Università degli Studi di Padova

DIPARTIMENTO DI MATEMATICA "TULLIO LEVI CIVITA"
Corso di Laurea Magistrale in Matematica

Tesi di Laurea Magistrale

Sard property in Carnot groups

Laureando:
Luca Nalon
Matricola 1236589

Relatore:
Prof. Davide Vittone

Anno Accademico 2020/2021
23 luglio 2021

Contents

Introduction

In this work we introduce the topic of sub-Riemannian geometry from an elementary viewpoint. Sub-Riemannian geometry is a quite modern field of differential geometry. The subject has been studied by that name from the 90s, see for instance [And96], however, several key ideas of Sub-Riemannian geometry are antecedent, e.g. the concept of sub-Riemannian distance, firstly denoted as Carnot-Carathéodory distance, which makes its appearance in [Mit85]. In [Mon06], Richard Montgomery depicts a wide representation of all scientific motivations behind the development of the theory in those years: from thermodynamics to robotics.

The main objective of this thesis is to provide a first description of abnormal curves, which are particular curves on a sub-Riemannian manifold which exhibit an anomalous (and hopefully rare) behaviour. Abnormal curves are related to many open problems in sub-Riemannian geometry such as the regularity of the sub-Riemannian distance, the homotopy of small sub-Riemannian balls and the study of the sub-Laplacian which is related to the heat diffusion on sub-Riemannian manifolds. Our main concern will be the study of the Sard problem for the end-point map: we will see how abnormal curves are related to critical values of a specific map from a functional space to a sub-Riemannian manifold, then we will be interested to determine the structure of this set of critical values and, in particular, whether or not its measure is always zero.

In chapter 1 we recall basic notions of differential geometry and we establish the relative notations used later on, moreover we state and prove Frobenius theorem which is a founding result for sub-Riemannian geometry.

In chapter 2 we start to familiarize ourselves with sub-Riemannian structures and their properties. Rashevskii-Chow theorem will motivate our interest for length-minimizers and their description. We introduce this topic in chapter 3 where we translate length-minimality in an optimal control problem and we finally define normal and abnormal extremals in terms of Pontryagin's maximum principle. We conclude chapter 3 introducing the end-point map and its relation with abnormal extremals.

In chapter 4 we momentarily abandon sub-Riemannian geometry for a quick algebraic introduction to nilpotent Lie groups. Lie groups are fundamental examples of differentiable manifolds and their algebraic structure proves to be handy for explicit computations. In particular, we will use Baker-Campbell-Hausdorff formula to provide a complete description of connected, simply-connected and nilpotent Lie groups which will be the protagonists of the last chapter.

In chapter 5 we introduces Carnot groups. Carnot groups are particular connected, simply connected, nilpotent Lie groups, they appear in many fields of mathematics. We will be interested in their sub-Riemannian structure. In a sense Carnot groups are the most elementary examples of sub-Riemannian manifolds which are not trivial (that is Riemannian), on the other hand, as explained in [Don16], every sub-Riemannian structure is locally very similar to a Carnot groups and this fact suggest their relevance in sub-Riemannian geometry. We will follow the work done in [Le +16] and [OV19] regarding the Sard property of the abnormal set in Carnot groups of step 2. Finally we provide an algebraic description for the abnormal set for free Carnot groups of step 2 through an explicit system of polynomials, which is not present in the literature. We will use use this result to fully answer the Sard problem for free Carnot groups of step 2 with rank 2,3,4 and 5.

Chapter 1

Elements of differential geometry

In this chapter we introduce classical notions and results of differential geometry. We will mainly focus on the tools that will be needed in the successive chapters. This brief summary will also be the occasion to settle the notation used later in the discussion. A detailed reference for this topic is [Lee13], the presentation of the subject partially follows [ABB20].

1.1 Differentiable manifolds and smooth maps

Definition 1.1.1. A *topological manifold* is a topological space M together with a collection $\{ (U_i, \varphi_i) \}_{i\in I}$ such that $\{ U_i \}_{i\in I}$ is an open covering of M and for every $i \in I$

$$\varphi_i \colon U_i \to \varphi_i(U_i) \quad \text{open subset of } \mathbb{R}^n \tag{1.1}$$

is a homeomorphism for some $n \in \mathbb{N}$. In this case (U_i, φ_i) are called *local charts* and the collection $\{ (U_i, \varphi_i) \}_{i\in I}$ is said to be an *atlas* for the manifold.

Remark 1.1.2. In a topological manifold M, whenever two local charts

$$\varphi_i \colon U_i \to \varphi_i(U_i) \subseteq \mathbb{R}^n \qquad \varphi_j \colon U_j \to \varphi_j(U_j) \subseteq \mathbb{R}^m$$

verify $U_i \cap U_j \neq \varnothing$, then $\eta_{ij} \colon \varphi_i \circ \varphi_j^{-1}$ defines a homeomorphism between $\varphi_j(U_i \cap U_j)$ and $\varphi_i(U_i \cap U_j)$ which implies $m = n$. The functions η_{ij} are called *transition maps* and they satisfy the following properties:

$$\eta_{ii} = \mathrm{id}_{\varphi(U_i)} \qquad \text{for any } i \in I \tag{T1}$$

$$\eta_{ji} = \eta_{ij}^{-1} \qquad \text{for any } i, j \in I \text{ such that } U_i \cap U_j \neq \varnothing \tag{T2}$$

$$\eta_{ik} = \eta_{ij} \circ \eta_{jk} \qquad \text{for any } i, j, k \in I \text{ such that } U_i \cap U_j \cap U_k \neq \varnothing \tag{T3}$$

The topology of a topological manifold M inherits the local properties of the real space $\mathbb{R}^n$. For instance, a topological manifold M is locally compact, locally connected and locally arc-wise connected. However, M may not be Hausdorff nor connected, in contrast to $\mathbb{R}^n$ which has both properties. We will turn our attention to Hausdorff and connected spaces.

Remark 1.1.3. Given a topological manifold M and a local chart (U_i, φ_i), we may define its dimension as the dimension of $\varphi_i(U_i)$. Then we can consider

$$M^{(n)} = \bigcup_{i \in I} \{ U_i \mid (U_i, \varphi_i) \text{ has dimension } n \}$$

which defines, by remark 1.1.2, a family of disjoint open sets which cover M. If M is connected, then $M = M^{(d)}$ for some $d \in \mathbb{N}$. We will refer to such a d as the dimension of the connected topological manifold M.

Definition 1.1.4. A *differentiable manifold* is a topological manifold M such that its transition maps are smooth $\mathcal{C}^\infty$.

Remark 1.1.5. Given a differentiable manifold M and an atlas $\{ (U_i, \varphi_i) \}_{i \in I}$ for M, a continuous function

$$f \colon M \to \mathbb{R}$$

defines a family $\left\{ \widetilde{f}_i = f \circ \varphi_i^{-1} \;\middle|\; i \in I \right\}$ of continuous real maps from open sets in $\mathbb{R}^n$. If $U_i \cap U_j \neq \varnothing$ then

$$\widetilde{f}_j = \widetilde{f}_i \circ \eta_{ij} \qquad \text{on } \varphi_j(U_i \cap U_j) \tag{1.2}$$

Definition 1.1.6. Let us consider a differentiable manifold M, $\{ (U_i, \varphi_i) \}_{i \in I}$ an atlas for M and a continuous function $f \colon M \to \mathbb{R}$.

We say f is continuous of class $\mathcal{C}^r$ at $x \in M$ if $\widetilde{f}_i$ are continuous of class $\mathcal{C}^r$ at $\varphi_i^{-1}(x)$ for any $i \in I$ such that $x \in U_i$, f is smooth at $x \in M$ (that is continuous of class $\mathcal{C}^\infty$) if $\widetilde{f}_i$ are smooth at $\varphi_i^{-1}(x)$ for any $i \in I$ such that $x \in U_i$. If f is continuous of class $\mathcal{C}^r$ (respectively smooth) at all points $x \in M$ we say that f is continuous of class $\mathcal{C}^r$ (respectively smooth).

Let us notice that, considering remark 1.1.5, the previous definitions are independent of the choice of the local chart since the transition maps are smooth.

We will apply the same approach into the study of continuous maps between differentiable manifolds.

Remark 1.1.7. Let us consider two differentiable manifolds M and N and a continuous function

$$F \colon M \to N\,.$$

Given a point $x \in M$, we may consider a local chart (U, φ) of M with dimension m such that $x \in U$ and a local chart (V, ψ) of N with dimension n such that $F(x) \in V$. Then we can consider

$$\widetilde{F} \colon \psi \circ F \circ \varphi^{-1} \qquad \text{on } \varphi(U \cap F^{-1}(V)) \tag{1.3}$$

which defines a function from an open set in $\mathbb{R}^m$ to an open set of $\mathbb{R}^n$.

Now we will investigate how those newly defined functions behave under change of local charts. This time we consider two local charts $(U_i, \varphi_i), (U_j, \varphi_j)$ of M such that

$x \in U_i \cap U_j$, and on the other hand $(V_h, \varphi_h), (U_k, \varphi_k)$ two local charts of N such that $F(x) \in V_h \cap V_k$. As above we consider

$$\widetilde{F}_{hi} = \psi_h \circ F \circ \varphi_i^{-1} \qquad \text{on } \varphi_i(U_i \cap F^{-1}(V_h)) \tag{1.4}$$

$$\widetilde{F}_{kj} = \psi_k \circ F \circ \varphi_j^{-1} \qquad \text{on } \varphi_j(U_j \cap F^{-1}(V_k)) \tag{1.5}$$

which lead to the following diagram:

$$\begin{array}{ccc} \varphi_i\big(U_i \cap U_j \cap F^{-1}(V_h \cap V_k)\big) & \xrightarrow{\widetilde{F}_{hi}} & \psi_h(V_h \cap V_k) \\ \varphi_i \uparrow & & \uparrow \psi_h \\ U_i \cap U_j \cap F^{-1}(V_h \cap V_k) & \xrightarrow{F} & V_h \cap V_k \\ \varphi_j \downarrow & & \downarrow \psi_k \\ \varphi_j\big(U_i \cap U_j \cap F^{-1}(V_h \cap V_k)\big) & \xrightarrow{\widetilde{F}_{kj}} & \psi_k(V_h \cap V_k) \end{array} \qquad \varphi_i \circ \varphi_j^{-1} = \eta_{ij}, \quad \theta_{kh} = \psi_k \circ \psi_h^{-1} \tag{1.6}$$

Hence

$$\widetilde{F}_{kj} = \theta_{kh} \circ \widetilde{F}_{hi} \circ \eta_{ij} \qquad \text{on } \varphi_j\big(U_i \cap U_j \cap F^{-1}(V_h \cap V_k)\big)\,.$$

Definition 1.1.8. Let us consider two differentiable manifold M, N and a continuous function $F\colon M \to N$. Using the notation in remark 1.1.7, we say that F is continuous of class $\mathcal{C}^r$ at $x \in M$ if $\widetilde{F}$ is continuous of class $\mathcal{C}^r$ at $\varphi(x)$, we say that F is smooth at $x \in M$ if $\widetilde{F}$ is smooth at $\varphi(x)$. If F is continuous of class $\mathcal{C}^r$ (respectively smooth) at all points $x \in M$ we say that F is continuous of class $\mathcal{C}^r$ (respectively smooth).

Definition 1.1.9. A continuous function $F\colon M \to N$ between two differentiable manifolds is a *diffeomorphism* (of class $\mathcal{C}^r$) if F is bijective, continuous of class $\mathcal{C}^r$ and its inverse $F^{-1}\colon N \to M$ is again continuous of class $\mathcal{C}^r$. If not specified, we will always assume continuous of class $\mathcal{C}^\infty$.

The classification of differential and topological manifold up to diffeomorphism is a widely studied problem in differential geometry. Later on it will be useful to refer to the set of inner diffeomorphism of a differentiable manifold M as

$$\mathrm{Diff}(M) = \{\, F\colon M \to M \mid F \text{ is a diffeomorphism} \,\}\,, \tag{1.7}$$

which has a natural group structure with respect to the composition of maps.

Definition 1.1.10. A smooth function $F\colon M \to N$ between two differentiable manifolds is a *local diffeomorphism* if, for any $x \in M$ there exists an open neighbourhood U of x such that $F\colon U \to F(U)$ is a diffeomorphism.

Definition 1.1.11. A smooth function $\pi\colon \widetilde{M} \to M$ between two differentiable manifolds is a *smooth covering map* if:

(i) π is surjective

(ii) For every $x \in M$ there exists an open neighbourhood U of x such that the restriction of π to any connected component $\widetilde{U}$ of $\pi^{-1}(U)$ is a diffeomorphism between $\widetilde{U}$ and U.

If $\widetilde{M}$ is simply connected, then π is said to be an *universal covering map*.

Definition 1.1.12. A *partition of the unity* over a differentiable manifold M is a family of $\mathcal{C}^\infty$ functions on M, $\{ \rho_\alpha \}_\alpha$, such that

1. $0 \leq \rho_\alpha(x) \leq 1$ for all $x \in X$ and α,
2. $\{ \operatorname{supp} \rho_\alpha \}$ is a locally finite open covering of M, that is, for each $x \in M$, $x \in \operatorname{supp} \rho_\alpha$ only for a finite number of α,
3. $\sum_\alpha \rho_\alpha(x) = 1$ for all $x \in M$.

We say that a partition of the unity $\{ \rho_\alpha \}_\alpha$ is *subordinate* to an open covering $\mathcal{U} = \{ U_\alpha \}_\alpha$ if $\operatorname{supp} \rho_\alpha \subseteq U_\alpha$ for each α.

Theorem 1.1.13. *Let M be a differentiable manifold which is Hausdorff and with countable basis. Then for each open covering $\{ U_\alpha \}_\alpha$ of M there exists a partition of the unity subordinate to the covering $\{ U_\alpha \}_\alpha$.*

1.2 Tangent vectors and vector bundles

Tangent vectors to a differentiable manifolds can be defined using different approaches. For instance one may identify tangent vectors at $q \in M$ as derivations in the algebra of the germs of function at q. Instead, we will mainly focus on curves defined on a manifold, which makes the definition through curves more natural in our discussion.

In what follows, I will denote an open interval in $\mathbb{R}$.

Definition 1.2.1. Let M be a differentiable manifold and let us consider the set of smooth curves based at q, that is

$$\mathbf{\Gamma}_q = \{ \gamma \colon I \to M \mid \gamma \text{ smooth}, 0 \in I \subseteq \mathbb{R}, \gamma(0) = q \} . \tag{1.8}$$

Let consider (U, φ) a local chart such that $q \in U$. Two curves $\gamma_1, \gamma_2 \in \mathbf{\Gamma}_q$ are said to be equivalent with respect to the chart (U, φ) if

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\varphi \circ \gamma_1)(t) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\varphi \circ \gamma_2)(t) \tag{1.9}$$

which defines an equivalence relation on $\mathbf{\Gamma}_q$. The *tangent vectors* at q are identified to be the equivalence classes of $\mathbf{\Gamma}_q$ with respect to this equivalence relation. If $\gamma \in \mathbf{\Gamma}_q$, the corresponding tangent vector is denoted by

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \gamma(t) \qquad \text{or} \qquad \dot{\gamma}(0) . \tag{1.10}$$

Remark 1.2.2. Let us notice that, in definition 1.2.1, the notion of equivalence on $\mathbf{\Gamma}_q$ does not depend on the choice of the local chart (U, φ). Indeed, if (U_i, φ_i) and (U_j, φ_j) are two local charts such that $q \in U_i \cap U_j$ and $\gamma \in \mathbf{\Gamma}_q$, then

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\varphi_j \circ \gamma_1)(t) = (\mathbf{J}\, \eta_{ji})_{\varphi^{-1}(q)} \left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\varphi_i \circ \gamma_1)(t) \right) \tag{1.11}$$

where $(\mathbf{J}\,\eta_{ji})_{\varphi^{-1}(q)}$ denotes the Jacobian matrix of the transition map η_{ji} at $\varphi^{-1}(q)$, which is invertible since the transition maps are diffeomorphisms. This means that the notion of tangent vector is intrinsic to the manifold M.

Definition 1.2.3. Let M be a differentiable manifold and $q \in M$, we define

$$T_qM = \left\{ \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \gamma(t) \;\middle|\; \gamma \text{ smooth, } 0 \in I \subseteq \mathbb{R},\ \gamma(0) = q \right\} \tag{1.12}$$

that is the quotient space of $\mathbf{\Gamma}_q$ under the equivalence relation defined in definition 1.2.1.

Remark 1.2.4. The tangent space to a point T_qM is naturally endowed with vector space structure. Indeed, given $\gamma_1, \gamma_2 \in \mathbf{\Gamma}_q$, a real number $a \in \mathbb{R}$ and a local chart (U_q, φ_q), we may define on $\mathbf{\Gamma}_q$ the following operations:

(i) $\big(a \cdot \gamma_1\big)(t) = \gamma(at)$

(ii) $\big(\gamma_1 + \gamma_2\big)(t) = \varphi_q^{-1}\big((\varphi_q \circ \gamma_1)(t) + (\varphi_q \circ \gamma_2)(t)\big)$ on a neighbourhood of zero.

One can check that these operations are faithful with respect to the equivalence relation on $\mathbf{\Gamma}_q$ so that they induce a vector space structure on T_qM. Moreover, the vector field structure on T_qM is independent of the choice of the local chart. Finally, a transition map also induces a linear change of coordinates on T_qM via the Jacobian matrix $(\mathbf{J}\,\eta_{ji})_{\varphi^{-1}(q)}$.

Proposition 1.2.5. *Let M be a differentiable manifold of dimension d and (U, φ) a local chart, let $(e_1, \dots, e_d)$ be the canonical base on $\mathbb{R}^d \supseteq \varphi(U)$. For every $q \in U$ and $k \in \{1, \dots, d\}$ we can consider a curve based on q:*

$$\gamma_q^{(k)}(t) = \varphi^{-1}\big(\varphi(q) + te_k\big) \tag{1.13}$$

defined in some neighbourhood of zero. For every $q \in U$ the set

$$\left\{ \dot\gamma_q^{(1)}(t) = \partial_{x^1}\big|_q, \dots, \dot\gamma_q^{(d)}(t) = \partial_{x^d}\big|_q \right\} \tag{1.14}$$

is a basis for T_qM which has dimension d.

The basis in proposition 1.2.5 is called *local trivialization* of the tangent space.

Definition 1.2.6. Let $F\colon M \to N$ be a smooth map between two differentiable manifolds M and N, let $q \in M$. The *differential* of F at the point q is a map

$$\mathrm{d}F_q\colon T_qM \to T_{F(q)}N \tag{1.15}$$

defined as:

$$\mathrm{d}F_q(v) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (F \circ \gamma)(t) \qquad \text{if} \quad v = \dot\gamma(0)\,, \quad \gamma \in \mathbf{\Gamma}_q\,. \tag{1.16}$$

One can check that this definition depends only on the equivalence class of γ and that the differential map is linear.

Definition 1.2.7. Let M be a differential manifold. The *tangent bundle* of M is the set

$$TM = \bigcup_{q \in M} (q, T_q M) \,. \tag{1.17}$$

We may refer to it also with $T(M)$.

Remark 1.2.8. By proposition 1.2.5, the vector bundle TM is endowed with a structure of differentiable manifold. An atlas $\{\, U_i, \varphi_i \,\}_{i \in I}$ for M can be extended to an atlas for TM:

$$\left\{ \bigcup_{q \in U_i} (q, T_q M), (\varphi_i, \mathrm{d}x^1\big|_q, \dots, \mathrm{d}x^d\big|_q) \right\}_{i=I} \tag{1.18}$$

where $\left(\mathrm{d}x^1\big|_q, \dots, \mathrm{d}x^d\big|_q\right)$ is the dual base of $\left(\partial_{x^1}\big|_q, \dots, \partial_{x^d}\big|_q\right)$. Transition maps are defined as in proposition 1.2.5.

Le us notice that, by remark 1.2.8, each local chart φ of M defines a local chart for TM, which is a system of local coordinates. We will call them *local coordinates* on TM induced by φ. The previous differentiable structure is the standard example of a more general notion, namely that of a vector bundle. We can think of it as a smooth family of vector spaces parametrized by points in a manifold M.

Definition 1.2.9. If $F\colon M \to N$ is a smooth map, we define the *tangent map* of F as

$$TF\colon TM \to TN \qquad (q, v) \mapsto (F(q), \mathrm{d}F_q(v)) \,. \tag{1.19}$$

We may refer to it also with $T(F)$.

Definition 1.2.10. Let M be a differentiable manifold. A *(smooth) vector bundle* of rank k over M is a differentiable manifold E with a surjective differentiable map $\pi\colon E \to M$ such that

(i) the set $E_q = \pi^{-1}(q)$, named *fiber* of E at $q \in M$, is a k-dimensional vector space,

(ii) for every $q \in M$ there exist a neighbourhood U_q of q and a linear-on-fibers diffeomorphism (named *local trivialisation*) $\psi\colon \pi^{-1}(U_q) \to U_q \times \mathbb{R}^k$) such that the following diagram commutes

$$\begin{array}{ccc} \pi^{-1}(U_q) & \xrightarrow{\;\psi\;} & U_q \times \mathbb{R}^k \\ & {\searrow}^{\pi} & \downarrow{\scriptstyle \pi_1} \\ & & U_q \end{array} \tag{1.20}$$

where π_1 is the canonical projection on the first component.

The space E is called *total space* and M is the *base* of the vector bundle. We will refer to π as the *canonical projection* and $\mathrm{rank}(E)$ will denote the rank of the vector bundle.

Remark 1.2.11. As a differentiable manifold, a vector bundle E has dimension

$$\dim(E) = \dim(M) + \operatorname{rank}(E)\,. \tag{1.21}$$

If there exists a global trivialisation map (that is a local trivialisation with $U_q = M$ for every $q \in M$), then E is diffeomorphic to $M \times \mathbb{R}^k$ and we say that E is *trivializable*.

Definition 1.2.12. A *morphism* $f\colon E \to E'$ between two vector bundles E, E' with canonical projections π, π' on the same base M is a smooth map such that the following diagram is commutative

$$\begin{array}{ccc} E & \xrightarrow{\;f\;} & E' \\ & \searrow^{\pi} & \downarrow{\scriptstyle \pi'} \\ & & M \end{array} \tag{1.22}$$

and f is linear on fibers.

Example 1.2.13. For any differentiable manifold M of dimension d, the tangent bundle TM has a natural $2d$ dimensional vector bundle structure with respect to the canonical projection $\pi(q, v) = q$, as explained in remark 1.2.8. In the same way we can consider the *cotangent bundle* T^*M, defined as

$$T^*M = \bigcup_{q \in M} (q, T_q^*M = (T_qM)^*) \tag{1.23}$$

which has again a $2d$ dimensional vector bundle structure, again with respect to the canonical projection. Indeed, a local trivialisation of a vector bundle induces a local trivialisation in its dual vector bundle.

Given a local chart (U, φ) for M, we consider the coordinate function $(x^1, \ldots, x^d)$ induced by the local charts. The differentials of the maps

$$(\mathrm{d}x^1\big|_q, \ldots, \mathrm{d}x^d\big|_q) \qquad q \in U \tag{1.24}$$

form a basis of T_q^*M [1], which is precisely the dual basis of $(\partial_{x^1}, \ldots, \partial_{x^d})$, that is

$$\langle \mathrm{d}x^i\big|_q, \partial_{x^j}\big|_q \rangle = \delta_{ij}\,. \tag{1.25}$$

As the differential of a smooth map yields a linear map between tangent spaces, so the dual of the differential is a linear map between cotangent spaces.

Definition 1.2.14. A *differential 1-form* on a smooth manifold M is a section of T^*M, that is a smooth map

$$\omega\colon T^*M \to M \qquad q \mapsto \omega(q) \in T^*M\,. \tag{1.26}$$

We denote with $\Lambda^1(M)$ the set of differential 1-forms defined on M.

[1] with the identification of $T_{x^i(q)}\mathbb{R}$ with $\mathbb{R}$ such that $\mathrm{d}x^i(\partial_{x^i}) \mapsto 1$

Definition 1.2.15. Let $F\colon M \to N$ be a smooth map and let $q \in M$. The *pullback* of F at $F(q)$ is the adjoint of the differential map

$$F^*\colon T^*_{F(q)}N \to T^*_q M \qquad \lambda \mapsto F^*\lambda\,, \tag{1.27}$$

defined by duality:

$$\langle F^*\lambda, v\rangle = \langle \lambda, \mathrm{d}F(v)\rangle \qquad \forall v \in T_qM \quad \forall \lambda \in T^*_{F(q)}N\,. \tag{1.28}$$

In particular we can "pull back" differential 1-forms.

Example 1.2.16. Given a vector bundle E with base M we can consider its dual bundle $E^*$and then a tensor bundle

$$T^k_h(E) = \underbrace{E \otimes \cdots \otimes E}_{h \text{ times}} \otimes \underbrace{E^* \otimes \cdots \otimes E^*}_{k \text{ times}} \tag{1.29}$$

where the tensorial operation acts at fiber level. In $T^k(E) = T^k_0(E)$ we stress two important vector sub-bundles:

$$S^kE = \left\{ \mathbf{t} \in T^k(E) \;\middle|\; \mathbf{t}(\alpha_1, \ldots, \alpha_k) = \mathbf{t}(\alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(k)}),\, \sigma \in S_k \right\} \tag{1.30}$$

$$\Lambda^kE = \left\{ \mathbf{t} \in T^k(E) \;\middle|\; \mathbf{t}(\alpha_1, \ldots, \alpha_k) = \mathrm{sgn}(\sigma)\mathbf{t}(\alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(k)}),\, \sigma \in S_k \right\}\,, \tag{1.31}$$

respectively the symmetric and the anti-symmetric (or alternating) vector bundle.

1.3 Vector fields and their flow

Definition 1.3.1. Let $\pi\colon E \to M$ a vector bundle over M. A *local section* of E is a smooth map $\sigma\colon A \subseteq M \to E$ such that $\pi \circ \sigma = \mathrm{id}_A$, where A is an open set in M. That means that $\sigma(q) \in E_q$ and it changes smoothly with respect to q. If σ is defined on all of M it is said to be a *global section*.

Definition 1.3.2. A *smooth vector field* is a global section of TM. We denote by $\mathrm{Vec}(M)$ the set of smooth vector fields on M.

Using local coordinates, we can locally write $X = X^i\,\partial_{x^i}$. We will denote the value of a vector field at $q \in M$ either with $X(q)$ or $X\big|_q$.

Definition 1.3.3. Let M be a differentiable manifold and $X \in \mathrm{Vec}(M)$. The equation

$$\dot{q} = X(q) \tag{1.32}$$

is called an *ordinary differential equation* (or *ODE*) on M. A *solution* to eq. (1.32) is a smooth curve $\gamma\colon I \to M$, where $0 \in I \subseteq \mathbb{R}$ is an open interval, and

$$\dot{\gamma}(t) = X(\gamma(t)) \qquad \forall t \in I\,.$$

We also say that γ is an *integral curve* of the vector field X.

Theorem 1.3.4. *Let $X \in \mathrm{Vec}(M)$ and let us consider the Cauchy problem*

$$\begin{cases} \dot{q}(t) = X(q(t)) \\ q(0) = q_0 \end{cases} \tag{1.33}$$

For any point $q_0 \in M$ there exists $\delta > 0$ and a solution $\gamma\colon (-\delta, \delta) \to M$ of eq. (1.33), denoted with $\gamma(t; q_0)$.

If there exists two solutions of eq. (1.33) defined in two open intervals I_1 and I_2 containing zero, then they coincide in their intersection. This means that a solution of eq. (1.33) is essentially unique and we can introduce the notion of *maximal solution* with respect to its domain of definition.

Definition 1.3.5. A vector field $X \in \mathrm{Vec}(M)$ is called *complete* if, for evry $q_0 \in M$, the maximal solution to eq. (1.33) is defined over all $t \in \mathbb{R}$.

Theorem 1.3.6. *If a maximal solution γ of eq. (1.33) is defined on a bounded interval $I = (a, b)$, then for every compact $K \subseteq M$ there exists $t_K < b$ such that $\gamma(t; q_0) \notin K$ for every $t > t_k$. In other words: such solutions leave every compact of M after finite time.*

Remark 1.3.7. As the previous theorem suggests, there are conditions that ensure completeness of a vector field. For instance

(i) if M is compact, or more generally X has compact support in M,

(ii) if $M = \mathbb{R}^n$ and X has sub-linear growth at infinity, that is

$$|X(q)| \leq C_1|x| + C_2 \qquad \forall x \in \mathbb{R}^n$$

for some constants $C_1, C_2 > 0$.

Theorem 1.3.8. *Let M be a differentiable manifold and $X \in \mathrm{Vec}(M)$, then there exists a unique open neighbourhood $\mathcal{U}$ of $\{\,0\,\} \times M \subseteq \mathbb{R} \times M$ and a unique smooth function $\Phi^X : \mathcal{U} \to M$ satisfying the following properties:*

(i) for every $q \in M$, the set $\mathcal{U}^q = \{\, t \in \mathbb{R} \mid (t, q) \in \mathcal{U} \,\}$ is an open neighbourhood of zero,

(ii) for every $q \in M$

$$\gamma(t; q)\colon \mathcal{U}^q \to M \qquad t \mapsto \Phi^X(t, q)$$

is the unique maximal integral curve of X with initial condition q,

(iii) for every $t \in \mathbb{R}$ the set $\mathcal{U}_t = \{\, q \in M \mid (t, q) \in \mathcal{U} \,\}$ is an open set of M,

(iv) if we define $\Phi^X_t(q) = \Phi^X(t, q)$ then

(a) $\Phi^X_s \circ \Phi^X_t = \Phi^X_{t+s}$ when defined

(b) $\Phi^X_0 = \mathrm{id}_M$

(c) $\Phi^X_t : \mathcal{U}_t \to \mathcal{U}_{-t}$ is a diffeomorphism with inverse Φ^X_{-t},

(d) $\mathrm{d}\big(\Phi^X_t\big)_q(X(q)) = X(\Phi^X_t(q))$.

The function Φ^X is called flow *of X.*

The previous results about integral curves and flows of vector fields are discussed and proved in [Lee13, chapter 9]

Remark 1.3.9. If X is complete, then $\{ \Phi_t^X \}_{t\in\mathbb{R}}$ is a one parametric subgroup of $\mathrm{Diff}(M)$. It is convenient to introduce the exponential notation for the vector flow

$$\Phi_t^X = \exp(tX) \tag{1.34}$$

A vector field $X \in \mathrm{Vec}(M)$ induces an action on the algebra of functions $\mathcal{C}^\infty(M)$, defined as

$$X\colon \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M) \qquad f \mapsto Xf\,,$$

where

$$(Xf)(q) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \big(f \circ \exp(tX)\big)(q) \qquad q \in M\,. \tag{1.35}$$

If we fix $f_t = f \circ \exp(tX)$, its Taylor expansion with Lagrange remainder leads to

$$f_t(q) = f(q) + t(Xf)(q) + O(t^2) \tag{1.36}$$

and the reminder is locally uniform with respect to q, hence we can write

$$f_t = f + t(Xf) + O(t^2)\,. \tag{1.37}$$

This newly defined action is a derivation on the algebra, that is the Leibniz rule is satisfied:

$$X(fg) = X(f)g + fX(g) \qquad \forall f, g \in \mathcal{C}^\infty(M)\,. \tag{1.38}$$

Definition 1.3.10. Let $X \in \mathrm{Vec}(M)$ and $F\colon M \to N$ be a diffeomorphism between differentiable manifold. The *pushforward* $F_*X \in \mathrm{Vec}(N)$ is a vector field defined on N:

$$(F_*X)(F(q)) = \mathrm{d}F(X(q)) \qquad \forall q \in M\,. \tag{1.39}$$

When $F \in \mathrm{Diff}(M)$ is a diffeomorphism on M, we can rewrite eq. (1.39) as

$$(F_*X)(q) = \mathrm{d}F(X(F^{-1}(q)))\,. \tag{1.40}$$

Let us notice that, if F is not bijective, its pushforward may not be everywhere well defined. Moreover we stress the difference between $\mathrm{d}F$, which acts on TM, and F_* that is an operator defined from $\mathrm{Vec}(M)$ to itself.

