ELEMENTS OF SUB-RIEMANNIAN GEOMETRY AND ITS APPLICATIONS

Julián Jiménez Cárdenas

A thesis presented for the degree of
Master of Mathematics

Departamento de Matemáticas
Universidad de los Andes
Bogotá, Colombia
April 14, 2022

ABSTRACT. The goal of this work is to introduce the main theorems in sub-Riemannian geometry, and for that we will cover first the topic of distributions, stating the Frobenius' theorem and introducing the idea of foliation. Later on, we focus on the sub-Riemannian structures, and prove Chow's theorem, central in the study of connectivity between horizontal geodesics. The last part of this work focuses on G-principal bundles endowed with compatible sub-Riemannian structures. We use this construction with ideas of symplectic geometry to prove the Theorem on normal geodesics of bundle type sub-Riemannian metrics and to study Wong's equations.

INTRODUCTION

Riemannian geometry is an intense field of study due to the fact that it is the generalization of the idea of metric and distance onto general geometric spaces. This characteristic makes Riemannian geometry one of the preferred formalisms to work in physics' topics like general relativity and quantum mechanics. In this field we start with a manifold, endowed with a metric, which is a section of the $(0, 2)$-tensor bundle which is symmetric and non-degenerate at all points of the manifold. The non-degeneracy condition can be relaxed, and in this case some important examples raise up, like the Lorentz/Minkowski metric, the most widely used metric in special relativity. In principle, all proofs that depend only on the invertibility of the metric tensor, such as existence and uniqueness of the Riemannian connection and geodesics, work fine in the pseudo-Riemannian setting, while proofs that use positivity in an essential way, such as those involving distance-minimizing properties of geodesics, do not [1, p. 31]. An extensive mathematical development of pseudo-Riemannian geometry can be found on [2].

Another alternative formulation of Riemannian geometry arises when we consider a sub-bundle (called distribution) of the tangent bundle of a given manifold: in this case, the metric will be a section of the $(0, 2)$-tensor bundle of this sub-bundle, with the same properties than in the case of Riemannian geometry (non-degeneracy and symmetry). This formulation receives the name of sub-Riemannian geometry, and it will be the main focus of this text. Before introducing the concept of sub-Riemannian structure, we focus on explaining the idea of distribution, their different definitions and associated results. To do this, we define a distribution in terms of local frames and 1-forms, we explain when a distribution is involutive, integrable and completely integrable, and use the Frobenius's theorem to connect these three concepts, while discussing when the points of a given manifold are connected via horizontal curves, *i.e.*, curves whose tangent vector is always contained in the distribution.

Later on, we introduce the concept of sub-Riemannian structure, and define the horizontal distance as the infimum of the length of all horizontal curves that connect two given points in the manifold, in analogy with Riemannian distance. We then introduce the bracket

generating distributions as a special class of distribution such that there is a local frame that locally spans the tangent bundle through successive applications of Lie brackets, and we state and show the Chow's theorem, which states that a sufficient condition for all points to be connected through horizontal curves is that the distribution is bracket generating. Therefore, the distance between two points in a sub-Riemannian structure is always well-defined when the subjacent distribution is bracket generating.

In order to work with principal G-bundles endowed with metric of bundle type, we start introducing what is a submersion, and we endow it with a sub-Riemannian structure in the total space (equivalently, a Riemannian structure in the base space). With these tools, we start the topic of metrics on principal bundles explaining aspects of Lie groups and their actions, and give some examples of principal G-bundles.

Following the steps of Riemannian geometry, we proceed studying the normal geodesics associated to sub-Riemannian structures, and in order to do that we make a brief introduction to symplectic geometry. This will allow to use the Hamiltonian formalism to characterize the horizontal geodesics as integral curves of the vector field associated to the sub-Riemannian Hamiltonian, which is defined in the same way than the Riemannian Hamiltonian, using the sub-Riemannian metric instead. With this formalism we proceed proving the Theorem on normal geodesics of bundle type sub-Riemannian metrics, which provides a way to identify normal (horizontal) geodesics in the (base) total space for a principal G-bundle whose metric satisfies some special properties.

In case that the conditions of the Theorem on normal geodesics of bundle type sub-Riemannian metrics do not apply, we can characterize the normal horizontal geodesics of a principal G-bundle using the Wong's equations. These are a family of differential equations over the quotient of the cotangent space and the action of the Lie group G over it.

1. Distributions

1.1. **Distributions and Horizontal Curves.**

Definition 1. ([3], section 1.3) The triple $(E, M, \pi : E \to M)$ is a *smooth vector bundle* if E, M are smooth manifolds, π is a smooth surjection; the fibers $\pi^{-1}(p)$ for $p \in M$ have the structure of finite dimensional vector spaces; and for every point $p \in M$, there is an open neighborhood U of p, a natural number k and a diffeomorphism $\phi : U \times \mathbb{R}^k \to \pi^{-1}(U)$, such that for all $q \in U$ the following propositions hold:

- $(\pi \circ \phi)(q, v) = q$ for all $v \in \mathbb{R}^k$, and

- the map $v \mapsto \phi(p, v)$ is a linear isomorphism between the vector spaces $\pi^{-1}(p)$ and $\mathbb{R}^k$.

The first condition means that the following diagram is commutative,

$$
\begin{array}{ccc}
U \times \mathbb{R}^k & \xrightarrow{\ \ \phi\ \ } & \pi^{-1}(U) \\
\ \ \downarrow{\scriptstyle p} & & \ \ \downarrow{\scriptstyle \pi} \\
U & \xrightarrow[\ \ Id\ \ ]{} & U
\end{array}
$$

where $p : U \times \mathbb{R}^k \to U$, $p(q, v) = q$ is the projection onto the first factor.

We take the disjoint union of the tangent spaces at every point of the manifold in order to define the tangent bundle, $TM = \bigsqcup_{p \in M} T_p M$, equipped with manifold structure given by the one of M. It is called a bundle because the triple $(TM, M, \pi : TM \to M)$ satisfies the definition of smooth vector bundle, where $\pi : TM \to M$ is the projection function, that assigns each tangent vector to the point where it is based. Each fiber of this bundle is a tangent space based on a point, and the tangent bundle is in particular a vector bundle because each fiber has structure of finite dimensional real vector space.

There is another important class of smooth vector bundles, namely, the tensor bundles, that are defined following [4], examples 2.4 and 2.5.

Definition 2. If $(E, M, \pi : E \to M)$ is a smooth vector bundle, the dual bundle of this bundle is another smooth vector bundle $(E^*, M, \pi^* : E^* \to M)$, known as *dual bundle*, whose fibers are the dual spaces to the fibers of E, that is,

$$
(\pi^*)^{-1}(p) = (\pi^{-1}(p))^*, \text{ for all } p \in M,
$$

and

$$
E^* = \bigsqcup_{p \in M} (\pi^*)^{-1}(p).
$$

Definition 3. If $(E, M, \pi : E \to M)$ is a smooth vector bundle, the $(k, l)-tensor\ bundle$ associated to this bundle, with $k, l \geq 0$, is defined as the smooth vector bundle $(T_l^k E, M, \pi_{kl} : T_l^k E \to M)$, whose fibers are the vector spaces of $(k, l)-$tensors of the fibers of E, that is,

$$
\pi_{kl}^{-1}(p) = T_l^k \pi^{-1}(p) = \left(\bigotimes_{j=1}^{l} (\pi^*)^{-1}(p) \right) \otimes \left(\bigotimes_{j=1}^{k} \pi^{-1}(p) \right), \text{ for all } p \in M,
$$

and

$$
T_l^k E = \bigsqcup_{p \in M} \pi_{kl}^{-1}(p).
$$

Remark 1. If $k = 1, l = 0$, $T_1^0 E = E$, and if $k = 0, l = 1$, $T_0^1 = E^*$. We associate $T_0^0 = M \times \mathbb{R}$ to the *trivial line bundle* if both $k = l = 0$.

The formal proof that both definitions are consistent definitions of smooth vector bundles can be found in [3], section 1.3.

Definition 4. A *smooth section* of a smooth vector bundle $(E, M, \pi : E \to M)$ is a (smooth) application $s : M \to E$ such that $\pi \circ s = \mathbb{1}_M$.

For an arbitrary smooth vector bundle bundle $(E, B, \pi : E \to B)$, a sub-bundle of this bundle $(E', B', \pi' : E' \to B')$ is a smooth vector bundle bundle that satisfies $E' \subset E$, $B' \subset B$ and $\pi|_{E'} = \pi'$.

Definition 5. A *distribution* $\mathcal{H}$ over a manifold M is a vector sub-bundle of the tangent bundle TM. If this vector sub-bundle is in addition smooth, the distribution is said to be smooth, and the rank k of this distribution is defined as the dimension of its fibers.

The fiber of $\mathcal{H}$ at a point $p \in M$ is denoted as H_p.

The concept of (smooth) distribution can be defined in an alternative but equivalent (see [5], Ch. 19, lemma 19.5) fashion using 1-forms, with the help of the following proposition.

Proposition 1 (Smooth Distribution (using 1-forms)). *Suppose that M is a smooth $m-$dimensional manifold, and $\mathcal{H} \subset TM$ is a distribution of rank k. Then $\mathcal{H}$ is smooth if and only if for every point $p \in M$ there is a neighborhood U of that point where exist $n - k$ smooth 1-forms $\omega^1, \ldots, \omega^{n-k}$, such that for all $q \in U$,*

$$\mathcal{H}_q = \operatorname{Ker} \omega^1|_q \cap \cdots \cap \operatorname{Ker} \omega^{n-k}|_q.$$

The following proposition gives an alternative way to represent a smooth distribution in terms of local frames ([5], Ch. 19, page 491).

Proposition 2 (Smooth Distribution (using local frames)). *It is said that $\mathcal{H} \subset TM$ is a smooth distribution over M if and only if for each point $p \in M$, there is an open neighborhood U of it on which there are smooth vector fields $X_1, \ldots, X_n : U \to TM$ that at each point of U form a base for the fiber of that point respect to the distribution.*

Some examples of distributions over manifolds are given below.

Example 1 (Tangent Bundle). For any manifold M, the tangent bundle TM is by definition a distribution over it.

Example 2 (Heisenberg Group). Take $\mathbb{R}^3$ as the manifold, and define over it the distribution $\mathcal{H} \subset T\mathbb{R}^3$, whose fiber at an arbitrary $(x, y, z) \in \mathbb{R}^3$ is

$$\mathcal{H}_{(x,y,z)} = \operatorname{Ker} \omega|_{(x,y,z)},$$

where $\omega = dz - (xdy - ydx)/2 \in \Omega^1(\mathbb{R}^3)$. It is important to notice that $\omega \neq 0$ for all $(x, y, z) \in \mathbb{R}^3$.

Remark 2. The last example is of particular interest because of its connection with the isoperimetrical problem. For more information see [6], Ch. 1.

Example 3 (Vector field over a manifold). Any smooth vector field $X : M \to TM$ (which does not vanish out at any point) on a manifold determines a distribution $\mathcal{H}$ over the manifold M whose fiber at an arbitrary $p \in M$ is the span of X_p.

Given a distribution $\mathcal{H}$ over a manifold M, a curve or vector field is said to be *horizontal* if it is tangent to $\mathcal{H}$. This means that, in the case of a curve, its tangent vector at every point is contained in the fibers of $\mathcal{H}$; and in the case of a vector field, its value at every point is contained in the fibers of $\mathcal{H}$.

For example, every curve over a manifold with distribution given by its tangent bundle is horizontal. In the example 2, the liftings of the solutions of the isoperimetrical problem are horizontal. In example 3, and in particular in the case of systems of linear differential equations, the solution curves of the system of differential equations are examples of horizontal curves for that distribution.

1.2. **Frobenius' Theorem.** Now the natural question that arises is if, given two points on a manifold with a given horizontal distribution, is there a smooth horizontal curve that joins them?

Definition 6. Given a smooth distribution $\mathcal{H} \subset TM$, we say that a nonempty immersed submanifold $N \subseteq M$ is an *integral manifold* of $\mathcal{H}$ if $T_pN = \mathcal{H}_p$ for all $p \in N$, *i.e.*, if the tangent space of all points of the submanifold is the fiber of the horizontal distribution.

As an example, in the distribution given in example 3, the image of any integral curve of the vector field that determines the distribution over the manifold is an integral manifold for that distribution. A particular example (taken from [5], example 19.1) of a distribution and its family of integral manifolds is given below.

Example 4 (Orthogonal complement of a given vector field). Let $\mathcal{H}$ be the distribution over $\mathbb{R}^n$ determined by the radial vector field $x^i \partial/\partial x^i$, and let $\mathcal{H}^\perp$ be its perpendicular bundle, *i.e.,* the distribution whose fibers are the orthogonal complement of the fibers of $\mathcal{H}$. $\mathcal{H}^\perp$ is a distribution over $\mathbb{R}^n$, and the sphere centered at 0, of radius $|x|$, for all $x \neq 0$ is an integral submanifold of $\mathcal{H}^\perp$.

There are distributions over a manifold that do not have integral manifolds, as is shown in the following example (taken from [5], Ch. 19, example 19.1).

Example 5 (Distribution without integral manifolds). Consider the distribution $\mathcal{H}$ over $\mathbb{R}^3$ spanned by the following vector fields:

$$X = \frac{\partial}{\partial x} + y\frac{\partial}{\partial z}, \quad Y = \frac{\partial}{\partial y}.$$

This distribution does not have integral manifolds. To see this, suppose that N is an integral manifold that contains an arbitrary point $\mathbf{x} \in \mathbb{R}^3$. Because X and Y are tangent to N, any integral curve of X or Y that starts in N stays in N, at least for a short time (*i.e.*, the integral curve γ must be defined over a sufficiently small interval such that its image is fully contained in N).

Therefore, as the integral curve of X is a straight line contained in the plane that is parallel to the $xz-$plane, there is a segment of this line that is contained in N, and for every point in this segment, the integral curves of Y that passes through them are straight
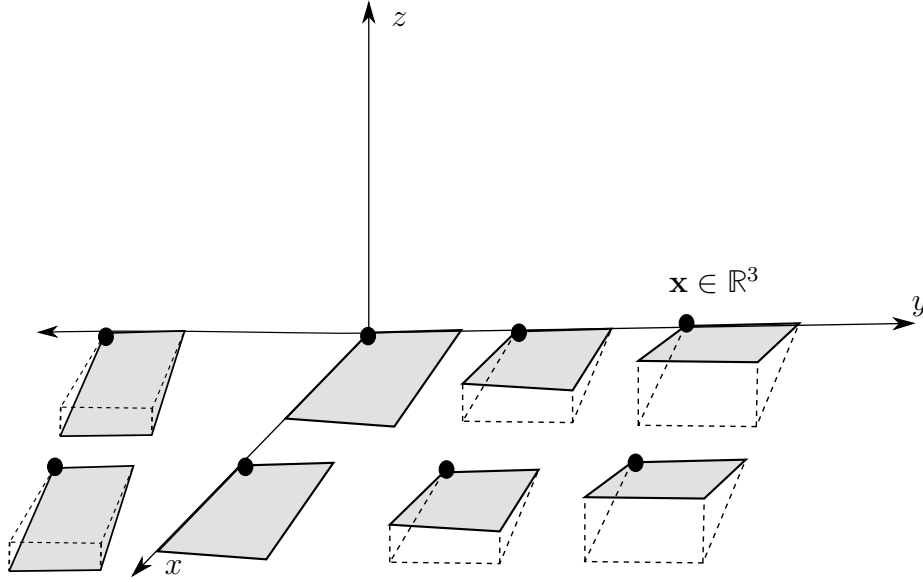
FIGURE 1. Illustration of the planes generated at each point by the integral curves of the smooth vector fields that determine the distribution given in example 5.

lines parallel to the $y-$axis, so there must be an open neighborhood of $\mathbf{x}$ that contains a plane generated by the segments of the integral curves of X and Y that passes through $\mathbf{x}$ (see figure 1). However, the tangent plane at any point p of this plane off of the segment induced by the integral curve of X is not equal to $\mathcal{H}_p$, since X depends of y.

Definition 7. A smooth distribution $\mathcal{H}$ is said to be *involutive* if given any pair of smooth vector fields X, Y defined on a open subset of M that satisfy $X_p, Y_p \in \mathcal{H}_p$, their Lie bracket also satisfies the same condition, *i.e.*, it is tangent to the distribution in the given open subset.

Let $\Gamma(\mathcal{H})$ be the set of all smooth vector fields globally defined, tangent to $\mathcal{H}$. It is clear from the definition of involutivity that if $\mathcal{H}$ is involutive, $\Gamma(\mathcal{H}) \subset \mathfrak{X}(M)$ is a Lie subalgebra. The converse statement (if $\Gamma(\mathcal{H})$ is a Lie algebra, then $\mathcal{H}$ is involutive) is also true, and can be proved extending any pair of smooth vector fields locally defined, tangent to $\mathcal{H}$, to a pair of vector fields in $\Gamma(\mathcal{H})$, using an adequate bump function.

A smooth distribution $\mathcal{H}$ over a manifold M is called *integrable* if each point of M is contained in an integral manifold of $\mathcal{H}$. It is clear then that *every integrable distribution is involutive*, because every pair of smooth vectors tangent to $\mathcal{H}$, defined over an open set $U \subset M$, satisfy that their Lie bracket is also tangent to $\mathcal{H}$ in that open set, since there exists a integral manifold N of $\mathcal{H}$ for every point of U such that the pair of smooth vectors are tangent to N, and in consequence, their Lie bracket is also tangent to N.
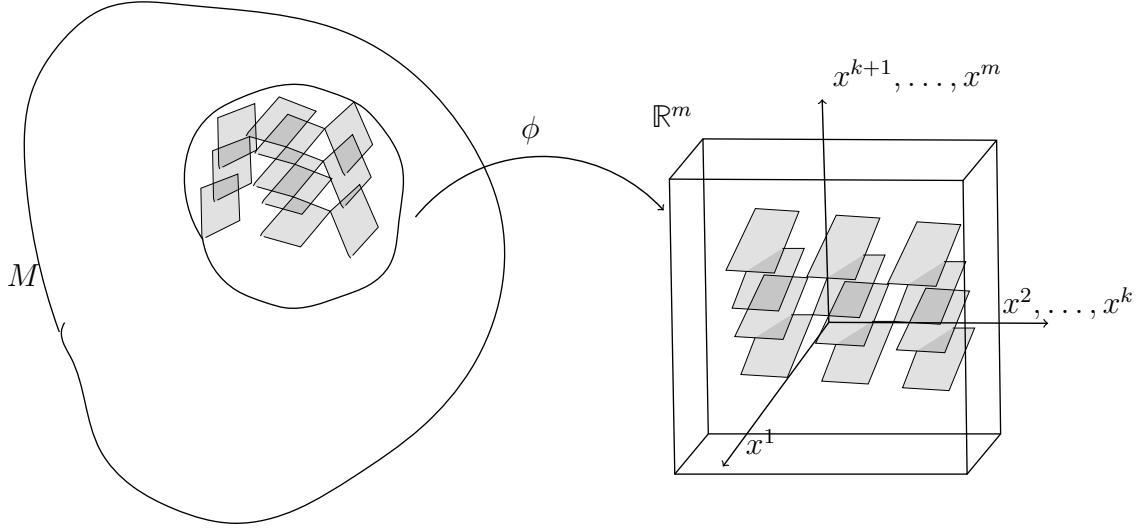
FIGURE 2. Flat chart for a distribution.

