ESAIM: COCV 19 (2013) 274–287
DOI: 10.1051/cocv/2012006

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

SUBRIEMANNIAN GEODESICS OF CARNOT GROUPS OF STEP 3 *

Kanghai Tan[1] and Xiaoping Yang[2]

Abstract. In Carnot groups of step ≤ 3, all subriemannian geodesics are proved to be normal. The proof is based on a reduction argument and the Goh condition for minimality of singular curves. The Goh condition is deduced from a reformulation and a calculus of the end-point mapping which boils down to the graded structures of Carnot groups.

Mathematics Subject Classification. 53C17, 49K30.

Received May 10, 2011. Revised November 2, 2011.
Published online June 12, 2012.

1. Introduction

This paper is inspired by the smoothness problem of subriemannian geodesics, one of the fundamental problems in subriemannian geometry. We first study the case of Carnot group with step ≤ 3. In this case we proved that all subriemannian geodesics are normal and thus smooth.

To state the subriemannian geodesic problem, we first recall some basic facts on subriemannian geometry. We refer to the book [27] for detail. A subriemannian manifold is a smooth n-dimensional manifold M with a k-dimensional subbundle or distribution $\triangle \subset TM$ on which a smooth inner product g_c is endowed. $(\triangle, g_c)$ is called a subriemannian structure on M and $\triangle$ horizontal bundle. In this paper, we always assume M is connected and $\triangle$ satisfies the so-called Chow-Hömander condition which means that vector fields of $\triangle$ together with all their commutators span the tangent space at each point on M. Carnot groups are important examples of subriemannian manifolds. A Carnot group $\mathbb{G}$ is a connected, simply connected Lie group with a graded Lie algebra

$$\flat = V^1 \oplus \ldots \oplus V^r, \quad \text{with } V^i = [V^1, V^{i-1}], \quad [V^1, V^r] = 0, \quad i = 2, \ldots, r. \tag{1.1}$$

The integer r is called the step of $\mathbb{G}$. Since $\mathbb{G}$ is connected and simply connected, the exponential map from $\flat$ to $\mathbb{G}$ gives a global chart for $\mathbb{G}$. Carnot groups are tangent spaces (in the sense of Gromov-Hausdorff) of equiregular subriemannian manifolds, see [9, 25]. It is believed that the role played by Carnot groups in subriemannian geometry is similar to the role of Euclidean Spaces in Riemannian geometry.

Keywords and phrases. Subriemannian geometry, geodesics, calculus of variations, Goh condition, generalized Legendre-Jacobi condition.

* *The first author is supported by NSF of China (No. 10801073) and a grant from China Scholarship Council for study abroad.*

[1] Department of Applied Mathematics, Nanjing University of Science & Technology, Nanjing 210094, P.R. China.
khtan@mail.njust.edu.cn

[2] School of Science, Nanjing University of Science & Technology, Nanjing 210094, P.R. China. yangxp@mail.njust.edu.cn

It follows from the Chow-Rashevskii connectivity theorem that for any given points $p, q \in M$ there always exists at least a horizontal curve connecting p and q, see [16, 29]. Here a horizontal curve is by definition an absolutely continuous curve $\gamma : [0,1] \to M$ such that $\dot\gamma(t) \in \Delta_{\gamma(t)}M$ whenever $\dot\gamma(t)$ exists. Thus one can define a natural distance:

$$\mathrm{d}_{sr}(p,q) = \inf \int_0^1 \sqrt{g_c(\dot\gamma,\dot\gamma)}\mathrm{d}t$$

where the infimum is taken among the set $\Omega(p,q)$ of all horizontal curves γ such that $\gamma(0) = p$ and $\gamma(1) = q$. d_{sr} is called the Carnot-Carathéodory distance of $(M, \triangle, g_c)$. A subriemannian geodesic is a horizontal curve locally realizing d_{sr}. We will reserve the terminology "minimizing geodesic" or "minimizer" for those globally distance-realizing subriemannian geodesics. It is not difficult to prove that any two sufficiently close points can be joined by a minimizing geodesic. If (M, d_{sr}) is complete, there is a minimizing geodesic connecting any two points. Before Montgomery [26] (in 1991) discovered a smooth singular minimizer, it was taken for grant (see *e.g.* [30]) that each subriemannaian geodesic similar to a Riemannian geodesic could satisfy a Hamilton-Jacobi equation:

$$\dot x^i = \frac{\partial H}{\partial \lambda_i}, \qquad \dot\lambda_i = -\frac{\partial H}{\partial x^i} \tag{1.2}$$

where (x^i, λ_i) is a coordinate system of $T^\star M$, $H(x,\lambda) = \max_{v\in\triangle_x}\{\lambda(v) - \frac{1}{2}g_c(v,v)\}$ $(\lambda \in T^\star_x M)$ is the subriemannian Hamiltonian. A horizontal curve $\gamma(t) = (x^i(t))$ (denoted by a local coordinate) satisfying (1.2) almost everywhere for some lift $\lambda(t) = (\lambda_i(t))$ can be proved to be locally minimizing and smooth, and is called a normal geodesic. Montgomery's example shows that not all subriemannian geodesics are normal. The subriemannian geodesic problem is a special case of geometric control problems. In fact singular curves or abnormal extremals play a very important role in optimal control theory, see *e.g.* [4, 12]. It is well known that the Pontryagin maximum principle (or the Lagrange Multiplier Rule in the Lagrangian formulation) gives the first order necessary condition of optimality for optimal control problems. This first order condition is hardly considered to be satisfactory when one studies abnormal extremals. Recently experts developed necessary/sufficient second order conditions of optimality, *i.e.*, Goh condition and generalized Legendre-Jacobi condition, see *e.g.* [3–7]. These conditions were derived from the finiteness of the generalized Morse index of critical points of the end-point mapping. We refer to [4, 5, 7] for the finiteness of the generalized Morse index.

Let $\Omega(p)$ be the set of all horizontal curves $\gamma : [0,1] \to M$ such that $\gamma(0) = p$. It is well known that $\Omega(p)$ is a smooth Banach manifold, see [10]. The end-point mapping is the smooth mapping $\mathscr{E} : \Omega(p) \to M$ defined by taking $\gamma \in \Omega(p)$ to $\gamma(1)$. Thus $\Omega(p,q) = \mathscr{E}^{-1}(q)$. In general $\mathscr{E}$ is not regular at all curves in $\Omega(p)$ and thus $\Omega(p,q)$ possibly has no smooth structure if q is a critical value of $\mathscr{E}$. If $\gamma \in \Omega(p)$ is a critical point of $\mathscr{E}$, we call γ a singular curve. After Montgomery's example, Liu and Sussmann in [23] gave more examples of singular curves which are locally minimizing. All these examples found on rank two distributions are in fact C^1-rigid curves which by definition are locally isolated curves in $\Omega(p,q)$ with respect to the C^1-topology, see also Bryant and Hsu [13]. For every rank two distribution satisfying $\triangle^3_p \neq \triangle^2_p$ at $p \in M$ they proved that there exists a rigid curve locally length-minimizing and emanating from p, see also Agrachev and Sarychev in [5], Theorem 6.2. Here $\triangle^1 = \triangle, \triangle^i = \triangle^{i-1} + [\triangle^1, \triangle^{i-1}]$ for $i = 2, \ldots$ The research of such curves first appeared in the work of Carathéodory, Engel, and Hilbert, see [11, 32]. Classical calculus of variations can not fully deal with the subriemannian geodesic problem when $\Omega(p,q)$ contains singular curves, because there possibly exist no smooth variations of such curves. But singular curves could be subriemannian geodesics as shown by the above mentioned work. A fundamental problem is whether all subriemannian geodesics are smooth. This problem is equivalent to the question whether all singular geodesics are smooth, since normal (regular) geodesics are always smooth. There are some substantial results so far, while the problem is still open for general cases. Agrachev and Sarychev [6] proved that there admit no strictly singular geodesics for medium fat distributions including strong-generating distributions (fat distributions) for which Strichartz [30] had already obtained the conclusion. For a class of equiregular subriemannian manifolds, Leonardi and Monti [22] showed that length-minimizing curves have no corner-like singularities which in particular implies that all singular geodesics in Carnot groups of rank two with step ≤ 4 are smooth. There are also some "generic" results which claims that for $3 \leq k < n$ there exists an open dense subset $\mathcal{O}_k$ of the space $\mathscr{D}_k$ consisting of all k-dimensional distributions on M (endowed with the Whitney C^∞ topology), such that each distribution in $\mathcal{O}_k$ admits no singular geodesics, see [3,15] and references therein.