Remark 1.3.11. For $X, Y \in \mathrm{Vec}(M)$, then

$$\begin{aligned} T_qM \ni \big(\exp(tX)_*Y\big)(q) &= \mathrm{d}\big(\exp(tX)\big)Y(\exp(-tX)(q)) && (1.41)\\ &= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \exp(tX) \circ \exp(sY) \circ \exp(-tX)(q) && (1.42)\end{aligned}$$

Lemma 1.3.12. *If $F \in \mathrm{Diff}(M)$, $X \in \mathrm{Vec}(M)$ and $f \in \mathcal{C}^\infty(M)$, then*

$$\begin{aligned} \exp(tF_*X) &= F \circ \exp(tX) \circ F^{-1}\,, && (1.43)\\ \big(F_*X\big)f &= \big(X(f \circ F)\big) \circ F^{-1} && (1.44)\end{aligned}$$

Proof. For eq. (1.43), it is sufficient to show that $t \mapsto F \circ \exp(tX) \circ F^{-1}$ is an integral curve of F_*X. The initial condition is the same and

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} F \circ \exp(tX) \circ F^{-1}(q) = \mathrm{d}F(X(F^{-1}(q))) = (F_*X)(q)\,.$$

For eq. (1.44)

$$\begin{aligned}(F_*X)f(q) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f \circ \exp(tF_*X)(q) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f \circ F \circ \exp(tX) \circ F^{-1}(q) \\ &= X(f \circ F) \circ F^{-1}(q)\,.\end{aligned}$$

□

1.4 Lie bracket of vector fields

Definition 1.4.1. Let $X, Y \in \mathrm{Vec}(M)$. We define their *Lie bracket* as the vector field

$$[X,Y](q) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \overbrace{\exp(-tX)_*Y(q)}^{\in T_qM}\,, \tag{1.45}$$

that is, using remark 1.3.11:

$$[X,Y](q) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \exp(-tX) \circ \exp(sY) \circ \exp(tX)(q)\,. \tag{1.46}$$

Remark 1.4.2. Let us notice that, using local coordinates, we obtain

$$\begin{aligned}&\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \exp(-sY) \circ \exp(-tX) \circ \exp(sY) \circ \exp(tX)(q) = \\ &\quad = -Y(q) + \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \exp(-tX) \circ \exp(sY) \circ \exp(tX)(q)\end{aligned} \tag{1.47}$$

and $Y(q)$ is independent of t, therefore we can also write

$$[X,Y](q) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \exp(-sY) \circ \exp(-tX) \circ \exp(sY) \circ \exp(tX)(q)\,. \tag{1.48}$$

Proposition 1.4.3. *As derivation on smooth functions, the Lie bracket acts as*

$$[X,Y] = X \circ Y - Y \circ X\,. \tag{1.49}$$

Proof. Using eq. (1.36) twice and considering eq. (1.44):

$$\begin{aligned}\big(\exp(-tX)_*Y\big)f &= Y(f \circ \exp(-tX)) \circ \exp(tX) \\ &= Y(f - tXf + O(t^2)) \circ \exp(tX) \\ &= (Yf - tYXf + O(t^2)) \circ \exp(tX) \\ &= (Yf - tYXf) + tX(Yf - tYXf) + O(t^2) \\ &= Yf + t(XY - YX)f + O(t^2)\,.\end{aligned}$$

This means

$$[X,Y]f = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \big(\exp(-tX)_* Y\big) f = (XY - YX) f \,. \qquad \square$$

Remark 1.4.4. Using proposition 1.4.3, the *Jacobi identity* easily follows:

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0 \,, \tag{1.50}$$

moreover using local coordinates for $X, Y \in \mathrm{Vec}(M)$

$$X = \sum_{i=1}^{d} X^i \partial_i \,, \qquad Y = \sum_{i=1}^{d} X^i \partial_i$$

we obtain

$$[X,Y] = \sum_{i=1}^{d} \sum_{j=1}^{d} \left(X^j \frac{\partial Y^i}{\partial x^i} - Y^j \frac{\partial X^i}{\partial x^i} \right) \partial_i \,. \tag{1.51}$$

Proposition 1.4.5. *Let* $F \in \mathrm{Diff}(M)$*. Then* F_* *is a Lie algebra homomorphism of* $\mathrm{Vec}(M)$*, that is*

$$F_*[X,Y] = [F_* X, F_* Y] \qquad \forall X, Y \in \mathrm{Vec}(X) \,. \tag{1.52}$$

Proof. We will show that they are equal as derivations of smooth functions. Let $f \in \mathcal{C}^\infty(M)$, then

$$\begin{aligned} F_* X(F_* Y f) &= F_* X(Y(f \circ F) \circ F^{-1}) \\ &= X(Y(f \circ F) \circ F^{-1} \circ F) \circ F^{-1} \\ &= X(Y(f \circ F)) \circ F^{-1} \,, \end{aligned}$$

and using this property

$$\begin{aligned} [F_* X, F_* Y] f &= F_*(F_* Y f) - F_* Y(F_* X f) \\ &= XY(f \circ F) \circ F^{-1} - YX(f \circ F) \circ F^{-1} \\ &= (XY - YX)(f \circ F) \circ F^{-1} \\ &= F_*[X,Y] f \,. \end{aligned} \qquad \square$$

Proposition 1.4.6. *Let* $X, Y \in \mathrm{Vec}(M)$*. The following are equivalent:*

(i) $[X,Y] = 0$,

(ii) $\exp(tX) \circ \exp(sX) = \exp(sX) \circ \exp(tX) \quad \forall t, s \in \mathbb{R}$.

Proof. Firstly we prove

$$[X,Y] = 0 \quad \Longrightarrow \quad \exp(-tX)_* Y = Y \quad \forall t \in \mathbb{R} \,.$$

It is sufficient to show that $\exp(-tX)_*Y$ is constant, since the equality trivially holds at $t=0$:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\exp(-tX)_*Y &= \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\exp(-(t+s)X)_*Y = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\exp(-tX)_*\exp(-sX)_*Y \\ &= \exp(-tX)_*\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\exp(-sX)_*Y = \exp(-tX)_*[X,Y] = 0\,.\end{aligned}$$

(i)⇒(ii). Fix $t \in \mathbb{R}$. The flow generated by Y is

$$\phi_s = \exp(-tX)\circ\exp(sY)\circ\exp(tX)\,.$$

Indeed

$$\frac{\mathrm{d}}{\mathrm{d}s}\phi_s = \exp(-tX)_*Y\circ\phi_s = Y\circ\phi_s\,.$$

Then, using uniqueness of the flow generated by a vector field as in theorem 1.3.8 we obtain

$$\exp(-tX)\circ\exp(sY)\circ\exp(tX) = \exp(sY) \qquad \forall t,s\in\mathbb{R}\,.$$

(ii)⇒(i). For every function $f \in \mathcal{C}^\infty$ we have

$$\begin{aligned}XYf &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} f\circ\exp(sY)\circ\exp(tX) \\ &= \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f\circ\exp(tX)\circ\exp(sY) = YXf\,.\end{aligned}$$

□

Proposition 1.4.7. *Let M be a differentiable d-dimensional manifold and $X_1,\dots,X_d$ be linearly independent vector fields in a neighbourhood of $q_0 \in M$. Then the map*

$$\Psi\colon \mathbb{R}^d \to M \qquad \Psi(t_1,\dots,t_d) = \exp(t_1X_1)\circ\cdots\circ\exp(t_dX_d)(q_0) \tag{1.53}$$

is a local diffeomorphism at $0\in\mathbb{R}^d$. Moreover, if $[X_i,X_j]=0$ for every $1\le i,j\le d$, then

$$\mathrm{d}\Psi_t(e_i) = X_i(\Psi(t))\,. \tag{1.54}$$

Proof. By theorem 1.3.8, Ψ is the composition of smooth maps in a neighbourhood $0\in\mathbb{R}^d$, hence a local smooth map. Let us compute its differential at zero. If $\gamma\colon s\mapsto se_i$, we have

$$\mathrm{d}\Psi_0(e_i) = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\exp(sX_i)(q_0) = X_i(q_0)\,.$$

This means that $\mathrm{d}\Psi_0$ sends a basis of $T_0\mathbb{R}^d$ to a basis of $T_{q_0}M$, which in turn means that it is invertible. By inverse function theorem we conclude that Ψ is a local diffeomorphism at zero.

Now we consider $t=(t_1,\dots,t_d)$ and a curve $\gamma_t\colon s\mapsto t+se_i$, then

$$\begin{aligned}\mathrm{d}\Psi_t(e_i) &= \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\exp(t_1X_1)\circ\cdots\circ\exp\big((s+t_i)X_i\big)\circ\cdots\circ\exp(t_dX_d)(q_0) \\ &= \mathrm{d}\big(\exp(t_1X_1)\circ\cdots\circ\exp(t_{i-1}X_{i-1})\big)X_i\big(\exp(t_iX_i)\circ\cdots\circ\exp(t_dX_d)\big)\,.\end{aligned}$$

If $[X_i, X_j] = 0$ for every $1 \leq i, j \leq d$, using proposition 1.4.6 we get

$$\begin{aligned} \mathrm{d}\Psi_t(e_i) &= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \exp(t_1X_1) \circ \cdots \circ \exp\big((s+t_i)X_i\big) \circ \cdots \circ \exp(t_dX_d)(q_0) \\ &= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \exp(sX_i) \circ \exp(t_1X_1) \circ \cdots \circ \exp(t_dX_d)(q_0) \\ &= X_i\big(\exp(t_1X_1) \circ \cdots \circ \exp(t_dX_d)(q_0)\big) = X_i\big(\Psi(t)\big) . \end{aligned}$$

□

1.5 Distributions and Frobenius theorem

Definition 1.5.1. Let M be a differentiable manifold. A *vector distribution* D of rank m on M is a family of vector subspaces $D_q \subseteq T_qM$ where $\dim D_q = m$ for every $\in M$.

A vector distribution D is said to be *smooth* if, for every $q_0 \in M$ there exists neighbourhood U_{q_0} of q_0 and a family of vector fields $X_1, \ldots, X_m$ such that

$$D_q = \operatorname{span}\{\, X_1(q), \ldots, X_m(q) \,\} \qquad \forall q \in U_{q_0} . \tag{1.55}$$

Definition 1.5.2. A smooth vector distribution D (of rank m) on M is said to be *involutive* if there exists a local basis $X_1, \ldots, X_m$ such that *eq.* (1.55) holds and smooth functions a_{ij}^k on M such that

$$[X_i, X_j] = \sum_{j=1}^{m} a_{ij}^k X_j \qquad \forall i, k = 1, \ldots, m . \tag{1.56}$$

If *eq.* (1.56) holds for one local basis D, then it holds for every local basis.

Definition 1.5.3. A smooth vector distribution D on M is said to be *flat* if, for every point $q_0 \in M$ there exists a local diffeomorphism $F\colon U_{q_0} \to \mathbb{R}^d$ such that $\mathrm{d}F_q(D_q) = \mathbb{R}^m \times 0$ for all $q \in U_{q_0}$.

Lemma 1.5.4. *If D is involutive vector distribution with local basis $X_1, \ldots, X_m$, then $\exp(tX_k)_*D = D$ for every $1 \leq k \leq m$.*

Proof. Let us define $Y_i^k(t) = \exp(tX_k)X_i$. Using *eq.* (1.56) and the same argument as in the first part of the proof of *eq.* (1.56) we obtain

$$\dot{Y}_i^k(t) = \exp(tX_k)_*[X_i, X_k] = \sum_{j=1}^{m} \exp(tX_k)_*\big(a_{ij}^k X_j\big) = \sum_{j=1}^{m} a_{ij}^k(t) Y_j^k(t) ,$$

where we define $a_{ij}^k(t) = a_{ij}^k \circ \exp(-tX_k)$. We write $A^k(t) = \big(a_{ij}^k(t)\big)_{i,j=1}^m$ and we consider the unique solution $\Gamma(t) = \big(\gamma_{ij}^k(t)\big)_{i,j=1}^m$ to the matrix Cauchy problem

$$\begin{cases} \dot{\Gamma}^k(t) = A^k(t)\Gamma^k(t) \\ \Gamma^k(0) = I \end{cases} .$$

Then we have

$$Y_i^k(t) = \sum_{j=1}^{m} \gamma_{ij}^k(t) Y_j^k(0)\,,$$

i.e. for every $1 \leq i, k \leq m$

$$\exp(tX_k)_* X_i = \sum_{j=1}^{m} \gamma_{ij}^k(t) X_j\,,$$

which proves the statement. □

Theorem 1.5.5 (Frobenius theorem)**.** *A smooth distribution D is involutive if and only if it is flat.*

Proof. Both involutiveness and flatness are local notions, hence, for every $q_0 \in M$ we need to verify the statement only for a proper neighbourhood of q_0.

($\Leftarrow$). If D is flat, then we consider a local diffeomorphism $F\colon U_{q_0} \to \mathbb{R}^d$ such that $D_q = (\mathrm{d}F)_q^{-1}(\mathbb{R}^m \times \{\,0\,\})$. It follows that, for every $q \in U_{q_0}$, we have

$$D_q = \operatorname{span}\{\, X_1(q), \dots, X_m(q) \,\}\,, \qquad X_i(q) = (\mathrm{d}F)_q^{-1}\partial_{x^i}\,.$$

For every $i, k = 1, \dots, m$ we obtain

$$[X_i, X_k] = \Big[(\mathrm{d}F)_q^{-1}\partial_{x^i}, (\mathrm{d}F)_q^{-1}\partial_{x^j}\Big] = (\mathrm{d}F)_q^{-1}[\partial_{x^i}, \partial_{x^j}] = 0\,.$$

($\Rightarrow$). As before, we consider a neighbourhood of q_0 such that *eq.* (1.55) and *eq.* (1.56) holds. We complete $X_1(q), \dots, X_m(q)$ to a basis for T_qM:

$$T_qM = \operatorname{span}\{\, X_1(q), \dots, X_m(q), Z_{m+1}(q), \dots, Z_d(q) \,\}\,,$$

in a neighbourhood of q_0 and let define $\Psi\colon \mathbb{R}^d \to M$ as in proposition 1.4.7:

$$\Psi(t_1, \dots, t_m, s_{m+1}, \dots, s_d) = \exp(t_1X_1) \circ \cdots \circ \exp(s_dX_d)(q_0)\,.$$

Then, following the proof to proposition 1.4.7, Ψ is a local diffeomorphism at $(t, s) = (0, 0)$ and in a neighbourhood of $(0, 0)$ we get

$$\mathrm{d}\Psi_{(t,s)}(e_i) = \big(\exp(t_1X_1)_* \dots \exp(t_iX_i)_* X_i\big)\big(\Psi(t, s)\big)$$

for every $1 \leq i \leq m$. These vectors are linear independent and, using lemma 1.5.4, they belong to D. Therefore

$$D_q = \mathrm{d}\Psi\big(\operatorname{span}\{\, e_1, \dots, e_m \,\}\big) \qquad q = \Psi(t, s)\,,$$

which means D in flat. □

Chapter 2

Introduction to sub-Riemannian geometry

In this chapter we start to dive into the real subject of this work, which is sub-Riemannian geometry. We will introduce sub-Riemannian structures using the formalism of vector bundles and this will grant a sufficient amount of generality. We will show the notion of equivalence of sub-Riemannian structures and we will explain how every sub-Riemannian structure is essentially free. The latter result will simplify the discussion later. The dissertation on the topic will follow [ABB20, chapter 3]. Hereafter we always assume that M is a differentiable manifold which is Hausdorff and has a countable basis.

2.1 Introductory definitions

Definition 2.1.1. Let M be a differentiable manifold and $\mathcal{F} \subseteq \mathrm{Vec}(M)$ a family of smooth vector fields. We denote with $\mathrm{Lie}\,\mathcal{F}$ the smallest Lie subalgebra of $\mathrm{Vec}(M)$ containing $\mathcal{F}$. We say that $\mathcal{F}$ is *bracket-generating* (or that satisfies the *Hörmander condition*) if

$$\mathrm{Lie}_q\,\mathcal{F} = \{\, X(q) \mid X \in \mathrm{Lie}\,\mathcal{F} \,\} = T_qM \quad \forall q \in M\,.$$

Definition 2.1.2. Let M be a differentiable connected manifold. A *sub-Riemannian structure* on M is a pair $(\mathbf{U}, f)$ where:

(i) $\mathbf{U}$ is an Euclidean bundle with base M and fibers U_q, that is, a vector bundle $\mathbf{U}$ equipped with a positively defined section $(\,\cdot\,|\,\cdot\,)$ of $S^k\,\mathbf{U}$. Explicitly, for every $q \in M$, the fiber U_q is a vector space equipped with a scalar product $(\,\cdot\,|\,\cdot\,)_q$ that changes smoothly with respect to q. For $u \in U_q$ we consider $|u|^2 = (u|u)_q$ the norm of u.

(ii) $f\colon \mathbf{U} \to TM$ is a morphism of vector bundles (we recall definition 1.2.12).

(iii) The *set of horizontal fields* $\mathcal{D} = \{\, f(\sigma) \mid \sigma \text{ section of } \mathbf{U} \,\}$ is a bracket-generating family of vector fields.

Whenever the vector bundle $\mathbf{U}$ admits a global trivialization, we say that $(\mathbf{U}, f)$ is a *free sub-Riemannian structure*. The triple $(M, \mathbf{U}, f)$ where $(\mathbf{U}, f)$ is a sub-Riemannian structure on M is called a *sub-Riemannian manifold*.

Definition 2.1.3. Let $(M, \mathbf{U}, f)$ be a sub-Riemannian manifold, the associated *distribution* is the family of subspaces

$$\{ \mathcal{D}_q \}_{q \in M} \quad \text{where} \quad \mathcal{D}_q = f(U_q) \subseteq T_q M\,. \tag{2.1}$$

We call $m \in \operatorname{rank}(\mathbf{U})$ the *bundle rank* of the sub-Riemannian structure, and $r(q) = \dim \mathcal{D}_q$ the *rank* of the sub-Riemannian structure at $q \in M$. We say that the sub-Riemannian structure $(\mathbf{U}, f)$ on M has *constant rank* if $r(q)$ is constant.

Remark 2.1.4. Let us notice that the distribution associated to a sub-Riemannian manifold is a smooth distribution. Indeed, for every $q \in M$ we can write

$$\mathcal{D}_q = \{ X(q) \mid X \in \mathcal{D} \}$$

and the image of a section of $\mathbf{U}$ via a vector bundle morphism f is again a section of TM.

Hereafter we denote points in $\mathbf{U}$ as a pairs (q, u), where $q \in M$ and $u \in U_q$. We will usually denote $f(q, u)$ with $f_u(q)$ to stress that it is a tangent vector at $q \in M$.

Remark 2.1.5. We recall that a curve defined of a metric space $\gamma \colon I \subseteq \mathbb{R} \to X$ is said to be *Lipschitz* if there exists $L \geq 0$ such that

$$d_X(\gamma(t), \gamma(s)) \leq L|t - s| \qquad \forall\, t, s \in I\,, \tag{2.2}$$

and L is said to be a Lipschitz constant for the curve γ. Moreover, γ is said to be *locally Lipschitz* if it is Lipschitz in every compact subset of I. However, the previous notion can be extended to curves defined on a differentiable manifold M even if M is not explicitly endowed with a metric space structure, provided that γ is defined only on a bounded interval $I = [a, b]$. In this case the image of γ is a compact. The curve is said to be Lipschitz if its restriction to every chart is locally Lipschitz. We know that a locally-Lipschitz curve defined on $\mathbb{R}^d$ is differentiable almost everywhere. Hence we can define a tangent vector for γ at t for almost every $t \in [a, b]$ through the local trivialization of TM. This definition is independent of the local trivialization. We denote with $\dot\gamma \in TM$ the derivative of γ, defined almost everywhere.

Definition 2.1.6. A Lipschitz curve $\gamma \colon [0, T] \to M$ is said to be *admissible* (or *horizontal*) for a sub-Riemannian structure if there exists a measurable (in terms of local trivialization) and essentially bounded function

$$u \colon t \in [0, T] \mapsto u(t) \in U_{\gamma(t)}\,, \tag{2.3}$$

called *control*, such that

$$\dot\gamma(t) = f(\gamma(t), u(t)) \tag{2.4}$$

for almost every $t \in [0, T]$. We say that $u(\cdot)$ is a *control corresponding* to γ. If f is not injective on fibers, then multiple controls may corresponds to the same admissible curve.

Definition 2.1.7. Let M be a differentiable manifold, a *nonautonomous vector field* is a family of vector fields $\{ X_t \}_{t\in\mathbb{R}} \subseteq \mathrm{Vec}(M)$ such that the map $X(t,q) = X_t(q)$ satisfies the following properties:

(a) the map $t \mapsto X(t,q)$ is measurable, for every fixed $q \in M$,

(b) the map $q \mapsto X(t,q)$ is smooth, for every fixed $t \in \mathbb{R}$,

(c) for every system of local coordinates defined on $\Omega \subseteq M$ and every compact $K \subseteq \Omega$ and compact interval $I \subseteq \mathbb{R}$ there exist two functions $c(t), k(t) \in L^\infty(I)$ such that for all $(t,x),(t,y) \in I \times K$

$$\|X(t,x)\| \leq c(t) \quad \text{and} \quad \|X(t,x) - X(t,y)\| \leq k(t)\|x-y\| \,. \tag{2.5}$$

Theorem 2.1.8 (Carathéodory theorem)**.** *Let $\{ X_t \}_{t\in\mathbb{R}}$ be a nonautonomous vector field defined on a differentiable manifold M. Then the Cauchy problem*

$$\begin{cases} \dot q(t) = X(t,q(t)) \\ q(t_0) = q_0 \end{cases} \tag{2.6}$$

has a unique solution $\gamma(t;t_0,q_0)$ defined on a open interval I containing t_0 such that eq. (2.6) *is satisfied for almost every $t \in I$ and $\gamma(t_0;t_0,q_0) = q_0$. Moreover the map $(t,q_0) \mapsto \gamma(t;t_0,q_0)$ is locally-Lipschitz with respect to t and smooth with respect to q_0.*

Remark 2.1.9. We assume that the nonautonomous vector fields X_t are *complete*, that is, for all $t_0 \in \mathbb{R}$ and $q_0 \in M$ the solution $\gamma(t;t_0,q_0)$ to eq. (2.6) is defined on $I = \mathbb{R}$. Let us denote $P_{t_0,t}(q) = \gamma(t;t_0,q_0)$. The family of maps $\{ P_{t,s} \}_{t,s\in\mathbb{R}}$ where $P_{t,s} \colon M \to M$ is the *nonautonomous flow* generated by X_t.

By definition, for every fixed $t_0 \in \mathbb{R}$, the nonautonomous flow $t \mapsto P_{t_0,t}$ associated to a nonautonomous vector field X_t is locally Lipschitz and satisfies the equation

$$\partial_t P_{t_0,t}(q) = X(t, P_{t_0,t}(q)) \qquad q \in M \,, \tag{2.7}$$

for almost every t. Moreover the following identities hold

$$P_{t,t} = \mathrm{id} \qquad \forall t \in \mathbb{R} \,, \tag{2.8}$$

$$P_{t_2,t_3} \circ P_{t_1,t_2} = P_{t_1,t_3} \qquad \forall t_1,t_2,t_3 \in \mathbb{R} \,, \tag{2.9}$$

$$\left(P_{t_1,t_2}\right)^{-1} = P_{t_2,t_1} \qquad \forall t_1,t_2 \in \mathbb{R} \,. \tag{2.10}$$

Conversely, for a family of smooth diffeomorphism $P_{t,s} \colon M \to M$ satisfying the previous relations, one can associate its *infinitesimal generator* X_t as follows:

$$X_t(q) = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} P_{t,t+s}(q) \qquad \forall q \in M \,. \tag{2.11}$$

We will use these results to characterize admissible curves.

Remark 2.1.10. Given a local trivialization $\mathcal{U}_q \times \mathbb{R}^m$ for the vector bundle $\mathbf{U}$, where $\mathcal{U}_q$ is a neighbourhood of $q \in M$, we can consider a basis in the fibers and the map f is expressed as $f(q,u) = \sum_{i=1}^m u_i f_i(q)$, where m is the rank of $\mathbf{U}$. In this local trivialization, a Lipschitz curve $\gamma\colon [0,T] \to M$ is admissible if there exists $u = (u_1, \dots, u_m) \in L^\infty([0,T], \mathbb{R}^m)$ such that

$$\dot{\gamma}(t) = \sum_{i=1}^m u_i(t) f_i(\gamma(t)) \qquad \text{for almost every } t \in [0,T]\,. \tag{2.12}$$

In the previous expression, f_i represents locally defined smooth vector fields. If we set $X_t(q) = \sum_{i=1}^m u_i(t) f_i(q)$, conditions (a) and (b) in definition 2.1.7 are trivially satisfied. Referring to (c), thanks to the smoothness of f_i, for each compact $K \subseteq \mathcal{U}_q$ there are positive constants C_K, L_K such that for all $i = 1, \dots, m$ and $j = 1, \dots, d$

$$\|f_i(x)\| \leq C_K \quad \text{and} \quad \left\| \frac{\partial f_i}{\partial x_j}(x) \right\| \leq L_K \quad \forall x \in K\,. \tag{2.13}$$

Therefore, for all $(t,x), (t,y) \in I \times K$ we get

$$\|X(t,x)\| \leq C_K \sum_{i=1}^m |u_i(t)| \qquad \|X(t,x) - X(t,y)\| \leq L_K \sum_{i=1}^m |u_i(t)| \cdot \|x - y\|\,. \tag{2.14}$$

Hence also the condition (c) holds. Theorem 2.1.8 thus ensures that there exists an admissible curve γ defined on a sufficiently small interval, such that u is a control associated γ and $\gamma(0) = q$.

The following result for this kind of nonautonomous vector fields will be useful:

Proposition 2.1.11. *Let $\{ X_t \}_{t \in \mathbb{R}}$ be a nonautonomous vector field on M that can be expressed as*

$$X_t(q) = \sum_{i=1}^m u_i(t) f_i(q) \tag{2.15}$$

for some control $u \in L^\infty(\mathbb{R}, \mathbb{R}^m)$, and suppose that for some $q_0 \in M$ there is a solution to the associated Cauchy problem with initial condition $q(0) = q_0$, defined on an interval $[0,T]$. Then:

(i) there is a neighbourhood O_{q_0} of q_0 such that for every $q' \in O_{q_0}$ the Cauchy problem associated to $\{ X_t \}_{t \in \mathbb{R}}$ and initial condition $q(0) = q'$ has a solution defined on $[0,T]$,

(ii) there is an L^∞ neighbourhood $\mathcal{V}$ of $u \in L^\infty(\mathbb{R}, \mathbb{R}^m)$ such that, for every $v \in \mathcal{V}$, the Cauchy problem

$$\begin{cases} \dot{q}(t) = \sum_{i=1}^m v_i(t) f_i(q) \\ q(0) = q_0 \end{cases} \tag{2.16}$$

has a solution defined on $[0,T]$.

The previous results regarding control theory are all proved and discussed in [BP07, chapters 2 and 3].

Example 2.1.12. There may be Lipschitz curves such that $\dot\gamma(t) \in \mathcal{D}_{\gamma(t)}$ for almost every $t \in [0,T]$ and still they are not admissible. For instance, we consider two different sub-Riemannian structures on $\mathbb{R}^2$ with bundle rank two and defined as follows:

$$f(x,y,u_1,u_2) = (x,y,u_1,u_2 x) \quad \text{and} \quad f'(x,y,u_1,u_2) = (x,y,u_1,u_2 x^2)\,.$$

We name $\mathcal{D}$ and $\mathcal{D}'$ the corresponding families of horizontal vector fields. The curve $\gamma\colon t \mapsto (t,t^2)$ satisfies both $\dot\gamma(t) \in \mathcal{D}_{\gamma(t)}$ and $\dot\gamma(t) \in \mathcal{D}'_{\gamma(t)}$. On the other hand, γ is admissible for f, since its corresponding control is $(u_1,u_2) = (1,2)$,but not for f' since its uniquely determined control would be $(u_1,u_2) = (1,2/t)$ which is not essentially bounded (nor integrable) for $t \in [0,1]$.

Remark 2.1.13. As we stressed in the previous example, two sub-Riemannian structures $(\mathbf{U},f)$ and $(\mathbf{U}',f')$ on the same manifold may have different horizontal distributions $\mathcal{D} \neq \mathcal{D}'$ despite the fact that $\mathcal{D}_q = \mathcal{D}'_q$ for every $q \in M$.

Proposition 2.1.14. *Let $\mathcal{D}$ and $\mathcal{D}'$ be two distributions with constant rank, then*

$$\mathcal{D}_q = \mathcal{D}'_q \quad \forall q \in M \qquad \Longleftrightarrow \qquad \mathcal{D} = \mathcal{D}'\,. \tag{2.17}$$

Proof. For $q_0 \in M$ we consider a basis $X_1(q_0),\dots,X_m(q_0)$ of D_{q_0}, then $X_1,\dots,X_m \in \mathcal{D}$ defines a basis for $\mathcal{D}_q$ for q in a neighbourhood of q_0, the same holds for $X'_1,\dots,X'_m \in \mathcal{D}'$. The change of variables is smooth in q. If $X \in \mathcal{D}$, then X is a linear combination of $X_1,\dots,X_m \in \mathcal{D}$ and, using the smooth change of variable, it is also a linear combination of $X'_1,\dots,X'_m \in \mathcal{D}'$, hence $X \in \mathcal{D}'$. Therefore $\mathcal{D} = \mathcal{D}'$, the inverse implication is trivial. □

2.2 Length of admissible curves

Definition 2.2.1. Let $v \in \mathcal{D}_q$. We define the *sub-Riemannian norm* of v as

$$\|v\| = \min\{\, |u| \mid u \in U_q \text{ such that } f(q,u) = v \,\}\,. \tag{2.18}$$

Since f is linear on fiber, the minimum in eq. (2.18) is always obtained and it is unique. The argument u that realises the minimum of $|u|$ is the orthogonal projection of the origin $0 \in U_q$ onto the affine plane $\{\, u \mid f(q,u) = v \,\}$.

Remark 2.2.2. Notice that $\ker f_q$ defines a subspace in U_q: each $u \in U_q$ that realizes a minimum in eq. (2.18) lies in the orthogonal V_q of $\ker f_q$ in U_q and there is a canonical isomorphism between $V_q \cong U_q/\ker f_q$ and $\mathcal{D}_q$. The sub-Riemannian norm is the induced Euclidean norm from V_q though this isomorphism. Therefore, the sub-Riemannian norm is actually a norm and it is Euclidean, that is induced by a scalar product.

Moreover, if f_q is injective, then $U_q \cong \mathcal{D}_q$ and f_q induces an isometry from U_q to $\mathcal{D}_q$ with the sub-Riemannian norm. In particular f_q sends an orthonormal basis into an orthonormal basis.

Definition 2.2.3. Given an admissible curve $\gamma\colon [0,T] \to M$, we define at every differentiability point of γ

$$u^*(t) = \operatorname{argmin} \{ \, |u| \mid u \in U_q \text{ such that } \dot\gamma(t) = f(\gamma(t), u) \, \} \,, \tag{2.19}$$

which is unique by definition 2.2.1. We say that $u^*(t)$ is the *minimal control* associated to γ.

We recall that, since every admissible curve is Lipschitz, it is differentiable almost everywhere on $[0,T]$. The following result ensures that the minimal control is actually a control.

Proposition 2.2.4. *Let $\gamma\colon [0,T] \to M$ be an admissible curve. Then its minimal control $u^*(\cdot)$ is measurable (in terms of local trivializations) and essentially bounded on $[0,T]$.*

The proof of the previous proposition is quite technical. We need some auxiliary lemmas. Let us consider $I = [a,b]$, a compact set $U \subseteq \mathbb{R}^m$ and two functions $g\colon I \times U \to \mathbb{R}^n$, $v\colon I \to \mathbb{R}^n$ such that

$$g(\cdot, u) \text{ is measurable in } t \text{ for every fixed } u \in U, \tag{M1}$$

$$g(t, \cdot) \text{ is continuous in } u \text{ for every fixed } t \in I, \tag{M2}$$

$$v(t) \text{ is measurable with respect to } t, \tag{M3}$$

$$\min \{ \, |u| \mid g(t,u) = v(t), u \in U \, \} \text{ has a unique solution for every } t. \tag{M4}$$

We denote with $u^*(t)$ the solution of (M4) for a fixed $t \in I$.

Lemma 2.2.5. *Under assumptions (M1)-(M4), the function $t \mapsto |u^*(t)|$ is measurable on I.*

Proof. We will show that for any fixed $r \geq 0$ the set

$$A = \{ \, t \in I \mid |u^*(t)| \leq r \, \}$$

is measurable in $\mathbb{R}$, which is sufficient for the measurability of $t \mapsto |u^*(t)|$. By definition

$$A = \{ \, t \in I \mid \exists u \in U \colon |u| \leq r, \, g(t,u) = v(t) \, \} \,.$$

We fix $r > 0$ and a countable dense set $\{ \, u_i \, \}_{i \in \mathbb{N}}$ in the ball of radius r contained in U, then we define

$$A_{i,n} = \{ \, t \in I \mid |g(t, u_i) - v(t)| < 1/n \, \} \qquad A_n = \bigcup_{i \in \mathbb{N}} A_{i,n} \,.$$

Let us show that $A = \bigcap_{i \in \mathbb{N}} A_n$:

- $\subseteq$ Let $t \in A$, then there exists $\overline{u} \in U$ such that $|\overline{u}| \leq r$ and $g(t, \overline{u}) = v(t)$. Since g is continuous with respect to u and $\{ \, u_i \, \}_{i \in \mathbb{N}}$ is dense, for each n we can find u_{i_n} such that $|g(t, u_{i_n}) - v(t)| < 1/n$, this means that $t \in A_n$ for all n.

$\supseteq$ Let us assume $t \in A_n$ for all n. Then for every n there exists i_n such that $|g(t, u_{i_n}) - v(t)| < 1/n$. Then, by compactness, we have convergence up to a subsequence $u_{i_n} \to \overline{u}$. By continuity of g with respect to u we get $g(t, \overline{u}) = v(t)$, that means $t \in A$.

Since all $A_{i,n}$ are measurable by construction, then also A_n is measurable for any n. By the proved equality, A is measurable. □

Lemma 2.2.6. *Under assumptions (M1)-(M4), the vector function $t \mapsto u^*(t)$ is measurable on I.*

Proof. Let us denote $\varphi(t) = |u^*(t)|$. We will show that for any closed ball O in $\mathbb{R}^n$ the set

$$B = \{ t \in I \mid u^*(t) \in O \}$$

is measurable, this is sufficient for the measurability of $t \mapsto u^*(t)$. Since the minimum in (M4) is unique, we can also write

$$B = \{ t \in I \mid \exists u \in O\colon |u| = \varphi(t),\, g(t, u) = v(t) \} .$$

We fix a closed ball O and a countable dense set $\{ u_i \}_{i \in \mathbb{N}}$ in O, then we define

$$B_{i,n} = \{ t \in I \mid |g(t, u_i) - v(t)| < 1/n,\, |u_i| < \varphi(t) + 1/n \} \qquad A_n = \bigcup_{i \in \mathbb{N}} A_{i,n} .$$

Let us show that $B = \bigcap_{i \in \mathbb{N}} B_n$:

$\subseteq$ Let $t \in B$, then there exists $\overline{u} \in O$ such that $|\overline{u}| = \varphi(t)$ and $g(t, \overline{u}) = v(t)$. Since g is continuous with respect to u and $\{ u_i \}_{i \in \mathbb{N}}$ is dense in O, for each n we can find u_{i_n} such that $|g(t, u_{i_n}) - v(t)| < 1/n$ and $|u_{i_n}| < \varphi(t) + 1/n$, this means that $t \in B_n$ for all n.

$\supseteq$ Let us assume $t \in B_n$ for all n. Then for every n there exists i_n such that $|g(t, u_{i_n}) - v(t)| < 1/n$ and $|u_{i_n}| < \varphi(t) + 1/n$. Then, by compactness of O, we have convergence up to a subsequence $u_{i_n} \to \overline{u}$. By continuity of g with respect to u we get $g(t, \overline{u}) = v(t)$, moreover $|\overline{u}| \leq \varphi(t)$. Hence $|\overline{u}| = \varphi(t)$ and therefore $t \in B$.

Since all $B_{i,n}$ are measurable by construction, also B_n is measurable for any n. By the proved equality, B is measurable. □

We are now ready to prove proposition 2.2.4.

Proof of proposition 2.2.4. We consider an admissible curve $\gamma\colon [0, T] \to M$, in terms of local trivialization TM is locally identified with $\mathcal{U}_q \times \mathbb{R}^n$ and the vector bundle $\mathbf{U}$ is locally identified with $\mathcal{U}_q \times \mathbb{R}^m$. Since γ is admissible, a generic control is essentially bounded and therefore its minimal control is bounded, hence contained in a compact set $U \subseteq \mathbb{R}^m$. We define

$$g\colon [0, T] \times U \to \mathbb{R}^n \qquad g(t, u) = f_{\gamma(t)}(u) ,$$
$$v\colon [0, T] \to \mathbb{R}^n \qquad v(t) = \dot{\gamma}(t) .$$

Assumptions (M1)-(M3) are satisfied since $g(t,u)$ is linear with respect to u and measurable in t. Condition (M4) follows from the linearity of f with respect to u as we previously observed. We can now apply lemma 2.2.6 and conclude that the minimal control $u^*(t)$ is measurable in t. □

Thanks to this measurability result, we can define the following:

Definition 2.2.7. Let $\gamma\colon [0,T] \to M$ be an admissible curve and $u^*(t)$ its minimal control. We define the *sub-Riemannian length* of γ as

$$\ell(\gamma) = \int_0^T \|\dot\gamma(t)\| \,\mathrm{d}t = \int_0^T |u^*(t)| \,\mathrm{d}t\,. \tag{2.20}$$

We say that γ is *parametrized by arc length* (or *arc length parametrized*) if $\|\dot\gamma(t)\| = 1$ for almost every $t \in [0,T]$.