In fact, the involutivity condition does not need to be checked for all smooth vector fields on $\Gamma(\mathcal{H})$, but it suffices to check if the Lie brackets of pairs of smooth vector fields of the local frame that determines the distribution are tangent to it, as the following lemma states ([5], lemma 19.4).

Lemma 1 (Local frame criterion for involutivity). *Let $\mathcal{H} \subset TM$ be a distribution. If for every point in M there is a neighborhood such that there exists a local smooth frame $(X_1, \ldots, X_k)$ for $\mathcal{H}$ such that $[X_i, X_j]$ is a section of $\mathcal{H}$ for each i, j, then $\mathcal{H}$ is involutive.*

Definition 8. Given a smooth distribution $\mathcal{H} \subset TM$ of rank k, it is said that a smooth coordinate chart (U, ϕ) of M is *flat for $\mathcal{H}$* if $\phi(U)$ is a cube in $\mathbb{R}^m$ (being m the dimension of M), and at points of U, $\mathcal{H}$ is spanned by the first k coordinate vector fields $\partial/\partial x^1, \ldots, \partial/\partial x^k$ (see figure 2).

The geometrical meaning of this property for a distribution is that, locally (in that chart), the manifold is homeomorphic to a cube in $\mathbb{R}^m$, and the vector subspaces induced by the distribution at each point fit together as affine spaces of dimension given by the rank of the distribution in $\mathbb{R}^m$. Also, each slice of the form $x^{k+1} = c^{k+1}, \ldots, x^m = c^m$, for constants $c^{k+1}, \ldots, c^m$ in the image of U of such a chart is an integral manifold of the distribution generated by $\partial/\partial x^1, \ldots, \partial/\partial x^k$, and the preimage of this integral manifold is an integral manifold of $\mathcal{H}$, contained in U.

The case when every point in M has a coordinate chart centered in it, flat for $\mathcal{H}$ is discussed in the following definition and geometrical interpretation.

Definition 9. A smooth distribution $\mathcal{H} \subset TM$ is said to be *completely integrable* if there exists a flat chart for $\mathcal{H}$ in a neighborhood of each point of M.

Therefore, if a distribution is completely integrable, is then integrable, since there exists an integral manifold for each point (that in particular is known as *maximal integral manifold*, as the dimension of this manifold coincides with the rank of the distribution), given by the integral manifold that contains this point in the image of the coordinate chart flat for $\mathcal{H}$. In this way, *if a distribution is completely integrable, it is integrable, and therefore involutive.* In fact, these implications are actually equivalences, thanks to Frobenius's theorem.

Theorem 1 (Frobenius). *Every involutive distribution is completely integrable.*

Proof. ([5], Ch. 19, theorem 19.12) First, it will be shown that any involutive distribution is spanned by independent smooth commuting vector fields, and as a consequence of this fact, the distribution is completely integrable.

Let $\mathcal{H}$ be an involutive distribution of rank k on an $m-$dimensional manifold M, and let $p \in M$. Let (U, φ) be a smooth coordinate chart centered in p. Then, $\varphi(U) \subset \mathbb{R}^m$, and let $X_1, \ldots, X_k$ be a smooth local frame for $\mathcal{H}$ in U.

Now, $\{d\varphi(X_i)\}_{i=1}^{k}$ is a set of linearly independent vector fields over $\varphi(U)$. One can reordinate the coordinates of $\mathbb{R}^m$ to make $\mathcal{H}'_{\varphi(p)}$ (understood as the fiber of the distribution whose local frame is given by $\{d\varphi(X_i)\}_{i=1}^{k}$) complementary to the subspace generated by $\left(\partial/\partial x^{k+1}|_{\varphi(p)}, \ldots, \partial/\partial x^m|_{\varphi(p)}\right)$.

Let $\pi : \mathbb{R}^m \to \mathbb{R}^k$ the projection onto the first k coordinates, that is, $\pi(x^1, \ldots, x^m) = (x^1, \ldots, x^k)$. This map induces a smooth (smooth because it is the composition $\mathcal{H}' \hookrightarrow T\varphi(U) \xrightarrow{d\pi} T\pi(\varphi(U))$) bundle homomorphism $d\pi : T\mathbb{R}^m \to \mathbb{R}^k$ that acts as follows:

$$d\pi \left(\sum_{i=1}^{m} v_i \frac{\partial}{\partial x^i}\bigg|_{\varphi(q)}\right) = \sum_{i=1}^{k} v_i \frac{\partial}{\partial x^i}\bigg|_{\pi(\varphi(q))}, \text{ for } q \in U.$$

By the choice of coordinates, $\mathcal{H}'_{\varphi(p)} \subset T_{\varphi(p)}\mathbb{R}^m$ is complementary to the kernel of $d\pi_{\varphi(p)}$, so the restriction $d\pi|_{\mathcal{H}'_{\varphi(p)}}$ is bijective. By continuity, the same is true for $d\pi|_{\mathcal{H}'_{\varphi(q)}}$, with $q \in U$, and therefore, the matrix entries of $\left(d\pi|_{\mathcal{H}'_{\varphi(q)}}\right)^{-1} : T_{\pi(\varphi(q))}\mathbb{R}^k \to \mathcal{H}'_{\varphi(q)}$ are smooth, which makes the map smooth. With this in mind, the set of vector fields $V_1, \ldots, V_k$, defined by

$$V_i|_{\varphi(q)} = \left(d\pi|_{\mathcal{H}'_{\varphi(q)}}\right)^{-1} \frac{\partial}{\partial x^i}\bigg|_{\pi(\varphi(q))}$$

are a local frame of $\mathcal{H}'$. It is in fact a smooth commuting local frame, as will be seen below.

First, notice that V_i and $\partial/\partial x^i$ are $\pi-$related, for $i = 1, \dots, k$, because

$$\left.\frac{\partial}{\partial x^i}\right|_{\pi(\varphi(q))} = d\pi|_{\mathcal{H}'_{\varphi(q)}}\left(V_i|_{\varphi(q)}\right) = d\pi_{\varphi(q)}(V_i|_{\varphi(q)}),$$

so by the naturality of Lie brackets,

$$d\pi_{\varphi(q)}\left([V_i, V_j]_{\varphi(q)}\right) = \left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right]_{\pi(\varphi(q))} = 0,$$

but as $d\pi|_{\mathcal{H}'_{\varphi(q)}}$ is injective and $[V_i, V_j]_{\varphi(q)} \in \mathcal{H}'_{\varphi(q)}$, since $\mathcal{H}$ is involutive, $[V_i, V_j]_{\varphi(q)} = 0$ for all $q \in U$, so $\{V_i\}_{i=1}^k$ form a smooth commuting frame for $\mathcal{H}'$, and $\{d\varphi^{-1}(V_i)\}_{i=1}^k$ are a smooth commuting local frame for $\mathcal{H}$.

Finally, to show that the existence of the commuting frame $\{V_i\}_{i=1}^k$ for $\mathcal{H}'$ is a sufficient condition for $\mathcal{H}$ to be completely integrable, let θ_i be the flow of V_i, $i = 1, \dots, k$. There is a neighborhood W of $\varphi(p)$, contained in $\varphi(U)$ such that the composition $(\theta_1)_{t_1} \circ (\theta_2)_{t_2} \circ \cdots \circ (\theta_k)_{t_k}$ is well defined, for sufficiently small $t_1, \dots, t_k$ ($|t_i| < \epsilon$, for an adequate $\epsilon > 0$). Define $\Omega \subset \mathbb{R}^{m-k}$ as

$$\Omega = \left\{(s^{k+1}, \dots, s^m) \in \mathbb{R}^{m-k} \ : \ (0, \dots, 0, s^{k+1}, \dots, s^m) \in W\right\},$$

and $\Phi : (-\epsilon, \epsilon) \times \Omega \to \varphi(U)$ as

$$\Phi(s^1, \dots, s^k, s^{k+1}, \dots, s^m) = (\theta_1)_{s^1} \circ \cdots \circ (\theta_k)_{s^k}(0, \dots, 0, s^{k+1}, \dots, s^m).$$

Notice that, by construction, $\Phi(\{0\}^k \times \Omega) = (0^k \times \mathbb{R}^{m-k}) \cap W$. Moreover, $\partial/\partial s^i$ and V_i are $\Phi-$related for $i = 1, \dots, k$, because, as the flows commute, for a given $s^0 = (s^1, \dots, s^m) \in (-\epsilon, \epsilon)^k \times \Omega$,

$$d\Phi_{s^0}\left(\left.\frac{\partial}{\partial s^i}\right|_{s^0}\right)f = \left.\frac{\partial}{\partial s^i}\right|_{s^0} f\left(\Phi\left(s^1, \dots, s^m\right)\right)$$

$$= \left.\frac{\partial}{\partial s^i}\right|_{s^0} f\left((\theta_1)_{s^1} \circ \cdots \circ (\theta_k)_{s^k}(0, \dots, 0, s^{k+1}, \dots, s^m)\right)$$

$$= \left.\frac{\partial}{\partial s^i}\right|_{s^0} f\left((\theta_i)_{s^i} \circ \cdots \circ (\theta_{i-1})_{s^{i-1}} \circ (\theta_{i+1})_{s^{i+1}} \circ \cdots \right.$$
$$\left. \circ (\theta_k)_{s^k}(0, \dots, 0, s^{k+1}, \dots, s^m)\right),$$

and for any $q \in \varphi(U)$, $t \mapsto (\theta_i)_t(q)$ is an integral curve of V_i, so the above expression is equal to $V_i|_{\Phi(s^0)}f$, which shows that $\partial/\partial s^i|_{s^0}$ and V_i are $\Phi-$related.

By the previous computations,

$$d\Phi_0\left(\left.\frac{\partial}{\partial s^i}\right|_0\right) = V_i|_{\varphi(p)}, \ i = 1, \dots, k,$$

and on the other hand, since $\Phi(0, \dots, 0, s^{k+1}, \dots, s^m) = (0, \dots, 0, s^{k+1}, \dots, s^m)$, it follows that

$$d\Phi_0\left(\left.\frac{\partial}{\partial s^i}\right|_0\right) = \left.\frac{\partial}{\partial x^i}\right|_{\varphi(p)}, \ i = k+1, \dots, m.$$
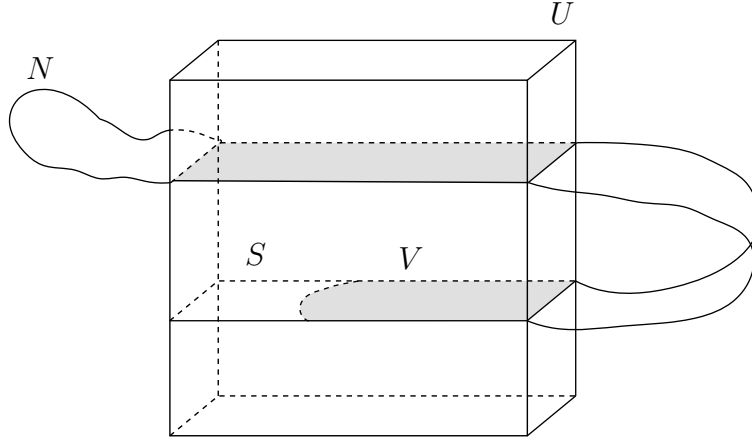
FIGURE 3. Local structure of an integral manifold.

Therefore, $d\Phi_0$ takes the basis $(\partial/\partial s^1|_0, \ldots, \partial/\partial s^m|_0)$ of $T_0\mathbb{R}^m$ to the basis $(V_1|_{\varphi(p)}, \ldots, V_k|_{\varphi(p)}, \partial/\partial x^{k+1}|_{\varphi(p)}, \ldots, \partial/\partial x^m|_{\varphi(p)})$ of $T_{\varphi(p)}\varphi(U)$. By the inverse function theorem, Φ is a local diffeomorphism (in a neighborhood of 0), and $\phi = \Phi^{-1}$ is a smooth coordinate chart that takes V_i to $\partial/\partial s^i$, for $i = 1, \ldots, k$, and takes $\partial/\partial x^i$ to $\partial/\partial s^i$, for $i = k+1, \ldots, m$. Thus, the smooth coordinate chart flat for $\mathcal{H}$ in a neighborhood of $p \in M$ is $\phi \circ \varphi$, and since this is independent of p, for every point there is a smooth coordinate chart flat for $\mathcal{H}$, what makes $\mathcal{H}$ a completely integrable distribution.

$\square$

The next proposition is one of the main consequences of Frobenius theorem, and it is fundamental to study foliations, topic that is going to be discussed below.

Proposition 3 (Local structure of integral manifolds). *Let $\mathcal{H}$ be an involutive distribution of rank k on a smooth manifold M, and let (U, φ) be a flat chart for $\mathcal{H}$. If N is any connected integral manifold of $\mathcal{H}$, then $\varphi(U \cap N)$ is the union of countably many disjoint open subsets of parallel $k-$dimensional slices of $\varphi(U)$, whose preimages are open in N and embedded in M.*

Proof. ([5], Ch. 19, proposition 19.16) Let N be an integral manifold of $\mathcal{H}$. Since the inclusion map $\iota : N \hookrightarrow M$ is continuous, $\varphi(\iota^{-1}(U)) = \varphi(N \cap U)$ is open in $\varphi(N)$. Then, $\varphi(U \cap N)$ is the union of open slices (*i.e.*, open subsets of a slice of $\varphi(U)$) with $x_i = $ constant, for $i = k+1, \cdots, m$. Moreover, this union is at most countable, because N is second countable, and since $\varphi(N \cap U)$ is a union of open slices, $\pi(\varphi(U \cap N))$ consists

of a countable number of points (recall that $\pi : \varphi(U) \to \mathbb{R}^{m-k}$ is the projection onto the last $m - k$ coordinates) in $\mathbb{R}^{m-k}$, which implies that the union of open slices is at most countable.

Finally, if V is an open slice in the slice S, the inclusion map $\varphi^{-1}(V) \hookrightarrow M$ is a smooth embedding, because it is the composition of smooth embeddings $\varphi^{-1}(V) \hookrightarrow \varphi^{-1}(S) \hookrightarrow M$. $\qquad\square$

The last proposition implies that one can put all the maximal integral manifolds of an involutive distribution of rank k together, to obtain a partition on M into $k-$dimensional submanifolds, that satisfy the following definition.

Definition 10. A smooth chart (U, φ) for M is called *flat for a collection $\mathcal{F}$ of $k-$dimensional submanifolds of M* if $\varphi(U)$ is a cube in $\mathbb{R}^m$, and the image of each submanifold via φ intersects $\varphi(U)$ in either the empty set or in a countable union of $k-$dimensional slices of the form $x^{k+1} = c^{k+1}, \ldots, x^m = c^m$.

Definition 11. A *foliation of dimension k on a smooth manifold M* is a collection $\mathcal{F}$ of disjoint, connected, nonempty, immersed $k-$dimensional submanifolds of M (called the *leaves* of the foliation), whose union is M, and such that in a neighborhood of each point $p \in M$ there is a flat chart for $\mathcal{F}$.

The following are examples of foliations over a manifold, mainly taken from [5], Ch. 19, example 19.18.

Example 6 (Collection of affine subspaces). The collection of all $k-$dimensional affine subspaces of $\mathbb{R}^m$ parallel to $\mathbb{R}^k \times \{0\}$ is a $k-$dimensional foliation for $\mathbb{R}^m$.

Example 7 (Spheres centered at the origin). The collection of all spheres centered at 0 is an $(m - 1)-$dimensional foliation of $\mathbb{R}^m \setminus \{0\}$.

Example 8 (Cartesian product of manifolds). If M and N are connected smooth manifolds, the collection of subsets of the form $M \times \{q\}$, with $q \in N$, is a foliation of $M \times N$, each of whose leaves is diffeomorphic to M.

Example 9 (Foliations on a torus). The torus $T = \mathbb{S}^1 \times \mathbb{S}^1$ can be endowed with the distribution induced by the cartesian product of manifolds (see example 8). In this case, the foliation is conformed by copies of $\mathbb{S}^1$. The horizontal curves are segments of this copies, and if two points lay in different copies, there is not a horizontal curve that connects them.

From the definition of foliations, it is clear that *if $\mathcal{F}$ is a foliation on a smooth manifold M, the collection of tangent spaces to the leaves of $\mathcal{F}$ forms an involutive distribution on M*. In a reciprocal way, the collection of maximal integral manifolds of an involutive distribution forms a foliation over the manifold, as it is stated in the global Frobenius theorem.

Theorem 2 (Global Frobenius theorem)**.** *Let $\mathcal{H}$ be an involutive distribution on a smooth manifold M. The collection of all maximal connected integral manifolds of $\mathcal{H}$ forms a foliation of M.*

The next lemma ([5], Ch. 19, lemma 19.22) is going to be used to prove this theorem.

Lemma 2. *Suppose that $\mathcal{H} \subset TM$ is an involutive distribution, and let $\{N_\alpha\}_{\alpha \in A}$ be any collection of connected integral manifolds of $\mathcal{H}$ with a point in common. Then, $N = \bigcup_\alpha N_\alpha$ has a unique smooth manifold structure making it into a connected integral manifold of $\mathcal{H}$.*

Proof of the global Frobenius theorem. ([5], Ch. 19, theorem 19.21) For each $p \in M$, let L_p be the union of all connected integral manifolds of $\mathcal{H}$ containing p. By lemma 2, L_p is a connected integral manifold of $\mathcal{H}$ containing p, and it is clearly maximal. By maximality, if L_p and $L_{p'}$ intersect, $L_p = L_{p'}$, because $L_p \cup L_{p'}$ is an integral manifold containing both p and p'. Thus, the maximal connected integral manifolds are either disjoint or identical.

