(\bar{t} - \underline{t}), \tag{6.11}$$
$$\|\psi_2(\bar{t})\| \geq \|\psi_2(\underline{t})\| - \|\psi(\underline{t})\| B(\bar{t} - \underline{t})) \tag{6.12}$$

with (probably different) constant B depending on $\mathcal{D}$ but neither depending on geodesic nor on $\rho < \rho_0$. In what follows we assume $B\rho_0 < 1$.

As far as medium fat distribution has degree of nonholonomy (at most) 3, then according to the Corollary 6.6 any geodesic control $\hat{u}(\cdot)$, which drives us from q^0 to a point of ρ-sphere, must satisfy the condition $\|\hat{u}(\cdot)\|_o \geq c\rho^3$, that means that for some $t_0, t_1 \in [0, \rho]$: $\| \int_{t_0}^{t_1} \hat{u}(\tau)d\tau\| \geq c\rho^3$. Obviously $t_1 - t_0 > c\rho^3$.

Coming back to the proof of the Lemma 6.7 let us choose some $\delta \in (0, 1/2)$; some additional restrictions on δ will be imposed later.

Take the point t_0. If
$$\|\psi^2(t_0)\| \geq B\delta\rho^3\|\psi(t_0)\| \geq \mu^{-1}(1 - B\rho_0)B\delta\rho^3\|\psi_T\|,$$
then the statement ii) of the lemma is proved.

Let now $\|\psi^2(t_0)\| \leq B\delta\rho^3\|\psi(t_0)\|$. Without loss of generality one may assume (provided A in (6.8) is big enough)
$$\|\psi^3(t_0)\| \geq 2^{-1}\|\psi(t_0)\|.$$
From the second one of the differential equations (6.5) we derive

$$\begin{aligned}\psi^2_{ij}(t_1) - \psi^2_{ij}(t_0) \quad &= \quad \int_{t_0}^{t_1} \hat{\psi}(\tau) \left[G\hat{u}(\tau), [g^i, g^j]\right] (\hat{q}(\tau))d\tau = \hat{\psi}(t_0) \left[G \int_{t_0}^{t_1} \hat{u}(\tau)d\tau, [g^i, g^j]\right] (\hat{q}(t_0)) \\ &+ \int_{t_0}^{t_1} \hat{\psi}(\tau) \left[G\hat{u}(\tau), [G \int_{\tau}^{t_1} \hat{u}(\xi)d\xi, [g^i, g^j]]\right] (\hat{q}(\tau))d\tau.\end{aligned} \tag{6.13}$$

As far as the distribution is medium fat then for some $i_0, j_0 \in \{1, \dots, r\}$ there holds

$$\hat{\psi}(t_0) \left[G \int_{t_0}^{t_1} \hat{u}(\tau)d\tau, [g^{i_0}, g^{j_0}]\right] (\hat{q}(t_0)) \geq b\|\psi^3(t_0)\| \left\|\int_{t_0}^{t_1} \hat{u}(\tau)d\tau\right\| \geq 2^{-1}b\|\psi(t_0)\| \cdot \left\|\int_{t_0}^{t_1} \hat{u}(\tau)d\tau\right\|.$$

For any i, j the norm of the second addend at the right-hand side of (6.13) admits an upper estimate

$$C\rho e^{B\rho}\|\psi(t_0)\| \left\|\int_{t_0}^{t_1} \hat{u}(\tau)d\tau\right\|.$$

Without loss of generality we may assume $C\rho_0 e^{B\rho_0} < 2^{-2}b$ thus coming to the lower estimate

$$|\psi^2_{i_0j_0}(t_1) - \psi^2_{i_0j_0}(t_0)| \geq 2^{-2}b\|\psi(t_0)\| \left\|\int_{t_0}^{t_1} \hat{u}(\tau)d\tau\right\| \geq 2^{-2}bc\|\psi(t_0)\|\rho^3.$$

Now recalling that $\|\psi^2(t_0)\| \leq B\delta\rho^3\|\psi(t_0)\|$ we obtain

$$|\psi^2_{i_0j_0}(t_1)| \geq -B\delta\rho^3\|\psi(t_0)\| + 2^{-2}bc\rho^3\|\psi(t_0)\|.$$

Choosing $\delta < bc/8B$ we conclude

$$\psi^2_{i_0j_0}(t_1) \geq 2^{-3}bc\rho^3\|\psi(t_0)\| \geq B\delta\rho^3\|\psi(t_0)\| \geq B\delta\rho^3(1 - B\rho_0)\mu^{-1}\|\hat{\psi}_T\|.$$

□

Now we may complete the proof of the Theorem.