In this paper we will concentrate on the case of Carnot groups. As mentioned above Carnot groups are very important in subriemannian geometry. Our study here will be instructive for later considerations of subriemannian geodesics of general distributions. Our starting point is the refined formulation of the end-point mapping which boils down to the graded structures of Carnot groups. The graded structure (1.1) implies that each horizontal curve $\Upsilon = (\gamma^1, \gamma^2, \ldots, \gamma^r)$ is uniquely determined by the first layer γ^1. Here we use the exponential mapping exp to identify the Carnot group $\mathbb{G}$ with its Lie algebra $\flat$ and $\gamma^i = \pi^i(\exp^{-1} \Upsilon) \in V^i$ where $\pi^i : \flat = V^1 \oplus \ldots \oplus V^r \to V^i$ is the projection to the i-th component. The subriemannian geodesic problem in Carnot groups can be formulated as a minimization problem with equality constraint. The end-point mapping $\mathscr{E}$ is different from the ordinary one which usually takes a control function to the end point of the trajectory uniquely determined by the control function (the initial point is fixed). The formula for the differential of $\mathscr{E}$ can be written out for Carnot groups with any step by a very tedious computation. In the case of step 3, the differential and the intrinsic Hessian of the end-point mapping are of simple form. We will get the second order necessary conditions for optimality of a singular curve, that is, if a singular curve $\Upsilon = (\gamma^1, \gamma^2, \ldots, \gamma^r)$ is locally energy-minimizing then γ^1 must satisfy the Goh condition and the generalized Legendre-Jacobi condition, see Propositions 4.3 and 4.5. From the Goh condition and the graded structure (1.1) we deduce that the first layer γ^1 of a singular geodesic Υ must be in a lower-dimension subspace. Thus singular geodesics must be in Carnot subgroups of rank 2 or step 2. We finally reduce the problem to the rank two case which (known to experts) is easy, see Theorems 5.2 and 5.4. The reduction to lower subgroups is a curious coincidence with Hamenstädt's suggestion for the smoothness problem, see [19].

The paper is organized into five sections. In the next section we give a formulation of the end-point mapping which is based on a characterization of horizontal curves in Carnot groups. Section 3 is devoted to the calculus of the end-point mapping. We will give the differential, the intrinsic Hessian for the step ≤ 3 case. In Section 4 we derive the second order necessary conditions of singular geodesics. We prove the main results in Section 5.

2. Horizontal curves and the end-point mapping

2.1. Basic structure of Carnot groups

Let $\mathbb{G}$ be a Carnot group with a Lie algebra $\flat$ satisfying (1.1) (we call such Lie algebras Carnot algebras). We endow on V^1 an inner product $\langle \cdot, \cdot \rangle$. Let $n_i = \dim(V_i)$, $n = \sum_{i=1}^r n_i$. The $r-$vector $(n_1, n_1 + n_2, \ldots, \Sigma_{j=1}^i n_j, \ldots, n)$ is called the growth vector of the Carnot group. We fix an orthonormal basis of V^1, $\{e_1, \ldots, e_{n_1}\}$, then arbitrarily extend it to a basis of $\flat$, $\{e_1, \ldots, e_{n_1}, e_{n_1+1} \ldots, e_n\}$, and extend $\langle \cdot, \cdot \rangle$ to an inner product g on $\flat$ making the basis orthonormal. *Via* the exponential mapping exp we identify $\mathbb{G}$ with $\flat$ or $\mathbb{R}^n$ with a group law determined by the Baker-Campbell-Hausdorff formula. For $p \in \mathbb{G}$, setting $X_i(p) = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \{p \cdot \exp(te_i)\}$, $i = 1 \ldots, n$, we get the basis of the space of left-invariant vector fields. The horizontal bundle $\triangle = \mathrm{span}\{X_1, \ldots, X_{n_1}\}$ satisfies the Chow-Hörmander condition by (1.1). Let $g_c((X_i(p), X_j(p)) = \langle e_i, e_j \rangle, i, j = 1, \ldots, n_1$. Thus we have a subriemannian structure $(\triangle, g_c)$ on $\mathbb{G}$. We also extend g_c to a left-invariant Riemannian metric g_r such that $\{X_1, \ldots, X_n\}$ is an orthonormal basis of $T\mathbb{G}$. We emphasize that subriemannian geodesics in $\mathbb{G}$ are independent of the choice of orthonormal bases and their extensions, that is, they are completely determined by $(\flat, V^1, \langle \cdot, \cdot \rangle)$ or equivalently by $(\mathbb{G}, \triangle, g_c)$.

Example 2.1.

1. The simplest Carnot group is the Heisenberg group $\mathbb{H}^m$ with the Heisenberg algebra (growth vector $= (2m, 2m+1)$) $\mathbb{h} = \mathrm{span}\{e_1, \ldots, e_m, f_1 \ldots, f_m\} \oplus \mathrm{span}\{g\}$ with the basis satisfying that $[e_i, f_i] = g$, $i = 1, \ldots, m$, and all other brackets vanish;

2. the Engel group is a Carnot group with the growth vector $(2,3,4)$. Its algebra is $\mathrm{span}\{e_1,e_2\}\oplus\mathrm{span}\{e_3\}\oplus\mathrm{span}\{e_4\}$ with $[e_1,e_2]=e_3, [e_1,e_3]=e_4$. Note that Carnot groups of rank two ($n_1=2$) has a special feature: the second layer has dimension 1 whatever its step. We will see later this feature make the study of its subriemannian geodesics very easy;
3. the free Carnot group with bi-dimension (k,r) has the maximal vector growth among all Carnot groups with k generators and step $=r$. The free Carnot groups play some particular roles in nilpotent analysis;
4. let $\flat$ be a Carnot algebra satisfying (1.1). If $W\subset V^1$ is a lower-dimensional subspace, then $\bar{\flat}=W^1\oplus\ldots\oplus W^{\bar{r}}$ is a Lie subalgebra of $\flat$, where $W^1=W, W^i=[W^1,W^{i-1}]$, $i=2,\ldots,\bar{r}$, and $\bar{r}\in[1,r]$ is the largest integer such that $W^{\bar{r}}\neq 0$. It is obvious that $\bar{\flat}$ is a Carnot algebra and $\bar{\mathbb{G}}:=\exp(\bar{\flat})$ is a Carnot subgroup of $\mathbb{G}$ (we regard Euclidean spaces as abelian Carnot groups). We use $\bar{\flat}(W)$ (resp. $\bar{\mathbb{G}}(W)$) to indicate the Carnot subalgebra (resp. Carnot subgroup) generated by W. The reduction to Carnot subgroups with lower-dimensional first layer is one of main tricks in this paper.

Recall that the differential of the exponential mapping is given by the following formula

$$\mathrm{d}\exp(e)=\mathrm{Id}-\sum_{m=2}^{r}\frac{(-1)^m}{m!}\mathrm{ad}(e)^{m-1}, \tag{2.1}$$

see *e.g.* [31]. For an absolutely continuous curve $\Upsilon:[0,1]\to\mathbb{G}$, we denote by γ the corresponding curve $\exp^{-1}(\Upsilon)$ with values in the Lie algebra $\flat$. We have $\gamma=\Sigma_{i=1}^{r}\gamma^i$ where $\gamma^i=\pi^i(\gamma)$, $\pi^i:\flat\to V^i$ is the projection to the i-th layer. It is obvious that γ is also absolutely continuous. From (2.1) we get for a.e. $t\in[0,1]$

$$\begin{aligned}\dot{\Upsilon}(t)&=\dot{\gamma}(t)-\sum_{m=2}^{r}\frac{(-1)^m}{m!}\mathrm{ad}(\gamma(t))^{m-1}(\dot{\gamma}(t))\\&=\dot{\gamma}(t)-\sum_{m=2}^{r}\frac{(-1)^m}{m!}[\gamma(t),\dot{\gamma}(t)]_{m-1}.\end{aligned} \tag{2.2}$$

In the last formula we used the iterated Lie bracket which is defined by

$$[e,f]_m=\underbrace{[e,[e,[\ldots,[e}_{m\text{ times}},f],],\ldots,]\quad\text{and}\quad[e,f]_0=f.$$

From (2.2) we obtain that Υ is horizontal if and only if for a.e. $t\in[0,1]$

$$\pi^i\left(\dot{\gamma}(t)-\sum_{m=2}^{r}\frac{(-1)^m}{m!}[\gamma(t),\dot{\gamma}(t)]_{m-1}\right)=0,\quad i=2,\ldots,r.$$

We summarize as

Lemma 2.2. *An absolutely curve Υ in $\mathbb{G}$ is horizontal if and only if for a.e. $t\in[0,1]$*

$$\dot{\gamma}^i(t)=\sum_{m=2}^{r}\frac{(-1)^m}{m!}\pi^i([\gamma(t),\dot{\gamma}(t)]_{m-1}),\quad i=2,\ldots,r.$$

