An introduction to sub-Riemannian structures

Yuxuan Song

May 2024

Prologue

This set of notes is aimed to be a gentle introduction to sub-Riemannian geometry and some of its metric-geometric aspects. The write-up is based on my own understanding of the subject which I have learnt from my master thesis supervisor Dr Karen Habermann and the excellent text by [Bel96]. The majority of the content here is adapted from a section in my masters thesis.

1 Sub-Riemannian metric compatible with distribution

Let M be a smooth n-manifold, $\mathcal{F} \subset TM$ a smooth distribution in the tangent bundle of M, which we refer to as the horizontal distribution. Locally, we can write $\mathcal{F} = \text{span}(\{X_i\}_{i=1}^k)$ for some locally generating vector fields $X_1, \ldots, X_k \in \mathfrak{T}_1(M)$. In general, we are not guaranteed with a globally generating frame for $\mathcal{F}$ for some global topological obstructions. An example would be an even-dimensional sphere $\mathbb{S}^{2m}$, which does not admit an everywhere non-vanishing continuous vector field as a consequence of the Hairy Ball theorem, see for instance [Mil78]. Notice that this construction does not necessarily require the horizontal distribution to be of constant rank, which would be necessary for some applications in geometric control theory. For simplicity and the purpose of a gentle introduction, we assume that our horizontal distribution has constant rank throughout. We write $X \in \Gamma(\mathcal{F})$ for vector field X on M such that $X|_x \in \mathcal{F}_x$ at any point $x \in M$. Similarly, for any $U \subset M$, we denote a local section $V \in \mathfrak{T}_1(U)$ with $V|_x \in \mathcal{F}_x$ for every $x \in U$ by $V \in \Gamma(U, \mathcal{F})$. When given a locally frame $(X_1, \ldots, X_k)$ for $\mathcal{F}$, we can construct a smooth fibre inner product on $\mathcal{F}$ by declaring $(X_1, \ldots, X_k)$ a local orthonormal frame:

Definition 1.1 (Local choice of sub-Riemannian metric)**.** *Let M be a smooth manifold with a rank k smooth distribution $\mathcal{F}$ in the tangent bundle TM. Choose smooth vector fields $X_1, \ldots, X_k \in \mathfrak{T}_1(M)$ such that locally $\mathcal{F} = \text{span}(\{X_1, \ldots, X_k\})$. Define for each $x \in M$ and $v \in T_xM$*

$$\tilde{g}_x(v) := \inf\{(\xi^1)^2 + \cdots + (\xi^k)^2 \mid \xi^i X_i|_x = v\}. \tag{1.1}$$

We adopt the convention that if $v \in T_xM \setminus \mathcal{F}_x$, then $\tilde{g}_x(v) = \infty$. Thus, for all $x \in M$, $\tilde{g}_x$ defines a positive-definite quadratic form on the vector subspace $\mathcal{F}_x \subset T_xM$. We can use the polarisation identity on each distribution plane $\mathcal{F}_x$ for $x \in M$ to define:

$$g_x(v, w) := \frac{1}{4}(\tilde{g}_x(v + w) - \tilde{g}_x(v - w)) \text{ for } v, w \in T_xM, \tag{1.2}$$

which gives rise to a smooth fibre inner product $g \in \Gamma(\mathcal{F}^* \otimes \mathcal{F}^*)$ on the distribution $\mathcal{F}$.

It is not immediate whether (1.1) defines a positive-definite quadratic form on $\mathcal{F}_x$ for each $x \in M$ as stated. We verify this by constructing the following linear map:

$$\mathcal{L}_x : \mathbb{R}^k \to T_xM, \quad (\xi_1, \ldots, \xi_k) \longmapsto \xi^i X_i|_x. \tag{1.3}$$

Note that $\mathcal{L}_x$ is a linear map between finite-dimensional vector spaces, thus $\ker \mathcal{L}_x$ defines a closed subspace. This induces an orthogonal decomposition of the Euclidean inner product space $(\mathbb{R}^k, \langle \cdot, \cdot \rangle_E)$ given by $\mathbb{R}^k = (\ker \mathcal{L}_x) \oplus (\ker \mathcal{L}_x)^\perp$, where $\langle \cdot, \cdot \rangle_E$

denotes the Euclidean inner product with induced norm $\|\cdot\|_E$ on $\mathbb{R}^k$. The restriction $\mathcal{L}_x|_{(\ker\mathcal{L}_x)^\perp}:(\ker\mathcal{L}_x)^\perp\to\mathcal{F}_x$ defines a linear isomorphism by construction. Now we consider its inverse $\mathcal{L}_x^{-1}:\mathcal{F}_x\to(\ker\mathcal{L}_x)^\perp$. Let $\xi\in(\ker\mathcal{L}_x)^\perp$ with $\mathcal{L}_x(\xi)=v$ for some $v\in\mathcal{F}_x$. Then by linearity and definition of ξ, we have $\xi-\mathcal{L}_x^{-1}(v)\in\ker\mathcal{L}_x$. The Pythagoras theorem and the orthogonal decomposition with respect to $\langle\cdot,\cdot\rangle_E$ then yield:

$$\|\mathcal{L}_x^{-1}(v)\|_E^2\leqslant\|\mathcal{L}_x^{-1}(v)\|_E^2+\|\xi-\mathcal{L}_x^{-1}(v)\|_E^2=\|\mathcal{L}_x^{-1}(v)+\xi-\mathcal{L}_x^{-1}(v)\|_E^2=\|\xi\|_E^2=(\xi^1)^2+\cdots+(\xi^k)^2.$$

