# Geodesics in geometry with constraints and applications

Irina Markina

**Abstract.** In this course we will carefully define the notion of a non-holonomic manifold wich is a manifold with a certain non-integrable distribution. We will define such concepts as horizontal distribution, the Ehresmann connection, bracket generating condition for a distribution, sub-Riemannian structure and sub-Riemannian metric, Hamiltonian system, normal and abnormal geodesics, principal bundle and others.

**Mathematics Subject Classification (2000).** Primary 53C17; Secondary 37K05, 58B25.

## 1. Introduction

These notes are based on the course of lectures presented at the summer school "Analysis - with Applications to Mathematical Physics" that took place at Georg - August - Universität, Göttingen, in August 29 - September 2, 2011. The main purpose of these notes is to give a flavor of the subject that during the last decade received the name *Sub - Riemannian Geometry* and that studies the geometry of manifolds with non - holonomic constraints and presence of a positive definite metric. This subject has attracted attention of scientists since 19-th century. We will not describe the history of development of this subject, we only mention that it was independently considered in several branches of mathematics such as non - holonomic mechanics, geometry of bundles, CR manifolds, geometric control theory and others.

It is supposed that the reader is familiar with basic notions of differential geometry, topology, and Lie groups. Nevertheless, in order to keep self - sufficiency of the notes, we present the main definitions and basic notions related to these

---

I acknowledge support through the Norwegian Research Council, project # 204726 /V30 and the Mittag-Leffler Institute, Stockholm, Sweden during fall 2012.

topics in the Appendix A. It is advisable to consult first the Appendix A, if the reader meets an unfamiliar notion.

The principal subject of the sub - Riemannain geometry, discussed in the notes, is the notion of geodesics related to sub - Riemannian Hamiltonian functions produced by a sub - Riemannian metric. We present basic models where the sub-Riemannian geometry appears rather naturally. Based on these examples we show main features and peculiarities of geodesics in geometry with non-holonomic constraints. The structure of these notes is as follows. Section 2 collects main definitions that reveal the similarity and difference between Riemannian and sub - Riemannian geometries. Carnot groups and their particular examples are presented in Section 3. We describe the sub-Riemannian structure of odd dimensional spheres in Section 4. Section 5 deals with principal bundles. After presenting main definitions we reconsider examples of Sections 3 and 4 from the point of view of principal bundles. Sections 6 is dedicated to a mechanical problem of rolling one manifold over another, where kinematic constraints are described in the language of a smooth sub - bundle of the tangent bundle of the configuration space. In the last Section 7 we generalize some results obtained for principal bundles on the infinite dimensional Lie group of orientation preserving diffeomorphisms of the unit circle. Appendix A collects a vast of definitions and concrete formulas used in the text. Some of them are well known, some of them are not widely presented in the literature. As it was noticed above, we recommend for a not very experienced reader to start reading the notes from Appendix A. Appendix B is short and very technical, where we wrote some of the expressions that are useful, but not necessary for the first reading.

## 2. Main definitions

### 2.1. Smooth manifolds, vector fields, tangent map

It is supposed that the reader is familiar with the notion of smooth or $C^\infty$ manifolds. We set up main definitions and notations. A smooth manifold is a Hausdorff, second countable topological space, where the smooth complete atlas is defined. We write $M$ for a smooth manifold, or rather $M^n$ if we want to emphasize the dimension $n$ of the manifold. Let $C^\infty(M)$ denote the space of smooth real valued functions defined on $M$.

The tangent space at a point $q \in M$ is denoted by $T_qM$. Recall that any element $v_q \in T_qM$ is a function $v_q \colon C^\infty(M) \to \mathbb{R}$ satisfying two properties

1. $\mathbb{R}$-linearity: $v_q(af+bg) = av_q(f) + bv_q(g)$,
2. Leibnizian property: $v_q(fg) = v_q(f)g(q) + f(q)v_q(g)$

for all $a, b \in \mathbb{R}$, $f, g \in C^\infty(M)$, $q \in M$. The space $T_qM$, $q \in M$, is a real vector space and therefore $v_q$ is called *tangent vector*.

Equivalently, a tangent vector $v_q$ at $q \in M$ can be defined as an equivalence class of parameterized curves through $q$, as follows. Let $\varphi \colon U \to V$, $U \subset M$, $V \subset \mathbb{R}^n$ be a coordinate chart with $q \in U$ and let $\gamma_1, \gamma_2$ be two smooth curves

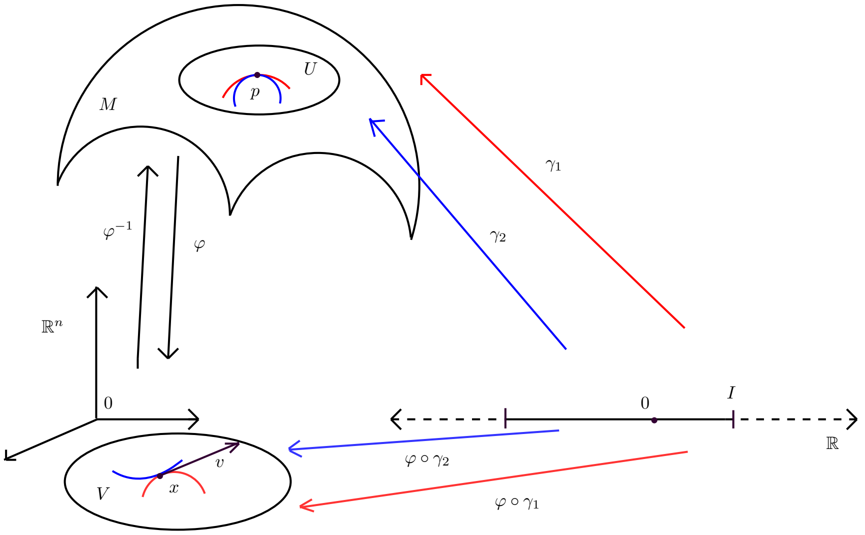

FIGURE 2.1. The notion of the tangent vector

defined on an interval $I \subset \mathbb{R}$ containing 0 such that $\gamma_i(0) = q$. We say that $\gamma_1$ and $\gamma_2$ are "equivalent" or have the same "velocity vector" at $t = 0$ if two smooth maps

$$I \ni t \mapsto \varphi(\gamma_i(t)) \in V, \quad i = 1, 2,$$

have the same first derivatives at $t = 0$. The set of all equivalence classes of curves through $q$ is called the *tangent vector space* $T_qM$. Note that $\varphi^{-1}$ induces a one-to-one correspondence between the model space $\mathbb{R}^n$ and the tangent space $T_qM$. In fact, if $\varphi^{-1}(x) = q$, $x \in V \subset \mathbb{R}^n$, then each vector $v \in \mathbb{R}^n$ corresponds to the equivalence class of the curve

$$\big[t \mapsto \varphi^{-1}(x + tv)\big] \in T_qM.$$

The visualisation is presented at Figure 2.1.

After the previous definitions one can also say, that the notion of a tangent vector $v$ is the generalization of the derivative of $C^\infty$-functions along the direction $v$. If the chart $\big(U, \varphi = (x^1, \dots, x^n)\big)$ is chosen, then the standard notation for the basis of $T_qM$ is $\left(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\right)$ or shortly $(\partial_1, \dots, \partial_n)$. Any vector $v \in T_qM$ will be written in coordinates as $v = \sum_{j=1}^n v^j \partial_j$. Notice the position of indices!

The dual space to $T_qM$ is denoted by $T_q^*M$ and the pairing is written as $\langle \cdot, \cdot \rangle_q$, where we usually omit the subscript "$q$", see Definition 42. The dual basis to $(\partial_1, \dots, \partial_n)$ with respect to the pairing is denoted by $(dx^1, \dots, dx^n)$ and, by definition, it satisfies $\langle dx^i, \partial_j \rangle = \delta_{ij}$, where $\delta_{ij}$ is the Kronecker symbol. Then any co-vector $\lambda \in T_q^*M$ is written in coordinates as $\lambda = \sum_{k=1}^n \lambda_k dx^k$. Notice the position of indices. The elements of $T_q^*M$ are usually called *co-vectors* in geometry and *momenta* in physics.

The tangent and co-tangent bundles are denoted by $TM$ and $T^*M$, correspondingly, consult Definition 59. Both vector bundles are $C^\infty$-smooth manifolds [23, 117]. The notations

$$\begin{array}{cccc} \mathrm{pr}_M\colon & TM & \to & M \\ & (q,v) & \to & q \end{array} \qquad \text{and} \qquad \begin{array}{cccc} \mathrm{pr}^*_M\colon & T^*M & \to & M \\ & (q,\lambda) & \to & q \end{array}$$

will be fixed for the canonical projections from the tangent and co-tangent bundles to the underlying manifold.

A *vector field* $X$ on a manifold $M$ is a function that assigns to each point $q \in M$ a tangent vector $X(q) \in T_qM$. We also write $X_q$ for the value of the vector field $X$ at the point $q \in M$. If $f \in C^\infty(M)$, then $Xf$ denotes a real valued function on $M$ given by

$$(Xf)(q) = X(q)f, \qquad \text{for all} \quad q \in M.$$

A vector field $X$ is called smooth if for any $f \in C^\infty(M)$ the function $Xf\colon M \to \mathbb{R}$ is an element in $C^\infty(M)$. If $\big(U, \varphi = (x^1, \dots, x^n)\big)$ is a coordinate chart, then any vector field $X$ can be written in terms of coordinates as $X(q) = \sum_{j=1}^n X^j(q)\partial_j$. Then the smoothness condition of the vector field $X$ on the neighborhood $U$ is equivalent to the requirement that all functions $X^j$, $j = 1, \dots, n$, are of class $C^\infty(U)$. If the functions $X^j$, $j = 1, \dots, n$, are analytic in $U$, then the corresponding vector field $X$ is called an *analytic* vector field.

Another way to define a vector field $X$ is to use the definition of a local section. Namely, a vector fields $X$ is a smooth map $X\colon U \to TM$, such that $\mathrm{pr}_M \circ X = \mathrm{id}_U$ for any open set $U \subset M$. The section is global if $U$ can be taken as entire $M$. We write $\operatorname{Vect} M$ ($\operatorname{Vect} U$) for the collection of smooth vector fields, defined on $M$ ($U$, $U \subset M$). Algebraically, $\operatorname{Vect} M$ is a module over the ring $C^\infty(M)$ and a vector space over the field $\mathbb{R}$ (or $\mathbb{C}$ if the manifold $M$ is modelled over $\mathbb{C}^n$). Moreover, an operation of multiplication of two vector fields can be defined. The multiplication $[\cdot\,,\cdot]$ (that received the name *commutator* or the *Lie product*) is defined by

$$[X,Y]f = X(Yf) - Y(Xf). \tag{2.1}$$

The Lie product is a map $[\cdot\,,\cdot]\colon \operatorname{Vect} M \times \operatorname{Vect} M \to \operatorname{Vect} M$ satisfying the three axioms of Definition 51. The set of smooth vector fields considered as a real vector space endowed with the Lie multiplication forms a Lie algebra.

**Definition 1.** *Let $M$ and $N$ be two smooth manifolds and $F\colon M \to N$ be a map. The map $F$ is smooth if the following holds. For any $q \in M$ and for any local charts $(U,\varphi)$ of $q \in M$ and $(V,\psi)$ of $F(q) \in N$, the composition $\psi \circ F \circ \varphi^{-1}$ is a smooth map*

$$\psi \circ F \circ \varphi^{-1}\colon \varphi(U) \to \psi(V)$$

*in the sense of smoothness defined in the Euclidean space $\mathbb{R}^n$.*

The definition is illustrated in Figure 2.2. A diffeomorphism between two manifolds is defined in a similar way.

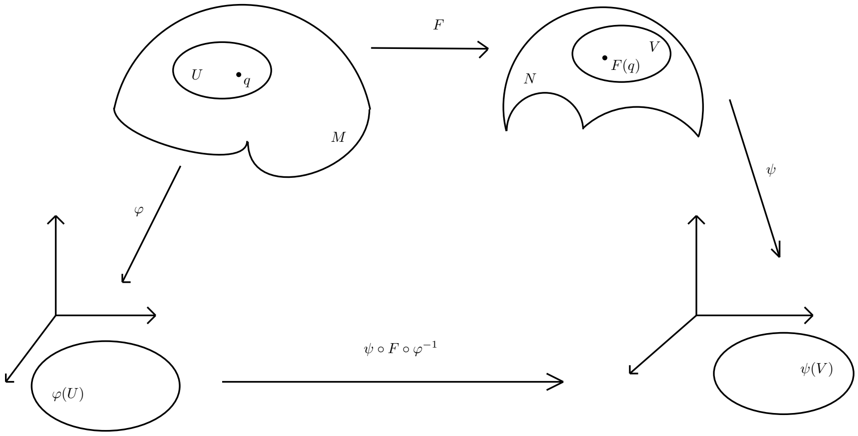

Figure 2.2. The smooth map $F$

**Definition 2.** *Let $F\colon M \to N$ be a smooth map. The differential of $F$ at $q \in M$ is the linear map $d_qF\colon T_qM \to T_{F(q)}N$ such that*

$$\big(d_qF(X_q)\big)f := X_q(f\circ F)$$

*for any $f \in C^\infty(N)$ and $X_q \in T_qM$.*

If the local charts $\big(U, \varphi = (x^1,\dots,x^m)\big)$ of $q \in M$ and $\big(V, \psi = (y^1,\dots,y^n)\big)$ of $F(q) \in N$ are chosen, then

$$d_qF(\partial_{x^j}) = \sum_{k=1}^{n} \frac{\partial}{\partial x^j}\Big(y^k\big(F(q)\big)\Big)\partial_{y^k}|_{F(q)}, \quad j = 1,\dots,m.$$

The matrix $\Big\{\frac{\partial}{\partial x^j}\Big(y^k\big(F(q)\big)\Big)\Big\}_{k,j}$ is called the *Jacobi matrix* of the map $F$ with respect to the given coordinate charts.

### 2.2. Distributions and non-holonomic constraints

**Definition 3.** *Let $M$ be a smooth manifold. A mapping $D$ that assigns to every point $q \in M$ a linear subspace $D_q$ of the tangent space $T_qM$ is called a singular distribution on $M$.*

**Definition 4.** *A distribution $D$ is called smooth on $M$, if for any $q \in M$ there is a neighborhood $U(q)$ and smooth linearly independent vector fields $X_1,\dots,X_k$, such that $D_x = \operatorname{span}\{X_1(x),\dots,X_k(x)\}$, for all $x \in U(q)$.*

A distribution $D$ is called *analytic* if the vector fields $X_1,\dots,X_k$ in Definition 4 can be chosen to be analytic. The smooth (analytic) distribution $D$ on $M$ is a smooth (analytic) sub-bundle of the tangent bundle $TM$, and its rank is equal to $k$ for all $q \in M$. In the case of a singular distribution the set of vector fields $X_1, \dots, X_k$ may not be necessarily linear independent, and therefore, the dimension of the linear subspace $D_q$ can vary from point to point.

From now on, we will work only with smooth distributions and smooth manifolds and therefore we omit the word "smooth". Analogous definition can be given for a map $D^*$ that assigns to any point $q \in M$ a linear subspace in the co-tangent space $T^*_q M$ and in this case it is called a *co-distribution*.

The notion of a smooth distribution naturally leads to the following question. When does a smooth distribution or a smooth sub-bundle $D \subset TM$ define a submanifold $N$ inside of the original manifold $M$? The answer was given by Frobenius [48].

**Definition 5.** *A smooth distribution $D$ on $M$ is called involutive or integrable if $[X, Y]$ is a smooth section of $D$ for any choice of smooth sections $X$ and $Y$ of $D$.*

**Definition 6.** *A smooth submanifold $N$ of a manifold $M$ is the integral manifold of a distribution $D$ if for any point $q \in N$ there is an open neighborhood $U(q) \subset N$ such that $T_x N = D_x$ for any $x \in U(q)$.*

**Theorem 2.1.** [48, 131] *A submanifold $N$ of a manifold $M$ is the integral manifold of a distribution $D$, if and only if, $D$ is involutive.*

In this case a foliation of the manifold $M$ by integral manifolds $N$ passing through different points $q \in M$ is produced. Somehow, one can not leave a chosen leaf of the foliation produced by the integral manifold $N$ of $D$ and being touched to the distribution $D$.

A smooth curve $c \colon I \to M$ can be considered as a smooth map between two manifolds. In this case the image of the tangent vector $\frac{\partial}{\partial r} \in T_t I$ under the tangent map $d_t c \colon T_t I \to T_{c(t)} M$ is denoted by $\dot{c}(t)$, i. e. $d_t c\big(\frac{\partial}{\partial r}\big) = \dot{c}(t)$, and is called the *velocity vector* of the curve $c$ at $t \in I$.

**Definition 7.** *We say that a smooth curve $c \colon I \to M$ is tangent to the distribution $D$ (or horizontal) if the tangent vector $\dot{c}(t)$ belongs to the vector space $D_{c(t)}$ for any $t \in I$, whenever $\dot{c}(t)$ is defined.*

One can release the condition of smoothness for the curve $c$ and require that the curve has derivative almost everywhere on the interval $I$. If a distribution $D$ is involutive then given a point $q$ on its integral manifold $N$ one can reach only the points on $N$ being tangent to $D$. Let us ask the opposite question: when can we reach any point (of the original manifold $M$) starting from a given one and always staying tangent to the prescribed distribution $D$?

To answer this question we introduce a flag of distributions. Let $X$ be a vector field such that $X_q \in D_q$ for all $q \in M$. We denote by $D + [X, D]$ the sub-bundle of $TM$ spanned by $D$ and all the vector fields $[X, Y]$, where the vector field $Y$ is such that $Y_q \in D_q$, $q \in M$. Thus

$$D_q + [X, D]_q = \operatorname{span}\{D_q, [X, Y]_q \mid \ \forall\, Y_q \in D_q, \quad q \in M\}.$$

We also drop the subscript $q$ and write $D+[X,D]$. We define the $k$-bracket $(k,X)$ inductively by

$$\text{bracket } (2,X)=D+[X,D],\ \dots,\ \text{bracket } (k,X)=D+[\text{bracket}(k-1,X),D].$$

More generally, changing the vector field $X$ to the entire distribution $D$, we set

$$\text{bracket } (2,D)=D+[D,D],\ \dots,\ \text{bracket } (k,D)=D+[\text{bracket}(k-1,D),D].$$

We get a flag of distributions

$$D\subset \text{bracket } (2,D)\subset\ldots\subset \text{bracket } (k,D)\subset\ldots.$$

A smooth section $X$ of $D$ is a *$k$-step generator* if

$$\text{bracket } (k,X_q)=T_qM \quad \text{for any} \quad q\in M.$$

Similarly, a distribution $D$ is said to be the *$k$-step bracket generating* (or *completely non-holonomic*) distribution if

$$\text{bracket } (k,D_q)=T_qM \quad \text{for every} \quad q\in M.$$

We say that a distribution $D$ is *strongly bracket generating* if $D_q+[X,D]_q=T_qM$ for all non-vanishing $X_q\in D_q$. If we do not emphasize the number of steps for $k$-step bracket generating distribution, then we simply say *bracket generating distribution.*

**Example 1.** *Consider vector fields in* $\mathbb{R}^4$ *written in coordinates* $(x,y,z,w)$*:*

$$X_1=\frac{\partial}{\partial x},\quad X_2=\frac{\partial}{\partial y},\quad X_3=\frac{\partial}{\partial z}+x\frac{\partial}{\partial w}.$$

*The distribution* $D=\operatorname{span}\{X_1,X_2,X_3\}$ *is 2-step bracket generating but not strongly bracket generating.*

If a distribution $D$ is bracket generating and the dimension of brackets $(k,D_q)$ does not depend on the point $q\in M$ for any $k$, then the distribution $D$ is called *regular.*

Now we are ready to formulate a sufficient condition for the connectivity problem. This condition was independently proved by P. K. Rashevskiĭ [119] and W. L. Chow [33].

**Theorem 2.2.** [33, 119] *If a manifold $M$ is topologically connected and if a distribution $D$ on $M$ is bracket generating, then any two points on $M$ can be connected by a piecewise smooth curve tangent to $D$.*

Necessary and sufficient conditions for the connectivity problem in the case of $C^\infty$-manifold and $C^\infty$-smooth distribution can be found in [128]. See also references therein.

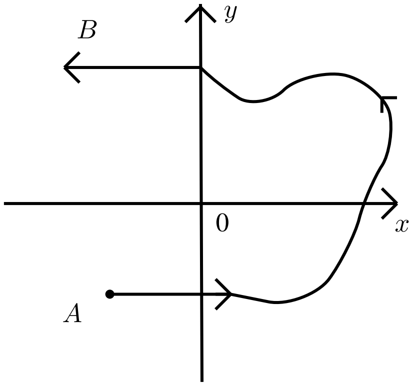

FIGURE 2.3. $(\mathbb{R}^2, D)$ is horizontally connected, but $D$ is not bracket generating.

**Example 2.** In the following example we show that the Chow-Rashevskiĭ condition is not necessary for connectivity. Let $M = \mathbb{R}^2$, $X_1 = \frac{\partial}{\partial x}$, $X_2 = \phi(x)\frac{\partial}{\partial y}$, where the $C^\infty$-function $\phi$ satisfies

$$\begin{array}{lll} \phi(x) > 0, & \text{if} & x > 0, \\ \phi(x) = 0, & \text{if} & x \leq 0. \end{array}$$

It is clear that one can not move vertically in the left half-plane, but one can move horizontally to the right half-plane, displace arbitrarily in the right half-plane and proceed to the left half-plane, see Figure 2.3. In this example one can connect any points in the plane being tangent to the distribution $D = \text{span}\{X_1, X_2\}$, but the vector fields definitely do not span the entire plane at points $q = (x, y)$ with $x \leq 0$.

**Example 3.** Another example of a bracket generating distribution is the Grušin distribution spanned by vector fields in $\mathbb{R}^2$

$$X_1 = \frac{\partial}{\partial x}, \quad X_2 = x\frac{\partial}{\partial y},$$

studied by M. S. Baouendi in his PhD thesis in early 70-s, and then by numerous authors, see for instance, [6, 25, 40, 52, 67, 129].

The latter two examples are based on non-smooth distributions.

**Example 4.** Historically the integrability condition was given in terms of one-forms, but not in terms of vector fields. Let a manifold $M$ be of dimension $n$ and we want to describe a distribution $D \subset TM$ of rank $k$, $k < n$. To achieve this we need to find $n-k$ one-forms $\Theta_1, \ldots, \Theta_{n-k}$, such that the distribution $D$ belongs to their common kernel. The forms $\Theta_j$, $j = 1, \ldots, n-k$, are called annihilators of $D$. It is equivalent to solve the system

$$\begin{cases}\Theta_1(x^1,\dots,x^n)=0\\ \dots\dots\dots\dots\dots\\ \Theta_{n-k}(x^1,\dots,x^n)=0,\end{cases}$$

that received the name Pfaffian equations. This system is integrable if the one-forms $\Theta_1,\dots,\Theta_{n-k}$ are exact forms:

$$\begin{cases}\Theta_1(x^1,\dots,x^n)=d\theta_1(x^1,\dots,x^n)=0\\ \dots\dots\dots\dots\dots\\ \Theta_{n-k}(x^1,\dots,x^n)=d\theta_{n-k}(x^1,\dots,x^n)=0.\end{cases}\tag{2.2}$$

After integrating the latter system we get $n-k$ functions describing a $k$-dimensional integral submanifold of $M$ defined by the integrable system (2.2) or by the involutive distribution $D$.

The Chow-Rashevskiĭ Theorem 2.2 for an analytic co-rank one distribution $D$, or for one Pfaffian equation was solved by C. Carathéodory. The result states as follows. *Let $M$ be a connected manifold endowed with an analytic co-rank one distribution $D$. If there exist two points $A$, $B\in M$ that cannot be connected by a horizontal curve, then the distribution $D$ is integrable.* Or, formulating the negation of the above statement *if for any points $A,B\in M$ there is a horizontal curve connecting these points, then the distribution $D$ is non-integrable* (*completely non-holonomic, bracket generating*).

C. Carathéodory developed this theory due to the question posted by M. Born to derive the second law of thermodynamics and the existence of the entropy function. Translating the problem into the geometric language we work with a manifold $M$ that is the set of all possible thermodynamical states of some isolated system. The admissible or horizontal curves are adiabatic curves, such curves that correspond to slow processes in time and such that during these processes (along the admissible curves) no heat $\Theta$ is exchanged. C. Carathéodory wrote the condition of an adiabatic process as a Pfaffian equation $\Theta=0$ on $M$. It was known at that moment from works by S. Carnot, J. P. Joule and others, that there are thermodynamical states $A,B\in M$, which cannot be connected by an adiabatic process (by an admissible curve). Carathéodory's theorem states in this case that the distribution defined by the Pfaffian equation $\Theta=0$ is integrable, that leads to the existence of two functions $T$ (temperature) and $S$ (entropy) that locally satisfy the relation $\Theta=TdS$. This proves the existence of the entropy function $S$, as well as that the adiabatic process remains in the leaf (hypersurface) of the state space $M$ corresponding to the entropy function. The entropy function $S$ tends not to decrease, being constant or increasing, according to the second law of thermodynamics.

Due to the names of S. Carnot and C. Carathéodory involved into this discovery, M. Gromov called the sub-Riemannian geometry as the *Carnot-Carathéodory geometry*.

Exercises.

Decide whether the following distributions $D = \operatorname{span}\{X_1, X_2\}$ in $\mathbb{R}^3$ are bracket generating and regular. Find one forms $\omega$ such that $D = \ker(\omega)$.

1. Heisenberg distribution: $X_1 = \frac{\partial}{\partial x}$, $X_2 = \frac{\partial}{\partial y} + x\frac{\partial}{\partial z}$.
2. Martinet distribution: $X_1 = \frac{\partial}{\partial x}$, $X_2 = \frac{\partial}{\partial y} + x^2\frac{\partial}{\partial z}$.

### 2.3. Riemannian and sub-Riemannian manifolds

Let us recall some notions from the Riemannian geometry and compare basic definitions in the Riemannian and sub-Riemannian settings.

**Definition 8.** *A Riemannian metric is a map $g\colon T_qM \times T_qM \to \mathbb{R}$, which is symmetric, bilinear, positively definite for any $q \in M$, and smoothly varying with respect to $q$.*

If the coordinate chart $\big(U, \varphi = (x^1, \dots, x^n)\big)$ is chosen and $(\partial_1, \dots, \partial_n)$ is the local basis of $T_qM$, $q \in U$, then $g_{ij} = g(\partial_i, \partial_j)$ is the associated matrix to the metric $g$. Smoothness of $g$ means that the matrix $g_{ij}(q) = g_{ij}(x^1, \dots, x^n)$ is a smooth function of $(x^1, \dots, x^n)$ in $\varphi(U)$.

The couple $(M, g)$ is called a Riemannian manifold. It would be more correct to say that the triplet $(M, TM, g)$ is called a Riemannian manifold.

**Definition 9.** *The distance $d(q_0, q_1)$ between two points $q_0, q_1 \in M$ related to the Riemannian metric $g$ is defined by the equality*

$$d(q_0, q_1) = \inf\Big\{\int_0^1 \Big(g\big(\dot c(t), \dot c(t)\big)\Big)^{1/2} dt\Big\},$$

*where the infimum is taken over all curves $c\colon [0,1] \to M$ differentiable almost everywhere in $[0,1]$, and such that $c(0) = q_0$, $c(1) = q_1$.*

We are ready now to define a sub-Riemannian manifold. Let $M$ be a smooth manifold and let $D$ be a smooth distribution (a smooth sub-bundle) of the tangent bundle $TM$.

**Definition 10.** *A map $g_D\colon D_q \times D_q \to \mathbb{R}$ which is symmetric, bilinear, positively definite for any $q \in M$ and smoothly varying with respect to $q$ is called a sub-Riemannian metric.*

**Definition 11.** *The couple $(D, g_D)$ is called a sub-Riemannian structure and the triplet $(M, D, g_D)$ is called a sub-Riemannian manifold.*

If $D = TM$, then Definition 11 is reduced to the definition of a Riemannian manifold. In this sense the sub-Riemannian geometry is a generalization of the Riemannian geometry. The distance function related to a sub-Riemannian metric $g_D$ is defined by

$$d_{c-c}(q_0, q_1) = \inf\Big\{\int_0^1 \Big(g_D\big(\dot c(t), \dot c(t)\big)\Big)^{1/2} dt\Big\}, \tag{2.3}$$

where the infimum is taken over all horizontal curves $c\colon [0,1] \to M$ differentiable almost everywhere in $[0,1]$ and such that $c(0) = q_0$, $c(1) = q_1$. Thus, we have added the horizontality condition, $\dot{c}(t) \in D_{c(t)}$, for the set of admissible curves. The set of admissible curves is smaller, therefore, the $d_{c-c}$-distance is, in general, bigger than the Riemannian distance if both metrics are defined on the manifold and coincide on $D_q$, $q \in M$. Theorem 2.2 guaranties that the set of horizontal curves is not empty and therefore, the function $d_{c-c}$ takes only finite values. The distance $d_{c-c}$ is called the Carnot-Carathéodory distance due to the impact by S. Carnot and C. Carathéodory described in Example 4. Let us suppose that a Riemannian metric $g$ and a sub-Riemannian metric $g_D$ are defined on a smooth manifold $M$, and the Riemannian distance $d$ and the Carnot-Carathéodory distance $d_{c-c}$ on $M$ are produced, respectively. As a result, two metric spaces $(M, d)$ and $(M, d_{c-c})$ and two topological spaces $(M, \tau_d)$ and $(M, \tau_{c-c})$ are defined, where the topology $\tau_d$ is generated by open balls in the $d$-metric and $\tau_{c-c}$ is generated by $d_{c-c}$-balls. It is established that the topological spaces $(M, \tau_d)$ and $(M, \tau_{c-c})$ are equivalent, but the metric spaces $(M, d)$ and $(M, d_{c-c})$ are not in general Lipschitz equivalent, see [111, p. 27], [13, 55, 63, 114]. Example 5 shows non-equivalence of the metric spaces $(M, d)$ and $(M, d_{c-c})$ in some particular cases.

#### 2.3.1. Riemannian and sub-Riemannian gradient.

At the end of the subsection we would like to say some words about the gradient vector field in sub-Riemannian geometry. Let us recall that the gradient on the Riemannian manifold $(M, g)$ is a vector field "grad" such that it is detected by its action on smooth functions by

$$g(\operatorname{grad} f, X) = Xf, \quad \text{for any} \quad X \in \operatorname{Vect} M \quad \text{and} \quad f \in C^\infty(M).$$

If a coordinate chart is chosen, then the gradient can be written as

$$\operatorname{grad} f = \sum_{ij} g^{ij} \frac{\partial f}{\partial x^i} \partial_j, \tag{2.4}$$

where $\{g^{ij}\}_{i,j=1}^n$ is the inverse matrix to $g_{ij} = g(\partial_i, \partial_j)$, $i, j = 1, \ldots, n$. More details about differential operators on Riemannian manifolds can be found in [117]. In the case of a sub-Riemannian manifold $(M, D, g_D)$ the definition is analogous. A sub-Riemannian gradient $\operatorname{grad}_D$ is a horizontal vector field, such that

$$g_D(\operatorname{grad}_D f, X) = Xf, \tag{2.5}$$

for any smooth section $X$ of $D$ and $f \in C^\infty(M)$.

### 2.4. Hamiltonian formalism and geodesics

Let us compare the problem of finding a curve realizing the distance between two points in the Riemannian and sub-Riemannian geometries.

**2.4.1. Geodesic on Riemannian manifolds.** Historically, a geodesic was defined as a curve $\gamma$ that locally realizes the distance between two points on a Riemannian manifold. The corresponding equation is

$$\nabla_{\dot\gamma(t)}\dot\gamma(t)=0,\quad \gamma\colon I\to M, \tag{2.6}$$

where $\nabla$ is the Levi-Civita connection, which is the generalization of the directional derivative of vector fields defined on a Riemannian manifold $(M,g)$, see Definitions 45, 46, and 47. The connection $\nabla$ is compatible with the Riemannian metric $g$, see Theorem 8.1 and [23, p. 53], [117, p. 59]. Geometrically, equation (2.6) also implies, that the corresponding first geodesic curvature of the solution vanishes. The physical interpretation asserts that solutions of equation (2.6) give trajectories of the motion of particles under the absence of any external force, motion of "free particles" or "free motion".

Given a coordinate chart $\big(U,\varphi=(x^1,\dots,x^n)\big)$, the Christoffel symbols of the Levi-Civita connection are introduced by $\nabla_{\partial_i}\partial_j=-\sum_{k=1}\Gamma^k_{ij}\partial_k$. Then equation (2.6) takes the form

$$\ddot x^k(t)=\sum_{ij=1}^n\Gamma^k_{ij}\dot x^i(t)\dot x^j(t),\quad k=1,\dots,n,\quad t\in I. \tag{2.7}$$

Given a Riemannian metric $g$ there is a predeterminant choice of the dual space $T^*_qM$ to the tangent space $T_qM$ given as follows. If $v\in T_qM$, then $v^*(\cdot)=g(v,\cdot)\colon T_qM\to\mathbb{R}$ is a continuous linear functional, and therefore, an element of $T^*_qM$. We write it in coordinates. Let $\{\partial_j\}_{j=1}^n$ be a basis of $T_qM$ and let $\{g_{ij}\}$ be the matrix associated with the metric $g$, then $dx^i=\sum_{j=1}^n g_{ij}\partial_j$, $i=1,\dots,n$, represent the basis of the dual $T^*_qM$. If $v=\sum_{j=1}^n v^j\partial_j$ then $v^*=\sum_{i=1}^n v^*_i dx^i$, where $v^*_i=\sum_{j=1}^n g_{ij}v^j$. This process is called "lowering indices" in physics.

We can say now that the Riemannian metric $g$ defines a map $\tilde g\colon T_qM\to T^*_qM$, which is an isomorphism between two vector spaces. Therefore, the inverse map $\tilde g^{-1}\colon T^*_qM\to T_qM$ is defined. The map $\tilde g^{-1}$ defines a metric on $T^*_qM$, called a *co-metric*, which we denote by $g^{-1}$. Thus, the co-metric is the map $g^{-1}\colon T^*_qM\times T^*_qM\to\mathbb{R}$ defined by

$$g^{-1}(v^*,w^*)=v^*\big(\tilde g^{-1}(w^*)\big)=g\big(\tilde g^{-1}(v^*),\tilde g^{-1}(w^*)\big).$$

We see that maps $\tilde g$ and $\tilde g^{-1}$ became linear isometries between $T_qM$ and $T^*_qM$ for all $q\in M$. The matrix corresponding to $g^{-1}$ is the inverse matrix to $\{g_{ij}\}$ and it is usually written as $\{g^{ij}\}$. The process that associates a vector $v=(v^1,\dots,v^n)$ to a given co-vector $\lambda=(\lambda_1,\dots,\lambda_n)$ by making use of the map $\tilde g^{-1}$ is called "rising indices":

$$v^i=\sum_{j=1}^n g^{ij}\lambda_j,\quad i=1,\dots,n.$$

We conclude that the Riemannian metric $g$ defines a pairing

$$\langle\cdot\,,\cdot\rangle\colon T_qM\times T^*_qM\to\mathbb{R}$$

by

$$\langle v, \lambda\rangle = g(v, \tilde{g}^{-1}(\lambda)) = g^{-1}(\tilde{g}(v), \lambda), \quad \lambda \in T_q^*M, \quad v \in T_qM.$$

Having the co-metric and a chosen coordinate chart, we define the Riemannian Hamiltonian function $H\colon T^*M \to \mathbb{R}$ by

$$H(q,\lambda) = \frac{1}{2}g^{-1}(\lambda_q, \lambda_q) = \frac{1}{2}\sum_{i,j=1}^{n} g^{ij}\lambda_i\lambda_j.$$

A solution of the Hamiltonian equations

$$\begin{aligned} \dot{x}^i(s) &= \frac{\partial H(q(s),\lambda(s))}{\partial \lambda_i}, \quad q(s) = (x^1(s), \dots, x^n(s)) \qquad (2.8)\\ \dot{\lambda}_i(s) &= -\frac{\partial H(q(s),\lambda(s))}{\partial x^i}, \quad \lambda(s) = (\lambda_1(s), \dots, \lambda_n(s)), \qquad i = 1, \dots, n, \end{aligned}$$

$s \in I$, is called the *bi-characteristic* curve. The projection of the bi-characteristic curve to the manifold $M$ is called *geodesic*. The vector field

$$\vec{H}(q,\lambda) = \Big(\frac{\partial H(q,\lambda)}{\partial \lambda_i}, -\frac{\partial H(q,\lambda)}{\partial x^i}\Big)$$

is called *Hamiltonian vector field*. The Hamiltonian function is constant along the bi-characteristic since

$$\begin{aligned} \frac{H(q(s),\lambda(s))}{ds} &= \sum_{i=1}^{n}\Big(\frac{\partial H(q,\lambda)}{\partial x^i}\dot{x}^i(s) + \frac{\partial H(q,\lambda)}{\partial \lambda_i}\dot{\lambda}_i(s)\Big)\\ &= \sum_{i=1}^{n}\big(-\dot{\lambda}_i\dot{x}^i(s) + \dot{x}^i\dot{\lambda}_i(s)\big) = 0. \end{aligned}$$

If a geodesic is parametrized by the arc length, then $H = 1/2$. Remark that the notions of a local length minimiser and of a geodesic as the projection of a bi-characteristic to the manifold coincide in the Riemannian geometry, see, for instance [7].

Denote by $\gamma_{q,v}$ a geodesic starting from $q \in M$ with the initial velocity $v \in T_qM$. The notion of a geodesic leads to the construction of a map associating to vectors from $T_qM$ points in a neighborhood of $q \in M$. The domain of definition for this map is

$$\mathcal{D}(q) = \{v \in T_qM \mid \ \exists \text{ a geodesic } \ \gamma_{q,v}\colon [0,1] \to M, \ \ \gamma(0) = q, \ \ \dot{\gamma}(0) = v\}. \qquad (2.9)$$

**Definition 12.** *The Riemannian exponential map* $\exp_q\colon \mathcal{D}(q) \to M$ *is defined by*

$$\exp_q(v) = \gamma_{q,v}(1) \quad \textit{for all} \quad v \in \mathcal{D}(q).$$

Actually the Riemannian exponential map is the composition of the following maps.

$$\begin{tikzcd} T_qM \arrow[r, "\iota"] \arrow[rrrr, bend right, "\exp"'] & TM \arrow[r, "\tilde{g}"] & T^*M \arrow[r, "\Phi"] & T^*M \arrow[r, "\mathrm{pr}^*_M"] & M. \end{tikzcd} \qquad (2.10)$$

Here we denote by $\iota$ the inclusion of the tangent space $T_qM$ into the tangent bundle, by $\tilde{g}$ the association of the tangent and co-tangent bundles, by using the metric $g$, by $\Phi$ the flow produced by the Hamiltonian vector field on the co-tangent bundle, see Definition 44, and by $\mathrm{pr}^*_M$ the canonical projection to the base manifold $M$. The concrete choice of the initial velocity $v$ at $q \in M$ gives the value of the dual momentum $\lambda_q \in T^*_qM$.

In the following proposition we collect some basic properties of the Riemannian exponential map.

**Proposition 1.** [23, 117] *Let $v \in \mathcal{D}(q)$ be as defined in (2.9). Then*

1. *the exponential map $\exp_q$ carries lines through the origin of $T_qM$ to geodesics on $M$ through $q$ in the following sense*
$$\exp_q(tv) = \gamma_{q,tv}(1) = \gamma_{q,v}(t), \qquad t \in [0,1];$$
2. *for each $q \in M$, there is a neighborhood $V$ of the origin in $T_qM$, such that the exponential map $\exp_q\colon V \to U$ is a diffeomorphism onto a neighborhood $U$ of $q \in M$;*
3. *if $U$ is a normal neighborhood of $q \in M$ ($U$ is the diffeomorphic image of a starlike neighborhood of the origin in $T_qM$), then for each point $x \in U$ there is a unique geodesic $\gamma_{q,v}\colon [0,1] \to U$ joining $q$ and $x$ in $U$ and $\dot{\gamma}_{q,v}(0) = v = \exp_q^{-1}(x)$.*

Exercises.

1. Show that equations (2.7) and (2.8) are equivalent if we introduce the co-vectors (called *momenta* in physics) $\lambda$ and the Christoffel symbols for the Levi-Civita connection by
$$\lambda_i = \sum_j^n g_{ij}\dot{x}^j, \qquad \Gamma^k_{ij} = \frac{1}{2}\sum_{m=1}^n g^{km}\Big(\frac{\partial g_{jm}}{\partial x^i} + \frac{\partial g_{im}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^m}\Big).$$

2 Suppose that a coordinate chart is chosen and $X_1(q), \dots, X_n(q)$ is an orthonormal basis of $T_qM$. If the collection $X_1(x), \dots, X_n(x)$ is smooth and orthonormal in a neighborhood $U$ of $q$, then the family of vector fields $X_1, \dots, X_n$ is called an *orthonormal frame* in $U$. Assume that an orthonormal frame is given. Show that the Hamiltonian function can be written as
$$H(q,\lambda) = \frac{1}{2}\sum_{i=1}^n \langle X_i(q), \lambda_q\rangle^2. \tag{2.11}$$

Hint. Write co-vectors in the form $\lambda = \sum_{i=1}^n \lambda_i\omega^i$, where $\{\omega^i\}_{i=1}^n$ is the dual basis to $(X_1, \dots, X_n)$: $\langle X_j, \omega^i\rangle = \delta_{ij}$.

3. Calculate the exponential map $\exp_q\colon T_q\mathbb{R}^n \to \mathbb{R}^n$, where $\mathbb{R}^n$ is considered as a Riemannian manifold with the Euclidean metric. Let the Euclidean metric be also defined on $T_q\mathbb{R}^n$. Show that the exponential map is an isometry.

**2.4.2. Geodesics on sub-Riemannian manifolds.** Let $(M, D, g_D)$ be a sub-Riemannian manifold. Let us assume that we are interested in finding the local minimizer of the length functional (2.3) over all almost everywhere differentiable horizontal curves, or in other words, we look for a curve that locally realizes the Carnot-Carathéodory distance. We need to define an analogue of the Levi-Civita connection, but there is no metric defined on the entire tangent bundle. We will not enter this question deeply, since it requires some amount of knowledge of differential geometry, see, for instance [37, 53]. Instead, we adapt the Hamiltonian approach, since it is more suitable for physical applications. We also distinguish length minimizers, curves realizing $d_{c-c}$-distance, and geodesics, which are projections of bi-characteristic curves of the Hamiltonian system onto the underlying manifold.

To use the Hamiltonian approach we still have to overcome the absence of a metric defined on the entire tangent bundle that was used for definition of the dual to $TM$. Therefore, we assume that we are given a dual $T^*M$ (as a set of all continuous linear functionals) and a pairing $\langle \cdot\,, \cdot\rangle\colon T_qM \times T^*_qM \to \mathbb{R}$ that is the evaluation of a functional over a vector.

**Definition 13.** *We define a linear map* $\tilde{g}^D\colon T^*_qM \to T_qM$ *by the following two conditions*

1. *the image of* $T^*_qM$ *is the linear space* $D_q \subset T_qM$*;*
2. *for* $\lambda_q \in T^*_qM$ *the image* $\tilde{g}^D(\lambda_q)$ *is a vector* $X_q \in D_q$*, such that* $\langle Y_q, \lambda_q\rangle = g_D(Y_q, X_q)$ *for all* $Y_q \in D_q$*.*

The map $\tilde{g}^D$ is an analogue of the map $\tilde{g}^{-1}$ in the Riemannian geometry. The map $\tilde{g}^D$ defines the *co-metric* $g^D\colon T^*_qM \times T^*_qM \to \mathbb{R}$ by the following rule

$$g^D(\xi, \lambda) = \langle \tilde{g}^D(\lambda), \xi\rangle = g_D(\tilde{g}^D(\xi), \tilde{g}^D(\lambda)).$$

We still can write the matrix $g^{ij}$ for the co-metric $g^D$ in local coordinates, but we have no analogue for $g_{ij}$ since the matrix $g^{ij}$ is not invertible in this case.

Let us introduce the notation $D^\perp_q = \ker(\tilde{g}^D)$, $q \in M$ for the kernel of the linear map $\tilde{g}^D$. The elements of the smooth sub-bundle $D^\perp \subset T^*M$ are called annihilators of the distribution $D$ since $\langle Y_q, \xi_q\rangle = g_D\big(Y_q, \tilde{g}^D(\xi_q)\big) = g_D(Y_q, \overrightarrow{0}) = 0$ for $Y_q \in D_q$, $\xi_q \in D^\perp_q$. Then also

$$g^D(\lambda_q, \xi_q) = \langle \tilde{g}^D(\xi_q), \lambda_q\rangle = \langle \overrightarrow{0}, \lambda_q\rangle = 0, \quad \forall\ \ \xi_q \in D^\perp_q, \quad \text{and}\ \forall\ \ \lambda_q \in T^*_qM.$$

As in the Riemannian case, having a co-metric, one can define the Hamiltonian function $H_{sR}\colon T^*M \to \mathbb{R}$ by

$$H_{sR}(q, \lambda) = \frac{1}{2} g^D(\lambda_q, \lambda_q) = \frac{1}{2}\sum_{ij=1}^{n} g^{ij}\lambda_i\lambda_j, \quad q \in M. \tag{2.12}$$

We call $H_{sR}$ the sub-Riemannian Hamiltonian function. Consider again the Hamiltonian equations (2.8). The first equation written in the form

$$\dot{x}^i(s) = \sum_{j=1}^{n} g^{ij}\lambda_j, \quad i = 1,\dots,n, \quad \text{or} \quad \dot{x}(t) = \tilde{g}^D(\lambda) \tag{2.13}$$

says that the velocity of the solution to (2.8) will be a horizontal vector field by Definition 13.

Since the co-metric in sub-Riemannian case is not strictly positive definite, it can happened that the sub-Riemannian Hamiltonian function vanishes. This leads to two different types of geodesics: normal and abnormal. Recall that the Hamiltonian function is constant along any bi-characteristic. If this constant is zero, then the projection to $M$ is called an abnormal geodesic. If the Hamiltonian function is not zero along the bi-characteristic, then the geodesic is called normal. If $X_1,\dots,X_k$ is an orthonormal frame of the distribution $D$, then the abnormal bi-characteristic is a solution of the Hamiltonian system for $k$ Hamiltonian functions

$$H_i(q,\lambda) = \langle \lambda_q, X_i(q)\rangle = 0, \qquad \lambda_q \in D_q^{\perp} \setminus \{0\}, \quad i = 1,\dots,k.$$

To find normal bi-characteristics we need to work with the Hamiltonian function

$$H(q,\lambda) = \sum_{i=1}^{k} \langle \lambda_q, X_i(q)\rangle^2.$$

We will mostly work with normal geodesics. The reader can find a lot of useful information about abnormal geodesics in [97, 110]. Here we only want to present a short description of $D^{\perp}$ and the cases when the abnormal geodesics are trivial.

**Proposition 2.** [97] *Let $D$ be a smooth distribution of rank $k$ on an $n$-dimensional manifold $M$. Then $D^{\perp}$ is a smooth $(2n-k)$-dimensional sub-bundle of $T^*M$. Locally, it can be described as a set of $(q,\lambda) \in T^*M$ such that $H_i(q,\lambda) = \langle \lambda_q, X_i(q)\rangle = 0$, $i = 1,\dots,k$, where $\{X_i\}_{i=1}^k$ is a local basis for $D$.*

We remark that the sub-bundle $D \subset TM$ defines the set of annihilators $D^{\perp} \subset T^*M$. The converse is also true. Given a smooth sub-bundle $D^{\perp} \subset T^*M$, the distribution $D \subset TM$ is defined by

$$D_q = \{v \in T_qM \ \mid \ \langle v, \lambda_q\rangle = 0, \ \ \text{for all} \ \ \lambda_q \in D_q^{\perp}\}, \quad q \in M.$$

**Theorem 2.3.** [97, 127] *Let $D$ be a smooth distribution on a smooth manifold $M$.*

1. *If $D = TM$, then there are no abnormal geodesics.*
2. *If $D$ is strongly bracket generating, but $D \neq TM$, then the abnormal geodesics are constant curves.*

*Proof.* For some additional information about symplectic manifolds check Subsection 8.2. Let $\Gamma\colon I \to T^*M \setminus \{0\}$ be an abnormal bi-characteristic curve for the distribution $D$ that we write as $\Gamma(t) = \big(\gamma(t),\lambda(t)\big)$. Then if $X$ is a smooth section of $D$, then

$$H_X(\Gamma(t)) = H_X\big(\gamma(t),\lambda(t)\big) = \big\langle \lambda(t), X\big(\gamma(t)\big)\big\rangle = 0.$$

If $D = TM$, then all possible Hamiltonians vanish, and since the pairing is non-degenerate, we get $\lambda(t) = 0$ for all $t \in I$. This contradicts the assumption that $\lambda(t) \in D^{\perp}_{\gamma(t)} \setminus \{0\}$.

Let us assume now that $D$ is strongly bracket generating and $\Gamma, \gamma$ are as above. It implies that $H_Y(\Gamma(t)) = 0$ for any $Y \in D$. Differentiating with respect to $t$ the latter equality, we obtain

$$\frac{dH_Y}{dt}(\Gamma(t)) = dH_Y(\dot{\Gamma}(t)) = 0. \tag{2.14}$$

Suppose that the bi-characteristic $\Gamma$ is the solution of the Hamiltonian system

$$\dot{\Gamma}(t) = \overrightarrow{H_X}(\Gamma(t)), \qquad t \in I,$$

for some smooth section $X$ of $D$. Then for any $Y \in D$ we get

$$H_{[X,Y]}(\Gamma(t)) = \{H_X, H_Y\}(\Gamma(t)) = \Omega\big(\overrightarrow{H_X}(\Gamma(t)), \overrightarrow{H_Y}(\Gamma(t))\big) = dH_Y(\dot{\Gamma}(t)) = 0$$

by Definition 49 and (2.14). We conclude that $\lambda(t)$ annihilates the tangent space

$$D_{\gamma(t)} + [X, D]_{\gamma(t)}, \quad t \in I$$

along $\gamma$. Since $\lambda(t) \neq 0$, the condition $H_X(\Gamma(t)) = \langle \lambda(t), X(\gamma(t)) \rangle = 0$ implies that $X(\gamma(t)) = 0$, i.e., $\dot{\gamma}(t) = 0$ by Corollary 9. We conclude that the curve $\gamma$ is constant. ☐

The relation between the length minimizing curves and the geodesics (projections of bi-characteristics of the Hamiltonian system) in sub-Riemannian geometry is expressed in the following theorem

**Theorem 2.4.** *Let $(M, D, g_D)$ be a sub-Riemannian manifold.*

1. *If $\gamma\colon [a, b] \to M$ is a length minimizer, parametrized by the arc length, then $\gamma$ is geodesic (normal or abnormal)* [97].
2. *Every normal geodesic is a local length minimizer* [15, 95].
3. *There are abnormal geodesics that are local length minimizers* [109, 110].
4. *There are abnormal geodesics that are not local length minimizers* [97].
5. *If $(M, d_{c-c})$ is a complete metric space for a Carnot-Carathéodory metric $d_{c-c}$, then any two points can be joined by a minimizing geodesic. In particular, this is true for compact $M$*[13, Theorem 2.7, p. 19 and Remark 2, p. 20].
6. *On a sub-Riemannian manifold with a bracket generating distribution of step 2, any length minimizing curve is $C^{\infty}$-smooth, or in other words there are no strictly abnormal minimizing geodesics in this case* [112, Theorem 4].
7. *For a generic (in the Whitney $C^{\infty}$ topology) bracket generating distribution of rank greater than or equal to three, there do not exist nontrivial minimizing singular curves* [30].

At the end of this section we note that the sub-Riemannian exponential map is produced in the same form as in (2.10), where the initial velocity vector is horizontal. It is reflected in the following scheme

$$\begin{tikzcd} D_q \arrow[r, "\iota"] \arrow[rrrr, bend right, "\exp"'] & D \arrow[r, "j"] & T^*M \arrow[r, "\Phi"] & T^*M \arrow[r, "\mathrm{pr}^*_M"] & M, \end{tikzcd} \tag{2.15}$$

where we have to change the metric dependent identification $\tilde{g}$ of $TM$ with $T^*M$ to any other map $j$ giving this identification. Unfortunately, not all good properties of the Riemannian exponential map are inherited. For instance, the sub-Riemannian exponential map is never a local diffeomorphism, since the map $j$ is not invertible for any $q \in M$.

EXERCISE.

1. Let $(M, D, g_D)$ be a sub-Riemannian manifold. Show that the co-metric $g^D$ is non-negative definite, symmetric, and smoothly varying with respect to the point $q \in M$.
2. Let $M = \mathbb{R}^3$ with coordinates $q = (x, y, z)$. Find a basis of the distribution $D = \ker\{\omega = x^2 dy - (1-x)dz\}$. (Check if the basis $X = \frac{\partial}{\partial x}$, $Y = (1-x)\frac{\partial}{\partial y} + x^2\frac{\partial}{\partial z}$ works.) Is $D$ bracket generating? Regular? Find the matrix of the sub-Riemannian metric $g_D$ making vector fields $X, Y$ orthonormal. Find the sub-Riemannian Hamiltonian function $H$ generated by $g_D$ and the corresponding Hamiltonian system. It was shown in [97] that the curve $\gamma\colon [a, b] \to \mathbb{R}^3$, $\gamma(t) = (0, t, 0)$ is a length minimizer for the Carnot-Carathéodory distance if $(b - a)$ is small enough. Show that the curve $\gamma$ is not a bi-charcteristic for the sub-Riemannian Hamiltonian function $H$. Conclude that the curve $\gamma$ is a length minimizer but not a normal geodesic.

## 3. Carnot groups

Let us consider a special example of smooth manifolds, where the sub-Riemannian structure appears naturally.

### 3.1. Short introduction to Lie groups

It is recommended for the reader who is not familiar with Lie group theory to start from Subsection 8.3 of the Appendix A. A Lie group is an object that nicely combine algebraic, geometric, and analytic properties. Namely, a Lie group $\mathbb{G}$ is a pair $(M, \rho)$, where

1. $M$ is a $C^\infty$-smooth manifold modeled on some (complete locally convex) vector space,
2. the map $\rho\colon M \times M \to M$ satisfies the axioms of the group product,
3. the map $\rho$ is compatible with the smooth manifold structure in the sense that the map $\rho\colon M \times M \to M$ is $C^\infty$-smooth as a map between the smooth manifold $M \times M$ and another smooth manifold $M$.

As usual in mathematics, we will write only $\mathbb{G}$ instead of $(M, \rho)$ to denote the group and the underlying manifold $M$.

Recall, that a Lie algebra is a pair $(V, [\cdot\,,\cdot])$, where $V = (V, +)$ is a vector space over the fields $\mathbb{R}$ or $\mathbb{C}$ and $[\cdot\,,\cdot]$ is the Lie product introduced in Definition 51, Appendix A.There is a close relation between Lie groups and Lie algebras.

From a Lie group to its Lie algebra. To define the Lie algebra $\mathfrak{g}$ of a Lie group $\mathbb{G}$ we consider special vector fields on $\mathbb{G}$. To describe this class of vector fields we introduce the action of the group on itself. We call the mappings

$$l_\tau(q) := \rho(\tau, q) = \tau q, \quad \tau \in \mathbb{G} \ \text{ fixed}, \quad q \in \mathbb{G} \ \text{ is arbitrary},$$

the *left action of* $\mathbb{G}$ *on itself* and

$$r_\tau(q) := \rho(q, \tau) = q\tau, \quad \tau \in \mathbb{G} \ \text{ fixed}, \quad q \in \mathbb{G} \ \text{ is arbitrary},$$

the *right action of* $\mathbb{G}$ *on itself.* Since the group multiplication and the inversion are smooth, the maps $l_\tau, r_\tau \colon \mathbb{G} \to \mathbb{G}$ are smooth diffeomorphisms of $\mathbb{G}$. Their differentials $d_q l_\tau \colon T_q\mathbb{G} \to T_{l_\tau(q)}\mathbb{G}$ and $d_q r_\tau \colon T_q\mathbb{G} \to T_{r_\tau(q)}\mathbb{G}$ are linear maps of the respective tangent spaces.

**Definition 14.** *A vector field* $X$ *on* $\mathbb{G}$ *satisfying*

$$d_q l_\tau(X(q)) = X\big(l_\tau(q)\big) = X(\tau q) \quad \Big(d_q r_\tau(X(q)) = X\big(r_\tau(q)\big) = X(q\tau)\Big),$$

*for all* $\tau, q \in \mathbb{G}$ *is called left- (right-) invariant vector field.*

The set of left invariant vector fields considered as a vector space over the field $\mathbb{R}$ with the Lie product defined by the commutator of vector fields (2.1) forms a real Lie algebra $\mathcal{L}$. Of course, one needs to verify that the commutator of left invariant vector fields is a left invariant vector field. Since any left invariant vector field is defined by its value at the identity of the group $e \in \mathbb{G}$, there is an isomorphism $\iota$ between the vector space $T_e\mathbb{G}$ and $\mathcal{L}$ defined by

$$\mathcal{L} \ni X \mapsto X(e) \in T_e\mathbb{G}, \qquad T_e\mathbb{G} \ni v \mapsto dl(v) \in \mathcal{L}.$$

This isomorphism $\iota$ can be extended to an isomorphism of Lie algebras if we define Lie brackets in $T_e\mathbb{G}$ as

$$[X(e), Y(e)] := [X, Y](e).$$

The Lie algebra $(T_e\mathbb{G}, [\cdot\,,\cdot])$ is denoted usually by $\mathfrak{g}$ and is called the Lie algebra of the Lie group $\mathbb{G}$. The Lie algebra $\mathcal{R}$ of right invariant vector fields is isomorphic to $\mathfrak{g}$ if we set

$$\mathcal{R} \ni [X, Y] \leftrightarrow \ -[X, Y](e) \in \mathfrak{g}.$$

The dual space to the space of left invariant vector fields consists of left invariant one-forms and they satisfy the Maurer-Cartan equations, see [131].

The next question is to find a map between a given Lie group $\mathbb{G}$ and its Lie algebra $\mathfrak{g}$. The answer is given in terms of the exponential map $\exp\colon \mathfrak{g} \to \mathbb{G}$. There are essentially two ways to introduce the exponential map. The first one uses the property that any homomorphism of Lie algebras can be lifted to a homomorphism of the groups [87, 131]. The second one uses properties of solutions of ordinary differential equations [42].

THE FIRST WAY. Let $(\mathbb{R},+)$ be the additive group of real numbers and $\mathfrak{r}$ be the corresponding Lie algebra with generator $\frac{d}{dr}$. Let $\mathbb{G}$ be a Lie group, $\mathfrak{g}$ be its Lie algebra, and $X \in \mathfrak{g}$ be an arbitrary element. Then the map $h\colon \mathfrak{r} \to \mathfrak{g}$

$$\mathfrak{r} \ni t\frac{d}{dr} \overset{h}{\mapsto} tX \in \mathfrak{g}, \quad t \in \mathbb{R},$$

is a homomorphism from the Lie algebra $\mathfrak{r}$ into the Lie algebra $\mathfrak{g}$. Theorems of Lie group theory [87, 131] ensures that there is a unique Lie group homomorphism $c_X$, such that

$$c_X : \mathbb{R} \to \mathbb{G}, \quad \text{and} \quad d\,c_X = h, \quad \text{or} \quad dc_X(t\frac{d}{dr}) = tX.$$

In other words, the curve $c_X : \mathbb{R} \to \mathbb{G}$ is a one-parametric subgroup of $\mathbb{G}$ and it is such that $c_X(0) = e$ and $\dot{c}_X(0) = X$.

THE SECOND WAY. Let $\mathbb{G}$ be a Lie group, $\mathfrak{g}$ be its Lie algebra, and let $X \in \mathfrak{g}$ be an arbitrary left invariant vector field. Then the theory of ordinary differential equations guaranties that the solution of the Cauchy problem

$$\begin{cases} \frac{dc_X(t)}{dt} = X(c_X(t)) \\ c_X(0) = e \end{cases}$$

is unique, possesses the properties of one parameter subgroup of $\mathbb{G}$, and $\dot{c}_X(0) = X(e)$ [42].

**Definition 15.** *The map* $\mathfrak{g} \ni X \to c_X(1) \in \mathbb{G}$ *is called the group exponential map and denoted by* exp. *Thus*

$$\begin{array}{cccc} \exp & : & \mathfrak{g} & \to & \mathbb{G} \\ & & X & \mapsto & c_X(1). \end{array}$$

We will call the curve $c_X(t)$, $t \in \mathbb{R}$, the *exponential curve* and it is customary to use also the notation $\exp(tX)$ instead of $c_X(t)$. The main properties of the exponential map are listed in the Appendix A, Subsection 8.3. We write in these notations

$$T_e\mathbb{G} \supset \frac{d}{dt}\Big|_{t=0} \exp(tX) = \dot{c}_X(0) = \widetilde{X}(e) \cong X \in \mathfrak{g}. \tag{3.1}$$

Let us assume now that the Lie algebra $\mathfrak{g}$ of a Lie group $\mathbb{G}$ is endowed with an inner product $(\cdot\,,\cdot)$. Then, by making use of left translations we can define a metric $g$ on the group. Namely, let $v_q, w_q \in T_q\mathbb{G}$, then $d_q l_{q^{-1}}(v_q), d_q l_{q^{-1}}(w_q) \in T_e\mathbb{G}$. We define

$$g(v_q, w_q) := \big(d_q l_{q^{-1}}(v_q), d_q l_{q^{-1}}(w_q)\big) \quad \text{for any} \quad q \in \mathbb{G}. \tag{3.2}$$

Using right translations we also can define a metric.

Conversely, if there is a Riemannian metric $g$ defined on a Lie group $\mathbb{G}$ considered as a smooth manifold, then it is compatible with the Lie structure if it is invariant under the action of the group on itself.

**Definition 16.** *A Riemannian metric $g$ on $\mathbb{G}$ is called left invariant (right invariant), if for any $v_q, w_q \in T_q\mathbb{G}$ the following holds:*

$$g(v_q, w_q) = g(d_q l_\tau(v_q), d_q l_\tau(w_q)) = g(v_{\tau q}, w_{\tau q}),$$
$$\Big(g(v_q, w_q) = g(d_q r_\tau(v_q), d_q r_\tau(w_q)) = g(v_{q\tau}, w_{q\tau})\Big).$$

EXERCISES

1. Show that the following pairs are Lie groups.
   a. $(\mathbb{R}, +)$.
   b. $(\mathbb{R}, \cdot)$, where "$\cdot$" is the usual product of real numbers.
   c. $(S^1, \cdot)$, where $S^1$ is the set of complex numbers of absolute value 1 and "$\cdot$" is the usual product of complex numbers. The group $(S^1, \cdot)$ is also denoted by $U(1)$ and it is called unitary one dimensional group.
   d. $(M, \cdot)$, where $M$ is the set of $(3 \times 3)$ upper triangular real matrices
   $$\begin{pmatrix} 1 & x & t \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}, \qquad x, y, t \in \mathbb{R}$$
   and "$\cdot$" stands for the usual matrix product.

   A group is called compact, if the underlying manifold is compact as a topological space. Which of the above mentioned groups are compact?
2. Show that if $X, Y$ are left invariant (right invariant) vector fields on $\mathcal{G}$, then the commutator $[X, Y]$ is also a left invariant (right invariant) vector field.
3. Find the Lie algebras corresponding to the Lie groups mentioned in the first exercise. Describe the left invariant vector fields.
4. Show that the metric from (3.2) is a left invariant metric on the group.

## 3.2. Heisenberg group

We start from the simplest example of a sub-Riemannian manifold that is called Heisenberg group.

**3.2.1. The Heisenberg sub-Riemannian manifold.** Consider the smooth manifold $\mathbb{R}^3$ with coordinates $q = (x, y, t)$. Then $T_q\mathbb{R}^3 = \operatorname{span}\{\partial_x, \partial_y, \partial_t\}$ and $T_q^*\mathbb{R}^3 = \operatorname{span}\{dx, dy, dt\}$. We define the smooth 2-dimensional distribution $D$ as the span of two vector fields

$$X = \partial_x - \frac{1}{2}y\partial_t, \quad Y = \partial_y + \frac{1}{2}x\partial_t. \tag{3.3}$$

See Figure 3.1. Let us find a Riemannian metric $g$ in coordinates $(x, y, t)$ making $X, Y$ and $T = [X, Y] = \partial_t$ orthonormal. So we have $g(X, X) = g(X, Y) = g(T, T) = 1$ and other values vanish. We express the basis $(\partial_x, \partial_y, \partial_t)$ in the form

$$\partial_x = X + \frac{1}{2}yT, \quad \partial_y = Y - \frac{1}{2}xT, \quad \partial_t = T.$$

Then by making use of the bi-linearity of $g$ we get

$$g_{11} = g(\partial_x, \partial_x) = 1 + \frac{y^2}{4}, \quad g_{12} = g(\partial_x, \partial_y) = -\frac{xy}{4}, \quad g_{13} = g(\partial_x, \partial_t) = \frac{y}{2} \quad \dots .$$

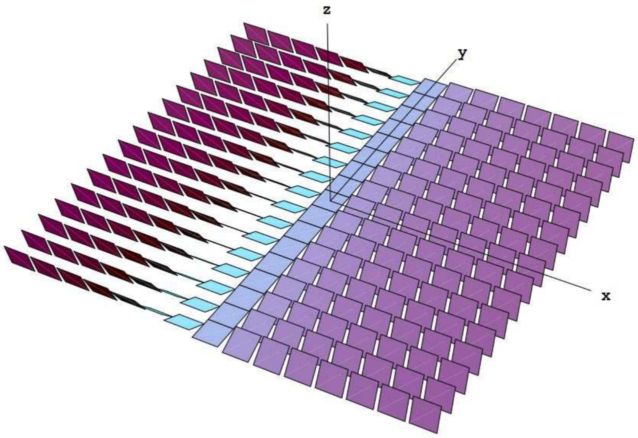

Figure 3.1. The Heisenberg distribution.

The matrix $\{g_{ij}\}$ takes the form

$$g_{ij}=\begin{pmatrix} 1+\frac{y^2}{4} & -\frac{xy}{4} & \frac{y}{2}\\ -\frac{xy}{4} & 1+\frac{x^2}{4} & -\frac{x}{2}\\ \frac{y}{2} & -\frac{x}{2} & 1\end{pmatrix}. \tag{3.4}$$

Notice that $\det g = 1$. It implies that the volume form in $(\mathbb{R}^3, g)$ is given by the standard Lebesgue measure: $dx\wedge dy\wedge dt$.

The distribution $D$ is bracket generating of step 2 since $[X,Y]=\partial_t := T$ and $T_q\mathbb{R}^3=\mathrm{span}\{X,Y,T\}$. Moreover, the distribution is strongly bracket generating and we will be interested in normal geodesics only (by Theorem 2.3). The dual basis to $X,Y,T$ is

$$dx,\quad dy,\quad \omega=dt-\frac{1}{2}xdy+\frac{1}{2}ydx.$$

(Verify it!) The form $\omega$ is the annihilator of the distribution $D$ and $D^{\perp}=\mathrm{span}\{\omega\}$, therefore, we can also define the distribution $D$ as $D=\ker(\omega)=\{v=(x,y,z)\in\mathbb{R}^3 \mid \omega(v)=0\}$. Define the sub-Riemannian metric $g_D$ as the restriction of the metric $g$ on the planes $D_q$ for all $q\in\mathbb{R}^3$. Then $(\mathbb{R}^3,D,g_D)$ is the Heisenberg sub-Riemannian manifold.

To find the normal geodesics on the Heisenberg sub-Riemannian manifold we write $\lambda=\xi dx+\eta dy+\theta dt$ for any co-vector $\lambda$. Since the basis $X,Y$ is orthonormal, the Hamiltonian function is

$$H(q,\lambda)=\frac{1}{2}\Big(\langle\lambda,X\rangle^2+(\langle\lambda,Y\rangle^2\Big)=\frac{1}{2}\Big((\xi-\frac{1}{2}\theta y)^2+(\eta+\frac{1}{2}\theta x)^2\Big).$$

The Hamiltonian system and the initial conditions are

$$\begin{cases} \dot{x} = \xi - \frac{1}{2}\theta y \\ \dot{y} = \eta + \frac{1}{2}\theta x \\ \dot{t} = \frac{1}{2}(\eta x - \xi y) + \frac{1}{4}\theta(x^2 + y^2) \\ \dot{\xi} = -\frac{1}{2}\eta\theta - \frac{1}{4}\theta^2 x \\ \dot{\eta} = -\frac{1}{2}\xi\theta - \frac{1}{4}\theta^2 y \\ \dot{\theta} = 0, \end{cases} \qquad \begin{cases} x(0) = y(0) = t(0) = 0, \\ \xi(0) = \xi_0,\ \eta(0) = \eta_0,\ \theta(0) = \theta_0. \end{cases} \tag{3.5}$$

We need projections of the bi-characteristics of $H$ onto $\mathbb{R}^3$, therefore, we try to reduce the Hamiltonian system to a system containing only the variables $(x, y, t)$. We differentiate the first two equations and replace $\dot{\xi}$ and $\dot{\eta}$ from the fourth and fifth equations. It gives

$$\begin{cases} \ddot{x} &= -\theta_0 \dot{y} \\ \ddot{y} &= \theta_0 \dot{x}, \end{cases} \qquad or \qquad \begin{pmatrix} \ddot{x} \\ \ddot{y} \end{pmatrix} = \theta_0 \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix}. \tag{3.6}$$

Then multiplying the first equation by $y$ and the second by $x$ we notice that the third equation is equivalent to the condition

$$\dot{t}(s) = \frac{1}{2}(x(s)\dot{y}(s) - y(s)\dot{x}(s)). \tag{3.7}$$

The solution $\gamma(s) = \big(x(s), y(s), t(s)\big)$ is

$$\begin{array}{lll} x(s) & = \frac{\xi_0}{|\theta_0|}\sin(|\theta_0|s) - \frac{\eta_0}{|\theta_0|}(\cos(|\theta_0|s) - 1), & \\ y(s) & = -\frac{\xi_0}{|\theta_0|}(\cos(|\theta_0|s) - 1) - \frac{\eta_0}{|\theta_0|}\sin(|\theta_0|s), & \text{if} \quad \theta_0 \neq 0 \\ t(s) & = \frac{(\xi_0^2 + \eta_0^2)}{2|\theta_0|^2}(|\theta_0|s - \sin(|\theta_0|s)), & \end{array} \tag{3.8}$$

and

$$x(s) = \xi_0 s, \quad y(s) = \eta_0 s, \quad t(s) = 0, \qquad \text{if} \quad \theta_0 = 0. \tag{3.9}$$

The graph is in the Figure 3.2.

Let us look at the condition (3.7) of the Hamiltonian equation. To understand better this equation let us calculate the velocity vector of any curve in the basis $X, Y, T$. Let $c(s) = (x(s), y(s), t(s))$, $s \in I$, be a curve, then

$$\begin{aligned} \dot{c}(s) &= \dot{x}(s)\partial_x + \dot{y}(s)\partial_y + \dot{t}(s)\partial_t = \dot{x}(s)\big(\partial_x - \frac{1}{2}y(s)\partial_t\big) + \dot{y}(s)\big(\partial_y + \frac{1}{2}x(s)\partial_t\big) \\ &+ \big(\dot{t}(s) + \frac{1}{2}\dot{x}(s)y(s) - \frac{1}{2}\dot{y}(s)x(s)\big)\partial_t \\ &= \dot{x}(s)X(c(s)) + \dot{y}(s)Y(c(s)) + \big(\dot{t}(s) + \frac{1}{2}\dot{x}(s)y(s) - \frac{1}{2}\dot{y}(s)x(s)\big)T(c(s)). \end{aligned}$$

To be a horizontal curve, the coordinate of $\dot{c}$ in front of the vector field $T$ have to vanish for all $s \in I$, that leads to the equation (3.7). Thus, the third equation of the Hamiltonian system (3.5) is just the horizontality condition, and it is not surprising due to the general fact (2.13).

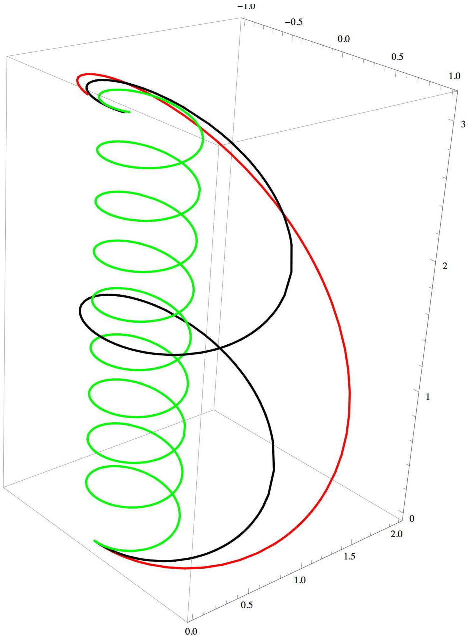

FIGURE 3.2. Geodesics on the Heisenberg group

Observe the following relations between the initial values for the Hamiltonian system and the initial velocity vector: $\xi_0 = \dot{x}(0)$, $\eta_0 = \dot{y}(0)$. The values $\dot{x}(0)$ and $\dot{y}(0)$ and the initial point completely define the initial velocity $\dot{t}(0)$. So, fixing the initial velocity we still have countably many geodesics starting from the origin, that are parametrized by the parameter $\theta_0$ and connecting a fixed point $(0,0,t)$. To see this we parametrize all geodesics, starting from the origin with a fixed initial velocity $(\xi_0,\eta_0)$, in the unit interval $[0,1]$. Denote by $x_1 = x(1)$, $y_1 = y(1)$ the final point of these geodesics. We assume $x_1^2+y_1^2 = 0$, that means that geodesics connect the origin with a point $(0,0,t_1)$ on $t$-axis, as it is shown in Figure 3.2. Then,

$$0 = x_1^2 + y_1^2 = \frac{\sin^2(|\theta_0|)}{|\theta_0|^2}(\xi_0^2 + \eta_0^2) \quad \Longrightarrow \quad |\theta_0| = 2\pi n, \quad n \in \mathbb{N}$$

by solutions (3.8). The corresponding value of $t_1$ is $\frac{1}{4}(\xi_0^2 + \eta_0^2)$. The complete description of geodesics on the Heisenberg sub-Riemannian manifold can be found in [20, 77, 88].