Lemma 2.2.8. *The length of an admissible curve is invariant under Lipschitz reparametrization.*

Proof. Let $\gamma\colon [0,T] \to M$ be an admissible curve and $\varphi\colon [0,T'] \to [0,T]$ a Lipschitz reparametrization, that is a Lipschitz monotone surjective map. Consider the reparametrized curve

$$\gamma_\varphi\colon [0,T'] \to M \qquad \gamma_\varphi = \gamma \circ \varphi\,.$$

Let us notice that γ_φ is again Lipschitz, since it is the composition of Lipschitz maps. Moreover, since f is linear on fibers, its minimal control is $(u^* \circ \varphi)\dot\varphi$ where u^* is the minimal control of γ. Since $(u^* \circ \varphi)\dot\varphi$ is essentially bounded, γ_φ is admissible. Using the change of variable $t = \varphi(s)$ and recalling the chain rule for Sobolev spaces, we can compute

$$\begin{aligned}\ell(\gamma_\varphi) &= \int_0^{T'} \|\dot\gamma_\varphi(s)\| \,\mathrm{d}s = \int_0^{T'} |u^*(\varphi(s))| \cdot |\dot\varphi(s)| \,\mathrm{d}s \\ &= \int_0^T |u^*(t)| \,\mathrm{d}t = \int_0^T \|\dot\gamma(t)\| \,\mathrm{d}t = \ell(\gamma)\,.\end{aligned}$$

□

Lemma 2.2.9. *Every admissible curve of positive length is a Lipschitz reparametrization of an arc length parametrized admissible one.*

Proof. Let $\gamma\colon [0,T] \to M$ be an admissible curve with $\ell(\gamma) > 0$ and minimal control u^*. Consider the Lipschitz monotone function $\varphi\colon [0,T] \to [0,\ell(\varphi)]$ defined by

$$\varphi(t) = \int_0^t |u^*(\tau)| \,\mathrm{d}\tau\,.$$

Notice that if $\varphi(t_1) = \varphi(t_2)$, then $\gamma(t_1) = \gamma(t_2)$ since φ is monotone. Hence we can define the curve $\zeta\colon [0,\ell(\gamma)] \to M$ by

$$\zeta(s) = \gamma(t) \qquad \text{if } s = \varphi(t) \text{ for some } t \in [0,T].$$

Therefore we have $\gamma = \zeta \circ \varphi$. We need to show that ζ is Lipschitz, let us consider t_0, t_1 such that $\gamma(t_0)$ and $\gamma(t_1)$ lies in the same local chart. Representing the Euclidean norm in local coordinate by $|\cdot|$, we would like to prove

$$|\gamma(t_1) - \gamma(t_0)| \le C \int_{t_0}^{t_1} |u^*(\tau)| \,\mathrm{d}\tau\,.$$

To do so, we fix $K \subseteq M$ a compact set such that $\gamma([0,T]) \subseteq K$ and we define $C = \max_{x\in K} \left(\sum_{i=1}^m |f_i(x)|^2 \right)^{1/2}$. Then

$$\begin{aligned}
|\gamma(t_1) - \gamma(t_0)| &\le \int_{t_0}^{t_1} \sum_{i=1}^m |u_i^*(t) f_i(\gamma(t))| \,\mathrm{d}t \\
&\le \int_{t_0}^{t_1} \left(\sum_{i=1}^m |u_i^*(t)|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m |f_i(\gamma(t))|^2 \right)^{\frac{1}{2}} \mathrm{d}t \\
&\le C \int_{t_0}^{t_1} |u^*(t)| \,\mathrm{d}t\,.
\end{aligned}$$

Hence if $s_1 = \varphi(t_1)$ and $s_0 = \varphi(t_0)$ one has

$$|\zeta(s_1) - \zeta(s_0)| = |\gamma(t_1) - \gamma(t_0)| \le C \int_{t_0}^{t_1} |u^*(\tau)| \,\mathrm{d}\tau = C|s_1 - s_0|\,, \tag{2.21}$$

thus ζ is Lipschitz. In particular $\dot\zeta(s)$ is defined almost everywhere on $[0, \ell(\gamma)]$. Finally we prove that ζ is admissible and its minimal control has norm one. Firstly we notice that the set

$$C_\varphi = \{\, s \in \mathbb{R} \mid s = \varphi(t),\ \dot\varphi(t) \text{ exists},\ \dot\varphi(t) = 0 \,\}$$

has measure zero. Indeed we consider a decreasing sequence of open sets $(A_n)_n$ containing C_φ and such that $\mu(A_n) \to \mu(C_\varphi)$. Since each A_n is open, it is a countable union of open intervals. Let us notice that $\mu(\varphi(I)) \le 2\int_I |\dot\varphi|$ and this holds also for countable union of intervals, therefore

$$\mu(\varphi(A_n)) \le 2 \int_{A_n} |\dot\varphi| = 2 \int_{A_n \setminus C_\varphi} |\dot\varphi| + 2 \int_{C_\varphi} |\dot\varphi| \le 2L\,\mu(A_n \setminus C_\varphi) \underset{n\to\infty}{\to} 0\,. \tag{2.22}$$

Thus we can define for every s such that $s = \varphi(t)$, $\dot\varphi$ exists and $\dot\varphi(t) \neq 0$, the control

$$v(s) = \frac{u^*(t)}{\dot\varphi(t)} = \frac{u^*(t)}{|u^*(t)|}\,.$$

Notice that the control is defined almost everywhere on $s \in [0, \ell(\gamma)]$. Moreover by construction $|v(s)| = 1$ almost everywhere and v is the minimal control associated to ζ. $\square$

Remark 2.2.10. In this section, we developed the theory of Lipschitz admissible curves, one may pursue other approaches using $W^{1,2}$ admissible curves (corresponding to L^2 controls) or absolutely continuos admissible curves (corresponding to L^1 controls). In the final part we will explain in what extent these approaches are equivalent. We remark again that the notion of absolutely continuous curves from a compact interval to a differentiable manifold is well defined (similarly to Lipschitz curves).

Definition 2.2.11. An absolutely continuos curve on a sub-Riemannian manifold M, $\gamma\colon [0,T] \to M$, is said to be *AC-admissible* if there exists an L^1 function $u\colon t \in [0,T] \mapsto u(t) \in U_{\gamma(t)}$ such that $\dot\gamma(t) = f(\gamma(t), u(t))$ for almost every $t \in [0,T]$.

An absolutely continuos curve $\gamma\colon [0,T] \to M$ on a sub-Riemannian manifold M is said to be $W^{1,2}$*-admissible* if there exists an L^2 function $u\colon t \in [0,T] \mapsto u(t) \in U_{\gamma(t)}$ such that $\dot\gamma(t) = f(\gamma(t), u(t))$ for almost every $t \in [0,T]$.

Hence every Lipschitz admissible curve is $W^{1,2}$-admissible and every $W^{1,2}$-admissible curve is AC-admissible. However, thanks to the following statements, these admissible notions are equivalent up to reparametrization.

In the context of AC-admissible curves, lemma 2.2.8 can be rephrased as

Lemma 2.2.12. *The length of an AC-admissible curve is invariant by AC reparametrization.*

and lemma 2.2.9 becomes

Lemma 2.2.13. *Any AC-admissible curve of a positive length is AC reparametrization of an arc length parametrized admissible one.*

The proof of these lemmas follow the ones for lemma 2.2.8 and lemma 2.2.9 (every instance of L^∞ is replaced with L^1). As a consequence of these results, if we define

$$d_{AC}(q_0, q_1) = \inf \{\, \ell(\gamma) \mid \gamma \text{ is AC-admissible, } \gamma(0) = q_0,\ \gamma(T) = q_1 \,\}$$
$$d_{W^{1,2}}(q_0, q_1) = \inf \{\, \ell(\gamma) \mid \gamma \text{ is } W^{1,2}\text{-admissible, } \gamma(0) = q_0,\ \gamma(T) = q_1 \,\}$$

we get the following:

Proposition 2.2.14.

$$d_{AC}(q_0, q_1) = d_{W^{1,2}}(q_0, q_1) = d(q_0, q_1) \qquad \textit{for all } q_0, q_1 \in M\,. \tag{2.23}$$

2.3 Equivalence of sub-Riemannian structures

Definition 2.3.1. Let $(\mathbf{U}, f)$ and $(\mathbf{U}', f')$ be two sub-Riemannian structures on a differentiable manifold M. They are said to be *equivalent as distributions* if there exist an Euclidean bundle $\mathbf{V}$ and two surjective vector bundle morphism $p\colon \mathbf{V} \to \mathbf{U}$ and

$p' \colon \mathbf{V} \to \mathbf{U}$ such that the following diagram is commutative:

$$\begin{array}{ccccc} & & \mathbf{U} & & \\ & \overset{p}{\nearrow} & & \overset{f}{\searrow} & \\ \mathbf{V} & & & & TM \\ & \underset{p'}{\searrow} & & \underset{f'}{\nearrow} & \\ & & \mathbf{U}' & & \end{array} \tag{2.24}$$

The structures $(\mathbf{U}, f)$ and $(\mathbf{U}', f')$ are said to be *equivalent as sub-Riemannian structures* (or simply *equivalent*) if they are equivalent as distributions and moreover p, p' are compatible with the scalar product, that is

$$|u| = \min \{ |v| \mid p(v) = u \} \qquad \forall u \in \mathbf{U}\,, \tag{2.25}$$

$$|u'| = \min \{ |v| \mid p'(v) = u \} \qquad \forall u' \in \mathbf{U}\,. \tag{2.26}$$

Remark 2.3.2. If $(\mathbf{U}, f)$ and $(\mathbf{U}', f)$ are equivalent as sub-Riemannian structures on M, then

(a) the distributions $\mathcal{D}_q$ and $\mathcal{D}'_q$ defined by f and f' coincide, since $f(U_q) = f'(U'_q)$ for every $q \in M$,

(b) for each $w \in \mathcal{D}_q$ we have $\|w\| = \|w\|'$, where $\|\cdot\|$ and $\|\cdot\|'$ are norms induced by $(\mathbf{U}, f)$ and $(\mathbf{U}', f')$ respectively.

In particular the length of an admissible curve for two equivalent sub-Riemannian structures is equal.

Lemma 2.3.3. *Let* $\mathbf{V}$ *and* $\mathbf{U}$ *be vector bundles on the differentiable manifold* M*. If* $p \colon \mathbf{V} \to \mathbf{U}$ *is a morphism of vector bundles which is surjective on fibers, then each section* $\sigma \in \Gamma(\mathbf{U})$ *of* $\mathbf{U}$ *is the image though* p *of a section* $\mu \in \Gamma(\mathbf{V})$ *of* $\mathbf{V}$*.*

Proof. Firstly let us show the following: if $p \colon \mathbf{V} \to \mathbf{U}$ is a morphism of vector bundles over M which is surjective on fibers, then for each section σ of $\mathbf{U}$ and $q \in M$ there exists an open neighbourhood O_q of q and a section μ such that $p(\mu)\big|_{O_q} = \sigma\big|_{O_q}$.

Given $q \in M$, there is a proper open neighbourhood of q such that both $\mathbf{U}$ and $\mathbf{V}$ are trivializable on that open neighbourhood. If $\mathbf{V}$ has rank n and $\mathbf{U}$ has rank m, we can consider (restrictively to that neighbourhood) $V_1, \dots, V_n$ coordinate sections of $\mathbf{V}$ (which is a base of $\Gamma(V)$ as a $\mathcal{C}^\infty$-module) and $X_1, \dots, X_m$ coordinate sections of $\mathbf{U}$. Then

$$\begin{aligned} p(V_1) &= a_1^1 X_1 + \cdots + a_1^m X_m \\ &\ \vdots \\ p(V_n) &= a_n^1 X_1 + \cdots + a_n^m X_m \end{aligned}$$

where a_i^j are all $\mathcal{C}^\infty$ functions. Then $A = \big(a_i^j\big)_{ij}$ is a matrix of $\mathcal{C}^\infty$ function which has always maximum rank, since p is surjective on fibers. Thus we can select m vector fields in $\{ V_1, \dots, V_n \}$ such that $p(V_{i_1}(q)), \dots, p(V_{i_m}(q))$ generates the fiber of $\mathbf{U}$ at q entirely, that is the $m \times m$ matrix $A' = \big(a^j_{i_k}\big)_{kj}$ is invertible at q. Since A' is $\mathcal{C}^\infty$, then A' is invertible also in an open neighbourhood O_q of q. For each $j = 1, \dots, m$ we have

$$X_j = \sum_{k=1}^{m} (A^{-1})_{jk} p(V_{i_k})$$

and therefore each $\sigma \in \Gamma(\mathbf{U})$ restricted to O_q can be written as $X = b_1 X_1 + \cdots + b_m X_m$ where b_j are $\mathcal{C}^\infty$ functions and

$$X = p\Bigg(\sum_{j=1}^{m} \sum_{K=1}^{m} b_j (A^{-1})_{jk} V_{i_k} \Bigg) .$$

Then we can show that each section of $\mathbf{U}$ is globally the image of a section of $\mathbf{V}$. Indeed for each $q \in M$ we have an open neighbourhood O_q of q with the stated properties. Then we can consider a partition of the unity $\{ \rho_q \}_{q \in M}$ subjected to the open covering $\{ O_q \}_{q \in M}$ and we can write X, a section of $\mathbf{U}$, as

$$X = \sum_{q \in M} \rho_q X\big|_{O_q} . \qquad \square \qquad (2.27)$$

The terminology "equivalent as distributions" is justified by by the following proposition:

Proposition 2.3.4. *Two sub-Riemannian structures $(M, \mathbf{U}, f)$ and $(M, \mathbf{U}', f')$ over the same differentiable manifold M are equivalent as distributions if and only if the respective modules of vector fields $\mathcal{D}$ and $\mathcal{D}'$ coincides.*

Proof. ($\Rightarrow$). Let $X \in \mathcal{D}$, then $X = f(\sigma)$ for some σ section of $\mathbf{U}$ and by lemma 2.3.3 $\sigma = p(\mu)$ for some μ section of $\mathbf{V}$. Then

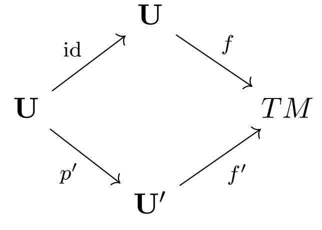
$$X = f \circ p(\mu) = f' \circ p'(\mu) = f'\big(p'(\mu)\big)$$

and therefore $X \in \mathcal{D}'$. The opposite inclusion is similar.

($\Leftarrow$). For each $q \in M$ we consider an open neighbourhood O_q of q that trivializes both $\mathbf{U}$ and $\mathbf{U}'$. Then we consider $U_1, \dots, U_m \in \Gamma(\mathbf{U})$ and $U'_1, \dots, U'_{m'} \in \Gamma(\mathbf{U}')$ coordinates vector fields which are both basis as $\mathcal{C}^\infty$-modules. Since $\mathcal{D} = \mathcal{D}'$, we get

$$f(U_j) = f'\big(a^1_j U'_1 + \cdots + a^{m'}_j U'_{m'}\big)$$

for some $\mathcal{C}^\infty$ functions a^k_i. So that we can define $p' \colon \mathbf{U} \to \mathbf{U}'$ a morphism of vector bundles simply imposing $p'(U_j) = a^1_j U'_1 + \cdots + a^{m'}_j U'_{m'}$ for each $j = 1, \dots, m$, thus $f = f' \circ p'$ on O_q. Considering a proper partition of the unity we can extend the equality to all M. Therefore the following diagram commutes:



that is $\mathbf{U}$ and $\mathbf{U}'$ are equivalent as distributions. □

Definition 2.3.5. Let M be a sub-Riemannian manifold. We define the *minimal bundle rank* of M as the minimum of rank of bundles that induces equivalent structures on M. Given $q \in M$ the *local minimum bundle rank* of M at q is the minimal bundle rank of the structure restricted on a sufficiently small neighbourhood O_q of q

Example 2.3.6. Let $\mathbf{U} = M \times \mathbb{R}^m$ be the trivial Euclidean bundle of rank m on M. An element of $\mathbf{U}$ is written as (q, u) where $q \in M$ and $u \in \mathbb{R}^m$. We consider $\{ e_1, \dots, e_m \}$ an orthonormal basis of $\mathbb{R}^m$. Then we can define m global vector fields on M by $f_i(q) = f(q, e_i)$ for $i = 1, \dots, m$. Then we obtain

$$f(q,u) = \sum_{i=1}^{m} u_i f_i(q) \qquad q \in M\,, \quad u = (u_1, \dots, u_m) \in \mathbb{R}^m\,. \tag{2.28}$$

A sub-Riemannian structure whose Euclidean bundle is globally trivial is a free sub-Riemannian structure. The set of the defined vector fields $f_1, \dots, f_m$ is called a *generating family.* They may not be orthonormal if f is not injective.

Remark 2.3.7. In the setting of the previous example, a curve $\gamma\colon [0,T] \to M$ define as $\gamma(t) = e^{tf_i}$, which is an integral curve for the horizontal vector field f_i, is admissible and $\ell(\gamma) \le T$ since e_i is a possible control for γ.

Lemma 2.3.8. *Let M be a d-dimensional differentiable manifold and $\pi\colon E \to M$ a vector bundle of rank m. Then there exists a vector bundle $\pi_0\colon E_0 \to M$ with* $\operatorname{rank} E_0 \le 2d + m$ *such that $E \oplus E_0$ is a trivial vector bundle.*

Proof. The vector bundle E, as a differentiable manifold, has dimension $d + m$. We consider the map $i\colon M \to E$ which embeds M into the vector bundle E as the zero section. We denote $T_M E = i^*(TE)$ the pulled back vector bundle, which is the restriction of TE to the section M_0. Let us notice that every fiber E_q, since it is a vector space, is canonically isomorphic to its tangent space $T_q E_q$ at zero. Then we can write

$$T_q E = T_q E_q \oplus T_q M \cong E_e \oplus T_q M \quad \forall q \in M_0 \cong M$$

and therefore

$$T_M E \cong E \oplus TM\,.$$

By Whitney's theorem we can embed k-dimensional differentiable manifolds in $\mathbb{R}^{2k}$. Thus, for $N = 2(d + m)$, we can consider an immersion

$$\Psi\colon E \to \mathbb{R}^N \qquad \Psi_*\colon T_M E \subseteq TE \to T\mathbb{R}^N\,,$$

where Ψ_* is an injective bundle map. Then we can interpret $T_M E$ as a sub-bundle of $T\mathbb{R}^N \cong \mathbb{R}^N \times \mathbb{R}^N$. We now consider the orthogonal bundle E' to $T_E M$ (on the base M) with respect to the Euclidean metric $\mathbb{R}^N$, that is

$$E' = \bigcup_{q \in M} E'_q \qquad E'_q = (T_q \oplus T_q M)^{\perp}\,.$$

Finally, we consider $E_0 = T_M E \oplus E' \cong E \oplus (TM \oplus E')$. It is trivial since its fibers are trivially identified with $\mathbb{R}^N$. The rank of $(TM \oplus E')$ is $d + 2(d+m) - (d+m) = 2d + m$. □

A reference for Whitney's theorem is [Lee13, chapter 6].

Theorem 2.3.9. *Every sub-Riemannian structure* $(\mathbf{U}, f)$ *on* M *is equivalent to a free sub-Riemannian structure.*

Proof. By lemma 2.3.8 there exists a vector bundle $\mathbf{U}'$ such that the direct sum $\overline{\mathbf{U}} = \mathbf{U} \oplus \mathbf{U}'$ is a trivial vector bundle. We fix a Riemannian metric g' on $\mathbf{U}'$. We then define a metric $\overline{g}$ on $\overline{\mathbf{U}}$ such that $\overline{g}(u+u', v+v') = g(u,v) + g'(u',v')$ on each fiber $\overline{U}_q = U_q \oplus U'_q$. In this way U_q and U'_q are orthogonal subspaces og $\overline{\mathbf{U}}_q$ with respect to $\overline{g}$.

We consider the projection $p_1 \colon \mathbf{U} \oplus \mathbf{U}' \to \mathbf{U}$ on the first factor and we define a sub-Riemannian structure $(\overline{\mathbf{U}}, \overline{f})$ on M by

$$\overline{f} \colon \overline{\mathbf{U}} \to TM \qquad \overline{f} = f \circ p_1 \,.$$

By construction, the following diagram is commutative:

$$\begin{array}{ccccc} & & \overline{\mathbf{U}} & & \\ & \nearrow^{\mathrm{id}} & & \searrow^{\overline{f}} & \\ \mathbf{U} \oplus \mathbf{U}' & & & & TM \\ & \searrow_{p_1} & & \nearrow_{f} & \\ & & \mathbf{U} & & \end{array} ,$$

which means $\mathbf{U}$ and $\overline{\mathbf{U}}$ are equivalent as distributions. Finally, since for every $\overline{u} = u + u'$ where $u \in U_q$ and $u' \in U'_q$ we have $|\overline{u}|^2 = |u'|^2$, hence $|u| = \min \{ |\overline{u}| \mid p_1(\overline{u}) = \overline{u} \}$, which proves that $\mathbf{U}$ and $\overline{U}$ are equivalent as sub-Riemannian structures. ☐

2.4 Rashevskii-Chow theorem

Definition 2.4.1. Let M be a sub-Riemannian structure and $q_0, q_1 \in M$. The *sub-Riemannian distance* (or *Carathéodory distance*) between q_0 and q_1 is

$$d(q_0, q_1) = \inf \{ \ell(\gamma) \mid \gamma \colon [0,T] \to M \text{ admissible}, \gamma(0) = q_0, \gamma(T) = q_1 \} \,. \tag{2.29}$$

In what follows $B(q,r)$ (or $B_r(q)$) is the (open) sub-Riemannian ball of radius r and center q, namely

$$B(q,r) = \left\{ q' \in M \;\middle|\; d(q,q') < r \right\} \,. \tag{2.30}$$

The entire section is devoted to the following result:

Theorem 2.4.2 (Rashevskii-Chow)**.** *Let* M *be a* sub-Riemannian manifold*. Then*

(i) (M, d) *is a metric space,*

(ii) the topology induced by (M, d) *is equivalent to the manifold topology.*

In particular $d \colon M \times M \to \mathbb{R}$ *is continuous.*

The explicit statements enclosed in theorem 2.4.2 are:

(a) $0 \leq d(q_0, q_1) \leq +\infty$ for all $q_0, q_1 \in M$,

(b) $d(q_0, q_1) = 0$ if and only if $q_0 = q_1$,

(c) $d(q_0, q_1) = d(q_1, q_0)$, $d(q_0, q_2) \leq d(q_0, q_1) + d(q_1, q_2)$ for all $q_0, q_1, q_2 \in M$,

(d) for every $\varepsilon > 0$ there is O_{q_0} neighbourhood of q_0 such that $O_{q_0} \subseteq B(q_0, \varepsilon)$,

(e) for every neighbourhood O_{q_0} of q_0 there is $\delta > 0$ such that $B(q_0, \delta) \subseteq O_{q_0}$.

Proof of (c). If $\gamma\colon [0, T] \to M$ is admissible, then also $\overline{\gamma}\colon [0, T] \to M$ defined by $\overline{\gamma}(t) = \gamma(T - t)$ is admissible and $\ell(\overline{\gamma}) = \ell(\gamma)$. This proves that d is symmetric. For the triangular inequality we consider two admissible curves $\gamma_1\colon [0, T_1] \to M$ and $\gamma_2\colon [0, T_2] \to M$ such that $\gamma_1(T_1) = \gamma(0)$, then we can define their concatenation:

$$\gamma\colon [0, T_1 + T_2] \to M \qquad \gamma(t) = \begin{cases} \gamma_1(t) & t \in [0, T] \\ \gamma_2(t - T_1) & t \in [T_1, T_1, +T_2] \end{cases},$$

which is again admissible and $\ell(\overline{\gamma}) = \ell(\gamma_1) + \ell(\gamma_2)$. This proves the triangular inequality. □

Lemma 2.4.3. *Let $N \subseteq M$ be a submanifold and $\mathcal{F} \subseteq \mathrm{Vec}(M)$ be a family of vector fields tangent to N, that is for every $X \in \mathcal{F}$ and $q \in N$ we have $X(q) \in T_qN$. Then for all $q \in M$ we have $\mathrm{Lie}_q\mathcal{F} \subseteq T_qN$, in particular $\dim \mathrm{Lie}_q\mathcal{F} \leq \dim N$.*

Proof. Let $X \in \mathcal{F}$. We then consider the two Cauchy problems

$$\begin{cases} \dot{q} = X(q) & q \in M \\ q(0) = q_0 & q_o \in N \end{cases} \qquad \text{and} \qquad \begin{cases} \dot{q} = X\big|_N(q) & q \in N \\ q(0) = q_0 \,, \end{cases}$$

from local existence and uniqueness of their solutions (theorem 1.3.4), it follows that $e^{tX}(q) \in N$ for every $q \in N$ and t close enough to zero. If we consider eq. (1.46) for the definition of Lie bracket of vector fields, we get that, if X, Y are tangent to N, then also $[X, Y]$ is again tangent to N. This implies that $\mathrm{Lie}_q\mathcal{F} \subseteq T_qN$ for all $q \in M$. □

Let us recall that, thanks to theorem 2.3.9, we can always assume a sub-Riemannian structure to be free. Therefore we can state the following lemma in this way:

Lemma 2.4.4. *Let M be a d-dimensional sub-Riemannian manifold with a (global) generating family $\mathcal{F} = \{ f_1, \ldots, f_m \}$. For every $q_0 \in M$ and every neighbourhood V of the origin in $\mathbb{R}^d$ the exists $\widehat{s} = (\widehat{s_1}, \ldots, \widehat{s_d}) \in V$, and a choice of (possibly repeating) vector fields $f_{i_1}, \ldots, f_{i_d} \in \mathcal{F}$, such that $\widehat{s}$ is a regular point of the map*

$$\psi\colon \mathbb{R}^d \to M \qquad \psi(s_1, \ldots, s_d) = e^{s_d f_{i_d}} \circ \cdots \circ e^{s_1 f_{i_1}}(q_0)\,. \tag{2.31}$$

Proof. Fixing $q_0 \in M$, there exists a vector field $f_{i_1} \in \mathcal{F}$ such that $f_{i_1}(q_0) \neq 0$, otherwise $\dim \mathrm{Lie}_{q_0}\mathcal{F} = 0$ and that would contradict the bracket-generating condition in the definition of a sub-Riemannian structure. Then, for $|s_1|$ small enough, the map

$$\phi_1 \colon s_1 \mapsto e^{s_1 f_{i_1}}(q_0)$$

is a local diffeomorphism onto its image Σ_1. If $\dim M = 1$ the lemma is proved. Otherwise we iterate the following argument until we reach the dimension d of the manifold M.

After n iterations we obtained a point $t^{(n)} = (t_1, \dots, t_{n-1}, 0) \in \mathbb{R}^n$ arbitrarily close to zero and $f_{i_1}, \dots, f_{i_n} \in \mathcal{F}$ such that

$$\phi_n \colon (s_1, \dots, s_n) \mapsto e^{s_n f_{i_n}} \circ \dots \circ e^{s_1 f_{i_1}}(q_0)$$

is a local diffeomorphism at $t^{(n)}$ onto its image Σ_n which has dimension n near $q_n = \phi_k(t^{(n)})$.

We suppose that $n < \dim M$, then there exists $q_{n+1} \in \Sigma_n$ arbitrarily close to q_n and $f_{i_{n+1}} \in \mathcal{F}$ such that $f_{i_{n+1}}$ is not tangent to Σ_n at q_{n+1}. Otherwise there would be a neighbourhood of q_n in Σ_n such that all vector fields in $\mathcal{F}$ are tangent to Σ_n in that neighbourhood. Thus lemma 2.4.3 would then imply that $\mathrm{Lie}_q\mathcal{F} \subseteq \Sigma_n$ for every q in that neighbourhood of Σ_n, but we assumed $\dim \Sigma_n = n < d$ and that would contradict the bracket-generating condition.

Then we define

$$\phi_{n+1} \colon (s_1, \dots, s_n, s_{n+1}) \mapsto e^{s_{n+1} f_{i_{n+1}}} \circ \dots \circ e^{s_1 f_{i_1}} .$$

If we have $q_{n+1} = \phi_n(\tau_1, \dots, \tau_n)$, then we observe that ϕ_n is a local diffeomorphism at $t^{(n+1)} = (\tau_1, \dots, \tau_n, 0) \in \mathbb{R}^{n+1}$ onto its image Σ_{n+1}. Indeed

$$\left.\frac{\partial \phi_{n+1}}{\partial s_k}\right|_{t^{(n+1)}} \qquad k = 1, \dots, k = n$$

are n linearly independent vectors in $T_{q_{n+1}}\Sigma_n$ and on the other hand

$$\left.\frac{\partial \phi_{n+1}}{\partial s_{n+1}}\right|_{t^{(n+1)}} = f_{i_{n+1}}(q_{n+1})$$

is, by construction, independent of $T_{q_{n+1}}\Sigma_n$. After d iterations we obtain the thesis. □

Proof of (d). Let V be a neighbourhood of $0 \in \mathbb{R}^d$, by lemma 2.4.4 there exists $\widehat{s} \in V$ and $\widehat{V} \subset V$ a neighbourhood of $\widehat{s}$ such that ψ (as defined in lemma 2.4.4) is a diffeomorphism from $\widehat{V}$ to $\psi(\widehat{V})$ which is a neighbourhood of $\psi(\widehat{s})$. We then consider a map

$$\widehat{\psi} \colon \mathbb{R}^d \to M \qquad \widehat{\psi}(s_1, \dots, s_d) = e^{-\widehat{s}_1 f_{i_1}} \circ \dots \circ e^{-\widehat{s}_n f_{i_n}} \circ \psi(s_1, \dots, s_n) ,$$

where $\widehat{s} = (\widehat{s}_1, \dots, \widehat{s}_d)$. The map $\widehat{\psi}$ is again a diffeomorphism from a proper neighbourhood of $\widehat{s}$, that we may suppose to be again $\widehat{V}$, and a neighbourhood of $\widehat{\psi}(\widehat{s}) = q_0$.

Let us now fix $\varepsilon > 0$, consider $V = \{ s \in \mathbb{R}^d \mid |s_i| < \varepsilon \text{ for } i = 1, \dots, d \}$ and apply the previous construction. We now show that in this configuration $O_{q_0} = \widehat{\psi}(\widehat{s})$ is contained in "a small sub-Riemannian ball" whose radius is controlled by ε.

Indeed, let $q \in \widehat{\psi}(\widehat{s})$. We set $q = \widehat{\psi}(s_1, \dots, s_d)$, then an admissible curve joining q_0 and q is, by definition of $\widehat{\psi}$, the concatenation of the admissible curves

$$\gamma_k \colon [0, s_k] \to M \qquad t \mapsto e^{tf_{i_k}} \qquad \text{for } k = 1, \dots, d$$

and

$$\widehat{\gamma}_k \colon [0, \widehat{s}_k] \to M \qquad t \mapsto e^{-tf_{i_k}} \qquad \text{for } k = d, \dots, 1\,.$$

Let γ be such a concatenation. Considering remark 2.3.7 and the fact that $s, \widehat{s} \in V$, we obtain

$$d(q, q_0) \le \ell(\gamma) \le |s_1| + \dots + |s_d| + |\widehat{s}_1| + \dots + |\widehat{s}_d| \le 2d\varepsilon\,,$$

which completes the proof. □

Proof of (a). We need to show that d is always finite. We consider in M the binary relation

$$q_1 \sim q_2 \quad \Longleftrightarrow \quad d(q_1, q_2) < +\infty\,.$$

It is immediate, using the triangular inequality, that the defined relation is an equivalence relation. From (d) follows that the equivalence classes are open, since each point $q \in M$ has a neighbourhood with finite sub-Riemannian distance. Then each equivalence class represents a disconnection. Since M is connected, there can be only one equivalence class. This means that d is finite. □

Remark 2.4.5. We can already prove the continuity of d. Indeed, let $(q_1, q_2) \in M \times M$ and $\varepsilon > 0$, then we consider a neighbourhood O_{q_1} of q_1 and a neighbourhood O_{q_2} of q_2 such that $d(q_1, O_{q_1}) < \varepsilon/2$ and $d(q_2, O_{q_2}) < \varepsilon/2$. For $(q_1', q_2') \in O_{q_1} \times O_{q_2}$ we have, using the triangular inequality

$$|d(q_1', q_2') - d(q_1, q_2)| \le d(q_1, q_1') + d(q_2, q_2') \le \varepsilon/2 + \varepsilon/2 = \varepsilon\,.$$

Lemma 2.4.6. *Let $q_0 \in M$ and $K \subseteq M$ a compact set such that $q_0 \in \operatorname{int} K$. Then there exists $\delta_K > 0$ such that every admissible curve γ starting from q_0 and with $\ell(\gamma) \le \delta_K$ is contained in K.*

Proof. Since the statement is more restrictive as the compact K becomes smaller and smaller, we can assume K to be small enough so that it is contained in a unique local chart of M. We denote by $|\cdot|$ the Euclidean norm in the coordinate chart. Let us define

$$C_K = \max_{x \in K} \left(\sum_{i=1}^{m} |f_i(x)|^2 \right)^{1/2},$$

and fix $\delta_K > 0$ such that $\operatorname{dist}(q_0, \partial K) > C_K \delta_K$, where dist denotes the Euclidean distance from a point to a set, in coordinates.