Taking infimum over the set $\{(\xi^1)^2+\cdots+(\xi^k)^2\mid\xi^iX_i|_x=v\}$ in the above inequality and recalling that $\mathcal{L}_x^{-1}$ also defines a linear isomorphism, we obtain $\tilde{g}_x(v)\geqslant\|\mathcal{L}_x^{-1}(v)\|_E^2$, thus proving the positive-definiteness of $\tilde{g}_x$. The isomorphism also guarantees the existence of minimal coefficients $(\xi^i)_{i=1}^k$, thus the infimum in 1.1 is in fact a minimum. We should also check that $\tilde{g}$ satisfies the parallelogram identity on each fibre of $\mathcal{F}$, so that it is indeed induced from an inner product. Observe that for horizontal tangent vectors $v,w\in\mathcal{F}_x$, $(\xi^i)_{i=1}^k\in\mathbb{R}$ satisfies $\xi^iX_i|_x=v+w$ if and only if there are some choices of $(v^i)_{i=1}^k,\ (w^i)_{i=1}^k\in\mathbb{R}$ such that $v^iX_i|_x=v,\quad w^iX_i|_x=w$ and $v^i+w^i=\xi^i$ for each i. The analogous result also holds for $v-w=\mu^iX_i|_x$ using the same approach. We thus have for choices of constants as above:

$$\sum_{i=1}^k(\xi^i)^2+\sum_{j=1}^k(\mu^j)^2=\sum_{i=1}^k(v^i+w^i)^2+\sum_{j=1}^k(v^j-w^j)^2=2\sum_{i=1}^k((v^i)^2+(w^i)^2)$$

Taking the infimum over the set $\aleph=\{((v^i)_{i=1}^k,(w^i)_{i=1}^k)\mid v^iX_i|x=v,\ w^iX_i|_x=w\}$ gives:

$$\tilde{g}_x(v+w)+\tilde{g}_x(v-w)=\inf_{\aleph}\sum_{i=1}^k(v^i+w^i)^2+\inf_{\aleph}\sum_{j=1}^k(v^j-w^j)^2=\inf_{\aleph}2\sum_{i=1}^k((v^i)^2+(w^i)^2)=2(\tilde{g}_x(v)+\tilde{g}_x(w))$$

Thus, $\tilde{g}_x$ satisfies the parallelogram law, which can therefore be polarised to obtain an inner product on each fibre. Now, for a given open neighbourhood U in a smooth manifold M and a smooth distribution $F\subset TM$, we can choose some fibre inner product $\tilde{g}\in\Gamma(\mathcal{F}^*\otimes\mathcal{F}^*)$, locally defined relative to some a choice of generating frame $X_1,\ldots,X_k\in\Gamma(U,\mathcal{F})$. This can be extended to the whole $T^*M\otimes T^*M$ by zero, thus giving rise to a global section.
The above construction is valid in any open neighbourhood $U\subset M$. Now take an open cover (U_{ii}) for M and a partition of unity $\{\varphi_i\}_i$ subordinate to this cover. Choosing a smooth fibre inner product g_i for each U_i and extending them to a global section $g_i\in\Gamma(\mathcal{F}^*\otimes\mathcal{F}^*)$ by zero, we can now patch them together using the partition of unity to obtain a sensible choice of sub-Riemannian metric $g:=\sum_i\varphi_ig_i\in\Gamma(\mathcal{F}^*\otimes\mathcal{F}^*)$. Now we provisionally refer to a smooth distribution $\mathcal{F}$ in TM with an associated choice of smooth fibre inner product g a sub-Riemannian structure.

2 Sub-Riemannian distance and Chow's condition

Let M^n be a smooth manifold with equipped with a sub-Riemannian structure $(\mathcal{F},g)$. The associated fibre inner product g gives rise to a notion of length of vectors and distance between points on the manifold. Indeed, restricted to any fibre $\mathcal{F}_x$, g is obtained by polarising a norm on the distribution plane. The length of a tangent vector $v\in T_xM$ can thus be defined as $|v|_{g_x}:=\sqrt{g_x(v)}$.
An absolutely continuous path $c:[0,T]\to M$ is a horizontal path with respect to the distribution $\mathcal{F}$ if:

$$\dot{c}(t)\in\mathcal{F}_x\text{ a.e. }t\in[0,T]\text{ and }t\mapsto|\dot{c}(t)|_{g_{c(t)}}\in L^1([0,T]),\text{ where }\dot{c}(t)=c_*\frac{\partial}{\partial t}.\tag{2.1}$$