Integrating the second addend of (6.1) by parts we transform the second variation into

$$-\int_0^\rho \langle u(\tau), u(\tau)\rangle d\tau + \hat{\psi}_T \int_0^\rho \left[Y_\tau \int_0^\tau u(\xi)d\xi, Y_\tau u(\tau)\right](q^0)d\tau - \hat{\psi}_T \int_0^\rho \left[\int_0^\tau \dot{Y}_\xi \int_0^\xi u(\theta)d\theta d\xi, Y_\tau u(\tau)\right](q^0)d\tau.$$

Let us take the point t_0 whose existence has been established by the Lemma 6.7. Take a subinterval of length $\delta\rho^\alpha$ joining (for example beginning at) this point. Put $\alpha = 1$, when dealing with a 2-generating distribution, and choose any $\alpha > 3$, when dealing with a medium fat distribution.

We restrict our consideration to controls $u(\cdot)$ of magnitudes ≤ 1 with supports in the subinterval $[t_0, t_0+\delta\rho^\alpha]$. Then the second variation takes form

$$\begin{aligned} -\int_{t_0}^{t_0+\delta\rho^\alpha} \langle u(\tau), u(\tau)\rangle d\tau \quad &+ \quad \hat{\psi}_T \int_{t_0}^{t_0+\delta\rho^\alpha} \left[Y_\tau \int_{t_0}^\tau u(\xi)d\xi, Y_\tau u(\tau)\right](q^0)d\tau \\ &- \quad \hat{\psi}_T \int_{t_0}^{t_0+\delta\rho^\alpha} \left[\int_{t_0}^\tau \dot{Y}_\xi \int_{t_0}^\xi u(\theta)d\theta d\xi, Y_\tau u(\tau)\right](q^0)d\tau. \end{aligned} \qquad (6.14)$$

Let us note that for fixed $\hat{\psi}_T$ and $\rho \to +0$ each integration in (6.14) adds a factor ρ^α to the order of smallness.

The norm of the first addend admits an upper[4] estimate $\bar{C}\rho^\alpha$. The last addend admits an upper estimate $\bar{C}\|\hat{\psi}_T\|\rho^{3\alpha}$. According to the formula (6.11)

$$\forall t \in [t_0, t_0+\delta\rho^\alpha]: \ \left| \hat{\psi}_T \left[Y_t v, Y_t u](q^0) - \hat{\psi}_T [Y_{t_0} v, Y_{t_0} u\right](q^0) \right| \leq \bar{C}\|\hat{\psi}_T\|\rho^\alpha$$

and hence the second addend in (6.14) can be represented as

$$\hat{\psi}_T \int_{t_0}^{t_0+\delta\rho^\alpha} \left[Y_{t_0} \int_{t_0}^\tau u(\xi)d\xi, Y_{t_0} u(\tau)\right](q^0)d\tau + R(\|\hat{\psi}_T\|\rho^{3\alpha}),$$

where again $R(\|\hat{\psi}_T\|\rho^{3\alpha}) \leq \bar{C}\|\hat{\psi}_T\|\rho^{3\alpha}$.

Here $\hat{\psi}_T[Y^i_{t_0}, Y^j_{t_0}](q^0)$ are the components of the covector $\psi^2(t_0)$. According to the conclusion of the Lemma 6.7, $\|\psi^2(t_0)\|$ can be estimated from below by $C\delta\|\hat{\psi}_T\|$ for a 2-generating distribution, or, by $C\delta\rho^3\|\hat{\psi}_T\|$ for a medium fat distribution.

