

Mathematics Area - PhD course in
Mathematical Analysis, Modelling, and Applications

**Second-order optimality conditions
and other problems
in geometric optimal control**

Candidate:
Motta Michele

Advisor:
Andrei A. Agrachev
Co-advisors:
Luca Rizzi

Academic Year 2025-26



Mathematik. — *Wir wollen die Feinheit und Strenge der Mathematik in alle Wissenschaften hineintreiben, so weit diess nur irgend möglich ist, nicht im Glauben, dass wir auf diesem Wege die Dinge erkennen werden, sondern um damit unsere menschliche Relation zu den Dingen festzustellen. Die Mathematik ist nur das Mittel der allgemeinen und letzten Menschenkenntniss.*

— *Friedrich Nietzsche, Die fröhliche Wissenschaft, §246*

Mathematics. — *We want to bring the subtlety and rigor of mathematics into all the sciences, as far as this is at all possible, not in the belief that in this way we shall come to know things themselves, but in order thereby to determine our human relation to things. Mathematics is only a means toward a general and ultimate human knowledge.*

English translation

Abstract

This thesis is primarily devoted to second-order optimality conditions in geometric optimal control. After introducing the necessary background, we present a unified framework for second-order necessary and sufficient conditions, based on the second variation, the associated Jacobi curve, conjugate times, and the Maslov index. This preliminary part reorganizes and adapts results available in the literature, with the aim of providing a coherent set of tools for the subsequent applications.

We then apply this framework to several classes of extremals and derive the corresponding optimality conditions. First, we study singular extremals for control-affine systems with an L^1 cost, establishing second-order necessary conditions and sufficient conditions for strong local optimality. Next, we consider piecewise-regular extremals obtained by concatenating regular arcs. This framework includes, as particular cases, regular extremals, that is extremals satisfying the strong Legendre condition, and bang-bang extremals. For this class, we characterize the second variation and the associated Jacobi curve, and formulate optimality criteria in terms of the Maslov index of the Jacobi curve. Finally, we partially extend the analysis to nice extremals, obtained by concatenating regular and singular arcs, and derive the corresponding optimality conditions.

The last two chapters contain two further contributions in geometric optimal control, independent of the study of second-order optimality. The first concerns the precise computation of the asymptotic motion complexity of a non-admissible curve for a control-affine system. The second provides an explicit formula for the Lyapunov exponents of a linear switched system alternating between two matrices that generate the Lie algebra $\mathfrak{sl}_2(\mathbb{R})$.

Contents

1	Introduction	11
2	Preliminaries of Geometric Control Theory	15
2.1	Notation and conventions	15
2.2	Symplectic and Hamiltonian preliminaries	16
2.2.1	Symplectic vector spaces	16
2.2.2	The Lagrangian Grassmannian	17
2.2.3	The cotangent bundle	18
2.2.4	Hamiltonian vector fields and Poisson brackets	18
2.3	Control systems and optimal control problems	20
2.3.1	Control systems	20
2.3.2	Optimal control problems	20
2.3.3	Local notions of optimality	21
2.3.4	Cost augmentation	22
2.4	Pontryagin extremals	22
2.4.1	The Pontryagin Maximum Principle	22
2.4.2	Regular and singular arcs	23
2.4.3	Switching times and the switching Legendre condition	24
3	Second-order necessary optimality conditions: general principles	25
3.1	Constrained optimization problems	27
3.1.1	OCP as a constrained optimization problem	27
3.1.2	First order condition: Lagrange multipliers	28
3.1.3	Second derivatives for constrained maps	29
3.1.4	$\mathcal{L}$ -derivative: definition and general properties	31
3.2	Endpoint map, its differential and Hessian in optimal control	33
3.2.1	Hamiltonian lifts	33
3.2.2	Transported-back dynamics	34
3.2.3	First differential	35
3.2.4	Second differential and Hessian	37

3.3	Jacobi curve and Jacobi equation	38
3.3.1	Jacobi curve	38
3.3.2	Regular extremal arcs and Jacobi equation	41
3.4	Morse and Maslov index	46
3.4.1	Maslov index for smooth curves	46
3.4.2	Maslov index for increasing piecewise smooth curves	49
3.4.3	Relation between the Maslov and the Morse index	49
4	Second-order sufficient conditions: general principles	51
4.1	Sufficient conditions	52
4.1.1	The Poincaré–Cartan form	53
4.1.2	The free-time field of extremals and the super-Hamiltonian	54
4.1.3	Sufficient optimality criterion	56
4.1.4	Invertibility of $\pi_{\mathcal{M}}$ and the Maslov index	59
4.2	Conjugate times and conjugate points	63
5	Singular extremals of optimal control problems with L^1 cost	67
5.1	Introduction	67
5.2	Description of extremal trajectories	70
5.2.1	Definitions and notation	70
5.2.2	Description of extremal trajectories	71
5.3	The super-Hamiltonian function and strong local optimality of short arcs	73
5.4	Jacobi equation and sufficient conditions for optimality	78
5.4.1	Jacobi curve and conjugate points	78
5.4.2	Sufficient condition for optimality	82
5.5	Examples	86
5.5.1	An example: the Heisenberg group with a drift	86
5.5.2	A Martinet-type example	88
5.5.3	Left-invariant problem in $SU(2)$	90
5.6	Second variation and the generalized Legendre condition	98
5.6.1	Definition of the endpoint map and the Hessian	98
5.6.2	Explicit expression for the second variation form	98
5.6.3	Proof of Theorem 5.1.1	102
5.6.4	Derivation of the Jacobi equation and proof of Theorem 5.1.3	104
5.7	Proof of formula (5.6.8)	106
6	Piecewise regular extremals - part 1: main results and a non-convex one-dimensional example	109
6.1	Introduction	109
6.1.1	Formulation of the problem and motivations	109
6.1.2	Second order optimality conditions for piecewise regular extremals	112

6.2	A one-dimensional non-convex Lagrangian example	115
6.2.1	Statements of the results	117
6.2.2	Generalized harmonic oscillator	120
6.2.3	Proof of Theorem 6.2.1	126
6.2.4	Proof of Proposition 6.2.2	131
6.2.5	Proof of Proposition 6.2.3	132
7	Piecewise regular extremals - part 2: proofs of main results	137
7.1	The endpoint map	138
7.2	Optimal control problems with fixed final time	140
7.3	Differential and Hessian of the endpoint map	142
7.3.1	Useful formulas	142
7.3.2	Differential and Hessian of the endpoint map	143
7.4	Jacobi curve of a piecewise regular extremal	148
7.4.1	Proof of Theorem 6.1.6	148
7.4.2	Hamiltonian characterization of the regular arcs	153
7.5	The maximized Hamiltonian flow and the field of extremals	154
7.5.1	The maximized Hamiltonian and its flow	155
7.5.2	Structure of the set of Lagrange multipliers	158
7.6	Sufficient condition for optimality	162
8	Nice extremals	173
8.1	Analysis of first-order singular extremals	174
8.2	Second variation of first-order singular extremals	182
8.2.1	Endpoint map and admissible variations	183
8.2.2	First and second variation of the endpoint map	184
8.2.3	Goh transformation	187
8.2.4	Augmented second variation	189
8.2.5	Generalized Legendre condition	191
8.2.6	Jacobi equation and Jacobi curve of a first-order singular extremal	196
8.3	Super-Hamiltonian and Jacobi equation	198
8.4	Proof of Proposition 8.3.2	201
8.5	Jacobi curve for a nice extremal	208
8.5.1	Endpoint map near a nice extremal	209
8.5.2	First and second differential of the endpoint map	210
8.5.3	Structure of the Jacobi curve for a nice extremal	212
8.6	The super-Hamiltonian near regular-singular junctions	214

9	Asymptotics of motion planning complexity for control-affine systems	223
9.1	Introduction	223
9.1.1	Setting and main results	224
9.2	Preliminaries and normal forms	229
9.2.1	Continuity of motion complexity	229
9.2.2	Decomposition of the drift and proof of (9.1.7)	231
9.2.3	Normal forms	232
9.3	Step 2 case: the reduced motion complexity	234
9.3.1	Proof of Theorem 9.3.3	235
9.3.2	Proof of Theorem 9.3.4	237
9.4	Step 2 case: proof of Theorem 9.1.2	241
9.4.1	Proofs of the auxiliary lemmata	242
9.5	Distributions of generic type	247
9.5.1	Case of a single Martinet point	247
9.5.2	Case of many Martinet Points	252
9.6	Intrinsic definition of κ	253
10	Lyapunov exponents of linear switched systems	255
10.1	Introduction	255
10.2	Reduction to OC Problem and application of PMP	258
10.2.1	Pontryagin Maximum Principle	259
10.3	General properties of extremals	262
10.3.1	Invariant Hyperboloid	263
10.3.2	Switching function, <i>bang-bang</i> and singular extremals	264
10.4	<i>Bang-bang</i> extremals	264
10.4.1	Properties of <i>bang-bang</i> extremals	264
10.4.2	Hyperboloid foliation	267
10.4.3	<i>Bang-bang</i> periodic controls	271
10.4.4	Principal Lyapunov exponent for periodic controls	273
10.5	Singular extremals	278
10.6	Final remarks	280
A	Appendix	285
A.1	Chronological calculus and the variation formula	285
A.2	Lipschitz maps and piecewise smooth manifolds	287
	Ringraziamenti	297

Chapter 1

Introduction

The mathematical theory of control concerns the study of dynamical systems governed by equations involving some parameters, the *controls*, which can be changed during the evolution of the system. By changing these parameters, one can influence the evolution of the system and steer it in a desired direction.

This is a mathematical framework for modelling a vast number of different situations. A typical example is driving a car: the driver controls the motion of the car by accelerating, braking, and turning the steering wheel. In this case, the dynamical system is described by the physical laws governing the motion of the car, while the controls describe the actions available to the driver.

A large variety of phenomena, ranging from everyday situations to advanced technological applications, are modelled following the same philosophy. Concrete applications of control theory appear in engineering, economics, biology, medicine, and many other fields.

Most importantly, in many of these applications one does not only want to control the dynamics of the system, but wants to do so in an *optimal* way. For instance, when driving a car from the city centre of Trieste to SISSA, one would like not only to find a path reaching the desired destination, but perhaps the shortest path, the quickest one, or the one requiring the least amount of fuel.

The mathematical theory of optimal control is the branch of mathematics devoted to understanding the properties of this class of optimization problems.

Although control problems are ubiquitous, optimal control is a relatively recent field of mathematics. With a few notable exceptions, until the twentieth century such problems were not developed within a systematic mathematical theory and were mainly studied on a case-by-case basis. It was only around the 1950s, with fundamental contributions by figures such as Bellman, Pontryagin, and Kalman, that control theory fully developed as an independent mathematical discipline.

One of the milestones of this theory is the Pontryagin Maximum Principle (PMP), a very flexible theorem providing necessary conditions that optimal trajectories must sat-

isfy. In many situations, PMP reduces the original optimization problem to the study of a Hamiltonian dynamical system together with a maximization condition, and therefore provides a way to characterize candidate optimal trajectories. Trajectories satisfying these conditions, together with their corresponding Hamiltonian lifts, are called *extremals*; informally, they are the candidate optimal trajectories selected by PMP.

However, the conditions provided by PMP are only necessary. They can be degenerate, may leave a large family of candidate extremals, and an extremal satisfying PMP need not be locally optimal. For this reason, stronger necessary conditions and sufficient conditions have been developed in order to distinguish locally optimal trajectories among the extremals selected by PMP. The theory of sufficient optimality conditions is substantially less universal, and there is no single general principle with the scope and flexibility of PMP.

