# Geometry and feedback classification of low-dimensional non-linear control systems

Ulysse Serres

# Thèse

présentée par

**Ulysse Serres**

en vue d'obtenir le titre de

Docteur de L'université de Bourgogne

Spécialité : Mathématiques et automatique

## *Géométrie et classification par feedback des systèmes de contrôle non linéaires de basse dimension*

soutenue publiquement à Dijon le 24 mars 2006 devant le jury composé de

| | | |
|---|---|---|
| Robert | Moussu | Président du jury |
| Bronisław | Jakubczyk | Rapporteur |
| Andrey | Sarychev | Rapporteur |
| Michel | Paindavoine | Examinateur |
| Andrei | Agrachev | Directeur de thèse |
| Jean-Paul | Gauthier | Directeur de thèse |

UNIVERSITÉ DE BOURGOGNE

Laboratoire Electronique Informatique et Image

# PhD THESIS

presented by

**Ulysse Serres**

submitted for the degree of

Docteur de L'université de Bourgogne

Special field: Mathematics et automatics

# *Geometry and feedback classification of low-dimensional non-linear control systems*

publicly defended in Dijon on the 24th of march 2006 before the examining board formed by

| | | |
|---|---|---|
| Robert | Moussu | Foreman |
| Bronisław | Jakubczyk | Referee |
| Andrey | Sarychev | Referee |
| Michel | Paindavoine | Examinator |
| Andrei | Agrachev | Supervisor |
| Jean-Paul | Gauthier | Supervisor |

**Acknowledgments**

I wish to thank my both supervisors Professor Andrei A. Agrachev and Professor Jean-Paul Gauthier for having offered me the opportunity to do this work. I also thank them a lot for their constant support with great patient and humanity during all these years.

I am also most grateful to the International School for Advanced Studies of Trieste for its hospitality during the long period of my PhD preparation.

I am indebted to Igor Zelenko for numerous fruitful discussions. I am thankful to Nicolas Boivin, Ugo Boscain, Željka Bošnjak, Natalia Chtcherbakova, Niccolò Desenzani, Matteo Facchin, Fabrice Méron, Adeline Nonciaux, Mario Sigalotti and Sofian Teber who helped me in many ways.

In this moment of joy and satisfaction my warmest thanks go to my family.

*Florence, 17th February 2006.*

# Contents

**Introduction** **9**

**1 Preliminaries** **15**
1.1 Chronological calculus . . . 15
1.2 Elements of geometry . . . 18
1.3 Pontryagin Maximum Principle . . . 23

**2 Curvature** **27**
2.1 Feedback equivalence of control systems . . . 28
2.2 Curvature of two-dimensional optimal control problems . . . 28
2.3 Evaluation of the invariant $b$ . . . 35
2.4 Projectivised Hamiltonians . . . 36
2.5 Curvature, first examples . . . 38
2.5.1 Two-dimensional Riemannian problem . . . 39
2.5.2 Two-dimensional Lorentzian manifold . . . 41
2.5.3 Problem of controlling the angular momentum of a rigid body . . . 42
2.6 Jacobi equation . . . 45

**3 Zermelo's problems** **51**
3.1 Zermelo navigation problem . . . 51
3.1.1 Description of the problem . . . 51
3.1.2 Hamiltonian function and associated Hamiltonian vector field . . . 52
3.1.3 Curvature of Zermelo problem . . . 55
3.2 Dual Zermelo navigation problem . . . 61
3.2.1 Definition of the problem . . . 61
3.2.2 Hamiltonian function . . . 62
3.3 Link between Zermelo and dual to Zermelo problems . . . 62

**4 Microlocal normal forms for control systems** **67**
4.1 Principal invariants of the equivalence problem . . . 68
4.1.1 Counting the invariants . . . 68
4.1.2 Relation between the principal invariants . . . 68
4.2 Microlocal normal forms . . . 70

4.2.1 Normal case . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 70
4.2.2 Abnormal case . . . . . . . . . . . . . . . . . . . . . . . . . . . . 76
4.3 Flat systems . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 77
4.3.1 Two fundamental theorems . . . . . . . . . . . . . . . . . . . . . . 77
4.3.2 Examples . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 82

**5 Some global results** **85**
5.1 A Gauss-Bonnet theorem for special control problems . . . . . . . . . . 85
5.2 A Hopf theorem for control systems . . . . . . . . . . . . . . . . . . . . 88
5.2.1 Jacobi curves . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 88
5.2.2 The Hopf theorem . . . . . . . . . . . . . . . . . . . . . . . . . . . 89
5.2.3 A natural question . . . . . . . . . . . . . . . . . . . . . . . . . . 91

**Closing** **95**

**Bibliography** **98**

# Introduction

The aim of this thesis is the study of the local and global geometry of fully nonlinear time-optimal control problems on two-dimensional smooth manifolds. In particular we are interested in the study of the feedback-invariants of such a system. Consider a generic time-optimal control problem on a smooth manifold $M$:

$$\dot{q} = \boldsymbol{f}(q,u), \quad q \in M, \quad u \in U, \tag{1}$$
$$q(0) = q_0, \quad q(t_1) = q_1, \tag{2}$$
$$t_1 \to \min, \tag{3}$$

where the control set $U$ is also a smooth manifold and the points $q_0$, $q_1 \in M$ are fixed. The geometry of control system (1) is studied up to feedback transformations that is, up to diffeomorphism of $M \times U$ of the form

$$(\tilde{q}, \tilde{u}) = (\phi(q), \psi(q,u)) .$$

After a preliminary chapter which is mostly an introduction to the vocabulary and the classical results that will be used in the present thesis, our first goal will be to determined the feedback-invariants of system (1) by making an analogy with the classical Riemannian geometry of surfaces. This is the purpose of the first part of Chapter 2 in which we will construct the control analogue to the classical Gaussian curvature of a surface. Why to look for such a similar invariant? First of all because in Riemannian geometry the Gaussian curvature of a surface reflects intrinsic properties of the geodesic flow, i.e., properties that do not depend on the choice of local coordinates. In particular, the geodesics issued from a point on the surface tend to "diverge" if the curvature at this point is negative whereas they tend to "converge" if the curvature at this point is positive. The work in this direction began with the paper [5] by A. A. Agrachev and R. V. Gamkrelidze in which the authors, using a purely variational approach, generalized the notion of Ricci curvature tensor of classical Riemannian geometry to smooth optimal control problems. We will neither adopt nor adapt the variational point of view in order to generalize the Gaussian curvature of a surface but instead we will use the Cartan's moving frame method which offers a very geometrical point of view.

For such a purpose we first write the maximized (normal) Hamiltonian of Pontryagin Maximum Principle (PMP) of the optimal control problem (1)–(3)

$$h(\lambda) = \max_{u \in U} \langle \lambda, \boldsymbol{f}_u(q) \rangle , \quad q \in M, \quad \lambda \in T_q^*M,$$

which is a function on the cotangent bundle, one-homogeneous on fibers. When this Hamiltonian function is smooth the associated Hamiltonian vector field $\vec{\boldsymbol{h}}$ is well-defined. The PMP asserts that the trajectories of the corresponding Hamiltonian system are the extremals of our control problem. The flow generated by the Hamiltonian field on the cotangent bundle is actually a direct generalization of the Riemannian geodesic flow in spite of the fact that the last flow was originally defined on the tangent bundle (or more precisely on the unitary tangent bundle). We now fix the level set $\mathcal{H} = h^{-1}(1)$ (the Hamiltonian analogue to the unitary Riemannian tangent bundle). Then, $\mathcal{H}_q = \mathcal{H} \cap T_q^*M$ is a curve in $T^*M$ and under certain regularity assumption (that we will be precised later on), this curve admits a natural parameter providing us with a vector field $\boldsymbol{v}_q$ on $\mathcal{H}_q$ and by consequence with a vertical vector field $\boldsymbol{v}$ on $\mathcal{H}$. Vector fields $\vec{\boldsymbol{h}}$ and $\boldsymbol{v}$ are feedback-invariant and $(\vec{\boldsymbol{h}}, \boldsymbol{v}, [\boldsymbol{v}, \vec{\boldsymbol{h}}])$ forms a moving frame on $\mathcal{H}$. It is thus quite natural to expect that the curvature may arise from some commutator relation of these fields. Indeed, the control curvature comes from the first nontrivial commutator relation between them: it is the coefficient $\kappa$ in the identity

$$\left[\vec{\boldsymbol{h}}, \left[\boldsymbol{v}, \vec{\boldsymbol{h}}\right]\right] = \kappa \boldsymbol{v}.$$

The above identity, due to A. Agrachev, is not trivial at all. It will be presented in Theorem 2.2.3.

In the second part of Chapter 2 we investigate some specific optimal control problems for which the curvature will be explicitly computed. In particular, we will see that the curvature $\kappa$ generalizes the Gaussian curvature (see Theorem 2.5.2) of a surface in the following sense: if the optimal control problem corresponds to the Riemannian geodesic problem then, $\kappa$ is the Gaussian curvature. The main difference between the Gaussian curvature of a surface and its control analogue is the following: whereas the Gaussian curvature is a function on the base manifold $M$, the control curvature $\kappa$ is a function on the level surface $\mathcal{H}$ which is a three dimensional manifold therefore $\kappa$ is a more complicated invariant. Anyway, as we shall see in the third part of Chapter 2, the classical comparison theorem on the occurrence of conjugate points remain valid in the control situation. That is, if the curvature is negative the optimal control problem does not admit conjugate points and if the curvature is positive then, as for Riemannian geometry, the Sturm comparison theorem leads to some estimates on the occurrence of conjugate points.

We want to point out that the knowledge of such an invariant is fundamental for the local and even global study of the Hamiltonian flow. Indeed, the curvature reveals very important information about the behavior of extremal trajectories and about the optimal synthesis of the problem WITHOUT solving any differential equation. Namely, the computation of the curvature requires only to compute certain polynomial of the partial derivative of the right-hand side of (1).

Chapter 3 deals with specific geometrical problems: Zermelo navigation problem and the corresponding dual problem.

In a first part we investigate Zermelo navigation which consists of finding the shortest path in time in a Riemannian manifold $(M, g)$ under the influence of a wind which is represented by a smooth vector field $\boldsymbol{X}$ on the considered manifold. This problem was firstly investigated by E. Zermelo itself in [43]. Dynamics of this problem are given by

$$\dot{q} = \boldsymbol{X}(q) + |u|_g, \quad q \in M, \quad u \in T_qM, \quad |u|_g < 1,$$

where $|\cdot|_g$ denotes the Riemannian norm for tangent vectors. The corresponding Hamiltonian function of PMP is

$$h(\lambda) = \langle \lambda, \boldsymbol{X} \rangle + |\lambda|_g,$$

where $|\cdot|_g$ has to be understood here as the co-norm on $T^*M$ dual to the Riemannian norm. Our goal in this problem is neither the study of the controllability, nor the detailed study of its optimal synthesis, but we want to make some "curvature" investigations. We explain in details the construction of the vertical vector field $\boldsymbol{v}$ and the local coordinate computation of the curvature for Zermelo navigation problem on the Euclidean plane $\mathbb{R}^2$. The coordinates expression of the curvature of Zermelo navigation problem on the Euclidean plane $\mathbb{R}^2$ is given by formula (3.7) of Chapter 3 which gives an idea of how much more complicated can be the control curvature compared to the Gaussian one. We will study in more details the case of a linear wind, i.e., the case in which the vector field that represents the wind distribution is a linear vector field. We will then see that the knowledge of the curvature is enough to prove the non-existence of conjugate point for certain particular linear fields. The total answer to the question about the existence of conjugate point in the case of a linear drift is contained in Theorem 3.1.4 which asserts that if the drift is linear then, there is no conjugate points along the extremal of Zermelo navigation problem on the $n$-dimensional Euclidean space $\mathbb{R}^n$. The two-dimensional version of this theorem was presented with a sketch of proof in the paper by U. Serres [40].

In a second part, we deal with the dual problem to Zermelo navigation problem. Without entering into details, we can say that the dual problem to Zermelo problem is the time-optimal for which the level set $\mathcal{H}$ of the Hamiltonian $h$ is the dual Riemannian unitary tangent bundle drifted by some one-form on $\Upsilon \in \Lambda^1(M)$. In other words, whereas Zermelo navigation problem was defined by its dynamics, i.e., by its curves of admissible velocities $\{u \mapsto \boldsymbol{f}(q, u) \mid u \in U\} \subset T_qM$, the dual problem to Zermelo is defined by its curves $\mathcal{H}_q \in T_q^*M$. The Hamiltonian function $h$ of the dual to Zermelo problem is given as the solution of the equation

$$|\lambda - h(\lambda)\Upsilon|_g = h(\lambda).$$

These two Zermelo problems are very important when they are considered together because, if the drift is different from one then, Zermelo navigation problem and its dual problem are equivalent, as it is shown in Proposition 3.3.1. This simple proposition shows in particular that these two problems have the same invariants.

In Chapter 4 we will return to more general features. We will give some new results for a problem as hold as the existence of geometric control theory itself: the problem of classification by feedback.

There are several previous approaches to this problem in particular, I wish to mention the approach of R. B. Gardner, W. F. Shadwick and G. R. Wilkens via Cartan's equivalence method see [23], [25], [24], [42]. Another point of view on this subject has been proposed by I. Kupka in [32]. B. Jakubczyk in [28] studies some "critical Hamiltonians" that allow to analyse feedback equivalence using tools of complex analysis. His method is particulary efficient in low dimensions. Other references are given at the beginning of Chapter 4.

In a first part we will build two new microlocal normal forms for two-dimensional control systems having a strongly convex sets of admissible velocities. One of these normal forms will be given around a regular extremal and the other around an abnormal. In particular, we will see how the microlocal normal form around the abnormal extremal shows that the control curvature, which was defined along normal extremals, can be smoothly extend to abnormals.

In a second part we present two new theorems dealing with flatness of control systems. We say that a control system is flat if it is feedback-equivalent to a system of the form $\dot{q} = \boldsymbol{f}(u)$. Whereas the flat Riemannian manifolds are characterized by the fact that their Gaussian curvature vanishes identically, we will see that a control system whose curvature is identically zero can be far from being flat. The first theorem presented there (Theorem 4.3.2) characterizes (in terms of feedback-invariants) the control systems that are feedback-equivalent to a control system having the property that the vector fields $\boldsymbol{f}$ and $\frac{\partial \boldsymbol{f}}{\partial u}$ commute. These systems, which are not necessarily flat, will also be parametrized. Theorem 4.3.3 characterizes flat controls systems. Theorem 4.3.2 and Theorem 4.3.3 are both stated in a completely intrinsic setting, however they provide checkable conditions for the characterization they give. This chapter will be conclude by some applications of these theorems on examples.

Chapter 5 is devoted to the study of some global properties of control systems of type (1). In a first part we will see how the classical Gauss-Bonnet theorem for Riemannian surfaces generalizes to two-dimensional control control systems. As we shall see this generalization does not holds for Zermelo problems. In a second part we will first generalize to control systems the following theorem due to E. Hopf ([26], 1948).

**Theorem.** *Let $M$ be a closed Riemannian surface of class $C^3$. If no geodesically conjugate points exist on $M$ the total curvature of $M$ must be negative or zero. In the latter case the Gaussian curvature must vanish everywhere on $M$.*

Then, we will try to answer the following natural question: *Considering a control system without conjugate points, does there exists a global reparametrization of the system such that the reparametrized curvature is negative?*

We will conclude this thesis with a discussion on a work in progress dealing with extremal flows on three-dimensional tori having negative control curvature.

# Chapter 1

# Preliminaries

The present chapter is mostly an introduction to the language, the notations and the classical results that will be used in this thesis.

In Section 1.1 we fix some notations of chronological calculus, a tool first developed by A. A. Agrachev and R. V. Gamkrelidze in [4], which is an operator calculus that allows, at least at the formal level, to work with nonlinear systems and flows as with linear ones.

Section 1.2 contains some basic elements of classical differential and symplectic geometry that will be used all along this thesis.

Finally, Section 1.3.1 formulates the Pontryagin Maximal Principle (PMP) which is the a first order optimality condition of optimal control theory. This is the analogue (actually the generalization) of the Euler Lagrange equation of the classical calculus of variations.

## 1.1 Chronological calculus

All results of this section can be found in [4, 6], so we will recall them without proof.

Let $M$ be a finite dimensional smooth manifold. Denote by $C^\infty(M)$ its algebra of smooth functions and by $\operatorname{Vec} M$ the space of smooth vector fields on $M$. We do assume that all smooth objects are of class $C^\infty$, unless otherwise specified.

The operator calculus, called chronological calculus, is based on the exponential representation of flows and thus, it essentially reflects their group properties. In order to build it, the main idea is to replace the smooth manifold $M$, which is a nonlinear object by its algebra of smooth functions $C^\infty(M)$, which is a linear object (although infinite-dimensional). Let us first recall that in chronological calculus a point $q$ on a smooth manifold $M$ is identified with a linear functional $q$, denoted by the same letter, on $C^\infty(M)$ which acts on functions as follows

$$\begin{aligned} q : C^\infty(M) &\to \mathbb{R} \\ a &\mapsto q \circ a = a(q). \end{aligned} \tag{1.1}$$

With this identification the topology and the smooth structure of the manifold $M$ are recovered from $C^\infty(M)$. In the same way, any diffeomorphism $P \in \operatorname{Diff} M$ is identified with the following automorphism $P$ of $C^\infty(M)$:

$$\begin{aligned} P : C^\infty(M) &\to C^\infty(M) \\ a &\mapsto Pa = a \circ P. \end{aligned} \tag{1.2}$$

Remark that the identification (1.2) is contravariant, i.e., the diffeomorphism $P_1 \circ P_2$ corresponds to the automorphism $P_2 \circ P_1$.

Consider now the following nonautonomous ODE on $M$

$$\begin{aligned} \dot q(t) &= \boldsymbol{f}_t(q(t)), \\ q(t_0) &= q_0, \quad q_0 \in M, \end{aligned} \tag{1.3}$$

where $t \mapsto \boldsymbol{f}_t$ is a locally bounded nonautonomus vector field on $M$. It is well known that for every $q_0$ in $M$ and every $t_0$ in $\mathbb{R}$ such a Cauchy problem admits a unique local Carathéodory solution $q(t, q_0)$. Namely, there exist an interval $I_{q_0} \subset \mathbb{R}$ neighborhood of $t_0$ and a unique absolutely continuous curve $t \mapsto q(t, q_0)$ such that the equation $\dot q(t, q_0) = \boldsymbol{f}_t(q(t, q_0))$ holds almost everywhere in $I_{q_0}$ and the initial condition $q(t_0, q_0) = q_0$ holds. The map

$$P_t : q_0 \mapsto q(t, q_0) \tag{1.4}$$

which associates with a $q_0 \in M$ the value of the solution of (1.3) evaluated at a fixed time $t \in I_{q_0}$ is a local diffeomorphism of $M$ (in a neighborhood of $q_0$) called the flow of $\boldsymbol{f}_t$ at time $t$.

Remark that the ODE (1.3) can be rewritten as a linear equation for absolutely continuous (with respect to $t$) families of functionals on $C^\infty(M)$ as

$$\begin{aligned} \dot q(t) &= q(t) \circ \boldsymbol{f}_t \\ q(t_0) &= q_0, \end{aligned}$$

which admits $q(t, q_0)$ as a unique solution. This implies that the flow $P_t$ defined in (1.4) is the unique solution (in the class of absolutely continuous flows on $M$) of the operator Cauchy problem

$$\dot P_t = P_t \circ \boldsymbol{f}_t, \tag{1.5}$$

$$P_{t_0} = \mathrm{Id}, \tag{1.6}$$

where Id denotes the identity operator. We called *right chronological exponential* this flow and denote it by

$$P_t = \overrightarrow{\exp} \int_{t_0}^t \boldsymbol{f}_\tau d\tau.$$

According to [4, 6], this notation is justified by the asymptotic Voltera series expansion

$$\overrightarrow{\exp} \int_{t_0}^t \boldsymbol{f}_\tau d\tau = \mathrm{Id} + \sum_{n=1}^{\infty} \int \cdots \int_{\Delta_n(t)} \boldsymbol{f}_{\tau_n} \circ \cdots \circ \boldsymbol{f}_{\tau_1} d\tau_n \cdots d\tau_1,$$

where
$$\Delta_n(t) = \{(\tau_1, \dots, \tau_n) \in \mathbb{R}^n) \mid t_0 \leqslant \tau_1 \leqslant \dots \leqslant \tau_n \leqslant t\}$$
is the $n$-dimensional simplex. Notice that by definition hold the following rules for the inverse and the composition of flows

$$\left(\overrightarrow{\exp}\int_{t_0}^{t} \boldsymbol{f}_\tau d\tau\right)^{-1} = \overrightarrow{\exp}\int_{t}^{t_0} \boldsymbol{f}_\tau d\tau,$$

$$\overrightarrow{\exp}\int_{t_0}^{t_1} \boldsymbol{f}_\tau d\tau \circ \overrightarrow{\exp}\int_{t_1}^{t} \boldsymbol{f}_\tau d\tau = \overrightarrow{\exp}\int_{t_0}^{t} \boldsymbol{f}_\tau d\tau,$$

which underscore their group properties. In the special case of an autonomous vector field $\boldsymbol{f} \in \operatorname{Vec} M$, its flow is called *exponential* and is denoted by $e^{t\,\boldsymbol{f}}$. In this case the asymptotic series for the exponential take the form

$$e^{t\,\boldsymbol{f}} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \boldsymbol{f}^n.$$

As usual, a smooth vector field $\boldsymbol{f} \in \operatorname{Vec} M$ is identified in a natural way with a derivation of the algebra $C^\infty(M)$. Namely we do the identification

$$\boldsymbol{f} \equiv L_{\boldsymbol{f}},$$

where the derivation $L_{\boldsymbol{f}}$ denotes the Lie derivative along the vector field $\boldsymbol{f}$. Using this identification, the *Lie bracket (or commutator)* of two vector fields $\boldsymbol{f}$, $\boldsymbol{g} \in \operatorname{Vec} M$ can be written
$$[\boldsymbol{f}, \boldsymbol{g}] = \boldsymbol{f} \circ \boldsymbol{g} - \boldsymbol{g} \circ \boldsymbol{f}. \tag{1.7}$$
As we know, the commutator $[\boldsymbol{f}, \boldsymbol{g}]$ is also a vector field. If a system of local coordinates is fixed on the manifold, the Lie bracket of two vector fields reads

$$[\boldsymbol{f}, \boldsymbol{g}] = \frac{\partial \boldsymbol{g}}{\partial q}\boldsymbol{f} - \frac{\partial \boldsymbol{f}}{\partial q}\boldsymbol{g},$$

where $\frac{\partial \boldsymbol{f}}{\partial q}$ and $\frac{\partial \boldsymbol{g}}{\partial q}$ denote the Jacobian matrix of $\boldsymbol{f}$ and $\boldsymbol{g}$ respectively in the chosen system of local coordinates. Remark that the Lie Bracket equips $\operatorname{Vec} M$ with a structure of Lie algebra.

For a vector field $\boldsymbol{f} \in \operatorname{Vec} M$ we define the operator $\operatorname{ad} \boldsymbol{f}$ from the space of vector fields onto itself by
$$\operatorname{ad} \boldsymbol{f}(\boldsymbol{g}) = [\boldsymbol{f}, \boldsymbol{g}].$$

The group $\operatorname{Diff} M$ of smooth diffeomorphism of $M$ acts naturally on $\operatorname{Vec} M$ associating to any vector field $\boldsymbol{f}$ a vector field denoted by $\operatorname{Ad} P\,\boldsymbol{f}$ according to the formula
$$\operatorname{Ad} P\,\boldsymbol{f} = P_*^{-1}\boldsymbol{f},$$

where $P_*^{-1}$ denotes the standard pushforward operator, i.e., the operator defined by

$$(P_*^{-1}(\boldsymbol{f}))(q) = P_{*P(q)}^{-1}\boldsymbol{f}(P(q)),$$

where the notation $P_{*q}$ is used, along with $P_*$ when no confusion is possible, for the differential of a mapping $P$ at a point $q$. It is easily seen from (1.7) that

$$\mathrm{Ad}\, P\,[\boldsymbol{f}, \boldsymbol{g}] = [\mathrm{Ad}\, P\, \boldsymbol{f}, \mathrm{Ad}\, P\, \boldsymbol{g}].$$

The relation

$$\frac{d}{dt}\mathrm{Ad}\,\overrightarrow{\exp}\int_{t_0}^{t} \boldsymbol{f}_\tau d\tau = \mathrm{Ad}\,\overrightarrow{\exp}\int_{t_0}^{t} \boldsymbol{f}_\tau d\tau \circ [\boldsymbol{f}_t, \boldsymbol{g}],$$

which holds for almost every $t$, justifies the relation

$$\overrightarrow{\exp}\int_{t_0}^{t} \mathrm{ad}\, \boldsymbol{f}_\tau d\tau = \mathrm{Ad}\,\overrightarrow{\exp}\int_{t_0}^{t} \boldsymbol{f}_\tau d\tau.$$

For the particular case of an autonomous vector field, we write

$$e^{t\,\mathrm{ad}\,\boldsymbol{f}} = \mathrm{Ad}\, e^{t\,\boldsymbol{f}}.$$

Notice that, for $\boldsymbol{f}$, $\boldsymbol{g} \in \mathrm{Vec}\, M$, $a \in C^\infty(M)$ and $P \in \mathrm{Diff}\, M$, we have

$$\begin{aligned}\mathrm{ad}\, \boldsymbol{f}(a\boldsymbol{g}) &= (L_{\boldsymbol{f}}a)\mathrm{ad}\, \boldsymbol{f}(\boldsymbol{g}),\\ \mathrm{Ad}\, P(a\boldsymbol{f}) &= (Pa)\mathrm{Ad}\, P\boldsymbol{f}.\end{aligned}$$

The formalism described in this section for non autonomous vector field will essentially be used in Section 4.3 of Chapter 4.

## 1.2 Elements of geometry

This section presents some basics elements of exterior calculus, symplectic and Riemannian geometry and two very classical theorems that will be needed in the sequel.

### Basic elements of exterior calculus and symplectic geometry

If $M$ is a $n$-dimensional smooth manifold, we respectively denote by $T_qM$ and $T_q^*M$ the tangent and cotangent linear spaces to $M$ at point $q \in M$. The sets,

$$TM = \bigcup_{q\in M} T_qM, \quad T^*M = \bigcup_{q\in M} T_q^*M,$$

are respectively called the tangent and cotangent bundle of $M$. These sets have a natural structure of $2n$-dimensional smooth manifold and the canonical projections from $TM$ and $T^*M$ to $M$ are linear fibration maps.

Denote by $\Lambda^k(M)$ the super-commutative algebra of differential forms of degree $k$ on $M$. We shall denote by the angle bracket, that is as a scalar product the action of one-forms on vector fields. Namely, if $\omega \in \Lambda^1(M)$ and $\boldsymbol{f} \in \operatorname{Vec} M$ we have

$$\langle \omega, \boldsymbol{f} \rangle = \omega(\boldsymbol{f}) = i_{\boldsymbol{f}}\omega,$$

where we have denoted by $i_{\boldsymbol{f}}\omega$ the interior product of the one-form $\omega$ with the vector field $\boldsymbol{f}$. Recall that the cotangent bundle $T^*M$ is endowed with a canonical structure of symplectic manifold given by the exterior derivative of the Liouville one-form. Denote by $\pi$ the canonical projection from $T^*M$ to $M$

$$\begin{aligned} \pi : T^*M &\to M \\ \lambda &\mapsto q, \quad \lambda \in T_q^*M. \end{aligned}$$

The *Liouville (or tautological) one-form* $s \in \Lambda^1(T^*M)$ is defined as follows:

$$s_\lambda = \lambda \circ \pi_*,$$

where $\pi_*$ denotes the differential of $\pi$. More precisely, if $w \in T_\lambda(T^*M)$ is a tangent vector to $T^*M$ at $\lambda$ the action of the Liouville one-form at $\lambda$ on $w$ is

$$\langle s_\lambda, w \rangle = \langle \lambda, \pi_* w \rangle .$$

The *canonical symplectic structure* $\sigma \in \Lambda^2(T^*M)$ on $T^*M$ is defined to be the exterior differential of the Liouville one-form:

$$\sigma = ds,$$

which is a non degenerated closed two-form. Recall that if $\lambda = (p, q)$ is a system of canonical coordinates on $TM$, the Liouville form is given by $s_\lambda = \sum_{i=1}^n p_i dq_i$, and the symplectic structure by $\sigma_\lambda = \sum_{i=1}^n dp_i \wedge dq_i$. The cotangent bundle $T^*M$ with the canonical symplectic structure $\sigma \in \Lambda^2(T^*M)$ is the most important example of a symplectic manifold.

The following two identities, well known as *Cartan's formulas*, are very important for computations. The first one shows how to compute the Lie derivative of a differential form of an arbitrary order in a very simple manner. The second gives the action of the exterior derivative of a one-form on a pair of vector fields. These formulas read

$$L_{\boldsymbol{f}} = d \circ i_{\boldsymbol{f}} + i_{\boldsymbol{f}} \circ d \tag{1.8}$$

$$d\omega(\boldsymbol{f}, \boldsymbol{g}) = L_{\boldsymbol{f}} \langle \omega, \boldsymbol{g} \rangle - L_{\boldsymbol{g}} \langle \omega, \boldsymbol{f} \rangle - \langle \omega, [\boldsymbol{f}, \boldsymbol{g}] \rangle . \tag{1.9}$$

**Definition 1.2.1.** *Two local basis $(\omega_1, \ldots, \omega_n) \subset \Lambda^1(M)$ and $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n) \subset \operatorname{Vec} M$ are said to be dual basis if $\langle \omega_i, \boldsymbol{f}_j \rangle = \delta_j^i$ for every $i$, $j \in \{1, \ldots, n\}$, where the $\delta_j^i$ denote the Kronecker symbols.*

The next lemma exhibits the duality between $\operatorname{Vec} M$ and $\Lambda^1(M)$.

**Lemma 1.2.2.** *Let* $(\omega_1, \dots, \omega_n) \subset \Lambda^1(M)$ *and* $(\boldsymbol{f}_1, \dots, \boldsymbol{f}_n) \subset \operatorname{Vec} M$ *be two local dual basis. If for every* $k \in \{1, \dots, n\}$ *we have*

$$d\omega_k = \sum_{i<j} c_{ij}^k \, \omega_i \wedge \omega_j,$$

*then, for every* $i,\ j \in \{1, \dots, n\}$ *we have*

$$\left[\boldsymbol{f}_i, \boldsymbol{f}_j\right] = -\sum_{k=1}^{n} c_{ij}^k \boldsymbol{f}_k.$$

***Proof.*** Since the vector fields $\boldsymbol{f}_1, \dots, \boldsymbol{f}_n$ form a basis, we have, for all $i, j \in \{1, \dots, n\}$

$$[\boldsymbol{f}_i, \boldsymbol{f}_j] = \sum_{l=1}^{n} a_{ij}^l \boldsymbol{f}_l.$$

On the one hand, a direct computation shows that

$$d\omega_k(\boldsymbol{f}_i, \boldsymbol{f}_j) = c_{ij}^k,$$

On the other hand, using Cartan's formula (1.9) we get

$$\begin{aligned} d\omega_k(\boldsymbol{f}_i, \boldsymbol{f}_j) &= L_{\boldsymbol{f}_i} \left\langle \omega_k, \boldsymbol{f}_j \right\rangle - L_{\boldsymbol{f}_j} \left\langle \omega_k, \boldsymbol{f}_i \right\rangle - \left\langle \omega_k, [\boldsymbol{f}_i, \boldsymbol{f}_j] \right\rangle \\ &= - \left\langle \omega_k, [\boldsymbol{f}_i, \boldsymbol{f}_j] \right\rangle = -\langle \omega_k, \sum_{l=1}^{n} a_{ij}^l \boldsymbol{f}_l \rangle = a_{ij}^k, \end{aligned}$$

hence, $a_{ij}^k = -c_{ij}^k$ for all $k$ which completes the proof. ■

The next lemma describes the action of the flow of a vector field on a moving frame.

**Lemma 1.2.3.** *Let the vector fields* $\boldsymbol{f}_1, \dots, \boldsymbol{f}_n$ *form a moving frame on* $M$*. Take a vector field* $\boldsymbol{g} \in \operatorname{Vec} M$*. Let the operator* $\operatorname{ad} \boldsymbol{g}$ *have the matrix* $A = (a_j^i)$*:*

$$\operatorname{ad} \boldsymbol{g}(\boldsymbol{f}_j) = \left[\boldsymbol{g}, \boldsymbol{f}_j\right] = \sum_{i=1}^{n} a_j^i \boldsymbol{f}_i, \quad a_j^i \in C^\infty(M).$$

*Then the matrix* $\Gamma(t) = (\gamma_j^i(t))$ *of the operator* $e^{t \operatorname{ad} \boldsymbol{g}}$ *in the moving frame* $(\boldsymbol{f}_1, \dots, \boldsymbol{f}_n)$*:*

$$e^{t \operatorname{ad} \boldsymbol{g}} \boldsymbol{f}_j = \sum_{i=1}^{n} \gamma_j^i(t) \boldsymbol{f}_i, \tag{1.10}$$

*is the solution of the Cauchy problem*

$$\dot{\Gamma}(t) = \Gamma(t) A(t), \tag{1.11}$$

$$\Gamma(0) = \operatorname{Id}, \tag{1.12}$$

*where* $A(t) = (e^{t \boldsymbol{g}} a_j^i)$.