If (U, φ) is any flat chart for $\mathcal{H}$, then $\varphi(L_p \cap U)$ is a countable union of open subsets of slices (by proposition 3). For any such slice S, if $\varphi(U \cap L_p) \cap S$ is neither empty nor all of S, then $L_p \cup \varphi^{-1}(S)$ is a connected integral manifold properly containing L_p, which contradicts the maximality of L_p. Therefore, $\varphi(L_p \cap U)$ is precisely a countable union of slices, so the collection $\{L_p \ : \ p \in M\}$ is the desired foliation. $\qquad\square$

With the global Frobenius theorem, one can answer the question at the start of this subsection: in general, for an given distribution there is no smooth horizontal curve that connects an arbitrary pair of points, because the points can be in different leaves of the foliation given by the distribution.

2. Sub-Riemannian Geometry

2.1. **Sub-Riemannian Structure and Geodesics.**

Definition 12. A *sub-Riemannian structure over a manifold M* is a pair $(\mathcal{H}, \langle \cdot, \cdot \rangle)$, where $\mathcal{H} \subset TM$ is a distribution and $\langle \cdot, \cdot \rangle$ is a section of the bundle $T_2^0 \mathcal{H} \xrightarrow{\pi} M$, whose values are positive definite symmetric bilinear forms (see [7], Ch. 1, page 12).

The distribution $\mathcal{H}$ is called *horizontal*, and the tensor field $\langle \cdot, \cdot \rangle$ is called *metric*.

Some examples of sub-Riemannian structures are given below.

Example 10 (Riemannian Structure)*.* Every Riemannian structure is in particular a sub-Riemannian structure, where the distribution is the entire tangent bundle.

Example 11 (Heisenberg Group)*.* The distribution of the Heisenberg group is commented in example 2. The inner product over a fiber $\mathcal{H}_{(x,y,z)}$, with $(x, y, z) \in \mathbb{R}^3$ is given by $\langle \cdot, \cdot \rangle : \mathcal{H}_{(x,y,z)} \times \mathcal{H}_{(x,y,z)} \to \mathbb{R} : (v, w) \mapsto v_1 w_1 + v_2 w_2$, where $v = (v_1, v_2, v_3)$ and $w = (w_1, w_2, w_3)$.

Example 12 (Vector Field over a Manifold). As seen in example 3, any smooth vector field $X : M \to TM$ that does not cancel out in any point determines a distribution. The fiber inner-product $\langle \cdot, \cdot \rangle : \mathcal{H}_p \times \mathcal{H}_p \to \mathbb{R}$ for $p \in M$ is given by $\langle \lambda_1 X_p, \lambda_2 X_p \rangle = \lambda_1 \lambda_2$.

In the particular case of the horizontal curves, one can define the *length of a horizontal smooth curve* γ (denoted by $\ell(\gamma)$) as in the case of Riemannian geometry:

$$\ell(\gamma) = \int ||\dot{\gamma}|| dt,$$

since the tangent vectors of the curve at all points are in the fibers of the horizontal distribution, where there is a inner product defined, and $||\dot{\gamma}|| = \sqrt{\langle \dot{\gamma}, \dot{\gamma} \rangle}$.

Definition 13. The *distance between two points* $p, q \in M$, denoted by $d(p, q)$, is defined as the infimum of the lengths of all absolutely continuous horizontal curves that begin in p and end in q, that is,

$$d(p, q) = \inf \left\{ \ell(\gamma) \mid \gamma : [0, 1] \to M \text{ is absolutely continuous and } \gamma(0) = p, \ \gamma(1) = q \right\}.$$

The distance between two points is said to be infinite if there is no horizontal curve joining them.

The curves that are considered in the previous definition of distance are required to be absolutely continuous, that is a weaker condition of smoothness, and refers to a curve that is differentiable in almost all points of its domain. Nevertheless, the distance remains invariant if we consider instead the set of all smooth curves (see [6], page 23), but the motivation to consider the bigger set of absolutely continuous curves is the fact that the curves that realize the distance between two points are not always smooth.

Definition 14. Given a sub-Riemannian structure $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ over a manifold M (or Riemannian structure, if $\mathcal{H} = TM$), it is said that an absolutely continuous horizontal curve $\gamma : [a, b] \to M$, with $\gamma(a) = p, \gamma(b) = q$ is a *geodesic* if it realizes the distance between p and q, that is,

$$\ell(\gamma) = d(p, q).$$

2.2. **Chow's Theorem.** The section 1.2 provides a class of (involutive) distributions where it is not true that every two points can be connected through a horizontal curve. Now, it is worthwhile to ask for a sufficient condition for a distribution to have the connectedness property through horizontal curves. To discuss such a condition, we need to cover some preliminaries.

Definition 15. A distribution $\mathcal{H} \subset TM$ is called *bracket generating* if for every $p \in M$, there is a local frame $X_1, \ldots, X_k : U \to TM$ of $\mathcal{H}$ such that

$$TU = \operatorname{span} \left\{ [X_{i_1}, \ldots, [X_{i_{j-1}}, X_{i_j}]] \ : \ i_1, \ldots, i_j = 1, \ldots, k; \ j \in \mathbb{N} \right\}.$$

Lemma 3. *Let X_1, X_2 be two smooth vector fields over the manifold M, with respective local flows $\Phi_1, \Phi_2 : (-\epsilon, \epsilon) \times U \to U$, and let $p \in U$. Then, in any coordinate system the following relation holds:*

$$[\Phi_1(t), \Phi_2(t)]\,(p) := \left(\Phi_1(t) \circ \Phi_2(t) \circ \Phi_1(t)^{-1} \circ \Phi_2(t)^{-1}\right)(p) = p + t^2\,[X_1, X_2]\,(p) + O(t^3).$$

Proof. By Taylor's theorem, we have

$$\Phi_i(t)(p) = p + tX_i(p) + \frac{t^2}{2}X_i(X_i(p)) + O(t^3), \text{ for } i = 1, 2.$$

In the other hand, as $\{\Phi_i(t)\}_{t \in (-\epsilon, \epsilon)}$ is the family of diffeomorphisms associated with the flow of X_i, $\Phi_i(t)^{-1} = \Phi_i(-t)$. Then,

$$\begin{aligned}
(\Phi_1(t)^{-1} \circ \Phi_2(t)^{-1})(p) = p &- t\,(X_1(p) + X_2(p)) + t^2 X_1(X_2(p)) \\
&+ t^2\,(X_1(X_1(p)) + X_2(X_2(p))) + O(t^3),
\end{aligned}$$

and evaluating $\Phi_1(t) \circ \Phi_2(t)$ in this point gives

$$\begin{aligned}
\left(\Phi_1(t) \circ \Phi_2(t) \circ \Phi_1(t)^{-1} \circ \Phi_2(t)^{-1}\right)(p) = p &+ t^2\,[X_1, X_2]\,(p) + t^2\,(X_1(X_1(p)) + X_2(X_2(p))) \\
&- t(X_1(p) + X_2(p)) - t^2\,(X_1(X_1(p)) + X_2(X_2(p))) \\
&+ t(X_1(p) + X_2(p)) + O(t^3) \\
= p &+ t^2\,[X_1, X_2]\,(p) + O(t^3),
\end{aligned}$$

which is the desired result. □

Chow's theorem guarantees that any bracket generating over a connected manifold has the connectedness property through horizontal curves, and we going to derive this result in the rest of this section, following closely the procedure of [6], Section 2.4. First of all, choose a base point $p \in M$ and a local orthonormal frame X_i, $i = 1, \ldots, k$ for the distribution $\mathcal{H}$. Let Φ_i be their respective flows. These flows can be used to move in the horizontal directions (points connected to p through integral curves of the orthonormal frame), using the formula $\Phi_i(t)(q) = q + tX_i(q) + O(t^2)$ for sufficiently small t.

The assignments $t \mapsto \Phi_i(t)(q)$ are horizontal curves for the distribution $\mathcal{H}$, and are called *simple horizontal curves.* As the frame is orthonormal, the length of a simple horizontal curve with $0 \le t \le \epsilon$ is ϵ.

Applying $\Phi_k(t_k) \circ \cdots \circ \Phi_2(t_2) \circ \Phi_1(t_1)$ to p and letting $t_1, \ldots, t_k$ vary over the cube $|t_i| \le \epsilon$, we move from p to another point inside the $k-$dimensional cube of volume $(2\epsilon)^k$ in the coordinate chart, whose sides are determined by the local frame in the image of p. Each point of this cube can be reached from p through the concatenation of k or fewer flows, and this cube is inside of the sub-Riemannian ball of radius $k\epsilon$, that is,

$$\begin{aligned}
\text{Box}_p(\epsilon) := \{\Phi_k(t_k) \circ \cdots &\circ \Phi_2(t_2) \circ \Phi_1(t_1)(p) \mid |t_i| < \epsilon, \ i = 1, \ldots, k\} \\
&\subset \text{B}(k\epsilon, p) := \{q \in M \mid d(p, q) < k\epsilon\}.
\end{aligned}$$

We can move in the remaining directions along horizontal paths by successive applications of the commutators of flows, based on the bracket generating property of the distribution in the following way. Let $\mathcal{H}$ denote also the sheaf of smooth vector fields tangent to $\mathcal{H}$, that is, the association between open sets $U \subset M$ and horizontal smooth vector fields $\mathcal{H}(U)$ defined locally on U, and define the new vector sheaves

$$\mathcal{H}^2 := \mathcal{H} \oplus [\mathcal{H}, \mathcal{H}], \quad \mathcal{H}^{r+1} := \mathcal{H}^r \oplus [\mathcal{H}, \mathcal{H}^r],$$

where

$$\left[\mathcal{H}, \mathcal{H}^j\right] := \mathrm{span}\left\{[X, Y] \mid X \in \mathcal{H}, \ Y \in \mathcal{H}^j\right\}, \ \text{for } j = 1, 2, \ldots$$

The movement in the $\mathcal{H}^2/\mathcal{H}$ directions can be made along horizontal paths by applying the commutators $\Phi_{ij}(t) := [\Phi_i(t), \Phi_j(t)]$ to p. As $\Phi_{ij}(t)(p) = p + t^2[X_i, X_j](p)$ (by lemma 3) if $|t| \leq \epsilon$, we will move by an amount ϵ^2 in the $\mathcal{H}^2/\mathcal{H}$ directions. This process is realized inductively until exhausting the tangent space. Let $I = (i_1, i_2, \ldots, i_n)$, $1 \leq i_j \leq k$, and $X_I := [X_{i_1}, X_J]$, where $J = (i_2, \ldots, i_n)$, so that X_I is the iteration of n Lie brackets. Similarly, define flows $\Phi_I(t) := [\Phi_{i_1}(t), \Phi_J(t)]$. By successive application of lemma 3, we have that

$$\Phi_I(t) = \mathbb{1} + t^n X_I + O(t^{n+1}).$$

Due to the bracket generating condition over any local frame of the distribution, we can select a local frame for the entire tangent bundle amongst the X_I, letting n and $i_1, \ldots, i_n$ vary accordingly. We choose such a frame and relabel it as Y_i, $i = 1, \ldots, m$, in a way that it satisfies: $Y_1 = X_1, \ldots, Y_k = X_k$ span $\mathcal{H}$ near p, $\{Y_1, \ldots, Y_{n_2}\}$ span $\mathcal{H}^2$ near p, $\{Y_1, \ldots, Y_{n_2}, \ldots, Y_{n_3}\}$ span $\mathcal{H}^3$ near p, and so on. The tuple $(k, n_2, n_3, \ldots, n_r)$ is called the *growth vector of the distribution at* p, and the smallest integer r is called the *step or degree of nonholonomy of the distribution at* p.

Definition 16 ([6], page 49.)**.** Let $\{Y_1, \ldots, Y_m\}$ be the local frame of the tangent bundle in a neighborhood of $p \in M$ constructed as stated before using the local frame of a bracket generating distribution. Let $Y_i = X_I$, $I = (i_1, \ldots, i_{n_I})$. We denote the length of I as $w_i := |I|$, and define the assignment $i \mapsto w_i$, called the *weighting associated to the growth vector*.

We relabel the flows Φ_I in a similar fashion, so that Φ_i is the flow associated to Y_i. In this case, each point $\Phi_i(t)(p)$ is the endpoint of the concatenation of w_i simple horizontal curves, each one of length t. If we impose that $|t| \leq \epsilon$, then $\Phi_i(t)(p)$ lies in the ball of radius $w_i \epsilon$ centered at p. Moreover, in coordinates we have that

$$\Phi_i(t)(p) = p + t^{w_i} X_i(p) + O(t^{w_i+1}),$$

so $\Phi_i(t)(p)$ lies in the Euclidean box of volume ϵ^{w_i}, in the $\mathcal{H}^{w_i}$ directions. From this we can deduct that the sub-Riemannian ball $B(\epsilon, p)$ contains an Euclidean coordinate box whose sides are of order ϵ^{w_i} in the i-th direction. This result is known as the ball-box theorem, and it will be stated and proved after the following definitions.

Definition 17 ([6], definition 2.4.1)**.** Coordinates $y_1, \ldots, y_m$ are said to be *linearly adapted to the distribution* $\mathcal{H}$ *at* p if $\mathcal{H}^i(p) := \{X(p) \mid X$ is a section of $\mathcal{H}$ defined on $p\}$ is

annihilated by the differentials $dy_{n_i+1}, \ldots, dy_m$ at p, where $n_i = n_i(p)$ are the coordinates of the growth vector at p.

Definition 18. The w- weighted box of size ϵ is the point set

$$\mathrm{Box}^w(\epsilon) := \{y \in \mathbb{R}^m \mid |y_i| \leq \epsilon^{w_i}; \ i = 1, \ldots, m\}.$$

Set y_i as the coordinates centered at p such that $dy_i(p)$ are the dual basis to the $Y_i(p)$, so that these coordinates are linearly adapted to the distribution by definition.

Theorem 3 (Ball-Box Theorem). *There exist linearly adapted coordinates $y_1, \ldots, y_m$ and positive constants $c < C$, $\epsilon_0 > 0$, such that for all $\epsilon < \epsilon_0$,*

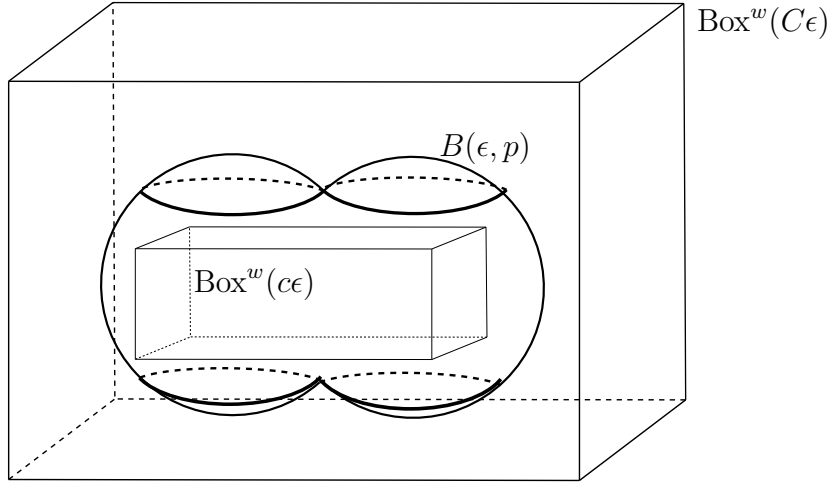$$\mathrm{Box}^w(c\epsilon) \subset B(\epsilon, p) \subset \mathrm{Box}^w(C\epsilon).$$

FIGURE 4. Ball box theorem.

Proof of $\mathrm{Box}^w(c\epsilon) \subset B(\epsilon, p)$. ([6], theorem 2.4.2). As the approximation $\Phi_{ij}(t) = \mathbb{1} + t^2[X_i, X_j] + O(t^3)$ has the coefficient t^2 that is always positive, we can not move in the negative $[X_i, X_j]$ direction. To circunvent this, we define

$$\Psi_{ij}(t) = \begin{cases} [\Phi_i(t), \Phi_j(t)] & t \geq 0 \\ [\Phi_j(t), \Phi_i(t)] & t < 0, \end{cases}$$

so that

$$\Psi_{ij}(t) = \begin{cases} I + t^2 X_{ij} + O(t^3) & t \geq 0 \\ I - t^2 X_{ij} + O(t^3) & t < 0. \end{cases}$$

This problem occurs wheneve the number w_i associated to the flow Φ_I of Y_i is even. To solve this in general, we set

$$\Psi_I(t) = \begin{cases} \Phi_I(t) & t \geq 0 \\ [\Phi_J(t), \Phi_{i_1}(t)] & t < 0. \end{cases}$$

If w_i is odd, we keep $\Psi_I = \Phi_I$, and we relabel Ψ_I as Ψ_i accordingly. Now, we introduce the functions

$$\sigma_i(t) = \begin{cases} t^{w_i} & w_i \text{ even}, t \geq 0 \\ -t^{w_i} & w_i \text{ even}, t < 0 \\ t^{w_i} & w_i \text{ odd} \end{cases}$$

to simplify the notation of the approximation

$$\Psi_I(t) = \mathbb{1} + \sigma_i(t)Y_i + O(t^{w_i+1}).$$

Define the map $F : \mathbb{R}^m \to M$ as $F(t_1, \ldots, t_m) = (\Psi_m(t_m) \circ \cdots \circ \Psi_1(t_1))(p)$, and let y_i be linearly adapted coordinates for which the dy_i are dual to the Y_i of the basis of the tangent space at the base point p. Then, in coordinates,

$$F_i(t_1, \ldots, t_m) = \sigma_i(t_i) + o(t_i^{w_i}) = \pm t_i^{w_i} + o(t_i^{w_i}),$$

where the error $o(t_i^{w_i})$ corresponds to the composition of the previous $\Psi_1(t), \ldots, \Psi_{i-1}(t)$ flows evaluated in the big $O-$errors of the previous sequential arguments. It remains to understand and bound these error terms, and for that we introduce the new variables $s_i = \sigma_i(t)$, so that $F_i(t_1, \ldots, t_m) = s_i + o(s_i)$.

The map $\sigma = (\sigma_1, \ldots, \sigma_m)$ is a homeomorphism between neighborhoods of 0 in $\mathbb{R}^m$, so let S be the inverse of σ that is given by $S = (s_1, \ldots, s_m)$, where $s_i(t_i) = \pm|t_i|^{1/w_i}$. Then, it follows that

$$y \circ F \circ S(s_1, \ldots, s_m) = (s_1, \ldots, s_m) + o(|s|),$$

from where we deduce that $F \circ S$ is differentiable at the origin because of the differentiability of both y and $y \circ G \circ S$, and its derivative is the identity in the $s - y$ coordinates, so it is also C^1 near 0. We will proceed applying the inverse function theorem to invert $F \circ S$ in a neighborhood of p.