The main aim of this thesis is to contribute to both directions through the study of second-order optimality conditions in optimal control. The adjective *geometric* in the title reflects the fact that, in many applications, the natural state space of the system is not a Euclidean space but a smooth manifold. For example, the configuration of a mechanical arm with two rotating joints is naturally described by two angles and therefore belongs to a two-dimensional torus $\mathbb{S}^1 \times \mathbb{S}^1$, or to a subset of it if the joints have limited ranges. More generally, using manifolds makes it possible to describe the configuration space of a system without introducing artificial coordinates. At the same time, one of the central tools of the theory is symplectic geometry, whose intrinsic formulation naturally lives on the cotangent bundle of a manifold. Working intrinsically has an important practical advantage: results formulated in coordinate-free terms apply directly to any new problem satisfying the required assumptions, without having to adapt the underlying formulas case by case. For general introductions to this viewpoint, we refer to [13, 5].

More precisely, our approach is based on the second variation, the Jacobi curve, conjugate times, the Maslov index, and the method of fields of extremals. As in ordinary finite-dimensional optimization, second-order information provides a natural way to refine first-order necessary conditions. Here the role of the Hessian is played by the second variation, which describes how the cost changes, to second order, under admissible perturbations of an extremal. The central idea is that second-order information, naturally expressed analytically through the Hessian of the optimal control problem, can be encoded geometrically by a curve of Lagrangian subspaces and studied using symplectic methods. This viewpoint has been developed in several forms in the literature; some references particularly relevant for the framework used here are [4, 3, 6].

This thesis is organized as follows. In Chapter 2, we introduce the general background material: the basic notions of symplectic and Hamiltonian geometry, optimal control problems on smooth manifolds, and the Pontryagin Maximum Principle.

In Chapter 3, we introduce the general principles of second-order necessary optimality conditions. Starting from constrained optimization and the linearization of the Lagrange multiplier rule, we introduce the $\mathcal{L}$ -derivative and the Jacobi curve associated with an extremal. The main result relates the Morse index of the constrained Hessian to the

Maslov index of the Jacobi curve, connecting analytic second-order information with a finite-dimensional symplectic invariant.

Chapter 4 provides the complementary approach yielding sufficient optimality conditions. We recall the method of fields of extremals and its Hamiltonian formulation through suitable super-Hamiltonians and formulate a version adapted to the applications developed later in the thesis. Related Hamiltonian approaches to sufficient optimality conditions can be found, for instance, in [14, 71] and references therein.

The material in Chapters 3 and 4 mostly adapts results already present in the literature, with the purpose of providing a unified framework for the subsequent applications.

In Chapters 5, 6, 7, and 8, we apply these principles to increasingly general classes of extremals. There is a natural progression among these chapters. Chapter 5 serves as a first singular model and as a testing ground in which we developed both the super-Hamiltonian technique and the analysis of singular second variations. Chapters 6 and 7 then study piecewise regular extremals, extending the classical bang-bang setting to trajectories whose maximizing Hamiltonian changes among smooth branches. Finally, Chapter 8 combines the two constructions to treat trajectories formed by regular and first-order singular arcs, a broad class of regular–singular patterns frequently encountered in applications.

Chapter 5 covers the joint work [7] and studies singular extremals for the problem of minimizing the L^1 norm of the control in a control-affine system. We derive a generalized Legendre–Clebsch condition, determine the corresponding Jacobi equation, and, adapting the approach in [70], we construct a super-Hamiltonian adapted to the singular dynamics. This leads to necessary conditions and to sufficient conditions for strong local optimality, first for short arcs and then in the absence of conjugate points.

Chapters 6 and 7 cover the joint work [12] and are devoted to piecewise regular extremals. We characterize their second variation and Jacobi curve by combining the smooth evolution along regular arcs with discontinuities at switching times, obtaining necessary and sufficient optimality criteria in terms of the Maslov index of this curve. Chapter 6 contains the main results and a detailed analysis of a non-convex one-dimensional example, where piecewise regular extremals arise as candidate minimizers. The complete proofs of our optimality conditions for this family of extremals are developed in Chapter 7.

Chapter 8 is devoted to an ongoing joint work and constitutes the last step of this progression. We introduce *nice extremals*, obtained by concatenating regular and first-order singular arcs. Although the ingredients developed in the preceding chapters extend naturally to this setting, combining them is non-trivial, in particular at regular–singular junctions. We develop the second-order theory and the super-Hamiltonian construction for first-order singular arcs, characterize the Jacobi curve of nice extremals, and derive the corresponding necessary optimality conditions. On the sufficient side, adapting the technique used in [64], we prove strong local optimality in the absence of conjugate points for nice extremals containing at most one singular arc. We expect that the restriction to one singular arc is technical: we conjecture that the same result holds for an arbitrary finite number of first-order singular arcs.

The material contained in Chapters 9 and 10 consists of two further contributions in geometric optimal control which are independent of the second-order optimality theory developed in the central part of the thesis. Chapter 9 covers the joint work [58] with Dario Prandi and gives precise asymptotic formulas, with explicit constants, for the motion-planning complexity of tracking a non-admissible curve in control-affine systems. This work extends the previous results in [69, 40] for control-linear systems. Chapter 10 covers the joint work [11] with Andrei Agrachev and gives an explicit formula for the maximal Lyapunov exponent of switched systems generated by two matrices spanning a Lie algebra isomorphic to $\mathfrak{sl}_2(\mathbb{R})$, together with the switching strategies realizing it.

The results of the thesis also suggest several directions for further research. The most immediate one is to complete the sufficient optimality theory for nice extremals with an arbitrary finite number of singular arcs. A second natural objective is to turn the theoretical criteria into an effective algorithmic procedure for testing optimality and locating conjugate times, in the spirit of the numerical implementation developed for regular extremals in [24]. It would also be interesting to extend the theory to include higher-order singular extremals and Fuller-type phenomena, where infinitely many switching times may accumulate in finite time. Other extensions concern problems with free endpoints and costs depending on the initial and final states, as well as optimal control problems with state constraints.

The results obtained in the different classes considered in this thesis suggest that the combination of techniques developed throughout this work provides a robust framework for the analysis of more general problems. Therefore, on the basis of this experience, we think that a similar approach can be pushed beyond the cases treated here to address the directions mentioned above.

As already pointed out, Chapters 2, 3, and 4 mainly rely on material already present in the literature. The original research carried out in this PhD thesis appears in the following preprints and publications:

- [11] Andrei A. Agrachev and Michele Motta, *Lyapunov Exponents of Linear Switched Systems*. Regul. Chaot. Dyn. 30, 481–503 (2025).
- [7] Andrei Agrachev, Ivan Beschastnyi, Michele Motta, *Singular extremals of optimal control problems with L^1 cost*, arXiv:2511.21527.
- [58] Michele Motta and Dario Prandi, *Asymptotics of motion planning complexity for control-affine systems*, arXiv:2511.17130.
- [12] Andrei A. Agrachev and Michele Motta, *Second order optimality conditions for piecewise regular extremals in Optimal Control*, arXiv:2607.10434.

Finally, the material in Chapter 8 is part of an ongoing work, still in preparation:

- Andrei A. Agrachev and Michele Motta, *Second order optimality conditions for nice extremals in Optimal Control*, in preparation.

Chapter 2

Preliminaries of Geometric Control Theory

This chapter collects the notation and the standard notions in geometric control theory used throughout the thesis. The presentation is deliberately concise and is mainly based on [13, 5]. More specific assumptions, as well as the second-order constructions required in the subsequent chapters, are introduced when needed.

2.1 Notation and conventions

In this thesis, all manifolds are assumed to be Hausdorff, second countable, and smooth; throughout the thesis, *smooth* means C^∞ -smooth. We always denote by M a finite-dimensional smooth manifold and, for $q \in M$, we denote by $T_q M$ and $T_q^* M$ its tangent and cotangent spaces at q , and by TM and T^*M the corresponding bundles. The canonical projection of the cotangent bundle is denoted by

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

The spaces of smooth functions and smooth vector fields on M are denoted by $C^\infty(M)$ and $\text{Vec}(M)$, respectively. The duality pairing between tangent vectors and covectors is written as $\langle \lambda, v \rangle$.

For a smooth map $F : M \rightarrow N$, its differential at $q \in M$ is denoted by $d_q F : T_q M \rightarrow T_{F(q)} N$. When no confusion is possible, we also use F_* for the push-forward and F^* for the pull-back. For a subset A of a vector space, $\text{span } A$ denotes its linear span; for a linear map B , its kernel and image are denoted by $\ker B$ and $\text{Im } B$. The identity map is denoted by Id .

Curves are defined on compact intervals unless stated otherwise. An equality involving a measurable control or the derivative of a Lipschitz continuous trajectory is understood

to hold almost everywhere. We frequently write q_t , u_t , and λ_t in place of $q(t)$, $u(t)$, and $\lambda(t)$. The restriction of a curve q to an interval I is denoted by $q|_I$.

The control set is denoted by $U \subset \mathbb{R}^m$. The precise assumptions on U , on the dynamics, and on the admissible-control space will be specified in each chapter.

2.2 Symplectic and Hamiltonian preliminaries

The symplectic framework introduced below provides the natural language for describing Pontryagin extremals and, later, their second-order variations through Jacobi curves and the Maslov index. We recall only the symplectic notions that will be used repeatedly. For a more complete account, we refer to [5, Chapters 4 and 14] and [57].

2.2.1 Symplectic vector spaces

A *symplectic vector space* is a pair (V, σ) , where V is a finite-dimensional vector space and σ is a non-degenerate skew-symmetric bilinear form on V . In particular, $\dim V = 2n$ for some $n \in \mathbb{N}$. For a linear subspace $\Gamma \subset V$, its symplectic orthogonal is

$$\Gamma^\perp := \{z \in V \mid \sigma(z, w) = 0 \text{ for every } w \in \Gamma\}.$$

A subspace Γ is called *isotropic* if $\Gamma \subset \Gamma^\perp$, *coisotropic* if $\Gamma^\perp \subset \Gamma$, and *symplectic* if $\Gamma \cap \Gamma^\perp = \{0\}$. A subspace $\Lambda \subset V$ is *Lagrangian* if it is isotropic and has dimension n . Equivalently,

$$\Lambda \text{ is Lagrangian} \iff \Lambda = \Lambda^\perp.$$

We use the notation

$$\mathfrak{L}(V) := \{\Lambda \subset V \mid \Lambda \text{ is Lagrangian}\}. \quad (2.2.1)$$

The set $\mathfrak{L}(V)$ is called *Lagrangian Grassmannian* and is studied in more detail in the next subsection.

Given $\Pi \in \mathfrak{L}(V)$, consider another Lagrangian subspace Δ such that

$$V = \Pi \oplus \Delta.$$

It is always possible to choose a basis $\{e_1, \dots, e_n\}$ of Π and a basis $\{f_1, \dots, f_n\}$ of Δ such that

$$\sigma(e_i, f_j) = \delta_{ij}. \quad (2.2.2)$$

Every vector $v \in V$ can be written in these coordinates: $v = \sum_{i=1}^n p^{(i)} e_i + q^{(i)} f_i$. Consequently, denoting by $(p, q) = (p^{(1)}, \dots, p^{(n)}, q^{(1)}, \dots, q^{(n)})$, the symplectic form takes the Darboux form

$$\sigma((p_1, q_1), (p_2, q_2)) = \langle p_1, q_2 \rangle - \langle p_2, q_1 \rangle.$$

A linear isomorphism $A : (V_1, \sigma_1) \rightarrow (V_2, \sigma_2)$ is a *symplectomorphism* if

$$\sigma_2(Av, Aw) = \sigma_1(v, w)$$

for every $v, w \in V_1$. Symplectomorphisms preserve symplectic orthogonals and, in particular, map Lagrangian subspaces to Lagrangian subspaces.

2.2.2 The Lagrangian Grassmannian

In this subsection, we recall some properties of the Lagrangian Grassmannian $\mathfrak{L}(V)$, defined in (2.2.1). It is a compact smooth submanifold of the Grassmannian of n -dimensional subspaces of V , of dimension

$$\dim \mathfrak{L}(V) = \frac{n(n+1)}{2}.$$

We briefly recall its standard local coordinates and the identification of its tangent spaces with quadratic forms (see [5, Chapter 14]).