***Proof.*** The initial condition (1.12) is obvious. In order to derive the matrix equation (1.11), we differentiate identity (1.10) with respect to $t$. This gives

$$\begin{aligned}\sum_{i=1}^{n} \dot{\gamma}_j^i \boldsymbol{f}_i &= \frac{d}{dt}\left(e^{t\,\mathrm{ad}\,\boldsymbol{g}} \boldsymbol{f}_j\right) = e^{t\,\mathrm{ad}\,\boldsymbol{g}} \left[\boldsymbol{g}, \boldsymbol{f}_j\right] = e^{t\,\mathrm{ad}\,\boldsymbol{g}} \left(\sum_{i=1}^{n} a_j^i \boldsymbol{f}_i\right) \\ &= \mathrm{Ad}\, e^{t\boldsymbol{g}} \left(\sum_{i=1}^{n} a_j^i \boldsymbol{f}_i\right) = \sum_{i=1}^{n} \left(e^{t\boldsymbol{g}} a_j^i\right) e^{t\,\mathrm{ad}\,\boldsymbol{g}} \boldsymbol{f}_i = \sum_{i,k=1}^{n} \left(e^{t\boldsymbol{g}} a_j^i\right) \gamma_i^k \boldsymbol{f}_k,\end{aligned}$$

from which it follows the ODE. ■

### Riemannian structures

A *Riemannian structure* or *Riemannian metric* on a smooth manifold $M$ is a covariant two-tensor field $g$ that is symmetric and positive definite. A Riemannian structure thus determines an inner product on each tangent space $T_qM$, which we write

$$\langle v, w\rangle_g = g(v, w), \quad v,\ w \in T_qM.$$

As in Euclidean geometry, if $q$ is a point in a Riemannian manifold $(M, g)$, we define the *norm* of any tangent vector $v \in T_qM$ to be

$$|v|_g = \sqrt{g(v, v)}.$$

One elementary but important property of Riemannian structures is that they allow us to convert vectors into covectors and vice versa. Given a Riemannian structure $g$ on $M$ we identify $TM$ and $T^*M$ as follows. We first define a map from $TM$ to $T^*M$ by

$$v \mapsto v^{\flat_g} = \langle v, \cdot\rangle_g\,. \tag{1.13}$$

Since $g$ is definite the operator $\flat_g$ is invertible and we denote its inverse by $\lambda \mapsto \lambda^{\sharp_g}$. When no confusion is possible we may also write $\flat$ (respectively $\sharp$) instead of $\flat_g$ (respectively $\sharp_g$). This identification naturally induces an inner product (and consequently a norm) on each cotangent space $T_q^*M$ by

$$\langle \lambda, \mu\rangle_g = \left\langle \lambda^\sharp, \mu^\sharp\right\rangle_g, \quad \lambda,\ \mu \in T_q^*M.$$

Let $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) \subset \mathrm{Vec}\, M$ be a local orthonormal frame for $g$ (i.e., $\langle \boldsymbol{e}_i, \boldsymbol{e}_j\rangle_g = \delta_{ij}$). Then the coframe $(\boldsymbol{e}_1^*, \ldots, \boldsymbol{e}_2^*) \subset \Lambda^1(M)$ dual to $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ is also orthonormal for $g$ and for $\lambda \in T_q^*M$ we have

$$|\lambda|_g = \sqrt{\langle \lambda, \boldsymbol{e}_1\rangle^2 + \cdots + \langle \lambda, \boldsymbol{e}_n\rangle^2}.$$

### Euler's theorem

Let us recall here a technical ingredient that is fundamental in the study of homogeneous Hamiltonians. The next result is well-known as *Euler's theorem.*

**Theorem 1.2.4.** *Let $h$ be a real-valued function on a $n$-dimensional vector space differentiable away from the origin. Then the two following statements are equivalent:*

- *$h$ is positively homogeneous of degree $r$. That is,*

$$h(c\lambda) = c^r h(\lambda) \quad for\ all\ c > 0.$$

- *The radial directional derivative of $h$ is $r$ times $h$. That is,*

$$d_\lambda h(\lambda) = rh(\lambda). \tag{1.14}$$

Relation (1.14) is well-known as Euler Identity.

### Sturm Comparison Theorem

In this section we just state and recall the proof a classical theorem due to Sturm (1836): the Sturm Comparison Theorem for ordinary differential equations. This theorem will be used in the next chapter in order to get estimate about conjugate points along the extremals of optimal control problems.

**Theorem 1.2.5** (Sturm Comparison Theorem)**.** *Let $u(t)$ and $v(t)$ be respectively solutions to*

$$\begin{aligned} u''(t) + A(t)u(t) = 0, \quad u(0) = 0,\ u'(0) = 1, \\ v''(t) + B(t)v(t) = 0, \quad v(0) = 0,\ v'(0) = 1. \end{aligned}$$

*Suppose that $A(t) \geqslant B(t)$. If $a$ and $b$ are the first zeros, after $t = 0$, of $u(t)$ and $v(t)$, respectively, then*

$$a \leqslant b. \tag{$i$}$$

*Furthermore, for $t_0$, $t_1$ satisfying $0 < t_0 < t_1 < a$,*

$$v(t_1)u(t_0) \geqslant u(t_1)v(t_0) \text{ and } v(t_1) \geqslant u(t_1). \tag{$ii$}$$

*((iii) If $A(t) > B(t)$, then $a < b$, $v(t_1)u(t_0) > u(t_1)v(t_0)$, and $v(t_1) > u(t_1)$.)*

***Proof.*** $(i)$. Let us prove it by contradiction. Since $u'(0) = v'(0) = 1$, $u(t), v(t) \geqslant 0$ for all $t$, $0 < t < a$. Assume that $a > b$. We have

$$0 = \int_0^b u(v'' + Bv) - v(u'' + Au)\,dt = [uv' - vu']_0^b + \int_0^b (B - A)uv\,dt.$$

Since $A(t) \geqslant B(t)$, $(B-A)uv \leqslant 0$ on $[0,b]$, so $[uv' - vu']_0^b = u(b)v'(b) \leqslant 0$. But $u(t) > 0$ and $v'(b) < 0$ (otherwise $v(b) = v'(b) = 0$, and $v(t)$ is identically equal to zero which is impossible in view of the hypothesis $v'(0) = 1$) which leads to a contradiction. Thus $a \leqslant b$.
$(ii)$. Since

$$\begin{aligned} 0 &= \int_0^t u(v'' + Bv) - v(u'' + Au)\, dt \\ &= [uv' - vu']_0^t + \int_0^t (B-A)uv\, dt \leqslant [uv' - vu']_0^t, \end{aligned}$$

we get by integrating

$$\frac{d}{dt}(\log v(t)) \geqslant \frac{d}{dt}(\log u(t)).$$

Thus, if $0 < t_0 \leqslant t_1 < a$,

$$\log\left(v(t_1)/v(t_0)\right) = \int_{t_0}^{t_1} \frac{d}{dt}(\log v(t)) \geqslant \int_{t_0}^{t_1} \frac{d}{dt}(\log u(t)) = \log\left(u(t_1)/u(t_0)\right),$$

i.e., $v(t_1)u(t_0) \geqslant u(t_1)v(t_0)$. We already know that $u(0) = v(0) = 0$. Using the rule of de L'Hôpital, we get that $\lim_{t_0 \to 0} v(t_0)/u(t_0) = v'(0)/u'(0) = 1$ which combined with condition $(i)$ implies $v(t_1) \geqslant u(t_1)$.
$(iii)$. An analogous reasoning with strict inequalities gives the result. ■

## 1.3 Pontryagin Maximum Principle

In this section we present a geometric version of the fundamental first order and necessary condition for optimality in control problems: the Pontryagin Maximum Principle (PMP). The PMP is due to the Russian mathematicians V. G. Boltyanskij, R. V. Gamkrelidze and L. S. Pontryagin who presented the first proof in [16] in the case of optimal control problems in $\mathbb{R}^n$.

A time-optimal control problem on a manifold $M$ is a dynamical system with boundary conditions whose dynamic laws depend on a parameter $u$ belonging to a set $U$. The parameter $u$ is called the control parameter and the set $U$ the set of admissible control parameters. Such a time-optimal control problem with general boundary condition for the initial point takes the form

$$\dot{q} = \boldsymbol{f}(q,u), \quad q \in M, \quad u \in U \tag{1.15}$$
$$q(0) \in N_0, \quad q(t_1) \in N_1, \tag{1.16}$$
$$t_1 \to \min, \tag{1.17}$$

where $N_0$ and $N_1$ are given immersed submanifolds of the state space $M$ and $\boldsymbol{f}(\cdot,u) \in \operatorname{Vec} M$. In particular $N_0$ and $N_1$ can be two fixed points of the base manifold. An admissible control function is a locally bounded mapping

$$u : t \mapsto u(t) \in U.$$

In order to insure the existence of Carathéodory solutions to control system (1.15), we assume, for the right-hand side of the control system (1.15) that:

$$(u, q) \mapsto \boldsymbol{f}(u, q) \in C^0(M \times U).$$

In words, the time-optimal control problem (1.15)−(1.17) asks to minimize the final time $t_1$ among all admissible controls functions $t \mapsto u(t)$, $t \in [0, t_1]$, for which the solution to the Cauchy problem (1.15) satisfies the boundary conditions (1.16).

We associate to the control system (1.15) a control dependent Hamiltonian function according to the formula

$$h_u(\lambda) = \langle \lambda, \boldsymbol{f}(q, u) \rangle , \quad \lambda \in T_q^*M.$$

We denote by $\vec{\boldsymbol{h}}_u$ the Hamiltonian vector field on $T^*M$ associated to the control dependent Hamiltonian $h_u$. Recall that this vector fields is defined by the rule

$$i_{\vec{\boldsymbol{h}}_u} \sigma = -dh_u,$$

where $\sigma$ is the standard symplectic two-form on $T^*M$.

Suppose now that we want to solve the time-optimal problem (1.15)−(1.17), then the following holds.

**Theorem 1.3.1** (PMP)**.** *Let an admissible control $u^*(t)$ be time-optimal. Then, there exists a Lipschitzian curve*

$$\lambda_t \in T^*M, \quad \lambda_t \neq 0, \quad t \in [0, t_1],$$

*such that the following conditions hold for almost all $t \in [0, t_1]$:*

$$\dot{\lambda}_t = \vec{\boldsymbol{h}}_{u^*(t)}(\lambda_t), \tag{1.18}$$

$$h_{u^*(t)}(\lambda_t) = \max_{u \in U} h_u(\lambda_t), \tag{1.19}$$

$$h_{u^*(t)}(\lambda_t) = 0 \quad \text{or} \quad > 0, \tag{1.20}$$

$$\lambda_0 \perp T_{\pi(\lambda_0)} N_0, \quad \lambda_{t_1} \perp T_{\pi(\lambda_{t_1})} N_1. \tag{1.21}$$

**Remark 1.3.2.** (1) Condition (1.18) of PMP says that the solutions $q_{u^*(\cdot)}(\cdot)$ of the optimal control problem (1.15)−(1.17) on $M$ are just projections of the solutions of the Hamiltonian system $\dot{\lambda} = \vec{\boldsymbol{h}}_{u^*}(\lambda)$ on $T^*M$.
(2) There are two distinct possibilities for condition (1.20) of PMP:

- if $h_{u^*(t)}(\lambda_t) > 0$, then the curve $\lambda_t$ is called a *normal (or regular) extremal.* In this case, one can normalize $\lambda_t$ in such a way that condition (1.20) of PMP becomes $h_{u^*(t)}(\lambda_t) = 1$,

- if $h_{u^*(t)}(\lambda_t) = 0$, then the curve $\lambda_t$ is called an *abnormal extremal*.

(3) Conditions (1.21) are called *transversality conditions*. Since any linear functional acts naturally on subspaces by restriction, transversality conditions (1.21) respectively read

$$\langle \lambda_0, v \rangle = 0, \quad \forall\ v \in T_{\pi(\lambda_0)} N_0,$$
$$\langle \lambda_{t_1}, v \rangle = 0, \quad \forall\ v \in T_{\pi(\lambda_{t_1})} N_1.$$

**Remark 1.3.3.** If instead of minimizing time we ask to maximize time then, the PMP holds with condition (1.20) replaced by

$$h_{u^*(t)}(\lambda_t) = 0 \quad \text{or} \quad < 0.$$

# Chapter 2

# Curvature

In classical Riemannian geometry the Ricci curvature tensor of a manifold reflects intrinsic properties of the geodesic flow, i.e., properties that do not depend on the choice of local coordinates. For example, the geodesics of the surface have no conjugate points if the curvature is non-positive. Indeed, these geodesics are just extremals of the very particular time-optimal control problem whose dynamics are given by $\dot{q}(t) = \sum_{i=1}^{n} u_i \boldsymbol{e}_i(q)$, with $\sum_{i=1}^{n} u_i^2 = 1$, where $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ forms a local orthonormal basis of the Riemannian structure of the manifold.

Our goal in this chapter is to generalize the classical notion of Gaussian curvature of surfaces to two-dimensional smooth optimal control problems using the Cartan's moving frame method. The notion of curvature tensor along regular extremals of Hamiltonian and control systems was first introduced in [5] by A. A. Agrachev and R. V. Gamkrelidze with a purely variational approach by means of Jacobi curves which are curves in the Lagrangian Grassmannian. Then, still using a variational approach, the case of singular extremal was treated in [1] by A. A. Agrachev and the geometry of Jacobi curves for regular and abnormal extremals was studied in [8], [9].

Here we will not deal with Jacobi curves but use instead the moving frame method in order to construct a feedback-invariant frame associated to our optimal control problem and provide a very geometric definition of the curvature function for two-dimensional control systems by means of Lie brackets. We will then see that the "control" analogue to Gaussian curvature reflects similar properties and give some examples for the computation of the curvature in local coordinates.

We want to point out that the knowledge of such an invariant is really advantageous because it reveals fundamental information about the behavior of extremal trajectories and about the optimal synthesis without solving any differential equation.

## 2.1 Feedback equivalence of control systems

Consider the following time-optimal control problem:

$$\dot{q} = \boldsymbol{f}(q, u), \quad q \in M, \quad u \in U, \tag{2.1}$$

$$q(0) = q_0, \quad q(t_1) = q_1, \tag{2.2}$$

$$t_1 \to \min \text{ (or } \max), \tag{2.3}$$

where $M$ and $U$ are finite dimensional, connected, smooth manifolds. Let $\tilde{\boldsymbol{f}}(\tilde{q}, \tilde{u})$, $(\tilde{q}, \tilde{u}) \in \tilde{M} \times \tilde{U}$, be the right-hand side of another control system of type (2.1), where $\tilde{M}$ and $\tilde{U}$ are smooth connected manifold such that $\dim \tilde{M} = \dim M$ and $\dim \tilde{U} = \dim U$. We say that the two considered systems are *feedback-equivalent* if there exists a diffeomorphism (called feedback transformation) $\Theta : M \times U \to \tilde{M} \times \tilde{U}$ of the form

$$\Theta(q, u) = (\phi(q), \psi(q, u)) \tag{2.4}$$

which transforms the first system into the second, i.e., such that

$$\phi_{*q}(\boldsymbol{f}(q, u)) = \tilde{\boldsymbol{f}}(\phi(q), \psi(q, u)).$$

The diffeomorphism $\phi$ plays the role of a change of coordinates in the state space $M$, and $\psi$ called *pure feedback transformation* reparametrizes the set of controls $U$ in a way depending on the state variable $q \in M$. Our aim is to provide feedback invariants for the control system (2.1), i.e., invariants of the action of the group of feedback transformations, when the manifold $M$ is of dimension two and the control set $U$ of dimension one.

Unless otherwise specified, we suppose from now that $\dim M = 2$, $\dim U = 1$.

## 2.2 Curvature of two-dimensional optimal control problems

For the control system (2.1) we do the following regularity assumptions on the curves of admissible velocities

$$\boldsymbol{f}(q, u) \wedge \frac{\partial \boldsymbol{f}(q, u)}{\partial u} \neq 0 \tag{2.5}$$

$$\frac{\partial \boldsymbol{f}(q, u)}{\partial u} \wedge \frac{\partial^2 \boldsymbol{f}(q, u)}{\partial u^2} \neq 0, \quad q \in M, \ u \in U. \tag{2.6}$$

Condition (2.5) implies that control system (2.1) does not admit abnormal extremals and condition (2.6) means that the *curves of admissible velocities* (also called *indicatrix*) at the point $q \in M$

$$\mathcal{S}_q = \{\boldsymbol{f}(q, u) \mid u \in U\} \subset T_q M$$

are either strongly convex or strongly concave.

A classical example of such a system is given by the control system corresponding to the geodesic problem on a two-dimensional Riemannian manifold:

$$\dot{q} = \cos u \, \boldsymbol{e}_1(q) + \sin u \, \boldsymbol{e}_2(q), \quad u \in S^1, \tag{2.7}$$

where $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ forms a local orthonormal frame of the Riemannian structure on the manifold $M$. Recall that for such a Riemannian manifold the associated Riemannian metric is given in local coordinates $q = (q_1, q_2)$ on $M$ by $(\boldsymbol{e}\boldsymbol{e}^\dagger)^{-1}$ where $\boldsymbol{e}$ is the matrix representation of the vector fields $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ in the coordinates system $(\frac{\partial}{\partial q_1}, \frac{\partial}{\partial q_2})$.

We introduce the linear in fibers control-dependent Hamiltonian

$$h_u(\lambda) = \langle \lambda, \boldsymbol{f}_u(q) \rangle, \quad q \in M, \quad \lambda \in T_q^* M,$$

and the maximized Hamiltonian of PMP

$$h(\lambda) = \max_{u \in U} h_u(\lambda), \tag{2.8}$$

which is a function on the cotangent bundle $T^*M$.

We suppose that the maximized Hamiltonian $h$ is defined in a domain in $T^*M$ under consideration and that for any $\lambda$ in this domain the maximum in (2.8) is attained for a unique $u \in U$ so that the convexity condition (2.6) implies the smoothness of $h$ in this domain and strong convexity (or concavity) on the fibers of $T^*M$.

Because the Hamiltonian functions $h_u$ are linear on fibers, $h$ is positively homogeneous of degree one on fibers so that the we can restrict our study to the level surface

$$\mathcal{H} = h^{-1}(\epsilon), \quad \epsilon = \pm 1.$$

We also denote the intersection with a fiber

$$\mathcal{H}_q = \mathcal{H} \cap T_q^* M.$$

Notice that, because of being independent of $u$, the Hamiltonian $h$ is a feedback-invariant function. Thus, all objects constructed from Hamiltonian $h$ through intrinsic relations, i.e., relations that do not depend on the choice of local coordinates in $\mathcal{H}$, will also be feedback invariants. As a consequence the level surface $\mathcal{H}$ and the fibers $\mathcal{H}_q$, $q \in M$ are also feedback-invariant.

The following result is well known, but for convenience of the reader we shall supply a simple proof. Denote by $\pi$ the canonical projection

$$\pi : T^*M \to M.$$

Let $\omega$ be the restriction on $\mathcal{H}$ of the Liouville one form $s_\lambda = \lambda \circ \pi_*$ of $T^*M$ and let $\vec{\boldsymbol{h}}$ be the Hamiltonian vector field associated to $h$.

**Lemma 2.2.1.** *The Hamitonianity of vector field $\vec{\boldsymbol{h}}$ on the hypersurface $\mathcal{H} \subset T^*M$ is characterized by:*

$$\left\langle \omega, \vec{\boldsymbol{h}} \right\rangle = \epsilon, \tag{2.9}$$

$$L_{\vec{\boldsymbol{h}}}\omega = 0. \tag{2.10}$$

***Proof.*** Relation (2.9) is a direct consequence of Euler identity (1.14) for homogeneous functions. Namely, for $\lambda \in \mathcal{H}$ we have:

$$\left\langle \omega, \vec{\boldsymbol{h}} \right\rangle\Big|_{\lambda} = \left\langle \lambda \circ \pi_*, \vec{\boldsymbol{h}}(\lambda) \right\rangle = d_\lambda h(\lambda) = h(\lambda) = \epsilon.$$

For relation (2.10) we use the first Cartan's identity (1.8) and the previous relation which gives:

$$L_{\vec{\boldsymbol{h}}}\omega = (i_{\vec{\boldsymbol{h}}} \circ d + d \circ i_{\vec{\boldsymbol{h}}})\omega = \sigma|_{\mathcal{H}}(\vec{\boldsymbol{h}}, \cdot) + d(\langle \omega, \vec{\boldsymbol{h}} \rangle) = -dh|_{\mathcal{H}} = 0,$$

which completes the proof. ■

**Remark 2.2.2.** Relation (2.10), well-known as Liouville's lemma, shows that the flow of a Hamiltonian field associated to a homogeneous Hamiltonian function preserves the Liouville one-form restricted to a regular level set of the Hamiltonian, i.e., $e^{t\vec{\boldsymbol{h}}*}\omega = \omega$.

We are now ready to construct a feedback-invariant moving frame on the three-dimensional surface $\mathcal{H}$ which will enable us to derive an ODE − the Jacobi equation in the moving frame − on conjugate time of our two-dimensional optimal control problem.

We start with the construction of a vertical vector field tangent to the curve $\mathcal{H}_q$. To do so, let us introduce some local trivialization map $\mathcal{H} \to M$ of fiber $\mathcal{H}_q$, i.e., a parametrization of $\mathcal{H}_q$ by angle $\theta$ providing us with an identification:

$$\mathcal{H} \cong \{\theta\} \times M = \mathcal{H}_q \times M. \tag{2.11}$$

Through this identification, the restriction $\omega$ on $\mathcal{H}$ of the Liouville form of $T^*M$ can then be viewed as a family of one-forms $\{\omega_\theta\}_{\theta \in \mathcal{H}_q}$ on the manifold $M$ parametrized by the angle $\theta$. Because $h$ is smooth and satisfies the condition of strong convexity on fibers, the level set $\mathcal{H}$ is regular. Therefore, $\omega$ is a contact form on $\mathcal{H}$ which implies that

$$\omega \wedge d\omega \neq 0.$$

But the differential $d\omega$ can be rewritten in terms of parameter $\theta$ as

$$d\omega = d\theta \wedge \frac{\partial \omega_\theta}{\partial \theta} + d\omega_\theta. \tag{2.12}$$

Thus,

$$\omega \wedge d\omega = \omega_\theta \wedge d\theta \wedge \frac{\partial \omega_\theta}{\partial \theta} + \underbrace{\omega_\theta \wedge d\omega_\theta}_{=0} \neq 0, \tag{2.13}$$

which shows in particular that $(\omega_\theta, \frac{\partial \omega_\theta}{\partial \theta})$ is a frame of horizontal one-forms on the base manifold $M$. The decomposition of the second derivative $\frac{\partial^2 \omega_\theta}{\partial \theta^2}$ in this frame reads

$$\frac{\partial^2 \omega_\theta}{\partial \theta^2} = -a(\theta)\omega_\theta + b(\theta)\frac{\partial \omega_\theta}{\partial \theta}.$$

The curve $\mathcal{H}_q$ being strongly convex (or concave), we have

$$a(\theta) \neq 0.$$

If we make a change of parameter in $\mathcal{H}_q$ via the relation $\theta = \Theta(\alpha)$, the formula above reads:

$$\frac{\partial^2 \omega_\alpha}{\partial \alpha^2} = -a\big(\Theta(\alpha)\big)\left(\frac{\partial \Theta}{\partial \alpha}\right)^2 \omega_\alpha + \left(b\big(\Theta(\alpha)\big)\frac{\partial \Theta}{\partial \alpha} + \frac{\frac{\partial^2 \Theta}{\partial \alpha^2}}{\frac{\partial \Theta}{\partial \alpha}}\right)\frac{\partial \omega_\alpha}{\partial \alpha}, \tag{2.14}$$

which shows in particular that, up to translation and orientation, i.e., up to a feedback transformation of the form $\theta \mapsto \pm\theta + \psi(q)$, there exists a unique parameter $\theta$ such that

$$\frac{\partial^2 \omega_\theta}{\partial \theta^2} = -\epsilon\omega_\theta + b(\theta)\frac{\partial \omega_\theta}{\partial \theta}, \quad \epsilon = \pm 1.$$

We fix such a parameter $\theta$. Hence, it provides us with a vector field $\boldsymbol{v}_q = \frac{\partial}{\partial \theta}$ on $\mathcal{H}_q$. By consequence, fixing such a parameter $\theta$ in each fiber $\mathcal{H}_q$ we can define the corresponding vertical vector field $\boldsymbol{v}$ on $\mathcal{H}$ by:

$$\boldsymbol{v} = \frac{\partial}{\partial \theta}.$$

In invariant terms vector field $\boldsymbol{v}$ is characterized by the fact that it is, up to sign, the unique vector field on $\mathcal{H}$ such that

$$L_{\boldsymbol{v}}^2 \omega = -\epsilon\omega + bL_{\boldsymbol{v}}\omega,$$

where $b$ is by definition a feedback-invariant smooth function on $\mathcal{H}$. Actually the function $b$ is the feedback invariant of our control system that characterizes Riemannian and Lorentzian problems. Namely control problem (2.1) defines a Riemannian (respectively Lorentzian) geodesic problem if and only if the invariant $b$ is identically equal to zero and the curves $\mathcal{H}_q$ are strongly convex curves surrounding the origin (respectively strongly concave curves).

From now, if $\theta$ is a parameter such that the above formula holds, we will denote $L_{\boldsymbol{v}} = L_{\frac{\partial}{\partial \theta}} = {}'$. The above formula then reads

$$\omega'' = -\epsilon\omega + b\,\omega'. \tag{2.15}$$

Define the moving frame $\mathcal{F}$ on $\mathcal{H}$ as follows:

$$\mathcal{F} = \left(\vec{\boldsymbol{h}}, \boldsymbol{v}, \left[\boldsymbol{v}, \vec{\boldsymbol{h}}\right]\right). \tag{2.16}$$

Observe that these vector fields are linearly independent since $\boldsymbol{v}$ is a vertical field while $\vec{\boldsymbol{h}}$ and $[\boldsymbol{v}, \vec{\boldsymbol{h}}]$ have linearly independent horizontal parts. Indeed,

$$\begin{aligned}\vec{\boldsymbol{h}}|_{\text{hor}} &= \pi_* \vec{\boldsymbol{h}} = \boldsymbol{f},\\ \left[\boldsymbol{v}, \vec{\boldsymbol{h}}\right]\Big|_{\text{hor}} &= \pi_* \left[\boldsymbol{v}, \vec{\boldsymbol{h}}\right] = \frac{\partial \boldsymbol{f}}{\partial u}\frac{du}{d\theta}.\end{aligned}$$

In the above formula $u(\theta)$ denotes the maximizing control of PMP on the fiber $\mathcal{H}_q$, i.e., $h_{u(\theta)} \geqslant h_u$ for all $u \in U$. Moreover, observe that

$$\frac{du}{d\theta} \neq 0, \tag{2.17}$$

which can be shown as follows. Maximizing condition (1.19) of PMP implies that

$$\left\langle \omega_\theta, \frac{\partial \boldsymbol{f}}{\partial u}\Big|_{u=u(\theta)} \right\rangle \equiv 0, \tag{2.18}$$

according to which the derivative of (2.9) with respect to $\theta$ reads

$$\begin{aligned}0 &= \langle \omega_\theta', \boldsymbol{f}\rangle + \langle \omega_\theta, \boldsymbol{f}'\rangle = \langle \omega_\theta', \boldsymbol{f}\rangle + \left\langle \omega_\theta, \frac{\partial \boldsymbol{f}}{\partial u}\frac{du}{d\theta}\right\rangle\\ &= \langle \omega_\theta', \boldsymbol{f}\rangle\,. \end{aligned}\tag{2.19}$$

One thus infers that

$$0 = \langle \omega_\theta', \boldsymbol{f}\rangle' = \langle -\epsilon\omega_\theta + b\,\omega_\theta', \boldsymbol{f}\rangle + \langle \omega_\theta', \boldsymbol{f}'\rangle = -1 + \langle \omega_\theta', \boldsymbol{f}'\rangle\,,$$

or, equivalently that

$$1 = \langle \omega_\theta', \boldsymbol{f}'\rangle = \left\langle \omega_\theta', \frac{\partial \boldsymbol{f}}{\partial u}\frac{du}{d\theta}\right\rangle, \tag{2.20}$$

from which it follows that $du/d\theta \neq 0$.

Since the vector fields $\vec{\boldsymbol{h}}$ and $\boldsymbol{v}$ are feedback-invariant, it is natural to expect that the principal feedback invariant of our system (the curvature) may arise from a commutator relation of these fields. Indeed, the following theorem due to A. A. Agrachev confirms this intuition.

**Theorem 2.2.3.** *Vector fields* $\boldsymbol{v}$ *and* $\vec{\boldsymbol{h}}$ *satisfy the following nontrivial commutator relation:*

$$\left[\vec{\boldsymbol{h}}, \left[\boldsymbol{v}, \vec{\boldsymbol{h}}\right]\right] = \kappa \boldsymbol{v}. \tag{2.21}$$

***Proof.*** We fix a parameter $\theta$ so that (2.15) holds. This provides us with an identification (2.11) so that the tangent spaces to $\mathcal{H}$ decompose into a direct sum of horizontal and vertical subspaces. The decomposition of the field $\vec{\boldsymbol{h}}$ into horizontal and vertical part is

$$\vec{\boldsymbol{h}} = \underbrace{\boldsymbol{f}}_{\text{horizontal}} + \underbrace{x\frac{\partial}{\partial\theta}}_{\text{vertical}},$$

where $x = x(\theta, q)$ is a smooth function on $\mathcal{H}$.

Notice that, even if $\theta$ is such that (2.15) holds the trivialization (2.11) is not feedback-invariant since the parameter $\theta$ is defined only up to a translation. It implies that the property of a subspace to be horizontal is not feedback-invariant, in particular the function $x(\theta, q)$ in the above decomposition of $\vec{\boldsymbol{h}}$ is not feedback-invariant.

Consider on $\mathcal{H}$ the following coframe of differential one-forms:

$$(\epsilon\omega_\theta, d\theta, \omega'_\theta)\,.$$

These forms are clearly linearly independent since $d\theta$ is vertical while $\omega_\theta$ and $\omega'_\theta$ are linearly independent by (2.13). From (2.9), (2.18), (2.19) and (2.20) it follows that $(\boldsymbol{f}, \boldsymbol{f}')$ is an horizontal family (parametrized by $\theta$) of moving frames on $M$ dual to the horizontal family of coframes $(\epsilon\omega_\theta, \omega'_\theta)$ and since $\boldsymbol{v}$ is a vertical field dual to the vertical one-form $d\theta$, it follows that the frame $(\vec{\boldsymbol{h}}, \boldsymbol{v}, \vec{\boldsymbol{h}}') \subset \operatorname{Vec}\mathcal{H}$ is dual to the coframe $(\epsilon\omega_\theta, d\theta, \omega'_\theta) \subset \Lambda^1(\mathcal{H})$.

We now complete the proof of the theorem computing the required Lie bracket using these frames.