We denote by $\mathcal{H}^1$ the Sobolev type space of all horizontal curves $\Upsilon:[0,1]\to\mathbb{G}$ with square integrable derivatives. In the rest of the paper we assume all horizontal curves in $\mathbb{G}$ are in $\mathcal{H}^1$. For our purpose this assumption is not restrictive since all rectifiable curves can be arc-length parameterized. Combining (2.2) with Lemma 2.2, we have for a.e. $t\in[0,1]$

$$\dot{\Upsilon}(t)=\sum_{i=1}^{n_1}\dot{x}^i(t)X_i(\Upsilon(t)) \tag{2.3}$$

where $\gamma^1(t) = \Sigma_{i=1}^{n_1} x^i(t) e_i$. Define $\mathcal{P} : \mathbb{G} \to V^1$, $\mathcal{P}(p) = \pi^1(\exp^{-1}(p))$. The mapping $\mathcal{P}$ is just the projection $\mathbb{R}^n \ni (x_1, \ldots, x_{n_1}, \ldots, x_n) \to (x_1, \ldots, x_{n_1}) \in \mathbb{R}^{n_1}$ when we identify $\mathbb{G}$ as $(\mathbb{R}^n, \cdot)$. The formula (2.3) in particular implies that $\mathcal{P}$ is a Riemannian submersion from $(\mathbb{G}, g_r)$ to $(V^1, \langle\cdot,\cdot\rangle)$ (or equivalently from $(\mathbb{R}^n, g_r)$ to $(\mathbb{R}^{n_1}, \langle\cdot,\cdot\rangle)$) with the property that for any $p \in G$, $\mathcal{P}_{\star,p}(X_i(p)) = e_i$, $i = 1, \ldots, n_1$, and $\mathcal{P}_{\star,p}(X_j(p)) = 0$ for $j = n_1 + 1, \ldots, n$.

Note that the graded condition (1.1) for the Lie algebra $\flat$ is equivalent to the following condition

$$V^i = \underbrace{[V^1, [\ldots, [V^1, V^1]]]}_{i \text{ times}}, \quad i = 2, \ldots r, \quad \text{and } V^j = 0 \text{ for } j > r, \tag{2.4}$$

which together with Lemma 2.2 implies that

$$\dot{\gamma}^i = \sum_{m=2}^{i} \frac{(-1)^m}{m!} \left(\sum_{j_1+j_2+\ldots+j_m=i} \left[\gamma^{j_1}, \left[\gamma^{j_2}, \left[\ldots, \left[\gamma^{j_{m-1}}, \dot{\gamma}^{j_m}\right]\right]\right]\right] \right) \quad \text{for } i = 2, \ldots, r, \text{ a.e.}, \tag{2.5}$$

which means that γ^1 determines $\gamma^2, \gamma^3, \ldots, \gamma^r$ recursively. We list $\dot{\gamma}^2$, $\dot{\gamma}^3$, $\dot{\gamma}^4$ as functions of γ^1:

$$\begin{cases} \dot{\gamma}^2 = \frac{1}{2}\left[\gamma^1, \dot{\gamma}^1\right] \\ \dot{\gamma}^3 = \frac{1}{2}\left\{\left[\gamma^1, \dot{\gamma}^2\right] + \left[\gamma^2, \dot{\gamma}^1\right]\right\} - \frac{1}{6}\left[\gamma^1, \left[\gamma^1, \dot{\gamma}^1\right]\right] \\ \dot{\gamma}^4 = \frac{1}{2}\left\{\left[\gamma^1, \dot{\gamma}^3\right] + \left[\gamma^2, \dot{\gamma}^2\right] + \left[\gamma^3, \dot{\gamma}^1\right]\right\} - \frac{1}{6}\left\{\left[\gamma^1, \left[\gamma^1, \dot{\gamma}^2\right]\right] + \left[\gamma^1, \left[\gamma^2, \dot{\gamma}^1\right]\right]\right. \\ \qquad \left. + \left[\gamma^2, \left[\gamma^1, \dot{\gamma}^1\right]\right]\right\} + \frac{1}{24}\left[\gamma^1, \dot{\gamma}^1\right]_3. \end{cases} \tag{2.6}$$

Sometimes we will abuse the notation $\Upsilon = (\gamma^1, \ldots, \gamma^r)$ or $\Upsilon = \Sigma_{i=1}^r \gamma^i$.

Proposition 2.3. (1) *Given $p \in \mathbb{G}$. Every absolutely continuous curve $\gamma^1 : [0,1] \to V^1$ has a unique horizontal lift $\Upsilon = (\gamma^1, \ldots, \gamma^r) : [0,1] \to \mathbb{G}$ determined by* (2.5) *with $\Upsilon(0) = p$. They have the same length and same regularity or smoothness;* (2) *Horizontal lifts of every straight line (or its interval) in V^1 are subriemannian minimizing geodesics.*

Proof. (1) Note that the class of absolutely continuous curves in $\flat$ is just the Sobolev class $W^{1,1}([0,1], \flat)$. This implies γ^1 is continuous and thus bounded in $[0,1]$ with $\dot{\gamma}^1 \in L^1([0,1], \flat)$, see *e.g.* [14], Chapter 2. So there exists $\gamma^2 \in W^{1,1}([0,1], \flat)$ such that $\gamma^2(t) = \frac{1}{2}\int_0^t [\gamma^1, \dot{\gamma}^1] \mathrm{d}\tau + \pi^2(\exp^{-1} p)$. Continuing this process, and noting that each summand in the right hand side of (2.5) contains only one term with derivative and other terms are bounded, that is, the right hand side of (2.5) is in $L^1([0,1], \flat)$. Thus the function γ^i satisfying

$$\gamma^i(t) = \int_0^t \sum_{m=2}^{i} \frac{(-1)^m}{m!} \left(\sum_{j_1+j_2+\ldots+j_m=i} \left[\gamma^{j_1}, \left[\gamma^{j_2}, \left[\ldots, \left[\gamma^{j_{m-1}}, \dot{\gamma}^{j_m}\right]\right]\right]\right] \right) \mathrm{d}\tau + \pi^i(\exp^{-1} p)$$

is in $W^{1,1}([0,1], \flat)$. From (2.3) γ^1 and its lift above have the same length.

To see (2), we recall by definition that in $(\mathbb{G}, g_r)$ Riemannian geodesics which are horizontal must be subriemannian geodesics. Since the geodesics of Euclidean space $(V^1, \langle\cdot,\cdot\rangle)$ are straight lines (or their intervals) , their horizontal lifts are Riemannian geodesics because $\mathcal{P}$ is a Riemannian submersion, see [28]. The minimizing property is obvious. $\square$

By Proposition 2.3 we sometimes do not distinguish a horizontal curve $\Upsilon = (\gamma^1, \ldots, \gamma^r)$ with its projection to the first layer γ^1.

Note that horizontal lifts of straight lines in V^1 are not necessarily still a line (looking in $\flat$) if $\mathbb{G}$ with step ≥ 3. In fact, let $\gamma^1(t) = vt + v_0$ with $v, v_0 \in V^1$. By the formula (2.6) we have

$$\begin{cases} \dot{\gamma}^2(t) = \dfrac{1}{2}[v, v_0] \\ \dot{\gamma}^3(t) = \dfrac{1}{6}[v,[v,v_0]]t - \dfrac{5}{12}[v_0,[v_0,v]] + \dfrac{1}{4}\left[\gamma^2(0), v\right]. \end{cases}$$

So the third layer of the lift is not a line unless $[v,[v,v_0]] = 0$. While if $v_0 = 0$, that is, the line passes through the origin, its lift is just itself.