The geodesics (3.8) (3.9) are local length minimizers by Theorem 2.3. The length of a sub-Riemannian geodesic $\gamma\colon I = [0,T] \to \mathbb{R}^3$ is

$$\text{length}(\gamma) = \int_0^T (\dot{x}(s)^2 + \dot{y}(s)^2)^{1/2} ds = (\xi_0^2 + \eta_0^2)^{1/2} T.$$

If there are several sub-Riemannian geodesics connecting the origin with a point $q \in \mathbb{R}^3$ and they are parametrized by the arc length, then the curve of the smallest length is the minimizing curve, realizing the Carnot-Carathéodory distance between the origin and $q \in \mathbb{R}^3$.

**3.2.2. Heisenberg sub-Riemannian manifold as a Lie group.** Let us consider the following non-commutative group on the smooth manifold $\mathbb{R}^3$. Define the product for $\tau = (x, y, t)$ and $q = (x_1, y_1, t_1)$ by

$$\tau q = (x, y, t)(x_1, y_1, t_1) = (x + x_1, y + y_1, t + t_1 + \frac{1}{2}(xy_1 - x_1 y)). \tag{3.10}$$

As a motivation for this law one can consider the product of $(4 \times 4)$ real matrices

$$\begin{pmatrix} 1 & x & y & t \\ 0 & 1 & 0 & \frac{y}{2} \\ 0 & 0 & 1 & -\frac{x}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & x_1 & y_1 & t_1 \\ 0 & 1 & 0 & \frac{y_1}{2} \\ 0 & 0 & 1 & -\frac{x_1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & x + x_1 & y + y_1 & t + t_1 + \frac{1}{2}(xy_1 - x_1 y) \\ 0 & 1 & 0 & \frac{y+y_1}{2} \\ 0 & 0 & 1 & -\frac{x+x_1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

that leads to formula (3.10). It is an easy exercise to verify that the product (3.10) satisfies the group axioms. The identity $e$ of the obtained group has coordinates $(0, 0, 0)$ with respect to this multiplication and the inverse element to $(x, y, t)$ is $(-x, -y, -t)$. The pair, consisting of the smooth manifold $\mathbb{R}^3$ and the introduced group law, is called the *Heisenberg group* and is denoted by $\mathbb{H}^1$. This group law defines the left translation: $l_\tau(q) = \tau q$. The left translation $l_\tau$ by $\tau = (x, y, z)$ has the differential at $e = (0, 0, 0)$ written in coordinates $(x, y, t)$ as

$$d_e l_\tau = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{2}y & \frac{1}{2}x & 1 \end{pmatrix}.$$

The action of $d_e l_\tau$ on the basis $(\partial_x, \partial_y, \partial_t)$, that coincides with $(X, Y, T)$ at $e$, gives the basis $(X, Y, T)$ at $\tau$. We conclude that the basis $(X, Y, T)$ is just the basis of left invariant vector fields on the group $\mathbb{H}^1$. They form the famous Heisenberg algebra $\mathfrak{h}^1$ which is by definition a 3-dimensional Lie algebra with only one non-trivial commutator: $[X, Y] = T$ and all other commutators vanish. We use the identification of the Lie algebra of left-invariant vector fields with $T_e\mathbb{H}^1$. The exponential map is a global diffeomorphism [50] in this case. The coordinates on the group $\mathbb{H}^1$ are of the first kind and are given by

$$\mathbb{H}^1 \ni q = (x, y, t) = \exp(xX + yY + tT), \quad xX + yY + tT \in \mathfrak{h}^1.$$

The inverse map to the exponential restores the group multiplication law from the commutation relations of the Heisenberg algebra in the following way. Let $V = xX + yY + tT \in \mathfrak{h}^1$, $V_1 = x_1 X + y_1 Y + t_1 T \in \mathfrak{h}^1$ and $\tau = \exp(V)$, $q = \exp(V_1)$, then by the Baker-Campbell-Hausdorff formula (8.1) (BCH-formula for short) we obtain

$$\begin{aligned}\tau q &= \exp(V)\exp(V_1) = \exp(V + V_1 + \frac{1}{2}[V, V_1] + ...) \\ &= \exp\big((x + x_1)X + (y + y_1)Y + (t + t_1)T + \frac{1}{2}(xy_1 - x_1 y)T\big) \\ &= \big(x + x_1, y + y_1, t + t_1 + \frac{1}{2}(xy_1 - x_1 y)\big),\end{aligned}$$

that coincides with (3.10).

There is a norm $\|\cdot\|_{\mathbb{H}^1}$ on the group $\mathbb{H}^1$ which is a direct analogue of the Euclidean norm in $\mathbb{R}^3$. It is defined by

$$\|\tau\|_{\mathbb{H}^1} = \Big(\big(x^2 + y^2\big)^2 + t^2\Big)^{1/4}, \quad \tau = (x, y, t). \tag{3.11}$$

If we stretch the basic elements $X$ and $Y$ of the Heisenberg algebra by a number $s > 0$, then the bi-linearity of the commutator implies $[sX, sY] = s^2 T$. Making use of the BCH-formula we get the dilatation $\delta_s$ on the group:

$$\delta_s(\tau) = \delta_s(x, y, t) = (sx, sy, s^2 t). \tag{3.12}$$

This dilation, which is called the *homogeneous dilation* is compatible with the norm in the sense that the norm becomes a homogeneous of order one function:

$$\|\delta_s(\tau)\|_{\mathbb{H}^1} = \|(sx, sy, s^2 t)\|_{\mathbb{H}^1} = s\|\tau\|_{\mathbb{H}^1}.$$

Compare this situation with the Euclidean norm and the usual dilation in $\mathbb{R}^3$! The Heisenberg distance function $d_{\mathbb{H}^1}$ is

$$d_{\mathbb{H}^1}(\tau, q) = \|\tau^{-1} q\|_{\mathbb{H}^1}.$$

By Exercise 2, the Heisenberg distance $d_{\mathbb{H}^1}$ and the Carnot-Carathéodory distance $d_{c-c}$ are equivalent.

**Example 5.** Let us show that the Heisenberg distance and the Euclidean distance $d_E$ are not Lipschitz equivalent, even locally in $\mathbb{R}^3$. Take two points $e = (0, 0, 0)$ and $q = (0, 0, t)$. Then

$$d_{\mathbb{H}^1}(e, q) = \sqrt{|t|}, \quad d_E(e, q) = |t|,$$

which shows non-equivalence of the distance functions. This also proves that the metric spaces $(\mathbb{R}^3, d_E)$ and $(\mathbb{R}^3, d_{\mathbb{H}^1})$ are not equivalent. But the topological spaces $(\mathbb{R}^3, \tau_E)$ and $(\mathbb{R}^3, \tau_{\mathbb{H}^1})$ are equivalent since any Heisenberg ball contains an Euclidean ball and vice versa. We also present the picture of balls in different metrics in Figure 3.4.

The metric with the matrix (3.4) is a left invariant metric on $\mathbb{H}^1$. The distribution $D = \operatorname{span}\{X, Y\}$, where $X, Y$ are defined in (3.3), itself can be called left

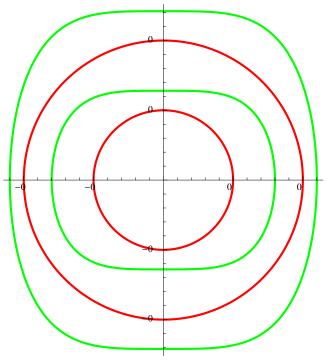

Figure 3.3. The Heisenberg and the Euclidean balls inside each other.

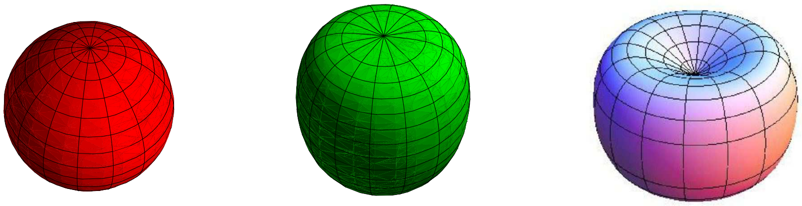

Figure 3.4. The Euclidean, Heisenberg, and Carnot-Carathéodory balls.

invariant since it is completely defined by $X(e) = \partial_x$, $Y(e) = \partial_y$, and $D_\tau = dl_\tau D_e$. The differential operator

$$\Delta_{sR} = X^2 + Y^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{1}{2}(x^2+y^2)\frac{\partial^2}{\partial t^2} - \Big(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\Big)\frac{\partial}{\partial t} \qquad (3.13)$$

is called sub-Laplacian. It is an analogue of the Laplace-Beltrami operator $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial t^2}$ in $\mathbb{R}^3$ with respect to the Euclidean metric. Observe that the homogeneous function $N(\tau) = \big((x^2+y^2)^2 + t^2\big)^{1/4}$ for $\tau = (x,y,t) \in \mathbb{H}^1$ is connected to the fundamental solution $\Gamma(\tau)$ to the sub-Laplacian (3.13) as follows $\Gamma(\tau) = \frac{c(Q)}{N(\tau)^{Q-2}}$. The constant $c(Q) < 0$ can be calculated explicitly and $Q = 4$ is the Hausdorff dimension of the metric space $(\mathbb{H}^1, d_{\mathbb{H}^1})$, see [49].

Let us present the formulas for the gradient on $\mathbb{H}^1$ in coordinates. In order to use formula (2.4), we calculate the inverse matrix to (3.4):

$$g^{ij} = \begin{pmatrix} 1 & 0 & -\frac{y}{2} \\ 0 & 1 & \frac{x}{2} \\ -\frac{y}{2} & \frac{x}{2} & 1+\frac{x^2+y^2}{4} \end{pmatrix}$$

Then

$$g^{ij}\begin{pmatrix} \partial_x f \\ \partial_y f \\ \partial_t f \end{pmatrix} = \begin{pmatrix} \partial_x f - \frac{y}{2}\partial_t f \\ \partial_y f + \frac{x}{2}\partial_t f \\ \frac{x\partial_y f - y\partial_x f}{2} + \Big(1+\frac{x^2+y^2}{4}\Big)\partial_t f \end{pmatrix} = \begin{pmatrix} Xf \\ Yf \\ \frac{x}{2}Yf - \frac{y}{2}Xf + Tf \end{pmatrix}.$$

Thus

$$\operatorname{grad} f = g^{ij}\begin{pmatrix} \partial_x f \\ \partial_y f \\ \partial_t f \end{pmatrix}\cdot\begin{pmatrix} \partial_x \\ \partial_y \\ \partial_t \end{pmatrix} = Xf\,X + Yf\,Y + Tf\,T.$$

The horizontal gradient “$\operatorname{grad}_D$” is the projection of “grad” onto $D = \operatorname{span}\{X, Y\}$ and it is written as $\operatorname{grad}_D f = (Xf, Yf)$ in the left invariant basis $X, Y$ of $D$.

**3.2.3. Heisenberg group and isoperimetric problem.** Let us recall the ansient story of Dido, or Elissa in Greek version, the founder and the first Queen of Carthage (in modern-day Tunisia). She was daughter of the king of Tyre and after the dangerous for her life intrigues of her brother Pygmalion she had to leave her land. Eventually Elissa and her followers arrived at the coast of North Africa where Elissa asked the local inhabitants for a small piece of land for a temporary refuge until she could continue her journey. She was allowed to have only as much land as could be encompassed by an oxhide. Elissa cut the oxhide into thin strips so that she had enough to encircle an entire nearby hill. According to this legend, Elissa was the first person who solved the isoperimetric problem of enclosing the maximum area within a boundary of a fixed length.

The dual problem is to find a minimal length curve enclosing the fixed area. Let us formulate this problem mathematically. Introduce the coordinates $(x, y)$ on the plane $\mathbb{R}^2$ and let $c(s) = (x(s), (y(s)))$, $s \in I$, be a closed curve in $\mathbb{R}^2$ that encloses a bounded domain $\Omega$. Then the area $A$ of $\Omega$ can be calculated as $\int_\Omega dA = \int_c dt$ by the Stokes theorem. Here the area form $dA = dx \wedge dy$ is the differential of the one form $dt = \frac{1}{2}(xdy - ydx)$. The variational problem with constraint is formulated as follows:

*Find a closed curve* $c\colon I \to \mathbb{R}^2$ *of minimal length* $\int_c \sqrt{\dot{x}^2(s) + \dot{y}^2(s)}ds$, *such that the area* $A = \int_c dt = \frac{1}{2}\int_c (xdy - ydx)$ *enclosed by this curve is fixed.*

Let us introduce the third coordinate $t$ that will reflect the change of the area swept by the curve $c(s) = (x(s), y(s))$, $s \in I$, i. e.,

$$\dot{t}(s) = \frac{1}{2}\big(x(s)\dot{y}(s) - y(s)\dot{x}(s)\big) \quad \text{for all} \quad s \in I. \tag{3.14}$$

We associate the family of curves $\gamma\colon I \to \mathbb{R}^3$, $\gamma(s) = \big(x(s), y(s), t(s)\big)$ to a single planar curve $c(s) = (x(s), y(s))$, $s \in I$, in such a way that we obey the constraint (3.14). Integrating condition (3.14), we get

$$t - t_0 = \frac{1}{2}\int_c \big(x(s)\dot{y}(s) - y(s)\dot{x}(s)\big)\, ds,$$

which means that the area enclosed by the planar curve $c$ and the straight line connecting the end of $c$ with the origin, is equal to the change of the vertical coordinate of $\gamma$ (here we assumed $t_0 = 0$), see the Figure 3.5. Another desirable condition is *to find a Riemannian metric $g$ in $\mathbb{R}^3$ such that the length of $\gamma\colon I \to \mathbb{R}^3$ is equal to the length of the planar curve $c$.* In order to satisfy it, we find a distribution $D$ of planes in $\mathbb{R}^3$ such that $\gamma$ will be tangent to $D$ and the length of the vector $\dot{c}(s) = (\dot{x}(s), \dot{y}(s))$ in $\mathbb{R}^2$ coincides with the length of the vector $\dot{\gamma}(s) = (\dot{x}(s), \dot{y}(s), \dot{t}(s)) \in D_{\gamma(s)} \subset \mathbb{R}^3$. In this case we only need the restriction $g_D$ of the Riemannian metric $g$ to planes $D_{\gamma(s)}$ that will be the sub-Riemannian metric. Thus the distribution $D$ has to be annihilated by the form dual to the additional velocity coordinate $\dot{t}$ of the spacial curve $\gamma$. So

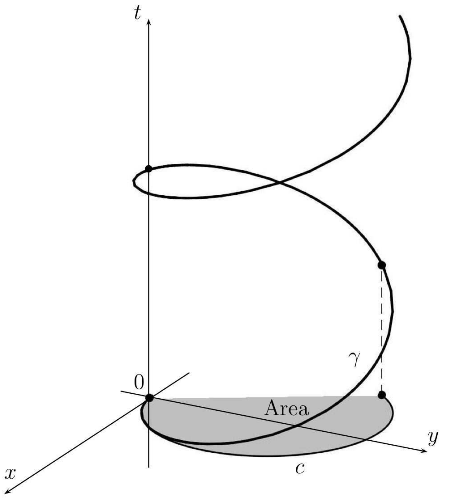

FIGURE 3.5. The sub-Riemannian length of the curve is equal to the area of the projection

$$D(x, y, t) = \ker(\omega) = \ker(dt - \frac{1}{2}(xdy - ydx)),$$

and the sub-Riemannian metric $g_D$ is just the Euclidean metric on $D$ making the basis of $D$ orthonormal. The reader may recognize the Heisenberg manifold in the space $(\mathbb{R}^3, D, g_D)$ described in the first part of Subsection 3.2. More information about the relation between the isoperimetric problems and the Heisenberg groups the reader can find in [6, 22].

**3.2.4. Variational equation on the Heisenberg group.** For the sake of completeness we would like to mention the variational equation for geodesics on $\mathbb{H}^1$ obtained in [122]. Let $g$ be a left invariant Riemannian metric on the Heisenberg group, such that the left invariant vector fields $X, Y, T$ are orthonormal at each point $q \in \mathbb{H}^1$. We emphasize that it is a Riemannian, but not a sub-Riemannian metric. Let $\nabla$ be the Levi-Civita connection associated with $g$. Let $J$ be an almost complex structure on $D$ defined as in (3.17). To formulate the result we also need to introduce the set of admissible curves for the variational problem. Recall that we are looking for a horizontal curve $c$ connecting $q_0 \in \mathbb{H}^1$ and $q_1 \in \mathbb{H}^1$ and minimizing the length functional

$$l(c) = \int_c \Big[g(\dot{c}(s), \dot{c}(s))\Big]^{1/2}\, ds. \tag{3.15}$$

The set of admissible curves for variation is just the set of horizontal curves connecting the points $q_0$ and $q_1$. Since all considered curves are horizontal we can change the Riemannian metric $g$ in (3.15) to a sub-Riemannian metric $g_D$ obtained by restriction of $g$ on $D$ at each point of the Heisenberg group.

**Theorem 3.1.** [122] *Let $\gamma\colon I \to \mathbb{H}^1$ be a smooth $(C^2)$ horizontal curve parametrized by the arc length. The curve $\gamma$ is a critical point of the length functional* (2.3) *for any admissible variation, if and only if, there is $\varkappa \in \mathbb{R}$ such that $\gamma$ satisfies the second order ordinary differential equation*

$$\nabla_{\dot{\gamma}}\dot{\gamma} + 2\varkappa J(\dot{\gamma}) = 0. \tag{3.16}$$

This result was recently extended to higher dimensional Heisenberg groups [121]. The parameter $\varkappa$ is called the curvature of the sub-Riemannian geodesic $\gamma$ since the projection of $\gamma$ to $\mathbb{R}^2$ is a curve with the curvature $\varkappa$. The value $\varkappa = 0$ corresponds to straight lines parallel to $\mathbb{R}^2$.

**3.2.5. Heisenberg algebra and inner product.** Finishing discussions on the Heisenberg group, we want to show that the commutation relations of the Heisenberg algebra induce a natural inner product on the Heisenberg algebra under an additional condition. Let us use the notations: $X, Y, T$ for the basis of the Heisenberg algebra $\mathfrak{h}^1$, $U = \operatorname{span}\{X, Y\}$, $V = \operatorname{span}\{T\}$, and $\mathfrak{h}^1 = U \oplus V$. The one-dimensional vector space $V$ is naturally isomorphic to $\mathbb{R}^1$ by fixing the basis element $T$. Therefore, $V$ possesses the metric, such that the length $|T| = 1$. Thus, $V$ becomes a normed vector space $(V, |\cdot|)$, and therefore, a metric space.

The commutator $[\cdot\,,\cdot]$ on the Heisenberg algebra $\mathfrak{h}^1$ produces a bi-linear skew symmetric form $\varpi(\cdot\,,\cdot)\colon U \times U \to V$ by $\varpi(u_1, u_2) := [u_1, u_2]$, $u_1, u_2 \in U$. Let $J$ be an almost complex structure on $U$, that is, a linear map $J\colon U \to U$ defined on the basis by

$$J^2 = -Id_U, \quad \text{and} \quad J(X) = -Y, \quad J(Y) = X. \tag{3.17}$$

Observe that $J$ is compatible with the commutator structure in the sense that the brackets are invariant under the transformation $J$. We have

$$\varpi(JX, JY) = [JX, JY] = [-Y, X] = [X, Y]$$

for the basis of $U$, and thus by bi-linearity of $\varpi$, for any vector from $U$. Verify, that $-J$ also possesses the same properties.

The *skew symmetric* bi-linear form $\varpi$ and the compatible almost complex structure $J$ define a *symmetric* bi-linear form $g(\cdot\,,\cdot)\colon U \times U \to V$ by $g(Z, W) = \varpi(JZ, W)$. Indeed, linearity is obvious and symmetry follows from

$$g(Z, W) = \varpi(JZ, W) = \varpi(JJZ, JW) = -\varpi(Z, JW) = \varpi(JW, Z) = g(W, Z).$$

On the basis elements of $U$ we get

$$g(X, X) = \varpi(JX, X) = T, \quad g(Y, Y) = \varpi(JY, Y) = T, \quad g(X, Y) = 0.$$

Recalling that we use the isomorphism $V$ with $\mathbb{R}^1$, we conclude that $g$ is a metric making $X, Y$ orthonormal.

Consider now the commutator as an adjoint map

$$\mathrm{ad}_X : U \to V : \mathrm{ad}_X(Z) = [X, Z],$$

(for the definition of the adjoint map and details see Appendix A, Subsection 8.3). Then $\mathrm{ad}_X\colon (\ker(\mathrm{ad}_X))^{\perp} \to V$ is an isomorphism, where the orthogonal complement is taken with respect to the metric $g$. Moreover, $\mathrm{ad}_X$ is an isometry form $(\ker(\mathrm{ad}_X))^{\perp}$ to $V$, since the length 1 basis element $Y$ is mapped to the length 1 basis element $T$. The same holds for $\mathrm{ad}_Y$. In the construction here we used the almost complex structure $J$.

Let $\mathbb{H}^1$ be a Heisenberg group. Left translations of the plane $U = \mathrm{span}\{X, Y\}$ defines the distribution

$$D_{(x,y,z)} = \mathrm{span}\{X = \partial_x - \frac{1}{2}y\partial_t, \quad Y = \partial_y + \frac{1}{2}x\partial_t\}$$

on $\mathbb{H}^1$. Left translations of the inner product $g$ becomes the left invariant sub-Riemannian metric $g_D$. We conclude that the commutation relations on the Heisenberg algebra and the presence of the compatible almost complex structure $J$ naturally lead to the left invariant sub-Riemannian structure $(D, g_D)$ on the group $H^1$. More about this see [77].

EXERCISES.

1. Show that the Carnot-Carathéodory distance function on the Heisenberg sub-Riemannian manifold is homogeneous with respect to the dilation (3.12).
2. Show that any two homogeneous with respect to the dilation (3.12) distance functions $d_1$ and $d_2$ are equivalent on the Heisenberg group; that is, there are constants $C, \widetilde{C} > 0$, such that
$$C d_1(\tau, q) \le d_2(\tau, q) \le \widetilde{C} d_1(\tau, q), \quad \tau, q \in \mathbb{H}^1.$$
3. Verify that the Heisenberg distance function $d_{\mathbb{H}^1}$ is symmetric and satisfies the triangle inequality. If you did not succeed see [89].

4. Show that geodesics (3.8) and (3.9) are invariant under the left translation defined by the multiplication (3.10).
5. Prove that all geodesics on the Heisenberg group can be obtained by left translations of geodesics (3.8) and (3.9) starting from $e = (0, 0, 0)$.

### 3.3. $\mathbb{H}$-type groups

The Heisenberg type groups ($\mathbb{H}$-type for shortness) were introduced by A. Kaplan [77] and have been studied extensively by many mathematicians, see for instance [21, 26, 36, 78, 88, 120].

The Heisenberg-type groups $\mathbb{H}$ are the groups whose Lie algebras $\mathfrak{h}$ are generalizations of the Heisenberg algebra in the following sense. Let a vector space $\mathfrak{h}$ endowed with a commutator $[\cdot\,,\cdot]$ and an inner product $(\cdot\,,\cdot)$. We suppose that the commutator defines the decomposition

$$\mathfrak{h} = U \oplus V, \quad [U, U] \subseteq V, \quad [U, V] = [V, V] = \{0\},$$

and, moreover, this decomposition is orthogonal with respect to the inner product. The next assumption is the compatibility between $[\cdot\,,\cdot]$, $(\cdot\,,\cdot)$, and an almost complex structure. We assume that there is a map

$$\begin{array}{rcrcl} J & : & V & \to & \mathrm{End}(U) \\ & & T & \mapsto & J_T, \end{array}$$

that satisfies

$$(J_T X, Y) = (T, [X, Y]), \quad \text{for any} \quad X, Y \in U, \ \text{and any} \ \ T \in V. \tag{3.18}$$

This immediately implies the skew-symmetry of $J_T$ for any $T \in V$:

$$(J_T X, Y) = -(X, J_T Y). \tag{3.19}$$

Here $J_T^{tr}$ denotes dual map to $J_T$. For any element $X \in U$ the adjoint map $\mathrm{ad}_X(\cdot) = [X, \cdot]\colon U \to V$ gives the decomposition

$$U = \ker(\mathrm{ad}_X) \oplus \big(\ker(\mathrm{ad}_X)\big)^{\perp},$$

where the orthogonal complement is taken with respect to the inner product $(\cdot\,,\cdot)$. We say that the Lie algebra $\mathfrak{h} = \big(U \oplus_{\perp} V, [\cdot\,,\cdot], (\cdot\,,\cdot)\big)$ is of $\mathbb{H}$-type if for any $X \in U$, $(X, X) = 1$ the map

$$\mathrm{ad}_X\colon \big(\ker(\mathrm{ad}_X)\big)^{\perp} \to V$$

is an isometry onto $V$. The last condition is equivalent to

$$J_T^2 = -|T|^2\,\mathrm{Id}_U, \quad \text{for all} \quad T \in V, \tag{3.20}$$

where $\mathrm{Id}_U$ denotes the identity mapping in $\mathrm{End}(U)$ [36]. The conditions (3.19), (3.20) imply

$$J_T J_{T'} + J_{T'} J_T = -2(T, T')\,\mathrm{Id}_U, \qquad \text{for all} \quad T, T' \in V, \tag{3.21}$$

see [36]. When there exists a linear mapping $J\colon V \to \mathrm{End}(U)$ satisfying (3.19) and (3.20) or (3.21), $U$ is called the Clifford module over $V$. The relation between $\mathbb{H}$-type groups and Clifford modules was carefully studied in [36]. Some interesting generalization where an inner product is changed to an arbitrary non-degenerate scalar product can be found in [34, 35, 59].

In Theorem 3.2 a result on the classification of $\mathbb{H}$-type algebras is presented. We need some definitions.

**Definition 17.** *The algebra $\mathfrak{h}$ satisfies the $J^2$ condition if, whenever $X \in U$ and $T, T' \in V$ with $(T, T') = 0$, then there exists $T'' \in V$, such that*

$$J_T J_{T'} X = J_{T''} X. \tag{3.22}$$

Denote by $\mathfrak{h}_0^n$ the Euclidean $n$-dimensional space, by $\mathfrak{h}_1^n$ the $n$-dimensional Heisenberg algebra, by $\mathfrak{h}_3^n$ the $n$-dimensional quaternion $\mathbb{H}$-type algebra, and by $\mathfrak{h}_7^1$ the octonion $\mathbb{H}$-type algebra. The lower index corresponds to the topological dimension of $V$ and the upper index reflects the real, complex, quaternion and octonion topological dimensions of $U$.

**Theorem 3.2** ([36]). *Suppose that $\mathfrak{h}$ is an $\mathbb{H}$-type algebra satisfying the $J^2$ condition. Then $\mathfrak{h}$ is isometrically isomorphic to $\mathfrak{h}_0^n$, $\mathfrak{h}_1^n$, $\mathfrak{h}_3^n$ or to $\mathfrak{h}_7^1$.*

This classification is intimately related to Clifford algebras. The first three of $\mathbb{H}$-type algebras are also connected to division algebras of real, complex, and quaternion numbers since all of these algebras are isomorphic to some Clifford algebras. The last one $\mathfrak{h}_7^1$ is not related to division algebra of octonion numbers, since the algebra of octonion numbers is not isomorphic to a Clifford algebra due to non associative product of octonions. The groups related to division algebras were studied in [21], where the parametric formulas of geodesics and other questions also were obtained. We present their construction in the following subsection.

**3.3.1. Constructions of groups related to division algebras.** Before we describe the general construction of groups $\mathbb{H}_0^n$, $\mathbb{H}_1^n$, $\mathbb{H}_3^n$, and $\mathbb{H}_7^1$, we would like to recall the Cayley-Dickson construction of division algebras $\mathbb{R}$ (real numbers), $\mathbb{C}$ (complex numbers), $\mathbb{Q}$ (quaternion numbers), and $\mathbb{O}$ (octonion numbers). The Cayley-Dickson construction explains why each algebra fits neatly inside the next one. Recall that the division algebra means that each non-zero element has a unique inverse. The Cayley-Dickson construction is given nicely in [8].

The complex number, as well known, can be thought of as a pair $(a, b)$ of real numbers $a, b \in \mathbb{R}$. We define the conjugate to a real number as $a^* = a$ and the conjugate to the pair as

$$(a, b)^* = (a^*, -b). \tag{3.23}$$

Then the *Cayley-Dickson product* is defined by

$$(a, b)(c, d) = (ac - db^*, a^*d + cb). \tag{3.24}$$

Now we can think of a pair $(a, b)$ as a quaternion, where $a, b \in \mathbb{C}$. The conjugate is defined as in (3.23) and the product as in (3.24). We obtain the quaternion numbers $\mathbb{Q}$ that form a non-commutative algebra with respect to (3.24). Finally, we define an octonion as a pair $(a, b)$ with $a, b \in \mathbb{Q}$, the conjugate as in (3.23), and the product (3.24). The octonions with the multiplication (3.24) form a non-commutative, non-associative algebra. Actually, we can continue the *Cayley-Dickson construction* doubling the dimension and getting a bit worse algebras. First we loose the fact that every element is its own conjugate, then we loose commutativity, associativity, and finally we loose the division algebra property. An algebra possesses a division property if

$$xy = 0 \quad \text{implies} \quad x = 0 \quad \text{or} \quad y = 0.$$

Using the Cayley-Dickson product, we first describe the following groups: Euclidean $n$-dimensional space $\mathbb{H}_0^n = \mathbb{R}^n$, the $n$-dimensional Heisenberg group $\mathbb{H}_1^n$, the $n$-dimentional quaternion $\mathbb{H}$-type group $\mathbb{H}_3^n$, and the octonion group $\mathbb{H}_7^1$. The corresponding Lie algebras $\mathfrak{h}_0^n$, $\mathfrak{h}_1^n$, $\mathfrak{h}_3^n$, and $\mathfrak{h}_7^1$ are infinitesimal representations of these groups.

We recall that the definitions of $\mathbb{H}_7^1$ and $\mathfrak{h}_7^1$ differ from the ones in [36]. In our construction we used the octonion product which is not associative, therefore, it can not give a Clifford algebra, where the product is associative by definition. The group corresponding to the classification of Theorem 3.2 is essentially the same as we present, where the product of octonions has to be changed to an associative multiplication, presented in [36].

THE EUCLIDEAN SPACE. The group $\mathbb{H}_0^n = \big(\mathbb{R}^n, +\big)$ is a trivial example of an $\mathbb{H}$-type group. We have the identifications

$$\mathfrak{h}_0^n = T_e\mathbb{R}^n = \mathbb{R}^n = \operatorname{span}\{\partial_{x^1}, \dots, \partial_{x^n}\}.$$

Left invariant vector fields are linear combinations of $(\partial_{x^1}, \dots, \partial_{x^n})$ with constant coefficients. The exponential map is the identity map. Since all commutators $[\partial_{x^i}, \partial_{x^j}]$, $i, j = 1, \dots, n$, vanish, we get

$$\mathfrak{h}_0^n = U \oplus V, \qquad U = \mathbb{R}^n, \quad V = \{0\}.$$

THE HEISENBERG GROUP $\mathbb{H}_1^n$. In the notation of the Heisenberg group $\mathbb{H}_1^n = \big(\mathbb{C}^n \times \mathbb{R}, \circ\big)$ the upper index stands for the complex dimension of $\mathbb{C}^n$ that corresponds to the real dimension $2n$ of the space $U$ in $\mathfrak{h}_1^n = U \oplus V$. The lower index reflects the real dimension of $\mathbb{R}$ that is isomorphic to the center $V$ of the algebra. We start from $n = 1$, and then generalize it to an arbitrary $n \in \mathbb{N}$. Complex numbers considered as a vector space have 2 basis vectors that we call units, since their squares have absolute value 1:

$$\text{real} \;\; 1 = (1, 0), \quad 1^2 = 1, \quad \text{and imaginary} \quad i = (0, 1), \quad i^2 = -1.$$

Take a complex number $z = (x_1, x_2)$, $x_1, x_2 \in \mathbb{R}$, and a real number $t$. Define a new non-commutative law between the elements $\tau = [z, t]$, $q = [z', t'] \in \mathbb{C} \times \mathbb{R}$ by

$$\tau \circ q = \tau q = [z, t][z', t'] = [z + z', t + t' + \frac{1}{2}(zi) \cdot z'], \tag{3.25}$$

where firstly we take the Cayley-Dickson product $zi = (x_1, x_2)(0, 1)$ and then the inner product "$\cdot$" of vectors $z, z' \in \mathbb{R}^2$. We write $\tau q$ instead of $\tau \circ q$ for the $\mathbb{H}_1^n$

group product to simplify the notation. If we use the representation of $i$ as the $(2 \times 2)$ matrix

$$\mathbf{i} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

then the group law can be written as

$$\tau q = [z,t][z',t'] = [z+z', t+t' + \frac{1}{2}(\mathbf{i}z) \cdot z']. \tag{3.26}$$

Using the algebraic form of a complex number $z = x_1 + ix_2 = \mathrm{Re}z + i\,\mathrm{Im}z$, we can write (3.25) in the form

$$\tau q = [z,t][z',t'] = [z+z', t+t' + \frac{1}{2}\mathrm{Im}(z^* z')],$$

where $z^* z'$ is the Cayley-Dickson product of $z^*$ by $z'$. The non-commutativity of the new multiplication law in $\mathbb{C} \times \mathbb{R}$ is seen for the last variable $t \in \mathbb{R}$. The dimension one of the second slot of coordinates reflects the existence of only one imaginary unit. The reader easily recognizes in (3.26) the Heisenberg group $\mathbb{H}^1$ multiplication law (3.10).

In order to present an $n$-dimensional analogue of the Heisenberg group we take two $n$-dimensional vectors of complex numbers $w = (z_1, \dots, z_n)$ and $w' = (z'_1, \dots, z'_n)$ where $z_l = x^{1l} + ix^{2l}$, $z'_l = (x^{1l})' + i(x^{2l})'$, $l = 1, \dots, n$. The matrix $\mathbf{i}$ is changed to a block diagonal matrix $\mathbf{J} = \operatorname{diag} \mathbf{i}$ with $n$ matrices $\mathbf{i}$ on the diagonal. The multiplication law between the elements $\tau = [w,t]$ and $q = [w',t'] \in \mathbb{C}^n \times \mathbb{R}$ is transformed into the following one

$$\begin{aligned} \tau q = [w,t][w',t'] &= [w+w', t+t' + \frac{1}{2}\sum_{l=1}^{n}(z_l i) \cdot z'_l] \\ &= [w+w', t+t' + \frac{1}{2}(\mathbf{J}w) \cdot w'] \\ &= [w+w', t+t' + \frac{1}{2}\mathrm{Im}(w^* w')], \end{aligned}$$

where $w^* w' = \sum_{l=1}^{n} z_l^* z'_l$. The unit element is $e = (0,0)$ and $(-w - t) = (w,t)^{-1}$ is the inverse element to $(w,t)$.

A left invariant basis of the Heisenberg algebra $\mathfrak{h}_1^n$, $n \in \mathbb{N}$, is obtained as in the one dimensional case by translation of the basis vectors $\{\partial_{x^{1l}}, \partial_{x^{2l}}, \partial_t\}_{l=1}^n$ by $d_e l_\tau$. With $l = 1, \dots, n$ we get

$$X_{1l} = \partial_{x_{1l}} - \frac{1}{2}x_{2l}\partial_t, \quad X_{2l} = \partial_{x_{2l}} + \frac{1}{2}x_{1l}\partial_t, \quad \text{and} \quad T = \partial_t. \tag{3.27}$$

Let us introduce the notations $U = \mathrm{span}\{X_{1l}, X_{2l}\}_{l=1}^n$, $V = \mathrm{span}\{T\}$. Since $[X_{1l}, X_{2l}] = T$ and other commutators vanish, we get $\mathfrak{h}_1^n = U \oplus V$. Let the inner product $(\cdot\,,\cdot)$ in $\mathfrak{h}_1^n$ be such that basis vectors become orthonormal. The condition (3.18) holds due to the commutation relations. The endomorphism $J_T$ is represented by the matrix $\mathbf{J}$, which possesses properties (3.19), (3.20). The $J^2$ condition holds trivially, since $\dim V = 1$. The space $(U, J, (\cdot\,,\cdot))$ is isomorphic to the space of complex numbers $(\mathbb{C}^n, i, |\cdot|)$, with usual multiplication, which is the simplest example of a Clifford algebra with generators $1, i$.

Quaternion group $\mathbb{H}^n_3$. In the notation of the quaternion group $\mathbb{H}^n_3 = \big(\mathbb{Q}^n \times \mathbb{R}, \circ\big)$ the upper index denotes the quaternion dimension of the space of quaternions $\mathbb{Q}^n$ that corresponds to the real dimension $4n$ of the horizontal distribution. The lower index in this case reflects the real dimension 3 of the center of the Lie algebra $\mathfrak{h}^n_3$. As previously, we start from the one-dimensional case: $n = 1$, and then consider its multidimensional analogue. Quaternion numbers $\mathbb{Q}$, which we think of as pairs of complex numbers, have one real unity $1 = (1, 0)$, $1^2 = 1$, and three imaginary units

$$i_1 = (i, 0), \quad i_2 = (0, 1), \quad i_3 = (0, i), \quad \text{such that} \quad i_1^2 = i_2^2 = i_3^2 = i_1 i_2 i_3 = -1.$$

The Cayley-Dickson product is no longer commutative, for example,

$$i_1 i_2 = -i_2 i_1 = -i_3, \quad i_2 i_3 = -i_3 i_2 = -i_1, \quad i_3 i_1 = -i_1 i_3 = -i_2. \tag{3.28}$$

In order to construct the quaternion $\mathbb{H}$-type group $\mathbb{H}^1_3$, we take a quaternion $q = (z_1, z_2)$, $z_1, z_2 \in \mathbb{C}$, and three real numbers $t_1, t_2, t_3$ that reflects the three dimensional nature of the space of the imaginary quaternions. Define a new non-commutative law between the elements $h = [q, t_1, t_2, t_3] \in \mathbb{Q} \times \mathbb{R}^3$ and $p = [q', t'_1, t'_2, t'_3] \in \mathbb{Q} \times \mathbb{R}^3$ by

$$\begin{aligned} hp &= [q, t_1, t_2, t_3][q', t'_1, t'_2, t'_3] \\ &= [q + q', t_1 + t'_1 + \frac{1}{2}(q i_1) \cdot q',\ t_2 + t'_2 + \frac{1}{2}(q i_2) \cdot q',\ t_3 + t'_3 + \frac{1}{2}(q i_3) \cdot q'], \end{aligned} \tag{3.29}$$

where $q i_k$, $k = 1, 2, 3$ is the Cayley-Dickson product for the quaternions and "$\cdot$" is the inner product in $\mathbb{R}^4$. As in the case of the Heisenberg group we can use the matrix representation of the imaginary units

$$\mathbf{i}_1 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \ \mathbf{i}_2 = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \tag{3.30}$$

$$\mathbf{i}_3 = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

and rewrite the group law (3.29) in the form

$$hp = [q + q', t_1 + t'_1 + \frac{1}{2}(\mathbf{i}_1 q) \cdot q',\ t_2 + t'_2 + \frac{1}{2}(\mathbf{i}_2 q) \cdot q',\ t_3 + t'_3 + \frac{1}{2}(\mathbf{i}_3 q) \cdot q'].$$

Using the imaginary units we can represent a quaternion $q$ in the algebraic form as

$$q = \alpha + i_1 \beta + i_2 \gamma + i_3 \delta = \alpha + i_1 \mathrm{Im}_1 q + i_2 \mathrm{Im}_3 q + i_3 \mathrm{Im}_3 q.$$

Then the multiplication law (3.29) admits the form

$$\begin{aligned} hp &= [q, t_1, t_2, t_3][q', t_1', t_2', t_3'] \qquad (3.31)\\ &= [q+q', t_1+t_1'+\frac{1}{2}\mathrm{Im}_1(q^*\,q'),\ t_2+t_2'+\frac{1}{2}\mathrm{Im}_2(q^*\,q'),\ t_3+t_3'+\frac{1}{2}\mathrm{Im}_3(q^*\,q')], \end{aligned}$$

where $q^*\,q'$ is the Cayley-Dickson product of $q^*$ by $q'$.

To give an $n$-dimensional analogue of the quaternion $\mathbb{H}-$type group, we take the $n$-dimensional vectors of quaternion numbers $w = (q_1, \dots, q_n)$, $w' = (q_1', \dots, q_n')$. Each of the matrices $\mathbf{i}_m$, $m = 1, 2, 3$, is changed to the block diagonal matrix $\mathbf{J}_m = \mathrm{diag}\,\mathbf{i}_m$ with $n$ $(4\times 4)$-matrices $\mathbf{i}_m$ on the main diagonal. The multiplication law between the elements $h = [w, t_1, t_2, t_3]$, $p = [w', t_1', t_2', t_3'] \in \mathbb{Q}^n \times \mathbb{R}^3$ is

$$\begin{aligned} hp &= [w, t_1, t_2, t_3][w', t_1', t_2', t_3']\\ &= [w+w', t_1+t_1'+\frac{1}{2}\sum_{l=1}^{n}(q_l i_1)q_l', t_2+t_2'+\frac{1}{2}\sum_{l=1}^{n}(q_l i_2)q_l', t_3+t_3'+\frac{1}{2}\sum_{l=1}^{n}(q_l i_3)q_l']\\ &= [w+w', t_1+t_1'+\frac{1}{2}(\mathbf{J}_1 w)\cdot w', t_2+t_2'+\frac{1}{2}(\mathbf{J}_2 w)\cdot w', t_3+t_3'+\frac{1}{2}(\mathbf{J}_3 w)\cdot w']\\ &= [w+w', t_1+t_1'+\frac{1}{2}\mathrm{Im}_1(w^*\,w'), t_2+t_2'+\frac{1}{2}\mathrm{Im}_2(w^*\,w'), t_3+t_3'+\frac{1}{2}\mathrm{Im}_3(w^*\,w')], \end{aligned}$$

where $w^*\,w' = \sum_{l=1}^{n} q_l^*\,q_l'$. The unit of the group $\mathbb{H}_3^n$ is $e = (0,0) \in \mathbb{Q}^n \times \mathbb{R}^3$ and the inverse element to $(w, t_1, t_2, t_3)$ is $(-w, -t_1, -t_2, -t_3)$, $w \in \mathbb{Q}^n$, $t_1, t_2, t_3 \in \mathbb{R}$.

The quaternion algebra $\mathfrak{h}_3^n$, $\in \mathbb{N}$, is the direct sum of $U \oplus V$, where

$$U = \mathrm{span}(X_{11}, X_{21}, X_{31}, X_{41}, \dots, X_{1n}, X_{2n}, X_{3n}, X_{4n})$$

with

$$\begin{aligned} X_{1l}(w,t) &= \partial_{x_{1l}} + \frac{1}{2}\big(-x_{2l}\partial_{t_1} - x_{3l}\partial_{t_2} - x_{4l}\partial_{t_3}\big),\\ X_{2l}(w,t) &= \partial_{x_{2l}} + \frac{1}{2}\big(x_{1l}\partial_{t_1} + x_{4l}\partial_{t_2} - x_{3l}\partial_{t_3}\big),\\ X_{3l}(w,t) &= \partial_{x_{3l}} + \frac{1}{2}\big(-x_{4l}\partial_{t_1} + x_{1l}\partial_{t_2} + x_{2l}\partial_{t_3}\big),\\ X_{4l}(w,t) &= \partial_{x_{4l}} + \frac{1}{2}\big(x_{3l}\partial_{t_1} - x_{2l}\partial_{t_2} + x_{1l}\partial_{t_3}\big), \end{aligned} \qquad l = 1, \dots n, \qquad (3.32)$$

and $w = (q_1, \dots, q_n) = (x_{11}, x_{21}, x_{31}, x_{41}, \dots, x_{1\,n}, x_{2\,n}, x_{3\,n}, x_{4\,n})$. The subspace $V$ is spanned by $\{T_1, T_2, T_3\}$ with $T_k = \partial_{t_k}$. The following commutator relations

$$\begin{gathered} [X_{1l}, X_{2l}] = T_1, \quad [X_{1l}, X_{3l}] = T_2, \quad [X_{1l}, X_{4l}] = T_3,\\ [X_{2l}, X_{3l}] = T_3, \quad [X_{2l}, X_{4l}] = -T_2, \quad [X_{3l}, X_{4l}] = T_1, \end{gathered}$$

hold for $l = 1, \dots, n$ and all other brackets vanish. Thus, the condition (3.18) is verified with respect to the inner product making the bases of $U$ and $V$ orthonormal. The endomorphisms $J_{T_m}$ are represented by matrices $\mathbf{J}_m$, $m = 1, 2, 3$. The $J^2$ condition holds due to relation (3.28).

**Remark 1.** If we involve only two imaginary units into the construction, then we obtain the quaternion $\mathbb{H}$-type group with two dimensional center $V$. Taking into consideration only one of the $i_k$, $k = 1, 2, 3$, we get a group isomorphic to the Heisenberg group $\mathbb{H}^n_1$.

Octonion $\mathbb{H}$-type group $\mathbb{H}^1_7$. Octonions or Caley numbers, which we think of as pairs of quaternion numbers, have one real unit $1 = (1, 0)$, $1^2 = 1$ and 7 imaginary units

$$\begin{aligned} &j_1 = (i_1, 0), \quad j_2 = (i_2, 0), \quad j_3 = (i_3, 0), \quad j_4 = (0, 1), \\ &j_5 = (0, i_1), \quad j_6 = (0, i_2), \quad j_7 = (0, i_3), \end{aligned}$$

whose squares equal $(-1)$. The rule of multiplication is presented in Table 9 in Appendix B. The product of octonions is not associative, for example,

$$j_1(j_2j_4) = -j_7, \qquad (j_1j_2)j_4 = j_7.$$

We take an octonion $w = (q_1, q_2)$, $q_1, q_2 \in \mathbb{Q}$ and $t \in \mathbb{R}^7$, corresponding to the 7-dimensional space of imaginary octonions. Define a new non-commutative product law for elements

$$h = [w, t] \in \mathbb{O} \times \mathbb{R}^7 \quad \text{and} \quad p = [w', t'] =\in \mathbb{O} \times \mathbb{R}^7$$

by

$$\begin{aligned} hp &= [w, t][w', t'] = [w, t_1, \dots, t_7][w', t'_1, \dots, t'_7] \\ &= [w + w', t_1 + t'_1 + \frac{1}{2}(wj_1) \cdot w', \ \dots \ , t_7 + t'_7 + \frac{1}{2}(wj_7) \cdot w'] \qquad (3.33) \\ &= [w + w', t_1 + t'_1 + \frac{1}{2}\mathrm{Im}_1(w^* \, w'), \ \dots \ , t_7 + t'_7 + \frac{1}{2}\mathrm{Im}_7(w^* \, w')], \end{aligned}$$

where $wj_m$, $m = 1, \dots, 7$, and $w^* \, w'$ are the Cayley-Dickson products and "$\cdot$" is the inner product in $\mathbb{R}^8$. As in the previous cases $e = (0, 0)$ is the unit element of the group $\mathbb{H}^1_7$, $(w, t)^{-1} = (-w, -t)$.

There is no matrix representation of $j_k$ since the multiplication between $j_k$ is not associative, in contrast to the matrix multiplication. Nevertheless, it is possible to associate a matrix $\mathbf{J}_m$ with any imaginary unit $j_m$ which can be considered as a replacement of endomorphism $J_{Z_m}$, $m = 1, \dots, 7$. The matrices $\mathbf{J}_m$ are given in the Appendix B. Using $\mathbf{J}_m$ we write the multiplication law (3.33) as follows

$$hp = [w, t][w', t'] = [w+w', t_1+t'_1-\frac{1}{2}(w\mathbf{J}_1)\cdot w', \ \dots \ , t_7+t'_7-\frac{1}{2}(w\mathbf{J}_7)\cdot w']. \qquad (3.34)$$

Notice some properties of the matrices $\mathbf{J}_m$:

$$\mathbf{J}^2_m = -U, \ \ \mathbf{J}^T_m = -\mathbf{J}_m, \ \ \mathbf{J}^{-1}_m = \mathbf{J}_m, \ \ m = 1, \dots, 7, \qquad (3.35)$$

where $U$ is the $(7 \times 7)$ identity matrix. The product of the matrices $\mathbf{J}_m$ does not correspond to the product of the corresponding imaginary unities $j_m$, for example,

$$j_1j_2 = -j_3, \quad \text{but} \quad \mathbf{J}_1\mathbf{J}_2 \neq -\mathbf{J}_3.$$

The matrices $\mathbf{J}_m$ do not represent the imaginary units in octonion algebra $\mathbb{O}$, but they can be used to write the group law and the left invariant basis of the corresponding algebra.

The algebra $\mathfrak{h}_7^1$ is the direct sum $U \oplus V$, where $U = \operatorname{span}(X_1, \dots, X_8)$ with

$$X_l(w,t) = \partial_{x_l} + \frac{1}{2}\sum_{m=1}^{7}(x\mathbf{J}_m)_l \partial_{t_m}, \qquad l = 1, \dots, 8, \tag{3.36}$$

where $w = (x_1, \dots, x_8)$ and $(x\mathbf{J}_m)_l$ is the $l$-th coordinate of the vector $x\mathbf{J}_m$. We give the coefficients $(x\mathbf{J}_m)_l$ in the Table 1. For instance, to write the vector field $X_1$ we take the first line from the Table 1 and get

$$X_1(w,t) = \partial_{x_1} + \frac{1}{2}\Big( - x_2\partial_{t_1} - x_3\partial_{t_2} - x_4\partial_{t_3} - x_5\partial_{t_4} - x_6\partial_{t_5} - x_7\partial_{t_6} - x_8\partial_{t_7}\Big).$$

The subspace $V$ is spanned by $\{T_1, \dots, T_7\}$ with $T_m = \partial_{t_m}$. The non-vanishing commutators are given in Table 2 showing that condition (3.18) still holds.

| | $\partial_{t_1}$ | $\partial_{t_2}$ | $\partial_{t_3}$ | $\partial_{t_4}$ | $\partial_{t_5}$ | $\partial_{t_6}$ | $\partial_{t_7}$ |
|---|---|---|---|---|---|---|---|
| $X_1$ | $-x_2$ | $-x_3$ | $-x_4$ | $-x_5$ | $-x_6$ | $-x_7$ | $-x_8$ |
| $X_2$ | $x_1$ | $x_4$ | $-x_3$ | $x_6$ | $-x_5$ | $-x_8$ | $x_7$ |
| $X_3$ | $-x_4$ | $x_1$ | $x_2$ | $x_7$ | $x_8$ | $-x_5$ | $-x_6$ |
| $X_4$ | $x_3$ | $-x_2$ | $x_1$ | $x_8$ | $-x_7$ | $x_6$ | $-x_5$ |
| $X_5$ | $-x_6$ | $-x_7$ | $-x_8$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
| $X_6$ | $x_5$ | $-x_8$ | $x_7$ | $-x_2$ | $x_1$ | $-x_4$ | $x_3$ |
| $X_7$ | $x_8$ | $x_5$ | $-x_6$ | $-x_3$ | $x_4$ | $x_1$ | $-x_2$ |
| $X_8$ | $-x_7$ | $x_6$ | $x_5$ | $-x_4$ | $-x_3$ | $x_2$ | $x_1$ |

Table 1. The product $x\mathbf{J}_m$

| | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_7$ | $X_8$ |
|---|---|---|---|---|---|---|---|---|
| $X_1$ | 0 | $T_1$ | $T_2$ | $T_3$ | $T_4$ | $T_5$ | $T_6$ | $T_7$ |
| $X_2$ | $-T_1$ | 0 | $T_3$ | $-T_2$ | $T_5$ | $-T_4$ | $-T_7$ | $T_6$ |
| $X_3$ | $-T_2$ | $-T_3$ | 0 | $T_1$ | $T_6$ | $T_7$ | $-T_4$ | $-T_5$ |
| $X_4$ | $-T_3$ | $T_2$ | $-T_1$ | 0 | $T_7$ | $-T_6$ | $T_5$ | $-T_4$ |
| $X_5$ | $-T_4$ | $-T_5$ | $-T_6$ | $-T_7$ | 0 | $T_1$ | $T_2$ | $T_3$ |
| $X_6$ | $-T_5$ | $T_4$ | $-T_7$ | $T_6$ | $-T_1$ | 0 | $-T_3$ | $T_2$ |
| $X_7$ | $-T_6$ | $T_7$ | $T_4$ | $-T_5$ | $-T_2$ | $T_3$ | 0 | $-T_1$ |
| $X_8$ | $-T_7$ | $-T_6$ | $T_5$ | $T_4$ | $-T_3$ | $-T_2$ | $T_1$ | 0 |

Table 2. Non-vanishing commutators

Using the normal coordinates $(w, t)$ for the elements, we identify the elements of the group with the elements of the algebra via the exponential map $\exp\colon \mathfrak{h}_7^1 \to \mathbb{H}_7^1$:

$$\mathbb{H}_7^1 \ni (x_1, \dots, x_8, t_1, \dots, t_7) \overset{\exp}{\longleftarrow} \Big(\sum_{k=1}^{8} x_k X_k + \sum_{m=1}^{7} t_m T_m\Big) \in \mathfrak{h}_7^1.$$

**3.3.2. Groups related to division algebras considered as sub-Riemannian manifolds.** The group $\mathbb{H}_0^n$ is a usual Euclidean space, since $V = \{0\}$. The sub-Riemannian metric is the Riemannian metric given as a left translation of the Euclidean product by abilian group law "+". Further,

$$\mathbb{H}_1^n = (\mathbb{R}^{2n+1}, D, g_D), \quad \mathbb{H}_3^n = (\mathbb{R}^{4n+3}, D, g_D), \quad \mathbb{H}_7^8 = (\mathbb{R}^{15}, D, g_D),$$

where the left invariant distributions $D$ are such that $D_q = \operatorname{span}\{X_1(q), \dots, X_\alpha(q)\}$. Here left invariant vector fields $X_j$, $j = 1, \dots, \alpha = \dim U$, are given by (3.27), (3.32), and (3.36). We use as the underlying smooth real manifolds for groups the real vector spaces $\mathbb{R}^{2n+1}$, $\mathbb{R}^{4n+3}$, and $\mathbb{R}^{15}$ that isomorphic to $\mathbb{C}^n \times \mathbb{R}$, $\mathbb{Q}^n \times \mathbb{R}^3$, and $\mathbb{O} \times \mathbb{R}^7$, respectively. The metric $g_D$ is such that the basis of $D_q$ becomes orthonormal in each case. The distributions $D$ are of step 2, strongly bracket generating and regular.

In order to present normal geodesics on $\mathbb{H}$-type groups related to division algebras, we denote all of them by $\mathbb{H}$. We also use the notations $\alpha = \dim U_q$, $\beta = \dim V_q$ for all $q = (x, t) \in \mathbb{H} \cong \mathbb{R}^\alpha \times \mathbb{R}^\beta$, $\lambda = (\xi, \theta) \in T_q^*\mathbb{H}$. Then the Hamiltonian function is

$$\begin{aligned} H(q, \lambda) &= H(x, t, \xi, \theta) \\ &= \frac{1}{2}\sum_{l=1}^{\alpha} \langle \lambda, X_l \rangle^2 = \sum_{l=1}^{\alpha} \xi_l^2 + \frac{1}{4}\Big(\sum_{l=1}^{\alpha} (x^l)^2\Big)\Big(\sum_{m=1}^{\beta} \theta_m^2\Big) + \mathbf{M}x \cdot \xi \end{aligned}$$

where "$\cdot$" denotes the usual inner product in $\mathbb{R}^\alpha$ and $\mathbf{M} = \sum_{m=1}^{\beta} \theta_m \mathbf{J}_m$. The matrix $\mathbf{M}$ is skew symmetric.

The corresponding Hamiltonian system is

$$\begin{cases} \dot{x} &= \frac{\partial H}{\partial \xi} = 2\xi + \mathbf{M}x \\ \dot{t}_m &= \frac{\partial H}{\partial \theta_m} = \frac{\theta_m}{2}|x|^2 + \mathbf{J}_m x \cdot \xi, \quad m = 1, \dots, \beta. \\ \dot{\xi} &= -\frac{\partial H}{\partial x} = -\frac{1}{2}|\theta|^2 x - \mathbf{M}\xi \\ \dot{\theta}_m &= -\frac{\partial H}{\partial t_m} = 0 \end{cases} \tag{3.37}$$

and the initial conditions are $x(0) = t(0) = 0$, $\xi(0) = \xi_0$, $\theta(0) = \theta_0$. We see from the last equation that the $\theta$-coordinates of the momentum are constants $\theta_0$. The system (3.37) is reduced to

$$\begin{cases} \ddot{x} = 2\mathbf{M}\dot{x} \\ \dot{t} = -\frac{1}{2}\mathbf{J}_m x \cdot \dot{x}, \quad m = 1, \dots, \beta. \end{cases} \tag{3.38}$$

The solutions of the system (3.38) for $\theta_0 \neq 0$ are

$$\begin{cases} x(s) = \frac{1-\cos(2s|\theta_0|)}{2|\theta_0|^2}\mathbf{M}\dot{x}(0) + \frac{\sin(2s|\theta_0|)}{2|\theta_0|}\,\mathrm{Id}\,\dot{x}(0) \\ t^m(s) = \frac{(\theta_0)_m|\dot{x}(0)|^2}{4|\theta_0|^2}\big(s - \frac{\sin(2s|\theta_0|)}{2|\theta_0|}\big), \qquad m = 1,\dots,\beta, \end{cases} \tag{3.39}$$

where Id is the identity matrix in $\mathbb{R}^\alpha$. If $\theta_0$ vanishes, then the geodesics are straight lines starting from $(0,0)$ with the initial velocity $\xi_0$, $\theta_0 = 0$, remaining in the space $t = 0$. Analysing the solutions, we obtain the information about the behaviour of geodesics similar to the Heisenberg group. Projections to any 3-dimensional subspace containing two coupled $x$-coordinates and one $t$-coordinate give the Heisenberg picture, see Figure 3.2. The detailed study of this question can be found in [21].

In the work [58] the authors obtained an analogue of the variational equation (3.1) for the $\mathbb{H}$-type group $\mathbb{H}_3^1 = (\mathbb{R}^7, \rho)$, where the group law $\rho$ is given by (3.31).

**Theorem 3.3.** *Let $\gamma\colon [a,b] \to \mathbb{H}_3^1$ be a horizontal curve, parameterized by arc length. Then $\gamma$ is a critical point of the functional for the Carnot-Carathéodory distance, if and only if, there exist numbers $\varkappa_1, \varkappa_2, \varkappa_3 \in \mathbb{R}$ such that $\gamma$ satisfies the second order differential equation*

$$\nabla_{\dot{\gamma}}\dot{\gamma} - 2\sum_{m=1,2,3}\varkappa_m J_m(\dot{\gamma}) = 0. \tag{3.40}$$

Here $J_m$, $m = 1,2,3$, are the almost complex structures given by the endomorphisms $J_{T_m}$, $m = 1,2,3$ in the definition of $\mathbb{H}_3^1$. The Levi-Civita connection $\nabla$ is compatible with the extension $g$ of the sub-Riemannian metric $g_D$ from $D$ onto the entire tangent bundle $T\mathbb{H}_3^1$. The Riemannian metric $g$ is left invariant. The question whether there are similar equations for other $\mathbb{H}$-type groups is still open. See an analogue of equation (3.40) in Subsection 5.5.

We also present here formulas for horizontal gradients on groups $\mathbb{H}$, written in the left invariant bases $(X_1,\dots,X_\alpha)$ given in (3.27), (3.32), (3.36) and in the standard basis $\{\partial_{x^l}, \partial_{t^m}\}$, $l = 1,\dots,\alpha$, $m = 1,\dots,\beta$.

$$\mathrm{grad}_D = (X_1,\dots,X_\alpha) = (\mathrm{grad}_x + \frac{1}{2}\sum_{m=1}^{\beta}(\mathbf{J}_m x)\partial_{t_m}), \tag{3.41}$$

with $\mathrm{grad}_x = (\partial_{x^1},\dots,\partial_{x^\alpha})$.

### 3.3.3. Action of groups related to division algebras on the Siegel upper half spaces.

Let $\mathbb{C}^{n+1}$ be the $(n+1)$-dimensional complex space. We use the notation $z = (z', z_{n+1})$, where $z' = (z_1,\dots,z_n) \in \mathbb{C}^n$. The set

$$\mathcal{U}_n = \{(z_1,\dots,z_{n+1}) \in \mathbb{C}^{n+1} :\ 4\mathrm{Re}\,(z_{n+1}) > |z'|^2 = \sum_{l=1}^{n}|z_l|^2\}$$

defines the Siegel upper half space in $\mathbb{C}^{n+1}$. Let $B_{\mathbb{C}}$ denotes the unit ball in $\mathbb{C}^{n+1}$:

$$B_{\mathbb{C}} = \{(w_1, \dots, w_{n+1}) \in \mathbb{C}^{n+1} : \sum_{l=1}^{n+1} |w_l|^2 < 1\}.$$

Then the Cayley transformation

$$w_{n+1} = \frac{1 - z_{n+1}}{1 + z_{n+1}}, \quad w_l = \frac{z_l}{1 + z_{n+1}}, \qquad l = 1, \dots, n,$$

and its inverse

$$z_{n+1} = \frac{1 - w_{n+1}}{1 + w_{n+1}}, \quad z_l = \frac{2w_l}{1 + w_{n+1}}, \qquad l = 1, \dots, n,$$

show that the unit ball $B_{\mathbb{C}}$ and the Siegel upper half space $\mathcal{U}_n$ are biholomorphically equivalent.

Let $\mathbb{Q}^{n+1}$, $n+1 \in \mathbb{N}$, be an $(n+1)$-dimensional quaternion vector space. The elements of $\mathbb{Q}^{n+1}$ are $(n+1)$-tuples of quaternions that we denote by $q = (q', q_{n+1})$, $q' = (q_1, \dots, q_n) \in \mathbb{Q}^n$, with the norm $|q|^2 = \sum_{l=1}^{n+1} |q|_l^2$. The Siegel upper half space in $\mathbb{Q}^{n+1}$ can be defined by analogy with the complex case as:

$$\mathcal{U}_n = \{(q_1, \dots, q_{n+1}) \in \mathbb{Q}^{n+1} : \ 4\mathrm{Re}\,(q_{n+1}) > \sum_{l=1}^{n} |q_l|^2 = |q'|^2\}.$$

The unit ball $B_{\mathbb{Q}}$ in $\mathbb{Q}^{n+1}$ is

$$B_{\mathbb{Q}} = \{(h_1, \dots, h_{n+1}) \in \mathbb{Q}^{n+1} : \sum_{l=1}^{n+1} |h_l|^2 < 1\}.$$

Since the multiplication of quaternions is not commutative there are two forms of Cayley transformation that give the symmetric geometry. The (left) Cayley transformation, mapping the Siegel upper half space $\mathcal{U}_n$ onto the unit ball $B_{\mathbb{Q}}$, has the form

$$q_{n+1} = (1 + h_{n+1})^{-1}(1 - h_{n+1}) = \frac{(1 + h_{n+1}^*)(1 - h_{n+1})}{|1 + h_{n+1}|^2},$$

$$q_l = 2h_l(1 + h_{n+1})^{-1} = \frac{2h_l(1 + h_{n+1}^*)}{|1 + h_{n+1}|^2},$$

for $l = 1, \dots, n$. The inverse transformation from $B_{\mathbb{Q}}$ onto $\mathcal{U}_n$ is

$$h_l = q_l(1 + q_{n+1})^{-1} = \frac{q_l(1 + q_{n+1}^*)}{|1 + q_{n+1}|^2},$$

$$h_{n+1} = (1 - q_{n+1})(1 + q_{n+1})^{-1} = \frac{(1 - q_{n+1})(1 + q_{n+1}^*)}{|1 + q_{n+1}|^2}$$

for $l = 1, \dots, n$. The Cayley transformation is biholomorphic in the quaternion sense, where the notion of a quaternion holomorphic function is not a direct generalization of a holomorphic complex function, it requires some additional inputs, see, for instance [39, 68].

Continuing by analogy, we define the Siegel upper half space in $\mathbb{O}^2$. We consider only 2-dimensional octonion vector space: $\mathbb{O}^2 = \{o = (o', o_2) : \ o', o_2 \in \mathbb{O}\}$ with the norm $|o|^2 = |o'|^2 + |o_2|^2$. The Siegel upper half space $\mathcal{U}_1$ and the unit disk $B_{\mathbb{O}}$ is defined as

$$\begin{aligned}\mathcal{U}_1 &= \{(o', o_2) \in \mathbb{O}^2 : \ 4\mathrm{Re}\,(o_2) > |o'|^2\},\\ B_{\mathbb{O}} &= \{(r', r_2) \in \mathbb{O}^2 : \ |r'|^2 + |r_2|^2 < 1\}.\end{aligned}$$

The transformations

$$o' = 2r'(1 + r_2)^{-1}, \qquad o_2 = (1 + r_2)^{-1}(1 - r_2),$$

map the Siegel upper half space onto the unit ball and the map

$$r' = o'(1 + o_2)^{-1}, \qquad r_2 = (1 - o_2)(1 + o_2)^{-1}$$

acts from the unit ball to the Siegel upper half space. To verify this we observe that

$$(xy)^* = y^* x^*, \quad x^{-1} = \frac{x^*}{|x|^2},$$

where $x$ is a complex number, a quaternion or an octonion, and $xy$ is the Cayley-Dickson product. Let us denote by $(q', q)$ an element of one of the above mentioned Siegel upper half spaces, and by $(h', h)$ a point from the corresponding unit ball. Then

$$\begin{aligned}|h'|^2 &= h'(h')^* = |q'|^2 |(1 + q)^{-1}|^2,\\ |h|^2 &= hh^* = |1 - q|^2 |(1 + q)^{-1}|^2,\end{aligned}$$

and

$$|h'|^2 + |h|^2 = (|q'|^2 + |1 - q|^2)|(1 + q)^{-1}|^2 < 1.$$

Since $|(1 + q)^{-1}|^2 = |(1 + q)|^{-2}$, we have

$$|q'|^2 + |1 - q|^2 = |q'|^2 + 1 + |q|^2 - 2\mathrm{Re}\,(q) < |(1 + q)|^2 = 1 + |q|^2 + 2\mathrm{Re}\,(q),$$

that yields $|q'|^2 < 4\mathrm{Re}\,(q)$.

Let $\mathbb{K}$ be one of the following spaces $\mathbb{C}^{n+1}$, $\mathbb{H}^{n+1}$ or $\mathbb{O}^2$. We denote by $p = (q', q)$ a point from the Siegel upper half space $\mathcal{U}_n$ of $\mathbb{K}$. The boundary of $\mathcal{U}_n$ is

$$\partial\mathcal{U}_n = \{(q', q) \in \mathbb{K} : \ 4\mathrm{Re}\,(q) = |q'|^2 = \sum_{l=1}^{n} |q_l|^2\}.$$

We mention here three automorphisms of the domain $\mathcal{U}_n$: dilation, rotation and translation.

**Dilation.** Let $p = (q', q) \in \mathcal{U}_n$. For every positive number $\delta$ we define a *dilation* $\delta_s(p)$ by

$$\delta_s(p) = \delta_s(q', q) = (sq', s^2 q).$$

The non-isotropy of the dilation comes from the definition of $\mathcal{U}_n$.

**Rotation.** For every unitary linear transformation $U$ that acts on $\mathbb{C}^n$, $\mathbb{H}^n$ or $\mathbb{O}$, we define the *rotation* $\mathrm{Rot}(p)$ on $\mathcal{U}_n$ by

$$\mathrm{Rot}(p) = \mathrm{Rot}(q', q) = (U(q'), q).$$

Both the dilation and the rotation are extended to mappings on the boundary $\partial\mathcal{U}_n$.

**Translation.** We use the notation $\mathbb{H}$ for the groups $\mathbb{H}_1^n$, $\mathbb{H}_3^n$, $\mathbb{H}_7^1$. To every element $[w,t]$ of $\mathbb{H}$ we associate the following affine self-map of $\mathcal{U}_n$. Notice that it is a holomorphic map for the cases $\mathbb{C}^{n+1}$, $\mathbb{Q}^{n+1}$. This map is the action on the left of the group $\mathbb{H}$ on $\mathcal{U}_n$:

$$[w,t].(q',q) \mapsto (q'+w, q+\frac{|w|^2}{4}+\frac{1}{2}w^*q'+i\cdot t). \tag{3.42}$$

Here $i\cdot t=\sum_{l=1}^{\beta} i_k t_k$. This mapping preserves the level sets, given by the function

$$r(p)=4\mathrm{Re}\,(q)-|q'|^2. \tag{3.43}$$

In fact, since $|q'+w|^2=|q'|^2+|w|^2+2\mathrm{Re}\,(w^*q_1)$, we obtain

$$4\mathrm{Re}\,(q+\frac{|w|^2}{4}+\frac{1}{2}w^*q')-|q'+w|^2=4\mathrm{Re}\,(q)-|q'|^2.$$

Hence, the transformation (3.42) maps $\mathcal{U}_n$ onto itself and preserves the boundary $\partial\mathcal{U}_n$.

Let us check that the mapping (3.42) defines an action of the group $\mathbb{H}$ on the space $\mathcal{U}_n$. If we compose the mappings (3.42), corresponding to elements $[w,t]$ and $[\omega,s]\in\mathbb{H}$, we get

$$\begin{aligned}&[w,t].\big([\omega,s].(q',q)\big)\\ &\quad=\Big(w+\omega+q', q+\frac{|w|^2}{4}+\frac{|\omega|^2}{4}+\frac{1}{2}(w+\omega)^*q'+\frac{1}{2}w^*\omega+i\cdot(s+t)\Big).\end{aligned} \tag{3.44}$$

On the other hand, the transformation corresponding to the element $[w,t][\omega,s]$ is

$$\begin{aligned}&[w,t][\omega,s].(q',q)\\ &\quad=\Big(w+\omega+q', q+\frac{|w+\omega|^2}{4}+\frac{1}{2}(w+\omega)^*q'+\frac{1}{2}i\cdot\mathrm{Im}\,w^*\omega+i\cdot(s+t)\Big).\end{aligned} \tag{3.45}$$

Observing that

$$\begin{aligned}\frac{|w+\omega|^2}{4}+\frac{1}{2}i\cdot\mathrm{Im}\,w^*\omega&=\frac{|w|^2}{4}+\frac{|\omega|^2}{4}+\frac{1}{2}\mathrm{Re}\,(w^*\omega)+\frac{1}{2}i\cdot\mathrm{Im}\,w^*\omega\\ &=\frac{|w|^2}{4}+\frac{|\omega|^2}{4}+\frac{1}{2}w^*\omega,\end{aligned}$$

we conclude that (3.44) and (3.45) give the same result. Thus, (3.42) gives us a realization of $\mathbb{H}$ as a group of affine ($q$-holomorphic) bijections of $\mathcal{U}_n$. We can identify the elements of $\mathcal{U}_n$ with the boundary via its action at the origin

$$h(0)=[w,t].(0,0)\mapsto(w,|w|^2+i\cdot t),$$

where $h=[w,t]$. Thus,

$$\mathbb{H}\ni[w,t]\mapsto(w,|w|^2+i\cdot t)\in\partial\mathcal{U}_n.$$

We may use the following coordinates $(q', t, r) = (q', t_1, \dots, t_{\dim V_2}, r)$ on $\mathcal{U}_n$:

$$\mathcal{U}_n \ni (q', q) = (q', t, r), \quad t_k = \mathrm{Im}_k q, \; k = 1, \dots, \beta, \quad r = r(q', q) = 4\mathrm{Re}\,(q) - |q'|^2.$$

If $4\mathrm{Re}\,(q) = |q'|^2$, then we get coordinates on the boundary $\partial\mathcal{U}_n$ of the Siegel upper half space

$$\partial\mathcal{U}_n \ni (q', q) = (q', t_1, \dots, t_{\dim V_2}),$$

where $t_k$ are as above and $r = r(q', q) = 0$.

### 3.4. Carnot groups

The following example includes connected simply connected Lie groups $\mathbb{G}$ whose Lie algebras are the direct sum of their subspaces

$$\mathfrak{g} = V_1 \oplus V_2 \oplus \dots V_m,$$

such that $[V_1, V_k] = V_{k+1}$, $k = 1, 2, \dots, m-1$, and $[V_1, V_m] = 0$. Since the commutators have finite length the algebras and the groups are nilpotent. The Lie algebras are also graded: $[V_l, V_k] \subseteq V_{l+k}$ and stratified

$$0 \in V_1 \subset V_1 \oplus V_2 \subset \dots \subset \oplus_{k=1}^m V_k.$$

Such kind of groups received the name the *Carnot groups* in literature.

**3.4.1. Two step Carnot groups.** The two step Carnot groups $\mathbb{G}$ are those possessing Lie algebras $\mathfrak{g}$ which are nilpotent of step 2, graded, stratified:

$$\mathfrak{g} = V_1 \oplus V_2, \qquad [V_1, V_1] \subseteq V_2, \quad [V_1, V_2] = [V_2, V_2] = \{0\}.$$

The group underlying manifold is $\mathbb{R}^{\alpha+\beta}$, $\alpha = \dim V_1$, $\beta = \dim V_2$. The group multiplication law can be written by making use of a $\mathbb{R}^\beta$-valued skew symmetric form $\Omega\colon \mathbb{R}^\alpha \times \mathbb{R}^\alpha \to \mathbb{R}^\beta$. Namely, if we write $(v_1, v_2), (v_1', v_2') \in V_1 \oplus V_2$ for the Lie algebra elements, then

$$\big[(v_1, v_2), (v_1', v_2')\big] = (0, \Omega(v_1, v_1')).$$

If we write $\tau = (x, t)$, $q = (x_1, t_1)$ for the elements of $\mathbb{G}$, then

$$\tau q = (x, t)(x_1, t_1) := (x + x_1, t + t_1 + \frac{1}{2}\Omega(x, x_1)) \tag{3.46}$$

by the BCH-formula. All $\mathbb{H}$-type groups and groups related to division algebras are examples of 2-step Carnot groups. Another treatment of 2 step nilpotent groups by making use of metric, see [43, 44].