Let us consider the quadratic form

$$G(u(\cdot)) = \hat{\psi}_T \int_{t_0}^{t_0+\delta\rho^\alpha} \left[Y_{t_0} \int_{t_0}^\tau u(\xi)d\xi, Y_{t_0} u(\tau)\right](q^0)d\tau. \qquad (6.15)$$

The values of this quadratic form depend only on $\psi_2(t_0)$ (linearly) and on the length of the interval. Introducing a function $w(\theta) = u(t_0+\theta\delta\rho^\alpha), \ \theta \in [0,1]$, we arrive to the equality

$$G(u(\cdot)) = \delta^2\rho^{2\alpha}\hat{\psi}_T \int_0^1 \left[Y_{t_0} \int_0^\tau w(\xi)d\xi, Y_{t_0} w(\tau)\right](q^0)d\tau.$$

The latter quadratic form possesses a symmetry: the involution $J : w(\theta) \mapsto w(1-\theta)$ changes the signs of its values. This implies that for any subspace on which the form is positive definite there exists a subspace

[4]We can choose common constant $\bar{C} > 0$ in this and the subsequent upper estimates. This constant does not depend on geodesic.

of the same dimension on which the form is negative definite. As long as the kernel of this form has infinite codimension both of those dimensions must be infinite.

It follows from the aforesaid (see also [9] Appendix 2), that

- in the medium fat case for each integer $N > 0$ there exist N-dimensional subspaces L_N^+ (respectively L_N^-) and constants $A_N > 0$ such that the quadratic form (6.15) is positive (respectively negative) definite on L_N^+ (respectively on L_N^-) and admit a lower estimate $b_N = A_N \rho^{2\alpha+3} \|\hat{\psi}_T\| \|u(\cdot)\|^2_{L\infty}$ (respectively an upper estimate $-b_N$);
- in the 2 -generating case these estimates change to $\pm A_N \rho^2 \|\hat{\psi}_T\| \|u(\cdot)\|^2_{L\infty}$.

Therefore if $\|\hat{\psi}_T\| \to +\infty$ then the values of (6.15) grow at least as $\|\hat{\psi}_T\| \rho^{2\alpha+3}$ or, respectively in the 2-generating case, as $\|\hat{\psi}_T\| \rho^2$. Hence we may choose ρ_0 in such a way that this principal term dominates the terms which grow at most as $\|\hat{\psi}_T\| \rho^{3\alpha}$, as $\rho \to +0$. Also for sufficiently large $\|\hat{\psi}_T\|$ (larger than $c\rho^{-\alpha-3}$ in the medium fat case, or larger than $c\rho^{-1}$ in the 2-generating case) this principal term will dominate the addend $-\int_{t_0}^{t_0+\delta\rho^\alpha} \langle u(\tau), u(\tau) \rangle d\tau$ which is $O(\rho^\alpha)$. In this case the second variation will have large positive and negative indices and hence corresponding geodesic path would not be length minimizer.

Therefore for length minimizers of 2-generating sub-Riemannian structures there must hold $\|\hat{\psi}_T\| = O(\rho^{-1})$ as $\rho \to 0$, while for medium fat distribution there must hold $\|\hat{\psi}_T\| = O(\rho^{-\alpha-3})$, as $\rho \to 0$ for any $\alpha > 3$.

To obtain the respective results for the geodesics parametrized by the interval $[0, 1]$ instead of length we only need to introduce a time substitution $t \mapsto \rho^{-1} t$ which results in the transformation of other variables $u \mapsto \rho u, \ \psi \mapsto \rho \psi$. □