Write $F \circ S(s_1, \ldots, s_m)$ in coordinates as $(y_1, \ldots, y_m)$. Then, there exists a constant ϵ_0 and constants $c = c(\epsilon_0)$, $C = C(\epsilon_0)$ such that $c|s_i| \leq |y_i(s_1, \ldots, s_m)| \leq C|s_i|$ whenever $|s_i| \leq \epsilon_0$. Apply the constraint $|s_i| \leq \epsilon^{w_i}$ to the s_i. It is clear that the inverse image of the point set $|s_i| \leq \epsilon^{w_i}$ under the map S is the $\epsilon-$cube box $\text{Box}(\epsilon) = \{(t_1, \ldots, t_m) \mid |t_i| \leq \epsilon\}$, so that applying F to the $\epsilon-$cube gives

$$\text{Box}^w(c\epsilon) \subset F(\text{Box}(\epsilon)) \subset \text{Box}^w(C\epsilon).$$

Each curve in $F(\text{Box}(\epsilon))$ is the endpoint of a horizontal curve starting at p whose length is less than $M\epsilon$, where $M = M(w)$ counts the number of concatenations involved in F. Consequently,

$$F(\text{Box}(\epsilon)) \subset B(M\epsilon, p),$$

from where we get that $\text{Box}^w(c\epsilon) \subset B(\epsilon, p)$. $\qquad \square$

The remaining containment in the Ball-Box theorem will not be proved in this text, as the proof of Chow's theorem only requires the first containment. The proof of the remaining containment can be found in [6], section 2.7.

With the Ball-Box theorem, Chow's theorem can be derived easily.

Theorem 4 (Chow's Theorem). *If $\mathcal{H}$ is a bracket generating distribution on a connected manifold M, then any two points of M can be connected by a horizontal path.*

Proof. ([6], page 52). The containment $\text{Box}^w(c\epsilon) \subset B(\epsilon, p)$ shows that $B(\epsilon, p)$ is a neighborhood of p, which means that the accesible points from p through horizontal curves form a neighborhood of p. Let $q \in M$ be another arbitrary point on the connected manifold, and take a smooth curve γ connecting p and q.

The image of γ is compact, so we can cover it with finitely many succesive box neighborhoods, where the Ball-Box theorem holds. Denote these neighborhoods as $U_1, \ldots, U_n$, with each one of these boxes centered on succesive points p_i along the image of γ, and $p_1 = p$, $p_n = q$. Moreover, these open sets can be chosen in orden to satisfy $U_i \cap U_{i+1} \neq \emptyset$, so that there exists $q_i \in U_i \cap U_{i+1}$. By the Ball-Box theorem, we have paths connecting p_i with q_i, and q_i with p_{i+1}. Concatenating these paths yields a piecewise horizontal path connecting p and q.

$\square$

Example 13 (Cartan's Distribution). Consider the distribution over $\mathbb{R}^3$ determined by the 1-form $\omega = dy - zdx$, or in dual manner, by the smooth vector fields $\partial/\partial x + z\partial/\partial y, \partial/\partial z$. From its definition it is clear that Cartan's distribution is of rank 2, so we can not know beforehand if any two pair of points can be connected through a horizontal curve. Nevertheless, as

$$\left[\frac{\partial}{\partial z}, \frac{\partial}{\partial x} + z\frac{\partial}{\partial y}\right] = \frac{\partial}{\partial y}$$

and $T\mathbb{R}^3 = \text{span}\{\partial/\partial x + z\partial/\partial y, \partial/\partial z, \partial/\partial y\}$, Cartan's distribution is bracket generating, so by Chow's theorem any two points in $\mathbb{R}^3$ can be connected by a horizontal path.

3. Sub-Riemannian metrics on Bundles

3.1. **Ehresmann connection.** Let $\pi : Q \to M$ be a submersion, which means that for all $q \in Q$, $d\pi_q : T_qQ \to T_{\pi(q)}M$ is a surjective map. The fiber $Q_m = \pi^{-1}(m)$, with $m = \pi(q)$ is a submanifold because m is a regular value of π.

Definition 19. Given a submersion $\pi : Q \to M$, the *vertical space at $q \in Q$*, denoted as V_q, is the tangent space to the fiber Q_m, with $m = \pi(q)$, that is,

$$V_q = \text{Ker}(d\pi_q) = T_qQ_m.$$

The collection of vertical spaces is a distribution $V \subset TQ$ that assigns to each $q \in Q$ the space V_q. The distribution is by construction integrable, and its integral manifolds are the fibers Q_m, $m \in \pi(Q)$.

Definition 20 ([6], definition 11.1.1). A *connection for a submersion* $\pi : Q \to M$ is a distribution that is everywhere transverse to the vertical one, that is,

$$V_q \oplus \mathcal{H}_q = T_q Q, \text{ for all } q \in Q.$$

The restriction of $d\pi_q$ to $\mathcal{H}_q$ is a linear isomorphism between $\mathcal{H}_q$ and $T_{\pi(q)}M$ because $d\pi_q$ is surjective, and by the isomorphism theorem we have that $\mathcal{H}_q \cong T_{\pi(q)}M$.

Definition 21. If M is endowed with a metric g, we can use the previous linear isomorphism to pull the inner product back to the horizontal distribution $\mathcal{H}$: the sub-Riemannian structure $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ induced by the Riemannian structure (M, g) is given by

$$\langle v, w \rangle_q = g_{\pi(q)} \left(d\pi_q(v), d\pi_q(w) \right), \text{ for all } q \in Q \text{ and } v, w \in \mathcal{H}_q.$$

This sub-Riemannian metric varies smoothly with $q \in Q$ ([6], page 194), and is referred as the *induced sub-Riemannian structure* over Q by the submersion $\pi : Q \to M$ and the Riemannian metric over M.

Definition 22. The *horizontal lift* of a curve $c : I \to M$ starting at $m \in M$ is defined as the unique curve $\gamma : I \to Q$ that is tangent to $\mathcal{H}$, starts at $q \in Q_m$, and projects to c, that is, $\pi \circ \gamma = c$.

The uniqueness of the horizontal lift follows from the fact that $\gamma : I \to Q$ is solution of the initial value problem

$$\frac{d\gamma(t)}{dt} = d\pi_{c(t)}^{-1} \left(\frac{dc(t)}{dt} \right), \quad \gamma(0) = q.$$

Example 14 (Canonical projection from $\mathbb{R}^{m+n}$ to $\mathbb{R}^m$). Consider the canonical projection $\pi : \mathbb{R}^{m+n} \to \mathbb{R}^m$ that maps the first m-coordinates of a point in $\mathbb{R}^{m+n}$ to $\mathbb{R}^m$. This map is a submersion, and its vertical space at a point $q \in \mathbb{R}^{m+n}$ is $V_q = \{(0, \ldots, 0, v_1, \ldots, v_n) : v_1, \ldots, v_n \in \mathbb{R}\} \cong \mathbb{R}^n$. A connection for this submersion is $\mathcal{H}_q = \{(v_1, \ldots, v_m, 0, \ldots, 0) : v_1, \ldots, v_m \in \mathbb{R}\} \cong \mathbb{R}^m$. Given an arbitrary curve $c : I \to \mathbb{R}^m$, starting at $\boldsymbol{p} = (p_1, \ldots, p_m)$, its horizontal lift passing through the point $(p_1, \ldots, p_m, q_1, \ldots, q_n) \in \pi^{-1}(\boldsymbol{p})$ is the curve $\gamma : I \to \mathbb{R}^{m+n}$, $t \mapsto (c(t), q_1, \ldots, q_n)$.

Remark 3. The connection $\mathcal{H}$ is called *complete* or is referred as an *Ehresmann connection* if every smooth curve $c : I \to M$ has a horizontal lift.

Proposition 4. *The induced sub-Riemannian structure satisfies the following properties:*

(1) *The sub-Riemannian length of a horizontal path on Q equals the Riemannian length of its projection to M.*

(2) *The horizontal lift of a Riemannian geodesic in M is a sub-Riemannian geodesic in Q.*

(3) *The projection π is distance decreasing, that is, $d_M(\pi(q_1), \pi(q_2)) \leq d_Q(q_1, q_2)$, for all $q_1, q_2 \in Q$.*

Proof. (1) Let $\gamma : I \to Q$ be a horizontal path to $\mathcal{H}$ starting at $q \in Q$. Then, its longitude is

$$\ell_Q(\gamma) = \int_0^1 ||\dot{\gamma}(t)||_Q dt = \int_0^1 ||d\pi_{\gamma(t)}(\dot{\gamma}(t))||_M dt,$$

having in mind that $d\pi_{\gamma(t)}$ induces an isomorphism between $\mathcal{H}_{\gamma(t)}$ and $T_{\pi(\gamma(t))}M$, and this isomorphism determines the induced sub-Riemannian structure over Q. As $c = \pi \circ \gamma$ and $d/dt(\pi \circ \gamma) = d\pi_{\gamma(t)}(\dot{\gamma}(t))$, $\ell_Q(\gamma) = \ell_M(c)$.

(2) Suppose that $c : I \to M$ is a Riemannian geodesic, and let $\gamma : I \to Q$ be its respective horizontal lift. Take any other horizontal curve $\eta : I \to Q$ with the same endpoints than γ. Then, by the first affirmation of this proposition and the fact that c is a geodesic, we have

$$\ell_Q(\eta) = \ell_M(\pi \circ \eta) \geq \ell_M(c) = \ell_Q(\gamma),$$

from where we get that $\gamma : I \to Q$ is a sub-Riemannian geodesic.

(3) Let $q_1, q_2 \in Q$. If $d_Q(q_1, q_2) = \infty$ the result is immediate. Else, let $\{\gamma_i\}_{i \in \mathbb{N}}$ be a minimizing sequence of $d_Q(q_1, q_2)$. Now consider the sequence of projected curves $\{\pi \circ \gamma_i\}_{i \in \mathbb{N}}$, from where there are two possibilities: if this last sequence is minimizing, $d_Q(q_1, q_2) = d_M(\pi(q_1), \pi(q_2))$, as $\ell_Q(\gamma_i) = \ell_M(\pi \circ \gamma)$; if it is not minimizing, $d_Q(q_1, q_2) = \lim_{i \to \infty} \ell_Q(\gamma_i) = \lim_{i \to \infty} \ell_M(\pi \circ \gamma_i) > d_M(\pi(q_1), \pi(q_2))$ by definition. Therefore, $d_Q(q_1, q_2) \geq d_M(\pi(q_1), \pi(q_2))$.

$\square$

Definition 23. A Riemannian metric on Q is said to be *compatible* with the induced sub-Riemannian metric if the algebraic splitting $TQ = V \oplus \mathcal{H}$ is an orthogonal decomposition with respect to the Riemannian metric on Q.

Remark 4. Thanks to the orthogonal algebraic splitting, a compatible Riemannian metric is determined by its restriction to V, and the induced sub-Riemannian structure restricted to $\mathcal{H}$. This fact can be turned around in the following sense. Take a Riemannian metric g_Q on Q. Let $V_q^\perp$ be the orthogonal complement of the vertical space V_q in $q \in Q$ with respect to the Riemannian metric, and set $\mathcal{H}_q = V_q^\perp$. The Riemannian metric defines inner products in $V_q^\perp$ and $T_m M$, being defined in the first as the restriction of the Riemannian metric, and in the second is given by the metric and the isomorphism $\mathcal{H}_q \cong T_m M$. Nevertheless, this inner product on $T_m M$ can depend on the point $q \in Q_m$, as the metric could vary with the point of the fiber. If that is the case, we cannot construct a Riemannian metric on M, as there is no consistent way to choose which $q \in Q_m$ is going to be used to define the inner product ([6], page 195).

Definition 24. Given a Riemannian manifold (Q, g_Q), $\pi : Q \to M$ a submersion, we say that the pair (π, g_Q) is a *Riemannian submersion* if for every $m \in M$ and $q \in Q_m$, the isomorphism $d\pi_q|_{\mathcal{H}_q}$ is an isometry between $\mathcal{H}_q$ and $T_m M$.

Example 15. Note that the canonical projection $\pi : \mathbb{R}^{m+n} \to \mathbb{R}^m$ given in example 14 is a Riemannian submersion if $\mathbb{R}^{m+n}$ is endowed with the cartesian metric.

3.2. **Metrics on Principal Bundles.**

Definition 25 ([5], page 268)**.** Let M, G, Q be manifolds. A *(locally trivial) fiber bundle* over M with fiber G is a manifold Q, together with a smooth surjective map $\pi : Q \to M$ with the property that for each $x \in M$ there is a neighborhood U of x and a diffeomorphism $\phi : \pi^{-1}(U) \to U \times G$, called a local trivialization of Q over U, such that the following diagram commutes:

$$\pi^{-1}(U) \xrightarrow{\phi} U \times G$$
$$\pi \searrow \qquad \swarrow \text{proj}_U$$
$$U$$

Q is called the *total space of the bundle*, M is the *base*, G is the *standard fiber* and π the *projection*.

The following definition will require some definitions related to group actions over manifolds, that we will cover next.

Definition 26. For a Lie group G and a manifold Q, a *group action* is a smooth application $\alpha : Q \times G \to Q$, denoted by $\alpha(q, g) = q \cdot g$, such that:

(1) For all $q \in Q$, $q \cdot e = q$, where e is the identity element of G.

(2) For all $g, h \in G$ and $q \in Q$, $(q \cdot g) \cdot h = q \cdot (gh)$.

If this action exists, we say that G acts on Q through α. As a consequence of that definition, we can see that the map $\alpha_g : Q \to Q : q \mapsto q \cdot g$ is a diffeomorphism for all $g \in G$, as it is smooth and its inverse is $\alpha_{g^{-1}}$.

We say that the action α *is free (or that G acts freely on Q)* if for all $q \in Q$, $q\dot{g} = q$ implies that $g = e$. The *orbit G_q of the action* for a point $q \in Q$ is defined as the image of the restriction $\alpha|_{\{q\} \times G}$, that is, $G_q := \{q \cdot g \mid g \in G\}$. We can construct an equivalence relation over Q in the following way: $q \sim p$ if and only if $G_p = G_q$, that is, if both points belong to the same orbit, or written explicitly, if there is an element $g \in G$ such that $q = p \cdot g$. The set of all equivalence classes is denoted as Q/G, and is known as the *orbit space of the action*.

Remark 5. The orbit space of an action is not, in general, a manifold, but because of the *quotient manifold theorem* (see [5], theorem 21.10), it is sufficient to have a free, proper and smooth action in order to guarantee that Q/G is a topological manifold.

Definition 27. The submersion $\pi : Q \to M$ is called a *principal G-bundle* if its is a fiber bundle whose fiber G is a Lie group, and this group acts on Q in such a way that the following properties hold:

(1) G acts freely on Q.

(2) The action orbits are the fibers of $\pi : Q \to M$, that is, for all $m \in M$, $\pi^{-1}(m) = G_q$, with $q \in Q_m$.

(3) The local trivializations of the fiber bundle are equivariant with respect to the action, that is, for every local trivialization $\phi : U \times G \to \pi^{-1}(U)$, the following property holds:
$$\phi(q, g \cdot h) = \phi(q, g) \cdot h,$$
for every $q \in U$, and $g, h \in G$.

From this we can conclude that if $\pi : Q \to M$ is a principal $G-$bundle, then Q/G is a topological manifold, and M is homeomorphic to Q/G. This homeomorphism $\phi : M \to Q/G$ is given by $\phi(m) = [q]$, with $\pi(q) = m$.

Definition 28 ([6], page 196). Given a principal G-bundle $\pi : Q \to M$, we say that a horizontal distribution $\mathcal{H}$ is a *connection for this principal G-bundle* if it is a connection for π and the G-action on Q preserves the horizontal distribution $\mathcal{H}$ in the sense that $\mathcal{H}_{q \cdot g} = d(\alpha_g)_q(\mathcal{H}_q)$, for all $q \in Q$, $g \in G$. In this case, we say that the pair $(\pi, \mathcal{H})$ is a *principal G-bundle with connection*.

Remark 6. An extremely useful property that any principal G-bundle with connection is that its connection is an Ehresmann connection, *i.e.*, it is complete, as it is stated in the following result.

Proposition 5 ([8], proposition 3.1). *Let $\pi : Q \to M$ be a principal G-bundle with connection and $c : I \to M$ a smooth curve. For an arbitrary $q \in Q_{c(0)} = G_q$, there is a unique horizontal lift $\gamma : I \to Q$ of c ($\pi \circ \gamma = c$, $\dot{\gamma}(t) \in \mathcal{H}_{\gamma(t)}$ for all $t \in I$) that starts from q.*

Definition 29 ([6], definition 11.2.1). A sub-Riemannian structure $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ on the total space of the principal G-bundle $\pi : Q \to M$ is called a *metric of bundle type* if it is an induced sub-Riemannian structure with respect to the submersion π and the metric on M (see definition 21).

3.3. Examples of principal G-bundles.

3.3.1. *Trivial bundle.* Let M be a manifold and G a Lie group. Define the projection onto the first factor $\pi : M \times G \to M$. It is clear that this map is surjective and a submersion. Moreover, it is a fiber bundle, whose local (global) trivialization is the identity map $\iota : M \times G \to M \times G$.

If we define the action $\alpha : (M \times G) \times G \to M \times G$ as $((p, g), q) \mapsto (p, gq)$, it is clear that G acts freely on $M \times G$, $\iota(p, gh) = \iota(p, g) \cdot h$ (the local trivialization is equivariant with respect to the action) and $\pi^{-1}(p) = \{p\} \times G = G_{(p,g)}$, so that π is a principal G-bundle. A connection for this bundle is the distribution $\mathcal{H}_{(p,g)} = \operatorname{Ker} d\tilde{\pi}(p, g) \subset T(M \times G)$, with $\tilde{\pi} : M \times G \to G$ the projection onto the second factor.