Any absolutely continuous path $c\colon[0,T]\to M$ can be rewritten as an absolutely continuous reparametrisation of a Lipschitz curve (cf. [ABBZ19, Chapter 3.6]). Thus by the Rademacher theorem, we have c is differentiable a.e. on $[0,T]$, which implies $\dot{c}(t)$ is well-defined for a.e. $t\in[0,T]$. We now define the sub-Riemannian length functional acting on absolutely continuous paths c on M:

$$\mathcal{L}_g(c) := \int_0^T |\dot{c}(t)|_{g_{c(t)}} \mathrm{d}t. \tag{2.2}$$

By construction, the above integral in (2.2) is finite whenever we have $|\dot{c}(t)|_{g_{c(t)}} \in \mathcal{F}_{c(t)}$ a.e. $t \in [0,T]$. Thus, it is clear that an absolutely continuous path $c\colon [0,T] \to M$ has finite length measured by the sub-Riemannian length functional if and only if c is a horizontal path. Similar to the construction in Riemannian geometry, a notion of distance between points $p, q \in M$ can be defined by using the length functional:

$$d_g(p,q) := \inf\{\mathcal{L}_g(c) \mid c\colon [0,T] \to M \text{ absolutely continuous path}, c(0) = p, c(T) = q\} \in [0,\infty]. \tag{2.3}$$

The criterion for an absolutely continuous path being horizontal, given an initial position, can be rephrased into an ODE problem. Indeed, consider a local g-orthonormal frame $(X_1, \dots, X_k)$ on $U \subset M$ for $\mathcal{F}$. Then, using (1.1) with our choice of orthonormal frame $(X_1, \dots, X_k)$, we can characterise an absolutely continuous path $c\colon [0,T] \to M$ as the solution c to the following differential equation with some measurable functions $(\xi^i)_{i=1}^k$ on $[0,T]$:

$$\dot{c}(t) = \xi^i(t) X_i|_{c(t)} \text{ a.e. } t \in [0,T]. \tag{2.4}$$

Then, assuming that the image of a given path $c\colon [0,T] \to M$ is contained in the neighbourhood where $(X_1, \dots, X_k)$ defines an orthonormal frame, the length of c can be equivalently formulated as:

$$\mathcal{L}_g^{loc}(c) = \int_0^T \sqrt{\xi^1(t)^2 + \cdots + \xi^k(t)^2} \mathrm{d}t.$$

Thus, the path c has finite length, and thus defines a horizontal path if and only if $xi^i \in L^1([0,T];\mathbb{R})$ for each $i = 1, \dots, k$. If $c([0,T]) \not\subset U$, the fixed choice of generating frame for $\mathcal{F}$ may not be linearly independent at each point on the path. This leads us to a different system in terms of the choice of measurable coefficient functions $(\xi^i)_{i=1}^k$, for which $|\dot{c}(t)|^2_{g_{c(t)}} = \sum_{i=1}^k \xi^i(t)^2$ should still hold for a.e. $t \in [0,T]$. Then for a given absolutely continuous path $c\colon [0,T] \to M$, we say c has finite length if there exists a choice of $(\xi^i)_{i=1}^k \in L^1([0,T];\mathbb{R})$ such that c defines a solution to (2.4) with our choice of $(xi^i)_{i=1}^k$. In particular, the length of the horizontal path c is given by:

$$\mathcal{L}_g^{loc}(c) = \inf\{\int_0^T \sqrt{\xi^1(t)^2 + \cdots + \xi^k(t)^2} \mathrm{d}t \mid (\xi^i)_{i=1}^k \in L^1([0,T];\mathbb{R}) \text{ satisfying (2.4)}\}. \tag{2.5}$$

Taking the infimum over the solution set to (2.4), subject to the boundary data $c(0) = p$ and $c(T) = q$, we recover the sub-Riemannian distance between two points. We can think of horizontal paths as curves in a manifold with restricted directions in which we can move in. Due to the constrained motion, our definition for the distance between two points can be infinite since the existence of a horizontal path connecting any two points $p, q \in M$ is not apriori guaranteed. We thus briefly discuss the notion of accessibility and introduce a condition on $\mathcal{F}$ to guarantee accessibility between two arbitrary points on the manifold. Let $p \in M$ be arbitrary. A point $q \in M$ is accessible from p if there exists a horizontal path $\gamma\colon I \to M$ with endpoints p and q. For a point $p \in M$ and a fixed time $T > 0$, we define the accessibility component A_p from each point $p \in M$:

$$\begin{cases} \mathcal{C}_{p,T} := \{c\colon [0,T] \to M \text{ horizontal path} \mid c(0) = p\}, \\ A_p := \{q \in M \mid \text{there exists a horizontal path } \gamma\colon [0,T] \to M \text{ with } \gamma(0) = p, \gamma(T) = q\}. \end{cases}$$

Every $c \in \mathcal{C}_{p,T}$ can be obtained as the solution to the ODE problem coupled with $(\xi^i)_{i=1}^k \in L^1([0,T])$:

$$\begin{cases} \dot{c}(t) = \xi^i(t) X_i|_{c(t)} \text{ on } [0,T] \\ c(0) = p. \end{cases} \tag{2.6}$$

By existence theory of ODEs, there exists an open neighbourhood $\mathcal{U}_{p,T} \subset L^1([0,T];\mathbb{R}^k)$ of the origin such that a well-defined solution $c_\xi(t)$ exists for $\xi = (\xi^1, \dots, \xi^k) \in L^1([0,T];\mathbb{R}^k)$. This motivates the following definition, which is a crucial element in our proof of the Chow-Rashevskii theorem:

Definition 2.1 (End-point map [Bel96])**.** *Fix $p \in M$ and a terminal time $T > 0$. We define the end-point map at p for time $T > 0$ by:*

$$E = E_{p,T} \colon \mathcal{U}_{p,T} \to M, (\xi^1, \dots, \xi^k) \longmapsto c_\xi(T).$$

Since we can reparametrise the paths by scaling the time variable, the set of accessible points from $p \in M$ is thus precisely the image of the end-point map for a fixed $T > 0$. Thus, in order to show that the sub-Riemannian distance between any two points is finite, it suffices to show that the accessibility component A_p is the whole manifold M for every point $p \in M$. This is the consequence of the celebrated Chow-Rashevskii theorem (see for instance [Bel96, Theorem 2.4]). We give a short proof relying on the Stefan-Sussmann theorem, see [Bel96, Theorem 2.3], whose statement will be given here without proof:

Theorem 2.1 (Stefan-Sussmann Theorem [Bel96])**.** *Let M be a smooth manifold, $\mathcal{F} \subset TM$ a smooth distribution in the tangent bundle. Then for any $p \in M$, we have $A_p \subset M$ is an embedded submanifold.*

An extra ingredient is needed for proving the Chow-Rashevskii theorem, which is the fact that $q \in A_p$ induces an equivalence relation. We will demonstrate this fact below:

Lemma 2.1 (Equivalence relation)**.** *Let $p, q \in M$. Then the relation*

$$p \sim q \textit{ if and only if } q \in A_p,$$

in fact defines an equivalence relation on M.

Proof. We verify the axioms of equivalence relation in the sequel. Clearly, $p \in A_p$ via the constant path $c(t) = p$ for all $t \in [0,T]$, which satisfies $\dot{c}(t) = 0 \in \mathcal{F}_{c(t)}$ for all $t \in [0,T]$. Thus $\sim$ is reflexive. The symmetry property is also rather clear. Indeed, if $p \sim q$, then there exists a horizontal path $c \colon [0,T] \to M$ with $c(0) = p$ and $c(T) = q$. Its reversal $\gamma \colon [0,T] \to M$ given by $\gamma \colon t \mapsto \gamma(t) := c(T-t)$ is a horizontal path satifying $\gamma(0) = c(T) = q$ and $\gamma(T) = c(0) = p$. Thus it yields $q \sim p$. Now transitivity of $\sim$ follows from the fact that the concatenation of horizontal path remains a horizontal path. Hence $\sim$ defines an equivalence relation. □

Theorem 2.2 (Chow-Rashevskii Theorem, [Bel96])**.** *Suppose $(M, \mathcal{F}, g)$ a connected sub-Riemannian manifold such that for any $x \in M$, there exists a neighbourhood $\mathcal{U} \subset M$ of x and a local orthonormal frame $(X_1, \dots, X_k)$ such that:*

$$T_x M = \mathit{span}_{\mathbb{R}}(\{[X_i, X_j], [[X_i, X_j], X_k], \dots\}) = \mathit{span}_{\mathbb{R}}(\bigcup_l \{[X_{i_1}, [X_{i_2}, [X_{i_3}, \dots]]]\}_{1 \leqslant i_1, \dots, i_l \leqslant k}). \tag{2.7}$$

i.e. The tangent space is spanned by iterated brackets of the orthonormal basis $(X_1|_x, \dots, X_k|_x)$.
Then any two points $p, q \in M$ can be connected by a horizontal path with respect to the distribution $\mathcal{F}$. Thus, the notion of sub-Riemannian distance is well-defined.

Proof. By Theorem 2.1, A_p defines an embedded submanifold of M. Let $(X_1, \dots, X_k)$ be a local orthonormal frame around p. We observe that A_p is stable under the local flow $(t,p) \mapsto \Phi_i(t,p) := \exp(tX_i)$ for each $i = 1, \dots, k$, where the exponential map here takes each X_i to the integral curves of X_i starting from p. Then, we must have X_i is tangent to A_p. From this, it follows that the iterated Lie brackets are also tangent to the accessible set A_p from $p \in M$. Now, (2.7) implies for each $q \in A_p$, we have $T_q A_p = T_q M$ by comparing dimensions and pointwise linear independence. In particular, by the invariance of domain, each A_p is an n-dimensional immersed submanifold.
Recall that the inclusion map $\iota_p \colon A_p \hookrightarrow M$ is an embedding, in particular an immersion. This implies that its differential $\mathrm{d}\iota_p|_q \colon T_q A_p \to T_q M$ is injective at each point $q \in A_p$. Since $\dim T_q A_p = \dim T_q M = n$, the differential is thus necessarily

bijective. By the inverse function theorem, it now follows that A_p is an open submanifold. Now for fixed $p \in M$, we have $p \in A_p$ by reflexivity of the equivalence relation from Lemma 2.1 and thus $A_p \neq \emptyset$. Then we can write:

$$M = A_p \cup (\bigcup_{q \notin A_p} A_q).$$

A consequence of Lemma 2.1 is that for $q \notin A_p$, we have $A_p \cap A_q = \emptyset$. By above, we have A_q is open for any $q \in M$, thus $\bigcup_{q \notin A_p} A_q$ is open. Thus the above gives a disjoint union of open sets in M. By connectedness of M, we must have $\bigcup_{q \notin A_p} A_q = \emptyset$ and thus $A_p = M$, which is the desired result. □