6.3. Proof of the Theorem 3.8 on lack of strictly abnormal minimizers for medium fat distributions

Proof. Let us consider an abnormal geodesic $\left(\hat{u}(\cdot), \hat{q}(\cdot), \hat{\psi}(\cdot)\right)$ and the absolutely continuous covector-function $\psi^2(t)$ introduced by (6.3). It can not vanish identically. Indeed if $\psi^2(t_0) = 0$, and in addition t_0 is a Lebesgue point of the function $t \mapsto \hat{u}(t)$ then according to (6.5) the components of the derivative $\dot{\psi}^2(t_0)$ equal $\hat{\psi}(t_0)[G\hat{u}(t_0), [g^i, g^j]](\hat{q}(t_0))$. Due to the medium fatness of the distribution this derivative is non vanishing. Therefore $\|\psi^2(t)\| \neq 0$ on some interval $[t_0, t_0 + \nu]$, *i.e.* Goh condition fails for the abnormal extremal. Then according to the Theorem 3.7 this strictly abnormal geodesic path $(\hat{u}(\cdot), \hat{q}(\cdot))$ can not be a length minimizer. □

6.4. Proof of the Theorem 5.2 on generic medium fat distributions

Proof. For a generic germ of r-distribution we may assume $\dim \mathcal{D}_q^2 = \frac{r(r+1)}{2}$ (see Sect. 6, the proof of the Prop. 5.2).

Let $f^1, \ldots, f^r$ be a basis of the distribution. Let us enumerate the pairs (i, j) with $1 \leq i < j \leq r$, by means of one-to-one map $n : (i, j) \mapsto n_{ij}, \ n_{ij} \in \{1, \ldots, r(r-1)/2\}$. Now we introduce a $(\frac{r(r-1)}{2} \times r)$-matrix $C = (c_{ks}) : \ c_{ks} = [f^s[f^i, f^j]]$, where $(i, j) = n^{-1}(k)$.

Obviously a germ of distribution fails to be medium fat if and only if for some covector ψ from the annihilator of $\mathcal{D}^2$ in $T_q^* M$ the matrix $C^\psi = (c_{ks}^\psi) = (\langle \psi, [f^s[f^i, f^j]] \rangle)$ has rank $< r$. Indeed in this case there exists a nonzero vector $v = (v_1, \ldots, v_r)$ such that $C^\psi v = 0$ what means that for $X = \sum_{i=1}^r f^i v_i$ the covector ψ annihilates also $[X, \mathcal{D}^2](q)$.

The codimension of the (maximal stratum of the stratified) manifold $\mathcal{S}$ of $(\frac{r(r-1)}{2} \times r)$-matrices of rank $< r$ equals $\frac{r(r-1)}{2} - (r-1) = \frac{(r-2)(r-1)}{2}$. The codimension of $\mathcal{D}_q^2$ in $T_q M$ is $\leq r + (r-1)^2 - \frac{r(r+1)}{2} = \frac{(r-2)(r-1)}{2}$. These codimensions coincide and were the entries of the matrix C independent, then we could say that for a generic distribution $\mathcal{D}$ there is no nonzero $\psi \in (\mathcal{D}_q^2)^\perp$ for which the corresponding matrix C^ψ has rank $< r$ (due to homogeneity we may think that ψ belongs to the unit sphere in $(\mathcal{D}_q^2)^\perp$).

Clearly the 3rd-order iterated Lie brackets which appear in the matrix C are *not* independent, since for example they must satisfy the Jacobi identities which will implicate the equality relations for the elements of each C^ψ. This means that the range of the map $\psi \mapsto C^\psi$ is not the whole space $\mathbb{R}^{\frac{r(r-1)}{2}\times r}$ but some linear subspace L of this space. This change would not affect our reasoning if we could prove that the codimension of $\mathcal{S}\bigcap L$ in L coincides with the codimension of $\mathcal{S}$ in $\mathbb{R}^{\frac{r(r-1)}{2}\times r}$. To do this it is enough to verify that L is transversal to $\mathcal{S}$ in $\mathbb{R}^{\frac{r(r-1)}{2}\times r}$.

To establish the transversality we will have to describe the relations between the elements of the matrices C^ψ (the relations which determine L). To proceed with it we consider a Hall basis (see [33]) in the free Lie algebra generated by $f^1, \ldots, f^r$ and select from this basis the subbasis consisting of third-order iterated Lie brackets. If we put in order the generators $f^1 \prec f^2 \prec \cdots \prec f^r$; then one may think that this subbasis consists of all Lie brackets $[f^s[f^i, f^j]]$ for which $i \le s$ and $i < j$.