3.3.2. *Hopf Fibration.* All rotations in $\mathbb{R}^3$ can be expressed using quaternions. The set of quaternions is denoted as $\mathbb{H} := \{a + bi + cj + dj \mid a, b, c, d \in \mathbb{R}\}$, and in it we can define sum and product operations. The sum of two quaternions $a_1 + b_1 i + c_1 j + d_1 k$, $a_2 + b_2 i + c_2 j + d_2 k$ is $(a_1 + a_2) + (b_1 + b_2)i + (c_1 + c_2)j + (d_1 + d_2)k$. The product is defined in order to satisfy the distributive property with respect to the sum, and it also satisfies the following rules:

$$i^2 = j^2 = k^2 = -1, \quad ij = -ji = k, jk = -kj = i, ki = -ik = j.$$

In analogy with the complex plane and its isomorphism with $\mathbb{R}^4$, we define the norm of $z = a + bi + cj + dk$ as $|z| = \sqrt{a^2 + b^2 + c^2 + d^2}$, and the conjugate of z as $z^* = a - bi - cj - dk$. Moreover, every quaternion $z \neq 0$ has a multiplicative inverse, given by $z^{-1} = z^*/|z|^2$. One remarkable property of quaternions is the relation $|z_1 z_2| = |z_1||z_2|$. Every rotation around an axis given by a vector $\vec{v} = (b, c, d)$ and an angle θ can be identified with an unitary quaternion $r = a + bi + cj + dk$ ($|z| = 1$), where $\theta = 2\arccos(a)$ (see [9], page 6). This rotation $R_r : \mathbb{R}^3 \to \mathbb{R}^3$ is given by $x \mapsto rxr^{-1} = rxr^*$, and x identified as a pure quaternion, that is, $x = x_1 i + x_2 j + x_3 k$. To show explicitly that this map is the desired rotation, one can first show that this linear map preserves norm and the vector $bi + cj + dk$, that is, it is an eigenvector with eigenvalue 1, so it must be a rotation with respect the axis given by (b, c, d). The idea of the proof that the rotation angle is $2\arccos(a)$ can be found in [9], page 6, investigation step 3.

The set of all unitary quaternions is $\mathbb{S}^3$, so that $\mathbb{S}^3$ acts on $\mathbb{R}^3$ via rotations. The Hopf projection is defined as the map $\pi : \mathbb{S}^3 \to \mathbb{S}^2$ that sends the point $(1, 0, 0)$ (identified as the quaternion i) to its respective image after rotating $\mathbb{R}^3$ with the associated quaternion, as an element of $\mathbb{S}^3$. That is, $\pi(a, b, c, d) = R_r(1, 0, 0)$, where $r = a + bi + cj + dk$. The function π previously defined is surjective, as every point $x \in \mathbb{S}^2$ can be reached from $(1, 0, 0)$ through an adequate rotation. Explicitly, it suffices to take a normal vector to both $(1, 0, 0), x$ as the axis of rotation, and the angle of rotation is the angle between both vectors.

If we identify $\mathbb{R}^4 = \mathbb{C} \times \mathbb{C}$ and $\mathbb{R}^3 = \mathbb{R} \times \mathbb{C}$, Hopf projection from $\mathbb{S}^3 = \{(z_1, z_2) \mid z_1, z_2 \in \mathbb{C}, |z_1|^2 + |z_2|^2 = 1\}$ to $\mathbb{S}^2 = \{(x, z) \mid z \in \mathbb{C}, x \in \mathbb{R}, |z|^2 + x^2 = 1\}$ can be written as $\pi : \mathbb{S}^3 \to \mathbb{S}^2$, $\pi(z_1, z_2) = (|z_1|^2 - |z_2|^2, 2z_1 z_2^*)$, because writing explicitly rir^{-1} with $r = a + bi + cj + dk$ gives $(a^2 + b^2 - c^2 - d^2)i + 2(ad + bc)j + 2(bd - ac)k$. In this form one can show that this map is a submersion analyzing the minors of the jacobian matrix in cartesian coordinates (see [10], proof of theorem 3.11).

To recognize the fibers of this submersion, notice that $\pi^{-1}(1, 0) = \{(w_1, 0) \mid |w_1| = 1\} = \{(e^{i\theta}, 0) \mid \theta \in (0, 2\pi]\}$. Now, in quaternion notation, if $s, r \in \pi^{-1}(x)$, $rir^{-1} = sis^{-1} = x$, that is, $(s^{-1}r)i(r^{-1}s) = i$. Then, $s^{-1}r \in \pi^{-1}(1, 0)$, and $r = se^{i\theta}$, so $\pi^{-1}(x) = \{re^{i\theta} \mid \theta \in (0, 2\pi]\}$, and each fiber is diffeomorphic to $\mathbb{S}^1$.

With this, $\mathbb{S}^3$ can be seen as the disjoint union of these circles, and endowing this disjoint union with the standard manifold structure of $\mathbb{S}^3$, we identify $\mathbb{S}^2 = \mathbb{C}P^1$, and take the open sets $U_i = \{[z_0 : z_1] \mid z_i \neq 0\}$, with $i = 0, 1$. The local trivializations are (see [3], page

68) $\varphi_i : U_i \times \mathbb{S}^1 \to \pi^{-1}(U_i)$,

$$\varphi_i([z_0 : z_1], \lambda) = \left(\frac{\lambda}{\sqrt{1 + |\omega_i|^2}}, \frac{\lambda \omega_i}{\sqrt{1 + |\omega_i|^2}} \right),$$

with $\omega_0 = z_1/z_0$ and $\omega_1 = z_0/z_1$. Now, we define an action $\alpha : \mathbb{S}^3 \times \mathbb{S}^1 \to \mathbb{S}^3$, given by $\alpha(z_1, z_2, \lambda) = (z_1 \lambda, z_2 \lambda)$, where $(z_1, z_2) \in \mathbb{S}^3 \subset \mathbb{C}^2$ and $\lambda \in \mathbb{S}^1 \subset \mathbb{C}$. This action acts freely on $\mathbb{S}^3$, and its orbits coincide with the fibers of π, so that $\mathbb{S}^3/\mathbb{S}^1 \cong \mathbb{S}^2$. From this point it is clear that the Hopf bundle is a fiber bundle, and its local trivializations are equivariant with respect to the action, as $\varphi_i([z_0 : z_1], \lambda_1 \lambda_2) = \varphi_i([z_0 : z_1], \lambda_1)\lambda_2$. Therefore, $\pi : \mathbb{S}^3 \to \mathbb{S}^2$ is a principal $\mathbb{S}^1$-bundle.

To find a connection for this submersion, we recall that, by the preimage of regular value theorem, $\pi^{-1}(m) = \mathbb{S}_q^1$ is a submanifold of $\mathbb{S}^3$, of dimension $\dim(\mathbb{S}^3) - \dim(\mathbb{S}^2) = 1$, and its tangent space (vertical space of the submersion) $T_q \mathbb{S}_q^1$ is $\mathrm{Ker}(d\pi_q)$. Now, if $\mathbb{S}^3$ is endowed with the standard Riemannian structure, there is only one connection compatible with the metric, and is given by $\mathcal{H}_q = \mathrm{Ker}(d\pi_q)^\perp$. The vector that determines $\mathrm{Ker}(d\pi_q)$ can be found using the action if we consider the tangent vector of the curve $\gamma : (-\epsilon, \epsilon) \to \mathbb{S}^3$, $t \mapsto \gamma(z_1, z_2, e^{it}) = (z_1 e^{it}, z_2 e^{it})$ at $t = 0$, where $q = (z_1, z_2)$. With this, $\mathrm{Ker}(d\pi_q) = \langle (iz_1, iz_2) \rangle$, and the connection at $q \in \mathbb{S}^3$, $\mathcal{H}_q$, will be the orthogonal complement (with respect to $\mathbb{R}^4$ provided with the usual norm) of the vector space $\langle (iz_1, iz_2), (z_1, z_2) \rangle = \langle (x_1, y_1, x_2, y_2), (-y_1, x_1, -y_2, x_2) \rangle$, that is, $\mathcal{H}_q = \langle (x_2, -y_2, -x_1, y_1), (y_2, x_2, -y_1, -x_1) \rangle$, where $q = (x_1 + iy_1, x_2 + iy_2)$.

This principal $\mathbb{S}^1$-bundle has connection because the kernel of $d(\pi \circ \alpha_g)_q$ is the same of $d\pi_q$, as $d\alpha_g$ is an isomorphism. Therefore, $d(\alpha_g)_q(\mathcal{H}_q) = \mathcal{H}_{q \cdot g}$, and the $\mathbb{S}^1$-action preserves $\mathcal{H}$. By proposition 5, this connection $\mathcal{H}$ is complete.

3.4. Theorem on normal geodesics of bundle type sub-Riemannian metrics.
Before proceeding with one of the principal results on sub-Riemannian geometry and metrics on bundles, we need to recall some definitions and important results of symplectic geometry. Therefore, this subsection is dedicated to the study of these prerequisites and to the statement and proof of the theorem on normal geodesics of bundle type sub-Riemannian metrics. To start with, we define the concept of symplectic structure, and continue to explain the ideas of cotangent bundle and Hamilton's equations.

Definition 30 ([11], section 8A). Let M be an even-dimensional manifold. A *symplectic structure on M* is a closed non-degenerate differential 2-form ω on M, that is,

(1) $d\omega = 0$,

(2) $\forall \xi \neq 0 :\ \omega_m(\xi, \cdot) \neq 0\ (\xi \in T_m M),\ \forall m \in M$.

The non-degeneracy condition is equivalent to $\det \omega_m \neq 0, \forall m \in M$. The pair (M, ω) is called a *symplectic manifold*.

Example 16 (Canonical symplectic structure on $\mathbb{R}^{2n}$). Consider $\mathbb{R}^{2n}$ with canonical coordinates $(q_1, q_2, \ldots, q_n, p_1, p_2, \ldots, p_n)$. The 2-form $\omega = dp_i \wedge dq_i$ is closed and non-degenerate, so that the pair $(\mathbb{R}^{2n}, dp_i \wedge dq_i)$ is a symplectic manifold.

Definition 31 ([6], definition A.4.1). Given a symplectic manifold (M, ω), an *observable* is a smooth function $H : M \rightarrow \mathbb{R}$. Depending of the context, a given observable can induce some structure on the symplectic manifold, and in these cases the observable will be referred as *Hamiltonian*.

Historically, the definition of a function over a symplectic manifold receives the name of observable because the physical space has been modeled through history as a manifold, and the dual tangent space of this manifold represents the position and momentum space, which is a symplectic manifold (as will be discussed below). In principle, all observables in a physical phenomena can be calculated with the variables of the dual tangent space of the manifold that represents physical space, and they are functions of the position and momentum. From this fact, a real function over the symplectic space receives the name of observable. Over all the physical observables, one of the most important is the Hamiltonian, as it describes the evolution of the system through Hamilton's equations as a function of time. Thus, a real function over the symplectic manifold that is important in the sense of giving additional structure to the manifold is called Hamiltonian.

Definition 32 ([11], page 203). For each vector $\xi \in T_m M$ tangent to a symplectic manifold (M, ω), we associate a 1-form $\omega_\xi \in T_m^* M$ by the formula

$$\omega_\xi(\eta) = \omega(\xi, \eta).$$

This last definition induces an isomorphism between $T_m M$ and $T_m^* M$, given by the map $\xi \mapsto \omega_\xi$. If $H : M \rightarrow \mathbb{R}$ is a Hamiltonian, then $dH = \partial_i H dq^i$ is a differential 1-form on M, and by the previous isomorphism, at every point $m \in M$ there is a tangent vector $(X_H)_m \in T_m M$ such that $(dH)_m = \omega_{(X_H)_m}$.

Definition 33. For a given Hamiltonian $H : M \rightarrow \mathbb{R}$, the vector field $X_H \in \mathfrak{X}(M)$ that satisfies

$$dH = \omega(X_H, \cdot)$$

is called *Hamiltonian vector field* for H, or *symplectic gradient* of H.

From any manifold M we can construct a symplectic structure: the cotangent bundle. Suppose that M is an n-dimensional manifold. $T^* M = \bigsqcup_{p \in M} T_p^* M$ is then a $2n$-dimensional manifold. There is a canonical choice of coordinates for this space, namely, given $m \in M$, there is a coordinate chart (U_m, ϕ) on M such that $q = \phi(m) = (q_1, \ldots, q_n) \in \mathbb{R}^n$, and for a given 1-form $\alpha \in T_m^* M$, in these coordinates we have that $\alpha = p_i dq^i$; the cross product of these two coordinates in a suitable neighborhood of an element $\alpha_m \in T^* M$ gives the *canonical coordinates of $T^* M$*, where the element $\alpha_m \in T^* M$ can be represented locally as the pair of coordinates (q, p).

Theorem 5 ([11], page 202). *The cotangent bundle T^*M has a natural symplectic structure given in coordinates by $\omega = dq^i \wedge dp_i$.*

Proof. Let $v \in T_\alpha(T^*M)$ be a vector tangent to the cotangent bundle at $\alpha \in T_m^*M$. If $\pi : T^*M \to M$ is the natural projection, the differential $d\pi : T(T^*M) \to TM$ takes v to $d\pi(v) = d\pi_\alpha(v)$, which is a vector tangent to M at m. We define a 1-form θ on T^*M by the relation $\theta_\alpha(v) = \alpha(d\pi_\alpha(v))$. In local coordinates this form is $\theta = p_i dq^i$, and the form $\omega = d\theta$ is non-degenerate. $\qquad\square$

Given the symplectic structure $(T^*M, \omega = dq^i \wedge dp_i)$, for a Hamiltonian $H : T^*M \to \mathbb{R}$, there is a vector field $X_H \in \mathfrak{X}(T^*M)$ such that $dH = \omega(X_H, \cdot)$. Writing X_H in coordinates, $X_H = X_i \partial/\partial q^i + X^i \partial/\partial p_i$, and evaluating $\partial/\partial q^j$ and $\partial/\partial p_j$ in dH gives

$$dH(\partial/\partial q^j) = \frac{\partial H}{\partial q^j} = \omega(X_H, \partial/\partial q^j) = -\omega(\partial/\partial q^j, X_H) = -X^j,$$

$$\text{and } dH(\partial/\partial p_j) = \frac{\partial H}{\partial p_j} = \omega(X_H, \partial/\partial p_j) = X_j,$$

so that $X_H = (\partial H/\partial p_j)\partial/\partial q^j - (\partial H/\partial q^j)\partial/\partial p_j$, and its integral curves are defined by the following system of differential equations:

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q^i}. \tag{1}$$

This system of differential equations is called the *Hamilton equations for H*. Despite this formulation on the cotangent bundle came after Hamilton, this method turns out to be a generalization of Hamiltonian mechanics to symplectic manifolds, as one obtains the classical Hamiltonian equations of the classical mechanics using $M = \mathbb{R}^n$, and taking the Hamiltonian as the energy of the system. That is the reason why they are referred as Hamilton equations.

Theorem 6. *The Hamiltonian $H : T^*M \to \mathbb{R}$ is constant along the integral curves of the vector field X_H.*

Proof. Suppose that $c : I \subseteq \mathbb{R} \to T^*M$ is an integral curve of X_H. Then, the composition $(H \circ c)(t)$ can be differentiated: let $t \in I$, and $(U_{c(t)}, \varphi)$ be a coordinate chart, then,

$$\frac{d}{dt}(H \circ c)(t) = \frac{d}{dt}\left((H \circ \varphi^{-1}) \circ (\varphi \circ c)\right)(t) = \frac{d(\varphi \circ c)^i}{dt} \cdot \frac{\partial(H \circ \varphi^{-1})}{\partial q^i} = X_H^i \partial_i H$$

$$= dH(X_H) = \omega(X_H, X_H) = 0,$$

and H is constant throughout the curve c, given the skew-symmetry of ω. $\qquad\square$

Definition 34 ([6], example 2, section A.4). For any vector field $\xi \in \mathfrak{X}(M)$, we define its *momentum function* $P_\xi : T^*M \to \mathbb{R}$ by

$$P_\xi(p_q) = p(\xi_q).$$

Example 17 ([6], example 1, section A.4). If $\xi = \partial/\partial q^i \in \mathfrak{X}(\mathbb{R}^n)$, its momentum function is

$$P_{\partial/\partial q^i}(p_q) = p(\partial/\partial q^i) = p_i.$$

On the other hand, if we take the Hamiltonian $H(p,q) = p_i$, we get that $X_H = \partial/\partial q^i$.

In local coordinates, $P_\xi = p_i \xi_q^i$, hence, computing its Hamilton equations gives

$$\dot{q}^i = \xi_q^i, \quad \dot{p}_i = -p_j \frac{\partial \xi_q^j}{\partial q^i}.$$

The first Hamilton equation says that the flow of the Hamiltonian vector field X_{P_ξ} on T^*M projects onto the flow of the original vector field ξ on M ([6], page 278). To understand the second Hamilton equation, recall that the *cotangent lift* $\Phi_{T^*} : T^*M \to T^*M$ of any diffeomorphism $\Phi : M \to M$ is given by

$$\Phi_{T^*}(p) = (\Phi(q), (\Phi^{-1})_{\Phi(q)}^*(p)), \ p \in T_q^*M,$$

where $(\Phi^{-1})^*$ is

$$(\Phi^{-1})_{\Phi(q)}^*(p)(\eta) = p\left(d\Phi_{\Phi(q)}^{-1}(\eta)\right)$$

for every $\eta \in T_{\Phi(q)}M$. If $\Phi_t : M \to M$ is the flow of the vector field ξ, then the flow of X_{P_ξ} is $(\Phi_t)_{T^*}$, as

$$(\Phi_t)_{T^*}(p) = \left(\Phi_t(q), p\left(d(\Phi_{-t})_{\Phi(q)}(\cdot)\right)\right),$$

and the first component represents the flow of ξ, while the second component is in coordinates $p_i(\partial(\Phi_{-t})^i/\partial q^j)dq^j$, so differentiating this expression with respect to t at $t = 0$ gives $-p_i(\partial \xi_q^i/\partial q^j)dq^j$, which is the 1-form whose coordinates are equal to the coordinates on p of X_{P_ξ}. Therefore, the second Hamilton equation for P_ξ asserts that the Hamiltonian flow is the cotangent lift of this flow on the configuration space M ([6], page 278). Summarizing, we have the formula

$$X_{P_\xi} = \left.\frac{d(\Phi_t)_{T^*}}{dt}\right|_{t=0}.$$

Definition 35. Let $f, g : T^*M \to \mathbb{R}$ be two observables. We define the *Poisson bracket of f and g* as

$$\{f, g\} = X_g(f),$$

where X_g is the Hamiltonian vector field associated with g.