2.2. The end-point mapping

Given $p, q \in \mathbb{G}$, deonte by $\Omega(p)$ the Hilbert manifold of all horizontal curves $\Upsilon \in \mathcal{H}^1$ with $\Upsilon(0) = p$. Let $\Omega(p,q) = \{\Upsilon \in \Omega(p) : \Upsilon(1) = q\}$. Since $\Omega(p) = p \cdot \Omega(0) := \{p \cdot \Upsilon : \Upsilon \in \Omega(0)\}$, $\Omega(p,q) = p \cdot \Omega(0, p^{-1} \cdot q)$ and the metric g_c is left-invariant, it suffices to consider horizontal curves emanating from the unit. Here we abuse 0 to denote the unit of $\mathbb{G}$. From Proposition 2.3, we see that the projection $\mathcal{P}$ gives a bijective mapping (still denoted by $\mathcal{P}$) from $\Omega(0)$ to $H^1(0) := \{\gamma^1 \in H^1([0,1], V^1) : \gamma^1(0) = 0\}$ with the mapping of horizontal lift as its inverse, where $H^1([0,1], V^1)$ denote the Sobolev space of all absolutely curves $\gamma^1 : [0,1] \to V^1$ with square integrable derivatives. Note from the formula (2.5), we have for $\Upsilon \in \Omega(0), t \in [0,1]$,

$$\Upsilon(t) = \left(\gamma^1(t), \gamma^2(t), \ldots, \gamma^r(t)\right)$$

where $\gamma^i(t), i = 2, \ldots, r$ is regard as a mapping $F^{i,t}$ (defined recursively) from $H^1(0)$ to V^i:

$$F^{i,t}\left(\gamma^1\right) = \int_0^t \sum_{m=2}^{i} \frac{(-1)^m}{m!} \left(\sum_{j_1+j_2+\ldots+j_m=i} \left[\gamma^{j_1}, \left[\gamma^{j_2}, \left[\ldots, \left[\gamma^{j_{m-1}}, \dot{\gamma}^{j_m}\right]\right]\right]\right] \right) \mathrm{d}\tau. \tag{2.7}$$

Now the original end-point mapping

$$\mathrm{end} : \Omega(0) \ni \Upsilon \to \Upsilon(1) \in \mathbb{G} \tag{2.8}$$

can be interpreted as

$$\mathscr{E} : H^1(0) \ni \gamma \to \left(F^1\left(\gamma^1\right), F^2\left(\gamma^1\right), \ldots, F^r\left(\gamma^1\right)\right) \in \flat \tag{2.9}$$

where $F^1(\gamma^1) = \gamma^1(1), F^i = F^{i,1}, i = 2, \ldots, r$.

Noting that given $\exp \xi = q \in \mathbb{G}$ for $\xi \in \flat$, $\Omega(0,q) = \exp(\mathscr{E}^{-1}(\xi))$, the subriemannian geodesic problem in $\mathbb{G}$

$$\min_{\Upsilon \in \Omega(0,q)} \frac{1}{2} \int_0^1 g_c(\dot{\Upsilon}, \dot{\Upsilon}) \mathrm{d}t$$

is equivalent to the minimizing problem with equality constraint

$$\min_{\mathscr{E}(\gamma^1)=\xi} \frac{1}{2} \int_0^1 \left|\dot{\gamma}^1\right|^2 \mathrm{d}t. \tag{2.10}$$

By the Cauchy-Schwarz inequality the problem of minimizing the energy functional is equivalent to that of minimizing the length functional. The existence of the subriemannian geodesic problem even for general subriemannian manifolds can be obtained by an argument of direct method in calculus of variations, see *e.g.* Appendix D in [27] where one also will find the ordinary formulation of the end-point mapping. What we are concerned with is their smoothness. The refined mapping $\mathscr{E} : H^1(0) \to \flat$ from a Hilbert space to a vector space will help us much.

3. The calculus of the end-point mapping

3.1. The generalized morse index theorem

To begin some computation, let us first see what we need according to the theory of generalized Morse index which we will later resort to. Let $(X, \|\cdot\|)$ be a Banach space, $L : X \to \mathbb{R}$, and $F : X \to Y$ with Y a finite dimensional vector space, be C^2 Fréchet differentiable mappings. Given $y_0 \in Y$, consider the minimizing problem with equality constraint

$$\min_{F(x)=y_0} L(x). \tag{3.1}$$

The Lagrange Multiplier Rule states that if $x \in X$ is a solution of (3.1) then there exists a nontrivial couple $(\lambda^0, \lambda^\star) \in \mathbb{R} \times Y^\star$ such that $\lambda^0 \mathrm{d}_x L + \lambda^\star \mathrm{d}_x F = 0$, where $\mathrm{d}_x L$ (resp. $\mathrm{d}_x F$) denotes the Fréchet derivative of L (resp. F) at the point x. In other words, the point x is a singular point of the augmented end-point mapping $\mathcal{L} : X \ni y \to (L(y), F(y)) \in \mathbb{R} \times Y$ and

$$\tilde{\lambda} \cdot \mathrm{d}_x \mathcal{L} = \lambda^0 \mathrm{d}_x L + \lambda^\star \mathrm{d}_x F = 0 \text{ with } \tilde{\lambda} = (\lambda^0, \lambda^\star). \tag{3.2}$$

The abnormal case $\lambda^0 = 0$ arises exactly when $\mathrm{Im}(\mathrm{d}_x F) \neq Y$, *i.e.*, x is a singular point of F. In this case we call $(x, \lambda^\star)$ an abnormal extremal. In the regular case $\mathrm{Im}(\mathrm{d}_x F) = Y$, we take $\lambda^0 = 1$. For $x \in X$ if there exists $\lambda^\star$ such that

$$\mathrm{d}_x L + \lambda^\star \mathrm{d}_x F = 0, \tag{3.3}$$

we call $(x, \lambda^\star)$ a normal extremal. The definition of normal extremals is equivalent to the one given in the Introduction. In the theory of subriemannian geodesics there exists a correspondence between $\lambda^\star$ in (3.3) and the Hamiltonian lift $\lambda(t)$ in (1.2), see [20] or [27], Chapter 5. We remark that an abnormal extremal may be normal by choosing a suitable multiplier (or a Hamiltonian lift). Those abnormal extremals which can not be normal for any multiplier are called strictly abnormal extremals.

The corank of x is defined as the codimension of $\mathrm{Im}(\mathrm{d}_x \mathcal{L})$. The following theorem, which gives necessary/sufficient conditions of optimality for a singular point is enough for our purpose, for the general versions see [4], Chapter 20.

Theorem 3.1 ([2, 4–6])**.** *If x is a local minimizer in X of the minimizing problem (3.1), of corank N, then for the nontrivial pair of Lagrange multiplier $\tilde{\lambda} = (\lambda^0, \lambda^\star)$ ($\lambda^0 = 0$ or 1) satisfying (3.2), the Morse index of the quadratic form $\tilde{\lambda} \cdot \mathrm{d}_x^2 \mathcal{L}$ restricted to $\ker \mathrm{d}_x F$ is less than or equal to $N - 1$.*

We recall that the Morse index of a quadratic form is the maximal dimension of subspaces on which the quadratic form is negative definite. Theorem 3.1 is classical for the regular case for which (3.3) is the Euler-Lagrange equation, see [24].

3.2. The differential of the end-point mapping

In the next section we will derive second order necessary conditions for minimality of abnormal extremals of the problem (2.10) from the finiteness of the Morse index of the quadratic form $\tilde{\lambda} \cdot \mathrm{d}_x^2 \mathcal{L}$ stated in Theorem 3.1. In the following we do some computation of the differential of the end-point mapping.

Lemma 3.2. *Given $\Upsilon = (\gamma^1, \ldots, \gamma^r) \in \Omega(0)$, then $T_\Upsilon \Omega(0) = T_{\gamma^1} H^1(0) = H^1(0)$.*

Proof. $H^1(0)$ is a Hilbert space. Let $\phi \in H^1(0)$. The family of horizontal lifts of $\gamma^1 + \epsilon\phi$ ($\epsilon \in [-\epsilon_0, \epsilon_0]$),

$$\Upsilon_\epsilon(t) = \left(\gamma^1(t) + \epsilon\phi(t), F^{2,t}(\gamma^1 + \epsilon\phi), \ldots, F^{r,t}(\gamma^1 + \epsilon\phi)\right)$$

where $F^{i,t}, i = 2 \ldots, r$, is defined as in (2.7), is a smooth curve in $\Omega(0)$ with $\Upsilon_0 = \Upsilon$.