Remark 2.1. We call $r(x) := \dim \mathcal{F}_x$ the rank of the sub-Riemannian structure of M at $x \in M$. By the discussion at the beginning of this chapter, we can relax the constraint on the rank of sub-bundle to allow rank-varying distributions, i.e. $r_{\mathcal{F}}(x) \neq$ const. is permitted. If the sub-Riemannian structure $(\mathcal{F}, g)$ on M satisfies $r_{\mathcal{F}} \equiv$ const., then we have a global choice if the bundle rank of $\mathcal{F}$ and we say such a sub-Riemannian structure has constant rank.

Now, we have the following characterisation of the length of a horizontal path $c : [0, T] \to M$:

$$\mathcal{L}_g(c) = \sup\{\sum_{i=1}^{N} d_g(c(t_{i-1}), c(t_i)) \mid 0 = t_0 < t_1 < \cdots < t_N = T\}. \tag{2.8}$$

We can now extend $\mathcal{L}_g$ to act on all continuous paths, which gives rise to a lower-semicontinuous functional $\mathcal{L}_g : C^0([0,T]; M) \to [0, \infty]$. Recall we have defined a horizontal path $c\colon [0,T] \to M$ to be absolutely continuous and thus pointwise differentiable a.e. $t \in [0,T]$. Since c is everywhere tangent to the horizontal distribution, we have $|\dot{c}(t)|_g < \infty$ a.e. $t \in [0,T]$. Hence, horizontal paths are rectifiable with respect to the sub-Riemannian distance function. It is thus natural to assume Chow's condition in the definition of a sub-Riemannian manifold. Hence, we can now give a formal definition to a sub-Riemannian manifold:

Definition 2.2 (Sub-Riemannian Manifold)**.** *Let M^n be a smooth manifold, $\mathcal{F}$ a smooth distribution in the tangent bundle TM of M. We call the triple $(M, \mathcal{F}, g)$ a sub-Riemannian manifold if $\mathcal{F}$ satisfies the Chow's condition (2.7) and $g \in \Gamma(\mathcal{F}^* \otimes \mathcal{F}^*)$ is the smoothly varying fibre inner product obtained from polarising (1.1) with respect to $\mathcal{F}$.*

3 Length space structure and metric geometry

The Chow-Rashevkii theorem guarantees the finiteness of sub-Riemannian distance between any two points. Similar to the construction in Riemannian geometry, we can show that (M, d_g) defines a metric space whose metric topology is equivalent to the original manifold topology of M.

Theorem 3.1 (Metric space structure, [Bel96])**.** *Let $(M, \mathcal{F}, g)$ be a sub-Riemannian manifold where $\mathcal{F}$ satisfies Chow's condition (2.7). Then (M, d_g) defines a metric space whose metric induced topology coincides with the manifold topology on M.*

Proof. Indeed, a useful consequence of Theorem 2.2 is that the end-point map E is open, for details see [Bel96, Theorem 2.5]. Thus, any ε-ball with respect to metric d_g is the open image of $\mathbb{B}(0, \varepsilon) \subset L^1([0,T]; \mathbb{R}^k)$, which defines an open set in M. By continuity of E at $0 \in L^1([0,T]; \mathbb{R}^k)$, any neighbourhood of p contains a d_g-ball of sufficiently small radius around p, which is the image of an open ball $B \subset L^1([0,T]; \mathbb{R}^k)$. Thus, the topology induced by the distance function d_g and the manifold topology on M are equivalent.

We will now show that d_g defines a metric. By Theorem 2.2, we have $d_g(x, y) \in [0, \infty)$ for any two points $x, y \in M$. The symmetry property of d_g follows from the invariance of length under reparametrisation. Indeed, for any given horizontal path $c\colon [0,T] \to M$, we have

$$\gamma\colon [0,T] \to M, t \longmapsto c(T - t),$$

defines a horizontal path with $\mathcal{L}_g(\gamma) = \mathcal{L}_g(c)$ and switched endpoints, i.e. $\gamma(0) = c(T), \gamma(T) = c(0)$. If $c_1\colon [0, T_1] \to M$ and $c_2\colon [0, T_2] \to M$ are two horizontal paths with $c_1(T_1) = c_2(0) = p$, their concatenation given by

$$c\colon [0, T_1 + T_2] \to M, t \longmapsto \begin{cases} c_1(t), t \in [0, T_1] \\ c_2(t - T_1), t \in [T_1, T_1 + T_2] \end{cases} ,$$

remains an absolutely continuous path with $\dot{c}(t) \in \mathcal{F}_{c(t)}$ a.e. $t \in [0, T_1 + T_2]$. In particular, c defines a horizontal path with $d_g(c_1(0), c_2(T)) \leqslant \mathcal{L}_g(c) = \mathcal{L}_g(c_1) + \mathcal{L}_g(c_2)$. Taking infimum over the set of horizontal paths connecting $c_1(0), p$ and $p, c_2(T)$ respectively gives: $d_g(c_1(0), c_2(T)) \leqslant d_g(c_1(0), p) + d_g(p, c_2(T))$, thus establishing the triangle inequality.