In coordinates, the Poisson bracket is

$$\{f, g\} = \frac{\partial g}{\partial p_i}\frac{\partial f}{\partial q^i} - \frac{\partial g}{\partial q^i}\frac{\partial f}{\partial p_i}.$$

Another way to write the Poisson bracket of f and g using the symplectic form ω of T^*M is the following:

$$\{f, g\} = df(X_g) = \omega(X_f, X_g).$$

Proposition 6. *The Poisson bracket satisfies the following properties:*

    (1) $\{f, g\} = -\{g, f\}$,

(2) $\{af + bg, h\} = a\{f, g\} + b\{g, h\}$,

(3) $\{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} = 0$,

(4) $\{f, gh\} = g\{f, h\} + \{f, g\}h$,

where $f, g, h : T^*M \to \mathbb{R}$ are observables, and $a, b \in \mathbb{R}$.

Proof. ([12], proof of proposition 1.3) Properties 1,2 and 4 are trivial. The third property, known as the *Jacobi identity*, is equivalent to the fact that ω is a closed form, as will be shown next. Recall that

$$d\chi(X, Y, Z) = X\chi(Y, Z) - Y\chi(X, Z) + Z\chi(X, Y) - \chi([X, Y], Z) + \chi([X, Z], Y) - \chi([Y, Z], X)$$

for any 2-form χ and $X, Y, Z \in \mathfrak{X}(T^*M)$. Therefore,

$$\begin{aligned}
d\omega(X_f, X_g, X_h) &= X_f\{g, h\} - X_g\{f, h\} + X_h\{f, g\} + [X_f, X_g]h - [X_f, X_h]g + [X_g, X_h]f \\
&= \{f, \{g, h\}\} - \{g, \{f, h\}\} + \{h, \{f, g\}\} + \{f, \{g, h\}\} - \{g, \{f, h\}\} \\
&\quad - \{f, \{h, g\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} - \{h, \{g, f\}\} \\
&= 3\left(\{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\}\right).
\end{aligned}$$

From the last expression it is clear that $d\omega = 0$ is equivalent to the Jacobi identity. $\square$

The first three properties of last proposition guarantee that $(C^\infty(T^*M), \{\cdot, \cdot\})$ is a Lie algebra. The last property is known as Leibniz identity, because due to it the operator $f \mapsto \{f, H\}$, with H a Hamiltonian, is a derivation of $C^\infty(T^*M)$:

$$\{fg, H\} = f\{g, H\} + g\{f, H\}.$$

Therefore, $f \mapsto \{f, H\}$ determines a vector field over T^*M: the Hamiltonian vector field of H, X_H ([6], page 279).

Proposition 7 (Hamilton equations in bracket form)**.** *If $c(t)$ is an integral curve for X_H and $f \in C^\infty(T^*M)$ is an observable, then*

$$\frac{d}{dt}(f \circ c)(t) = \{f, H\}(c(t)). \tag{2}$$

Proof. Following the steps of the proof of theorem 6, we obtain that

$$\frac{d}{dt}(f \circ c)(t) = \omega(X_f, X_H) = \{f, H\}(c(t)).$$

$\square$

We write $\dot{f} = \{f, H\}$ to refer to the expression (2), valid for any observable f. Equation (2) also receives the name of Hamilton equations, and when f runs over the canonical coordinates q^i, p_i, we get the Hamilton equations previously stated in (1).

Proposition 8 (Poisson bracket of momentum functions)**.** *For any two vector fields $\xi, \eta \in \mathfrak{X}(M)$ with momentum functions $P_\xi, P_\eta : T^*M \to \mathbb{R}$, we have that*

$$\{P_\xi, P_\eta\} = -P_{[\xi,\eta]},$$

where $[\xi, \eta]$ is the Lie bracket between ξ and η, and $P_{[\xi,\eta]}$ is the momentum function of $[\xi, \eta] \in \mathfrak{X}(M)$.

Proof. Follows directly from calculus in coordinates. $\qquad\qquad\qquad\qquad\square$

If (M, g) is a Riemannian manifold, we define the *Riemannian Hamiltonian* $H_R : T^*M \to \mathbb{R}$ as

$$H_R(q, p) = \frac{1}{2} g_q(p, p).$$

In local coordinates, $H_R(q, p) = g^{ij}(q) p_i p_j / 2$. This Hamiltonian has the following Hamilton equations:

$$\dot{q}^i = \frac{\partial H_R}{\partial p_i} = g^{ij} p_j, \quad \dot{p}_i = -\frac{\partial H}{\partial q^i} = -\frac{1}{2} \frac{\partial g^{jk}(q)}{\partial q^i} p_j p_k.$$

From the first Hamilton equation we deduce that $p_i = g_{ij} \dot{q}^j$, and substituting this expression in the second Hamilton equation gives

$$\dot{p}_i = \frac{1}{2} \frac{\partial g_{ml}}{\partial q^i} \dot{q}^m \dot{q}^l,$$

so that

$$\ddot{q}^i = \frac{\partial g^{ij}(q)}{\partial q^k} \dot{q}^k p_j + g^{ij} \dot{p}_j = \frac{1}{2} g^{ij} \left(\frac{\partial g_{kl}}{\partial q^j} - 2 \frac{\partial g_{jl}}{\partial q^k} \right) \dot{q}^k \dot{q}^l = -\Gamma^i_{kl} \dot{q}^k \dot{q}^l,$$

and q satisfies the geodesic equation, from where we deduce that the integral curves of X_{H_R} are geodesics. Moreover, any geodesic $q(t)$ satisfies Hamilton equations, given that $p_i = g_{ij} \dot{q}^j$. From this discussion we derive the following proposition.

Proposition 9. *If (M, g) is a Riemannian manifold, and $H_R : T^*M \to \mathbb{R}$ is its Riemannian Hamiltonian, $q(t)$ is a geodesic if and only if $(q^i(t), g_{ij}(q(t)) \dot{q}^j(t))$ is an integral curve of X_{H_R}.*

Recall that given a Riemannian manifold (M, g), we can construct a metric g_q^* over every cotangent fiber T_q^*M, using the identification

$$g_q^*(v^*, w^*) := g_q(v, w),$$

where $v^*, w^* \in T_q^*M$, and $v, w \in T_qM$ are the preimages of v^*, w^* via the isomorphism $T_q^*M \cong T_qM$ induced by the metric g, that is, $v^* = g_q(v, \cdot)$ and $w^* = g_q(w, \cdot)$. In coordinates, if $g_q = g_{ij}(q)$, then $g_q^* = g^{ij}(q)$, that is, the representation in coordinates of the metric in the cotangent fiber is equivalent to the inverse matrix of the representation in coordinates of the metric in the tangent fiber.

Now, if M is a manifold with a sub-Riemannian structure $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ over it, we can define a *cometric*, that is just a metric over the fibers of the cotangent bundle as follows. Given that $\langle \cdot, \cdot \rangle_q$ is a metric for $\mathcal{H}_q$, for any $v^* \in T_q^*M$ there exists a single $V \in \mathcal{H}_q$ such that

$v^*|_{\mathcal{H}_q} = \langle V, \cdot \rangle_q$. With this property, we can define the cometric $(\cdot, \cdot) \in \mathfrak{X}_2^0(M)$ as a bilinear form (in general it is not a metric) over the fibers of the cotangent space as follows:

$$(v^*, w^*)_q = \langle V, W \rangle_q,$$

where $v^*, w^* \in T_q^* M$, and $V, W \in \mathcal{H}_q$ are such that $v^*|_{\mathcal{H}_q} = \langle V, \cdot \rangle_q$, $w^*|_{\mathcal{H}_q} = \langle W, \cdot \rangle_q$.

Definition 36 ([6], definition 1.5.2). If M is a manifold, and $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ is a sub-Riemannian structure over it, we define the sub-Riemannian Hamiltonian $H_{sR} : T^* M \to \mathbb{R}$ as

$$H_{sR}(p, q) = \frac{1}{2}(p, p)_q,$$

where $(\cdot, \cdot)$ is the cometric induced by $\langle \cdot, \cdot \rangle$.

If $\pi : Q \to M$ is a submersion endowed with a connection, and (M, g) is a Riemannian manifold, we know that the submersion induces a sub-Riemannian structure on Q, $(\mathcal{H}, \langle \cdot, \cdot \rangle)$, where $\mathcal{H}$ is the Ehresmann connection associated with π, and $\langle \cdot, \cdot \rangle$ is defined by the isomorphism $d\pi_q|_{\mathcal{H}_q} : \mathcal{H}_q \to T_{\pi(q)}M$ and the metric g on M.

The sub-Riemannian Hamiltonian $H_{sR} : T^* Q \to \mathbb{R}$ can be written in terms of the Riemannian Hamiltonian $H_R : T^* M \to \mathbb{R}$ in the following way: given the splitting $TQ = \mathcal{H} \oplus V$, we have that $T^* Q = \mathcal{H}^* \oplus V^*$, and so we have the canonical projection $\mathrm{pr}_1 : T^* Q \to \mathcal{H}^*$. Now, we define a second projection $\mathrm{pr}_2 : \mathcal{H}^* \to T^* M$ by the formula $\mathrm{pr}_2(\alpha)(v) = \alpha(h_q v)$, where $\alpha \in \mathcal{H}_q^*$, $v \in T_{\pi(q)}M$, and $h_q v$ is the horizontal lift of v into $\mathcal{H}_q$ (well-defined and unique, because $\mathcal{H}$ is an Ehresmann connection).

Proposition 10 ([6], proposition 11.1.5). *The induced sub-Riemannian Hamiltonian is given by* $H_{sR} = H_R \circ \mathrm{pr}_2 \circ \mathrm{pr}_1$.

Proof. We can horizontally lift any $Y \in \mathfrak{X}(M)$ in a unique way to a vector field $hY \in \mathfrak{X}(Q)$ that satisfies the following properties: $(hY)_q \in \mathcal{H}_q$ and $d\pi_q((hY)_q) = Y_{\pi(q)}$ for all $q \in Q$. If Y_a is an orthonormal frame for the Riemannian structure on M, then hY_a is a horizontal orthonormal frame for the induced sub-Riemannian structure on Q, and therefore

$$H_R = \frac{1}{2} \sum_a P_{Y_a}^2 \text{ and } H_{sR} = \frac{1}{2} \sum_a P_{(hY)_a}^2,$$

where P_{Y_a} and $P_{(hY)_a}$ are the momentum functions associated to Y_a and $(hY)_a$, respectively. Now, for any $Y \in \mathfrak{X}(M)$, $\mathrm{pr}^* P_Y = P_{hY}$ because if $\alpha : T^* Q \to \mathbb{R}$, then

$$(\mathrm{pr}^* P_Y)(\alpha) = P_Y(\mathrm{pr}(\alpha)) = \mathrm{pr}(\alpha)(Y_{\pi(q)}) = \mathrm{pr}(\alpha)(d\pi_q((hY)_q))$$
$$= \mathrm{pr}_1(\alpha)((hY)_q) = \alpha((hY)_q) = P_{hY}(\alpha).$$

With this,

$$H_R \circ \mathrm{pr} = \frac{1}{2} \sum_a (P_{Y_a} \circ \mathrm{pr})^2 = \frac{1}{2} \sum_a (\mathrm{pr}^* P_{Y_a})^2 = \frac{1}{2} \sum_a P_{(hY)_a}^2 = H_{sR}.$$

$\square$

Definition 37. Let (M, g) be a Riemannian manifold, and G a Lie group that acts on M via the application $\alpha : M \times G \to M$, $\alpha(m, h) = m \cdot h$. We say that *the metric g is G-invariant* if
$$g_{m \cdot h}((d\alpha_h)_m(v), (d\alpha_h)_m(w)) = g_m(v, w),$$
for all $m \in M$, $v, w \in T_m M$ and $h \in G$.

Let $\pi : Q \to M$ be a principal G-bundle with connection $\mathcal{H}$. If $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ is a metric of bundle type, we say that a metric h on Q is a *compatible Riemannian metric* if h is G-invariant and compatible (see definition 23). From this definition, it is clear that $h|_{G_q} : TG_q \oplus TG_q \to \mathbb{R}$ is a right-invariant metric, and we can use it to define a metric on G as follows:

Let $\mathfrak{g} = \mathrm{Lie}(G) = T_e G$, and $\exp : \mathfrak{g} \to G$ be the exponential application. *The infinitesimal generator map for the group action $\sigma_v : Q \to T_q Q$, for $v \in \mathfrak{g}$, is defined as*
$$\sigma_v(q) = \frac{d}{d\epsilon}\Big|_{\epsilon=0} (q \cdot \exp(\epsilon v)).$$

Definition 38 ([6], definition 11.2.3)**.** For a given $q \in Q$, the bilinear form on $\mathfrak{g}$, $\mathbb{I}_q : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$, given by
$$\mathbb{I}_q(v, w) = h_q(\sigma_v q, \sigma_w q)$$
is called the *moment of inertia tensor at q*.

Recall that for a Lie Group G, the map $\mathrm{ad}_g : G \to G$ given by $h \mapsto ghg^{-1}$ is a diffeomorphism, and its differential at identity, $d(\mathrm{ad}_g)_e$, is denoted as $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$.

Proposition 11. *The family of vector fields $\{\sigma_v\}_{v \in \mathfrak{g}}$ satisfies the following property:*
$$\sigma_v(q \cdot g) = d(\alpha_g)_q \left(\sigma_{\mathrm{Ad}_g(v)}(q) \right),$$
where $\alpha : Q \times G \to Q$ is the action of G on Q.

Proof. Let $\gamma : (-\epsilon, \epsilon) \to G$ be the one parameter subgroup such that $\gamma(0) = e$, $\dot{\gamma}(0) = \xi \in \mathfrak{g}$. Then, $\mathrm{Ad}_g(\xi) = d/dt|_{t=0}(g\gamma(t)g^{-1})$, and evaluating this on σ_q gives

$$\sigma_\xi(q) = \frac{d}{d\epsilon}\Big|_{\epsilon=0} q \cdot \exp\left(\epsilon \frac{d}{dt}\Big|_{t=0} g\gamma(t)g^{-1} \right) = \frac{d}{d\epsilon}\Big|_{\epsilon=0} q \cdot g\gamma(\epsilon)g^{-1}.$$

Finally, if we evaluate this on $d(\alpha_g)_q$, we get that

$$\begin{aligned}
d(\alpha_g)_q \left(\frac{d}{d\epsilon}\Big|_{\epsilon=0} q \cdot g\gamma(\epsilon)g^{-1} \right) &= \frac{d}{d\epsilon}\Big|_{\epsilon=0} q \cdot g\gamma(\epsilon) \\
&= \frac{d}{d\epsilon}\Big|_{\epsilon=0} q \cdot g \exp(\epsilon\xi) \\
&= \frac{d}{d\epsilon}\Big|_{\epsilon=0} (q \cdot g) \cdot \exp(\epsilon\xi) \\
&= \sigma_\xi(q \cdot g).
\end{aligned}$$

□

Thanks to the last proposition and the fact that h is G-invariant, we have that

$$
\begin{aligned}
\mathbb{I}_{q \cdot g}(v, w) &= h_{q \cdot g}\left(d(\alpha_g)_q\left(\sigma_{\mathrm{Ad}_g(v)}(q)\right), d(\alpha_g)_q\left(\sigma_{\mathrm{Ad}_g(w)}(q)\right)\right) \\
&= h_q\left(\sigma_{\mathrm{Ad}_g(v)}(q), \sigma_{\mathrm{Ad}_g(w)}(q)\right) \\
&= \mathbb{I}_q\left(\mathrm{Ad}_g(v), \mathrm{Ad}_g(w)\right).
\end{aligned}
$$

Definition 39. Let G be a Lie group. A bilinear form $\beta \in \Lambda^2(\mathfrak{g})$ is called *bi-invariant* or *adjoint invariant* if

$$
\beta\left(\mathrm{Ad}_g v, \mathrm{Ad}_g w\right) = \beta(v, w) \text{ for all } g \in G, \ v, w \in \mathfrak{g}.
$$

Any bilinear form $\beta \in \Lambda^2(\mathfrak{g})$ can be extended to all TG using left or right translations $L_g : G \to G, h \mapsto gh$, $R_g : G \to G, h \mapsto hg$ in the following way:

$$
\beta_g(v, w) = \beta\left(d(L_{g^{-1}})_g(v), d(L_{g^{-1}})_g(w)\right), \text{ or}
$$

$$
\beta_g(v, w) = \beta\left(d(R_{g^{-1}})_g(v), d(R_{g^{-1}})_g(w)\right),
$$

for all $g \in G$, $v, w \in T_g G$. If β is bi-invariant, both extensions coincide, because

$$
\begin{aligned}
\beta\left(d(R_{g^{-1}})_g(v), d(R_{g^{-1}})_g(w)\right) &= \beta\left(d(R_{g^{-1}})_g \circ d(L_g)_e \circ d(L_{g^{-1}})_g(v), d(R_{g^{-1}})_g \circ d(L_g)_e\right. \\
&\qquad\qquad\qquad \left. \circ d(L_{g^{-1}})_g(w)\right) \\
&= \beta\left(\mathrm{Ad}_g \circ d(L_g^{-1})_g(v), \mathrm{Ad}_g \circ d(L_g^{-1})_g(w)\right) \\
&= \beta\left(d(L_{g^{-1}})_g(v), d(L_{g^{-1}})_g(w)\right).
\end{aligned}
$$

Therefore, an adjoint invariant positive definite bilinear form on the Lie algebra defines a *bi-invariant Riemannian metric on G*.

Definition 40 ([6], definition 11.2.4)**.** The G-invariant Riemannian metric (or pseudo-metric) h on Q is said to be of *constant bi-invariant type* if its inertia tensor $\mathbb{I}_q$ equals the same bi-invariant bilinear form on $\mathfrak{g}$, for all $q \in Q$.

Note that if G is a Lie group acting on itself via the multiplication map $\alpha : G \times G \to G$, $(k, h) \mapsto kh$, $\alpha_h = R_h$, and if g is a G-invariant metric on G we have that

$$
g_{kh}\left(d(R_h)_k(v), d(R_h)_k(w)\right) = g_k(v, w)
$$

for all $k, h \in G$, $v, w \in T_k G$. Therefore, a metric on G is G-invariant under the canonical action of G on itself if and only if this metric is right invariant. Moreover, g is of constant bi-invariant type if and only if it is bi-invariant, because for any $v, w \in T_k G$, there exist $\xi, \eta \in \mathfrak{g}$ such that $d(L_k)_e \xi = v, d(L_k)_e \eta = w$, and

$$
\mathbb{I}_k(\xi, \eta) = g_k(v, w) = \mathbb{I}_{hk}(\xi, \eta) = g_{hk}\left(d(L_h)_k v, d(L_h)_k w\right)
$$

as $\sigma_\chi(k) = d(L_k)_e \chi$ for all $k \in G$, $\chi \in \mathfrak{g}$.