Fix a pair of transverse Lagrangian subspaces $\Pi, \Delta \in \mathfrak{L}(V)$, so that

$$V = \Pi \oplus \Delta.$$

Every $\Lambda \in \mathfrak{L}(V)$ transverse to Δ is the graph of a linear map $S_\Lambda : \Pi \rightarrow \Delta$:

$$\Lambda = \{(p, S_\Lambda p) \mid p \in \Pi\}.$$

Once a basis of Π and Δ as in (2.2.2) is fixed, we can identify S_Λ with its matrix. A direct check shows that Λ is Lagrangian if and only if S_Λ is symmetric.

Lagrangian subspaces transversal to Δ form an open set

$$\Delta^\natural := \{\Lambda \in \mathfrak{L}(V) \mid \Lambda \cap \Delta = \{0\}\}$$

and is naturally identified with the vector space of linear maps on from Π to Δ . These open sets, as Δ varies in $\mathfrak{L}(V)$, form an atlas of $\mathfrak{L}(V)$. For fixed transversal $\Pi, \Delta \in \mathfrak{L}(V)$, we call these local coordinates the (Π, Δ) -coordinates of $\Delta^\natural$. Since the space of symmetric $n \times n$ matrices has dimension $n(n+1)/2$, this chart also gives the dimension formula above.

Now, let $t \mapsto \Lambda(t)$ be a smooth curve in $\mathfrak{L}(V)$ and fix t_0 . Given $v \in \Lambda(t_0)$, choose a smooth curve $v(t) \in \Lambda(t)$ satisfying $v(t_0) = v$. The *velocity* of Λ at t_0 is the quadratic form

$$\dot{\Lambda}(t_0)[v] := \sigma \left(v, \frac{d}{dt} v(t) \Big|_{t=t_0} \right).$$

This definition is independent of the chosen extension $v(t)$. To show this, we can fix any $\Delta \in \mathfrak{L}(V)$, $\Delta \cap \Lambda_{t_0} = \{0\}$, take Darboux coordinates such that $\Lambda_{t_0} = \{(p, 0)\}$ and

$\Delta = \{(0, q)\}$, and compute $\sigma\left(v(t_0), \frac{d}{dt}\big|_{t=t_0} v(t)\right)$ in these coordinates. Let $S(t)$ be the linear map corresponding to $\Lambda(t)$. We have that $v(t) = (p(t), S(t)p(t))$ and

$$\sigma\left((p, Sp), (\dot{p}, \dot{S}p + S\dot{p})\right) = p^\top \dot{S}p,$$

where $p, \dot{p}, S, \dot{S}$ are all evaluated at t_0 . Thus, we see that the quantity depends only on $p(t_0)$. Furthermore, from this computation we see that, if $\Lambda(t)$ is represented by a family of symmetric matrices $S(t)$, then $\dot{\Lambda}(t)$ is represented by $\dot{S}(t)$.

A smooth curve $\Lambda : [0, T] \rightarrow \mathfrak{L}(V)$ is called *monotone increasing* if

$$\dot{\Lambda}(t) \geq 0, \text{ for every } t \in [0, T],$$

and *strictly monotone increasing* if the inequality is strict. Finally, we say that a curve Λ is *simple* if there is $\Delta \in \mathfrak{L}(V)$ such that $\Lambda(t) \in \Delta^\natural$ for every $t \in [0, T]$.

2.2.3 The cotangent bundle

The cotangent bundle T^*M carries a canonical symplectic structure. The tautological, or Liouville, one-form ϑ is defined by

$$\vartheta_\lambda(v) := \langle \lambda, d_\lambda \pi[v] \rangle, \quad \lambda \in T^*M, \quad v \in T_\lambda(T^*M).$$

We use the convention

$$\sigma = d\vartheta.$$

Thus, σ is a skew-symmetric non-degenerate 2-form, which is called the *canonical symplectic form* of T^*M . In local coordinates (p, q) on T^*M , this reads

$$\vartheta = p dq, \quad \sigma = dp \wedge dq.$$

For every $\lambda \in T^*M$, the vertical subspace is

$$\Pi_\lambda := \ker d_\lambda \pi = T_\lambda(T_{\pi(\lambda)}^*M).$$

It is elementary to check that Π_λ is a Lagrangian subspace of the symplectic vector space $T_\lambda(T^*M)$.

2.2.4 Hamiltonian vector fields and Poisson brackets

Given $h \in C^\infty(T^*M)$, its Hamiltonian vector field $\vec{h}$ is defined by

$$dh = \sigma(\cdot, \vec{h}).$$

In local coordinates, the corresponding Hamiltonian system is

$$\begin{aligned}\dot{p} &= -\frac{\partial h}{\partial q}, \\ \dot{q} &= \frac{\partial h}{\partial p}.\end{aligned}$$

The Poisson bracket of $a, b \in C^\infty(T^*M)$ is

$$\{a, b\} := \sigma(\vec{a}, \vec{b}) = \vec{a}(b).$$

Hence, if λ solves $\dot{\lambda} = \vec{a}(\lambda)$, then

$$\frac{d}{dt}b(\lambda_t) = \{a, b\}(\lambda_t).$$

We will repeatedly use the standard identities

$$\{a, b\} = -\{b, a\}, \quad \{a, bc\} = \{a, b\}c + b\{a, c\}, \quad [\vec{a}, \vec{b}] = \overrightarrow{\{a, b\}}.$$

To every vector field $f \in \text{Vec}(M)$ we associate the fiber-linear Hamiltonian

$$h_f(\lambda) := \langle \lambda, f(\pi(\lambda)) \rangle.$$

With the preceding conventions, Lie brackets of vector fields and Poisson brackets of fiber-linear Hamiltonians are related by

$$\{h_f, h_g\} = h_{[f, g]}.$$

Let $\Phi_{s,t}^h$ denote the flow of a possibly time-dependent Hamiltonian h_t . In the autonomous case we write Φ_t^h or $e^{t\vec{h}}$. Hamiltonian flows preserve the symplectic form:

$$(\Phi_{s,t}^h)^* \sigma = \sigma.$$

In particular, the differential

$$d_\lambda \Phi_{s,t}^h : T_\lambda(T^*M) \longrightarrow T_{\Phi_{s,t}^h(\lambda)}(T^*M)$$

is a linear symplectomorphism. We will often use this differential to transport Lagrangian subspaces between tangent spaces based at different points of T^*M .

2.3 Control systems and optimal control problems

2.3.1 Control systems

Let M be a smooth manifold and let $U \subset \mathbb{R}^m$. A control system on M is described by a map

$$f : M \times U \longrightarrow TM, \quad f(q, u) \in T_q M.$$

Given $T > 0$ and a measurable control $u : [0, T] \rightarrow U$, an *admissible trajectory* is a Lipschitz continuous curve $q : [0, T] \rightarrow M$ satisfying

$$\dot{q}(t) = f(q(t), u(t)) \tag{2.3.1}$$

for almost every $t \in [0, T]$. When it is useful to emphasize the initial point and the control, the corresponding trajectory is denoted by $q(t; q_0, u)$. Existence and uniqueness of solutions of (2.3.1) are usually obtained by Carathéodory's theorem, see [5, Theorem 2.15] and references therein. In the following, we will always assume that f is regular enough to satisfy the hypotheses of Carathéodory's theorem.

We denote by $\mathfrak{U}_T$ the chosen space of admissible controls on $[0, T]$ and by

$$q(\cdot; q_0, u)$$

the maximal trajectory with initial condition q_0 . If u is an admissible control and q is its corresponding trajectory, We say that (u, q) is an admissible pair.

In this thesis, we will often fix a reference control and the corresponding reference trajectory and compare them with nearby admissible controls and trajectories. Whenever this is the case, we will denote by $\tilde{u}, \tilde{q}$ the reference control and its corresponding trajectory.

Moreover, since in all the cases that we will study all trajectories under consideration are contained in a fixed relatively compact neighbourhood of $\tilde{q}$, without loss of generality we will always assume completeness of all the relevant vector fields.

2.3.2 Optimal control problems

The common model considered in the thesis is the fixed-endpoint optimal control problem with integral cost. That is, given a continuous function $L : M \times U \rightarrow \mathbb{R}$, we want to minimize

$$J(u) = \int_0^T L(q(t), u(t)) dt \longrightarrow \min, \tag{2.3.2}$$

subject to

$$\dot{q}(t) = f(q(t), u(t)), \quad q(0) = q_0, \quad q(T) = q_1. \tag{2.3.3}$$

The final time T may be fixed or free, according to the problem under consideration. We always specify this choice explicitly when it affects the statement of a result.

The same formalism extends with minor modifications to problems in which the initial and final points are not fixed and to Bolza costs containing endpoint terms. Since these extensions play no essential role in the arguments developed here, we formulate the preliminary results only for fixed endpoints. The different optimization problem used in Chapter 10 will be introduced directly there.

2.3.3 Local notions of optimality

Let $\tilde{u}$ be an admissible control, let $\tilde{q}$ be the corresponding trajectory, and assume that the boundary conditions are fixed. The control $\tilde{u}$ is a *global minimizer* if

$$J(\tilde{u}) \leq J(u)$$

for every admissible competitor u satisfying the same boundary conditions.

Local optimality depends on the topology chosen on the admissible controls or trajectories. The conventional distinction between weak and strong local optimality is the following.

Definition 2.3.1 (weak-local optimality). Assume that the set $\mathcal{U}$ is equipped with a topology. An admissible control $\tilde{u} \in \mathcal{U}$ is *weakly locally optimal* for the OCP (2.3.2),(2.3.3) if there is a neighbourhood $\mathcal{U}_0 \subset \mathcal{U}$ of $\tilde{u}$ such that

$$J(\tilde{u}) \leq J(u) \tag{2.3.4}$$

for every $u \in \mathcal{U}_0$ satisfying $q(0; u) = \tilde{q}(0)$, $q(T; u) = \tilde{q}(T)$.

In the OCPs considered in this thesis, $\mathcal{U}$ will be either a subset of $L^\infty([0, T]; U)$ or of $L^1([0, T]; U)$ and we consider weak-local optimality with respect to the induced topology.

Definition 2.3.2 (Local-strong-optimality). Let $\tilde{u}$ be an admissible control for the optimal control problem (2.3.2),(2.3.3) with fixed final time T and denote by $\tilde{q}$ its corresponding admissible trajectory. We say that $\tilde{q}$ is *strongly locally optimal* if there exists a neighbourhood $\mathcal{O}$ of its image such that

$$J(\tilde{u}) \leq J(u) \tag{2.3.5}$$

for every admissible competitor u with

$$q(0; u) = \tilde{q}(0), \quad q(T; u) = \tilde{q}(T) \quad \text{and} \quad q(t; u) \in \mathcal{O}, \quad t \in [0, T].$$

If the final time T is free, we say that $\tilde{q}$ is *strongly locally optimal* if there exists a neighbourhood $\mathcal{O}$ of $\tilde{q}$ and $\varepsilon > 0$ such that (2.3.5) holds for every admissible trajectory $q(\cdot; u)$ satisfying

$$q(0; u) = \tilde{q}(0), \quad q(T'; u) = \tilde{q}(T) \quad \text{and} \quad q(t; u) \in \mathcal{O}, \quad t \in [0, T']$$

with $|T' - T| < \varepsilon$.

A local minimizer is called *strict* if the inequalities in (2.3.5),(2.3.4) are strict for $u \neq \tilde{u}$.

2.3.4 Cost augmentation

An integral cost can be incorporated into the state space by introducing an additional variable q^0 satisfying

$$\dot{q}^0(t) = L(q(t), u(t)), \quad q^0(0) = 0.$$

Then

$$q^0(T) = J(u).$$

Thus the original optimal control problem can be viewed as an endpoint problem for the extended system on $\mathbb{R} \times M$. This elementary device underlies the geometric interpretation of the cost multiplier in the Pontryagin Maximum Principle.