First of all notice that because the horizontal one-forms $\epsilon\omega_\theta$ and $\omega'_\theta$ are linearly independent, the horizontal two-form $d\omega_\theta$ decomposes as

$$d\omega_\theta = c\,\epsilon\omega_\theta \wedge \omega'_\theta, \quad c = c(\theta, q) \in C^\infty(\mathcal{H}), \tag{2.22}$$

which, in view of (2.12), implies

$$d\omega = d\theta \wedge \omega'_\theta + c\,\epsilon\omega_\theta \wedge \omega'_\theta.$$

And since

$$i_{\vec{\boldsymbol{h}}}(d\omega) = 0,$$

then,

$$\begin{aligned} 0 = i_{\vec{\boldsymbol{h}}}(d\omega) &= i_{\boldsymbol{f}+x\frac{\partial}{\partial\theta}}(d\theta \wedge \omega'_\theta + c\,\epsilon\omega_\theta \wedge \omega'_\theta) \\ &= x\,\omega'_\theta - \langle \omega'_\theta, \boldsymbol{f}\rangle\, d\theta + c\,\epsilon\,\langle \omega_\theta, \boldsymbol{f}\rangle\, \omega'_\theta - c\,\epsilon\,\langle \omega'_\theta, \boldsymbol{f}\rangle\, \omega_\theta \\ &= (x + c)\,\omega'_\theta. \end{aligned}$$

Thus the decomposition of $\vec{\boldsymbol{h}}$ is

$$\vec{\boldsymbol{h}} = \boldsymbol{f} - c\frac{\partial}{\partial\theta}.$$

Consequently,

$$\vec{\boldsymbol{h}}' = \left[\frac{\partial}{\partial\theta}, \vec{\boldsymbol{h}}\right] = \boldsymbol{f}' - c'\frac{\partial}{\partial\theta},$$

and we can now compute the required Lie bracket, which gives:

$$\begin{aligned}
\left[\vec{\boldsymbol{h}}, \left[\frac{\partial}{\partial\theta}, \vec{\boldsymbol{h}}\right]\right] &= \left[\vec{\boldsymbol{h}}, \vec{\boldsymbol{h}}'\right] \\
&= \left[\boldsymbol{f} - c\frac{\partial}{\partial\theta}, \boldsymbol{f}' - c'\frac{\partial}{\partial\theta}\right] \\
&= \underbrace{[\boldsymbol{f}, \boldsymbol{f}'] + c'\boldsymbol{f}' - c\boldsymbol{f}''}_{\text{horizontal part}} + \underbrace{\left(L_{\vec{\boldsymbol{h}}'}c - L_{\vec{\boldsymbol{h}}}c'\right)\frac{\partial}{\partial\theta}}_{\text{vertical part}}.
\end{aligned}$$

In order to conclude the proof, we have to show that the horizontal part of the bracket $[\vec{\boldsymbol{h}}, [\boldsymbol{v}, \vec{\boldsymbol{h}}]]$ vanishes. By duality of the horizontal frames $(\boldsymbol{f}, \boldsymbol{f}')$ and $(\epsilon\omega_\theta, \omega'_\theta)$, equality $\omega'' = -\epsilon\omega + b\,\omega'$ implies that

$$\boldsymbol{f}'' = -\epsilon\boldsymbol{f} - b\boldsymbol{f}'.$$

Furthermore, applying Proposition 1.2.2 to the horizontal frames $(\omega_1, \omega_2) = (\epsilon\omega_\theta, \omega'_\theta)$ and $(\boldsymbol{f}_1, \boldsymbol{f}_2) = (\boldsymbol{f}, \boldsymbol{f}')$, we get

$$[\boldsymbol{f}, \boldsymbol{f}'] = -c\epsilon\boldsymbol{f} - c^2_{12}\boldsymbol{f}'. \tag{2.23}$$

where $c^2_{12}$ is the structural constant defined by

$$d\left(\omega'_\theta\right) = c^2_{12}\,\epsilon\omega_\theta \wedge \omega'_\theta.$$

But,

$$\begin{aligned}
d\left(\omega'_\theta\right) &= (d\omega_\theta)' = (c\,\epsilon\omega_\theta \wedge \omega'_\theta)' \\
&= c'\epsilon\omega_\theta \wedge \omega'_\theta + c\,\epsilon\omega_\theta \wedge \omega''_\theta = c'\epsilon\omega_\theta \wedge \omega'_\theta + cb\,\epsilon\omega_\theta \wedge \omega'_\theta.
\end{aligned} \tag{2.24}$$

Summing up, the horizontal part of the field $[\vec{\boldsymbol{h}}, [\boldsymbol{v}, \vec{\boldsymbol{h}}]]$ is

$$[\boldsymbol{f}, \boldsymbol{f}'] + c'\boldsymbol{f}' - c\boldsymbol{f}'' = -c\,\epsilon\boldsymbol{f} - (c' + cb)\boldsymbol{f}' + c'\boldsymbol{f}' - c\,(-\epsilon\boldsymbol{f} - b\boldsymbol{f}') = 0,$$

which proves that

$$\left[\vec{\boldsymbol{h}}, \left[\boldsymbol{v}, \vec{\boldsymbol{h}}\right]\right] = \kappa\boldsymbol{v},$$

where $\kappa$ is evaluated as follows:

$$\kappa(\theta, q) = L_{\vec{\boldsymbol{h}}'}c - L_{\vec{\boldsymbol{h}}}c'. \tag{2.25}$$

This ends the proof of the theorem. ■

The function $\kappa = \kappa(\lambda)$, $\lambda \in \mathcal{H}$, in relation (2.21) is defined to be the *curvature* of our optimal control problem and since the fields $\vec{\boldsymbol{h}}$ and $\boldsymbol{v}$ are feedback-invariant, the curvature $\kappa$ is also feedback-invariant.

Let us now make some remarks about the evaluation of the curvature of control systems.

**Remark 2.2.4.** The vertical field $\boldsymbol{v}$ which satisfies $L_{\boldsymbol{v}}^2\omega = -\epsilon\omega + bL_{\boldsymbol{v}}\omega$ is unique up to sign while the vertical that satisfies (2.21) is unique up to a factor constant along the trajectories of $\vec{\boldsymbol{h}}$. Consequently, if $\boldsymbol{v}$ is such that (2.15) holds true, any vector field $\boldsymbol{w}$ of the form

$$\boldsymbol{w} = g\boldsymbol{v},$$

where $g$ is a first integral of $\vec{\boldsymbol{h}}$ can be use to compute the curvature $\kappa$ via formula (2.21).

**Remark 2.2.5.** Let $\boldsymbol{w}$ be any vertical vector field on $\mathcal{H}$. It is easy to see that in this case

$$\Big[\vec{\boldsymbol{h}}, \Big[\boldsymbol{w}, \vec{\boldsymbol{h}}\Big]\Big] = x\boldsymbol{w} + y\Big[\boldsymbol{w}, \vec{\boldsymbol{h}}\Big], \quad x,\ y \in C^\infty(\mathcal{H}).$$

The curvature can be evaluated in terms of the structure constants $x$ and $y$ of the frame $(\vec{\boldsymbol{h}}, \boldsymbol{w}, [\vec{\boldsymbol{h}}, \boldsymbol{w}])$ as follows:

$$\kappa = x - \frac{y^2}{4} + \frac{L_{\vec{\boldsymbol{h}}}y}{2}.$$

Indeed, if $\boldsymbol{w}$ is vertical there exists a non-vanishing function $a \in C^\infty(\mathcal{H})$ such that $\boldsymbol{v} = a\boldsymbol{w}$. Thus, we have

$$\begin{aligned}
\kappa a\boldsymbol{w} = \Big[\vec{\boldsymbol{h}}, \Big[a\boldsymbol{w}, \vec{\boldsymbol{h}}\Big]\Big] &= \Big[\vec{\boldsymbol{h}}, a\Big[\boldsymbol{w}, \vec{\boldsymbol{h}}\Big] - L_{\vec{\boldsymbol{h}}}a\boldsymbol{w}\Big] \\
&= a\Big[\vec{\boldsymbol{h}}, \Big[\boldsymbol{w}, \vec{\boldsymbol{h}}\Big]\Big] + 2L_{\vec{\boldsymbol{h}}}a\Big[\boldsymbol{w}, \vec{\boldsymbol{h}}\Big] - L_{\vec{\boldsymbol{h}}}^2a\boldsymbol{w} \\
&= (ax - L_{\vec{\boldsymbol{h}}}^2a)\boldsymbol{w} + (2L_{\vec{\boldsymbol{h}}}a + ay)\Big[\boldsymbol{w}, \vec{\boldsymbol{h}}\Big],
\end{aligned}$$

where the function $a$ has to satisfies

$$\kappa = x - \frac{L_{\vec{\boldsymbol{h}}}^2a}{a}, \quad 2L_{\vec{\boldsymbol{h}}}a + ay = 0,$$

from which it follows the required expression for $\kappa$.

## 2.3 Evaluation of the invariant $b$

Notice that the two equations

$$\omega'' = -\epsilon\omega + b\,\omega', \quad \boldsymbol{f}'' = -\epsilon\boldsymbol{f} - b\boldsymbol{f}',$$

show that the feedback-invariant $b$ can be evaluated as follows in the system of local coordinates $(\theta, q)$:

$$\begin{aligned} b = & \frac{\det(\omega'', \omega)}{\det(\omega', \omega)} = \left(\log|\det(\omega', \omega)|\right)' \\ = & -\frac{\det(\boldsymbol{f}'', \boldsymbol{f})}{\det(\boldsymbol{f}', \boldsymbol{f})} = -\left(\log|\det(\boldsymbol{f}', \boldsymbol{f})|\right)'. \end{aligned}$$

If we plug the parameter $u = u(\theta)$ of maximizing control in the previous equations we get the expression of $b$ in any system of local coordinates:

$$\begin{aligned} b = & -\frac{1}{2}\frac{du}{d\theta}\frac{d}{du}\log\frac{\left|\det(\frac{\partial^2\omega}{\partial u^2}, \frac{\partial\omega}{\partial u})\right|}{\left|\det(\frac{\partial\omega}{\partial u}, \omega)\right|^3} \\ = & \ \frac{1}{2}\frac{du}{d\theta}\frac{d}{du}\log\frac{\left|\det(\frac{\partial^2\boldsymbol{f}}{\partial u^2}, \frac{\partial\boldsymbol{f}}{\partial u})\right|}{\left|\det(\frac{\partial\boldsymbol{f}}{\partial u}, \boldsymbol{f})\right|^3}, \end{aligned}$$

where the factor $\frac{du}{d\theta}$ is easily recovered from the formula of change of parameter (2.14) as follows

$$\frac{du}{d\theta} = \sqrt{\left|\frac{\det(\frac{\partial^2\omega}{\partial u^2}, \frac{\partial\omega}{\partial u})}{\det(\frac{\partial\omega}{\partial u}, \omega)}\right|} = \sqrt{\left|\frac{\det(\frac{\partial^2\boldsymbol{f}}{\partial u^2}, \frac{\partial\boldsymbol{f}}{\partial u})}{\det(\frac{\partial\boldsymbol{f}}{\partial u}, \boldsymbol{f})}\right|}.$$

Summing up, one can compute the feedback invariant $b$ directly from the control system (2.1) written in any system of local coordinates using one of the formulas

$$\begin{aligned} b = & -\frac{1}{2}\sqrt{\left|\frac{\det(\frac{\partial^2\omega}{\partial u^2}, \frac{\partial\omega}{\partial u})}{\det(\frac{\partial\omega}{\partial u}, \omega)}\right|}\frac{d}{du}\log\frac{\left|\det(\frac{\partial^2\omega}{\partial u^2}, \frac{\partial\omega}{\partial u})\right|}{\left|\det(\frac{\partial\omega}{\partial u}, \omega)\right|^3}, \\ = & \ \frac{1}{2}\sqrt{\left|\frac{\det(\frac{\partial^2\boldsymbol{f}}{\partial u^2}, \frac{\partial\boldsymbol{f}}{\partial u})}{\det(\frac{\partial\boldsymbol{f}}{\partial u}, \boldsymbol{f})}\right|}\frac{d}{du}\log\frac{\left|\det(\frac{\partial^2\boldsymbol{f}}{\partial u^2}, \frac{\partial\boldsymbol{f}}{\partial u})\right|}{\left|\det(\frac{\partial\boldsymbol{f}}{\partial u}, \boldsymbol{f})\right|^3}. \end{aligned}$$

## 2.4 Projectivised Hamiltonians

In the previous section we defined the curvature for two-dimensional control systems looking at the curvature function as a function on the regular level set $\mathcal{H} = h^{-1}(1)$ of the maximized Hamiltonian function of PMP. This was possible because the Hamiltonian function of PMP is one-homogeneous on fibers of $T^*M$. Actually one-homogeneous on fibers Hamiltonians can be regarded as functions on the projectivised cotangent bundle over $M$:

$$\mathbb{P}(T^*M) = \bigcup_{q\in M}\mathbb{P}(T_q^*M \setminus 0),$$

where $\mathbb{P}(T_q^*M \setminus 0)$ is the projective space of dimension one.

The two bundles over $M$, $\mathbb{P}(T^*M)$ and $\mathcal{H}$ are in fact contact manifolds of dimension three thus, by Darboux's theorem (see e.g. [12], [15], [21]) these two manifolds are locally diffeomorphic (if the curves $\mathcal{H}_q = \mathcal{H} \cap T_q^*M$ are strongly convex curves surrounding the origin the two considered bundles are actually isometric with respect to the Finsler norm defined by the Hamiltonian $h$, see [14] for the proof of this fact and [13] for the definition of Finsler structures).

For certain control systems it may be convenient to consider the maximized Hamiltonian as a function on the contact manifold $\mathbb{P}(T^*M)$. Let $(p,q) = (p_1,p_2,q_1,q_2)$ be a system of local coordinates on the cotangent bundle $T^*M$. The manifold $\mathbb{P}(T^*M)$ can be described through the two charts:

$$(p,q) \mapsto (\xi,q) = \left(\frac{p_2}{p_1}, q\right), \quad (p,q) \mapsto (\xi,q) = \left(\frac{p_1}{p_2}, q\right).$$

In other words, $\mathbb{P}(T^*M)$ is the bundle of all directions with $(p_1,p_2)$ regarded as homogeneous coordinates.

Let $h_u(p,q) \in C^\infty(T^*M)$ be the control dependent Hamiltonian function of PMP for a control system of type (2.1) with regularity assumptions (2.5) and (2.6). If $u^*$ is a maximizing control for $h_u(p,q)$ then, it follows that

$$\left.\frac{\partial h_u}{\partial u}\right|_{u=u^*} = 0.$$

From regularity assumption (2.6) one infers that

$$\left.\frac{\partial^2 h_u}{\partial u^2}\right|_{u=u^*} \neq 0.$$

Hence, by the implicit function theorem one can reconstruct the control $u$ as a smooth function $u = u(p,q)$.

Since the covector $p$ of PMP is different from zero along extremal, we can consider the maximized Hamiltonian of PMP $h(p,q) = \max_{\{u\in U\}} h_u(p,q)$ as a function on $\mathbb{P}(T^*M)$. Suppose for example that the coordinate $p_2$ is different from zero. Then, looking through the chart $\xi = \frac{p_1}{p_2}$ we form the contact Hamiltonian corresponding to $h(p,q)$ via the formula

$$h(\xi,q) = h\left(\frac{p_1}{p_2}, 1, q\right) = h_{u\left(\frac{p_1}{p_2},1,q\right)}\left(\frac{p_1}{p_2}, 1, q\right).$$

The corresponding equations for the associated contact vector field $\vec{\boldsymbol{h}}$ are (see [12] for the detailed computation)

$$\dot{q}_1 = \frac{\partial h}{\partial \xi} \tag{2.26}$$

$$\dot{q}_2 = h - \xi \frac{\partial h}{\partial \xi} \tag{2.27}$$

$$\dot{\xi} = -\frac{\partial h}{\partial q_1} + \xi \frac{\partial h}{\partial q_2}. \tag{2.28}$$

In order to be able to compute the curvature in such coordinates, we now need to find the coordinate expression of the vertical vector field $\boldsymbol{v}$. Let $\bar{\omega}$ be the contact form on $\mathbb{P}(T^*M)$ corresponding to the Liouville one-form of $T^*M$. In our local coordinates $\bar{\omega}$ has the normal form

$$\bar{\omega} = \xi dq_1 + dq_2,$$

and we have

$$\left\langle \bar{\omega}, \vec{\boldsymbol{h}} \right\rangle = h(\xi, q).$$

Thus the contact form $\omega_\xi$ on $\mathbb{P}(T^*M)$ satisfying $\langle \omega_\xi, \vec{\boldsymbol{h}} \rangle = 1$ is

$$\omega_\xi = \frac{1}{h(\xi, q)} \bar{\omega}.$$

A straightforward computation shows that the decomposition of $\frac{\partial^2 \omega_\xi}{\partial \xi^2}$ in the basis $(\omega_\xi, \frac{\partial \omega_\xi}{\partial \xi})$ is

$$\frac{\partial^2 \omega_\xi}{\partial \xi^2} = -\frac{\frac{\partial^2 h}{\partial \xi^2}(\xi, q)}{h(\xi, q)} \omega_\xi - 2 \frac{\frac{\partial h}{\partial \xi}(\xi, q)}{h(\xi, q)} \frac{\partial \omega_\xi}{\partial \xi}.$$

Then, it follows immediately from the formula of change of parameter (2.14) that the vertical vector field $\boldsymbol{v}$ takes the form

$$\boldsymbol{v} = \sqrt{\frac{h(\xi, q)}{\frac{\partial^2 h}{\partial \xi^2}(\xi, q)}} \frac{\partial}{\partial \xi}. \tag{2.29}$$

Consequently, we can apply formula (2.21) of Theorem 2.2.3 in order to compute the curvature of the control system. We will use this setting later on.

## 2.5 Curvature, first examples

We give in this section some examples of computation of curvature for control systems of type (2.1). First of all observe that if for such a system we have to minimize an integral cost of the form

$$\int_0^{t_1} \varphi(q(t), u(t)) dt,$$

all the results of the previous section remain valid as long as the normal Hamiltonian of PMP

$$h(\lambda) = \max_{u \in U} (\langle \lambda, \boldsymbol{f}(q,u) \rangle - \varphi(q,u))$$

is smooth (in the domain under consideration) and such that the curves $\mathcal{H}_q = h^{-1}(e) \cap T^*M$ are strongly convex (or strongly concave). This type of optimal control problems is indeed equivalent to a time optimal control problem under a certain reparametrization of time.

### 2.5.1 Two-dimensional Riemannian problem

Let $(M, g)$ be a two-dimensional Riemannian manifold and let $(\boldsymbol{e}_1, \boldsymbol{e}_2) \subset \operatorname{Vec} M$ be a local orthonormal frame for the Riemannian structure $g$.

The Riemannian geodesic problem consists in finding the shortest curve in $M$ that connects two fixed points $q_0$, $q_1$. The corresponding optimal control problem is stated as follows:

$$\begin{aligned} &\dot{q} = u_1 \boldsymbol{e}_1 + u_2 \boldsymbol{e}_2, \quad q \in M, \quad (u_1, u_2) \in \mathbb{R}^2, \\ &q(0) = q_0, \quad q(t_1) = q_1, \\ &\int_0^{t_1} \langle \dot{q}, \dot{q} \rangle_g^{\frac{1}{2}} \, dt = \int_0^{t_1} (u_1^2 + u_2^2)^{\frac{1}{2}} dt \to \min . \end{aligned} \tag{2.30}$$

It is well-known that after a reparametrization of trajectories of control system (2.30) by arc length, this problem with unbounded control $u \in \mathbb{R}^2$ and fixed time interval $[0, t_1]$ is equivalent to the time optimal control problem with the compact set of control parameters $U = \{u \in \mathbb{R}^2 \mid |u| \leqslant 1\}$. Hence, the optimal control problem can equivalently be stated as follows:

$$\begin{aligned} &\dot{q} = u_1 \boldsymbol{e}_1 + u_2 \boldsymbol{e}_2, \quad q \in M, \quad u_1^2 + u_2^2 \leqslant 1, \\ &q(0) = q_0, \quad q(t_1) = q_1, \\ &t_1 \to \min . \end{aligned}$$

The maximized Hamiltonian of PMP for the above time-minimum problem is

$$h(\lambda) = \max_{u_1^2 + u_2^2 \leqslant 1} (\langle \lambda, \boldsymbol{e}_1 \rangle u_1 + \langle \lambda, \boldsymbol{e}_2 \rangle u_2),$$

and the maximizing condition leads to

$$u_1 = \frac{\langle \lambda, \boldsymbol{e}_1 \rangle}{\sqrt{\langle \lambda, \boldsymbol{e}_1 \rangle^2 + \langle \lambda, \boldsymbol{e}_2 \rangle^2}}, \quad u_2 = \frac{\langle \lambda, \boldsymbol{e}_2 \rangle}{\sqrt{\langle \lambda, \boldsymbol{e}_1 \rangle^2 + \langle \lambda, \boldsymbol{e}_2 \rangle^2}}.$$

Thus, the extremals of this problem (the Riemannian geodesics) are projection onto $M$ of the integral curves of the Hamiltonian field:

$$\begin{aligned} &\dot{\lambda} = \vec{\boldsymbol{h}}(\lambda), \quad \lambda \neq 0, \\ &h(\lambda) = \sqrt{\langle \lambda, \boldsymbol{e}_1 \rangle^2 + \langle \lambda, \boldsymbol{e}_2 \rangle^2}. \end{aligned}$$

It is obvious that this problem does not admit abnormal extremals.

The level surface $\mathcal{H} = h^{-1}(1)$ is a spherical bundle over $M$ whose fiber

$$\mathcal{H}_q = \mathcal{H} \cap T_q^*M = \{\lambda \in T_q^*M \mid \langle \lambda, \boldsymbol{e}_1(q) \rangle^2 + \langle \lambda, \boldsymbol{e}_2(q) \rangle^2 = 1\} \cong S^1$$

can be parametrized by an angle $\theta$ defined by

$$\langle \lambda, \boldsymbol{e}_1(q) \rangle = \cos\theta, \quad \langle \lambda, \boldsymbol{e}_2(q) \rangle = \sin\theta.$$

Let $\omega$ be the restriction on $\mathcal{H}$ of the Liouville one-form of $T^*M$ and let $(\boldsymbol{e}_1^*, \boldsymbol{e}_2^*) \subset \Lambda^1(M)$ be the coframe dual to the orthonormal frame $(\boldsymbol{e}_1, \boldsymbol{e}_2) \subset \operatorname{Vec} M$. In the system of local coordinates $\lambda = (\theta, q)$ on $\mathcal{H}$ the one-form $\omega$ reads

$$\begin{aligned} \omega_\lambda &= \langle \lambda, \boldsymbol{e}_1(q) \rangle \, \boldsymbol{e}_1^*(q) + \langle \lambda, \boldsymbol{e}_2(q) \rangle \, \boldsymbol{e}_2^*(q), \quad \lambda \in \mathcal{H}_q \\ &= \cos\theta \, \boldsymbol{e}_1^*(q) + \sin\theta \, \boldsymbol{e}_2^*(q). \end{aligned}$$

It is obvious that the one-form $\omega$ satisfies

$$\frac{\partial^2 \omega}{\partial \theta^2} = -\omega,$$

which shows that the feedback-invariant frame (constructed in the previous section) for the Riemannian problem is

$$\left( \vec{\boldsymbol{h}}, \frac{\partial}{\partial \theta}, \vec{\boldsymbol{h}}' \right), \quad \vec{\boldsymbol{h}}' = \left[ \frac{\partial}{\partial \theta}, \vec{\boldsymbol{h}} \right].$$

The decomposition into horizontal and vertical part of the Hamiltonian field $\vec{\boldsymbol{h}}$ for Riemannian problem reads

$$\vec{\boldsymbol{h}} = \boldsymbol{f} - c_{\text{riem}} \frac{\partial}{\partial \theta}, \quad \boldsymbol{f} = \cos\theta \, \boldsymbol{e}_1 + \sin\theta \, \boldsymbol{e}_2,$$

with the function $c_{\text{riem}}$ defined by $d_q\omega = c_{\text{riem}} \, \omega \wedge \frac{\partial \omega}{\partial \theta}$, where we have denoted by $d_q\omega$ the exterior differential of the form $\omega$ with respect to the horizontal coordinates.

Denote by $c_1$, $c_2$ the structure constants of the frame $(\boldsymbol{e}_1, \boldsymbol{e}_2)$:

$$[\boldsymbol{e}_1, \boldsymbol{e}_2] = c_1 \boldsymbol{e}_1 + c_2 \boldsymbol{e}_2, \quad c_1, \; c_2 \in C^\infty(M),$$

The structure constant $c_{\text{riem}}$ can easily be computed via the second Cartan formula (1.9). Namely we have:

$$\begin{aligned} -c_{\text{riem}} &= d_q\omega\big(\boldsymbol{f}, \tfrac{\partial \boldsymbol{f}}{\partial \theta}\big) = \omega\big(\big[\boldsymbol{f}, \tfrac{\partial \boldsymbol{f}}{\partial \theta}\big]\big) = \omega([\boldsymbol{e}_1, \boldsymbol{e}_2]) = (\cos\theta \, \boldsymbol{e}_1^* + \sin\theta \, \boldsymbol{e}_2^*)(c_1 \boldsymbol{e}_1 + c_2 \boldsymbol{e}_2) \\ &= \cos\theta \, c_1 + \sin\theta \, c_2. \end{aligned} \tag{2.31}$$

Using the above decomposition of $\vec{\boldsymbol{h}}$ and the formula

$$\left[ \vec{\boldsymbol{h}}, \left[ \boldsymbol{v}, \vec{\boldsymbol{h}} \right] \right] = \kappa \boldsymbol{v}$$

of Theorem 2.2.3, we can easily compute the curvature of the Riemannian problem. A straightforward computation leads to

$$\kappa = -c_1^2 - c_2^2 + L_{\boldsymbol{e}_1} c_2 - L_{\boldsymbol{e}_2} c_1. \tag{2.32}$$

**Remark 2.5.1.** Because the tangent and cotangent bundles of a Riemannian manifold can be identified via the Riemannian structure, $\mathcal{H} \subset T^*M$ is identified with the spherical bundle

$$\mathcal{S} = \{v \in TM \mid \langle v, v \rangle_g = 1\}$$

so that the flow generated by the Hamiltonian field $\vec{\boldsymbol{h}}$ can be considered as the Riemannian geodesic flow on $\mathcal{S}$.

Let us write the Riemannian orthonormal basis of $M$ in geodesic coordinates (see e.g. [41]). Namely, we have

$$(\boldsymbol{e}_1, \boldsymbol{e}_2) = \left(\frac{\partial}{\partial q_1}, \frac{1}{J}\frac{\partial}{\partial q_2}\right), \quad J \in C^\infty(M), \quad J \neq 0.$$

Then equation (2.32) of the curvature reduces to

$$\kappa = -\frac{1}{J}\frac{\partial^2 J}{\partial q_1^2},$$

and we immediately recognize the expression of the Gaussian curvature in geodesic coordinates. Thus, we have proved the following theorem.

**Theorem 2.5.2.** *For the two-dimensional Riemannian problem the control curvature coincides with the Gaussian curvature of the manifold.*

For the Riemannian problem the curvature $\kappa = \kappa(q)$ depends only on the base point $q \in M$ as one can see from formula (2.32) but in general this is not the case: the curvature $\kappa$ depends also on the coordinate in the fiber $\mathcal{H}_q$ and thus is a function on the whole three-dimensional manifold $\mathcal{H}$.

## 2.5.2 Two-dimensional Lorentzian manifold

A Lorentzian structure on a smooth $n$-dimensional manifold $M$ is a symmetric two-tensor $g$ that is non degenerate at each point $q \in M$ and has signature $(n-1, 1)$. Namely, given a Lorentzian metric $g$ and a point $q \in M$, one can construct a basis $(v_1, \ldots, v_n)$ for $T_qM$ in which the metric $g$ has the normal form

$$g = (dq_1)^2 + \cdots + (dq_{n-1})^2 - (dq_n)^2.$$

Let $M$ be a two-dimensional Lorentzian manifold and let $(\boldsymbol{e}_1, \boldsymbol{e}_2) \subset \operatorname{Vec} M$ be an orthonormal frame for the Lorentzian structure. For such a manifold a similar construction as the one made for the Riemannian problem can be done in order to compute the curvature of the Lorentzian manifold. In this case the dynamics on the manifold $M$ are given by

$$\dot{q} = \cosh u\, \boldsymbol{e}_1(q) + \sinh u\, \boldsymbol{e}_2(q), \quad q \in M, \quad u \in \mathbb{R}. \tag{2.33}$$

While the curves of admissible velocities were centered ellipses in the Riemannian case, they are centered hyperbolas in the Lorentzian case. These curves are thus strictly concave which means that the time-optimal problem whose dynamics are given by (2.33) is a time-maximum problem.

For this problem the curvature is evaluated as follows

$$\kappa = c_1^2 - c_2^2 + L_{\boldsymbol{e}_1} c_2 + L_{\boldsymbol{e}_2} c_1,$$

where $c_1$, $c_2$ are the structure constants of the frame $(\boldsymbol{e}_1, \boldsymbol{e}_2)$.

We did not repeat details of the construction for the feedback invariant moving for this problem since all computations are similar to the Riemannian case.

### 2.5.3 Problem of controlling the angular momentum of a rigid body

We compute here the curvature for a control system that describes the control of the angular velocity by one torque of an asymmetric rigid body.

Recall that the rotation of rigid body fixed at its center of gravity and free to move around this pivotal point is described by the Euler equation (see e.g. [6], [12]):

$$\dot{x} = x \times \beta x,$$

where $x \in \mathbb{R}^3$ is the angular momentum vector in a coordinate system connected with the body, $\beta$ is the symmetric operator inverse to the inertia tensor of the body, and "$\times$" denotes the vector product in $\mathbb{R}^3$. We suppose that the body is asymmetric, i.e., that the operator $\beta$ has three distinct eigenvalues which we denote by $J_1$, $J_2$ and $J_3$. If we apply to the rigid body a controlling angular momentum along an axis passing through its center of gravity, then the dynamics of the angular momentum vector $x$ include these external force and they are given by

$$\dot{x} = x \times \beta x + u\boldsymbol{b}, \tag{2.34}$$

where $\boldsymbol{b} \in \mathbb{R}^3$ is a unit vector along which the torque is applied.

For such a controllable system (see e.g. [30] for the controllability of such a system) we consider the time optimal problem

$$\begin{aligned} &x(0) = x_0, \quad x(t_1) = x_1, \\ &t_1 \to \min . \end{aligned}$$

System (2.34) is an affine in control system of the form

$$\dot{x} = \boldsymbol{f}(x) + u\boldsymbol{g}, \quad x \in \mathbb{R}^3, \tag{2.35}$$

where $\boldsymbol{g} \in \operatorname{Vec} \mathbb{R}^3$ is a non zero constant vector field. It as been proved by A. A. Agrachev and A. V. Sarychev (see [6], [7]) that such a system can be reduced to a nonlinear system on $\mathbb{R}^2$. Namely, since the vector field $\boldsymbol{g}$ is non zero and constant on the whole manifold $\mathbb{R}^3$, one can define globally on the manifold $\mathbb{R}^3$ an equivalence relation that puts in a single class all the points lying on a single trajectory of the vector field $\boldsymbol{g}$. Denote by $(\mathbb{R}^3)^{\boldsymbol{g}}$ the quotient manifold defined by this equivalence relation. Under this equivalence relation system (2.35) is equivalent to the partial system

$$\dot{q} = e^{u\,\mathrm{ad}\,\boldsymbol{g}}\boldsymbol{f}(q), \quad q \in (\mathbb{R}^3)^{\boldsymbol{g}}.$$

Notice that the family of vector fields $e^{u\,\mathrm{ad}\,\boldsymbol{g}}\boldsymbol{f}$, $u \in \mathbb{R}$ is well defined on the quotient manifold $(\mathbb{R}^3)^{\boldsymbol{g}}$. This follows from the fact that under the action of the diffeomorphism $e_*^{t\boldsymbol{g}}$ the vector field $e^{u\,\mathrm{ad}\,\boldsymbol{g}}\boldsymbol{f}$ passes into a vector field of the same family. Indeed, under the action of $e_*^{t\boldsymbol{g}}$ we have

$$e_*^{t\boldsymbol{g}} e^{u\,\mathrm{ad}\,\boldsymbol{g}}\boldsymbol{f} = \mathrm{Ad}\, e^{-t\boldsymbol{g}} e^{u\,\mathrm{ad}\,\boldsymbol{g}}\boldsymbol{f} = e^{-t\,\mathrm{ad}\,\boldsymbol{g}} e^{u\,\mathrm{ad}\,\boldsymbol{g}}\boldsymbol{f} = e^{(u-t)\,\mathrm{ad}\,\boldsymbol{g}}\boldsymbol{f},$$

which shows that the group of diffeomorphisms $e_*^{t\boldsymbol{g}}$ carries the family $e^{u\,\mathrm{ad}\,\boldsymbol{g}}\boldsymbol{f}$ into itself.

The reduction procedure applied to the rigid body problem (2.34) reads

$$\begin{aligned}\dot{q} &= e^{u\,\mathrm{ad}\,\boldsymbol{b}}(q \times \beta q) \\ &= (q + u\boldsymbol{b}) \times \beta(q + u\boldsymbol{b}), \quad q \in (\mathbb{R}^3)^{\boldsymbol{b}}, \quad u \in \mathbb{R}. \end{aligned} \tag{2.36}$$

The quotient manifold $(\mathbb{R}^3)^{\boldsymbol{b}}$ can be identified with the plane $\mathbb{R}^2$ passing through the origin and orthogonal to the vector $\boldsymbol{b}$. After projection on this plane system (2.36) reads

$$\dot{q} = (q + u\boldsymbol{b}) \times \beta(q + u\boldsymbol{b}) - \langle q \times \beta(q + u\boldsymbol{b}), \boldsymbol{b}\rangle \boldsymbol{b}, \quad q \in (\mathbb{R}^3)^{\boldsymbol{b}}, \quad u \in \mathbb{R}. \tag{2.37}$$

We introduce Cartesian coordinates $(q_1, q_2)$ on $(\mathbb{R}^3)^{\boldsymbol{b}}$ such that the field $\frac{\partial}{\partial q_1}$ has the direction of the vector $\boldsymbol{b} \times \beta\boldsymbol{b}$ and the field $\frac{\partial}{\partial q_2}$ has the direction of the orthogonal vector $\boldsymbol{b} \times (\boldsymbol{b} \times \beta\boldsymbol{b})$. In this coordinate system (2.37) takes the form

$$\begin{aligned}\dot{q}_1 &= \beta_{13}q_2^2 + (-\beta_{23}q_1 + (\beta_{11} - \beta_{33})q_2)u + u^2 \\ \dot{q}_2 &= -\beta_{13}q_1q_2 - ((\beta_{11} - \beta_{22})q_1 - \beta_{23}q_2)u, \quad u \in \mathbb{R}. \end{aligned} \tag{2.38}$$

Here we have denoted by $\beta_{ij}$ the coefficients of the symmetric operator $\beta$ in the basis $(\boldsymbol{b}, \boldsymbol{b} \times \beta\boldsymbol{b}, \boldsymbol{b} \times (\boldsymbol{b} \times \beta\boldsymbol{b}))$ of $\mathbb{R}^3$. A direct computation gives also that $\beta_{13} < 0$ and $\beta_{11} - \beta_{22} \neq 0$.