Proposition 12. *If g is a bi-invariant metric on a Lie group G, its geodesics through the identity coincide with its one-parameter subgroups.*

Proof. If $m : G \times G \to G$ is the multiplication map, let $h, k \in G$, and $\gamma, \eta : (-\epsilon, \epsilon) \to G$ be two curves such that $\gamma(0) = h, \dot{\gamma}(0) = v \in T_h G, \eta(0) = k, \dot{\eta}(0) = w \in T_k G$. The differential $dm_{(h,k)} : T_{(h,k)}G \times G \cong T_h G \oplus T_k G \to T_{hk}G$ is bilinear, so that if $l, s : (-\epsilon, \epsilon) \to H$ are constant maps with $l(t) = h, s(t) = k$, then

$$dm_{(h,k)}(v, 0) = \frac{d}{dt}\Big|_{t=0} (\gamma(t)k) = d(R_k)_h(v),$$

and

$$dm_{(h,k)}(0, w) = \frac{d}{dt}\Big|_{t=0} (h\eta(t)) = d(L_h)_k(w).$$

Therefore, $dm_{(h,k)}(v, w) = d(R_k)_h(v) + d(L_h)_k(w)$ by linearity. Using this result, if $\gamma : (-\epsilon, \epsilon)$ is a curve such that $\gamma(0) = 0, \dot{\gamma}(0) = v \in T_h G$, differentiating the expression $\gamma(t) [\gamma(t)]^{-1} = e$ gives

$$0 = d(R_{h^{-1}})_h(v) + d(L_h)_{h^{-1}} (d\iota_h(v)),$$

where $\iota : G \to G$ is the inversion map $h \mapsto h^{-1}$. From last expression we get that $d\iota_h(v) = -d(L_{h^{-1}})_e \circ d(R_{h^{-1}})_h(v)$. As g is a bi-invariant metric, $g_{h^{-1}}(d\iota_h(v), d\iota_h(w)) = g_h(v, w)$, and as $\iota : G \to G$ is a diffeomorphism, it is an isometry.

Now, if $\gamma : (-\epsilon, \epsilon) \to G$ is a geodesic with $\gamma(0) = e, \dot{\gamma}(0) = v \in T_e G$, as ι is an isometry, $\iota \circ \gamma(t) = [\gamma(t)]^{-1}$ is also a geodesic. Moreover, $d\iota_e(v) = -v$, so that $\gamma(t)$ and $[\gamma(t)]^{-1}$ are the same geodesic, with $[\gamma(t)]^{-1} = \gamma(-t)$. For small $t_0 \in \mathbb{R}$ (sufficiently small in order to have that $\operatorname{Im} \gamma|_{[0,t_0]} \subset U$, where U is a totally normal neighborhood of e), consider the curve $\eta : (-t_0, 0) \to G, t \mapsto \gamma(t_0)\gamma(t) = L_{\gamma(t_0)}(\gamma(t))$. This curve is a geodesic because g is bi-invariant, with $\gamma(0) = \gamma(t_0), \eta(-t_0) = e$. Due to the uniqueness of minimizing geodesics in a totally normal neighborhood, $\eta(t) = \gamma(t_0 + t)$, that is, $\gamma(t_0)\gamma(t) = \gamma(t_0 + t)$, for $t \in (-t_0, 0)$. Using the relation $\gamma(s+t) = \gamma(s)\gamma(t)$ and the local uniqueness of minimizing geodesics we can extend this property over all the domain of the geodesic γ. With this, γ is a one-parameter subgroup, and

$$\exp(v) = \gamma(1) = (\exp_R)_e(v),$$

where $(\exp_R)_e : T_e G \to G$ is the Riemannian exponential. $\qquad \square$

Proposition 13. *If G is a Lie group endowed with a bi-invariant metric g, the Riemannian Hamiltonian $H_R : T^*G \to \mathbb{R}, (q, p) \mapsto g_q(p, p)/2$ commutes with every right or left invariant observable $f \in C^\infty(T^*G)$.*

Proof. From proposition 7, if $f : T^*G \to G$ is an observable, then $d/dt\,(f \circ c)\,(t) = \{f, H_R\}\,(c(t))$, where $c(t)$ is an integral curve of the Hamiltonian vector field associated to H_R. By proposition 9, $c(t) = (q(t), g_{q(t)}(\cdot, \dot{q}(t)))$, where $q(t)$ is a geodesic. As G is complete (because g is bi-invariant), there is a geodesic that connects e with every $h \in G$, and by last proposition, this geodesic is a one-parameter subgroup, which implies

that $d/dt(f \circ c)(t) = 0$, because the quantity $f \circ c(t)$ is constant since f is left or right invariant. $\qquad\square$

Theorem 7 (Theorem on normal geodesics of bundle type sub-Riemannian metrics, [6], theorem 11.2.5)**.** *Let $\pi : Q \to M$ be a principal G-bundle with connection $\mathcal{H}$. Suppose that Q is a Riemannian manifold such that its metric is a Riemannian submersion (see definition 24), compatible with $\mathcal{H}$, G-invariant, and its associated vertical metric has constant bi-invariant type. Therefore, given any geodesic $\gamma : I \to Q$, the horizontal lift of $\pi \circ \gamma$ is a sub-Riemannian geodesic. All sub-Riemannian geodesics can be obtained in this manner. Moreover, the projected curve $\pi \circ \gamma$ is a geodesic in the base space M if and only if γ is a horizontal geodesic in $(Q, \mathcal{H})$.*

Remark 7. The assumptions of this theorem are equivalent to suppose that $\pi : Q \to M$ is a principal G-bundle endowed with a sub-Riemannian metric of bundle type, and that this metric can be extended to a compatible Riemannian metric whose associated vertical metric has constant bi-invariant type.

Proof. Let H_R be the Riemannian Hamiltonian of Q, H_{sR} its sub-Riemannian Hamiltonian, and let H_G be the vertical part of the Riemannian Hamiltonian, that is, $H_R = H_{sR} + H_G$. We will need the following two claims for this proof:

(1) H_{sR} and H_G Poisson-commute, that is, $\{H_{sR}, H_G\} = 0$.

(2) The projection to Q of any integral curve in T^*Q for the Hamiltonian vector field defined by H_G is of the form $q \exp(tv)$, for some fixed $v \in \mathfrak{g}$, which depends of the initial conditions of the integral curve.

Proof of the claims. It suffices to show that the claims hold over any locally trivialized neighborhood $\pi^{-1}(U)$ in Q. Let $\phi_U : \pi^{-1}(U) \to U \times G$ be the corresponding local trivialization, which induces a diffeomorphism $T^*Q \cong T^*U \times T^*G$.

Let $(x, p) \in T^*U$ and $(g, \mu) \in T^*G$, being $\mu \in \mathfrak{g}^*$. Then, we have that

$$H_G(x, p, g, \mu) = \beta^*(\mu) \tag{3}$$

and

$$H_{sR}(x, p, g, \mu) = f(x, p, \mu). \tag{4}$$

Equations (3) and (4) come from the fact that the metric is of constant bi-invariant type and G-invariant. Using these equations, the Poisson bracket of H_{sR} and H_G is

$$\{f, \beta^*\} = \underbrace{\{f, \beta^*\}_U}_{0} + \{f, \beta^*\}_G = \sum_a \frac{\partial f}{\partial g^a} \frac{\partial \beta^*}{\partial \mu_a} = \sum_a \frac{\partial f}{\partial \mu_a} \frac{\partial \mu_a}{\partial g^a} \frac{\partial \beta^*}{\partial \mu_a} = \sum_a \frac{\partial f}{\partial \mu_a} \{\mu_a, \beta^*\}_G,$$

using both the Leibniz rule and the product Poisson structure, μ_a are coordinates on $\mathfrak{g}^*$, and g^a are coordinates on G. As the μ_a are right-invariant functions on T^*G, by proposition 13, $\{\mu_a, \beta^*\}_G = 0$, and this proves the first claim.

To prove the second claim, note that the flow of H_G is of the form $\Phi_G(t)(x, p, g, \mu) = (x, p; \Phi_\beta(t)(g, \mu))$, where $\Phi_\beta(t) : T^*G \to T^*G$ is the flow associated with β^*, that is, $\Phi_\beta(t)(g, \mu) = (\exp_g(tv), \mathfrak{v}_{\exp_g(tv)}(\cdot, d/dt \exp_g(tv)))$, where $\mathfrak{v}$ is the vertical metric, and $v \in \mathfrak{g}$ satisfies $\mu = \mathfrak{v}_g(\cdot, v)$. With this, the projection to Q of $\Phi_G(t)(x, p, g, \mu)$ is $x \exp_g(tv)$. $\quad\square$

With the claims proved, we proceed with the proof of the theorem. The Riemannian exponential map is the composition of four maps:

$$T_qQ \xrightarrow{\iota} TQ \longrightarrow T^*Q \xrightarrow{\Phi_R(1)} T^*Q \xrightarrow{\text{proj}} Q,$$

where the first map ι is the inclusion, the second map is the metric induced identification, $\Phi_R(1)$ is the time 1-flow of H_R, and proj is the canonical projection. The Riemannian geodesics are obtained by the same sequence of compositions, changing $\Phi_R(1)$ to $\Phi_R(t)$. Since $H_{sR} = H_R - H_G$, and the flows for H_R, H_{sR} and H_R all commute according to claim (1), we have that $\Phi_{sR}(t) = \Phi_R(t) \circ \Phi_G(-t)$. Applying the projection proj $: T^*Q \to Q$ to this formula, and considering the claim (2), we find that the normal sub-Riemannian geodesics all have the form

$$\exp_R(tv) \exp_G(-tw). \tag{5}$$

These sub-Riemannian geodesics must be horizontal, which leads to $w = A(v) \in \mathfrak{g}$, where $A \in \Omega^1(Q, \mathfrak{g})$ is the connection one-form annihilating the horizontal space.

As the flow Φ_G is purely vertical ($\pi(qg) = \pi(q)$ for any $g \in G$), (5) projects to $\pi(\exp_R(tv))$ in the quotient $M = Q/G$, and its horizontal lift will be (5). This proves that the lifting of any projected geodesic for the extended metric is a normal sub-Riemannian geodesic.

Finally, observe that if $\gamma : I \to Q$ is a horizontal geodesic, its projection $\pi \circ \gamma$ is a geodesic in M. Reciprocally, if $\pi \circ \gamma$ is a geodesic, as the distance between points in M is the distance between fibers in Q, $\pi(\gamma(t))$ is a geodesic if and only if $\gamma(t)$ minimizes distances between fibers, but this implies that this curve is normal to the fibers (if it were not normal, we could construct a shorter curve, leading to a contradiction), and therefore it is a sub-Riemannian geodesic. $\quad\square$

4. Classical particles in Yang-Mills fields

What we aim to show in this section is that, given a principal G-bundle $\pi : Q \to M$ with metric of bundle type $(\mathcal{H}, \langle \cdot, \cdot \rangle)$, there is a system of differential equations that describes the projected normal geodesics $\pi \circ \gamma$, where $\gamma : I \to Q$ is a horizontal geodesic. When the fiber is a circle, these projected equations arise the *Lorentz equations of electromagnetism* in a concrete system. When the fiber is a general Lie group G, these equations are called the *nonabelian Lorentz equations,* also known as *Wong's equations* ([6], p. 209). Moreover, the results of this section are useful to find the sub-Riemannian geodesics in cases where Theorem 7 can not be applied, which may happen when the Riemannian geodesics on M are not known, or are harder to compute than the sub-Riemannian geodesics.

The sub-Riemannian geodesic equations are ordinary differential equations on T^*Q. For metrics of bundle type, the Hamiltonian $H : T^*Q \to \mathbb{R}$ is G-invariant, and hence the

action of G on T^*Q takes sub-Riemannian geodesics to sub-Riemannian geodesics because of Theorem 6. Therefore, we should be able to find differential equations for projections of the sub-Riemannian geodesics on T^*Q/G.

4.1. Associated bundles.

Definition 41. Let G be a Lie group. A *representation of G on a vector space E* is a Lie group homomorphism $\rho : G \to \mathrm{GL}(E)$.

Definition 42 ([6], p. 210). Let $\pi : Q \to M$ be a principal G-bundle, and $\rho : G \to \mathrm{GL}(E)$ be a representation of G on a vector space E. Consider the right action on $Q \times E$ given by $(q, v) \cdot g = (q \cdot g, \rho(g^{-1})v)$. The orbit space associated to this action is denoted as $(Q \times E)/G$, $Q \times_\rho E$ or $\rho(Q)$. A point of $\rho(Q)$ corresponding to the orbit of (q, v) is written as $[q, v]$ or qv.

The associated bundle of $\pi : Q \to M$ is the vector bundle $\pi_{\rho(Q)} : \rho(Q) \to M$ with fiber E, where $\pi_{\rho(Q)}(qv) = \pi(q)$. The sum between two elements qv, qw in the same fiber is $q(v + w)$.

The associated bundle is a vector bundle because the fibers are isomorphic to E, as $\pi_{\rho(Q)}^{-1}(q) = \{[q, v] : v \in E\}$. Moreover, the local trivialization property holds: as $\pi : Q \to M$ is a principal G- bundle, for every $m \in M$, there is an open neighborhood U of m and a diffeomorphism $\phi : \pi^{-1}(U) \to U \times G$, $q \mapsto (m, \varphi(q))$ that realizes the local trivialization property of Q; with this, the diffeomorphism that realizes the local trivialization property of $\rho(Q)$ is $\Phi : \pi_{\rho(Q)}^{-1}(U) \to U \times E$, $[q, v] \mapsto (m, \rho(\varphi(q))v)$ ([13], p. 37).

Example 18 ([6], p. 210). Let $\xi : R \to M$ be a real vector bundle of rank n, and $\pi : Q \to M$ is the principal $\mathrm{GL}(n, \mathbb{R})$-bundle whose fibers are

$$\pi^{-1}(m) = \left\{ q : \mathbb{R}^n \to \xi^{-1}(m) \mid q \in \mathrm{Isom}(\mathbb{R}^n, \xi^{-1}(m)) \right\},$$

that is, each fiber consists of all frames for $\xi^{-1}(m)$, and each element $q \in Q_m$ is a linear isomorphism between $\mathbb{R}^n$ and R_m, with the given basis being the image of the standard basis of $\mathbb{R}^n$. The principal $\mathrm{GL}(n, \mathbb{R})$-bundle π is called *frame bundle* of the vector bundle ξ.

The action of $\mathrm{GL}(n, \mathbb{R})$ on Q_m is given by $q \mapsto qg := q \circ g$, and $\mathrm{GL}(n, \mathbb{R})$ acts on $\mathbb{R}^n$ in the standard way. Now, let us show that the associated bundle of $\pi : Q \to M$ is $\xi : R \to M$. We consider the standard representation of $\mathrm{GL}(n, \mathbb{R})$, that is, the identity map $\rho : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$, so that the action on $Q \times \mathbb{R}^n$ is given by $(q, v) \cdot g = (q \circ g, g^{-1}(v))$, where $q \in Q$, $v \in \mathbb{R}^n$ and $g \in \mathrm{GL}(n, \mathbb{R})$. Now, the orbit space associated to this action, $\rho(Q)$, is diffeomorphic to R through the maps

$$\psi : \rho(Q) \to R, \ qv \mapsto q(v), \quad \psi^{-1} : R \to \rho(Q), \ r \mapsto [q, q^{-1}(r)], \ \text{where } q \in \pi^{-1}(\xi(r)),$$

and the following diagram commutes because $\xi \circ \psi(qv) = \xi(q(v)) = \pi(q) = \phi(qv)$. Explicitly, $\xi(q(v)) = \pi(q)$ because $q \in \pi^{-1}(m)$ and $q(v) \in \xi^{-1}(m)$ for a given $m \in M$, so that $\xi(q(v)) = m = \pi(q)$. Therefore, the associated bundle of the frame bundle $\pi : Q \to M$ is $\xi : R \to M$.

$$\begin{array}{ccc} \rho(Q) & \overset{\psi}{\longrightarrow} & R \\ {\scriptstyle\phi}\downarrow & & \downarrow{\scriptstyle\xi} \\ M & \underset{1_M}{\longrightarrow} & M \end{array}$$

This example can be extended to vector bundles endowed with additional structure. For example, for complex vector bundles we take Q_m as the space of complex linear maps $\mathbb{C}^n \to R_m$; for a vector bundle ξ with a fiber metric we take Q_m to be the space of linear isometries $\mathbb{R}^n \to R_m$; and if ξ is complex with a fiber-Hermitian structure we take Q_m to consist of the unitary maps $\mathbb{C}^n \to R_m$.

The adjoint representation of a Lie group G on its Lie algebra $\mathfrak{g}$ is given by the adjoint map $\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g})$, where $\mathrm{Ad}(g) = \mathrm{Ad}_g = d(\mathrm{ad}_g)_e : \mathfrak{g} \to \mathfrak{g}$. Its dual representation is $\mathrm{Ad}^* : G \to \mathrm{Aut}(\mathfrak{g}^*)$, $\mathrm{Ad}^*(g) = \mathrm{Ad}(g^{-1})^*$, defined on the dual Lie algebra $\mathfrak{g}^*$. From these representations, given a principal G-bundle $\pi : Q \to M$, we can construct the associated bundles $\mathrm{Ad}(Q)$ and $\mathrm{Ad}^*(Q)$, called *adjoint and co-adjoint bundles of $\pi : Q \to M$,* respectively. These two bundles are *dual vector bundles over M,* which means that fibers of $\mathrm{Ad}^*(Q)$ are the dual spaces to the fibers of $\mathrm{Ad}(Q)$.

Example 19 ([6], p. 211). If $\pi : Q \to M$ is the frame bundle of some vector bundle $\xi : R \to M$, then $\mathrm{Ad}(Q)$ can be identified with the bundle whose fiber over $m \in M$ is the space of linear maps on ξ_m in the sense that $\mathrm{Ad}(Q)_m \cong \mathrm{Hom}(\xi_m, \xi_m)$, because $\mathrm{Ad}(Q)_m \cong \mathfrak{gl}(n, \mathbb{R})$.

Definition 43. Let Q be a manifold, G a Lie group that acts on Q, E a vector space and $\rho : G \to GL(E)$ a representation of G on E. We say that $F : Q \to E$ is *G-equivariant* if $F(q \cdot g) = \rho(g^{-1})F(q)$. The set of all G-equivariant maps will be denoted as $C_G(Q, E)$.