On the other hand, if Υ_ϵ is a smooth family in $\Omega(0)$ with $\Upsilon_0 = \Upsilon$, then by Lemma 2.2 and (2.5), (2.7) we have

$$\Upsilon_\epsilon(t) = \left(\gamma^1_\epsilon, F^{2,t}(\gamma^1_\epsilon), \ldots, F^{r,t}(\gamma^1_\epsilon)\right)$$

where $\gamma^1_\epsilon = \mathcal{P}(\Upsilon_\epsilon)$ is a smooth family in $H^1(0)$ with $\gamma^1_0 = \gamma^1 = \mathcal{P}(\Upsilon)$. $\square$

For $\gamma^1, \phi \in H^1(0)$, the differential at γ^1 of the end-point mapping $\mathscr{E}$ is

$$\mathrm{d}_{\gamma^1}\mathscr{E} : H^1(0) \ni \phi \to \mathrm{d}_{\gamma^1}\mathscr{E}(\phi) = \big(\phi(1), \mathrm{d}_{\gamma^1}F^2(\phi), \dots, \mathrm{d}_{\gamma^1}F^r(\phi)\big) \in T_q\mathbb{G} \tag{3.4}$$

where $q = (\gamma^1(1), F^2(\gamma^1), \dots, F^r(\gamma^1))$ and F^i is shortened for $F^{i,1}$, $i = 2, \dots, r$. Noting that the differential $\mathrm{d}_{\gamma^1}F^i(\phi)$, $i = 2, \dots, r$, recursively depends on $\mathrm{d}_{\gamma^1}\dot{F}^{j,t}(\phi)$, $\mathrm{d}_{\gamma^1}F^{j,t}(\phi), j = 2, \dots, i-1$, $t \in (0, 1]$, its computation is complicated for $i \geq 5$. We restrict to the case of step ≤ 3 partly also because of technical difficulties in the next two sections. Observe that $\frac{\mathrm{d}}{\mathrm{dt}}\mathrm{d}_{\gamma^1}F^{i,t}(\phi) = \mathrm{d}_{\gamma^1}\dot{F}^{i,t}(\phi)$. From (2.6)–(2.7) we have for step=2

$$\begin{cases} F^{2,t}(\gamma^1) = \dfrac{1}{2}\displaystyle\int_0^t [\gamma^1, \dot{\gamma}^1]\mathrm{d}\tau \\ \mathrm{d}_{\gamma^1}F^{2,t}(\phi) = \displaystyle\int_0^t [\phi, \dot{\gamma}^1]\mathrm{d}\tau + \dfrac{1}{2}[\gamma^1(t), \phi(t)] \\ \mathrm{d}_{\gamma^1}\dot{F}^{2,t}(\phi) = \dfrac{1}{2}[\phi(t), \dot{\gamma}^1(t)] + \dfrac{1}{2}[\gamma(t), \dot{\phi}(t)]. \end{cases} \tag{3.5}$$

For step $= 3$, from (2.6) we first have

$$\dot{F}^{3,t} = \dot{\gamma}^3(t) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{dt}}\left[F^{2,t}, \gamma^1(t)\right] + \frac{1}{3}\left[\gamma^1(t), \left[\gamma^1(t), \dot{\gamma}^1(t)\right]\right],$$

then using (3.5) get

$$\begin{cases} \mathrm{d}_{\gamma^1}F^{3,t}(\phi) = \displaystyle\int_0^t \left[\gamma^1, \left[\phi, \dot{\gamma}^1\right]\right]\mathrm{d}\tau + \frac{1}{2}\left(\left[F^{2,t}, \phi\right] + \left[\mathrm{d}_{\gamma^1}F^{2,t}(\phi), \gamma^1\right]\right) + \frac{1}{3}\left[\gamma^1(t), \phi(t)\right]_2 \\ \qquad = \displaystyle\int_0^t \left[\gamma^1, \left[\phi, \dot{\gamma}^1\right]\right]\mathrm{d}\tau + \frac{1}{2}\left[\int_0^t \left[\phi, \dot{\gamma}^1\right]\mathrm{d}\tau, \gamma^1(t)\right] + \frac{1}{4}\left[\int_0^t \left[\gamma^1, \dot{\gamma}^1\right]\mathrm{d}\tau, \phi(t)\right] \\ \qquad\quad + \dfrac{1}{12}\left[\gamma^1(t), \left[\gamma^1(t), \phi(t)\right]\right]. \end{cases} \tag{3.6}$$

In the above computation we used integration by parts, skew-symmetry and Jacobi identity of Lie brackets to arrange terms.

Lemma 3.3. *In the case $r = 3$, by (3.4)–(3.6) we have $\phi \in \ker\left(\mathrm{d}_{\gamma^1}\mathscr{E}\right)$ if and only if*

$$\left.\begin{aligned} \phi(1) = 0 \\ \int_0^1 [\phi, \dot{\gamma}^1]\mathrm{d}t = 0 \\ \int_0^1 [\gamma^1, [\phi, \dot{\gamma}^1]]\mathrm{d}t = 0 \end{aligned}\right\}. \tag{3.7}$$

Now we compute the second Fréchet derivative of the end-point mapping $\mathscr{E}$ for $r = 3$. From (3.5) we have

$$\mathrm{d}^2_{\gamma^1}F^2(\phi, \phi) = \int_0^1 [\phi, \dot{\phi}]\mathrm{d}t. \tag{3.8}$$

For $\phi \in \ker\left(\mathrm{d}_{\gamma^1}\mathscr{E}\right)$ it follows from (3.6) and (3.7) that

$$\mathrm{d}^2_{\gamma^1}F^3(\phi, \phi) = \int_0^1 [\phi, [\phi, \dot{\gamma}^1]]\mathrm{d}t + \int_0^1 \left[\gamma^1 - \frac{1}{2}\gamma^1(1), [\phi, \dot{\phi}]\right]\mathrm{d}t. \tag{3.9}$$

So the intrinsic quadratic mapping (see *e.g.* [4], pp. 294–296) of $\mathscr{E}$ for $r = 3$

$$\mathrm{d}^2_{\gamma^1}\mathscr{E} : \ker\left(\mathrm{d}_{\gamma^1}\mathscr{E}\right) \times \ker\left(\mathrm{d}_{\gamma^1}\mathscr{E}\right) \to \flat$$

is

$$\mathrm{d}^2_{\gamma^1}\mathscr{E}(\phi,\phi) = \left(0, \int_0^1 [\phi,\dot\phi]\mathrm{d}t, \int_0^1 [\phi,[\phi,\dot\gamma^1]]\mathrm{d}t + \int_0^1 \left[\gamma^1 - \frac{1}{2}\gamma^1(1), [\phi,\dot\phi]\right]\mathrm{d}t\right). \tag{3.10}$$

4. Goh condition and Legendre-Jacobi condition

In the rest of the paper we assume the Carnot group has step ≤ 3. From (3.5) and (3.6) we have for $\phi \in H^1(0)$ with $\phi(1) = 0$

$$\mathrm{d}_{\gamma^1}\mathscr{E}(\phi) = \left(0, \int_0^1 [\phi,\dot\gamma^1]\mathrm{d}t, \int_0^1 \left[\gamma^1 - \frac{1}{2}\gamma^1(1), [\phi,\dot\gamma^1]\right]\mathrm{d}t\right). \tag{4.1}$$

Applying the Lagrange Multiplier Rule to the problem (2.10) we get

Proposition 4.1. *If $\gamma^1 \in H^1(0)$ is a minimizer of the problem (2.10), then there exists $\tilde\lambda = (\lambda^1,\lambda^2,\lambda^3) \in \flat^\star$ with $\lambda^i \in (V^i)^\star$, $i = 1,2,3$, such that for any $\phi \in H^1(0)$*

$$\int_0^1 < \dot\gamma^1, \dot\phi > + \lambda^2\left[\phi,\dot\gamma^1\right] + \lambda^3\left[\gamma^1 - \frac{1}{2}\gamma^1(1), \left[\phi,\dot\gamma^1\right]\right]\mathrm{d}t = 0 \quad \textit{with} \quad \phi(1) = 0 \tag{4.2}$$

or

$$\tilde\lambda \mathrm{d}_{\gamma^1}\mathscr{E}(\phi) = \lambda^1\phi(1) + \lambda^2\mathrm{d}_{\gamma^1}F^2(\phi) + \lambda^3\mathrm{d}_{\gamma^1}F^3(\phi) = 0 \quad \textit{with} \quad \tilde\lambda \neq 0. \tag{4.3}$$

Proof. Taking L as the energy functional $L(\gamma^1) = \frac{1}{2}\int_0^1 |\dot\gamma^1|^2\mathrm{d}t$, by the Lagrange Multiplier Rule there exists nontrivial $(\lambda^0,\tilde\lambda) \in \mathbb{R}\times\flat^\star$ with $\lambda^0 = 0$ or 1 and $\tilde\lambda = (\lambda^1,\lambda^2,\lambda^3)$ such that for any $\phi \in H(0)$

$$\lambda^0\mathrm{d}_{\gamma^1}L(\phi) + \lambda^1\phi(1) + \lambda^2\mathrm{d}_{\gamma^1}F^2(\phi) + \lambda^3\mathrm{d}_{\gamma^1}F^3(\phi) = 0.$$

When $\lambda^0 = 1$, (4.2) follows from the last formula, (4.1) and $\mathrm{d}_{\gamma^1}L(\phi) = \int_0^1\langle\dot\gamma^1,\dot\phi\rangle\mathrm{d}t$. □

We call γ^1 (or its horizontal lift) satisfying (4.2) for some (λ^2,λ^3) a normal geodesic. By a standard argument from the theory of (elliptic) differential equations normal geodesics are smooth, see *e.g.* [14]. Singular geodesics are local minimizers γ^1 (or their horizontal lifts) of the problem (2.10) satisfying (4.3) for some $\tilde\lambda = (\lambda^1,\lambda^2,\lambda^3) \in \flat^\star$. For singular geodesics the following result follows from Theorem 3.1 and (3.10).