2.4 Pontryagin extremals

We recall the Pontryagin Maximum Principle in the fixed-endpoint setting. The statement below is sufficient for the applications in this thesis; more general variants can be found in [13].

2.4.1 The Pontryagin Maximum Principle

For $\nu \in \mathbb{R}$, $\lambda \in T_q^*M$, and $u \in U$, define the Pontryagin Hamiltonian

$$h_u^\nu(\lambda) := \langle \lambda, f(q, u) \rangle + \nu L(q, u).$$

Whenever the maximum exists, the maximized Hamiltonian is

$$H^\nu(\lambda) := \max_{u \in U} h_u^\nu(\lambda).$$

Theorem 2.4.1 (Pontryagin maximum principle (PMP)). *Let $\tilde{u}$ be a minimizer for the OCP (2.3.2), (2.3.3) and let $\tilde{q}$ be its corresponding admissible trajectory. Then, there exists a non-trivial pair*

$$(\nu, \tilde{\lambda}_t) \neq 0, \quad \tilde{\lambda}_t \in T_{\tilde{q}(t)}^*M, \quad \nu \in \mathbb{R}, \quad t \in [0, T],$$

such that the following conditions hold:

$$\begin{aligned} \dot{\tilde{\lambda}}_t &= \vec{h}_{\tilde{u}(t)}^\nu(\tilde{\lambda}_t), \quad \text{for a.e. } t \in [0, T], \\ h_{\tilde{u}(t)}^\nu(\tilde{\lambda}_t) &= H^\nu(\tilde{\lambda}_t) \quad \text{for a.e. } t \in [0, T], \\ \nu &\leq 0. \end{aligned} \tag{2.4.1}$$

Moreover, $[0, T] \ni t \mapsto h_{\tilde{u}(t)}^\nu(\tilde{\lambda}_t)$ is constant and, in the case of free final time, the constant is equal to 0.

The pair $(\tilde{u}, \tilde{\lambda})$ is called an *extremal pair*, $\tilde{\lambda}$ an *extremal lift*, and $\tilde{q} = \pi \circ \tilde{\lambda}$ an *extremal trajectory*. Since the multiplier is defined up to multiplication by a positive constant, there are two cases. If $\nu < 0$, it is normalized to $\nu = -1$ and the extremal is called *normal*. If $\nu = 0$, the extremal is called *abnormal*. In the normal case, we usually omit ν and write

$$h_u(\lambda) = \langle \lambda, f(q, u) \rangle - L(q, u).$$

Notice that in this thesis we will deal essentially only with normal extremals. For this reason, unless it is specifically stated, we will always assume $\nu = -1$. The Maximum Principle is a first-order necessary condition: in general curves satisfying PMP need not to be minimizers, not even locally.

2.4.2 Regular and singular arcs

Let $(\tilde{u}, \tilde{\lambda})$ be an extremal pair on an interval I . Suppose that there is a sufficiently small open ball $B \subset \mathbb{R}^m$ such that $U_0 := U \cap B$ is a smooth manifold and that the control $\tilde{u}|_I$ takes values in U_0 . Assume moreover that $h|_{T^*M \times U_0} \in C^2$. The maximization condition implies the stationarity relation

$$\left. \frac{\partial h_u}{\partial u} \right|_{u=\tilde{u}(t), \lambda=\tilde{\lambda}(t)} [v] = 0, \quad v \in T_{\tilde{u}(t)}U_0, \quad t \in I.$$

The *Legendre form* is the quadratic form on $T_{\tilde{u}(t)}U_0$ defined by

$$\mathbb{L}_t(v) := \left. \frac{\partial^2 h_u}{\partial u^2} \right|_{u=\tilde{u}(t), \lambda=\tilde{\lambda}(t)} [v, v]. \quad (2.4.2)$$

Because the Pontryagin Hamiltonian is maximized with respect to u , every optimal extremal satisfies the ordinary Legendre condition

$$\mathbb{L}_t \leq 0, \quad \text{for a.e. } t \in I.$$

The *strong Legendre condition* is

$$\mathbb{L}_t < 0, \quad \text{for a.e. } t \in I.$$

Under the strong Legendre condition, the implicit function theorem locally determines the maximizing control as a smooth function of the covector,

$$u = \hat{u}(\lambda),$$

and the corresponding maximized Hamiltonian is smooth:

$$H(\lambda) = h_{\hat{u}(\lambda)}(\lambda).$$

On such an interval, the extremal is governed by the ordinary Hamiltonian system

$$\dot{\lambda} = \vec{H}(\lambda).$$

We call an extremal arc *regular* when the maximizing control is locally unique and is determined smoothly in this way, with non-degenerate negative Legendre form. An arc is called *singular* when the Legendre form has a non-trivial kernel.

2.4.3 Switching times and the switching Legendre condition

In the simple case where U is a finite collection of isolated points, which includes case of bang-bang controls, the set U_0 considered in the previous subsection is reduced to a single point and the tangent space $T_{\tilde{u}(t)}U_0$ is $\{0\}$. Thus, there are no non-trivial variations of $\tilde{u}$ which are close to the original control in the L^∞ norm. In particular, every control is L^∞ -weakly-locally optimal. This motivates the following considerations.

Let $(\tilde{u}, \tilde{\lambda})$ be an extremal pair on an interval I and suppose that there are two sufficiently small disjoint balls $B_0, B_1 \subset \mathbb{R}^m$ such that $U_0 = U \cap B_0$, $U_1 = U \cap B_1$ are two smooth manifolds and there is $t_1 \in I$ such that $\tilde{u}(t) \in U_0$ for $t \leq t_1$ and $\tilde{u}(t) \in U_1$ for $t > t_1$. We say that t_1 is a *switching time*.

We describe the analogous non-degeneracy condition at a switching time between two regular arcs. Let H^0 and H^1 be the smooth maximized Hamiltonians on U_0 and U_1 respectively. At the switching time,

$$H^0(\tilde{\lambda}_{t_1}) = H^1(\tilde{\lambda}_{t_1}).$$

Introduce the switching function

$$\chi := H^1 - H^0.$$

Its derivative at the switching time is

$$\left. \frac{d}{dt} \chi(\tilde{\lambda}_t) \right|_{t=t_1} = \{H^0, H^1\}(\tilde{\lambda}_{t_1}).$$

If the maximizing Hamiltonian changes from H^0 to H^1 , then χ changes from non-positive to non-negative. Consequently, the switching Legendre condition is

$$\{H^0, H^1\}(\tilde{\lambda}_{t_1}) \geq 0.$$

The *strong switching Legendre condition* is

$$\{H^0, H^1\}(\tilde{\lambda}_{t_1}) > 0.$$

In this case, both Hamiltonian vector fields $\vec{H}^0, \vec{H}^1$ are transverse to the switching hypersurface $\{H^0 = H^1\}$ and the switching time is isolated.

Chapter 3

Second-order necessary optimality conditions: general principles

The Pontryagin Maximum Principle (PMP) provides first-order necessary conditions for optimality and identifies a class of candidate minimizers, called extremals. Although PMP encodes the distinction between minimization and maximization through the sign of the cost multiplier, it does not in general distinguish locally optimal extremals from non-optimal ones, such as saddle-type extremals. Moreover, even if an extremal arc is minimizing, PMP does not by itself determine when it ceases to be minimizing. This is the same phenomenon encountered in Riemannian geometry: sufficiently short segments of a geodesic are locally minimizing, but the geodesic equation alone does not determine the interval on which this minimizing property persists. Understanding this loss of optimality requires studying the problem at the second order.

The purpose of this chapter is to present the general strategy underlying second-order necessary conditions for optimality. From the analytic viewpoint, an optimal control problem with fixed boundary conditions is regarded as an infinite-dimensional constrained optimization problem, in which the endpoint map plays the role of the constraint. With this viewpoint, the PMP encodes the first-order information, while the Hessian of the constrained functional contains the corresponding second-order information.

The abstract criterion of Theorem 3.1.7 states that, if the negative Morse index of this Hessian is at least the corank of the reference control, then the extended endpoint map is locally open and the control cannot be a weak local minimizer. The problem is therefore reduced to computing the Morse index of the constrained Hessian.

The geometric viewpoint provides an effective way to perform this computation. We introduce the $\mathcal{L}$ -derivative, which is a Lagrangian subspace encoding the relevant information carried by the second variation. If the set of the Lagrange multipliers is a smooth manifold, its tangent space coincides with the $\mathcal{L}$ -derivative. The strength of the $\mathcal{L}$ -derivative lies in the fact that it is a well defined Lagrangian subspace even when the set of Lagrange

multipliers fails to be a smooth manifold.

By restricting the admissible variations to successive time intervals, one obtains a curve of Lagrangian subspaces, called the Jacobi curve. Its intersections with the vertical subspace are measured by the Maslov index, and the Morse–Maslov formula of Theorem 3.4.11 identifies the Maslov index of the Jacobi curve, up to an explicit correction term, with the negative Morse index of the constrained Hessian. The main aim of the chapter is thus to construct a bridge between the analytic description of second-order variations and their symplectic representation through the Jacobi curve.

The abstract framework presented in this chapter is largely based on the results of [6] and on the earlier references cited therein. Since these techniques are quite involved and are not usually presented in a unified form, we provide a detailed account adapted to the notation and applications developed in this thesis.

Recall that an admissible control $\tilde{u}$ satisfying PMP is said to have *corank* r if the maximal number of pointwise linearly independent extremal lifts associated with $\tilde{u}$ is r . Equivalently,

$$r = \text{codim Im } D_{\tilde{u}}(F, J).$$

where we denote by F the endpoint map defined in Definition 3.1.1 and by J the minimized functional in (2.3.2). We let μ_{Π} be the Maslov index with respect to $\Pi \in \mathfrak{L}(T_{\tilde{\lambda}_0}^*(T^*M))$, see Section 3.4. We assume that the set of admissible controls $\mathcal{U}$ has a structure of Banach manifold modelled on a separable Banach space. The general criterion developed in this chapter can be summarized as follows.

Theorem 3.0.1 (Second-order necessary optimality condition). *Let $\tilde{u}$ be an extremal control for the optimal control problem (2.3.2), (2.3.3), and assume that $\tilde{u}$ has corank r .*

For every extremal lift $\tilde{\lambda}$ associated with $\tilde{u}$, let $\Lambda^{\tilde{\lambda}}$ be the corresponding Jacobi curve and let

$$\Pi^{\tilde{\lambda}} := T_{\tilde{\lambda}_0}^*(T_{q_0}^*M).$$

Assume that $-(\tilde{\lambda}, \nu) \text{Hess}_{\tilde{u}}(F, J)$ has finite negative Morse index for every such lift. If

$$\mu_{\Pi^{\tilde{\lambda}}}(\Lambda^{\tilde{\lambda}}) - \frac{1}{2} \left(d - \dim \left(\Pi^{\tilde{\lambda}} \cap \bigcap_{t \in [0, T]} \Lambda_t^{\tilde{\lambda}} \right) \right) \geq r$$

for every extremal lift $\tilde{\lambda}$ associated with $\tilde{u}$, then $\tilde{u}$ is not a weak-local minimizer in with respect to the topology of $\mathcal{U}$.

Thus, starting from a candidate minimizing control $\tilde{u}$ selected by PMP, one must determine its corank, construct the corresponding Jacobi curve and compute its Maslov index. Although these computations may be difficult in general, in the subsequent chapters we carry them out for several classes of extremals.

This chapter is organized as follows: Section 3.1 develops the abstract constrained-optimization framework, Section 3.2 computes the first differential and the Hessian of the endpoint map in Hamiltonian form, and Section 3.3 introduces the Jacobi curve through the $\mathcal{L}$ -derivative. Finally, Section 3.4 introduces the Maslov index and establishes its relation with the Morse index, completing the passage from the analytic to the geometric viewpoint.