For simplicity of computations we shall suppose that the coefficient $\beta_{23}$ of the tensor $\beta$ is zero which means that the torque $\boldsymbol{b}$ lies in one of the planes formed by two of the principal axis of inertia of the body. Indeed, since $\beta_{23}$ vanishes, the decomposition of the vector $\beta(\boldsymbol{b} \times \beta\boldsymbol{b})$ in the basis $(\boldsymbol{b}, \boldsymbol{b} \times \beta\boldsymbol{b}, \boldsymbol{b} \times (\boldsymbol{b} \times \beta\boldsymbol{b}))$ takes the form

$$\beta(\boldsymbol{b} \times \beta\boldsymbol{b}) = \lambda\boldsymbol{b} + \mu(\boldsymbol{b} \times \beta\boldsymbol{b}),$$

for some real constants $\lambda$ and $\mu$. Thus, the vector $\beta(\boldsymbol{b} \times \beta\boldsymbol{b})$ is orthogonal to the vector $\boldsymbol{b} \times (\boldsymbol{b} \times \beta\boldsymbol{b})$ and we have

$$\begin{aligned} 0 &= \langle \boldsymbol{b} \times (\boldsymbol{b} \times \beta\boldsymbol{b}), \beta(\boldsymbol{b} \times \beta\boldsymbol{b})\rangle = \langle \langle \boldsymbol{b}, \beta\boldsymbol{b}\rangle \boldsymbol{b} - \langle \boldsymbol{b}, \boldsymbol{b}\rangle \beta\boldsymbol{b}, \beta(\boldsymbol{b} \times \beta\boldsymbol{b})\rangle \\ &= \langle \boldsymbol{b}, \beta\boldsymbol{b}\rangle \underbrace{\langle \boldsymbol{b}, \beta(\boldsymbol{b} \times \beta\boldsymbol{b})\rangle}_{=\langle \beta\boldsymbol{b}, \boldsymbol{b}\times\beta\boldsymbol{b}\rangle = 0} - \underbrace{\langle \boldsymbol{b}, \boldsymbol{b}\rangle}_{=1} \langle \beta\boldsymbol{b}, \beta(\boldsymbol{b} \times \beta\boldsymbol{b})\rangle \\ &= \left\langle \beta^2\boldsymbol{b}, \boldsymbol{b} \times \beta\boldsymbol{b}\right\rangle = \det(\boldsymbol{b}, \beta\boldsymbol{b}, \beta^2\boldsymbol{b}). \end{aligned}$$

In the basis formed by the unitary vectors along the principal axis of the rigid body, the operator $\beta$ is diagonal and we have

$$\det(\boldsymbol{b}, \beta\boldsymbol{b}, \beta^2\boldsymbol{b}) = \begin{vmatrix} \ell_1 & \ell_1 J_1 & \ell_1 J_1^2 \\ \ell_2 & \ell_2 J_2 & \ell_2 J_2^2 \\ \ell_3 & \ell_3 J_3 & \ell_3 J_3^2 \end{vmatrix} = \ell_1\ell_2\ell_3(J_2 - J_1)(J_1 - J_3)(J_3 - J_2).$$

But, since the eigenvalues of the operator $\beta$ are distinct, we conclude that one of the $\ell_i$'s is zero which exactly means that the torque lies in one of the planes formed by two axis of inertia of the body.

For the reduced system (2.38) we form the Hamiltonian of PMP

$$h_u(p, q) = p_1(\beta_{13}q_2^2 + (\beta_{11} - \beta_{33})q_2u + u^2) - p_2(\beta_{13}q_1q_2 + (\beta_{11} - \beta_{22})q_1u),$$

which attains a maximum at $u = u^*$ when

$$((\beta_{11} - \beta_{33})q_2 + 2u^*)p_1 = p_2(\beta_{11} - \beta_{22})q_1. \tag{2.39}$$

Because the unitary bundle $\mathcal{H} = \{p \in T^*M \mid \max_{u\in\mathbb{R}} h_u(p, q) = 1\}$ is diffeomorphic to the projective bundle $\mathbb{P}(T^*M)$, we regard $(p_1, p_2)$ as homogeneous coordinates on $T^*M \setminus 0$. Supposing that $p_1 \neq 0$ we can look at the Hamiltonian $h_u$ as a function on $\mathbb{P}(T^*M)$ through the chart $\xi = \frac{p_2}{p_1}$. The maximality condition (2.39) implies that

$$u^* = -\frac{1}{2}((\beta_{11} - \beta_{33})q_2 - (\beta_{11} - \beta_{22})q_1\xi).$$

Consequently, the maximized projectivised Hamiltonian corresponding to $h_u(p, q)$ is

$$h(\xi, q) = h_{u^*}(1, \xi, q) = -\frac{1}{4}\left((\beta_{11} - \beta_{33})q_2 - (\beta_{11} - \beta_{22})q_1\xi\right)^2 + \beta_{13}(q_2^2 - q_1q_2\xi),$$

and the corresponding Hamiltonian system is

$$\begin{aligned} \dot{q}_1 &= \frac{1}{4}\left(\left(4\beta_{13} - (\beta_{11} - \beta_{33})^2\right) q_2^2 + (\beta_{11} - \beta_{22})^2 q_1^2\xi^2\right) \\ \dot{q}_2 &= \frac{1}{2}\left(\left((\beta_{11} - \beta_{22})(\beta_{11} - \beta_{33}) - 2\beta_{13}\right) q_1q_2 - (\beta_{11} - \beta_{22})^2 q_1^2\xi\right) \\ \dot{\xi} &= \frac{1}{2}\Big( \left((\beta_{11} - \beta_{33})^2 - 4\beta_{13}\right) q_2 - \left((\beta_{11} - \beta_{22})(\beta_{11} - \beta_{33}) - 2\beta_{13}\right) q_1\xi \\ &\qquad + \left((\beta_{11} - \beta_{22})(\beta_{11} - \beta_{33}) - 2\beta_{13}\right) q_2\xi^2 - (\beta_{11} - \beta_{22})^2 q_1\xi^3\Big). \end{aligned}$$

Notice that, because we are using here the chart $\xi = \frac{p_2}{p_1}$, the above equations do not correspond to the equations (2.26), (2.27) and (2.28) of Section 2.4. In order to recover the equations for the Hamiltonian field in the chart $\xi = \frac{p_2}{p_1}$ from the equations in the chart $\xi = \frac{p_1}{p_2}$, it suffices to permute the roles of the base variables $q_1$ and $q_2$.

The last thing we need for the computation of the curvature is the vertical vector field that satisfies $L^2_{\boldsymbol{v}} s|_{\mathbb{P}(T^*M)} = -s|_{\mathbb{P}(T^*M)} + bL_{\boldsymbol{v}} s|_{\mathbb{P}(T^*M)}$. Using formula (2.29) we easily see that this field takes the form

$$\boldsymbol{v} = \sqrt{\frac{(\beta_{11}-\beta_{33})^2 - 4\beta_{13}}{2(\beta_{11}-\beta_{22})^2}\frac{q_2^2}{q_1^2} + \left(\frac{2\beta_{13}}{(\beta_{11}-\beta_{22})^2} - \frac{\beta_{11}-\beta_{33}}{\beta_{11}-\beta_{22}}\right)\frac{q_2}{q_1}\xi + \frac{\xi^2}{2}}\,\frac{\partial}{\partial\xi}.$$

Then, formula (2.21) of Theorem 2.2.3 leads to

$$\left[\vec{\boldsymbol{h}}, \left[\boldsymbol{v}, \vec{\boldsymbol{h}}\right]\right] = (a_0 + a_1\xi + a_2\xi^2 + a_3\xi^3)\boldsymbol{v},$$

so that the curvature

$$\kappa(\xi, q) = a_0 + a_1\xi + a_2\xi^2 + a_3\xi^3$$

is a polynomial of degree three in the variable $\xi$ with coefficients

$$\begin{aligned}
a_0 &= -\beta_{13}\left(\beta_{13} - (\beta_{11}-\beta_{22})(\beta_{22}-\beta_{33})\right) q_1^2 - \left((\beta_{11}-\beta_{33})^2 - 4\beta_{13}\right)^2 \frac{q_2^4}{q_1^2},\\
a_1 &= -\frac{3}{8}\left(4\beta_{13} - (\beta_{11}-\beta_{33})^2\right)\left((\beta_{11}-\beta_{22})(\beta_{11}-\beta_{33}) - 2\beta_{13}\right)\frac{q_2^3}{q_1}\\
&\quad +3\beta_{13}\left(\beta_{13} - (\beta_{11}-\beta_{22})(\beta_{22}-\beta_{33})\right) q_1 q_2,\\
a_2 &= -\frac{3}{8}\left(\left((\beta_{11}-\beta_{22})(\beta_{11}-\beta_{33}) - 2\beta_{13}\right)^2 - 2\beta_{13}\left(\beta_{13} - (\beta_{11}-\beta_{22})(\beta_{22}-\beta_{33})\right)\right) q_2^2,\\
a_3 &= -\frac{1}{8}(\beta_{11}-\beta_{22})^2\left(2\beta_{13} - (\beta_{11}-\beta_{22})(\beta_{11}-\beta_{33})\right) q_1 q_2.
\end{aligned}$$

Of course one can substitute $\xi$ by its value involving the optimal control in order to get an expression of the curvature as a function of $q$ and $u^*$.

## 2.6 Jacobi equation

In this section we use the moving frame previously constructed to derive an ODE on conjugate time of our two-dimensional optimal control problem. This ODE, Jacobi equation in the moving frame, will show that the control analogue to the Gaussian curvature enjoys similar properties.

Fix a point $q_0 \in M$ and define

$$\mathcal{L}_0 = e^{t\vec{\boldsymbol{h}}}(\mathcal{H}_{q_0}), \quad t \in \mathbb{R}\setminus 0,$$

where $e^{t\vec{\boldsymbol{h}}}$ denotes the flow of the Hamiltonian field $\vec{\boldsymbol{h}}$.

**Definition 2.6.1.** *We call conjugate points to $q_0$ the set of critical value of the canonical projection $\pi : \mathcal{L}_0 \to M$.*

**Definition 2.6.2.** *We say that time $t \neq 0$ is conjugate to zero if there exists $\lambda_0 \in \mathcal{H}_{q_0}$ such that $(t, \lambda_0)$ is critical for $\pi$.*

Obviously when time $t$ is conjugate to zero, the point $q(t) = \pi(e^{t\vec{\boldsymbol{h}}}(\lambda_0))$ is conjugate to $q_0$.

Notice that the two above definitions do not depend on the dimension of the manifold $M$. The next proposition gives a characterization of conjugate of conjugate times when $M$ is of dimension two.

**Proposition 2.6.3.** *Time $t$ is conjugate to zero if and only if*

$$e^{t\,\mathrm{ad}\,\vec{\boldsymbol{h}}}\boldsymbol{v}(\lambda_0) \in \mathrm{span}\left(\vec{\boldsymbol{h}}(\lambda_0), \boldsymbol{v}(\lambda_0)\right).$$

***Proof.*** Observe that the tangent space $T_{\lambda_t}\mathcal{L}_0$, $\lambda_t = e^{t\vec{\boldsymbol{h}}}(\lambda_0)$ is spanned by the vectors $\vec{\boldsymbol{h}}(\lambda_t)$ and $e^{t\vec{\boldsymbol{h}}}_*\boldsymbol{v}(\lambda_0)$. Since the projection onto $T_{q(t)}M$, $q(t) = \pi(\lambda_t)$, of $\vec{\boldsymbol{h}}(\lambda_t)$ is $\pi_*\vec{\boldsymbol{h}}(\lambda) = \boldsymbol{f}(q(t), u(t)) \neq 0$, time $t$ is conjugate to zero if and only if the projection onto $T_{q(t)}M$ of $e^{t\vec{\boldsymbol{h}}}_*\boldsymbol{v}(\lambda_0)$ is parallel to $\boldsymbol{f}(q(t), u(t))$, i.e., if and only if

$$e^{t\vec{\boldsymbol{h}}}_*\boldsymbol{v}(\lambda_0) \in \mathrm{span}\left(\vec{\boldsymbol{h}}(\lambda_t), \boldsymbol{v}(\lambda_t)\right),$$

or, since $e^{t\vec{\boldsymbol{h}}}_*\vec{\boldsymbol{h}} = \vec{\boldsymbol{h}}$, equivalently, if and only if

$$e^{t\,\mathrm{ad}\,\vec{\boldsymbol{h}}}\boldsymbol{v}(\lambda_0) \in \mathrm{span}\left(\vec{\boldsymbol{h}}(\lambda_0), \boldsymbol{v}(\lambda_0)\right).$$

The proposition is proved. ■

By the previous proposition, an instant $t$ is conjugate to zero if and only if the coefficient $\gamma(t)$ in the decomposition

$$e^{t\,\mathrm{ad}\,\vec{\boldsymbol{h}}}\boldsymbol{v}(\lambda_0) = \alpha(t)\vec{\boldsymbol{h}}(\lambda_0) + \beta(t)\boldsymbol{v}(\lambda_0) + \gamma(t)\left[\boldsymbol{v}, \vec{\boldsymbol{h}}\right](\lambda_0) \tag{2.40}$$

satisfies the equality:

$$\gamma(t) = 0.$$

According to Lemma 1.2.3, the matrix of the operator $e^{t\,\mathrm{ad}\,\vec{\boldsymbol{h}}}$ in the moving frame $(\vec{\boldsymbol{h}}, \boldsymbol{v}, [\vec{\boldsymbol{h}}, \boldsymbol{v}])$ is solution of the Cauchy problem (1.11), (1.12) with the matrix

$$A(t) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \kappa_t \\ 0 & -1 & 0 \end{pmatrix}, \quad \kappa_t = \kappa(\lambda_t),$$

which implies that the coefficient $\gamma(t)$ in the decomposition (2.40) is solution of the Cauchy problem

$$\ddot{\gamma} + \kappa_t \gamma = 0, \quad \gamma(0) = 0, \quad \dot{\gamma}(0) = 1.$$

Summing up, we get that time $t$ is conjugate to zero if and only if there exists a non trivial solution to the boundary value problem

$$\ddot{\gamma} + \kappa_t \gamma = 0, \quad \gamma(0) = \gamma(t) = 0. \tag{2.41}$$

We call the boundary value problem (2.41) *Jacobi equation for system (2.1) in the moving frame.* As for classical two-dimensional Riemannian geometry, Sturm's comparison Theorem 1.2.5 for second order ODEs leads to the following theorem about the occurrence of conjugate points for system (2.1).

**Theorem 2.6.4.** *Let $q(t)$, $q(0) = q_0$, be a solution of an optimal two-dimensional control problem and let $\kappa_t$ be the value of the curvature along an extremal $\lambda(t)$, $\pi(\lambda(t)) = q(t)$.*

(*i*) *If $\kappa_t \leqslant 0$ for all $t \geqslant 0$, then $q_0$ has no conjugate points for $t \in [0, +\infty]$.*

(*ii*) *If $\kappa_t \leqslant K$ (resp. $\kappa_t < K$) for all $t \geqslant 0$, and some constant $K > 0$, then $q_0$ has no conjugate points along $q(\cdot)$ for $t \in [0, \pi/\sqrt{K}\,[$ (resp. for $t \in [0, \pi/\sqrt{K}\,]$).*

(*iii*) *If $0 < k \leqslant \kappa_t$ (resp. $0 < k < \kappa_t$), for all $t \geqslant 0$, then $q_0$ must have at least a conjugate point for $t \in \,]0, \pi/\sqrt{k}\,]$ (resp. for $t \in \,]0, \pi/\sqrt{k}\,[$ ).*

***Proof.*** (*i*) Apply Sturm comparison Theorem 1.2.5 for $A(t) = 0$ and $B(t) = \kappa_t$. Since the solution $u(t)$ is equal to $t$, the solution $v(t)$ cannot vanish for $t > 0$, which means that the coefficient $\gamma(t)$ in the decomposition of 2.40 cannot vanish, hence, $q_0$ has no conjugate points on $[0, +\infty[$.
(*ii*) Apply Sturm comparison theorem for $A(t) = K$ and $B(t) = \kappa_t$. Since the solution $u(t)$ is equal to $\sin\left(t\sqrt{K}\right)$, the solution $v(t)$ cannot vanish for $t < \pi/\sqrt{K}$, as in the proof of (*i*), the same argument shows that $q_0$ has no conjugate points for $t \in [0, \pi/\sqrt{K}[$.
The other cases of (*i*) and the case (*ii*) are proved analogously. ■

As a direct consequence of this theorem, we have the following

**Corollary 2.6.5.** *Let $q(t)$ be an extremal trajectory on which $k \leqslant \kappa_t \leqslant K$, $t \geqslant 0$. Then $q_0$ has no conjugate points along $q(\cdot)$ for $t \in [0, \pi/\sqrt{K}\,[$ and at least one conjugate point in $[\pi/\sqrt{K}, \pi/\sqrt{k}\,]$.*

The following theorem gives sufficient condition for a trajectory $q(t)$ on $M$ to be strongly locally optimal in terms of conjugate points (see [6] for the definition of strong local optimality and the proof of the following theorem).

**Theorem 2.6.6.** *Let the trajectory $q(t)$ be as in Theorem 2.6.4. If the time interval $]0, t_1]$ does not contain conjugate points, then the trajectory $q(t)$ is strongly locally optimal for $t \in [0, t_1]$.*

On the other hand if an instant $t_c \in ]0, t_1]$ is conjugate to zero, then there exists an instant $\tilde{t} \in ]0, t_1]$ where the trajectory $q(t)$, $t \in ]0, t_1]$, ceases to be locally optimal.

In practice, it is sometimes more easy for computations to not to consider the curvature itself but some reparametrization of the curvature. We will thus see how the curvature $\kappa$ changes under a reparametrization of time. So let $\tau = \varphi(t)$ be a reparametrization of time. Under this reparametrization the ODE $\dot{\lambda} = \vec{\boldsymbol{h}}(\lambda)$ changes as follows:

$$\frac{d\lambda}{d\tau} = \frac{d\lambda}{dt}\frac{d\varphi^{-1}}{d\tau} = \left(\frac{d\varphi}{dt}\right)^{-1}\vec{\boldsymbol{h}}(\lambda).$$

Thus, reparametrizing time just means to consider the field $\vec{\boldsymbol{h}}$ in the form

$$\vec{\boldsymbol{h}} = \frac{\hat{\boldsymbol{h}}}{\varrho}, \tag{2.42}$$

where $\varrho \in C^\infty(\mathcal{H})$ is a non-vanishing function whose primitive along the trajectories of $\vec{\boldsymbol{h}}$ is the time reparametrization function. Under the reparametrization (2.42) we define a new moving frame on $\mathcal{H}$ by

$$\hat{\mathcal{F}} = \left(\hat{\boldsymbol{h}}, \hat{\boldsymbol{v}}, \left[\hat{\boldsymbol{v}}, \hat{\boldsymbol{h}}\right]\right),$$

where $\hat{\boldsymbol{v}}$ is the vertical field defined by

$$\boldsymbol{v} = \sqrt{\varrho}\,\hat{\boldsymbol{v}}.$$

Vector fields of this frame satisfy the following non trivial commutator relation:

$$\left[\hat{\boldsymbol{h}}, \left[\hat{\boldsymbol{v}}, \hat{\boldsymbol{h}}\right]\right] = \kappa_\varrho \hat{\boldsymbol{v}} + \xi \hat{\boldsymbol{h}}, \quad \kappa_\varrho,\ \xi \in C^\infty(\mathcal{H}). \tag{2.43}$$

Indeed, denote for simplicity $\psi = \sqrt{\varrho}$ then, we have:

$$\begin{aligned}
\left[\hat{\boldsymbol{h}}, \left[\hat{\boldsymbol{h}}, \hat{\boldsymbol{v}}\right]\right] &= \left[\psi^2\vec{\boldsymbol{h}}, \left[\psi^2\vec{\boldsymbol{h}}, \psi^{-1}\boldsymbol{v}\right]\right] = \left[\psi^2\vec{\boldsymbol{h}}, \psi\left[\vec{\boldsymbol{h}}, \boldsymbol{v}\right] + \psi^2 L_{\vec{\boldsymbol{h}}}(\psi^{-1})\boldsymbol{v} - \psi^{-1}L_{\boldsymbol{v}}(\psi^2)\vec{\boldsymbol{h}}\right] \\
&= \left[\psi^2\vec{\boldsymbol{h}}, \psi\left[\vec{\boldsymbol{h}}, \boldsymbol{v}\right]\right] - \left[\psi^2\vec{\boldsymbol{h}}, L_{\vec{\boldsymbol{h}}}\psi\boldsymbol{v}\right] + \xi_1\vec{\boldsymbol{h}} \\
&= \psi^3\left[\vec{\boldsymbol{h}}, \left[\vec{\boldsymbol{h}}, \boldsymbol{v}\right]\right] + \psi^2 L_{\vec{\boldsymbol{h}}}\psi\left[\vec{\boldsymbol{h}}, \boldsymbol{v}\right] - \psi^2 L_{\vec{\boldsymbol{h}}}\psi\left[\vec{\boldsymbol{h}}, \boldsymbol{v}\right] - \psi^2 L^2_{\vec{\boldsymbol{h}}}\psi\boldsymbol{v} + \xi_2\vec{\boldsymbol{h}} \\
&= -(\psi^4\kappa + \psi^3 L^2_{\vec{\boldsymbol{h}}}\psi)\hat{\boldsymbol{v}} + \xi\hat{\boldsymbol{h}} \\
&= -\kappa_\varrho\hat{\boldsymbol{v}} + \xi\hat{\boldsymbol{h}},
\end{aligned}$$

where $\xi_1, \xi_2, \xi \in C^\infty(\mathcal{H})$. From the previous computation one infers that the curvature and ts reparametrization $\kappa_\varrho$ satisfy the following relation:

$$\kappa = \frac{\kappa_\varrho - \mathcal{S}(\varrho)}{\varrho^2}, \tag{2.44}$$

where the function

$$\mathcal{S}(\varrho) = L_{\hat{\boldsymbol{h}}}\left(\frac{L_{\hat{\boldsymbol{h}}}\varrho}{2\varrho}\right) - \left(\frac{L_{\hat{\boldsymbol{h}}}\varrho}{2\varrho}\right)^2$$

is called the *Schwartzian derivative* along $\hat{\boldsymbol{h}}$ of the time reparametrization function.

We call the function $\kappa_\varrho$ defined by the relation (2.43) the *$\varrho$-curvature* of control system (2.1).

Since the fields $\hat{\boldsymbol{h}}$ and $\hat{\boldsymbol{v}}$ have the same directions than the fields $\vec{\boldsymbol{h}}$ and $\boldsymbol{v}$ respectively, it follows from Proposition 2.6.3 that a point $q = \pi(e^{t\hat{\boldsymbol{h}}})(\lambda_0)$, $t \neq 0$, will be conjugate to $q_0 = \pi(\lambda_0)$ if and only if the coefficients in the decomposition

$$e^{t\,\mathrm{ad}\,\hat{\boldsymbol{h}}}\hat{\boldsymbol{v}} = \hat{\alpha}(t)\hat{\boldsymbol{h}} + \hat{\beta}(t)\hat{\boldsymbol{v}} + \hat{\gamma}(t)\left[\hat{\boldsymbol{v}}, \hat{\boldsymbol{h}}\right].$$

are such that

$$\hat{\alpha}(t) = \hat{\gamma}(t) = 0.$$

According to Lemma 1.2.3, the matrix of the operator $e^{t\,\mathrm{ad}\,\hat{\boldsymbol{h}}}$ in the moving frame $\hat{\mathcal{F}}$ is solution of the Cauchy problem (1.11), (1.12) with the matrix

$$A(t) = \begin{pmatrix} 0 & 0 & \xi_t \\ 0 & 0 & \kappa_{\varrho t} \\ 0 & -1 & 0 \end{pmatrix}, \quad \kappa_{\varrho t} = \kappa_\varrho\left(e^{t\hat{\boldsymbol{h}}}(\lambda_0)\right), \quad \xi_t = \xi\left(e^{t\hat{\boldsymbol{h}}}(\lambda_0)\right), \tag{2.45}$$

from which its follows that the point $q = \pi(e^{t\hat{\boldsymbol{h}}}(\lambda_0))$ is conjugate to $q_0 = \pi(\lambda_0)$ if and only if there exists a non trivial solution to the boundary value problem

$$\ddot{\hat{\gamma}} + \kappa_{\varrho t}\hat{\gamma} = 0, \quad \hat{\gamma}(0) = \hat{\gamma}(t) = 0.$$

This boundary value problem is the same as problem (2.41) thus, Theorem 2.6.4, Corollary 2.6.5 and Theorem 2.6.6 are valid with $\kappa_t$ replaced by $\kappa_{\varrho t}$. Notice that the fields $\vec{\boldsymbol{h}}$ and $\hat{\boldsymbol{h}}$ have the same conjugate points, only their conjugate times change. Consequently, one only needs to compute a reparametrization of the curvature in order to get some information about the occurrence conjugate points.

# Chapter 3

# Zermelo's problems

In this chapter we will apply the theory of curvature for smooth optimal control systems described in the previous chapter to the study of two time-optimal control problems. The first problem that will be described is the classical Zermelo navigation problem the aim of which is to find the shortest path in time in a Riemannian manifold under the influence of a wind (or a current) which is represented by a vector field on the considered manifold. When the Riemannian norm of the drift is strictly less than one, Zermelo navigation problem defines a Finsler metric on the manifold, geodesics of which are time-optimal solution of Zermelo navigation problem. The second problem that will be described will be the time-optimal problem dual to Zermelo navigation problem for the pairing "one-forms, vector fields". More than being dual we will see that these problems arise one from the other under the assumption that the Riemannian norm of the drift is strictly less than one.

## 3.1 Zermelo navigation problem

### 3.1.1 Description of the problem

We begin with a short description of the Zermelo navigation problem. We refer the reader to the book of Carathéodory ([19]) for a detailed description of the problem. Zermelo in [43] set the problem as follows.

*"In an unbounded plane where the wind distribution is given by a vector field as a function of position and time, a ship moves with constant velocity relative to the surrounding air mass. How must the ship be steered in order to come from a starting point to a given goal in the shortest time?"*

In this problem the sea surface is modeled by the Euclidean plane $\mathbb{R}^2$ and the wind distribution by a non autonomous vector field $\boldsymbol{X}_t(q) \in \operatorname{Vec}\mathbb{R}^2$. Here we will limit ourselves to the case of a stationary wind distribution, i.e., when $\boldsymbol{X}_t(q) = \boldsymbol{X}(q)$.

For what is going on, we shall assume that the vector field $\boldsymbol{X}$ is at least of class $C^2$. In addition, we assume that the ship can be oriented in each planar direction and that its speed cannot exceed a constant $k$, say $k = 1$ for simplicity. Then, the dynamics of Zermelo problem is given by

$$\dot{q} = \boldsymbol{X}(q) + u, \quad u \in \mathbb{R}^2, \quad |u| \leqslant 1.$$

Notice that Zermelo navigation problem can easily be generalized on any $n$-dimensional Riemannian manifold $M$. The problem is set as a time-optimal control problem in the following manner.

$$\begin{aligned} &\dot{q} = \boldsymbol{X}(q) + u, \quad q \in M, \quad u \in T_qM, \quad |u|_g \leqslant 1, \\ &q(0) = q_0, \quad q(t_1) = q_1 \\ &t_1 \to \min, \end{aligned} \tag{3.1}$$

where $q_0$, $q_1 \in M$ are fixed and $|u|_g$ denotes the Riemannian norm of vector $u$.

Our purpose here is not the study of the controllability of Zermelo problem which, in general, requires a subtitled analysis of the dynamics. In general this problem is not completely controllable. For instance, if the drift is too strong, the ship cannot return to its departure point. Nevertheless, because the attainable sets are compact, if a point $q_1$ is reachable from a point $q_0$, Filippov theorem (see e.g. [6] Chapter 10) implies that this point is reachable in minimum time.

### 3.1.2 Hamiltonian function and associated Hamiltonian vector field

Let us apply the PMP to the optimal control problem (3.1). The control-dependent Hamiltonian function of PMP is

$$h_u(\lambda) = \langle \lambda, \boldsymbol{X}(q) \rangle + \langle \lambda, u \rangle, \quad \lambda \in T_q^*M,$$

and the maximized Hamiltonian is

$$h(\lambda) = \max_{|u|_g \leqslant 1} h_u(\lambda) = \langle \lambda, \boldsymbol{X}(q) \rangle + \max_{|u|_g \leqslant 1} \langle \lambda, u \rangle.$$

It is easy to see that the maximality condition leads to $u(t) = \lambda_t^\sharp / |\lambda_t|_g$, where $\lambda_t^\sharp$ denotes the preimage of $\lambda_t$ by the diffeomorphism (1.13). Therefore the maximized Hamiltonian reads

$$h(\lambda) = \langle \lambda, \boldsymbol{X}(q) \rangle + |\lambda|_g, \tag{3.2}$$

From now suppose that $M$ is a two-dimensional Riemannian manifold. Let $(\boldsymbol{e}_1, \boldsymbol{e}_2) \subset \operatorname{Vec} M$ be a local orthogonal frame for the Riemannian structure on $M$ and denote by $(\boldsymbol{e}_1^*, \boldsymbol{e}_2^*) \subset \Lambda^1(M)$ the corresponding dual coframe, i.e., $\langle \boldsymbol{e}_i^*, \boldsymbol{e}_j \rangle = \delta_{ij}$. In terms of the frame $(\boldsymbol{e}_1, \boldsymbol{e}_2)$, the maximized Hamiltonian (3.2) reads

$$h(\lambda) = \langle \lambda, \boldsymbol{X}(q) \rangle + \sqrt{\langle \lambda, \boldsymbol{e}_1 \rangle^2 + \langle \lambda, \boldsymbol{e}_2 \rangle^2}.$$

Denote by $\mathcal{H}$ the hypersurface $h^{-1}(1) \subset T^*M$.

In order to get the expression of the Hamiltonian vector field $\vec{\boldsymbol{h}}$ associated to $h$ on the level $\mathcal{H}$, we introduce some polar coordinates $(r, \theta)$ on the fiber $T_q^*M$:

$$\langle \lambda, \boldsymbol{e}_1 \rangle = r \cos\theta, \quad \langle \lambda, \boldsymbol{e}_2 \rangle = r \sin\theta. \tag{3.3}$$

In these coordinates we have

$$h(r, \theta, q) = r(1 + \langle \boldsymbol{V}, \boldsymbol{X} \rangle_g),$$

where $\boldsymbol{V} \in \operatorname{Vec} M$ is the vector field defined by

$$\boldsymbol{V} = \cos\theta\, \boldsymbol{e}_1 + \sin\theta\, \boldsymbol{e}_2,$$

and $\langle \cdot, \cdot \rangle_g$ denotes the Riemannian scalar product on $M$. In this way chosen, $\theta$ is the angle formed by $\boldsymbol{e}_1(q)$ and the direction navigated by the ship.

Observe that because $\lambda$ is the covector of PMP, we have $r = |\lambda|_g \neq 0$ along any extremal. Thus, an optimal trajectory $q(t)$ starting from the point $q_0$ is the projection onto $M$ of an abnormal extremal if and only if $h(r, \theta, q) = 0$, i.e., if and only if the quantity

$$\varrho = 1 + \langle \boldsymbol{V}, \boldsymbol{X} \rangle_g$$

is identically zero along the extremal.

Let us analyze more carefully the existence of abnormal trajectories. For this, we write $\varrho$ in the form

$$\varrho(\theta, q) = |\boldsymbol{X}(q)|_g \cos(\theta - \theta^{\boldsymbol{X}}(q)) + 1,$$

where $\theta^{\boldsymbol{X}}(q)$ is the angle defined by

$$\begin{cases} \theta^{\boldsymbol{X}}(q) = 0 & \text{if } \boldsymbol{X}(q) = 0, \\ \cos\theta^{\boldsymbol{X}}(q) = \dfrac{\langle \boldsymbol{X}(q), \boldsymbol{e}_1(q) \rangle_g}{|\boldsymbol{X}(q)|_g}, \quad \sin\theta^{\boldsymbol{X}}(q) = \dfrac{\langle \boldsymbol{X}(q), \boldsymbol{e}_2(q) \rangle_g}{|\boldsymbol{X}(q)|_g} & \text{if } \boldsymbol{X}(q) \neq 0. \end{cases} \tag{3.4}$$

The above expression for $\varrho$ shows immediately that there are three different cases to distinguish for the resolution of the equation

$$\varrho(\theta, q_0) = 0. \tag{3.5}$$

- If $|\boldsymbol{X}(q_0)|_g < 1$. Then, equation (3.5) has no solution. In this case the point $q_0$ belongs to the interior of the Riemannian disc $\mathcal{D}_{q_0} = \{v \in T_{q_0}M \mid |v|_g \leqslant 1\}$. In other words, the curve of admissible velocities $\theta \mapsto \cos\theta\, \boldsymbol{e}_1(q_0) + \sin\theta\, \boldsymbol{e}_2(q_0)$ is strongly convex for all values of $\theta$ so that all optimal trajectories starting from $q_0$ are time-minimum trajectories.

- If $|\boldsymbol{X}(q_0)|_g = 1$. Then, equation (3.5) admits the unique solution $\theta = \pi + \theta^{\boldsymbol{X}}(q_0)$ which implies the existence of a unique abnormal trajectory starting from the point $q_0$. In this case, the point $q_0$ lies on the boundary of the Riemannian disc $\mathcal{D}_{q_0}$ and the curve of admissible velocities is strongly convex except for $\theta = \theta_0$.

- If $|\boldsymbol{X}(q_0)|_g > 1$. Then, equation (3.5) admits the two distinct solutions $\theta_0 = \pm \arccos(1/|\boldsymbol{X}(q_0)|_g) + \theta^{\boldsymbol{X}}(q_0)$ corresponding to two different abnormal trajectories starting from $q_0$. in this case the curve of admissible velocities is divided in two parts: a strongly convex one which corresponds to minimum time trajectories and a strongly concave one corresponding to maximum time trajectories.