**3.4.2. Engel group.** The Engel group is an example of a 3-step Carnot group. The underlying manifold is $\mathbb{R}^4$. We use coordinates $q = (x, y, z, w)$. Let us calculate the Lie group multiplication law by making use of the BCH-formula for a nilpotent group of step 3:

$$\begin{aligned}\exp(F_1)\exp(F_2) &= \\ &= \exp(F_1 + F_2 + \frac{1}{2}[F_1, F_2] + \frac{1}{12}[F_1, [F_1, F_2]] - \frac{1}{12}[F_2, [F_1, F_2]]).\end{aligned} \tag{3.47}$$

The Lie algebra for the Engel group has to satisfy the relations

$$[X,Y]=Z,\quad [X,Z]=aW,\quad [Y,Z]=bW,\quad a,b\in\mathbb{R}.$$

For example, if we choose a slight modification of the Heisenberg vector fields

$$X=\partial_x-\frac{1}{2}y\partial_z+z\partial_w,\quad Y=\partial_y+\frac{1}{2}x\partial_z-z\partial_w,$$

then we get

$$[X,Y]=\partial_z:=Z,\quad [X,Z]=-\partial_w:=W,\quad [Y,Z]=\partial_w=W. \tag{3.48}$$

If we write $F_i=x_iX+y_iY+z_iZ+w_iW$, $i=1,2$ then the BCH-formula (3.47) and the commutation relations (3.48) lead to the group law

$$\begin{aligned} &(x_1,y_1,z_1,w_1)(x_2,y_2,z_2,w_2)=\\ =\ &\big(x_1+x_2,y_1+y_2,z_1+z_2+\frac{1}{2}(x_1y_2-x_2y_1), \qquad (3.49)\\ &w_1+w_2-\frac{1}{2}(x_1z_2-x_2z_1)+\frac{1}{2}(y_1z_2-y_2z_1)+\frac{1}{12}(y_1-y_2)(x_1y_2-x_2y_1)\big). \end{aligned}$$

Another coordinate representation of the Engel group can be found in [37].

EXERCISES

1. Find the matrices of left invariant Riemannian metrics for groups related to division algebras. These metrics should also make the left invariant basis

   $$\{X_1,\dots,X_\alpha,T_1,\dots,T_\beta\},\quad \alpha=\dim U,\quad \beta=\dim V$$

   orthonormal.
2. Find gradients and sub-Laplacian operators.

## 4. Sub-Riemannian spheres

In this section we will consider sub-Riemannian manifolds whose underlying smooth manifolds are odd dimensional unit spheres. For the beginning we pay special attention to the spheres $S^3$ and $S^7$. We will see how the same sub-Riemannian structure is defined by considering $S^3$ as a group, as a CR-manifold, and as a principal $U(1)$-bundle. We also compare construction of sub-Riemannian structures on $S^3$ and $S^7$. The reason why we present the examples of $S^3$ and $S^7$ is that we can consider these spheres globally as manifolds endowed with a globally non-vanishing linearly independent basis. The structure of the basis on any sphere is given in the following theorem.

**Theorem 4.1.** [1] *Let* $S^{n-1}=\{x\in\mathbb{R}^n \mid \|x\|_E^2=1\}$ *be the unit sphere in* $\mathbb{R}^n$*, with respect to the usual Euclidean norm* $\|\cdot\|_E$*. Then* $S^{n-1}$ *has precisely* $\varrho(n)-1$ *linearly independent, globally defined and non-vanishing vector fields, where* $\varrho(n)$ *is defined in the following way: if* $n=(2a+1)2^b$ *and* $b=c+4d$*, where* $0\le c\le 3$*, then* $\varrho(n)=2^c+8d$*.*

In particular, two classical consequences follow: $S^1$, $S^3$ and $S^7$ are the only spheres with a maximal number of linearly independent globally defined non-vanishing vector fields, and all even dimensional spheres have no globally defined and non-vanishing vector fields. Rephrasing the property of a manifold $M$ to have maximal number of linearly independent globally defined non-vanishing vector fields one says that $M$ is *parallelizable*. The fact that $S^1$, $S^3$ and $S^7$ are the only parallelizable spheres was proved in [17]. Even dimensional spheres have no globally defined and non-vanishing vector fields which is a consequence of the Hopf index theorem, see [130].

### 4.1. Sub-Riemannian structures on $S^3$

**4.1.1. $S^3$ as a Lie group.** Consider the smooth manifold $S^3$:

$$S^3 = \{x = (x^0, x^1, x^2, x^3) \in \mathbb{R}^4 \mid \ \|x\|_E^2 = 1\}.$$

In order to introduce the multiplication between the point of $S^3$, we consider the set $S^3$ as a subset of the quaternion numbers $\mathbb{Q}$ of norm one. Recall that $\mathbb{Q} = (\mathbb{R}^4, +, \cdot)$, where $+$ stands for the usual coordinate-wise addition in $\mathbb{R}^4$ and "$\cdot$" is a non-commutative product g iven by the formula

$$\begin{aligned}(x^0 + \sum_{k=1}^{3} x^k i_k) \cdot (y^0 + \sum_{k=1}^{3} y^k i_k) \quad &= \quad (x^0y^0 - x^1y^1 - x^2y^2 - x^3y^3) \\ &+ \quad (x^1y^0 + x^0y^1 - x^3y^2 + x^2y^3)i_1 \\ &+ \quad (x^2y^0 + x^3y^1 + x^0y^2 - x^1y^3)i_2 \qquad (4.1) \\ &+ \quad (x^3y^0 - x^2y^1 + x^1y^2 + x^0y^3)i_3.\end{aligned}$$

The conjugate of $q = (x^0 + \sum_{k=1}^{3} x^k i_k)$, is given by $\bar{q} = (x^0 - \sum_{k=1}^{3} x^k i_k)$ and the norm $|q|$ of $q \in \mathbb{Q}$ is defined by $|q|^2 = q\bar{q}$. The realization of the sphere $S^3$ as the set of unit quaternions with the multiplication (4.1), gives the Lie group $S^3 = (S^3, \cdot)$.

The multiplication rule (4.1) induces a right translation $r_\tau(q)$ of an element $q = x^0 + \sum_{k=1}^{3} x^k i_k$ by the element $\tau = y^0 + \sum_{k=1}^{3} y^k i_k$. The matrix corresponding to the tangent map $dr_\tau(q)$, obtained by the multiplication rule, becomes

$$dr_\tau = \begin{pmatrix} y^0 & y^1 & y^2 & y^3 \\ -y^1 & y^0 & -y^3 & y^2 \\ -y^2 & y^3 & y^0 & -y^1 \\ -y^3 & -y^2 & y^1 & y^0 \end{pmatrix}.$$

Calculating the action of $dr_\tau(q)$ on the basis of the unit vectors $(\partial_0, \partial_1, \partial_2, \partial_3)$, we get four vector fields

$$\begin{aligned} N_\tau &= y^0\partial_0 + y^1\partial_1 + y^2\partial_2 + y^3\partial_3, & V_\tau &= -y^1\partial_0 + y^0\partial_1 - y^3\partial_2 + y^2\partial_3, \\ X_\tau &= -y^2\partial_0 + y^3\partial_1 + y^0\partial_2 - y^1\partial_3, & Y_\tau &= -y^3\partial_0 - y^2\partial_1 + y^1\partial_2 + y^0\partial_3. \end{aligned} \tag{4.2}$$

It is easy to see that $N_\tau$ is the unit normal to $S^3$ at $\tau \in S^3$ with respect to the Euclidean inner product $(\cdot\,,\cdot)$ in $\mathbb{R}^4$. Moreover, for any $\tau \in S^3$

$$(N_\tau, V_\tau) = (N_\tau, X_\tau) = (N_\tau, Y_\tau) = 0$$

and

$$(N_\tau, N_\tau) = (V_\tau, V_\tau) = (X_\tau, X_\tau) = (Y_\tau, Y_\tau) = 1.$$

Since the matrix

$$\begin{pmatrix} -y^1 & y^0 & -y^3 & y^2 \\ -y^2 & y^3 & y^0 & -y^1 \\ -y^3 & -y^2 & y^1 & y^0 \end{pmatrix}$$

has rank three, we conclude that the vector fields $\{V(\tau), X(\tau), Y(\tau)\}$ form an orthonormal basis of $T_\tau S^3$ with respect to $(\cdot\,,\cdot)_\tau$, for any $\tau \in S^3$.

Observing that $[X, Y] = 2V$, we see that the distribution $D = \operatorname{span}\{X, Y\}$ is bracket generating, strongly bracket generating, and regular, therefore it satisfies the hypotheses of Theorem 2.2 and Theorem 2.3. Notice that the distribution $D = \operatorname{span}\{X, Y\}$ can also be defined as the kernel of the contact one form

$$\omega = -y_1\, dy^0 + y_0\, dy^1 - y_3\, dy^2 + y_2\, dy^3. \tag{4.3}$$

**Remark 2.** It is easy to see that $[V, Y] = 2X$ and $[X, V] = 2Y$, therefore the distributions $\operatorname{span}\{Y, V\}$ and $\operatorname{span}\{X, V\}$ are also bracket generating. The corresponding contact forms are

$$\theta = -y_2\, dy^0 + y_3\, dy^1 + y_0\, dy^2 - y_1\, dy^3, \qquad \eta = -y_3\, dy^0 - y_2\, dy^1 + y_1\, dy^2 + y_0\, dy^3,$$

respectively. This means that there is a priori no natural choice of a sub-Riemannian structure on $S^3$ generated by the Lie group action of multiplication of quaternions. Any choice that can be made, will produce essentially the same geometry.

Exercises

1. See that the constructed group $(S^3, \cdot)$ coincides with the matrix group $Sp(1)$.
2. Show that the constructed group $(S^3, \cdot)$ is isomorphic to the group $SU(2)$ of matrices

$$\begin{pmatrix} z_1 & z_2 \\ -\bar{z}_2 & \bar{z}_1 \end{pmatrix}, \qquad |z_1|^2 + |z_2|^2 = 1, \quad z_1, z_2 \in \mathbb{C}.$$

Use the correspondence

$$q = x^0 + x^1 i_1 + x^2 i_2 + x^3 i_3 \quad \leftrightarrow \quad z_1 = x^0 + ix^1,\ z_2 = x^2 + ix^3.$$

**4.1.2. $S^3$ as a CR manifold.** Consider $S^3$ as the boundary of the unit ball $B_{\mathbb{C}} \in \mathbb{C}^2$, or the hypersurface

$$S^3 := \{(z, w) \in \mathbb{C}^2 : z\bar{z} + w\bar{w} = 1\}.$$

The sphere $S^3$ cannot be endowed with a complex structure since it has 3 dimensional tangent space. Nevertheless it possesses a differentiable structure compatible with the natural complex structure of the ball $B_{\mathbb{C}} = \{(z, w) \in \mathbb{C}^2 : z\bar{z} + w\bar{w} < 1\}$ as an open set in $\mathbb{C}^2$. We show that this differentiable structure over the sphere $S^3$, called CR structure, is equivalent to the sub-Riemannian one considered in the previous subsection. We begin by recalling the definition of a CR structure, according to [14].

In the case $W = T_q\mathbb{R}^{2n}$, $q = (x^1, y^1, \dots, x^n, y^n) \in \mathbb{R}^{2n}$, we say that the *standard almost complex structure* for $W$ is defined by setting

$$J(\partial_{x^j}) = \partial_{y^j}, \qquad J(\partial_{y^j}) = -\partial_{x^j}, \qquad 1 \le j \le n.$$

For a smooth real submanifold $M$ of $\mathbb{C}^n$ and a point $q \in M$, in general, the tangent space $T_qM$ is not invariant under the standard almost complex structure map $J\colon T_q\mathbb{C}^n \to T_q\mathbb{C}^n$, $T_q\mathbb{C}^n \cong T_q\mathbb{R}^{2n}$. We are interested in the largest subspace invariant under the action of $J$.

**Definition 18.** *The holomorphic tangent space $H_qM$ of $M$ at $q$ is the vector space*

$$H_qM = T_qM \cap J(T_qM) \quad \textit{for a point} \quad q \in M.$$

A real submanifold $M$ of $\mathbb{C}^n$ is said to have a CR structure if $\dim_{\mathbb{R}} H_qM$ does not depend on $q \in M$. A result of [14] implies that every smooth real hypersurface $S$ embedded in $\mathbb{C}^n$ satisfies $\dim_{\mathbb{R}} H_qS = 2n-2$, therefore, $S$ is a CR manifold. This fact applies to every odd dimensional sphere, considered as an embedded manifold to $\mathbb{C}^n$.

Let us describe the holomorphic tangent space $H_qS^3$. The space $H_qS^3$ can be seen as a complex vector space of complex dimension one. This description is achieved by considering the differential form $\omega = \bar{z}dz + \bar{w}dw$ and observing that $\ker(\omega)$ is precisely the set we are looking for. Straightforward calculations show that

$$\ker(\omega) = \operatorname{span}\{\bar{w}\partial_z - \bar{z}\partial_w\}.$$

In real coordinates this corresponds to

$$\bar{w}\partial_z - \bar{z}\partial_w = \frac{1}{2}(-X + iY),$$

where $X$ and $Y$ were defined in (4.2). It is important to remark that this is precisely the maximal 1-complex-dimensional $J$-invariant subspace of $T_qS^3$, namely

$$J(X) = Y, \qquad J(Y) = -X.$$

Then $J(\operatorname{span}\{X, Y\}) = \operatorname{span}\{X, Y\}$, but $J(V) = -N \notin T_qS^3$ for any point $q \in S^3$. Therefore, the right invariant distribution corresponding to the left action of $S^3$ over itself coincides with the 1-complex-dimensional holomorphic tangent space.

**Remark 3.** Essentially the same almost complex structure can be obtained by means of the Levi-Civita connection $\nabla$ on $S^3$ considered as a smooth Riemannian manifold embedded into $\mathbb{R}^4$. Namely, in [74] it is introduced the mapping $J_V(W) = \nabla_W V$ for $W \in D$, and the vector field $V$ defined in (4.2).

EXERCISE

1. Show that the distribution $D = HS^3$ at $q \in S^3$ can be also defined as a set of complex 2-dimensional vectors that are orthogonal to $n = z\partial_z + w\partial_w$ with respect to standard Hermitian product $(v, n)_H = \bar{v}_1 n_1 + \bar{v}_2 n_2$ at each point $q \in S^3$:

$$D_q = \{v \in T_q\mathbb{C}^2 \mid (v, n)_H = 0, \quad n = z\partial_z + w\partial_w, \quad z, w \in \mathbb{C}\}.$$

**4.1.3. $S^3$ as a principal $U(1)$-bundle.** In this part we describe how the structure of a principal $U(1)$-bundle over $S^3$ induces a bracket generating distribution on $S^3$. More details about the relation between principal bundle and sub-Riemannian geometry will be given in Section 5. The group $U(1)$, consisting of complex numbers of absolute value 1, acts on the right on the manifold $S^3$ by

$$\begin{array}{cccc} \mu\colon & S^3 \times U(1) & \to & S^3 \\ & (z, w).v & \mapsto & (zv, wv). \end{array}$$

Here $v \in U(1) = \{v \in \mathbb{C} : |v|^2 = 1\}$ and $(z, w) \in S^3 \subset \mathbb{C}^2$.

Consider the Hopf map $h : S^3 \to S^2$ [71, 98], given explicitly by

$$h(z, w) = (|z|^2 - |w|^2, 2z\bar{w}),$$

where $S^2 = \{(x, \zeta) \in \mathbb{R} \times \mathbb{C} : x^2 + |\zeta|^2 = 1\}$. Clearly, $h$ is a submersion of $S^3$ onto $S^2$, and it is a bijection between $S^3/U(1)$ and $S^2$, where $S^3/U(1)$ is understood as the orbit space of the $U(1)$-action over $S^3$.

Let $p = (x_0, \zeta_0) \in S^2$. Consider the great circle

$$\gamma_p(s) = (z_0, w_0)e^{2\pi i s}, \qquad s \in [0, 1],$$

in $S^3$, that projects to $p$ under the Hopf map. Here $(z_0, w_0)$ is a point in the pre-image of $p$ under $h$. Consider the tangent vector field, defined by

$$\dot{\gamma}_p(s) = 2\pi i(z_0, w_0)e^{2\pi i s} \in T_{\gamma_p(s)}S^3.$$

We write the curve $\gamma_p$ and the map $d_{\gamma_p(s)}h$ in real coordinates. Then

$$\begin{aligned} \gamma_p(s) = (z(s), w(s)) &= (x^0(s) + ix^1(s), x^2(s) + ix^3(s)) \\ &= (x^0(s), x^1(s), x^2(s), x^3(s)) \end{aligned}$$

and

$$d_{\gamma_p(s)}h = 2\begin{pmatrix} x^0(s) & x^1(s) & -x^2(s) & -x^3(s) \\ x^2(s) & x^3(s) & x^0(s) & x^1(s) \\ -x^3(s) & x^2(s) & x^1(s) & -x^0(s) \end{pmatrix}. \qquad (4.4)$$

Thus, the Hopf map induces the following action over the vector field $\dot{\gamma}_p(s) = i\gamma_p(s)$:

$$d_{\gamma_p(s)}h\big(\dot{\gamma}_p(s)\big) = 2\begin{pmatrix} x^0(s) & x^1(s) & -x^2(s) & -x^3(s) \\ x^2(s) & x^3(s) & x^0(s) & x^1(s) \\ -x^3(s) & x^2(s) & x^1(s) & -x^0(s) \end{pmatrix}\begin{pmatrix} \dot{x}^0(s) \\ \dot{x}^1(s) \\ \dot{x}^2(s) \\ \dot{x}^3(s) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Therefore, if $d_{\gamma_p(s)}h$ is a full rank matrix, then we would have characterized the kernel of it, by

$$\ker\left(d_{\gamma_p(s)}h\right) = \operatorname{span}\{\dot\gamma_p(s)\} = \operatorname{span}\{V_{\gamma_p(s)}\}, \tag{4.5}$$

since $\dot\gamma_p(s) = 2\pi V_{\gamma_p(s)}$ by (4.2). In order to see that the matrix (4.4) is full rank we observe that $[d_{\gamma_p(s)}h][d_{\gamma_p(s)}h]^{tr} = 4I_3$, where $I_3$ denotes the identity $(3\times 3)$-matrix. This implies that $d_{\gamma_p(s)}h$ is full rank.

Now we describe how the Hopf map induces the distribution $D$ constructed in the previous subsections. Define the distribution $D$ as the orthogonal complement to the kernel of $dh$ with respect to the Euclidean inner product $(\cdot\,,\cdot)$ in $\mathbb{R}^4$. More precisely,

$$D_q = \{v \in T_qS^3 \mid (v,w) = 0 \quad \forall \quad w \in \ker(d_qh)\}.$$

Since we know that $\ker(d_qh) = \operatorname{span}\{V(q)\}$, and moreover,

$$(X_q, V_q) = (Y_q, V_q) = (X_q, Y_q) = 0,$$

we see that

$$D_q = \operatorname{span}\{X_q, Y_q\}. \tag{4.6}$$

In the literature the distribution obtained by this way is called the *Ehresmann connection* and $\ker(d_qh)$ is called the *vertical space*. We give a general definition of the Ehresmann connection in Subsection 5.1.

The action of the group $U(1)$ on the manifold $S^3$ satisfies the definition of the principal bundle, see Definition 21. We conclude that the Hopf fibration is a principal $U(1)$-bundle. Moreover the distribution $D$ is invariant under the right action of $U(1)$:

$$d_qr_\tau(D_q) = D_{r_\tau(q)} = D_{q.\tau}, \quad \tau \in U(1), \quad q \in S^3.$$

Thus, the Hopf map, written in coordinates, indicates, in a topological way, how one makes a natural choice of the horizontal distribution $D$ that was not obvious when we considered the left action of $S^3$ over itself.

The sub-Riemannian metric $g_D$ is defined by restricting the usual Riemannian metric on $S^3$ to the distribution $D$. Summarizing the last three subsections we conclude that all presented constructions lead to the sub-Riemannian manifold $(S^3, D, g_D)$, where

$$D_q = \operatorname{span}\{X_q, Y_q\}, \qquad q = (x^0, x^1, x^2, x^3),$$

$$X_q = -x^2\partial_{x^0} + x^3\partial_{x^1} + x^0\partial_{x^2} - x^1\partial_{x^3}, \qquad Y_q = -x^3\partial_{x^0} - x^2\partial_{x^1} + x^1\partial_{x^2} + x^0\partial_{x^3},$$

and the sub-Riemannian metric $g_D = (\cdot\,,\cdot)|_D$ is the restriction of the Euclidean inner product $(\cdot\,,\cdot)$ in $\mathbb{R}^4$ to the distribution $D$.

Let us find an analogue of the horizontality condition (3.7) for $S^3$. A smooth curve $c\colon I \to S^3$ is horizontal if $\dot c(s) \in D_{c(s)}$ for all $s \in I$ or if the third coordinate in the decomposition

$$\dot c = \alpha(s)X(c(s)) + \beta(s)Y(c(s)) + \delta(s)V(c(s))$$

vanishes. Write $c(s) = \Big(x^0(s), x^1(s), x^3(s), x^4(s)\Big)$ and $\dot{c}(s) = \sum_{k=0}^{4} x^k \partial_k$. Then, since

$$\delta(s) = \big(T(c(s)), \dot{c}(s)\big) = -x^1\dot{x}^0 + x^0\dot{x}^1 - x^3\dot{x}^2 + x^2\dot{x}^3$$

by (4.2), we conclude that the curve is horizontal if it satisfies the differential equation

$$-x^1(s)\dot{x}^0(s) + x^0(s)\dot{x}^1(s) - x^3(s)\dot{x}^2(s) + x^2(s)\dot{x}^3(s) = 0, \quad s \in I. \tag{4.7}$$

It is a reformulation of the condition $\dot{c} \in \ker(\omega)$, where $\omega$ is the one-form from (4.3): $\omega = -x^1dx^0 + x^0dx^1 - x^3dx^2 + x^2dx^3$. This form can be written as $\omega = dA_{01} - dA_{32}$, where $dA_{01} = x^0dx^1 - x^1dx^0$, $dA_{32} = x^3dx^2 - x^2dx^3$ are the area forms on the planes $(x^0, x^1)$ and $(x^3, x^2)$, up to the factor 1/2, respectively. Let us denote by $A_{01}$ the area swept by the projection of the curve $c$ onto $(x^0, x^1)$-plane, and by $A_{32}$ the area swept by the projection of the curve $c$ onto $(x^3, x^2)$-plane. Then the curve $c$ is horizontal if and only if $A_{01} = A_{32}$, see Figure 4.1. Compare this with the isoperimetric property of a horizontal curve on the Heisenberg group $\mathbb{H}^1$. Observe that the considered vector fields $X, Y, V$ are right invariant vector fields produced by the left action of $S^3$ on itself. The right action of the group $S^3$ on itself leads to left invariant vector fields on $S^3$. This phenomenon is general for the action of groups, see Appendix A.

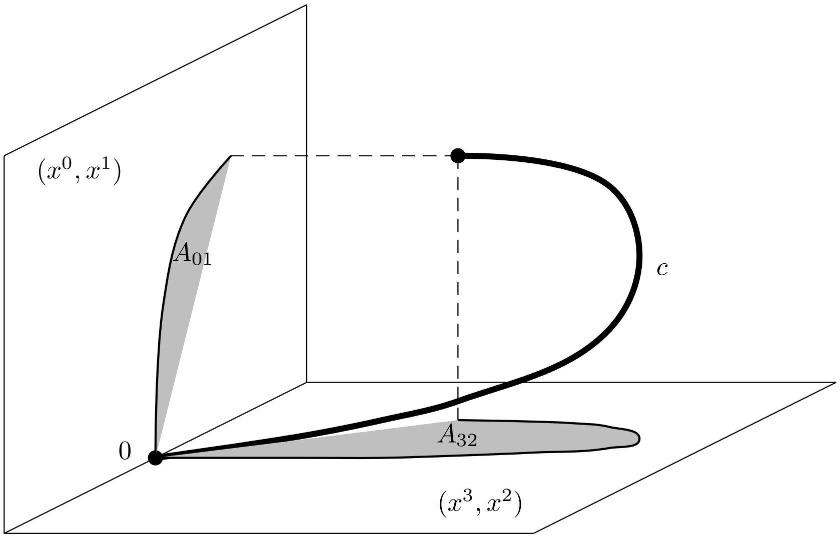

FIGURE 4.1. Projections of $c$ to the planes $(x_0, x_1)$ and $(x_3, x_2)$.

The geodesics can be found by making use of the Hamiltonian approach as in the case of the Heisenberg group, but the Hamiltonian equations in this case are much more difficult. This method was exploited in [28, 29], where the sub-Riemannian structure was defined by left invariant vector fields.

The authors of [74] showed that the result of Theorem 3.1 remains true for the sub-Riemannian manifold $S^3$, where they used the complex structure described in Remark 3. They considered the horizontal distribution defined by the right invariant vector fields on $S^3$.

We will present a different method to find sub-Riemannian geodesics on $S^3$. This method is valid for all odd dimensional spheres and even for all principal bundles with the appropriate choice of metrics. It will be one of the main points of consideration in Section 5.

## 4.2. Sub-Riemannian structures on $S^7$

**4.2.1. Tangent vector fields for $S^7$.** In this section we obtain two structurally different types of horizontal distributions on $S^7$. One of them is of rank 6 and the other is of rank 4. We start from the construction of a convenient basis of tangent vector fields on the sphere $S^7$.

The multiplication of unit octonions is not associative, therefore $S^7$ is not a group in contrast with $S^3$. Nevertheless, we are still able to use the multiplication law in order to find global non-vanishing tangent vector fields. In calculations we use a slightly different multiplication table of unit octonions from what we considered in Subsection 3.3 and that leads to a different product. It is more convenient for our purpose in this subsection. Both the multiplication table of unit octonions and the product of two arbitrary octonions are presented in Appendix B. The multiplication rule induces a matrix representation of the right octonion multiplication, given explicitly by:

$$dr_\tau = \begin{pmatrix} y^0 & -y^1 & -y^2 & -y^3 & -y^4 & -y^5 & -y^6 & -y^7 \\ y^1 & y^0 & y^3 & -y^2 & y^5 & -y^4 & -y^7 & y^6 \\ y^2 & -y^3 & y^0 & y^1 & y^6 & y^7 & -y^4 & -y^5 \\ y^3 & y^2 & -y^1 & y^0 & y^7 & -y^6 & y^5 & -y^4 \\ y^4 & -y^5 & -y^6 & -y^7 & y^0 & y^1 & y^2 & y^3 \\ y^5 & y^4 & -y^7 & y^6 & -y^1 & y^0 & -y^3 & y^2 \\ y^6 & y^7 & y^4 & -y^5 & -y^2 & y^3 & y^0 & -y^1 \\ y^7 & -y^6 & y^5 & y^4 & -y^3 & -y^2 & y^1 & y^0 \end{pmatrix}.$$

for $\tau = y^0 + \sum_{k=1}^{7} y^k j_k$.

We are able to find globally defined tangent vector fields which are invariant under the right multiplication rule. We proceed by analogy with the constructions made for $S^3$. The explicit formulas of vector fields are given in Appendix B. The vector fields $\{Y_1, \dots, Y_7\}$ form a frame for $TS^7$ and $Y_0$ is the normal to $S^7$. More explicitly

$$\big(Y_i(\tau), Y_j(\tau)\big)_\tau = \delta_{ij}, \qquad \tau \in S^7, \qquad i, j \in \{0, 1, \dots, 7\},$$

where $(\cdot\,,\cdot)$ is the standard inner product in $\mathbb{R}^8$, and $\delta_{ij}$ stands for Kronecker's delta.

**4.2.2. CR structure and the Hopf map on $S^{2n+1}$.** Before we go further in studying structures on $S^7$, we present general relations between the CR structures on odd dimensional spheres and the higher dimensional Hopf fibration.

Consider $S^{2n+1} = \{z \in \mathbb{C}^{n+1} \mid \|z\|^2 = 1\}$. Then the right $U(1)$-action on $S^{2n+1}$ given by

$$(z_0, \dots, z_n).\upsilon = (z_0\upsilon, \dots, z_n\upsilon),$$

for $\upsilon \in U(1)$ and $(z_0, \dots, z_n) \in S^{2n+1}$, induces the principal $U(1)$-bundle $U(1) \to S^{2n+1} \xrightarrow{h} \mathbb{C}P^n$ given explicitly by

$$S^{2n+1} \ni (z_0, \dots, z_n) \mapsto h(z_0, \dots, z_n) = [z_0 : \dots : z_n] \in \mathbb{C}P^n,$$

where $[z_0 : \dots : z_n]$ denotes homogeneous coordinates. This map is called *higher Hopf fibration*. The kernel of the map $h\colon S^{2n+1} \to \mathbb{C}P^n$ gives the vertical space at each point of $S^{2n+1}$. The horizontal distribution or the Ehresmann connection $D$ is given by the orthogonal complement to the vertical distribution $V$ with respect to the inner product of $\mathbb{R}^{2n+2}$. We show that the vertical space is always given by the action of standard almost complex structure in $\mathbb{C}^{n+1}$ on the normal vector field to $S^{2n+1}$, and the Ehresmann connection coincides with the holomorphic tangent space at each point of $S^{2n+1}$.

Theorem 4.1 asserts that any odd dimensional sphere has at least one globally defined non-vanishing tangent vector field. If the dimension of the sphere is of the form $4n+1$, then it has only one globally defined non-vanishing tangent vector field. If the dimension of the sphere is of the form $4n+3$, then the sphere admits at least three globally defined non-vanishing vector fields. Any sphere $S^{2n+1}$ possesses the vector field

$$V = -y^1\partial_0 + y^0\partial_1 - y^3\partial_2 + \ldots - y^{2n+2}\partial_{2n+1} + y^{2n+1}\partial_{2n+2}. \tag{4.8}$$

Observe that this vector field has appeared already in two cases: as the vector field $V$ for $S^3$, and as the vector field $Y_1$ for $S^7$.

The vector field $V$ encloses valuable information concerning the CR structure of $S^{2n+1}$. A result of [14] states that the sphere $S^{2n+1}$, as a smooth hypersurface in $\mathbb{C}^{n+1}$, admits a holomorphic tangent space of dimension $\dim_{\mathbb{R}}(H_qS^{2n+1}) = 2n$ for any point $q \in S^{2n+1}$. The following lemma implies a description of the holomorphic tangent space $H_qS^{2n+1}$ as the orthogonal complement to $V$.

**Lemma 1.** *Let $W$ be an Euclidean space of dimension $k+2$, $k \geq 1$, with an inner product $(\cdot\,,\cdot)_W$ and let $X, Y$ be two vectors from $W$. Consider an orthogonal decomposition $W = \operatorname{span}\{X, Y\} \oplus_\perp \widetilde{W}$ with respect to $(\cdot\,,\cdot)_W$ and an orthogonal endomorphism $A : W \to W$ such that*

$$A(\operatorname{span}\{X, Y\}) = \operatorname{span}\{X, Y\},$$

*then* $\widetilde{W}$ *is an invariant space under the action of* $A$*, i.e.*

$$A(\widetilde{W}) = \widetilde{W}.$$

*Proof.* Let $v \in \widetilde{W}$, then for any $\alpha, \beta \in \mathbb{R}$ it is clear that

$$\big(Av, \alpha X + \beta Y\big)_W = \big(v, A^{tr}(\alpha X + \beta Y)\big)_W = \big(v, A^{-1}(\alpha X + \beta Y)\big)_W.$$

Since $A(\operatorname{span}\{X, Y\}) = \operatorname{span}\{X, Y\}$, there exist $a, b \in \mathbb{R}$ such that

$$A^{-1}(\alpha X + \beta Y) = aX + bY,$$

and therefore, $\big(Av, \alpha X + \beta Y\big)_W = \big(v, aX + bY\big)_W = 0$, which implies $Av \in \widetilde{W}$. □

As an application of Lemma 1, it is possible to obtain an explicit characterization of the space $H_qS^{2n+1}$.

**Lemma 2.** *The vector space* $H_qS^{2n+1}$ *is the orthogonal complement to the vector* $V_q \in T_qS^{2n+1}$ *from* (4.8) *for any* $q \in S^{2n+1}$.

*Proof.* Consider the vector space

$$W_q = \operatorname{span}\{N(q)\} \oplus_\perp T_qS^{2n+1} \cong T_q\mathbb{R}^{2n+2},$$

where $N(q)$ is the normal vector to $S^{2n+1}$ at the point $q$. The standard almost complex structure map $J \colon W_q \to W_q$ is orthogonal. Moreover, $J(V(q)) = -N(q)$, $J(N(q)) = V(q)$. Using the decomposition $W_q = \widetilde{W_q} \oplus_\perp \operatorname{span}\{V(q), N(q)\}$, it is possible to apply Lemma 1 in order to conclude that $\widetilde{W_q}$, which is the orthogonal complement to $V(q)$ in $T_qS^{2n+1}$, is invariant under $J$. Since $\dim_{\mathbb{R}} \widetilde{W_q} = 2n$, we conclude that $\widetilde{W_q} = H_qS^{2n+1}$. □

The space $HS^{2n+1}$ can also be described as the kernel of the one-form

$$\theta = \bar{z}_0 dz_0 + \ldots + \bar{z}_n dz_n.$$

Indeed, consider $X \in HS^{2n+1}$. Then by straightforward calculations we have

$$\theta(X) = (X, N) + i(X, V) = 0. \tag{4.9}$$

Lemma 2 provides a horizontal distribution of rank $2n$ for the spheres $S^{2n+1}$, by considering the holomorphic tangent bundle: $D = HS^{2n+1}$. The bracket generating property follows from the following general result for an arbitrary contact manifold.

**Definition 19.** *Let* $M$ *be a* $(2n + 1)$*-dimensional manifold. A smooth one form* $\omega$ *is called contact if it satisfies the condition*

$$\omega_q \wedge (d\omega_q)^n \neq 0 \quad \textit{for any} \quad q \in M.$$

*The pair* $(M, \omega)$ *is called a contact manifold.*

**Lemma 3.** *Let* $M$ *be a* $(2n + 1)$*-dimensional contact manifold with contact form* $\omega$*, then* $D = \ker(\omega)$ *is a bracket generating distribution of rank* $2n$ *and step* $2$.

*Proof.* Recall Cartan's formula for a differential one-form $\omega$, namely

$$d\omega(X,Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X,Y]), \tag{4.10}$$

for all $X, Y \in TM$. It follows from (4.10) that $D$ is Frobenius integrable if and only if $d\omega(X,Y) = 0$ for all $X, Y \in D$. Thus, if $\omega$ is a contact form, then $d\omega(X,Y) \neq 0$ for all $X, Y \in TM$ and, therefore $D$ is not integrable. This implies the bracket generating property for $D$, since if $[X,Y]_q \notin D_q$ at any point $q \in M$ for some $X_q, Y_q \in D_q$ then $\operatorname{span}\{[X,Y]_q\} \oplus D_q = T_qM$. □

By Lemma 3, in order to prove that $HS^{2n+1}$ is bracket generating, it is sufficient to find a contact one-form $\omega$ such that $HS^{2n+1} = \ker(\omega)$. To achieve this, let us consider

$$\omega = \operatorname{Im}\theta = -y_1dy^0 + y_0dy^1 - \ldots - y_{2n+1}dy^{2n} + y_{2n}dy^{2n+1} \tag{4.11}$$

defined on $S^{2n+1}$. By (4.9), the relation $HS^{2n+1} = \ker(\omega)$ holds immediately.

**Theorem 4.2.** [57] *The one-form $\omega$ defined in* (4.11) *is a contact form. More specifically, $\omega$ satisfies $(d\omega)^n \wedge \omega = n! \cdot 2^n \mathrm{dvol}_{S^{2n+1}}$, where $\mathrm{dvol}_{S^{2n+1}}$ is the volume form for $S^{2n+1}$.*

The following corollary holds by Lemma 3 and Theorem 4.2.

**Corollary 1.** *The holomorphic tangent bundle $HS^{2n+1}$ is a bracket generating distribution of step* 2 *and rank* $2n$.

An important consequence of Theorem 4.2 follows by considering a classical result by G. Darboux, see [38]. In modern terms, this theorem asserts that every $(2n+1)-$dimensional contact manifold is locally the $n-$dimensional Heisenberg group. This means precisely that the tangent cone of $S^{2n+1}$, as a sub-Riemannian manifold with distribution $HS^{2n+1}$ and metric induced by the usual Euclidean metric in $\mathbb{R}^{2n+2}$, is isomorphic to the $n-$dimensional Heisenberg group. See [13, 63] for the definition of the tangent cone to a sub-Riemannian manifold.

It is necessary to remark, that in general, there is no globally defined basis for $HS^{2n+1}$. By Theorem 4.1, this is only possible for $S^3$ and $S^7$. A basis for the distribution in the case of $S^3$ has already been discussed. An explicit proof that shows the bracket generating property of the basis of $HS^7$ can be found in [9, 10, 57].

We conclude this section by proving that the line $\operatorname{span}\{V\}$ from (4.8) forms the kernel of $dh$, where $h$ is the Hopf fibration $U(1) \to S^{2n+1} \xrightarrow{h} \mathbb{C}P^n$. The orthogonal complement to $V$ is the horizontal distribution $D = HS^{2n+1}$. To achieve this, we recall that the charts defining the holomorphic structure of $\mathbb{C}P^n$ are given by the open sets

$$U_k = \{[z_0 : \cdots : z_n] : z_k \neq 0\},$$

together with the homeomorphisms

$$\begin{array}{cccc} \varphi_k & : & U_k & \to & \mathbb{C}^n \\ & & [z_0 : \ldots : z_n] & \mapsto & (\frac{z_0}{z_k}, \ldots, \frac{z_{k-1}}{z_k}, \frac{z_{k+1}}{z_k}, \cdots, \frac{z_n}{z_k}). \end{array}$$

Then, without loss of generality we assume that $n = 3$ and perform explicit calculations for $k = 0$. Other cases can be treated similarly.

Using the chart $(U_0, \varphi_0)$ defined above, we have the map

$$\begin{array}{rccc} \varphi_0 \circ h & : & S^7 & \to & \mathbb{C}^3 \\ & & (z_0, z_1, z_2, z_3) & \mapsto & \left(\frac{z_1}{z_0}, \frac{z_2}{z_0}, \frac{z_3}{z_0}\right), \end{array}$$

which in real coordinates can be written as

$$\begin{aligned} \varphi_0 \circ h(x_0, \dots, x_7) \quad = \quad & \Bigg( \frac{x_0x_2 + x_1x_3}{x_0^2 + x_1^2}, \frac{x_0x_3 - x_1x_2}{x_0^2 + x_1^2}, \frac{x_0x_4 + x_1x_5}{x_0^2 + x_1^2}, \\ & \frac{x_0x_5 - x_1x_4}{x_0^2 + x_1^2}, \frac{x_0x_6 + x_1x_7}{x_0^2 + x_1^2}, \frac{x_0x_7 - x_1x_6}{x_0^2 + x_1^2} \Bigg). \end{aligned}$$

The differential of this mapping is given by the matrix

$$d(\varphi_0 \circ h) = \big(A, B\big) \in \mathbb{R}^{6\times 8}$$

where $A \in \mathbb{R}^{6\times 2}$ and $B \in \mathbb{R}^{6\times 6}$ have the following forms

$$A = \begin{pmatrix} \frac{(x_1^2-x_0^2)x_2-2x_0x_1x_3}{(x_0^2+x_1^2)^2} & \frac{(x_0^2-x_1^2)x_3-2x_0x_1x_2}{(x_0^2+x_1^2)^2} \\ \frac{(x_1^2-x_0^2)x_3+2x_0x_1x_2}{(x_0^2+x_1^2)^2} & \frac{(x_1^2-x_0^2)x_2-2x_0x_1x_3}{(x_0^2+x_1^2)^2} \\ \frac{(x_1^2-x_0^2)x_4-2x_0x_1x_5}{(x_0^2+x_1^2)^2} & \frac{(x_0^2-x_1^2)x_5-2x_0x_1x_4}{(x_0^2+x_1^2)^2} \\ \frac{(x_1^2-x_0^2)x_5+2x_0x_1x_4}{(x_0^2+x_1^2)^2} & \frac{(x_1^2-x_0^2)x_4-2x_0x_1x_5}{(x_0^2+x_1^2)^2} \\ \frac{(x_1^2-x_0^2)x_6-2x_0x_1x_7}{(x_0^2+x_1^2)^2} & \frac{(x_0^2-x_1^2)x_7-2x_0x_1x_6}{(x_0^2+x_1^2)^2} \\ \frac{(x_1^2-x_0^2)x_7+2x_0x_1x_6}{(x_0^2+x_1^2)^2} & \frac{(x_1^2-x_0^2)x_6-2x_0x_1x_7}{(x_0^2+x_1^2)^2} \end{pmatrix}$$

and

$$B = \begin{pmatrix} \frac{x_0}{x_0^2+x_1^2} & \frac{x_1}{x_0^2+x_1^2} & 0 & 0 & 0 & 0 \\ -\frac{x_1}{x_0^2+x_1^2} & \frac{x_0}{x_0^2+x_1^2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{x_0}{x_0^2+x_1^2} & \frac{x_1}{x_0^2+x_1^2} & 0 & 0 \\ 0 & 0 & -\frac{x_1}{x_0^2+x_1^2} & \frac{x_0}{x_0^2+x_1^2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{x_0}{x_0^2+x_1^2} & \frac{x_1}{x_0^2+x_1^2} \\ 0 & 0 & 0 & 0 & -\frac{x_1}{x_0^2+x_1^2} & \frac{x_0}{x_0^2+x_1^2} \end{pmatrix}.$$

By straightforward calculations, we know that

$$\det([d(\varphi_0 \circ h)][d(\varphi_0 \circ h)]^{tr}) = (x_0^2 + x_1^2)^{-8} = |z_0|^{-16} \neq 0,$$

therefore, the matrix $d(\varphi_0 \circ h)$ has rank 6 or equivalently $\dim_{\mathbb{R}} \big(\ker d(\varphi_0 \circ h)\big) = 2$. Moreover, since $d(\varphi_0 \circ h)(N) = d(\varphi_0 \circ h)(V) = 0$, by direct calculations, we conclude

$$\ker \big(d(\varphi_0 \circ h)\big) = \operatorname{span}\{N, V\}.$$

This implies $\ker(dh) = \operatorname{span}\{V\}$.

**4.2.3. Application of the first quaternionic Hopf map.** We wish to find with the help of the quaternionic Hopf bundle $S^3 \to S^7 \to S^4$ a natural choice of horizontal distributions of rank 4 on $S^7$. The right action of $S^3$ on $S^7$ is defined in the following way. Represent any point of $S^7$ as a pair of quaternions $(q_1, q_2)$ of the norm $|(q_1, q_2)| = 1$. Any point in $S^3$ is also a quaternion $\upsilon \in S^3$ of unit norm. Then the right action is

$$\begin{array}{rccc} \mu\colon & S^7 \times S^3 & \to & S^7 \\ & (q_1, q_2).\upsilon & \mapsto & (q_1\upsilon, q_2\upsilon), \end{array}$$

where $q_1\upsilon$ and $q_2\upsilon$ are the usual products of quaternions. We consider the quaternionic Hopf map given by

$$\begin{array}{rcccc} h & : & S^7 & \to & S^4 \\ & & (z, w) & \mapsto & (|z|^2 - |w|^2, 2z\bar{w}) \end{array}, \tag{4.12}$$

which can be written in real coordinates as:

$$\begin{aligned} h(x_0, \dots, x_7) = (x_0^2 + x_1^2 + x_2^2 + x_3^2 - x_4^2 - x_5^2 - x_6^2 - x_7^2, \\ 2(x_0x_4 + x_1x_5 + x_2x_6 + x_3x_7), 2(-x_0x_5 + x_1x_4 - x_2x_7 + x_3x_6), \\ 2(-x_0x_6 + x_1x_7 + x_2x_4 - x_3x_5), 2(-x_0x_7 - x_1x_6 + x_2x_5 + x_3x_4)). \end{aligned} \tag{4.13}$$

The differential map $dh$ is the following:

$$dh = 2\begin{pmatrix} x_0 & x_1 & x_2 & x_3 & -x_4 & -x_5 & -x_6 & -x_7 \\ x_4 & x_5 & x_6 & x_7 & x_0 & x_1 & x_2 & x_3 \\ -x_5 & x_4 & -x_7 & x_6 & x_1 & -x_0 & x_3 & -x_2 \\ -x_6 & x_7 & x_4 & -x_5 & x_2 & -x_3 & -x_0 & x_1 \\ -x_7 & -x_6 & x_5 & x_4 & x_3 & x_2 & -x_1 & -x_0 \end{pmatrix}.$$

Since none of the commutators $[Y_i, Y_j]$, $i, j = 1, \dots, 7$ coincides with $Y_k$ for $k = 1, \dots, 7$, we look for the kernel of $dh$ among the commutators $Y_{ij}$, $i, j = 1, \dots, 7$. The precise form of commutators is given in Appendix B. After tedious calculations we find that $dh(Y_{45}) = dh(Y_{46}) = dh(Y_{56}) = 0$. Define $V = \{Y_{45}, Y_{46}, Y_{56}\}$. Notice that the commutation relation between $Y_{45}, Y_{46}, Y_{56}$ are:

$$[Y_{45}, Y_{46}] = Y_{56}, \quad [Y_{46}, Y_{56}] = Y_{45}, \quad [Y_{56}, Y_{45}] = Y_{46}$$

reflecting that they form a Lie algebra for the Lie group $S^3 \cong SU(2)$. As in the previous cases we define the horizontal distribution $D$ or the Ehresmann connection as an orthogonal complement to $V$ with respect to the usual inner product in $\mathbb{R}^8$. The sub-Riemannian metric $g_D$ is the restriction of the inner product from $\mathbb{R}^8$ to $D$.

**Remark 4.** Another way to construct the horizontal distribution was proposed in [57] The authors presented several bracket generating distributions transversal to $V$. None of them has a globally defined non-vanishing basis, meanwhile the vertical space $V$ has such kind of the basis. The authors of the paper [9], see also [10], presented another horizontal distribution of rank 4 that was constructed considering the Clifford algebra structure of $S^7$ and that possesses a globally defined non-vanishing basis. However, in this case a globally defined basis of the vertical space, different from $V$ constructed above, was not found. We would like to draw the attention to the paper [10], where complete description of trivializable sub-Riemannian structures on $S^{2n+1}$ induced by Clifford algebras is given.

## 5. Principal bundles

In this section we show the explicit formulas for geodesics on any odd dimensional spheres. They were found by making use of a result from [111]. To present this result we give all necessary definitions, prove the theorem of existence of geodesics and apply it to odd dimensional spheres. At the end, we will give some applications of the geometry of principal bundles to physics and will illustrate them by exploiting the results from previous sections. We recommend the book [76] as an introduction to the theory of smooth bundles.

### 5.1. Ehresmann connection

In this subsection we describe two possible ways to introduce sub-Riemannian structures on a smooth manifold $M$, provided that there exists a submersion $\pi\colon M \to B$ to another smooth manifold $B$. We call the map $\pi$ projection and the manifold $B$ the base space. For $q \in M$ we call the pre-image $F_b = \pi^{-1}(b)$, $b = \pi(q)$, the *fiber through a point* $q \in M$. The set $F_b$ is a smooth submanifold of $M$. Since the differential map $d_q\pi$ is surjective for all $q \in M$, the kernel $\ker(d_q\pi)$ is non-empty. We denote it by $V_q$ and call it the *vertical space* at $q \in M$. The collection of all vertical spaces is called *vertical distribution* or vertical sub-bundle $V \subset TM$. The vertical space is actually the tangent space $V_q = T_qF$ to the fiber $F$ passing through $q$.

**Definition 20.** *An Ehresmann connection for a submersion $\pi\colon M \to B$ is a distribution $D \subset TM$ that is everywhere transverse and of complementary dimension to the vertical distribution $V$:*

$$T_qM = D_q \oplus V_q. \qquad q \in M.$$

Notice that $\oplus$ denotes only transversality of two vector spaces at $q \in M$, but not orthogonality, because there was no any kind of metric defined on $M$ up to the moment. The vector space $D_p$ is an example of a horizontal vector space and the Ehresmann connection is an example of a horizontal distribution. Notice that given a submersion we always have a vertical distribution and the construction of the horizontal distribution or the Ehresmann connection is the question of mathematical art.

There are two ways to introduce the sub-Riemannian structure on a manifold $M$ by making use a given submersion $\pi\colon M \to B$.

CASE 1. SUB-RIEMANNIAN STRUCTURE BY RESTRICTION. Suppose that we have a submersion $\pi\colon M \to B$ and that the manifold $M$ is endowed with a Riemannian metric $g_M$. Let $V_q = \ker(d_q\pi)$. Define the horizontal vector space $D_q$ as the orthogonal complement to the vertical space $V_q$ at each $q \in M$ with respect to the given metric $g_M$: $D_q \oplus_\perp V_q = T_qM$. The obtained horizontal distribution $D$ will be the Ehresmann connection, since it is orthogonally transversal to $V$ at each point. If we denote by $g_{M_D}$ the restriction of the Riemannian metric $g_M$ to the distribution $D$, then $(M, D, g_{M_D})$ is the *sub-Riemannian manifold defined by the submersion* $\pi$ *and the Riemannian metric* $g_M$ *on* $M$. In this case the sub-Riemannian length of a horizontal curve is equal to the Riemannian length, since the vertical components vanish.

CASE 2. SUB-RIEMANNIAN STRUCTURE BY LIFTING. Suppose now that for the submersion $\pi\colon M \to B$ the Ehresmann connection $D$ is defined, and moreover, the base manifold $B$ is endowed with a Riemannian metric $g_B$. Since the restriction $d_q\pi|_{D_q}\colon D_q \to T_{\pi(q)}B$ is an isomorphism, we can pullback the metric $g_B$ to the horizontal space $D_q$ at each $q \in M$. Denote the obtained metric by $g_{B_D}$. Thus,

$$g_{B_D}(v, w) := g_B(d_q\pi(v), d_q\pi(w)), \quad v, w \in D_q, \quad q \in M.$$

The obtained metric varies smoothly with $q \in M$. The triplet $(M, D, g_{B_D})$ is called a *sub-Riemannian manifold induced by the Ehresmann connection* $D$ *on* $M$ *and by the Riemannian metric* $g_B$ *on* $B$. In this case we get the following properties.

1. If a horizontal curve $\gamma$ is given on $M$, then the sub-Riemannian length of $\gamma$ is equal to the Riemannian length of its projection to $B$. This is obvious, since the vertical component of the velocity vector of a horizontal curve $\gamma$ in $T_qM$ is absent, and moreover, the vertical space is projected to the 0-subspace in each tangent space $T_{\pi(q)}B$.
2. What happens if we are given a curve on the base space and we pull it back to the manifold $M$? Define the *horizontal lift* of a curve $c\colon I \to B$ to $M$. The horizontal lift of the curve $c$ is a curve $\gamma\colon I \to M$ such that

$$(1)\ \ \dot\gamma(t) \in D_{\gamma(t)} \quad \text{and} \quad (2)\ \ \pi(\gamma(t)) = c(t) \quad \text{for all} \ \ t \in I.$$

   The horizontal lift of a Riemannian geodesic in $B$ is a sub-Riemannian geodesic in $M$. If the Riemannian geodesic in $B$ is a length minimizers between its end points, then its horizontal lift is a sub-Riemannian length minimizer between the corresponding fibers.

**Remark 5.** It is natural to ask, when the horizontal lift exists. We will not discuss it here. But if it exists, then given a point $q \in M$ and a curve $c$ starting at $\pi(q)$, the horizontal lift $\gamma$ of $c$ starting from $q \in M$ is unique.

Now, having these two ways of constructing sub-Riemannian structures on $M$, we can ask when these structures coincide. Suppose we are given a submersion $\pi$ of a Riemannian manifold $(M, g_M)$ to a Riemannian manifold $(B, g_B)$ and the Ehresmann connection $D$ is orthogonal to $\ker(d\pi)$ everywhere. If the restriction $d_q\pi|_{D_q}\colon (D_q, d_{M_D}) \to (T_{\pi(q)}B, g_B)$ is a linear isometry for corresponding vector spaces for all $q \in M$, then $(M, D, g_{M_D}) = (M, D, g_{B_M})$. In this case the submersion $\pi$ is actually a Riemannian submersion, see Definition 41.

### 5.2. Metrics on principal bundles

A definition of a fiber bundle is given in Appendix A, Definition 59. We present the definition of the principal bundle in the smooth setting. Let $M$ and $B$ be smooth manifolds and let $\mathbb{G}$ be a Lie group. Recall that an action of a Lie group on a smooth manifold is a smooth map by definition.

**Definition 21.** *Let $M, B$, and $\mathbb{G}$ be as above. A fiber bundle $(F, M, B, \pi)$ is a smooth principal bundle if the typical fiber $F$ has the structure of a Lie group $\mathbb{G}$ and, moreover, the group $\mathbb{G}$ acts freely and transitively on each fiber, see (Definitions* 56, 57).

In the case of principal $\mathbb{G}$-bundle we have the following properties.

**Proposition 3.** [75, 125]

1. *The action $\mu_\tau \colon M \to M$ of the group $\mathbb{G}$ is a proper map for any $\tau \in \mathbb{G}$, that is, the pre-image of a compact set is compact.*
2. *The base space $B$ is diffeomorphic to the space $M/\mathbb{G}$ of orbits of the group $\mathbb{G}$. The space $M/\mathbb{G}$ becomes a smooth homogeneous manifold.*
3. *The map $\pi$ is a natural projection $\pi \colon M \to M/\mathbb{G}$ onto the quotient space and it is a smooth submersion.*

We assume from now on that the group $\mathbb{G}$ acts on the right: $q \mapsto q.\tau$ for $\tau \in \mathbb{G}$, $q \in M$ and we will omit the word “smooth” in the notion of principal bundle.

**Definition 22.** *Let $\pi \colon M \to B$ be a principal $\mathbb{G}$-bundle, $D$ be the Ehresmann connection and let $g_D$ be a sub-Riemannian metric on $M$ associated with $D$. If $g_D$ is invariant under the right action of $\mathbb{G}$ on $M$, that is,*

$$g_D(v_q, w_q) = g_D\big(dr_\tau(v_q), dr_\tau(w_q)\big) = g_D(v_{q.\tau}, w_{q.\tau}), \quad \tau \in \mathbb{G}, \quad q \in M,$$

*then the metric $g_D$ is said to be of bundle type.*

**Example 6.** Let us suppose that $\pi \colon M \to B$ is a principal $\mathbb{G}$-bundle and suppose that $M$ is endowed with a Riemannian metric $g_M$ which is invariant with respect to the right action of $\mathbb{G}$. Let $D$ be the distribution from the Case 1. Then the restriction $g_{M_D}$ of $g_M$ to $D$ gives a bundle type metric.

In what follows we want to present the special situation, where geodesics can be calculated by making use of Riemannian and sub-Riemannian metrics. Assume that $\pi \colon M \to B$ is a principal $\mathbb{G}$-bundle, $g_M$ is a Riemannian metric on $M$, and $D_q$ is the horizontal space orthogonal to the vertical space $V_q = \ker(d_q\pi)$ at each $q \in M$ with respect to $g_M$. We also suppose that the Riemannian metric $g_M$ is right $\mathbb{G}$-invariant. In addition to the above described structure we assume that the distribution $D$ is furnished with a sub-Riemannian metric $g_D$ of bundle type. We say that the metric $g_M$ is *compatible* with $g_D$ if the restriction $g_{M_D} := g_M|_D$ coincides with $g_D$ on $M$, that is

$$g_{M_D}(v_q, w_q) = g_D(v_q, w_q) \quad \text{for all} \quad v_q, w_q \in D_q \quad \text{and all} \quad q \in M.$$

Now let us consider the restriction $g_{M_V} := g_M|_V$ of $g_M$ to the vertical subspace $V_q \subset T_qM$, $q \in M$. The metric $g_{M_V}$ is defined on the tangent space $V_q$ to the fiber $\mathbb{G}$ at each $q$. Since there is an isomorphism between $V_q$ and the Lie algebra $\mathfrak{g}$ of the Lie group $\mathbb{G}$, the metric $g_{M_V}$ defines a bilinear symmetric form $\mathbb{I}_q \colon \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$, which is called the *moment of inertia tensor* at $q \in M$. We can express it in the following way. By making use of the group exponential map $\exp_{\mathbb{G}} \colon \mathfrak{g} \to \mathbb{G}$, we introduce the infinitesimal generator map $\sigma_q$ for the right $\mathbb{G}$-group action on $M$. Namely,

$$\sigma_q \colon \mathfrak{g} \to V_q \subset T_qM \quad \text{is such that} \quad \mathfrak{g} \ni \xi \mapsto \frac{d}{d\epsilon}\Big|_{\epsilon=0} q.\exp(\epsilon\xi). \tag{5.1}$$

Let us observe the following feature. Fixing fiber, by choosing a point $q \in M$ and a local trivialization, we identify the fiber with the Lie group $\mathbb{G}$. Now we can consider $q.\tau$ not only as a right action of the Lie group $\mathbb{G}$ ($\tau \in \mathbb{G}$) on $M$ along the fiber, but also as the action on the left of $q$ (considered as an element of $\mathbb{G}$) on the Lie group $\mathbb{G}$, or

$$q.\tau = l_q(\tau). \tag{5.2}$$

So, it is convenient to think of the infinitesimal generator of the right action as of a locally left invariant vector field, since

$$\sigma_q(\xi) = \frac{d}{d\epsilon}\Big|_{\epsilon=0} q.\exp(\epsilon\xi) = \frac{d}{d\epsilon}\Big|_{\epsilon=0} l_q\big(\exp(\epsilon\xi)\big) = dl_q(\xi)$$

by the property 1 in Theorem 8.2 or by the definition of the exponential curve in Subsection 3.1, see (3.1). Then we define the bilinear symmetric tensor $\mathbb{I}_q \colon \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ by

$$\mathbb{I}_q(\xi, \eta) = g_{M_V}(\sigma_q(\xi), \sigma_q(\eta)), \qquad \xi, \eta \in \mathfrak{g}, \quad q \in M. \tag{5.3}$$

We conclude that if a right $\mathbb{G}$-invariant Riemannian metric $g_M$ is given on a principal $\mathbb{G}$-bundle $\pi \colon M \to B$, then we define the sub-Riemannian structure $(D, g_{M_D})$, with $D = V^{\perp}$, and the moment of inertia tensor $\mathbb{I}_q$ (5.3) by means of restrictions of $g_M$.

Conversely, if we have a sub-Riemannian structure $(D, g_D)$ with the metric $g_D$ of bundle type and a moment of inertia tensor $\mathbb{I}_q \colon \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$, $q \in M$, then we can define a Riemannian metric $g_M$ as follows. Let us write any vector $v_q \in T_qM$ according to the transversal decomposition $D_q \oplus V_q$ as $v_q = v_{D_q} + v_{V_q}$. Then the Riemannian metric $g_M$ is defined by

$$g_M(v_q, w_q) := g_D(v_{D_q}, w_{D_q}) + \mathbb{I}_q\big(\sigma_q^{-1}(v_{V_q}), \sigma_q^{-1}(w_{V_q})\big),$$

where the inverse map $\sigma_q^{-1}$ is well defined. In order to check that the obtained Riemannian metric $g_M$ is compatible with the bundle type metric $g_D$ we observe that

- $D_q$ and $V_q$ become orthogonal with respect to $g_M$, since if $v_q \in D_q$ and $w_q \in V_q$, then $g(v_q, w_q) = g_D(v_q, 0) + \mathbb{I}_q\big(0, \sigma_q^{-1}(w_q)\big) = 0$;
- the restriction of $g_M$ to the distribution $D$ coincides with $g_D$,

- to be right $\mathbb{G}$-invariant the metric $g_M$ has to satisfy the relation
$$g_M(v_q, w_q) = g_M(dr_\tau(v_q), dr_\tau(w_q)) = g_M(v_{q.\tau}, w_{q.\tau}). \tag{5.4}$$
Let us reformulate the last condition in terms of the symmetric bi-linear tensor $\mathbb{I}_q$. The left-hand side of (5.4) yields
$$g_M(v_q, w_q) = g_D(v_{D_q}, w_{D_q}) + \mathbb{I}_q(\xi, \eta), \tag{5.5}$$
where $\sigma_q(\xi) = v_{V_q}$, and $\sigma_q(\eta) = w_{V_q}$. The right hand side of (5.4) leads to
$$g_M(v_{q.\tau}, w_{q.\tau}) = g_D(dr_\tau(v_{D_q}), dr_\tau(w_{D_q})) + \mathbb{I}_{q.\tau}\big(\zeta, \chi\big), \tag{5.6}$$
where $\sigma_{q.\tau}(\zeta) = dr_\tau\sigma_q(\xi)$ and $\sigma_{q.\tau}(\chi) = dr_\tau\sigma_q(\eta)$. Let us calculate $dr_\tau\sigma_q(\xi)$. We get
$$\begin{aligned} dr_\tau\sigma_q(\xi) &= \frac{d}{d\epsilon}\Big|_{\epsilon=0} q\exp(\epsilon\xi)\tau = \frac{d}{d\epsilon}\Big|_{\epsilon=0} q\tau\big(\tau^{-1}\exp(\epsilon\xi)\tau\big) \\ &= \frac{d}{d\epsilon}\Big|_{\epsilon=0} q\tau\exp\big(\epsilon\,\mathrm{Ad}_{\tau^{-1}}(\xi)\big) = \sigma_{q.\tau}\big(\,\mathrm{Ad}_{\tau^{-1}}(\xi)\big). \end{aligned} \tag{5.7}$$
where Ad is the adjoint action of $\mathbb{G}$ over $\mathfrak{g}$, for definition see Example 13. For the relation of the adjoint map and the exponential map see Appendix A. Analogously, we get $dr_\tau\sigma_q(\eta) = \sigma_{q.\tau}\big(\,\mathrm{Ad}_{\tau^{-1}}(\eta)\big)$. Since $g_D$ is right invariant by definition of a bundle type metric, the equalities (5.4), (5.5), and (5.6) imply
$$\mathbb{I}_q(\xi, \eta) = \mathbb{I}_{q\tau}\big(\,\mathrm{Ad}_{\tau^{-1}}(\xi), \mathrm{Ad}_{\tau^{-1}}(\eta)\big) \quad \text{or} \quad \mathbb{I}_{q\tau}(\xi, \eta) = \mathbb{I}_q\big(\,\mathrm{Ad}_\tau(\xi), \mathrm{Ad}_\tau(\eta)\big)$$
by changing $q$ to $q\tau^{-1}$, and then, $\tau^{-1}$ to $\tau$. We conclude that to make the constructed Riemannian metric $g_M$ invariant under the right action of the Lie group $\mathbb{G}$, we have to require that the given inertia tensor $\mathbb{I}_q$ should be invariant with respect of the adjoint action of $\mathbb{G}$ on its Lie algebra $\mathfrak{g}$.

After all these discussions, we define a metric that we will work with. We will use the terminology of [111].

**Definition 23.** *A Riemannian metric $g_M$ on a smooth manifold $M$ is said to be of constant bi-invariant type if*

1. *$g_M$ is right $\mathbb{G}$-invariant,*
2. *its inertia tensor $\mathbb{I}_q$ is independent of $q \in M$.*

The word “constant” refers to the independence of the moment of inertia tensor from the points of the manifold. The bi-invariance reflects the fact that the inertia tensor defines a bi-invariant metric along the fiber $\mathbb{G}$. We discuss it in the following remark.

**Remark 6.** Bi-invariant metrics on Lie group. Since the Lie algebra $\mathfrak{g}$ is related to the tangent spaces $T_q\mathbb{G}$ to its group by left translations, any tensor $\Theta$ on the algebra $\mathfrak{g}$ corresponds to a left invariant tensor $T$ on its group $\mathbb{G}$, because $T$ is defined by making use of left translations to bring $\Theta$ to each point $q \in \mathbb{G}$. If, moreover, the tensor $\Theta$ on $\mathfrak{g}$ is invariant under the adjoint action of $\mathbb{G}$ on its Lie algebra $\mathfrak{g}$, then the left invariant tensor $T$ on $\mathbb{G}$ becomes right invariant. Applying this considerations to the bi-linear symmetric non-degenerate form $\mathbb{I}_q$ on $\mathfrak{g}$, we obtain that it generates a bi-invariant metric on the fiber $\mathbb{G}$ through the point $q$. We see that the adjoint invariance on $\mathfrak{g}$ means that $\mathbb{I}_q$ is constant along the fiber through the point $q$.

### 5.3. Geodesics theorem

Before we state and prove one of the principle theorems, we formulate an auxiliary statement.

Let $\mathbb{G}$ be a Lie group, $\mathfrak{g}$ be its Lie algebra and let $g_{\mathbb{G}}$ be a bi-invariant Riemannian metric on the group, that may be given by an adjoint invariant bilinear symmetric form on $\mathfrak{g}$ as was noticed in Remark 6. Thus, the Lie group is also considered as a Riemannian manifold $(\mathbb{G}, g_{\mathbb{G}})$. There are two exponential maps defined in this case: the group exponential $\exp_{\mathbb{G}}\colon \mathfrak{g} \to \mathbb{G}$ and the Riemann exponential map $(\exp_R)_e\colon T_e\mathbb{G} \sim \mathfrak{g} \to \mathbb{G}$.

**Proposition 4.** [106] *In the above stated notations the two exponential maps coincide.*

In other words the Riemannian geodesic through the identity of the group $\mathbb{G}$ coincides with the one-parameter subgroup produced by the group exponential map.

Let $\pi\colon M \to B$ be a principal $\mathbb{G}$-bundle and let $g_M$ be a Riemannian metric of constant bi-invariant type. Let $D$ be the Ehresmann connection which is orthogonal with respect to $g_M$ to the vertical space $V_q$ at each $q \in M$. Let $g_D$ and $g_V$ denote restrictions of $g_M$ to $D$ and $V$, respectively. Recall that the infinitesimal action $\sigma_q\colon \mathfrak{g} \to V_q$ is an isomorphism by (5.1). Denote by $proj_q$ the projection from $T_qM$ to $V_q$ at each $q \in M$. The composition $A = \sigma_q^{-1} \circ proj_q$ is called the $\mathfrak{g}$-*valued connection form*, see diagram (5.8).

$$\begin{tikzcd}
\mathfrak{g} \arrow[r, bend left, "\sigma_q"] \arrow[r, phantom, "\cong"] & \ker(d_q\pi) = V_q \arrow[l, bend left, "\sigma_q^{-1}"] & T_qM \arrow[l, "proj."'] \arrow[ll, bend right, "A"']
\end{tikzcd} . \tag{5.8}$$

Let $\exp_R$ be the Riemannian exponential map generated by $g_M$ and let $\gamma_{v,R}(t) = \exp_R(tv)$ be the Riemannian geodesic passing through $q \in M$ with the initial velocity vector $v \in T_qM$. We project this Riemannian geodesic to the base manifold $B$ obtaining a curve $\pi(\gamma_{v,R})$. Then we lift horizontally $\pi(\gamma_{v,R})$ to $M$ and obtain a curve that we denote by $\gamma_{sR}$.

**Theorem 5.1.** [111] *In the above mentioned notations the curve $\gamma_{sR}$ is a normal sub-Riemannian geodesic starting at $q \in M$. It is given by the formula*

$$\gamma_{sR}(t) = \gamma_{v,R}(t)\exp_G(-tA_q(v)), \quad v \in T_qM. \tag{5.9}$$

*Proof.* We follow the ideas in [111]. Since the decomposition of $D \oplus_{\perp} V$ is orthogonal with respect to the Riemannian metric $g_M$, and $g_D$ and $g_V$ are defined by the restriction of $g_M$ to the corresponding distributions, we can define three Hamiltonian functions $H_R$, $H_{sR}$ and $H_V$. Here we denote by $H_R$ the Riemannian Hamiltonian function related to the Riemannian metric $g_M$, by $H_{sR}$ the sub-Riemannian Hamiltonian function related to the metric $g_D$, see (2.12) and by $H_V$ the vertical Hamiltonian function related to $g_V$ and constructed by the same rule as in (2.12). Then the orthogonality of the composition $D \oplus_{\perp} V$ implies that $H_{sR} = H_R - H_V$. Let us also use the notations

$$\exp_R\colon T_qM \to M, \quad \exp_{sR}\colon D_q \to M, \quad \exp_V\colon V_q \to M, \quad q \in M.$$

The rough idea of the proof is to show that if these Hamiltonian functions Poisson commute, then the corresponding flows on $T^*M$ produced by its Hamiltonian vector fields also commute. Therefore, if $v = v_D + v_V$ is the initial velocity vector written according to the decomposition $D_q \oplus_{\perp} V_q = T_qM$, then the flow commutativity property leads to the commutativity of the exponential maps, that is

$$\exp_{sR}(tv_D) = \exp_R(tv)\exp_V(-tv_V).$$

In the last step of the proof we observe that $\exp_V(-tv_V)$ coincides with the group $\mathbb{G}$ exponential map because the metric $g_V$ is bi-invariant along the fiber through $q \in M$.

The first step in the proof of the theorem is to show that the Hamiltonian functions $H_R$, $H_{sR}$, and $H_V$ Poisson commute. Actually we only need to show that $\{H_{sR}, H_V\} = 0$. We use the local trivialization for the bundle $\pi\colon M \to B$. Let $U \subset B$ be a neighborhood of $\pi(q)$, then $\pi^{-1}(U)$ is diffeomorphic to $\mathbb{G} \times U$. At the level of cotangent bundles it leads to the diffeomorphism

$$T^*_{\pi^{-1}(U)}M = T^*(\mathbb{G} \times U) \cong T^*\mathbb{G} \times T^*U.$$

Let us use the coordinates $\pi^{-1}(U) \ni q = (\tau, b) = (\tau^1, \dots \tau^l, b^1, \dots, b^k) \in \mathbb{G} \times U$, $l + k = n = \dim M$ for points, and $T_q^*M \ni \lambda = (\mu, p) = (\mu_1, \dots, \mu_l, p_1, \dots, p_k) \in T_q^*\mathbb{G} \times T_b^*U$ for momenta. Since the moment of inertia tensor $\mathbb{I}$ is independent of $q \in M$ and it is independent of the horizontal part of any vector, the dual tensor $\mathbb{I}^*\colon \mathfrak{g}^* \times \mathfrak{g}^* \to \mathbb{R}$ is also independent of $q \in M$ and $p \in T_b^*U$. This implies that $H_V(\tau, b, \mu, p)$ is only a function of the $\mu$-variables: $H_V = H_V(\mu)$. The right invariant property of the metric $g_D$ leads to independence of the corresponding Hamiltonian function $H_{sR}$ from $\tau$-slot of variables: $H_{sR} = H_{sR}(b, \mu, p)$. The Poisson brackets are

$$\begin{aligned}\{H_{sR}, H_V\} = \Big(\sum_{j=1}^{l} \frac{\partial H_{sR}}{\partial \mu_j}\frac{\partial H_V}{\partial \tau_j} + \sum_{j=1}^{k} \frac{\partial H_{sR}}{\partial p_j}\frac{\partial H_V}{\partial b_j}\Big) \\ - \Big(\sum_{j=1}^{l} \frac{\partial H_{sR}}{\partial \tau_j}\frac{\partial H_V}{\partial \mu_j} + \sum_{j=1}^{k} \frac{\partial H_{sR}}{\partial b_j}\frac{\partial H_V}{\partial p_j}\Big).\end{aligned}$$

In the last sum we have $\frac{\partial H_V(\mu)}{\partial \tau_j} = \frac{\partial H_V(\mu)}{\partial b_j} = \frac{\partial H_V(\mu)}{\partial p_j} = 0$ and $\frac{\partial H_{sR}(b,\mu,p)}{\partial \tau_j} = 0$, that implies $\{H_{sR}, H_V\} = 0$. We conclude that flows $\Phi_R, \Phi_{sR}$, and $\Phi_V$ on $T^*M$ corresponding to $\overrightarrow{H}_R$, $\overrightarrow{H}_{sR}$, and $\overrightarrow{H}_V$, respectively, commute. Recall that the exponential map is a composition of the following maps, see (2.10).

$$\begin{tikzcd} T_qM \arrow[r, "\iota"] \arrow[rrrr, bend right, "\exp"'] & TM \arrow[r, "duality"] & T^*M \arrow[r, "\Phi"] & T^*M \arrow[r, "\mathrm{pr}^*_M"] & M, \end{tikzcd}$$

where we have to change the corresponding flows and $T_qM$ to $D_q$ and $V_q$ respectively. We see that the commutation of the flows leads to the commutation of the exponential maps and we have

$$\exp_{sR}(tv) = \exp_R(tv)\exp_V(-tv), \quad t \in I \tag{5.10}$$

for $H_{sR} = H_R - H_V$, where $v \in T_qM$ is the initial velocity vector that corresponds to the choice of $\lambda$ in the flow. The curve $\exp_R(tv) = \gamma_{v,R}(t)$, $t \in I$, is the Riemannian geodesic on $M$, starting from $q \in M$ with the initial velocity $v \in T_qM$. Recall, that the exponential curve $\exp_V(-tv)$ produced by the Hamiltonian function $H_V(\mu)$ is independent of $p$ variables in the momentum slot. It gives $\exp_V(-tv) = \exp_V(-tv_V)$. Now we exploit the fact that the Riemannian geodesic $\exp_V(-tv_V)$ coincides with the geodesic (or one-parametric subgroup) given by the group exponential map $\exp_G(-t\sigma^{-1}(v_V))$. The composition of the projection of $v$ to the vertical space $V_q$ and of the map $\sigma_q^{-1}$ is called a $\mathfrak{g}$-valued connection form $A_q \colon T_qM \to \mathfrak{g}$, see (5.8). So equation (5.10) takes the form (5.9).