Clearly, for any point $x \in M$, $d_g(x, x) = 0$ since the constant path $\gamma(t) \equiv x$ has length zero. Thus, it suffices to show $d_g(x, y) > 0$ whenever $x \neq y \in M$. We choose a compact neighbourhood $K \subset M$ with $x \in \text{int}(K)$, $y \in M \setminus K$. Now without loss of generality, we can pick K sufficiently small such that it is contained in a coordinate chart of M. Thus, we have the identification $K \simeq K_{Euc} \subset \mathbb{R}^n$ via our choice of coordinate maps and K can be endowed with the Euclidean metric d_E on $\mathbb{R}^n$.

Denote the local coordinates on K by $(x^1, \dots, x^n)$. By reordering if necessary, we can suppose $(\frac{\partial}{\partial x^i})_{i=1}^k$ forms a local frame for the distribution $\mathcal{F}$. Thus we can identify each horizontal vector $V = V^i \frac{\partial}{\partial x^i}$ with $\tilde{V} = (V^1 g_{11}, \dots, V^k g_{kk}) \in \mathbb{R}^k$, where $g_{ii} = \langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^i} \rangle_g$. The Einstein summation sums over $i \in \{1, \dots, k\}$, which will also be used in the remainder of the proof without further comment. We will now compare the Euclidean and sub-Riemannian distances in K.

Note for each $i = 1, \dots, k$, $x \mapsto g_{ii}(x)$ is a smooth function. By compactness, there exists $C \in (0, \infty)$ with

$$\sup_{x_0 \in K} \sqrt{\sum_{i=1}^k g_{ii}(x_0)^2} = \max_{x_0 \in K} \sqrt{\sum_{i=1}^k g_{ii}(x_0)^2} \leqslant C$$

Since $x \in \text{int}(K)$, we must have $\text{dist}_E(x, \partial K) := \inf_{x_0 \in \partial K} d_E(x, x_0) > 0$, thus we can pick some $\varepsilon > 0$ sufficiently small such that:

$$\text{dist}_E(x, \partial K) > C\varepsilon \geqslant \varepsilon \sqrt{\sum_{i=1}^k g_{ii}(x_0)^2}, \text{ for all } x_0 \in K$$

We claim that, for any horizontal path $c : [0, T] \to M$ with $c(0) = x$, if $\mathcal{L}_g(c) \leqslant \varepsilon$, then the image $c([0, T])$ is contained in the compact neighbourhood K. Assuming the claim, observe that, since $y \notin K$, any continuous path $\gamma\colon I \to M$ connecting the two points x, y must pass leave K at least once. Thus, it necessarily holds $\mathcal{L}_g(c) \geqslant \varepsilon > 0$, which then leads to:

$$d_g(x, y) = \inf\{\mathcal{L}_g(c) \mid c\colon [0, T] \to M \text{ horizontal path, } c(0) = x, c(T) = y\} \geqslant \varepsilon > 0,$$

thus finishing the proof that d_g defines a metric. Now it remains to prove our claim to conclude the proof. We recall the construction of a local sub-Riemannian metric (cf. Definition 1.1) from polarisation. Since K is contained in some coordinate chart, we may as well assume that $(\frac{\partial}{\partial x^i})_{i=1}^k$ forms a local frame for the distribution $\mathcal{F}$ around K. In particular, horizontal paths starting from $x \in K$ within the compact set K can be given as solutions to the differential equations:

$$\begin{cases} \tilde{\dot{c}}^i(t) = \xi^i(t) g_{ii}(c(t)) \\ c(0) = x \end{cases} \quad \text{a.e. } t \in [0, T] \text{ for all } i = 1, \dots, k, \text{ with } (\xi^i)_{i=1}^k \in L^1([0, T]; \mathbb{R}).$$

Suppose for contradiction that the claim is not true. Then there exists a horizontal path $c\colon [0, T] \to M$ starting from $x \in K$ with $\mathcal{L}_g(c) \leqslant \varepsilon$ and $c(t)$ leaves K eventually. We can choose minimal coefficients $(\xi^i)_{i=1}^k \in L^1([0, T]; \mathbb{R})$ such that $g_{c(t)}(\dot{c}(t), \dot{c}(t)) = \sum_{i=1}^k \xi^i(t)^2$ by discussion in the previous section. By assumption, it yields $t_0 := \sup\{t \in [0, T] \mid c([0, t]) \subset$