3.1 Constrained optimization problems

In this first section, we recall some general facts and geometric objects from the theory of constrained optimization problems. More precisely, we are interested in the following class of problems. Let $\mathcal{U}$ be a smooth Banach manifold and M a finite dimensional smooth manifold. We consider the functional $J : \mathcal{U} \rightarrow \mathbb{R}$ and a function $F : \mathcal{U} \rightarrow M$. We assume that both J and F are smooth. Given a point $q_1 \in M$, we are interested in the minimization problem:

$$\min\{J(u) \mid u \in F^{-1}(q_1)\}. \quad (3.1.1)$$

In the case of an optimal control problem, $\mathcal{U}$ is the space of admissible controls, M is the state manifold, F is the endpoint map and J the minimized functional.

First, in Subsection 3.1.1 we show that optimal control problems are particular instance of constrained optimization problems on an infinite dimensional space. Then, in Subsections 3.1.2, 3.1.3, 3.1.4, we recall some concepts in the general abstract framework of constrained optimization that we will apply to OCPs in Section 3.3.

3.1.1 OCP as a constrained optimization problem

First, we want to reformulate an OCP as a constrained optimization problem, motivating the subsequent subsections.

Let M be a smooth manifold and consider the control system

$$\dot{q}(t) = f(q(t), u(t)), \quad q(0) = q_0, \quad (3.1.2)$$

together with the cost functional

$$J(u) = \int_0^T L(q(t; u), u(t)) dt. \quad (3.1.3)$$

Here, $\mathcal{U}$ denotes a Banach manifold of admissible controls and $q(\cdot; u)$ is the admissible trajectory corresponding to u .

Throughout this chapter, the interval on which a control is defined is regarded as part of the information carried by the control itself. Thus, if u is defined on $[0, T]$, the corresponding evolution is understood to stop at time T .

Definition 3.1.1 (Endpoint map). The endpoint map based at q_0 is the map

$$F : \mathcal{U} \longrightarrow M, \quad F(u) = q(T; u).$$

The fixed-endpoint constraint $q_u(T) = q_1$ can therefore be written as

$$u \in F^{-1}(q_1).$$

Definition 3.1.2 (Extended endpoint map). The extended endpoint map is the map

$$E : \mathcal{U} \longrightarrow M \times \mathbb{R}, \quad E(u) = (F(u), J(u)).$$

The extended endpoint map can itself be regarded as an ordinary endpoint map. Indeed, introduce the augmented state space

$$\widehat{M} := M \times \mathbb{R}$$

and the augmented control system

$$\begin{aligned} \dot{q}(t) &= f(t, q(t), u(t)), \\ \dot{z}(t) &= L(t, q(t), u(t)), \\ (q(0), z(0)) &= (q_0, 0). \end{aligned}$$

Its endpoint is

$$(q(T; u), z(T; u)) = (F(u), J(u)) = E(u).$$

Thus, the optimal control problem of minimizing (3.1.3) subject to (3.1.2) and $q(T; u) = q_1$ is rephrased as the following constrained optimization problem:

$$\min\{J(u) \mid u \in F^{-1}(q_1)\}.$$

3.1.2 First order condition: Lagrange multipliers

We recall the statement of the Lagrange multiplier rule. This is a classical consequence of the open mapping theorem and can be found, for instance, in [15, 5].

Proposition 3.1.3 (The Lagrange multipliers rule). *Suppose that $u \in \mathcal{U}$ is a critical point of the function $J|_{F^{-1}(q_1)}$. Then, there is $\lambda \in T_{q_1}^* M$ and $\nu \in \{0, -1\}$ such that*

$$\langle \lambda, D_u F[v] \rangle + \nu D_u J[v] = 0, \quad \forall v \in T_u \mathcal{U}. \quad (3.1.4)$$

The pair (λ, ν) is called a *Lagrange multiplier*. If $\nu = -1$, we say that (λ, ν) is a normal Lagrange multiplier. If instead $\nu = 0$, we say that (λ, ν) is an abnormal Lagrange multiplier. In the following, we consider only the normal case. Finally, if u is a critical

point of the function $J|_{F^{-1}(q_1)}$ and $\lambda \in T_{q_1}^*M$ is a normal Lagrange multiplier for u , we say that the pair (u, λ) is a *Lagrange point*.

Let $F^*(T^*M)$ be the pullback bundle of T^*M over $\mathcal{U}$, that is

$$F^*(T^*M) = \{(u, \lambda) \in T^*M \times \mathcal{U} \mid \lambda \in T_{F(u)}^*M\} \subset T^*M \times \mathcal{U}.$$

We define the map

$$\Psi : F^*(T^*M) \rightarrow T^*\mathcal{U}, \quad \Psi(u, \lambda) = \langle \lambda, D_u F \rangle - D_u J. \quad (3.1.5)$$

Definition 3.1.4 (set of normal Lagrange multipliers). We define

$$\mathcal{L} = \{\lambda \in T^*M \mid \text{there is } u \in \mathcal{U} \text{ such that } \langle \lambda, D_u F \rangle = D_u J\}.$$

The set $\mathcal{L}$ is called the *set of normal Lagrange multipliers*.

3.1.3 Second derivatives for constrained maps

We now introduce the second-order object associated with a constrained optimization problem at a Lagrange point. We first briefly recall the meaning of first and second derivatives for maps between smooth manifolds.

Let $F : \mathcal{U} \rightarrow M$ be a smooth map and let $u \in \mathcal{U}$. Its first differential at u is the intrinsically defined linear map

$$D_u F : T_u \mathcal{U} \longrightarrow T_{F(u)} M,$$

characterized by

$$D_u F[v] = \left. \frac{d}{ds} \right|_{s=0} F(\gamma(s)),$$

where γ is any smooth curve such that $\gamma(0) = u$ and $\dot{\gamma}(0) = v$.

Concerning second derivatives, in general, if the codomain is a manifold, there is no intrinsic definition of second differential in $T_{F(u)}M$. After choosing local coordinates, one can define a bilinear map $D_u^2 F$, but its value depends on the chosen coordinates. However, if $v, w \in \ker D_u F$, a change of coordinates modifies $D_u^2 F[v, w]$ only by an element of $\text{Im } D_u F$. Therefore

$$[D_u^2 F[v, w]] \in T_{F(u)}M / \text{Im } D_u F$$

is intrinsically defined for $v, w \in \ker D_u F$. For a real-valued function, the usual Hessian is intrinsically defined at a critical point. We refer to Chapter 20 in [13] for a more detailed discussion about the definitions of second order derivatives of a map with values in a manifold.

Now, following [6], let $u \in F^{-1}(q_1)$ be a critical point of $J|_{F^{-1}(q_1)}$ and let (λ, ν) be a Lagrange multiplier associated with u . In local coordinates, we define the Hessian of the extended endpoint map $E = (F, J)$ by

$$(\lambda, \nu) \text{ Hess}_u E[v, w] := \lambda D_u^2 F[v, w] + \nu D_u^2 J[v, w], \quad v, w \in \ker D_u F. \quad (3.1.6)$$

The Lagrange multiplier relation and the restriction to $\ker D_u F$ imply that the right-hand side of (3.1.6) is independent of the choice of local coordinates, and hence defines an intrinsic symmetric bilinear form on $\ker D_u F$.

In the rest of this chapter we only consider normal Lagrange multipliers, normalized by $\nu = -1$. We therefore omit ν from the notation and write

$$\lambda \text{Hess}_u E := (\lambda, -1) \text{Hess}_u E.$$

If F is a submersion at u , then $F^{-1}(q_1)$ is a smooth Banach submanifold near u and

$$T_u F^{-1}(q_1) = \ker D_u F.$$

In this regular case, the Hessian can be identified with the ordinary Hessian of the restriction of J to the constraint.

Proposition 3.1.5. *Let $J : \mathcal{U} \rightarrow \mathbb{R}$ and $F : \mathcal{U} \rightarrow M$ be smooth maps. Fix $q_1 \in M$ and let $u \in F^{-1}(q_1)$ be a critical point of $J|_{F^{-1}(q_1)}$. Assume that F is a submersion at u and let $\lambda \in T_{q_1}^* M$ be a normal Lagrange multiplier associated with u . Then*

$$\text{Hess}_u (J|_{F^{-1}(q_1)}) [v, w] = -\lambda \text{Hess}_u E[v, w] = D_u^2 J[v, w] - \lambda D_u^2 F[v, w]$$

for every $v, w \in \ker D_u F$.

Before concluding this part, we recall the following second-order necessary condition, which is a reformulation in the present setting of the local-openness criterion in [13, Theorem 20.3]. To state the result, we need the notion of Morse index.

Definition 3.1.6 (negative Morse index). Let V be a vector space and $Q : V \rightarrow \mathbb{R}$ a quadratic form on V . The negative Morse index of Q is

$$\text{ind}^- Q = \sup\{\dim L \mid L \text{ linear subspace of } V, \dim L < +\infty, Q|_{L \setminus \{0\}} < 0\}.$$

Theorem 3.1.7. *Let $F : \mathcal{U} \rightarrow M$ and $J : \mathcal{U} \rightarrow \mathbb{R}$ be smooth maps and let $u \in F^{-1}(q_1)$ be a critical point of $J|_{F^{-1}(q_1)}$. Set $r := \text{codim Im } D_u(F, J)$. If*

$$\text{ind}^- ((-\nu D_u^2 J[v, w] - \lambda D_u^2 F[v, w])|_{\ker D_u F}) \geq r,$$

for every Lagrange multiplier $(\lambda, \nu) \in T_{q_1}^* M \times \{0, -1\}$ associated with u , then E is locally open at u . In particular, u is not a local-weak minimizer with respect to the topology of $\mathcal{U}$.

The integer r defined above is called the *corank* of the control u .

3.1.4 $\mathcal{L}$ -derivative: definition and general properties

In the previous subsection, we saw that if a control $u \in \mathcal{U}$ is a critical point of the functional $J|_{F^{-1}(q_1)}$, then there is a Lagrange multiplier λ satisfying Equation (3.1.4). The aim of this subsection is to introduce the $\mathcal{L}$ -derivative, which is a Lagrangian subspace obtained by the linearization at a Lagrange point of the Lagrange multipliers equation.

From a geometrical point of view, suppose that, locally around the Lagrange point (u, λ) , the solutions of the Lagrange multiplier equation form a smooth manifold and that their projection onto T^*M is a local embedding, with image $\mathcal{L}$ near λ . Under these regularity assumptions, the $\mathcal{L}$ -derivative coincides with the tangent space

$$L(u, \lambda) = T_\lambda \mathcal{L}.$$

In general, however, the set $\mathcal{L}$ need not have such a regular structure, while the $\mathcal{L}$ -derivative is still intrinsically well defined. First, we give an intrinsic definition of the $\mathcal{L}$ -derivative, which does not rely on any choice of coordinates. Then, we give an explicit formula for it in local coordinates.

Let (u, λ) be a Lagrange point, so that $\Psi(u, \lambda) = 0_u$, where Ψ was defined in (3.1.5) and 0_u denotes the zero covector in $T_u^*\mathcal{U}$. Let

$$\mathbf{0}_\mathcal{U} := \{0_{u'} \in T^*\mathcal{U} \mid u' \in \mathcal{U}\}$$

be the zero section of $T^*\mathcal{U}$. The differential of Ψ at (u, λ) is a linear map

$$D_{(u, \lambda)} \Psi : T_{(u, \lambda)}(F^*(T^*M)) \longrightarrow T_{0_u}(T^*\mathcal{U}).$$

Let $\mathcal{W}$ be a finite-dimensional submanifold of $\mathcal{U}$ containing u . We denote by $F|_\mathcal{W} := F|_\mathcal{W}$ the restriction of F to $\mathcal{W}$, and use analogous notation for J and Ψ . We also denote by $\mathbf{0}_\mathcal{W} \subset T^*\mathcal{W}$ the zero section.