Let us make some observations. The vector $v$ is just an initial velocity vector at $q \in M$. The element $-tA_q(v) \in \mathfrak{g}$, and therefore, the vector $tv - \sigma(tA_q(v))$ is horizontal for any $t \in I$. So the velocity vector of the resulting curve in the right-hand side of (5.10) is horizontal for any moment $t$, and the resulting curve in the left hand side is a horizontal geodesic. □

It is shown in [111] that, moreover, all normal sub-Riemannian geodesics are given by the formula (5.9).

**5.3.1. Geodesics on odd dimensional spheres.** In the case of odd dimensional spheres

$$S^{2n+1} = \{(z_0, z_1, \dots, z_n) \in \mathbb{C}^{n+1} \mid \sum_{j=0}^{n} |z_j|^2 = 1\},$$

there is a natural action of $U(1)$ given by $q.v = (z_0v, z_1v, \dots, z_nv)$, where $v \in U(1)$. This action induces the Hopf fibration $U(1) \to S^{2n+1} \to \mathbb{C}P^n$, which forms a principal $U(1)-$bundle with connection $D$ given by the orthogonal complement to the vector field

$$V_q = -y^0\partial_{x_0} + x^0\partial_{y_0} - \dots - y^n\partial_{x_n} + x^n\partial_{y_n} \tag{5.11}$$

at each $q = (x^0, y^0, \dots, x^n, y^n) \in S^{2n+1}$, $\{z_j = x^j + iy^j\}_{j=0}^n$, with respect to the usual inner product in $\mathbb{R}^{2n+2}$. As it was shown, this distribution can be also given by $\ker\omega$ with respect to the contact form $\omega = -y_0dx^0+x_0dy^0-\ldots-y_ndx^n+x_ndy^n$. Note that $V_q = qi$, $q \in S^{2n+1}$, where $q$ is thought of as a radial vector at the origin to the unit sphere, $i$ is the complex imaginary unit, and $q.i = qi$, $i \in \mathfrak{u}(1)$ is the $\mathfrak{u}(1)$ action.

Consider the sphere $(S^{2n+1}, D, g_D)$ as a sub-Riemannian manifold with the sub-Riemannian metric $g_D$ obtained by the restriction of the usual Riemannian metric $g$ on $TS^{2n+1}$ to the distribution $D$. As a direct application of Theorem 5.1, it is possible to describe all normal sub-Riemannian geodesics for $S^{2n+1}$. The Lie algebra $\mathfrak{u}(1)$ is one dimensional, its typical elements are purely imaginary numbers: $\xi = i\alpha$. The $\mathfrak{u}(1)-$valued connection form is $A_q(v) = ig(v, V_q)$, $v \in T_qS^{2n+1}$ and $g(v, V_q)$ is just the projection of $v$ to the vertical space $V_q$ by making use of the Riemannian metric $g$. The Riemannian metric $g$ on $S^{2n+1}$ is of constant bi-invariant type, because we have

$$\sigma_q(\xi) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} q\exp_{U(1)}(\epsilon\xi) = q.i\alpha = \alpha V_q,$$

for any $q \in S^{2n+1}$ and $\xi = i\alpha \in \mathfrak{u}(1)$. Therefore, the moment of inertia tensor is given by

$$\mathbb{I}_q(\xi, \tilde{\xi}) = \mathbb{I}_q(i\alpha, i\tilde{\alpha}) = g(\alpha V_q, \tilde{\alpha} V_q) = \alpha\tilde{\alpha},$$

which does not depend on $q \in M$. By Theorem 5.1, we have the following result.

**Proposition 5.** *Let $q \in S^{2n+1}$ and $v \in T_qS^{2n+1}$. If $\gamma_R(t) = (z_0(t), \ldots, z_n(t))$ is the great circle satisfying $\gamma_R(0) = q$ and $\dot{\gamma}_R(0) = v$, then the corresponding sub-Riemannian geodesic $\gamma_{sR}$ is given by*

$$\gamma_{sR}(t) = \Big(z_0(t)e^{-itg(v,V_q)}, \ldots, z_n(t)e^{-itg(v,V_q)}\Big), \quad t \in \mathbb{R}. \tag{5.12}$$

To analyze formula (5.12) we recall that the Riemannian geodesic starting at $q \in S^n$ with a velocity $v \in T_qS^n$ for any sphere $S^n$ is a submanifold of $\mathbb{R}^{n+1}$ is given by:

$$\gamma_R(t) = q\cos(\|v\|t) + \frac{v}{\|v\|}\sin(\|v\|t), \quad \text{where} \quad \|v\| = \sqrt{g(v,v)}. \tag{5.13}$$

The great circle $\gamma_R(t)$ on $S^{2n+1}$, considered as a submanifold of $\mathbb{R}^{2(n+1)} \cong \mathbb{C}^{n+1}$, will be written in complex notation as $\gamma_R(t) = (z_0(t), \ldots, z_n(t))$. Observe that $V(\gamma(t)) = \gamma(t).i$ and $V_q = V(\gamma(0))$.

The following corollary can be thought of as a sort of Pythagoras theorem for sub-Riemannian spheres.

**Corollary 2.** *For a horizontal sub-Riemannian geodesic on $S^{2n+1}$ of the form* (5.12) *the following equation holds*

$$\|\dot{\gamma}_{sR}(t)\|^2 + g^2(v, V_q) = \|v\|^2.$$

*Thus, its sub-Riemannian velocity is constant and its sub-Riemannian length for $t \in [a,b]$ is $\ell(\gamma) = (b-a)\sqrt{\|v\|^2 - g^2(v, V_q)}$.*

*Proof.* Denote by $(\cdot\,,\cdot)_H$ the standard Hermitian product in $\mathbb{C}^{n+1}$, $\mathrm{Re}\,(\cdot\,,\cdot)_H = g(\cdot\,,\cdot)$. By straightforward calculations, we have

$$\begin{aligned}
&(\dot\gamma_{sR},\dot\gamma_{sR})_H = \\
&\quad= \Big((-ig(v,V_q)\gamma_R+\dot\gamma_R)e^{-itg(v,V_q)},(-ig(v,V_q)\gamma_R+\dot\gamma_R)e^{-itg(v,V_q)}\Big)_H \\
&\quad= g^2(v,V_q)(\gamma_R,\gamma_R)_H+(\dot\gamma_R,\dot\gamma_R)_H+g(v,V_q)\big(i(\dot\gamma_R,\gamma_R)_H-i(\gamma_R,\dot\gamma_R)_H\big) \\
&\quad= g^2(v,V_q)+\|v\|^2-2g^2(v,V_q).
\end{aligned}$$

The assertion follows. □

**Corollary 3.** *If a curve $\gamma_{sR}(t)=\gamma_R(t)e^{-itg(v,V_q)}$ is parameterized by arc length then $\|v\|^2=1+g^2(v,V_q)$.*

**Corollary 4.** *The set of sub-Riemannian geodesics arising from the great circles $\gamma_R(t)$, such that $\dot\gamma_R(0)\in D$ is diffeomorphic to $\mathbb{C}P^n$.*

*Proof.* In this case, any sub-Riemannian geodesic starting at $q\in S^{2n+1}$ with the initial velocity $v\in D\subset T_qS^{2n+1}$ coincides with the corresponding great circle, because the condition $\dot\gamma_R(0)\in D$ is equivalent to $g(v,V_q)=0$, thus

$$\gamma_{sR}(t)=p\cos(\|v\|t)+\frac{v}{\|v\|}\sin(\|v\|t),$$

whose loci is uniquely determined by the point $[v]\in\mathbb{C}P^n$. □

Observe that the manifold $\mathbb{C}P^n$ can be seen as a submanifold of $S^{2n+1}$ which is transversal to $V$ along the fiber containing $q$ and it can be thought of as a sophisticated analogue of the horizontal space at the identity in the $(2n+1)-$dimensional Heisenberg group.

**5.3.2. Curvature or charge of sub-Riemannian geodesics on $S^3$.** The following equation

$$\nabla_{\dot\gamma_{sR}}\dot\gamma_{sR}+2\varkappa J(\dot\gamma_{sR})=0, \tag{5.14}$$

obtained by variational method, is true for length minimizers in $S^3$ [74]. Here $\nabla$ is the Levi-Civita connection associated with the Riemannian metric on $S^3$ and $J$ is an almost complex structure on $S^3$ satisfying $J(X)=-Y$, $J(Y)=X$.

The geometers call the parameter $\varkappa$ in (5.14) the *curvature* of $\gamma_{sR}$, since after projecting the curve $\gamma_{sR}$ via the Hopf fibration, $\varkappa$ becomes precisely the curvature of the projected curve in $S^2$. Note that curves of zero curvature are the horizontal great circles. Physicists call the parameter $\varkappa$ charge or phase and denote it by $\lambda$. We return to the notion of a charge later in Subsection 5.5.

Since on $S^3$ all length minimizers are given by normal geodesics, we conclude that solutions of (5.14) coincides with (5.9). Let us see closer on this relation.

**Proposition 6.** *The curvature of the normal sub-Riemannian geodesic*

$$\gamma_{sR}(t)=\gamma_R(t)e^{-itg(v,V_q)}$$

*in $S^3$, starting from $q \in S^3$ with an initial velocity $v \in T_qS^3$, parameterized by arc length, equals the value $g(v, V_q)$.*

*Proof.* Recall that the Lie group structure of $S^3$ as of the set of unit quaternions, induces the globally defined vector fields (4.2). Let $q = (x^0, x^1, x^2, x^3) = \gamma(0) \in S^3$ be an initial point of $\gamma$ and let $v = (v^0, v^1, v^2, v^3) = \dot{\gamma}_R(0) \in T_qS^3$ be an initial velocity of the corresponding great circle $\gamma_R$. By direct calculation, we have

$$\dot{\gamma}(t) = f_X(t)X(\gamma(t)) + f_Y(t)Y(\gamma(t)), \tag{5.15}$$

where, denoting $\alpha = g(v, X), \beta = g(v, Y)$, we have

$$\begin{aligned} f_X(t) &= \alpha\cos(2tg(v, V)) + \beta\sin(2tg(v, V)), \\ f_Y(t) &= \beta\cos(2tg(v, V)) - \alpha\sin(2tg(v, V)). \end{aligned}$$

It follows from this decomposition that

$$J(\dot{\gamma}(t)) = -f_Y(t)X(\gamma(t)) + f_X(t)Y(\gamma(t)). \tag{5.16}$$

It remains to determine the term $\nabla_{\dot{\gamma}}\dot{\gamma}$. As it is well-known for submanifolds of $\mathbb{R}^n$, the vector field $\nabla_{\dot{\gamma}}\dot{\gamma}$ corresponds to the projection of the second derivative $\ddot{\gamma}$ to the tangent space of the submanifold. In this case, differentiating (5.15) we obtain

$$\nabla_{\dot{\gamma}}\dot{\gamma} = 2g(v, V)(f_Y(t)X(\gamma(t)) - f_X(t)Y(\gamma(t))) = -2g(v, V)\, J(\dot{\gamma}(t)).$$

This finishes the proof. □

In [74] the problem of existence of closed sub-Riemannian geodesics is also discussed. Their result states that a complete geodesic $\gamma$ in $S^3$ parameterized by arc length, with curvature $\varkappa$ is closed, if and only if, $\varkappa/\sqrt{1 + \varkappa^2} \in \mathbb{Q}$. This result can be generalized to any odd dimensional sphere.

**Proposition 7.** *Let $\gamma_{sR} \colon \mathbb{R} \to S^{2n+1}$ be a complete sub-Riemannian geodesic parameterized by arc length, with an initial velocity $v \in T_qS^{2n+1}$. Then $\gamma_{sR}$ is closed if and only if*

$$\frac{g(v, V_q)}{\sqrt{1 + g^2(v, V_q)}} \in \mathbb{Q}.$$

*Proof.* The curve $\gamma_{sR} \colon \mathbb{R} \to S^{2n+1}$ is closed, if and only if,

$$q = e^{-iTg(v,V_q)}\left(q\cos(\|v\|T) + \frac{v}{\|v\|}\sin(\|v\|T)\right)$$

for some $T > 0$. Since $v \in T_qS^{2n+1}$, we know that $v$ is orthogonal to the vector joining $0 \in \mathbb{R}^{2n+2}$ and $q$, with respect to $g$. Thus, $\sin(\|v\|T) = 0$, which forces $T = k\pi/\|v\|$, $k \in \mathbb{Z}$. To complete the argument, we only need to see that $\pm e^{-i\pi k\left(\frac{g(v,V_q)}{\|v\|}\right)}q = q$ for some $k \in \mathbb{Z}$, if and only if,

$$\frac{g(v, V_q)}{\|v\|} = \frac{g(v, V_q)}{\sqrt{1 + g^2(v, V_q)}} \in \mathbb{Q},$$

where we have used Corollary 3. □

Exercise.

1. Calculate directly that the curve (5.9) is horizontal.
2. Write the equation of geodesics starting from the point $q = (1, 0, 0, 0)$. What is the value of $\varkappa$ at $q = (1, 0, 0, 0)$?

**5.3.3. Sub-Riemannian geodesics on $S^{4n+3}$.** Let us consider the sphere

$$S^{4n+3} = \{(q_0, \dots, q_n) \in \mathbb{Q}^{n+1} \ \mid \ \sum_{j=0}^{n} |q_j|^2 = 1\}.$$

The right action of the group $Sp(1) \cong SU(2) \cong S^3$ on $S^{4n+3}$ is defined by $q.\upsilon = (q_0, \dots, q_n).\upsilon = (q_0\upsilon, \dots, q_n\upsilon)$, $q \in S^{4n+3}$. This action induces a quaternionic Hopf fibration $S^3 \to S^{4n+3} \to \mathbb{H}P^n$, given by

$$\begin{array}{cccc} h & : & S^{4n+3} & \to & \mathbb{H}P^n \\ & & (q_0, \dots, q_n) & \mapsto & [q_0 : \dots : q_n]. \end{array} \tag{5.17}$$

This map forms a principal $S^3-$bundle with the Ehresmann connection given by the orthogonal complement to the vector fields

$$V_q^1 = -y^0\partial_{x^0} + x^0\partial_{y^0} + w^0\partial_{z^0} - z^0\partial_{w^0} - \dots - y^n\partial_{x^n} + x^n\partial_{y^n} + w^n\partial_{z^n} - z^n\partial_{w^n},$$

$$V_q^2 = -z^0\partial_{x^0} - w^0\partial_{y^0} + x^0\partial_{z^0} + y^0\partial_{w^0} - \dots - z^n\partial_{x^n} - w^n\partial_{y^n} + x^n\partial_{z^n} + y^n\partial_{w^n},$$

$$V_q^3 = -w^0\partial_{x^0} + z^0\partial_{y^0} - y^0\partial_{z^0} + x^0\partial_{w^0} - \dots - w^n\partial_{x^n} - z^n\partial_{y^n} + y^n\partial_{z^n} + x^n\partial_{w^n},$$

at each $q = (x^0, y^0, z^0, w^0, \dots, x^n, y^n, z^n, w^n) \in S^{4n+3}$, with respect to the usual Riemannian metric $g$ on $S^{4n+3}$. It is easy to see that the following commutation relations hold for $V^1, V^2, V^3$

$$[V^1, V^2] = V^3, \quad [V^2, V^3] = V^1, \quad [V^3, V^1] = V^2.$$

Thus one recovers the fact that $\operatorname{span}\{V_q^1, V_q^2, V_q^3\}$ considered as the Lie algebra $\mathfrak{sp}(1)$ is isomorphic to the Lie algebra of the Lie group $S^3$.

All in all, the studied sub-Riemannian manifold is $(S^{4n+3}, D, g_D)$, where $D^{\perp} = V = \operatorname{span}\{V^1, V^2, V^3\}$ with respect to the usual Euclidean metric $g$ in $TS^{4n+3}$ and $g_D$ is the restriction of $g$ to $D$. Compare it with the sub-Riemannian manifold $S^{2n+1}$.

It is an established fact that the distribution $D$ is bracket generating. The geometry of spheres $S^{4n+3}$ is known to be a quaternionic analogue of CR-geometry, see [5]. Note that the vectors $V_q^1$, $V_q^2$, $V_q^3$ coincide with $q.i_1$, $q.i_2$, $q.i_3$, respectively. Here $q.i_k$ is the action of $\mathfrak{sp}(1)$.

To apply Theorem 5.1 in this situation, it is necessary to specify the $\mathfrak{sp}(1)$-valued connection form associated to the Hopf map $h$ from (5.17). In this case, the connection form is given by

$$A(v) = i_1 g(v, V_q^1) + i_2 g(v, V_q^2) + i_3 g(v, V_q^3),$$

where $v \in T_qS^{4n+3}$. The Riemannian metric $g$ is of constant bi-invariant type, since for any $q \in S^{4n+3}$ and $\xi = i_1\alpha_1 + i_2\alpha_2 + i_3\alpha_3 \in \mathfrak{sp}(1)$, $\alpha_k \in \mathbb{R}$, $k = 1, 2, 3$, ($\xi$ is a pure imaginary quaternion) we have

$$\sigma_q(\xi) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} q\exp_{Sp(1)}(\epsilon\xi) = \alpha_1 q.i_1 + \alpha_2 q.i_2 + \alpha_3 q.i_3 = \alpha_1 V_q^1 + \alpha_2 V_q^2 + \alpha_3 V_q^3.$$

Therefore, the moment of inertia tensor, given by

$$\mathbb{I}_q(\xi,\xi) = \mathbb{I}_q\Big(\sum_{k=1}^{3} i_k\alpha_k, \sum_{k=1}^{3} i_k\tilde{\alpha}_k\Big) = g\Big(\sum_{k=1}^{3}\alpha_k V_q^k, \sum_{k=1}^{3}\tilde{\alpha}_k V_q^k\Big) = \sum_{k=1}^{3}\alpha_k\tilde{\alpha}_k,$$

does not depend on the point. As for Proposition 5, we have the following result.

**Proposition 8.** *If $\gamma_R(t) = (q_0(t), \dots, q_n(t))$ is the great circle satisfying $\gamma_R(0) = q$ and $\dot{\gamma}_R(0) = v \in T_qS^{4n+3}$, then the corresponding sub-Riemannian geodesic is given by*

$$\gamma_{sR}(t) = \Big(q_0(t)\cdot e^{-tA(v)}, \dots, q_n(t)\cdot e^{-tA(v)}\Big). \tag{5.18}$$

In Proposition 8, the quaternionic exponential is defined by

$$e^{ai_1+bi_2+ci_3} = \cos\sqrt{a^2+b^2+c^2} + \sin\sqrt{a^2+b^2+c^2}\cdot\frac{ai_1+bi_2+ci_3}{\sqrt{a^2+b^2+c^2}},$$

for $a, b, c \in \mathbb{R}$. Note that the curve $e^{-tA(v)}$ is simply the Riemannian geodesic in $S^3$ starting at the identity of the group $e = (1, 0, 0, 0) \in S^3$, with initial velocity vector $\big(0, -g(v, V_q^1), -g(v, V_q^2), -g(v, V_q^3)\big)$.

**Corollary 5.** *The set of sub-Riemannian geodesics in $S^{4n+3}$ arising from great circles $\gamma_R(t)$, such that $\dot{\gamma}_R(0)$ is orthogonal to* $\mathrm{span}\{V_q^1, V_q^2, V_q^3\}$ *is diffeomorphic to $\mathbb{H}P^n$.*

**Corollary 6.** *Let $\gamma_{sR}\colon \mathbb{R} \to S^{4n+3}$ be a complete sub-Riemannian geodesic parameterized by arc length, with the initial velocity $v \in T_qS^{4n+3}$. Then $\gamma$ is closed if and only if*

$$\frac{g(v, V_q^1)}{\|v\|}, \frac{g(v, V_q^2)}{\|v\|}, \frac{g(v, V_q^3)}{\|v\|} \in \mathbb{Q}.$$

**Corollary 7.** *For the horizontal sub-Riemannian geodesic of the form* (5.18) *the equality $\|\dot{\gamma}_{sR}(t)\|^2 + \|A(v)\|^2 = \|v\|^2$ holds, where $\|A(v)\|^2 = g^2(v, V_q^1) + g^2(v, V_q^2) + g^2(v, V_q^3)$.*

We leave the proofs of Corollaries 5 - 7 as an exercise.

### 5.4. Geodesics related to Yang-Mills fields.

This subsection is aimed at a description of sub-Riemannian geodesics produced by a principal $\mathbb{G}$-bundle $\pi\colon M \to B$ as was described in the Case 2 of Subsection 5.1. Recall that in this case the Ehresmann connection or the horizontal distribution $D$ transversal to the vertical distribution $V = \ker(d\pi)$ is given. Moreover, the sub-Riemannian metric $g_D$ is given as a pullback of the Riemannian metric $g_B$ from

$TB$ to the distribution $D$. We also require that the sub-Riemannian structure $(D, g_D)$ is invariant under the right action of the structure group $\mathbb{G}$. We want to write geodesic equations for the sub-Riemannian manifold $(M, D, g_D)$.

To describe sub-Riemannian geodesics and explain their physical meaning, we need to introduce more definitions related to the notion of a principal bundle. Let the base space $B$ be endowed with a Riemannian metric $g_B$. A Riemannian metric is a positively definite quadratic form and in physics it represents the *kinetic energy* of a system in the space $B$. Consider electromagnetic charged particles (or color-charged particle, or particle with other characteristics) moving in $B$. The information about charges is encoded in a compact Lie group $\mathbb{G}$, that is usually $SU(n)$ in physics. External forces also can be presented by their action on $B$. The motion of the particle is not free since it must respect some symmetries, such as the isometry group on the base space and transformations of the structure group $\mathbb{G}$. Let us avoid constraints as we did solving the Dido problem. We add more variables that allows to inherit the information about the presence of charges encoded in $\mathbb{G}$, so that to each point $b \in B$ we associate a copy of the group $\mathbb{G}_b$. In the enlarged space $M$ we assume that the structure group $\mathbb{G}$ acts freely and transitively such that $M$ receives the structure of a principal $\mathbb{G}$-bundle. The choice of the horizontal distribution is dictated by the external forces acting on the particle on the base space $B$ and it is expressed through the curvature of the connection one form annihilating the horizontal distribution.

The motion on the total space $M$ is governed by a Lagrangian. If the Lagrangian has redundant degrees of freedom or gauges, then the transformations between possible gauges, given by observed physical laws, are called gauge transformations, or gauge symmetries. So gauge transformations are automorphisms of the principal bundle, and they form a group with respect to the composition of the bundle automorphisms. To the principal $\mathbb{G}$-bundle we associate a vector bundle with the same base space where the typical fiber is the representation of $\mathbb{G}$. The obtained vector bundle is called an *associated bundle* and the gauge group $\mathrm{Gau}(M)$ consists of all smooth sections of the associated bundle. The trivial section $b \to \mathrm{id}_b$ of the associated bundle corresponds to the identity bundle automorphism, see [81, 108].

The electromagnetic charged particles are described by the theory of principal $U(1)$-bundles. We can think of charges as of elements of the space $\mathfrak{g}^*$ dual to the Lie algebra $\mathfrak{g}$ corresponding to one dimensional structure groups $U(1)$ or $\mathbb{R}$. In mathematical theory charges are elements of dual Lie algebras, because they fit better to the situation in which the geodesics are produced by bi-characteristics on the co-tangent bundle $T^*M$ of $M$. Yang and Mills [105] proposed the theory generalizing the gauge theory from principal $U(1)$-bundles to principal $U(n)$-bundles. For instance, $SU(2)$ symmetry group is used in the isospin model, $SU(2) \times U(1)$ symmetry group describes electroweak interaction, and $SU(3)$ symmetry group is the subject of quantum cromodynamics.

The $\mathbb{G}$-group action on $M$ with $\mathbb{G}$ invariant Riemannian metric, produces an action on $T^*M$ and both the Hamiltonian function and the flow on $T^*M$ corresponding to the Hamiltonian vector field are invariant under this action. Thus, it seems natural to reduce the space $T^*M$ to the space of orbits $T^*M/\mathbb{G}$ and consider the flow on the reduced space $T^*M/\mathbb{G}$. This reduction is called the Poisson reduction. The idea to consider the reduction of spaces endowed with some structures (Poisson, symplectic, Kähler) comes from works [100, 101]. The dynamics on the reduced space is related to sub-Riemannian geodesics on $M$, whose projections on the base space $B$ are trajectories of the motion of charged particles in Yang - Mills fields and the corresponding equations are called the Wong equations.

In the case of a one dimensional structure group the dynamic is quite well known, since the reduced space $T^*M/\mathbb{G}$ is diffeomorphic to $T^*B\oplus\mathbb{R}$ and for each fixed value of the charge we get its level set in $T^*M/\mathbb{G}$ that is diffeomorphic to $T^*B$. This level sets are glued together to form the entire reduced space $T^*M/\mathbb{G}$.

For a non-abelian group $\mathbb{G}$ acting on $M$ the structure of the reduced space $T^*M/\mathbb{G}$ is more complicated and it is isomorphic to $T^*B\oplus Ad^*(M)$, where we have to change the quite simple component $\mathbb{R}$ representing the abelian charge to the bundle $Ad^*(M)$ over the same base space $B$ that is associated with the principal $\mathbb{G}$-bundle $\pi\colon M\to B$.

The motion of a "free" particle on the base space $B$ means absence of forces acting there and the trajectory of "free motion" is the geodesic given by the equation

$$\nabla_{\dot c}\dot c=0,\quad \nabla\quad \text{is the Levi-Civita connection on}\quad B.$$

If a force $F$ is present, then the equation changes to the Newton equation

$$\nabla_{\dot c}\dot c=F$$

of geodesics on the manifold $B$ for a particle of constant charge and unit mass. If the charge is non-abelian, then it is encoded in the bundle $Ad^*(M)$ and the right hand side of the last equation depends in a complicated way on charge. Moreover, the condition on the level sets of an abelian charge changes to the requirement to be a "co-variantly constant" charge. To formulate the Wong equation supplemented by the conservation condition for the charge we start from necessary definitions.

**5.4.1. Structure of the reduced space.** INDUCED ACTION OF GROUP ON TANGENT AND CO-TANGENT BUNDLES. Let $\pi\colon M\to B$ be a principle $\mathbb{G}$-bundle. The right action of $\mathbb{G}$ on $M$ produces right actions on $TM$ and $T^*M$. They defined by the following:

$$\begin{array}{rccc}\mu\colon & TM\times\mathbb{G} & \to & TM\\ & (q,v).\tau & \mapsto & \big(q.\tau, dr_{\tau}(v)\big),\end{array}\tag{5.19}$$

and

$$\begin{array}{rccc}\mu\colon & T^*M\times\mathbb{G} & \to & T^*M\\ & (q,\omega).\tau & \mapsto & \big(q.\tau, \big(dr_{\tau}\big)^*(\omega)\big),\end{array}\tag{5.20}$$

where $(dr_\tau)^*$ is the dual operator to the differential $dr_\tau$ of the right translation $r$ by $\tau\in\mathbb{G}$.

The factorization of $TM$ by the action (5.19) leads to the factor space $TM/\mathbb{G}$ with elements $[q,v]$. Define the projection $\widetilde{\pi}\colon TM/\mathbb{G}\to B$ by

$$\widetilde{\pi}([q,v]) := \pi(q)\in B, \qquad [q,v]\in TM/\mathbb{G}.$$

We get a vector bundle over the base space $B$, where we will denote the projection $\widetilde{\pi}$ simply by $\pi$. Thus, we have $\pi\colon TM/\mathbb{G}\to B$. Analogously, taking the factor of $T^*M$ by the action (5.20) of $\mathbb{G}$, we get a vector bundle $\pi\colon T^*M/\mathbb{G}\to B$.

We aim to construct THE BUNDLE MAP $TM/\mathbb{G}\to TB$. The principal $\mathbb{G}$-bundle

$$\pi\colon M\to B$$

after differentiating leads to the bundle map

$$d\pi\colon TM\to\ TB.$$

Let us take the factor by the action of $\mathbb{G}$ of both parts. The action of $\mathbb{G}$ over $TB$ is trivial: $TB/\mathbb{G}=TB$. Thus we get a bundle map

$$\begin{tikzcd}
W_b \arrow[r, "d_q\pi"] \arrow[d, leftrightarrow] & T_bB \arrow[d, leftrightarrow] \\
TM/\mathbb{G} \arrow[r, "d\pi"] \arrow[d, "\mathrm{pr}_B"'] & TB \arrow[d, "\mathrm{pr}_B"] \\
B & B.
\end{tikzcd}$$

By making use of the dual map $(d\pi)^*$, we get an analogous bundle map

$$\begin{tikzcd}
W_b^* \arrow[d, leftrightarrow] & T_b^*B \arrow[l, "(d_q\pi)^*"'] \arrow[d, leftrightarrow] \\
T^*M/\mathbb{G} \arrow[d, "\mathrm{pr}_B^*"'] & T^*B \arrow[l, "(d\pi)^*"'] \arrow[d, "\mathrm{pr}_B^*"] \\
B & B.
\end{tikzcd}$$

Actually, we need to verify that the maps $d\pi$ and $(d\pi)^*$ are equivariant with respect to the action of the group $\mathbb{G}$:

$$d\pi\big((q,v_q).\tau\big) = d\pi(q,v_q).\tau = (\pi(q), d\pi(v_q)) = (b,w_b),$$

where $(q,v_q)\in T_qM$ and $(b,w_b)\in T_{\pi(q)}B$. The group $\mathbb{G}$ acts on $TM$ on the right and since the action of $\mathbb{G}$ on $TB$ is trivial, we also can suppose that it acts on the right. To show the equivariance we recall that the decomposition $D\oplus V=TM$ is preserved under the action of $\mathbb{G}$. Therefore

$$\begin{aligned}
d\pi((q,v_q).\tau) &= d\pi((q.\tau, dr_\tau(v_q|_D+v_q|_V)) = \big(\pi(q.\tau), d\pi(v_{q.\tau}|_D)+d\pi(v_{q.\tau}|_V)\big)\\
&= (\pi(q), w_{\pi(q.\tau)}) = (b,w_b),
\end{aligned}$$

where $d\pi(v_{q.\tau}|_V)$ vanishes. The proof for the bundle map $(d\pi)^*\colon T^*B \to T^*M/\mathbb{G}$ is similar.

Observe that $\ker(d_q\pi) = V_q$, $q \in M$, $b = \pi(q)$, and the typical fiber $W_b$ of the bundle $TM/\mathbb{G}$ splits into parts isomorphic to the vertical $V_q$ and horizontal $D_q$ spaces. Moreover, $D_q$ is isomorphic to the typical fiber $T_bB$ of the bundle $TB$. Sections of the bundle $\mathrm{pr}_B\colon TB \to B$ are vector fields on $B$. Sections of the bundle $\mathrm{pr}_B\colon TM/\mathbb{G} \to B$ are right invariant vector fields with respect to the action of $\mathbb{G}$. We finish the construction of the bundle map $TM/\mathbb{G} \to TB$.

Now we find the BUNDLE MAP $TB \to TM/\mathbb{G}$. We start from recalling that there exists a bundle map $h\colon TB \to TM$ that we called the horizontal lift such that the image $h_q(T_bB)$ is $D_q \subset T_qM$, where $D_q \oplus V_q = T_qM$. To show that the map $h$ is equivariant under $\mathbb{G}$ we take a point $(q, v_q|_D) \in T_qM$ and its pre-image $(b, w_b) = (\pi(q), w_{\pi(q)})$. Then on the one hand

$$h((b,w_b).\tau) = h((b,w_b)) = (q, v_q|_D) \tag{5.21}$$

since the action of $\mathbb{G}$ on $TB$ is trivial. On the other hand

$$\begin{aligned} h(b,w_b).\tau &= (q, v_q|_D).\tau = (q.\tau, dr_\tau(v_q|_D)) \\ &= h(\pi(q.\tau), w_{\pi(q.\tau)}) = h((b,w_b)) = (q, v_q|_D). \end{aligned} \tag{5.22}$$

The chains of equalities (5.21) and (5.22) show that $h$ is an equivariant map. Roughly speaking, the horizontal lift $h$ is the inverse map for $d\pi|_D$:

$$\begin{tikzcd} V_q \oplus D_q = T_qM \arrow[r, "d_q\pi"] & T_{\pi(q)}B. \arrow[l, bend right=60, "h_q"'] \end{tikzcd}$$

Since the map $h$ is equivariant, we can take the factor of $h\colon TB \to TM$ by the action of $\mathbb{G}$ and get the induced bundle map, that we again call $h$:

$$\begin{tikzcd} W_b \sim V_q \oplus D_q \arrow[d, leftrightarrow] & T_bB \arrow[l, "h_q"'] \arrow[d, leftrightarrow] \\ TM/\mathbb{G} \arrow[d, "\mathrm{pr}_B"'] & TB \arrow[l, "h"'] \arrow[d, "\mathrm{pr}_B"] \\ B & B. \end{tikzcd}$$

The bundle map

$$\begin{tikzcd}
W_b^* \sim V_q^* \oplus D_q^* \arrow[r, "h_q^*"] \arrow[d, leftrightarrow] & T_b^* B \arrow[d, leftrightarrow] \\
T^*M/\mathbb{G} \arrow[r, "h^*"] \arrow[d, "\mathrm{pr}_B^*"'] & T^*B \arrow[d, "\mathrm{pr}_B^*"] \\
B & B.
\end{tikzcd}$$

is produced similarly. Here $D_q = \mathrm{Im}(h_q)$ is isomorphic to $T_bB$, $b = \pi(q)$, and $V_q = \ker(d_q\pi)$. At the co-tangent bundles level we get $D_q^* = \mathrm{Im}(d_q\pi^*)$, $V_q^* = \ker(h_q^*)$ and $D_q^*$ is isomorphic to $T_b^*B$.

Resuming the discussion of the two last parts, we conclude that we constructed two maps between bundles $TM/\mathbb{G}$ and $TB$, and $T^*M/\mathbb{G}$ and $T^*B$:

$$\begin{tikzcd}
TM/\mathbb{G} \arrow[r, bend left, "d\pi"] & TB \arrow[l, bend left, "h"]
\end{tikzcd}, \qquad \begin{tikzcd}
T^*M/\mathbb{G} \arrow[r, bend left, "h^*"] & T^*B. \arrow[l, bend left, "d\pi^*"]
\end{tikzcd}$$

In these maps the horizontal distribution $D$ is the image of some map, meanwhile the vertical part is the kernel of some other mapping. In the next step we change the role of $D$ and $V$.

ADJOINT AND CO-ADJOINT BUNDLES, ASSOCIATED WITH THE PRINCIPAL BUNDLE $\pi\colon M \to B$. Let us suppose that a principal $\mathbb{G}$-bundle $\pi\colon M \to B$ is given. Recall, that in this case

- the typical fiber $F$ is isomorphic to the group $\mathbb{G}$,
- the group $\mathbb{G}$ acts on $F$ by right (or left) translations.

Then it is possible to define the associate bundle, where

- the typical fiber $F$ is isomorphic to some vector space $E$,
- some other action of $\mathbb{G}$ on $E$ is defined.

It is achieved in general through the representation of $\mathbb{G}$ on $E$. We will give the definition only in the particular case, when $E = \mathfrak{g}$ (or $E = \mathfrak{g}^*$) and the action of the Lie group $\mathbb{G}$ is the adjoint action on its Lie algebra $\mathfrak{g}$ (or the co-adjoint action on its dual Lie algebra $\mathfrak{g}^*$).

The adjoint fiber bundle $\mathrm{Ad}(M)$ to a principal $\mathbb{G}$-bundle $\pi\colon M \to B$ is the vector bundle $\widehat{\pi}\colon \mathrm{Ad}(M) \to B$ with a typical fiber isomorphic to $\mathfrak{g}$. The action of the group $\mathbb{G}$ on the fiber $\mathfrak{g}$ is defined by the adjoint action $\mathfrak{g} \ni \xi \to \mathrm{Ad}_\tau(\xi) \in \mathfrak{g}$ for all $\tau \in \mathbb{G}$. To construct the adjoint bundle $\mathrm{Ad}(M)$ one starts from the direct product $M \times \mathfrak{g}$ and then, taking factor by the right action of $\mathbb{G}$ defined on $M \times \mathfrak{g}$ by

$$\begin{array}{rccc} \mu\colon & (M \times \mathfrak{g}) \times \mathbb{G} & \to & M \times \mathfrak{g} \\ & (q, \xi).\tau & \mapsto & \big(q.\tau, Ad_{\tau^{-1}}(\xi)\big). \end{array} \tag{5.23}$$

Here we used that the group $\mathbb{G}$ acts on the right on $M$ and that the adjoint action on $\mathfrak{g}$ is the left action since it comes as a differential of the left action $a$

(by conjugation), see (8.2) and (8.5). This definition of the action is compatible with the definition of an equivariant (right-left) map. The adjoint bundle $\mathrm{Ad}(M)$ is produced by factoring $M \times \mathfrak{g}$ by the right action (5.23). The standard notations in the literature are $\mathrm{Ad}(M)$ or $M \times_{\mathrm{Ad}} \mathfrak{g}$. The equivalence class $[q, \xi] \in \mathrm{Ad}(M)$ of the representative $(q, \xi) \in M \times \mathfrak{g}$ is also often written as $q\xi$ due to the mnemonic cancelation rule $\big(q.\tau, Ad_{\tau^{-1}}(\xi)\big) = q\xi$, where $\tau$ is canceled. The projection map $\widehat{\pi}$ from $\mathrm{Ad}(M)$ to the base space $B$ is defined by $\widehat{\pi}([q, \xi]) := \pi(q)$.

The *co-adjoint bundle* $\mathrm{Ad}^*(M)$ is the vector bundle $\check{\pi}\colon \mathrm{Ad}^*(M) \to B$ with the typical fiber $\mathfrak{g}^*$. The bundle $\mathrm{Ad}^*(M)$ is obtained by division of $M \times \mathfrak{g}^*$ by the right action of $\mathbb{G}$ on $M \times \mathfrak{g}^*$:

$$\begin{array}{cccc} \mu\colon & (M \times \mathfrak{g}^*) \times \mathbb{G} & \to & M \times \mathfrak{g}^* \\ & (q, \omega).\tau & \mapsto & \big(q.\tau, Ad^*_{\tau^{-1}}(\omega)\big) \end{array}, \tag{5.24}$$

The next step is to reveal relations between the adjoint bundle $\mathrm{Ad}(M)$ and the vector bundle $TM/\mathbb{G}$.

Find the BUNDLE MAP $\mathrm{Ad}(M) \to TM/\mathbb{G}$.

As usual, we start from the principal $\mathbb{G}$-bundle, where the group $\mathbb{G}$ acts on the right on $M$. The right translation $r$ generates the infinitesimal generator

$$\sigma_q \colon \mathfrak{g} \to V_q \subset T_qM,$$

(see Example 12). Let us vary $q \in M$ and we get a bundle map

$$\begin{tikzcd}
\mathfrak{g} \arrow[r, "\sigma_q"] \arrow[d, leftrightarrow] & V_q \subset T_qM \arrow[d, leftrightarrow] \\
M \times \mathfrak{g} \arrow[r, "\sigma"] \arrow[d, "\mathrm{pr}_M"'] & V \subset TM \arrow[d, "\mathrm{pr}_M"] \\
M & M.
\end{tikzcd}$$

The map $\sigma$ is equivariant:

$$\sigma((q, \xi).\tau) := \sigma(q.\tau, \mathrm{Ad}_{\tau^{-1}}(\xi)) = dr_\tau \sigma(q, \xi) := \sigma(q, \xi).\tau, \quad \tau \in \mathbb{G}.$$

Indeed,

$$\begin{aligned} \sigma(q.\tau, \mathrm{Ad}_{\tau^{-1}}(\xi)) &= \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} q\tau\, \mathrm{Ad}_{\tau^{-1}}(\exp(\varepsilon\xi)) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} q\tau(\tau^{-1}\exp(\varepsilon\xi)\tau) \\ &= dr_\tau dl_q(\xi) = dr_\tau \sigma(q, \xi). \end{aligned}$$

Dividing by actions (5.23) and (5.24) we come to the bundle map

$$\begin{tikzcd}
\mathfrak{g} \arrow[r, "\sigma_q"] \arrow[d, leftrightarrow] & V_q \subset T_qM \arrow[d, leftrightarrow] \\
\big(M \times \mathfrak{g}\big)/\mathbb{G} \arrow[r, "\sigma"] \arrow[d, "\mathrm{pr}_M"'] & V/\mathbb{G} \subset TM/\mathbb{G} \arrow[d, "\mathrm{pr}_M"] \\
M/\mathbb{G} & M/\mathbb{G}.
\end{tikzcd}$$

or

$$\begin{tikzcd}
\mathfrak{g} \arrow[r, "\sigma_q"] \arrow[d, leftrightarrow] & V_q \subset T_qM \arrow[d, leftrightarrow] \\
\mathrm{Ad}(M) \arrow[r, "\sigma"] \arrow[d, "\mathrm{pr}_B"'] & V/\mathbb{G} \subset TM/\mathbb{G} \arrow[d, "\mathrm{pr}_B"] \\
B & B.
\end{tikzcd}$$

Thus, the image $\mathrm{Im}(\sigma) = V/\mathbb{G}$ is the collection of right invariant vertical vector spaces.

Construction of the BUNDLE MAP $TM/\mathbb{G} \to \mathrm{Ad}(M)$. The auxiliary map here is the $\mathfrak{g}$-valued connection one-form $A_q \colon T_qM \to \mathfrak{g}$ introduced in (5.8). The connection form is uniquely defined by two conditions

1. $\ker(A_q) = D_q$ and
2. $A_q \circ \sigma_q = \mathrm{Id}_{\mathfrak{g}}$ for any $q \in M$.

The second condition says that after the projection on the vertical space $V$ the connection form is the canonical identification between $V$ and the Lie algebra $\mathfrak{g}$. The map $A \colon TM \to M \times \mathfrak{g}$ is equivariant:

$$\begin{aligned} A\big((q, v|_V).\tau\big) : &= A(q.\tau, dr_\tau(v|_V)) \\ &= A(q, v|_V).\tau = (q, \xi).\tau := \big(q.\tau, \mathrm{Ad}_{\tau^{-1}}(\xi)\big), \end{aligned} \tag{5.25}$$

where we consider only the vertical part $v|_V$ of a vector $v \in T_qM$, since the horizontal part belong to the kernel of $A_q$. To prove (5.25), we note that $v|_V = \sigma_q(\xi)$ and $A_q(v|_V) = \xi$ from the Property 2. This and equivariance of $\sigma$ imply the equivariance of $A$ as follows

$$\begin{aligned} A(q.\tau, dr_\tau(v|_V)) &= A\big(q.\tau, dr_\tau \sigma(q, \xi)\big) = A\big(q.\tau, \sigma(q.\tau, \mathrm{Ad}_{\tau^{-1}}(\xi)\big) \\ &= \big(q.\tau, \mathrm{Ad}_{\tau^{-1}}(\xi)\big) = (q, \xi).\tau = A(q, v|_V).\tau. \end{aligned}$$

The factorization by the action of $\mathbb{G}$ leads to the bundle map

$$\begin{tikzcd}
\mathfrak{g} \arrow[d, leftrightarrow] & T_qM \arrow[l, "A_q"'] \arrow[d, leftrightarrow] \\
\mathrm{Ad}(M) \arrow[d, "\mathrm{pr}_B"'] & TM/\mathbb{G} \arrow[l, "A"'] \arrow[d, "\mathrm{pr}_B"] \\
B & B,
\end{tikzcd}$$

where $\ker(A) = D/\mathbb{G}$. The construction of dual bundle maps is straightforward.

The isomorphism of bundles $TM/\mathbb{G}$ and $TB \oplus \mathrm{Ad}(M)$. Summarizing everything for the constructed bundle maps, we get

$$\begin{tikzcd}
\mathfrak{g} \arrow[r, bend left, "\sigma_q"] \arrow[d, leftrightarrow] & W_q \arrow[l, bend left, "A_q"] \arrow[r, bend left, "d_q\pi"] \arrow[d, leftrightarrow] & T_bB \arrow[l, bend left, "h_q"] \arrow[d, leftrightarrow] & \text{for} \quad b = \pi(q). \\
\mathrm{Ad}(M) \arrow[r, bend left, "\sigma"] & TM/\mathbb{G} \arrow[l, bend left, "A"] \arrow[r, bend left, "d\pi"] & TB \arrow[l, bend left, "h"] &
\end{tikzcd}$$

Moreover,

$$\mathrm{Im}(h_q) = D_q = \ker(A_q), \qquad \mathrm{Im}(\sigma_q) = V_q = \ker(d_q\pi),$$

$$A_q \circ \sigma_q = \mathrm{Id}_{\mathfrak{g}}, \qquad d\pi_q \circ h_q = \mathrm{Id}_{T_bB}\,.$$

So, we conclude that the typical fiber $W_q$ is isomorphic to the product $\mathfrak{g} \times T_bB$ that leads to the isomorphism of bundles

$$\begin{array}{ccccc}
TM/\mathbb{G} & \cong & TB & \oplus & \mathrm{Ad}(M) \\
\big[q, \big(h_q(w) + \sigma_q(\xi)\big)\big] & \longleftarrow & (b, w) & \oplus & [q, \xi] \\
\big[q, v\big] & \longrightarrow & \big(\pi(q), d_q\pi(v)\big) & \oplus & [q, A_q(v)].
\end{array}$$

Here $h_q(w) \in D_q$ is the horizontal lift of $w \in T_{\pi(q)}B$ and $\sigma_q(\xi) \in V_q$ is a vector field. The result $d_q\pi(v)$ represents the horizontal part of $v \in T_qM$ and $A_q(w)$ is the missing vertical part regarded as an element of $\mathfrak{g}$. Let us emphasize one more time that the constructed bundle isomorphism is induced by chosen Ehresmann connection $D$ that is invariant under the action of the structural group $\mathbb{G}$.

**Remark 7.** Let us present the gauge group $\mathrm{Gau}(M)$ acting on the principal $\mathbb{G}$-bundle $\pi\colon M \to B$. To do this, let us recall that a section of a fiber bundle $\varphi\colon E \to B$ is any function $f\colon B \to E$ satisfying $\varphi \circ f = \mathrm{Id}_B$. The bundle isomorphism $\sigma\colon \mathrm{Ad}(M) \leftrightarrow V/\mathbb{G} \subset TM/\mathbb{G}$ gives a correspondence between the set $\mathcal{S}$ of all sections of the adjoint bundle $\pi\colon \mathrm{Ad}(M) \to B$ and the set $\mathcal{L}$ of $\mathbb{G}$ invariant vertical vector fields on $M$, since the set $\mathcal{L}$ is the set of sections of the bundle $\pi\colon V/\mathbb{G} \to B$. Thus the bundle map $\sigma$ induces the isomorphism $\widetilde{\sigma}\colon \mathcal{S} \leftrightarrow \mathcal{L}$.

Another observation: if there is a $\mathbb{G}$-invariant map $\Phi\colon M \to M$, then factorization by the action of the group $\mathbb{G}$ induces the map $\widetilde{\Phi}\colon M/\mathbb{G} = B \to M/\mathbb{G} = B$.

**Definition 24.** *The gauge group* $\mathrm{Gau}(M)$ *of the principal* $\mathbb{G}$*-bundle* $\pi\colon M \to B$ *is the set of all* $\mathbb{G}$*-invariant maps* $\Phi\colon M \to M$ *such that the induced map* $\widetilde{\Phi}$ *is the identity on* $B$ : $\widetilde{\Phi} = \mathrm{Id}_B$.

The Lie algebra $\mathfrak{gau}(M)$ for the gauge group $\mathrm{Gau}(M)$ is the set of $\mathbb{G}$-invariant vector fields on $M$, which is exactly the set $\mathcal{L}$, which is $\widetilde{\sigma}$-isomorphic to $\mathcal{S}$. Conclusion: *the sections of the adjoint bundle* $\mathrm{Ad}(M)$ *form the Lie algebra of the gauge group.*

The physicists are actually interested in working with the dual isomorphism $T^*M/\mathbb{G} \cong T^*B \oplus \mathrm{Ad}^*(M)$, where in the construction of the co-adjoint action on $\mathfrak{g}^*$ and dual maps $\sigma^*$, $A^*$ to $\sigma$, $A$ are used. Roughly speaking, the bundle isomorphism $T^*M/\mathbb{G} \cong T^*B \oplus \mathrm{Ad}^*(M)$ is the splitting of the reduced phase space into the vertical part encoded in $\mathrm{Ad}^*(M)$ (generated by the action of $\mathbb{G}$) and the complementary horizontal part isomorphic to $T^*B$.

### 5.4.2. The Wong equation.

The main goal of this part is to introduce the geodesic equation on the reduced space $T^*M/\mathbb{G}$. To do this we need to introduce the curvature form $\Omega$ of the Ehresmann connection $D$ and the covariant derivative for sections on the co-adjoint bundle $\mathrm{Ad}^*(M)$.

Curvature form. A curvature two-form $\Omega$ on $M$ associated with the Ehresmann connection $D$ measures the behavior of two horizontal vector fields $X, Y \in D$ with respect to each other, or more precisely, $\Omega$ maps a pair $X, Y \in \bigwedge^2 D$ to a vector $(-[X,Y] \mod D) \in V$. Since the connection one-form $A$ gives an identification of $V$ with the Lie algebra $\mathfrak{g}$, we define the $\mathfrak{g}$-valued two-form $\Omega_q$ by

$$\Omega_q(X,Y) = \begin{cases} -A_q([X,Y]) & \text{if} \quad X, Y \in D_q, \\ 0 & \text{othewise,} \end{cases} \qquad \text{for all} \quad q \in M.$$

Since the form $A$ is defined on $\mathbb{G}$-invariant vector fields, we get $-A([X,Y]) = dA(X,Y) - [A(X), A(Y)]$ and we conclude that

$$\Omega = dA - [A, A].$$

The form $\Omega$ is $\mathbb{G}$-equivariant, since the form $A$ is so. The equivariance of $\Omega$ allows to extend the definition of $\Omega$ to the bundle map

$$\begin{tikzcd}
\bigwedge^2 D_q \arrow[r, "\Omega_q"] \arrow[d, leftrightarrow] & \mathfrak{g} \arrow[d, leftrightarrow] & \text{factoring by} \quad \mathbb{G} & \bigwedge^2 D_q \arrow[r, "\Omega_q"] \arrow[d, leftrightarrow] & \mathfrak{g} \arrow[d, leftrightarrow] \\
\bigwedge^2 D \arrow[r, "\Omega"] \arrow[d, "\mathrm{pr}_M"'] & M \times \mathfrak{g} \arrow[d, "\mathrm{pr}_M"] & & \left(\bigwedge^2 D\right)/\mathbb{G} \arrow[r, "\Omega"] \arrow[d, "\mathrm{pr}_B"'] & \mathrm{Ad}(M) \arrow[d, "\mathrm{pr}_B"] \\
M & M & & B & B.
\end{tikzcd}$$

The last property is reflected in the name of the curvature two-form $\Omega$ as $\mathrm{Ad}(M)$-*valued curvature form.*

Covariant derivative for sections on the co-adjoint bundle $\mathrm{Ad}^*(M)$. Let $\psi\colon B \to \mathrm{Ad}(M)$ be a section, and $X \in \mathrm{Vect}\, B$, then we would like to define the covariant derivative $\mathcal{D}_X\psi$ that for any chosen $\psi$ and $X$ gives a section of the adjoint bundle $\mathrm{Ad}(M)$ that is

$$\mathcal{D}_X\psi\colon B \to \mathrm{Ad}(M).$$

To define it we first relate a map $F\colon M\to\mathfrak{g}$ to $\psi$ by the following

$$B\ni b\xrightarrow{\psi}(q,\xi)=(q,F(q))\in\mathrm{Ad}(M),$$

where $b=\pi(q)$, $q\in M$, and $\xi=F(q)\in\mathfrak{g}$. The map $F$ should be equivariant: $F(q.\tau)=F(q).\tau=\mathrm{Ad}_{\tau^{-1}}(F(q))$, since $\psi$ is equivariant by definition. The differential map

$$\begin{array}{rccc} d_qF\colon & T_qM & \to & T_{F(q)}\mathfrak{g}\cong\mathfrak{g} \\ & (q,v) & \mapsto & (q,d_qF(v)) \end{array}$$

leads to the bundle map $dF\colon TM\to M\times\mathfrak{g}$. If we show that $dF$ is equivariant, then by taking factor by the action of group, we get a bundle map

$$dF\colon TM/\mathbb{G}\to\mathrm{Ad}(M)$$

that sends any point $(q,Y)$, with $q\in M$ and $Y$ being a $\mathbb{G}$-invariant vector field, to the point $(q,d_qF(Y))\in\mathrm{Ad}(M)$, where $F\colon M\to\mathfrak{g}$ is equivariant. But $q\in M$ is such that $\pi(q)=b\in B$, and $Y$ is the horizontal lift by $h$ of some vector field $X\in\operatorname{Vect}B$. So given $b\in B$, $X\in\operatorname{Vect}B$, the composition $dF(h\circ X)$ is the desired section

$$\psi\colon B\overset{X}{\to}TB\overset{h}{\to}TM/\mathbb{G}\overset{dF}{\to}\mathrm{Ad}(M).$$

The following chain of equalities shows that $dF\colon TM\to M\times\mathfrak{g}$ is equivariant

$$dF(q.\tau,dr_\tau h(X))=\big(q.\tau,dF_{q.\tau}(hX(q.\tau))\big)=dF(q,h(X)).\tau$$

due to the equivarience of $F$.

THE WONG EQUATION. Let $c\colon I\to B$ be a curve on the base space, that represents the trajectory of the motion of some charged particle. Then $\dot c(t)\in\operatorname{Vect}B$ is the vector field along $c$ and the horizontal lift $h$ sends this vector field to $D_{\gamma(t)}\subset T_{\gamma(t)}M$, or

$$T_{(c(t))}B\ni(c(t),\dot c(t))\overset{h}{\to}(\gamma(t),h(\dot c(t)))\in D_{\gamma(t)}\subset T_{\gamma(t)}M,\quad \pi(\gamma(t))=c(t).$$

The contraction $i_{\dot c}\Omega$ of $\mathrm{Ad}(M)$-valued two-form $\Omega$ on $h(\dot c)$ is $\mathrm{Ad}(M)$-valued one-form along $c$:

$$i_{\dot c}\Omega(\cdot):=\Omega(h(\dot c),\cdot)\colon TB\to\mathrm{Ad}(M).$$

Take any section $\lambda(t)=\lambda(c(t))$ of the co-adjoint bundle $\mathrm{Ad}^*(M)$ along the curve $c$ on the base space $B$. The section $\lambda$ represents the charge (electromagnetic or color-charge) of a particle moving in $B$. The duality of $\mathrm{Ad}(M)$ and $\mathrm{Ad}^*(M)$ produces a momentum in $T^*B$, by

$$\Lambda_{\dot c,\Omega}(\cdot):=\langle\lambda(c(t)),i_{\dot c}\Omega(\cdot)\rangle.$$

The pairing $\langle\cdot\,,\cdot\rangle$ associates to each vector field $X\in TB$ a real number given by the pairing between $\lambda(c(t))\in\mathrm{Ad}^*(M)$ and $i_{\dot c(t)}\Omega(h(X))\in\mathrm{Ad}(M)$ along the curve $c$. Thus $\Lambda_{\dot c,\Omega}\in T^*B$. Now we exploit the Riemannian metric $g_B$ given on the base space $B$ and find the canonical dual $\Lambda^{\#}_{\dot c,\Omega}\in TB$ to $\Lambda_{\dot c,\Omega}$. The vector field $\Lambda^{\#}_{\dot c,\Omega}$ represents a force acting on the base space $B$ produced by the charge $\lambda$ of the particle. This force is called the *Lorentz force.*

The presence of the non-abelian Lorentz force $\Lambda^{\#}_{\dot{c},\Omega}$ leads to the equation called the *Wong equation*,

$$\nabla_{\dot{c}}\dot{c} = \Lambda^{\#}_{\dot{c},\Omega}, \tag{5.26}$$

where the non-abelian charge $\lambda$ has to satisfy the condition

$$\mathcal{D}_{\dot{c}(t)}\lambda(c(t)) = 0, \quad \text{for any} \quad t \in I. \tag{5.27}$$

expressing the property that the charge has to be "co-variantly constant" along $c$.

The *second order* differential equations (5.26) and (5.27) have the solution that is the curve $(c(t), \lambda(c(t))) \in \mathrm{Ad}^*(M)$, $t \in I$, in the co-adjoint bundle lying over the curve $c$ in the base space $B$. The system of equations can be rewritten as the first order system on $T^*B \oplus \mathrm{Ad}^*(M)$ by introducing the momentum $p(t) = p(c(t)) \in T^*_{c(t)}B$ along the base curve $c$. We again use the metric tensor $g_B$ and define the co-vector $p(t)(\cdot) = g_B(\dot{c}(t), \cdot)$. Then $(p, \lambda) \in T^*B \oplus \mathrm{Ad}^*(M)$.

Since we have the isomorphism $TM/\mathbb{G} \cong T^*B \oplus \mathrm{Ad}^*(M)$, the solution $(p, \lambda)$ on $T^*B \oplus \mathrm{Ad}^*(M)$ is also solution on $TM/\mathbb{G}$. The following theorem is an analogue of Theorem 5.1 produced for the Case 2 and it expresses the relation between sub-Riemannian geodesics in Case 2 and the solutions of the Wong equations.

**Theorem 5.2.** [111] *Let $\Gamma = (\gamma, p)$ be a normal sub-Riemannian bi-characteristic for $(M, D, g_D)$, produced by the principal $\mathbb{G}$-bundle $\pi\colon M \to B$, where the sub-Riemannian structure $(D, g_D)$ is $\mathbb{G}$-invariant and $g_D$ is the pullback of the Riemannian metric $g_B$ on $B$. The projection of $\Gamma$ onto $TM/\mathbb{G} \cong T^*B \oplus \mathrm{Ad}^*(M)$ is the solution of the Wong equations* (5.26) *and* (5.27).

*Conversely, let $c(t) \in B$, $t \in I$, be a solution of the Wong equation* (5.26) *complemented by* (5.27) *and $h(c(t))$ be its horizontal lift to $M$. Then $h(c(t))$ is a normal sub-Riemannian geodesic for $(M, D, g_D)$ described as above. The non-abelian charge $\lambda \in \mathrm{Ad}^*(M)$ corresponds to the co-vector $p \in T^*M$.*

Instead of presenting the proof of the theorem, that can be found in [111], we show the relation between this theorem and the examples of Carnot groups considered in Section 3.

### 5.5. Examples of solutions to the Wong equations

Before we present the examples, let us do some observations about the Wong equation (5.26).

CONSERVATION OF ENERGY. By definition, $\Lambda^{\#}_{\dot{c},\Omega}(v) = \langle \lambda(c(t)), \Omega(h(\dot{c}), h(v))\rangle$. Since the form $\Omega$ is skew symmetric, then for $v = \dot{c}$ we get $\Lambda^{\#}_{\dot{c},\Omega}(\dot{c}) = 0$ that leads to $\langle \nabla_{\dot{c}}\dot{c}, \dot{c}\rangle = 0$ by the Wong equation (5.26). Since the derivative of the kinetic energy is $\frac{d}{dt}\Big(\frac{1}{2}\langle \dot{c}, \dot{c}\rangle\Big) = \langle \nabla_{\dot{c}}\dot{c}, \dot{c}\rangle = 0$, we conclude that the kinetic energy is constant along solutions of (5.26). This energy is equal to the value of the sub-Riemannian Hamiltonian function along the corresponding sub-Riemannian geodesic.

RELATION TO PHYSICS. If the group $\mathbb{G}$ acting on $M$ is abelian, then the Wong equations are known as the Lorentz equations. In this case the adjoint bundle

$\mathrm{Ad}^*(M)$ is the trivial bundle $M \times \mathfrak{g}^*$, where elements of the fiber $\mathfrak{g}^*$ represent charges. The condition of the covariant constancy (5.27) asserts that the charge $\lambda(t)$ is constant. The rest (5.26) is the family of equations, parametrized by the charge. These sub-Riemannian geodesics, corresponding to the bundle type structures are projected to the motion of a particle on the base manifold $B$ under the influence of the magnetic field $\Omega$, defined by the curvature of the horizontal distribution $D$. We get a family of curves parametrized by the charge $\lambda$. This observation is one of the main parts of the Kaluza-Klein theory. The Lorentz equations play an important role in classical electrodynamics and both the classical and quantum versions of electromagnetism are highly successful physical theories.

The quantum version of non-abelian gauge theory are quite successful and actively developing subject as we discussed at the beginning of the section. The interesting peculiarity is, that in contrast to the abelian electromagnetic case, the non-abelian quantum theory has no physically meaningful classical analogue, or in other words, there are no such thing as a classical quark (non-abelian electron), or classical gluon (non-abelian photon). It seems that the non-abelian Lorentz equations have no useful physical applications in high energy particle theory, but they found their impact on the mechanical systems such as falling, swimming, orbiting, and rolling. In the next section we get a description of the rolling system of two bodies, leaving out other applications apart. The reader can find interesting examples of principal bundles associated to mechanical systems in [111].

**5.5.1. Heisenberg group.** THE HEISENBERG MANIFOLD. Let $B = \mathbb{R}^2$ be the base space, where an electrically charged particle will move. Assume that $B$ is endowed with the usual Euclidean metric and we will use the standard coordinate system $b = (x, y)$. Let $\pi\colon M = \mathbb{R}^3 \to B$, be the principle $\mathbb{R}$-bundle, where $\pi = \mathrm{pr}_{1,2}$ is the projection on the plane formed by two first coordinates. The action of the structure group $\mathbb{G} = (\mathbb{R}, +)$ is defined by

$$\begin{array}{cccc} \mu\colon & \mathbb{R}^3 \times \mathbb{R} & \to & \mathbb{R}^3 \\ & (x, y, t).\tau & \mapsto & (x, y, t + \tau). \end{array}$$

The vertical space is $V_q = \ker(\pi) = \mathrm{span}\{\partial_t\}$. Let us chose the horizontal distribution $D = \mathrm{span}\{\partial_x, \partial_y\}$. Then the connection $\mathbb{R}$-valued one-form is just $A = dt$ and the curvature form $\Omega = -dA \equiv 0$. We see that there is no magnetic field and the motion is a free motion on the base space, or its copy. Geodesics are straight lines. The sub-Riemannian metric $g_D$, which is the pullback of the Euclidean metric, is the Euclidean metric on $D$. The distribution $D$ is orthogonal to $V$ with respect to the Euclidean metric in $\mathbb{R}^3$, it is not bracket generating, the sub-Riemannian manifold $(\mathbb{R}^3, D, g_D)$ is the foliation by planes $t = \mathit{constant}$, and the motion is possible only inside of a plane defined by the initial position of the particle.

We choose now the horizontal distribution in another, non-trivial way, for instance $D = \mathrm{span}\{X, Y\}$, where

$$X = \partial_x, \qquad Y = \partial_y + x\partial_t,$$

or in a more symmetric way

$$X = \partial_x - \frac{1}{2}y\partial_t, \qquad Y = \partial_y + \frac{1}{2}x\partial_t.$$

In this case the reader recognizes the Heisenberg distribution. Since

$$d\pi(X) = \partial_x, \quad d\pi(Y) = \partial_y,$$

the pullback of the Euclidean metric from $\mathbb{R}^2$ to $D$ is represented by the identity matrix and makes the vector fields $X, Y$ into the orthonormal basis of $D$. The connection form is $A = dt - \frac{1}{2}xdy + \frac{1}{2}ydx$ and, as we remember, it is the dual form to the vertical vector field $T = \partial_t$. The vector fields $X, Y, T$ and their commutation relations define the Heisenberg group structure in $\mathbb{R}^3$ through the BCH-formula (8.1).

The curvature form $\Omega = -dA = dx \wedge dy$ is constant, non-vanishing and is equal to the volume form on $\mathbb{R}^2$. Any form $\Omega = F(x,y)dx \wedge dy$ represents a magnetic field in the base space $B = \mathbb{R}^2$, that also can be thought as a field $0\,dy\wedge dt + 0\,dt\wedge dx + F(x,y)dx\wedge dy$ orthogonal to the base space $B$. As it is known, in order to be a magnetic field the form $\Omega$ has to satisfy the Maxwell equation $d\Omega = 0$, which is true in this case. Observe that in the presence of the abelian structure group the curvature form given by $\Omega = -dA$ (since $[A,A] = 0$ in this case) automatically satisfies the Maxwell equation. In the case $F(x,y) = \mathit{constant}$ the magnetic field coincides up to a constant with the Heisenberg curvature form.

Let us have a look on the Lorentz equation (5.26). The Lie algebra of the structure group is $\mathbb{R}$ and the dual to it is also $\mathbb{R}$. Thus, the co-adjoint bundle is a trivial bundle with the typical fiber $\mathbb{R}$. The covariant derivative is the usual derivative and we get that the charge $\lambda$ has to be constant by (5.27). The horizontal lift $h$ maps the velocity vector $\dot{c} = \dot{x}\partial_x + \dot{y}\partial_y \in T_c\mathbb{R}^2$ to $\dot{c} = \dot{x}X(c) + \dot{y}Y(c) \in D_c$. Moreover,

$$\nabla_{\dot{c}}\dot{c} = \ddot{c}, \quad i_{\dot{c}}\Omega = \Omega(\dot{c}),$$

and the Lorentz equation becomes

$$\ddot{c} = -\lambda\Omega(\dot{c}).$$

The last equation coincides with the Hamiltonian equation (3.6) with $\theta_0 = \lambda$. So, geodesics produced by these equations are the Heisenberg geodesics described in Subsection 3.2.

HEISENBERG MANIFOLD AND $S^3$ WITH THE KÄHLER STRUCTURE. Suppose that we are given as the base space a manifold $B$, endowed with the Kähler structure. Let us recall the definition of a Kähler manifold.

**Definition 25.** *Let $M$ be a complex integrable manifold with corresponding complex structure $J$. We say that a Riemannian metric $g_M$ is compatible with $J$ if*

$$g_M(v,w) = g_M(Jv, Jw), \qquad \textit{for all} \quad v, w \in T_qM, \quad q \in M.$$

The triplet $(M, J, g_M)$ is called an Hermitian manifold. The compatible $J$ and $g_M$ defines a skew symmetric form $\omega$ by

$$\omega(v, w) := g_M(Jv, w) \qquad \text{for all} \quad v, w \in T_qM, \quad q \in M \tag{5.28}$$

and it is called the associated Kähler form. We can retrieve $g_M$ from $\omega$ also.

**Definition 26.** *A Kähler manifold is a complex integrable manifold $M$ endowed with the complex structure $J$, compatible Riemannian metric $g_M$, associated Kähler form $\omega$, such that $d\omega = 0$.*

It is important for us to know that the Kähler form $\omega$ and the Riemannian metric $g_M$ are related by (5.28). We suppose that the curvature form $\Omega$ is given by the Kähler form $\omega$. The Lorentz equation becomes

$$\nabla_{\dot{c}}\dot{c} = -\lambda J_c(\dot{c}).$$

We recognize the variational equations (3.16) and (5.14) obtained by variational methods for the Heisenberg group and for the Hopf fibration on $S^3$. Observe that the Levi-Civita connections in (3.16) and (5.14) are connections related to Riemannian metrics on the total space $M$ and the solutions $\gamma$ are the sub-Riemannian geodesics on $M$. Meanwhile, the Levi-Civita connection in the Lorentz equation is the connection related to the Riemannian metric on the base space $B$ and the solution is a curve $c$ in the base space, parameterized by the charge. However, projections of $\gamma$ coincide with $c$, as was asserted in Theorem 5.2.

CURVATURE OF GEODESICS ON SURFACES. Let $B$ be an oriented surface (two dimensional manifold) furnished with a Riemannian metric $g_B$. Let $\omega$ be its area form. The equation $\omega(v, w) = \omega(Jv, w)$ for $v, w \in T_bB$ defines an almost complex structure $J_b \colon T_bB \to T_bB$ on $B$.

Consider a principle $U(1)$-bundle $\pi \colon M \to B$ over $B$. The curvature form $\Omega$ can be written as $\Omega = F\omega$ for some $F \in C^\infty(B)$. The scalar field $F$ or the endomorphism $FJ \colon TB \to TB$ defines the magnetic field acting on the surface $B$. The Lorentz equation becomes

$$\nabla_{\dot{c}}\dot{c} = -FJ_c(\dot{c})$$

for a particle of unit charge. Suppose that the geodesic $c$ is parametrized by arc length: $g_B(\dot{c}, \dot{c}) = 1$. The geodesic curvature $k_{geod}(t)$ along $c$ is defined by $k_{geod}(t) = g_B(J_c(\dot{c}), \nabla_{\dot{c}}\dot{c})$ and the Lorentz equation is equivalent to

$$k_{geod}(t) = F(c(t)).$$

Suppose now that the magnetic field $F$ is constant, that leads to the constant geodesic curvature. Thus, if the base space $B$ is the plane then the geodesics are circles, which coincides with the Heisenberg case. If $B$ is the two dimensional sphere $S^2$, then geodesics will be great circles and this is reflected in the picture of the Hopf fibration $S^3$ over $S^2$.

**5.5.2. Quaternionic $\mathbb{H}$-type group with the Lorentzian metric.** In this example we take as the base space $B$ the Minkowski space $\mathbb{R}^{3,1} = (\mathbb{R}^4, g_B)$, where $g_B$ is a non-degenerate metric tensor of index 1 having the associated matrix

$$\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Define the principal $\mathbb{R}^3$-bundle $\pi\colon \mathbb{R}^{4+3} \to \mathbb{R}^4$ with the standard projection into the four dimensional subspace, where the abelian group $(\mathbb{R}^3, +)$ acts by

$$\begin{array}{cccc} \mu\colon & \mathbb{R}^7 \times \mathbb{R}^3 & \to & \mathbb{R}^7 \\ & (b, t_1, t_2, t_3).(\tau_1, \tau_2, \tau_3) & \to & (b, t_1 + \tau_1, t_2 + \tau_2, t_3 + \tau_3), \end{array} \qquad b \in \mathbb{R}^4.$$

Analogously to the Heisenberg manifold we find the vertical space $V_q$, $q = (b, t_1, t_2, t_3)$ by $V = \operatorname{span}\{\partial_{t_1}, \partial_{t_2}, \partial_{t_3}\}$. We chose the horizontal distribution $D$ as the span of $X_{11}$, $X_{21}$, $X_{31}$, $X_{41}$ in (3.32). With this we get the Ehresmann connection and recuperate the structure of quaternionic $\mathbb{H}$-type group of the lowest dimension. The choice of the basis for $D$ is completely defined by the choice of almost complex structures (3.30). The pullback of the Minkowski metric is now a non-degenerate metric of index 1, converting the vector fields $X_{11}, X_{21}, X_{31}, X_{41}$ into an orthonormal basis and defining the first one $X_{11}$ as a globally defined timelike vector field. The corresponding geometry can be called the *sub-Lorentzian geometry*. Different types of sub-Lorentzian manifolds were studied in [27, 61, 62, 90, 91, 92].

The curvature form $\Omega$ defines an $\mathbb{R}^3$-valued constant magnetic field. The projections of sub-Lorentzian geodesics into the Minkowski base space are divided into three different causal type and were considered in details in [91].

The change of principle $\mathbb{R}^3$-bundle to the bundle $\pi\colon B \times S^1 \to B$, $B = \mathbb{R}^{3,1}$, where the abelian group $U(1)$ acts on the slot $S^1$ by the standard multiplication of complex numbers, leads to the classical Kaluza-Klein model.

EXERCISE

1. Consider a round version of the Heisenberg group $\mathbb{H}^1 \cong \mathbb{R}^3 = \mathbb{R}^2 \times S^1$ as principal $U(1)$-bundle, where the action is defined by

$$\begin{array}{ccc} (\mathbb{R}^2 \times S^1) \times U(1) & \to & \mathbb{R}^2 \times S^1 \\ (x, y, e^{i\varphi}).v & \mapsto & (x, y, e^{i\varphi}.v). \end{array}$$

   Find a horizontal distribution with non-trivial curvature form. Write basic vector fields for the horizontal distribution and find the commutation relations between them. Do they form the Heisenberg algebra? If yes, find the group multiplication law in $\mathbb{R}^3$ by making use the BCH-formula.
2. Construct the Kaluza-Klein model, associated with two-step Carnot groups in Subsection 3.4.

## 6. Rolling manifolds

Rolling surfaces without slipping or twisting is one of the classical kinematic problems that in recent years has again attracted attention of mathematicians due to its geometric and analytic richness. The kinematic conditions of rolling without slipping or twisting are described by means of motion on a configuration space being tangential to a smooth sub-bundle of the tangent bundle ot the configuration space that we call, as before, the horizontal distribution. The precise definition of the mentioned motion in the case of two $n$-dimensional manifolds imbedded in $\mathbb{R}^N$, given for example in [125], involves studying the behavior of the tangent bundles of the manifolds and the normal bundles induced by the embeddings. This extrinsic point of view, which depends on the embeddings, has been successfully applied, for instance in [72, 73]. The drawback of the extrinsic approach is that the geometric descriptions depend strongly on the embedding under consideration.

So far, few attempts have been made to formulate this problem intrinsically. An early enlightening formulation is given in [18], that is achieved by means of an intrinsic version of the moving frame method of Élie Cartan, see [24, 126]. One of the important results established there is the bracket generating property of the rank two distribution corresponding to no-twisting and no-slipping restrictions, namely, if the two surfaces have different Gaussian curvature, then the distribution is bracket generating, see [18]. A control theoretic approach to the same problem, studied in [3], has the advantage that the kinematic restrictions are written explicitly as vector fields on the appropriate configuration space.

We present here a short description of a generalization of the kinematic problem for two $n$-dimensional abstract manifolds rolling without twisting or slipping via an intrinsic formulation. We define the configuration space of the system, present an extrinsic definition of rolling for manifolds imbedded into the Euclidean space, several equivalent definitions of rolling, involving intrinsic characteristics, and discuss their relations. The intrinsic approach permits to determine the embedding- independent information contained in the extrinsic definition.

### 6.1. Rolling of embedded manifolds

Let $M$ and $\widehat{M}$ be oriented, connected, $n$-dimensional Riemannian manifolds isometrically imbedded into $\mathbb{R}^{n+\nu}$, equipped with the standard Euclidean metric and standard orientation. Isometrical embeddings always exist due to a result of Nash [115] and we denote them by $\iota$ and $\widehat{\iota}$, respectively. The corresponding Riemannian metrics on $M$ and $\widehat{M}$ coincide with the restrictions of the Euclidean metric from $\mathbb{R}^{n+\nu}$ and they will be denoted by $g_M$ and $g_{\widehat{M}}$.