Proposition 4.2. *If $\gamma^1 \in H^1(0)$ is a singular minimizer of the problem (2.10), there exists a nontrivial $\tilde\lambda = (\lambda^1,\lambda^2,\lambda^3)$ satisfying (4.3) such that the Morse index of the quadratic form*

$$\tilde\lambda\mathrm{d}^2_{\gamma^1}\mathscr{E}(\phi,\phi) = \int_0^1 \lambda^2[\phi,\dot\phi]\mathrm{d}t + \int_0^1 \lambda^3\left[\gamma^1 - \frac{1}{2}\gamma^1(1), [\phi,\dot\phi]\right]\mathrm{d}t + \int_0^1 \lambda^3[\phi,[\phi,\dot\gamma^1]]\mathrm{d}t \qquad \left(\phi \in \ker\left(\mathrm{d}_{\gamma^1}\mathscr{E}\right)\right) \tag{4.4}$$

is finite.

Proposition 4.3. *Let $\gamma^1, \tilde\lambda = (\lambda^1,\lambda^2,\lambda^3)$ be as in Proposition 4.2. Assume γ^1 is parameterized proportionally to arc-length. Then*

$$\begin{cases} \lambda^2 = 0 \\ \lambda^3\left[\gamma^1(t), [a,b]\right] = 0, \quad \forall a,b \in V^1, \quad \forall t \in [0,1]. \end{cases} \tag{4.5}$$

Proof. The argument is similar to [4], Proposition 20.13. The idea is a type of scaling or blowing up method.

Let $\bar{\tau} \in [0,1]$ be a Lebesgue point of $\dot{\gamma}^1$. Take a smooth mapping $c : \mathbb{R} \to V^1$ with support on $[0, 2\pi]$ such that $\int_0^{2\pi} c(s)\mathrm{d}s = 0$. Let $\phi_{c,\epsilon}(\tau) = \int_0^{\tau} c(\frac{\tilde{\tau}-\bar{\tau}}{\epsilon})\mathrm{d}\tilde{\tau}$ with ϵ small. This certainly implies $\dot{\phi}_{c,\epsilon}(\tau) = c(\frac{\tau-\bar{\tau}}{\epsilon})$ for $\tau \in [0,1]$ and $\phi_{c,\epsilon} \in H^1(0)$ with $\phi_{c,\epsilon}(1) = 0$. Letting $w(s) = \int_0^s c(\tilde{s})\mathrm{d}\tilde{s}$, we have

$$\int_0^1 \lambda^2[\phi_{c,\epsilon}, \dot{\phi}_{c,\epsilon}]\mathrm{d}\tau = \epsilon^2 \int_0^{2\pi} \lambda^2\left[w(s), c(s)\right]\mathrm{d}s \tag{4.6}$$

and similarly

$$\int_0^1 \lambda^3\left[\gamma^1 - \frac{1}{2}\gamma^1(1), \left[\phi_{c,\epsilon}, \dot{\phi}_{c,\epsilon}\right]\right]\mathrm{d}t = \epsilon^2 \int_0^{2\pi} \lambda^3\left[\gamma^1(\bar{\tau}) - \frac{1}{2}\gamma^1(1), [w(s), c(s)]\right]\mathrm{d}s + \epsilon^3 O(1) \tag{4.7}$$

where we used the fact $\dot{\gamma} \in L^\infty$ and thus $\gamma^1(\bar{\tau} + \epsilon s) = \gamma^1(\bar{\tau}) + \epsilon O(1)$ for ϵ small enough. For the last term in (4.4) we have

$$\int_0^1 \lambda^3\left[\phi_{c,\epsilon}, \left[\phi_{c,\epsilon}, \dot{\gamma}^1\right]\right]\mathrm{d}t = \epsilon^3 \int_0^{2\pi} \lambda^3\left[w(s), \left[w(s), \dot{\gamma}^1(\bar{\tau} + \epsilon s)\right]\right]\mathrm{d}s = \epsilon^3 O(1). \tag{4.8}$$

From (4.6)–(4.8), we get

$$\tilde{\lambda}\mathrm{d}^2_{\gamma^1}\mathscr{E}(\phi_{c,\epsilon}, \phi_{c,\epsilon}) = \epsilon^2 \int_0^{2\pi} \omega(w(s), c(s))\mathrm{d}s + \epsilon^3 O(1) \tag{4.9}$$

where $\omega(a,b) = \lambda^2[a,b] + \lambda^3\left[\gamma^1(\bar{\tau}) - \frac{1}{2}\gamma^1(1), [a,b]\right]$ is a skew-symmetric bilinear form on V^1.

We claim that $\omega \equiv 0$. In fact, if $\omega \neq 0$, then $\mathrm{rank}\,\omega = 2l_0 > 0$ and we can change the basis of V^1 such that

$$\omega(a,b) = \sum_{i=1}^{l_0}\left(x^i y^{i+l_0} - x^{i+l_0} y^i\right)$$

for any $b = (x^1, \ldots, x^{n_1}), a = (y^1, \ldots, y^{n_1}) \in V^1$. Now we take

$$c(s) = \left(x^1(s), 0, \ldots, 0, x^{l_0+1}(s), 0, \ldots, 0\right)$$

where $x^1(s) = \sum_{k=1}^\infty \xi_k \cos ks, x^{l_0+1}(s) = \sum_{k=1}^\infty \eta_k \sin ks$, and $(\xi_k)_{k=1}^\infty, (\eta_k)_{k=1}^\infty \in l^1$. Putting $c(s), w(s) = \int_0^s c(\tilde{s})\mathrm{d}\tilde{s}$ into (4.9) we get

$$\tilde{\lambda}\mathrm{d}^2_{\gamma^1}\mathscr{E}(\phi_{c,\epsilon}, \phi_{c,\epsilon}) = -\left(2\pi \sum_{k=1}^\infty \frac{1}{k}\xi_k\eta_k\right)\epsilon^2 + \epsilon^3 O(1).$$

From the last formula and the construction of $\phi_{c,\epsilon}$ it follows that there exists an infinite dimensional space $\mathcal{K}$ such that $\tilde{\lambda}\mathrm{d}^2_{\gamma^1}\mathscr{E}(\phi,\phi) < 0$ for each $\phi \in \mathcal{K}$. Note that $\mathcal{K} \cap \ker\left(\mathrm{d}_{\gamma^1}\mathscr{E}\right)$ is also infinite dimensional, since the rank of $\mathscr{E}$ is less than n. It implies the Morse index of $\tilde{\lambda}\mathrm{d}^2_{\gamma^1}\mathscr{E}$ is infinite. This is impossible by Proposition 4.2, so $\omega \equiv 0$.

We have proved that if $t \in [0,1]$ is a Lebesgue point of $\dot{\gamma}^1$, then

$$\lambda^2[a,b] + \lambda^3\left[\gamma^1(t) - \frac{1}{2}\gamma^1(1), [a,b]\right] = 0 \quad \text{for any } a, b \in V^1. \tag{4.10}$$

Since almost all points in $[0,1]$ are Lebesgue points of $\dot{\gamma}^1 \in L^\infty$ by the Lebesgue differentiation theorem, it follows from the continuity of γ^1 that (4.10) holds for any $t \in [0,1]$. In (4.10) letting $t = 0$ we get $\lambda^2[a,b] + \lambda^3\left[-\frac{1}{2}\gamma^1(1), [a,b]\right] = 0$ (since $\gamma^1(0) = 0$). Combing the last identity with (4.10), we obtain $\lambda^3\left[\gamma^1(t), [a,b]\right] = 0$ for any $t \in [0,1]$ and any $a, b \in V^1$. Applying the identity (4.10) again, we finally have $\lambda^2\xi = 0$ for any $\xi \in V^2$, since $[V^1, V^1] = V^2$. $\square$

The condition of (4.5) is called the Goh condition which first appeared in references of singular control theory, see [17]. A curve $\gamma^1 \in H^1(0)$ (or its horizontal lift) together with some nonzero $\tilde{\lambda} = (\lambda^1, \lambda^2, \lambda^3) \in \flat^\star$ satisfying the Goh condition is called a Goh curve and the pair $(\gamma^1, \tilde{\lambda})$ is called a Goh extremal. Note that for a Goh extremal $(\gamma^1, \tilde{\lambda})$, $\lambda^3 \neq 0$ and (4.3) automatically holds by choosing $\lambda^1 = 0$. The following fact is instructive for the study of subriemannian geodesics even for general case.