$K\} < T$. However:

$$\begin{aligned}
d_E(c(t_0), c(0)) &\leqslant \int_0^{t_0} |\dot{\tilde{c}}(t)|_E \mathrm{d}t = \int_0^{t_0} \sqrt{\xi^i(t)^2 g_{ii}(c(t))^2} \mathrm{d}t \\
&\leqslant \int_0^{t_0} \|(\xi^1(t) g_{11}(c(t)), \ldots, \xi^k(t) g_{kk}(c(t)))\|_{l^1} \mathrm{d}t = \int_0^{t_0} |\xi^i(t) g_{ii}(c(t))| \mathrm{d}t \\
&\leqslant \int_0^{t_0} \sqrt{\sum_{i=1}^k \xi^i(t)^2} \sqrt{\sum_{i=1}^k g_{ii}(c(t))^2 \mathrm{d}t} \\
&\leqslant C \int_0^{t_0} |\dot{c}(t)|_g \mathrm{d}t = C\mathcal{L}_g(c) \leqslant C\varepsilon < \mathrm{dist}_E(x, \partial K),
\end{aligned}$$

where the second inequality uses the fact $\|\|_{l^p}$ decreases in p and the third inequality is the consequence of Cauchy-Schwarz inequality. This is clearly a contradiction since the path c leaves K at time t_0. $\square$

Remark 3.1. Furthermore, we can deduce from the above proof that (M, d_g) is locally compact. Since the metric space (M, d_g) and the sub-Riemannian manifold $(M, \mathcal{F}, g)$ are homeomorphic, $\overline{B}_r(p) := \{x \in M \mid d_g(p, x) \leqslant r\}$ is a closed neighbourhood of $p \in M$ for any $r > 0$. We refer to this as a closed metric ball in M. We can choose a close metric ball with sufficiently small radius $r > 0$ such that the it is bounded and contained in some coordinate chart around p. Then $\overline{B}_r(p)$ is diffeomorphic to some closed and bounded set in a Euclidean space, thus compact.

In fact, this is simply a special case of smooth length structures, which are, loosely speaking, a metric space arising as adjoining an admissible class of paths.

Definition 3.1 (Length structure)**.** *A length structure on a topological space X is a class $\mathcal{A}$ of admissible continuous paths in X, together with a length functional $\mathcal{L}\colon \mathcal{A} \to [0, \infty]$ such that:*

(i) $\mathcal{A}$ is closed under restriction of paths: if $\gamma\colon [0,1] \to X$ is an admissible path and $[c, d] \subset [0, 1]$, then $\gamma|_{[c,d]} \in \mathcal{A}$;

(ii) $\mathcal{A}$ is closed under concatenation: if $\gamma\colon [0,1] \to X$ and $\sigma\colon [1,2] \to X$ are both admissible paths with $\gamma(1) = \sigma(1)$, then the product $\gamma \cdot \sigma$ is admissible;

(iii) $\mathcal{A}$ is closed under (linear) reparametrisation.

For a metric space (X, d) where X admits a length structure $(\mathcal{A}, \mathcal{L})$, we say that d is intrinsic if it can be obtained as the distance function for $\mathcal{A}$: i.e. for all $x, y \in X$, we have:

$$d_{\mathcal{L}}(x, y) := \inf \{L(\gamma) \mid \gamma \in \mathcal{A},\ \gamma(0) = x,\ \gamma(1) = y\} = d(x, y). \tag{3.1}$$

We say that a metric space (X, d) defines a length space if d is intrinsic to some length structure on X. A length structure $(\mathcal{A}, \mathcal{L})$ is said to be complete if for any pair of points $x, y \in X$, we have $d_{\mathcal{L}}(x, y) < \infty$ and there exists some $\gamma \in \mathcal{A}$ connecting x and y such that $\mathcal{L}(\gamma) = d_{\mathcal{L}}(x, y)$.

Remark 3.2. Any length space is locally connected and locally compact (as a metric space). The theorems in this section can be modified to hold for all length spaces. The reason for working with a length structure is to allow synthetic methods from metric geometry, where we can define the notion of, for instance curvatures, without an a priori smooth structure. Some contemporary research in non-smooth geometric analysis deals with synthetic methods in geometry, which connects different fields such as geometric measure theory, metric geometry, and optimal transport.

We will finish with a Hopf-Rinow theorem for sub-Riemannian manifolds, which links the completeness of M as a metric space and the geodesic completeness of M. The local existence of sub-Riemannian length minimisers would then be guaranteed. We call a horizontal path $c\colon [0, T] \to M$ *a minimising sub-Riemannian geodesic* if $\mathcal{L}_g(c) = d_g(c(0), c(T))$. We will first state a general fact from metric geometry which we will use in proving the existence of horizontal length minimiser.

Remark 3.3. Careful readers, with some experience in differential geometry, would have realised that I never actually defined the notion of a sub-Riemannian geodesic. It is rather cumbersome to define the approriate notion of geodesics in the world of sub-Riemannian geometry. Recall that in Riemannian geometry, we can characterise geodesics locally as the solution to the geodesic equations, which allows us to parametrise geodesics based on their initial position and velocity. On the contrary, a dimension argument tells us that this is impossible in sub-Riemannian geometry.
Indeed, let $(M^n, \mathcal{F}, g)$ be a sub-Riemannian manifold and $x \in M$ with $\dim \mathcal{F}_x < n$. Then the set of admissible directions for a horizontal curve $c\colon I \to M$ with $c(0) = x$ has dimension strictly smaller than the dimension of the ambient manifold M. This contradicts the Chow-Rashevskii theorem since we should have been able to cover a local neighbourhood of x by horizontal curves emanating from x. The canonical approach is to parametrise them on the cotangent bundle, which uses the so-called sub-Riemannian Hamiltonian and borrows analogy from control theory. We will not delve further into these, aside from remarking that as usual, any geodesic is locally length-minimising and any sub-Riemannian length-minimiser is necessarily a geodesic.