We would like to show that for any admissible curve $\gamma\colon [0,T] \to M$ such that $\gamma(0) = q_0$ and $\ell(\gamma) \le \delta_K$ we have $\gamma([0,T]) \subseteq K$. Let $\gamma\colon [0,T] \to M$ be an admissible curve with $\ell(\gamma) \le \delta_K$ and let define

$$t^* = \sup \{ \, t \in [0,T] \mid \gamma([0,t]) \subseteq K \, \} \, .$$

Then

$$\begin{aligned}
|\gamma(t^*) - \gamma(0)| &\le \int_0^{t^*} |\dot\gamma(t)| \,\mathrm{d}t \le \int_0^{t^*} \sum_{i=1}^m |u_i^*(t) f_i(\gamma(t))| \,\mathrm{d}t \\
&\le \int_0^{t^*} \left(\sum_{i=1}^m |f_i(\gamma(t))|^2 \right)^{1/2} \left(\sum_{i=1}^m u_i^*(t)^2 \right)^{1/2} \mathrm{d}t \\
&\le C_K \int_0^{t^*} \left(\sum_{i=1}^m u_i^*(t)^2 \right)^{1/2} \mathrm{d}t \le C_K \ell(\gamma) \\
&\le C_K \delta_K < \operatorname{dist}(q_0, \partial K) \, .
\end{aligned}$$

Then γ cannot leave the compact K in $[0,T]$ and therefore $t^* = T$ which implies the lemma. □

Proof of (b). The fact that $d(q,q) = 0$ for all $q \in M$ is trivial. On the other hand, let us consider points $q_0 \neq q_1$ in M, we can now consider a compact K which is a neighbourhood of q_0 and that it does not contain q_1, (we can do it because we assume M to be Hausdorff). By lemma 2.4.6, each admissible curve joining q_0 and q_1 has length at least δ_K, hence $d(q_0, q_1) \ge \delta_K > 0$. □

Proof of (e). Let us fix $\varepsilon > 0$ and a compact neighbourhood K of q_0 which is contained in a unique local chart. We then define C_K and δ_K as in the proof of lemma 2.4.6. We set $\delta = \min \{ \, \delta_K, \varepsilon / C_K \, \}$ and we want to show that $|q - q_0| < \varepsilon$ whenever $d(q, q_0) < \varepsilon$, where again $|\cdot|$ is the Euclidean norm in the local coordinates so that balls for this distance defines a base for the neighbourhoods of q.

Let $\gamma_n\colon [0,T] \to M$ be a minimizing sequence of admissible curves joining q_0 and q such that $\ell(\gamma_n) \to d(q_0, d)$ as $n \to +\infty$. Eventually $\ell(\gamma_n) \le \delta$ starting from a certain point in the sequence, by lemma 2.4.6 we then get $\gamma_n([0,T]) \subseteq K$ eventually.

Finally, we can repeat the same estimates as in the proof of lemma 2.4.6 and we obtain that $|q - q_0| = |\gamma_n(T) - \gamma_n(0)| \le C_K \ell(\gamma_n)$ eventually from a certain n. Passing to the limit for $n \to \infty$, we get

$$|q - q_0| \le C_K \, d(q_0, q) \le C_K \delta < \varepsilon \, . \qquad \square$$

We proved every statement of theorem 2.4.2, the proof is then concluded.

Corollary 2.4.7. *The metric space (M, d) is locally compact, that is, for any $q \in M$ there exists $\varepsilon > 0$ such that the closed sub-Riemannian ball $\overline{B}(q, r)$ is compact for all $0 \le r \le \varepsilon$.*

Proof. By the continuity of d, the set $\overline{B}(q,r) = \{ d(q,\cdot) \leq r \}$ is closed for all $q \in M$ and $r \geq 0$. Let a consider a compact neighbourhood K of q, since the sub-Riemannian distance d induces the manifold topology, then for sufficiently small radius R we have $\overline{B}(q,r) \subseteq K$. Then these ball are closed inside a compact, thus they are compact. □

Remark 2.4.8. In the proof of theorem 2.4.2 we rely on the bracket-generating condition, however eq. (2.29) may define a distance even if the structure is not bracket-generating. In these cases the resulting metric may not induce a topology equivalent to the manifold topology.

Chapter 3

Length-minimizers and their characterization

In this chapter we discuss the existence of length-minimizers and will provide a first order necessary condition. The study of length-minimizer has a vast interest in sub-Riemannian geometry and many questions on the topic has remained open still today, such as the regularity conjecture. We will introduce normal and abnormal extremals and we will provide two equivalent characterizations of abnormal extremals that will be useful in the last chapter. In the presentation of the topic we will refer to [ABB20, chapter 3 and 8].

3.1 Existence of length-minimizers

Definition 3.1.1. Let $\gamma\colon [0,T] \to M$ be an admissible curve for the sub-Riemannian manifold $(M, \mathbf{U}, f)$, we say that γ is e *length-minimizer* if $\ell(\gamma) = d(\gamma(0), \gamma(T))$. This means that γ has the least possible sub-Riemannian length among all admissible curves connecting $\gamma(0)$ with $\gamma(T)$.

Remark 3.1.2. Let us notice that the existence of length-minimizer between two points is not always guaranteed. For instance, let $M = \mathbb{R}^2 \setminus \{ 0 \}$ with the classical Riemannian structures. There are no length-minimizers from $x \in M$ to $-x \in M$. On the other hand, length-minimizers may not be unique, take for instance two antipodal point on the sphere $\mathbb{S}^2$.

Theorem 3.1.3. *Let $\gamma_n\colon [0,T] \to M$ be a sequence of admissible curves parametrized with constant speed such that $\gamma_n \to \gamma$ uniformly on $[0,T]$ and $\liminf_{n\to\infty} \ell(\gamma_n) < +\infty$. Then γ is admissible and*

$$\ell(\gamma) \leq \liminf_{n\to\infty} \ell(\gamma_n)\,. \tag{3.1}$$

Proof. Let $L = \liminf_{n\to\infty} \ell(\gamma_n) < +\infty$ and choose a subsequence, denoted again with $(\gamma_n)_n$, such that $\ell(\gamma_n) \to L$.

We fix $\delta > 0$, we may suppose that $\ell(\gamma_n) \leq L + \delta$ for every n. Moreover it is not restrictive to assume, by uniform convergence, that the image of all γ_n is contained in a common compact K. Indeed we know that for every $\varepsilon > 0$ the image of γ_n is eventually inside an ε-neighbourhood of $\gamma([0,T])$. We only need to prove that these sets a compact for ε small enough. We thus consider, using corollary 2.4.7, the open covering

$$\mathcal{B} = \left\{ B(q,r) \;\middle|\; q \in \gamma([0,T]),\, \overline{B}(q,r) \text{ is compact} \right\} .$$

Since $\gamma([0,T])$ is compact, it is covered by a finite subset $\mathcal{B}'$ of $\mathcal{B}$, let

$$K = \left\{ \overline{B}(q,r) \;\middle|\; B(q,r) \in \mathcal{B}' \right\}$$

which is compact since it is a finite union of compact sets. Then the interior of K contains $\gamma([0,T])$ and therefore

$$d(\partial K, \gamma([0,T])) = \inf \{ d(x,y) \mid x \in \partial K,\, y \in \gamma([0,T]) \} > 0\,.$$

We obtained that K is a compact ε-neighbourhood of $\gamma([0,T])$ where $\varepsilon = d(\partial K, \gamma([0,T])) > 0$.

Up to a common time rescaling, we can assume that the curves are parametrized with constant speed on the interval $[0,1]$. Under all these assumptions, we get that $\dot\gamma_n(t) \in V_{\gamma_n(t)}$ for almost every t, where

$$V_q = \{ f_u(q) \mid |u| \le L + \delta \} \subseteq T_q M\,.$$

Let us notice that, since f is linear in u, that V_q is convex foe every $q \in M$. We need to prove that γ is admissible and satisfies $\ell(\gamma) \le L + \delta$. Since δ is arbitrary, this will imply $\ell(\gamma) \le L$.

Let $\varepsilon > 0$ and $t \in [0,1]$. For ε small enough we may suppose that $\gamma_n(t+\varepsilon)$ and $\gamma_n(t)$ are in the same local chart, thus

$$\frac{1}{\varepsilon}\big(\gamma_n(t+\varepsilon) - \gamma_n(t)\big) = \frac{1}{\varepsilon}\int_t^{t+\varepsilon} f_{u_n(\tau)}(\gamma_n(\tau))\,\mathrm{d}\tau \in \operatorname{conv}\left\{ V_{\gamma_n(\tau)} \;\middle|\; \tau \in [t, t+\varepsilon] \right\}, \quad (3.2)$$

where $\operatorname{conv} S$ denotes the convex hull of a set S. For n large enough, uniform convergence grants $|\gamma_n(t) - \gamma(t)| < \varepsilon$ and an estimates similar to the one in the proof of lemma 2.4.6 leads to

$$|\gamma_n(t) - \gamma_n(\tau)| \le \int_t^\tau |\dot\gamma_n(s)|\mathrm{d}s \le C_K(L+\delta)|t-\tau| \le C_K(L+\delta)\varepsilon \quad (3.3)$$

for $\tau \in [t, t+\varepsilon]$. Hence, for $\tau \in [t, t+\varepsilon]$ and n large enough

$$|\gamma_n(\tau) - \gamma(t)| \le |\gamma_n(t) - \gamma_n(\tau)| + |\gamma_n(t) - \gamma(t)| \le C'\varepsilon$$

where C' is independent of n and ε. Since the manifold topology is equivalent to the metric topology, for all $\tau \in [t, t+\varepsilon]$ and n large enough, $\gamma_n(\tau) \in B_{\gamma(t)}(r_\varepsilon)$ where $r_\varepsilon \to 0$ as $\varepsilon \to 0$. Therefore

$$\operatorname{conv}\left\{ V_{\gamma_n(\tau)} \;\middle|\; \tau \in [t, t+\varepsilon] \right\} \subseteq \operatorname{conv}\left\{ V_q \;\middle|\; q \in B_{\gamma_n(t)}(r_\varepsilon) \right\} .$$

Combining this inclusion with eq. (3.2) and passing to the limit for $n \to \infty$ we get

$$\frac{1}{\varepsilon}\big(\gamma(t+\varepsilon)-\gamma(t)\big) \in \operatorname{conv}\left\{ V_q \;\middle|\; q \in B_{\gamma_n(t)}(r_\varepsilon) \right\} . \tag{3.4}$$

Passing to the limit for $n \to \infty$ in eq. (3.3) we obtain that γ is Lipschitz. Hence the left hand side in eq. (3.4) exists for almost $t \in [0,1]$ and for $\varepsilon \to 0$ we obtain $\dot\gamma_n(t) \in \operatorname{conv} V_{\gamma(t)} = V_{\gamma(t)}$, therefore γ is admissible. By construction, the minimal control $u^*(t)$ associated with γ is essentially bounded and $|u^*(t) \leq| L+\delta$ for almost every $t \in [0,1]$, this implies $\ell(\gamma) \leq L+\delta$ since γ is defined on $[0,1]$ and we concluded. □

Corollary 3.1.4. *Let $\gamma_n\colon [0,T] \to M$ be a sequence of length-minimizers parametrized with constant speed and such that $\gamma_n \to \gamma$ uniformly on $[0,T]$. Then γ is a length-minimizer.*

Proof. Since γ_n is a length-minimizer, we have $\ell(\gamma_n) = d(\gamma_n(0),\gamma_n(T))$. By uniform convergence $\gamma_n(t) \to \gamma(t)$ for every $t \in [0,T]$ and, recalling the continuity of the distance and the semicontinuity of the length

$$\ell(\gamma) \leq \liminf_{n\to\infty} \ell(\gamma_n) = \liminf_{n\to\infty} d(\gamma_n(0),\gamma_n(T)) = d(\gamma(0),\gamma(T)) ,$$

that implies $\ell(\gamma) = d(\gamma(0),\gamma(T))$ and thus γ is a length-minimizer. □

Theorem 3.1.5. *Let M be a sub-Riemannian manifold and $q_0 \in M$. Assume that $\overline{B}_{q_0}(r)$ is compact for some $r > 0$. Then for all $q_1 \in B_{q_0}(r)$ there exists a length minimizer joining q_0 and q_1.*

Proof. Fix $q_1 \in B_{q_0}(r)$ and consider a minimizing sequence $\gamma_n\colon [0,1] \to M$ of admissible curves, parametrized with constant speed, joining q_0 and q_1 and such that $\ell(\gamma_n) \to d(q_0,q_1)$.

Since $d(q_0,q_1) < r$, we have $\ell(\gamma_n) \leq r$ eventually for n large enough. Thus we can assume that the images of γ_n are all contained in the compact $K = \overline{B}_{q_0}(r)$. We consider a finite partition of $[0,1]$ such that each partition is contained in a local chart. Using these local chart, the same argument leading to eq. (3.3) gives

$$|\gamma_n(t)-\gamma_n(\tau)| \leq \int_t^\tau |\dot\gamma_n(s)| \mathrm{d}s \leq C_K r |t-\tau| .$$

for all t,τ in the same partition. In each partition the sequence is equicontinuous and uniformly bounded. By Ascoli-Arzelà theorem there exists a subsequence γ_{n_k} and a Lipschitz curve $\gamma\colon [0,T] \to M$ such that $\gamma_{n_k} \to \gamma$ uniformly in each partition (since there is a finite number of them). By theorem 3.1.3, the curve γ satisfies $\ell(\gamma) \leq \liminf_{k\to\infty} \ell(\gamma_{n_k}) = d(q_0,q_1)$, that is $\ell(\gamma) = d(q_0,q_1)$ and γ is a length-minimizer. □

Combining corollary 2.4.7 and theorem 3.1.5 we obtain the following:

Corollary 3.1.6. *Let $q_0 \in M$. There exists $\varepsilon > 0$ such that for every $q_1 \in B_{q_0}(\varepsilon)$ there is a length-minimizer joining q_0 and q_1.*

Lemma 3.1.7. *Let M be a sub-Riemannian manifold. For every $\varepsilon > 0$ and $x \in M$*

$$B(x, r+\varepsilon) = \bigcup_{y \in B(x,r)} B(y, \varepsilon) \, . \tag{3.5}$$

Proof. Let us prove $\subseteq$. Fix $z \in B(x, r+\varepsilon)$, if $z \in B(x, \varepsilon)$ we have concluded, otherwise $z \in B(x, r+\varepsilon) \setminus B(x, \varepsilon)$. Then we consider a length-parametrized curve joining x and z such that $\ell(\gamma) = t + \varepsilon$ where $o \leq t < r$. Let $t' \in (t, r)$, thus $\gamma(t') \in B(x, r)$ and $z \in B(\gamma(t'), \varepsilon)$.

The inclusion $\supseteq$ is an immediate consequence of the triangular inequality. □

Proposition 3.1.8. *Let M be a sub-Riemannian manifold. The following are equivalent:*

(i) *(M, d) is complete,*

(ii) *$\overline{B}(x, r)$ is compact for every $x \in M$ and $r > 0$,*

(iii) *there exists $\varepsilon > 0$ such that $\overline{B}(x, \varepsilon)$ is compact for every $x \in M$.*

Proof. (iii)⇒(i). We consider a Cauchy sequence $\{ x_n \}_{n \in \mathbb{N}}$ in M. We fix $\varepsilon >$ satisfying the assumption, then there exists $N \in \mathbb{N}$ such that $d(x_n, x_m) < \varepsilon$ for all $n, m \geq 0$. In particular, for $m = N$, we obtain that for all $n \geq N$ we have $x_n \in \overline{B}(x_N, \varepsilon)$ which is compact. Therefore $\{ x_n \}_{n \in \mathbb{N}}$ is Cauchy sequence that admits a convergent subsequence, hence it is convergent.

(ii)⇒(iii). This in immediate.

(i)⇒(ii). Let us fix $x \in M$, the we define

$$A = \left\{ r > 0 \;\middle|\; \overline{B}(x, r) \text{ is compact} \right\} \qquad R = \sup A \, .$$

Since (M, d) is locally compact, then A is non-empty and $R > 0$.

Firstly we prove that A is open, it is enough to show that if $r \in A$ then also $r + \delta \in A$ for some δ. Indeed A is convex since closed ball inside a compact are compact. For this purpose, for each $y \in B(x, r)$ there is $\rho(y) < \varepsilon$ small enough such that $\overline{B}(y, \rho(y))$ is compact. We obtain

$$\overline{B}(x, r) \subseteq \bigcup_{y \in \overline{B}(x,r)} B(y, \rho(y)) \, .$$

By compactness of $\overline{B}(x, r)$ we can select a finite number of points $\{ y_i \}_{i=1}^{N}$ and the correspondent $\rho_i = \rho(y_i)$ such that

$$\overline{B}(x, r) \subseteq \bigcup_{i=1}^{N} B(y_i, \rho_i) \; = \mathcal{B}.$$

The set $\overline{B}(x, r)$ has a positive distance δ from $\partial \mathcal{B}$ since the distance is pointwise positive and $\overline{B}(x, r)$ is compact. By lemma 3.1.7

$$\mathcal{B} \supseteq \bigcup_{y \in \overline{B}(x,r)} B(y, \delta) \supseteq \bigcup_{y \in B(x,r)} B(y, \delta) = B(x, r+\delta) \, ,$$

and finally

$$\overline{B}(x, r+\delta) \subseteq \overline{\mathcal{B}} = \bigcup_{i=1}^{N} \overline{B}(y_i, \rho_i)\,.$$

This proves that $r + \delta \in A$, since a finite union of compact sets is compact.

Secondly we prove that $R = +\infty$, assume by contradiction that $R < +\infty$ and let us prove that $B = \overline{B}(x, R)$ is compact, this will lead to a contradiction since A is open. Is sufficient to prove that for every $\varepsilon > 0$ there exists an ε-net for B, that is a finite set S such that $d(y, S) < \varepsilon$ for every $y \in B$. With a standard diagonal argument this implies that each subsequence admits a Cauchy subsequence, which is convergent by completeness. Indeed we consider $\{ x_n \}_n$ a sequence in B, then for each $n \in \mathbb{N}$ there is a finite set S_n such that the open balls of radius 2^{-n} centred at the points of S_n cover B. D_0 is finite, therefore there exists $y_0 \in S_0$ such that infinitely many points of the sequence are in $B(y_0, 2^0)$. We define

$$A_0 = \left\{ n \in \mathbb{N} \;\middle|\; x_n \in B(y_0, 2^0) \right\} .$$

and recursively, assuming A_k is infinite, there exist $y_{k+1} \in S_{k+1}$ such that infinitely many point of $\{ x_n \}_{n \in A_k}$ are in $B(y_{k+1}, 2^{-k-1})$. Then we define

$$A_{k+1} = \left\{ n \in A_k \;\middle|\; x_n \in B(y_{k+1}, 2^{-k-1}) \right\} .$$

We then choose an increasing sequence $\{ n_k \}_k$ such that $n_k \in A_k$, thus $\{ x_{n_k} \}_k$ is a Cauchy sequence., which is convergent by completeness. For this purpose, let $\varepsilon > 0$ and consider an $(\varepsilon/3)$-net S for the ball $B' = B(x, R - \varepsilon/3)$, that exists by compactness. By lemma 3.1.7 we have for every $y \in B$ that $d(y, B') \leq \varepsilon/3$. Then if follows that

$$d(y, S) \leq d(y, B') + \varepsilon/3 < \varepsilon$$

that is S is an ε-net for B. □

Corollary 3.1.9. *Let (M, d) be a complete sub-Riemannian manifold. Then for every $q_0, q_1 \in M$ there exists a length-minimizer joining q_0 and q_1.*

Proof. Simply combine theorem 3.1.5 with proposition 3.1.8. □

3.2 Pontryagin extremals

In optimal control theory, a main problem is to provide necessary conditions for a given optimal control problem. Pontryagin's maximum principle states first-order necessary conditions for a wide variety of optimal control problems, see [BP07, chapter 6] for details. In this section we reformulate the length-minimality problem of sub-Riemannian geometry as an optimal control problem and then we will state the respective instance of the Pontryagin's maximum principle.

Theorem 3.2.1. *Let $\gamma\colon [0,T] \to M$ be an admissible curve which is a length-minimizer parametrized by constant speed. Let $\overline{u}(\cdot)$ be the corresponding minimal control. We denote by $P_{0,t}$ the flow*[1] *of the nonautonomous vector field*

$$f_{\overline{u}(t)} = \sum_{i=1}^{m} \overline{u}_i(t) f_i \,.$$

Then there exists $\lambda_0 \in T^*_{\gamma(0)}M$ such that defining

$$\lambda(t) = \left(P_{0,t}^{-1}\right)^* \lambda_0 \,, \qquad \lambda(t) \in T^*_{\gamma(t)}M \,, \tag{3.6}$$

we have that at least one of the following conditions is satisfied:

(N) $\overline{u}_i(t) = \langle \lambda(t), f_i(\gamma(t)) \rangle$ *for all* $i = 1, \dots, m$ *and* $t \in [0,T]$,

(A) $\lambda_0 \neq 0$ *and* $0 = \langle \lambda(t), f_i(\gamma(t)) \rangle$ *for all* $i = 1, \dots, m$ *and* $t \in [0,T]$.

Definition 3.2.2. Let $\gamma\colon [0,T] \to M$ be an admissible curve with minimal control $\overline{u} \in L^\infty([0,T], \mathbb{R}^m)$. We define $\gamma(t)$ as in eq. (3.6) for a fixed $\lambda_0 \in T^*_{\gamma(0)}M$, then

(i) if $\lambda(t)$ satisfies (N) then it is called a *normal extremal* and $\gamma(t)$ is a *normal extremal trajectory*,

(ii) if $\lambda(t)$ satisfies (A) then it is called an *abnormal extremal* and $\gamma(t)$ is an *abnormal extremal trajectory*.

Remark 3.2.3. If $\gamma(t)$ is a normal extremal trajectory, then condition (N) is theorem 3.2.1 implies that $\gamma(t)$ is smooth. Indeed, by construction, $\lambda(t)$ is Lipschitz continuous (both in the normal and abnormal case) and therefore $\overline{u}_i(t) = \langle \lambda(t), f_i(\gamma(t)) \rangle$ is continuous for all i, thus $\gamma(t)$ is $\mathcal{C}^1$ since its control is continuous. Inductively, if $\gamma(t)$ is $\mathcal{C}^k$, then both $\lambda(t)$ and $f_i(\gamma(t))$ are $\mathcal{C}^k$ by construction and therefore $\overline{u}_i(t)$ is $\mathcal{C}^k$. This implies that $\gamma(t)$ is $\mathcal{C}^{k+1}$.

Remark 3.2.4. For a given lift $\lambda(t)$, conditions (N) and (A) are mutually exclusive, unless the minimal control $\overline{u}(t) = 0$ for almost every $t \in [0,T]$ and the sub-Riemannian structure is not Riemannian at q_0, namely

$$\mathcal{D}_{q_0} = \operatorname{span}_{q_0} \{\, f_1, \dots, f_m \,\} \neq T_{q_0}M \,.$$

Indeed the trivial trajectory, corresponding to $\overline{u}(t) = 0$, is always normal with associated $\lambda_0 = 0$. On the other hand it is also abnormal with associated $\lambda_0 \in \mathcal{D}_{q_0}^{\perp}$. In general there are no abnormal extremals in the Riemannian case. Indeed λ_0 should be orthogonal to $\mathcal{D}_{q_0} = \operatorname{span}_{q_0} \{\, f_1, \dots, f_m \,\} = T_{q_0}M$ and therefore $\lambda_0 = 0$, which is a contradiction.

However, even a non-trivial admissible trajectory γ can be both normal and abnormal, since some lifts may satisfy (N) while others satisfy (A).

[1] that, for theorem 2.1.8, is defined for $t \in [0,T]$ in a neighbourhood of $\gamma(0)$.

Definition 3.2.5. Let $\gamma\colon [0,T] \to M$ be an admissible curve. We define the *energy functional* J of γ as

$$J(\gamma) = \frac{1}{2}\int_0^T \|\dot\gamma(t)\|^2 \, \mathrm{d}t\,, \tag{3.7}$$

and we notice that $J(\gamma)$ is finite since γ is admissible.

Remark 3.2.6. Let us notice that, as opposed to ℓ, the functional J is not invariant by reparametrization. Indeed if we consider scalar re-parametrized curve

$$\gamma_\alpha\colon [0,T/\alpha] \to M \qquad \gamma_\alpha(t) = \gamma(\alpha t)\,, \tag{3.8}$$

we get $J(\gamma_\alpha) = \alpha J(\gamma)$. Therefore, if we don't fix the final time, the infimum of J among all admissible curves joining two fixed point is always zero.

Lemma 3.2.7. *Fix $T > 0$ and let Ω_{q_0,q_1} be the set of all admissible curves from $[0,T]$ to a sub-Riemannian manifold M. An admissible curve $\gamma\colon [0,T] \to M$ is a minimizer of J on Ω_{q_0,q_1} if and only if it is a minimizer of ℓ on Ω_{q_0,q_1} and has constant speed.*

Proof. We recall the Cauchy-Swartz inequality in its integral form

$$\left(\int_0^T f(t)g(t)\mathrm{d}t\right)^2 \le \int_0^T f(t)^2 \, \mathrm{d}t \int_0^T g(t)^2 \, \mathrm{d}t$$

and if we set $f(t) = \|\dot\gamma(t)\|$ and $g(t) = 1$ we obtain

$$\ell(\gamma)^2 \le 2J(\gamma)T\,.$$

Moreover, the equality holds if and only if f is proportional to g, that is, $\|\dot\gamma(t)\|$ is constant in $[0,T]$.

Hence, if $\ell(\gamma)$ is the least possible and the equality holds then the minimum value of $J(\gamma)$ is obtained. This is the case if and only if γ is a length minimizer defined on $[0,T]$ and with constant speed. Conversely, the minimum of $J(\gamma)$ is obtained if and only if the equality holds, since all the other terms in the inequality are constant in Ω_{q_0,q_1}. □

Proof of theorem 3.2.1, first part. By lemma 3.2.7 we can assume that the minimizer $\gamma\colon [0,T] \to M$ is parametrized with constant speed so that it is also a minimizer of the functional J among all admissible curves joining $q_0 = \gamma(0)$ and $q_1 = \gamma(T)$ in a fixed time $T > 0$. In particular, we can define a new functional

$$\widetilde{J}(u(\cdot)) = \frac{1}{2}\int_0^T |u(t)|^2 \, \mathrm{d}t$$

on the space of controls $u(\cdot) \in L^\infty([0,T],\mathbb{R}^m)$. Thus we get that the minimal control $\overline{u}(\cdot)$ of γ is a minimizer for the energy functional $\widetilde{J}$ among all controls corresponding to trajectories joining $q_0, q_1 \in M$. Hereafter we will refer to $\widetilde{J}$ as J with abuse of notation.

We consider now a variation $u(\cdot) = \overline{u}(\cdot) + v(\cdot)$ of the minimum control $\overline{u}(\cdot)$, and its associated trajectory $q(t)$ (that is defined on $[0, T]$ for proposition 2.1.11) that is a solution to the equation

$$\dot{q}(t) = f_{u(t)}(q(t)) \qquad q(0) = q_0 \,.$$

We recall that $P_{0,t}$ is the local flow associated with the optimal control $\overline{u}(\cdot)$ and that $\gamma(t) = P_{0,t}(q_0)$ is the optimal admissible curve. We define the curve as follows:

$$x(t) = P_{0,t}^{-1}(q(t))$$

that is again well defined, provided that v is small enough, thanks to proposition 2.1.11. Notice that if $v(\cdot) = 0$, then $x(t)$ is constant q_0.

Then we can write $q(t) = P_{0,t}(x(t))$ and differentiating we get

$$\begin{aligned}\dot{q}(t) &= f_{\overline{u}(t)}(q(t)) + \mathrm{d}(P_{0,t})(\dot{x}(t)) \\ &= f_{\overline{u}(t)}(P_{0,t}(x(t))) + \mathrm{d}(P_{0,t})(\dot{x}(t))\,,\end{aligned}$$

and inverting $\dot{x}(t) = \mathrm{d}(P_{0,t}^{-1})\big[\dot{q}(t) - f_{\overline{u}(t)}(P_{0,t}(x(t)))\big]$. We can then substitute $\dot{q}(t) = f_{u(t)}(q(t)) = f_{u(t)}(P_{0,t}(x(t)))$:

$$\begin{aligned}\dot{x}(t) &= \mathrm{d}(P_{0,t}^{-1})\big[(f_{u(t)} - f_{\overline{u}(t)})(P_{0,t}(x(t)))\big] \\ &= \big[(P_{0,t}^{-1})_*\big(f_{u(t)} - f_{\overline{u}(t)}\big)\big](x(t)) \\ &= \big[(P_{0,t}^{-1})_* f_{v(t)}\big](x(t))\,.\end{aligned}$$

If we define the nonautonomous vector field $g^t_{v(t)} = (P_{0,t}^{-1})_* f_{v(t)}$ we obtain a Cauchy problem for $x(t)$

$$\begin{cases}\dot{x}(t) = g^t_{v(t)}(x(t)) \\ x(0) = q_0\end{cases} \tag{3.9}$$

and $x(t)$ is defined on $[0, T]$ provided v is small enough. Notice that the vector field $g^t_{v(t)}$ is linear with respect to v, since f_u is linear with respect to u. Then we fix a general control $v(t)$ and we consider the following map, defined in a neighbourhood of zero:

$$s \mapsto \begin{pmatrix} J(\overline{u} + sv) \\ x(T; \overline{u} + sv)\end{pmatrix} \in \mathbb{R} \times M$$

where here $x(T; \overline{u} + sv)$ denotes the solution at time T of *eq.* (3.9) corresponding to $sv(\cdot)$ in place of $v(\cdot)$, and $J(\overline{u} + sv)$ is the cost associated to $q(t)$ corresponding to the control $\overline{u} + sv$. □

This first part of the proof justifies the need for the following lemma:

Lemma 3.2.8. *With the notations introduced in the previous proof, there exist* $\overline{\lambda} \in (\mathbb{R} \oplus T_{q_0}M)^*$, *with* $\overline{\lambda} \neq 0$, *such that for all* $v \in L^\infty([0, T], \mathbb{R}^m)$

$$\left\langle \overline{\lambda}, \left(\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} J(\overline{u} + sv), \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} x(T; \overline{u} + sv) \right) \right\rangle = 0\,. \tag{3.10}$$

Proof. Assume by contradiction that eq. (3.10) does not hold for any $\overline{\lambda}$ different from zero, then there are $v_0, \dots, v_d \in L^\infty([0,T], \mathbb{R}^m)$ such that

$$\begin{pmatrix} \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} J(\overline{u} + sv_0) \\ \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} x(T; \overline{u} + sv_0) \end{pmatrix}, \dots, \begin{pmatrix} \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} J(\overline{u} + sv_d) \\ \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} x(T; \overline{u} + sv_d) \end{pmatrix}$$

are $d+1$ linearly independent vectors in $\mathbb{R} \oplus T_{q_0}M$. We then consider the map

$$\Phi \colon \mathbb{R}^{d+1} \to \mathbb{R} \times M \qquad \Phi(s_0, \dots, s_d) = \begin{pmatrix} J(\overline{u} + \sum_{i=0}^d s_i v_i) \\ x(T; \overline{u} + \sum_{i=0}^d s_i v_i) \end{pmatrix}.$$

As we will compute later in the proof of theorem 3.2.1, the map Φ is differentiable at $s = 0$ and by assumption the differential at $s = 0$ is surjective, therefore Φ is surjective from a neighbourhood of $s = 0$ to a neighbourhood of $(J(\overline{u}), q_0)$ in $\mathbb{R} \times M$. As a result we can find $v(\cdot) = \sum_{i=0}^d s_i v_i(\cdot)$ such that

$$x(T, \overline{u} + v) = x(T, \overline{u} + v) q_0 \qquad J(\overline{u} + v) < J(\overline{u}).$$

Therefore the curve $t \mapsto q(t; \overline{u} + v)$ joins $q(0; \overline{u} + v) = q_0$ to

$$q(T; \overline{u} + v) = P_{0,T}(q_0) = q_1,$$

and with a smaller cost of $\gamma(t) = q(t; \overline{u})$, which is a contradiction. □

Remark 3.2.9. We notice that $\overline{\lambda}$ provided in eq. (3.10) is defined up to a multiplicative constant, thus we may suppose $\overline{\lambda}$ to be either $(-1, \lambda_0)$ or $(0, \lambda_0)$ where $\lambda_0 \in T^*_{q_0}M$ and $\lambda_0 \neq$ in the second case, since $\overline{\lambda}$ is not zero.

Proof of theorem 3.2.1, second part. By eq. (3.10) there exists $\lambda_0 \in T^*_{q_0}M$ such that one of the following holds for every $v \in L^\infty([0,T], \mathbb{R}^m)$:

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} J(\overline{u} + sv) = \left\langle \lambda_0, \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} x(T; \overline{u} + sv) \right\rangle \tag{3.11}$$

$$0 = \left\langle \lambda_0, \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} x(T; \overline{u} + sv) \right\rangle \tag{3.12}$$

where $\lambda_0 \neq 0$ in the second case considering remark 3.2.9. Now we compute explicitly the terms involved in these equality. From the definition of J (actually $\widetilde{J}$) we get

$$\begin{aligned} \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} J(\overline{u} + sv) &= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \frac{1}{2} \int_0^T |\overline{u} + sv|^2 \,\mathrm{d}t \\ &= \int_0^T \sum_{i=1}^m \overline{u}_i(t) v_i(t) \,\mathrm{d}t. \end{aligned}$$

On the other hand, since gv is linear with respect to v, we can compute the other term using local coordinates (since $x(T;\overline{u}+sv)$ is in a unique local chart for all $t \in [0,T]$ when s is small enough):

$$\begin{aligned}
\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} x(T;\overline{u}+sv) &= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \left(q_0 + \int_0^T g^t_{sv(t)}(x(t;\overline{u}+sv))\,\mathrm{d}t \right) \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \left(q_0 + s\int_0^T g^t_{v(t)}(x(t;\overline{u}+sv))\,\mathrm{d}t \right) \\
&= \int_0^T g^t_{v(t)}(x(t;\overline{u}))\,\mathrm{d}t = \int_0^T g^t_{v(t)}(q_0)\,\mathrm{d}t \\
&= \int_0^T \sum_{i=1}^m \left((P^{-1}_{0,t})_* f_i \right)(q_0) v_i(t)\,\mathrm{d}t\,.
\end{aligned}$$

Finally

$$\begin{aligned}
\left\langle \lambda_0, \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} x(T;\overline{u}+sv) \right\rangle &= \int_0^T \sum_{i=1}^m \left\langle \lambda_0, \left((P^{-1}_{0,t})_* f_i \right)(q_0) \right\rangle v_i(t)\,\mathrm{d}t \\
&= \int_0^T \sum_{i=1}^m \left\langle \lambda_0, \mathrm{d}(P^{-1}_{0,t})(f_i(\gamma(t)) \right\rangle v_i(t)\,\mathrm{d}t \\
&= \int_0^T \sum_{i=1}^m \left\langle (P^{-1}_{0,t})^* \lambda_0, f_i(\gamma(t)) \right\rangle v_i(t)\,\mathrm{d}t \\
&= \int_0^T \sum_{i=1}^m \left\langle \lambda(t), f_i(\gamma(t)) \right\rangle v_i(t)\,\mathrm{d}t\,.
\end{aligned}$$

Therefore eq. (3.11) can be rewritten as

$$\int_0^T \sum_{i=1}^m \overline{u}_i(t) v_i(t)\,\mathrm{d}t = \int_0^T \sum_{i=1}^m \left\langle \lambda(t), f_i(\gamma(t)) \right\rangle v_i(t)\,\mathrm{d}t$$

and since it holds for every $v \in L^\infty([0,T],\mathbb{R}^m)$ we obtain $\overline{u}_i(t) = \left\langle \lambda(t), f_i(\gamma(t)) \right\rangle$ which is (N). On the other hand eq. (3.12) can be rewritten as

$$0 = \int_0^T \sum_{i=1}^m \left\langle \lambda(t), f_i(\gamma(t)) \right\rangle v_i(t)\,\mathrm{d}t$$

and since it holds for every $v \in L^\infty([0,T],\mathbb{R}^m)$ we obtain $0 = \left\langle \lambda(t), f_i(\gamma(t)) \right\rangle$ which is (A). □

3.3 End-point map and its differential

In this section we will always assume the sub-Riemannian structure $(M, \mathbf{U}, f)$ to be free, that is $\mathbf{U} \cong M \times \mathbb{R}^m$ for some $m \in \mathbb{N}$. We recall from theorem 2.3.9 that every sub-Riemannian structure is equivalent to a free one, hence the previous assumption is not restrictive. Thus we can assume $\{ f_1, \dots, f_m \}$ to be a global generating family. We fix $q_0 \in M$, for $u \in L^\infty(\mathbb{R}, \mathbb{R}^m)$ we will denote with γ_u the unique maximal solution to the Cauchy problem

$$\begin{cases} \dot\gamma(t) = \sum_{i=1}^{m} u_i(t) f_i(\gamma(t)) \\ \gamma(0) = q_0 \,. \end{cases} \tag{3.13}$$

We define $\mathcal{U}_{q_0} \subseteq L^\infty([0,1], \mathbb{R}^m)$ as the set of control u such that the corresponding trajectory γ_u from q_0 is defined on the interval $[0,1]$. From proposition 2.1.11 we know that $\mathcal{U}_{q_0}$ is an open set of $L^\infty([0,1], \mathbb{R}^m)$.