Objects (points, curves, ...) related to the manifold $\widehat{M}$ will be marked by a hat $(\widehat{\ })$ on top, objects related to $M$ will be free of it, while those related to the ambient space $\mathbb{R}^N$, $N = n + \nu$, will carry a bar $(\,\bar{}\,)$.

Note that for any manifold $M$ imbedded in $\mathbb{R}^{n+\nu}$, there is a natural splitting of the tangent space of $\mathbb{R}^{n+\nu}$ into a direct sum:

$$T_x\mathbb{R}^{n+\nu} = T_xM \oplus T_xM^{\perp}, \quad x \in M, \tag{6.1}$$

where $T_xM$ is the tangent space and $T_xM^{\perp}$ is the normal space to $M$ at $x$. According to the splitting (6.1), any vector $v \in T_x\mathbb{R}^{n+\nu}$, $x \in M$, can be written uniquely as the sum $v = v^{\top} + v^{\perp}$, where $v^{\top} \in T_xM$, $v^{\perp} \in T_xM^{\perp}$. Analogous projections can be defined for $\widehat{M}$.

Let $\nabla$ denote the Levi-Civita connection on $M$ or on $\widehat{M}$. The "ambient" Levi-Civita connection on $\mathbb{R}^{n+\nu}$ is denoted by $\overline{\nabla}$. Note that if $X$ and $Y$ are tangent vector fields on $M$, and $\Upsilon$ is a normal vector field to $M$, then

$$\nabla_X Y(x) = \big(\overline{\nabla}_{\bar{X}}\bar{Y}(x)\big)^{\top}, \quad \nabla^{\perp}_X \Upsilon(x) := \big(\overline{\nabla}_{\bar{X}}\bar{\Upsilon}(x)\big)^{\perp}, \quad x \in M,$$

where $\bar{X}$, $\bar{Y}$ and $\bar{\Upsilon}$ are any local extensions to $\mathbb{R}^{n+\nu}$ of the vector fields $X$, $Y$ and $\Upsilon$, respectively. Equivalent statements hold for $\widehat{M}$.

If $Z$ and $\Psi$ are vector fields along a smooth curve $x\colon I \to \mathbb{R}^{n+\nu}$, we use $\frac{D}{dt}Z(t)$ to denote the covariant derivative of $Z$ along the curve $x$ and $\frac{D^{\perp}}{dt}\Psi$ for the normal covariant derivative of $\Psi$ along $x$ (these notations are according [117, p. 119]), see also Appendix A, Subsection 8.1. Observe that an isometric embedding of $M$ into $\mathbb{R}^{n+\nu}$ induces the equalities

$$\frac{D}{dt}Z = \left(\frac{d}{dt}Z\right)^{\top}, \qquad \frac{D^{\perp}}{dt}\Psi = \left(\frac{d}{dt}\Psi\right)^{\perp}.$$

A tangent vector $Z$ along a smooth curve $x$ is *parallel* if $\frac{D}{dt}Z(t) = 0$ for every $t \in I$. We say that a normal vector field $\Psi$ along $x$ is *normal parallel* if $\frac{D^{\perp}}{dt}\Psi(t) = 0$ for every $t$.

Definition 27 is a reformulation of the definition of a rolling map contained in [125, Appendix B]. The group $SE(N)$ of orientation preserving Riemannian isometries of $\mathbb{R}^N$ will play an important role. For the definition of the group $SE(N)$ see Appendix A, Subsection 8.3.

**Definition 27.** *A rolling of $M$ on $\widehat{M}$ without slipping or twisting is a smooth curve $(x, \mathcal{R})\colon\ [0,\tau] \to M \times SE(n+\nu)$ satisfying the following conditions:*

(i) $\widehat{x}(t) := \mathcal{R}(t)\,x(t) \in \widehat{M}$,
(ii) $d_{x(t)}\mathcal{R}(t)\,T_{x(t)}M = T_{\widehat{x}(t)}\widehat{M}$,
(iii) $d_{x(t)}\mathcal{R}(t)|_{T_{x(t)}M} : T_{x(t)}M \to T_{\widehat{x}(t)}\widehat{M}$ *is orientation preserving.*
(iv) *No slip condition:* $\dot{\widehat{x}}(t) = d_{x(t)}\mathcal{R}(t)\,\dot{x}(t)$, *for every* $t$.
(v) *No twist condition* (*tangential part*) :

$$d_{x(t)}\mathcal{R}(t)\,\frac{D}{dt}Z(t) = \frac{D}{dt}\,d_{x(t)}\mathcal{R}(t)\,Z(t),$$

*for any tangent vector field $Z(t)$ along $x(t)$ and every $t$.*

(vi) *No twist condition* (*normal part*) :

$$d_{x(t)}\mathcal{R}(t)\,\frac{D^{\perp}}{dt}\,\Psi(t)=\frac{D^{\perp}}{dt}\,d_{x(t)}\mathcal{R}(t)\,\Psi(t),$$

*for any normal vector field* $\Psi(t)$ *along* $x(t)$ *and every* $t$.

From now on we omit words “without slipping or twisting” just writing “a rolling”. Condition (v) is equivalent to the requirement that any tangent vector field $Z$ is parallel along the curve $x$ if and only if $d_x\mathcal{R}\,Z$ is parallel along $\widehat{x}$. As a consequence, this condition is automatically satisfied in the case of manifolds of dimension one. Similarly, condition (vi) is equivalent to the statement that any normal vector field $\Psi$ is normal parallel along the curve $x$ if and only if $d_x\mathcal{R}\,\Psi$ is normal parallel vector field along $\widehat{x}$. Thus, for embeddings of co-dimension one, condition (vi) holds automatically.

**Example 7.** Consider the submanifolds of $\mathbb{R}^3$, defined by

$$M=\left\{(\bar{x}_1,\sin\theta,1-\cos\theta)\in\mathbb{R}^3|\ \bar{x}_1\in\mathbb{R},\ \theta\in[0,2\pi)\right\},$$

$$\widehat{M}=\left\{(\bar{x}_1,\bar{x}_2,0)\in\mathbb{R}^3|\ \bar{x}_1,\bar{x}_2\in\mathbb{R},\right\}.$$

These are a cylinder and a plane. The rolling map

$$\mathcal{R}(t):\bar{x}=\begin{pmatrix}\bar{x}_1\\ \bar{x}_2\\ \bar{x}_3\end{pmatrix}\mapsto\begin{pmatrix}\bar{x}_1\\ \bar{x}_2\cos t+(\bar{x}_3-1)\sin t+t\\ -\bar{x}_2\sin t+(\bar{x}_3-1)\cos t+1\end{pmatrix},$$

describes the rolling of the infinite cylinder $M$ on $\widehat{M}$ along the $\bar{x}_2$-axis with constant speed 1. Any choice of a smooth curve $x\in M$, given by

$$x(t)=(\bar{x}_1,\sin t,1-\cos t),\quad \bar{x}_1\in\mathbb{R},\quad t\in I\subset\mathbb{R},$$

defines the rolling

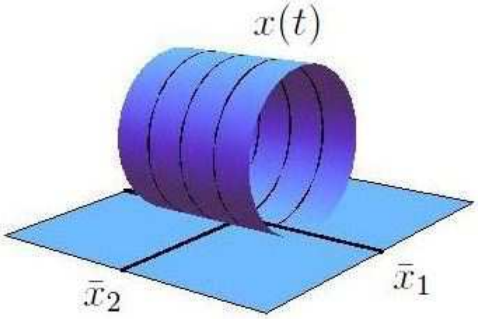

FIGURE 6.1. Rolling of the cylinder over the plane

Notice that Definition 27 ignores physical restrictions given by the actual shapes of the imbedded manifolds. If we think of $M$ and $\widehat{M}$ as touching along the curves $x$ and $\hat{x}$ and rolling according to the isometry $\mathcal{R}$, then we cannot rule out the possibility that there might be transverse intersections between the manifolds other than the contact points.

### 6.2. Intrinsic rolling

In this section we introduce a new object called intrinsic rolling.

**6.2.1. Frame bundles and bundles of isometries.** Let $V$ and $\widehat{V}$ be two oriented inner product $n$-dimensional spaces. We denote by $SO(V,\widehat{V})$ the collection of all linear orientation-preserving isometries between $V$ and $\widehat{V}$. When $V=\widehat{V}$, we write $SO(V)$ instead of $SO(V,V)$. Note that $SO(V)$ is a group. Given any choice of the basis in $V$, we can write an element of $SO(V)$ as an $(n\times n)$-matrix. However, since there is no canonical choice of the basis on $V$, the group $SO(V)$ is not canonically isomorphic to $SO(n)$.

For any pair $M$ and $\widehat{M}$, we introduce the space $Q$ of all relative positions in which $M$ can be tangent to $\widehat{M}$

$$Q=\Big\{q\in SO(T_xM,T_{\widehat{x}}\widehat{M})\Big|\,x\in M,\widehat{x}\in\widehat{M}\Big\}. \tag{6.2}$$

This space is a manifold with a structure of an $SO(n)-$fiber bundle over $M\times\widehat{M}$ and can be considered as the configuration space of the rolling. Its dimension is $\frac{n(n+3)}{2}$. Notice, that it is not a principal $SO(n)-$bundle since the action of $SO(n)$ on the fiber depends on the choice of coordinates in $M$ and $\widehat{M}$. To see this in more detail, we describe the space $Q$ in terms of frame bundles. Let $F$ and $\widehat{F}$ be the oriented orthonormal frame bundles of $M$ and $\widehat{M}$, respectively. As we know from Subsection 8.5.1 of Appendix A, $F$ and $\widehat{F}$ are principal $SO(n)$-bundles.

Consider $F\times\widehat{F}$ as a bundle over $M\times\widehat{M}$. The group $SO(n)$ acts on the right on $F$ and $\widehat{F}$ and we can divide by this action diagonally on fibers. Then, we can identify $Q$ with $(F\times\widehat{F})/SO(n)$ by the map assigning to each equivalence class $[f,\hat{f}]$ the mapping $q\in Q$, such that

$$\hat{f}_j=q\,f_j,\qquad \text{for}\quad j=1,\dots,n. \tag{6.3}$$

Clearly, this construction does not depend on the choice of a representative of an equivalence class of $(F\times\widehat{F})/SO(n)$. Conversely, given an isometry $q\in Q$, there exists a unique equivalence class of frames satisfying (6.3).

As we see, we can define the right action by $SO(T_xM)$ or the left action by $SO(T_{\hat{x}}\widehat{M})$ on $(F\times\widehat{F})/SO(n)$. Since both groups are not canonically isomorphic to $SO(n)$ (except for the case when $n=2$), the configuration space $Q=(F\times\widehat{F})/SO(n)$ does not have the structure of a principal $SO(n)-$bundle. However, since $Q$ is an $SO(n)-$fiber bundle we can exploit its local properties and think that it looks locally like the product $M\times\widehat{M}\times SO(n)$. Let $U$ be a neighborhood in

$M$ such that $F|_U$ is trivial and let $v$ be a section of $F|_U$: $v(x) = (v_1(x), \dots, v_n(x))$, $x \in M$. Each section determines a left action of $SO(n)$ on $F|_U$. To see this, recall that for each $x \in U$, the frame $v(x)$ can be considered as an isometry $v(x) : \mathbb{R}^n \to T_xM$. The left action takes the following form: if $f \in F_x$ is any other frame at $x \in U$, written in terms of the frame $v$ as

$$f_j = \sum_{i=1}^{n} f_{ij} v_i(x),$$

then $\tau = (\tau_{ij})_{i,j=1}^n \in SO(n)$ acts on the left on $f$ via the equation

$$\tau.f_j = \sum_{i,k=1}^{n} f_{ij}\tau_{ki}v_k, \qquad j = 1, \dots, n.$$

Observe that this action depends on the choice of the frame $v$.

This defines local left and right actions of $SO(n)$ on $Q$ as follows. Let $U$ and $\widehat{U}$ be neighborhoods in $M$ and $\widehat{M}$ respectively, so that both frame bundles trivialize over these neighborhoods. Let $v : U \to F|_U$ and $\hat{v} : \widehat{U} \to \widehat{F}|_{\widehat{U}}$ be sections. We define the left action of $\tau \in SO(n)$ on $Q$ with respect to $\hat{v}$ by

$$\tau.\hat{f}_j = (\tau.q)f_j,$$

where the left action of $\tau$ on $\hat{f}_j$ is defined with respect to $\hat{v}$ and $\hat{f}_j = qf_j$ for $j = 1, \dots, n$. Similarly, the right action of $SO(n)$ on $Q$ with respect to $v$ is defined by

$$\hat{f}_j = (q.\tau)\big(\tau^{-1}.f_j\big).$$

Remark that if we have a matrix representation of an element $\tau_0 \in SO(n)$ in coordinates of $\widehat{M}$ by $\tau_0 = \big\{g_{\widehat{M}}(\hat{v}_i, qv_j)\big\}_{i,j=1}^n$, then we have

$$\big\{g_{\widehat{M}}(\hat{v}_i, (\tau.q)v_j)\big\}_{i,j=1}^n = \tau\tau_0, \text{ and } \big\{g_{\widehat{M}}(\hat{v}_i, (q.\tau)v_j)\big\}_{i,j=1}^n = \tau_0\tau, \quad \tau \in SO(n).$$

**6.2.2. Reformulation of rolling in terms of bundles.** Both formulations of rolling surfaces given in [3, 18] define the configuration space as a manifold of isometries of tangent spaces of $M$ and $\widehat{M}$, as we did before, without taking into account the embedding into the ambient space. The condition (vi) imposed over a rolling $(x, \mathcal{R})$ by Definition 27 is non-trivial whenever the codimension $\nu$ of the imbedded manifolds is greater than 1. So, it is natural to suppose that the total configuration space of the rolling system will have a normal component which takes care of the action of $\mathcal{R}$ on the normal bundle. Therefore, by analogy with the construction of $Q$, we define a fiber bundle over $M \times \widehat{M}$ of isometries of the normal tangent space. Let $\iota : M \to \mathbb{R}^{n+\nu}$ and $\widehat{\iota} : \widehat{M} \to \mathbb{R}^{n+\nu}$ be two embeddings, given as initial data. Let $\Phi$ be the principal $SO(\nu)-$bundle over $M$, such that the fiber over a point $x \in M$ consists of all positively oriented orthonormal frames $\{\epsilon_\lambda(x)\}_{\lambda=1}^{\nu}$ spanning $T_xM^\perp$. Let $\widehat{\Phi}$ be the similarly defined principal $SO(\nu)-$bundle on $\widehat{M}$.

As it was done previously, we identify the manifold $(\Phi \times \widehat{\Phi})/SO(\nu)$ with

$$P_{\iota,\widehat{\iota}} := \left\{ p \in SO(T_xM^{\perp}, T_{\widehat{x}}\widehat{M}^{\perp}) \Big| \, x \in M, \widehat{x} \in \widehat{M} \right\}. \tag{6.4}$$

The space $P_{\iota,\widehat{\iota}}$ is not in general a principal $SO(\nu)-$bundle, but there are local left and right actions defined similarly as on $Q$. We notice and reflect it in notations that $Q$ is invariant of embeddings, while $P_{\iota,\widehat{\iota}}$ is not. We obtain $\dim(P_{\iota,\widehat{\iota}}) = 2n + \frac{\nu(\nu-1)}{2}$. We form the direct sum $Q \oplus P_{\iota,\widehat{\iota}}$ for the fiber bundle over $M \times \widehat{M}$, so that the fiber over $(x, \widehat{x}) \in M \times \widehat{M}$, is $Q_{(x,\widehat{x})} \times P_{\iota,\widehat{\iota}(x,\widehat{x})}$. Thus $\dim(Q \oplus P_{\iota,\widehat{\iota}}) = \frac{n(n+3)+\nu(\nu-1)}{2}$. The following proposition allows to reformulate Definition 27.

**Proposition 9.** *If a curve* $(x, \mathcal{R}) : [0, \tau] \to M \times SE(n + \nu)$ *satisfies the conditions* (i)-(vi) *in Definition* 27*, then the mapping*

$$t \mapsto (d_{x(t)}\mathcal{R}(t)|_{T_{x(t)}M}, d_{x(t)}\mathcal{R}(t)|_{T_{x(t)}M^{\perp}}) =: (q(t), p(t)),$$

*defines a curve in* $Q \oplus P_{\iota,\widehat{\iota}}$ *with the following properties:*

- (I) *no slip condition:* $\dot{\widehat{x}}(t) = q(t)\dot{x}(t)$ *for every* $t$.
- (II) *no twist condition, tangential part:* $q(t)\frac{D}{dt}Z(t) = \frac{D}{dt}q(t)Z(t)$ *for any tangent vector field* $Z(t)$ *along* $x(t)$ *and every* $t$.
- (III) *no twist condition, normal part:* $p(t)\frac{D^{\perp}}{dt}\Psi(t) = \frac{D^{\perp}}{dt}p(t)\Psi(t)$ *for any normal vector field* $\Psi(t)$ *along* $x(t)$ *and every* $t$.

*Conversely, if* $(q, p) : [0, \tau] \to Q \oplus P_{\iota,\widehat{\iota}}$ *is a smooth curve satisfying* (I)-(III), *then there exists a unique rolling*

$$(x, \mathcal{R})\colon [0, \tau] \to M \times SE(n + \nu),$$

*such that* $d_{x(t)}\mathcal{R}(t)|_{T_{x(t)}M} = q(t)$ *and* $d_{x(t)}\mathcal{R}(t)|_{T_{x(t)}M^{\perp}} = p(t)$.

*Proof.* Assume that $(x, \mathcal{R})\colon [0, \tau] \to M \times SE(n + \nu)$ is a rolling map satisfying (i)-(vi). The conditions (i) and (ii) assure that

$$\begin{aligned} &d_{x(t)}\mathcal{R}(t)|_{T_{x(t)}M} \in SO(T_{x(t)}M, T_{\widehat{x}(t)}\widehat{M}) \quad \text{and} \\ &d_{x(t)}\mathcal{R}(t)|_{T_{x(t)}M^{\perp}} \in SO(T_{x(t)}M^{\perp}, T_{\widehat{x}(t)}\widehat{M}^{\perp}). \end{aligned} \tag{6.5}$$

Since $d_{x(t)}\mathcal{R}(t)$ must be orientation preserving in $\mathbb{R}^{n+\nu}$ for any $t \in [0, \tau]$ we conclude that both of the mappings (6.5) are either orientation reversing or orientation preserving. The additional requirement (iii) implies that $(q, p)$ is orientation preserving. The conditions (I)-(III) correspond to the conditions (iv)-(vi).

Conversely, if we have a curve $(q, p)$ in $Q \oplus P_{\iota,\widehat{\iota}}$ with projection $(x, \widehat{x})$ into $M \times \widehat{M}$, then we have an isometry $\mathcal{R} \in SE(n+\nu)$ in the following way: $\mathcal{R}(t)\colon \bar{x} \mapsto \bar{A}(t)\bar{x} + \bar{a}(t), \bar{A}(t) \in SO(n + \nu)$, where $\bar{A}(t) = d_{x(t)}\mathcal{R}(t)$ is determined by the conditions

$$d_{x(t)}\mathcal{R}(t)|_{T_{x(t)}M} = q(t)|_{T_{x(t)}M}, \qquad d_{x(t)}\mathcal{R}(t)|_{T_{x(t)}M^{\perp}} = p(t)|_{T_{x(t)}M^{\perp}}.$$

Then for images of $d_{x(t)}\mathcal{R}(t)$ we have

$$\operatorname{Im}\Big(d_{x(t)}\mathcal{R}(t)|_{T_{x(t)}M}\Big) = T_{\widehat{x}(t)}\widehat{M}, \qquad \operatorname{Im}\Big(d_{x(t)}\mathcal{R}(t)|_{T_{x(t)}M^{\perp}}\Big) = T_{\widehat{x}(t)}\widehat{M}^{\perp}.$$

The vector $\bar{a}(t)$ is determined by $\bar{a}(t) = \widehat{x}(t) - \bar{A}(t)x(t)$ for any $t \in [0, \tau]$. $\square$

The one-to-one correspondence between rolling maps and smooth curves in $Q \oplus P_{\iota,\widehat{\iota}}$, satisfying (I)-(III), naturally leads to a definition of a rolling map in terms of these bundles.

**Definition 28.** *A rolling of $M$ on $\widehat{M}$ without slipping or twisting is a smooth curve $(q, p)\colon [0, \tau] \to Q \oplus P_{\iota,\widehat{\iota}}$ such that $(q(t), p(t))$ satisfies*

- (I) *no slip condition: $\dot{\widehat{x}}(t) = q(t)\dot{x}(t)$ for every $t$,*
- (II) *no twist condition, tangential part: $q(t)\frac{D}{dt}Z(t) = \frac{D}{dt}q(t)Z(t)$ for every $t \in [0, \tau]$ and for any tangent vector field $Z$ along $x$,*
- (III) *no twist condition, normal part: $p(t)\frac{D^\perp}{dt}\Psi(t) = \frac{D^\perp}{dt}p(t)\Psi(t)$ for every $t \in [0, \tau]$ and for any normal vector field $\Psi$ along $x$.*

Proposition 9 implies that the bundle $Q \oplus P_{\iota,\widehat{\iota}}$ can be seen as the configuration space for a rolling of two isometrically embedded manifolds $\iota\colon M \to \mathbb{R}^{n+\nu}$ and $\widehat{\iota}\colon \widehat{M} \to \mathbb{R}^{n+\nu}$. According to [125], the dimension $\frac{n(n+3)+\nu(\nu-1)}{2}$ corresponds to the degrees of freedom of the system.

A purely intrinsic definition of a rolling is deduced from Definition 28, by restricting it to the bundle $Q$. This concept naturally generalizes the definition given in [3] for 2-dimensional Riemannian manifolds imbedded into $\mathbb{R}^3$ and we use the term *intrinsic rolling* for this object.

**Definition 29.** *An intrinsic rolling of two $n$-dimensional oriented Riemannian manifolds $M$ on $\widehat{M}$ without slipping or twisting is a smooth curve $q\colon [0, \tau] \to Q$, with projections $x(t) = \mathrm{pr}_M\, q(t)$ and $\widehat{x}(t) = \mathrm{pr}_{\widehat{M}}\, q(t)$, satisfying the following conditions:*

- (I') *no slip condition: $\dot{\widehat{x}}(t) = q(t)\dot{x}(t)$ for all $t$,*
- (II') *no twist condition: $Z$ is a parallel tangent vector field along the curve $x$, if and only if $q\,Z$ is parallel along $\widehat{x}$.*

**Remark 8.** If $n$-dimensional manifolds $M$ and $\widehat{M}$ are imbedded into Euclidean space $\mathbb{R}^{n+1}$, then for each pair of points $(x, \widehat{x}) \in M \times \widehat{M}$, there is a unique orientation preserving isometry $p\colon T_xM^\perp \to T_{\widehat{x}}\widehat{M}^\perp$. Hence, since $P_{\iota,\widehat{\iota}}$ is an $SO(1)$-bundle, it can be identified with $M \times \widehat{M}$, and so $Q \oplus P_{\iota,\widehat{\iota}} \cong Q$. In this case we see that the notion of rolling in Definition 28 coincides with the intrinsic rolling in Definition 29.

**6.2.3. Extrinsic and intrinsic rollings along the same curves.** Let $(x, \widehat{x})\colon [0, \tau] \to M \times \widehat{M}$ be a given pair of curves. We aim to give an answer to the following questions:

- If $q_1$ and $q_2$ are two intrinsic rollings of $M$ on $\widehat{M}$, along curves $x$ and $\widehat{x}$, how are they related? What properties of the rolling are defined by fixing the paths $x$ and $\widehat{x}$?
- Suppose an intrinsic rolling $q$ and embeddings $\iota\colon M \to \mathbb{R}^{n+\nu}$ and $\widehat{\iota}\colon \widehat{M} \to \mathbb{R}^{n+\nu}$ are given. Is it possible to extend $q$ to a rolling $(q, p)$? Is this extension unique?

The following example clarifies the situation for one dimensional manifolds, where different embeddings are easy to describe.

**Example 8.** Consider $\widehat{M} = \mathbb{R}$, with the usual Euclidean structure, and $M = S^1$, with the usual round metric and positive orientation counterclockwise. Let $x : [0, \tau] \to S^1$ be written as $x(t) = e^{i\varphi(t)}$, where $\varphi : [0, \tau] \to \mathbb{R}$ is an absolutely continuous function. Since $SO(1) = \{\mathbf{1}\}$, the configuration space $Q$ for the intrinsic rolling is just $M \times \widehat{M}$. The no-slipping condition implies that

$$\widehat{x}(t) = \widehat{x}(0) + \varphi(t) - \varphi(0),$$

and we may assume $\widehat{x}(0) = \varphi(0) = 0$. We consider different rollings of $M$ on $\widehat{M}$ under various embeddings. Without loss of generality, we may assume that $\mathcal{R}(0) = \mathrm{id}_{\mathbb{R}^{1+\nu}}$ is the identity map in $\mathbb{R}^{1+\nu}$. We will use $r = (r_1, \dots, r_{1+\nu})$ for coordinates of $\mathbb{R}^{1+\nu}$.

*Case 1:* Consider the embeddings

$$\begin{array}{cccc} \iota_1 : & M & \to & \mathbb{R}^2 \\ & e^{i\varphi} & \mapsto & (\sin\varphi, 1 - \cos\varphi) \end{array}, \qquad \begin{array}{cccc} \widehat{\iota}_1 : & \widehat{M} & \to & \mathbb{R}^2 \\ & \widehat{x} & \mapsto & (\widehat{x}, 0) \end{array}.$$

Simple calculations show that there is only one possible rolling.

*Case 2:* Consider the embeddings

$$\begin{array}{cccc} \iota_2 : & M & \to & \mathbb{R}^3 \\ & e^{i\varphi} & \mapsto & (\sin\varphi, (1 - \cos\varphi)\cos\theta_0, (1 - \cos\varphi)\sin\theta_0) \end{array},$$

$$\begin{array}{cccc} \widehat{\iota}_2 : & \widehat{M} & \to & \mathbb{R}^3 \\ & \widehat{x} & \mapsto & (\widehat{x}, 0, 0) \end{array},$$

where $\theta_0$ is any fixed angle from $(0, \frac{\pi}{2})$. Conditions (ii), (iii) and (iv) of Definition 27 imply that the differential $d_{x(t)}\mathcal{R}(t)$ of $\mathcal{R}(t)$, $t \in [0, \tau]$, in matrix form can be written uniquely as

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varkappa(t) & \sin\varkappa(t) \\ 0 & -\sin\varkappa(t) & \cos\varkappa(t) \end{pmatrix} \begin{pmatrix} \cos\varphi(t) & \sin\varphi(t) & 0 \\ -\sin\varphi(t) & \cos\varphi(t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_0 & \sin\theta_0 \\ 0 & -\sin\theta_0 & \cos\theta_0 \end{pmatrix},$$

for some smooth function $\varkappa\colon [0, \tau] \to \mathbb{R}$. To satisfy the normal no-twist condition, $d_{x(t)}\mathcal{R}(t)$ must map the normal parallel vector fields on $M$

$$\epsilon_1 = -\sin\varphi(t)\frac{\partial}{\partial r_1} + \cos\varphi(t)\cos\theta_0\frac{\partial}{\partial r_2} + \cos\varphi(t)\sin\theta_0\frac{\partial}{\partial r_3},$$

$$\epsilon_2 = -\sin\theta_0\frac{\partial}{\partial r_2} + \cos\theta_0\frac{\partial}{\partial r_3},$$

to normal parallel vector fields on $\widehat{M}$. Calculating the covariant derivative of $d_{x(t)}\mathcal{R}(t)\epsilon_1$ and $d_{x(t)}\mathcal{R}(t)\epsilon_2$, we conclude that $\varkappa(t)$ is constant and the assumption $\mathcal{R}(0) = \mathrm{id}_{\mathbb{R}^{1+\nu}}$ implies that the constant is 0. Hence, the circle will roll along the line with a constant tilt given by $\theta_0$, see Figure 6.2.

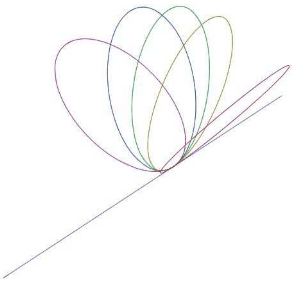

FIGURE 6.2. Case 2: $S^1$ rolling on $\mathbb{R}$. Different tilting angles give different embeddings, but equivalent rollings.

*Case 3:* Consider the isometric embedding of $\widehat{M}$ as a spiral.

$$\begin{array}{cccc} \widehat{\iota}_3 & : & \widehat{M} & \to & \mathbb{R}^3 \\ & & \widehat{x} & \mapsto & \frac{1}{\sqrt{2}}(\cos\widehat{x}, \sin\widehat{x}, \widehat{x}), \end{array}$$

and $\iota_2$ from the previous case. In this situation, the circle $M$ will rotate along the spiral $\widehat{M}$. Checking the normal no-twist condition we came to the same conclusion that the path is uniquely determined by the initial angle $\theta_0$.

Note that in all the cases above, the intrinsic rolling $t \mapsto (e^{i\varphi(t)}, \varphi(t))$ either uniquely induces a rolling, or the rolling is determined by an initial configuration of the normal tangent spaces, which corresponds to the initial tilting angle $\theta_0$. In fact it is also possible to find a choice of basis, consisting of normal parallel vector fields, so that the normal component of the rolling $p$ is constant with respect to this basis. We show in Lemma 4 below that this holds generally.

Let $x\colon [0,\tau] \to M$ and $\widehat{x} : [0,\tau] \to \widehat{M}$ be two fixed curves. We denote by $\{e_j(t)\}_{j=1}^n$ a collection of parallel tangent vector fields along $x(t)$ forming an orthonormal basis for $T_{x(t)}M$ and by $\{\epsilon_\lambda(t)\}_{\lambda=1}^\nu$ a collection of normal parallel vector fields along $x(t)$ forming an orthonormal basis for $T_{x(t)}M^\perp$. Such vector fields can be constructed by parallel transport and normal parallel transport along $x(t)$. Similarly, along $\widehat{x}(t)$, we define parallel frames $\{\hat{e}_i\}_{i=1}^n$ and $\{\hat{\epsilon}_\kappa\}_{\kappa=1}^\nu$. Recall that Latin indices $i, j, \ldots$ vary from 1 to $n$, while Greek ones $\kappa, \lambda, \ldots$ vary from 1 to $\nu$.

The following lemma shows that the image of a parallel frame over $M$ has constant coordinates in a parallel frame over $\widehat{M}$. This reflects the fact that rolling preserves parallel vector fields.

**Lemma 4.** *A curve $(q,p)\colon [0,\tau] \to Q \oplus P_{\iota,\widehat{\iota}}$ satisfies (II) and (III) if and only if the matrices*

$$A(t) = (a_{ij}(t)) = (g_{\widehat{M}}(\hat{e}_i, q(t)e_j)), \quad B(t) = (b_{\kappa\lambda}(t)) = (g_{\widehat{M}}(\hat{\epsilon}_\kappa(t), p(t)\epsilon_\lambda(t))),$$

*are constant for any $t \in [0,\tau]$.*

*Proof.* Check that the derivatives $\dot a_{ij}(t)$ and $\dot b_{ij}(t)$ vanish, see also [56]. □

The following theorem gives an answer to the first question raised at the beginning of Subsection 6.2.3.

**Theorem 6.1.** [56] *Let* $q\colon [0,\tau]\to Q$ *be a given intrinsic rolling map without slipping or twisting with projection* $\mathrm{pr}_{M\times\widehat M}\, q(t) = (x(t),\widehat x(t))$, $t\in[0,\tau]$. *Define the vector spaces*

$$V=\Big\{v \textit{ is a parallel vector field along } \ x,\ \textit{ and } g_M(v,\dot x)=0 \textit{ for all } t\Big\},$$

$$\widehat V=\Big\{\widehat v \textit{ is a parallel vector field along } \ \widehat x,\ \textit{ and } g_{\widehat M}(\widehat v,\dot{\widehat x})=0 \textit{ for all } t\Big\}.$$

*Then* $\dim V=\dim\widehat V$ *and, if we denote this dimension by* $k$*, the following holds.*

(a) *The map* $q$ *is the unique intrinsic rolling of* $M$ *on* $\widehat M$ *along curves* $x$ *and* $\widehat x$ *if and only if* $k\le 1$.

(b) *If* $k\ge 2$*, all the rollings along* $x$ *and* $\widehat x$ *differ from* $q$ *by an element in* $SO(\widehat V)$.

In particular, if the curve $x\colon [0,\tau]\to M$ is a geodesic, we have the following consequence of Theorem 6.1.

**Corollary 8.** *Assume that the curve* $x$ *is a geodesic in* $M$*. Then there exists an intrinsic rolling of* $M$ *on* $\widehat M$ *along* $(x,\widehat x)$ *if and only if* $\widehat x$ *is a geodesic with the same speed as* $x$*. Moreover, if* $n\ge 2$*, and if* $\widehat V$ *is defined as in Theorem* 6.1*, then* $\dim\widehat V=n-1$*, and all the rollings along* $x$ *and* $\widehat x$ *differ by an element in* $SO(\widehat V)$.

Concerning the problem of extending intrinsic rollings to extrinsic ones, the following theorem gives a complete answer to the question posed at the beginning of Subsection 6.2.3.

**Theorem 6.2.** *Let* $q\colon [0,\tau]\to Q$ *be an intrinsic rolling and let* $\iota\colon M\to\mathbb R^{n+\nu}$ *and* $\widehat\iota\colon \widehat M\to\mathbb R^{n+\nu}$ *be given embeddings. Then, given an initial normal configuration*

$$p_0\in(P_{\iota,\widehat\iota})_{(x_0,\widehat x_0)},\ \textit{ where } (x_0,\widehat x_0)=\mathrm{pr}_{M\times\widehat M}\,q(0),$$

*there exists a unique rolling* $(q,p)\colon [0,\tau]\to Q\oplus P_{\iota,\widehat\iota}$ *satisfying* $p(0)=p_0$.

*Proof.* Let $\{\epsilon_\lambda\}_{\lambda=1}^{\nu}$ and $\{\hat\epsilon_\kappa\}_{\kappa=1}^{\nu}$ be normal parallel frames along curves $x$ and $\widehat x$, respectively. Let $B_0\in SO(\nu)$ be defined by

$$B_0=\{b_{\kappa\lambda}\}_{\kappa,\lambda=1}^{\nu}=\big\{\bar g\big(\widehat\epsilon_\kappa(0),p_0\,\epsilon_\lambda(0)\big)\big\}_{\kappa,\lambda=1}^{\nu},$$

where $\bar g$ is the Euclidean metric in $\mathbb R^{n+\nu}$. Then $p(t)$ must satisfy

$$b_{\kappa\lambda}=\bar g\big(\widehat\epsilon_\kappa(t),p(t)\,\epsilon_\lambda(t)\big)\quad\text{for any}\quad t\in[0,\tau],$$

by Lemma 4, and it is uniquely determined by this. □

We already gave the answer about the uniqueness of the intrinsic rolling $q$ in Theorem 6.1. Then the extension $(q, p)$ was proposed in Theorem 6.2. Now the natural question arises: whether the extrinsic part $p$ is unique. In order to answer this question we define the vector spaces

$$E = \big\{\epsilon(t) \text{ is a normal parallel vector field along } x(t) \text{ and } \bar{g}\big(\dot{x}(t), \epsilon(t))\big) = 0\big\}\,,$$

$$\widehat{E} = \Big\{\widehat{\epsilon}(t) \text{ is a normal parallel vector field along } \widehat{x}(t) \text{ and } \bar{g}\big(\dot{\widehat{x}}(t), \widehat{\epsilon}(t))\big) = 0\Big\}\,,$$

with inner product $\bar{g}$ and orientation induced on $TM^{\perp}$ and $T\widehat{M}^{\perp}$ from embeddings. Both vector spaces have dimension $\nu$. An extrinsic rolling $(q, p)$ extending an intrinsic rolling $q$ is determined up to a left action of $SO(\widehat{E})$ or, equivalently, up to a right action of $SO(E)$. Both $SO(E)$ and $SO(\widehat{E})$ are isomorphic to $SO(\nu)$, but not canonically.

### 6.3. Distributions for extrinsic and intrinsic rolling

The aim of this subsection is to formulate the kinematic conditions of no-slipping and no-twisting in terms of a distribution. In this setting, a rolling will be a smooth curve almost everywhere tangent to this distribution.

**6.3.1. Local trivializations of $Q$.** Let $\pi\colon Q \oplus P_{\iota,\widehat{\iota}} \to M \times \widehat{M}$ denote the canonical projection. Consider a rolling $\mathcal{R}(t) = (q, p)\colon [0, \tau] \to Q \oplus P_{\iota,\widehat{\iota}}$, then $\pi \circ \mathcal{R} = (x, \widehat{x})$. Given an arbitrary $t_0$ in the domain of $\mathcal{R}$, let $U$ and $\widehat{U}$ denote neighborhoods of $x(t_0)$ and $\widehat{x}(t_0)$ in $M$ and $\widehat{M}$, respectively, such that both bundles $TM$ and $TM^{\perp}$ are trivialized being restricted to $U$. In the same way we chose $\widehat{U}$, such that both $T\widehat{M}$ and $T\widehat{M}^{\perp}$ are trivialized when they are restricted to $\widehat{U}$. This implies that the bundle $\pi\colon Q \oplus P_{\iota,\widehat{\iota}} \to M \times \widehat{M}$, trivializes when it is restricted to $U \times \widehat{U}$.

Each of the requirements (I)-(III) can be written as restrictions to $\dot{\mathcal{R}}$. We show, that all admissible values of $\dot{\mathcal{R}}$ form a distribution; that is a smooth subbundle of $T(Q \oplus P_{\iota,\widehat{\iota}})$. We will use the local trivializations to describe this distribution.

**6.3.2. The tangent space of $SO(n)$.** Let $U$ and $\widehat{U}$ be as in Subsection 6.3.1. The tangent space $T\pi^{-1}(U \times \widehat{U})$ is isomorphic to the following direct sum under the trivializaition

$$T\pi^{-1}(U \times \widehat{U}) = TU \times T\widehat{U} \times TSO(n) \times TSO(\nu).$$

The decomposition requires to know a detailed description of the tangent spaces $TSO(n)$ and $TSO(\nu)$ in terms of left and right invariant vector fields.

We start by considering the embedding of $SO(n)$ in $GL(n)$. Denote the matrix entries of a matrix $A$ by $(a_{ij})$ and the transpose matrix by $A^{tr}$. Then, differentiating the condition $A^t A = \mathbf{1}$, we obtain

$$TSO(n) = \bigcap_{i \leq j} \ker \omega_{ij}, \qquad \omega_{ij} = \sum_{r=1}^{n} \left(a_{rj}\, da_{ri} + a_{ri}\, da_{rj}\right).$$

It is clear that the tangent space at the identity 1 of $SO(n)$ is spanned by

$$W_{ij}(1) := \frac{\partial}{\partial a_{ij}} - \frac{\partial}{\partial a_{ji}}, \qquad 1 \le i < j \le n.$$

We denote $\mathfrak{so}(n) = \operatorname{span}\{W_{ij}(1)\}$ following the classical notation. We use left translations of these vectors to define

$$W_{ij}(A) := dl_A W_{ij}(1) = \sum_{r=1}^{n} \left( a_{ri} \frac{\partial}{\partial a_{rj}} - a_{rj} \frac{\partial}{\partial a_{ri}} \right) \tag{6.6}$$

as global left invariant basis of $TSO(n)$. Note that the left and right action in $TSO(n)$ is described by

$$dl_A \left( \frac{\partial}{\partial a_{ij}} \right) = \sum_{r=1}^{n} a_{ri} \frac{\partial}{\partial a_{rj}}, \qquad dr_A \left( \frac{\partial}{\partial a_{ij}} \right) = \sum_{s=1}^{n} a_{js} \frac{\partial}{\partial a_{is}}.$$

We have the following formula to switch from left to right translation and the other way around,

$$\begin{aligned} dl_A \left( \frac{\partial}{\partial a_{ij}} \right) &= \sum_{r,s=1}^{n} a_{ri} a_{si} \left( dr_A \left( \frac{\partial}{\partial a_{rs}} \right) \right), \\ dr_A \left( \frac{\partial}{\partial a_{ij}} \right) &= \sum_{r,s=1}^{n} a_{js} a_{ir} \left( dl_A \left( \frac{\partial}{\partial a_{rs}} \right) \right). \end{aligned}$$

Therefore, the right invariant basis of $TSO(n)$ can be written as

$$dr_A\big(W_{ij}(1)\big) = \sum_{r<s} (a_{ir} a_{js} - a_{jr} a_{is}) W_{rs}(A) = \operatorname{Ad}_{A^{-1}} \big(W_{ij}(A)\big).$$

If $W_{ij}(A)$ is defined by (6.6) and $i > j$, (so $W_{ij}(A) = -W_{ji}(A)$) then the bracket relations are given by

$$[W_{ij}, W_{kl}] = \delta_{j,k} W_{il} + \delta_{i,l} W_{jk} - \delta_{i,k} W_{jl} - \delta_{j,l} W_{ik}.$$

The detailed calculation presented in this subsection can be found in [56].

**6.3.3. Distributions.** Now we are ready to rewrite the kinematic conditions (I)-(III) as a distribution. Let $\mathcal{R}\colon [0.\tau] \to Q \oplus P_{\iota,\widehat{\iota}}$ be a rolling satisfying the conditions (I)-(III). Consider it image under the trivializations. Then

$$\dot{\mathcal{R}} = \dot{x} + \dot{\widehat{x}} + \sum_{i,j=1}^{n} \dot{a}_{ij} \frac{\partial}{\partial a_{ij}} + \sum_{\kappa,\lambda=1}^{\nu} \dot{b}_{\kappa\lambda} \frac{\partial}{\partial b_{\kappa\lambda}}. \tag{6.7}$$

Condition (I) holds if and only if $\dot{\widehat{x}}(t) = q(t)\dot{x}(t)$, $t \in [0, \tau]$.

We want to write the last two terms in (6.7) in the right invariant basis of corresponding tangent spaces of $SO(n)$ and $SO(\nu)$, based on conditions (II) and (III).

Satisfying (II), we obtain

$$\sum_{i,j=1}^{n} \dot a_{ij} \frac{\partial}{\partial a_{ij}} =$$
$$= \sum_{i<j} \big( g_M\big(\nabla_{\dot x(t)} q^{-1}\hat e_j, q^{-1}\hat e_i\big) - g_{\widehat M}\big(\nabla_{q\dot x(t)}\hat e_j, \hat e_i\big)\big) \operatorname{Ad}_{A^{-1}} \big(W_{ij}(A)\big).$$

Similarly, (III) holds if and only if

$$\sum_{\kappa,\lambda=1}^{\nu} \dot b_{\kappa\lambda} \frac{\partial}{\partial b_{\kappa\lambda}} =$$
$$= \sum_{\kappa<\lambda} \Big( g^{\perp}_{M}\big(\nabla^{\perp}_{\dot x(t)} p^{-1}\hat\epsilon_\lambda, p^{-1}\hat\epsilon_\kappa\big) - g^{\perp}_{\widehat M}\big(\nabla^{\perp}_{q\dot x(t)}\hat\epsilon_\lambda, \hat\epsilon_\kappa\big)\Big) \operatorname{Ad}_{B^{-1}} \big(W_{\kappa\lambda}(B)\big).$$

Here $e_j, \hat e_j, \epsilon_\kappa, \hat\epsilon_\kappa$, $j = 1, \dots, n$, $\kappa = 1, \dots, \nu$, are orthonormal bases of corresponding tangent and normal spaces and $g_M$, $g_{\widehat M}$, $g^{\perp}_M$, $g^{\perp}_{\widehat M}$ are restrictions of the Euclidean metric on corresponding tangent and normal spaces of $M$ and $\widehat M$.

Note that given an orthonormal bases $e_j, \hat e_j, \epsilon_\kappa, \hat\epsilon_\kappa$ and the vector field $\dot x$ along $x$ on $M$ we used the rolling map $(q, p)$, satisfying no-slipping condition (I) and no-twisting conditions (II), (III), to obtain all the terms of the vector field $\dot{\mathcal R}$ defined on the bundle $Q \oplus P_{\iota,\hat\iota}$. Thus the vector field $\dot x$ was lifted from $M$ to the vector field $\dot{\mathcal R}$ on $Q \oplus P_{\iota,\hat\iota}$. We generalize this notion in the following definition.

**Definition 30.** *If $X$ is a vector field on $M$, then we define $\mathcal V(X)$ and $\mathcal V^{\perp}(X)$ as the vector fields on $Q \oplus P_{\iota,\hat\iota}$, such that under any local trivialization $h$ and any $(q,p) \in \pi^{-1}(x)$ they satisfy*

$$dh\,(\mathcal V(X)(q,p)) = \sum_{i<j} \big( g_M\big(\nabla_{X(x)} e_j, e_i\big) - g_{\widehat M}\big(\nabla_{qX(x)} q e_j, q e_i\big)\big) W_{ij}(A). \tag{6.8}$$

$$dh\,\big(\mathcal V^{\perp}(X)(q,p)\big) = \sum_{\kappa<\lambda} \Big( g^{\perp}_{M}\big(\nabla^{\perp}_{X(x)} \epsilon_\lambda, \epsilon_\kappa\big) - g^{\perp}_{\widehat M}\big(\nabla^{\perp}_{qX(x)} p\epsilon_\lambda, p\epsilon_\kappa\big)\Big) W_{\kappa\lambda}(B). \tag{6.9}$$

Notice that if $Y(x) = X(x) = v \in T_xM$, then $\mathcal V(Y)(q,p) = \mathcal V(X)(q,p)$ for every $(q,p) \in (Q \oplus P_{\iota,\hat\iota})_x$. Hence, we may define $\mathcal V(v)(q,p)$ whenever $v \in T_xM$ and $(q,p) \in (Q \oplus P_{\iota,\hat\iota})_x$. Also notice that the map $X \mapsto \mathcal V(X)$ is linear. The same holds for $\mathcal V^{\perp}$.

At first glance, it may seem that all of the coefficients of $W_{ij}(A)$ and $W_{\kappa\lambda}(A)$ in (6.8) and (6.9) vanish from conditions (II) and (III), that is not true, however. Even though $\frac{D}{dt}X(x(t)) = \nabla_{\dot x(t)} X(x(t))$ for any tangential vector field $X$ along $x$, in general, $\nabla_{q\dot x(t)} q(t) e_j$ does not coincide with $\frac{D}{dt} q(t) e_j(x(t))$. To see this, notice that

$$\begin{aligned}\frac{D}{dt}a_{sj}\hat{e}_s(\widehat{x}(t)) &= \dot{a}_{sj}\hat{e}_s(\widehat{x}(t)) + a_{sj}\nabla_{\dot{\widehat{x}}(t)}\hat{e}_s(\widehat{x}(t))\\ &= \dot{a}_{sj}\hat{e}_s(\widehat{x}(t)) + a_{sj}\nabla_{q\dot{x}(t)}\hat{e}_s(\widehat{x}(t)),\end{aligned}$$

while $\nabla_{q\dot{x}(t)}a_{sj}\hat{e}_s(x(t)) = a_{sj}\nabla_{q\dot{x}(t)}\hat{e}_s(x(t))$. Similar relations hold for $\frac{D^{\perp}}{dt}$.

We sum up our considerations in this section and define distributions

$$Q\oplus P_{\iota,\widehat{\iota}} \ni (q,p)\mapsto E_{(q,p)}\subset T_{(q,p)}\big(Q\oplus P_{\iota,\widehat{\iota}}\big) \quad\text{and}\quad Q\ni q\mapsto D_q\subset T_q(Q)$$

**Proposition 10.** *A curve $(q(t),p(t))$ in $Q\oplus P_{\iota,\widehat{\iota}}$ is a rolling if and only if it is a horizontal curve with respect to the distribution $E$, defined by*

$$E_{(q,p)} = \left\{v + qv + \mathcal{V}(v)(q,p) + \mathcal{V}^{\perp}(v)(q,p) \,|\, v\in T_xM\right\},$$

*where $(q,p)\in(Q\oplus P_{\iota,\widehat{\iota}})_{(x,\widehat{x})}$.*

Using the same symbol to denote the restriction of $\mathcal{V}(X)$ to $Q$, we have

**Proposition 11.** *A curve $q(t)$ in $Q$ is an intrinsic rolling if and only if it is a horizontal curve with respect to the distribution $D$, defined by*

$$D_q = \{v + qv + \mathcal{V}(v)(q) \,|\, v\in T_xM\}, \quad q\in Q_x.$$

We mention the book [19], where the strong approach to the control theory from the viewpoint of differential geometry is presented. Moreover, we recommend preprint [31] and [32, 60, 64] as a most complete sources of information about rolling of $n$-dimensional manifolds and the controllability of this process.

### 6.4. Examples of rollings and their controllability

In this section, we show two examples of rolling configurations, namely the sphere $S^n$ rolling on $\mathbb{R}^n$ and the special Euclidean group $SE(3)$ rolling on its Lie algebra $\mathfrak{se}(3)$. The first case is controllable, which follows from the fact that the distribution $D$, encoding rolling without slipping and twisting, is bracket generating and thus the Chow-Rashevskiĭ theorem holds. The second example is not controllable, which follows from the orbit theorem and a strong version of it for the case of analytic manifolds, see [3]. For the second example we study a particular intrinsic rolling, which we extend to an extrinsic rolling. This exemplifies the result obtained in Theorem 6.2.

A CONTROLLABLE EXAMPLE: $S^n$ ROLLING ON $\mathbb{R}^n$. We want to illustrate the properties of the distribution $D$ from Proposition 11, by proving that the unit sphere $S^n$ in $\mathbb{R}^{n+1}$ rolling over $\mathbb{R}^n$ is a completely controllable system. This result was obtained in [132] by different methods. The aim of this subsection is to show the controllability of the system directly from the Chow-Rashevskiĭ theorem.

Consider the unit sphere $S^n$ as the submanifold of $\mathbb{R}^{n+1}$,

$$S^n = \Big\{(x^0,\dots,x^n)\in\mathbb{R}^{n+1} \;|\; \sum_{k=0}^{n}(x^k)^2 = 1\Big\},$$

with the induced metric $g$. For an arbitrary point $\tilde{x} = (\tilde{x}^0, \dots, \tilde{x}^n) \in S^n$, at least one of the coordinates $\tilde{x}^0, \dots, \tilde{x}^n$ does not vanish. Without loss of generality, we may assume that $\tilde{x}^n \neq 0$, and consider the neighborhood

$$U = \{(x^0, \dots, x^n) \in S^n |\ \pm x^n > 0\},$$

where the choice of the ± sign depends on the sign of $\tilde{x}^n$. To simplify the notation, we define the functions $s_j(x) = \sum_{r=j}^n (x^r)^2$ on $U$, for $j = 1, \dots, n$, that are strictly positive on $U$. We use them to define an orthonormal basis of $TU$. Define the vector fields $e_j$ on $U$

$$e_j = \sqrt{\frac{s_j}{s_{j-1}}} \left( -\frac{\partial}{\partial x^{j-1}} + \frac{x^{j-1}}{s_j} \sum_{r=j}^{n} x^r \frac{\partial}{\partial x^r} \right), \qquad j = 1, \dots, n. \tag{6.10}$$

These vector fields form an orthonormal basis of the tangent space over $U$ and we denote by $\Gamma_{ij}^k$ the Christoffel symbols with respect to the basis $\{e_i\}_{i=1}^n$. We set $\hat{e}_i = \dfrac{\partial}{\partial \widehat{x}^i}$ to be the standard basis of $T\mathbb{R}^n$. Let us state two technical lemmas whose proofs can be obtained by direct calculations or found in [56].

**Lemma 5.** *Let $1 \leq i < j \leq n$. Then $\Gamma_{kj}^i = -\dfrac{x^{i-1}}{\sqrt{s_{i-1} s_i}} \delta_{k,j}$ for all $k = 1, \dots, n$.*

For convenience, we denote $\Gamma^k = \frac{x^{k-1}}{\sqrt{s_{k-1} s_k}} = -\Gamma_{aa}^k$, $a = 1, \dots, n$. The properties of the Levi-Civita connection $\nabla$ have the following consequences:

- The compatibility of $\nabla$ with the metric $g$ and $g(e_i, e_j) = \delta_{i,j}$, imply that $\Gamma_{kj}^i = -\Gamma_{ki}^j$. In particular, $\Gamma_{ki}^i = 0$.
- The symmetry of $\nabla$ implies that, if $l < k$, then

$$[e_k, e_l] = \nabla_{e_k} e_l - \nabla_{e_l} e_k = \sum_{i=1}^n g(\nabla_{e_k} e_l - \nabla_{e_l} e_k, e_i) e_i = \Gamma^l e_k.$$

**Lemma 6.** *The vector fields $e_k$ on the sphere satisfy*

$$e_k \left( \Gamma^l \right) = \begin{cases} 0\,, & k > l, \\ -1/s_k\,, & k = l, \\ -\Gamma^l\, \Gamma^k\,, & k < l. \end{cases} \qquad k, l = 1, 2, \dots, n$$

Observe that $\nabla_{\hat{e}_k} \hat{e}_l = 0$, and $[\hat{e}_k, \hat{e}_l] = 0$ for all $k, l = 1, \dots, n$.

Consider the vector fields $X_k = e_k + q e_k + \mathcal{V}(e_k)$, restricted to $U$, which span the distribution $D$ introduced in Proposition 11. In this case, we have the explicit form

$$X_k(x, \hat{x}, A) = e_k(x) + \sum_{i=1}^n a^{ik} \hat{e}_i(\widehat{x}) - \sum_{i=1}^{k-1} \Gamma^i\, W_{ik}(A).$$

In order to determine $[X_k, X_l]$, we assume that $k > l$. Then

$$[X_k, X_l] = \Gamma^l\, X_k - W_{lk}.$$

Define the vector fields $Y_{lk}$, for $l < k$, by $Y_{lk} := [X_l, X_k] + \Gamma^l X_k = W_{lk}$. Finally, let

$$Z_1 = [Y_{12}, X_2] = \sum_{i=1}^{n} a^{i1}\hat{e}_i, \quad Z_k = [X_1, Y_{1k}] = \sum_{i=1}^{n} a^{ik}\hat{e}_i\,, \qquad k = 2, \dots, n.$$

We conclude that the entire tangent space is spanned by the vector fields $\{X_k\}_{k=1}^n$, $\{Y_{lk}\}_{1\le l<k\le n}$ and $\{Z_k\}_{k=1}^n$. Hence, $D$ is a regular bracket generating distribution of step 3, which implies that the system of rolling $S^n$ over $\mathbb{R}^n$ is completely controllable.

**6.4.1. A non-controllable example: $SE(3)$ rolling on $\mathbb{R}^6$.** We consider the case of $SE(3)$, endowed with a left invariant metric defined later, rolling over its tangent space at the identity $T_1SE(3) = \mathfrak{se}(3)$, with metric obtained by restricting the left invariant metric on $SE(3)$ to the identity. Our goal is to determine whether the system is controllable.

We give $SE(3)$ coordinates as follows. For any $x \in \mathrm{SE}(3)$ there exist $C = (c_{ij}) \in SO(3)$ and $r = (r_1, r_2, r_3) \in \mathbb{R}^3$, such that $x = (C, r)$ acts on $\mathbb{R}^3$ via

$$x(y) = Cy + r, \qquad \text{for all } y \in \mathbb{R}^3.$$

The tangent space of $SE(3)$ at $x = (C, r)$ is spanned by the left invariant vector fields

$$e_1 = Y_1 = \frac{1}{\sqrt{2}}\left(C\cdot\frac{\partial}{\partial c_{12}} - C\cdot\frac{\partial}{\partial c_{21}}\right) = \frac{1}{\sqrt{2}}\sum_{j=1}^{3}\left(c_{j1}\frac{\partial}{\partial c_{j2}} - c_{j2}\frac{\partial}{\partial c_{j1}}\right)$$

$$e_2 = Y_2 = \frac{1}{\sqrt{2}}\left(C\cdot\frac{\partial}{\partial c_{13}} - C\cdot\frac{\partial}{\partial c_{31}}\right) = \frac{1}{\sqrt{2}}\sum_{j=1}^{3}\left(c_{j1}\frac{\partial}{\partial c_{j3}} - c_{j3}\frac{\partial}{\partial c_{j1}}\right)$$

$$e_3 = Y_3 = \frac{1}{\sqrt{2}}\left(C\cdot\frac{\partial}{\partial c_{23}} - C\cdot\frac{\partial}{\partial c_{32}}\right) = \frac{1}{\sqrt{2}}\sum_{j=1}^{3}\left(c_{j2}\frac{\partial}{\partial c_{j3}} - c_{j3}\frac{\partial}{\partial c_{j2}}\right)$$

$$e_{k+3} = X_k = C\cdot\frac{\partial}{\partial r_k} = \sum_{j=1}^{3} c_{jk}\frac{\partial}{\partial r_j} \qquad\qquad k = 1, 2, 3.$$

Define a left invariant metric on SE(3) by declaring the vectors $e_1, \dots, e_6$ to form an orthonormal basis. The mapping

$$\sum_{j=1}^{6}\widehat{x}_j e_j(1) \mapsto (\widehat{x}_1, \widehat{x}_2, \widehat{x}_3, \widehat{x}_4, \widehat{x}_5, \widehat{x}_6) \in \mathbb{R}^6$$

permits to identify $\mathfrak{se}(3)$, endowed with the induced metric, with $\mathbb{R}^6$ endowed with the Euclidean metric. We write $\hat{e}_k = \frac{\partial}{\partial \widehat{x}_k}$ on $\mathbb{R}^6$ and study the behavior of the intrinsic rollings of SE(3) on $\mathbb{R}^6$. Note that the configuration space $Q$ for intrinsic rolling is $\mathrm{SE}(3)\times\mathbb{R}^6\times\mathrm{SO}(6)$, because both manifolds SE(3) and $\mathbb{R}^6$ are Lie groups, so their tangent bundles are trivial, and $\dim Q = 27$.

The covariant derivatives $\nabla_{e_i} e_j$ are non-zero only in the following cases

$$\nabla_{Y_1} Y_2 = -\frac{1}{2\sqrt{2}}\, Y_3, \quad \nabla_{Y_1} Y_3 = \frac{1}{2\sqrt{2}}\, Y_2, \quad \nabla_{Y_2} Y_3 = -\frac{1}{2\sqrt{2}}\, Y_1,$$

$$\nabla_{Y_1} X_k = \frac{1}{\sqrt{2}} \left(\delta_{2,k} X_1 - \delta_{1,k} X_2\right), \quad \nabla_{Y_2} X_k = \frac{1}{\sqrt{2}} \left(\delta_{3,k} X_1 - \delta_{1,k} X_3\right),$$

$$\nabla_{Y_3} X_k = \frac{1}{\sqrt{2}} \left(\delta_{3,k} X_2 - \delta_{2,k} X_3\right),$$

where $\delta_{i,j}$ denotes the Kronecker symbol. On the other hand, it is well-known that $\nabla_{\hat{e}_i} \hat{e}_j = 0$ for any $i, j$. Proposition 11 and Definition 30 show that the distribution $D$ over $Q$ is spanned by

$$Z_1 = Y_1 + qY_1 + \frac{1}{2\sqrt{2}} W_{23} + \frac{1}{\sqrt{2}} W_{45}, \qquad Z_2 = Y_2 + qY_2 - \frac{1}{2\sqrt{2}} W_{13} + \frac{1}{\sqrt{2}} W_{46},$$

$$Z_3 = Y_3 + qY_3 + \frac{1}{2\sqrt{2}} W_{12} + \frac{1}{\sqrt{2}} W_{56}, \tag{6.11}$$

$$K_1 = X_1 + qX_1, \qquad K_2 = X_2 + qX_2, \qquad K_3 = X_3 + qX_3.$$

In the case of $D$, defined by the vector fields (6.11) straightforward calculations yield that the flag associated to $D$ is of the form

$$\begin{array}{rcl} D^2 & = & D \oplus \operatorname{span}\{W_{12}, W_{13}, W_{23}\}, \\ D^3 & = & D^2 \oplus \operatorname{span}\{qY_1, qY_2, qY_3\}, \\ D^4 & = & D^3, \end{array} \tag{6.12}$$

and so $\dim D^2 = 9$, $\dim D^k = 12$ for all $k \geq 3$ and the step of $D$ is 3.

It follows from (6.12) that the distribution is not bracket generating and we can not produce all the tangent space to the configuration space, that has dimension 27 at each point. We conclude that the system is not controllable by orbit theorem [3].

**6.4.2. Concrete example of an extrinsic and intrinsic rolling.** We present the construction of the intrinsic rolling of the manifold $M = \mathrm{SE}(3)$ over $\mathbb{R}^6$: $q(t) = (x(t), \widehat{x}(t), A(t))$, with initial conditions

$$x(0) = \mathrm{Id}_{\mathbb{R}^3}, \quad \widehat{x}(0) = 0, \quad A(0) = \mathbf{1}$$

and the curve $x\colon [0, \tau] \to \mathrm{SE}(3)$

$$x(t)y = \begin{pmatrix} \cos\theta(t) & \sin\theta(t) & 0 \\ -\sin\theta(t) & \cos\theta(t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \psi(t) \end{pmatrix}, \qquad y \in \mathbb{R}^3, \tag{6.13}$$

where $\theta(t)$ and $\psi(t)$ are smooth functions with $\theta(0) = \psi(0) = 0$. Then

$$\dot{x} = \sqrt{2}\, \dot{\theta}(t) Y_1 + \dot{\psi}(t) X_3$$

for every $t$, and the intrinsic rolling has the form $\dot{q} = \sqrt{2}\,\dot{\theta}(t)Z_1 + \dot{\psi}(t)K_3$. This implies that

$$\dot{\widehat{x}}(t) = \sqrt{2}\,\dot{\theta}(t)\,qY_1 + \dot{\psi}(t)\,qX_3, \qquad \dot{A}(t) = \dot{\theta}(t)\left(\frac{1}{2}W_{23}(A) + W_{45}(A)\right) \tag{6.14}$$

for almost every $t$. It follows from the first equation in (6.14) that

$$\widehat{x}(t) = \left(\sqrt{2}\,\theta(t), 0, 0, 0, 0, \psi(t)\right)^{tr},$$

where $^{tr}$ denotes transposition. By exponentiating the second equation in (6.14), we obtain

$$A(t) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \cos\left(\frac{\theta(t)}{2}\right) & \sin\left(\frac{\theta(t)}{2}\right) & 0 & 0 & 0 \\ 0 & -\sin\left(\frac{\theta(t)}{2}\right) & \cos\left(\frac{\theta(t)}{2}\right) & 0 & 0 & 0 \\ 0 & 0 & 0 & \cos\theta(t) & \sin\theta(t) & 0 \\ 0 & 0 & 0 & -\sin\theta(t) & \cos\theta(t) & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

EMBEDDING OF $SE(n)$ INTO EUCLIDEAN SPACE. In the next step we want to extend the intrinsic rolling to the extrinsic rolling. We choose the following isometric embedding of $SE(n)$ into the Euclidean space $\mathbb{R}^{(n+1)^2}$. Identify an element $\overline{C} \in \mathbb{R}^{(n+1)^2}$ with the matrix

$$\overline{C} = \begin{pmatrix} \overline{c}_{11} & \cdots & \overline{c}_{1,n+1} \\ \vdots & \ddots & \vdots \\ \overline{c}_{n+1,1} & \cdots & \overline{c}_{n+1,n+1} \end{pmatrix}.$$

Define the inner product on $\mathbb{R}^{(n+1)^2}$ by $\langle \overline{C}_1, \overline{C}_2 \rangle = \operatorname{trace}\left((\overline{C}_1)^t\,\overline{C}_2\right)$. Note that since $\langle \overline{C}, \overline{C} \rangle = \sum_{i,j=1}^{n+1} |\overline{c}_{ij}|^2$, the metric $\langle \cdot\,, \cdot \rangle$ coincides with the Euclidean metric in $\mathbb{R}^{(n+1)^2}$. From this we get that $\left\{\frac{\partial}{\partial \overline{c}_{ij}}\right\}_{i,j=1}^{n+1}$ is an orthonormal basis for the tangent bundle $T\mathbb{R}^{(n+1)^2}$ with respect to $\langle \cdot\,, \cdot \rangle$.

We define the embedding of $SE(n)$ into $\mathbb{R}^{(n+1)^2}$ by

$$\begin{array}{rccl} \iota: & SE(n) & \to & \mathbb{R}^{(n+1)^2} \\ & x = (C, r) & \mapsto & \overline{C} = \begin{pmatrix} C & r \\ 0 & 1 \end{pmatrix}. \end{array}$$

This mapping is in fact an isometry of $SE(n)$ onto its image. To see this, notice that the metrics coincide at the identity, and that the metric of $\mathbb{R}^{(n+1)^2}$, restricted to the image $\mathrm{Im}(\iota)$ of $\iota$, is left invariant under the action of $SE(n)$. Hence, the metrics on $SE(n)$ and $\mathrm{Im}(\iota)$ coincide, and $\iota$ defines an isometric embedding.

EXTRINSIC ROLLING OF $SE(3)$ OVER $\mathfrak{se}(3)$. We will use the constructed embedding to build an extrinsic rolling of $SE(3)$ over $\mathfrak{se}(3)$ in $\mathbb{R}^{16}$ and we write $M$ for the image of $SE(3)$ in $\mathbb{R}^{16}$ under this embedding. Denote $\partial_{ij} = \frac{\partial}{\partial \overline{c}_{ij}}$, then the vector fields spanning $TM$ are

$$e_1 = Y_1 = \frac{1}{\sqrt{2}} \sum_{i=1}^{3} \left( \overline{c}_{i1} \partial_{i2} - \overline{c}_{i2} \partial_{i1} \right), \quad e_2 = Y_2 = \frac{1}{\sqrt{2}} \sum_{i=1}^{3} \left( \overline{c}_{i1} \partial_{i3} - \overline{c}_{i3} \partial_{i1} \right),$$

$$e_3 = Y_3 = \frac{1}{\sqrt{2}} \sum_{i=1}^{3} \left( \overline{c}_{i2} \partial_{i3} - \overline{c}_{i3} \partial_{i2} \right), \quad e_{3+k} = X_k = \sum_{i=1}^{3} \overline{c}_{ik} \partial_{i4}, \quad k = 1, 2, 3, \tag{6.15}$$

where we suppressed $d\iota$ in the notation. We introduce an orthonormal basis of $TM^{\perp}$

$$\Upsilon_1 = \frac{1}{\sqrt{2}} \sum_{j=1}^{3} \left( \overline{c}_{j1} \partial_{j2} + \overline{c}_{j2} \partial_{j1} \right), \quad \Upsilon_2 = \frac{1}{\sqrt{2}} \sum_{j=1}^{3} \left( \overline{c}_{j1} \partial_{j3} + \overline{c}_{j3} \partial_{j1} \right),$$

$$\Upsilon_3 = \frac{1}{\sqrt{2}} \sum_{j=1}^{3} \left( \overline{c}_{j2} \partial_{j3} + \overline{c}_{j3} \partial_{j2} \right),$$

$$\Psi_\lambda = \sum_{j=1}^{3} \overline{c}_{j\lambda} \partial_{j\lambda}, \quad \lambda = 1, 2, 3, \qquad \Xi_\mu = \partial_{4\mu}, \quad \mu = 1, 2, 3, 4. \tag{6.16}$$

We denote by $\widehat{M}$ the image of $\mathbb{R}^6$ into $\mathbb{R}^{16}$ by the embedding

$$(\widehat{x}_1, \widehat{x}_2, \widehat{x}_3, \widehat{x}_4, \widehat{x}_5, \widehat{x}_6) \overset{\widehat{\iota}}{\mapsto} \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} \widehat{x}_1 & \frac{1}{\sqrt{2}} \widehat{x}_2 & \widehat{x}_4 \\ -\frac{1}{\sqrt{2}} \widehat{x}_1 & 0 & \frac{1}{\sqrt{2}} \widehat{x}_3 & \widehat{x}_5 \\ -\frac{1}{\sqrt{2}} \widehat{x}_2 & -\frac{1}{\sqrt{2}} \widehat{x}_3 & 0 & \widehat{x}_6 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

We have the following orthonormal basis of $T\widehat{M}$,

$$\hat{e}_1 = \frac{1}{\sqrt{2}} (\partial_{12} - \partial_{21}), \quad \hat{e}_2 = \frac{1}{\sqrt{2}} (\partial_{13} - \partial_{31}), \quad \hat{e}_3 = \frac{1}{\sqrt{2}} (\partial_{23} - \partial_{32}),$$

$$\hat{e}_{3+k} = \partial_{k4}, \qquad k = 1, 2, 3,$$

while the vector fields spanning $T\widehat{M}^{\perp}$ are

$$\hat{\epsilon}_1 = \frac{1}{\sqrt{2}} (\partial_{12} + \partial_{21}), \quad \hat{\epsilon}_2 = \frac{1}{\sqrt{2}} (\partial_{13} + \partial_{31}), \quad \hat{\epsilon}_3 = \frac{1}{\sqrt{2}} (\partial_{23} + \partial_{32}),$$

$$\hat{\epsilon}_{3+\kappa} = \partial_{\kappa\kappa}, \qquad \kappa = 1, 2, 3, \qquad \hat{\epsilon}_{6+\varkappa} = \partial_{4\varkappa}, \qquad \varkappa = 1, 2, 3, 4.$$

In order to extend an intrinsic rolling $q$ with $\pi(q) = (x, \widehat{x})$, we find an orthonormal frame of normal parallel vector fields along curves $x$ and $\widehat{x}$. Along $\widehat{x}$, we may use the restriction of $\{\hat{\epsilon}_\kappa\}_{\kappa=1}^{10}$. For the curve $x$ the answer is more complicated.

We first study the value of $\nabla^{\perp}$ for different choices of vector fields.

(1) $\nabla^{\perp}_{X} \Xi_\mu = 0$, for any tangential vector field $X$, and $\Xi_\mu$ as in equation (6.16).

(2) $\nabla^{\perp}_{X_k} \Upsilon = 0$, for any normal vector field $\Upsilon$, and $X_k$ as in equation (6.15).

(3) Otherwise, the results are presented in the following table

| | $\Upsilon_1$ | $\Upsilon_2$ | $\Upsilon_3$ | $\Psi_1$ | $\Psi_2$ | $\Psi_3$ |
|---|---|---|---|---|---|---|
| $\nabla^{\perp}_{Y_1}$ | $\frac{1}{2}(\Psi_1-\Psi_2)$ | $-\frac{1}{2\sqrt{2}}\Upsilon_3$ | $\frac{1}{2\sqrt{2}}\Upsilon_2$ | $-\frac{1}{2}\Upsilon_1$ | $\frac{1}{2}\Upsilon_1$ | $0$ |
| $\nabla^{\perp}_{Y_2}$ | $-\frac{1}{2\sqrt{2}}\Upsilon_3$ | $\frac{1}{2}(\Psi_1-\Psi_3)$ | $\frac{1}{2\sqrt{2}}\Upsilon_1$ | $-\frac{1}{2}\Upsilon_2$ | $0$ | $\frac{1}{2}\Upsilon_2$ |
| $\nabla^{\perp}_{Y_3}$ | $-\frac{1}{2\sqrt{2}}\Upsilon_2$ | $\frac{1}{2\sqrt{2}}\Upsilon_1$ | $\frac{1}{2}(\Psi_2-\Psi_3)$ | $0$ | $-\frac{1}{2}\Upsilon_3$ | $\frac{1}{2}\Upsilon_3$ |

.

We use the relations above to construct an extrinsic rolling by making use of the curve (6.13). Since $\dot{x}(t)=\sqrt{2}\,\dot{\theta}(t)Y_1(x(t))+\dot{\psi}(t)X_3(x(t))$, the vector field

$$\Psi(t)=\sum_{\lambda=1}^{3}\left(\upsilon_{\lambda}(t)\Upsilon_{\lambda}(x(t))+\upsilon_{3+\lambda}(t)\Psi_{\lambda}(x(t))\right),$$

is normal parallel along $x(t)$ if

$$\left(\dot{\upsilon}_1-\frac{\dot{\theta}}{\sqrt{2}}(\upsilon_4-\upsilon_5)\right)\Upsilon_1+\left(\dot{\upsilon}_2+\frac{\dot{\theta}}{2}\upsilon_3\right)\Upsilon_2+\left(\dot{\upsilon}_3-\frac{\dot{\theta}}{2}\upsilon_2\right)\Upsilon_3$$
$$+\left(\dot{\upsilon}_4+\frac{\dot{\theta}}{\sqrt{2}}\upsilon_1\right)\Psi_1+\left(\dot{\upsilon}_5-\frac{\dot{\theta}}{\sqrt{2}}\upsilon_1\right)\Psi_2+\upsilon_6\Psi_3=0.$$

Hence we define a parallel orthonormal frame along $x(t)$ by

$$\epsilon_1(t)=\cos\theta\Upsilon_1(x(t))-\frac{1}{\sqrt{2}}\sin\theta\Psi_1(x(t))+\frac{1}{\sqrt{2}}\sin\theta\Psi_2(x(t)),$$

$$\epsilon_2(t)=\cos\left(\frac{\theta}{2}\right)\Upsilon_2(x(t))+\sin\left(\frac{\theta}{2}\right)\Upsilon_3(x(t)),$$

$$\epsilon_3(t)=-\sin\left(\frac{\theta}{2}\right)\Upsilon_2(x(t))+\cos\left(\frac{\theta}{2}\right)\Upsilon_3(x(t)),$$

$$\epsilon_4(t)=\frac{1}{\sqrt{2}}\sin\theta\Upsilon_1(x(t))+\frac{\cos\theta+1}{2}\Psi_1(x(t))+\frac{1-\cos\theta}{2}\Psi_2(x(t)),$$

$$\epsilon_5(t)=-\frac{1}{\sqrt{2}}\sin\theta\Upsilon_1(x(t))+\frac{1-\cos\theta}{2}\Psi_1(x(t))+\frac{1+\cos\theta}{2}\Psi_2(x(t)),$$

$$\epsilon_6(t)=\Psi_3(x(t))\quad\text{and}\quad\epsilon_{6+\lambda}(t)=\Xi_{\lambda}(x(t)),\qquad\lambda=1,2,3,4.$$

Thus $p(t)$ is represented by a constant matrix in the bases $\{\epsilon_{\lambda}(t)\}_{\lambda=1}^{10}$ and $\{\hat{\epsilon}_{\kappa}(t)\}_{\kappa=1}^{10}$. Let us choose $p(t)$ to be the identity in these bases, due to the given embedding.