Each Lie bracket of this latter form is involved into Jacobi identity $[f^s[f^i, f^j]] - [f^j[f^i, f^s]] + [f^i[f^j, f^s]] = 0$. This identity degenerates into a trivial relation of antisymmetry (*i.e.* does not provide a true relation) if $i = s$ or $j = s$. Otherwise the first and the second summands are elements of the Ph.Hall basis. The third summand, written the way it is, if $j < s$, or written as $-[f^i[f^s, f^j]]$, if $s < j$, is a linear combination of the first and the second summands.

Let us consider a maximal stratum of $\mathcal{S}$ (consisting of rank-(r-1) matrices) and assume, not loosing generality that, this stratum consists of the matrices C whose left upper $(r-1)\times(r-1)$-minor is non vanishing: det $A_{r-1} \neq 0$. Then

$$C = \begin{pmatrix} A_{r-1} & a \\ B & b \end{pmatrix}$$

and the subcolumn b is determined by the other elements of the matrix as $b = BA_{r-1}^{-1}a$.

The elements of this subcolumn correspond to the iterated Lie brackets $[f^r[f^i, f^j]]$ with all $(i,j) \in \Gamma \subset n^{-1}(\{r, \ldots, \frac{r(r-1)}{2}\})$, and all these brackets are the elements of the Ph.Hall basis. Therefore they are linearly independent as elements of free Lie algebra, but we have to prove that for a generic germ of distribution the linear span of their values at q is transversal to $\mathcal{D}_q^2$.

To this end let us choose an adapted coordinate system $(x,y) : \mathcal{O} \to \mathbb{R}^n$ such that:

i) $(x,y)(q) = 0_{R^n}$;
ii) $(x,y)_*|_q f_i(q)$ $(i = 1, \ldots, r)$ coincide with the first r elements $e_i, i = 1, \ldots, r$, of the standard basis of $\mathbb{R}^n$;
iii) $(x,y)_*|_q [f_i, f_j](q)(1 \le i < j \le r)$ coincide with the next $r(r-1)/2$ elements e_{ij} $(i < j)$ of this standard basis. (Obviously $(x,y)_*|_q \mathcal{D}_q^2 = \mathbb{R}^r \oplus \mathbb{R}^{r(r-1)/2}$.)
iv) $\Delta_q^3 = \mathrm{span}\{\mathcal{D}_q^2, [f^r, [f^i, f^j]](q), |(ij) \in \Gamma\}$ is mapped isomorphically by the differential $x_*|_q$ onto a subspace of $\mathbb{R}^n$, while $y_*|_q \Delta_q^3 = 0$.

Evidently if $\Delta_q^3 = T_q M$, this would mean that $\{[f^r, [f^i, f^j]](q)|\ (ij) \in \Gamma\}$ span a subspace transversal to $\mathcal{D}_q^2$ in $T_q M$ and then we are done. Otherwise choose any minimal subset $\bar{\Gamma} \subseteq \Gamma$ such that Δ_q^3 is spanned by $\{[f^r, [f^i, f^j]](q), |(ij) \in \bar{\Gamma}\}$ and $\mathcal{D}_q^2$. Let $y = (y_1, \ldots, y_s)$; $s = \mathrm{codim}\ \Delta_q^3$. Obviously

$$s = \mathrm{codim}\ \mathcal{D}_q^2 - \sharp\bar{\Gamma} \le (r-1)(r-2)/2 - \sharp\bar{\Gamma} = \sharp(\Gamma \setminus \bar{\Gamma}).$$

Let us choose an arbitrary injection $\sigma : \{1, \ldots, s\} \to \Gamma \setminus \bar{\Gamma}$ and define vector field

$$\tilde{f^r} = f^r - \sum_{\alpha=1}^{s} \varepsilon_\alpha x_{\sigma(\alpha)} \partial/\partial y_\alpha,\ (\varepsilon_\alpha \in \mathbb{R},\ \alpha = 1, \ldots, s).$$