Definition 3.3.1. Let $(M, \mathbf{U}, f)$ be a free sub-Riemannian manifold and $q_0 \in M$. The *end-point map* based at q_0 is the map

$$\mathrm{End}_{q_0} \colon \mathcal{U}_{q_0} \to M \qquad \mathrm{End}_{q_0}(u) = \gamma_u(1)\,, \tag{3.14}$$

where $\gamma_u \colon [0,1] \to M$ is the unique solution of the Cauchy problem 3.13.

Proposition 3.3.2. *The end-point map* End_{q_0} *is Fréchet differentiable on* $\mathcal{U}_{q_0}$*. For every* $u \in \mathcal{U}_{q_0}$ *the differential* $D_u\mathrm{End}_{q_0} \colon L^\infty([0,1], \mathbb{R}^m) \to T_{\gamma_u(1)}M$ *satisfies*

$$D_u\mathrm{End}_{q_0}(v) = \int_0^1 \left[(P^u_{t,1})_* f_{v(t)} \right] \big|_{\gamma_u(1)} \,\mathrm{d}t \qquad v \in L^\infty([0,1], \mathbb{R}^m)\,. \tag{3.15}$$

Proof. Firstly we prove the statement for $u = 0$. For $\|v\|_{L^\infty([0,1],\mathbb{R}^m)}$ small enough, integral curves associated to the controls V are contained in a compact neighbourhood K of q_0 inside a local chart. Using that local chart, we need to prove that

$$\left| \mathrm{End}_{q_0}(v) - \int_0^1 f_{v(t)}(q_0)\,\mathrm{d}t \right|$$

is $o(\|v\|_\infty)$ uniformly in v. We compute explicitly:

$$\begin{aligned} \left| \mathrm{End}_{q_0}(v) - \int_0^1 f_{v(t)}(q_0)\,\mathrm{d}t \right| &= \left| \int_0^1 f_{v(t)}(\gamma_v(t))\,\mathrm{d}t - \int_0^1 f_{v(t)}(q_0)\,\mathrm{d}t \right| \\ &\le \sum_{i=1}^{m} \int_0^1 |v_i(t)| \big| f_i(\gamma_v(t)) - f_i(q_0) \big| \,\mathrm{d}t \end{aligned}$$

Since the topology induced by the sub-Riemannian distance is equivalent to the manifold topology, the smooth vector fields f_i are locally-Lipschitz with respect to the sub-Riemannian distance, hence Lipschitz in K with Lipschitz constant C_i, let $C = \max C_i$. Then

$$\begin{aligned} \sum_{i=1}^{m} \int_0^1 |v_i(t)| \big| f_i(\gamma_v(t)) - f_i(q_0) \big| \,\mathrm{d}t &\le \sum_{i=1}^{m} \int_0^1 |v_i(t)| C d(\gamma_v(t), q_0)\,\mathrm{d}t \\ &\le \sum_{i=1}^{m} \int_0^1 |v_i(t)| C \|v\|_\infty \,\mathrm{d}t \\ &\le mC\|v\|_\infty^2\,, \end{aligned}$$

that is $o(\|v\|_\infty)$ uniformly in v.

Let now consider u generic. Following the proof of theorem 3.2.1, we get that

$$\mathrm{End}_{q_0}(u+v) = P^u_{0,1} \circ G^u_{q_0}(v)$$

where $G^u_{q_0}$ is the end-point map associated to vector field

$$\{ (P^u_{0,1})^{-1}_* f_1, \dots, (P^u_{0,1})^{-1}_* f_m \}$$

defined in a proper neighbourhood of q_0. Therefore $G^u_{q_0}$ is defined provided $\|v\|_\infty$ is small enough. We thus obtain, using the previous part of the proof, that

$$\begin{aligned} D_u \mathrm{End}_{q_0}(v) &= \mathrm{d}(P^u_{0,1}) \circ D_0 G^u_{q_0}(v) \\ &= \mathrm{d}(P^u_{0,1}) \int_0^1 (P^u_{0,1})^{-1}_* f_{v(t)}(q_0)\,\mathrm{d}t \\ &= \int_0^1 (P^u_{0,1})_* f_{v(t)}(\mathrm{End}_{q_0}(u))\,\mathrm{d}t\,. \end{aligned}$$

□

As we will explain in the next proposition, the differential of end-point map is particularly useful to characterize abnormal extremal.

Proposition 3.3.3. *With the same notations as in theorem 3.2.1, we have the following*

(N) $(u(t),\lambda(t))$ *is a normal extremal if and only if there exists* $\lambda_1 \in T^*_{q_1}M$*, where* $q_1 = \mathrm{End}_{q_0}(u)$*, such that* $\lambda(t) = \left(P^u_{t,1}\right)^* \lambda_1$ *for all* t*, and* u *satisfies*

$$\langle \lambda_1 \,|\, D_u \mathrm{End}_{q_0}(v) \rangle = (u \,|\, v)_{L^2} \tag{3.16}$$

for all $v \in L^\infty([0,1],\mathbb{R}^m)$*.*

(A) $(u(t),\lambda(t))$ *is an abnormal extremal if and only if there exists* $0 \neq \lambda_1 \in T^*_{q_1}$*, where* $q_1 = \mathrm{End}_{q_0}(u)$*, such that* $\lambda(t) = \left(P^u_{t,1}\right)^* \lambda_1$ *for all* t*, and* u *satisfies*

$$\lambda_1 D_u \mathrm{End}_{q_0} = 0\,, \tag{3.17}$$

as a linear operator $L^\infty([0,1],\mathbb{R}^m) \to \mathbb{R}$*.*

Proof. Firstly we prove (N). We recall that the pair $(u(t),\lambda(t))$ is a normal extremal if the curve $\lambda(t)$ satisfies $\lambda(t) = \left(P^{-1}_{0,t}\right)^*\lambda_0$ and $\langle \lambda(t), f_i(\gamma(t))\rangle = u_i(t)$ for all $i = 1,\dots,m$ and $\gamma(t) = \pi(\lambda(t))$. We prove that $\lambda_1 = \lambda(1)$ satisfies proposition 3.3.3. We observe that $\lambda(t) = \left(P^u_{t,1}\right)^* \lambda_1$. Let $v \in L^\infty([0,1],\mathbb{R}^m)$, thanks to proposition 3.3.2 we get

$$\begin{aligned} \langle \lambda_1, D_u \mathrm{End}_{q_0}(v)\rangle &= \int_0^1 \langle \lambda_1, (P^u_{t,1})_* f_{v(t)}(q_1)\rangle\ \mathrm{d}t = \int_0^1 \langle (P^u_{t,1})^*\lambda_1, f_{v(t)}(\gamma(t))\rangle\ \mathrm{d}t \\ &= \int_0^1 \langle \lambda(t), f_{v(t)}(\gamma(t))\rangle\ \mathrm{d}t = \int_0^1 \sum_{i=1}^m \langle \lambda(t), f_i(\gamma(t))\rangle\, v_i(t)\,\mathrm{d}t \end{aligned}$$

where we used $\gamma(t) = \left(P^u_{t,1}\right)^{-1}(q_1)$. Since $(u(t), \lambda(t))$ is a normal extremal, we get by definition

$$\int_0^1 \sum_{i=1}^m \langle \lambda(t), f_i(\gamma(t)) \rangle \, v_i(t) \, \mathrm{d}t = \int_0^1 \sum_{i=1}^m u_i(t) v_i(t) \, \mathrm{d}t \, .$$

Therefore $\lambda_1 D_u \mathrm{End}_{q_0} = u$ as a linear operator $L^\infty([0,1], \mathbb{R}^m) \to \mathbb{R}$. Conversely if λ_1 satisfies proposition 3.3.3, then the previous equation holds for all $v \in L^\infty([0,1], \mathbb{R}^m)$. The fundamental lemma of calculus of variations implies $\langle \lambda(t), f_i(\gamma(t)) \rangle = u_i(t)$ for all $i = 1, \dots, m$.

The case (A) is analogous. □

We are finally ready to state the following result:

Corollary 3.3.4. *The pair $(u(t), \lambda(t))$ is an abnormal extremal if and only if the differential of the end-point map* End_{q_0} *at u is not surjective.*

Remark 3.3.5. The previous corollary stresses that abnormal extremals depends only on the distribution $\mathcal{D}$ and not on the metric structure of the sub-Riemannian manifolds. Therefore, two sub-Riemannian structures on a manifold M which are equivalent as distributions leads to the same abnormal extremals.

Remark 3.3.6. The discussion of this section can be rephrased in the context of $L^1([0,1], \mathbb{R}^m)$ controls and $L^2([0,1], \mathbb{R}^m)$ controls, which lead to AC-admissible curves and $W^{1,2}$-admissible curves. The last setting is particularly useful since $L^2([0,1], \mathbb{R}^m)$ is an Hilbert space.

Chapter 4

Nilpotent Lie groups

In this chapter we present Lie algebras and Lie groups which will be the main object of study in the next chapter. Here we will concentrate on the algebraic aspects of the topic. The main purpose of this discussion is to justify the Baker-Campbell-Hausdorff formula and to frame the structure of simply connected nilpotent Lie groups, these results will be fundamental for developing calculus and reasonable coordinates in the context of sub-Riemannian geometry on Carnot groups. For a more detailed dissertation on the subject we refer to [HN12, chapters 5, 9 and 13].

4.1 Lie algebras and nilpotency

Definition 4.1.1. Let $\mathfrak{g}$ be a vector space. A *Lie bracket* on $\mathfrak{g}$ is a bilinear map $[\cdot,\cdot]\colon \mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$ satisfying the following properties:

$$[x,y] = -[y,x] \qquad \text{for } x,y\in\mathfrak{g}, \tag{L1}$$

$$[x,[y,z]]+[y,[z,x]]+[z,[x,y]] = 0 \qquad \text{for } x,y,z\in\mathfrak{g}\ (\textit{Jacobi identity}). \tag{L2}$$

For any Lie bracket on $\mathfrak{g}$, the pair $(\mathfrak{g},[\cdot,\cdot])$ is called a *Lie algebra.*

Example 4.1.2. A vector space $\mathcal{A}$ with a bilinear map "$\cdot$": $\mathcal{A}\times\mathcal{A}\to\mathcal{A}$ is called an *associative algebra* if

$$a\cdot(b\cdot c) = (a\cdot b)\cdot c \quad \text{for any } a,b,c\in\mathcal{A}\,. \tag{4.1}$$

The commutator

$$[a,b] = a\cdot b - b\cdot a \tag{4.2}$$

defines a Lie bracket on $\mathcal{A}$, since eq. (L1) and eq. (L2) follow immediately. We denote such Lie algebra with $\mathcal{A}_L$.

Example 4.1.3. We consider the most fundamental examples of Lie algebras:

(a) Let V be a finite dimensional vector space and $\mathrm{End}(V)$ be the set of linear endomorphism of V. Then $\mathrm{End}(V)$ is an associative algebra under composition, we consider $\mathfrak{gl}(V) = \mathrm{End}(V)_L$ to be the corresponding Lie algebra.

(b) The space $M_n(\mathbb{K})$ of $(n \times n)$-matrices with entries in $\mathbb{K}$ is an associative algebra under matrix multiplication. We consider $\mathfrak{gl}_n(\mathbb{K}) = M_n(\mathbb{K})_L$ to be the corresponding Lie algebra.

Definition 4.1.4. Let $\mathfrak{g}$ and $\mathfrak{h}$ be Lie algebras. A linear map $\alpha \colon \mathfrak{g} \to \mathfrak{h}$ is a *homomorphism* if

$$\alpha\big([x, y]\big) = \big[\alpha(x), \alpha(y)\big] \quad \text{for } x, y \in \mathfrak{g}\,. \tag{4.3}$$

An *isomorphism* of Lie algebras is a bijective homomorphism.

A *representation* of a Lie algebra $\mathfrak{g}$ over a vector space V is a homomorphism $\alpha \colon \mathfrak{g} \to \mathfrak{gl}(V)$, also denoted with (α, V).

Definition 4.1.5. Let us consider particular subspaces in a Lie algebra:

(a) Let U, V be subesets of a Lie algebra $\mathfrak{g}$. We denote by

$$[U, V] = \operatorname{span}\left\{\, [u, v] \,\middle|\, u \in U, v \in V \,\right\} \tag{4.4}$$

the smallest linear subspace spanned by Lie brackets of elements of V with elements in U.

(b) A linear subspace $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$ is a *Lie subalgebra* if $[\mathfrak{h}, \mathfrak{h}] \subseteq \mathfrak{h}$.

(c) A linear subspace $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$ is an *ideal* of $\mathfrak{g}$ if $[\mathfrak{g}, \mathfrak{h}] \subseteq \mathfrak{h}$.

(d) A Lie algebra $\mathfrak{g}$ is said to be *Abelian* if $[\mathfrak{g}, \mathfrak{g}] = \{\, 0 \,\}$, which means that all brackets vanish.

Proposition 4.1.6. *The image of a Lie algebra homomorphism $\alpha \colon \mathfrak{g}_1 \to \mathfrak{g}_2$ is a Lie subalgebra of g_2.*

Proof. $\big[\alpha(x), \alpha(y)\big] = \alpha\big([x, y]\big)$ for any $x, y \in \mathfrak{g}_1$. □

Proposition 4.1.7. *Let $\alpha \colon \mathfrak{g}_1 \to \mathfrak{g}_2$ be a Lie algebra homomorphism. Then, whenever $\mathfrak{h}$ is an ideal of $\mathfrak{g}_2$ then also $\alpha^{-1}(\mathfrak{h})$ is an ideal of $\mathfrak{g}_1$. In particular* $\ker \alpha$ *is always an ideal.*

Proof. Let $x \in \alpha^{-1}(\mathfrak{h})$, then for any $y \in \mathfrak{g}_1$

$$\alpha\big([x, y]\big) = \big[\alpha(x), \alpha(y)\big] \in \mathfrak{h}$$

since $\alpha(x) \in \mathfrak{h}$ and $\mathfrak{h}$ is an ideal. This means $[x, y] \in \alpha^{-1}(\mathfrak{h})$, hence $\alpha^{-1}(\mathfrak{h})$ is an ideal. Finally, $\{\, 0 \,\}$ is always an ideal of $\mathfrak{g}_2$, therefore $\ker \alpha = \alpha^{-1}(\{\, 0 \,\})$ is an ideal of $\mathfrak{g}_1$. □

Example 4.1.8. Let consider some examples of Lie subalgebras and ideals:

(i) Let $\mathfrak{g}$ be a Lie algebra. Then the *center* of $\mathfrak{g}$

$$\mathfrak{z}(\mathfrak{h}) = \left\{\, x \in \mathfrak{g} \,\middle|\, [x, y] = 0 \quad \forall y \in \mathfrak{g} \,\right\} \tag{4.5}$$

is an ideal of $\mathfrak{g}$.

(ii) Let V be a subspace of a Lie algebra $\mathfrak{g}$. The *normalizer* of V in $\mathfrak{g}$

$$\mathfrak{n}_{\mathfrak{g}}(V) = \left\{\, x \in \mathfrak{g} \,\middle|\, [x, V] \subseteq V \,\right\} \tag{4.6}$$

is a Lie subalgebra og $\mathfrak{g}$.

Definition 4.1.9. Let $\mathfrak{g}$ be a Lie algebra. A linear map $\delta\colon \mathfrak{g} \to \mathfrak{g}$ is called a *derivation* if

$$\delta\big([x,y]\big) = [\delta(x), y] + [x, \delta(y)] \quad \text{for any } x, y \in \mathfrak{g}\,. \tag{4.7}$$

The set of all derivations is denoted by $\operatorname{der}(\mathfrak{g})$.

Proposition 4.1.10. *Let $\mathfrak{g}$ be a Lie algebra and $x \in \mathfrak{g}$. Then the linear map*

$$\operatorname{ad} x\colon \mathfrak{g} \to \mathfrak{g} \qquad y \mapsto [x, y] \tag{4.8}$$

is a derivation. The map

$$\operatorname{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}) \qquad x \mapsto \operatorname{ad} x \tag{4.9}$$

is a representation of the Lie algebra $\mathfrak{g}$.

Proof. Using Jacoby identity

$$\begin{aligned}(\operatorname{ad} x)[y,z] &= [x,[y,z]] = [[x,y],z] + [y,[x,z]] \\ &= [(\operatorname{ad} x)y, z] + [y, (\operatorname{ad} x)z]\end{aligned}$$

hence $\operatorname{ad} x$ is a derivation. Moreover

$$\begin{aligned}\big(\operatorname{ad}[x,y]\big)z &= [[x,y,z]] = [x,[y,z]] - [y,[x,z]] \\ &= (\operatorname{ad} x) \circ (\operatorname{ad} y)z - (\operatorname{ad} y) \circ (\operatorname{ad} x)z \\ &= [\operatorname{ad} x, \operatorname{ad} y]\end{aligned}$$

hence ad is a Lie algebra homomorphism between $\mathfrak{g}$ and $\mathfrak{gl}(\mathfrak{g})$, that is a representation of Lie algebra. □

Definition 4.1.11. A derivation of the type $\operatorname{ad} x$ for some $x \in \mathfrak{g}$ is called *inner derivation.* The set of inner derivations is denoted with $\operatorname{ad}(\mathfrak{g})$. The representation $(\operatorname{ad}, \mathfrak{g})$ is called *inner representation.*

Proposition 4.1.12. *For any Lie algebra $\mathfrak{g}$:*

(i) $\operatorname{der}(\mathfrak{g})$ *is a Lie subalgebra of* $\mathfrak{gl}(\mathfrak{g})$ *and* $\operatorname{ad}(\mathfrak{g})$ *is an ideal of* $\operatorname{der}(\mathfrak{g})$. *In particular*

$$[\delta, \operatorname{ad} x] = \operatorname{ad} \delta(x) \qquad \textit{for any } \delta \in \operatorname{der}(\mathfrak{g}) \textit{ and } x \in \mathfrak{g}\,, \tag{4.10}$$

(ii) $\ker(\operatorname{ad}) = \mathfrak{z}(\mathfrak{g})$.

Proof. (i) Linear combinations of derivations are trivially derivations. Moreover, let $\delta, \eta \in \operatorname{der}(\mathfrak{g})$ and $x, y \in \mathfrak{g}$, then

$$\begin{aligned}
\big([\delta,\eta]\big)[x,y] &= (\delta\circ\eta)[x,y] - (\eta\circ\delta)[x,y] \\
&= \delta\big([\eta(x),y] + [x,\eta(y)]\big) - \eta\big([\delta(x),y] + [x,\delta(y)]\big) \\
&= [(\delta\circ\eta)x,y] + [\eta(x),\delta(y)] + [\delta(x),\eta(y)] + [x,(\delta\circ\eta)y] - \\
&\quad - [(\eta\circ\delta)x,y] - [\delta(x),\eta(y)] - [\eta(x),\delta(y)] - [x,(\eta\circ\delta)y] \\
&= [(\delta\circ\eta)x - (\eta\circ\delta)x, y] + [x, (\delta\circ\eta)y - (\eta\circ\delta)y] \\
&= \big[[\delta,\eta]x, y\big] + \big[x, [\delta,\eta]y\big]
\end{aligned}$$

that means $[\delta, \eta]$ is again a derivation and $\operatorname{der}(\mathfrak{g})$ is a Lie subalgebra of $\mathfrak{gl}(\mathfrak{g})$.

Linear combinations of inner derivations are inner derivation, indeed $\operatorname{ad}(x+y) = \operatorname{ad} x + \operatorname{ad} y$. Moreover, if $\delta \in \operatorname{der}(\mathfrak{g})$ and $x \in \mathfrak{g}$, then

$$\begin{aligned}
[\delta, \operatorname{ad} x]z &= \delta\circ(\operatorname{ad} x)z - (\operatorname{ad} x)\circ\delta(z) = \delta\big([x,z]\big) - [x,\delta(z)] \\
&= [\delta(x), z] = \big(\operatorname{ad}\delta(x)\big)z
\end{aligned}$$

that means $\operatorname{ad}(\mathfrak{g})$ is an ideal of $\operatorname{der}(\mathfrak{g})$ and eq. (4.10) holds.

(ii)

$$\operatorname{ad} x = 0_{\mathfrak{gl}(\mathfrak{g})} \iff [x,y] = 0 \quad \forall y \in \mathfrak{g} \iff x \in \mathfrak{z}(\mathfrak{g})\,.$$

□

Proposition 4.1.13. *Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{n}$ an ideal of $\mathfrak{g}$. Then the quotient space $\mathfrak{g}/\mathfrak{n} = \{\, x + \mathfrak{n} \mid x \in \mathfrak{g} \,\}$ is a Lie algebra with respect to the bracket*

$$[x + \mathfrak{n}, y + \mathfrak{n}] = [x, y] + \mathfrak{n}\,. \tag{4.11}$$

The quotient map $\pi\colon \mathfrak{g} \to \mathfrak{g}/\mathfrak{n}$ is a surjective Lie algebra homomorphism of Lie algebras with kernel $\mathfrak{n}$.

Proof. The Lie bracket is well defined since $\mathfrak{n}$ is an ideal. All other properties follows from the respective properties of the Lie bracket on $\mathfrak{g}$. □

Definition 4.1.14. Let $\mathfrak{g}$ be a Lie algebra, its *descending (lower) central series* is the sequence of subsets of $\mathfrak{g}$ defined inductively:

$$C^0(\mathfrak{g}) = \mathfrak{g} \quad \text{and} \quad C^{n+1}(\mathfrak{g}) = [\mathfrak{g}, C^n(\mathfrak{g})]\,. \tag{4.12}$$

In particular $C^2(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}]$ is the *commutator algebra* of $\mathfrak{g}$. The Lie algebra $\mathfrak{g}$ is called *nilpotent* if $C^d(\mathfrak{g}) = \{\,0\,\}$ for some integer d. If d is minimal with this property, then it is called the *nilpotence degree* of $\mathfrak{g}$.

Remark 4.1.15. By induction, we can easily check that $C^n(\mathfrak{g})$ in an ideal of $\mathfrak{g}$ and $C^{n+1}(\mathfrak{g}) \subseteq C^n(\mathfrak{g})$. Then, if $\mathfrak{g}$ is finite dimensional, $\mathfrak{g}$ is nilpotent if and only if

$$C^\infty(\mathfrak{g}) = \bigcap_{n\in\mathbb{N}} C^n(\mathfrak{g}) = \{\,0\,\}\,. \tag{4.13}$$

Proposition 4.1.16. *Let $\mathfrak{g}$ be a Lie algebra.*

(i) If $\mathfrak{g}$ is nilpotent, then all subalgebras and all homomorphic images of $\mathfrak{g}$ are nilpotent.

(ii) If $\mathfrak{n} \subseteq \mathfrak{z}(\mathfrak{g})$ is an ideal and $\mathfrak{g}/\mathfrak{n}$ is nilpotent, then $\mathfrak{g}$ is nilpotent.

(iii) If $\mathfrak{g} \neq \{ 0 \}$ is nilpotent, then $\mathfrak{z}(\mathfrak{g}) \neq \{ 0 \}$.

(iv) If $\mathfrak{g}$ is nilpotent, then there exists $n \in \mathbb{N}$ such that $(\operatorname{ad} x)^n = 0$ for all $x \in \mathfrak{g}$, that is $\operatorname{ad}(x)$ is nilpotent as linear maps.

Proof. (i) If $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, then $[\mathfrak{h}, \mathfrak{h}] \subseteq [\mathfrak{g}, \mathfrak{g}]$ and $C^n(\mathfrak{h}) \subseteq C^n(\mathfrak{g})$ by induction. Therefore each subalgebra of a nilpotent Lie algebra is nilpotent.

If $\alpha\colon \mathfrak{g} \to \mathfrak{h}$ is a homomorphism of Lie algebras, by induction we obtain

$$C^n\big(\alpha(\mathfrak{g})\big) = \alpha\big(C^n(\mathfrak{g})\big) \quad \forall n \in \mathbb{N}\,. \tag{4.14}$$

(ii) If $\mathfrak{g}/\mathfrak{n}$ is nilpotent, then $C^n(\mathfrak{g}/\mathfrak{n}) = \{ 0 \}$ for some n. Using eq. (4.14) with the quotient map $\pi\colon \mathfrak{g} \to \mathfrak{g}/\mathfrak{n}$ we get $C^n(\mathfrak{g}) \subseteq \ker(\pi) = \mathfrak{n} \subseteq \mathfrak{z}(\mathfrak{g})$. Then $C^{n+1}(\mathfrak{g}) \subseteq [\mathfrak{g}, \mathfrak{z}(\mathfrak{g})] = \{ 0 \}$.

(iii) If $\mathfrak{g} \neq \{ 0 \}$ is nilpotent, then its nilpotence degree $d \neq 0$ and $C^{d-1}(\mathfrak{g}) \neq \{ 0 \}$. Then $[\mathfrak{g}, C^{d-1}(\mathfrak{g})] = \{ 0 \}$ means that $C^{d-1}(\mathfrak{g})$ is a non-trivial ideal in the center.

(iv) If $C^n(\mathfrak{g}) = \{ 0 \}$, then $(\operatorname{ad} x)^n \mathfrak{g} \subseteq C^n(\mathfrak{g}) = \{ 0 \}$. □

We conclude this section with an important characterization theorem for nilpotent Lie algebra, that relates the global property of nilpotency with a local property.

Theorem 4.1.17. *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. Then $\mathfrak{g}$ is nilpotent if and only if, for each $x \in \mathfrak{g}$, the operator $\operatorname{ad} x$ is nilpotent as linear map.*

4.2 Introduction to Lie groups

Definition 4.2.1. A *Lie group* is a group G endowed with a structure of differentiable manifold such that the group operations

$$m_G\colon G \times G \to G \qquad (x, y) \mapsto xy \tag{4.15}$$

$$\iota_G\colon G \to G \qquad x \mapsto x^{-1} \tag{4.16}$$

are differentiable. A *morphism of Lie groups* is a differentiable homomorphism of Lie groups $\varphi\colon G_1 \to G_2$.

Remark 4.2.2. We will denote the identity element of a Lie group G with $\mathbf{1}_G$, or simply $\mathbf{1}$ if there is no ambiguity. For $g \in G$ we define

$$L_g\colon G \to G \qquad x \mapsto gx \tag{4.17}$$

$$R_g\colon G \to G \qquad x \mapsto xg \tag{4.18}$$

$$C_g\colon G \to G \qquad x \mapsto gxg^{-1}\,. \tag{4.19}$$

They are diffeomorphism of G, moreover C_g defines a group automorphism of G, so we obtain a group homomorphism

$$C\colon G \to \mathrm{Aut}(G) \qquad g \mapsto C_g \tag{4.20}$$

where $\mathrm{Aut}(G)$ is the *group of differentiable automorphism* of a Lie group G. Elements in the image of C are called *inner automorphism*, the image of C is denoted by $\mathrm{Inn}(G)$.

Example 4.2.3. The additive group $G = (\mathbb{R}^n, +)$ is a Lie group with respect to the natural n-dimensional manifold structure on $\mathbb{R}^n$. Addition and inverse are differentiable.

Example 4.2.4. We consider the group $G = \mathrm{GL}_n(\mathbb{R})$ of invertible $(n \times n)$-matrices. If $\det\colon M_n(\mathbb{R}) \to \mathbb{R}$ denotes the determinant function, then det is polynomial, hence differentiable and $\mathrm{GL}_n(\mathbb{R}) = \det^{-1}(\mathbb{R} \setminus \{ 0 \})$ is an open set of $M_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$. Hence G carries a natural differentiable manifold structure.

Both matrix multiplication and matrix inverse are polynomial, hence differentiable. This means that $\mathrm{GL}_n(\mathbb{R})$ is a Lie group.

We recall the notion of pushforward from eq. (1.39).

Definition 4.2.5. A vector field defined on a Lie group $X \in \mathrm{Vec}(G)$ is called *left invariant* if

$$X = (L_g)_* X = g.X \qquad \forall g \in G\,. \tag{4.21}$$

We refer to set of left invariant vector fields on G with $\mathrm{Vec}^\ell(G)$. From proposition 1.4.5 it follows that the Lie bracket of left invariant vector fields is again left invariant. Then $\mathrm{Vec}^\ell(G)$ is a Lie subalgebra of $\mathrm{Vec}(G)$.

Moreover, the value of a left invariant vector field is uniquely determined by its value at $\mathbf{1}$ (or any other point), indeed we have

$$X(g) = \mathrm{d}(L_g)_{\mathbf{1}}(X(\mathbf{1})) \quad \forall g \in G \tag{4.22}$$

and, on the other hand, for $x \in T_{\mathbf{1}}G$ then $g \mapsto \mathrm{d}(L_g)_{\mathbf{1}}x$ defines a left invariant vector field. Then there is a correspondence between $T_{\mathbf{1}}G$ and $\mathrm{Vec}^\ell(G)$ so that $T_{\mathbf{1}}G$ inherits a Lie algebra structure.

The Lie algebra

$$\mathbf{L}(G) = \bigl(T_{\mathbf{1}}G, [\cdot,\cdot]\bigr) \cong \mathrm{Vec}^\ell(G) \tag{4.23}$$

is the *Lie algebra* associated with G.

Hereafter, whenever X is a left invariant vector field, we will also use X for the corresponding tangent vector at $\mathbf{1}$.

Let us consider an integral curve γ of a left invariant vector field $X \in \mathrm{Vec}^\ell(G)$, then $t \mapsto g\gamma(t)$ is again an integral curve of X since X is left invariant. This means that an integral curve defined on a bounded set $]a, b[\subseteq \mathbb{R}$ can always be extended with a proper multiplication by $g \in \gamma(]a, b[)$ to an integral curve define on $]a - \varepsilon, b + \varepsilon[$ and ε is independent of a and b. The previous argument can be formalised to the following statement:

Proposition 4.2.6. *Each left invariant vector field X on a Lie group G is complete.*

Definition 4.2.7. We define the *exponential function*

$$\exp_G\colon \mathbf{L}(G) \to G \qquad \exp_G(X) = \exp(X)(\mathbf{1}_G)\,. \tag{4.24}$$

We use the same notation for both the exponential function of a Lie group and the flow of a vector field, the following lemma guarantees there is no ambiguity (up to an exchange) between composition of flows and group multiplication of elements expressed as exponential.

Lemma 4.2.8. *For each $X \in \mathbf{L}(G)$, the curve $\gamma_X(t) = \exp_G(tX)$ is a morphism of Lie group and $\gamma'_X(0) = X$. If $X \in \mathrm{Vec}(G)$ is left invariant, then*

$$\exp(tX)(g) = g\exp_G(tX)\,. \tag{4.25}$$

Proof. Since $\gamma_X(t)$ and the group multiplication are differentiable, then also $\exp_G(tX)g$ is a differentiable map $\mathbb{R}\times G \to G$. If we define $\gamma(t) = g\exp_G(tX)$ then

$$\gamma'(t) = \mathrm{d}L_g\gamma'_X(t) = \mathrm{d}L_g X\big(\gamma_X(t)\big) \overset{*}{=} X\big(\gamma_X(t)g\big) = X(\gamma(t))\,,$$

where in $(*)$ we used the left invariance of X. Therefore γ is an integral curve of X and $\gamma(0) = g$. In particular

$$\begin{aligned}\gamma_X(t+s) &= \exp((t+s)X)(g) = \exp(tX)(\exp_G(sX))\\ &= \exp_G(sX)\exp_G(tX) = \gamma_X(s)\gamma_X(t)\end{aligned}$$

and $\gamma_X\colon t \to G$ is a group homomorphism. □

Remark 4.2.9. Let G, H be a Lie group and $X \in \mathrm{Vec}^\ell(G)$, if $\varphi\colon G \to H$ is a Lie group morphism then we consider $y = \mathrm{d}\varphi\big(X(\mathbf{1}_G)\big)$ and we can define a left invariant vector field on H:

$$Y = \varphi_* X\colon H \to TH \qquad h \mapsto \big(\mathrm{d}L_h\big)y\,. \tag{4.26}$$

Moreover

$$\begin{aligned}\mathrm{d}\varphi_g\big(X(g)\big) &= (\mathrm{d}\varphi_g \circ \mathrm{d}L_g)X(\mathbf{1}_G) = \mathrm{d}(\varphi\circ L_g)_{\mathbf{1}_G}X(\mathbf{1}_G)\\ &= \mathrm{d}(L_{\varphi(g)}\circ\varphi)_{\mathbf{1}_G}X(\mathbf{1}_G) = \mathrm{d}L_{\varphi(g)}\big(\mathrm{d}\varphi\big(X(\mathbf{1}_G)\big)\big)\\ &= \varphi_* X(\varphi(g))\end{aligned}$$

so that φ_* is a generalised pushforward for left invariant vector fields, defined even if φ is neither injective nor surjective. Finally

$$\varphi(\exp_G(tX)) = \exp_H(t\varphi_* X)\,.$$

Proposition 4.2.10. *If $\varphi\colon G \to H$ is a morphism of Lie groups, then the map*

$$\mathbf{L}(\varphi) = \mathrm{d}\varphi_{\mathbf{1}}\colon \mathbf{L}(G) \to \mathbf{L}(H) \tag{4.27}$$

is a homomorphism of Lie algebras.