Corollary 4.4. *Assume $\mathbb{G}$ is a Carnot group of step 3, with a Carnot algebra $\flat = V^1 \oplus V^2 \oplus V^3$. Let W be a lower-dimensional subspace of V^1.*

(1) *If $[W, V^2] \subsetneq V^3$, then any curve $(H^1(0) \ni)\gamma^1 \subset W$ is a Goh curve;*
(2) *if $\flat$ is a free Carnot algebra, then any curve $(H^1(0) \ni)\gamma^1 \subset W$ is a Goh curve.*

Proof. (1) By assumption $\dim(V^3 \backslash [W, V^2]) \geq 1$. For any curve $(H^1(0) \ni)\gamma^1 \subset W$ choosing $\lambda^1 = 0, \lambda^2 = 0$ and $\lambda^3 \neq 0$ annihilating $[W, V^2]$, we conclude that $(\gamma^1, \tilde{\lambda})$ is a Goh extremal. The statement of (2) follows from (1) and the fact that for a free Carnot algebra, $[W, V^2] \subsetneq V^3$ always holds when $\dim W < \dim V^1$. □

Corollary 4.4 implies that each Carnot group of step 3 admits Goh curves. The following necessary condition is not used in this paper, but we include it for completeness.

Proposition 4.5. *Let $\gamma^1, \tilde{\lambda} = (\lambda^1, \lambda^2, \lambda^3)$ be as in Proposition 4.3. Then*

$$\lambda^3 \left[a, [a, \dot{\gamma}^1(t)]\right] \geq 0 \text{ for a.e. } \quad t \in [0,1] \quad \text{ and any } a \in V^1 \tag{4.11}$$

and

$$\tilde{\lambda} \mathrm{d}^2_{\gamma^1} \mathscr{E}(\phi, \phi) = \int_0^1 \lambda^3 \left[\phi, [\phi, \dot{\gamma}^1]\right] \mathrm{d}t \geq 0, \quad \phi \in \ker\left(\mathrm{d}_{\gamma^1}\mathscr{E}\right), \tag{4.12}$$

changing $\tilde{\lambda}$ to $-\tilde{\lambda}$ if necessary.

Proof. It suffices to prove that (4.11) holds for all Lebesgue points of $\dot{\gamma}^1$.

Let $\bar{\tau} \in [0,1]$ be a Lebesgue point of $\dot{\gamma}^1$. Assume that $\lambda^3 \left[\bar{a}, [\bar{a}, \dot{\gamma}^1(\bar{\tau})]\right] < 0$ for some $\bar{a} \in V^1$. We choose a suitable basis of V^1 to diagonalize the quadratic form

$$\lambda^3 \left[a, [a, \dot{\gamma}^1(\bar{\tau})]\right] = \sum_{j=1}^{n_1} \sigma_i (x^j)^2, \quad a = (x^1, \ldots, x^{n_1})$$

with at least one term $\sigma_i < 0$. For any smooth $x : \mathbb{R} \to \mathbb{R}^{n_1}$ with support in [0,1], let

$$\phi_x(t) = \left(\underbrace{0, \ldots, 0}_{(i-1) \text{ terms}}, x(t), 0, \ldots, 0 \right),$$

then we have

$$\tilde{\lambda} \mathrm{d}^2_{\gamma^1} \mathscr{E}(\phi_x, \phi_x) = \sigma_i \int_0^1 x^2(t) \mathrm{d}t < 0. \tag{4.13}$$

Denote by Π the set of all smooth mappings $x : [0,1] \to \mathbb{R}^{n_1}$ with support in [0,1] and satisfying (4.13). Π is infinitely dimensional, so is $\left\{\phi_x : \phi_x \in \ker\left(\mathrm{d}_{\gamma^1}\mathscr{E}\right), x \in \Pi\right\}$. It is a contraction by Proposition 4.2. □

(4.11) and (4.12) are called generalized Legendre-Jacobi condition.

5. Subriemannian geodesics of step 3

In this section we assume $\mathbb{G}$ is of step $= 3$.

Lemma 5.1. *Any line (or its interval) through 0 is a normal geodesic.*

Proof. For $\gamma^1(t) = Ctv_0$, where C a constant and $v_0 \in V^1$, we take $\lambda^2 = 0, \lambda^3 = 0$, then (4.2) holds for any $\phi \in H^1(0)$ with $\phi(1) = 0$. □

The following result on rank 2 case is well known, see *e.g.* [8]. For completeness we give a self-contained proof.

Theorem 5.2 (rank 2 case)**.** *Let G be of rank 2, i.e., $n_1 = 2$. Assume $V^1 = \mathrm{span}\{e_1, e_2\}$.*

(1) *In the case of the Engel group whose algebra $\flat = \mathrm{span}\{e_1, e_2\} \oplus \mathrm{span}\{e_3\} \oplus \mathrm{span}\{e_4\}$ with $[e_1, e_2] = e_3, [e_1, e_3] = e_4$, there is a unique arc-length parameterized singular geodesic γ^1 which is normal and tangent to e_2, that is, $\gamma^1(t) = te_2$;*
(2) *to the other case where $\flat = \mathrm{span}\{e_1, e_2\} \oplus \mathrm{span}\{e_3\} \oplus \mathrm{span}\{e_4, e_5\}$ with $[e_1, e_2] = e_3, [e_1, e_3] = e_4, [e_2, e_3] = e_5$, singular geodesics are exactly those lines (or their intervals) in V^1 through the origin.*

Proof. Let γ^1 be a singular geodesic parameterized proportionally to arc-length and satisfying (4.3) with some $\tilde{\lambda} = (\lambda^1, \lambda^2, \lambda^3)$. From (4.5) $\lambda^2 = 0$. This together with (4.3) implies that $\lambda^3 \neq 0$.

We claim that λ^1 must be in a line. In fact, if this is not true, there must exist $t_1, t_2 \in (0, 1]$ such that $V^1 = \mathrm{span}\{\gamma^1(t_1), \gamma^1(t_2)\}$ which together with (4.5) implies that $\lambda^3 = 0$ because $[V^1, V^2] = V^3$. A contradiction! So $\gamma^1(t) = (c^1e_1 + c^2e_2)t$ for constants c^1, c^2 with $(c^1)^2 + (c^2)^2 \neq 0$. In Proposition 2.3 we see they are shortest subriemannian geodesics.

(1) If $\flat$ is the Engel algebra, from $\lambda^3 \neq 0$ and $\lambda^3[c^1e_1t + c^2e_2t, e_3] = 0, \forall t \in [0, 1]$ we get $c^1 = 0$, since $[e_1, e_3] = e_4$ and $[e_2, e_3] = 0$. From (3.5) and (3.6) by direct computation we verify that $\gamma^1(t) = c^2e_2t$ ($c^2 \neq 0$) is a singular curve (choosing *e.g.* $\tilde{\lambda} = (0, 0, 1)$). By Lemma 5.1 it is normal;
(2) when $\flat$ is the free case, by direct computation we have $\mathrm{Im}\left(\mathrm{d}_{\gamma^1}\mathscr{E}\right) \neq \flat$. In fact, by (3.6) for any $\phi \in H^1(0)$ there exists a constant δ such that $\mathrm{d}_{\gamma^1}F^3(\phi) = \delta(c^1e_4 + c^2e_5)$. So all $\gamma^1(t) = (c^1e_1 + c^2e_2)t$ are singular geodesics. By Lemma 5.1 they are also normal. □

Lemma 5.3. *Let γ^1 (or its horizontal lift) be a subriemannian geodesic in $\mathbb{G}$ and be contained in a lower-dimensional subspace $W \subset V^1$. If γ^1 is a normal geodesic in the Carnot subgroup $\bar{\mathbb{G}}(W)$ of step 2 or 3, then γ^1 is also normal in $\mathbb{G}$.*

Proof. Assume $\bar{\mathbb{G}}(W)$ has step 3. Because $\gamma^1 \subset W$ is normal in $\bar{\mathbb{G}}(W)$, by (4.2) there exist $\mu \in [W, W]^\star$ and $\nu \in [W, [W, W]]^\star$ such that

$$\int_0^1 \left\langle \dot{\gamma}^1, \dot{\phi} \right\rangle + \mu[\phi, \dot{\gamma}] + \nu \left[\gamma^1 - \frac{1}{2}\gamma^1(1), [\phi, \dot{\gamma}^1]\right] \mathrm{d}t = 0$$

holds for any $\phi \in H^1([0, 1], W)$ with $\phi(0) = \phi(1) = 0$. Let V^2 (resp. V^3) be orthogonally decomposed as $[W, W] \oplus U^2$ (resp. $[W, [W, W]] \oplus U^3$). Now we take $\lambda^2 = \mu \in \left(V^2\right)^\star, \lambda^3 = \nu \in \left(V^3\right)^\star$, that is, we extend μ (resp. ν) to V^2 (resp. V^3) by annihilating U^2 (resp. U^3). It is obvious that (λ^2, λ^3) satisfies (4.2) for any $\phi \in H^1([0, 1], V^1)$ with $\phi(0) = \phi(1) = 0$. The case when $\bar{\mathbb{G}}(W)$ has step 2 is similar. □

Theorem 5.4 (general case)**.** *All subriemannian minimizers in* $\mathbb{G}$ *are normal.*

Proof.