Direct computation gives us $[\tilde{f^r}, [f^i, f^j]] = [f^r, [f^i, f^j]]$, if $(ij) \notin \sigma(\{1, \ldots, s\})$, in particular if $(ij) \in \bar{\Gamma}$. Hence the values (mod$\mathcal{D}_q^2$) at q of these iterated Lie brackets span $\Delta_q^3/\mathcal{D}_q^2$. For the pairs $(ij) \in \sigma(\{1, \ldots, s\})$ we obtain $[\tilde{f^r}, [f^i, f^j]] = [f^r, [f^i, f^j]] + \varepsilon_{\sigma^{-1}(ij)} \partial/\partial y_{\sigma^{-1}(ij)}$. As we know $[f^r, [f^i, f^j]](q) \in \Delta_q^3$ and the span of $\partial/\partial y_\alpha|_q$, $\alpha = 1, \ldots, s$ is transversal (in fact complement) to Δ_q^3. Therefore we constructed a distribution

$\tilde{\mathcal{D}}$ generated by $f^1, \dots, f^{r-1}, \tilde{f}^r$ which meets the transversality property and is arbitrarily close to $\mathcal{D}$ in C^∞ topology provided $|\varepsilon_\alpha|$ $\alpha = 1, \dots, s$ are sufficiently small. □

References

[1] A.A. Agrachev, *Quadratic mappings in geometric control theory,* in: Itogi Nauki i Tekhniki, Problemy Geometrii, VINITI, Acad. Nauk SSSR, Moscow **20** (1988) 11-205. English transl. in *J. Soviet Math.* **51** (1990) 2667-2734.

[2] A.A. Agrachev, The second-order optimality condition in the general nonlinear case. *Matem. Sbornik* **102** (1977) 551-568. English transl. in: *Math. USSR Sbornik* **31** (1977).

[3] A.A. Agrachev, *Topology of quadratic mappings and Hessians of smooth mappings,* in: Itogi Nauki i Tekhniki, Algebra, Topologia, Geometria; VINITI, Acad. Nauk SSSR **26** (1988) 85-124.

[4] A.A. Agrachev, B. Bonnard, M. Chyba and I. Kupka, Sub-Riemannian spheres in Martinet flat case. *ESAIM: Contr., Optim. and Calc. Var.* **2** (1997) 377-448.

[5] A.A. Agrachev and R.V. Gamkrelidze, Second-order optimality condition for the time-optimal problem. *Matem. Sbornik* **100** (1976) 610-643. English transl. in: *Math. USSR Sbornik* **29** (1976) 547-576.

[6] A.A. Agrachev and R.V. Gamkrelidze, Exponential representation of flows and chronological calculus. *Matem. Sbornik* **107** (1978) 467-532. English transl. in: *Math. USSR Sbornik* **35** (1979) 727-785.

[7] A.A. Agrachev, R.V. Gamkrelidze and A.V. Sarychev, Local invariants of smooth control systems. *Acta Appl. Math.* **14** (1989) 191-237.

[8] A.A. Agrachev and A.V. Sarychev, On abnormal extremals for Lagrange variational problems. (summary). *J. Mathematical Systems, Estimation and Control* **5** (1995) 127-130. Complete version: *J. Mathematical Systems, Estimation and Control* **8** (1998) 87-118.

[9] A.A. Agrachev and A.V. Sarychev, Abnormal sub-Riemannian geodesics: Morse index and rigidity. *Ann. Inst. H. Poincaré* **13** (1996) 635-690.

[10] A.A. Agrachev and A.V. Sarychev, Strong minimality of abnormal geodesics for 2-distributions. *J. Dynamical Control Systems* **1** (1995) 139-176.

[11] V.I. Arnol'd, A.N. Varchenko and S.M. Gusein-Zade, *Singularities of differentiable maps* **1** Birkhäuser, Boston (1985).

[12] P. Brunovsky, Existence of regular synthesis for general problems. *J. Differential Equations* **38** (1980) 317-343.