From now we shall assume that $\varrho \neq 0$ in order to be allowed to compute the curvature of the problem. For simplicity we denote by the sign "'" the derivative with respect to the parameter $\theta$.

The Liouville one-form of $T^*M$ in restriction to $\mathcal{H}$ takes the form

$$\omega_\theta = = \langle \lambda, \boldsymbol{e}_1 \rangle \, \boldsymbol{e}_1^* + \langle \lambda, \boldsymbol{e}_2 \rangle \, \boldsymbol{e}_2^* = \frac{\cos\theta}{\varrho} \boldsymbol{e}_1^* + \frac{\sin\theta}{\varrho} \boldsymbol{e}_2^*,$$

and its derivative with respect to $\theta$ reads

$$\omega_\theta' = \frac{-\sin\theta - X_2}{\varrho^2} \boldsymbol{e}_1^* + \frac{\cos\theta + X_1}{\varrho^2} \boldsymbol{e}_2^*,$$

where we have denoted by $X_1$ and $X_2$ the coordinates of the drift vector field $\boldsymbol{X}$ in the frame $(\boldsymbol{e}_1, \boldsymbol{e}_2)$. Because $\varrho \neq 0$ we easily check that $(\omega_\theta, \omega_\theta')$ is a frame of horizontal one-forms. Indeed, we have

$$\omega_\theta \wedge \omega_\theta' = \frac{\boldsymbol{e}_1^* \wedge \boldsymbol{e}_2^*}{\varrho^2} \neq 0.$$

We can then compute the vertical part (in coordinates $(\theta, q)$) of the Hamiltonian field $\vec{\boldsymbol{h}}$ associated to $h$ using formula (2.22). We have

$$\begin{aligned} d\omega_\theta &= d\left( \frac{\cos\theta}{\varrho} \boldsymbol{e}_1^* + \frac{\sin\theta}{\varrho} \boldsymbol{e}_2^* \right) \\ &= \left( -L_{\boldsymbol{e}_2} \left( \frac{\cos\theta}{\varrho} \right) + L_{\boldsymbol{e}_1} \left( \frac{\sin\theta}{\varrho} \right) \right) \boldsymbol{e}_1^* \wedge \boldsymbol{e}_2^* + \frac{1}{\varrho} (\cos\theta d\boldsymbol{e}_1^* + \sin\theta d\boldsymbol{e}_2^*). \end{aligned}$$

Notice that the term $\cos\theta d\boldsymbol{e}_1^* + \sin\theta d\boldsymbol{e}_2^*$ in the above formula corresponds to the Riemannian geodesic problem. It follows from formula (2.31) (see section 2.5.1 of the previous chapter) that

$$\cos\theta d\boldsymbol{e}_1^* + \sin\theta d\boldsymbol{e}_2^* = \langle \boldsymbol{V}, [\boldsymbol{e}_1, \boldsymbol{e}_2] \rangle_g \, \boldsymbol{e}_1^* \wedge \boldsymbol{e}_2^*.$$

Consequently,

$$d\omega_\theta = (L_{\boldsymbol{V}'} \varrho + \langle \boldsymbol{V}, [\boldsymbol{e}_1, \boldsymbol{e}_2] \rangle_g \, \varrho) \omega_\theta \wedge \omega_\theta',$$

from which we deduce that the Hamiltonian field of PMP restricted to $\mathcal{H}$ takes the form:

$$\vec{\boldsymbol{h}} = \boldsymbol{X} + \boldsymbol{V} - \left( L_{\boldsymbol{V}'} \varrho + \langle \boldsymbol{V}, [\boldsymbol{e}_1, \boldsymbol{e}_2] \rangle_g \, \varrho \right) \frac{\partial}{\partial \theta}, \tag{3.6}$$

in system $(\theta, q)$ of local coordinates.

### 3.1.3 Curvature of Zermelo problem

In order to compute the curvature of Zermelo problem we need the coordinate expression of the feedback invariant vertical field $\boldsymbol{v}$ on $\mathcal{H}$. But the decomposition of $\omega_\theta''$ in the frame $(\omega_\theta, \omega_\theta')$ reads

$$\omega_\theta'' = -\frac{1}{\varrho}\omega_\theta - 2\frac{\varrho'}{\varrho}\omega_\theta'.$$

Hence, using the formula (2.14) of a change of parameter we conclude that the vertical field $\boldsymbol{v}$ has the coordinate expression

$$\boldsymbol{v} = \sqrt{\epsilon\varrho}\frac{\partial}{\partial\theta}, \quad \epsilon = \operatorname{sign}\varrho.$$

Consequently, the curvature of the system can easily be computed using formula (2.21) of Theorem 2.2.3. A straightforward computation shows that the curvature takes the form

$$\kappa = \kappa_{\text{riem}} + \psi(\theta, q),$$

where $\kappa_{\text{riem}}$ denotes the Gaussian curvature of the manifold $M$ and $\psi$ is a smooth function on the manifold $\mathcal{H}$. We will not give here a coordinate version of this formula which is rather complicated in the general case.

**Remark 3.1.1.** Notice that, a priori, the curvature of Zermelo navigation problem is a defined only where the Hamiltonian functions does not vanish or, equivalently where $\varrho \neq 0$. It turns out that this function can be smoothly extended to $\{\lambda \in T^*M \setminus 0 \mid \varrho(\lambda) = 0\}$, that is on the submanifold of $T^*M$ formed by the abnormal extremals, since the function $\varrho$ satisfies the non autonomous linear ODE

$$L_{\vec{\boldsymbol{h}}}\varrho = \Gamma\varrho,$$

where,

$$\Gamma = \cos^2\theta L_{\boldsymbol{e}_1}X_1 + \cos\theta\sin\theta(L_{\boldsymbol{e}_2}X_1 + L_{\boldsymbol{e}_1}X_2) + \sin^2\theta L_{\boldsymbol{e}_2}X_2 + \langle \boldsymbol{V}, [\boldsymbol{e}_1, \boldsymbol{e}_2]\rangle_g \varrho'.$$

This property is not specific to Zermelo navigation problem as we shall see in the next chapter.

We now return to the navigation problem of Zermelo on the Euclidean plane $\mathbb{R}^2$. Recall that in this case the optimal dynamics are given by:

$$\begin{aligned} \dot{q}_1 &= X_1(q) + \cos\theta \\ \dot{q}_2 &= X_2(q) + \sin\theta, \quad q \in \mathbb{R}^2, \quad \theta \in S^1, \end{aligned}$$

and the Hamiltonian fields has the form

$$\vec{\boldsymbol{h}} = (X_1 + \cos\theta)\frac{\partial}{\partial q_1} + (X_2 + \sin\theta)\frac{\partial}{\partial q_2} - \left\langle D_q\boldsymbol{X}\left(\begin{smallmatrix}-\sin\theta\\ \cos\theta\end{smallmatrix}\right), \left(\begin{smallmatrix}\cos\theta\\ \sin\theta\end{smallmatrix}\right)\right\rangle\frac{\partial}{\partial\theta},$$

where formula

$$\begin{aligned}\dot\theta &= -\langle D_q\boldsymbol{X}\left(\begin{smallmatrix}-\sin\theta\\ \cos\theta\end{smallmatrix}\right), \left(\begin{smallmatrix}\cos\theta\\ \sin\theta\end{smallmatrix}\right)\rangle\\ &= -\cos^2\theta\frac{\partial X_1}{\partial q_2} + \cos\theta\sin\theta\left(\frac{\partial X_1}{\partial q_1} - \frac{\partial X_2}{\partial q_2}\right) + \sin^2\theta\frac{\partial X_2}{\partial q_1},\end{aligned}$$

that gives the evolution of the direction $\theta$ with respect to time is the so-called navigation formula of Zermelo.

The coordinate expression of the control curvature of this problem is given by

$$\begin{aligned}\kappa(q,\theta) = &-\frac{Q(\theta)}{32}\left(\frac{\partial X_1}{\partial q_1},\frac{\partial X_1}{\partial q_2},\frac{\partial X_2}{\partial q_1},\frac{\partial X_2}{\partial q_2}\right) + \frac{3}{4}\sum_{i=1}^{2}\sin 2\theta X_i\frac{\partial}{\partial q_i}\left(\frac{\partial X_1}{\partial q_2}+\frac{\partial X_2}{\partial q_1}\right)\\ &-\frac{1}{4}\sum_{1\leqslant i,k\leqslant 2}\left(1+(-1)^k 3\cos 2\theta\right)X_i\frac{\partial^2 X_k}{\partial q_i\partial q_k} + \sum_{1\leqslant i,j,k\leqslant 2}A_{ij}^k(\theta)\frac{\partial^2 X_k}{\partial q_i\partial q_j},\end{aligned} \tag{3.7}$$

where $Q(\theta)$ is the quadratic form whose matrix is

| | | | |
|---|---|---|---|
| $23-12\cos 2\theta$ $-3\cos 4\theta$ | $-18\sin 2\theta$ $-3\sin 4\theta$ | $6\sin 2\theta$ $-3\sin 4\theta$ | $-19+3\cos 4\theta$ |
| $-18\sin 2\theta$ $-3\sin 4\theta$ | $21+24\cos 2\theta$ $+3\cos 4\theta$ | $21+3\cos 4\theta$ | $6\sin 2\theta$ $+3\sin 4\theta$ |
| $6\sin 2\theta$ $-3\sin 4\theta$ | $21+3\cos 4\theta$ | $21-24\cos 2\theta$ $+3\cos 4\theta$ | $-18\sin 2\theta$ $+3\sin 4\theta$ |
| $-19+3\cos 4\theta$ | $6\sin 2\theta$ $+3\sin 4\theta$ | $-18\sin 2\theta$ $+3\sin 4\theta$ | $23+12\cos 2\theta$ $-3\cos 4\theta$ |

,

and the coefficients $A_{ij}^k(\theta)$ satisfy:

$$\begin{aligned}A_{ij}^k &= A_{ji}^k \quad \forall\, i,j,k\ ,\\ 2A_{11}^1 &= -2A_{12}^1 = \cos^3\theta\ , \qquad 8A_{22}^1 = 9\cos\theta - \cos 3\theta,\\ 2A_{22}^2 &= -2A_{12}^2 = \sin^3\theta\ , \qquad 8A_{11}^2 = 9\sin\theta + \sin 3\theta.\end{aligned}$$

We now restrict ourselves to the case of a linear drift current, i.e., when $\boldsymbol{X}$ is given by

$$\boldsymbol{X}(q_1,q_2) = \begin{pmatrix} a & b\\ c & d\end{pmatrix}\begin{pmatrix} q_1\\ q_2\end{pmatrix} = Aq.$$

In this very special case, one can immediately see from formula (3.7) the curvature does not depend on the point $q\in\mathbb{R}^2$ but only on the angle $\theta$ between $\boldsymbol{e}_1$ and $(u_1\boldsymbol{e}_1+u_2\boldsymbol{e}_2)$ and is given by the quadratic form

$$\kappa(a,b,c,d,\theta) = Q(\theta)(a,b,c,d).$$

Looking at the matrix of the quadratic form $Q(\theta)$ we see that it remains quite difficult to study this quadratic form for a general matrix $A$. So, we need to find normal forms for the matrix $A$. Namely, we want to find the linear transformations of $\mathbb{R}^2$ that preserve the control system. For instance, it is easy to see that a linear transformation $P \in \mathrm{GL}(2,\mathbb{R})$ does not change the dynamics of the control system if and only if the basis of constant vector fields $\{P\boldsymbol{e}_1, P\boldsymbol{e}_2\}$ is orthonormal, i.e., if $P \in \mathrm{O}(2,\mathbb{R})$.

**Proposition 3.1.2.** *For every matrix $A \in \mathrm{M}(2,\mathbb{R})$, there exists a matrix $U \in \mathrm{O}(2,\mathbb{R})$ such that $U^\dagger A U = \left(\begin{smallmatrix} a & b \\ -b & c\end{smallmatrix}\right)$.*

***Proof.*** Let us write $A = A_{\text{sym}} + A_{\text{ant}}$ where $A_{\text{sym}}$ and $A_{\text{ant}}$ are the symmetric and antisymmetric parts of the matrix $A$, respectively. It follows from the symmetry of $A_{\text{sym}}$ that there exists a matrix $U \in \mathrm{O}(2,\mathbb{R})$ such that $U^\dagger A_{\text{sym}} U$ is diagonal. Now remark that $U^\dagger A_{\text{ant}} U$ is antisymmetric. The matrix $U^\dagger A U = U^\dagger A_{\text{sym}} U + U^\dagger A_{\text{ant}} U$ is of the desired form. ∎

Putting the matrix $A$ in the normal form modulo $U \in \mathrm{O}(2,\mathbb{R})$, we easily see that

$$\kappa(\theta) = -\frac{1}{32}\big(23a^2 - 38ac + 23c^2 - 12(a^2-c^2)\cos 2\theta \\ -3(a-c)^2\cos 4\theta - 48b(a-c)\sin 2\theta\big).$$

Let us look at this formula in some special cases.
*First case.* The matrix $A$ is diagonal equal to $\left(\begin{smallmatrix} a & 0 \\ 0 & b\end{smallmatrix}\right)$. In this case the curvature becomes

$$\kappa(a,b,\theta) = \frac{1}{32}\big(-23a^2 + 38ab - 23b^2 + 12(a^2-b^2)\cos 2\theta + 3(a-b)^2\cos 4\theta\big)$$

and we have the following proposition:

**Proposition 3.1.3.** *If the matrix $A$ is symmetric, then the curvature of the system is non positive.*

***Proof.*** Since every real symmetric matrix is equivalent to a diagonal matrix modulo the orthogonal linear group, it is enough to make the proof when $A$ is diagonal. But, the curvature is given the quadratic form of $a$ and $b$ whose matrix is

$$M = -\frac{1}{32}\begin{pmatrix} 23 - 12\cos 2\theta - 3\cos 4\theta & -19 + 3\cos 4\theta \\ -19 + 3\cos 4\theta & 23 + 12\cos 2\theta - 3\cos 4\theta \end{pmatrix}.$$

But

$$23 - 12\cos 2\theta - 3\cos 4\theta \geqslant 23 - 12 - 3 > 0,$$

and

$$\begin{aligned} \det M &= (23 - 3\cos 4\theta - 12\cos 2\theta)(23 - 3\cos 4\theta + 12\cos 2\theta) - (19 - 3\cos 4\theta)^2 \\ &= (23 - 3\cos 4\theta)^2 - 144\cos^2 2\theta - (19 - 3\cos 4\theta)^2 \\ &= 4(42 - 6\cos 4\theta) - 144\cos^2 2\theta \geqslant 4(42-6) - 144 \geqslant 0. \end{aligned}$$

Then using Sylvester's criterion one concludes that $Q(a,0,0,b)$ is a positive quadratic form, i.e., that $\kappa(a,b) \leqslant 0$. ■

*Second case.* The matrix $A$ is the similitude $\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$. In this case the curvature is a negative constant of the element $a$:

$$\kappa = -a^2/4.$$

*Third case.* The matrix is the Jordan bloc $\left(\begin{smallmatrix} a & 1 \\ 0 & a \end{smallmatrix}\right)$. In this case the curvature is

$$\kappa(a,\theta) = -\frac{1}{32}(21 + 8a^2 + 24\cos 2\theta + 3\cos 4\theta - 24a\sin 2\theta)$$

and it can be positive if, for example, $a = 1$ and $\theta = 3\pi/8$.

Since the curvature is non positive for the first two cases, we can conclude, using Theorem 2.6.4 that there is no conjugate points along regular extremals. For the third case the estimates given by Sturm comparison theorem are not precise enough in order to conclude something on the occurrence of conjugate points. Nevertheless the following theorem holds. This theorem was firstly proved in [39] for the two-dimensional case.

**Theorem 3.1.4.** *In the case of a linear drift $\boldsymbol{X}(q) = Aq$, Zermelo navigation problem on $\mathbb{R}^n$ has no conjugate points.*

***Proof.*** Let $q_0 \in \mathbb{R}^n$ and $p_0 \in T^*_{q_0}\mathbb{R}^n \cap \mathcal{H}$. Remark that the linearity of $\boldsymbol{X}(q)$ reduces the adjoint equation of PMP $\dot{p} = -p\partial \boldsymbol{X}/\partial q$ to the linear equation on $\mathbb{R}^n$ $\dot{p} = -pA$ and the equation $\dot{q} = Aq + p^\dagger/|p|$ to a non autonomous linear equation in $\mathbb{R}^n$. Hence we can easily compute the map $\Gamma : (t, p_0) \mapsto q(t, q_0, p_0)$ and its differential which gives:

$$\Gamma(t,p_0) = q(t,q_0,p_0) = e^{tA}q_0 + \int_0^t e^{(t-s)A}\frac{p^\dagger(s)}{|p(s)|}ds \tag{3.8}$$

$$D_{(t,p_0)}\Gamma = \left(\frac{\partial \Gamma}{\partial t}, \frac{\partial \Gamma}{\partial p_0}\right) \tag{3.9}$$

where

$$\frac{\partial \Gamma}{\partial t} = \dot{q} = \pi_*\vec{\boldsymbol{h}} \tag{3.10}$$

and

$$\frac{\partial \Gamma}{\partial p_0} = e^{tA}\int_0^t e^{-sA}D_{p^\dagger(s)}(p^\dagger/|p|)\frac{\partial p^\dagger(s)}{\partial p_0}ds = e^{tA}\int_0^t e^{-sA}D_{p^\dagger(s)}(p^\dagger/|p|)e^{-sA^\dagger}ds$$

with

$$D_{p^\dagger(s)}(p^\dagger/|p|) = \frac{1}{|p(s)|}\left(\mathrm{Id}_n - \frac{p^\dagger(s)p(s)}{|p(s)|^2}\right) = \frac{1}{|p(s)|}\pi_{p^{\dagger\perp}} \tag{3.11}$$

where $\pi_{p^{\dagger\perp}}$ is the projection on the orthogonal complement to $\mathbb{R}p^\dagger$. This shows that $D_{p^\dagger(s)}(p^\dagger/|p|)$ is a positive symmetric matrix of rank $n-1$.
Denote by $N$, $N_1, \dots, N_n$ the maps $p \mapsto p^\dagger/|p|$, $p \mapsto p_1/|p|, \dots, p \mapsto p_n/|p|$ respectively. All these maps are homogeneous of degree zero which implies (think to Euler identity) that

$$p^\dagger \in \ker D_p N_i \quad \forall i \in \{1, \dots, n\}. \tag{3.12}$$

It follows from (3.12) and the symmetry of $D_p N$ that:

$$\left\langle p(t), \frac{\partial \Gamma}{\partial p_{0i}} \right\rangle = 0 \quad \forall i \in \{1, \dots, n\}. \tag{3.13}$$

Since the Hamiltonian $h$ is one-homogeneous, it follows from Lemma (2.9) and (3.10) that

$$\left\langle p(t), \frac{\partial \Gamma}{\partial t}(t, p_0) \right\rangle = 1. \tag{3.14}$$

Now remark that (3.13) and (3.14) together imply that $\partial\Gamma/\partial t$ and $\partial\Gamma/\partial p_{0i}$ are linearly independent whenever $\partial\Gamma/\partial p_{0i} \neq 0$ which reduces the proof of the theorem to show that $\partial\Gamma/\partial p_0$ has rank $n-1$ for all $t > 0$ but this will be done if we can prove that $\xi \partial\Gamma/\partial p_0 \xi^\dagger \neq 0$ (or equivalently $\xi e^{-tA}\partial\Gamma/\partial p_0 \xi^\dagger \neq 0$ since $e^{-tA}$ is a diffeomorphism) holds true for $n-1$ independent covectors $\xi$.

Observe that it follows from the regularity of trajectories we are considering that the covector $p_0$ is transverse to $T^*_{p_0} H_{q_0}$ so the covector $p(s) = p_0 e^{-sA}$ is transverse to $e^{t\vec{h}*} T^*_{p_0} H_{q_0} = \text{span}\{\xi e^{-sA}, \xi \in T^*_{p_0} H_{q_0}\} \subset T^*_{(t,p_0)} S_{q_0}$. In addition to (3.11) this last remark implies that

$$\xi e^{-sA} D_{p^\dagger(s)} N e^{-sA^\dagger} \xi^\dagger > 0 \quad \forall \xi \in T^*_{p_0} H_{q_0} \setminus 0.$$

Integration of the previous inequality leads to

$$\xi e^{-tA} \frac{\partial \Gamma}{\partial p} \xi^\dagger = \int_0^t \xi e^{-sA} D_{p^\dagger(s)} N e^{-sA^\dagger} \xi^\dagger ds > 0 \quad \forall \xi \in T^*_{p_0} H_{q_0} \setminus 0 \quad \forall t > 0$$

which shows that $e^{-tA}\partial\Gamma/\partial p_0$ has rank $n-1$ and thus completes the proof. ■

We now conclude our study of Zermelo's navigation problem by giving the optimal synthesis when the drift current $\boldsymbol{X}$ is linear and its matrix a similitude.

Now assume that the drift current $\boldsymbol{X}$ is defined by the by the formula

$$\boldsymbol{X}(q_1, q_2) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}.$$

The navigation formula of Zermelo then reads

$$\theta' = -b,$$

which leads to

$$\theta(t) = \theta_0 - b\,t.$$

Set $z = q_1 + iq_2$, and $\lambda = a - ib$, where $i$ is the complex number such that $i^2 = -1$. Then the system can be written as

$$z'(t) = \lambda z(t) + e^{i(\theta_0 - bt)},\ z(0) = z_0.$$

Hence

$$z(t) = e^{\lambda t} z_0 + \int_0^t e^{\lambda(t-s)} e^{i(\theta_0 - bs)} ds = e^{at} e^{-ibt} \left( z_0 + e^{i\theta_0} \int_0^t e^{-as} ds \right),$$

which gives

$$z(t) = \begin{cases} e^{-ibt}\left(z_0 + e^{i\theta_0} t\right) & \text{if } a = 0, \\ e^{at} e^{-ibt} \left( z_0 + e^{i\theta_0} \dfrac{1 - e^{-at}}{a} \right) & \text{if } a \neq 0. \end{cases}$$

Below are drawn the optimal synthesis when $a$ and $b$ are different from zero for the three different cases of existence of abnormal trajectories passing through the departure point $q_0$. The bold lines represent the boundary of the attainable set starting from $q_0$, i.e., the abnormal trajectories. Notice that the region where the speed of the ship is equal to the speed of the drift current corresponds to the ellipse $a^2 q_1^2 + b^2 q_2^2 = 1$.

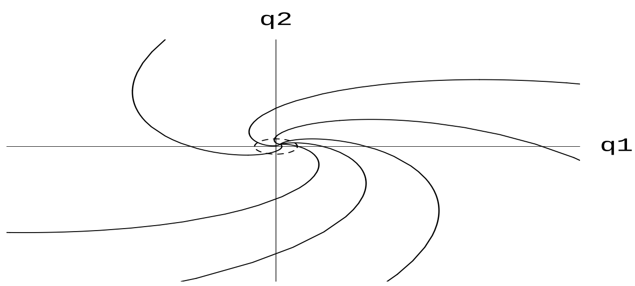


Figure 3.1: $X_1^2(q_0) + X_2^2(q_0) < 1$.

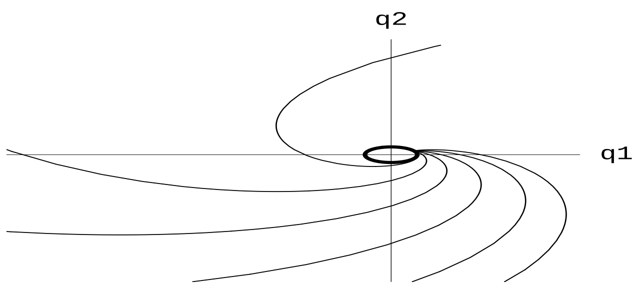


Figure 3.2: $X_1^2(q_0) + X_2^2(q_0) = 1$.

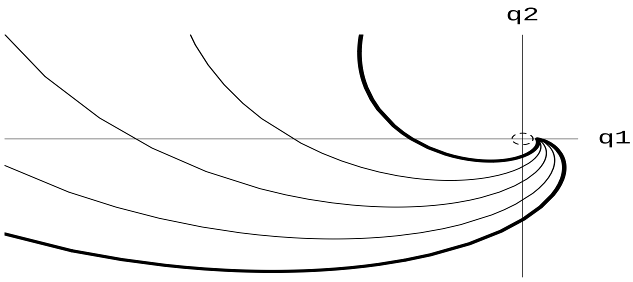


Figure 3.3: $X_1^2(q_0) + X_2^2(q_0) > 1$.

## 3.2 Dual Zermelo navigation problem

### 3.2.1 Definition of the problem

Recall that for Zermelo navigation problem the optimal dynamics on the manifold $M$ are given by

$$\dot{q} = \boldsymbol{X}(q) + \cos u\, \boldsymbol{e}_1(q) + \sin u\, \boldsymbol{e}_2(q), \quad q \in M, \quad u \in S^1, \quad \boldsymbol{X} \in \operatorname{Vec} M,$$

so that the curves of admissible velocities are the curves

$$F_q = \boldsymbol{X}(q) + \mathcal{S}_q \in T_q M,$$

where $\mathcal{S}_q$ is the fiber at point $q$ of the Riemannian spherical bundle over $M$:

$$\mathcal{S} = \bigcup_{q \in M} \mathcal{S}_q = \{v \in TM \mid \langle v, v \rangle_g = 1\}.$$

Roughly speaking, dual Zermelo navigation problem will be the time optimal problem for which the curves $\mathcal{H}_q$ are of the form $\Upsilon(q) + \mathcal{S}_q^*$, where $\Upsilon \in \Lambda^1(M)$ and $\mathcal{S}_q^*$ is the fiber at point $q \in M$ of the dual spherical bundle to $\mathcal{S}$ with respect to the diffeomorphism (1.13).

Let us begin with the following definition.

**Definition 3.2.1.** *If $\mathcal{H}$ is a subbundle of $T^*M$, we say that the subbundle $\mathcal{F} \subset TM$ defined by $\mathcal{F} = \{\boldsymbol{f} \in TM \mid \max_{\lambda \in F} \langle \lambda, \boldsymbol{f} \rangle = 1\}$ is dual to $\mathcal{H}$.*

Then we define dual Zermelo navigation problem as follows.

**Definition 3.2.2.** *We call dual problem to Zermelo navigation problem on the Riemannian manifold $(M, g)$ the minimum time problem defined by the differential inclusion $\dot{q}(t) \in \mathcal{F}_q$ where $\mathcal{F} = \cup_{q \in M} \mathcal{F}_q \subset TM$ is the subbundle of $TM$ dual to the subbundle $\mathcal{H} = \Upsilon + \mathcal{S}^{g*} \subset T^*M$ where $\Upsilon$ is a one-form on $(M, g)$ such that $|\Upsilon|_g < 1$.*

### 3.2.2 Hamiltonian function

Let $h \in C^\infty(M)$ be the maximized Hamiltonian function of PMP for a dual to Zermelo problem. Consequently, the function $h$ is one-homogeneous on fibers $T_q^*M$. Denote by $\mathcal{H}$ the hypersurface $h^{-1}(1)$ and by $\mathcal{H}_q$ the fiber $\mathcal{H} \cap T_q^*M$. By definition of a dual Zermelo navigation problem the curves $\mathcal{H}_q$ are characterized by

$$\langle \lambda - \Upsilon(q), \lambda - \Upsilon(q) \rangle_g = 1, \quad \forall \lambda \in \mathcal{H}_q. \tag{3.15}$$

Suppose now that $\lambda \in T_q^*M$ is a non zero covector such that $h(\lambda) \neq 0$. Then, using the homogeneity of $h$ we get

$$\frac{\lambda}{h(\lambda)} \in \mathcal{H}_q.$$

Consequently, the covector $\lambda/h(\lambda)$ has to satisfy equation (3.15). Plugging this covector in equation (3.15) leads to

$$\langle \lambda - h(\lambda)\Upsilon(q), \lambda - h(\lambda)\Upsilon(q) \rangle_g^2 = h(\lambda)^2, \quad \lambda \in T_q^*M \tag{3.16}$$

or, equivalently to

$$\left(1 - |\Upsilon(q)|_g^2\right) h(\lambda)^2 + 2 \langle \lambda, \Upsilon(q) \rangle_g h(\lambda) - |\lambda|_g^2 = 0, \tag{3.17}$$

which gives an implicit definition for the Hamiltonian function $h$. Solving equation (3.17) for $h(\lambda)$ gives

$$h(\lambda) = \frac{-\langle \lambda, \Upsilon(q) \rangle_g + \sqrt{\langle \lambda, \Upsilon(q) \rangle_g^2 + \left(1 - |\Upsilon(q)|_g^2\right) |\lambda|_g^2}}{1 - |\Upsilon(q)|_g^2}.$$

## 3.3 Link between Zermelo and dual to Zermelo problems

In this section we prove a proposition which asserts the equivalence between Zermelo navigation problem and dual to Zermelo problem under the assumption that the Riemannian norm of the drift is strictly less than one.

Let $(M, g)$ be a Riemannian manifold and fix an orthonormal frame $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ for $g$.

If $\boldsymbol{X} \in \operatorname{Vec} M$, we define the local orthonormal frame for $g$ associated to the vector field $\boldsymbol{X}$ with respect to the frame $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ by

$$\begin{aligned} \boldsymbol{e}_1^{\boldsymbol{X}} &= \cos\theta^{\boldsymbol{X}} \boldsymbol{e}_1 + \sin\theta^{\boldsymbol{X}} \boldsymbol{e}_2 \\ \boldsymbol{e}_2^{\boldsymbol{X}} &= -\sin\theta^{\boldsymbol{X}} \boldsymbol{e}_1 + \cos\theta^{\boldsymbol{X}} \boldsymbol{e}_2 \,, \end{aligned}$$

where $q \mapsto \theta^{\boldsymbol{X}}(q)$ is the angle defined by

$$\begin{cases} \theta^{\boldsymbol{X}}(q) = 0 & \text{if } \boldsymbol{X}(q) = 0, \\ \cos\theta^{\boldsymbol{X}}(q) = \dfrac{\langle \boldsymbol{X}(q), \boldsymbol{e}_1(q)\rangle_g}{|\boldsymbol{X}(q)|_g}, \quad \sin\theta^{\boldsymbol{X}}(q) = \dfrac{\langle \boldsymbol{X}(q), \boldsymbol{e}_2(q)\rangle_g}{|\boldsymbol{X}(q)|_g} & \text{if } \boldsymbol{X}(q) \neq 0. \end{cases}$$

In the same way if $\Upsilon \in \Lambda^1(M)$ we define the orthonormal frame associated to the one-form $\Upsilon$ with respect to $\boldsymbol{e}_1, \boldsymbol{e}_2$ by

$$\begin{aligned} \boldsymbol{e}_1^{\Upsilon} &= \quad \cos\theta^{\Upsilon}\boldsymbol{e}_1 + \sin\theta^{\Upsilon}\boldsymbol{e}_2 \\ \boldsymbol{e}_2^{\Upsilon} &= -\sin\theta^{\Upsilon}\boldsymbol{e}_1 + \cos\theta^{\Upsilon}\boldsymbol{e}_2 \ , \end{aligned}$$

where $q \mapsto \theta^{\Upsilon}(q)$ is the angle defined by

$$\begin{cases} \theta^{\Upsilon}(q) = 0 & \text{if } \Upsilon(q) = 0, \\ \cos\theta^{\Upsilon}(q) = \dfrac{\langle \Upsilon(q), \boldsymbol{e}_1(q)\rangle}{|\Upsilon(q)|_g}, \quad \sin\theta^{\Upsilon}(q) = \dfrac{\langle \Upsilon(q), \boldsymbol{e}_2(q)\rangle}{|\Upsilon(q)|_g} & \text{if } \Upsilon(q) \neq 0. \end{cases}$$

Notice that in this frames

$$\boldsymbol{X} = \langle \boldsymbol{X}, \boldsymbol{e}_1^{\boldsymbol{X}}\rangle_g \, \boldsymbol{e}_1^{\boldsymbol{X}} = |\boldsymbol{X}|_g \boldsymbol{e}_1^{\boldsymbol{X}}, \quad \Upsilon = \langle \Upsilon, \boldsymbol{e}_1^{\Upsilon}\rangle \, \boldsymbol{e}_1^{\Upsilon *} = |\Upsilon|_g \boldsymbol{e}_1^{\Upsilon *}.$$

Suppose for now that the Riemannian norm of the drift in our Zermelo navigation is strictly less than one.