Proof. Using remark 4.2.9, it is sufficient to prove that $\varphi_*[X,Y] = [\varphi_* X, \varphi_* Y]$ for every $X, Y \in \mathrm{Vec}^\ell(G)$. Since φ is a group homomorphism

$$\begin{aligned}\varphi\big(\exp(-tX)\circ\exp(sY)\circ\exp(tX)(\mathbf{1}_G)\big) = \\ &= \varphi\big(\exp_G(tX)\exp(sY)\exp(-tX)\big)\\ &= \varphi(\exp_G(tX))\varphi(\exp_G(sY))\varphi(\exp_G(-tX))\\ &= \exp_H(t\varphi_* X)\exp_H(s\varphi_* Y)\exp_H(-t\varphi_* X)\\ &= \exp(-t\varphi_* X)\circ\exp(s\varphi_* Y)\circ\exp(t\varphi_* X)(\mathbf{1}_H)\,.\end{aligned}$$

Differentiating both sides with respect to both t and s and evaluating at zero, we obtain by definition $\varphi_*[X,Y] = [\varphi_* X, \varphi_* Y]$. □

Proposition 4.2.11. *Let $\varphi\colon G \to H$ be a morphism of Lie groups and let $\mathbf{L}(\varphi)$ as in proposition 4.2.10. Then*

$$\exp_G \circ \mathbf{L}(\varphi) = \varphi \circ \exp_G\,, \tag{4.28}$$

thus the following diagram commutes

$$\begin{array}{ccc} \mathbf{L}(G) & \xrightarrow{\mathbf{L}(\varphi)} & H \\ {\scriptstyle \exp_G}\big\downarrow & & \big\downarrow{\scriptstyle \exp_H} \\ G & \xrightarrow{\varphi} & H \end{array}\,. \tag{4.29}$$

Example 4.2.12. Let us consider the Lie group of matrices $G = \mathrm{GL}_n(\mathbb{R})$ and its Lie algebra $\mathbf{L}(G) = T_{\mathbf{1}_G}G = M_n(\mathbb{R})$. A left invariant vector field X such that $X(\mathbf{1}) = A$ is given by

$$X(g) = gA\,. \tag{4.30}$$

The unique solution $\gamma_A\colon \mathbb{R} \to \mathrm{GL}_n(\mathbb{R})$ of the initial value problem

$$\gamma(0) = \mathbf{1}_G \quad \gamma'(t) = X\big(\gamma(t)\big) = \gamma(t)A$$

is given by the exponential matrix

$$\gamma_A(t) = e^{tA} = \sum_{k=0}^{\infty}\frac{1}{k!}t^k A^k\,.$$

So that we can write $\exp_G(X) = e^X$. The Lie algebra structure on $\mathbf{L}(G)$ is given by

$$\begin{aligned}[X,Y] &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0}\exp(-tX)\circ\exp(sY)\circ\exp(tX)(\mathbf{1}_G)\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} e^{tX}e^{sY}e^{-tX} = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} e^{tX}Ye^{-tX} = XY - YX\,.\end{aligned}$$

Then $\mathbf{L}(G) = \mathfrak{gl}_n(\mathbb{R})$.

Definition 4.2.13. We recall that for any $g \in G$ we define the inner automorphism $C_g \colon x \mapsto gxg^{-1}$, we then define

$$\mathrm{Ad}\colon G \to \mathrm{Aut}(\mathbf{L}(G)) \qquad \mathrm{Ad}(g) = \mathbf{L}(C_g)\,. \tag{4.31}$$

It is called the *adjoint representation* and it is differentiable, moreover

$$\mathbf{L}(\mathrm{Ad})\colon \mathbf{L}(G) \to \mathfrak{gl}(\mathbf{L}(G))\,, \tag{4.32}$$

which is a representation of $\mathbf{L}(G)$ on $\mathbf{L}(G)$.

Lemma 4.2.14. $\mathbf{L}(\mathrm{Ad}) = \mathrm{ad}$*, that is* $\mathbf{L}(\mathrm{Ad})(X)Y = [X, Y]$.

Proof. $X, Y \in \mathrm{Vec}^\ell(G)$, then

$$\begin{aligned}[X,Y] &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \exp(-tX) \circ \exp(sY) \circ \exp(tX)(\mathbf{1}_G) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \exp_G(tX)\exp_G(sY)\exp_G(-tX) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} C_{\exp_G(tX)}\big(\exp_G(sY)\big) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \mathrm{Ad}(\exp_G(tX))Y = \mathbf{L}(\mathrm{Ad})(X)Y\,. \end{aligned}$$

□

Remark 4.2.15. Combining proposition 4.2.11 and lemma 4.2.14 we obtain

$$\mathrm{Ad} \circ \exp_G = \exp_{\mathrm{Aut}(\mathbf{L}(G))} \circ \mathrm{ad} \tag{4.33}$$

that is

$$\mathrm{Ad}(\exp_G(X)) = e^{\mathrm{ad}\,X} \qquad \forall x \in \mathbf{L}(G)\,. \tag{4.34}$$

4.3 Baker-Campbell-Hausdorff formula

Proposition 4.3.1. *For a Lie group* G*, the exponential function*

$$\exp_G\colon \mathbf{L}(G) \to G$$

is differentiable and satisfies

$$(\mathrm{d}\exp_G)_0 = \mathrm{id}_{\mathbf{L}(G)}\,. \tag{4.35}$$

In particular $\exp_G$ *is a local diffeomorphism at* 0*. We define* $\log_G$*, the inverse of* $\exp_G$ *defined on* $\exp_G(U_0)$ *such that* $\exp_G$ *is a diffeomorphism on the neighbourhood* U_0 *of* 0.

Remark 4.3.2. From proposition 1.4.7 we also know that, if $X_1, \ldots, X_d$ is a basis for $\mathbf{L}(G)$, then

$$\begin{aligned}\Psi\colon \mathbf{L}(G) \cong \mathbb{R}^d \to G \qquad \Psi(t_1, \ldots, t_d) &= \exp(t_1X_1) \circ \cdots \circ \exp(t_dX_d)(\mathbf{1}) \\ &= \exp_G(t_dX_d)\ldots\exp_G(t_1X_1)\end{aligned}$$

is again a local diffeomorphism at 0. Then we can consider, for $X, Y \in U_0$, a tangent vector at $\mathbf{1}$ satisfying

$$X * Y = \log_G \left(\exp_G(X) \exp_G(Y) \right), \tag{4.36}$$

and we would like to express $X * Y$ in terms of X and Y, this is the aim of the Baker-Campbell-Hausdorff formula.

Firstly we consider the case $G = \mathrm{GL}_n(\mathbb{R})$. In this case we have seen that the exponential function coincides with the exponential of matrix. Its inverse is the matrix logarithm:

$$\log(X) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} (X - \mathrm{id})^{k+1}$$

defined for $\|X - \mathrm{id}\| < 1$. For $\|X\| < \log 2$ we have $\log(e^X) = X$.

Lemma 4.3.3. *Let define the following convergent series*

$$\Phi(z) = \frac{1 - e^{-z}}{z} = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{z^{k-1}}{k!} \qquad z \in \mathbb{C} \tag{4.37}$$

$$\Psi(z) = \frac{z \log z}{z - 1} = z \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} (z-1)^k \qquad |z - 1| < 1\,. \tag{4.38}$$

Then we have $\Psi(e^z)\Phi(z) = 1$ for $z \in \mathbb{C}$ such that $|z| < \log 2$.

Proof. Since $\log(e^z) = z$, we get

$$\Psi(e^z)\Phi(z) = \frac{e^z z}{e^z - 1} \frac{1 - e^{-z}}{z} = 1\,.$$

□

Remark 4.3.4. The same identity of lemma 4.3.3 holds also if we insert matrix inputs $L \in \mathrm{End}\big(M_n(\mathbb{R})\big)$ such that $\|L\| < \log 2$. In which case we obtain

$$\Psi(\exp L)\Phi(L) = \mathrm{id}_{M_n(\mathbb{R})}\,. \tag{4.39}$$

Proposition 4.3.5. *Let $X \in M_n(\mathbb{R})$, then*

$$\mathrm{d}\exp(X) = L_{\exp(X)} \circ \Phi(\mathrm{ad}X) \colon M_n(\mathbb{R}) \to M_n(\mathbb{R})\,. \tag{4.40}$$

Proof. Let $\alpha \colon [0, 1] \to M_n(\mathbb{R})$ be a differentiable curve. Then we define

$$\gamma(t, s) = \exp(-s\alpha(t)) \frac{\mathrm{d}}{\mathrm{d}t} \exp(s\alpha(t)) \tag{4.41}$$

which is a $\mathcal{C}^1$ map $[0,1]^2 \to M_n(\mathbb{R})$ and $\gamma(t,0) = 0$ for each t. We calculate

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\gamma(t,s) =&\exp(-s\alpha(t))(-\alpha(t))\frac{\mathrm{d}}{\mathrm{d}t}\exp(s\alpha(t))+\\ &+\exp(-s\alpha(t))\frac{\mathrm{d}}{\mathrm{d}t}\big(\alpha(t)\exp(s\alpha(t))\big)\\ =&\exp(-s\alpha(t))(-\alpha(t))\frac{\mathrm{d}}{\mathrm{d}t}\exp(s\alpha(t))+\\ &+\exp(-s\alpha(t))\left(\alpha'(t)\exp(s\alpha(t))+\alpha(t)\frac{\mathrm{d}}{\mathrm{d}t}\exp(s\alpha(t))\right)\\ =&\mathrm{Ad}\big(\exp(-s\alpha(t))\big)\alpha'(t) = e^{-s\,\mathrm{ad}\,\alpha(t)}\alpha'(t)\,.\end{aligned}$$

We integrate over $[0,1]$ with respect to s and we obtain

$$\gamma(t,1) = \gamma(t,0) + \int_0^1 e^{-s\,\mathrm{ad}\,\alpha(t)}\alpha'(t)\,\mathrm{d}s = \int_0^1 e^{-s\,\mathrm{ad}\,\alpha(t)}\alpha'(t)\,\mathrm{d}s\,.$$

For $X \in M_n(\mathbb{R})$ we notice that

$$\begin{aligned}\int_0^1 e^{-s\,\mathrm{ad}\,X}\,\mathrm{d}s &= \int_0^1\sum_{k=0}^{\infty}\frac{(-\mathrm{ad}\,X)^k}{k!}s^k\,\mathrm{d}s = \sum_{k=0}^{\infty}(-\mathrm{ad}\,X)^k\int_0^1\frac{s^k}{k!}\,\mathrm{d}s\\ &= \sum_{k=0}^{\infty}\frac{(-\mathrm{ad}\,X)^k}{(k+1)!} = \Phi(\mathrm{ad}\,X)\,.\end{aligned}$$

If we consider $\alpha(t) = X + tY$, then $\alpha(0) = X$ and $\alpha'(0) = Y$, thus we obtain

$$\exp(-x)\,\mathrm{d}\exp(X)Y = \gamma(0,1) = \int_0^1 e^{-s\,\mathrm{ad}\,X}Y\,\mathrm{d}s = \Phi(\mathrm{ad}\,X)Y\,,$$

therefore $\mathrm{d}\exp(X) = L_{\exp(X)} \circ \Phi(\mathrm{ad}\,X)$. □

We now define $F(t) = \log(\exp(X)\exp(tY))$ defined for $t \in [0,1]$ and X, Y small enough. The idea is to obtain the BCH formula integrating $F'(t)$ as in the proof of proposition 4.3.5. Firstly we compute

$$\begin{aligned}\mathrm{d}\exp(F(t))F'(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\exp(F(t)) = \frac{\mathrm{d}}{\mathrm{d}t}\exp(X)\exp(tY)\\ &= \big(\exp(X)\exp(tY)\big)Y = \big(\exp(F(t))\big)Y\end{aligned}$$

and using proposition 4.3.5 we obtain

$$\begin{aligned}Y &= \big(\exp(F(t))\big)^{-1}\mathrm{d}\exp\big(F(t)\big)F'(t)\\ &= \Phi(\mathrm{ad}\,F(t))F'(t)\,.\end{aligned}$$

Finally, for X, Y small enough, then also $\|\mathrm{ad}\,F(t)\|$ is small enough and we can apply lemma 4.3.3:

$$F'(t) = \Psi\big(\exp(\mathrm{ad}\,F(t))\big)Y\,. \tag{4.42}$$

Proposition 4.3.6. *For $X, Y \in M_n(\mathbb{R})$ such that $\|X\|$ and $\|Y\|$ are small enough, then*

$$\log(\exp(X)\exp(Y)) = X + \int_0^1 \Psi\big(\exp(\operatorname{ad} X)\exp(t\operatorname{ad} Y)\big)Y\,\mathrm{d}t\,. \tag{4.43}$$

Proof. Using eq. (4.42) and eq. (4.34) we obtain

$$\begin{aligned} F'(t) &= \Psi\big(\exp(\operatorname{ad} F(t))\big)Y = \Psi\big(\operatorname{Ad}(\exp\, F(t))\big)Y \\ &= \Psi\big(\operatorname{Ad}(\exp X \cdot \exp tY)\big)Y = \Psi\big(\operatorname{Ad}(\exp X)\operatorname{Ad}(\exp tY)\big)Y \\ &= \Psi\big(\exp(\operatorname{ad} X)\exp(\operatorname{ad} tY)\big)Y\,. \end{aligned}$$

Moreover $F(0) = \log(\exp X) = X$, So that we can integrate from 0 to 1:

$$\log(\exp X \cdot \exp Y) = X + \int_0^1 \Psi\big(\exp(\operatorname{ad} X)\exp(\operatorname{ad} tY)\big)Y\,\mathrm{d}t\,. \qquad \square$$

We can expand the integral in proposition 4.3.6:

$$\begin{aligned} &\int_0^1 \Psi\big(\exp(\operatorname{ad} X)\exp(t\operatorname{ad} Y)\big)Y\,\mathrm{d}t \\ &= \int_0^1 \sum_{k=0}^{\infty} \frac{(-1)^k\big(\exp(\operatorname{ad} X)\exp(\operatorname{ad} tY) - \mathrm{id}\big)^k}{k+1}\big(\exp(\operatorname{ad} X)\exp(\operatorname{ad} tY)\big)Y\,\mathrm{d}t \\ &= \int_0^1 \sum_{\substack{k\geq 0\\ p_i+q_i>0}} \frac{(-1)^k(\operatorname{ad} X)^{p_1}(\operatorname{ad} tY)^{q_1}\cdots(\operatorname{ad} X)^{p_k}(\operatorname{ad} tY)^{q_k}}{(k+1)p_1!q_1!\cdots p_k!q_k!}\exp(\operatorname{ad} X)Y\,\mathrm{d}t \\ &= \sum_{\substack{k,m\geq 0\\ p_i+q_i>0}} \frac{(-1)^k(\operatorname{ad} X)^{p_1}(\operatorname{ad} Y)^{q_1}\cdots(\operatorname{ad} X)^{p_k}(\operatorname{ad} Y)^{q_k}(\operatorname{ad} X)^m}{(k+1)p_1!q_1!\cdots p_k!q_k!m!}Y\int_0^1 t^{\sum q_i}\,\mathrm{d}t \\ &= \sum_{\substack{k,m\geq 0\\ p_i+q_i>0}} \frac{(-1)^k(\operatorname{ad} X)^{p_1}(\operatorname{ad} Y)^{q_1}\cdots(\operatorname{ad} X)^{p_k}(\operatorname{ad} Y)^{q_k}(\operatorname{ad} X)^m}{(k+1)(q_1+\cdots+q_k+1)p_1!q_1!\cdots p_k!q_k!m!}Y\,. \end{aligned}$$

We are now ready to generalise this result to abstract Lie groups.

Definition 4.3.7. Let G be a Lie group and M a differentiable manifold, the *logarithmic derivative* of a differentiable function $f\colon M \to G$ is a differentiable 1-form with values in $\mathfrak{g} = \mathbf{L}(G)$, that is $\delta(f) \in \Lambda^1(M, \mathfrak{g})$, such that

$$\delta(f)_m v = (\mathrm{d}L_{f(m)^{-1}})\mathrm{d}f_m v \quad \forall m \in M \quad \forall v \in T_m(M)\,. \tag{4.44}$$

If $\alpha \in \Lambda^1(M, \mathfrak{g})$ and $f\colon M \to G$ differentiable, then we define $\operatorname{Ad}(f)\alpha \in \Lambda^1(M, \mathfrak{g})$ pointwise $\big(\operatorname{Ad}(f)\alpha\big)_m = \operatorname{Ad}(f(m))\alpha_m$.

Lemma 4.3.8. *For two differentiable maps $f, h\colon M \to G$ the logarithmic derivative of the pointwise product fh is given by*

$$\delta(fh) = \delta(h) + \operatorname{Ad}(h^{-1})\delta(f) \tag{4.45}$$

where $h^{-1} = \iota_G \circ h$.

Proof. We rephrase $fh = m_G \circ (f, h)$ so that we get

$$\mathrm{d}(fh) = \mathrm{d}R_h \circ \mathrm{d}f + \mathrm{d}L_f \circ \mathrm{d}h \colon TM \to TG\,,$$

then

$$\begin{aligned}\delta(fh) &= \mathrm{d}L_{(fh)^{-1}} \circ \mathrm{d}(fh) = \mathrm{d}L_{h^{-1}} \circ \mathrm{d}L_{f^{-1}} \circ \big(\mathrm{d}R_h \circ \mathrm{d}f + \mathrm{d}L_f \circ \mathrm{d}h\big) \\ &= \mathrm{d}L_{h^{-1}} \circ \mathrm{d}R_h \circ (\mathrm{d}L_{f^{-1}} \circ \mathrm{d}f) + \mathrm{d}L_{h^{-1}} \circ \mathrm{d}h \\ &= \mathrm{Ad}(h^{-1})\delta(f) + \delta(h)\,. \end{aligned}$$ □

Proposition 4.3.9. *The logarithmic derivative of* $\exp_G$ *is given by*

$$\delta(\exp_G)_X = \Phi(\mathrm{ad}\, X) \colon \mathfrak{g} = \mathbf{L}(G) \to \mathfrak{g}\,. \tag{4.46}$$

Proof. We fix $t, s \in \mathbb{R}$ and we define the functions $f, f_t, f_s \colon \mathbf{L}(G) \to G$ as

$$f(X) = \exp_G((t+s)X) \quad f_t(X) = \exp_G(tX) \quad f_s = \exp_G(sX)\,,$$

which satisfy $f = f_t f_s$ pointwise on $\mathbf{L}(G)$. By lemma 4.3.8 we get

$$\delta(f) = \delta(f_s) + \mathrm{Ad}(f_s)^{-1}\delta(f_t)\,.$$

Now we define the smooth curve $\psi \colon \mathbb{R} \to \mathbf{L}(G)$, $\psi = \delta(\exp_G)_{tX}(tY)$, thus we obtain

$$\begin{aligned}\psi(t+s) &= \delta(f)_X(Y) = \delta(f_s)_X(Y) + \mathrm{Ad}(f_s)^{-1}\delta(f_t)_X(Y) \\ &= \psi(s) + \mathrm{Ad}(\exp_G(-sX))\psi(t)\,.\end{aligned}$$

We have $\psi(0) = 0$ and

$$\psi'(0) = \lim_{t\to 0} \delta(\exp_G)_{tX}(Y) = \delta(\exp_G)_0(Y) = Y\,.$$

So we take derivatives with respect to t at $t = 0$ and we obtain

$$\psi'(s) = \mathrm{Ad}(\exp_G(-sX))Y = e^{-\mathrm{ad}\, sX} Y\,.$$

Now, if we proceed as in the proof of proposition 4.3.5, with an integration

$$\delta(\exp_G)_X(Y) = \psi(1) = \int_0^1 e^{-\mathrm{ad}\, sX} Y \,\mathrm{d}s = \Phi(\mathrm{ad}\, X)Y\,.$$ □

We now consider a neighbourhood U of $0 \in \mathbf{L}(G)$ such that $\exp_G|_{U_0}$ is a diffeomorphism and another neighbourhood V of 0 such that

$$\exp_G(V)\exp_G(V) \subseteq \exp_G(U)\,, \tag{4.47}$$

and we can define, for $X, Y \in V$:

$$X * Y = \log_U(\exp_G(X)\exp_G(Y))\,. \tag{4.48}$$

The curve $F(t) = X * tY$ satisfies $\exp_G(F(t)) = \exp_G(X)\exp_G(tY)$, if we consider the logarithmic derivative of both sides (as functions $\mathbb{R} \to G$) we get

$$\delta(\exp_G)_{F(t)} F'(t) = Y\,. \tag{4.49}$$

We can restrict U such that eq. (4.39) holds. If we recall also proposition 4.3.9 we obtain

$$F'(t) = \Psi\big(e^{\operatorname{ad} F(t)}\big) Y \tag{4.50}$$

and we can follow the same argument of proposition 4.3.6 which leads us to the following:

Proposition 4.3.10. *Let G a Lie group, then there exists a neighbourhood $V \subseteq \mathbf{L}(G)$ of 0 such that, for $X, Y \in V$ the Hausdorff series*

$$X * Y = X + \sum_{\substack{k,m \geq 0 \\ p_i + q_i > 0}} \frac{(-1)^k (\operatorname{ad} X)^{p_1} (\operatorname{ad} Y)^{q_1} \cdots (\operatorname{ad} X)^{p_k} (\operatorname{ad} Y)^{q_k} (\operatorname{ad} X)^m}{(k+1)(q_1 + \cdots + q_k + 1) p_1! q_1! \cdots p_k! q_k! m!} Y \tag{4.51}$$

converges and satisfies

$$\exp_G(X * Y) = \exp_G(X)\exp_G(Y)\,. \tag{4.52}$$

4.4 Covering theory for Lie groups

The study of coverings for Lie groups leads to a complete description of Lie groups classification and reduces the problem to the classification of finite-dimensional Lie algebras. The existence of a universal covering space for topological manifolds is a well-known fact in homotopy theory, the details are discussed in [HN12, appendice A]. In the case of Lie groups we can extend these results further.

Theorem 4.4.1. *If G is a connected Lie group and $\pi_G \colon \widetilde{G} \to G$ its universal covering map, then $\widetilde{G}$ carries a unique Lie group structure for which π_G is a Lie group morphism. We call this Lie group $\widetilde{G}$* the universal covering group *of G.*

Ado's theorem (see [HN12, chapter 7]) ensures that each finite-dimensional Lie algebra is isomorphic to a proper subalgebra of some matrix Lie algebra $\mathfrak{gl}(V)$ where V is a finite dimensional vector space. Moreover if G is a Lie group and $\mathfrak{h} \subseteq \mathbf{L}(G)$ is a subalgebra of its Lie algebra, then the subgroup generated by the elements of $\{\exp_G(X) \mid X \in \mathfrak{h}\}$, referred as $\langle \exp_G(\mathfrak{h}) \rangle$, can be endowed with a Lie group structure such that its Lie algebra is $\mathfrak{h}$. This procedure lead us to the following result, known as Lie's third theorem.

Theorem 4.4.2. *Each finite-dimensional Lie algebra $\mathfrak{g}$ is the Lie algebra of a connected Lie group G.*

Proposition 4.4.3. *A surjective morphism $\varphi \colon G \to H$ of Lie groups is a covering map (we recall definition 1.1.11) if and only if $\mathbf{L}(\varphi) \colon \mathbf{L}(G) \to \mathbf{L}(H)$ is a linear isomorphism.*

Theorem 4.4.4. *Let $\pi\colon G \to H$ be a covering morphism of Lie groups. If $f\colon L \to H$ is a morphism of Lie groups and L is simply connected, then there exist a unique lift $\widehat{f}\colon L \to G$ which is a morphism of Lie groups.*

Theorem 4.4.5. *Let G be a connected Lie group and $\pi_G\colon \widetilde{G} \to G$ a universal covering homomorphism. Then $\ker \pi_G$ is a discrete central subgroup (that is contained in the center of $\widetilde{G}$) and $G \cong \widetilde{G}/\ker \pi_G$.*

Moreover, for any discrete central subgroup $\Gamma \subseteq \widetilde{G}$, the group $\widetilde{G}/\Gamma$ is a connected Lie group with the same universal covering group as G.

Proof. π_G is a covering map, hence a local diffeomorphism, then for each $g \in \widetilde{G}$ there is a neighbourhood U of g that is diffeomorphic to $\pi_G(U)$, therefore $\ker \pi_G \cap U = \{\, g \,\}$. Moreover $\ker \pi_G$ is a normal subgroup. Indeed let $h \in \ker \pi_G$, since $\widetilde{G}$ is arc-wise connected, we consider for each $g \in \widetilde{G}$ a smooth curve γ_g such that $\gamma(0) = \mathbf{1}$ and $\gamma(1) = g$. Then

$$\gamma(t)h\gamma(t)^{-1}$$

is a smooth curve contained on $\ker \pi_G$ (since it is normal) and connecting h with ghg^{-1}. Then this curve must be constant since $\ker \pi_G$ is discrete, therefore $ghg^{-1} = h$ for every $h \in \ker \pi_G$ and $g \in \widetilde{G}$, that is $\ker \pi_G$ is central.

Since π_G is surjective, then $G \cong \widetilde{G}/\ker \pi_G$ as groups. Moreover π_G is also an open map and it is easy to see that, in this case, also the induced map $\widetilde{\pi}\colon \widetilde{G}/\ker \pi_G \to G$ is bicontinuous. Continuous homomorphisms between Lie groups are automatically smooth, therefore $G \cong \widetilde{G}/\ker \pi_G$ as Lie groups.

Finally, if $\Gamma \subseteq \widetilde{G}$ is a discrete central subgroup, then it normal and the image of a smooth curve connecting to point in $\widetilde{G}$ via the quotient map is the image of a smooth curve connecting two points in $\widetilde{G}/\Gamma$. Since the quotient map is surjective, this proves that $\widetilde{G}/\Gamma$ is connected. Since the quotient map via a discrete central subgroup is a local diffeomorphism, then $\widetilde{G}$ is the universal covering group of $\widetilde{G}/\Gamma$. $\square$

Theorem 4.4.6. *Let G be a connected and simply connected Lie group, H another Lie group and $\psi\colon \mathbf{L}(G) \to \mathbf{L}(H)$ a Lie algebra morphism. Then there exists a unique morphism $\varphi\colon G \to H$ with $\mathbf{L}(\varphi) = \psi$.*

Theorem 4.4.7. *Two connected Lie group G and H have isomorphic Lie algebras if and only if their universal covering groups $\widetilde{G}$ and $\widetilde{H}$ are isomorphic as Lie groups.*

Proof. If $\widetilde{G}$ and $\widetilde{H}$ are isomorphic, then using proposition 4.4.3

$$\mathbf{L}(G) \cong \mathbf{L}(\widetilde{G}) \cong \mathbf{L}(\widetilde{H}) \cong \mathbf{L}(H)\,.$$

Conversely, let $\psi\colon \mathbf{L}(G) \to \mathbf{L}(H)$ be an isomorphism. By theorem 4.4.6 there is a unique morphism morphism $\varphi\colon \widetilde{G} \to \widehat{H}$ such that $\mathbf{L}(\varphi) = \psi$ and also a unique morphism $\widehat{\varphi}\colon \widetilde{H} \to \widetilde{G}$ with $\mathbf{L}(\widehat{\varphi}) = \psi^{-1}$. Then $\mathbf{L}(\varphi \circ \widehat{\varphi}) = \mathrm{id}_{\mathbf{L}(\widetilde{G})}$ and theorem 4.4.6 implies $\varphi \circ \widehat{\varphi} = \mathrm{id}_{\widetilde{G}}$, and similarly $\widehat{\varphi} \circ \varphi = \mathrm{id}_{\widetilde{H}}$. Therefore $\widetilde{G}$ and $\widetilde{H}$ are isomorphic as Lie groups. $\square$

If we combine the previous result with theorem 4.4.5, we conclude:

Theorem 4.4.8. *Let G be a connected Lie group and $\pi_G\colon \widetilde{G} \to G$ the universal covering morphism of connected Lie groups. Then for each discrete central subgroup $\Gamma \subseteq \widetilde{G}$, the group $\widetilde{G}/\Gamma$ is a connected Lie group with $\mathbf{L}(\widetilde{G}/\Gamma) \cong \mathbf{L}(G)$.*

Conversely, each Lie group with the same Lie algebra as G is isomorphic to $\widetilde{G}/\Gamma$ for some Γ discrete central subgroup of $\widetilde{G}$.

These theorems together states that one can define a functor from the category of finite-dimensional Lie algebras to the category of connected and simply connected Lie groups that associate to a Lie algebra the unique (up to isomorphism) connected and simply connected Lie group with that Lie algebra.

4.5 Structure of nilpotent Lie groups

Definition 4.5.1. Let G be a group. For two subsets $A, B \subseteq G$ we define

$$[A,B] = \langle aba^{-1}b^{-1} \mid a \in A, b \in B \rangle \ . \tag{4.53}$$

If we set inductively

$$C^0(G) = G \qquad C^{n+1} = [G, C^n(G)]\,, \tag{4.54}$$

then $\big(C^n(G)\big)_n$ is called the *lower central series* of G. A group G is said to be *nilpotent* if $C^d(G) = \{\,\mathbf{1}\,\}$ for some $d \in \mathbb{N}$. We define again inductively

$$D^0(G) = G \qquad D^{n+1} = [D^n(G), D^n(G)] \tag{4.55}$$

and $\big(D^n(G)\big)_n$ is called *derived series* of G. A group G is said to be *solvable* if $D^k(G) = \{\,\mathbf{1}\,\}$ for some $k \in \mathbb{N}$.

Proposition 4.5.2. *If G is a Lie group and $\mathfrak{g}, \mathfrak{h} \subseteq \mathbf{L}(G)$ are subalgebras of $\mathbf{L}(G)$, then*

$$\langle \exp_G[\mathfrak{g},\mathfrak{h}]\rangle = [\langle \exp_G \mathfrak{g}\rangle \,, \langle \exp_G \mathfrak{h}\rangle]\,. \tag{4.56}$$

An immediate consequence of the previous proposition is the following theorem that characterise nilpotent and solvable Lie groups.

Theorem 4.5.3. *A connected Lie group G is abelian, nilpotent or solvable if and only if its Lie algebra is abelian, nilpotent or solvable respectively.*

Proposition 4.5.4. *Let G be a Lie group with Lie algebra $\mathbf{L}(G)$. Then for $X, Y \in \mathbf{L}(G)$:*

$$\exp_G([X,Y]) = \lim_{k\to\infty} \big(\exp_G(\tfrac{1}{k}X)\exp_G(\tfrac{1}{k}Y)\exp_G(-\tfrac{1}{k}X)\exp_G(-\tfrac{1}{k}Y)\big)^{k^2} \,. \tag{4.57}$$

Proof. For k big enough we have

$$\begin{aligned}\big(\exp_G(\tfrac{1}{k}X)\exp_G(\tfrac{1}{k}Y)\exp_G(-\tfrac{1}{k}X)\exp_G(-\tfrac{1}{k}Y)\big)^{k^2} = \\ = \exp_G\left(k^2\big(\tfrac{1}{k}X * \tfrac{1}{k}Y * -\tfrac{1}{k}X * -\tfrac{1}{k}Y\big)\right),\end{aligned}$$

so, since the exponential map is continuous, it is sufficient to prove that

$$\lim_{k\to\infty} k^2\big(\tfrac{1}{k}X * \tfrac{1}{k}Y * -\tfrac{1}{k}X * -\tfrac{1}{k}Y\big) = [X,Y]\,. \tag{4.58}$$

Now we define $F(t,s) = tX * sY * (-tX) * (-sY)$ and we recall eq. (1.47). Therefore, since $\exp_G$ is a local diffeomorphism at zero and $(\mathrm{d}\exp_G)_0 = \mathrm{id}_{\mathfrak{g}}$, we can compute

$$\begin{aligned}\frac{\partial^2 F}{\partial s\partial t}(0,0) &= \frac{\partial^2}{\partial s\partial t}\bigg|_{(0,0)} \exp_G\big(F(s,t)\big)\\ &= \frac{\partial^2}{\partial s\partial t}\bigg|_{(0,0)} \exp_G(tX)\exp_G(sY)\exp_G(-tX)\exp_G(-sY)\\ &= \frac{\partial^2}{\partial s\partial t}\bigg|_{(0,0)} \exp(-sY)\circ\exp(-tX)\circ\exp(sY)\circ\exp(tX)(\mathbf{1})\\ &= [X,Y](\mathbf{1})\,.\end{aligned}$$

Moreover $F(0,0) = 0$ and

$$\frac{\partial F}{\partial t}(0,0) = \frac{\partial F}{\partial s}(0,0) = \frac{\partial^2 F}{\partial t^2}(0,0) = \frac{\partial^2 F}{\partial s^2}(0,0) = 0\,.$$

If we define $f(t) = F(t,t)$, then $f(t)$ is smooth and

$$\begin{aligned} f'(t) &= \frac{\partial F}{\partial t}(t,t) + \frac{\partial F}{\partial s}(t,t) \quad\Rightarrow\quad f'(0) = 0\\ f''(t) &= \frac{\partial^2 F}{\partial t^2}(t,t) + 2\frac{\partial^2 F}{\partial s\partial t}(t,t) + \frac{\partial^2 F}{\partial s^2}(t,t) \quad\Rightarrow\quad f''(0) = 2[X,Y]\,.\end{aligned}$$

Finally

$$\lim_{k\to\infty} k^2\big(\tfrac{1}{k}X * \tfrac{1}{k}Y * -\tfrac{1}{k}X * -\tfrac{1}{k}Y\big) = \lim_{k\to\infty} k^2 f\left(\frac{1}{k}\right) = \frac{1}{2}f''(0) = [X,Y]\,. \qquad \square$$

Theorem 4.5.5. *If $\mathfrak{g}$ is a nilpotent Lie algebra, then the Hausdorff series eq.* (4.51) *defines a polynomial map*

$$*\colon \mathfrak{g}\times\mathfrak{g}\to\mathfrak{g} \qquad (X,Y)\mapsto X + Y + \frac{1}{2}[X,Y] + \cdots\,. \tag{4.59}$$

We thus obtain a Lie group structure $(\mathfrak{g}, *)$ *with* $\exp_{\mathfrak{g}} = \mathrm{id}_{\mathfrak{g}}$ *and* $\mathbf{L}(\mathfrak{g}, *) = \mathfrak{g}$.