[13] R.L. Bryant and L. Hsu, Rigidity of integral curves of rank 2 distributions. *Invent. Math.* **114** (1993) 435-461.

[14] W-L. Chow, Über Systeme von linearen partiellen Differentialgleichungen erster ordnung. *Match. Ann.* **117**, (1940/41) 98-105.

[15] A.F. Filippov, On certain questions in the theory of optimal control. Vestnik Moskov. Univ., *Ser. Matem., Mekhan., Astron.* **2** (1959) 25-32.

[16] A. Gabrielov, Projections of semianalytic sets. *Funct. Anal Appl.* **2** (1968) 282-291.

[17] R.V. Gamkrelidze, *Principles of optimal control theory.* Plenum Press, New York (1978).

[18] Zhong Ge, Horizontal path space and Carnot-Caratheodory metric. *Pacific J. Math.* **161** (1993) 255-286.

[19] V.Ya. Gershkovich, Bilateral estimates for metrics, generated by completely nonholonomic distributions on Riemannian manifolds. *Doklady AN SSSR* **278** (1984) 1040-1044.

[20] B.S. Goh, Necessary conditions for singular extremals involving multiple control variables. *SIAM J. Control* **4** (1966) 716-731.

[21] M. Goresky and R. MacPherson, *Stratified Morse Theory.* Springer-Verlag, N.Y. (1988) Ch.1.

[22] R. Hardt, Stratifications of real analytic maps and images. *Inventiones Math.* **28** (1975) 193-208.

[23] G.W. Haynes and H. Hermes, Nonlinear Controllability *via* Lie Theory. *SIAM J. Control* **8** (1970) 450-460.

[24] H. Hironaka, *Subanalytic sets*, Lecture Notes Istituto Matematico "Leonida Tonelli", Pisa, Italy (1973).

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

[26] A.J. Krener, The high-order maximum principle and its applications to singular extremals. *SIAM J. Control and Optim.* **15** (1977) 256-293.

[27] W. Liu and H.J. Sussmann, *Shortest paths for sub-Riemannian metrics on rank-2 distributions*, Memoirs of AMS, No. 564 (1995).

[28] S. Lojasiewicz Jr. and H.J. Sussmann, Some examples of reachable sets and optimal cost functions that fail to be subanalytic. *SIAM J. Control and Optim.* **23** (1985) 584-598.

[29] R. Montgomery, *Geodesics, which do not satisfy geodesic equations*, Preprint (1991).

[30] R. Montgomery, A survey on singular curves in sub-Riemannian geometry. *J. Dynamical and Control Systems* **1** (1995) 49-90.

[31] P.K. Rashevsky, About connecting two points of a completely nonholonomic space by admissible curve. *Uchen. Zap. Ped. Inst. Libknechta* **2** (1938) 83-94.

[32] C.B. Rayner, *The exponential map for the Lagrange problem on differentiable manifolds. Philos. Trans. Roy. Soc. London Ser. A, Math. Phys. Sci.* **262** (1967) 299-344.

[33] J.P. Serre, *Lie algebras and lie groups*, Benjamin, New York (1965).

[34] H.J. Sussmann, Subanalytic sets and feedback control. *J. Differential Equations* **31** (1979) 31-52.

[35] H.J. Sussmann, *A cornucopia of four-dimensional abnormall sub-Riemannian minimizers*, A. Bellaïche, J.-J. Risler, Eds., Sub-Riemannian Geometry, Birkhäuser, Basel (1996) 341-364.

[36] H.J. Sussmann, *Optimal control and piecewise analyticity of the distance function.* A. Ioffe, S. Reich, Eds., Pitman Research Notes in Mathematics, Longman Publishers (1992) 298-310.

[37] A.M. Vershik and V.Ya. Gershkovich, *Nonholonomic dynamical systems, geometry of distributions and variational problems.* V.I. Arnol'd, S.P. Novikov, Eds., Dynamical systems VII, *Encyclopedia of Mathematical Sciences* **16**, Springer-Verlag, NY (1994).

[38] L.C. Young, *Lectures on the calculus of variations and optimal control theory,* Chelsea, New York (1980).