**Proposition 3.3.1.** *Any Zermelo navigation problem on a Riemannian manifold can be seen as a dual to Zermelo navigation problem on a Riemannian manifold and vice versa.*

***Proof.*** Consider a Zermelo navigation problem (3.1) on a the Riemannian manifold $(M, g)$ such that $|\boldsymbol{X}|_g < 1$ and let $(\boldsymbol{e_1}, \boldsymbol{e}_2)$ be an orthonormal frame for the metric $g$. Recall that, in the polar coordinates $\lambda = (r, \theta)$ defined by relations (3.3), the Hamiltonian function of PMP takes the form

$$h(r, \theta, q) = r\left(|\boldsymbol{X}(q)|_g \cos(\theta - \theta^{\boldsymbol{X}}(q)) + 1\right),$$

where $\theta^{\boldsymbol{X}}(q)$ is the angle defined by relations (3.4). Thus, the curve $\mathcal{H}_q = h^{-1}(1) \cap T_q^*M$ has the polar equation

$$r(\theta) = \frac{1}{|\boldsymbol{X}(q)|_g \cos(\theta - \theta^{\boldsymbol{X}}(q)) + 1}. \tag{3.18}$$

Since $|\boldsymbol{X}|_g < 1$, the curve $\mathcal{H}_q$ is an ellipse centered at a focus. Moreover this ellipse has for $g$ a focal distance

$$c = \frac{r(\pi + \theta^{\boldsymbol{X}}) - r(\theta^{\boldsymbol{X}})}{2} = \frac{|\boldsymbol{X}|_g}{1 - |\boldsymbol{X}|_g^2},$$

a semimajor distance

$$a = r(\theta^{\boldsymbol{X}}) + r(\pi + \theta^{\boldsymbol{X}}) = \frac{1}{1 - |\boldsymbol{X}|_g^2},$$

and a semiminor distance

$$b = \sqrt{a^2 - c^2} = \frac{1}{\sqrt{1 - |\boldsymbol{X}|_g^2}}.$$

In order to transform Zermelo navigation problem in dual Zermelo problem, we consider the curve $\mathcal{H}_q$ as the drifted dual Riemannian sphere at point $q$ for a new Riemannian structure $\tilde{g}$ on the manifold. In other words, we ask the one-forms

$$\tilde{\boldsymbol{e}}_1^* = \frac{1}{1 - |\boldsymbol{X}|_g^2} \boldsymbol{e}_1^{\boldsymbol{X}*}, \quad \tilde{\boldsymbol{e}}_2^* = \frac{1}{\sqrt{1 - |\boldsymbol{X}|_g^2}} \boldsymbol{e}_2^{\boldsymbol{X}*}$$

to be a dual orthonormal Riemannian basis for the new Riemannian structure $\tilde{g}$ on the manifold and the one-form

$$\Upsilon = -c\, \boldsymbol{e}_1^{\boldsymbol{X}*} = -\frac{|\boldsymbol{X}|_g}{1 - |\boldsymbol{X}|_g^2} \boldsymbol{e}_1^{\boldsymbol{X}*}$$

to be the drift one-form of dual to Zermelo problem. The corresponding (new) orthonormal frame $(\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_2)$ is characterized by

$$\langle (\tilde{\boldsymbol{e}}_1^*, \tilde{\boldsymbol{e}}_2^*), (\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_2) \rangle = \mathrm{Id},$$

which leads to

$$\tilde{\boldsymbol{e}}_1 = \left(1 - |\boldsymbol{X}|_g^2\right) \boldsymbol{e}_1^{\boldsymbol{X}}, \quad \tilde{\boldsymbol{e}}_2 = \sqrt{1 - |\boldsymbol{X}|_g^2}\, \boldsymbol{e}_2^{\boldsymbol{X}}.$$

Notice that we have

$$\Upsilon = -\boldsymbol{X}^{\flat_{\tilde{g}}}, \quad (\tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_2) = (\tilde{\boldsymbol{e}}_1^{-\Upsilon}, \tilde{\boldsymbol{e}}_2^{-\Upsilon}),$$

which shows in particular that

$$|\boldsymbol{X}|_g = |\Upsilon|_{\tilde{g}}.$$

In order to complete the proof it remains to check that our original Zermelo problem on $(M, g)$ with drift vector field $\boldsymbol{X}$ and the dual to Zermelo problem on $(M, \tilde{g})$ with drift form $\Upsilon$ have the same Hamiltonians. Denote by $h_{Zg}$ the Hamiltonian of our original Zermelo navigation problem and by $h_{DZ\tilde{g}}$ the Hamiltonian function of the associated dual to Zermelo problem. For simplicity we also denote $y = |\Upsilon|_{\tilde{g}} = |\boldsymbol{X}|_g$.

We have

$$
\begin{aligned}
h_{Zg}(\lambda) &= \langle \lambda, \boldsymbol{X} \rangle + |\lambda|_g = \langle \lambda, y\boldsymbol{e}_1^{\boldsymbol{x}} \rangle + \sqrt{\langle \lambda, \boldsymbol{e}_1^{\boldsymbol{x}} \rangle^2 + \langle \lambda, \boldsymbol{e}_2^{\boldsymbol{x}} \rangle^2} \\
&= \left\langle \lambda, y\frac{\tilde{\boldsymbol{e}}_1}{1-y^2} \right\rangle + \sqrt{\left\langle \lambda, \frac{\tilde{\boldsymbol{e}}_1}{1-y^2} \right\rangle^2 + \left\langle \lambda, \frac{\tilde{\boldsymbol{e}}_2}{\sqrt{1-y^2}} \right\rangle^2} \\
&= \frac{\langle \lambda, y\tilde{\boldsymbol{e}}_1 \rangle + \sqrt{\langle \lambda, \tilde{\boldsymbol{e}}_1 \rangle^2 + (1-y^2) \langle \lambda, \tilde{\boldsymbol{e}}_2 \rangle^2}}{1-y^2} \\
&= \frac{\langle \lambda, y\tilde{\boldsymbol{e}}_1 \rangle + \sqrt{\langle \lambda, \tilde{\boldsymbol{e}}_1 \rangle^2 + \langle \lambda, \tilde{\boldsymbol{e}}_2 \rangle^2 - y^2 \langle \lambda, \tilde{\boldsymbol{e}}_2 \rangle^2 - y^2 \langle \lambda, \tilde{\boldsymbol{e}}_1 \rangle^2 + y^2 \langle \lambda, \tilde{\boldsymbol{e}}_1 \rangle^2}}{1-y^2} \\
&= \frac{-\langle \lambda, -y\tilde{\boldsymbol{e}}_1 \rangle + \sqrt{\left(\langle \lambda, \tilde{\boldsymbol{e}}_1 \rangle^2 + \langle \lambda, \tilde{\boldsymbol{e}}_2 \rangle^2\right)(1-y^2) + (-y \langle \lambda, \tilde{\boldsymbol{e}}_1 \rangle)^2}}{1-y^2} \\
&= \frac{-\langle \lambda, \langle \Upsilon, \tilde{\boldsymbol{e}}_1 \rangle \tilde{\boldsymbol{e}}_1^* \rangle_{\tilde{g}} + \sqrt{|\lambda|_{\tilde{g}} (1-y^2) + (\langle \Upsilon, \tilde{\boldsymbol{e}}_1 \rangle \langle \lambda, \tilde{\boldsymbol{e}}_1 \rangle)^2}}{1-y^2} \\
&= \frac{-\langle \lambda, \Upsilon \rangle_{\tilde{g}} + \sqrt{\left(1 - |\Upsilon|_{\tilde{g}}^2\right) |\lambda|_{\tilde{g}} + \langle \lambda, \Upsilon \rangle_{\tilde{g}}^2}}{1 - |\Upsilon|_{\tilde{g}}^2} \\
&= h_{DZ\tilde{g}}(\lambda).
\end{aligned}
$$

This shows that the Zermelo navigation problem on $(M, g)$ with drift vector field $\boldsymbol{X}$ is feedback equivalent to the dual to Zermelo on $(M, \tilde{g})$ with drift one-form $\Upsilon = -\boldsymbol{X}^{\flat_{\tilde{g}}}$. In order to prove the converse, one has just to permute the roles of vector fields and one forms in the previous considerations. ■

The previous proposition can be reformulate in the following manner.

**Corollary 3.3.2.** *If $\kappa$ is the curvature of a Zermelo navigation (respectively dual to Zermelo) problem on $(M, g)$ then, there exists a metric $\tilde{g}$ on $M$ and a dual to Zermelo (respectively Zermelo navigation) problem on $(M, \tilde{g})$ having curvature $\kappa$.*

**Remark 3.3.3.** Proposition easily generalizes for the case in which the Riemannian of the drift is strictly grater than one. In this case Proposition 3.3.1 sounds a bit different because a Zermelo navigation problem on a Riemannian manifold will be transformed in a dual to Zermelo problem on a Lorentzian manifold and vice versa. The change of structure (from Riemannian to Lorentzian) does not matter since Zermelo problems can obviously be defined on Lorentzian manifold. This change of structure is actually very easy to understand since a Zermelo's navigation problem whose drift has a Riemannian norm strictly bigger than one admits abnormal extremals.

# Chapter 4

# Microlocal normal forms for control systems

The present chapter deals with the feedback classification of nonlinear two-dimensional control systems with scalar input.

The feedback classification of control systems has already been studied a lot beginning with the equivalence problem for pencils of matrices presented by F. R. Gantmacher in [22]. This classification gave rise to the well-known Brunovsky normal forms for linear control systems with constant coefficients (presented by Brunovsky in [17]). Then, the feedback classification problem for control-affine systems with scalar input was heavily studied in [1, 27, 28, 29, 33, 37, 38] where the authors also gave list of normal forms. Finally, in [10], A. A. Agrachev and I. Zelenko completely solved the problem of the local classification generic control-affine systems on a $n$-dimensional manifold with scalar input for any $n \geqslant 4$ and with two inputs for $n = 4$ and $n = 5$ by giving a complete set of invariants for these equivalence problems.

In the first part of this chapter we construct two microlocal normal forms for generic nonlinear two-dimensional control systems with scalar input satisfying the convexity condition $\frac{\partial \boldsymbol{f}}{\partial u} \wedge \frac{\partial^2 \boldsymbol{f}}{\partial u^2} \neq 0$. The first microlocal normal form will be given around a normal extremal and the second around an abnormal extremal.

In Section 4.3 we present and prove two theorems (firstly presented without proof in [40]). The first theorem gives a characterization of control systems satisfying the following property: there exists a feedback transformation such that the vector fields $\boldsymbol{f}$ and $\frac{\partial \boldsymbol{f}}{\partial u}$ commute. The second theorem gives a characterization of flat control systems. These two theorems provide also checkable conditions in terms of the feedback invariants of the system for the characterization they give.

## 4.1 Principal invariants of the equivalence problem

### 4.1.1 Counting the invariants

Consider a smooth control system of the form

$$\dot{q} = \boldsymbol{f}(q, u), \quad q \in M, \quad u \in U, \tag{4.1}$$

where $M$ is a two-dimensional manifold and $U$ is a one-dimensional manifold. For such a system we want to find some microlocal form around a point $(q_0, u_0) \in TM$ modulo feedback transformations. In other words, the problem is to transform the nonlinear system (4.1) into a simpler form by a feedback transformation of the form

$$\Theta(q, u) = (\phi(q), \psi(q, u)).$$

First of all, let us roughly estimate the number of parameters (invariants) in this equivalence problem. In this case, if the coordinates on the manifold are fixed, a control system of type (2.1) is parametrized by two functions of three variables, and the group of feedback transformations of type (2.4) is parametrized by two functions of two variables and one function of three variables. Indeed, in any local coordinate chart $q = (q_1, q_2)$ on the base manifold $M$, the control system reads

$$\begin{aligned} \dot{q}_1 &= f_1(q, u) \\ \dot{q}_2 &= f_2(q, u), \end{aligned}$$

and an element of the group of feedback transformations takes the form

$$\Theta(q, u) = (\phi_1(q, u), \phi_2(q, u), \psi(q, u)),$$

where $f_1$, $f_2$, $\phi_1$, $\phi_2$ and $\psi$ are real valued functions. Therefore, we can a priori normalize only one function among the two functions defining control system (2.1). Thus, we expect to have only $2 - 1 = 1$ "principal" feedback invariant, i.e., a function of three variables and a certain number of feedback-invariant functions of less than three variables, in this equivalence problem.

### 4.1.2 Relation between the principal invariants

In Chapter 2, we constructed for control systems of type (4.1) under the regularity assumptions on the curves of admissible velocities

$$\boldsymbol{f}(q, u) \wedge \frac{\partial \boldsymbol{f}(q, u)}{\partial u} \neq 0 \tag{4.2}$$

$$\frac{\partial \boldsymbol{f}(q, u)}{\partial u} \wedge \frac{\partial^2 \boldsymbol{f}(q, u)}{\partial u^2} \neq 0, \quad q \in M, \; u \in U, \tag{4.3}$$

a feedback-invariant moving frame on a three-dimensional manifold. This construction led to the construction of two feedback-invariants of three variables: the feedback-invariant $b$ defined by the relation (2.15) and the curvature $\kappa$ defined by the bracket relation (2.21). Both $b$ and $\kappa$ are functions on the three-dimensional level surface $\mathcal{H}$, so that they are principal feedback invariants of our control system. Since our feedback equivalence problem admits only one invariant these functions are not "independent". Indeed we have the following proposition.

**Proposition 4.1.1.** *The feedback invariants $b$ and $\kappa$ satisfy the following second order PDE:*

$$L_{\boldsymbol{v}}\kappa + b\kappa + L^2_{\vec{\boldsymbol{h}}}b = 0. \tag{4.4}$$

Before proving the proposition we need an auxiliary lemma.

**Lemma 4.1.2.** *Suppose that we are given a parameter $\theta$ on the fiber $\mathcal{H}_q$ such that $\boldsymbol{v} = \frac{\partial}{\partial\theta}$. Then, the structure constant $c$ defined by $d\omega_\theta = c\,\epsilon\omega_\theta \wedge \omega'_\theta$ satisfies the following second order PDE:*

$$c'' + bc' + \epsilon c = L_{\vec{\boldsymbol{h}}}b. \tag{4.5}$$

***Proof.*** From (2.24) one infers that

$$d\omega'_\theta = (c' + bc)\epsilon\omega_\theta \wedge \omega'_\theta.$$

Differentiating this equality with respect to $\theta$ leads, one the one hand, to

$$\begin{aligned}(d\omega'_\theta)' &= (c'' + b'c + bc')\epsilon\omega_\theta \wedge \omega'_\theta + (c' + bc)\epsilon\omega_\theta \wedge \omega''_\theta\\ &= (c'' + b'c + bc')\epsilon\omega_\theta \wedge \omega'_\theta + (c' + bc)\epsilon\omega_\theta \wedge (-\epsilon\omega_\theta + b\,\omega'_\theta)\\ &= (c'' + 2bc' + b'c + b^2c)\epsilon\omega_\theta \wedge \omega'_\theta.\end{aligned}$$

On the other hand,

$$\begin{aligned}(d\omega'_\theta)' &= d\omega''_\theta = d(-\epsilon\omega_\theta + b\,\omega'_\theta)\\ &= -c\,\omega_\theta \wedge \omega'_\theta + d_q b\,\omega'_\theta + b(c' + bc)\epsilon\omega_\theta \wedge \omega'_\theta\\ &= (-\epsilon c + L_{\boldsymbol{f}}b + bc' + b^2c)\epsilon\omega_\theta \wedge \omega'_\theta.\end{aligned}$$

Summing up, we get

$$c'' + 2bc' + b'c + b^2c = -\epsilon c + L_{\boldsymbol{f}}b + bc' + b^2c,$$

or equivalently,

$$c'' + bc' + \epsilon c = L_{\boldsymbol{f}}b - b'c = L_{\vec{\boldsymbol{h}}}b,$$

which ends the proof of the lemma. ■

***Proof of Proposition 4.1.1.*** From the previous lemma it follows that

$$\left[\frac{\partial}{\partial\theta},\left[\vec{\boldsymbol{h}},\frac{\partial}{\partial\theta}\right]\right] = -\vec{\boldsymbol{h}}'' = -\boldsymbol{f}'' + c''\frac{\partial}{\partial\theta} = \epsilon\boldsymbol{f} + b\boldsymbol{f}' + (L_{\vec{\boldsymbol{h}}}b - bc' - \epsilon c)\frac{\partial}{\partial\theta}$$
$$= \epsilon\vec{\boldsymbol{h}} + b\vec{\boldsymbol{h}}' + L_{\vec{\boldsymbol{h}}}b\frac{\partial}{\partial\theta}.$$

If we now compute the Lie bracket of the previous relation with $\vec{\boldsymbol{h}}$, we get for the right hand side

$$\left[\vec{\boldsymbol{h}},\epsilon\vec{\boldsymbol{h}} + b\vec{\boldsymbol{h}}' + L_{\vec{\boldsymbol{h}}}b\frac{\partial}{\partial\theta}\right] = L_{\vec{\boldsymbol{h}}}b\vec{\boldsymbol{h}}' + b\left[\vec{\boldsymbol{h}},\vec{\boldsymbol{h}}'\right] - L_{\vec{\boldsymbol{h}}}b\vec{\boldsymbol{h}}' + L^2_{\vec{\boldsymbol{h}}}b\frac{\partial}{\partial\theta} = (b\kappa + L^2_{\vec{\boldsymbol{h}}}b)\frac{\partial}{\partial\theta},$$

and using Jacobi identity, we get for the left hand side

$$\left[\vec{\boldsymbol{h}},\left[\frac{\partial}{\partial\theta},\left[\vec{\boldsymbol{h}},\frac{\partial}{\partial\theta}\right]\right]\right] = -\left[\frac{\partial}{\partial\theta},\left[\left[\vec{\boldsymbol{h}},\frac{\partial}{\partial\theta}\right],\vec{\boldsymbol{h}}\right]\right] - \left[\left[\vec{\boldsymbol{h}},\frac{\partial}{\partial\theta}\right],\left[\vec{\boldsymbol{h}},\frac{\partial}{\partial\theta}\right]\right]$$
$$= -\left[\frac{\partial}{\partial\theta},\kappa\frac{\partial}{\partial\theta}\right] = \kappa'\frac{\partial}{\partial\theta}.$$

and the equation follows. ∎

Notice that equation (4.4) shows that in the special case of Riemannian problems, the curvature $\kappa$ is a function on the base manifold $M$ without any computation. Indeed, since Riemannian problems are characterized by the vanishing of function $b$, (4.4) reduces to $L_{\boldsymbol{v}}\kappa = 0$.

## 4.2 Microlocal normal forms

In this section we present two microlocal (i.e. local in the cotangent bundle over the manifold) normal forms for control systems of type (4.1) under the regularity assumption (4.3). Since the feedback-invariants of such a system are functions on a three-dimensional bundle over the manifold $M$, the microlocalization of the problem is clearly reasonable. Actually, under the considered genericity assumption we may not expect better normal forms. These two normal forms will enable us to get a nice expression for the curvature in restriction to the extremal along which the normalization is done.

### 4.2.1 Normal case

Let $\pi : T^*M \to M$ be the canonical projection. Let

$$h(\lambda) = \max_{u\in U}\langle\lambda,\boldsymbol{f}(q,u)\rangle\,, \quad \lambda \in T^*_qM,$$

be the normal Hamiltonian function of PMP. Recall that if $u_0 \in U$ is a maximized control then, it follows from the regularity assumption (4.3) and the implicit function theorem that we can reconstruct $u_0$ as a smooth function

$$u_0 = u_0(\lambda).$$

Denote by $\vec{\boldsymbol{h}}$ the Hamiltonian vector field corresponding to $h$ and by $\boldsymbol{v}$ the canonical vertical vector field on $T^*M$ that satisfies relation (2.15). Fix a point $q_0$ on the base manifold $M$ and an optimal control $u_0 \in U$. Let $\lambda_0 \in T^*_{q_0}M$ be the covector of PMP associated to the control $u_0$.

Define the following curve on $M$:

$$\begin{aligned} \sigma : \mathbb{R} \to M \\ t \mapsto \pi \circ e^{\tau\,[\boldsymbol{v}, \vec{\boldsymbol{h}}]}(\lambda_0). \end{aligned}$$

The image $N_0$ of $\sigma$ is canonically defined on the manifold $M$ and is transverse to the projection of the integral curves of $\vec{\boldsymbol{h}}$. Indeed, $N_0$ is just the projection onto $M$ of an integral curve of the dynamical system $\dot{\lambda} = [\boldsymbol{v}, \vec{\boldsymbol{h}}]$ on $T^*M$. We use the curve $N_0$ in order to define the horizontal lines of our system of local coordinate on $M$. We then define the vertical lines of our system of local coordinate on $M$ as the time-optimal paths that connect the points belonging to $M$ to $N_0$. In other words, vertical lines are the projection onto $M$ of the extremals of the following time-optimal control problem:

$$\begin{aligned} &\dot{q} = \boldsymbol{f}(q, u), \quad \in M, \quad u \in U, \\ &q(0) \in N_0, \quad q(t_1) = q_1, \\ &t_1 \to \min . \end{aligned}$$

It follows from the PMP with general boundary conditions that the covector $\lambda(0) \in T^*_{q(0)}M$ has to satisfy

$$\lambda(0) \perp T_{q(0)}N_0, \tag{4.6}$$

or equivalently,

$$\left\langle p(0), \frac{\partial \boldsymbol{f}}{\partial u}(q(0), u) \right\rangle = 0.$$

Let $\phi : \mathbb{R}^2 \to M$ be the following mapping:

$$\phi(q_1, q_2) = \pi \circ e^{q_2 \vec{\boldsymbol{h}}}\Big(\xi, \pi \circ e^{q_1\,[\boldsymbol{v}, \vec{\boldsymbol{h}}]}(\lambda_0)\Big), \tag{4.7}$$

where the covector $\xi$ satisfies the boundary condition (4.6), i.e.,

$$\left\langle \xi, \pi_* \Big[\boldsymbol{v}, \vec{\boldsymbol{h}}\Big](\lambda_0) \right\rangle = 0.$$

From the regularity assumption (4.2) it follows that the differential of the map $\phi$ at $(0, 0)$ is bijective. Hence, $\phi^{-1}$ defines canonical microlocal coordinates in a neighborhood of $q_0$ and we can reconstruct the form of the system in these coordinates. Denote by $f_1$, $f_2$ the components of vector field $\boldsymbol{f}$ in the coordinate system defined by the mapping $\phi$. Thus, in the local coordinates defined by $\phi$ the control system (4.1) reads:

$$\begin{aligned} \dot{q}_1 &= f_1(q,u) \\ \dot{q}_2 &= f_2(q,u), \end{aligned}$$

where, by (4.7), $f_1$, $f_2$ have to satisfy:

$$f_1(q,u_0) = 0 \tag{4.8}$$

$$\frac{\partial f_1}{\partial u}(0,u_0) = 1 \tag{4.9}$$

$$f_2(q,u_0) = 1 \tag{4.10}$$

$$\frac{\partial f_2}{\partial u}(0,u_0) = 0. \tag{4.11}$$

Since $\frac{\partial f_1}{\partial u}(q_0,u_0) \neq 0$, the feedback transformation

$$(q,u) \mapsto \tilde{u} = f_1(q,u)$$

is well-defined in a neighborhood of $(q_0,u_0)$ and it brings the system to

$$\begin{aligned} \dot{q}_1 &= \tilde{u} \\ \dot{q}_2 &= \tilde{f}_2\left(q,\tilde{u}\right), \end{aligned} \tag{4.12}$$

where, by (4.10) $\tilde{f}_2$ satisfies

$$\tilde{f}_2(q,0) = 1. \tag{4.13}$$

Equation (4.13) shows that the function $\tilde{f}_2$ can be written in the form

$$\tilde{f}_2(q_1,q_2,u) = 1 - \psi(q_1,q_2,u)u. \tag{4.14}$$

Let $(p_1,p_2,q_1,q_2)$ be a system of local coordinates on $T^*\mathbb{R}^2$. Taking into account (4.14), the control dependent Hamiltonian function for the control system (4.12) reads:

$$h_u(p,q) = p_1 u + p_2(1 - \psi(q,u)u).$$

We now prove that the function $\psi(q,u)$ satisfies $\psi(q,0) = 0$. By construction, for any fixed $q_{01}$ the vertical line $\ell_{q_{01}} = \{(q_{01},t) \mid t \in [0,t_1]\} \subset \mathbb{R}^2$ of our system of local coordinates $(q_1,q_2)$ is the solution corresponding to the optimal control $u = 0$ of the time-optimal control problem

$$\begin{aligned} &\dot{q}_1 = u \\ &\dot{q}_2 = 1 - \psi(q_1,q_2,u)u, \\ &q(0) = (q_{01},0) \in \phi^{-1}(N_0), \quad q(t_1) = (q_{11},q_{12}) \text{ fixed}, \\ &t_1 \to \min. \end{aligned}$$

By consequence, the solution of the time-optimal control problem

$$\begin{aligned}
&\dot{q}_1 = u\\
&\dot{q}_2 = 1 - \psi(q_1, q_2, u)u,\\
&q(0) \in \phi^{-1}(N_0), \quad q(t_1) \in N_1 = \{(s, q_{12}) \mid s \in \mathbb{R}\},\\
&t_1 \to \min,
\end{aligned}$$

is the set of vertical lines $\{\ell_{q_{01}} \mid q_{01} \in \mathbb{R}\}$. Applying the PMP with general boundary conditions to the above time-optimal problem implies that, for the optimal control $u = 0$, the covector of PMP for the time-optimal problem associated to the system (4.12) is solution to:

$$\begin{aligned}
\dot{p_1} &= -\frac{\partial h_{u=0}}{\partial q_1} = 0\\
\dot{p_2} &= -\frac{\partial h_{u=0}}{\partial q_2} = 0,
\end{aligned} \tag{4.15}$$

$$(p_1(0), p_2(0)) \perp T_{(q_{01},0)}\phi^{-1}(N_0), \quad (p_1(t_1), p_2(t_1)) \perp T_{(q_{11},q_{12})}N_1. \tag{4.16}$$

By (4.7), we have

$$T_{(q_{10},0)}\phi^{-1}(N_0) = \{(v, 0),\ v \in \mathbb{R}\},$$

and by definition of $N_1$

$$T_{(q_{11},q_{12})}N_1 = \{(v, 0),\ v \in \mathbb{R}\}.$$

Thus, the transversality conditions (4.16) respectively read

$$0 = p_1(0), \quad 0 = p_1(t_1). \tag{4.17}$$

Taking into account that the covector of PMP never vanishes, (4.17) implies that the covectors $(p_1(0), p_2(0))$ and $(p_1(t_1), p_2(t_1))$ can be chosen to be $(0, 1)$. And since $q_{12} = t_1$ is arbitrary one infer that the covector $p(t)$ corresponding to the optimal control $u = 0$, i.e., the solution to (4.15) with transversality conditions (4.16) is:

$$(p_1(t), p_2(t)) = (0, 1), \quad \forall\, t \in \mathbb{R}. \tag{4.18}$$

Equation (4.18) implies in particular that the maximality condition $\frac{\partial h_u}{\partial u} = 0$ for $u = 0$ reads

$$\psi(q, 0) = 0 \quad \forall q, \tag{4.19}$$

from which it follows immediately that the function $\psi$ can be written in the form

$$\psi(q, u) = a(q, u)u. \tag{4.20}$$

Summing up, we have proved that that, in a neighborhood of $(q_0, u_0)$, the control system (4.1) can be put in the feedback-equivalent microlocal normal form:

$$\begin{aligned}
\dot{q}_1 &= u\\
\dot{q}_2 &= 1 - a(q_1, q_2, u)u^2.
\end{aligned} \tag{4.21}$$

We now prove that the function $a(q,u)$ in the above normal form never vanishes. From the regularity assumptions (4.2) and (4.3), it follows that $\boldsymbol{f}\circ\phi$ has to satisfy

$$\frac{\partial^2 \boldsymbol{f}\circ\phi}{\partial u^2} = -\alpha \boldsymbol{f}\circ\phi - \beta\frac{\partial \boldsymbol{f}\circ\phi}{\partial u}, \tag{4.22}$$

where $\alpha = \alpha(q,u)$ is a non vanishing function. Equation (4.22) implies in particular that

$$\det\left(\frac{\partial^2 \boldsymbol{f}\circ\phi}{\partial u^2}, \frac{\partial \boldsymbol{f}\circ\phi}{\partial u}\right)\Bigg|_{u=0} = -\alpha \det\left(\boldsymbol{f}\circ\phi, \frac{\partial \boldsymbol{f}\circ\phi}{\partial u}\right)\Bigg|_{u=0}. \tag{4.23}$$

In terms of the function $a(q,u)$, the above equation reads

$$2a(q,0) = \alpha(q,0), \tag{4.24}$$

which proves that the function $a(q,u)$ never vanishes at least in a small enough neighborhood of zero.

We now look for some boundary conditions on the function $a(q,u)$. To do so we first write the normal Hamiltonian function of PMP for the control system (4.21) in projective coordinates as explained in Section 2.4 of Chapter 2. The control dependent Hamiltonian for this system reads

$$h_u(p,q) = p_1 u + p_2(1 - a(q,u)u^2).$$

And for a maximizing control $u$ we have:

$$0 = \frac{\partial h_u}{\partial u} = p_1 - p_2\left(2a(q,u)u - \frac{\partial a}{\partial u}(q,u)u^2\right), \tag{4.25}$$

from which it follows that

$$\xi = \frac{p_1}{p_2} = u\left(2a(q,u) + \frac{\partial a}{\partial u}(q,u)\right) = u(2a(q,0) + O(u)).$$

from the previous expansion we deduce the following expansion for $u(\xi,q)$:

$$u = \frac{\xi}{2a(q,0)} + O(\xi^2). \tag{4.26}$$

Substituting (4.26) in 4.25 leads to the following expansion for the projectivised Hamiltonian $h(\xi,q) = h_u(\frac{p_1}{p_2},1,q)$:

$$h(\xi,q) = 1 + \frac{\xi^2}{4a(q,0)} + \xi^3 g(\xi,q),$$

where $g(\xi, q)$ is a smooth function of $(\xi, q)$. According to Section 2.4 of Chapter 2, we easily deduce that the equations of vector field $[\boldsymbol{v}, \vec{\boldsymbol{h}}]$ are given by

$$\dot{q}_1 = \quad \varrho \left( \frac{1}{2a(q,0)} + 6\xi g(\xi, q) + 6\xi^2 \frac{\partial g}{\partial \xi}(\xi, q) + \xi^3 \frac{\partial^2 g}{\partial \xi^2}(\xi, q) \right), \tag{4.27}$$

$$\dot{q}_2 = -\xi \varrho \left( \frac{1}{2a(q,0)} + 6\xi g(\xi, q) + 6\xi^2 \frac{\partial g}{\partial \xi}(\xi, q) + \xi^3 \frac{\partial^2 g}{\partial \xi^2}(\xi, q) \right), \tag{4.28}$$

$$\dot{\xi} = -\varrho \left( \frac{1}{2a(q,0)} \frac{\partial a}{\partial q_2}(q, 0) + O(\xi) \right), \tag{4.29}$$

where $\varrho$ is the non-vanishing function that satisfies $\boldsymbol{v} = \varrho \frac{\partial}{\partial \xi}$. By (4.7), it follows that on restriction to the line $\{q_2 = 0\}$ equations (4.27) and (4.28) read

$$\varrho \left( \frac{1}{2a(q_1,0,0)} + 6\xi g(\xi, q_1, 0) + 6\xi^2 \frac{\partial g}{\partial \xi}(\xi, q_1, 0) + \xi^3 \frac{\partial^2 g}{\partial \xi^2}(\xi, q_1, 0) \right) = 1,$$
$$-\xi \varrho \left( \frac{1}{2a(q_1,0,0)} + 6\xi g(\xi, q_1, 0) + 6\xi^2 \frac{\partial g}{\partial \xi}(\xi, q_1, 0) + \xi^3 \frac{\partial^2 g}{\partial \xi^2}(\xi, q_1, 0) \right) = 0,$$

which implies that

$$\xi|_{\{q_2=0\}} = 0.$$

Thus, on restriction to the line $\{q_2 = 0\}$ equation (4.29) reads

$$-\varrho \left( \frac{1}{2a(q_1,0,0)} \frac{\partial a}{\partial q_2}(q_1, 0, 0) \right) = 0,$$

which is equivalent to

$$\frac{\partial a}{\partial q_2}(q_1, 0, 0) = 0.$$

Summing up, we have proved the following

**Theorem 4.2.1.** *Under the regularity assumptions (4.2) and (4.3) control system (4.1) can be put into the microlocal normal form*

$$\dot{q}_1 = u$$
$$\dot{q}_2 = 1 \pm a(q_1, q_2, u) u^2,$$

*where the function $a$ is a strictly positive function meeting the boundary condition*

$$\frac{\partial a}{\partial q_2}(q_1, 0, 0) = 0, \tag{4.30}$$

*and the $\pm$ sign depends on whether the curves of admissible velocities of system (4.1) are strongly convex or strongly concave.*

The curvature of the control system in the normal form (4.21) is also easily computed using the formula (2.21) of Theorem 2.2.3. This leads to

$$\kappa(q,u) = -\frac{\partial^2}{\partial q_2^2}\frac{\log a(q,0)}{2} - \left(\frac{\partial}{\partial q_2}\frac{\log a(q,0)}{2}\right)^2 + O(u).$$

**Example 4.2.2.** Consider the control system

$$\begin{aligned}\dot{q}_1 &= u\\ \dot{q}_2 &= 1 - a(q)u^2, \quad u \in \mathbb{R}.\end{aligned}$$

This system is just the particular case of the normal form (4.21) when the function $a$ only depends on the base point $q \in M$. The curvature of this system takes the nice polynomial expression

$$\kappa(q,u) = -\frac{\partial^2}{\partial q_2^2}\frac{\log a}{2} - \left(\frac{\partial}{\partial q_2}\frac{\log a}{2}\right)^2 - 3a\left(\frac{\partial}{\partial q_2}\frac{\log a}{2}\right)u^2 - a\frac{\partial^2}{\partial q_2^2}\frac{\log a}{2}u^3. \tag{4.31}$$

It turns out that, if we ask the curvature to be constant then, this system is feedback-equivalent to the normal form

$$\begin{aligned}\dot{q}_1 &= u\\ \dot{q}_2 &= 1 - e^{2q_2\sqrt{-\kappa}}u^2, \quad u \in \mathbb{R}, \quad \kappa \leqslant 0.\end{aligned}$$

One easily get the above normal form from the resolution of the equations asking for the vanishing of the coefficients of the polynomial (4.31) and the use of boundary condition (4.30).