The curve $\mathcal{R}(t) = (q(t), p(t))$ in $SE(16)$ is given by $\mathcal{R}(t)(t)\overline{x} = \overline{A}\overline{x} + \overline{r}(t)$, where $\overline{A}(t)$ equals

$$\left(\begin{array}{cccccccccc|c}
\cos^2\frac{\theta}{2} & -\frac{\sin\theta}{2} & 0 & 0 & \frac{\sin\theta}{2} & \frac{\cos\theta-1}{2} & 0 & 0 & 0 & 0 & \\
\frac{\sin\theta}{2} & \cos^2\frac{\theta}{2} & 0 & 0 & \sin^2\frac{\theta}{2} & \frac{\sin\theta}{2} & 0 & 0 & 0 & 0 & \\
0 & 0 & \cos\frac{\theta}{2} & 0 & 0 & 0 & \sin\frac{\theta}{2} & 0 & 0 & 0 & \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \\
-\frac{\sin\theta}{2} & \sin^2\frac{\theta}{2} & 0 & 0 & \cos^2\frac{\theta}{2} & -\frac{\sin\theta}{2} & 0 & 0 & 0 & 0 & \mathbf{0}_{10\times 6}\\
\frac{\cos\theta-1}{2} & -\frac{\sin\theta}{2} & 0 & 0 & \frac{\sin\theta}{2} & \cos^2\frac{\theta}{2} & 0 & 0 & 0 & 0 & \\
0 & 0 & -\sin\frac{\theta}{2} & 0 & 0 & 0 & \cos\frac{\theta}{2} & 0 & 0 & 0 & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sin\frac{\theta}{2} & \cos\frac{\theta}{2} & \\
\hline
\multicolumn{10}{c|}{\mathbf{0}_{6\times 10}} & \mathbf{1}_6
\end{array}\right),$$

and $\overline{r}(t) = \left(-1, \frac{\theta}{\sqrt{2}}, 0, 0, \frac{\theta}{\sqrt{2}}, -1, 0, 0, 0, 0, -1, 0, 0, 0, 0, 0\right)^t$. Here, $\mathbf{0}_{m\times n}$ denotes the zero matrix of size $m \times n$ and $\mathbf{1}_6$ is the identity matrix of size $6 \times 6$.

## 7. Group of diffeomorphisms of the circle

In this section we consider one of the simplest examples of an infinite dimensional manifold possessing rather good structure and having various applications. We start from the definitions, then we give motivations and we finish with some results concerning sub-Riemannian structure of this example. The material of this section can be found in [81, 93, 99, 108, 116].

### 7.1. Manifold and group structure of Diff $S^1$.

Let $S^1$ be the unit circle with the standard counterclockwise orientation that we will think of as $\mathbb{R} \mod 2\pi\mathbb{Z}$. Denote by Diff $S^1$ the set of all $C^\infty$-diffeomorphisms of $S^1$ preserving the orientation. This is an open subset of the infinite dimensional space $C^\infty(S^1, S^1)$ of all $C^\infty$-smooth maps of the circle, see Exercise 1. The topology defined on the space $C^\infty(S^1, S^1)$ is the $C^\infty$ topology, which corresponds to the uniform convergence of all derivatives of $h \in C^\infty(S^1, S^1)$ and it can be described by the countable system of semi-norms

$$\|h\|_m = \sup_{\theta\in S^1} |h^m(\theta)|.$$

The set Diff $S^1$, considered as an open set of $C^\infty(S^1, S^1)$, inherits the $C^\infty$-topology.

Define the group operation in Diff $S^1$ as the composition of two diffeomorphisms and the inverse element of the group as the inverse map. The identity map $\mathrm{id}\colon S^1 \to S^1$ becomes the unit of the group. We would like to consider the defined group as an infinite dimensional analogue of a Lie group. To do this we need to define a manifold structure on the topological space Diff $S^1$. We start from the choice of a model vector space $\mathcal{V}$, where the coordinate charts $(U, \varphi)$ with $U$ being an open set on Diff $S^1$ and $\varphi\colon U \to V \subset \mathcal{V}$ being a diffeomorphism, will be constructed. The space $\mathrm{Vect}(S^1)$ of all $C^\infty$ real vector fields on the circle is a good candidate for the model space $\mathcal{V}$ since it possesses the Frechét topology and it is one of the examples of a locally convex complete topological vector space. The vector fields $S^1 \ni \theta \to v(\theta)\frac{d}{d\theta} = v(\theta)\partial_\theta \in \operatorname{Vect}(S^1)$ can be associated with the space $C^\infty(S^1, \mathbb{R})$ of real functions $v$. A sketch of the construction of the chart is given below. Let us consider the following neighborhood $V_0$ of $0 \in \operatorname{Vect}(S^1)$: $V_0 = \{v \in \operatorname{Vect}(S^1) \mid |v| < \pi\}$, where $|v|$ is the absolute value in $\mathbb{R}$. Then there is a homeomorphic map $\psi\colon V_0 \to U_0 \subset C^\infty(S^1, S^1)$ to a neighborhood

$$U_0 = \{f \in C^\infty(S^1, S^1) \mid f(\theta) \neq -\theta, \quad \text{for all} \quad \theta \in S^1\},$$

of the identity map $\mathrm{id} \in C^\infty(S^1, S^1)$. Construction of $\psi$ see in Exercise 2. Choose an open set $U \subset U_0$ consisting of diffeomorphisms. Then $U$ is a neighborhood of $\mathrm{id} \in \operatorname{Diff} S^1$. Then the set $V = \psi^{-1}(U)$ will be an open subset of $V_0 \in \operatorname{Vect}(S^1)$. Thus, we constructed the chart $(U, \varphi)$, $\varphi = \psi^{-1}|_U$ in a neighborhood of the identity map $\mathrm{id} \in \operatorname{Diff} S^1$. To construct a complete atlas we exploit the group structure of $\operatorname{Diff} S^1$. If $U$ is a neighborhood of $\mathrm{id} \in \operatorname{Diff} S^1$, then $f.U = \{\phi \in \operatorname{Diff} S^1 \mid \phi = f \circ h,\ h \in U\}$ is a neighborhood of $f$. Having the map $\psi$, we define $\psi_f\colon V \to f.U$ for any $f \in \operatorname{Diff} S^1$ as the composition $f \circ \psi$. Then the chart $(U_f, \varphi_f)$, where $U_f = f.U$ and $\varphi_f = \psi_f^{-1}|_{f.U}$ is the corresponding local chart about any group element $f \in \operatorname{Diff} S^1$. The reader has to verify that any composition $\varphi_h \circ \varphi_f^{-1}$ is a smooth map in $\operatorname{Vect}(S^1)$.

As we know, the space $\operatorname{Vect}(S^1)$, endowed with usual brackets for vector fields, forms a Lie algebra, that we shall denote by the same symbol $\operatorname{Vect}(S^1)$. On the other hand, the space $T_{\mathrm{id}} \operatorname{Diff} S^1$ furnished with brackets for tangent vectors is the Lie algebra $\mathfrak{diff}$ of left invariant vector fields. As we discussed in Subsection 2.1, a tangent vector is the equivalence class of all curves $f\colon \mathbb{R} \to \operatorname{Diff} S^1$ such that $f(0) = \mathrm{id}$ and that have the same initial velocity. Such curves can be seen as smooth functions

$$f\colon \mathbb{R} \times S^1 \to S^1, \quad f(0, \theta) = \theta,$$

where $f(t_0, \theta) \in \operatorname{Diff} S^1$ for any fixed value $t_0 \in \mathbb{R}$. The velocity vector at $t = 0$, corresponding to the equivalence class $[f]$, is a vector field

$$v(\theta) = \frac{\partial}{\partial t}\Big|_{t=0} f(t, \theta), \quad f \in [f],$$

defined on $S^1$. By this we identify $\operatorname{Vect}(S^1)$ and $\mathfrak{diff}$ with the Lie product $[X, Y] = YX - XY$. Now we explain why the sign in the latter commutation relation is opposite. The action of the group $\operatorname{Diff} S^1$ on the circle $S^1$, considered as a smooth compact oriented manifold, is defined as the natural left action

$$\begin{array}{cccc} \mu\colon & \operatorname{Diff} S^1 \times S^1 & \to & S^1 \\ & f.\theta & \mapsto & f(\theta). \end{array}$$

The differential $d\mu$ acts at the level of tangent spaces: $d\mu\colon T(\operatorname{Diff} S^1 \times S^1) \cong T\operatorname{Diff} S^1 \times TS^1 \to TS^1$. In order to make the following diagram commutative

$$\begin{tikzcd}
T(\operatorname{Diff} S^1 \times S^1) \ni w(f,\theta) \arrow[r, "d\mu"] & v(f(\theta)) \in TS^1 \\
\operatorname{Diff} S^1 \times S^1 \ni (f,\theta) \arrow[u, "w"] \arrow[r, "\mu"'] & f(\theta) \in S^1 \arrow[u, "v"']
\end{tikzcd}$$

or $d\mu(w(f,\theta)) = v(\mu(f,\theta)) = v(f(\theta))$ we have to extend the elements of $\mathfrak{diff}$ as right invariant vector fields on $\operatorname{Diff} S^1$, or $w(f,\cdot) = \frac{d}{d\epsilon}\Big|_{\epsilon=0}(\exp \epsilon w_0)f(\cdot)$, where $w_0 \in \mathfrak{diff} = T_{\mathrm{id}}\operatorname{Diff} S^1 \cong \operatorname{Vect}(S^1)$ and the exponential map is defined in (7.1). In this way we associate the elements of $\mathfrak{diff}$ with right invariant vector fields on $\operatorname{Diff} S^1$. The bracket in $\mathfrak{diff}$ is defined as in $\operatorname{Vect}(S^1)$, but with the opposite sign, see the discussion in Subsection 3.1. Thus

$$\mathfrak{diff} = \Big(\operatorname{Vect}(S^1), [\ ,\ ]\Big) \quad \text{with} \quad [u\partial_\theta, v\partial_\theta] = (-uv' + u'v)\partial_\theta.$$

The group exponential map is defined in a standard way through one parametric subgroup in $\operatorname{Diff} S^1$. The one parametric subgroup $\mathbb{R} \ni t \mapsto f(t,\theta) \in \operatorname{Diff} S^1$ of diffeomorphisms for any fixed value $\theta_0 \in S^1$ is the solution of the Cauchy problem

$$\begin{cases} \frac{d\theta}{dt} = v(\theta) \in T_\theta S^1, \quad \text{where} \quad v(\theta) = \frac{\partial}{\partial t}\Big|_{t=0} f(t,\theta) \\ \theta(0) = \theta_0. \end{cases}$$

Thus, the solution $\theta(t) = \exp_{\theta_0}(tv)$ is the exponential curve that carries each line $t \mapsto tv$ through the origin in $\operatorname{Vect}(S^1)$ to the one parametric subgroup in $\operatorname{Diff} S^1$ and

$$\begin{array}{rccc} \exp\colon & \operatorname{Vect}(S^1) & \to & \operatorname{Diff} S^1 \\ & v & \mapsto & \exp(v) \end{array} \tag{7.1}$$

is the $\operatorname{Diff} S^1$-group exponential map.

The interesting feature is that, in contrast to finite dimensional groups, there is no neighborhood of $\mathrm{id} \in \operatorname{Diff} S^1$, where the exponential map would be diffeomorphism. The map *is not locally surjective.* No matter how small the neighborhood $U$ of the identity map $\mathrm{id} \in \operatorname{Diff} S^1$, there is $f \in U$ such that $f$ does not belong to any one-parametric subgroup, or $f \notin \operatorname{Im}(\exp)$, see Exercise 3.

Moreover, the map exp *is not injective.* Let $f_n$ be a rotation by the angle $\frac{2\pi}{n}\colon \theta \overset{f_n}{\mapsto} \theta + \frac{2\pi}{n}$. The map $f_n$ belongs to the closed subgroup $S^1 \subset \operatorname{Diff} S^1$ of rotations in $\operatorname{Diff} S^1$. Any one-parameter element of $S^1$ is generated by a constant vector field $v$ in (7.1). Let

$$\mathbb{H} = \{\phi \in \operatorname{Diff} S^1 \mid \phi\Big(\theta + \frac{2\pi}{n}\Big) = \phi(\theta) + \frac{2\pi}{n}\}$$

be the subgroup of Diff $S^1$ of all periodic diffeomorphisms with period $\frac{2\pi}{n}$. It easy to see that $f_n$ commutes with $\mathbb{H}$. Then since

$$S^1 \ni f_n = \phi f_n \phi^{-1} \in \phi S^1 \phi^{-1}, \quad \phi \in \mathbb{H},$$

we conclude that $f_n$ belongs to all one-parametric subgroups from $\phi S^1 \phi^{-1}$, $\phi \in \mathbb{H}$.

Finally, we state a property of the group Diff $S^1$ and its Lie algebra Vect$(S^1)$ whose proof can be found in [70].

**Proposition 12.** *The group* Diff $S^1$ *is a simple group.*

For practical purposes it is convenient to introduce a basis of Vect$(S^1)$. Since we identify the vector fields $v(\theta)\partial_\theta$ with smooth real functions $v(\theta)$ on $S^1$ and the latter can be developed into the Fourier series, therefore, the Fourier basis $\cos(n\theta)$, $\sin(n\theta)$, $n = 0, 1, \dots$ is a natural choice of a basis for Vect$(S^1)$.

We observe also that all $f \in$ Diff $S^1$ are periodic functions with the period $2\pi$ in the following sense

$$f(\theta + 2\pi) = f(\theta) + 2\pi.$$

EXERCISES

1. Define
$$\eta(f) = \inf_{\theta,\vartheta \in S^1, \theta \neq \vartheta} \frac{\|f(\theta) - f(\vartheta)\|}{|\theta - \vartheta|}, \quad \text{where} \quad f\colon S^1 \to S^1, \quad f \in C^\infty(S^1).$$
Here $\|\cdot\|$ means the Euclidean distance in $\mathbb{R}^2$ and $S^1$ considered as an embedded manifold to $\mathbb{R}^2$. Verify that $\eta$ is a continuous function of $f$. Conclude that since $\eta(f) > 0$ if and only if $f$ is a diffeomorphism, the set Diff $S^1$ is an open set in $C^\infty(S^1, S^1)$.
2. Let
$$V_0 = \{v \in \mathrm{Vect}(S^1) \mid \ |v| < \pi\},$$
with $\|\cdot\|$ being the absolute value in $\mathbb{R}$ and
$$U_0 = \{f \in C^\infty(S^1, S^1) \mid \ f(\theta) \neq -\theta, \ \forall\ z \in S^1\}$$
be neighborhoods of $0 \in \mathrm{Vect}(S^1)$ and the identity map id $\in C^\infty(S^1, S^1)$, respectively. We construct a map $\psi\colon V_0 \to U_0$ by the following. Define a map $\psi_v\colon S^1 \to S^1$ for $v \in V_0$ as a map that sends a point $\theta \in S^1$ to the end point of the arc of length $|v|$, which starts at $\theta$ with the initial velocity $v(\theta)$. Show that $\psi$ is the homeomorphism from $V_0 \ni v \to \psi_v \in U_0$.
3. Let $n \in \mathbb{N}$ be big enough, and $\varepsilon \in (0, \frac{1}{n})$. We think of $S^1 = \{\theta \in \mathbb{R} \bmod 2\pi\mathbb{Z}\}$. Consider the diffeomorphism
$$f(\theta) = \theta + \frac{\pi}{n} + \varepsilon \sin^2(n\theta).$$
Show that by choosing $n$ and $\varepsilon$ we can make $f(\theta)$ so close in $C^\infty$-topology to the identity map as we want. Show that $f$ has only one periodic orbit of period $2n$. (Hint: start from $\theta = 0$.) Show that starting from any other value $\theta \in (0, \pi/n)$ we get a non-periodic orbit. Conclude that $f$ cannot belong to

any one-parameter subgroup of Diff $S^1$. Indeed if it belonged to $\exp(v)$ for some $v \in \mathrm{Vect}(S^1)$, then it would be a rotation $f(\theta) = \theta + \frac{\pi}{n}$, since after $2n$ repetitions we get $\Big(f(0)\Big)^{2n} = 2\pi = f(0 + 2\pi)$. But we know that it is not true since all other points of $S^1$ do not belong to the same orbit of $f$. See also [81, 108, 116].

4. Calculate the commutator in $\mathrm{Vect}(S^1)$ for its basis $\cos(n\theta)$, $\sin(n\theta)$, $n = 0, 1, \ldots$.

### 7.1.1. Central extensions of Diff $S^1$ and $\mathrm{Vect}(S^1)$.

To introduce the central extensions of the Lie-Frechét group Diff $S^1$ and its Lie algebra $\mathrm{Vect}(S^1)$ we start from the linear object, i. e. from the Lie algebra.

**Definition 31.** *The central extension* $\widetilde{\mathfrak{g}}$ *of a Lie algebra* $\mathfrak{g}$ *by the Lie algebra* $\mathbb{R}$ *of real numbers is the set* $\mathfrak{g}\times\mathbb{R}$ *and new Lie brackets* $[(\xi, a), (\eta, b)]_{\widetilde{g}}$, $\xi, \eta \in \mathfrak{g}$, $a, b \in \mathbb{R}$, *satisfying the axioms of Definition* 51.

In this case $\mathbb{R}$ becomes the center of the extended Lie algebra. The simplest trivial example of a central extension is the direct product $\mathfrak{g} \times \mathbb{R}$ with the Lie brackets defined by

$$[(\xi, a), (\eta, b)]_{\widetilde{\mathfrak{g}}} := ([\xi, \eta]_{\mathfrak{g}}, ab - ba) = ([\xi, \eta]_{\mathfrak{g}}, 0).$$

We are interested in a non-trivial extension. In order to get it we need to find an invariant skew symmetric bi-linear form $\omega\colon \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$, such that the new Lie bracket $[(\xi, a), (\eta, b)]_{\widetilde{\mathfrak{g}}} := ([\xi, \eta]_{\mathfrak{g}}, \omega(\xi, \eta))$ in the extended Lie algebra $\widetilde{\mathfrak{g}}$ satisfies the axioms of Definition 51. It leads to the condition on $\omega$, which is called *cocycle condition*:

$$\omega([\xi, \eta], \zeta) + \omega([\eta, \zeta], \xi) + \omega([\zeta, \xi], \eta) = 0. \tag{7.2}$$

The form $\omega$ is called a *2-cocycle*, and the terminology comes from the cohomology theory. It was shown in [54] that there is an essentially unique non-trivial 2-cocycle for the Lie algebra $\mathrm{Vect}(S^1)$, which is called the *Gelfand-Fuchs Lie algebra cocycle.* It is given by the following form $\omega$.

$$\begin{aligned}\omega\big(u(\theta)\partial_\theta, v(\theta)\partial_\theta\big) &= \frac{1}{2\pi}\int_{S^1} u'(\theta)v''(\theta)\,d\theta = \frac{1}{2\pi}\int_{S^1} u'(\theta)dv' \\ &= \frac{1}{4\pi}\int_{S^1} \det\begin{pmatrix} u'(\theta) & v'(\theta) \\ u''(\theta) & v''(\theta)\end{pmatrix} d\theta.\end{aligned} \tag{7.3}$$

The central extension of the Lie algebra $\mathrm{Vect}(S^1)$ by $\mathbb{R}$ is called the Virasoro algebra, is denoted by $\mathfrak{vir}$ and it is unique up to an isomorphism. The name of the algebra is coming from the name of the Argentinian physicist Miguel Angel Virasoro, who invented the idea of central extension for $\mathrm{Vect}(S^1)$ in his physical calculations. In physics, actually, a more general 2-cocycle is used. Let us explain the mathematical background of this more general 2-cocycle.

**Definition 32.** *A 2-cocycle* $\omega\colon \mathfrak{g}\times\mathfrak{g}\to\mathbb{R}$ *is called a 2-coboundary if there exists a linear map* $\eta\colon \mathfrak{g}\to\mathbb{R}$ *such that* $\omega(\xi,\upsilon)=\eta([\xi,\upsilon])$ *for all* $\xi,\upsilon\in\mathfrak{g}$.

Let us assume that the central extension $\widetilde{\mathfrak{g}}$ of a Lie algebra $\mathfrak{g}$ is given by a 2-coboundary $\eta$, e. g. $(\xi,a)\in\widetilde{\mathfrak{g}}$ and $[(\xi,a),(\upsilon,b)]=([\xi,\upsilon],\eta([\xi,\upsilon]))$. Then the change of the variables $(\xi,a)\mapsto(\xi,a-\eta(\xi))$ leads to

$$([\xi,\upsilon],\eta([\xi,\upsilon]))\mapsto([\xi,\upsilon],0).$$

We again obtained the trivial extension. By this observation, in describing different central extensions one is interested only in the 2-cocycles modulo 2-coboundaries. Meanwhile, in physical applications the following general form of the 2-cocycle $\omega_{h,c}$ for some positive constants $h,c$ is used:

$$\omega_{h,c}\big(u(\theta)\partial_\theta,v(\theta)\partial_\theta\big)=\frac{1}{2\pi}\int_{S^1}\Big(\big(h-\frac{c}{12}\big)v'(\theta)-\frac{c}{12}v'''(\theta)\Big)u(\theta)\,d\theta. \tag{7.4}$$

The constant $c$ received the name *central charge* in physics and its value depends on the underlying physical theory. The cocycle $\omega$ from (7.3) is obtained, up to the normalization factor $\frac{c}{12}$, by setting $h=\frac{c}{12}$ and is often called the *classical Gelfand-Fuchs 2-cocycle*. The following question arises: is there a group, whose Lie algebra is $\mathfrak{vir}$? The answer is positive. The corresponding group is called the real Virasoro-Bott group and is denoted by Vir. The Virasoro-Bott group as a set is the direct product of $\operatorname{Diff}S^1$ and $\mathbb{R}$: $\mathrm{Vir}=\operatorname{Diff}S^1\times\mathbb{R}$. In this case the group multiplication law in Vir can be defined as follows

$$(f,a)(h,b)=(f\circ h,a+b+\lambda(f,h)),\quad f,h\in\operatorname{Diff}S^1,\quad a,b\in\mathbb{R}, \tag{7.5}$$

where $\lambda\colon \operatorname{Diff}S^1\times\operatorname{Diff}S^1\to\mathbb{R}$ is a smooth function that makes the multiplication law (7.5) associative. The associativity of (7.5) corresponds to the *group cocycle identity*:

$$\lambda(f\circ h,g)+\lambda(f,h)=\lambda(f,h\circ g)+\lambda(h,g). \tag{7.6}$$

The infinitesimal version of the Lie group 2-cocycle is the Lie algebra 2-cocycle. The following theorem gives the group 2-cocycle for the classical Gelfand-Fuchs 2-cocycle and was obtained for the first time in [16].

**Proposition 13.** *The map*

$$\begin{array}{rccc} B\colon & \operatorname{Diff}S^1\times\operatorname{Diff}S^1 & \to & S^1\\ & (f,h) & \mapsto & \frac{1}{4\pi}\int_{S^1}\log(f\circ h)'\,d\log h' \end{array}$$

*is a continuous 2-cocycle on the group* $\operatorname{Diff}S^1$. *Here* $f,h\in\operatorname{Diff}S^1$ *and* $f',h'$ *are their derivatives with respect to* $\theta\in S^1$.

*Proof.* For the proof we need to verify the group 2-cocycle condition (7.6) and then to check that the infinitesimal version coincides with the classical Gelfand-Fuchs 2-cocycle $\omega$. The details can be found in [81]. ☐

As in the case of the central extension of a Lie algebra, any central extension of a Lie group is defined up to a 2-coboundary.

**Definition 33.** *A smooth 2-cocycle* $\lambda\colon \mathbb{G}\times\mathbb{G}\to\mathbb{R}$ *is called a 2-coboundary (on* $\mathbb{G}$*) if there exists a smooth map* $F\colon\mathbb{G}\to\mathbb{R}$ *such that*

$$\lambda(f,h)=F(f)+F(h)-F(f\circ h).$$

Two group 2-cocycles define isomorphic extensions if they differ by a 2-coboundary. Let us construct a group 2-coboundary for the group Diff $S^1$. First, we write the general Lie algebra cocycle $\omega_{h,c}$ in the following form

$$\begin{aligned}\omega_{h,c}\big(u(\theta)\partial_\theta,v(\theta)\partial_\theta\big) &= \frac{1}{2\pi}\int_{S^1}\Big(\big(h-\frac{c}{12}\big)v'(\theta)-\frac{c}{12}v'''(\theta)\Big)u\,d\theta\\ &= \frac{1}{2\pi}\int_0^{2\pi}\big(\alpha uv'+\beta u'v''\big)\,d\theta=:\alpha a(u,v)+\beta b(u,v),\end{aligned}$$

where $\alpha=h-\frac{c}{12}$, $\beta=\frac{c}{12}$ and we used integration by parts of periodic functions $u$ and $v$ on $S^1$. Let us verify that $a(u,v)=\frac{1}{2\pi}\int_0^{2\pi}uv'\,d\theta$ is the Lie algebra 2-coboundary. Introduce the functional

$$\eta(u\partial_\theta)=\frac{1}{2\pi}\int_0^{2\pi}u(\theta)\,d\theta,$$

expressing the mean value of $u$ on $S^1$. Observe that

$$\begin{aligned}\eta([u,v])&=\frac{1}{2\pi}\int_0^{2\pi}u'(\theta)v(\theta)-v'(\theta)u(\theta)\,d\theta\\ &=-2\frac{1}{2\pi}\int_0^{2\pi}u(\theta)v'(\theta)\,d\theta=-2a(u,v)\end{aligned}$$

by integration by parts. Thus $a$ is a group 2-coboundary. The second part $b(u,v)=\frac{1}{2\pi}\int_0^{2\pi}u'v''\,d\theta$ is the Gelfand-Fuchs algebra 2-cocycle. The multiplication law (7.5) takes the form

$$(f,a)(h,b)=(f\circ h,a+b+\alpha A(f,h)+\beta B(f,h)),\quad f,h\in\text{Diff}\,S^1,\quad a,b\in\mathbb{R}.$$

Here $B(f,h)$ is the Bott 2-cocycle given by Proposition 13 and $A(f,h)$ is the group 2-coboundary satisfying Definition 33. To find the smooth function $F$ from Definition 33 we verify the following two properties.

1. Since the identity element on Vir has the form $(\text{id},0)$ we get $A(f,\text{id})=0\implies F(\text{id})=0$, $\text{id}\in\text{Diff}\,S^1$ by the multiplication law (7.5).
2. In order to obtain the inverse element to $(f,a)\in$ Vir in the form $(f^{-1},-a)$ we require $A(f,f^{-1})=0\implies F(f)+F(f^{-1})=0$.

The function $F(f)=\int_0^{2\pi}\big(f(\theta)-\theta\big)\,d\theta$ obviously satisfies the first property. Let us show that it also satisfies the second property.

Step 1. First we assume that $f(0)=0$. Then $f(2\pi)=2\pi$ and $f$ is strictly increasing on $[0,2\pi]$. Thus

$$\begin{aligned} F(f)+F(f^{-1}) &= \int_0^{2\pi}\big(f(\theta)+f^{-1}(\theta)-2\theta\big)\,d\theta \\ &= \int_0^{2\pi}\int_0^{f} d\theta dy+\int_0^{2\pi}\int_f^{2\pi} d\theta dy-4\pi^2=0. \end{aligned}$$

Step 2. Now we assume that $f$ is an arbitrary element of $\operatorname{Diff} S^1$. Define $\hat{f}(\theta)=f(\theta+f^{-1}(0))$. Then $\hat{f}(\theta)$ satisfies three properties:

a) $\hat{f}(0)=0$,
b) $\hat{f}^{-1}(\theta)=f^{-1}(\theta)-f^{-1}(0)$,
c) $\int_0^{2\pi}\hat{f}(\theta+s)\,d\theta=\int_0^{2\pi}\big(\hat{f}(\theta)+s\big)\,d\theta$ for any $s\in\mathbb{R}$.

Indeed, the properties a), b) are obvious and to show c) we observe that $\hat{f}(\theta)-\theta$ is a periodic function with the period $2\pi$ by $\hat{f}(\theta+2\pi k)=\hat{f}+2\pi k$, $k\in\mathbb{N}$. Then

$$\begin{aligned} \int_0^{2\pi}\hat{f}(\theta+s)\,d\theta &= \int_s^{2\pi+s}\hat{f}(\theta)-\theta\,d\theta+\int_s^{2\pi+s}\theta\,d\theta \\ &= \int_0^{2\pi}\hat{f}(\theta)-\theta\,d\theta+\frac{1}{2}\big((2\pi+s)^2-s^2\big) \\ &= \int_0^{2\pi}\hat{f}(\theta)\,d\theta+2\pi s=\int_0^{2\pi}\big(\hat{f}(\theta)+s\big)\,d\theta. \end{aligned}$$

We continue to prove the second property for an arbitrary $f\in\operatorname{Diff} S^1$ and deduce

$$\begin{aligned} F(f)+F(f^{-1}) &= \int_0^{2\pi}\Big(f(\theta)+f^{-1}(\theta)-2\theta\Big)\,d\theta \\ &= \int_0^{2\pi}\Big(\hat{f}\big(\theta-f^{-1}(0)\big)+\hat{f}^{-1}(\theta)+f^{-1}(0)-2\theta\Big)\,d\theta \\ &= \int_0^{2\pi}\Big(\hat{f}(\theta)-f^{-1}(0)+\hat{f}^{-1}(\theta)+f^{-1}(0)-2\theta\Big)\,d\theta \\ &= F(\hat{f})+F(\hat{f}^{-1})=0 \end{aligned}$$

by the Step 1 and a), b), c).

We conclude, that

$$A(f,h)=\int_0^{2\pi}\big(f(\theta)+h(\theta)-(f\circ h)(\theta)-\theta\big)\,d\theta. \tag{7.7}$$

The last step in defining group 2-coboundary $A$ is to verify that the infinitesimal version of $A$ from (7.7) coincides with the algebra 2-coboundary $a$. We will use the following.

**Proposition 14.** [81] *Let* $A\colon \mathbb{G}\times\mathbb{G}\to\mathbb{R}$ *be a group 2-cocycle defining a central extension by* $\mathbb{R}$ *of the group* $\mathbb{G}$*. Then the algebra 2-cocycle* $a\colon \mathfrak{g}\times\mathfrak{g}\to\mathbb{R}$ *defining the corresponding central extension of the Lie algebra* $\mathfrak{g}$ *of* $\mathbb{G}$ *is given by*

$$a(u,v)=\frac{d^2}{dtds}\Big|_{t=0,s=0}A(f_t,h_s)-\frac{d^2}{dtds}\Big|_{t=0,s=0}A(h_s,f_t),$$

*where* $f_t$ *and* $h_s$ *are smooth curves in* $\mathbb{G}$ *such that*

$$\frac{d}{dt}\Big|_{t=0}f_t=u,\qquad \frac{d}{ds}\Big|_{s=0}h_s=v.$$

Now, differentiating (7.7) we obtain

$$\begin{aligned}&\textstyle\int_0^{2\pi}\Big(-\frac{\partial}{\partial\theta}\frac{\partial}{\partial t}f_t\frac{\partial}{\partial s}h_s+\frac{\partial}{\partial\theta}\frac{\partial}{\partial s}h_s\frac{\partial}{\partial t}f_t\Big)\,d\theta\Big|_{t=0,s=0}\\&=\textstyle\int_0^{2\pi}\big(-v\partial_\theta u+u\partial_\theta v\big)\,d\theta=-2\pi\eta([u,v])=4\pi a(u,v).\end{aligned}$$

Normalizing the 2-coboundary (7.7), we finally deduce

$$A(f,h)=\frac{1}{4\pi}\int_0^{2\pi}\big(f(\theta)+h(\theta)-(f\circ h)(\theta)-\theta\big)\,d\theta.$$

**Example 9.** Here we present the Heisenberg algebra from the central extension point of view. Take the abelian Lie algebra $\mathfrak{g}=\mathbb{R}^2$ and an arbitrary skew symmetric bilinear form $\omega\colon\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}$. Since the Lie algebra $\mathfrak{g}$ is abelian, the 2-cocycle condition (7.2) is trivial and the form $\omega$ satisfies it. The resulting central extension of $\mathbb{R}^2$ by $\mathbb{R}$ is the set $\mathfrak{h}=\mathbb{R}^2\oplus\mathbb{R}$ endowed with the brackets

$$[h_1,h_2]=[(v_1,r_1),(v_2,r_2)]=\big(0,\omega(v_1,v_2)\big).$$

Note that the choice of any other skew-symmetric bilinear form leads to an isomorphic Lie algebra $\widetilde{\mathfrak{h}}$. The Lie algebra $\mathfrak{h}$ with a non-degenerate form $\omega$ is a representative of this isomorphism class and it is called the three-dimensional Heisenberg algebra.

The $n$-dimensional analogue can be obtained by taking $\mathfrak{g}=\mathbb{R}^{2n+1}$ and an arbitrary skew symmetric bilinear form $\omega\colon\mathbb{R}^{2n+1}\times\mathbb{R}^{2n+1}\to\mathbb{R}$.

We even can continue and present an infinite dimensional version of the Heisenberg algebra as a central extension of the space

$$\mathfrak{g}=\{f\in C(S^1,S^1)\ \mid\ \eta(f)=\frac{1}{2\pi}\int_{S^1}f(\theta)\,d\theta=0\}.$$

The space $\mathfrak{g}$ is considered as an abelian algebra. Since a function with vanishing mean value can be written as a Fourier series

$$f(\theta)=\sum_{n=1}^{\infty}x_n\cos(n\theta)+y_n\sin(n\theta),$$

it can be interpreted as a point in an infinite dimensional space. The 2-cocycle $\omega\colon\mathfrak{g}\times\mathfrak{g}\to\mathbb{R}$ is given by

$$\omega(f,g)=\int_{S^1}f'(\theta)g(\theta)\,d\theta. \tag{7.8}$$

Exercises

1. Prove the cocycle condition (7.2) for the Gelfand-Fuchs cocycle (7.3).
2. Calculate the Heisenberg group 2-cocycle that corresponds to the algebra 2-cocycle (7.8).

### 7.1.2. Complexification of $\mathrm{Vect}(S^1)$ and the Virasoro algebra.

The next step is to consider complexifications of $\mathrm{Vect}(S^1)$ and $\mathfrak{vir}$ and their relations to $\mathrm{Diff}\, S^1$ and Vir. The reader who is not familiar with the general construction of complexification can find all the necessary definitions in Appendix A, Subsection 8.4.

The complexification $\mathrm{Vect}(S^1) \otimes \mathbb{C}$ consists of smooth complex-valued vector fields $v(\theta)\partial_\theta$ defined on $S^1$, that can be identified with the space $C^\infty(S^1, \mathbb{C})$. The natural basis is the complex valued Fourier basis $e_k := -ie^{ik\theta}\partial_\theta$, $k \in \mathbb{Z}$, produced from the real Fourier basis. The commutation relations for basic vector fields on $\mathrm{Vect}(S^1) \otimes \mathbb{C}$ are

$$[e_m, e_n] = (n-m)e_{m+n}, \quad m, n \in \mathbb{Z}.$$

These relations are known under the name of Witt, and the Lie algebra whose basis satisfies the Witt relations is called the Witt algebra. So the complexification of the Lie algebra $\mathrm{Vect}(S^1)$ is the Witt algebra. Actually, the complex valued vector fields $e_k = -ie^{ik\theta}\partial_\theta$ can be extended to meromorphic vector fields $L_k = z^{k+1}\frac{d}{dz}$ on $\mathbb{C}\cup\{\infty\}$, that are holomorphic vector fields in $\mathbb{C}\setminus\{0\}$. The extended to the Riemann sphere $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ algebra of complex valued vector fields $\{z^{k+1}\frac{d}{dz}\}_{k=-\infty}^{\infty}$ is also called the Witt algebra.

The complexification of the Lie algebra $\mathrm{Vect}(S^1)$ does not correspond to any Lie group. The explanation can be the following. First, we observe that the Lie algebra $\mathrm{Vect}(S^1)$ contains Lie sub-algebras $\mathfrak{g}_k = \mathrm{span}\{\partial_\theta, \sin(k\theta)\partial_\theta, \cos(k\theta)\partial_\theta\}$, $k = 1, 2, \ldots$. All these sub-algebras are isomorphic to the Lie algebra $\mathfrak{sl}(2,\mathbb{R})$. The algebras $\mathfrak{g}_k$ can be integrated to subgroups $\mathbb{G}_k$ of $\mathrm{Diff}\, S^1$. The subgroup $\mathbb{G}_1$ is the projective special linear group $PSL(2,\mathbb{R})$ and $\mathbb{G}_2 = SL(2,\mathbb{R})$. All other groups $\mathbb{G}_k$ are $k$-fold coverings of $\mathbb{G}_1$. It is known that to the algebra $\mathfrak{sl}(2,\mathbb{C})$, which is the complexification of $\mathfrak{sl}(2,\mathbb{R})$, there correspond only 2 complex groups $\mathbb{G}_1^{\mathbb{C}} = PSL(2,\mathbb{C})$ and $\mathbb{G}_2^{\mathbb{C}} = SL(2,\mathbb{C})$. All other groups $\mathbb{G}_k$, $k > 2$, do not admit complexifications.

The second observation is that the real Lie algebra $\mathrm{Vect}(S^1)$ is contained in the complex Lie algebra $\mathrm{Vect}(S^1) \otimes \mathbb{C}$ and therefore the real Lie sub-algebras $\mathfrak{g}_k$ are contained in the corresponding complex Lie sub-algebras $\mathfrak{g}_k \otimes \mathbb{C}$. If the complex group $(\mathrm{Diff}\, S^1)^{\mathbb{C}}$ existed then it would contain the real group $\mathrm{Diff}\, S^1$ and the real subgroups $\mathbb{G}_k$ would belong to the complex subgroups $\mathbb{G}_k^{\mathbb{C}}$. It is known that $PSL(2,\mathbb{R}) = \mathbb{G}_1 \subset \mathbb{G}_1^{\mathbb{C}} = PSL(2,\mathbb{C})$, $SL(2,\mathbb{R}) = \mathbb{G}_2 \subset \mathbb{G}_2^{\mathbb{C}} = SL(2,\mathbb{C})$, and no other complex subgroups of $(\mathrm{Diff}\, S^1)^{\mathbb{C}}$ containing $\mathbb{G}_k$, $k > 2$. The rigorous proof can be found in [118].

Instead of a complex structure, the group Diff $S^1$ admits a left invariant CR structure according to [96], that is constructed as follows. Define a subalgebra

$$\mathfrak{h}^{(1,0)} = \Big\{ \sum_{n=1}^{\infty} a_n e^{in\theta} \partial_\theta \in \mathrm{Vect}(S^1) \oplus \mathbb{C}, \quad a_n \in \mathbb{C} \Big\} \tag{7.9}$$

of $\mathrm{Vect}(S^1) \oplus \mathbb{C}$. The set $\mathfrak{h}^{(1,0)}$ is just the set of all vector fields having vanishing mean value on the circle. The sum $\mathfrak{h}^{(1,0)} \oplus \overline{\mathfrak{h}^{(1,0)}}$ is of complex co-rank 1 in $\mathrm{Vect}(S^1) \oplus \mathbb{C}$, whence we obtain a left invariant CR structure on Diff $S^1$. In addition,

$$\Big[ \sum_{n=1}^{\infty} a_n e^{in\theta} \partial_\theta, \sum_{n=1}^{\infty} \bar{a}_n e^{-in\theta} \partial_\theta \Big] = -i \sum_{k,n=1}^{\infty} (k+n) a_n \bar{a}_k e^{i(n-k)\theta} \partial_\theta$$

which is not in $\mathfrak{h}^{(1,0)} \oplus \overline{\mathfrak{h}^{(1,0)}}$ unless all $a_n = 0$. Thus $(\mathrm{Diff}\, S^1, \mathfrak{h}^{(1,0)})$ is strongly pseudoconvex.

We now move to the complexification $\mathfrak{vir} \otimes \mathbb{C}$ of the central extension of $\mathrm{Vect}(S^1)$ that is the Virasoro algebra $\mathfrak{vir}$. As a vector space $\mathfrak{vir} \otimes \mathbb{C}$ is the complex vector space generated by $e_k, \mathfrak{c}$, where $e_k$ is the basis of $\mathrm{Vect}(S^1) \otimes \mathbb{C}$ and the generator $\mathfrak{c}$ is called the central element. To define the bracket on $\mathfrak{vir} \otimes \mathbb{C}$ we extend the real Lie algebra 2-cocycle

$$\omega_{\alpha\beta}(u,v) = \frac{1}{2\pi} \int_0^{2\pi} \big( \alpha u(\theta) v'(\theta) + \beta u'(\theta) v''(\theta) \big)\, d\theta \tag{7.10}$$

to the complex valued cocycle

$$\omega^{\mathbb{C}}_{\alpha\beta} \colon \mathrm{Vect}(S^1) \otimes \mathbb{C} \times \mathrm{Vect}(S^1) \otimes \mathbb{C} \to \mathbb{C}$$

by the standard procedure of extension of the integral to complex valued functions. Then the commutator becomes

$$[(v, \mu\mathfrak{c}), (u, \nu\mathfrak{c})] = \big( [v,u], \omega^{\mathbb{C}}_{\alpha\beta}(v,u)\mathfrak{c} \big), \qquad [(v,0),(0,\mathfrak{c})] = 0,$$

$v, u \in \mathrm{Vect}(S^1) \otimes \mathbb{C}$, $\mu, \nu \in \mathbb{C}$. The value of $\omega^{\mathbb{C}}_{\alpha\beta}$ on the Witt basis $e_k = -ie^{ik\theta}\partial_\theta$, $k \in \mathbb{Z}$ is given by

$$\omega^{\mathbb{C}}_{\alpha\beta}(-ie^{im\theta}\partial_\theta, -ie^{in\theta}\partial_\theta) = \begin{cases} -i(\alpha n + \beta n^3) & \text{if} \quad n+m=0, \\ 0 & \text{if} \quad n+m \neq 0. \end{cases}$$

The complexification $\mathfrak{vir} \otimes \mathbb{C}$ of $\mathfrak{vir}$ is also called the Virasoro algebra and in physics it is used more than the real Virasoro algebra.

Are there complex groups that correspond to $\mathfrak{vir} \otimes \mathbb{C}$? In the work [96] the author proved that the Virasoro-Bott group Vir admits a left invariant complex structure. It means that the complexified Virasoro algebra $\mathfrak{vir} \oplus \mathbb{C}$ admits the splitting

$$\mathfrak{vir} \oplus \mathbb{C} = \mathfrak{vir}^{(1,0)} \oplus \mathfrak{vir}^{(0,1)}$$

and the manifold $(\mathrm{Vir}, \mathfrak{vir}^{(1,0)})$ can be considered as a complex manifold. To prove that Vir is a complex group one has to verify that the multiplication and inversion become holomorphic maps, but we leave it to the reader. L. Lempert shows in his work [96] that there is actually a family of complex structures, that are defined as follows. Fix any purely imaginary complex number $i\varkappa$, $\varkappa \in \mathbb{R}$, and define subalgebras

$$\mathfrak{vir}^{(1,0)} = \Big\{ \big( \sum_{n=0}^{\infty} a_n e^{in\theta} \partial_\theta \, , \, i\varkappa a_0 \mathfrak{c} \big) \in \mathfrak{vir} \oplus \mathbb{C}, \quad \{a_n\}_{n=0}^{\infty} \in \mathbb{C}, \ \varkappa \in \mathbb{R} \Big\} \tag{7.11}$$

and $\mathfrak{vir}^{(0,1)} = \overline{\mathfrak{vir}^{(1,0)}}$. Since $\mathfrak{vir} \oplus \mathbb{C} = \mathfrak{vir}^{(1,0)} \oplus \mathfrak{vir}^{(0,1)}$, we obtain a family of left invariant complex structures on Vir parametrized by $\varkappa \in \mathbb{R}$.

**7.1.3. Homogeneous manifold** Diff $S^1/S^1$**.** Let us denote by $S^1$ the closed subgroup of Diff $S^1$ generated by rotations of the unit circle: $\tau(\theta) = \theta + b$, $b \in \mathbb{R} \mod 2\pi\mathbb{Z}$ if $\tau \in S^1$. Suppose that $S^1$ acts on Diff $S^1$ on the right

$$\begin{array}{cccc} \mu\colon & \operatorname{Diff} S^1 \times S^1 & \to & \operatorname{Diff} S^1 \\ & f.\tau & \mapsto & f \circ \tau = f(\tau). \end{array}$$

Then the right quotient Diff $S^1/S^1$ has a manifold structure. Since the group $S^1$ is not a normal subgroup, then the manifold Diff $S^1/S^1$ has no group structure. The CR structure $\mathfrak{h}^{(1,0)}$ of the group Diff $S^1$ presented at (7.9) is invariant under the right action of $S^1$. The action of $S^1$ is transversal to CR structure, therefore the quotient Diff $S^1/S^1$ inherits the complex structure from the CR structure.

The space Diff $S^1/S^1$ can be considered as a homogeneous space, where the group Diff $S^1$ acts on the left by composition

$$\begin{array}{cccc} \mu\colon & \operatorname{Diff} S^1 \times \operatorname{Diff} S^1/S^1 & \to & \operatorname{Diff} S^1/S^1 \\ & f.h & \mapsto & f \circ h = f(h). \end{array}$$

Indeed, if $h_1, h_2$ belong to the same equivalence class on Diff $S^1/S^1$, then $h_1^{-1} \circ h_2$ is a rotation. Therefore the images $f(h_1)$, $f(h_2)$ belong also to the same equivalence class on Diff $S^1/S^1$ because the composition $[f(h_1)]^{-1} \circ f(h_2) = h_1^{-1} \circ h_2$ is the rotation.

The manifold

$$(\operatorname{Diff} S^1/S^1, D^{(1,0)}) \quad \text{with} \quad D_f^{(1,0)} = d_{\mathrm{id}}\mu_f(\mathfrak{h}^{(1,0)}), \tag{7.12}$$

where $d\mu_f$ is the differential of the left action of the group Diff $S^1$ on Diff $S^1/S^1$, is a strongly pseudo-convex CR manifold due to the properties of $\mathfrak{h}^{(1,0)}$.

The subalgebra $\mathfrak{h}^{(1,0)}$ can be also obtained in the following way. To the Lie-Frechét group Diff $S^1$ there corresponds a Lie algebra Vect$(S^1)$. Let us denote the Lie algebra of the group of rotations $S^1$ by $\mathfrak{s}^1$. The space $\mathfrak{s}^1$ consists of constant vector fields on the circle. Then the tangent bundle of the quotient Diff $S^1/S^1$ has sections that are vector fields invariant under the right action of $S^1$, or, in other words, $T_f(\operatorname{Diff} S^1/S^1)$ is isomorphic to Vect$(S^1)/\mathfrak{s}^1$. This space is the space of vector fields with vanishing mean value on the circle. The almost complex structure on Vect$(S^1)/\mathfrak{s}^1$is given by the Hilbert transform $J$, which is easier to describe through the Fourier basis as

$$
\begin{aligned}
Jv(\theta) &= \sum_{n=1}^{\infty} -a_n \sin(n\theta) + b_n \cos(n\theta) \quad \text{for} \qquad (7.13)\\
v(\theta) &= \sum_{n=1}^{\infty} a_n \cos(n\theta) + b_n \sin(n\theta).
\end{aligned}
$$

Then

$$
\begin{aligned}
\big(\mathrm{Vect}(S^1)/\mathfrak{s}^1\big)^{(1,0)} &= \{v\partial_\theta - iJ(v\partial_\theta)\}\\
&= \Big\{\sum_{n=1}^{\infty} c_n e^{in\theta}\partial_\theta, \quad c_n = a_n - ib_n\Big\} = \mathfrak{h}^{(1,0)}.
\end{aligned}
$$

**7.1.4. The groups** Vir **and** Diff $S^1$ **as principal bundles.** The aim of this subsection is to explain the bundle structures of the groups Vir and Diff $S^1$ over the base space Diff $S^1/S^1$.

**Proposition 15.** *The following bundle structures exist.*

1. *The bundle* $\pi\colon$ Diff $S^1 \to$ Diff $S^1/S^1$ *is a principal* $U(1)$*-bundle*
2. *The bundle* $\Pi\colon$ Vir $\to$ Diff $S^1/S^1$ *is a trivial* $\mathbb{C}^*$*-bundle.*

Here $\mathbb{C}^*$ is the multiplicative group of complex numbers $\mathbb{C} \setminus \{0\}$. To prove the proposition one can show that the manifolds Vir, Diff $S^1$, and Diff $S^1/S^1$, considered with their complex and CR structures, are bi-holomorphically equivalent to some spaces of univalent functions, where the bundle structure is more transparent. We start from the definitions of these spaces.

Let $\mathrm{Hol}(B_{\mathbb{C}})$ denote the space of holomorphic functions in the unit disk $B_{\mathbb{C}} \subset \mathbb{C}$. The subspaces $\mathcal{A}_0$ and $\widetilde{\mathcal{A}}_0$ of $\mathrm{Hol}(B_{\mathbb{C}})$ are defined by

$$
\mathcal{A}_0 = \{f \in C^{\infty}(\overline{B}_{\mathbb{C}}) \ \mid\ f \in \mathrm{Hol}(B_{\mathbb{C}}),\ f(0) = 0\}, \qquad \widetilde{\mathcal{A}}_0 = \{f \in \mathcal{A}_0 \mid\ f'(0) = 0\},
$$

where $\overline{B}_{\mathbb{C}}$ is the closure of the unit disk $B_{\mathbb{C}}$. The classes $\mathcal{A}_0$ and $\widetilde{\mathcal{A}}_0$ are complex Frechét vector spaces, where the topology is defined by the semi-norms

$$
\|f\|_m = \sup\{|f^{(m)}(z)| \ \mid\ z \in \overline{B}_{\mathbb{C}}\}.
$$

The topology is equivalent to the uniform convergence of all derivatives in $\overline{B}_{\mathbb{C}}$. Notice that both $\mathcal{A}_0$ and $\widetilde{\mathcal{A}}_0$ can be considered as complex manifolds, where the real tangent space is naturally isomorphic to the holomorphic part of the splitting under the induced almost complex structure from $\mathbb{C}^{\mathbb{N}}$. Then we define

$$
\mathcal{F} = \{f \in \mathcal{A}_0 \ \mid\ f \ \text{ is univalent in } \ B_{\mathbb{C}} \ \text{ and injective on the boundary } \ \partial B_{\mathbb{C}}\}.
$$

Geometrically, the class $\mathcal{F}$ defines all differentiable embeddings of the closed disk $\overline{B}_{\mathbb{C}}$ to $\mathbb{C}$ and analytically it is represented by functions $f = cz(1 + \sum_{n=1}^{\infty} c_n z^n)$, $c, c_n \in \mathbb{C}$. As a subset of $\mathcal{A}_0$ the space of univalent functions $\mathcal{F}$ forms an open subset inheriting the Frechét topology of complex vector space $\mathcal{A}_0$. Next we consider the class

$$
\mathcal{F}_1 = \{f \in \mathcal{F} \ \mid\ |f'(0)| = 1\},
$$

whose elements can be written as $f = e^{i\phi}z(1+\sum_{n=1}^{\infty} c_n z^n)$, $\phi \in \mathbb{R} \mod 2\pi\mathbb{Z}$. The set $\mathcal{F}_1$ is a pseudo- convex surface of real codimension 1 in the complex open set $\mathcal{F} \subset \mathcal{A}_0$. The last class of functions is

$$\mathcal{F}_0 = \{f \in \mathcal{F} \ \mid \ f'(0) = 1\}.$$

The elements of this class have the form $f = z(1 + \sum_{n=1}^{\infty} c_n z^n)$. It is obvious that $\mathcal{F}_0$ can be considered both as the quotient $\mathcal{F}_1/S^1$ and as the quotient $\mathcal{F}/\mathbb{C}^*$, $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$. In the latter case, $\mathcal{F}$ is the holomorphic trivial principal $\mathbb{C}^*$-bundle over the base space $\mathcal{F}_0$ (since the projection is just dividing by a non-zero complex number). The topological structure of the circle bundle $\mathcal{F}_1$ over the base space $\mathcal{F}_0$ is more complicated.

Since the set $\mathcal{F}_0$ can be also considered as an open subset of the affine space $v + \widetilde{\mathcal{A}}_0$, where $v(z) = z$, the tangent space $T_f\mathcal{F}_0$ inherits the natural complex structure of the complex vector space $\widetilde{\mathcal{A}}_0$ [4]. The real tangent space $T_f\mathcal{F}_0$ with the induced almost complex structure from $\widetilde{\mathcal{A}}_0$ is isomorphic to the complex vector space $T_f^{(1,0)}\mathcal{F}_0$ of the splitting $T\mathcal{F}_0 \otimes \mathbb{C} = T^{(1,0)}\mathcal{F}_0 \oplus T^{(0,1)}\mathcal{F}_0$. Moreover, the affine coordinates can be introduced so that to every $f \in \mathcal{F}_0$, written in the form $f(z) = z(1 + \sum_{n=1}^{\infty} c_n z^n)$ there will correspond the sequence $\{c_n\}_{n=1}^{\infty}$.

**Theorem 7.1.** *The following statements are true.*

1. *The Virasoro-Bott group $Vir$ with the left invariant complex structure*

$$(\mathrm{Vir}, \mathfrak{vir}^{(1,0)})$$

*defined by* (7.11) *is bi-holomorphic to* $\mathcal{F}$ [96].
2. *The group* Diff $S^1$ *with its left invariant CR structure* (Diff $S^1, \mathfrak{h}^{(1,0)}$) (7.9) *is isomorphic to the strongly convex hypersurface* $\mathcal{F}_1 \subset \mathcal{F}$ [96].
3. *The homogeneous space* Diff $S^1/S^1$ *with its complex structure*

$$(\mathrm{Diff}\, S^1/S^1, D^{(1,0)})$$

*introduced by* (7.12) *is bi-holomorphic to* $\mathcal{F}_0$ [96, 84].

It can be shown that Diff $S^1/S^1$ admits not only a complex but even a Kählerian structure [4, 84]. Proposition 15 follows from Theorem 7.1 and the known bundles in space of holomorphic functions: principal $\mathbb{C}^*$-bundle $\mathcal{F} \to \mathcal{F}_0$ and the circle bundle $\mathcal{F}_1 \to \mathcal{F}_0$. Recall that the group $\mathbb{C}^*$ is the multiplicative group of complex numbers $\mathbb{C} \setminus \{0\}$. The bundle maps of Theorem 7.1 are expressed in the following diagram.

$$\begin{tikzcd}
C^{\star} \arrow[d, "\mathrm{pr}_{\mathrm{by}}\,\mathbb{R}"'] & \mathcal{F} \arrow[l, hook'] \arrow[r, leftrightarrow, "F"] \arrow[d, "\mathrm{pr}_{\mathrm{by}}\,\mathbb{R}"'] & \mathrm{Vir} \arrow[r, hook] \arrow[d, "\mathrm{pr}_{\mathrm{by}}\,\mathbb{R}"] & C^{\star} \arrow[d, "\mathrm{pr}_{\mathrm{by}}\,\mathbb{R}"] \\
S^1 & \mathcal{F}_1 \arrow[l, hook'] \arrow[r, leftrightarrow, "F_1"] \arrow[d, "\mathrm{pr}_{\mathcal{F}_0}"'] & \mathrm{Diff}\, S^1 \arrow[r, hook] \arrow[d, "\mathrm{pr}_{\mathrm{Diff}\, S^1/S^1}"] & S^1 \\
 & \mathcal{F}_0 \arrow[r, leftrightarrow, "F_0"] & \mathrm{Diff}\, S^1/S^1. &
\end{tikzcd}$$

Here $F, F_0$ are corresponding bi-holomorphic maps from Theorem 7.1 and $F_1$ gives the isomorphism of CR structures. The left and right hand side extremes represent the typical fibers and the central part shows the projections of total spaces to the base spaces $\mathcal{F}_0$ and $\operatorname{Diff} S^1/S^1$.

Now we briefly describe the bijective map between $\mathcal{F}_0$ and $\operatorname{Diff} S^1/S^1$. Let $\overline{B}^c_{\mathbb{C}}$ be the complement to the closure $\overline{B}_{\mathbb{C}}$ of the unit disk $B_{\mathbb{C}}$. For any $f \in \mathcal{F}_0$, we define a *matching function* $g\colon \overline{B}^c_{\mathbb{C}} \to \mathbb{C}$, such that the image $g(\overline{B}^c_{\mathbb{C}})$ coincides with the complement to closure $\overline{f(B_{\mathbb{C}})}$. Assume also that the map $g$ satisfies the normalization $g(\infty) = \infty$. Note that such $g$ exists by the Riemann mapping theorem. Since both sets $B_{\mathbb{C}}$ and $\overline{B}^c_{\mathbb{C}}$ have a common boundary $S^1$ and the functions $f$ and $g$ have smooth extensions through $S^1$, the images $g(S^1)$ and $f(S^1)$ are defined uniquely and represent the same smooth contour in $\mathbb{C}$. If $g$ and $\widetilde{g}$ are two matching functions to $f$, then they are related by a rotation

$$\widetilde{g}(\zeta) = g(\zeta w), \quad \zeta \in \overline{B}^c_{\mathbb{C}}, \quad |w| = 1.$$

Thus for an arbitrary matching function $g$ to $f \in \mathcal{F}_0$ the diffeomorphism $h \in \operatorname{Diff} S^1$, given by

$$e^{ih(\theta)} = (f^{-1} \circ g)(e^{i\theta}), \tag{7.14}$$

is uniquely defined by $f$ up to the right superposition with a rotation. The relation

$$\operatorname{Diff} S^1/S^1 \ni [h] \quad \leftrightarrow \quad f \in \mathcal{F}_0$$

given by (7.14) defines a holomorphic bijection $\operatorname{Diff} S^1/S^1 \cong \mathcal{F}_0$. The composition $f^{-1} \circ g$ is often called a *conformal welding*.

The left action of $\operatorname{Diff} S^1$ on $\operatorname{Diff} S^1/S^1$ is transferred to the left action over $\mathcal{F}_0$. It was shown in [4] that the action $\mu_f\colon \mathcal{F}_0 \to \mathcal{F}_0$ is a holomorphic map for any fixed $f \in \operatorname{Diff} S^1$. The infinitesimal generator of this action $\sigma_f\colon \operatorname{Vect}(S^1) \to T_f\mathcal{F}_0$ is given by the variational formula of A. C. Schaeffer and D. C. Spencer [124, page 32]

$$\sigma_f[u\partial_\theta](z) := \frac{f^2(z)}{2\pi} \int\limits_{S^1} \left(\frac{wf'(w)}{f(w)}\right)^2 \frac{u(w)\,dw}{w(f(w)-f(z))} \in T_f\mathcal{F}_0$$

defined for $f \in \mathcal{F}_0$, $u\partial_\theta \in \operatorname{Vect}(S^1)$. It extends by linearity to a map

$$L[f,\cdot]\colon\ \operatorname{Vect}(S^1) \otimes \mathbb{C} \to T_f\mathcal{F}_0 \otimes \mathbb{C} = T_f^{(1,0)}\mathcal{F}_0 \oplus T_f^{(0,1)}\mathcal{F}_0.$$

The variation $L[f,\cdot]$ defines also an isomorphism of vector spaces $\mathfrak{h}^{(1,0)} \leftrightarrow T_f^{(1,0)}\mathcal{F}_0$, which is given explicitly by

$$L_k[f] = L[f,e_k](z) = z^{k+1}f'(z), \tag{7.15}$$
$$e_k = -ie^{ik\theta}\partial_\theta \in \mathfrak{h}^{(1,0)}, \quad L_k[f] \in T_f^{(1,0)}\mathcal{F}_0,\ k = 1,2,\dots$$

by making use of the residue calculus, see e.g.; [4, 83]. Taking the antiholomorphic part of the basis $e_{-k} = ie^{-ik(\theta)}\partial_\theta$, $k = 1,2,\dots$, we obtain expressions for $L_{-k}[f] \in$

$T^{(1,0)}_f\mathcal{F}_0$ which are rather difficult. The first two of them are

$$L_{-1}[f](z) = f'(z) - 2c_1 f(z) - 1,$$
$$L_{-2}[f](z) = \frac{f'(z)}{z} - \frac{1}{f(z)} - 3c_1 + (c_1^2 - 4c_2)f(z),$$

and others can be obtained by the Witt commutation relations [4, 84]

$$[L_k, L_n] = (n-k)L_{k+n}, \quad k, n \in \mathbb{Z}. \tag{7.16}$$

Here we use the affine coordinates $f = z(1 + \sum_{n=1}^{\infty} c_n z^n) \ \leftrightarrow\ (c_1, c_2, \dots)$. The vector fields $L_k$, $k \in \mathbb{Z}$, are right invariant since they are produced by infinitesimal generator of the left action of the diffeomorphism group on the space $\mathcal{F}_0$. The constant vector $u_0 = -i$ is mapped to $L_0[f](z) = zf'(z) - f(z)$. The right invariant vector fields $L_k$, $k \in \mathbb{Z}$ were obtained in [84] and received the name of Kirillov's vector fields, see also [4]. We have

$$T^{(1,0)}_{\mathrm{id}}\mathcal{F}_0 = \operatorname{span}\{L_0[id], L_1[id], L_2[id], \dots\} = \operatorname{span}\{z^2, z^3, \dots\}.$$

**7.1.5. KdV and Virasoro-Bott group.** Let us impose an $L^2$-inner product $(\cdot\,,\cdot)_{L^2}$ on the algebra $\mathfrak{vir}$ by

$$\big((u(\theta)\partial_\theta, a), (v(\theta)\partial_\theta, b)\big)_{L^2} = \frac{1}{2\pi}\int_{S^1} u(\theta)v(\theta)\,d\theta + ab. \tag{7.17}$$

Then by right translations we define a right invariant $L^2$-metric on the Virasoro-Bott group. We are interested in finding a geodesic equation on the group Vir with respect to the $L^2$-metric. First we present the Hamiltonian equation on Lie groups. In this case it can be rewritten as an equation on its dual Lie algebra $\mathfrak{g}^*$. We start from the Poisson structure on the group naturally defined by a Lie algebra structure, and then present the Hamiltonian equation on $\mathfrak{g}^*$ corresponding to this Poisson structure.

**Definition 34.** *The natural Lie-Poisson (or Kirillov-Kostant-Poisson) structure* $\{\cdot\,,\cdot\}$ *defined on the dual Lie algebra* $\mathfrak{g}^*$ *is*

$$\begin{array}{rccc} \{\cdot\,,\cdot\}\colon & C^\infty(\mathfrak{g}^*)\times C^\infty(\mathfrak{g}^*) & \to & C^\infty(\mathfrak{g}^*) \\ & (f(w), g(w)) & \mapsto & \langle [d_w f, d_w g], w\rangle. \end{array}$$

Here, as usual, $\langle\cdot\,,\cdot\rangle$ is the pairing between $\mathfrak{g}$ and $\mathfrak{g}^*$. The functions $f, g$ are from $C^\infty(\mathfrak{g}^*)$, $w \in \mathfrak{g}^*$ and $d_w f, d_w g \in T^*_w(\mathfrak{g}^*)$. We identify elements of $T^*_w(\mathfrak{g}^*)$ with $\mathfrak{g}$ and think of $df, dg$ as elements of the Lie algebra $\mathfrak{g}$.

Now if $f \in C^\infty(\mathfrak{g}^*)$, then the Hamiltonian equation takes the form

$$\begin{aligned} \frac{df(w)}{dt} &= \{H, f\}(w) = \langle [d_w H, d_w f], w\rangle = \langle \mathrm{ad}_{d_w H}(d_w f), w\rangle \\ &= \langle d_w f, \mathrm{ad}^*_{d_w H}(w)\rangle. \end{aligned}$$

Since the left hand side can be written as $\langle d_w f, \dot w\rangle$, we get that the Hamiltonian equation on $\mathfrak{g}^*$ takes the form

$$\dot w(t) = \mathrm{ad}^*_{d_{w(t)}H}(w(t)). \tag{7.18}$$

Let $A\colon \mathfrak{g} \to \mathfrak{g}^*$ be any invertible self adjoint operator, e. g. $\langle A\xi, w\rangle = \langle \xi, Aw\rangle$, where $\langle \cdot\,,\cdot\rangle$ is the pairing between $\mathfrak{g}$ and $\mathfrak{g}^*$. Such operators are often called "inertia operator". Define the Hamiltonian function $H\colon \mathfrak{g}^* \to \mathbb{R}$ by

$$H(w) := \frac{1}{2}\langle w, A^{-1}w\rangle, \quad w \in \mathfrak{g}^*.$$

Then $d_w H(w) = A^{-1}w$, and the Hamiltonian equation takes the form

$$\dot{w}(t) = \mathrm{ad}^*_{A^{-1}w(t)}(w(t)). \tag{7.19}$$

Let us apply this calculus to the Virasoro-Bott group. As we know, the Lie algebra $\mathfrak{vir}$ consists of vector fields $(v(\theta)\partial_\theta, a)$. The dual space $\mathfrak{vir}^*$ for infinite dimensional space $\mathfrak{vir}$ is too large, therefore one usually consideres its "smooth part" in the following sense: for every non-zero element $v\partial_\theta \in \mathfrak{vir}$ there is an element $w \in \mathfrak{vir}^*$ such that $\langle v\partial_\theta, w\rangle \neq 0$ and the converse is also true. The dual space $\mathfrak{vir}^*$ can be identified with so called smooth quadratic differentials $(u(\theta)(d\theta)^2, a)$, $u \in C^\infty(S^1, \mathbb{R})$, see [82]. The pairing is defined by

$$\Big\langle \big(u(\theta)(d\theta)^2, a\big), \big(v(\theta)\partial_\theta, b\big)\Big\rangle = \frac{1}{2\pi}\int_{S^1} v(\theta)u(\theta)\,d\theta + ab. \tag{7.20}$$

The co-adjoint action of the Lie algebra $\mathfrak{vir}$ on its dual $\mathfrak{vir}^*$ is the following

$$\mathrm{ad}^*_{\left(v(\theta)\partial_\theta, b\right)}\big(u(\theta)(d\theta)^2, a\big) = \big((-2v'u - vu' - av''')(d\theta)^2, 0\big). \tag{7.21}$$

Generally, the presence of any inner product $(\cdot\,,\cdot)$ on a Lie algebra $\mathfrak{g}$ allows to construct the inertia operator $A$ by $(u,v) = \langle A(u), v\rangle$ for all $u, v \in \mathfrak{g}$. It is analogous to the situation when a metric on a Riemannian manifold $M$ produces the identification $T_qM$ with its dual $T^*_qM$, $q \in M$. The $L_2$ product (7.17) and the pairing (7.20) define the following inertia operator

$$\begin{array}{rccc} A\colon & \mathfrak{vir} & \to & \mathfrak{vir}^* \\ & (u(\theta)\partial_\theta, a) & \mapsto & (u(\theta)(d\theta)^2, a). \end{array}$$

The Hamiltonian function defined by the product (7.17) and the pairing (7.20) is

$$H\big((u(\theta)(d\theta)^2, a)\big) = \frac{1}{2}\int_{S^1} u^2(\theta)\,d\theta + a^2$$

and $d_{\left(u(\theta)(d\theta)^2, a\right)}H = (u(\theta)\partial_\theta, a)$. Then substituting $\big(u(\theta)(d\theta)^2, a\big)$ for $w$ in (7.19) we get

$$\frac{d}{dt}\big(u(\theta)(d\theta)^2, a\big) = \big((-3uu' - au''')(d\theta)^2, 0\big).$$

The last equation is reduced to the system

$$\begin{array}{rcl} \dot{u} & = & -3uu' - au''', \\ \dot{a} & = & 0. \end{array} \tag{7.22}$$

The first equation is the Karteweg-de Vries (KdV) non-linear evolution equation that describes traveling waves in a shallow canal. The second equation is just saying that the parameter $a$ is a real constant.

Remark that the Euler equation for the $L^2$ metric on the group $\operatorname{Diff} S^1$ is called the Hopf or inviscid Burgers equation. We will obtained it in Subsection 7.2.

EXERCISES

1. Prove the formula for the co-adjoint action (7.21) using the pairing (7.20) and the definition $\langle \operatorname{ad}^*_\xi \omega, \eta\rangle = -\langle \omega, \operatorname{ad}_\xi \eta\rangle$ for $\xi, \eta$ from a Lie algebra and $\omega$ from the dual to the Lie algebra.

OTHER INTERESTING EQUATIONS. On the groups $\operatorname{Diff} S^1$ and Vir more metrics can be defined. Let us describe them and write the corresponding geodesic equations.

On the Virasoro algebra Vir and on $\operatorname{Vect}(S^1)$ the following weighted family of metrics $(\cdot\,,\cdot)_{H^1_{\alpha,\beta}}$ can be defined

$$\big((v\partial_\theta, a), (u\partial_\theta, b)\big)_{H^1_{\alpha,\beta}} = \int_{S^1} \big(\alpha v u + \beta v' u'\big) d\theta + ab. \tag{7.23}$$

**Theorem 7.2.** [80] *The Euler equations for the right invariant metric* $(\cdot\,,\cdot)_{H^1_{\alpha,\beta}}$, $\alpha \neq 0$ *on the Virasoro-Bott group are given by the following system*

$$\begin{aligned} \alpha(\dot u + 3uu') - \beta\big(\dot u'' + 2u'u'' + uu''' + au'''\big) &= 0 \\ \dot a &= 0, \end{aligned} \tag{7.24}$$

*for* $(u(\theta, t)\partial_\theta, a(t)) \in \mathrm{Vir}$ *for each* $t \in I$.

For $\alpha = 1$, $\beta = 0$ equation (7.24) is the KdV equation (7.22). For $\alpha = \beta = 1$ one recovers *Camassa-Holm equation.* For $\alpha = 0$, $\beta = 1$ equation (7.24) becomes the *Hunter-Saxton equation.* Note that in the case $\alpha = 0$ the metric $(\cdot\,,\cdot)_{H^1_{\alpha,\beta}}$ becomes the homogeneous degenerate $(\cdot\,,\cdot)_{\dot H^1}$ metric and therefore to define the Euler equation one has to consider the homogeneous space $\operatorname{Diff} S^1/S^1$ and define the geodesic flow on it (for details see [80]).

CONTOUR DYNAMICS AND VIRASORO ALGEBRA. The relations between contour dynamics, stochastic evolution equations, conformal field theory and the groups Vir and $\operatorname{Diff} S^1$ were described in [11, 12, 46, 47, 99].

### 7.2. Sub-Riemannian geodesics on $\operatorname{Diff} S^1$.

In this section we present some results for sub-Riemannian geodesics on the groups $\operatorname{Diff} S^1$ and Vir. First we describe the horizontal sub-bundles and metrics on them. Then we present some formulas for normal geodesics and discuss the controllability of these groups with respect to the chosen horizontal sub-bundles.

**7.2.1. Horizontal sub-bundles.** Recall that the linear map $\eta\colon \operatorname{Vect}(S^1) \to \mathbb{R}$ given by

$$\eta(u\partial_\theta) = \frac{1}{2\pi}\int_0^{2\pi} u(\theta)\, d\theta, \tag{7.25}$$

associates to each vector field from $\operatorname{Vect}(S^1)$ its mean value on the circle. The kernel of $\eta$, consisting of all vector fields with zero mean value, is isomorphic to

$\mathrm{Vect}(S^1)/\mathfrak{s}^1$, where $\mathfrak{s}^1$ as before denotes the subalgebra of $\mathrm{Vect}(S^1)$ of constant vector fields, corresponding to the abelian group of rotations $S^1$. We use the notation $\mathrm{Vect}_0(S^1) = \mathrm{Vect}(S^1)/\mathfrak{s}^1$. Then $\mathrm{Vect}(S^1) = \mathrm{Vect}_0(S^1) \oplus \mathfrak{s}^1$.

Define a horizontal sub-bundle $\mathcal{H}$ of $T\,\mathrm{Diff}\,S^1$ by left translations of $\mathrm{Vect}_0(S^1)$. A horizontal sub-bundle $\mathcal{E}$ of $T\,\mathrm{Vir}$ is left translations of $\big(\mathrm{Vect}_0(S^1), 0\big)$ on Vir. Then the complement of $\big(\mathrm{Vect}_0(S^1), 0\big)$ in $\mathfrak{vir}$ is given by

$$\widehat{\mathfrak{s}}^1 = \{(a_0\partial_\theta, a) \in \mathfrak{vir} \,:\, a_0, a \in \mathbb{R}\}$$

and we have $\mathfrak{vir} = \big(\mathrm{Vect}_0(S^1), 0\big) \oplus \widehat{\mathfrak{s}}^1$. The algebra $\widehat{\mathfrak{s}}^1$ is an abelian sub-algebra of $\mathfrak{vir}$ corresponding to the abelian sub-group

$$\widehat{S}^1 = \{(\theta \mapsto \theta + b_0, b) \in \mathrm{Vir} \,:\, b_0, b \in \mathbb{R}\}.$$

**Proposition 16.** *The sub-bundle $\mathcal{H}$ of $T\,\mathrm{Diff}\,S^1$ is invariant under the action of rotations $S^1$ and the sub-bundle $\mathcal{E}$ of $T\,\mathrm{Vir}$ is invariant under the action of $\widehat{S}^1$.*

*Proof.* If $\rho : \theta \to \theta + b_0$ is a rotation, then

$$\mathrm{Ad}_\rho(u)(\theta) = \frac{d}{d\epsilon}\Big|_{\epsilon=0}\big(b_0 + \exp(\epsilon u(\theta - b_0)\big) = u(\theta - b_0).$$

Therefore, $\eta(\mathrm{Ad}_\rho(u)) = \eta(u)$, which means that $\mathcal{H}$ is invariant under the action of $S^1$. By similar arguments $\mathcal{E}$ is invariant under $\widehat{S}^1$. ☐

**7.2.2. Sub-Riemannian metrics and normal geodesics.** Let us describe left-invariant metrics on $\mathcal{H}$ and $\mathcal{E}$. We start with $\mathcal{H}$. Let $(\cdot\,,\cdot)^{1,0}$ denote the standard $L^2$ inner product on $\mathrm{Vect}(S^1)$

$$\big(u\partial_\theta, v\partial_\theta\big)^{1,0} = \frac{1}{2\pi}\int_0^{2\pi} u(\theta)v(\theta)d\theta.$$

Let $\mathbf{g}^{1,0}$ be the Riemannian metric obtained by left translation of $(\cdot\,,\cdot)^{1,0}$, and let $\mathbf{h}^{1,0}$ be its restriction to $\mathcal{H}$.