Proof. Since $\mathfrak{g}$ is nilpotent, then $C^d(\mathfrak{g}) = \{\,0\,\}$ for some $d\in\mathbb{N}$. Hence the terms of order grater that d vanish and only finitely many terms remain. Therefore, $*\colon \mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$ is a polynomial map. Moreover

$$0 * x = x * 0 = x \qquad x * (-x) = (-x) * x = 0 \quad \forall x\in\mathfrak{g}\,.$$

Let us consider a group G whose Lie algebra is $\mathfrak{g}$, it exists by theorem 4.4.2. By proposition 4.3.1 there exists an open neighbourhood $U\subseteq\mathfrak{g}$ of zero such that

$$\exp_G(X * Y) = \exp_G(X)\exp_G(Y) \quad \forall X, Y\in U$$

and $\exp_G\big|_U$ is a diffeomorphism onto an open set of G. We now consider another open neighbourhood $V \subseteq U$ of zero such that

$$\big(V * V\big) * V \cup V * \big(V * V\big) \subseteq U\,.$$

For $X, Y, Z \in V$, we thus have $X * Y, Y * Z \in U$ and

$$\exp_G\big(X * (Y * Z)\big) = \exp_G(X)\exp_G(Y)\exp_G(Z) = \exp_G\big((X * Y) * Z\big)\,,$$

and by the injectivity of $\exp_G\big|_U$ we obtain

$$X * (Y * Z) = (X * Y) * Z \quad \forall X, Y, Z \in V\,.$$

We then observe that both maps

$$(X, Y, Z) \mapsto X * (Y * Z) \qquad (X, Y, Z) \mapsto (X * Y) * Z$$

are polynomial, since $\mathfrak{g}$ is nilpotent, and they coincides on a non-empty open set V, hence they are equal on all $\mathfrak{g}^3$. This implies that $*$ is associative, therefore $(\mathfrak{g}, *)$ is a Lie group with identity 0 and $X^{-1} = -X$.

Hereafter we identify $T_0(\mathfrak{g})$ with $\mathfrak{g}$ and we consider the exponential function

$$\exp_{(\mathfrak{g},*)} \colon \mathbf{L}(\mathfrak{g}, *) \cong \big(\mathfrak{g}, [\cdot,\cdot]\big) \to (\mathfrak{g}, *)\,.$$

The explicit form of the Hausdorff series grants the relation $tX * sX = (t+s)X$ for any $t, s \in \mathbb{R}$ and $X \in \mathfrak{g}$, hence

$$\exp_{(\mathfrak{g},*)}(X) = X \quad \forall X \in \mathfrak{g}\,,$$

and $\exp_{(\mathfrak{g},*)} = \mathrm{id}_{\mathfrak{g}}$. Finally, the Lie bracket in $\mathbf{L}(\mathfrak{g}, *)$ can be calculated with proposition 4.5.4:

$$[X, Y]_{\mathbf{L}(\mathfrak{g},*)} = \lim_{k\to\infty} k^2\big(\tfrac{1}{k}X * \tfrac{1}{k}Y * -\tfrac{1}{k}X * -\tfrac{1}{k}Y\big)$$

and if we compute explicitly $k^2\big(\tfrac{1}{k}X * \tfrac{1}{k}Y * -\tfrac{1}{k}X * -\tfrac{1}{k}Y\big)$, we notice that each term with more than one bracket appears with a factor k^{-n} with $n \geq 1$ and they tend to 0 uniformly as $k \to \infty$ (since there are only finitely many terms). Precisely we obtain

$$\begin{aligned} k^2\left(\frac{1}{k}X * \frac{1}{k}Y * -\frac{1}{k}X * -\frac{1}{k}Y\right) &= \cdots = \\ = k^2\left(\frac{1}{k}X + \frac{1}{k}Y - \frac{1}{k}X - \frac{1}{k}Y + \frac{1}{2k^2}[X, Y] + \frac{1}{2k^2}[X, Y] + \mathcal{O}(k^{-3})\right) \end{aligned}$$

and we conclude $[X, Y]_{\mathbf{L}(\mathfrak{g},*)} = [X, Y]_{\mathfrak{g}}$, that is $\mathfrak{g} \cong \mathbf{L}(\mathfrak{g}, *)$. □

Corollary 4.5.6. *Let G be a connected nilpotent Lie group with Lie algebra $\mathfrak{g}$. Then $\mathfrak{g}$ is nilpotent and $\exp_G \colon (\mathfrak{g}, *) \to G$ is the universal covering morphism of G. In particular, the exponential function of G is surjective.*

Proof. From theorem 4.5.3 we know that $\mathfrak{g}$ is nilpotent and we know from theorem 4.5.5 that $(\mathfrak{g},*)$ is a simply connected group with Lie algebra $\mathfrak{g}$, and it is unique up to isomorphism. Let π_G the unique morphism of Lie group with $\mathbf{L}(\pi_G) = \mathrm{id}_{\mathfrak{g}}$ (theorem 4.4.6). Then

$$\pi_G(X) = \pi_G\big(\exp_{(\mathfrak{g},*)}(X)\big) = \exp_G\big(\mathbf{L}(\pi_G)X\big) = \exp_G(X)\,,$$

which implies $\pi_G = \exp_G$. □

Lemma 4.5.7. *If $\mathfrak{g}$ is a nilpotent Lie algebra, then the center of the group $(\mathfrak{g},*)$ coincide with the center $\mathfrak{z}(\mathfrak{g})$ of the Lie algebra $\mathfrak{g}$.*

Proof. From the definition of $*$ in terms of Hausdorff series, we see that $X*Z = X+Z = Z*X$ whenever $Z \in \mathfrak{z}(\mathfrak{g})$, which leads to $\mathfrak{z}(\mathfrak{g}) \subseteq Z(\mathfrak{g},*)$. If conversely, $Z \in Z(\mathfrak{g},*)$, then $\mathrm{id}_{\mathfrak{g}} = \mathrm{Ad}(\exp_{(\mathfrak{g},*)} Z) = e^{\mathrm{ad}\, Z}$ and, since the exponential function is injective on the set of nilpotent endomorphism in $\mathrm{End}(\mathfrak{g})$, then $\mathrm{ad}\, Z = 0$ which means $Z \in \mathfrak{z}(\mathfrak{g})$. □

Proposition 4.5.8. *If G is a connected nilpotent Lie group with Lie algebra $\mathfrak{g}$, then $Z(G) = \exp_G(\mathfrak{z}(\mathfrak{g}))$ is connected.*

Proof. Since $\exp_G\colon (\mathfrak{g},*) \to G$ is the universal covering morphism of G, we obtain with lemma 4.5.7

$$Z(G) = \ker \mathrm{Ad}_G = \exp_G(\ker \mathrm{Ad}_{(\mathfrak{g},*)}) = \exp_G(\mathfrak{z}(\mathfrak{g}))\,.$$ □

An immediate consequence of the previous proposition combined with theorem 4.4.8 is the following:

Theorem 4.5.9. *Let G be a connected Lie group with Lie algebra $\mathfrak{g}$. Then $\mathfrak{g}$ is nilpotent as a Lie algebra and there exists a discrete subgroup $\Gamma \subseteq (\mathfrak{z}(\mathfrak{g}),+)$ such that*

$$G \cong (\mathfrak{g},*)/\Gamma\,. \tag{4.60}$$

Corollary 4.5.10. *Any compact connected nilpotent Lie group is abelian.*

Proof. If G is compact, then by theorem 4.5.9 also $\mathfrak{g}/\Gamma$ is compact for some $\Gamma \subseteq \mathfrak{z}(\mathfrak{g})$. This implies that $\mathfrak{g} = \mathrm{span}\,\Gamma \subseteq \mathfrak{z}(\mathfrak{g})$. Therefore $\mathfrak{g}$ is abelian and, using theorem 4.5.3, also G is abelian. □

Chapter 5

Sard property in Carnot groups

In this final chapter we will apply the developed theory of sub-Riemannian geometry and Lie groups in the context of Carnot groups. They were introduced under this name in the 80s and they appear as fundamental examples in many different contexts as explained in [Don16]. We will introduce sub-Riemannian structures of Carnot groups and we will characterize their admissible curves, with a main focus on abnormal curves. Finally, we will introduce the abnormal set which is both the set where abnormal curves lie and the set of critical values for the end-point map. A main open problem in sub-Riemannian geometry is to determine whether or not the abnormal set has zero measure, we will provide a solution for free Carnot groups of step 2. In this framework we will provide an algebraic description for that abnormal set. We will mostly refer to [Le +16] and [OV19].

5.1 Stratified Lie algebras and groups

Definition 5.1.1. A *stratified Lie algebra of rank r and step s* is a Lie algebra $\mathfrak{g}$ which admits a direct-sum decomposition

$$\mathfrak{g} = V_1 \oplus \cdots \oplus V_s \tag{5.1}$$

where $\dim(V_1) = r$, $[V_1, V_k] = V_{k+1}$ for each $k = 1, \ldots, s-1$ and $[V_1, V_s] = 0$ (and of course $V_s \neq 0$). In this decomposition, V_k is called the *layer of k-th degree* and the non-zero elements of V_k are said to have *degree k*.

Remark 5.1.2. We observe that a stratified Lie algebra is always nilpotent, indeed if $\mathfrak{g}$ is a stratified Lie algebra of step s as above, we get

$$C^k(\mathfrak{g}) = V_{k+1} \oplus \cdots \oplus V_s \tag{5.2}$$

and therefore s il also the nilpotence degree of $\mathfrak{g}$.

Example 5.1.3. Given $r, s \in N$ we can construct a stratified Lie algebra as follows: we consider formal elements $e_1, \ldots, e_r$ and (the span of) all their formal iterated Lie brackets up to length s, modulo the anti-commutativity relation and the Jacobi identity.

Lie bracket is defined formally and a Lie bracket whose length is greater that s is set to zero. The direct-sum decomposition is given by $\mathfrak{g} = V_1 \oplus \cdots \oplus V_s$ where V_k is the span of all formal iterated brackets of $e_1, \ldots, e_r$ of length k. We will refer to such a Lie algebra as *free nilpotent Lie algebra of rank r and step s* (here the term "free" has not to be confused with the one used to define free Sub-Riemannian structures).

When the step is greater than 3 it is not obvious how to determine the dimension of a free nilpotent Lie algebra. A general formula can be found in [Reu93]. When $s = 2$ one obtains that a free nilpotent Lie algebra $\mathfrak{g}$ of rank r and step 2 is given by

$$\mathfrak{g} = \operatorname{span}\{ e_1, \ldots, e_r \} \oplus \operatorname{span}\{ [e_i, e_j] \mid 1 \leq i < j \leq r \} , \tag{5.3}$$

therefore $\dim(\mathfrak{g}) = r + \binom{r}{2} = \frac{r(r+1)}{2}$.

The following proposition ensures that the notions of step and rank for a stratified Lie algebra are intrinsic to the Lie algebra and does not depends on the chosen direct-sum decomposition.

Proposition 5.1.4. *Let $\mathfrak{g}$ be a stratified Lie algebra with two stratifications,*

$$\mathfrak{g} = V_1 \oplus \cdots \oplus V_s = W_1 \oplus \cdots \oplus W_t . \tag{5.4}$$

Then $s = t$ and there exist a Lie algebra automorphism $\varphi \colon \mathfrak{g} \to \mathfrak{g}$ such that $\varphi(V_k) = W_k$ for all $k = 1, \ldots, s$.

Proof. Both s and t are equal to the nilpotency degree of $\mathfrak{g}$, therefore they are equal. We consider the quotient mappings

$$\pi_k \colon C^k(\mathfrak{g}) \to C^k(\mathfrak{g})/C^{k+1}(\mathfrak{g})$$

and we observe that the restrictions of π_k to V_{k+1} and W_{k+1} are both isomorphism. Therefore, for $v \in V_{k+1}$ we can define $\varphi(v)$ as the unique element $w \in W_{k+1}$ such that $\pi_k(v) = \pi_k(w)$. Finally we can extend $\varphi \colon \mathfrak{g} \to \mathfrak{g}$ using the direct-sum decomposition. We clearly obtained a linear isomorphism and by construction $\varphi(V_k) = W_k$ for $k = 1, \ldots, s$.

We only need to prove that φ is a Lie algebra morphism, that is $\varphi([a, b]) = [\varphi(a), \varphi(b)]$ for all $a, b \in \mathfrak{g}$. We have

$$a = \sum_{i=1}^{s} a_i \quad \text{and} \quad b = \sum_{i=1}^{s} b_i$$

for some $a_i, b_i \in V_i$. Then

$$\varphi([a, b]) = \sum_{i=1}^{s} \sum_{j=1}^{s} \varphi([a_i, b_j]) \quad \text{and} \quad [\varphi(a), \varphi(b)] = \sum_{i=1}^{s} \sum_{j=1}^{s} [\varphi(a_i), \varphi(b_j)] ,$$

which means that we only need to prove $\varphi([a_i, b_j]) = [\varphi(a_i), \varphi(b_j)]$ for $a_i \in V_i$ and $b_j \in V_j$. We notice that $[a_i, b_j] \in V_{i+j}$ and $[\varphi(a_i), \varphi(b_j)] \in W_{i+j}$, therefore $\varphi([a_i, b_j]) = [\varphi(a_i), \varphi(b_j)]$ if and only if $[a_i, b_j] - [\varphi(a_i), \varphi(b_j)] \in C^{i+j+2}(\mathfrak{g})$. We have

$$\begin{aligned} [a_i, b_j] \in V_{i+j} ,\ [\varphi(a_i), \varphi(b_j)] \in W_{i+j} &\Longrightarrow [a_i - \varphi(a_i), b_j] \in C^{i+j+2}(\mathfrak{g}) \\ \varphi(a_i) \in W_i ,\ \varphi(b_j) - b_j \in C^{j+2}(\mathfrak{g}) &\Longrightarrow [\varphi(a_i), \varphi(b_j) - b_j] \in C^{i+j+2}(\mathfrak{g}) \in C^{i+j+2}(\mathfrak{g}) . \end{aligned}$$

Finally we get

$$[a_i, b_j] - [\varphi(a_i), \varphi(b_j)] = [a_i - \varphi(a_i), b_j] - [\varphi(a_i), \varphi(b_j) - b_j] \in C^{i+j+2}(\mathfrak{g}) . \qquad \square$$

Definition 5.1.5. Let $\mathfrak{g}$ be a stratified Lie algebra, for $\lambda > 0$ we define the *dilation of factor* λ as the unique linear map $\delta_\lambda \colon \mathfrak{g} \to \mathfrak{g}$ such that

$$\delta_\lambda(v) = \lambda^k v \qquad \text{for } v \in V_k .$$

Dilations are Lie algebra isomorphism, more precisely the family of dilations are a one-parameter group of Lie algebra isomorphism, that is $\delta_\lambda \circ \delta_\eta = \delta_{\lambda\eta}$.

Definition 5.1.6. A *stratified group* is a connected, simply connected Lie group $\mathbb{G}$ whose associated Lie algebra $\mathbf{L}(\mathbb{G})$ is stratified.

Remark 5.1.7. Stratified groups are connected and simply connected Lie groups with nilpotent Lie algebra, therefore for a stratified group $\mathbb{G}$ the exponential map

$$\exp : \mathbf{L}(\mathbb{G}) \to \mathbb{G} \tag{5.5}$$

defines a diffeomorphism and we can identify $\mathbb{G}$ with its Lie algebra thanks to theorem 4.5.5 and the group multiplication is given by the Hausdorff series eq. (4.51). Therefore $\mathbb{G} \cong \mathbb{R}^n$ for some $n \in \mathbb{N}$ and we can consider an *adapted basis* $\{ X_1, \dots, X_n \}$ for $\mathfrak{g}$, which means it is coherent with the stratification:

$$\underbrace{X_1, \dots, X_r}_{\text{basis for } V_1}, \underbrace{X_{r+1}, \dots, X_{r_2}}_{\text{basis for } V_2}, \dots, \underbrace{X_{r_{s-1}+1}, \dots, X_n}_{\text{basis for } V_s} . \tag{5.6}$$

We recall that $X_1, \dots, X_n$ also denotes left invariant vector fields.

For $g \in \mathbb{G}$ we can consider its *exponential coordinates of first type*:

$$g = (x_1, \dots, x_n) \iff g = \exp(x_1 X_1 + \cdots + x_n X_n) \tag{5.7}$$

and its *exponential coordinates of second type*:

$$g = (x_1, \dots, x_n) \iff g = \exp(x_n X_n) \cdots \exp(x_2 X_2) \exp(x_1 X_1) , \tag{5.8}$$

or, using flows of vector fields $g = \exp(x_1 X_1) \circ \cdots \circ \exp(x_n X_n)(\mathbf{1}_{\mathbb{G}})$. We observe that if $g \in \mathbb{G}$ has exponential coordinates of the second type $(x_1, \dots, x_n)$, then the exponential coordinates of the first time of $\mathfrak{g}$ equals the coordinates of $x_n X_n * \cdots * x_1 X_1$ with respect to the basis $\{ X_1, \dots, X_n \}$.

Remark 5.1.8. For stratified Lie algebras $\mathfrak{g}$ we defined dilations $\delta_\lambda \colon \mathfrak{g} \to \mathfrak{g}$. Let $\mathbb{G}$ be a Carnot groups with Lie algebra $\mathfrak{g}$, since Carnot groups are simply connected, from theorem 4.4.6 there exists a unique Lie group automorphism $\Delta_\lambda \colon \mathbb{G} \to \mathbb{G}$ such that $\mathbf{L}(\Delta_\lambda) = \delta_\lambda$.

To stress the identification between stratified Lie algebras and stratified Lie groups, we will again refer to the automorphism Δ_λ as δ_λ.

Example 5.1.9. The *n-th Heisenberg group* is the stratified group $\mathbb{H}^n$ whose Lie algebra stratification $\mathfrak{g} = V_1 \oplus V_2$ is given by

$$V_1 = \operatorname{span}\{ X_1, \ldots, X_n, Y_1, \ldots, Y_n \} \,, \qquad V_2 = \operatorname{span}\{ T \} \,,$$

and the only non-zero Lie brackets between generators are

$$[X_i, Y_i] = T \qquad \text{for } i = 1, \ldots, n \,.$$

Example 5.1.10. The *Engel group* is the stratified group $\mathbb{E}$ associated to the Lie algebra of step 3 $\mathfrak{g} = V_1 \oplus V_2 \oplus V_3$ where

$$V_1 = \operatorname{span}\{ X_1, X_2 \} \,, \quad V_2 = \operatorname{span}\{ X_3 \} \,, \quad V_3 = \operatorname{span}\{ X_4 \} \,,$$

and the only non-trivial Lie brackets are

$$[X_2, X_1] = X_3 \,, \quad [X_3, X_1] = X_4 \,, \quad [X_3, X_2] = 0 \,.$$

The Engel group can be represented in exponential coordinates of second type as $\mathbb{E} \cong \mathbb{R}^4$ where

$$X_1 = \partial_1 \,, \quad X_2 = \partial_2 - x_1 \partial_3 + \frac{x_1^2}{2} \partial_4 \quad X_3 = \partial_3 - x_1 \partial_4 \quad X_4 = \partial_4 \,.$$

5.2 Sub-Riemannian structure of Carnot groups

Definition 5.2.1. Let $\mathbb{G}$ be a stratified Lie group and $\mathbf{L}(\mathbb{G}) = V_1 \oplus \cdots \oplus V_s$ a stratification for its Lie algebra and $\{ X_1, \ldots, X_r \}$ a basis for V_1. We consider the left-invariant Euclidean bundle $\mathbf{U}$ with base $\mathbb{G}$:

$$\mathbf{U} = \bigcup_{g \in \mathbb{G}} (g, \mathrm{d}L_g(V_1)) \tag{5.9}$$

equipped with a left-invariant positively defined section $(\,\cdot\,|\,\cdot\,)$ of $S^k\mathbf{U}$ such that the basis $\{ X_1, \ldots, X_r \}$ is an orthonormal frame. Explicitly $(\,\cdot\,|\,\cdot\,)_{\mathbf{1}}$ is a scalar product on V_1 such that

$$(X_i | X_j) = \delta^i_j \qquad \text{for } 1 \le i, j \le r \tag{5.10}$$

and for $v, w \in \mathrm{d}L_g(V_1)$ we have $\mathrm{d}L_{g^{-1}}(v), \mathrm{d}L_{g^{-1}}(w) \in V_1$ and

$$(v|w) = (\mathrm{d}L_{g^{-1}}(v) \,|\, \mathrm{d}L_{g^{-1}}(w))_{\mathbf{1}} \,. \tag{5.11}$$

The pair $(\mathbf{U}, f)$ where $f \colon \mathbf{U} \to T\mathbb{G}$ is the canonical inclusion is a (free) sub-Riemannian structure on $\mathbb{G}$, indeed the vector fields $X_1, \ldots, X_r$ are global generators for the horizontal distribution $f(\mathbf{U})$ and they are bracket-generating at every point $g \in \mathbb{G}$. A stratified Lie group $\mathbb{G}$ with such a sub-Riemannian structure is called *Carnot group*.

Remark 5.2.2. Both $T\mathbb{G}$ and the previously defined Euclidean bundle $\mathbf{U}$ are globally trivializable with the identifications between fibers given by $\mathrm{d}L_g$, then we can consider controls $u \in L^\infty([0,T], V_1)$ and the associated left-invariant non-autonomous vector fields

$$X_u(t) = \sum_{i=1}^{r} u_i(t) X_i\,. \tag{5.12}$$

By theorem 2.1.8, $X_u(t)$ admits a unique integral curve γ_u starting from a given $g \in \mathbb{G}$ and defined in a sufficiently small right-neighbourhood of 0. Since the horizontal vector fields $X_1, \dots, X_r$ are left invariant, an argument similar to proposition 4.2.6 concludes that the integral curve integral curve γ_u can be defined in the entire interval $[0,T]$ where T is arbitrary large. Moreover, since f is injective on fibers, u is the unique control associated to γ_u.

Definition 5.2.3. We will call *free Carnot group* whose Lie algebra is free nilpotent. The terminology must not be confused with the one used for free sub-Riemannian structures.

Proposition 5.2.4. *Let $\mathbb{G}$ be a Carnot group and d its sub-Riemannian distance. Then*

(i) d is left-invariant, that is $d(g\cdot x, g\cdot y) = d(x,y)$ for every $x,y,g \in G$,

(ii) if δ_λ is a dilation of factor $\lambda > 0$, then $d(\delta_\lambda(x), \delta_\lambda(y)) = \lambda d(x,y))$ for every $x,y \in \mathbb{G}$.

Proof. (i) If $\gamma\colon [0,T] \to \mathbb{G}$ is an admissible curve where $\gamma(0) = x$ and $\gamma(T) = y$, then $t \mapsto g\cdot\gamma(t)$ is again admissible since the sub-Riemannian structure of $\mathbb{G}$ is left-invariant. Moreover $g\cdot\gamma(0) = g\cdot x$ and $g\cdot\gamma(T) = g\cdot y$, finally $\ell(\gamma) = \ell(g\cdot\gamma)$ since the scalar product defined on $\mathbf{U}$ is again left-invariant. Analogously, for an admissible curve from $g\cdot x$ to $g\cdot y$ one can find and admissible curve from x to y with the same length (simply multiply by g^{-1}). This is sufficient to prove the statement since now $d(g\cdot x, g\cdot y)$ and $d(x,y)$ are infima of equal sets.

(ii) If $\gamma\colon [0,T] \to \mathbb{G}$ is an admissible curve where $\gamma(0) = x$ and $\gamma(T) = y$ parametrized with constant speed, then $t \mapsto \delta_\lambda(\gamma(t))$ is again admissible since

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=s} \delta_\lambda(\gamma(t)) = \mathrm{d}\big(\Delta_\lambda\big)\gamma'(s) = \delta_\lambda(\gamma'(s)) \tag{5.13}$$

and $\gamma'(s) \in \mathrm{d}L_{\gamma(s)}V_1$ since γ is admissible, therefore $\delta_\lambda(\gamma'(s)) = \lambda\gamma'(s) \in \mathrm{d}L_{\gamma(s)}V_1$ and $t \mapsto \delta_\lambda(\gamma(t))$ is admissible; eq. (5.13) also suggest that if u is the control associated to γ, then λu is the control associated to $\delta_\lambda(\gamma)$ and therefore $\ell(\delta_\lambda(\gamma)) = \lambda\ell(\gamma)$. Analogously, for an admissible curve from $\delta_\lambda(x)$ to $\delta_\lambda(y)$ one can find an admissible curve from x to y with length multiplied by λ^{-1} (simply apply $\delta_{\lambda^{-1}}$). This is sufficient to prove the statement since now $d(\delta_\lambda(x), \delta_\lambda(y))$ and $\lambda d(x,y)$ are infima of equal sets.

□

Proposition 5.2.5. *For a Carnot group $(\mathbb{G}, d)$ and for every $x, y \in \mathbb{G}$ there exists an admissible length minimizer γ connecting x to y.*

Proof. Since d is left-invariant, also the topology induced by d is left-invariant. From corollary 2.4.7 we know that there exists $\varepsilon > 0$ such that $\overline{B}_{\mathbf{1}}(\varepsilon)$ is compact. Then for each $g \in \mathbb{G}$ the closed ball $\overline{B}_g(\varepsilon) = g \cdot \overline{B}_{\mathbf{1}}(\varepsilon)$ is compact. By proposition 3.1.8 we find that $(\mathbb{G}, d)$ is complete and we conclude with corollary 3.1.9. □

In the context of Carnot groups proposition 3.3.2 can be formulated in a handier expression:

Theorem 5.2.6. *Let $\mathbb{G}$ be a Carnot group, then the differential of the end-point map from $\mathbf{1}_{\mathbb{G}}$ is given by*

$$D_u \mathrm{End}_{\mathbf{1}}(v) = \big(\mathrm{d}R_{\gamma_u(1)}\big)_{\mathbf{1}} \int_0^1 \mathrm{Ad}_{\gamma_u(t)} v(t)\, \mathrm{d}t\,, \qquad v \in L^\infty([0,1], V_1)\,. \tag{5.14}$$

Proof. From proposition 3.3.2

$$D_u \mathrm{End}_{\mathbf{1}}(v) = \int_0^1 \big[(P^u_{t,1})_* f_{v(t)}\big]\big|_{\gamma_u(1)}\, \mathrm{d}t\,, \qquad v \in L^\infty([0,1], \mathbb{R}^m)\,.$$

Firstly we rewrite the integrand $\big[(P^u_{t,1})_* f_{v(t)}\big]\big|_{\gamma_u(1)} = \mathrm{d}(P^u_{t,1}) f_{v(t)}\big((P^u_{t,1})^{-1}(\gamma_u(1))\big) = \mathrm{d}(P^u_{t,1}) f_{v(t)}(\gamma_u(t))$. Here $f_{v(t)}(\gamma_u(t)) = \mathrm{d}L_{\gamma_u(t)} v(t)$ and $P^u_{t,s} = R_{\gamma_u(s)} \circ R_{(\gamma_u(t))^{-1}}$. Indeed if $h \in \mathbb{G}$ then

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s} R_{\gamma_u(s)} \circ R_{(\gamma_u(t))^{-1}}(h) &= \frac{\mathrm{d}}{\mathrm{d}s} L_{h(\gamma_u(t))^{-1}}(\gamma_u(s)) = \mathrm{d}L_{h(\gamma_u(t))^{-1}}\left(\frac{\mathrm{d}}{\mathrm{d}s}\gamma_u(s)\right) \\
&= \mathrm{d}L_{h(\gamma_u(t))^{-1}}\, \mathrm{d}L_{\gamma_u(s)} u(s) = \mathrm{d}L_{h(\gamma_u(t))^{-1}\gamma_u(s)} u(s)
\end{aligned}$$

that is the value of the non-autonomous vector field associated to v evaluated at $P^u_{t,s}(h)$, since the non-autonomous vector field is left invariant. Therefore

$$\begin{aligned}
D_u \mathrm{End}_{\mathbf{1}}(v) &= \int_0^1 \big[(P^u_{t,1})_* f_{v(t)}\big]\big|_{\gamma_u(1)}\, \mathrm{d}t \\
&= \int_0^1 \mathrm{d}R_{\gamma_u(1)} \mathrm{d}R_{(\gamma_u(t))^{-1}} \mathrm{d}L_{\gamma_u(t)} v(t)\, \mathrm{d}t \\
&= \big(\mathrm{d}R_{\gamma_u(1)}\big)_{\mathbf{1}} \int_0^1 \mathrm{Ad}_{\gamma_u(t)} v(t)\, \mathrm{d}t\,.
\end{aligned}$$
□

Proposition 5.2.7. *Let $\mathbb{G}$ be a Carnot group and V_1 the first layer of its Lie algebra stratification. If $\gamma_u\colon [0,1] \to \mathbb{G}$ is an admissible curve leaving from the origin associated to the control u, then*

$$\mathrm{Im}\,(D_u \mathrm{End}_{\mathbf{1}}) = \big(\mathrm{d}R_{\gamma_u(1)}\big)_{\mathbf{1}} \big(\mathrm{span}\,\{\, \mathrm{Ad}_{\gamma_u(t)} V_1 \;\big|\; t \in [0,1] \,\}\big)\,. \tag{5.15}$$

Proof. We consider theorem 5.2.6 and the fact that $\big(\mathrm{d}R_{\gamma_u(1)}\big)_{\mathbf{1}}$ is a linear isomorphism from $\mathfrak{g}$ to $T_{\gamma_u(1)}\mathbb{G}$, so it is sufficient to show that

$$\left\{ \int_0^1 \mathrm{Ad}_{\gamma_u(t)} v(t)\, \mathrm{d}t \;\middle|\; v \in L^\infty([0,1], V_1) \right\} = \mathrm{span}\,\{\, \mathrm{Ad}_{\gamma_u(t)} V_1 \;\big|\; t \in [0,1] \,\}\,. \tag{5.16}$$

$\subseteq$ The right hand set contains any linear combination of terms $\mathrm{Ad}_{\gamma_u(t_i)} v_i$. Thus, integrals in the left hand set can be expressed as limits of sequences in the right hand set. Indeed for any $n \in \mathbb{N}$ we can define $t_n^k = \frac{k}{n}$ and $I_k^n = [\frac{k-1}{n}, \frac{k}{n}]$, then

$$a_n = \sum_{k=1}^{n} \int_{I_k^n} \mathrm{Ad}_{\gamma(t_n^k)} v(t)\,\mathrm{d}t = \frac{1}{n} \sum_{k=1}^{n} \mathrm{Ad}_{\gamma(t_n^k)} \underbrace{\left(\int_{I_n^k} v(t) \mathrm{d}t \right)}_{\in V_1},$$

and $\lim_{n\to\infty} a_n = \int_0^1 \mathrm{Ad}_{\gamma_u(t)} v(t)\,\mathrm{d}t$. Since the right hand side is closed, it contains the left hand side.

$\supseteq$ Firstly we notice that the left hand side is a vector space since its elements are linear with respect to v, therefore it is closed. It is sufficient to show that any element of the form $\mathrm{Ad}_{\gamma_u(t_i)} v_i$ lies in the left hand set. Let $\psi_n(t)$ be a family of mollifiers centered at t_i and converging to the Dirac delta δ_{t_i} as distributions. Then

$$\lim_{n\to\infty} \int_0^1 \mathrm{Ad}_{\gamma_u(t)} \psi_n(t) v_i = \mathrm{Ad}_{\gamma_u(t_i)} v_i$$

and since the left hand side is closed, $\mathrm{Ad}_{\gamma_u(t_i)} v_i$ falls in the left hand side. □

Remark 5.2.8. If we evaluate eq. (5.15) at $t = 0$ and $t = 1$ we get

$$\left(\mathrm{d}R_{\gamma_u(1)}\right)_{\mathbf{1}} V_1 + \left(\mathrm{d}L_{\gamma_u(1)}\right)_{\mathbf{1}} V_1 \subseteq \mathrm{Im}\left(D_u \mathrm{End}_{\mathbf{1}}\right). \tag{5.17}$$

We recall that corollary 3.3.4 relates abnormal extremals with critical points of the end-point map. This connection can be rephrased in the context of Carnot groups:

Corollary 5.2.9. *Let $\mathbb{G}$ be a Carnot group and V_1 the first layer of its Lie algebra stratification. If $\gamma\colon [0,1] \to \mathbb{G}$ is an admissible curve leaving from the origin, then the following are equivalent*

(i) *γ is abnormal,*

(ii) *there exists $\lambda \in \mathfrak{g}^* \setminus \{0\}$ such that $\lambda\left(\mathrm{Ad}_\gamma(t) V_1\right) = \{0\}$ for every $t \in [0,1]$,*

(iii) *there exists a non-zero right-invariant 1-form α on $\mathbb{G}$ such that $\alpha(\mathrm{d}L_{\gamma(t)} V_1) = \{0\}$ for every $t \in [0,1]$.*

Proof. (i)⇔(ii). It follows from corollary 3.3.4 and eq. (5.15).

(ii)⇔(iii). Let $\lambda \in \mathfrak{g}^* \setminus \{0\}$ such that $\lambda\left(\mathrm{Ad}_\gamma(t) V_1\right) = \{0\}$ for every $t \in [0,1]$. Then we consider the right-invariant 1-form α whose value at $\mathbf{1}$ is λ. Since $\left(R_g\right)^* \alpha = \alpha$ then for $v \in T_g\mathbb{G}$ we have $\alpha(v) = \alpha(\mathrm{d}R_{g^{-1}} v) = \lambda(\mathrm{d}R_{g^{-1}} v)$. Thus

$$\alpha(\mathrm{d}L_{\gamma(t)} V_1) = \lambda(\mathrm{d}R_{\gamma(t)^{-1}} \mathrm{d}L_{\gamma(t)} V_1) = \lambda\left(\mathrm{Ad}_\gamma(t) V_1\right) = \{0\} \quad \text{for every } t \in [0,1].$$

On the other hand, given α from (iii), its value at $\mathbf{1}$ satisfies (ii). □

Definition 5.2.10. Let $\mathbb{G}$ be a Carnot group, we define $\mathrm{Abn}_{\mathbb{G}}$ the *abnormal set* of $\mathbb{G}$ as the set of all critical values of the end-point map starting from the origin:

$$\mathrm{End}_{\mathbf{1}} \colon L^\infty([0,1], V_1) \to \mathbb{G} \qquad u \mapsto \gamma_u(1)\,. \tag{5.18}$$

Namely, $\mathrm{Abn}_{\mathbb{G}} = \left\{ \gamma_u(1) \;\middle|\; D_u\mathrm{End}_{\mathbf{1}}(L^\infty([0,1], V_1)) \neq T_{\gamma_u(1)}\mathbb{G} \right\}$.

A classical problem in singularity theory is to determine whether the set of singular values of a map has measure zero. We will refer to such question as *Sard (or Morse-Sard) problem.* As Arthur Sard proved in 1942, the answer is positive when the map f is sufficiently regular between differentiable manifold (of finite dimensions), see [Sar42]. The answer for the end-point map of Carnot group (and for Sub-Riemannian manifolds in general) is not trivial and it is connected with the regularity problem. Indeed, as shown in [Le +16], if we restrict our analysis to normal-abnormal extremals (that is extremals that are both normal and abnormal) then the problem can be reduced to the finite dimensional case through Hamiltonian formalism and thus we can conclude with the classical Sard theorem.