(1) Let γ^1 be a singular geodesic which is parameterized proportionally to arc-length and satisfies (4.3) for some $\tilde{\lambda} = (\lambda^1, \lambda^2, \lambda^3)$.
From (4.5) and (4.3) we have $\lambda^3 \neq 0$ because $[V^1, V^1] = V^2, [V^1, V^2] = V^3$. We claim that γ^1 is contained in a lower-dimensional subspace W of V^1. Otherwise, there are $t_1, \ldots, t_{n_1} \in (0, 1]$ such that $V^1 = \mathrm{span}\{\gamma^1(t_1), \ldots, \gamma^1(t_{n_1})\}$ which together with (4.5) implies $\lambda^3 = 0$.
Thus γ_1 (or its horizontal lift) is a subriemannian geodesic in the Carnot subgroup $\bar{\mathbb{G}}(W)$ whose algebra is $\bar{\flat} = W \oplus [W, W] \oplus [W, [W, W]]$ or $W \oplus [W, W]$ or W;

(2) if $\bar{\flat} = W$, this implies that the horizontal lift of γ^1 is itself. By Proposition 2.3 the line through 0 and $\gamma^1(t_0)$ for any $t_0 \in (0, 1]$ is the shortest subriemannian geodesic. So γ^1 must be an interval of a line through 0;

(3) if $\bar{\flat} = W \oplus [W, W]$, then γ^1 is a subriemannian geodesic in a Carnot group of step 2. So γ^1 is normal in $\bar{\mathbb{G}}(W)$. From Lemma 5.3, γ^1 is also normal in $\mathbb{G}$;

(4) if $\bar{\flat} = W \oplus [W, W] \oplus [W, [W, W]]$, then γ^1 is also a subriemannian minimizer in $\bar{\mathbb{G}}(W)$. If γ^1 is regular in $\bar{\mathbb{G}}(W)$, then by Lemma 5.3 γ^1 is also normal in $\mathbb{G}$. If γ^1 is a singular geodesic in $\bar{\mathbb{G}}(W)$ and $\dim W \geq 3$, we repeat the procedure from step (1), with $\mathbb{G}$ (resp. $\flat$) replaced by $\bar{\mathbb{G}}(W)$ (resp. $\bar{\flat}$).

By finite steps we arrive at the case of rank 2. Our statement follows from Theorem 5.2. □

Remark 5.5. The smoothness of subriemannian geodesics is very close to the regularity of the subriemannian distance. In fact, the pointwise smoothness of the subriemannian distance depends on the strict normalness and uniqueness of subriemannian geodesics. The subanalyticity of the subriemannian distance (or sphere) was usually derived from the exclusivity of Goh curves. We refer the readers to [3,21] and references therein for this topic. In our case of step 3, as pointed out in Corollary 4.4, there typically exist Goh curves which are smooth even normal if they are shortest. Theorem 10 in [3] proved that the subriemannian distances of free Carnot groups of step 3 are not subanalytic.

Acknowledgements. Part of this work was done when the first author visited Department of Mathematics, University of Notre Dame. He would thank Professor Jianguo Cao for his help and thank the staff for their hospitality. We also thank the referee for useful comments and suggestions.

Note added in proof. After the paper was accepted for publication, we generalized the smoothness result to subriemannian manifolds of step 3 with a nilpotent basis, see `http://arxiv.org/abs/1202.4287`. In the latter article we noticed that the example by [18] is not a strictly abnormal minimizer. In fact we proved that all abnormal minimizers in rank 2 Carnot groups of step 4 must be normal.

References

[1] A. Agrachev and R.V. Gamkrelidze, Second order optimality condition for the time optimal problem. *Matem. Sbornik* **100** (1976) 610–643.

[2] A. Agrachev and R.V. Gamkrelidze, Symplectic methods for optimization and control, in *Geometry of Feedback and Optimal Control*, edited by B. Jacubczyk and W. Respondek. Marcel Dekker, New York (1997).

[3] A. Agrachev and J.-P. Gauthier, On subanalyticity of Carnot-Carathéodory distances. *Ann. Inst. Henri Poincaré Anal. Non Linéaire* **18** (2001) 359–382.

[4] A. Agrachev and Y. Sachkov, Control Theory from the Geometric Viewpoint, edited by Springer. *Encycl. Math. Sci.* **87** (2004).

[5] A. Agrachev and A. Sarychev, Abnormal sub-Riemannian geodesics: morse index and rigidity. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **13** (1996) 635–690.

[6] A. Agrachev and A. Sarychev, On abnormal extremals for Lagrange variational problems. *J. Math. Syst. Estim. Control* **8** (1998) 87–118.

[7] A. Agrachev and A. Sarychev, Sub-Riemannian metrics: minimality of abnormal geodesics *versus* sub-analyticity. *ESAIM: COCV* **4** (1999) 377–403.

[8] A. Agrachev, B. Bonnard, M. Chyba and I. Kupka, Subriemannian sphere in martinet flat case. *ESAIM: COCV* **2** (1997) 377–448.

[9] A. Bellaïche, The tangent space in sub-Riemannian geometry. *Sub-Riemannian Geometry*, *Progr. Math.* **144** (1996) 1–78.

[10] J.-M. Bismut, Large deviations and the Malliavin calculus, *Progr. Math.* **45** (1984).

[11] G.A. Bliss, *Lectures on the calculus of variations*. University of Chicago Press (1946).

[12] B. Bonnard and M. Chyba, *Singular Trajectories and Their Role in Control Theory*. Springer, Berlin (2003).

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

[14] G. Buttazzo, M. Giaquinta and S. Hildebrandt, One-dimensional variational problems. An introduction, Oxford Lecture Series. Edited by Univ. of Oxford Press, New-York. *Math. App.* **15** (1998).

[15] Y. Chitour, F. Jean and E. Trélat, Genericity results for singular curves. *J. Differ. Geom.* **73** (2006) 45–73.

[16] W.L. Chow, Über systeme von linearen partiellen differentialgleichungen erster Ordnung. *Math. Ann.* **117** (1940) 98–105.

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

[18] C. Golé and R. Karidi, A note on Carnot geodesics in nilpotent Lie groups. *J. Dyn. Control Syst.* **1** (1995) 535–549.

[19] U. Hamenstädt, Some regularity theorems for Carnot-Carathéodory metrics. *J. Differ. Geom.* **32** (1990) 819–850.

[20] L. Hsu, Calculus of variations *via* the Griffiths formalism. *J. Differ. Geom.* **36** (1991) 551–591.

[21] S. Jacquet, Subanalyticity of the sub-Riemannian distance. *J. Dyn. Control Syst.* **5** (1999) 303–328.

[22] G.P. Leonardi and R. Monti, End-point equations and regularity of sub-Riemannian geodesics. *Geom. Funct. Anal.* **18** (2008) 552–582.

[23] W.S. Liu and H.J. Sussmann, Shortest paths for sub-Riemannian metrics of rank two distributions, edited by American Mathematical Society, Providence, RI. *Mem. Amer. Math. Soc.* **118** (1995) 104.

[24] J. Milnor, Morse Theory, edited by Princeton University Press, Princeton, New Jersey. *Annals of Mathematics Studies* **51** (1963).

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

[26] R. Montgomery, Abnormal minimizers. *SIAM J. Control Optim.* **32** (1994) 1605–1620.

[27] R. Montgomery, A tour of subriemannian geometries, their geodesics and applications, edited by American Mathematical Society, Providence, RI. *Mathematical Surveys and Monographs* **91** (2002).

[28] B. O'Neill, Submersions and geodesics. *Duke Math. J.* **34** (1967) 363–373.

[29] P.K. Rashevsky, About connecting two points of a completely nonholonomic space by admissible curve. *Uch. Zapiski Ped. Inst. Libknechta* **2** (1938) 83–94.

[30] R.S. Strichartz, Sub-Riemannian geometry. *J. Differ. Geom.* **24** (1986) 221–263. [Corrections to Sub-Riemannian geometry. *J. Differ. Geom.* **30** (1989) 595–596].

[31] V.S. Varadarajan, *Lie groups, Lie algebras and their representation*. Springer-Verlag, New York (1984).

[32] L.C. Young, *Lectures on the calculus of variations and optimal control theory*. W.B. Saunders Co., Philadelphia-London-Toronto, Ont. (1969).