Before we present the equations for sub-Riemannian normal geodesics with respect to the metric $\mathbf{h}^{1,0}$, we formulate a general result that can be found in [65, 66] and that defines the geodesic equation and the exact form of normal geodesics under some invariance conditions. We mention also *regular* infinite dimensional Lie groups, see [108, 116]. For the first reading the reader can pay a small attention to this, since up to now all known Lie groups are regular groups. Nevertheless, this condition ensures, particularly, the existence and smoothness of the group exponential map, that may be not bijective. The assumption about the existence of the map $\mathrm{ad}_\xi^\top$ is also nontrivial in the infinite dimensional case.

**Theorem 7.3.** [65] *Let $\mathbb{G}$ be an infinite-dimensional regular Lie group and $\mathbb{K}$ be its connected subgroup. Denote by $\mathfrak{g}$ and $\mathfrak{k}$ their respective Lie algebras. Let $(\cdot\,,\cdot)$ be an inner product in $\mathfrak{g}$ for which $\mathrm{ad}_\xi^\top$ exists for any $\xi \in \mathfrak{g}$. Assume that $\mathfrak{h} = \mathfrak{k}^\perp$ and $\mathfrak{g} = \mathfrak{h} \oplus_\perp \mathfrak{k}$. Define the horizontal distribution $\mathcal{H}$ by left translations of $\mathfrak{h}$. Let $\mathbf{g}$ be a Riemannian metric on $\mathbb{G}$ obtained by left translation of $(\cdot\,,\cdot)$ and $\mathbf{h} = \mathbf{g}|_{\mathcal{H}}$. The following statements hold.*

(a) *If* $(\cdot\,,\cdot)$ *is* $\mathrm{ad}(\mathfrak{k})$ *invariant and if* $\gamma_R\colon [0,1]\to\mathbb{G}$ *is a Riemannian geodesic with respect to* $\mathbf{g}$*, then*

$$\lambda(t)=\mathrm{pr}_{\mathfrak{k}}\,\kappa^{\ell}(\dot\gamma_R(t)),\qquad t\in[0,1]$$

*is constant. Here* $\mathrm{pr}_{\mathfrak{k}}\colon\ \mathfrak{g}\to\mathfrak{k}$ *is the orthogonal projection with respect to* $(\cdot\,,\cdot)$ *and* $\kappa^{\ell}(\dot\gamma(t))=d_{\gamma^{-1}(t)}l(\dot\gamma(t))$ *is the left logarithmic derivative.*

(b) *If* $(\cdot\,,\cdot)$ *is* $\mathrm{Ad}(\mathbb{K})$ *invariant and if* $\gamma_{sR}\colon[0,1]\to\mathbb{G}$ *is a sub-Riemannian geodesic with respect to* $\mathbf{h}$*, then* $\gamma_{sR}$ *is a normal geodesic, if and only if, it is of the form*

$$\gamma_{sR}(t)=\gamma_R(t)\cdot\exp_{\mathbb{G}}(-\lambda t),\qquad \lambda=\mathrm{pr}_{\mathfrak{k}}\,\kappa^{\ell}(\dot\gamma_R),\quad t\in[0,1],\tag{7.26}$$

*where* $\gamma_R\colon[0,1]\to\mathbb{G}$ *is a Riemannian geodesic with respect to* $\mathbf{g}$.

(c) *The left logarithmic derivative* $u_{sR}(t)=\kappa^{\ell}(\dot\gamma_{sR}(t))$ *of the curve* $\gamma_{sR}\colon[0,1]\to\mathbb{G}$ *satisfies the equation* $\dot u_{sR}(t)=\mathrm{ad}^{\top}_{u_{sR}(t)}(u_{sR}(t)+\lambda)$ *with a constant* $\lambda$ *and* $t\in[0,1]$.

To write the equation of an $\mathcal{H}$-horizontal normal geodesic on Diff $S^1$ we have to check the conditions of Theorem 7.3 for $\mathbb{G}=\mathrm{Diff}\,S^1$ and $\mathbb{K}=S^1$. We start from the calculation of the adjoint operator $\mathrm{ad}^{\top}_{u\partial_\theta}$ with respect to the inner product $(\cdot\,,\cdot)^{1,0}$ on $\mathrm{Vect}(S^1)$, that is defined by $(\mathrm{ad}^{\top}_{u\partial_\theta}v\partial_\theta,w\partial_\theta)^{1,0}=(v\partial_\theta,\mathrm{ad}_{u\partial_\theta}w\partial_\theta)^{1,0}=(v\partial_\theta,[u\partial_\theta,w\partial_\theta])^{1,0}$. We drop the symbol $\partial_\theta$ to simplify the notation. We calculate

$$\begin{aligned}(\mathrm{ad}^{\top}_u v,w)^{1,0}&=(v,-uw'+u'w)^{1,0}=\frac{1}{2\pi}\int_0^{2\pi}\big(-uvw'+u'vw\big)\,d\theta\\&=\frac{1}{2\pi}\int_0^{2\pi}\big(uv'+2u'v\big)w\,d\theta=(uv'+2u'v,w)^{1,0}\end{aligned}$$

using integration by parts. Remember that all the functions $u,v,w$ are periodic on $[0,2\pi]$, and therefore the term outside of the integral vanishes. We conclude

$$\mathrm{ad}^{\top}_u v=uv'+2u'v,\qquad u\partial_\theta,v\partial_\theta\in\mathrm{Vect}(S^1)\tag{7.27}$$

**Proposition 17.** *The inner product* $(\cdot\,,\cdot)^{1,0}$ *is* $\mathrm{Ad}(S^1)$*- and* $\mathrm{ad}(\mathfrak{s}^1)$*- invariant.*

*Proof.* As in general, the invariance with respect to the adjoin action of the Lie group implies the invariance with respect to the adjoin action of the Lie algebra, the $\mathrm{Ad}(S^1)$- invariance implies $\mathrm{ad}(\mathfrak{s}^1)$- invariance. Nevertheless, we present both proofs. It was shown in the proof of Proposition 16 that for the rotation $\rho(\theta)=\theta+b$ the adjoint action of $\rho\in S^1$ on $\mathrm{Vect}(S^1)$ is expressed as $\mathrm{Ad}_\rho u(\theta)=u(\theta-b)$. Then $(\mathrm{Ad}_\rho u,\mathrm{Ad}_\rho v)^{1,0}=(u,v)^{1,0}$, since $u,v$ are periodic with period $2\pi$.

Denote by $p_0=\partial_\theta$ the basis vector for $\mathfrak{s}^1$. Then

$$\begin{aligned}(\mathrm{ad}_{p_0}u,v)^{1,0}&=([p_0,u],v)^{1,0}=-(u',v)^{1,0}\\&=(u,v')^{1,0}=-(u,[p_0,v])^{1,0}=-(u,\mathrm{ad}^{\top}_{p_0},v)^{1,0}.\end{aligned}$$

It implies that the inner product $(\cdot\,,\cdot)^{1,0}$ is invariant under the adjoint action of the algebra $\mathfrak{s}^1$. □

Moreover, the subspaces $\mathrm{Vect}_0(S^1)$ and $\mathfrak{s}^1$ are orthogonal with respect to the inner product $(\cdot\,,\cdot)^{1,0}$, making the linear map $\eta$ from (7.25) an orthogonal projection onto $\mathfrak{s}^1$.

We see that all the hypotheses of Theorem 7.3 are satisfied. Thus, a normal $\mathcal{H}$-horizontal geodesic $\gamma\colon I \to \mathrm{Diff}\, S^1$ is the solution to the equations

$$\kappa^\ell(\dot\gamma) = u, \qquad \dot u = \mathrm{ad}_u^\top(u+\lambda) = 3uu' + 2\lambda u', \quad u \in \mathrm{Vect}_0(S^1),\ \lambda \in \mathbb{R}.$$

The Riemannian geodesics obtained for $\lambda = 0$ are solutions to inviscid Burgers' equation $\dot u = 3uu'$. For the map $\pi\colon\ \mathrm{Diff}\, S^1 \to \mathrm{Diff}\, S^1/S^1$ we denote the base space $\mathrm{Diff}\, S^1/S^1$ by $B$. Let $\mathbf{b}^{1,0}$ be the Riemannian metric on $B = \mathrm{Diff}\, S^1/S^1$ obtained as a pushforward of $\mathbf{h}^{1,0}$ by $\pi$. Then the Riemannian geodesics in $B$ with respect to $\mathbf{b}^{1,0}$ are projections $\pi(\gamma)$.

Now we consider a more general family of metrics than just the $L_2$-metric. Define a two-parameter family $(\cdot\,,\cdot)_0^{\alpha\beta}$ of scalar products on $\mathrm{Vect}_0(S^1)$ by the formula

$$(u,v)_0^{\alpha\beta} = \frac{1}{2\pi}\int_0^{2\pi} (\alpha u(\theta)v(\theta) + \beta u'(\theta)v'(\theta))d\theta, \quad u,v \in \mathrm{Vect}_0(S^1).$$

The scalar product is non-degenerate for $\alpha \neq -n^2\beta$, $n \in \mathbb{N}$, and is positively definite only if $\beta \geq 0$ and $\alpha > -\beta$. We extend the inner product $(\cdot\,,\cdot)_0^{\alpha\beta}$ to the entire Lie algebra $\mathrm{Vect}(S^1)$ by the formula

$$(u,v)^{\alpha\beta} = \big(u-\eta(u), v-\eta(v)\big)_0^{\alpha\beta} + \eta(u)\eta(v) \qquad u,v \in \mathrm{Vect}(S^1). \tag{7.28}$$

Let us define a Riemannian metric $\mathbf{g}^{\alpha\beta}$ by left translation of $(\cdot\,,\cdot)^{\alpha\beta}$, and let $\mathbf{h}^{\alpha\beta}$ be its restriction to $\mathcal{H}$. Theorem 7.3 can be applied also in this case and we deduce that an $\mathcal{H}$-horizontal normal geodesic $\gamma\colon I \to \mathrm{Diff}\, S^1$ with respect to the metric $\mathbf{h}^{\alpha\beta}$ is a solution to the equations

$$\kappa^\ell(\dot\gamma) = u, \qquad \beta\dot u'' - \alpha\dot u = \beta(uu''' + 2u'u'') - 3\alpha uu' + 2\lambda u', \quad u \in \mathrm{Vect}_0(S^1).$$

If $\mathbf{b}^{\alpha\beta}$ is the Riemannian metric on $B = \mathrm{Diff}\, S^1/S^1$ induced by $\mathbf{h}^{\alpha\beta}$ as a pushforward, then the Riemannian geodesics on $B$ are given as projections $\pi(\gamma)$ of solutions. The details can be found in [65].

Now we present metrics and normal geodesics for the Virasoro-Bott group $\mathrm{Vir}_{\mu\nu}$, where the sub-index corresponds to the 2-cocycle $\omega_{\mu\nu}$. We extend the inner product $(\cdot\,,\cdot)^{\alpha\beta}$ to the Virasoro algebra $\mathfrak{vir}_{\mu\nu}$. The extension is given by the formula

$$\Big((u\partial_\theta, a_1), (v\partial_\theta, a_2)\Big)_{\mu\nu}^{\alpha\beta} = (u,v)^{\alpha\beta} + a_1a_2.$$

Let $\mathbf{g}_{\mu\nu}^{\alpha\beta}$ be the Riemannian metric on $\mathrm{Vir}_{\mu\nu}$ obtained by left translations of $(\cdot\,,\cdot)_{\mu\nu}^{\alpha\beta}$, and let $\mathbf{h}_{\mu\nu}^{\alpha\beta}$ be its restriction to the sub-bundle $\mathcal{E}$.

Let us calculate the adjoint $\mathrm{ad}_{(u,a)}^\top$ of $\mathrm{ad}_{(u,a)}$ with respect to the metric $(\cdot\,,\cdot)_{\mu\nu}^{1,0}$. Notice that

$$\omega_{\mu\nu}(u,v) = \frac{1}{2\pi}\int_0^{2\pi} \big(\mu u(\theta)v'(\theta) + \nu u'(\theta)v''(\theta)\big)\, d\theta = -(u, L_{\mu\nu}v')^{1,0},$$

where we used the notation $L_{\mu\nu}v = \Big(-\mu + \nu\frac{\partial^2}{\partial\theta^2}\Big)v$ the operator $-\mu + \nu\frac{\partial^2}{\partial\theta^2}$ is also known as the Hill operator. Then we calculate

$$\begin{aligned}\Big(\operatorname{ad}^{\top}_{(u,a)}(v,b),(w,c)\Big)^{1,0}_{\mu\nu} &= \big(v,[u,w]\big)^{1,0} - b\omega_{\mu\nu}(w,u)\\ &= \big(\operatorname{ad}^{\top}_{u} v, w\big)^{1,0} + \big(w, bL_{\mu\nu}u'\big)^{1,0} \qquad (7.29)\\ &= \Big((uv' + 2u'v + bL_{\mu\nu}u', 0),(w,c)\Big)^{1,0}_{\mu\nu}\end{aligned}$$

by formula (7.27).

The conditions of Theorem 7.3 are satisfied. The left logarithmic derivative $(u(t),0) \in (\mathrm{Vect}_0(S^1),0) \subset \mathfrak{g}_{\mu\nu}$ of an $\mathcal{E}$-horizontal normal geodesic $(\gamma,b)\colon I \to \mathrm{Vir}_{\mu\nu}$ with respect to the metric $\mathbf{h}^{1,0}_{\mu\nu}$ is a solution to the equation $(\dot{u},0) = \operatorname{ad}^{\top}_{(u,0)}(u+\lambda_1,\lambda_2)$, $\lambda_1,\lambda_2 \in \mathbb{R}$. This means that the curve $(\gamma,b)$ is a solution to

$$\kappa^{\ell}(\dot{\gamma}) = u, \quad \text{with} \quad \dot{u} = 3uu' + (2\lambda_1 - \lambda_2\mu)u' + \lambda_2\nu u''', \qquad u \in \mathrm{Vect}_0(S^1). \quad (7.30)$$

The corresponding Riemannian geodesics with respect to $\mathbf{g}^{1,0}_{0,1}$ satisfy the KdV equation, as was shown in Subsection 7.1.5 for an analogous right invariant metric. The equations for a normal geodesic with respect to the general metric $\mathbf{h}^{\alpha\beta}_{\mu\nu}$ can be found in [65].

### 7.2.3. Metrics on $\mathcal{H}$ corresponding to invariant Kählerian metrics.

In this subsection we discuss metrics on the sub-bundle $\mathcal{H}$ of $T\,\mathrm{Diff}\,S^1$ obtained by the pullback of some Kählerian metrics defined on $B = \mathrm{Diff}\,S^1/S^1$, where we identify $B$ and $\mathcal{F}_0$ as it was made in Subsection 7.1.4. Recall that the left action of $\mathrm{Diff}\,S^1$ on $\mathcal{F}_0$ is well-defined. Let us choose an Hermitian metric on the base space $\mathcal{F}_0$ assuming that this metric is Kählerian and invariant under the action of $\mathrm{Diff}\,S^1$. All pseudo-Hermitian metrics on $\mathcal{F}_0$ are included into the two-parameter family $\mathbf{b}_{\alpha\beta}$, see [83, 85, 86]. It is sufficient to describe this metric only at $\mathrm{id}_{B_{\mathbb{C}}} \in \mathcal{F}_0$ because at other points of $\mathcal{F}_0$ the metric $\mathbf{b}_{\alpha\beta}$ are defined by the left action of $\mathrm{Diff}\,S^1$. Any smooth curve $f_t$ in $\mathcal{F}_0$ with $f_0 = \mathrm{id}_{B_{\mathbb{C}}}$ can be written as

$$f_t(z) = z + tzF(z) + o(t), \qquad F \in \mathcal{A}_0\,.$$

Hence, we can identify $T_{\mathrm{id}_{B_{\mathbb{C}}}}\mathcal{F}_0$ with $\mathcal{A}_0$ by relating the equivalence class $[t \mapsto f_t]$ to $F$. With this identification, the metric $\mathbf{b}_{\alpha\beta} \in \mathcal{F}_0$ can be written as

$$\begin{aligned}\mathbf{b}_{\alpha\beta}\big|_{\mathrm{id}_{B_{\mathbb{C}}}}(F_1,F_2) &= \frac{2}{\pi}\iint_{B_{\mathbb{C}}}\Big(\alpha F_1'\overline{F}_2' + \beta(zF_1')'\overline{(zF_2')'}\Big)d\sigma(z),\\ &= 2\sum_{n=1}^{\infty}(\alpha n + \beta n^3)a_n\overline{b}_n, \qquad (7.31)\end{aligned}$$

where $d\sigma(z)$ is the area element and $F_1(z) = \sum_{n=1}^{\infty} a_n z^n$, $F_2(z) = \sum_{n=1}^{\infty} b_n z^n$. If $\alpha \neq -n^2\beta, n \in \mathbb{Z}$, then the metric $\mathbf{b}_{\alpha\beta}$ is non-degenerate pseudo-Hermitian. Otherwise, $\mathbf{b}_{\alpha\beta}$ degenerates along a distribution of complex dimension 1. Moreover, we require $\beta \geq 0$ and $-\alpha < \beta$ in order to obtain a positively definite Hermitian metric. Since it is impossible to write the left action of Diff $S^1$ on $\mathcal{F}_0$ explicitly, it is not easy to describe $\mathbf{b}_{\alpha\beta}$ globally on $\mathcal{F}_0$. However, these metrics can be pulled back to $\mathcal{H}$ by projections $\pi\colon$ Diff $S^1 \to \mathcal{F}_0$. Consider the injective map

$$\begin{array}{rccl} d_{\mathrm{id}}\pi\colon & \mathrm{Vect}_0(S^1) & \to & T_{\mathrm{id}_{B_{\mathbb{C}}}}\mathcal{F}_0 \cong \ \mathcal{A}_0 \\ & u\partial_\theta & \mapsto & F \end{array}.$$

Then the elements $F$ and $u$ are related by the formula, see [85]

$$F(e^{i\theta}) = \frac{i}{2}\big(u(\theta) - iJu(\theta)\big).$$

where $J$ is from (7.13). Observe that

$$\begin{aligned} \mathbf{b}_{\alpha\beta}|_{\mathrm{id}_{B_{\mathbb{C}}}}(F_1, F_2) &= \frac{2}{\pi}\iint_{B_{\mathbb{C}}} \Big(\alpha F_1'\overline{F}_2' + \beta(zF_1')'\overline{(zF_2')'}\Big)d\sigma(z) \\ &= \frac{-i}{\pi}\iint_{B_{\mathbb{C}}} \Big(\alpha dF_1 \wedge d\overline{F}_2 + \beta d(zF_1') \wedge d\overline{(zF_2')}\Big) \\ &= \frac{-i}{\pi}\int_{S^1} \Big(\alpha F_1 d\overline{F}_2 + \beta(zF_1')d\overline{(zF_2')}\Big). \end{aligned}$$

So we conclude that for $u, v \in \mathrm{Vect}_0(S^1)$, and $F_1 = \frac{i}{2}\big(u - iJu\big)$, $F_2 = \frac{i}{2}\big(v - iJv\big)$

$$\begin{aligned} &\mathbf{b}_{\alpha\beta}|_{\mathrm{id}_{B_{\mathbb{C}}}}\big(d_{\mathrm{id}}\pi u, d_{\mathrm{id}}\pi v\big) = \\ &\quad = \frac{i}{4\pi}\int_{S^1} \Big(\alpha(u - iJu)\, d(v + iJv) + \beta(u' - iJu')\, d(v' + iJv')\Big) \\ &\quad = \frac{i}{4\pi}\int_{S^1} \Big(\alpha(u\, dv - iJu\, dv + iu\, dJv + Ju\, dJv) \\ &\qquad\qquad + \beta(u'\, dv' - iJu'\, dv' + iu'\, dJv' + Ju'\, dJv')\Big) \\ &\quad = \frac{i}{4\pi}\int_{S^1} \Big(\alpha(u\, dv + Ju\, dJv) + \beta(u'\, dv' + Ju'\, dJv')\Big) \\ &\qquad + \frac{1}{4\pi}\int_{S^1} \Big(\alpha(Ju\, dv - u\, dJv) + \beta(Ju'\, dv' - u'\, dJv')\Big) \\ &\quad = i\omega_{\alpha\beta}(u, v) + \omega_{\alpha\beta}(Ju, v), \end{aligned}$$

where $\omega_{\alpha\beta}$ is 2-cocycle (7.10) and we used $\int_{S^1} u\, dv = \int_{S^1} Ju\, d(Jv)$ in the last equation, that can be shown by Fourier expansions. The corresponding to the cocycle $\omega_{\alpha\beta}$ inner product on $\mathrm{Vect}_0(S^1)$ is obtained by

$$\langle\!\langle u, v \rangle\!\rangle_{\alpha\beta} = \omega_{\alpha\beta}(Ju, v).$$

Observe that

$$\langle\!\langle u, v \rangle\!\rangle_{\alpha\beta} = -(Ju', v)^{\alpha\beta}, \qquad u, v \in \mathrm{Vect}_0(S^1). \tag{7.32}$$

Extend $\langle\!\langle\ ,\ \rangle\!\rangle_{\alpha\beta}$ to an inner product on the whole algebra $\mathrm{Vect}(S^1)$ as in (7.28). Let $\mathbf{g}_{\alpha\beta}$ be the Riemannian metric obtained by left translation of $\langle\!\langle\ ,\ \rangle\!\rangle_{\alpha\beta}$, and let $\mathbf{h}_{\alpha\beta}$

be the metric restricted to $\mathcal{H}$. We apply Theorem 7.3 and deduce that a normal critical curve $\gamma\colon I \to \mathrm{Diff}\, S^1$ is the solution to

$$\kappa^{\ell}(\dot\gamma) = u, \qquad -\alpha J\dot u' + \beta \dot u''' = -\alpha(uJu'' + 2u'^2) + \beta(uJu'''' + 2u''') + 2\lambda u', \quad \lambda \in \mathbb{R}.$$

Here we used the property (7.32), see also [65]. We conclude that the geodesics for $\mathbf{b}_{\alpha\beta}$ can be found by solving the above equation for $\lambda = 0$ and then projecting them to $\mathcal{F}_0$.

For $(\alpha, \beta) = (1, 0)$, this is a special case of the modified Constantin-Lax-Majda (CLM) equation. For more information, see [45], where the Riemannian geometry of $\mathbf{g}_{1,0}$ is considered.

**7.2.4. Controllability on** Diff $S^1$**.** Before we formulate the main result in controllability, we describe some special subgroups of Diff $S^1$.

We start from subalgebras of Vect$(S^1)$. For each $n \in \mathbb{Z}$, let us define

$$p_n = \cos n\theta\, \partial_\theta, \qquad k_n = \sin n\theta\, \partial_\theta.$$

The Lie brackets are given by

$$[k_m, k_n] = \tfrac{m+n}{2} k_{m-n} + \tfrac{m-n}{2} k_{m+n}, \tag{7.33}$$

$$[p_m, p_n] = -\tfrac{m+n}{2} k_{m-n} - \tfrac{m-n}{2} k_{m+n}, \tag{7.34}$$

$$[p_m, k_n] = -\tfrac{m+n}{2} p_{m-n} + \tfrac{m-n}{2} p_{n+m}. \tag{7.35}$$

It is easy to see from (7.33–7.35) that $\mathfrak{h}_n = \mathrm{span}\{p_0, p_n, k_n\}$ are subalgebras of Vect$(S^1)$, and that $\mathfrak{h}_n$ is isomorphic to $\mathfrak{su}(1,1)$ for each $n$. To each Lie sub-algebra $\mathfrak{h}_n \subset \mathrm{Vect}(S^1)$ corresponds a subgroup $\mathbb{H}_n$ of Diff $S^1$.

To show that any two points on groups Diff $S^1$ or Vir can be connected by $\mathcal{H}$- or, respectively, $\mathcal{E}$-horizontal curve, we use the invariance of these horizontal sub-bundles under the corresponding group action. We start from a general result. Assume that a horizontal sub-bundle $\mathcal{H}$ is invariant under the action of some subgroup $\mathbb{K}$ of a given group $\mathbb{G}$. Then, if the tangent bundle $T\mathbb{K}$ is transversal to $\mathcal{H}$, the question of controllability is reduced to the question whether elements of $\mathbb{K}$ can be reached from the unity of $\mathbb{G}$ by an $\mathcal{H}$-horizontal curve.

**Lemma 7.** *Let $\mathbb{G}$ be a Lie group with the Lie algebra $\mathfrak{g}$, and let a left- (or right-) invariant horizontal sub-bundle $\mathcal{H}$ be obtained by left (or right) translations of a subspace $\mathfrak{h} \subseteq \mathfrak{g}$. Assume that there is a sub-group $\mathbb{K}$ of $\mathbb{G}$ with the Lie algebra $\mathfrak{k}$ such that $\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{k}$ for some $\mathfrak{p} \subseteq \mathfrak{h}$. Suppose also that $\mathfrak{h}$ is* Ad$(\mathbb{K})$*-invariant. Then any pair of elements in $\mathbb{G}$ can be connected by a smooth $\mathcal{H}$-horizontal curve, if and only if, for every $a \in \mathbb{K}$ there is an $\mathcal{H}$-horizontal smooth curve connecting $\mathbf{1} \in \mathbb{K}$ and $a$.*

*Proof.* We present the proof for the case of a left-invariant sub-bundle $\mathcal{H}$. Let $c\colon [0,1] \to \mathbb{G}$ be any curve (not necessarily horizontal), connecting the points $a_0$ and $a_1$, and having left logarithmic derivative $u$. Using left translation of $c$ by $a_0^{-1}$, we can assume that $a_0 = \mathbf{1}$. Let $\mathrm{pr}_{\mathfrak{k}}\colon \mathfrak{g} \to \mathfrak{k}$ be the projection with the kernel $\mathfrak{p} \subset \mathfrak{h}$.

Consider the projection $k(t) = \mathrm{pr}_{\mathfrak{k}}\, u(t)$, $t \in [0,1]$. Let $\vartheta$ be a curve in $\mathbb{K}$ with left logarithmic derivative $k$, starting at $\mathbf{1}$. Then the left logarithmic derivative of the curve $\vartheta(t)^{-1}$ is $-\operatorname{Ad}_{\vartheta} k$.

Let us show that the curve $\gamma_1(t) = c(t) \cdot \vartheta(t)^{-1}$, $t \in [0,1]$, is $\mathcal{H}$-horizontal. We calculate the left logarithmic derivative of $\dot{\gamma}_1(t)$ and find

$$\kappa^{\ell}(\partial_t(c(t) \cdot \vartheta(t)^{-1})) = \operatorname{Ad}_{\vartheta(t)}(u(t) - k(t)) \in \mathfrak{h},$$

since $\mathfrak{h}$ is $\operatorname{Ad}(\mathbb{K})$ invariant. Hence, we have constructed an $\mathcal{H}$-horizontal curve $\gamma_1$, from $\mathbf{1}$ to $a_1 \cdot \vartheta(1)^{-1}$. Applying the right translation by $\vartheta(1)$, that keeps the curve $\mathcal{H}$-horizontal because of the $\operatorname{Ad}(\mathbb{K})$-invariance of $\mathfrak{h}$, we get a curve from $\vartheta(1)$ to $a_1$. Moreover, by the hypothesis of the theorem, we can connect $\mathbf{1}$ with $\vartheta(1)$ by a smooth horizontal curve $\gamma_2$. Finally, we glue the curves $\gamma_1$ and $\gamma_2$ into one smooth curve by slowing exponentially down to zero speed at the connecting point. □

**Theorem 7.4.** *The following is true.*

(a) *Let $\mathcal{H}$ be a choice of a horizontal sub-bundle on* $\operatorname{Diff} S^1$ *defined as in Section* 7.2.1. *Then any pair of points can be connected by an $\mathcal{H}$-horizontal curve.*

(b) *Let $\mathcal{E}$ be a choice of a horizontal sub-bundle on* $\mathrm{Vir}_{\mu\nu}$ *defined as in Section* 7.2.1. *Then any two points on* $\mathrm{Vir}_{\mu\nu}$ *can be connected by an $\mathcal{E}$-horizontal curve.*

*Proof.* To prove (a), it is sufficient to show that any two points in $\operatorname{Diff} S^1$ can be connected by an $\mathcal{H}$-horizontal curve. Due to Lemma 7, we only need to verify that $\mathrm{id} \in \operatorname{Diff} S^1$ can be connected with any element in $S^1$ by an $\mathcal{H}$-horizontal curve. The subgroup $S^1$ is contained in $\mathbb{H}_n$ for any $n$, in particular, $S^1$ can be considered as a subgroup of $\mathbb{H}_1$. Any $\mathcal{H}$-horizontal curve in $\mathbb{H}_1$, will have left logarithmic derivative in $\mathfrak{h}_1 \cap \mathrm{Vect}_0(S^1) = \mathrm{span}\{k_1, p_1\}$. Since $[p_1, k_1] = p_0$, the horizontal distribution $\mathcal{H}$ restricted to $\mathbb{H}_1$ is bracket generating. The group $\mathbb{H}_1$ is finite-dimensional, therefore we can apply the Rashevskiĭ-Chow theorem to conclude that every point in $\mathbb{H}_1$, including points in $S^1$, can be reached by an $\mathcal{H}$-horizontal curve.

To prove (b), we need to show that any point in $\widehat{S}^1 = \{\theta \mapsto (\theta + a, b) \in \mathrm{Vir}_{\mu\nu}\}$ can be connected to $(\mathrm{id}, 0) \in \mathrm{Vir}_{\mu\nu}$ by an $\mathcal{E}$-horizontal curve. Let

$$\widehat{\mathbb{H}}_n = \{(\phi, a) \in \mathrm{Vir}_{\mu\nu} \, : \, \phi \in H_n, \; a \in \mathbb{R}\}.$$

which has Lie algebra

$$\widehat{\mathfrak{h}}_n = \mathrm{span}\,\{(p_0, 0), (p_n, 0), (k_n, 0), (0, 1)\}.$$

Unfortunately, the sub-bundle restricted to the group $\widehat{\mathbb{H}}_n$ is not bracket generating. We need to find a smaller subgroup. The Lie algebras $\widehat{\mathfrak{h}}_n$ have special subalgebras

$$\widehat{\mathfrak{t}}_n = \mathrm{span}\,\big\{(p_0, n^2\nu - \mu), (p_n, 0), (k_n, 0)\big\}.$$

Denote the corresponding subgroups by $\widehat{\mathbb{T}}_n$. On the contrary to what holds on $\widehat{\mathbb{H}}_n$, the distribution $\mathcal{E}$ restricted to any subgroup $\widehat{\mathbb{T}}_n$ is bracket generating, and so all elements in such a subgroup $\widehat{\mathbb{T}}_n$ can be reached by an $\mathcal{E}$-horizontal curve. It is clear that

$$\widehat{\mathbb{T}}_n \cap \widehat{S}^1 = \left\{ \big(\theta \mapsto \theta + r, r(n^2\nu - \mu)\big) \ : \ r \in \mathbb{R} \right\},$$

where $\widehat{S}^1$ is the subgroup of translations in $\widehat{\mathbb{H}_1}$. Since $\widehat{S}^1$ is isomorphic to $\mathbb{R}^2$ as a group and $\nu \neq 0$, we can find unique elements $g_j \in \widehat{\mathbb{T}}_j \cap \widehat{S}^1$, $j = 1, 2$ such that $g = g_1 \cdot g_2 = g_2 \cdot g_1$ for any $g \in \widehat{S}^1$. Denote by $c_1$ and $c_2$ curves that connect $(\mathrm{id}, 0) \in \mathrm{Vir}_{\mu\nu}$ with $g_1$ and $g_2$, respectively. We can reach $g$ by first following a curve $c_2$ and then moving to $g$ by a curve from $g_2 \circ c_1$ to $g_1$, that is the translation of $c_1$ by $g_2$. This finishes the proof. $\square$

The question of the controllability on infinite dimensional manifolds is very difficult and is not well studied. We mention the book [93], where the smooth calculus on most general complete topological locally convex vector spaces is presented and the theory of infinite dimensional manifolds is also developed, see also [108, 116] for the study of infinite dimensional Lie groups. The analogous of the Chow-Rashevskiĭ theorem for the Hilbert manifolds can be found in [69] for the Banach manifolds in [94] and for manifolds modelled on more general complete topological vector spaces see [79].

## 8. Appendix A

### 8.1. Smooth manifolds

**Definition 35.** *A topological space $S$ is second countable if its topology has a countable base, that is a countable collection $B$ of open sets such that every open set is a union of some sub-collection of $B$.*

**Definition 36.** *A set $P$ is a submanifold of a smooth manifold $M$ if:*

1. *$P$ is a smooth manifold,*
2. *the inclusion map $j \colon P \to M$ is smooth and at each point $p \in P$ its differential $d_q j \colon T_qP \to T_{j(q)}M$ is injective.*

Some authors require that $P$ is also a topological subspace of $M$.

**Definition 37.** *An immersion $\varphi \colon M^m \to N^n$ is a smooth map such that the differential $d_q\varphi \colon T_qM^m \to T_{\varphi(q)}N^n$ is injective for all $q \in M$. It is equivalent to say that the Jacobi matrix of $d_q\varphi$ has rank $m$ relatively to one (hence every) choice of coordinate system.*

**Definition 38.** *An embedding $\phi \colon P \to M$ of a manifold $P$ into a manifold $M$ is*

1. *an injective immersion, such that*
2. *the induced map $\tilde{\phi} \colon P \to \phi(P) \subset M$ is a homeomorphism onto the subspace $\phi(P)$*

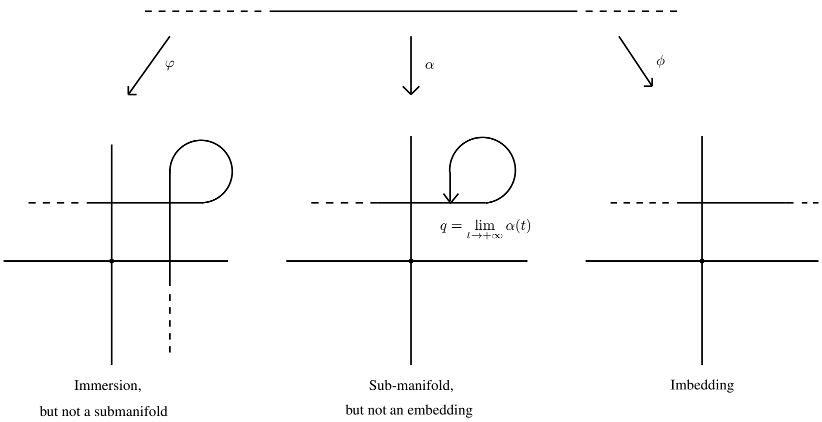

FIGURE 8.1. Difference between immersion, embedding and submanifold.

If $P$ is a submanifold of $M$, then the inclusion map $j\colon P \to M$ is an embedding. Conversely, if $\phi\colon P \to M$ is an embedding, then this map induces a manifold structure on the image $\phi(P) \subset M$ and the induced map $\tilde{\phi}\colon P \to \phi(P)$ is a diffeomorphism. The map $\phi \circ \tilde{\phi}^{-1}\colon \phi(P) \to M$ is the inclusion $j\colon \phi(P) \to M$ which is smooth and whose differential is injective as a composition of two injective differentials $d\phi \circ d(\tilde{\phi}^{-1})$. We conclude that $\phi(P)$ is a submanifold of $M$. In Figure 8.1 one can see the difference between embedding and submersion of the manifold $P = \mathbb{R}$ into the manifold $M = \mathbb{R}^2$.

**Definition 39.** *A submersion $\pi\colon M \to B$ is a smooth surjective map such that the differential $d_q\pi\colon T_qM \to T_{\pi(q)}B$ is surjective for all $q \in M$.*

**Definition 40.** *Let $(M, g_M)$ and $(N, g_N)$ be two Riemannian manifolds. A diffeomorphism $\iota\colon M \to N$ is called a Riemannian isometry, if*

$$g_M(v, w) = g_N(d_q\iota(v), d_q\iota(w)) \quad \text{for all} \;\; v, w \in T_qM \;\; \text{and all} \;\; q \in M.$$

**Definition 41.** *Let $(M, g_M)$ and $(B, g_B)$ be two Riemannian manifolds and let $\pi\colon M \to B$ be a submersion. Let $T_qM = \ker(d_q\pi) \oplus_\perp \Big(\ker(d_q\pi)\Big)^\perp$ be the orthogonal decomposition with respect to $g_M$. If the restriction $d\pi|_{\ker(d_q\pi)}\colon \ker(d_q\pi) \to T_{\pi(q)}B$ is a linear isometry for any $q \in M$, then the map $\pi$ is called the Riemannian submersion.*

**Definition 42.** *A pairing between the tangent and the co-tangent bundle is a map*

$$\langle \cdot\,, \cdot \rangle\colon T_qM \times T_q^*M \to \mathbb{R}$$

*which is bi-linear, non-degenerate, and smoothly varying with respect to $q \in M$. It is non-degenerate in the sense that if $\langle v, \lambda \rangle = 0$ for all $v \in T_qM$ and $\lambda \in T_q^*M$, then $\lambda \equiv 0$.*

**Definition 43.** *An absolutely continuous curve* $c\colon I \to M$ *is an integral curve of a vector field* $X \in \mathrm{Vect}(M)$ *if it satisfies the differential equation*

$$\dot{c}(t) = X(c(t)), \quad \textit{for all} \quad t \in I.$$

A vector field $X$ is called *complete* if each of its non-extendable integral curves (starting from different points $q \in M$) is defined on $I = \mathbb{R}$. Let us denote by $c_{q,X}$ the integral curve of a complete vector field $X$ starting at $q \in M$. Thus the curve $c_{q,X}$ is a solution of the Cauchy problem

$$\dot{c}(t) = X(c(t)), \quad c(0) = q, \quad t \in \mathbb{R}.$$

**Definition 44.** *The flow of a complete vector field* $X$ *on a smooth manifold* $M$ *is the map* $\widehat{X}\colon \mathbb{R} \times M \to M$ *given by*

$$\widehat{X}(t, q) = c_{q,X}(t),$$

*where* $c_{q,X}$ *is the non-extendable integral curve of* $X$ *starting at* $q \in M$.

**Proposition 18.** *The flow* $\widehat{X}$ *of a complete vector field satisfies the conditions:*

1. $\widehat{X}(0, \cdot)\colon M \to M$ *is the identity map of* $M$,
2. $\widehat{X}(s + t, \cdot) = \widehat{X}\big(s, (\widehat{X}(t, \cdot)\big) = \widehat{X}(s, \cdot) \circ \widehat{X}(t, \cdot)$ *for all* $s, t \in \mathbb{R}$. *As a corollary we conclude that flows commute for fixed times,*
3. *the map* $\widehat{X}(s, \cdot)\colon M \to M$ *is a diffeomorphism for any* $s \in \mathbb{R}$, *where the inverse map is* $\widehat{X}^{-1}(s, \cdot) := \widehat{X}(-s, \cdot)$.

We need the completeness assumption in order to work with the entire manifold $M$ and not only locally. For arbitrary vector fields one can define a local analogue of the flow.

Now we define a Levi-Civita connection. We start from the definition of the *affine connection.*

**Definition 45.** *An affine connection* $\nabla$ *on a smooth manifold* $M$ *is a map*

$$\begin{array}{cccc} \nabla\colon & \mathrm{Vect}(M) \times \mathrm{Vect}\, M & \to & \mathrm{Vect}(M) \\ & (X, Y) & \mapsto & \nabla_X Y, \end{array}$$

*satisfying the following properties*

1. $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$,
2. $\nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z$, *for all* $X, Y, Z \in \mathrm{Vect}(M)$ *and* $f, g \in C^\infty(M)$.
3. $\nabla_X(fY) = f\nabla_X Y + X(f)Y$,

The notion of the affine connection leads to the definition of the *covariant derivative* along a given curve. Namely, let an affine connection $\nabla$ be defined on a smooth manifold $M$. Suppose that $c\colon I \to M$ is a smooth curve and $X\colon I \to TM$ is a vector field along the curve $c$. Then there exists a unique correspondence which associates to a vector field $X$ another vector fields $\frac{DX}{dt}$ along $c$ by the rule

$$\frac{DX}{dt} := \nabla_{\dot{c}} X(c(t)).$$

The covariant derivative $\frac{D}{dt}$ satisfies the properties

$$\frac{D}{dt}(X+Y)=\frac{DX}{dt}+\frac{DY}{dt}, \qquad \frac{D}{dt}(fX)=\left(\frac{df}{dt}\right)X+f\frac{D}{dt}X,$$

where $f$ is a smooth function along the curve $c$.

**Definition 46.** *An affine connection $\nabla$ on a smooth manifold $M$ is symmetric if*

$$\nabla_X Y-\nabla_Y X=[X,Y] \quad \textit{for all} \quad X,Y\in \mathrm{Vect}(M).$$

**Definition 47.** *Let $(M,g)$ be a Riemannian manifold with an affine connection $\nabla$. We say, that the affine connection $\nabla$ is compatible with the Riemannian metric $g$ if*

$$X\big(g(Y,Z)\big)=g(\nabla_X Y,Z)+g(Y,\nabla_X Z) \quad \textit{for all} \quad X,Y,Z\in \mathrm{Vect}(M).$$

The following theorem asserts that the existence and uniqueness of the Riemannian metric guarantees the existence of an affine connection that is compatible with the Riemannian metric and symmetric.

**Theorem 8.1.** [23, 117] *Given a Riemannian manifold $(M,g)$ there is a unique affine connection $\nabla$, called the* Levi-Civita connection*, such that $\nabla$ is symmetric and compatible with the metric $g$.*

### 8.2. Symplectic manifolds

**Definition 48.** *A non-degenerate skew symmetric real valued 2-form $\Omega$ is called a symplectic form. The pair $(N,\Omega)$, where $N$ is a smooth manifold and $\Omega$ is a symplectic form, is called a symplectic manifold.*

In some literature it is also required that $\Omega$ is a closed form.

**Definition 49.** *Let $(N,\Omega)$ be a symplectic manifold and $H\in C^\infty(N)$ a function, then the associated with $H$ Hamiltonian vector field $\overrightarrow{H}$ is defined by*

$$\Omega(X,\overrightarrow{H}):=dH(X) \quad \textit{for all} \qquad X\in \mathrm{Vect}(N).$$

The Poisson brackets between functions $H,K\in C^\infty(N)$ is the directional derivative of one function in the direction of the Hamiltonian vector field, associated with another function. Namely,

$$\{H,K\}:=dK(\overrightarrow{H})=\Omega(\overrightarrow{H},\overrightarrow{K}).$$

As an example, consider a smooth manifold $M$ and its co-tangent bundle $T^*M$. Recall that if $\mathrm{pr}^*_M\colon T^*M\to M$ is the canonical projection, then

$$d(\mathrm{pr}^*_M)\colon T(T^*M)\to TM.$$

We use the notation $\langle\cdot\,,\cdot\rangle$ to denote the pairing between $TM$ and $T^*M$. Define a real valued one-form $\omega\colon T(T^*M)\to\mathbb{R}$ on the manifold $N=T^*M$ by

$$\omega_{(q,\lambda)}(v):=\langle d(\mathrm{pr}^*_M)(v),\lambda\rangle, \quad v\in T_{(q,\lambda)}(T^*M).$$

Then the 2-form $\Omega=d\omega$ is a symplectic form. Verify it!

If we chose a chart $\big(U,\varphi=(x^1,\dots,x^n)\big)$ on $M$, then it induces the chart

$$\big(T^*U,\Phi=(x^1,\dots,x^n,\lambda_1,\dots,\lambda_n)\big)$$

on $T^*M$. The canonical projection $\mathrm{pr}^*_M$ and its differential take the matrix form

$$\mathrm{pr}^*_M=\begin{pmatrix} I_{n\times n} & 0\\ 0 & 0\end{pmatrix}\qquad\Longrightarrow\qquad d\,\mathrm{pr}^*_M=\begin{pmatrix} I_{n\times n} & 0\\ 0 & 0\end{pmatrix}.$$

Therefore, the one-form $\omega$ at a point $(q,\lambda)=(x^1,\dots,x^n,\lambda_1,\dots,\lambda_n)$ is written as $\omega_{(q,\lambda)}=\sum_{j=1}^n\lambda_j dx^j$. The symplectic form $\Omega$ becomes $\Omega=\sum_{j=1}^n d\lambda_j\wedge dx^j$. The Hamiltonian vector field is

$$\overrightarrow{H}(q,\lambda)=\sum_{j=1}^n\Big(\frac{\partial H}{\partial\lambda_j}\frac{\partial}{\partial x^j}-\frac{\partial H}{\partial x^j}\frac{\partial}{\partial\lambda_j}\Big)$$

and the Poisson brackets are

$$\{H,K\}=\sum_{j=1}^n\Big(\frac{\partial H}{\partial\lambda_j}\frac{\partial K}{\partial x^j}-\frac{\partial H}{\partial x^j}\frac{\partial K}{\partial\lambda_j}\Big).$$

To each vector field $X\in\mathrm{Vect}(M)$ we associate the function $H_X\colon T^*M\to\mathbb{R}$ by

$$H_X(q,\lambda)=\langle X(q),\lambda_q\rangle.$$

Then one can associate the Hamiltonian vector field $\overrightarrow{H}_X$ to the function $H_X$ in the following way. If $X(q)=\sum_{k=1}^n X^k(x^1,\dots,x^n)\partial_k$ and $\lambda_q=\sum_{j=1}^n\lambda_j(x^1,\dots,x^n)dx^j$ then $H_X(q,\lambda)=\sum_{j=1}^n\lambda_jX^j(q)$ and

$$\overrightarrow{H}_X(q,\lambda)=\sum_{j=1}^n\Big(X^j(q)\frac{\partial}{\partial x^j}-\Big[\frac{\partial}{\partial x^j}\big(\sum_{k=1}^n\lambda_kX^k(q)\big)\Big]\frac{\partial}{\partial\lambda_j}\Big).$$

Now it is obvious that

$$d(\mathrm{pr}^*_M)(\overrightarrow{H}_X(q,\lambda))=X(q)\quad\text{for all}\quad(q,\lambda)\in T^*M.$$

**Corollary 9.** *Geodesics produced by the Hamiltonian function $H_X$ coincide with integral curves of the vector field $X$.*

If $X\in\mathrm{Vect}(M)$, then the vector field $\overrightarrow{H}_X$ is called the *Hamiltonian lift of $X$*.

EXERCISES.

1. Let $(N,\Omega)$ be a symplectic manifold. Verify that $(C^\infty(N),\{\cdot\,,\cdot\})$ is a Lie algebra and the map $H\mapsto\overrightarrow{H}$ is a Lie algebra homomorphism

$$(C^\infty(N),\{\cdot\,,\cdot\})\quad\longrightarrow\quad(\mathrm{Vect}(N),[\cdot\,,\cdot]).$$

2. Let $M$ be a smooth manifold. Show that $\{H_X,H_Y\}=H_{[X,Y]}$ for all $X,Y\in\mathrm{Vect}(M)$. Conclude that the map $X\mapsto H_X$ produces a homomorphism of Lie algebras

$$(\mathrm{Vect}(M),[\cdot\,,\cdot])\quad\longrightarrow\quad(C^\infty(T^*M),\{\cdot\,,\cdot\}).$$

### 8.3. Lie groups

The content of this subsection can be found in [42, 87, 108, 131].

**Definition 50.** *A Lie group $\mathbb{G}$ is a smooth (finite dimensional) manifold $M$ and a group such that the operations of*

$$\begin{array}{lllll} \textit{multiplication} & \rho\colon & M\times M & \to & M, \\ & & (\tau,q) & \to & \tau q \end{array}$$

*and*

$$\begin{array}{lllll} \textit{inversion} & \mathrm{in}\colon & M & \to & M, \\ & & x & \to & x^{-1} \end{array}$$

*are smooth maps between corresponding smooth manifolds: $M\times M\to M$ and $M\to M$, respectively.*

It is customary to use the letter $\mathbb{G}$ to denote the underlying manifold $M$ and the pair $(M,\rho)$ in the case of Lie groups.

**Definition 51.** *A Lie algebra $\mathfrak{g}$ over $\mathbb{R}$ ($\mathbb{C}$) is a real (complex) vector space $V$ together with an operation $[\cdot\,,\cdot]\colon V\times V\to V$ (called the bracket, commutator, or Lie product) satisfying the following three axioms:*

1. *skew symmetry:* $[X,Y]=-[Y,X]$,
2. *bi-linearity:* $[aX+bY,Z]=a[X,Z]+b[Y,Z]$, $a,b\in\mathbb{R}\,(\mathbb{C})$ *(and the same with respect to the second term),*
3. *Jacobi identity:* $[[X,Y],Z]+[[Z,X],Y]+[[Y,Z],X]=0$

*for any $X,Y,Z\in V$.*

We will use the letters $\mathfrak{g},\mathfrak{h},\dots$ to denote the also underlying vector spaces $V,U,\dots$ in the case of Lie algebras.

**Example 10.**

1. The general linear group $GL(n,\mathbb{R})=GL(n)$ is the Lie group of all $(n\times n)$-matrices $L$ with real entries such that $\det L\neq 0$. Since the determinant function $\det\colon\mathbb{R}^{n^2}\to\mathbb{R}$ is smooth, the underlying manifold of $GL(n)$ is an open subset in $\mathbb{R}^{n^2}$ defined by the complement to the inverse image of the function "det" of the value $0\in\mathbb{R}$. The group multiplication is the multiplication of matrices. The corresponding Lie algebra $\mathfrak{gl}(n,\mathbb{R})=\mathfrak{gl}(n)$ is formed by all $(n\times n)$-matrices and isomorphic to $\mathbb{R}^{n^2}$ as a vector space. The commutator in $GL(n)$ is the commutator of two matrices. The group $GL(n)$ is the group of all non-degenerate linear transformations of $\mathbb{R}^n$. In a similar way the group $GL(n,\mathbb{C})$ can be defined.
2. The orthogonal group $O(n,\mathbb{R})=O(n)$ is the subspace of $GL(n)$ such that $L^{tr}L=LL^{tr}=\mathrm{Id}$, where $L^{tr}$ is the transpose to $L\in O(n)$. Verify that in this case $\det^2 L=1$. The smooth underlying manifold for $O(n)$ is the level set of the function "det" inside of $\mathbb{R}^{n^2}$ and it consists of two connected components corresponding to the value 1 and -1 of "det". The special orthogonal

group $SO(n,\mathbb{R}) = SO(n)$ is the subset of $O(n)$ whose matrices have determinant 1 and the underlying manifold is the connected component containing the identity matrix. Both groups have the same Lie algebra $\mathfrak{o}(n)$ consisting of $(n \times n)$-matrices that are skew symmetric: $L^{tr} = -L$. The main feature of these groups is that under their transformations the Euclidean inner product in $\mathbb{R}^n$ is preserved. (Why?)

3. The unitary group $U(n)$ is the group of $(n\times n)$-matrices with complex entries such that $\bar{L}^{tr}L = L\bar{L}^{tr} = \mathrm{Id}$, where $\bar{L}^{tr}$ is the transposed and conjugate matrix to $L \in U(n)$. The special unitary group $SU(n)$ is the subset of $U(n)$ whose matrices have determinant 1. The Lie algebra $\mathfrak{u}(n)$ is the set of $(n\times n)$-matrices that are skew-Hermitian symmetric: $\bar{L}^{tr} = -L$. The Lie algebra $\mathfrak{su}(n)$, $n \geq 2$, is the subset of $\mathfrak{u}(n)$ having vanishing trace. The unitary and special unitary groups preserve the Hermitian product $(z,w) = \sum_{k=1}^{n} \bar{z}^k w^k$ in $\mathbb{C}^n$.
4. The symplectic group $Sp(n)$ is the group of $(n \times n)$-matrices with quaternion entries such that $\bar{L}^{tr}L = L\bar{L}^{tr} = \mathrm{Id}$, where $\bar{L}^{tr}$ is the transposed and quaternion conjugate matrix to $L$. The Lie algebra $\mathfrak{sp}(n)$ is the set of $(n\times n)$-matrices that are skew-Hermitian symmetric: $\bar{L}^{tr} = -L$. Symplectic groups preserve the Hermitian product in the $n$-dimensional quaternionic space $\mathbb{Q}^n$.
5. The special Euclidean group $SE(n)$, or the group of rigid motions in $\mathbb{R}^n$ is the group consisting of rotations and translations in $\mathbb{R}^n$. An element $\tau \in SE(n)$ is usually written as a pair $\tau = (A,a)$, where $A \in SO(n)$ and $a$ is a $n$-dimensional vector. The multiplication is given by $\tau v = (A,a)(B,b) := (AB, Ab + a)$. Thus the group $SE(n)$ is the group of all isometries in the Euclidean space. Its dimension is $\frac{n(n+1)}{2}$, where $n$ stands for translations and $\frac{n(n-1)}{2}$ is the dimension of $SO(n)$.

We define the exponential map and list its properties. Let $(\mathbb{R},+)$ be the additive group of real numbers and $\mathfrak{r}$ be the corresponding Lie algebra with generator $\frac{d}{dr}$. Let $\mathbb{G}$ be a Lie group, $\mathfrak{g}$ be its Lie algebra, and $X \in \mathfrak{g}$ be an arbitrary element. Then the map

$$\begin{array}{rccc} h\colon & \mathfrak{r} & \to & \mathfrak{g} \\ & t\frac{d}{dr} & \mapsto & tX \end{array}$$

is a homomorphism of the Lie algebra $\mathfrak{r}$ into the Lie algebra $\mathfrak{g}$. The theorems of Lie group theory [87, 131] ensure that there is a unique Lie group homomorphism $c_X$, such that

$$c_X\colon \mathbb{R} \to \mathbb{G}, \quad \text{and} \quad d\,c_X = h, \qquad \text{or} \quad dc_X(t\frac{d}{dr}) = tX.$$

In other words, the curve $c_X\colon \mathbb{R} \to \mathbb{G}$ is a one-parametric subgroup of $\mathbb{G}$ and it is such that $c_X(0) = e$ and $\dot{c}_X(0) = X$. The curve $c_X(t)$, $t \in \mathbb{R}$, is called the *exponential curve* and it is often denoted by $\exp(tX)$, $t \in \mathbb{R}$. The map

$$\begin{array}{rccc} \exp\colon & \mathfrak{g} & \to & \mathbb{G} \\ & X & \mapsto & \exp(X). \end{array}$$

is called the *exponential map*. We list the properties of the exponential map in the following theorem.

**Theorem 8.2.** [87, 131] *Let $X$ belong to the Lie algebra $\mathfrak{g}$ of a finite dimensional Lie group $\mathbb{G}$. Then the following properties hold.*

1. *The exponential curve* $\exp(tX) = c_X(t)$ *for each* $t \in \mathbb{R}$ *satisfies*
$$\frac{d}{dt}\Big|_{t=0} \exp(tX) = \frac{d}{dt}\Big|_{t=0} c_X(t) = \dot{c}_X(0) = X,$$
(*see also the definition of the exponential curve in Subsection* 3.1).
2. $\big(\exp(t_1+t_2)X\big) = \big(\exp(t_1X)\big)\big(\exp(t_2X)\big)$, *for all* $t_1, t_2 \in \mathbb{R}$.
3. $\exp(-tX) = \big(\exp(tX)\big)^{-1}$ *for each* $t \in \mathbb{R}$.
4. *The map* $\exp\colon \mathfrak{g} \to \mathbb{G}$ *is a* $C^\infty$*-map between two manifolds.*
5. *The differential at zero vector of the exponential map* $d_0 \exp\colon T_0\mathfrak{g} \to T_e\mathbb{G}$ *is the identity map* $\mathfrak{g} \to \mathfrak{g}$, *where we identify elements of* $\mathfrak{g}$ *with* $T_0\mathfrak{g}$ *and with* $T_e\mathbb{G}$. *An important corollary is that* $\exp$ *gives a diffeomorphism between a neighborhood of* $0 \in \mathfrak{g}$ *and a neighborhood of* $e \in \mathbb{G}$.
6. *The left translation of* $c_X$ *by* $\tau \in \mathbb{G}$ *given by* $\tilde{c} = l_\tau(c_X(t)) = \tau c_X(t)$ *is the unique integral curve of the left invariant vector field* $\widetilde{X}$ ($\widetilde{X}(e) = X$) *such that it starts at the point* $\tau$: $\tilde{c}(0) = \tau$. *As a particular consequence, left invariant vector fields are always complete.*
7. *The one-parametric flow of diffeomorphisms* $\widehat{X}(t,\tau)\colon \mathbb{G} \to \mathbb{G}$ *associated with a left invariant vector field* $\widetilde{X}$ ($\widetilde{X}(e) = X$) *is given by* $\widehat{X}(t,\tau) = \tau c_X(t) = r_{c_X(t)}\tau$, *where* $r_{c_X(t)}$ *is the right translation by* $c_X(t)$.
8. *In the neighborhoods of* $0 \in \mathfrak{g}$ *and* $e \in \mathbb{G}$, *where* $\exp$ *is a diffeomorphism, the inverse map is defined and is called logarithm. It expresses the product of two exponents through the Baker-Campbell-Hausdorff formula* [123], *whose first terms are given as follows*
$$\begin{aligned} \exp(X)\exp(Y) &= \exp\Big(X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] \\ &- \frac{1}{12}[Y,[X,Y]] + \dots\Big). \end{aligned} \tag{8.1}$$

It is useful to keep in mind the following diagram defining the exponential map as time-one value of the exponential curve. For chosen $X \in \mathfrak{g}$

$$\begin{tikzcd}
t\frac{d}{dr} \in \mathfrak{r} \arrow[r, "h=dc_X"] \arrow[d, "\mathrm{Id}"'] & \mathfrak{g} \ni tX \arrow[d, "\exp(tX)"] & \Longrightarrow & X \mapsto \exp X = c_X(1) \\
t \in \mathbb{R} \arrow[r, "c_X"'] & \mathbb{G} \ni c_X(t) & &
\end{tikzcd}$$

The straight line $tX \in \mathfrak{g}$ is mapped to the one-parametric subgroup $c_X(t) = \exp(tX) \in \mathbb{G}$.

**Definition 52.** *A subgroup $N$ of a group $\mathbb{G}$ is called a normal subgroup if it is invariant under conjugation; that is, for each element $n \in N$ and each $\tau \in \mathbb{G}$ the element $\tau n \tau^{-1} \in N$.*

**Definition 53.** *A group $\mathbb{G}$ is called simple if it is a non-trivial group and there are no other normal subgroups except the trivial subgroup and the group itself.*

A group that is not simple can be decomposed into two smaller groups, a normal subgroup and the corresponding quotient group, and the process can be repeated.

**8.3.1. Action of Lie groups on manifolds.** Let $M$ be a smooth manifold and let $\mathbb{G}$ be a Lie group.

**Definition 54.** *An action of $\mathbb{G}$ on $M$ on the left is a smooth map $\mu\colon \mathbb{G} \times M \to M$ such that*

$$\mu(\varsigma\tau, q) = \mu(\varsigma, \mu(\tau, q)), \qquad \mu(e, q) = q,$$

*for all $\varsigma, \tau \in \mathbb{G}$ and $q \in M$.*

If $\mu\colon \mathbb{G} \times M \to M$ is an action of $\mathbb{G}$ on $M$ on the left, then for a fixed $\varsigma \in \mathbb{G}$ the map $q \mapsto \mu(\varsigma, q)$ is a diffeomorphism of $M$ which we will denote by $\mu_\varsigma$. Similarly we define a right action.

**Definition 55.** *An action of $\mathbb{G}$ on $M$ on the right is a smooth map $\mu\colon M \times \mathbb{G} \to M$ such that*

$$\mu(q, \varsigma\tau) = \mu(\mu(q, \varsigma), \tau), \qquad \mu(q, e) = q$$

*for all $\varsigma, \tau \in \mathbb{G}$ and $q \in M$.*

We also use the notation $\tau.q$ instead of $\mu(\tau, q)$ for the left action and $q.\tau$ instead of $\mu(q, \tau)$ for the right action.

**Definition 56.** *We say that a Lie group $\mathbb{G}$ acts freely on the right on a smooth manifold $M$ if for all $q \in M$, $q.\tau = q.\varsigma$ if and only if $\tau = \varsigma$. Equivalently: if there exists $q \in M$ such that $q.\tau = q$ (that is, if $\tau$ has at least one fixed point), then $\tau$ is the identity element $e \in \mathbb{G}$.*

**Definition 57.** *We say that $\mathbb{G}$ acts transitively on the right on $M$ if for any $q, p \in M$ there is $\tau \in \mathbb{G}$ such that $q.\tau = p$.*

The same definitions can be given for the left action of the group $\mathbb{G}$ on a manifold $M$.

**Example 11.**

1. The flow on $M$ defined in Definition 44 is an example of the action of the group $(\mathbb{R}, +)$ on a smooth manifold $M$.
2. The action $\mu\colon GL(n) \times \mathbb{R}^n \to \mathbb{R}^n$ of the general linear group in $\mathbb{R}^n$ is defined as a product of an $(n \times n)$-matrix by an $n$-vector written as a column (or $(n \times 1)$-matrix).

3. Let $S^{n-1} = \{x \in \mathbb{R}^n |\ \|x\|_E = 1\}$. The action $\mu\colon O(n) \times S^{n-1} \to S^{n-1}$ is defined as a product of an $(n \times n)$-matrix from $O(n)$ by a vector from $\mathbb{R}^n$ of the length one.
4. The multiplication law in any Lie group $\mathbb{G}$ produces two actions on itself: *left and right translations*. Recall that the action
$$\begin{array}{rccc} l\colon & \mathbb{G}\times\mathbb{G} & \to & \mathbb{G} \\ & (\tau,\upsilon) & \mapsto & \tau\upsilon, \end{array}$$
or $l(\tau,\upsilon) = l_\tau(\upsilon) = \tau.\upsilon := \tau\upsilon$, is the left translation and
$$\begin{array}{rccc} r\colon & \mathbb{G}\times\mathbb{G} & \to & \mathbb{G} \\ & (\upsilon,\tau) & \mapsto & \upsilon\tau, \end{array}$$
or $r(\upsilon,\tau) = r_\tau(\upsilon) = \upsilon.\tau := \upsilon\tau$ is the right translation.
5. Define a left action $a$ of a Lie group $\mathbb{G}$ on itself by
$$\begin{array}{rccc} a\colon & \mathbb{G}\times\mathbb{G} & \to & \mathbb{G} \\ & (\tau,\upsilon) & \mapsto & \tau\upsilon\tau^{-1}, \end{array} \tag{8.2}$$
or $a(\tau,\upsilon) = a_\tau(\upsilon) = \tau.\upsilon := \tau\upsilon\tau^{-1}$. This action is called the *action by conjugation* or the inner automorphism. This action produces other very interesting actions of $\mathbb{G}$ on its the Lie algebra $\mathfrak{g}$ and even an action of the Lie algebra $\mathfrak{g}$ over itself, see Example (14).
6. Define a left action of a group $\mathbb{G}$ on its tangent bundle $T\mathbb{G}$ by
$$\begin{array}{rccc} \mu\colon & \mathbb{G}\times T\mathbb{G} & \to & T\mathbb{G} \\ & \tau.(q,v_q) & \mapsto & \big(\tau q, d_q l_\tau(v_q)\big). \end{array} \tag{8.3}$$
The right action of $\mathbb{G}$ on $T\mathbb{G}$ is defined analogously.
7. Define the left action of a group $\mathbb{G}$ on its co-tangent bundle $T^*G$ by
$$\begin{array}{rccc} \mu\colon & \mathbb{G}\times T^*\mathbb{G} & \to & T^*\mathbb{G} \\ & \tau.(q,\omega_q) & \mapsto & \big(\tau q, (d_q l_\tau)^*(\omega_q)\big), \end{array} \tag{8.4}$$
where $(d_q l_\tau)^*$ the dual map to the differential $d_q l_\tau$. The right action of $\mathbb{G}$ on $T^*\mathbb{G}$ is defined analogously.
8. The Adjoint action of a Lie group $\mathbb{G}$ on its Lie algebra $\mathfrak{g}$ is defined by
$$\begin{array}{rccc} \mu\colon & \mathbb{G}\times\mathfrak{g} & \to & \mathfrak{g} \\ & (\tau,\xi) & \mapsto & \mathrm{Ad}_\tau(\xi) \end{array} \tag{8.5}$$
The definition of the adjoint map $\mathrm{Ad}_\tau\colon \mathfrak{g}\to\mathfrak{g}$ is given in Example 13. The adjoint action uses the notion of the action $a$ by conjugation (8.2) which is a left action. Therefore, the action Ad is an action on the left.
9. The co-adjoint action of a Lie group $\mathbb{G}$ on the dual to its Lie algebra $\mathfrak{g}^*$ is defined by
$$\begin{array}{rccc} \mu\colon & \mathbb{G}\times\mathfrak{g}^* & \to & \mathfrak{g}^* \\ & (\tau,\omega) & \mapsto & \mathrm{Ad}^*_\tau(\omega) \end{array} \tag{8.6}$$
See the definition of the co-adjoint map $\mathrm{Ad}^*_\tau\colon \mathfrak{g}^*\to\mathfrak{g}^*$ in (8.9).

10. An action on the left of the special Euclidean group $SE(n)$ over $\mathbb{R}^n$ is defined by
$$\begin{array}{rccc}\mu\colon & SE(n)\times\mathbb{R}^n & \to & \mathbb{R}^n\\ & (A,a).x & \mapsto & Ax+a.\end{array}$$

To proceed with examples of the action of a Lie group $\mathbb{G}$ on its Lie algebra $\mathfrak{g}$ (the underlying manifold is just the vector space $T_e\mathbb{G}$) we make some observations about the differential map at the identity of an action $\mu$ on $\mathbb{G}$. Let
$$\mu\colon \mathbb{G}\times\mathbb{G}\to\mathbb{G}.$$
Fix one variable $\tau\in\mathbb{G}$ and consider
$$\mu_\tau\colon \mathbb{G}\to\mathbb{G}$$
as a diffeomorphism of the group $\mathbb{G}$. Then the differential at $e\in\mathbb{G}$ is the linear map
$$d_e\mu_\tau\colon T_e\mathbb{G}\to T_{\mu_\tau(e)}\mathbb{G}=T_\tau\mathbb{G}.$$
Let us consider three examples of an action $\mu$: translations $r, l$, and the action by conjugation $a$.

**Example 12.** Let $\mu$ be the right translation $r\colon \mathbb{G}\times\mathbb{G}\to\mathbb{G}$, $r_\tau(q)=q.\tau$. We want to understand how $q.\tau$ changes with respect to the variable $\tau$ near $\tau=e$. This variation is called the *infinitesimal generator* of the right action $r_\tau(q)$ at point $q$ and is denoted by $\sigma_q$. To calculate $\sigma_q$ we observe that the map $q.\tau$ with fixed $q$ and variable $\tau$ is just the left translation $q.\tau=l_q(\tau)$, so
$$\sigma_q=d_el_q\colon T_e\mathbb{G}\to T_{l_q(e)}\mathbb{G}=T_q\mathbb{G},$$
or it is customary to write
$$\sigma_q\colon \mathfrak{g}\to T_q\mathbb{G}.$$
We conclude that the map $\sigma_q$ generates a left invariant vector field on $\mathbb{G}$, since it translates an element $\xi$ of the Lie algebra $\mathfrak{g}$ to a vector field $X_q=d_el_q(\xi)$ that will be left invariant by definition. In a practical way the map $\sigma_q$ is calculated by making use of the exponential curve by the following
$$\sigma_q(\xi)=\frac{d}{d\varepsilon}\Big|_{\varepsilon=0}q\exp(\varepsilon\xi)=\frac{d}{d\varepsilon}\Big|_{\varepsilon=0}l_q(\exp(\varepsilon\xi))=dl_qd\big(\exp(\varepsilon\xi)|_{\varepsilon=0}\big)=dl_q(\xi). \tag{8.7}$$

Analogously, the left translation $l$ has its infinitesimal generator, that is the map $\mathfrak{g}\to T_q\mathbb{G}$ generating right invariant vector fields. The formula corresponding to (8.7) is
$$\sigma_q(\xi)=\frac{d}{d\varepsilon}\Big|_{\varepsilon=0}\exp(\varepsilon\xi))q=\frac{d}{d\varepsilon}\Big|_{\varepsilon=0}r_q(\exp(\varepsilon\xi))=dr_q(\xi). \tag{8.8}$$

**Example 13.** Now we consider $\mu=a$,
$$a(\tau,q)=a_\tau(q)=\tau q\tau^{-1}\colon \mathbb{G}\to\mathbb{G},$$

which is not only a diffeomorphic map of the underlying manifold of the group $\mathbb{G}$, but it also preserves the group structure, so the map $a_\tau$ is a group automorphism and we write $a_\tau \in \operatorname{Aut}(\mathbb{G})$. The differential of $a_\tau$ at the identity is

$$d_e a_\tau \colon T_e\mathbb{G} \to T_{a_\tau(e)}\mathbb{G} = T_e\mathbb{G}.$$

From general group theory [87, 123, 131] it is known that $d_e a_\tau$ preserves the Lie algebra structure of the vector space $T_e\mathbb{G}$ (since $a_\tau$ preserves the Lie group structure). So we can write

$$d_e a_\tau \colon \mathfrak{g} \to \mathfrak{g}$$

and conclude that it is an automorphism of Lie algebras. It is denoted by $\operatorname{Ad}_\tau := d_e a_\tau$ and is called the *adjoint map at* $\tau \in \mathbb{G}$. Thus $\operatorname{Ad}_\tau(\xi) \in \mathfrak{g}$ for any $\xi \in \mathfrak{g}$, or

$$\operatorname{Ad}_\tau \colon \mathfrak{g} \to \mathfrak{g}, \qquad \operatorname{Ad}_\tau \in \operatorname{Aut}(\mathfrak{g}).$$

Now let the variable $\tau$ vary and consider the adjoint map

$$\operatorname{Ad} \colon \mathbb{G} \to \operatorname{Aut}(\mathfrak{g})$$

as a homomorphism of groups, where to the product in $\mathbb{G}$ there corresponds a superpositions of linear maps in $\operatorname{Aut}(\mathfrak{g})$. This map is also called the *adjoint representation of the group* $\mathbb{G}$ *on its Lie algebra* $\mathfrak{g}$, or the *adjoint action of the group* $\mathbb{G}$ *on its Lie algebra* $\mathfrak{g}$.

Let the adjoint map $\operatorname{Ad}_\tau \colon \mathfrak{g} \to \mathfrak{g}$ at $\tau \in \mathbb{G}$ be given, and let $\langle \cdot\,, \cdot \rangle$ be the pairing between the Lie algebra $\mathfrak{g}$ and its dual $\mathfrak{g}^*$. The *co-adjoint map* or the dual representation $\operatorname{Ad}^*_\tau \colon \mathfrak{g}^* \to \mathfrak{g}^*$ at $\tau \in \mathbb{G}$ is defined by

$$\langle \operatorname{Ad}^*_\tau(\omega), \xi \rangle := \langle \omega, \operatorname{Ad}_{\tau^{-1}}(\xi) \rangle, \qquad \xi \in \mathfrak{g}, \quad \omega \in \mathfrak{g}^*. \tag{8.9}$$

**Example 14.** Since the map Ad sends $e \in \mathbb{G}$ to $\operatorname{Id} \in \operatorname{Aut}(\mathfrak{g})$, the differential $d_e \operatorname{Ad}$ at $e \in \mathbb{G}$ is the linear map

$$d_e \operatorname{Ad} \colon T_e\mathbb{G} \to T_{\operatorname{Id}} \operatorname{Aut}(\mathfrak{g})$$

and, moreover, it is a homomorphism of the Lie algebra $\mathfrak{g} = (T_e\mathbb{G}, [\cdot\,, \cdot])$ and the Lie algebra $\operatorname{End}(\mathfrak{g}) = (T_{\operatorname{Id}} \operatorname{Aut}(\mathfrak{g}), [\cdot\,, \cdot])$ of all linear transformations of $\mathfrak{g}$, where the Lie brackets are defined through the composition of linear maps from $\operatorname{End}(\mathfrak{g})$. The map $d_e \operatorname{Ad}$ is denoted by ad and is called the *adjoint representation of the Lie algebra* $\mathfrak{g}$ *over itself*, or the *adjoint action of the Lie algebra* $\mathfrak{g}$ *over itself.* The construction is reflected beautifully in the following commutative diagram:

$$\begin{tikzcd}
\mathbb{G} \arrow[r, "\operatorname{Ad}"] & \operatorname{Aut}(\mathfrak{g}) \\
\mathfrak{g} \arrow[u, "\exp_{\mathbb{G}}"] \arrow[r, "\operatorname{ad}"'] & \operatorname{End}(\mathfrak{g}) \arrow[u, "\exp"']
\end{tikzcd}$$

or, in other words,

$$\operatorname{Ad}_{\exp_{\mathbb{G}}(\xi)} = \exp(\operatorname{ad}_\xi), \quad \xi \in \mathfrak{g}, \quad \exp_{\mathbb{G}}(\xi) \in \mathbb{G}.$$

Notice that $\exp_{\mathbb{G}}$ in the left hand side is the exponential map from the Lie algebra $\mathfrak{g}$ to its Lie group $\mathbb{G}$. The right hand side exp is the exponential map from the Lie algebra $\mathrm{End}(\mathfrak{g})$ to its Lie group $\mathrm{Aut}(\mathfrak{g})$.