Definition 3.1.8 (Intrinsic $\mathcal{L}$ -derivative). Let $\mathcal{U}$ be a Banach manifold modelled on a separable Banach space B and let M be a finite-dimensional smooth manifold. Let $\mathcal{W}$ be a finite-dimensional submanifold of $\mathcal{U}$ containing u . We define

$$\Lambda_\mathcal{W} := \mathfrak{p} \left([D_{(u, \lambda)} \Psi|_\mathcal{W}]^{-1} (T_{0_u} \mathcal{Z}_\mathcal{W}) \right), \quad (3.1.7)$$

where

$$\mathfrak{p} : T_{(u, \lambda)}(F_\mathcal{W}^*(T^*M)) \longrightarrow T_\lambda(T^*M)$$

is the differential of the canonical projection

$$F_\mathcal{W}^*(T^*M) \ni (u', \lambda') \longmapsto \lambda' \in T^*M.$$

The space $\Lambda_\mathcal{W}$ is the $\mathcal{L}$ -derivative of the problem (3.1.1) at the Lagrange point (u, λ) restricted to the submanifold $\mathcal{W}$.

Now, take any sequence of finite-dimensional submanifolds $(\mathcal{W}_j)_{j \in \mathbb{N}}$ of $\mathcal{U}$ such that $\mathcal{W}_j \subset \mathcal{W}_{j+1}$ and

$$\overline{\bigcup_{j=1}^{\infty} \mathcal{W}_j} = \mathcal{U}.$$

We define

$$\Lambda := \lim_{j \rightarrow +\infty} \Lambda_{\mathcal{W}_j},$$

provided that the limit exists.

The use of the tangent space to the zero section in (3.1.7) is motivated by the fact that, if we consider a smooth family of Lagrange points (u_s, λ_s) , one has $\Psi(u_s, \lambda_s) = 0_{u_s}$, and therefore its derivative with respect to the parameter s is a vector tangent to the zero section. Consequently, the intrinsic definition above gives the linearization of the Lagrange multiplier equation.

This construction is intrinsic, since it is expressed entirely in terms of bundle maps, their differentials, and the zero section, and therefore does not depend on any choice of coordinates.

Now, we want to give a more explicit description of the $\mathcal{L}$ -derivative using local coordinates. Fix any two local charts $\mathcal{U} \rightarrow B$ and $M \rightarrow \mathbb{R}^d$ in a neighbourhood of u and $F(u)$ respectively. If we linearize (3.1.4) with respect to u and λ , we obtain

$$\langle \xi, D_u F[w] \rangle + \lambda D_u^2 F[v, w] - D_u^2 J[v, w] = 0, \quad D_u F[v] = x, \quad v, w \in T_u \mathcal{U}. \quad (3.1.8)$$

Definition 3.1.9 ($\mathcal{L}$ -derivative with local coordinates). Let $\mathcal{U}$ be a Banach manifold modelled on a separable Banach space and M a finite dimensional smooth manifold. Let $F : \mathcal{U} \rightarrow M$ and $J : \mathcal{U} \rightarrow \mathbb{R}$ be two smooth maps. Let $u \in F^{-1}(q_1)$ be a critical point of $J|_{F^{-1}(q_1)}$ and let $\lambda \in T_{q_1}^* M$ be its Lagrange multiplier. Take a finite dimensional subspace $V \subset T_u \mathcal{U}$. The $\mathcal{L}$ -derivative of (F, J) at (u, λ) over the subspace V is the vector space

$$L(u, \lambda)(V) = \{(\xi, x) \in T_\lambda(T^*M) \mid \exists v \in V, (\xi, v) \text{ satisfy (3.1.8)} \forall w \in V\}.$$

Moreover, let $(V_k)_{k \in \mathbb{N}}$ be an increasing sequence of finite dimensional subspaces of $T_u \mathcal{U}$, that is every V_k is finite dimensional, they satisfy $V_k \subset V_{k+1}$ and $\bigcup_{k \in \mathbb{N}} V_k = T_u \mathcal{U}$. We define the $\mathcal{L}$ -derivative of (F, J) over $T_u \mathcal{U}$ as

$$L(u, \lambda) := \lim_{k \rightarrow +\infty} L(u, \lambda)(V_k), \quad (3.1.9)$$

where the limit in the right-hand side is with respect to the topology of the Grassmannian of $T_\lambda(T^*M)$. One can check that the limit does not depend on the choice of the sequence $(V_k)_{k \in \mathbb{N}}$.

Remark 3.1.10. We recall some properties of the $\mathcal{L}$ -derivative, which are proved, for instance, in [6] and [15].

1. If $\dim V < +\infty$, then the space $L(u, \lambda)(V)$ is a Lagrangian subspace of $T_\lambda(T^*M)$, that is $\dim L(u, \lambda)(V) = d$ and, denoting by σ the standard symplectic form, we have $\sigma_\lambda|_{L(u, \lambda)(V)} = 0$.
2. It is necessary to define $L(u, \lambda)$ as a limit of the finite dimensional approximations $L(u, \lambda)(V_k)$ because Equation (3.1.8) might be ill-posed on infinite dimensional spaces and the resulting space might have dimension strictly less than $\dim M$.
3. The limit in (3.1.9) might not exist in general. However, if $D_u^2 J|_{F^{-1}(q_1)}$ has finite negative Morse index, then one can prove that the limit exists.
4. If $V \subset T_u \mathcal{U}$ is a dense subspace, then $L(u, \lambda) = L(u, \lambda)(V)$.

3.2 Endpoint map, its differential and Hessian in optimal control

The purpose of this section is to introduce the symplectic objects that are necessary to adapt the previous ideas from constrained optimization theory to the geometric control theory setting. To this aim, first we introduce the endpoint map and its cotangent lift. Then, we need to transport back the entire dynamics to the fixed initial symplectic vector space $T_{\lambda_0}(T^*M)$. This way, all the relevant quantities that we are going to introduce are defined in the same space, rather than in the moving space $T_{\lambda(t)}(T^*M)$. This fact allows us to compute integral and derivatives without invoking complicate non-linear objects. The same purpose could be achieved by introducing a local trivialization of the cotangent bundle T^*M in a neighbourhood of the reference extremal $\tilde{\lambda}(t)$. The advantage of our approach is that it is coordinate free.

Throughout this section, we assume that the range of $\tilde{u}$ is contained in a smooth finite-dimensional manifold $U_0 \subset U$ and that the dynamics f and the running cost L are smooth with respect to the state and control variables in a neighbourhood of the reference pair. We also assume that the admissible-control space $\mathcal{U}$ is a smooth Banach manifold with tangent space naturally identified with a space of essentially bounded measurable functions

$$v(t) \in T_{\tilde{u}(t)} U_0, \quad \text{for a.e. } t \in [0, T].$$

3.2.1 Hamiltonian lifts

For $u \in U$, define the Hamiltonian associated with the dynamics by

$$h(\lambda, u) := \langle \lambda, f(\pi(\lambda), u) \rangle, \quad \lambda \in T^*M.$$

Its Hamiltonian vector field projects onto the controlled vector field:

$$d_\lambda \pi(\vec{h}(\lambda, u)) = f(\pi(\lambda), u).$$

Hence, for every $\ell_0 \in T_{q_0}^*M$, the map

$$\mathcal{F}_{\ell_0}(u) := \overrightarrow{\exp} \int_0^T \vec{h}(\cdot, u(t)) dt \circ \ell_0$$

is a Hamiltonian lift of the endpoint map F defined in (3.1.1), in the sense that

$$\pi \circ \mathcal{F}_{\ell_0} = F.$$

The extended endpoint map has a canonical lift to $T^*\widehat{M}$. Identifying a covector in $T_{(q,z)}^*\widehat{M}$ with a pair

$$(\lambda, \nu) \in T_q^*M \times \mathbb{R},$$

the Hamiltonian of the augmented system is

$$\widehat{h}(\lambda, \nu, q, z, u) = \langle \lambda, f(q, u) \rangle + \nu L(t, q, u).$$

Since $\widehat{h}$ does not depend on z , the component ν is constant along the corresponding Hamiltonian flow. In this thesis, we are mainly interested in the case $\nu = -1$, thus from now on we assume $\nu = -1$ and, to simplify the notation, we drop all the superscript ν . The dynamics on $T^*\widehat{M}$ therefore reduces to the Hamiltonian system on T^*M generated by

$$h(\lambda, u) := \langle \lambda, f(\pi(\lambda), u) \rangle - L(t, \pi(\lambda), u).$$

For $\ell_0 \in T_{q_0}^*M$, we denote the corresponding reduced Hamiltonian lift by

$$\mathcal{E}_{\ell_0}(u) := \overrightarrow{\exp} \int_0^T \vec{h}(\cdot, u(t)) dt \circ \ell_0.$$

Its projection onto M is again the endpoint of the state trajectory:

$$\pi \circ \mathcal{E}_{\ell_0} = F.$$

3.2.2 Transported-back dynamics

Fix a reference control $\tilde{u} \in \mathcal{U}$. Let

$$\widetilde{\Phi}_{s,t} : T^*M \longrightarrow T^*M$$

be the non-autonomous Hamiltonian flow generated by

$$\vec{h}(\cdot, \tilde{u}(t)).$$

Starting from $\tilde{\lambda}_0 \in T_{q_0}^*M$, we write

$$\tilde{\lambda}(t) = \widetilde{\Phi}_{0,t}(\tilde{\lambda}_0).$$

The transported-back Hamiltonian is defined by

$$\mathcal{H}_t(u, \ell) := (h(\cdot, u) - h(\cdot, \tilde{u}(t))) \circ \tilde{\Phi}_{0,t}(\ell).$$

By the variation formula of chronological calculus,

$$\mathcal{E}_{\tilde{\lambda}_0}(u) = \tilde{\Phi}_{0,T} \circ \overrightarrow{\exp} \int_0^T \vec{\mathcal{H}}_t(u(t), \cdot) dt \circ \tilde{\lambda}_0.$$

It is therefore convenient to introduce the transported-back lifted endpoint map

$$\mathcal{G} := (\Phi_T)^{-1} \circ \mathcal{E}_{\tilde{\lambda}_0}.$$

It satisfies

$$\mathcal{G}(u) = \overrightarrow{\exp} \int_0^T \vec{\mathcal{H}}_t(u(t), \cdot) dt \circ \tilde{\lambda}_0,$$

and, in particular,

$$\mathcal{G}(\tilde{u}) = \tilde{\lambda}_0.$$

The advantage of this representation is that all infinitesimal variations are transported to the fixed symplectic space $T_{\tilde{\lambda}_0}(T^*M)$.

3.2.3 First differential

In this subsection, we compute the first differential of the maps F and E . In particular, we will see how PMP can be reinterpreted using the Lagrange multiplier rule.

For every $t \in [0, T]$, define the linear map

$$Z_t : T_{\tilde{u}(t)}U \longrightarrow T_{\tilde{\lambda}_0}(T^*M)$$

by

$$Z_tv := \left. \frac{\partial \vec{\mathcal{H}}_t}{\partial u}(u, \tilde{\lambda}_0)[v] \right|_{u=\tilde{u}(t)} = (\Phi_t)_*^{-1} \left(\left. \frac{\partial \vec{h}}{\partial u}(\tilde{\lambda}(t), u)[v] \right|_{u=\tilde{u}(t)} \right). \quad (3.2.1)$$

The map Z_t describes how an infinitesimal variation of the control acts on the dynamics. A variation of the control naturally produces a tangent vector at $\lambda(t)$. Thus, the resulting map Z_t is the variation-to-tangent map written in a fixed symplectic frame along the reference extremal. Since we have to compute integrals and time derivatives of Z_t in the following, we transport back the dynamics by applying the transformation $(\Phi_t^{-1})_*$. This way, the image of Z_t is contained in the fixed vector space $T_{\tilde{\lambda}_0}(T^*M)$.