### 4.2.2 Abnormal case

The construction of the microlocal normal form around a regular extremal can easily be adapted in order to get a micro local form around an abnormal extremal, that is, around an extremal along which the Hamiltonian function of PMP vanishes identically. To make sure that there exists an abnormal trajectory passing through the point $q_0 \in M$, we assume that there exists a control $u_0 \in U$ such that

$$\boldsymbol{f}(q_0,u_0) \wedge \frac{\partial \boldsymbol{f}}{\partial u}(q_0,u_0) = 0. \tag{4.32}$$

We do not repeat the detailed construction but only cite the following

**Theorem 4.2.3.** *Suppose that the regularity assumption (4.3) holds in a neighborhood of $(q_0,u_0) \in M \times U$ and that condition (4.32) is satisfied. Then, control system (4.1) can be put into the microlocal normal form*

$$\begin{aligned}\dot{q}_1 &= u\\ \dot{q}_2 &= a(q_1,q_2,u)(1-u)^2,\end{aligned} \tag{4.33}$$

*where the function $a$ is a strictly positive.*

The curvature of the control system in the normal form (4.33) is also easily computed using the formula (2.21) of Theorem 2.2.3. This leads to

$$\kappa(q,u) = -\frac{\partial^2}{\partial q_1^2}\frac{\log a(q,1)}{2} - \left(\frac{\partial}{\partial q_1}\frac{\log a(q,1)}{2}\right)^2 + O(u-1),$$

which shows in particular that the value $\kappa(q,1)$ is well defined so that the curvature can be smoothly extended along the abnormal trajectory.

## 4.3 Flat systems

In Riemannian geometry it is well-known that if the Gaussian curvature of the surface is nonzero then, one can not rectify simultaneously the geodesics by a change of coordinates. Only Riemannian flat systems, i.e., systems for which the geodesics are "straight lines" have this property. For control systems the situation is quite different, first of all because control systems with zero curvature are not necessarily flat. We present here a new theorem which gives a characterization of flat control systems in terms of the feedback invariants $\kappa$ and $b$. The control systems considered in this section are supposed to satisfy the regularity conditions (4.2), (4.3). We begin with the following definition.

**Definition 4.3.1.** *A control system $\dot{q} = \boldsymbol{f}(q,u)$ is said to be flat if it is feedback equivalent to a control system of the form $\dot{q} = \boldsymbol{f}(u)$.*

### 4.3.1 Two fundamental theorems

It is obvious that a flat system has zero curvature but the contrary is in general not true. For example a Zermelo problem defined on the Euclidean plane $\mathbb{R}^2$ with a nonzero linear drift term is never flat (see Example 4.3.6).

Suppose that a control system satisfies

$$L_{\vec{h}} b = 0. \tag{4.34}$$

The above property implies in particular that the plane curves $\mathcal{H}_q \subset T^*M$ are all of the same centro-affine length. Control systems of this type are very peculiar and have nice geometric properties that will be discussed in the next chapter. However such systems with zero curvature are characterized in the theorem below.

**Theorem 4.3.2.** *There exists a feedback transformation such that:*

$$\left[\boldsymbol{f}(\cdot,u), \frac{\partial \boldsymbol{f}(\cdot,u)}{\partial u}\right] = 0 \tag{4.35}$$

*if and only if the feedback invariants $\kappa$ and $L_{\vec{h}}b$ are identically equal to zero. Moreover if we fix local coordinates $q = (q_1, q_2)$ in $M$, then these systems can be parametrized by a one-parameter family of diffeomorphisms generated by the vector field:*

$$\boldsymbol{X}_u = (a_1(u) \pm q_2)\frac{\partial}{\partial q_1} + (a_2(u, q_2) - q_1)\frac{\partial}{\partial q_2}. \tag{4.36}$$

In the above theorem if $u$ is a control parameter such that the fields $\boldsymbol{f}$ and $\frac{\partial \boldsymbol{f}}{\partial u}$ commute then, vector field $\boldsymbol{X}_u$ is the infinitesimal generator of the diffeomorphism $P_u \in \mathrm{Diff}\,(M, \mathbb{R}^2)$ such that

$$P_{u*}\left(\boldsymbol{f}(\cdot, u), \frac{\partial \boldsymbol{f}(\cdot, u)}{\partial u}\right) = \left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right), \tag{4.37}$$

and the $\pm$ sign in the expression of $\boldsymbol{X}_u$ depends on whether the curve the curves of admissible velocities of system (4.1) are strongly convex or strongly concave.

Notice that commutativity between vector fields $\boldsymbol{f}$ and $\frac{\partial \boldsymbol{f}}{\partial u}$ is not a feedback-invariant property. When the curvature is identically zero the above theorem shows that the PDE (4.34) can be reduced to the nonautononous ODE

$$\frac{dq}{du} = \boldsymbol{X}_u(q).$$

***Proof.*** Suppose that $\kappa$ and $L_{\vec{h}}b$ are identically equal to zero for control system (4.1). Then relation (2.21) of Theorem 2.2.3 reduces to

$$\left[\vec{\boldsymbol{h}}, \left[\boldsymbol{v}, \vec{\boldsymbol{h}}\right]\right] = 0.$$

In particular, the flows $e^{t\vec{\boldsymbol{h}}}$ and $e^{t\,[\boldsymbol{v},\vec{\boldsymbol{h}}]}$ commute. Therefore, the vector fields $\vec{\boldsymbol{h}}$ and $[\boldsymbol{v}, \vec{\boldsymbol{h}}]$ are good candidates in order to define a system of local coordinates. Now fix a parameter $\theta$ in the fiber $\mathcal{H}_q$ such that

$$\boldsymbol{v} = \frac{\partial}{\partial \theta}.$$

As usual, denote by the sign "$\,'\,$" the Lie derivative along vector field $\boldsymbol{v}$. This choice of parameter $\theta$ defines a foliation of the three-dimensional manifold $\mathcal{H}$, the leaves of which are formed by the trajectories of the fields $\vec{\boldsymbol{h}}$ and $\vec{\boldsymbol{h}}'$, i.e.,

$$\mathcal{H} = \bigcup_{\lambda \in \mathcal{H}_q} \mathcal{C}_\lambda, \quad \mathcal{C}_\lambda = \left\{ e^{s\vec{\boldsymbol{h}}'} \circ e^{t\vec{\boldsymbol{h}}}(\lambda) \;\middle|\; (s,t) \in \mathbb{R}^2 \right\}.$$

Recall that this choice of $\theta$ is not feedback invariant. Indeed, the parameter $\theta$ is only fixed up to feedback transformations of the form

$$\theta \mapsto \pm\theta + g(q). \tag{4.38}$$

Now fix this parameter $\theta$ in such a way that its value on the leaf $\mathcal{C}_{\lambda_0}$ is constant. In other words we choose the function $g$ in (4.38) such that

$$\theta|_{\mathcal{C}_{\lambda_0}} = \theta_0. \tag{4.39}$$

Recall that in coordinates $(\theta, q)$ on $\mathcal{H}$ vector fields $\vec{\boldsymbol{h}}$ and $\vec{\boldsymbol{h}}'$ take the form

$$\vec{\boldsymbol{h}} = \boldsymbol{f} - c\frac{\partial}{\partial\theta}, \quad \vec{\boldsymbol{h}}' = \boldsymbol{f}' - c'\frac{\partial}{\partial\theta},$$

which, in addition with (4.39), implies that

$$c|_{\mathcal{C}_{\lambda_0}} = 0, \quad c'|_{\mathcal{C}_{\lambda_0}} = 0.$$

Because $L_{\vec{\boldsymbol{h}}}b = 0$ identically, it follows from Proposition 4.4 that $c$ is solution to the Cauchy problem:

$$c'' + bc' + \epsilon c = 0, \quad c|_{\mathcal{C}_{\lambda_0}} = 0, \quad c'|_{\mathcal{C}_{\lambda_0}} = 0.$$

By unicity of solution of a differential equation, it follows that

$$c = 0,$$

identically on $\mathcal{H}$. Hence, $\vec{\boldsymbol{h}} = \boldsymbol{f}$ and $\vec{\boldsymbol{h}}' = \boldsymbol{f}'$ from which it follows that

$$\left[\boldsymbol{f}(\cdot, u), \frac{\partial \boldsymbol{f}(\cdot, u)}{\partial u}\right] = 0.$$

The first implication is thus proved.

We now prove the converse. Let $u$ be a control parameter such that (4.35) holds. In particular,

$$\left\langle \omega, \left[\boldsymbol{f}(\cdot, u), \frac{\partial \boldsymbol{f}(\cdot, u)}{\partial u}\right] \right\rangle = 0,$$

where, as usual $\omega$ denotes the Liouville one-form in restriction to $\mathcal{H}$. According to (2.23), one infers that

$$\begin{aligned} 0 &= \left\langle \omega, \left[\boldsymbol{f}, \frac{d\theta}{du}\boldsymbol{f}'\right] \right\rangle = \left\langle \omega, \frac{d\theta}{du}\left[\boldsymbol{f}, \boldsymbol{f}'\right] + \left(L_{\boldsymbol{f}}\frac{d\theta}{du}\right)\boldsymbol{f}' \right\rangle \\ &= \left\langle \omega, \frac{d\theta}{du}\left(-c\epsilon\boldsymbol{f} - c_{12}^2\boldsymbol{f}'\right) + \left(L_{\boldsymbol{f}}\frac{d\theta}{du}\right)\boldsymbol{f}' \right\rangle \\ &= -c\frac{d\theta}{du}, \end{aligned}$$

which, according to (2.17) implies that $c = 0$ identically on $\mathcal{H}$. In this case, equations (2.25) and (4.5) obviously imply that $\kappa$ and $L_{\vec{\boldsymbol{h}}}b$ are zero identically. The first part of the theorem is thus proved.

In order to parametrize control systems with zero curvature that satisfy (4.35) we will use the well-known Moser's homotopic method. If a control system is such that (4.35) holds, it follows from Frobenius theorem that the vector fields $\boldsymbol{f}$ and $\boldsymbol{f}'$ can be rectified simultaneously. Thus for every $u \in U$ there exists a diffeomorphism $P_u \in \operatorname{Diff} M$ such that

$$P_{u*}\left(\boldsymbol{f}(\cdot,u), \frac{\partial \boldsymbol{f}(\cdot,u)}{\partial u}\right) = \left(\boldsymbol{f}(\cdot,u_0), \frac{\partial \boldsymbol{f}}{\partial u}(\cdot,u_0)\right). \tag{4.40}$$

In order to get the expression (4.36) we use Moser's homotopic method the key idea of which is to determine the diffeomorphisms $P_u$ by representing them as the flow of a family of vector fields $\boldsymbol{X}_u$ on $M$. We thus suppose that

$$\frac{d}{dt}P_u = P_u \circ \boldsymbol{X}_u, \quad P_{u_0} = \mathrm{Id},$$

or equivalently that

$$P_u = \overrightarrow{\exp} \int_{u_0}^{u} \boldsymbol{X}_v \, dv.$$

The expression of $\boldsymbol{X}_u$ in coordinates will follow from the differentiation with respect to $u$ of (4.40). But, after multiplication of both sides by $P_{u*}^{-1}$, (4.40) is equivalent to

$$\begin{aligned} \boldsymbol{f}(\cdot,u) &= \operatorname{Ad} \overrightarrow{\exp} \int_{u_0}^{u} \boldsymbol{X}_v \, dv \, \boldsymbol{f}(\cdot,u_0), \\ \frac{\partial \boldsymbol{f}(\cdot,u)}{\partial u} &= \operatorname{Ad} \overrightarrow{\exp} \int_{u_0}^{u} \boldsymbol{X}_v \, dv \, \frac{\partial \boldsymbol{f}}{\partial u}(\cdot,u_0), \end{aligned}$$

which, after differentiation with respect to $u$ gives

$$\begin{aligned} \frac{\partial \boldsymbol{f}}{\partial u}(\cdot,u) &= \operatorname{Ad} \overrightarrow{\exp} \int_{u_0}^{u} \boldsymbol{X}_v \, dv \operatorname{ad} \boldsymbol{X}_u \left(\boldsymbol{f}(\cdot,u_0)\right) = P_{u*}^{-1}[\boldsymbol{X}_u, \boldsymbol{f}(\cdot,u_0)], \\ \frac{\partial^2 \boldsymbol{f}}{\partial u^2}(\cdot,u) &= \operatorname{Ad} \overrightarrow{\exp} \int_{u_0}^{u} \boldsymbol{X}_v \, dv \operatorname{ad} \boldsymbol{X}_u \left(\frac{\partial \boldsymbol{f}}{\partial u}(\cdot,u_0)\right) = P_{u*}^{-1}\left[\boldsymbol{X}_u, \frac{\partial \boldsymbol{f}}{\partial u}(\cdot,u_0)\right], \end{aligned}$$

which, according to (4.40) is equivalent to

$$\frac{\partial \boldsymbol{f}}{\partial u}(\cdot,u_0) = [\boldsymbol{X}_u, \boldsymbol{f}(\cdot,u_0)], \tag{4.41}$$

$$P_{u*}\frac{\partial^2 \boldsymbol{f}}{\partial u^2}(\cdot,u) = \left[\boldsymbol{X}_u, \frac{\partial \boldsymbol{f}}{\partial u}(\cdot,u_0)\right]. \tag{4.42}$$

Fix a system of local coordinates $q = (q_1, q_2)$ on the base manifold such that

$$\boldsymbol{f}(\cdot,u_0) = \frac{\partial}{\partial q_1}, \quad \frac{\partial \boldsymbol{f}}{\partial u}(\cdot,u_0) = \frac{\partial}{\partial q_2},$$

and denote

$$\boldsymbol{X}_u = X_1(q,u)\frac{\partial}{\partial q_1} + X_2(q,u)\frac{\partial}{\partial q_2}.$$

In these coordinates, equation (4.41) reads

$$-\frac{\partial X_1}{\partial q_1} = 0, \quad -\frac{\partial X_2}{\partial q_1} = 1,$$

which implies that

$$X_1(q,u) = \alpha_1(q_2,u), \quad X_2(q,u) = \alpha_2(q_2,u) - q_1,$$

where $\alpha_1$, ans $\alpha_2$ are $C^\infty$ functions. Recall that $\boldsymbol{f}(\cdot,u)$ satisfies the second order ODE

$$\frac{\partial^2 \boldsymbol{f}}{\partial u^2}(\cdot,u) = -\epsilon \boldsymbol{f}(\cdot,u) - b(\cdot,u)\frac{\partial \boldsymbol{f}}{\partial u}(\cdot,u).$$

Thus, according to (4.40), equation (4.42) reads

$$\begin{aligned}\left[\boldsymbol{X}_u, \frac{\partial \boldsymbol{f}}{\partial u}(\cdot,u_0)\right] &= P_{u*}\frac{\partial^2 \boldsymbol{f}}{\partial u^2}(\cdot,u) = -\epsilon P_{u*}\boldsymbol{f}(\cdot,u) - P_{u*}\left(b(\cdot,u)\frac{\partial \boldsymbol{f}}{\partial u}(\cdot,u)\right)\\ &= -\epsilon \boldsymbol{f}(\cdot,u_0) - P_u b(\cdot,u) P_{u*}\frac{\partial \boldsymbol{f}}{\partial u}(\cdot,u)\\ &= -\epsilon \boldsymbol{f}(\cdot,u_0) - P_u b(\cdot,u)\frac{\partial \boldsymbol{f}}{\partial u}(\cdot,u_0),\end{aligned}$$

where, according to the identification (1.2),

$$P_u b(q,u) = b(P_u(q),u).$$

So in our system of local coordinates on $M$ this last equation reads

$$-\frac{\partial X_1}{\partial q_2} = -\frac{\partial \alpha_1}{\partial q_2} = -\epsilon, \quad -\frac{\partial X_2}{\partial q_2} = -\frac{\partial \alpha_2}{\partial q_2} = -P_u b,$$

from which it follows that

$$X_1(q_1,q_2) = a_1(u) + \epsilon q_2, \quad X_2(q_1,q_2) = a_2(q_2,u) - q_1,$$

which is the required expression for the field $\boldsymbol{X}_u$ and ends the proof. ∎

The following theorem characterizes flat control systems.

**Theorem 4.3.3.** *A control system of type (2.1) is flat if and only if its feedback invariants $\kappa$, $L_{\vec{\boldsymbol{h}}} b$ and $L_{[\boldsymbol{v},\vec{\boldsymbol{h}}]} b$ vanish identically.*

***Proof.*** Suppose that the system under consideration is flat. By definition this system is feedback equivalent to a system of the form $\dot{q} = \boldsymbol{f}(u)$. For such a system it is obvious that the feedback invariant $b$ depends only on the control parameter $u$ and that the Hamiltonian is horizontal. Therefore, the feedback invariants $\kappa$, $L_{\vec{\boldsymbol{h}}}b$ and $L_{[\boldsymbol{v},\vec{\boldsymbol{h}}]}b$ vanish identically.

We now prove the converse. It follows from Theorem 4.3.2 that the vanishing of $\kappa$ and $L_{\vec{\boldsymbol{h}}}b$ implies that, up to a feedback, the vector fields $\vec{\boldsymbol{h}}$ and $[\boldsymbol{v}, \vec{\boldsymbol{h}}]$ are horizontal. Therefore, the vanishing of $L_{\vec{\boldsymbol{h}}}b$ and $L_{[\boldsymbol{v},\vec{\boldsymbol{h}}]}b$ is equivalent to the vanishing of $L_{\boldsymbol{f}}b$ and $L_{[\boldsymbol{v},\boldsymbol{f}]}b$, from which it immediately follows that the invariant $b$ depends only on the control parameter $u$. In this case, the infinitesimal generator of the one-parameter family of diffeomorphisms defined by (4.37) is

$$\boldsymbol{X}_u = (a_1(u) \pm q_2)\frac{\partial}{\partial q_1} + (a_2(u) - q_1)\frac{\partial}{\partial q_2}.$$

Thus,

$$\left(\overrightarrow{\exp}\int_{u_0}^{u} \boldsymbol{X}_v \, dv\right)(q) = e^{(u-u_0)\left(\begin{smallmatrix} 0 & \pm 1 \\ -1 & 0\end{smallmatrix}\right)}(q) + \int_{u_0}^{u} e^{(u-v)\left(\begin{smallmatrix} 0 & \pm 1 \\ -1 & 0\end{smallmatrix}\right)}\left(\begin{smallmatrix} a_1(v) \\ a_2(v)\end{smallmatrix}\right) dv$$

from which it follows that

$$\boldsymbol{f}(q,u) = \mathrm{Ad}\,\overrightarrow{\exp}\int_{u_0}^{u} \boldsymbol{X}_v \, dv \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \boldsymbol{f}(u).$$

That's ends the proof. ■

### 4.3.2 Examples

**Example 4.3.4. Flat Riemannian manifolds**.
Both Theorems 4.3.2 and 4.3.3 imply the following classical theorem.

**Theorem 4.3.5.** *A two-dimensional Riemannian manifold is flat if and only if its Gaussian curvature vanishes identically.*

Indeed, we know form Section 2.5.1 that, in the case of a Riemannian manifold, the curvature $\kappa$ is the Gaussian curvature of the manifold and that the feedback invariant $b$ vanishes identically. If we denote by $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ a local orthonormal basis for the Riemannian structure on the manifold, we then see that Theorems 4.3.2 and 4.3.3 reduce to

$$\kappa \equiv 0 \quad \Leftrightarrow \quad \text{there exists a feedback such that } \left[\boldsymbol{f}, \frac{\partial \boldsymbol{f}}{\partial u}\right] = [\boldsymbol{e}_1, \boldsymbol{e}_2] = 0.$$

But if $[\boldsymbol{e}_1, \boldsymbol{e}_2] = 0$, according to the Frobenius theorem, one can find a system of local coordinates on $M$ such that $\boldsymbol{e}_1 = \frac{\partial}{\partial q_1}$, $\boldsymbol{e}_2 = \frac{\partial}{\partial q_2}$, i.e., such that the dynamic of the Riemannian problem read

$$\dot{q} = \cos u \frac{\partial}{\partial q_1} + \sin u \frac{\partial}{\partial q_2} = \boldsymbol{f}(u).$$

Consequently, the system is flat. Moreover if we fix local local coordinates on the base manifold and set $b = 0$ in the proof of Theorem 4.3.2 we see that flat Riemannian problems are parametrized by the vector field

$$\boldsymbol{X}_u = q_2 \frac{\partial}{\partial q_1} - q_1 \frac{\partial}{\partial q_2}.$$

**Example 4.3.6. Flat Zermelo navigation problem.** Consider the Zermelo navigation problem (3.1) on a Riemannian manifold. Theorem 4.3.2 implies that if the problem is flat then, there exists a feedback transformation such that the coordinate in $\frac{\partial}{\partial\theta}$ vanishes identically. But, for this problem we know from equation (3.6) that the vertical part of the corresponding Hamiltonian is

$$\begin{aligned} \dot{\theta} &= -\left(L_{\boldsymbol{V}'}\varrho + \langle \boldsymbol{V}, [\boldsymbol{e}_1, \boldsymbol{e}_2]\rangle\, \varrho\right) \qquad (4.43)\\ &= -\cos^2\theta L_{\boldsymbol{e}_2} X_1 + \cos\theta\sin\theta\,(L_{\boldsymbol{e}_1} X_1 - L_{\boldsymbol{e}_2} X_2) + \sin^2\theta L_{\boldsymbol{e}_1} X_2\\ &\quad - (1 + \cos\theta X_1 + \sin\theta X_2)(c_1\cos\theta + c_2\sin\theta), \qquad (4.44) \end{aligned}$$

where the notations are those of Section 3.1. Notice that equation (4.43) is intrinsic, i.e., that it does not depend on any local coordinate system on $M$. Thus the existence of a feedback transformation such that $\dot{\theta} = 0$ reduces to the vanishing of equation (4.44) which implies in particular that

$$c_1 = \frac{c(0,q) - c(\pi,q)}{2} = 0, \quad c_2 = \frac{c(\pi/2,q) - c(-\pi/2,q)}{2} = 0,$$

which shows that the Riemannian manifold must by flat. If we choose local coordinates on $M$ such that $[\boldsymbol{e}_1, \boldsymbol{e}_2] = 0$ then, the vanishing of (4.44) implies in particular that

$$\begin{aligned} L_{\boldsymbol{e}_1} X_1 &= \frac{c(\pi/4,q) + c(-\pi/4,q)}{2} = 0, \quad L_{\boldsymbol{e}_2} X_1 = c(0,q) = 0,\\ L_{\boldsymbol{e}_1} X_2 &= \frac{c(\pi/4,q) - c(-\pi/4,q)}{2} = 0, \quad L_{\boldsymbol{e}_2} X_2 = c(\pi/2,q) = 0, \end{aligned}$$

which trivially implies that the coordinates $X_1$, $X_2$ of the drift in $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ have to be constant. Summing up, we have proved the following

**Theorem 4.3.7.** *A Zermelo navigation problem on a Riemannian manifold is flat if and only if the Riemannian manifold is flat and the drift vector field is constant in any system of local coordinates such that* $\mathbf{e}_1$*,* $\mathbf{e}_2$ *commute.*

**Example 4.3.8. Normal forms for flat problems with $b$ constant.** In [42] G. R. Wilkens shows that if the feedback invariant $b$ is a real constant different from zero then, the right-hand side of the control system is feedback-equivalent to one of four normal forms. The particular form depends on whether the curves of admissible velocities are strongly convex, strongly concave and $b^2 - 4$ is positive, zero or negative. He fund that $\boldsymbol{f}(u)$ is respectively equivalent to either

- $\boldsymbol{f}(u) = \cosh\left(\sqrt{\frac{b^2+4}{2}}u\right)e^{\frac{b}{2}u}\dfrac{\partial}{\partial q_1} + \sinh\left(\sqrt{\frac{b^2+4}{2}}u\right)e^{\frac{b}{2}u}\dfrac{\partial}{\partial q_2}$,
- $\boldsymbol{f}(u) = \cosh\left(\sqrt{\frac{b^2-4}{2}}u\right)e^{\frac{b}{2}u}\dfrac{\partial}{\partial q_1} + \sinh\left(\sqrt{\frac{b^2-4}{2}}u\right)e^{\frac{b}{2}u}\dfrac{\partial}{\partial q_2}$,
- $\boldsymbol{f}(u) = e^{\pm u}\dfrac{\partial}{\partial q_1} + u\, e^{\pm u}\dfrac{\partial}{\partial q_2}$, or
- $\boldsymbol{f}(u) = \cos\left(\sqrt{\frac{b^2-4}{2}}u\right)e^{\frac{b}{2}u}\dfrac{\partial}{\partial q_1} + \sin\left(\sqrt{\frac{b^2-4}{2}}u\right)e^{\frac{b}{2}u}\dfrac{\partial}{\partial q_2}$.

Obviously, the case $b = 0$ corresponds to either the Euclidean flat structure, or the Lorentzian flat structure depending on whether the curves of admissible velocities are strongly convex or strongly concave.

# Chapter 5

# Some global results

This chapter is devoted to the study of some global properties of the time-optimal problem (2.1)–(2.3) on a compact base manifold $M$. In particular, we will see in which case the Gauss-Bonnet theorem can be generalized to such problem and we will see how the Hopf theorem stated in the Introduction generalizes for control systems.

## 5.1 A Gauss-Bonnet theorem for special control problems

In this section we give a generalization of the Gauss-Bonnet theorem for special two-dimensional control systems. In classical two-dimensional Riemannian geometry, the Gauss-Bonnet theorem is a local-global theorem *par excellence*, since it asserts the equality of two very different invariants: the integral of the Gaussian curvature which is determined by the local differential geometry of the Riemannian manifold and the Euler characteristic of the manifold which is a global topological invariant. In order to do such a generalization, we suppose that the two-dimensional manifold on which is defined the control system (2.1) is, not only connected but also compact and orientable. Obviously, we do keep the regularity assumption (2.6) of strong convexity for the curves of admissible velocities and we do assume moreover that these curves are simple closed curves surrounding the origin. In this case, we may view these curves as the unit vectors, or indicatrix, for a Finsler structure on the manifold.

The generalization of the Gauss-Bonnet theorem that we will give corresponds to the Gauss-Bonnet theorem for the special case of Landsberg surfaces in Finsler geometry (see e.g. [13]). We do make the proof of the theorem in order to bring to the fore the obstructions that unable such a generalization for more general control systems.

Let us begin with the statement of the theorem. The notations that are used here are those of Chapter 2.

**Theorem 5.1.1.** *Let $M$ be a compact, orientable, connected two-dimensional smooth manifold on which is defined a control system of type (2.1). Assume that this control system is such that:*

- *the curves of admissible velocities are strongly convex simple closed curves surrounding the origin,*
- *the feedback-invariant $b$ is a first integral of the Hamiltonian field (of PMP) $\vec{\boldsymbol{h}}$.*

*Then,*

$$\int_M \kappa\, \varsigma^*\omega \wedge \varsigma^* L_{\boldsymbol{v}}\omega = \ell\chi(M),$$

*where $\ell = \int_{H_q} d\theta$ and $\varsigma$ is a smooth section of the fibration $\mathcal{H} \to M$.*

Before going to the proof of this theorem, we first need to fix some preliminary material. Let $\mathcal{F}^* \subset \Lambda^1(\mathcal{H})$ be the coframe dual to the feedback invariant moving frame $\mathcal{F} \subset \operatorname{Vec} M$ defined by (2.16). We have

$$\mathcal{F}^* = (\omega, \mu, L_{\boldsymbol{v}}\omega)\,,$$

where the one-form $\mu$ takes the form

$$\mu = d\theta + c\,\omega_\theta + c'\,\omega'_\theta \tag{5.1}$$

in any system of local canonical coordinates $\lambda = (\theta, q)$ on $\mathcal{H}$, i.e., such that $\frac{\partial}{\partial\theta} = \boldsymbol{v}$. Indeed, denote $\mu = \mu_\theta\, d\theta + \mu_1\,\omega_\theta + \mu_2\,\omega'_\theta$. The duality of $\mathcal{F}^*$ and $\mathcal{F}$ implies that

$$\left\langle \mu, \frac{\partial}{\partial\theta} \right\rangle = \mu_\theta = 1, \quad \left\langle \mu, \vec{\boldsymbol{h}} \right\rangle = \mu_1 - c = 0, \quad \left\langle \mu, \vec{\boldsymbol{h}}' \right\rangle = \mu_2 - c' = 0,$$

which proves equality (5.1). Notice moreover that

$$\mu|_{H_q} = d\theta,$$

which shows that the one-form $\mu$ is a well-defined, i.e., closed, one-form on restriction to fibers. Moreover, a straightforward computation shows that the structure equation for $\mu$ reads

$$d\mu = -\kappa\,\omega \wedge L_{\boldsymbol{v}}\omega + L_{\vec{\boldsymbol{h}}}b\,\mu \wedge L_{\boldsymbol{v}}\omega. \tag{5.2}$$

In particular, (5.2) shows that the exterior derivative of $\mu$ is induced by a two-form $\eta$ on $M$ if and only if

$$L_{\vec{\boldsymbol{h}}}b = 0,$$

that is, if and only if $b$ is a first integral of the Hamiltonian field $\vec{\boldsymbol{h}}$.

***Proof of Theorem 5.1.1.*** Let $\varsigma : M \to \mathcal{H}$ be a smooth section of the fibration $\pi : \mathcal{H} \xrightarrow{H_q} M$ with isolated singularities. Since $M$ is compact, the number of singularities of such a section must be finite, possibly none. Denote by $q_1, \dots, q_k$, $k \in \mathbb{N}$ these singularities. It follows from the Stokes formula that

$$\int_{M \setminus \{q_1,\dots,q_k\}} \varsigma^* d\mu = -\sum_{i=k}^{k} \deg(\varsigma; q_i) \int_{H_{q_i}} \mu|_{H_{q_i}}. \tag{5.3}$$

In the very special situation when $L_{\vec{\boldsymbol{h}}} b = 0$, it follows from (5.2) that the two-form $d\mu \in \Lambda^2(\mathcal{H})$ is induced by the two-form $\eta = -\kappa\, \varsigma^*\omega \wedge \varsigma^* L_{\boldsymbol{v}}\omega \in \Lambda^2(M)$ and since $\pi \circ \varsigma = \mathrm{Id}$, we obtain from (5.3)

$$\int_{M \setminus \{q_1,\dots,q_k\}} \kappa\, \eta = -\sum_{i=k}^{k} \deg(\varsigma; q_i) \int_{H_{q_i}} \mu|_{H_{q_i}}. \tag{5.4}$$

The left-hand side of (5.4) does not depend on $q$. Hence, in particular $\int_{H_q} \mu|_{H_{q_i}}$ does not depend on $q$. Denote by $\ell$ the value of this integral. Equality (5.4) then reads

$$\int_M \kappa\, \varsigma^*\omega \wedge \varsigma^* L_{\boldsymbol{v}}\omega = \ell \sum_{i=k}^{k} \deg(\varsigma; q_i).$$

Then, by the Poincaré-Hopf theorem (see e.g. [36]), it follows that

$$\int_M \kappa\, \varsigma^*\omega \wedge \varsigma^* L_{\boldsymbol{v}}\omega = \ell \chi(M),$$

where $\chi(M)$ is the Euler characteristic of $M$. ■

Recall that $\mathcal{H}$ denotes the level $h^{-1}(1)$ where $h$ is the maximized Hamiltonian function of PMP associated to the time-optimal problem (2.1)–(2.3). The Liouville one-form restricted to $\mathcal{H}$ defines a non degenerated volume form, that we shall denote by $d\mathcal{L}$, on $\mathcal{H}$ according to the formula

$$d\mathcal{L} = -\omega \wedge d\omega. \tag{5.5}$$

It is immediate from (2.12) and (5.1) that

$$d\mathcal{L} = \mu \wedge \omega \wedge L_{\boldsymbol{v}}\omega = d\theta \wedge \omega_\theta \wedge \omega'_\theta. \tag{5.6}$$

From (5.5), (5.6) and Theorem 5.1.1, we immediately get the following

**Corollary 5.1.2.** *Assume the hypothesis of Theorem 5.1.1. Then,*

$$\int_{\mathcal{H}} \kappa\, d\mathcal{L} = \ell^2 \chi(M).$$

## 5.2 A Hopf theorem for control systems

Riemannian metrics on tori without conjugate points have to be flat. This theorem was proved by E. Hopf in 1948 for the two-dimensional case (see [26]) and for higher dimensional manifolds it was proved by D. Burago and S. Ivanov in 1994 (see [18]).

The aim of this section is to give a generalization of the Hopf's theorem given in the introduction for control problems.

### 5.2.1 Jacobi curves

We introduce here the *Jacobi curves* which are a generalization of the space of Jacobi fields along Riemannian geodesics. since the construction of Jacobi curves does not depend on the dimension of the manifold, we begin with the general case to then go to our special low-dimensional case.

Let $h$ be the Hamiltonian function of PMP for a time-optimal smooth control problem and $\mathcal{H}$ its hypersurface $h^{-1}(1)$. Let $e^{t\vec{\boldsymbol{h}}} : \mathcal{H} \to \mathcal{H}$ denotes the flow generated by the Hamiltonian field of PMP $\vec{\boldsymbol{h}}$. This flow defines a one-dimensional foliation $\mathcal{F}$ of $\mathcal{H}$ whose leaves, the trajectories of $\vec{\boldsymbol{h}}$, are transverse to the fibers $T_q^*M$, $q \in M$. This foliation enable us to make the following symplectic reduction.