Definition 5.2.11. We say that a Carnot group $\mathbb{G}$ satisfies the *Sard property* if $\mathrm{Abn}_{\mathbb{G}}(\mathbf{1})$ has zero measure.

We say that a Carnot group $\mathbb{G}$ satisfies the *algebraic (respectively analytic) Sard property* if $\mathrm{Abn}_{\mathbb{G}}(\mathbf{1})$ is contained in a proper real algebraic (respectively analytic) subvariety of $\mathbb{G}$.

To shorten notation, the following definition will be particularly useful:

Definition 5.2.12. Given an admissible curve $\gamma\colon [0,1] \to \mathbb{G}$ starting from the origin, we define

$$\mathscr{E}_\gamma = \mathrm{span}\left\{ \mathrm{Ad}_{\gamma(t)} V_1 \;\middle|\; t \in [0,1] \right\}\,. \tag{5.19}$$

Remark 5.2.13. Evaluating eq. (5.19) at $t = 0$ and $t = 1$ yields

$$V_1 \oplus \mathrm{Ad}_{\gamma(1)} V_1 \subseteq \mathscr{E}_\gamma\,. \tag{5.20}$$

Definition 5.2.14. Let $\mathbb{G}$ and $\mathbb{H}$ be Carnot groups with V_1 and W_1 as first layers of their Lie algebra's stratifications. A *morphism of Carnot groups* $\pi\colon \mathbb{G} \to \mathbb{H}$ is a morphism of Lie groups such that $\mathbf{L}(\pi)(V_1) \subseteq W_1$.

The following results describe how the abnormal set behaves under Cartesian product of Carnot groups and morphisms of Carnot groups.

Proposition 5.2.15. *Let $\mathbb{G}$ and $\mathbb{H}$ be Carnot groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ respectively, each admissible curve $\gamma\colon [0,1] \to \mathbb{G} \times \mathbb{H}$ can be uniquely expressed as $\gamma = (\alpha, \beta)$ for some $\alpha\colon [0,1] \to \mathbb{G}$ and $\beta\colon [0,1] \to \mathbb{H}$ admissible. Then γ is abnormal if either α is abnormal in $\mathbb{G}$ or β is abnormal in $\mathbb{H}$. Therefore*

$$\mathrm{Abn}_{\mathbb{G}\times\mathbb{H}} = (\mathrm{Abn}_{\mathbb{G}} \times \mathbb{H}) \cup (\mathbb{G} \times \mathrm{Abn}_{\mathbb{H}})\,. \tag{5.21}$$

Proof. Let V_1 and W_1 be the first layers of the stratifications of $\mathfrak{g}$ and $\mathfrak{h}$. If $u \in L^\infty([0,1], V_1)$ and $v \in L^\infty([0,1], W_1)$ are controls for α and β respectively, then (u,v) is a control for γ and

$$D_{(u,v)}\mathrm{End}_{\mathbb{G}\times\mathbb{H}} = D_u\mathrm{End}_{\mathbb{G}} \otimes D_v\mathrm{End}_{\mathbb{H}} \,. \tag{5.22}$$

Thus $D_{(u,v)}\mathrm{End}_{\mathbb{G}\times\mathbb{H}}$ is surjective if and only if both $D_u\mathrm{End}_{\mathbb{G}}$ and $D_v\mathrm{End}_{\mathbb{H}}$ are. This is sufficient to conclude. ☐

Proposition 5.2.16. *Let $\pi\colon \mathbb{G} \to \mathbb{H}$ be a surjective morphism of Carnot groups and $\gamma\colon [0,1] \to \mathbb{G}$ an admissible curve. If $\pi \circ \gamma\colon [0,1] \to \mathbb{H}$ is abnormal in $\mathbb{H}$, then γ is abnormal in $\mathbb{G}$. Therefore*

$$\mathrm{Abn}_{\mathbb{H}} \subseteq \pi\left(\mathrm{Abn}_{\mathbb{G}}\right) \,. \tag{5.23}$$

Proof. By hypothesis we know that $\mathscr{E}_{\pi\circ\gamma} = \operatorname{span}\left\{ \mathrm{Ad}_{\pi(\gamma(t))}W_1 \;\middle|\; t \in [0,1] \right\}$ is a proper subspace of $\mathbf{L}(\mathbb{H})$. Firstly we notice that that, since π is a group homomorphism, then $\pi \circ L_g = L_{\pi(g)} \circ \pi$ and $\pi \circ R_g = R_{\pi(g)} \circ \pi$ for every $g \in \mathbb{G}$, thus

$$\mathrm{d}\pi \circ \mathrm{d}L_g = \mathrm{d}L_{\pi(g)} \circ \mathrm{d}\pi \quad \text{and} \quad \mathrm{d}\pi \circ \mathrm{d}R_g = \mathrm{d}R_{\pi(g)} \circ \mathrm{d}\pi \,. \tag{5.24}$$

Then we compute

$$\begin{aligned}
\mathbf{L}(\pi)\,\mathscr{E}_\gamma &= \operatorname{span}\left\{ \mathbf{L}(\pi)\mathrm{Ad}_{\gamma(t)}V_1 \;\middle|\; t \in [0,1] \right\} \\
&= \operatorname{span}\left\{ \mathrm{d}\pi \circ \mathrm{d}L_{\gamma(t)} \circ R_{\gamma(t)^{-1}}V_1 \;\middle|\; t \in [0,1] \right\} \\
&= \operatorname{span}\left\{ \mathrm{d}L_{\pi(\gamma(t))} \circ \mathrm{d}R_{\pi(\gamma(t))^{-1}}\left(\mathbf{L}(\pi)V_1\right) \;\middle|\; t \in [0,1] \right\} \\
&\subseteq \operatorname{span}\left\{ \mathrm{d}L_{\pi(\gamma(t))} \circ \mathrm{d}R_{\pi(\gamma(t))^{-1}}W_1 \;\middle|\; t \in [0,1] \right\} \subseteq \mathscr{E}_{\pi\circ\gamma} \,.
\end{aligned}$$

Since π is surjective, then also $\mathbf{L}(\pi)$ is surjective, therefore $\mathscr{E}_\gamma$ must be a proper subspace of $\mathbf{L}(\mathbb{G})$, thus γ is abnormal.

In order to prove eq. (5.23) it is sufficient to recall that each curve c in $\mathbb{H}$ leaving from the identity admits a unique lift $\overline{c}$ leaving from the identity of $\mathbb{G}$. If c is admissible, then also $\overline{c}$ is admissible since $\mathbf{L}(\pi)V_1 \subseteq W_1$. ☐

Example 5.2.17. Examples of morphisms of Carnot groups are specific quotient maps. If $\mathbb{F}$ is a Carnot group and $\mathfrak{f} = V_1 \oplus \cdots \oplus V_s$ is a stratification for its Lie algebra, we can consider $\mathfrak{I} \subseteq V_2 \oplus \cdots \oplus V_s$ an ideal of $\mathfrak{f}$. Then the quotient $\mathfrak{f}/\mathfrak{I}$ is a stratified Lie algebra (here it is important that $\mathfrak{I}$ does not contain the first layer) and we can consider Carnot group $\mathbb{G}$ with $\mathfrak{f}/\mathfrak{I}$ as Lie algebra. In particular if $N = \exp(\mathfrak{I})$, then N is a normal subgroup of $\mathbb{F}$ and $\mathbb{G} \cong \mathbb{F}/N$.

The quotient map $\pi\colon \mathfrak{f} \to \mathfrak{f}/\mathfrak{I}$ admits a unique lift $\overline{\pi}\colon \mathbb{F} \to \mathbb{G}$ and it is a morphism of Lie group. Here eq. (5.23) becomes

$$\mathrm{Abn}_{\mathbb{G}} \subseteq \pi\left(\mathrm{Abn}_{\mathbb{F}}\right) \,. \tag{5.25}$$

The previous example is particularly useful since each stratified Lie algebra of rank r and step s is a quotient of the unique (up to isomorphisms) free nilpotent Lie algebra of rank r and step s. This is why free Carnot groups represent the starting point for the investigation of the Sard property in Carnot groups.

5.3 Abnormal set in free Carnot groups of step 2

In what follows we will focus on Carnot groups associated with a stratified Lie algebra $\mathfrak{g} = V_1 \oplus V_2$ of step 2. The map $\pi_1 \colon \mathfrak{g} \to V_1$ will denote the canonical projection.

In this framework, eq. (4.34) becomes

$$\mathrm{Ad}_{\exp(X)} Y = e^{\mathrm{ad}_X} Y = Y + [X, Y] = Y + [\pi_1(X), Y] \qquad \forall X, Y \in \mathfrak{g}\,. \tag{5.26}$$

Proposition 5.3.1. *Let $\mathbb{G}$ be a Carnot group of step 2, let γ be an horizontal curve in $\mathbb{G}$ leaving from the identity and define the linear space $P_\gamma \subseteq \mathfrak{g}$ by*

$$P_\gamma = \mathrm{span}\left\{\, \pi_1\left(\exp^{-1}\gamma(t)\right) \;\middle|\; t \in [0,1] \,\right\}. \tag{5.27}$$

Then

$$\mathscr{E}_\gamma = V_1 \oplus [P_\gamma, V_1] \tag{5.28}$$

and γ is abnormal if and only if $[P_\gamma, V_1] \neq V_2$.

Proof. Using *eq.* (5.27) and considering that from eq. (5.20) we have $V_1 \subseteq \mathscr{E}_\gamma$, we obtain

$$\begin{aligned}
\mathscr{E}_\gamma &= \mathrm{span}\left\{\, \mathrm{Ad}_{\gamma(t)} Y \;\middle|\; t \in [0,1],\, Y \in V_1 \,\right\} \\
&= \mathrm{span}\left\{\, \mathrm{Ad}_{\exp(\exp^{-1}\gamma(t))} Y \;\middle|\; t \in [0,1],\, Y \in V_1 \,\right\} \\
&= \mathrm{span}\left\{\, Y + [\pi_1(\exp^{-1}), Y] \;\middle|\; t \in [0,1],\, Y \in V_1 \,\right\} \\
&= V_1 \oplus \mathrm{span}\left\{\, [\pi_1(\exp^{-1}), Y] \;\middle|\; t \in [0,1],\, Y \in V_1 \,\right\} \\
&= V_1 \oplus [P_\gamma, V_1]\,.
\end{aligned}$$

□

Remark 5.3.2. We notice that if γ_u is the admissible curve leaving from the identity associated to the control u, then

$$\pi_1\left(\exp^{-1}\gamma_u(t)\right) = \int_0^t u(s)\mathrm{d}s\,.$$

Indeed, we recall that eq. (4.59) for Lie groups with nilpotency degree 2 becomes

$$X * Y = X + Y + \frac{1}{2}[X, Y]\,, \tag{5.29}$$

and the flow $P^u_{0,t}$ can be expressed as the limit of iterated right multiplication as follows:

$$\begin{aligned}
\gamma_u(t) &= \lim_{n\to\infty} \exp\left(\int_0^{\frac{t}{2^n}} u(s)\mathrm{d}s\right) \cdots \exp\left(\int_{\frac{kt}{2^n}}^{\frac{(k+1)t}{2^n}} u(s)\mathrm{d}s\right) \cdots \exp\left(\int_{\frac{(2^n-1)t}{2^n}}^{\frac{2^n t}{2^n}} u(s)\mathrm{d}s\right) \\
&= \lim_{n\to\infty} \exp\left(\underbrace{\int_0^t u(s)\mathrm{d}s}_{\in V_1} + \underbrace{v_2^{(n)}}_{\in V_2}\right) = \exp\left(\underbrace{\int_0^t u(s)\mathrm{d}s}_{\in V_1} + \underbrace{v_2}_{\in V_2}\right).
\end{aligned}$$

Therefore P_{γ_u} can be seen as the smallest subspace of V_1 such that $u(t) \in P_{\gamma_u}$ for almost every $t \in [0,1]$.

Proposition 5.3.3. *Let $\mathbb{G}$ be a Carnot group associated with a Lie algebra $\mathfrak{g}\colon V_1 \oplus V_2$ of step 2 and rank r, then*

$$\operatorname{Abn}_{\mathbb{G}} = \bigcup \{ \exp(P \oplus [P,P]) \mid P \leq V_1,\, \dim P \leq r-2,\, [P,V_1] \neq V_2 \} . \tag{5.30}$$

Proof.

- $\subseteq$ From remark 5.3.2 we deduce that, for every admissible curve γ, $\operatorname{Im}\gamma$ is contained in the subgroup of $\mathbb{G}$ associated to the Lie algebra generated by P_γ, that is

$$\operatorname{Im}\gamma \subseteq \exp\left(P_\gamma \oplus [P_\gamma, P_\gamma]\right) . \tag{5.31}$$

  If γ is abnormal, then by proposition 5.3.1 we can ask $[P_\gamma, V_1] \neq V_2$. If $\dim P_\gamma \in \{ r, r-1 \}$ we would have $[P_\gamma, V_1] = V_2$, therefore we can also ask $\dim P_\gamma \leq r-2$.
- $\supseteq$ If $P \leq V_1$ is such that $\dim P \leq r-2$ and $[P,P] \neq V_2$, then by Rashevskii-Chow connectivity theorem 2.4.2 any point in the subgroup $H = \exp\left(P \oplus [P,P]\right)$ can be connected to $\mathbf{1}_{\mathbb{G}}$ by an admissible curve γ entirely contained in H, and such curve γ must be abnormal by proposition 5.3.1. □

In the context of free Carnot groups of step 2, the condition $\dim P_\gamma \leq r-2$ is also sufficient to have $[P_\gamma, V_1] \neq V_2$, therefore eq. (5.30) becomes

$$\operatorname{Abn}_{\mathbb{G}} = \bigcup \{ \exp(P \oplus [P,P]) \mid P \leq V_1,\, \dim P = r-2 \} . \tag{5.32}$$

We are now ready to state the following:

Theorem 5.3.4. *Let $\mathbb{G}$ be a free Carnot group of step 2, then $\operatorname{Abn}_{\mathbb{G}}$ is contained in an affine algebraic subvariety of codimension 3.*

Proof. The set $\operatorname{Gr}(r-2, V) = \{ P \leq V_1 \mid, \dim P = r-2 \}$, known as *Grassmannian of rank* $r-2$ is an algebraic variety of dimension $2(r-2)$. For each $P \in \operatorname{Gr}(r-2,V)$ the set $\exp(P \oplus [P,P])$ is a free Carnot group of step 2 and rank $r-2$, therefore its dimension is $\frac{(r-1)(r-2)}{2}$.

As a result, each element in $\operatorname{Abn}_{\mathbb{G}}$ can be expressed with polynomial combinations of

$$2(r-2) + \frac{(r-1)(r-2)}{2} = \frac{r(r+1)}{2} - 3 = \dim \mathbb{G} - 3$$

parameters. □

Finally, we are interested in expressing $\operatorname{Abn}_{\mathbb{G}}$ in free Carnot groups of step 2 as zeros of a system of polynomials. Firstly we fix some notation:

Adapted basis and coordinates for free Carnot groups of step 2

Referring to the Lie algebra, we will denote with

- $\{ X_1, \dots, X_r \}$ a basis for the first layer V_1,

- $\{ X_{(i,j)} \mid 1 \leq i < j \leq r \}$ a basis for the second layer V_2, where $[X_s, X_t] = X_{(s,t)}$ whenever $s < t$.

We will denote with $x_u, x_{(s,t)}$ the coordinates relative to the basis $\{ X_1, \ldots, X_r \} \cup \{ X_{(i,j)} \mid 1 \leq i < j \leq r \}$, with the convention that, when $t < s$, $x_{(s,t)} = -x_{(t,s)}$. On the Carnot group $\mathbb{G}$ we will use exponential coordinates of the first type given from the coordinates $x_u, x_{(s,t)}$ on $\mathfrak{g}$.

Permutations on subsets of natural numbers

Let A be a finite set of natural numbers, we will denote its cardinality by $\#A$. The canonical total order on $\mathbb{N}$ induces a total order on A, thus we can denote the elements of A as

$$1_A < 2_A < \cdots < (\#A)_A \,. \tag{5.33}$$

We denote by $S(A)$ the set of permutations of A. Given $j \leq r$ positive integers, we define $A^r_j = \{ 1, 2, \ldots, r \} \setminus \{ j \}$ and we notice that $\#(A^r_j) = r - 1$.

Given a permutation $\sigma \in S(A)$, we define the set of "matchings induced by σ", denoted with C^σ_A, as follows:

- $C^\sigma_A = \{ (\sigma(1_A), \sigma(2_A)), (\sigma(3_A), \sigma(4_A)), \ldots, (\sigma((\#A-1)_A), \sigma((\#A)_A)) \}$ if $\#A$ even,
- $C^\sigma_A = \{ (\sigma(1_A), \sigma(2_A)), \ldots, (\sigma((\#A-2)_A), \sigma((\#A-1)_A)), \sigma((\#A)_A) \}$ if $\#A$ odd.

We notice that in the first case C^σ_A is a set of couples, in the second case there is an unmatched element. In both cases C^σ_A inherits a total order induced by the total order of A:

- $(\sigma(1_A), \sigma(2_A)) < (\sigma(3_A), \sigma(4_A)) < \cdots < (\sigma((\#A-1)_A), \sigma((\#A)_A))$,
- $(\sigma(1_A), \sigma(2_A)) < \cdots < (\sigma((\#A-2)_A), \sigma((\#A-1)_A)) < \sigma((\#A)_A)$.

Abnormal polynomials and combinatorial remarks

Let r be the rank of a free Carnot group $\mathbb{G}$ of step 2 and $1 \leq j \leq r$ an integer. We define the following polynomials:

$$\mathscr{P}^r_j = \sum_{\sigma \in S(A^r_j)} \operatorname{sgn}(\sigma) \Bigg(\prod_{c \in C^\sigma_{A^r_j}} x_c \Bigg) \,. \tag{5.34}$$

First of all, let us understand the combinatorial structures of these polynomials. It may happen that distinct permutations $\sigma_1, \sigma_2 \in S(A^r_h)$ produce the same monomial, that is

$$\prod_{c \in C^{\sigma_1}_{A^r_j}} x_c = \pm \prod_{c \in C^{\sigma_2}_{A^r_j}} x_c \,.$$

This happens if (and only if) the couples in the two expressions are the same, possibly in a different order and possibly with some of the elements in the same couple exchanged. For instance monomials $(\pm)x_{(1,2)}x_{(3,4)}$ and $(\pm)x_{(4,3)}x_{(1,2)}$ are essentially equal (up to sign) and therefore they can be summed. Indeed we recall that $x_{(3,4)}$ and $x_{(4,3)}$ corresponds to the same coordinates but with opposite sign.

The previous observation clearly induces an equivalence relation on $S(A^r_j)$, we now prove that equivalent permutations produces monomials with the same sign. Two equivalent permutations differs by a finite number of operations of the type:

- "Switch between couples" given by a double transposition $(2a \quad 2b)(2a-1 \quad 2b+1)$ where $1 \leq a, b \leq \lfloor \frac{r-1}{2} \rfloor$. This case refers to the commutativity of the factors x_c. In this case the double transposition does not change the sign of the permutation ($\mathrm{sgn}(\sigma)$ in the expression of eq. (5.34)).
- "Switch inside a couple" given by a single transposition $(2d \quad 2d-1)$ where $1 \leq d \leq \lfloor \frac{r-1}{2} \rfloor$. This case refers to the identification $x_{(a,b)} = -x_{(b,a)}$. In this case the single transposition changes $\mathrm{sgn}(\sigma)$ but it is corrected by the fact that, in the identification, the sign is again changed.

Therefore in the expression eq. (5.34) there are as many "essentially different monomials" as there are equivalence classes of permutations. Each of them appears with an integer coefficient equal (in absolute value) to the cardinality of the associated equivalence class. The total number of permutations is $(r-1)!$ while each class contains

$$\left(\lfloor \tfrac{r-1}{2} \rfloor\right)!\, 2^{\lfloor \frac{r-1}{2} \rfloor},$$

elements. Therefore the total number of equivalence classes is

$$\sum_{k=1}^{\lfloor \frac{r}{2} \rfloor} (2k-1)\,.$$

Theorem 5.3.5. *Let $\mathbb{G}$ be a free Carnot group of step 2 and rank r. Let $x_u, x_{(s,t)}$ be coordinates in $\mathbb{G}$ referring to an adapted basis, then for every $v \in \mathrm{Abn}_{\mathbb{G}}$ its coordinates x satisfies*

$$\mathscr{P}^r_j(x) = 0 \qquad \forall 1 \leq j \leq r\,. \tag{5.35}$$

Proof. If $v \in \mathrm{Abn}_{\mathbb{G}}$ then $v \in \exp(P \oplus [P,P])$ for some P subspace of V_1 with $\dim V_1 = r-2$. We then select $P_1, \dots, P_{r-2}$ a basis for P, thus

$$P_i = \sum_{u=1}^{r} w^u_i X_u \qquad i = 1, \dots, r-2 \tag{5.36}$$

and

$$[P_h, P_k] = \sum_{s<t} (w^s_h w^t_k - w^s_k w^t_h) X_{(s,t)} \qquad h, k = 1, \dots, r-2 \tag{5.37}$$

for some w^u_i. A generic element of $\exp(P \oplus [P,P])$ can be expressed in this new basis:

$$v = \sum_{i=1}^{r-2} \alpha_i P_i + \sum_{h<k\leq r-2} \beta_{(h,k)} [P_h, P_k]\,, \tag{5.38}$$

and in the original basis:

$$v = \sum_{u=1}^{r} \left(\sum_{i=1}^{r-2} \alpha_i w_i^u \right) X_u + \sum_{s<t} \left(\sum_{h<k\le r-2} \beta_{(h,k)} (w_h^s w_k^t - w_k^s w_h^t) \right) X_{(s,t)} \,. \tag{5.39}$$

Let us substitute the coordinates of v, written in the previous way, in the polynomial $\mathscr{P}_j^r$ (hereafter $A = A_j^r$).

If r is odd we obtain an algebraic sum of factors of the type

$$\beta_{(h_1,k_1)} \beta_{(h_2,k_2)} \dots \beta_{(h_{(r-1)/2},k_{(r-1)/2})} w_{p_1}^{\sigma(1_A)} w_{q_1}^{\sigma(2_A)} w_{p_2}^{\sigma(3_A)} w_{q_2}^{\sigma(4_A)} \dots w_{p_{(r-1)/2}}^{\sigma((r-2)_A)} w_{q_{(r-1)/2}}^{\sigma((r-1)_A)} \tag{5.40}$$

where $\sigma \in S(A_j^r)$ and (p_n, q_n) is a certain permutation of (h_n, k_n) (we recall that $h_n < K_n$) for every $1 \le n \le (r-1)/2$. Such a factor appears with a sign equal to $\operatorname{sgn}(\sigma)(-1)^m$ where m is the number of times (p_n, q_n) is inverted compared to (h_n, k_n).

Let us recall that $1 \le h_n, k_n \le r-2$ for every $1 \le n \le (r-1)/2$, therefore, for the pigeonhole principle, in each monomial there is at least one repeating subscript among the $r-1$ subscripts in that monomial. Therefore also among p_n, q_n there is a repetition.

Therefore, in the order given by eq. (5.40), we can consider the first repetition appearing in subscripts of factors w_b^a, starting from the left. We call w_z^f, w_z^g those factors corresponding to the first repeating subscript z. Let us now consider the product obtained switching those two factors in the expression eq. (5.40). Such new product refers to a new permutation σ' which differs from σ by the transposition $(f\,g)$, therefore $\operatorname{sgn}(\sigma') = -\operatorname{sgn}(\sigma)$. Since the subscripts of w_z^f and w_z^g were equal, the "m factor" of the new product is the same.

The described association is clearly a pairing and paired products appear with opposite signs. Since they are equal up to a commutations of factors, they can be simplified. Therefore

$$\mathscr{P}_j^r(x) = 0\,.$$

If r is even we obtain an algebraic sum of factors of the type

$$\beta_{(h_1,k_1)} \beta_{(h_2,k_2)} \dots \beta_{(h_{(r-2)/2},k_{(r-2)/2})} \alpha_i w_{p_1}^{\sigma(1_A)} w_{q_1}^{\sigma(2_A)} w_{p_2}^{\sigma(3_A)} w_{q_2}^{\sigma(4_A)} \dots w_{p_{(r-2)/2}}^{\sigma((r-3)_A)} w_{q_{(r-2)/2}}^{\sigma((r-2)_A)} w_i^{\sigma((r-1)_A)}$$

and the reasoning to simplify them is the same. □

Example 5.3.6. When $r = 2$, $\mathbb{F}_2$ is equal to the second Heisenberg group $\mathbb{H}^1$. In this case abnormal polynomials from eq. (5.34) are (up to constants)

$$\mathscr{P}_1^2 = x_2 \quad \text{and} \quad \mathscr{P}_2^2 = x_1\,.$$

On the other hand, from eq. (5.32), we get $\mathrm{Abn}_{\mathbb{F}_2} = \{\,0\,\}$ which is strictly bigger that $\{\, x_1 = 0,\, x_2 = 0\,\}$.

Example 5.3.7. When $r = 3$, the abnormal polynomials are (up to constants)

$$\mathscr{P}_1^3 = x_{12}\,, \quad \mathscr{P}_2^3 = x_{13} \quad \text{and} \quad \mathscr{P}_3^3 = x_{23}$$

while eq. (5.32) gives $\mathrm{Abn}_{\mathbb{F}_3} = \exp(V_1)$, therefore

$$\mathrm{Abn}_{\mathbb{F}_3} = \{\, x_{12} = 0,\, x_{13} = 0,\, x_{23} = 0 \,\}\ .$$

Example 5.3.8. When $r = 4$, the abnormal polynomials are (up to constants)

$$\begin{aligned}
\mathscr{P}_1^4 &= -x_2 x_{34} + x_3 x_{24} - x_4 x_{23}\,,\\
\mathscr{P}_2^4 &= x_1 x_{34} - x_3 x_{14} + x_4 x_{13}\,,\\
\mathscr{P}_3^4 &= -x_1 x_{24} + x_2 x_{14} - x_4 x_{12}\,,\\
\mathscr{P}_4^4 &= x_1 x_{23} - x_2 x_{13} + x_3 x_{12}\,.
\end{aligned}$$

They are related by the expression $\sum_{i=1}^4 x_i \mathscr{P}_i^4 = 0$. Therefore, if $x_i \neq 0$, then the three polynomials in $\left\{ \mathscr{P}_j^4 \;\middle|\; j \neq i \right\}$ being equal to zero implies $\mathscr{P}_i^4$ being zero as well. Thus, outside from $\{\, x_1 = 0,\, x_2 = 0,\, x_3 = 0,\, x = 4 \,\}$, which has codimension 4, the set of zeros can be described locally using only three of those polynomials. For instance, let us suppose $g \in \{\, x_1 \neq 0 \,\}$, then in a neighbourhood of g the set of zeros is described with

$$\begin{aligned}
x_1 x_{34} - x_3 x_{14} + x_4 x_{13} &= 0\,,\\
-x_1 x_{24} + x_2 x_{14} - x_4 x_{12} &= 0\,,\\
x_1 x_{23} - x_2 x_{13} + x_3 x_{12} &= 0\,.
\end{aligned}$$

The gradients of those polynomials are

$$\begin{gathered}
(x_{24}, 0, -x_{14}, x_{13}, 0, x_4, -x_3, 0, 0, x_1)\,,\\
(-x_{24}, x_{14}, 0, -x_{12}, -x_4, 0, x_2, 0, -x_1, 0)\,,\\
(x_{23}, -x_{13}, x_{12}, 0, x_3, -x_2, 0, x_1, 0, 0)\,,
\end{gathered}$$

and since $x_1 \neq 0$ in a neighbourhood of g, then those gradients are linearly independent and therefore they define locally an algebraic variety of codimension 3.

Finally we prove that a solution to this system of polynomials is actually in the abnormal set. For $r = 4$ we can express explicitly eq. (5.32) in exponential coordinates:

$$\begin{aligned}
\mathrm{Abn}_{\mathbb{F}_4} = &\left\{\, (0, x'') \;\middle|\; 0 \in \mathbb{R}^4,\, x'' \in \mathbb{R}^6 \,\right\} \cup \\
&\left\{\, (x', x'') \;\middle|\; x' \in \mathbb{R}^4 \setminus \{\, 0 \,\},\, x'' \in \mathbb{R}^6,\, \exists y \in \mathbb{R}^4 \text{ such that } [x', y] = x'' \,\right\}\ ,
\end{aligned}$$

where Lie brackets of coordinates indicates a Lie brackets between vectors with those coordinates (and zeros in the other coordinates). Then, if $g = (x', x'')$ is a solution to the system of polynomials and $x' \neq 0$, we need to prove that there exists $y \in \mathbb{R}^4$ such that $[x', y] = x''$. Firstly we observe that the set $\left\{ [x', y] \;\middle|\; y \in \mathbb{R}^4 \right\}$ depends linearly on x' and also the solution set to the system of linear equations in $\mathbb{R}^6$

$$\begin{aligned} -x'_2 x_{34} + x'_3 x_{24} - x'_4 x_{23} &= 0\,, \\ x'_1 x_{34} - x'_3 x_{14} + x'_4 x_{13} &= 0\,, \\ -x'_1 x_{24} + x'_2 x_{14} - x'_4 x_{12} &= 0\,, \\ x'_1 x_{23} - x'_2 x_{13} + x'_3 x_{12} &= 0\,. \end{aligned}$$

depends linearly on (x'_1, x'_2, x'_3, x'_4). In this way we only need to prove the equivalence of those sets when $x' = e_1, e_2, e_3, e_4$. When $x'_1 = e_1$ then the system becomes

$$x_{34} = 0\,, \quad x_{24} = 0\,, \quad x_{23} = 0\,,$$

and

$$\mathrm{Abn}_{\mathbb{F}_4} \cap \{ x'_1 = e_1 \} = \left\{ (e_1, x'') \;\middle|\; x'' \in \mathbb{R}^6,\, \exists y \in \mathbb{R}^4 \text{ such that } [e_1, y] = x'' \right\}\,,$$

and $\left\{ [e_1, y] \;\middle|\; y \in \mathbb{R}^4 \right\} = \operatorname{span}\{ X_{12}, X_{13}, X_{14} \} = \{ x_{34} = 0, x_{24} = 0, x_{23} = 0 \}$. the other cases follow analogously.

Example 5.3.9. When $r = 4$, the abnormal polynomials defined in eq. (5.34) are (up to constants)

$$\begin{aligned} \mathscr{P}^5_1 &= x_{23}x_{45} - x_{24}x_{35} + x_{25}x_{34}\,, \\ \mathscr{P}^5_2 &= -x_{13}x_{45} + x_{14}x_{35} - x_{15}x_{34}\,, \\ \mathscr{P}^5_3 &= x_{12}x_{45} - x_{14}x_{25} + x_{15}x_{24}\,, \\ \mathscr{P}^5_4 &= -x_{12}x_{35} + x_{13}x_{25} - x_{15}x_{34}\,, \\ \mathscr{P}^5_5 &= x_{12}x_{34} - x_{13}x_{24} + x_{14}x_{23}\,. \end{aligned}$$

In this case, polynomial relations are $\sum_{i=1}^5 x_{ji}\mathscr{P}^5_i$ for $j = 1, \ldots, 5$, with the usual convention that $x_{ij} = -x_{ji}$ and $x_{ii} = 0$. Therefore, in a neighbourhood of g contained in $\{ x_{ij} = 0 \mid 1 \le i, j \le 5 \}$ (and this set has codimension 10), the set of solutions to the system of polynomials can be again described using 3 polynomials. Analogously to the case $r = 4$, in this neighbourhood the selected polynomials have linearly independent gradients and therefore the set of solutions is locally an algebraic variety of codimension 3.

Bibliography

[ABB20] Andrei Agrachev, Davide Barilari, and Ugo Boscain. *A Comprehensive Introduction to Sub-Riemannian Geometry*. Cambridge Studies IN Advanced Mathematics 181. Cambridge University Press, 2020. ISBN: 9781108476355.

[And96] Jean-Jacques Risler (eds.) André Bellaïche (auth.) André Bellaïche. *Sub-Riemannian Geometry*. 1st ed. Progress in Mathematics 144. Birkhäuser Basel, 1996. ISBN: 978-3-0348-9946-8,978-3-0348-9210-0.

[BP07] Alberto Bressan and Benedetto Piccoli. *Introduction to the Mathematical Theory of Control*. First. Applied Mathematics. American Institute of Mathematical Sciences, 2007. ISBN: 9781601330024.

[Don16] Enrico Le Donne. *A primer on Carnot groups: homogenous groups, CC spaces, and regularity of their isometries*. 2016. arXiv: 1604.08579 [math.MG].

[HN12] Joachim Hilgert and Karl-Hermann Neeb. *Structure and Geometry of Lie Groups*. 1st ed. Springer Monographs in Mathematics. Springer-Verlag New York, 2012. ISBN: 9780387847948.

[Le +16] Enrico Le Donne et al. "Sard property for the endpoint map on some Carnot groups". In: *Annales de l'Institut Henri Poincaré C, Analyse non linéaire* 33.6 (Nov. 2016), pp. 1639–1666. ISSN: 0294-1449. DOI: 10.1016/j.anihpc.2015.07.004. URL: http://dx.doi.org/10.1016/j.anihpc.2015.07.004.

[Lee13] John M Lee. *Introduction to smooth manifolds*. 2nd ed. Graduate texts in mathematics 218. Springer, 2013. ISBN: 9781441999818.

[Mit85] John Mitchell. "On Carnot-Carathéodory metrics". In: *Journal of Differential Geometry* 21.1 (1985), pp. 35–45. DOI: 10.4310/jdg/1214439462. URL: https://doi.org/10.4310/jdg/1214439462.

[Mon06] Richard Montgomery. *A Tour of Subriemannian Geometries, Their Geodesics and Applications*. Mathematical Surveys and Monographs 91. American Mathematical Society, 2006. ISBN: 0821841653,9780821841655. URL: http://gen.lib.rus.ec/book/index.php?md5=a9427e12e8ebc55ed9028e32d8385163.

[OV19] Alessandro Ottazzi and Davide Vittone. *On the codimension of the abnormal set in step two Carnot groups*. 2019. DOI: https://doi.org/10.1051/cocv/2018002.

[Reu93] Christophe Reutenauer. *Free Lie algebras*. London Mathematical Society monographs new ser., 7. Clarendon Press; Oxford University Press, 1993. ISBN: 0198536798,9780198536796.

[Sar42] Arthur Sard. "The measure of the critical values of differentiable maps". In: *Bull. Amer. Math. Soc.* 48 (1942), pp. 883–890. DOI: https://doi.org/10.1090/S0002-9904-1942-07811-6.