The first term in the chronological expansion of $\mathcal{G}$ gives

$$D_{\tilde{u}}\mathcal{G}[v] = \int_0^T Z_tv(t) dt.$$

Consequently,

$$D_{\tilde{u}}\mathcal{E}_{\tilde{\lambda}_0}[v] = (\Phi_T)_* \left(\int_0^T Z_t v(t) dt \right).$$

Taking $\nu = 0$ and projecting onto the base manifold yields the first differential of F :

$$D_{\tilde{u}}F[v] = d_{\tilde{\lambda}_T}\pi \circ (\Phi_T)_* \left(\int_0^T Z_t v(t) dt \right).$$

Let

$$\Pi_{\tilde{\lambda}_0} := T_{\tilde{\lambda}_0}^*(T_{q_0}^*M)$$

be the vertical subspace at $\tilde{\lambda}_0$. Since the flow of a fibrewise linear Hamiltonian is the cotangent lift of the corresponding flow on M , it maps vertical spaces into vertical spaces. It follows that

$$v \in \ker D_{\tilde{u}}F \iff \int_0^T Z_t v(t) dt \in \Pi_{\tilde{\lambda}_0}.$$

An analogous formula holds for the extended endpoint map on the augmented cotangent bundle. Denoting by $\widehat{\Phi}_t$, $\widehat{Z}_t$, and $\widehat{\Pi}$ the corresponding objects on $T^*\widehat{M}$, one has

$$D_{\tilde{u}}E[v] = d\widehat{\pi} \circ (\widehat{\Phi}_T)_* \left(\int_0^T \widehat{Z}_t v(t) dt \right),$$

and therefore

$$v \in \ker D_{\tilde{u}}E \iff \int_0^T \widehat{Z}_t v(t) dt \in \widehat{\Pi}.$$

The first differential of E also admits a convenient scalar Hamiltonian representation. Let

$$(\tilde{\lambda}_T, \nu) \in T_{E(\tilde{u})}^*(M \times \mathbb{R})$$

and let $\tilde{\lambda}$ be the corresponding solution of the Hamiltonian system generated by $h(\cdot, \tilde{u}(t))$. Then

$$\begin{aligned} \langle (\tilde{\lambda}_T, \nu), D_{\tilde{u}}E[v] \rangle &= \langle \tilde{\lambda}_T, D_{\tilde{u}}F[v] \rangle + \nu D_{\tilde{u}}J[v] \\ &= \int_0^T \frac{\partial h}{\partial u}(\tilde{\lambda}(t), u)[v(t)] \Big|_{u=\tilde{u}(t)} dt. \end{aligned}$$

Equivalently, in transported-back coordinates, the latter equation reads

$$\langle (\tilde{\lambda}_T, \nu), D_{\tilde{u}}E[v] \rangle = \int_0^T \frac{\partial \mathcal{H}_t}{\partial u}(u, \tilde{\lambda}_0)[v(t)] \Big|_{u=\tilde{u}(t)} dt.$$

In particular,

$$(\tilde{\lambda}_T, \nu) \in (\text{im } D_{\tilde{u}}E)^\perp$$

if and only if

$$\left. \frac{\partial h}{\partial u}(\tilde{\lambda}(t), u) \right|_{u=\tilde{u}(t)} = 0 \quad (3.2.2)$$

for almost every t , whenever arbitrary localized control variations are admissible.

From this discussion, we can see the connection between PMP and the first order differential of the extended endpoint map E . Indeed, there exists a non-zero covector

$$(\tilde{\lambda}_T, \nu) \in (\text{im } D_{\tilde{u}}E)^\perp,$$

says that the endpoint map $E = (F, J)$ cannot be a submersion at $\tilde{u}$. This condition is equivalent to

$$D_{\tilde{u}}(J|_{F^{-1}(q_1)}) = 0 \iff \tilde{\lambda}_T D_{\tilde{u}}F + \nu D_{\tilde{u}}J = 0.$$

Defining $\tilde{\lambda}_t := (\tilde{\Phi}_{t,T})^{-1}(\tilde{\lambda}_T)$, we see that the corresponding covector curve $t \mapsto \tilde{\lambda}_t$ satisfies the adjoint Hamiltonian equation of PMP, while (3.2.2) correspond to the maximization principle. Notice that actually PMP say a bit more than what is done here: the control $\tilde{u}$ has to be a maximizer of h .

3.2.4 Second differential and Hessian

In this subsection, we recall the formula for the Hessian of the maps F and E , which will be crucial to compute the $\mathcal{L}$ -derivatives at a Lagrange point.

The second differential of the transported-back lifted endpoint map $\mathcal{G}$ can be obtained from the second term of the chronological expansion. To state it, let

$$\mathcal{Z}_t[v] := \left. \frac{\partial \vec{\mathcal{H}}_t}{\partial u}(u, \cdot)[v] \right|_{u=\tilde{u}(t)}$$

be the corresponding vector field near $\tilde{\lambda}_0$, so that

$$Z_t v = \mathcal{Z}_t[v](\tilde{\lambda}_0).$$

For every smooth function a on T^*M , the second differential of $a \circ \mathcal{G}$ is

$$\begin{aligned} D_{\tilde{u}}^2(a \circ \mathcal{G})[v, w] &= \int_0^T \left(\left. \frac{\partial^2 \vec{\mathcal{H}}_t}{\partial u^2}[v(t), w(t)] \right|_{u=\tilde{u}(t)} a \right) (\tilde{\lambda}_0) dt + \\ &+ \int_0^T \int_0^t \left(\mathcal{Z}_s[v(s)] \circ \mathcal{Z}_t[w(t)] + \mathcal{Z}_s[w(s)] \circ \mathcal{Z}_t[v(t)] \right) a(\tilde{\lambda}_0) ds dt. \end{aligned} \quad (3.2.3)$$

For the applications to second-order optimality conditions, it is more convenient to express the scalarized Hessian directly in Hamiltonian form. Let

$$(\lambda_T, \nu) \in (\text{im } D_{\tilde{u}}E)^\perp$$

and let $\tilde{\lambda}_t = \Phi_t(\tilde{\lambda}_0)$ be the associated extremal lift.

Following the convention introduced in Section 3.1, we define

$$Q(v, w) := -\nu D_{\tilde{u}}^2 J[v, w] - \tilde{\lambda}_T D_{\tilde{u}}^2 F[v, w], \quad v, w \in T_{\tilde{u}}\mathcal{U}.$$

With this notation, the linearized Lagrange multiplier equation becomes

$$\langle \xi, D_{\tilde{u}} F[w] \rangle - Q(v, w) = 0, \quad D_{\tilde{u}} F[v] = x, \quad v, w \in T_{\tilde{u}}\mathcal{U}.$$

By Equation (3.2.3), the bilinear form Q can be computed using the following formula:

$$\begin{aligned} Q(v, w) = & - \int_0^T \frac{\partial^2 \mathcal{H}_t}{\partial u^2}(\tilde{\lambda}(t), u)[v(t), w(t)] \Big|_{u=\tilde{u}(t)} dt \\ & - \frac{1}{2} \int_0^T \int_0^t \sigma_{\tilde{\lambda}_0}(Z_s v(s), Z_t w(t)) ds dt \\ & - \frac{1}{2} \int_0^T \int_0^t \sigma_{\tilde{\lambda}_0}(Z_s w(s), Z_t v(t)) ds dt. \end{aligned}$$

In particular, its associated quadratic form is

$$\begin{aligned} Q(v) = & - \int_0^T \frac{\partial^2 \mathcal{H}_t}{\partial u^2}(\tilde{\lambda}(t), u)[v(t), v(t)] \Big|_{u=\tilde{u}(t)} dt \\ & - \int_0^T \sigma_{\tilde{\lambda}_0} \left(\int_0^t Z_s v(s) ds, Z_t v(t) \right) dt. \end{aligned} \tag{3.2.4}$$

3.3 Jacobi curve and Jacobi equation

3.3.1 Jacobi curve

In this subsection, we are going to apply the theory of $\mathcal{L}$ -derivatives, which is discussed in Subsection 3.1.4, to optimal control problems. This leads to the notion of Jacobi curve, which is a generalization of the classical concept of Jacobi field from Riemannian geometry to a non-smooth setting.

We use the notion of $\mathcal{L}$ -derivative only in Definition 3.3.1, to highlight the connection with the second differential of the endpoint map and to justify some of its properties that would otherwise not be straightforward to prove.

For brevity, we introduce some notation: we let $\mathcal{V} = T_{\tilde{u}}\mathcal{U}$ be the space of admissible variations and

$$\begin{aligned} \mathcal{V}_\tau &:= \{w \in \mathcal{V} \mid \text{supp } w \subset [0, \tau]\}, \\ \mathcal{V}_{\tau_1, \tau_2} &:= \{w \in \mathcal{V} \mid \text{supp } w \subset [\tau_1, \tau_2]\}, \end{aligned}$$

and $\mathcal{K} := \ker D_{\tilde{u}}\mathcal{G}$, $\mathcal{K}_\tau = \mathcal{K} \cap \mathcal{V}_\tau$ and $\mathcal{K}_{\tau_1, \tau_2} = \mathcal{V}_{\tau_1, \tau_2} \cap \mathcal{K}_{\tau_2}$.

Definition 3.3.1 (Jacobi curve). Let $(\tilde{u}, \tilde{\lambda})$ be an extremal pair for the optimal control problem (2.3.2), (2.3.3) defined on the time interval $[0, T]$. Let E be the extended endpoint map as defined in (3.1.2). Let $L(u, \lambda)$ denote the $\mathcal{L}$ -derivative of E at the Lagrange point (u, λ) , see Definition 3.1.9. The *Jacobi curve* of $\tilde{u}, \tilde{\lambda}$ is the curve $\Lambda : [0, T] \rightarrow \mathfrak{L}(T_{\lambda_0}(T^*M))$ defined by

$$\Lambda_t := (\tilde{\Phi}_{0,t})_*^{-1} L(\tilde{u}|_{[0,t]}, \tilde{\lambda}(t)).$$

Moreover, given a subspace $V \subset \mathcal{V}$, we denote by $\Lambda_t(V) := (\tilde{\Phi}_{0,t})_*^{-1} L(\tilde{u}, \tilde{\lambda}(t))(V)$.

Since throughout the thesis we consider only the Jacobi curve of a fixed reference extremal pair $(\tilde{u}, \tilde{\lambda})$, we do not need to specify the dependence of Λ_t on $(\tilde{u}, \tilde{\lambda})$.

To justify the relation between the Jacobi curve and the Morse index of $\text{Hess}_{\tilde{u}} J|_{F^{-1}(q_1)}$, we recall the following result from [15] (see Lemma 1.2 in Chapter 1). This result shows that there is a deep link between the kernel of $\text{Hess}_{\tilde{u}} J|_{F^{-1}(q_1)}$ and the intersections of the Jacobi curve with the vertical fibre. This serve as motivation for introducing the Maslov index in the subsequent section, where we will recall a much stronger result.

Lemma 3.3.2. *Assume that $\tilde{u}$ is a regular point of F and that $\ker d_{\tilde{u}}F \cap \ker Q = \{0\}$. Then*

$$\ker \text{Hess}_{\tilde{u}} (J|_{F^{-1}(\tilde{q}(t))}) = \{0\} \iff \Lambda_t \cap \Pi_{\tilde{\lambda}_0} = \{0\}.$$

Equivalently, the constrained Hessian is non-degenerate if and only if Λ_t is transverse to the vertical $\Pi_{\tilde{\lambda}_0}$.

Remark 3.3.3. If $(\tilde{u}, \tilde{\lambda})$ is an extremal pair and $\tilde{\lambda}$ is a strictly normal extremal, i.e. $\tilde{u}$ admits no non-trivial abnormal multiplier, then $\tilde{u}$ is a regular point of F . Moreover, $\tilde{u}$ is a regular point of F and $(\xi, x) \in \Lambda_t$ satisfy the definition of $\mathcal{L}$ -derivative with $v \in \ker D_{\tilde{u}}F \cap \ker Q$, then $(\xi, x) = 0$. Indeed, in this case we have $x = D_{\tilde{u}}F[v] = 0$ and ξ must satisfy $\xi D_{\tilde{u}}F[w] = 0$ for every $w \in T_{\tilde{u}}\mathcal{U}$, which implies $\xi = 0$ since $D_{\tilde{u}}F$ is surjective.