Proposition 14. *If $\pi : Q \to M$ is a principal G-bundle and $\rho : G \to \mathrm{GL}(E)$ is a representation of G, the module of sections of $\rho(Q) \to M$ is in one-to-one correspondence with $C_G(Q, E)$.*

Proof. Assign to each $F \in C_G(Q, E)$ the section s defined by $s(m) = [q, F(q)] \in \rho(Q)_m$ whenever $\pi(q) = m \in M$. This map is well-defined because if $h = q \cdot g$, $[h, F(h)] = [q \cdot g, \rho(g^{-1})F(q)] = [q, F(q)]$. It is a section, since $\pi \circ s = 1_M$, and it is smooth because it can be seen as the quotient with respect to the G-action of the smooth application $1_Q \times F : Q \to Q \times E$.

On the other hand, if $s : M \to \rho(Q)$ is a section, we can construct a G-equivariant map F in the following way: the map $s \circ \pi : Q \rho(Q)$ sends an element $q \in Q$ to the fiber $\pi_{\rho(Q)}^{-1}(\pi(q))$, because s is a section; therefore, $s \circ \pi(q) = [q \cdot g, v]$, for some $g \in G$ and $v \in E$ that depend on $\pi(q)$. The element of this equivalence class whose first component is q is $(q, \rho(g)v)$, because $[q \cdot g, v] = [q, \rho(g)v]$. We set then $F(q) = \rho(g)v$, where $1_q \times (s \circ \pi)(q) = [q, \rho(g)v]$. This map is G-equivariant because $F(q \cdot h)$ is determined by the same equivalence class of $F(q)$, as $s \circ \pi(q \cdot h) = s \circ \pi(q)$, but in this case $[q \cdot g, v] = [(q \cdot g) \cdot (g^{-1}h), \rho(h^{-1}g)v] =$

$[q \cdot h, \rho(h^{-1}g)v]$, so that $F(q \cdot h) = \rho(h^{-1})F(q)$, as desired. This assignation is smooth because it is equivalent to to the smooth map $\Psi \circ \phi : Q \times (G \times E) \to E$, where $\phi : Q \times (G \times E) \to Q \times \rho(Q)$, $(q, (g, v)) \mapsto (q, [q \cdot g, v])$, and $\Psi : Q \times \rho(Q) \to E$, $(q, [q \cdot g, v]) = \rho(g)v$.

Finally, to show that this is a one-to-one correspondence, observe that the associated G-equivariant map of the section $s(m) = [q, F(q)]$ is F, while the associated section of the G-equivariant map $F(q) = \rho(g)v$ is $s(m) = [q, \rho(g)v] = [q \cdot g, v]$, as expected. $\qquad\square$

Example 20 ([6], p. 211). A function $\xi : Q \to \mathfrak{g}$ which is Ad-invariant, meaning that $\xi(q \cdot g) = \mathrm{Ad}_{g^{-1}} \xi(q)$, defines a section of the adjoint bundle.

This relation between bundles and sections extends to forms, as will be seen next.

Definition 44. A k- *form ω with values in the associated bundle $\rho(Q)$ is a multilinear totally skew assignation, taking $v_1, \ldots, v_k \in T_m M$ to a vector $\omega_m(v_1, \ldots, v_k) \in \rho(Q)_m$. We denote as $\Omega^k(M, \rho(Q))$ the space of these forms.*

Any k-form $\omega \in \Omega^k(M, \rho(Q))$ is *horizontal* if it annihilates any vertical vector, that is, $\omega_m(V, \cdot, \ldots) = 0$ whenever V is a vertical vector in Q. For such a form, $\omega_m(X_1 + V_1, \ldots, X_k + V_k) = \omega_m(X_1, \ldots, X_k)$ for V_i vertical.

Proposition 15. *The space of k-forms on Q with values in E which are G-equivariant and horizontal (denoted as $\Omega_G^k(Q, E)$) is isomorphic to $\Omega^k(M, \rho(Q))$.*

Proof. For a given $\tilde{\omega} \in \Omega_G^k(Q, E)$, we assign to it the form $\omega \in \Omega^k(M, \rho(Q))$ given by $\omega_m(v_1, \ldots, v_k) = [q, \tilde{\omega}(z_1, \ldots, z_k)]$, $v_i \in T_m M$, where the z_i are any k vectors in $T_q Q$ projecting onto the v_i, that is, $d\pi_q(z_i) = v_i$. $\qquad\square$

Given a projection $\pi : Q \to M$ and a Lie group G acting on the total space Q, the infinitesimal generators of the G-action $\sigma_v : Q \to T_q Q$, $v \in \mathfrak{g}$,

$$\sigma_v(q) = \frac{d}{d\epsilon}\Big|_{\epsilon=0} q \cdot \exp(\epsilon v),$$

satisfy $V_q = \mathrm{Ker}\, d\pi_q = \{\sigma_v(q) : v \in \mathfrak{g}\}$. From this, $\sigma_\cdot(q)$ is a linear isomorphism from $\mathfrak{g}$ to V_q. Moreover, the map $\eta_q(v) = \sigma_v(q)$ is equivariant (see definition 43)

$$\eta_{q \cdot g} \circ \mathrm{Ad}_{g^{-1}} = d\alpha_g \circ \eta_q$$

thanks to Proposition 11. Thinking σ as a map from $Q \times \mathfrak{g}$ to V, $\sigma : Q \times \mathfrak{g} \to V$, $\sigma(q, v) = \sigma_v(q)$, we have that it is an *equivariant bundle map* over Q, that is, it is a fiber bundle such that G acts on the total and base spaces, and the following equivariant property holds:

$$\sigma\left((q, v) \cdot g\right) = \sigma(q, v) \cdot g.$$

Proposition 16. *The map $\sigma : \mathrm{Ad}(Q) = (Q \times \mathfrak{g})/G \to V/G$ is an isomorphism between vector bundles over M.*

Proof. The group G acts on $V = \bigsqcup_{q \in Q} V_q$ through the assignment

$$w_q \cdot g = (d\alpha_g)_q(w_q),$$

where $w_q \in V_q$. From this, we can consider the equivalence classes of this action, denoted by $V/G := \{[w_q] : w_q \in V\}$. The map $\sigma \, \mathrm{Ad}(Q) \to V/G$ sends an element $[q, v] \in \mathrm{Ad}(Q)$ to $[\sigma_v(q)] \in V/G$. From the equivariant property of σ, this map is well defines. The inverse map $\sigma^{-1} : V/G \to \mathrm{Ad}(Q)$ is $\sigma^{-1}[w_q] = [q, \sigma_{w_q}^{-1}]$, given that $\sigma.(q) : \mathfrak{g} \to V_q$ is a linear isomorphism. To conclude this proof, we need to show that the following diagram is commutative,

$$\mathrm{Ad}(Q) \xrightarrow{\sigma} V/G \xrightarrow{\sigma^{-1}} \mathrm{Ad}(Q)$$

$$\pi_{\mathrm{Ad}(Q)} \searrow \quad \downarrow \pi_{V/G} \quad \swarrow \pi_{\mathrm{Ad}(Q)}$$

$$M$$

but this is immediate, since $\pi_{Ad(Q)}[q, v] = \pi_{V/G}[w_q] = \pi(q)$. $\qquad\qquad\square$

Under the isomorphism of this Proposition, given the one-to-one correspondence between sections of $\mathrm{Ad}(Q) \to M$ and G-equivariant maps $C_G(Q, \mathfrak{g})$, sections of the adjoint bundle correspond to G-invariant vertical vector fields. This is because such a vector field defines and is uniquely defined by a section of $V/G \to M$. To see this, let $X : Q \to V$ be a G-invariant vertical vector field, that is, $(d\alpha_g)_q(X_q) = X_{q \cdot g}$, and $X_q \in V_q$ for all $q \in Q$. The map $\phi : M \to V/G$, $\phi(m) = [X_q]$, where $q \in \pi^{-1}(m)$ is well defined because of the G-invariant property of X. Now, if $\phi : M \to V/G$ is a section of $V/G \to M$, we can construct the G-invariant vertical vector field $X : Q \to V$, $X_q = w_q$, where $\phi(m) = [w_q]$, that is, we choose the element of the equivalence class $\phi(m)$ that is based on q. This is a smooth vector field because the following diagram is commutative,

$$
\begin{array}{c}
Q \\
{\scriptstyle 1_Q \times 1_Q} \downarrow \qquad \searrow^{X} \\
Q \times Q \qquad\qquad V \\
{\scriptstyle 1_Q \times \pi} \downarrow \qquad\qquad \uparrow_{\eta} \quad \nwarrow^{\eta \circ \chi} \\
Q \times M \xrightarrow[1_Q \times \phi]{} Q \times V/G \xleftarrow[\chi]{} Q \times \mathfrak{g}
\end{array}
$$

where $\eta : Q \times V/G \to V$ maps $(q, [w_{\tilde{q}}])$ to the element of the equivalence class of $[w_{\tilde{q}}]$ that is based on q, and $\chi : Q \times \mathfrak{g} \to V/G$ is the map that assigns (q, v) to $(q, [\Psi(v)])$, where $\Psi : \mathfrak{g} \to V_q$ is the natural isomorphism. The composition $\eta \circ \chi : Q \times \mathfrak{g} \to V$ is the immersion of Ψ to V, from where we can conclude that η is smooth, and in consequence $X : Q \to V$ is smooth.

The isomorphism between sections of $\mathrm{Ad}(Q)$ and V/G will be denoted as $\sigma : \Gamma(\mathrm{Ad}(Q)) \to \Gamma(V)_G$, $\xi \mapsto \sigma\xi$, where Γ denotes the space of sections.

Definition 45 ([6], p. 212). *The gauge group* of $\pi : Q \to M$ consists of all the G-maps $\Phi : Q \to Q$ which induce the identity on M. That is, Φ is a diffeomorphism that satisfies $\Phi(q \cdot g) = \Phi(q) \cdot g$, and it makes the following diagram commutative.

$$
\begin{array}{ccc}
Q & \xrightarrow{\ \ \Phi\ \ } & Q \\
& \searrow{\scriptstyle \pi} \quad \swarrow{\scriptstyle \pi} & \\
& M &
\end{array}
$$

Despite of not being a (finite dimensional) Lie group, we can associate to it a Lie algebra to it, and this algebra corresponds to the space of equivariant vector fields: if $\Phi_t : Q \to Q$ is a one-parameter subgroup of the gauge group, we can consider the vector field

$$
X_q = \left.\frac{d\Phi_t(q)}{dt}\right|_{t=0} \in T_q Q.
$$

This vector field is equivariant, because

$$
(d\alpha_g)_q(X_q) = (d\alpha_g)_q\left(\left.\frac{d\Phi_t(q)}{dt}\right|_{t=0}\right) = d(\alpha_g \circ \Phi_t(q))_{t=0} = \left.\frac{d\Phi_t(q \cdot g)}{dt}\right|_{t=0} = X_{q \cdot g}
$$

thanks to the fact that Φ is a G-map. Moreover, if $X \in \mathfrak{X}(Q)$ is an equivariant vector field, its flow is an element of the gauge group. Therefore, sections of the adjoint bundle form the Lie algebra of the gauge group via the isomorphism $\sigma : \Gamma(\mathrm{Ad}(Q)) \to \Gamma(V)_G$.

Definition 46 ([6], p. 212). Let $\pi : Q \to M$ be a principal G-bundle with connection $\mathcal{H}$. *The connection one-form on Q*, denoted as A, is a Lie-algebra-valued one-form on Q, that is, $A_q : T_q Q \to \mathfrak{g}$, which satisfies the following conditions:

(1) $\mathrm{Ker}\, A_q = \mathcal{H}_q, \ \forall q \in Q$.

(2) $A_q \circ \eta_q = 1_{\mathfrak{g}}, \ \forall q \in Q$.

The second condition of previous definition is a normalization condition: as the image of η_q is V_q, when restricting to the vertical A_q is the canonical identification between V_q and $\mathfrak{g}$. This connection one-form is unique, and it satisfies the following equivariance condition:

$$
\mathrm{Ad}_{g^{-1}} \circ A_q = A_{q \cdot g} \circ (d\alpha_g)_q \circ \eta_q \circ A_q.
$$

As seen in Proposition 16, differentiating the group action on Q yields the infinitesimal generator map $\sigma : \mathrm{Ad}(Q) \to (TQ)/G$, whose image is V/G. $(TQ)/G \to M$ is a vector bundle over M, with projection $[q, w] \mapsto \pi(q)$. Similarly, $d\pi : TQ \to TM$ is equivariant with respect to the G-action, where G acts trivially on M, which means that $d\pi : (TQ)/G \to M$ is well-defined. As $\mathrm{Ker}\, d\pi = \mathrm{Im}\, \sigma$, these maps fit together in the following exact sequence

$$
\mathrm{Ad}(Q) \xrightarrow{\ \sigma\ } (TQ)/G \xrightarrow{\ d\pi\ } TM
$$

of vector bundles over M. This sequence is called the *Atiyah sequence*.

A connection one-form is a splitting of this exact sequence:

$$\mathrm{Ad}(Q) \xleftarrow{\quad A \quad} (TQ)/G \xleftarrow{\quad h \quad} TM,$$

where h is the collection of horizontal lifts $h_q : T_m M \to T_q Q$ composed with the canonical map $TQ \to (TQ)/G$, and A results when we consider the connection one form $A : TQ \to Q \times \mathfrak{g}$, $w_q \mapsto (q, A_q(v))$, and take its quotient by the G-action. These maps satisfy the following algebraic relations:

$$\mathrm{Im}\, h_q = \mathcal{H}_q = \mathrm{Ker}\, A_q,$$
$$\mathrm{Im}\, \eta_q = V_q = \mathrm{Ker}\, d\pi_q,$$
$$A_q \circ \eta_q = 1_{\mathfrak{g}}, \quad d\pi_q \circ h_q = 1_{T_m M}.$$

Proposition 17. *The vector bundles $(TQ)/G$ and $TM \oplus \mathrm{Ad}(Q)$ are isomorphic.*

Proof. Take an element $v \oplus [q, w] \in T_m M \oplus \mathrm{Ad}(Q)$ to $[h_q(v) + \sigma_w(q)] \in (TQ)/G$. The inverse sends $[w_q] \in (TQ)/G$ to $d\pi_q(w_q) \oplus [q, A_q(w_q)]$. The proof that these maps are inverses relies on the algebraic relations stated above, and the commutativity of the diagram

$$TM \oplus \mathrm{Ad}(Q) \longrightarrow (TQ)/G \longrightarrow TM \oplus \mathrm{Ad}(Q)$$
$$\searrow \qquad \downarrow \qquad \swarrow$$
$$M$$

is immediate from the definition. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Definition 47 ([6], definition 4.1.1)**.** *The curvature of a distribution $\mathcal{H}$ is the linear bundle map $F : \Lambda^2 \mathcal{H} \to TQ/\mathcal{H}$ defined by $F(X, Y) = -[X, Y] \bmod \mathcal{H}$.*

The connection one-form A provides an identification $(TQ)/\mathcal{H} \to \mathfrak{g}$, given that the kernel of A_q is precisely $\mathcal{H}_q$.

Definition 48 ([6], p. 214)**.** *The curvature form is a G-equivariant horizontal two-form Ω defined by the assignment $\Omega(X, Y) = -A([X, Y])$, if X, Y are horizontal, and $\Omega(X, Y) = 0$ if X or Y are vertical vectors.*

Definition 49 ([6], p. 214)**.** *Let $\rho(Q)$ be an associated bundle of $\pi : Q \to M$. If $\psi : M \to \rho(Q)$ is a section of this bundle, and if $X \in \mathfrak{X}(M)$, we define the covariant derivative $D_X \psi$ of ψ in the direction of X as follows. The section ψ corresponds to an equivariant map $F : Q \to E$, and hX is a G-equivariant horizontal vector field on Q, so that $dF(hX)$ is an equivariant map $Q \to E$, which corresponds to a section of $\rho(Q)$, $D_X \psi$.*

Definition 50 ([6], p. 214)**.** *Let $c : I \to M$ be a curve. A curve $\psi : \mathrm{Im}(c) \to \rho(Q)$ lying over c is said to be covariantly constant if $D_{\dot{c}} \psi = 0$.*

The G-invariant vector fields on Q can be identified with sections of $(TQ)/G$, and thanks to Proposition 17 we can split them up into horizontal and vertical vector fields, according to $(TQ)/G = TM \oplus \mathrm{Ad}(Q)$. Then, if $X \in \mathfrak{X}(M)$ and $\xi : M \to \rho(Q)$, we have the following relations:

$$[hX, hY] = h([X,Y]) \oplus -\eta(\Omega(X,Y)), \qquad \text{(horizontal-horizontal)}$$
$$[\sigma\xi, \sigma\chi] = -\sigma([\xi,\chi]), \qquad \text{(vertical-vertical)}$$
$$[hX, \sigma\xi] = \eta(D_X\xi). \qquad \text{(horizontal-vertical)}$$

References

[1] John M Lee. *Introduction to Riemannian manifolds*. Springer, 2018.

[2] Barrett O'neill. *Semi-Riemannian geometry with applications to relativity*. Academic press, 1983.

[3] Glenys Luke and Alexander S Mishchenko. *Vector bundles and their applications*, volume 447. Springer Science & Business Media, 2013.

[4] Chris Wendl. Lecture notes on bundles and connections. *Lecture notes, Humboldt-Universität zu Berlin*, 2008.

[5] J.M. Lee. *Introduction to Smooth Manifolds*. Graduate Texts in Mathematics. Springer, 2003.

[6] Richard Montgomery. *A tour of subriemannian geometries, their geodesics and applications*. Number 91. American Mathematical Soc., 2002.

[7] Allen Hatcher. Vector bundles and k-theory. *Im Internet unter http://www. math. cornell. edu/˜ hatcher*, 2003.

[8] Shoshichi Kobayashi and Katsumi Nomizu. *Foundations of differential geometry*, volume 1. New York, London, 1963.

[9] David W Lyons. An elementary introduction to the hopf fibration. *Mathematics magazine*, 76(2):87–98, 2003.

[10] Chenjia Lin. Introductory differential topology and an application to the hopf fibration.

[11] Vladimir Igorevich Arnol'd. *Mathematical methods of classical mechanics*, volume 60. Springer Science & Business Media, 2013.

[12] Hassan Najafi Yunjiang Jiang and Josef Pozny. Poisson geometry. lectures 1, 2, 3. "`https://webspace.science.uu.nl/~crain101/Poisson/poisson-lectures123.pdf`", 2014. [Online; accessed 29-November-2021].

[13] Richard W Sharpe. *Differential geometry: Cartan's generalization of Klein's Erlangen program*, volume 166. Springer Science & Business Media, 2000.

[14] Frank W Warner. *Foundations of differentiable manifolds and Lie groups*, volume 94. Springer Science & Business Media, 2013.