An interesting feature of the map $\mathrm{ad} \in \mathrm{End}(\mathfrak{g})$ is the following. If we fix $\xi \in \mathfrak{g}$, then

$$\mathrm{ad}_{\xi} \colon \mathfrak{g} \to \mathfrak{g}$$

is the map given by $\mathrm{ad}_{\xi}\, \eta = [\xi, \eta]$, where $[\cdot\,,\cdot]$ are Lie brackets on $\mathfrak{g}$, see [87, 131].

Notice here also the relation between the adjoint action $\mathrm{Ad}_{\tau}$ of the group $\mathbb{G}$ on its Lie algebra $\mathfrak{g}$, the action $a_{\tau}$ on the group $\mathbb{G}$ by conjugation, and the exponential map $\exp_{\mathbb{G}}$ reflected in the following commutative diagram:

$$\begin{tikzcd}
\mathbb{G} \arrow[r, "a_{\tau}"] & \mathbb{G} \\
\mathfrak{g} \arrow[u, "\exp_{\mathbb{G}}"] \arrow[r, "\mathrm{Ad}_{\tau}"'] & \mathfrak{g} \arrow[u, "\exp_{\mathbb{G}}"']
\end{tikzcd}$$

or in other words $\tau \exp(\xi)\tau^{-1} = \exp(\mathrm{Ad}_{\tau}\, \xi), \quad \xi \in \mathfrak{g}, \quad \tau \in \mathbb{G}$.

**Example 15.** [81] The co-adjoint action of a diffeomorphism $f \in \mathrm{Diff}\, S^1$ on the dual $\mathfrak{vir}^*$ of the Virasoro algebra is defined by the following formula

$$\begin{array}{lccc}
\mathrm{Ad}^*\colon & \mathrm{Diff}\, S^1 \times \mathfrak{vir}^* & \to & \mathfrak{vir}^* \\
 & f(\theta).\big(u(\theta)(d\theta^2), a\big) & \mapsto & \Big(\big(u(f)\cdot(f')^2(d\theta)^2 + aS(f)\big)(d\theta)^2, a\Big),
\end{array}$$

where

$$S(f) = \frac{f'f'' - 3/2(f'')^2}{(f')^2}$$

is the so-called *Schwarzian derivative* of the diffeomorphism $f$.

The following interesting observation concerns the tangent bundle of a group $\mathbb{G}$. The action, right or left, of a group $\mathbb{G}$ on itself induces an action on the tangent bundle $T\mathbb{G}$ making the last one into a group.

**Proposition 19.** [108] *The product map induces a smooth associative map*

$$\widetilde{\rho}\colon T(\mathbb{G}\times\mathbb{G}) \sim T\mathbb{G}\times T\mathbb{G} \to T\mathbb{G},$$

*that makes the tangent bundle of a Lie group $\mathbb{G}$ into a Lie group $T\mathbb{G}$.*

*Proof.* Let $(\tau, v_{\tau})$ and $(g, v_g)$ be two points on $T\mathbb{G}$, then we define the multiplication law $\widetilde{\rho}\colon T\mathbb{G}\times T\mathbb{G} \to T\mathbb{G}$ by

$$\widetilde{\rho}\Big((\tau, v_{\tau}), (g, v_g)\Big) = (\tau g, v_{\tau} \star v_g) = (\tau g, v_{\tau g}),$$

where the vector $v_{\tau g} \in T_{\tau g}\mathbb{G}$ is obtained in the following way. Consider smooth curves

$$\gamma_{\tau, v_{\tau}}\colon [-1,1] \to \mathbb{G}, \quad \gamma_{\tau, v_{\tau}}(0) = \tau, \quad \dot{\gamma}_{\tau, v_{\tau}}(0) = v_{\tau},$$

and

$$\gamma_{g, v_g}\colon [-1,1] \to \mathbb{G}, \quad \gamma_{g, v_g}(0) = g, \quad \dot{\gamma}_{g, v_g}(0) = v_g.$$

Then the product $\gamma_{\tau,v_\tau}(t) \star \gamma_{g,v_g}(t)$ in $\mathbb{G}$ is defined for any $t \in [-1,1]$ and defines a curve

$$\gamma_{\tau g,v_{\tau g}} \colon [-1,1] \to \mathbb{G}, \quad \gamma_{\tau g,v_{\tau g}}(0) = \tau g, \quad \dot{\gamma}_{\tau g,v_{\tau g}}(0) = v_{\tau g} := v_\tau \star v_g.$$

So the product $v_\tau \star v_g$ is the initial vector velocity of the product curve $\gamma_{\tau g,v_{\tau g}}$ obtained by multiplication of $\gamma_{\tau,v_\tau}$ by $\gamma_{g,v_g}$. □

The natural projection $\mathrm{pr} \colon T\mathbb{G} \to \mathbb{G}$ induces a group homomorphism

$$\widetilde{\mathrm{pr}} \colon (T\mathbb{G}, \widetilde{\rho}) \to (\mathbb{G}, \rho).$$

The kernel of this homomorphism is isomorphic to the abelian additive group of tangent vectors at the identity of the group $T_e\mathbb{G}$. In other words, there is a short exact sequence

$$0 \longrightarrow T_e\mathbb{G} \xrightarrow{\sigma} T\mathbb{G} \xrightarrow{\widetilde{\mathrm{pr}}} \mathbb{G} \longrightarrow 1$$

of smooth group homomorphisms. Here $\sigma$ is the infinitesimal generator of the right action of the group $\mathbb{G}$ on itself that associates a left invariant vector field to any vector $\xi \in T_e\mathbb{G}$. Let us define a map

$$z \colon \mathbb{G} \to T\mathbb{G}$$

that to any element $\tau$ associates $(\tau, \overrightarrow{0}_\tau) \in T_\tau\mathbb{G}$, i. e., the null section at tangent space $T_\tau\mathbb{G}$. Then the composition $\mathbb{G} \xrightarrow{z} T\mathbb{G} \xrightarrow{\mathrm{pr}} \mathbb{G}$ is the identity map on $\mathbb{G}$.

This means that the group $T\mathbb{G}$ can be decomposed as a semi-direct product

$$T\mathbb{G} = T_e\mathbb{G} \rtimes \mathbb{G},$$

where $T_e\mathbb{G}$ is a normal subgroup on $T\mathbb{G}$: $(e,\xi) \in T_e\mathbb{G} \subset T\mathbb{G}$, and $\mathbb{G}$ is a subgroup of $T\mathbb{G}$: $\mathbb{G} \ni \tau \overset{z}{\mapsto} (\tau, \overrightarrow{0}_\tau) \in T\mathbb{G}$. The semi-direct product can be written by making use of a homomorphism $h \colon \mathbb{G} \to \mathrm{Aut}(T_e\mathbb{G})$ [87] by

$$(\xi,\tau)(\eta,g) \mapsto (\xi h_\tau(\eta), \tau g),$$

where $h_\tau \colon T_e\mathbb{G} \to T_e\mathbb{G}$ is the action of the subgroup $\mathbb{G} \subset T\mathbb{G}$ over the normal subgroup $T_e\mathbb{G} \subset T\mathbb{G}$ [87]. Since the elements of $T_e\mathbb{G}$ can be considered as elements of a subgroup $T_e\mathbb{G}$ of the group $T\mathbb{G}$ and they are also elements of $\mathfrak{g}$, then it can be shown that Ad is the suitable homomorphism. Thus the semi-direct product can be written as

$$(\xi,\tau)(\eta,g) \mapsto (\xi \, \mathrm{Ad}_\tau(\eta), \tau g),$$

where the product of $\xi$ by $\mathrm{Ad}_\tau(\eta)$ is considered as the product in the normal subgroup of the group $T\mathbb{G}$.

**Definition 58.** *Let $\mathbb{G}$ be a group and $A, B$ be two sets where the group $\mathbb{G}$ acts. A map $F \colon A \to B$ is said to be equivariant, if*

$$\begin{array}{rclll} F(\tau.q) & = & \tau.F(q) & \textit{for} & \textit{left-left action}, \\ F(q.\tau) & = & F(q).\tau & \textit{for} & \textit{right-right action}, \\ F(q.\tau) & = & \tau^{-1}.F(q) & \textit{for} & \textit{right-left action}, \\ F(\tau.q) & = & F(q).\tau^{-1} & \textit{for} & \textit{left-right action}, \end{array}$$

*for all $\tau \in \mathbb{G}$ and all $q \in A$.*

The definition says that an equivariant map is a map that commutes with the action of the group in the domain of definition and on the target space. As we can see, the definition depends on whether right or left action is chosen on the domain of definition and on the target space of $F$. Equivariant maps are also known as $\mathbb{G}$-maps or $\mathbb{G}$-homomorphisms.

Exercises

1. We present here one more point of view on the map ad. Let us fix any vector $\xi \in \mathfrak{g}$ and see the difference between $\xi$ and the result of adjoint action Ad of $\mathbb{G}$ on $\xi \in \mathfrak{g}$:
$$\mathrm{Ad}_\tau \xi - \xi.$$
Then the differential of the map $\mathbb{G} \ni \tau \mapsto \mathrm{Ad}_\tau \xi - \xi \in \mathfrak{g}$ at the identity $\tau = e$ is denoted by $\mathrm{ad}_\xi$ and it is a linear map $\mathrm{ad}_\xi \colon T_e\mathbb{G} \to T_e\mathfrak{g}$. After identifications $T_e\mathbb{G} \sim T_e\mathfrak{g} \sim \mathfrak{g}$ and proving the correspondence of Lie brackets, we come to the previous definition of $\mathrm{ad}_\xi \colon \mathfrak{g} \to \mathfrak{g}$.
Show that $ad_\xi(\eta)$ is bilinear and satisfies $\mathrm{ad}_\xi(\eta) = -\mathrm{ad}_\eta(\xi)$.
2. Verify that the pair $(T\mathbb{G}, \widetilde{\rho})$ from Proposition 19 satisfies the definition of a Lie group.
3. Show that $T_e\mathbb{G}$ is a normal subgroup of $T\mathbb{G}$ and $\mathbb{G}$ is the complementary subgroup.
4. Show that $\mathrm{ad}_\xi(\eta) = [\xi, \eta]$.
5. Check that the dual representation $\mathrm{Ad}^* \colon \mathbb{G} \to \mathrm{Aut}(\mathfrak{g}^*)$ from (8.9) is a group homomorphism.
6. Define the co-adjoint map $\mathrm{ad}^*_\xi \colon \mathfrak{g}^* \to \mathfrak{g}^*$ by $\langle \mathrm{ad}^*_\xi \omega, \eta\rangle = -\langle \omega, \mathrm{ad}_\xi \eta\rangle$, $\xi, \eta \in \mathfrak{g}$, $\omega \in \mathfrak{g}^*$. Verify that the co-adjoint map $\mathrm{ad}^* \colon \mathfrak{g} \to \mathrm{End}(\mathfrak{g}^*)$ is an algebra homomorphism.

## 8.4. Complexifications

Here we present definitions of complexifications of real vector spaces, complex and CR structures on manifolds and Lie groups, including the infinite dimensional case.

**8.4.1. Complexification of a real vector space.** A *complexification* of a real vector space $V$ is the tensor product $V \otimes \mathbb{C}$ over $\mathbb{R}$, where the generators are $v \otimes 1$ and $v \otimes i$, $v \in V$. So, $V \otimes \mathbb{C}$ are all possible linear combinations of $v \otimes 1$ and $v \otimes i$, $v \in V$ with real coefficients, modulo the equivalence relations
$$(v_1 + v_2) \otimes z \sim v_1 \otimes z + v_2 \otimes z,$$
$$v \otimes (z_1 + z_2) \sim v \otimes z_1 + v \otimes z_2,$$
$$av \otimes z \sim v \oplus az, \quad a \in \mathbb{R}.$$
The real dimension of $V \otimes \mathbb{C}$ is $2 \dim V$. The multiplication by complex numbers is defined by
$$\alpha(v \otimes z) = v \otimes \alpha z, \quad \text{for} \quad \alpha, z \in \mathbb{C}, \quad \text{and} \quad v \in V.$$
It makes the space $V \otimes \mathbb{C}$ into a complex vector space of complex dimension $\dim V$. The generators for complex vector space $V \otimes \mathbb{C}$ are $v \otimes 1$ and $v \otimes i$. The real space

$V$ is naturally imbedded into $V \otimes \mathbb{C}$ by identifying $V$ with the space $V \otimes 1$ (any element $v \in V$ is identified with the element $v \otimes 1 \in V \otimes \mathbb{C}$). The *conjugation* for $V \otimes \mathbb{C}$ is defined by

$$\overline{v \otimes z} := v \oplus \bar{z}.$$

As an application we consider a complexification of a smooth real manifold $M$ of real dimension $n$. For any $q \in M$, the complex vector space $T_qM \otimes \mathbb{C}$ is called the *complexified tangent space* and $T_q^*M \otimes \mathbb{C}$ is called the *complexified co-tangent space*. The complex space $T_q^*M \otimes \mathbb{C}$ can also be regarded as the complex dual space of $T_qM \otimes \mathbb{C}$ by defining the pairing

$$\langle v \otimes z, \xi \oplus w\rangle := \langle v, \xi\rangle zw, \quad \text{for} \quad v \in T_qM,\ \xi \in T_q^*M,\ z, w \in \mathbb{C},$$

for any point $q \in M$. The *complexified tangent bundle* is $T^{\mathbb{C}}M = \cup_{q\in M}\big(T_qM \otimes \mathbb{C}\big)$ and the *complexified co-tangent bundle* is $T^{*\mathbb{C}}M = \cup_{q\in M}\big(T_q^*M \otimes \mathbb{C}\big)$. A *complexified vector field* $L$ on $M$ is a smooth section of $T^{\mathbb{C}}M$, which means that $L$ assigns to each $q \in M$ a vector $L_q \in T_qM \otimes \mathbb{C}$. In any smooth coordinate system $\big(U, \varphi = (x^1, \dots, x^n)\big)$ we can express $L$ as

$$L_q = \sum_{j=1}^{n} L^j(q)\partial_{x^j},$$

where $L^j$, $j = 1, \dots, n$ are smooth, complex valued functions defined on $U \subset M$.

If $M$ is complex manifold of complex dimension $n$, then it is important to distinguish between the real tangent bundle and the complexified tangent bundle. The real tangent bundle $TM$ corresponds to a smooth manifold $M$ of real dimension $2n$. Its fiber $T_qM$ is a real vector space and has real dimension $2n$. The fiber $T_qM \otimes \mathbb{C}$ of the complexified tangent bundle is a complex space of complex dimension $2n$.

**8.4.2. Complex structures.** If the real vector space $V$ is of even dimension, then it is possible to define an *almost complex structure* $J$, that is, a map $J\colon V \to V$, such that $J^2 = -\operatorname{Id}_V$.

**Example 16.** Let $V = T_q\mathbb{R}^{2n} \cong \mathbb{C}^n$. Take the coordinates $q = (x_1, y_1, \dots, x_n, y_n)$. The *standard almost complex* structure for $T_q\mathbb{R}^{2n}$ is defined by setting

$$J(\partial_{x^j}) = \partial_{y^j}, \qquad J(\partial_{y^j}) = -\partial_{x^j}, \qquad j = 1, \dots, n, \tag{8.10}$$

on the standard basis. Then $J$ extends by linearity to all $T_q\mathbb{R}^{2n}$. This almost complex structure is designed to simulate the multiplication by $i = \sqrt{-1}$.

*The standard almost complex structure $J^*$ on the co-tangent space $T_q^*\mathbb{R}^{2n}$ is the following*

$$J^*(dx^j) = -dy^j \qquad J^*(dy^j) = -dx^j, \qquad j = 1, \dots, n.$$

An almost complex structure can be defined on a real tangent space of a complex manifold $M$ by pushing forward the complex structure from $\mathbb{C}^n$ up to $M$

via a coordinate chart. For $q \in M$ and a holomorphic chart $(U, \zeta)$, $\zeta\colon U \to \mathbb{C}^n$, we define $J_q\colon T_qM \to T_qM$ by

$$J_q(L) := d_{\zeta(q)}\zeta^{-1} J(d_q\zeta(L)), \tag{8.11}$$

where $J$ in the right hand side is the standard almost complex structure in $\mathbb{C}^n$. The definition implies that if $\zeta = (z_1, \dots, z_n)$, $z_j = x^j + iy^j$, then $J_q(\partial_{x^j}) = \partial_{y^j}$ and $J_q(\partial_{y^j}) = -\partial_{x^j}$.

If $J$ is an almost complex structure on a real vector space $V$, then we can extend it to an almost complex structure $J_{\mathbb{C}}$ on the complexification $V \otimes \mathbb{C}$ by setting

$$J_{\mathbb{C}}(v \otimes z) := J(v) \otimes z, \quad v \in V, \quad z \in \mathbb{C}.$$

Then

$$J_{\mathbb{C}}(\overline{w}) = \overline{J_{\mathbb{C}} w}, \quad \text{for} \quad w \in V \otimes \mathbb{C}. \tag{8.12}$$

The linear map $J_{\mathbb{C}}$ has two eigenvalues $i$ and $-i$, since $J_{\mathbb{C}}^2 = -\operatorname{Id}_{V\otimes\mathbb{C}}$. The corresponding eigen spaces are denoted by $V^{(1,0)}$ and $V^{(0,1)}$. Thus we have

$$V \otimes \mathbb{C} = V^{(1,0)} \oplus V^{(0,1)}$$

from linear algebra. The property (8.12) implies $\overline{V^{(1,0)}} = V^{(0,1)}$. Let us construct bases for $V^{(1,0)}$ and $V^{(0,1)}$. First we observe that $v$ and $Jv$ are linearly independent over $\mathbb{R}$ in $V$, since $J$ has no real eigenvalues. Then

$$\{v_1 - iJv_1, \dots, v_n - iJv_n\} \tag{8.13}$$

is a basis for the complex $n$-dimensional vector space $V^{(1,0)}$ and

$$\{v_1 + iJv_1, \dots, v_n + iJv_n\} \tag{8.14}$$

is a basis for the complex $n$-dimensional vector space $V^{(0,1)}$. Recall, that $\dim V = 2n$.

Let see how it works for a complex $n$-dimensional manifold $M$. Let $(z_1, \dots, z_q)$ with $z_j = x^j + iy^j$ be a set of local holomorphic coordinates and the almost complex structure on $T_qM$, $q \in M$, is given by (8.11). Define the vector fields

$$\partial_{z_j} = \frac{1}{2}(\partial_{x^j} - i\partial_{y^j}) \qquad \partial_{\bar{z}_j} = \frac{1}{2}(\partial_{x^j} + i\partial_{y^j}), \quad j = 1, \dots, n.$$

Then in view of the above discussions, a basis for $T_q^{(1,0)}M$ is given by $\{\partial_{z_1}, \dots, \partial_{z_n}\}$ and a basis for $T_q^{(0,1)}M$ is given by $\{\partial_{\bar{z}_1}, \dots, \partial_{\bar{z}_n}\}$. Due to the form of the bases the spaces $T_q^{(1,0)}M$ and $T_q^{(0,1)}M$ received the names *holomorphic and antiholomorphic* tangent vector spaces. The Hermitian inner product on $T_qM \otimes \mathbb{C}$ is defined by declaring that $\{\partial_{z_1}, \dots, \partial_{z_n}, \partial_{\bar{z}_1}, \dots, \partial_{\bar{z}_n}\}$ is an orthonormal basis.

Let $M$ now be a real manifold, such that at each $q \in M$ the tangent space $T_qM$ admits an almost complex structure $J_q\colon T_qM \to T_qM$. Then it leads to the splitting $T^{\mathbb{C}}M = T^{(1,0)}M \oplus T^{(0,1)}M$ into the holomorphic and antiholomorphic bundles, and each of them is naturally isomorphic to the real tangent bundle of $M$, but now they are equipped with an additional structure $J_q$. If $T^{(1,0)}M$ is integrable, that is, $[T^{(1,0)}M, T^{(1,0)}M] \subset T^{(1,0)}M$, then the pair $(M, T^{(1,0)}M)$ is called a complex manifold.

**8.4.3. Lie groups, Lie algebras and complexification.** Let us impose a Lie algebra structure on $V$ and see how one can define a complexification $\mathfrak{g}\otimes\mathbb{C}$ of the Lie algebra $\mathfrak{g} = (V, [\cdot\,,\cdot])$. All that we need is to define the Lie bracket

$$[v\otimes\alpha, u\otimes\beta] := [v,u]\otimes\alpha\beta, \qquad v,u\in\mathfrak{g}, \quad \alpha,\beta\in\mathbb{C}. \tag{8.15}$$

Next we consider the relation between the almost complex structure and the Lie algebra structure. Let $\mathbb{G}$ be a Lie group and $\mathfrak{g}$ be its Lie algebra. Let $J\colon T_e\mathbb{G}\to T_e\mathbb{G}$ be an almost complex structure. It determines the splitting $T_e\mathbb{G}\otimes\mathbb{C} = \mathfrak{g}\otimes\mathbb{C} = \mathfrak{g}^{(1,0)}\oplus\mathfrak{g}^{(0,1)}$. If the subspace $\mathfrak{g}^{(1,0)}$ is a Lie subalgebra of $\mathfrak{g}\otimes\mathbb{C}$, then the pair $(\mathbb{G}, \mathfrak{g}^{(1,0)})$ is called a left invariant complex structure. This structure is also right invariant, if $\mathfrak{g}^{(1,0)}$ is adjoint invariant with respect to the adjoint action of the group $\mathbb{G}$, or $\mathrm{Ad}_\tau\,\xi\in\mathfrak{g}^{(1,0)}$ for all $\xi\in\mathfrak{g}^{(1,0)}$ and $\tau\in\mathbb{G}$.

We define now a CR-structure on a real manifold $N$. We follow a scheme, that is suitable for finite and infinite dimensional manifolds. Let $TN$ be the tangent bundle and $D$ be a corank 1 smooth sub-bundle, where an almost complex structure $J_q\colon D_q\to D_q$ is defined. Let $TN\otimes\mathbb{C}$ be the complexified tangent bundle, then $D\otimes\mathbb{C}$ is a complex corank 1 smooth sub-bundle, where the splitting $D\otimes\mathbb{C} = D^{(1,0)}\oplus D^{(0,1)}$ is defined. If $[D^{(1,0)}, D^{(1,0)}]\subset D^{(1,0)}$, then $D^{(1,0)}$ is called an integrable CR structure and the pair $(N, D^{(1,0)})$ is an integrable CR manifold.

**Example 17.** Let $N$ be a real hypersurface in a complex manifold $(M, T^{(1,0)}M)$. Define

$$D^{(1,0)} = T^{(1,0)}M|_N \cap TM\otimes\mathbb{C}.$$

Then $(N, D^{(1,0)})$ is a CR manifold with the CR structure $D^{(1,0)}$ inherited from the complex manifold $(M, T^{(1,0)}M)$. The manifolds $S^3\subset\mathbb{C}^2$ and the boundary of the Siegel upper half space in $\mathbb{C}^2$ have CR structures induced from $\mathbb{C}^2$.

A CR manifold is *strongly pseudo-convex* if $[L,\bar{L}]_q\notin D_q^{(1,0)}\oplus D_q^{(0,1)}$ for any local non-vanishing section $L$ of $D^{(1,0)}$.

If we have a Lie group $\mathbb{G}$ with a Lie algebra $\mathfrak{g}$, then a left invariant $CR$ structure is defined by a splitting $\mathfrak{h}^{(1,0)}\oplus\mathfrak{h}^{(0,1)}$ of a complex co-rank 1 subspace $\mathfrak{g}\otimes\mathbb{C}$ with subalgebras $\mathfrak{h}^{(1,0)}$ and $\mathfrak{h}^{(0,1)} = \overline{\mathfrak{h}^{(1,0)}}$. This structure is strongly pseudoconvex if $[\xi,\bar{\xi}]\notin\mathfrak{h}^{(1,0)}\oplus\mathfrak{h}^{(0,1)}$ holds for any non-zero $\xi\in\mathfrak{h}^{(1,0)}$.

Exercises

1. Show that the standard almost complex structure (8.10) in $T_q\mathbb{R}^{2n}$ is an isometry in $\mathbb{R}^{2n}$.
2. Show that the description of $J$ given in (8.11) does not depend on the choice of coordinate chart. Conclude that the push forward of the standard almost complex structure $J$ from $\mathbb{C}^n$ to a complex manifold is well defined.
3. Prove (8.12).
4. Show that $v - iJv\in V^{(1,0)}$ and $v + iJv\in V^{(0,1)}$ for any $v\in V$.

5. Prove that (8.13) and (8.14) are linearly independent systems.
6. Find the dual basis for $\{\partial_{z_1}, \dots, \partial_{z_n}, \partial_{\bar{z}_1}, \dots, \partial_{\bar{z}_n}\}$ with respect to the standard Hermitian product.
7. Verify that the Lie bracket defined by (8.15) is $\mathbb{C}$-linear, skew symmetric and satisfies the Jacobi identity.

### 8.5. Fiber bundles

**Definition 59.** *A fiber bundle is a collection $(F, E, B, \pi)$, where in general $F, E, B$ are topological spaces and $\pi$ is a continuous map $\pi\colon E \to B$. It can also be written as*

$$F \longrightarrow E \xrightarrow{\pi} B \qquad \text{or shortly} \qquad E \xrightarrow{\pi} B.$$

*It is required that for any $x \in E$, there is an open neighborhood $U \subset B$ of $\pi(x)$ (which will be called a trivializing neighborhood) such that $\pi^{-1}(U)$ is homeomorphic to the product space $F \times U$, in such a way that $\pi$ carries over to the projection onto the second factor. In other words, the following diagram should commute:*

$$\begin{tikzcd}
\pi^{-1}(U) \arrow[r, "\phi"] \arrow[d, "\pi"'] & F \times U, \arrow[dl, "\mathrm{pr}_2"] \\
U &
\end{tikzcd}$$

*where the map $\mathrm{pr}_2\colon F \times U \to U$ is the natural projection on the second coordinate and $\phi\colon \pi^{-1}(U) \to F \times U$ is a homeomorphism. The set of all $(U_i, \phi_i)$ is called a local trivialization of the bundle.*

Thus for any $b \in B$, the pre-image $\pi^{-1}(b)$ is homeomorphic to $F$ and is called the fiber over $b$. The set $B$ is called the base space, $E$ is the total space and the map $\pi$ is the projection map. Every fiber bundle $\pi\colon E \to B$ is an open map, since projection $\mathrm{pr}_2$ is an open map. Therefore, $B$ carries the quotient topology determined by the map $\pi$.

**Definition 60.** *A bundle $\pi'\colon E' \to B'$ is a sub-bundle of the bundle $\pi\colon E \to B$, provided $E'$ is a subspace of $E$, $B'$ is a subspace of $B$ and $\pi' = \pi|_{E'}\colon E' \to B'$.*

If we speak about smooth fiber bundle, we require that $F, E, B$ are smooth manifolds and all other maps are smooth. A smooth sub-bundle $(F', E', B, \pi|_{E'})$ of a bundle $(F, E, B, \pi)$ is a smooth bundle such that the inclusions

$$j_F\colon F' \to F, \quad j_E\colon E' \to E$$

are smooth.

**Definition 61.** *Let $\pi'\colon E' \to B$ and $\pi\colon E \to B$ be two bundles over the same base space $B$. A bundle map (bundle morphism) $u\colon \big(\pi'\colon E' \to B\big) \longrightarrow \big(\pi\colon E \to B\big)$ is a map $u\colon E' \to E$ such that $\pi' = \pi u$. The last equality is the requirement that the* *following diagram commutes*

$$\begin{tikzcd}
E' \arrow[r] \arrow[dr, "\pi'"'] & u \arrow[r] & E, \arrow[dl, "\pi"] \\
& B &
\end{tikzcd}$$

**Definition 62.** *The fiber product (direct sum or Whitney sum) of two bundles* $\pi'\colon E' \to B$ *and* $\pi\colon E \to B$ *over* $B$ *is the bundle* $\Pi\colon E' \oplus E \to B$*, where*

$$E' \oplus E = \{(q', q) \in E' \times E \ \mid \ \pi'(q') = \pi(q), \quad \textit{and} \ \ \Pi(q', q) = \pi'(q') = \pi(q)\}.$$

The fiber $\Pi^{-1}(b)$ of $\Pi\colon E' \oplus E \to B$ over $b \in B$ is

$$\pi'^{-1}(b) \times \pi^{-1}(b) \subset E' \times E.$$

**8.5.1. Frame bundle.** Let $(M, g_M)$ be an $n$-dimensional oriented Riemannian manifold. The frame bundle $\pi\colon F \to M$ is a fiber bundle, whose total space $F$ consists of collections $(q, v_1, \dots, v_n) \in M \times (T_qM)^n$ such that $g_M(v_i, v_j) = \delta_{ij}$. An element $(q, v_1, \dots, v_n)$ is an orthonormal basis of $T_qM$. Sections of the frame bundle are called orthonormal frame fields, and they are just assignments to any point $q \in M$ of some orthonormal basis $(v_1, \dots, v_n) \in T_qM$. If $SO(T_qM)$ is the group of orientation preserving isometries of the vector space $T_qM$, then there is a natural left action of $SO(T_qM)$ on the frame bundle $F$ given by

$$\begin{array}{rccc}
\mu\colon & SO(T_qM) \times F & \to & F \\
 & \tau(q, v_1, \dots, v_n) & \mapsto & (q, \tau.(v_1, \dots, v_n)),
\end{array}$$

where $\tau.(v_1, \dots, v_n)$ is just an isometrical transformation of the basis $(v_1, \dots, v_n)$. If $\tau$ is written as a matrix in the basis $(v_1, \dots, v_n)$, then it is a product of an $(n \times n)$-matrix by a $(n \times 1)$-column.

It is possible to think of the frame field (or just a frame) as a linear isomorphism $f_q\colon \mathbb{R}^n \to T_qM$ that assigns to any standard basic element

$$e_j = (0, \dots, 0, 1, 0, \dots, 0)$$

with 1 on $j$-th place of $\mathbb{R}^n$, the component $v_j$ of $(v_1, \dots, v_n)$. The map $f_q$ belongs to the space $SO(\mathbb{R}^n, T_qM)$ of all isometrical transformations from $\mathbb{R}^n$, with the standard Euclidean metric, to the vector space $T_qM$, endowed with some inner product. In this case it is possible to define the right action of the group $SO(n)$ of $F$ by

$$\begin{array}{rccc}
\mu\colon & F \times SO(n) & \to & F \\
 & (q, f).\tau & \mapsto & (q, f.\tau),
\end{array}$$

where $f.\tau = f \circ \tau = f(\tau)$ is the composition of the isometry in $\mathbb{R}^n$ and then the map $f_q$. It gives the principal $SO(n)$-bundle structure for the frame bundle $F$. Notice that there is no natural left action of $SO(n)$ on the fiber over $q \in M$, but only the action of $SO(T_qM)$. The group $SO(T_qM)$ is not canonically isomorphic to $SO(n)$ when $n \geq 3$.

## 9. Appendix B

The Table 3 represents products of unit octonions that were used in the construction of the octonion $\mathbb{H}$-type group $\mathbb{H}^1_7$.

| | $j_1$ | $j_2$ | $j_3$ | $j_4$ | $j_5$ | $j_6$ | $j_7$ |
|---|---|---|---|---|---|---|---|
| $j_1$ | $-1$ | $-j_3$ | $j_2$ | $-j_5$ | $j_4$ | $j_7$ | $-j_6$ |
| $j_2$ | $j_3$ | $-1$ | $-j_1$ | $-j_6$ | $-j_7$ | $j_4$ | $j_5$ |
| $j_3$ | $-j_2$ | $j_1$ | $-1$ | $-j_7$ | $j_6$ | $-j_5$ | $j_4$ |
| $j_4$ | $j_5$ | $j_6$ | $j_7$ | $-1$ | $-j_1$ | $-j_2$ | $-j_3$ |
| $j_5$ | $-j_4$ | $j_7$ | $-j_6$ | $j_1$ | $-1$ | $j_7$ | $-j_6$ |
| $j_6$ | $-j_7$ | $-j_4$ | $j_5$ | $j_2$ | $-j_7$ | $-1$ | $j_5$ |
| $j_7$ | $j_6$ | $-j_5$ | $-j_4$ | $j_3$ | $j_6$ | $-j_5$ | $-1$ |

Table 3. Multiplication table of unit octonions $j_m$

The precise forms of the matrices $\mathbf{J}_m$ for the product in the octonion $\mathbb{H}$-type group $\mathbb{H}^1_7$ are given below.

$$\mathbf{J}_1=\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{J}_2=\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

$$\mathbf{J}_3=\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix} \mathbf{J}_4=\begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{J}_5=\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{J}_6=\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{J}_7 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Vector fields on $S^7$. Table 4 gives products of unit octonions and allows to calculate the product of two arbitrary octonions. It was used to calculate the vector fields on $S^7$

| | $j_0$ | $j_1$ | $j_2$ | $j_3$ | $j_4$ | $j_5$ | $j_6$ | $j_7$ |
|---|---|---|---|---|---|---|---|---|
| $j_0$ | $j_0$ | $j_1$ | $j_2$ | $j_3$ | $j_4$ | $j_5$ | $j_6$ | $j_7$ |
| $j_1$ | $j_1$ | $-j_0$ | $j_3$ | $-j_2$ | $j_5$ | $-j_4$ | $-j_7$ | $j_6$ |
| $j_2$ | $j_2$ | $-j_3$ | $-j_0$ | $j_1$ | $j_6$ | $j_7$ | $-j_4$ | $-j_5$ |
| $j_3$ | $j_3$ | $j_2$ | $-j_1$ | $-j_0$ | $j_7$ | $-j_6$ | $j_5$ | $-j_4$ |
| $j_4$ | $j_4$ | $-j_5$ | $-j_6$ | $-j_7$ | $-j_0$ | $j_1$ | $j_2$ | $j_3$ |
| $j_5$ | $j_5$ | $j_4$ | $-j_7$ | $j_6$ | $-j_1$ | $-j_0$ | $-j_3$ | $j_2$ |
| $j_6$ | $j_6$ | $j_7$ | $j_4$ | $-j_5$ | $-j_2$ | $j_3$ | $-j_0$ | $-j_1$ |
| $j_7$ | $j_7$ | $-j_6$ | $j_5$ | $j_4$ | $-j_3$ | $-j_2$ | $j_1$ | $-j_0$ |

Table 4. Multiplication table for the basis elements of $\mathbb{O}$.

Let $o_1 = (x^0 j_0 + x^1 j_1 + x^2 j_2 + x^3 j_3 + x^4 j_4 + x^5 j_5 + x^6 j_6 + x^7 j_7)$ and $o_2 = (y^0 j_0 + y^1 j_1 + y^2 j_2 + y^3 j_3 + y^4 j_4 + y^5 j_5 + y^6 j_6 + y^7 j_7)$ be two octonions. Then we have according to Table 4:

$$\begin{gathered} o_1 \cdot o_2 = (x^0 j_0 + x^1 j_1 + x^2 j_2 + x^3 j_3 + x^4 j_4 + x^5 j_5 + x^6 j_6 + x^7 j_7) \circ \\ (y^0 j_0 + y^1 j_1 + y^2 j_2 + y^3 j_3 + y^4 j_4 + y^5 j_5 + y^6 j_6 + y^7 j_7) = \\ = (x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3 - x^4 y^4 - x^5 y^5 - x^6 y^6 - x^7 y^7) j_0 + \\ + (x^1 y^0 + x^0 y^1 - x^3 y^2 + x^2 y^3 - x^5 y^4 + x^4 y^5 + x^7 y^6 - x^6 y^7) j_1 + \\ + (x^2 y^0 + x^3 y^1 + x^0 y^2 - x^1 y^3 - x^6 y^4 - x^7 y^5 + x^4 y^6 + x^5 y^7) j_2 + \\ + (x^3 y^0 - x^2 y^1 + x^1 y^2 + x^0 y^3 - x^7 y^4 + x^6 y^5 - x^5 y^6 + x^4 y^7) j_3 + \\ + (x^4 y^0 + x^5 y^1 + x^6 y^2 + x^7 y^3 + x^0 y^4 - x^1 y^5 - x^2 y^6 - x^3 y^7) j_4 + \\ + (x^5 y^0 - x^4 y^1 + x^7 y^2 - x^6 y^3 + x^1 y^4 + x^0 y^5 + x^3 y^6 - x^2 y^7) j_5 + \\ + (x^6 y^0 - x^7 y^1 - x^4 y^2 + x^5 y^3 + x^2 y^4 - x^3 y^5 + x^0 y^6 + x^1 y^7) j_6 + \\ + (x^7 y^0 + x^6 y^1 - x^5 y^2 - x^4 y^3 + x^3 y^4 + x^2 y^5 - x^1 y^6 + x^0 y^7) j_7 . \end{gathered}$$

According to the multiplication table for octonions, we have the following unit vector fields in $\mathbb{R}^8$ arising as right translations of $\partial_{y^j}$, $j=0,\dots,7$ under the octonion product. If $q=(y^0,\dots,y^7)\in S^7$, then

$$\begin{array}{rcl}
Y_0(q) & = & y^0\partial_{y^0}+y^1\partial_{y^1}+y^2\partial_{y^2}+y^3\partial_{y^3}+y^4\partial_{y^4}+y^5\partial_{y^5}+y^6\partial_{y^6}+y^7\partial_{y^7}\\
Y_1(q) & = & -y^1\partial_{y^0}+y^0\partial_{y^1}-y^3\partial_{y^2}+y^2\partial_{y^3}-y^5\partial_{y^4}+y^4\partial_{y^5}-y^7\partial_{y^6}+y^6\partial_{y^7}
\end{array}$$

$$\begin{array}{rcl}
Y_2(q) & = & -y^2\partial_{y^0}+y^3\partial_{y^1}+y^0\partial_{y^2}-y^1\partial_{y^3}-y^6\partial_{y^4}+y^7\partial_{y^5}+y^4\partial_{y^6}-y^5\partial_{y^7}\\
Y_3(q) & = & -y^3\partial_{y^0}-y^2\partial_{y^1}+y^1\partial_{y^2}+y^0\partial_{y^3}+y^7\partial_{y^4}+y^6\partial_{y^5}-y^5\partial_{y^6}-y^4\partial_{y^7}
\end{array}$$

$$\begin{array}{rcl}
Y_4(q) & = & -y^4\partial_{y^0}+y^5\partial_{y^1}+y^6\partial_{y^2}-y^7\partial_{y^3}+y^0\partial_{y^4}-y^1\partial_{y^5}-y^2\partial_{y^6}+y^3\partial_{y^7}\\
Y_5(q) & = & -y^5\partial_{y^0}-y^4\partial_{y^1}-y^7\partial_{y^2}-y^6\partial_{y^3}+y^1\partial_{y^4}+y^0\partial_{y^5}+y^3\partial_{y^6}+y^2\partial_{y^7}
\end{array}$$

$$\begin{array}{rcl}
Y_6(q) & = & -y^6\partial_{y^0}+y^7\partial_{y^1}-y^4\partial_{y^2}+y^5\partial_{y^3}+y^2\partial_{y^4}-y^3\partial_{y^5}+y^0\partial_{y^6}-y^1\partial_{y^7}\\
Y_7(q) & = & -y^7\partial_{y^0}-y^6\partial_{y^1}+y^5\partial_{y^2}+y^4\partial_{y^3}-y^3\partial_{y^4}-y^2\partial_{y^5}+y^1\partial_{y^6}+y^0\partial_{y^7}.
\end{array}$$

The vector fields $Y_i$, $i=1,\dots,7$ form an orthonormal frame of $T_qS^7$, $q\in S^7$, with respect to restriction of the inner product $\langle\cdot\,,\cdot\rangle$ from $\mathbb{R}^8$ to the tangent space $T_qS^7$ at each $q\in S^7$.

### Commutators between vector fields

Let us denote by $Y_{ij}(q)=\frac{1}{2}[Y_i(q),Y_j(q)]$ the commutators between the constructed above vector fields $Y_j$, $j=0,\dots,7$. We have the following list:

$$\begin{array}{rcl}
Y_{12}(q) & = & y^3\partial_{y^0}+y^2\partial_{y^1}-y^1\partial_{y^2}-y^0\partial_{y^3}+y^7\partial_{y^4}+y^6\partial_{y^5}-y^5\partial_{y^6}-y^4\partial_{y^7}\\
Y_{13}(q) & = & -y^2\partial_{y^0}+y^3\partial_{y^1}+y^0\partial_{y^2}-y^1\partial_{y^3}+y^6\partial_{y^4}-y^7\partial_{y^5}-y^4\partial_{y^6}+y^5\partial_{y^7}\\
Y_{14}(q) & = & y^5\partial_{y^0}+y^4\partial_{y^1}-y^7\partial_{y^2}-y^6\partial_{y^3}-y^1\partial_{y^4}-y^0\partial_{y^5}+y^3\partial_{y^6}+y^2\partial_{y^7}\\
Y_{15}(q) & = & -y^4\partial_{y^0}+y^5\partial_{y^1}-y^6\partial_{y^2}+y^7\partial_{y^3}+y^0\partial_{y^4}-y^1\partial_{y^5}+y^2\partial_{y^6}-y^3\partial_{y^7}\\
Y_{16}(q) & = & y^7\partial_{y^0}+y^6\partial_{y^1}+y^5\partial_{y^2}+y^4\partial_{y^3}-y^3\partial_{y^4}-y^2\partial_{y^5}-y^1\partial_{y^6}-y^0\partial_{y^7}\\
Y_{17}(q) & = & -y^6\partial_{y^0}+y^7\partial_{y^1}+y^4\partial_{y^2}-y^5\partial_{y^3}-y^2\partial_{y^4}+y^3\partial_{y^5}+y^0\partial_{y^6}-y^1\partial_{y^7}
\end{array}$$

$$\begin{array}{rcl}
Y_{23}(q) & = & y^1\partial_{y^0}-y^0\partial_{y^1}+y^3\partial_{y^2}-y^2\partial_{y^3}-y^5\partial_{y^4}+y^4\partial_{y^5}-y^7\partial_{y^6}+y^6\partial_{y^7}\\
Y_{24}(q) & = & y^6\partial_{y^0}+y^7\partial_{y^1}+y^4\partial_{y^2}+y^5\partial_{y^3}-y^2\partial_{y^4}-y^3\partial_{y^5}-y^0\partial_{y^6}-y^1\partial_{y^7}\\
Y_{25}(q) & = & -y^7\partial_{y^0}+y^6\partial_{y^1}+y^5\partial_{y^2}-y^4\partial_{y^3}+y^3\partial_{y^4}-y^2\partial_{y^5}-y^1\partial_{y^6}+y^0\partial_{y^7}\\
Y_{26}(q) & = & -y^4\partial_{y^0}-y^5\partial_{y^1}+y^6\partial_{y^2}+y^7\partial_{y^3}+y^0\partial_{y^4}+y^1\partial_{y^5}-y^2\partial_{y^6}-y^3\partial_{y^7}\\
Y_{27}(q) & = & y^5\partial_{y^0}-y^4\partial_{y^1}+y^7\partial_{y^2}-y^6\partial_{y^3}+y^1\partial_{y^4}-y^0\partial_{y^5}+y^3\partial_{y^6}-y^2\partial_{y^7}
\end{array}$$

$$\begin{array}{rcl}
Y_{34}(q) & = & -y^7\partial_{y^0}+y^6\partial_{y^1}-y^5\partial_{y^2}+y^4\partial_{y^3}-y^3\partial_{y^4}+y^2\partial_{y^5}-y^1\partial_{y^6}+y^0\partial_{y^7}\\
Y_{35}(q) & = & -y^6\partial_{y^0}-y^7\partial_{y^1}+y^4\partial_{y^2}+y^5\partial_{y^3}-y^2\partial_{y^4}-y^3\partial_{y^5}+y^0\partial_{y^6}+y^1\partial_{y^7}\\
Y_{36}(q) & = & y^5\partial_{y^0}-y^4\partial_{y^1}-y^7\partial_{y^2}+y^6\partial_{y^3}+y^1\partial_{y^4}-y^0\partial_{y^5}-y^3\partial_{y^6}+y^2\partial_{y^7}\\
Y_{37}(q) & = & y^4\partial_{y^0}+y^5\partial_{y^1}+y^6\partial_{y^2}+y^7\partial_{y^3}-y^0\partial_{y^4}-y^1\partial_{y^5}-y^2\partial_{y^6}-y^3\partial_{y^7}
\end{array}$$

$$\begin{aligned}
Y_{45}(q) &= y^1\partial_{y^0} - y^0\partial_{y^1} - y^3\partial_{y^2} + y^2\partial_{y^3} + y^5\partial_{y^4} - y^4\partial_{y^5} - y^7\partial_{y^6} + y^6\partial_{y^7}\\
Y_{46}(q) &= y^2\partial_{y^0} + y^3\partial_{y^1} - y^0\partial_{y^2} - y^1\partial_{y^3} + y^6\partial_{y^4} + y^7\partial_{y^5} - y^4\partial_{y^6} - y^5\partial_{y^7}\\
Y_{47}(q) &= -y^3\partial_{y^0} + y^2\partial_{y^1} - y^1\partial_{y^2} + y^0\partial_{y^3} + y^7\partial_{y^4} - y^6\partial_{y^5} + y^5\partial_{y^6} - y^4\partial_{y^7}\\
\\
Y_{56}(q) &= -y^3\partial_{y^0} + y^2\partial_{y^1} - y^1\partial_{y^2} + y^0\partial_{y^3} - y^7\partial_{y^4} + y^6\partial_{y^5} - y^5\partial_{y^6} + y^4\partial_{y^7}\\
Y_{57}(q) &= -y^2\partial_{y^0} - y^3\partial_{y^1} + y^0\partial_{y^2} + y^4\partial_{y^6} - y^5\partial_{y^7}\\
Y_{67}(q) &= y^1\partial_{y^0} - y^0\partial_{y^1} - y^3\partial_{y^2} + y^2\partial_{y^3} - y^5\partial_{y^4} + y^4\partial_{y^5} + y^7\partial_{y^6} - y^6\partial_{y^7}.
\end{aligned}$$

## BASIC NOTATIONS

| | |
|---|---|
| $\mathbb{N}$ | is the set of positive integer numbers |
| $\mathbb{R}$ | is the set of real numbers |
| $\mathbb{C}$ | is the set of complex numbers |
| $\mathbb{Q}$ | is the set of quaternion numbers |
| $\mathbb{O}$ | is the set of the Caley numbers (octonions) |
| $M, N$ | are smooth manifolds |
| $T_qM$ | is the tangent space at the point $q \in M$ |
| $T_q^*M$ | is the cotangent space at the point $q \in M$ |
| $TM$ | is the tangent bundle for a manifold $M$ |
| $T^*M$ | is the cotangent bundle for a manifold $M$ |
| $\mathrm{pr}_M, \mathrm{pr}_M^*$ | are the canonical projections from $TM$, $T^*M$ to $M$ |
| $\mathrm{Vect}(M)$ | is the set of smooth vector fields on $M$ |
| $X, Y, Z$ | are smooth vector fields, elements of $\mathrm{Vect}(M)$ |
| $S^1$ | is the unit circle |
| $\mathrm{Diff}\, S^1$ | is the group of orientation preserving diffeomorphisms of $S^1$ |
| Vir | is the Virasoro-Bott group |
| $\mathfrak{vir}$ | is the Virasoro algebra |
| $D$ | is a smooth distribution on $M$, horizontal sub-bundle of $TM$ |
| $D^*$ | is a smooth co-distribution on $M$, smooth sub-bundle of $T^*M$ |
| $D^{\perp}$ | is the set of annihilators of $D$, horizontal sub-bundle of $TM$ |
| $g$ | is a Riemannian metric |
| $g_D$ | is a sub-Riemannian metric related to the distribution $D$ |
| $g^D$ | is a sub-Riemannian co-metric related to the distribution $D$ |
| $\Omega$ | is a curvature form of a horizontal distribution |
| $A$ | is a connection form of a horizontal distribution |
| $\sigma$ | is the infinitesimal generator of a group acting on a manifold |
| $\nabla$ | is the Levi-Civita connection |
| $H$ | is a Hamiltonian function |
| $\overrightarrow{H}$ | is the Hamiltonian vector field associated to a function $H$ |
| $\mathbb{G}, \mathbb{H}$ | are Lie groups |
| $\mathfrak{g}, \mathfrak{h}$ | are Lie algebras |
| $\mathrm{Aut}(E)$ | is the group of automorphisms of $E$ |
| $\mathrm{End}(E)$ | is the Lie algebra of endomorphisms of $E$ |
| $W, W^*$ | is a vector space and its dual |

| | |
|---|---|
| $\cdot$ | is the Euclidean inner product |
| $(\cdot\,,\cdot)$ | is an inner product |
| $\langle\cdot\,,\cdot\rangle$ | is a pairing between $W$ and $W^*$ |
| $\{\cdot\,,\cdot\}$ | are Poisson brackets |
| $[\cdot\,,\cdot]$ | is the commutator |
| $\Vert\cdot\Vert_E$ | is the Euclidean norm |
| $\Vert\cdot\Vert_{\mathbb{H}^1}$ | is the Heisenberg norm |
| $d_E$ | is the Euclidean distance function |
| $d_{\mathbb{H}^1}$ | is the Heisenberg distance function |
| $d_{c-c}$ | is the Carnot-Carathéodory distance function |
| $f\vert_S$ | is the restriction of the function $f$ to the set $S$ |
| $\operatorname{grad} f$ | is the Riemannian gradient of a function $f$ |
| $\operatorname{grad}_D f$ | is the sub-Riemannian gradient of a function $f$ |
| $\mathcal{R}$ | is the rolling map |

**Acknowledgment**

It is not a duty, but a pleasure to thank the organizers of the school "Analysis - with Applications to Mathematical Physics" in Göttingen, and in particular, Wolfram Bauer, for the invitation, warm hospitality, an extraordinarily interesting school, and the indispensable help in preparation of these notes.

I am grateful to the Analysis Group at the University of Bergen for the creative friendly atmosphere and the infinite source of ideas, questions and emotions. Particularly, I would like to express deep gratitude to my co-authors with whom I shared hard hours of work and lovely time of mathematical conversations.

Special sincere thanks to my lovely husband Alexander Vasiliev for his support of all my initiatives.

## References

[1] J. F. Adams, *Vector fields on spheres.* Ann. of Math. (2) **75** (1962), 603–632.

[2] A. Agrachev, M. Caponigro, *Controllability on the group of diffeomorphisms.* Ann. Inst. H. Poincaré. Anal. Non Linéaire **26** (2009), 2503–2509.

[3] A. Agrachev, Y. Sachkov, *Control theory from the geometric viewpoint.* Encyclopedia of Mathematical Sciences, 87. Control Theory and Optimization, II. Springer-Verlag, Berlin, 2004. 412 pp.

[4] H. Airault, P. Malliavin, *Unitarizing probability measures for representations of Virasoro algebra.* J. Math. Pures Appl. **80** (2001), no. 6, 627–667.

[5] D. Alekseevsky, Y. Kamishima, *Pseudo-conformal quaternionic CR structure on* $(4n+3)-$ *dimensional manifolds.* Ann. Mat. Pura Appl. (4) **187** (2008), no. 3, 487–529.

[6] N. Arcozzi, A. Baldi, *From Grushin to Heisenberg via an isoperimetric problem.* J. Math. Anal. Appl. **340** (2008), no. 1, 165–174.

[7] V. I. Arnold, *Mathematical methods of classical mechanics.* Translated from the 1974 Russian original by K. Vogtmann and A. Weinstein. Corrected reprint of the

second (1989) edition. Graduate Texts in Mathematics, 60. Springer-Verlag, New York, 516 pp.

[8] J. C. Baez, *The octonions*. Bull. Amer. Math. Soc. (N.S.) **39** (2002), no. 2, 145–205.

[9] W. Bauer, K. Furutani, *Spectral analysisand geometry of a sub-Riemannian structure on $S^3$ and $S^7$*. J. Geom. Phys. **58** (2008), 1693–1738.

[10] W. Bauer, K. Furutani, C. Iwasaki, *Trivializable sub-Riemannian structures on spheres*. Bull. Sci. Math. 137 (2013), no. 3, 361–385.

[11] M. Bauer, D. Bernard, *SLE martingales and the Virasoro algebra*. Phys. Lett. B 557 (2003), no. 3-4, 309–316.

[12] M. Bauer, D. Bernard, *Conformal field theories of stochastic Loewner evolutions*. Comm. Math. Phys. 239 (2003), no. 3, 493–521.

[13] A. Bellaïche, *The tangent space in Sub-Riemannian geometry*. in Sub-Riemannian geometry Edited by André Bellaïche and Jean-Jacques Risler. Progress in Mathematics, 144. Birkhäuser Verlag, Basel, 1996. 393 pp.

[14] A. Boggess, *CR manifolds and the tangential Cauchy-Riemann complex*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1991. 364 pp.

[15] V. G. Boltyanskii, *Sufficient conditions for optimality and the justification of the dynamic programming method*. SIAM J. Control **4** (1966), 326–361.

[16] R. Bott, *On the characteristic classes of groups of diffeomorphisms*. Enseignment Math. (2) **23** (1977), no. 3-4, 209–220.

[17] R. Bott, J. Milnor, *On the parallelizability of the spheres*. Bull. Amer. Math. Soc. **64** (1958), 87–89.

[18] R. Bryant, L. Hsu, *Rigidity of integral curves of rank 2 distributions*. Invent. Math. **114** (1993), no. 2, 435–461.

[19] F. Bullo, A. D. Lewis, *Geometric control of mechanical systems. Modeling, analysis, and design for simple mechanical control systems*. Texts in Applied Mathematics, 49. Springer-Verlag, New York, 2005. 726 pp.

[20] O. Calin, D. C. Chang, P. Greiner, *Geometric analysis on the Heisenberg group and its generalizations*. AMS/IP Studies in Advanced Mathematics, 40. American Mathematical Society, Providence, RI; International Press, Somerville, MA, 2007. 244 pp.

[21] O. Calin, D. C. Chang, I. Markina, *Geometric analysis on $H$-type groups related to division algebras*. Math. Nachr. **282** (2009), no. 1, 44–68.

[22] L. Capogna, D. Danielli, S. D. Pauls, J. T. Tyson, *An introduction to the Heisenberg group and the sub-Riemannian isoperimetric problem*. Progress in Mathematics, 259. Birkhäuser Verlag, Basel, 2007. 223 pp.

[23] M. P. do Carmo, *Riemannian geometry*. Mathematics: Theory & Applications. Birkhäuser Boston, Inc., Boston, MA, 1992. 300 pp.

[24] É. Cartan, *Les systèmes de Pfaff, à cinq variables et les équations aux dérivées partielles du second ordre*. Ann. Sci. École Norm. Sup. **27** (1910), no. 3, 109–192.

[25] C. H. Chang, D. C. Chang, B. Gaveau, P. Greiner, H. P. Lee, *Geometric analysis on a step 2 Grusin operator*. Bull. Inst. Math. Acad. Sin. (N.S.) **4** (2009), no. 2, 119–188.

[26] D. C. Chang, I. Markina, *Geometric analysis on quaternion $H$-type groups.* J. Geom. Anal. **16** (2006), no. 2, 266–294.

[27] D. C. Chang, I. Markina, A. Vasil'ev, *Sub-Lorentzian geometry on anti-de Sitter space.* J. Math. Pures Appl. **90** (9) (2008), no. 1, 82–110.

[28] D. C. Chang, I. Markina, A. Vasil'ev, *Sub-Riemannian geodesics on the 3-D sphere.* Complex Anal. Oper. Theory **3** (2009), no. 1, 44–68.

[29] D. C. Chang, I. Markina, A. Vasil'ev, *Modified action and differential operators on the 3-D sub-Riemannian sphere.* Asian J. Math. **14** (3) (2010), no. 4, 439–474.

[30] Y. Chitour, F. Jean, E. Trélat, *Genericity results for singular curves.* J. Diff. Geom. **73** (2006), no. 1, 45–73.

[31] Y. Chitour, P. Kokkonen, *Rolling Manifolds: Intrinsic Formulation and Controllability.* arXiv:1011.2925

[32] Y. Chitour, P. Kokkonen, *Rolling manifolds on space forms.* Ann. Inst. H. Poincaré Anal. Non Linéaire **29** (2012), no. 6, 927–954.

[33] W. L. Chow, *Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung.* Math. Ann. **117** (1939), 98–105.

[34] P. Ciatti, *Scalar products on Clifford modules and pseudo-$H$-type Lie algebras.* Ann. Mat. Pura Appl. **178** (2000), no. 4, 41–31.

[35] L. A. Cordero, P. E. Parker, *Isometry groups of pseudoriemannian 2-step nilpotent Lie groups.* Houston J. Math. **35** (2009), no. 1, 49–72.

[36] M. Cowling, A. H. Dooley, A. Korányi, F. Ricci, *$H$-type groups and Iwasawa decompositions.* Adv. Math. **87** (1991), no. 1, 1–41.

[37] D. Danielli, N. Garofalo, D. M. Nhieu, *Sub-Riemannian calculus on hypersurfaces in Carnot groups.* Adv. Math. **215** (2007), no. 1, 292–378.

[38] G. Darboux, *Sur le problème de Pfaff.* Bull. Sci. Math. **6** (1882), 14–36, 49–68.

[39] C. A. Deavours, *The quaternion calculus.* Amer. Math. Monthly, **80** (1973), 995–1008.

[40] J. Dou, P. Niu, J. Han, *Polar coordinates for the generalized Baouendi-Grushin operator and applications.* J. Partial Differential Equations **20** (2007), no. 4, 322–336.

[41] P. I. Dubnikov, S. N. Samborskii, *Controllability criterion for systems in a Banach space (Generalization of Chow's theorem).* Ukraine Math. J. **32** (1979), no. 5, 649–653.

[42] J. J. Duistermaat, J. A. C. Kolk, *Lie groups.* Universitext. Springer-Verlag, Berlin, 2000. 344 pp.

[43] P. Eberlein, *Riemannian submersion and lattices in 2-step nilpotent Lie groups.* Comm. Anal.Geom. **11**, (2003), no. 3, 441-488.

[44] P. Eberlein, *Geometry of 2-step nilpotent Lie groups.* Modern dynamical systems and applications,67–101, Cambridge Univ. Press, Cambridge, 2004.

[45] J. Escher, B. Kolev, and M. Wunch, *The geometry of a vorticity model equation.* Commun. Pure Appl. Anal. **11** (2012), no. 4, 1407–1419.

[46] R. Friedrich, W. Werner, *Conformal restriction, highest-weight representations and SLE.* Comm. Math. Phys. **243** (2003), no. 1, 105–122.

[47] R. M. Friedrich, *The global geometry of stochastic Löwner evolutions.* Probabilistic approach to geometry, 79117, Adv. Stud. Pure Math., **57**, Math. Soc. Japan, Tokyo, 2010.

[48] G. Frobenius, *Über die Pfaffsche probleme.* J. für Reine und Agnew. Math **82** (1877), 230–315.

[49] G. B. Folland, *A fundamental solution for a subelliptic operator.* Bull. Amer. Math. Soc. **79** (1973), 373–376.

[50] G. B. Folland, E. M. Stein, *Hardy spaces on homogeneous groups.* Mathematical Notes, 28. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1982. 285 pp.

[51] W. Fulton, J. Harris, *Representation theory.* A first course. Graduate Texts in Mathematics, 129. Readings in Mathematics. Springer-Verlag, New York, 1991. 551 pp.

[52] N. Garofalo, D. Vassilev, *Strong unique continuation properties of generalized Baouendi-Grushin operators.* Comm. Partial Differential Equations **32** (2007), no. 4-6, 643–663.

[53] Z. Ge, *Betti numbers, characteristic classes and sub-Riemannian geometry.* Illinois J. Math. **36** (1992), no. 3, 372–403.

[54] I. M. Gel'fand, D. B. Fuchs, *Cohomology of the Lie algebra of vector fields on the circle.* Functional Anal. Appl. **2** (1968), no. 4, 342–343.

[55] V. Gershkovich, A. Vershik, *Nonholonomic manifolds and nilpotent analysis.* J. Geom. Phys. **5** (1988), no. 3, 407–452.

[56] M. Godoy Molina, E. Grong, I. Markina, S. Leite, *An intrinsic formulation of the rolling manifolds problem.* J. Dyn. Control Syst. **18** (2012), no. 2, 181–214.

[57] M. Godoy, I. Markina, *Sub-Riemannian geometry of parallelizable spheres.* Rev, Mat. Iberoam. **27** (2011), no. 3, 997–1022.

[58] M. Godoy, I. Markina, *Sub-Riemannian geodesics and heat operator on odd dimensional spheres.* Anal. Math. Phys. **2** (2012), no. 2, 123–147.

[59] M. Godoy, A. Korolko, I. Markina, *Sub-semi-Riemannian geometry of general $H$-type groups.* Bull. Sci. Math. **137** (2013), no. 6, 805–835.

[60] M. Godoy Molina, E. Grong, *Geometric conditions for the existence of an intrinsic rolling.* arXiv:1111.0752

[61] M. Grochowski, *Geodesics in the sub-Lorentzian geometry.* Bull. Polish Acad. Sci. Math. **50** (2002), no. 2, 161–178.

[62] M. Grochowski, *Reachable sets for the Heisenberg sub-Lorentzian structure on $\mathbb{R}^3$. An estimate for the distance function.* J. Dyn. Control Syst. **12** (2006), no. 2, 145–160.

[63] M. Gromov, *Carnot-Carathéodory spaces seen from within.* in Sub-Riemannian geometry Edited by André Bellaïche and Jean-Jacques Risler. Progress in Mathematics, 144. Birkhäuser Verlag, Basel, 1996. 393 pp.

[64] E. Grong, *Controllability of rolling without twisting or slipping in higher dimensions.* SIAM J. Control Optim. **50** (2012), no. 4, 2462–2485.

[65] E. Grong, I. Markina, A. Vasil'ev, *Geometry with non-holonomic constraints on infinite dimensional manifolds.* arXiv:1201.2251.

[66] E. Grong, I. Markina, A. Vasil'ev, *Sub-Riemannian structures corresponding to Khlerian metrics on the universal Teichmller space and curve.* "60 years of analytic functions in Lublin" in memory of our professors and friends Jan G. Krzy?, Zdzis?aw Lewandowski and Wojciech Szapiel, 97116, Monogr. Univ. Econ. Innov. Lublin, Innovatio Press Sci. Publ. House Univ. Econ. Innov. Lublin, Lublin, 2012.

[67] V. V. Grušin, *A certain class of hypoelliptic operators.* (Russian) Mat. Sb. (N.S.) **83** (125) (1970), 456–473.

[68] K. Gürlebeck, W. Sprössig, *Quaternionic and Clifford Calculus for Physicists and Engineers.* Chichester: John Wiley and Sons, 1997. 371 pp.

[69] E. Heintze, X. Liu, *Homogeneity of infinite dimensional isoparametric submanifolds.* Ann. of Math. (2), **149** (1999), 149–181.

[70] M. R. Herman, *Simplicité du groupe des difféomorphismes de classe $C^\infty$, isotopes à l'identité, du tore de dimension n.* (French) C. R. Acad. Sci. Paris Sér. A-B **273** (1971), A232–A234.

[71] H. Hopf, *Über die Abbildungen von Sphären auf Sphären niedrigerer Dimension.* Math. Ann. **104** (1931), 637–665.

[72] K. Hüper, F. Silva Leite, *On the geometry of rolling and interpolation curves on $S^n$, $\mathrm{SO}_n$, and Grassmann manifolds.* J. Dyn. Control Syst. **13** (2007), no. 4, 467–502.

[73] K. Hüper, M. Kleinsteuber, F. Silva Leite, *Rolling Stiefel manifolds.* Internat. J. Systems Sci. **39** (2008), no. 9, 881–887.

[74] A. Hurtado, C. Rosales, *Area-stationary surfaces inside the sub-Riemannian three-sphere.* Math. Ann. **340** (2008), no. 3, 675–708.

[75] D. Husemoller, *Fibre bundles.* Third edition. Graduate Texts in Mathematics, 20. Springer-Verlag, New York, 1994. 353 pp.

[76] D. D. Joyce, *Riemannian holonomy groups and calibrated geometry.* Oxford Graduate Texts in Mathematics, 12. Oxford University Press, Oxford, 2007. 303 pp.

[77] A. Kaplan, *Fundamental solutions for a class of hypoelliptic PDE generated by composition of quadratics forms.* Trans. Amer. Math. Soc. **258** (1980), no. 1, 147–153.

[78] A. Kaplan, *On the geometry of groups of Heisenberg type.* Bull. London Math. Soc. **15** (1983), no. 1, 35–42.

[79] M. Khajeh Salehani, I. Markina, *Controllability on infinite-dimensional manifolds.* arXiv:1209.1796

[80] B. Khesin, G. Misiołek, *Euler equations on homogeneous spaces and Virasoro orbits.* Adv. Math. 176 (2003), no. 1, 116–144.

[81] B. Khesin, R. Wendt, *The geometry of infinite-dimensional groups.* A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series.], 51. Springer-Verlag, Berlin, 2009. 304 pp.

[82] A. A. Kirillov, *The orbits of the group of diffeomorphisms of the circle, and local Lie superalgebras.* (Russian) Funktsional. Anal. i Prilozhen. **15** (1981), no. 2, 75–76.

[83] A. A. Kirillov, *Geometric approach to discrete series of unirreps for vir.* J. Math. Pures Appl. **77** (1998), 735–746.

[84] A. A. Kirillov, D. V. Yuriev, *Representations of the Virasoro algebra by the orbit method.* J. Geom. Phys. **5** (1988), no. 3, 351–363.

[85] A. A. Kirillov, *Kähler structures on $K$-orbits of the group of diffeomorphisms of a circle*. Funct. Anal. Appl. **21** (1987), no. 2, 42–45.

[86] A. A. Kirillov, D. V. Yur'ev, *Kähler geometry and the infinite-dimensional homogenous space* $M = \mathrm{Diff}_+(S^1)/\mathrm{Rot}(S^1)$. Funct. Anal. Appl. **21** (1987), no. 4, 284–294.

[87] A. W. Knapp, *Lie groups beyond an introduction*. Second edition. Progress in Mathematics, 140. Birkhäuser Boston, Inc., Boston, MA, 2002. 812 pp.

[88] A. Korányi, *Geometric properties of Heisenberg-type groups*. Adv. in Math. **56** (1985), no. 1, 28–38.

[89] A. Korányi, H. M. Reimann, *Quasiconformal mappings on the Heisenberg group*. Invent. Math. **80** (1985), no. 2, 309–338.

[90] A. Korolko, I. Markina, *Nonholonomic Lorentzian geometry on some H-type groups*. J. Geom. Anal. **19** (2009), no. 4, 864–889.

[91] A. Korolko, I. Markina, *Geodesics on H-type groups with sub-Lorentzian metric and their physical interpretation*. Comp. Anal. Oper. Theory. **4** (2010), no. 3, 589–618.

[92] A. Korolko, I. Markina, *Semi-Riemannian geometry with nonholonomic constraints*. Taiwanese J. of Math., **15** (2011), no. 4, 1581–1616.

[93] A. Kriegl, P. W. Michor, *The convenient setting of global analysis*. Mathematical Surveys and Monographs, 53. American Mathematical Society, Providence, RI, 1997. 618 pp.

[94] Yu. S. Ledyaev, *On an infinite-dimensional variant of the Rashevski-Chow theorem*. Dokl. Akad. Nauk **398** (2004), no. 6, 735–737.

[95] E. B. Lee, L. Markus, *Foundations of optimal control theory*. John Wiley & Sons, Inc., New York-London-Sydney 1967. 576 pp.

[96] L. Lempert, *The Virasoro group as a complex manifold*. Math. Res. Lett. **2** (1995), no. 4, 479–495.

[97] W. Liu, H. J. Sussmann, *Shortest paths for sub-Riemannian metrics on rank-two distributions*. Mem. Amer. Math. Soc. **118** (1995), no. 564, 104 pp.

[98] D. W. Lyons, *An elementary introduction to the Hopf fibration*. Math. Mag. **76** (2003), no. 2, 87–98.

[99] I. Markina, A. Vasil'ev, *Virasoro algebra and dynamics in the space of univalent functions*. Five lectures in complex analysis, 85116, Contemp. Math., **525**, Amer. Math. Soc., Providence, RI, 2010.

[100] J. Marsden, A. Weinstein, *Reduction of symplectic manifolds with symmetry*. Rep. Mathematical Phys. **5** (1974), no. 1, 121–130.

[101] K. R. Meyer, *Symmetries and integrals in mechanics*. Dynamical systems (Proc. Sympos., Univ. Bahia, Salvador, 1971), pp. 259–272. Academic Press, New York, 1973.

[102] P. W. Michor, D. Mumford, *Riemannian geometries on spaces of plane curves*. J. Eur. Math. Soc. (JEMS) **8** (2006), no. 1, 1–48.

[103] P. W. Michor, D. Mumford, *An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach*. Appl. Comput. Harmon. Anal. 23 (2007), no. 1, 74–113.

[104] P. W. Michor, D. Mumford, J. Shah, L. Younes, *A metric on shape space with explicit geodesics.* Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **19** (2008), no. 1, 25–57.

[105] R. L. Mills, C. N. Yang, *Conservation of isotopic spin and isotopic gauge invariance.* Phys. Rev. **96** (1954), p. 191.

[106] J. Milnor, *Morse theory.* Annals of Mathematics Studies, No. 51 Princeton University Press, Princeton, N.J. 1963. 153 pp.

[107] J. Milnor, D. Husemoller, *Symmetric bilinear forms.* Ergebnisse der Mathematik und ihrer Grenz-gebiete, Band 73. Springer-Verlag, New York-Heidelberg, 1973. 147 pp.

[108] J. Milnor, *Remarks on infinite-dimensional Lie groups.* Relativity, groups and topology, II (Les Houches, 1983), 1007–1057, North-Holland, Amsterdam, 1984.

[109] R. Montgomery, *Abnormal minimizers.* SIAM J. Control Optim. **32** (1994), no. 6, 1605–1620.

[110] R. Montgomery, *A survey of singular curves in sub-Riemannian geometry.* J. Dynam. Control Systems **1** (1995), no. 1, 49–90.

[111] R. Montgomery, *A tour of subriemannian geometries, their geodesics and applications.* Mathematical Surveys and Monographs, 91. American Mathematical Society, Providence, RI, 2002. 259 pp.

[112] R. Monti, *The regularity problem for sub-Riemannian geodesics.* Springer, INDAM Series, 2013.

[113] T. Nagano, *Linear differential systems with singularities and an application to transitive Lie algebras.* J. Math. Soc. Japan **18** (1966), 398–404.

[114] A. Nagel, E. M. Stein, S. Wainger, *Balls and metrics defined by vector fields. I. Basic properties.* Acta Math. **155** (1985), no. 1-2, 103–147.

[115] J. Nash, *The imbedding problem for Riemannian manifolds.* Ann. of Math. **63** (1956), no. 2, 20–63.

[116] H. Omori, *Infinite-dimensional Lie groups.* Translated from the 1979 Japanese original and revised by the author. Translations of Mathematical Monographs, 158. American Mathematical Society, Providence, RI, 1997. 415 pp.

[117] B. O'Neill, *Semi-Riemannian geometry. With applications to relativity.* Pure and Applied Mathematics, 103. Academic Press, [Harcourt Brace Jovanovich, Publishers], New York, 1983. 468 pp.

[118] A. Pressley, G. Segal, *Loop groups.* Oxford Mathematical Monographs. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1986. 318 pp.

[119] P. K. Rashevskiĭ, *About connecting two points of complete nonholonomic space by admissible curve.* Uch. Zapiski Ped. Inst. K. Liebknecht **2** (1938), 83–94.

[120] F. Ricci, *The spherical transform on harmonic extensions of $H$-type groups.* Differential geometry (Turin, 1992). Rend. Sem. Mat. Univ. Politec. Torino, **50** (1992), no. 4, 381–392.

[121] M. Ritoré, *A proof by calibration of an isoperimetric inequality in the Heisenberg group $\mathbb{H}^n$.* Calc. Var. Partial Differential Equations **44** (2012), no. 1-2, 47–60.

[122] M. Ritoré, C. Rosales, *Area-stationary and stable surfaces in the sub-Riemannian Heisenberg group* $\mathbb{H}^1$. Mat. Contemp. **35** (2008), 185–203.

[123] A. A. Sagle, R. E. Walde, *Introduction to Lie Groups and Lie Algebras.* Pure and Appl. Mathematics, 51. Academic Press, New York-London, 1973. 361 pp.

[124] A. C. Schaeffer, D. C. Spencer, *Coefficient Regions for Schlicht Functions (With a Chapter on the Region of the Derivative of a Schlicht Function by Arthur Grad).* American Mathematical Society Colloquium Publications, Vol. 35. American Mathematical Society, New York, 1950.

[125] R. W. Sharpe, *Differential geometry. Cartan's generalization of Klein's Erlangen program.* Graduate Texts in Mathematics, 166. Springer-Verlag, New York, 1997. 421 pp.

[126] I. M. Singer, J. A. Thorpe, *Lecture notes on elementary topology and geometry.* Undergraduate Texts in Mathematics. Springer-Verlag, New York-Heidelberg, 1976. 232 pp.

[127] R. S. Strichartz, *Sub-Riemannian geometry.* J. Diff. Geom. **24** (1986), 221–263; Correction, ibid. **30** (1989), 595–596.

[128] H. J. Sussmann, *Orbits of families of vector fields and integrability of distributions.* Trans. Amer. Math. Soc. **180** (1973), 171–188.

[129] J. T. Tyson, *Sharp weighted Young's inequalities and Moser-Trudinger inequalities on Heisenberg type groups and Grushin spaces.* Potential Anal. **24** (2006), no. 4, 357–384.

[130] J. Vick, *Homology Theory: An Introduction to Algebraic Topology.* Graduate Texts in Mathematics, Vol 145. Springer-Verlag New York, 1994.

[131] F. W. Warner, *Foundations of differentiable manifolds and Lie groups.* Graduate Texts in Mathematics, 94. Springer-Verlag, New York-Berlin, 1983. 272 pp.

[132] J. A. Zimmerman, *Optimal control of the sphere* $S^n$ *rolling on* $E^n$. Math. Control Signals Systems **17** (2005), no. 1, 14–37.

Irina Markina
Department of Mathematics, University of Bergen, Norway
e-mail: `irina.markina@uib.no`