Lemma 3.1 (Lower semi-continuity of length functional)**.** *Let (X, d) be a metric space, $\gamma_j : I \to X$ a sequence of absolutely continuous paths. Suppose $\gamma_j \to \gamma$ pointwise in X and $\liminf_{j\to\infty} l(\gamma_j) < \infty$ where*

$$l\colon \mathcal{C}_X := \{\gamma\colon I \to X \mid \gamma \textit{ absolutely continuous path in } X\} \to [0, \infty], \tag{3.2}$$

denotes the induced length functional from the metric d. Then $l(\gamma) \leqslant \liminf_{j\to\infty} l(\gamma_j)$.

Proof. Note it suffices to show that for any $\varepsilon > 0$, we have $l(\gamma) \leqslant l(\gamma_j) + \varepsilon$ for j sufficiently large, which then implies $l(\gamma) \leqslant \liminf_{j\to\infty} l(\gamma_j) + \varepsilon$ for every $\varepsilon > 0$.
Fix $\varepsilon > 0$. We denote $L(\gamma, (t_i)_{i=1}^N) = \sum_{i=1}^N d(\gamma(t_{i-1}), \gamma(t_i))$, then $l(\gamma) = \sup\{L(\gamma, (t_i)_{i=1}^N) \mid (t_i)_{i=1}^N \in \mathcal{P}_{[0,T]}\}$, where $\mathcal{P}_{[0,T]}$ denotes the set of partitions for the interval $[0, T]$. By definition, we can choose partition $(t_i)_{i=1}^N$ such that $l(\gamma) \leqslant L(\gamma, (t_i)_{i=1}^N) + \frac{\varepsilon}{2}$. Now pointwise convergence implies that for j sufficiently large, we have $d(\gamma_j(t_i), \gamma(t_i)) \leqslant \frac{\varepsilon}{4N}$ for every $i = 1, :, N$. Thus we obtain:

$$d(\gamma(t_{i-1}), \gamma(t_i)) \leqslant d(\gamma(t_{i-1}), \gamma_j(t_i)) + d(\gamma_j(t_{i-1}), \gamma_j(t_i)) + d(\gamma(t_i), \gamma_j(t_i)) \leqslant d(\gamma_j(t_{i-1}), \gamma_j(t_i)) + \frac{\varepsilon}{2N}$$

Summing over i and taking the supremum over the set of partitions gives:

$$l(\gamma) = \sup_{(t_i)_{i=1}^N \in \mathcal{P}_{[0,T]}} \sum_{i=1}^N d(\gamma(t_{i-1}), \gamma(t_i)) \leqslant \sum_{i=1}^N d(\gamma_j(t_{i-1}), \gamma_j(t_i)) + \frac{\varepsilon}{2N} = L(\gamma_j, (t_i)_{i=1}^N) + \frac{\varepsilon}{2} \leqslant l(\gamma_j) + \varepsilon,$$

which is the desired result. □

Theorem 3.2 (Hopf-Rinow theorem for sub-Riemannian manifolds, [Bel96])**.** *Suppose (M, d_g) is a complete metric space. Then any two suffienctly close points p, $q \in M$ can be joined by a minimising sub-Riemannian geodesic. In particular, M is geodesically complete.*
Furthermore, any point $p \in M$ admits a geodesically complete neighbourhood.

Proof. As a consequence of Theorem 2.2, we have (M, d_g) a complete metric space with $d_g\colon M \times M \to [0, \infty)$ finite and continuous with respect to the natural product topology on $M \times M$. We note since M is a manifold, in particular locally diffeomorphic to a subset in an Euclidean space which is locally compact. Thus we have M is locally compact space.
Let $p, q \in M$ be arbitrary and denote for simplicity $L := d_g(p, q) < \infty$, where finiteness is a consequence of Theorem 2.2. By 2.2, we can choose a sequence of horizontal paths (γ_j) connecting p and q with $\mathcal{L}_g(\gamma_j) \to L$ as $j \to \infty$. Assume without loss of generality the image of each γ_j is contained inside some compact neighbourhood of p and q. By the Arzelá-Ascoli theorem, we can pass to a subsequence so that $\gamma_j \to \gamma$ uniformly on $[0, T]$, where $\gamma\colon [0, T] \to M$ is a continuous path with $\gamma(0) = p$ and $\gamma(T) = q$. By the lower semi-continuity of the length functional, we deduce that $\mathcal{L}_g(\gamma) = L = d_g(p, q)$. □

References

[ABBZ19] A. Agrachev, D. Barilari, U. Boscain, and I. Zelenko. *A Comprehensive Introduction to Sub-Riemannian Geometry*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2019.

[Bel96] André Bellaïche. *The tangent space in sub-Riemannian geometry*, pages 1–78. Birkhäuser Basel, Basel, 1996.

[Mil78] John Milnor. Analytic proofs of the ”hairy ball theorem” and the brouwer fixed point theorem. *The American Mathematical Monthly*, 85(7):521–524, 1978.