Consider the canonical projection

$$\varphi : \mathcal{H} \to \Sigma = \mathcal{H}/\mathcal{F}.$$

The quotient space $\Sigma$, space of trajectories of $\vec{\boldsymbol{h}}$, is, at least locally, a well-defined smooth manifold and carries a structure of symplectic manifold with symplectic form $\bar{\sigma}$ characterized by the property that its pull-back to $\mathcal{H}$ is the restriction $\sigma|_{\mathcal{H}}$.

Let $\Pi \subset T\mathcal{H}$ denote the vertical distribution, i.e., $\Pi_\lambda = T_\lambda \mathcal{H}_\lambda$, $\lambda \in \mathcal{H}$.

**Definition 5.2.1.** *We call Jacobi curve at* $\lambda$ *the curve*

$$\begin{aligned} J_\lambda : \mathbb{R} &\to T_{\varphi(\lambda)}\Sigma \\ t &\mapsto J_\lambda(t) = \varphi_* \circ e_*^{-t\vec{\boldsymbol{h}}} \Pi_{e^{t\vec{\boldsymbol{h}}}(\lambda)}. \end{aligned}$$

Because the Hamiltonian flow preserves the symplectic structure, it is easy to check that the spaces $J_\lambda(t)$, $t \in \mathbb{R}$, are Lagrangian subspaces of the symplectic space $T_{\varphi(\lambda)}\Sigma$ so that the Jacobi curves are curves in the Lagrangian Grassmannian $L(T_{\varphi(\lambda)}\Sigma)$.

Recall that the Lagrangian Grassmannian $L(T_{\varphi(\lambda)}\Sigma)$ of the symplectic space $T_{\varphi(\lambda)}\Sigma$ is defined by:

$$L(T_{\varphi(\lambda)}\Sigma) = \{\Lambda \subset T_{\varphi(\lambda)}\Sigma \mid \Lambda^\angle = \Lambda\},$$

where

$$\Lambda^\angle = \{\xi \in T_{\varphi(\lambda)}\Sigma \mid \bar{\sigma}(\xi, \Lambda) = 0\}.$$

The Lagrangian Grassmannian of a symplectic space is a well-defined smooth and compact manifold. In our particular case of a two-dimensional manifold $M$, the Lagrangian Grassmannian $L(T_{\varphi(\lambda)}\Sigma)$ is diffeomorphic to the one-dimensional real projective space $\mathbb{RP}(1)$. Moreover, since the vertical distribution $\Pi$ is generated by the vertical vector field $\boldsymbol{v}$ the Jacobi curve can written as

$$J_\lambda(t) = \mathbb{R}\left(\varphi_* e^{t\,\mathrm{ad}\,\vec{\boldsymbol{h}}}\boldsymbol{v}(\lambda)\right). \tag{5.7}$$

Thus, one can easily characterize the conjugate points in terms of Jacobi curves. It immediately follows from Proposition 2.6.3 that time $t$ is conjugate to time $\tau \neq t$ if and only if

$$J_\lambda(t) \cap J_\lambda(\tau) \neq \{0\}.$$

### 5.2.2 The Hopf theorem

In this section we prove the following

**Theorem 5.2.2.** *Consider a control system $\dot{q} = \boldsymbol{f}(q,u)$ on a compact surface $M$ without boundary. Assume that the curves of admissible velocities are strongly convex curves surrounding the origin. Then, if there is no conjugate points on $M$ the total curvature $\int_{\mathcal{H}} \kappa\, d\mathcal{L}$ must be negative or zero. In the latter case $\kappa$ must be zero.*

***Proof.*** Notice that because the curves of admissible velocities are strongly convex curves surrounding the origin, the manifold $\mathcal{H}$ is compact. Although the proof we make here essentially follows the one given by Hopf in [26], it will however be exposed in a more intrinsic and geometrical manner. The first step in the proof consists in the construction of a well-defined function on any extremal of our system, i.e., a function that does not depend on time but only on the point of the extremal. To do so we use the notion of Jacobi curve described in the previous section.

Let $\lambda$ be a point of the hypersurface $\mathcal{H} \subset T^*M$ and let $J_\lambda(t)$ be the Jacobi curve associated with the extremal $e^{t\vec{\boldsymbol{h}}}(\lambda)$. we have

$$J_\lambda(t) = \mathbb{R}\left(\varphi_* e^{t\,\mathrm{ad}\,\vec{\boldsymbol{h}}}\boldsymbol{v}(\lambda)\right) \in \mathbb{RP}(1),$$

with

$$e^{t\,\mathrm{ad}\,\vec{\boldsymbol{h}}}\boldsymbol{v}(\lambda) = \beta(t,\lambda)\boldsymbol{v}(\lambda) + \gamma(t,\lambda)\left[\boldsymbol{v}, \vec{\boldsymbol{h}}\right](\lambda).$$

Considering $(\beta : \gamma)$ as homogeneous coordinate in $\mathbb{RP}(1)$, we can identify the Jacobi curve with the curve

$$t \mapsto (\beta(t,\lambda) : \gamma(t,\lambda)).$$

From the non existence of conjugate points it follows that $\gamma(t,\lambda) \neq 0$ for $t \neq 0$ we can thus use the chart $(\beta : \gamma) \mapsto \frac{\beta}{\gamma}$ and thus make the identification

$$J_\lambda(t) = u_t(\lambda) = \frac{\beta(t,\lambda)}{\gamma(t,\lambda)}, \quad t \neq 0.$$

It follows from Lemma 1.2.3 that the coefficients $\beta$ and $\gamma$ are solutions of the Cauchy problems

$$\begin{aligned} \ddot{\beta} + \kappa\beta = 0, \quad \beta(0) = 1, \quad \dot{\beta}(0) = 0, \\ \ddot{\gamma} + \kappa\gamma = 0, \quad \gamma(0) = 0, \quad \dot{\gamma}(0) = 1, \end{aligned}$$

which shows in particular that $\beta$ and $\gamma$ are two linearly independent solutions of the Hill equation $\ddot{x} + \kappa x = 0$. The derivative with respect to time of the function $u_t$ is

$$\frac{du_t}{dt} = \frac{\dot{\beta}\gamma - \beta\dot{\gamma}}{\gamma^2}$$

and because the Wronskian

$$\dot{\beta}(0)\gamma(0) - \beta(0)\dot{\gamma}(0) = -1,$$

the function $u_t$ is strictly decreasing or, equivalently the Jacobi curve is strictly decreasing in $\mathbb{RP}(1)$. Since $u_t$ is strictly decreasing its limit as $t$ goes to infinity exists. Moreover, because of the non existence of conjugate points, this limit is finite. Indeed, notice that because of the initial conditions $\beta(0, \lambda) = 1$, $\gamma(0, \lambda) = 0$ and $\dot{\gamma}(0, \lambda) = 1$ we have for $t$ small enough

$$u_t(\lambda) > 0, \quad u_{-t}(\lambda) < 0. \tag{5.8}$$

So if we suppose that

$$\lim_{t\to+\infty} u_t(\lambda) = -\infty, \tag{5.9}$$

it would follow from Equations (5.8) and from the strict monotonicity of $u_t$ the existence of $t^+ > 0$ and $t^- < 0$ such that $u_{t^+}(\lambda) = u_{t^-}(\lambda)$. Then, the time reparametrization $\tau = t - t^-$ would imply that time $\tau = t^+ - t^-$ is conjugate to $\tau = 0$, which is a contradiction. Hence, (5.9) is false and $u_t$ takes real value. We denote

$$u = \lim_{t\to+\infty} u_t,$$

which is a well defined function on the manifold $\mathcal{H}$. Equivalently, the distribution $\Pi_\lambda^\infty \in T\mathcal{H}$ defined by

$$\Pi_\lambda^\infty = \lim_{t\to+\infty} J_\lambda(t)$$

is a well defined distribution on $\mathcal{H}$ transverse to the vertical distribution. This distribution $\Pi_\lambda^\infty$ is, by definition, invariant by the flow of $\vec{\boldsymbol{h}}$. In terms of the function $u$, this invariance reads

$$\left[\vec{\boldsymbol{h}}, u\boldsymbol{v} + \left[\boldsymbol{v}, \vec{\boldsymbol{h}}\right]\right] = \alpha\left(u\boldsymbol{v} + \left[\boldsymbol{v}, \vec{\boldsymbol{h}}\right]\right),$$

or, equivalently

$$L_{\vec{\boldsymbol{h}}} u\boldsymbol{v} + u\left[\boldsymbol{v}, \vec{\boldsymbol{h}}\right] + \left[\vec{\boldsymbol{h}}, \left[\boldsymbol{v}, \vec{\boldsymbol{h}}\right]\right] = \alpha u\boldsymbol{v} + \alpha\left[\boldsymbol{v}, \vec{\boldsymbol{h}}\right], \tag{5.10}$$

where $\alpha$ is function on $\mathcal{H}$. Solving (5.10) for $\alpha$ gives

$$\alpha = -u \quad \text{and} \quad L_{\vec{\boldsymbol{h}}} u + \kappa - \alpha u = 0,$$

which shows that $u$ satisfies the Riccati equation

$$L_{\vec{\boldsymbol{h}}} u + u^2 + \kappa = 0. \tag{5.11}$$

If we now integrate equation (5.11) over $\mathcal{H}$ with respect to the Liouville volume $d\mathcal{L}$, the first term in the left-hand side of (5.11) will disappear since the Liouville volume is invariant by the flow of $\vec{\boldsymbol{h}}$. As a result we obtain

$$\int_{\mathcal{H}} \kappa \, d\mathcal{L} = -\int_{\mathcal{H}} u^2 \, d\mathcal{L} \tag{5.12}$$

which immediately proves the validity of the first part of the theorem. If we now suppose that the total curvature $\int_{\mathcal{H}} \kappa \, d\mathcal{L}$ is zero it follows from (5.12) that the function $u$ must vanish everywhere on $\mathcal{H}$. According to (5.11) $\kappa$ must therefore vanish everywhere. ■

### 5.2.3 A natural question

In the proof of Theorem 5.2.2 we constructed a function $u$ well-defined on $\mathcal{H}$ that satisfies Riccati equation (5.11). This construction is valid along every regular extremal without conjugate points. Recall moreover that a control system with negative curvature does not admit conjugate points. A very natural question is thus the following: *Considering a control system without conjugate points, does there exists a global reparametrization of the system such that the reparametrized curvature is negative?* To answer this question let $f$ be a non vanishing function on $\mathcal{H}$ and consider the reparametrization

$$\vec{\boldsymbol{h}} = \frac{\hat{\boldsymbol{h}}}{f^2} \quad \text{and} \quad \boldsymbol{v} = f\hat{\boldsymbol{v}},$$

and we compute the new function $\hat{u}$:

$$\begin{aligned} u\boldsymbol{v} + \left[\boldsymbol{v}, \vec{\boldsymbol{h}}\right] &= uf\hat{\boldsymbol{v}} + \left[f\hat{\boldsymbol{v}} + \frac{1}{f^2}\hat{\boldsymbol{h}}\right] = uf\hat{\boldsymbol{v}} + \frac{1}{f}\left[\hat{\boldsymbol{v}}, \hat{\boldsymbol{h}}\right] - \frac{1}{f^2} L_{\hat{\boldsymbol{h}}} f \hat{\boldsymbol{v}} \quad \left(\text{mod } \vec{\boldsymbol{h}}\right) \\ &= \left(uf - L_{\vec{\boldsymbol{h}}} f\right)\hat{\boldsymbol{v}} + \frac{1}{f}\left[\hat{\boldsymbol{v}}, \hat{\boldsymbol{h}}\right] \quad \left(\text{mod } \vec{\boldsymbol{h}}\right). \end{aligned} \tag{5.13}$$

We thus have

$$\Pi^\infty = \mathbb{R}\Big(\hat{u}\hat{\boldsymbol{v}} + \left[\hat{\boldsymbol{v}}, \hat{\boldsymbol{h}}\right]\Big), \quad \hat{u} = uf^2 - fL_{\vec{\boldsymbol{h}}} f.$$

In the same way as for the function $u$ it is easy to see that the function $\hat{u}$ satisfies the Riccati equation

$$L_{\hat{\boldsymbol{h}}} \hat{u} + \hat{u}^2 + \hat{\kappa} = 0.$$

Now, the question is: *does there exists a non vanishing function $f$, say $f > 0$ for simplicity such that $L_{\hat{h}}\hat{u} = 0$, or equivalently such that $L_{\vec{h}}\hat{u} = 0$?* To answer, let us first compute $L_{\vec{h}}\hat{u}$.

$$L_{\vec{h}}\hat{u} = L_{\vec{h}}(uf^2 - fL_{\vec{h}}f) = f^2L_{\vec{h}}u + 2ufL_{\vec{h}}f - (L_{\vec{h}}f)^2 - fL^2_{\vec{h}}f,$$

so that (dividing by $f^2$) $L_{\vec{h}}\hat{u} = 0$ is equivalent to

$$L_{\vec{h}}u + 2u\left(\frac{L_{\vec{h}}f}{f}\right) - \left(\frac{L_{\vec{h}}f}{f}\right)^2 - \frac{L^2_{\vec{h}}f}{f} = 0,$$

i.e., to

$$L_{\vec{h}}u + 2uL_{\vec{h}}\log f - (L_{\vec{h}}\log f)^2 - \frac{L^2_{\vec{h}}f}{f} = 0. \tag{5.14}$$

Denote $g = \log f$. We have

$$L^2_{\vec{h}}g = L_{\vec{h}}(L_{\vec{h}}\log f) = L_{\vec{h}}\left(\frac{L_{\vec{h}}f}{f}\right) = \frac{(L^2_{\vec{h}}f)f - (L_{\vec{h}}f)^2}{f^2} = \frac{L^2_{\vec{h}}f}{f} - (L_{\vec{h}}g)^2,$$

or equivalently

$$\frac{L^2_{\vec{h}}f}{f} = L^2_{\vec{h}}g + (L_{\vec{h}}g)^2.$$

This implies that equation (5.14) is equivalent to

$$L_{\vec{h}}u + 2uL_{\vec{h}}g - 2(L_{\vec{h}}g)^2 - L^2_{\vec{h}}g = 0,$$

i.e., to the Riccati equation

$$L_{\vec{h}}y + 2y^2 - 2uy - L_{\vec{h}}u = 0, \tag{5.15}$$

where we have set $y = L_{\vec{h}}g$.

The function $y = u$ is solution to Riccati equation (5.15). Thus we will have the required reparametrization of $\vec{h}$ if we can solve the equation

$$L^2_{\vec{h}}\log f = u \tag{5.16}$$

on the three-dimensional manifold $\mathcal{H}$. The first thing we need for the resolution of equation (5.16) is the continuity of the function $u$ on $\mathcal{H}$. In the case of hyperbolic systems (see [31] for the definition), the function $u$ is easily seen to be continuous due to some "exponential estimates" along the stable distribution (see [31]). Also, for such systems $u$ is in general never differentiable and even never Lipschitz continuous but only Hölder continuous (see [31] Theorem 19.1.6 of Chapter 19). In the case of systems without conjugate points the situation is quite different because we do not have the exponential estimates and by consequence the continuity of the function $u$ is not so obvious. Anyway we do believe it and set the following

**Conjecture 5.2.3.** *The function $u$ defined above is continuous.*

Notice that the function $u$ is easily seen to be upper semi-continuous. Indeed, let $(\lambda_n)_{n\in\mathbb{N}} \subset \mathcal{H}$ be a converging sequence to $\lambda \in \mathcal{H}$. Since $u_t(\lambda_n)$ is decreasing in $t$, it follows that

$$u_t(\lambda_n) \geqslant u(\lambda_n) = \lim_{t\to+\infty} u_t(\lambda_n).$$

Taking the lim inf as $n$ tends to $+\infty$ in the previous relation, we get since $u_t(\lambda)$ is continuous in $(t, \lambda)$

$$u_t(\lambda) \geqslant \liminf_{\lambda_n\to\lambda} u,$$

and then, letting $t$ going to $+\infty$ leads to

$$u(\lambda) \geqslant \liminf_{\lambda_n\to\lambda} u,$$

which proves the upper semi-continuity of $u$.

Suppose that Conjecture 5.2.3 is true. It implies that we can solve locally equation (5.16). In order to solve this equation globally, the question is more delicate because the problem is closely related to the fact that the quotient manifold $\Sigma$, defined in Section 5.2.1 of this chapter, is globally defined. We do not want to discuss in details this problem here. However we can say the following. Let $\tilde{M}$ be the universal covering of $M$. because of the non existence of conjugate points, $\tilde{M}$ is diffeomorphic to $\mathbb{R}^2$. Let

$$\dot{\tilde{q}} = \tilde{\boldsymbol{f}}(\tilde{q}, u), \quad \tilde{q} \in \tilde{M}, \quad u \in U, \tag{5.17}$$

be the lift on $\tilde{M}$ of the control system $\dot{q} = \boldsymbol{f}(q, u)$, and $\tilde{\mathcal{H}}$ be the corresponding Hamiltonian hypersurface. Then, Conjecture 5.2.3 implies that when the control system $\dot{q} = \boldsymbol{f}(q, u)$ has no conjugate points then, there exists a reparametrization of $\vec{\boldsymbol{h}}$ or, equivalently globally defined function $f$ satisfying equation (5.16), such that the lifted system (5.17) has negative curvature.

# Closing

We conclude this thesis with some remarks on the possible generalizations of the exposed results and some perspective of studies. This thesis was dealing in part, with generic non linear two-dimensional control problems and in part with Zermelo problems and there is still a lot work to do for the understanding of such problems.

Let us begin with some obvious possible generalizations. Chapter 3 was dealing with Zermelo problems on Riemannian manifolds. Of course, Zermelo problems can be defined on any Finsler manifold even on any structure defined by an optimal control system on a manifold.

Although all the results of Chapter 5, excepted Theorem 5.1.1 of course, were proved, for the clarity of the exposure, in the case of Zermelo problems on Riemannian surfaces they do extend straightforwardly to the case of Zermelo problems on Landsberg surfaces, that is to geometrical structure defined by a control system for which the curves of admissible velocities are strictly convex closed curves surrounding the origin, with even no add of difficulties nor in proofs or in computations. This was for the direct generalizations.

We now turn our attention on some perspective concerning control systems having negative curvature. In the appendix to the paper [11] by D. V. Anosov and Ya. G. Sinai, Margulis proved the nonexistence of Anosov flows (see [31] for the definition) on three-dimensional tori. A particular case of Anosov hyperbolic systems is given by geodesic flows on Riemannian manifolds with negative sectional curvature. Hence, according to the result by Margulis such flows do not exist on three-dimensional tori. Even if no Anosov flow can be fund on the torus, there does exist, however, extremal flow on $\mathbb{T}^2$ with negative curvature. We fund the following example.

Consider the following Zermelo's navigation problem

$$\begin{aligned} \dot{q}_1 &= a\cos(q_1+q_2)+\cos\theta \\ \dot{q}_2 &= a\sin(q_1+q_2)+\sin\theta, \quad a\in\mathbb{R}. \end{aligned} \tag{5.18}$$

We will see that for $a \geqslant \sqrt{2}$ this problem has negative control curvature. Actually the curvature of this problem is

$$\begin{aligned} \kappa(q,\theta) = -\frac{a}{16}\Big(& a\big(18+6\cos(2q-2\theta)-5\sin(2q)-3\sin(2q-4\theta)-6\sin(2\theta)\big) \\ & -\big(-24\cos(q-\theta)+4\sin(q-3\theta)+12\sin(q+\theta)\big)\Big), \end{aligned}$$

where $q = q_1 + q_2$. The above formula can also be rewritten in the more symmetric form

$$\begin{aligned}\kappa(\Gamma, T) &= -\frac{a}{16}\left(a\left(18 + 6\cos\Gamma - 5\sin(\Gamma + T) - 3\sin(\Gamma - T) - 6\sin T\right)\right.\\ &\qquad \left.-4\left(-6\cos\left(\tfrac{\Gamma}{2}\right) + \sin\left(\tfrac{\Gamma}{2} - T\right) + 3\sin\left(\tfrac{\Gamma}{2} + T\right)\right)\right)\\ &= -\frac{a}{16}\left(aX(\Gamma, T) - Y(\Gamma, T)\right),\end{aligned}$$

where $\Gamma = 2q - 2\theta$ and $T = 2\theta$. In order to have $\kappa \leqslant 0$ we first of all make sure that $X$ is positive.

**Claim 5.2.4.** *The function $X(\Gamma, T)$ is positive.*

***Proof.***

$$X(\Gamma, T) = \big(18 + 6\cos\Gamma - 5\sin(\Gamma + T) - 3\sin(\Gamma - T) - 6\sin T\big)$$

so at minimum point of $X$ we have

$$\begin{aligned}0 &= \frac{\partial X}{\partial \Gamma} = -6\sin\Gamma - 5\cos(\Gamma + T) - 3\cos(\Gamma - T)\\ 0 &= \frac{\partial X}{\partial T} = \qquad\qquad - 5\cos(\Gamma + T) + 3\cos(\Gamma - T) - 6\cos T.\end{aligned}$$

Subtracting the second of the above equations to the first one we see that at critical point we have

$$6\cos T - 6\sin T = 6\cos(\Gamma - T),$$

so that

$$X = 18 - 5\sin(\Gamma + T) - 3\sin(\Gamma - T) + 6\cos(\Gamma - T) \geqslant 4 > 0,$$

which completes the proof. ■

It is now clear that we can choose the value of the number $a$ in order to have a negative (and non constant) curvature. Indeed it is sufficient to choose $a \geqslant \max\{0, \max_{\mathbb{T}^1} \frac{Y}{X}\}$. More precisely we have

$$\max_{\mathbb{T}^1} \frac{Y}{X} = \sqrt{2}.$$

In the above example, because $a \geqslant \sqrt{2} > 1$, the Zermelo problem admits abnormal extremal and thus it does not define a Finsler structure on the two-torus $\mathbb{T}^2$. Anyway, this problem still has negative curvature and thus, we may expect a kind of hyperbolic behavior for the Hamiltonian flow (maybe a weaker type of hyperbolicity than for Anosov flows).

Theorem 2 of [3] asserts that if a Hamiltonian system has negative curvature at every point of an invariant set for the Hamiltonian flow then, this set is a hyperbolic set. Unfortunately it is not possible to simply apply this theorem to our system, first of all because, due to the existence of abnormal extremals, the level set $\mathcal{H}$ is not compact. However, this lack of compacity is not a big problem. Indeed, we can easily get off from this obstruction doing a compactification of $\mathcal{H}$. The following construction works for any control system $\dot{q} = \boldsymbol{f}(q, u)$ on a smooth manifold $M$. Let $h^\nu$ be the maximized Hamiltonian of PMP defined by

$$h^\nu = \max_{u \in U} \langle \lambda, \boldsymbol{f}(q, u) \rangle - \nu, \quad \lambda \in T_q^* M, \quad \nu \in \mathbb{R}.$$

Instead of working as before on the level set $(h^1)^{-1}(0)$ ($= h^{-1}(1)$ in the notations of Chapter 2), we firstly consider the level $\mathcal{H}^\nu = (h^\nu)^{-1}(0)$ which enable us to unify normal and abnormal extremals. Notice in addition that $\mathcal{H}^\nu$ contains both the minimum time extremals and the maximum time extremal. Define the following equivalence relation on $\mathcal{H}^\nu = (h^\nu)^{-1}(0)$

$$\lambda_1 \sim \lambda_2 \quad \Leftrightarrow \quad \exists\, c \neq 0 \mid \lambda_1 = c\lambda_2,$$

which is well defined since, according to PMP, $\lambda \neq 0$. The quotient space

$$\mathcal{H}^\nu / \sim = \{(\lambda, \nu) \in \mathbb{P}(T^*M) \times \mathbb{R} \mid h^\nu(\lambda) = 0\}$$

has a natural well defined structure of smooth manifold. Moreover, if $M$ is compact then, $\mathcal{H}^\nu / \sim$ is compact. In the case of system (5.18) we have

$$\mathcal{H}^\nu / \sim \cong \mathbb{T}^3,$$

and easy computations show that the Hamiltonian field $\vec{\boldsymbol{h}}^{\boldsymbol{\nu}}$ associated to $h^\nu$ goes to a well defined vector field $\tilde{\boldsymbol{h}}^{\boldsymbol{\nu}}$ on the quotient manifold $\mathbb{T}^3$. But, be careful: the field $\tilde{\boldsymbol{h}}^{\boldsymbol{\nu}}$ is no more a Hamiltonian vector field. This gives a second reason why Theorem 2 of [3] does not apply here. Indeed, the proof of this theorem heavily uses the nondegenaracy of the symplectic structure which completely degenerates along abnormal extremals.

# Bibliography


[1] A. A. Agrachev, *Feedback-invariant optimal control theory and differential geometry-II. Jacobi curves for singular extremals*, Journal of Dynamical and Control Systems **4** (1998), 583-604.

[2] A. A. Agrachev, *A "Gauss-Bonnet formula" for contact sub-Riemannian manifolds*, Russian Math. Dokl. **381** (2001), 583-585.

[3] A. A. Agrachev, N. N. Chtcherbakova, *Hamiltonian systems of negative curvature are hyperbolic*, Russian Math. Dokl. **400** (2005), 295-298.

[4] A. A. Agrachev, R. V. Gamkrelidze, *The exponential representation of flows and the chronological calculus.* (Russian) *Mat. Sb.* **107** (1978), 467-532. (English translation) *Math. USSR. Sb.* **35** (1979), 727-785.

[5] A. A. Agrachev, R. V. Gamkrelidze, *Feedback-Invariant Optimal Control Theory and Differential Geometry-I. Regular Extremals*, Journal of Dynamical and Control Systems **3** (1997), 343-389.

[6] A. A. Agrachev, Yu. L. Sachkov, *Control Theory from the Geometric Viewpoint*, Springer-Verlag, 2004.

[7] A. A. Agrachev, A. V. Sarychev, *On reduction of a smooth system linear in the control*, Math. USSR Sbornik **58** (1987), 15-30.

[8] A. A. Agrachev, I. Zelenko, *Geometry of Jacobi curves I*, Journal of Dynamical and Control Systems **8** (2002), 93-140.

[9] A. A. Agrachev, I. Zelenko, *Geometry of Jacobi curves II*, Journal of Dynamical and Control Systems **8** (2002), 167-215.

[10] A. A. Agrachev, I. Zelenko, *On feedback classification of control-affine systems with one and two-dimensional inputs*, SIAM J. on Control and Optimization, submitted.

[11] D. V. Anosov, Ya. G. Sinai, *Some smooth ergodic systems*, Russian Mathematical Surveys **22** (1967), 103-167.

[12] V. I. Arnold, *Méthodes mathématiques de la mécanique classique*, Mir, 1974 (traduction française).

[13] D. Bao, S.-S. Chern, Z. Shen, *An Introduction to Riemann-Finsler Geometry*, Springer, 2000.

[14] D. Bao, Z. Shen, *On the volume of unit tangent spheres in a Finsler manifold*, Results in Math. **26** (1994), 1-17.

[15] R. Berndt, *An Introduction to Symplectic Geometry*, American Mathematical Society **26**, 2000.

[16] V. G. Boltyanskij, R. V. Gamkrelidze, L. S. Pontryagin, it The theory of optimal processes, I, The maximum principle, Izv. Akad. nauk SSSR, seriya matem. **24** (1960), No. 1, 3-42.

[17] P. Brunovsky, *A classification of linear controllable systems*, Kybernetika (Prague) **6** (1970), 173-188.

[18] D. Burago, S. Ivanov, *Riemannian tori without conjugate points are flat*, Geometric and Functional Analysis **4** (1994), no. 3, 259-269.

[19] C. Carathéodory, *Calculus of Variations*, Chelsea Publishing Company, New York, 1989 (third edition), **§276-§460**.

[20] H. Cartan, *Cours de Calcul Différentiel*, Hermann, Paris, 1967.

[21] D. McDuff, D. Salamon, *Introduction to symplectic topology*, Oxford University Press, 1995.

[22] F. R. Gantmacher, *The Theory of Matrices*, Chelsea Publishing Company, New York, 1959.

[23] R. B. Gardner, W. F. Shadwick, *Feedback equivalence for general control systems*, Systems Control Lett. **15** (1990), no. 1, 15-23.

[24] R. B. Gardner, W. F. Shadwick, G. R. Wilkens, *Feedback equivalence and symmetries of Brunowski normal forms*, Dynamics and control of multibody systems (Brunswick, ME, 1988), Contemp. Math. **97**, Amer. Math. Soc., Providence, RI, 1989 , 115-130.

[25] R. B. Gardner, G. R. Wilkens, *The fundamental theorems of curves and hypersurfaces in centro-affine geometry*, Bull. Belgian Math. Society **4** (1997), no. 3, 379-401.

[26] E. Hopf, *Closed surfaces without conjugate points*, Proc. Nat. Acad. Sci. U. S. A. **34** (1948), 47-51.

[27] B. Jakubczyk, *Feedback Equivalence of Nonlinear Control Systems*, Nonlinear Controllability and Optimal Control, H. J. Sussman Ed., Marcel Dekker, 1990.

[28] B. Jakubczyk, *Critical Hamiltonians and feedback invariants*, Geometry of feedback and optimal control, Inc. Pure Appl. Math., Marcel Dekker **207**, New York, 219-256, 1998.

[29] B. Jakubczyk, W. Respondek, *Feedback classification of analytic control systems in the plane*, Analysis of Controlled Dynamical Systems, B. Bonnard et al., eds., Birkhäuser, Boston, 262-273, 1991.

[30] V. Jurdjevic, *Geometric control theory*, Cambridge University Press, New York, 1997.

[31] A. Katok, B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, New York, 1995.

[32] I. Kupka, *On feedback equivalence*, Differential geometry, global analysis, and topology (Halifax, NS, 1990), Amer. Math. Soc., Providence, RI, CMS Conf. Proc. **12** (1991), 105-117.

[33] I. Kupka, J.-B. Pomet, *Global Aspects of Feedback Equivalence for a Parametrized Family of Systems*, Analysis of Controlled Dynamical Systems, B. Bonnard et al., eds., Birkhäuser, Boston, 337-346, 1991.

[34] John M. Lee, *Riemannian Manifolds – An Introduction to Curvature*, Springer, 1997.

[35] O. Mayer, A. Miller, *La géométrie centroaffine des courbes planes*, Annales Scientifiques de l'Université de Jassy **18** (1933), 234-280.

[36] J. W. Milnor, *Topology from the differentiable viewpoint*, Princeton University Press, 1997.

[37] W. Respondek, *Feedback classification of nonlinear control systems in* $\mathbb{R}^2$ *and* $\mathbb{R}^3$, Geometry of Feedback and Optimal Control, B. Jakubczyk and W. Respondek, eds., Marcel Dekker, New York, 347-382, 1998.

[38] W. Respondek, I. A. Tall, *Feedback classification of nonlinear single-input control systems with controllable linearization: normal forms, canonical forms, and invariants*, SIAM J. Control Optim. **41** (2003), 1498-1531.

[39] U. Serres, *On the curvature of two-dimensional optimal control systems and Zermelo's navigation problem*, to appear in Journal of Mathematical Sciences.

[40] U. Serres, *On curvature and feedback classification of two-dimensional optimal control systems*, Sovrem. Mat. Prilozh. **21** (2004), Geom. Zadachi Teor. Upr. 134-153.

[41] D. J. Struik, *Lectures on classical differential geometry*, 2nd edition, Dover, 1988.

[42] G. R. Wilkens, *Centro-affine geometry in the plane and feedback invariants of two-state scalar control systems*, Proceedings of Symposia in Pure Mathematics **64** (1999), 319-333.

[43] E. Zermelo, *Über das Navigationsproblem bei ruhender oder veränderlicher Windverteilung*, Z. Angew. Math. Mech. **11** (1931), 114-124.

## Abstract

The purpose of this thesis is the study of the local and global differential geometry of fully nonlinear smooth control systems on two-dimensional smooth manifolds. We are particularly interested in the feedback-invariants of such systems.

In a first part we will use the Cartan's moving frame method in order to determine these invariants and we will see that one of the most important feedback-invariants is the control analogue to the Gaussian curvature of a surface. As we will explain it, the control curvature reveals very precious information on the optimal synthesis of time optimal problems.

In a second part we will construct some microlocal normal forms for time optimal control system and we will characterize in an intrinsic manner the flat systems. Finally, we will deal with global features ; in particular we will see how to generalize the Gauss-Bonnet theorem for control systems on surfaces without boundary.



**Key-words :** control curvature, feedback-equivalence, control system, Pontryagin Maximum Principle.



## Résumé

L'objet de cette thèse est l'étude de la géométrie locale et globale des systèmes de contrôle non linéaires sur des variétés lisses de dimension deux. Nous nous intéressons particulièrement aux invariants par feedback de tels sytèmes.

Dans une première partie nous utiliserons la méthode du repère mobile de Cartan afin de déterminer ces invariants et nous verrons que l'un des plus importants invariants par feedback est l'analogue de contrôle de la courbure gaussienne d'une surface. Comme nous l'expliquerons, la courbure de contrôle révèle de trés précieuses informations sur la synthèse optimale des problèmes de temps minimal.

Dans une seconde partie nous construirons des formes normales microlocales pour les problèmes de temps minimal et nous caractériserons de manière intrinsèque les systèmes plats. Enfin, nous traiterons de propriétés globales ; nous verrons en particulier comment généraliser le théorème de Gauss-Bonnet aux systèmes de contrôle sur des surfaces compactes sans bord.



**Mots-clés :** courbure de contrôle, équivalence par feedback, système de contrôle, principe du maximum de Pontriaguine.