Proof. Suppose that $\ker \text{Hess}_{\tilde{u}} (J|_{F^{-1}(q(t))}) = \{0\}$, we want to prove $\Lambda_t \cap \Pi_{\tilde{\lambda}_0} = \{0\}$. Let $\eta \in \Lambda_t \cap \Pi_{\tilde{\lambda}_0}$, $\eta = (\xi, x)$. Then $x = 0$ and from the definition of the $\mathcal{L}$ -derivative, there is a vector $v \in T_{\tilde{u}}\mathcal{U}$ such that

$$d_{\tilde{u}}F[v] = 0$$

and

$$\langle \xi, d_{\tilde{u}}F[w] \rangle - Q(v, w) = 0 \quad \text{for every } w \in T_{\tilde{u}}\mathcal{U}.$$

Taking $w \in \ker d_{\tilde{u}}F$ gives

$$Q(v, w) = 0 \quad \text{for every } w \in \ker d_{\tilde{u}}F,$$

which means that $v \in \ker \text{Hess}_{\tilde{u}} (J|_{F^{-1}(q_1)})$. So, this implies $v = 0$, and hence ξ must satisfy $\langle \xi, D_{\tilde{u}}F[w] \rangle = 0$ for all $w \in T_{\tilde{u}}\mathcal{U}$. But this forces $\xi = 0$ since $D_{\tilde{u}}F$ is surjective.

Thus, we have proved that $(\xi, x) \in \Lambda_t \cap \Pi_{\tilde{\lambda}_0}$ implies $(\xi, x) = (0, 0)$. We point out that the hypothesis $\ker d_{\tilde{u}}F \cap \ker Q = \{0\}$ was not used to prove this implication.

Conversely, let

$$v \in \ker \text{Hess}_{\tilde{u}} \left(J|_{F^{-1}(\tilde{q}(t))} \right), \quad v \neq 0.$$

Then $v \in \ker d_{\tilde{u}}F$ and

$$Q(v, w) = 0 \quad \text{for every } w \in \ker d_{\tilde{u}}F.$$

Let us introduce the splitting $T_{\tilde{u}}\mathcal{U} = V \oplus \ker D_{\tilde{u}}F$, where V is any linear subspace of $T_{\tilde{u}}\mathcal{U}$ spanning $\text{im} D_{\tilde{u}}F$ of minimal dimension. Then, writing $w = w_0 + w_1$, $w \in T_{\tilde{u}}\mathcal{U}$, $w_0 \in \ker D_{\tilde{u}}F$, $w_1 \in V$, the equation defining the $\mathcal{L}$ -derivative reduces to the finite dimensional equation

$$\langle \xi, d_{\tilde{u}}F[w_1] \rangle - Q(v, w_1) = 0 \quad \text{for every } w_1 \in V.$$

Notice that $D_{\tilde{u}}F : V \rightarrow \text{im} D_{\tilde{u}}F$ is an isomorphism, hence also $(D_{\tilde{u}}F^\top)^{-1}$ is an isomorphism. The assumption $\ker d_{\tilde{u}}F \cap \ker Q = \{0\}$ yields that $Q(v, \cdot) \in V^*$ is non-zero, so $\xi := (D_{\tilde{u}}F^\top)^{-1}Q(v, \cdot)$ is also non-zero and satisfies the previous equation. This shows that $\eta = (\xi, 0) \in \Lambda_t \cap \Pi_{\tilde{\lambda}_0}$ and $\eta \neq 0$, proving the second implication. Notice that, in the proof of the second implication, we have not used that $\tilde{u}$ is a regular point of F . $\square$

Remark 3.3.4. The previous proof shows that if $\tilde{u}$ is a critical point of F , that is, in presence of non-trivial abnormal extremals, the intersection $\Lambda_t \cap \Pi_{\tilde{\lambda}_0}$ is non-trivial even if $\ker \text{Hess}_{\tilde{u}} \left(J|_{F^{-1}(q(t))} \right) = \{0\}$. On the other hand, $\ker \text{Hess}_{\tilde{u}} \left(J|_{F^{-1}(q(t))} \right) \neq \{0\}$ yields a non-trivial intersection $\Lambda_t \cap \Pi_{\tilde{\lambda}_0}$ also when $\tilde{u}$ is a critical point of F , provided that $\ker d_{\tilde{u}}F \cap \ker Q = \{0\}$. Thus, if abnormal extremals are present, one should separate carefully intersections coming from the fact that the extremal is abnormal from intersections due to the degeneracy of $\text{Hess}_{\tilde{u}} \left(J|_{F^{-1}(q(t))} \right)$.

We give the following alternative characterization, which is much more practical and can be used as working definition for the rest of this Section.

Proposition 3.3.5 (Jacobi curve, working definition). *Let $(\tilde{u}, \tilde{\lambda})$ be an extremal pair for the optimal control problem (2.3.2), (2.3.3). Let $V \subseteq \mathcal{V}_t$ be a linear subspace. Then, the Jacobi curve $\Lambda_t(V)$ is the vector space generated by the vectors of the form*

$$\eta_0 + \int_0^t Z_s w_1(s) ds,$$

where $\eta_0 \in T_{q_0}^*M$, $w_1 \in V$ satisfy the equation

$$Q(w_1, w_2) = \int_0^t \sigma(\eta_0, Z_s w_2(s)) ds \tag{3.3.1}$$

for all $w_2 \in V$.

For a proof, see [6]. This characterization allows us to compute the Jacobi curve by means of finite-dimensional approximations. The following Lemma was proved in [6]. It allows us to compute $\Lambda_{t+\varepsilon}$ from Λ_t .

Lemma 3.3.6. *Take any $\tau_1, \tau_2 \in (0, T)$, with $\tau_1 < \tau_2$, and suppose that the negative Morse index of the quadratic form $Q_{\tau_2}|_{\mathcal{K}_{\tau_2}}$ is finite. We denote by $V_2 \subset \mathcal{V}_T$ a subspace of admissible variations supported on (τ_1, τ_2) , that is $V_2 \subseteq \mathcal{V}_{\tau_1, \tau_2}$. Then, $\Lambda_{\tau_2}(\mathcal{V}_{\tau_1} \oplus V_2)$ is the Lagrangian subspace generated by vectors*

$$\eta + \int_{\tau_1}^{\tau_2} Z_t w_1(t) dt, \quad (3.3.2)$$

where $\eta \in \Lambda_{\tau_1}$ and $w_1 \in V_2$ satisfy

$$Q_{\tau_1, \tau_2}(w_1, w_2) = \int_{\tau_1}^{\tau_2} \sigma(\eta, Z_t w_2(t)) dt, \quad (3.3.3)$$

for all variations $w_2 \in V_2$. In particular, Λ_{τ_2} is generated by vectors as in (3.3.2) where $\eta \in \Lambda_{\tau_1}$ and $w_1 \in \mathcal{V}_{\tau_1, \tau_2}$ satisfy (3.3.3) for every $w_2 \in \mathcal{V}_{\tau_1, \tau_2}$.

Remark 3.3.7. Notice that this is possible thanks to the fact that the Jacobi curve Λ_t is computed using the quadratic form Q_t on all variations $\mathcal{V}_t$ and not just the Hessian of G_t , which is defined on the kernel of the first variation $\mathcal{K}_t$. Indeed, while $\mathcal{V}_{\tau_2} = \mathcal{V}_{\tau_1} \oplus \mathcal{V}_{\tau_1, \tau_2}$, in general it does not hold that $\mathcal{K}_{\tau_2} = \mathcal{K}_{\tau_1} \oplus \mathcal{K}_{\tau_1, \tau_2}$. This is the reason why we need to consider the $\mathcal{L}$ -derivatives and the form Q_t defined on the whole space $\mathcal{V}_t$ and not simply on $\ker DG_t$.

3.3.2 Regular extremal arcs and Jacobi equation

In this subsection we recall that for a normal regular extremal arcs, i.e. under the strong Legendre condition, the Jacobi curve is a smooth curve in the Lagrangian Grassmannian and that it can be obtained as by solving an ODE, which goes under the name of Jacobi equation. Since the same strategies for a regular extremal arc are reused and generalized in Chapters 5,7,8, we recall them in this simpler case as guiding principles for more complex extremals.

Let $(\tilde{u}, \tilde{\lambda})$ be a regular extremal pair on an interval $[0, T]$. The Jacobi curve Λ_t associated with the extremal arc $\tilde{\lambda}$ is the curve of Lagrangian subspaces of $T_{\tilde{\lambda}_0}^*(T^*M)$ introduced in Definition 3.3.1.

We now give an equivalent description of Λ_t in terms of a linear Hamiltonian equation. Recall the definition of

$$Z_t v := \left. \frac{\partial \vec{\mathcal{H}}_t}{\partial u}(u, \tilde{\lambda}_0)[v] \right|_{u=\tilde{u}(t)},$$

which was given in (3.2.1), and of the Legendre form

$$\mathbb{L}_t := \left. \frac{\partial^2 h_u}{\partial u^2} \right|_{u=\tilde{u}(t), \lambda=\tilde{\lambda}_t},$$

which was given in (2.4.2). Since the extremal arc is regular, $\mathbb{L}_t$ is non-degenerate for every $t \in [0, T]$. We identify the quadratic form $\mathbb{L}_t : T_{\tilde{u}}U \times T_{\tilde{u}}U \rightarrow \mathbb{R}$ with the linear map $\mathbb{L}_t : T_{\tilde{u}}U \rightarrow (T_{\tilde{u}}U)^*$, which is invertible as well and we denote its inverse by $\mathbb{L}^{-1}$.

Definition 3.3.8. The *Jacobi equation* along the regular extremal arc $(\tilde{u}, \tilde{\lambda})$ is the linear non-autonomous Hamiltonian equation on $T_{\tilde{\lambda}_0}(T^*M)$

$$\dot{\eta}(t) = Z_t \mathbb{L}_t^{-1} \sigma_{\tilde{\lambda}_0}(Z_t \cdot, \eta(t)). \quad (3.3.4)$$

The Jacobi equation is a linear Hamiltonian equation, whose Hamiltonian function is given by the bilinear function

$$b_t(\eta) := -\frac{1}{2} \left\langle \sigma_{\tilde{\lambda}_0}(Z_t \cdot, \eta), \mathbb{L}_t^{-1} \sigma_{\tilde{\lambda}_0}(Z_t \cdot, \eta) \right\rangle, \quad \eta \in T_{\tilde{\lambda}_0}(T^*M).$$

Following this notation, the Jacobi equation can be shortened as $\dot{\eta} = \vec{b}_t(\eta)$ and we denote by

$$B_t = \overrightarrow{\exp} \int_0^t \vec{b}_s ds,$$

the flow of $\vec{b}_t$. Notice that, by strong Legendre condition, $\mathbb{L}_t$ is negative definite, thus b_t is positive definite as well.

Theorem 3.3.9. *Let $(\tilde{u}, \tilde{\lambda})$ be a regular extremal pair on an interval $[0, T]$ for problem (2.3.2), (2.3.3) with fixed final time. Then*

$$\Lambda_t = B_t \Pi,$$

that is, the Jacobi curve is generated by the flow of the Jacobi equation starting from the vertical space.

Sketch of proof. Fix $t \in [0, T]$ and abbreviate $\sigma_{\tilde{\lambda}_0}$ by σ . By the working characterization of the Jacobi curve, a vector η_t belongs to Λ_t if and only if there exist $\eta_0 \in \Pi_0$ and $v \in \mathcal{V}_t$ such that

$$\eta_\tau = \eta_0 + \int_0^\tau Z_s v(s) ds, \quad \tau \in [0, t],$$

and

$$Q_t(v, w) = \int_0^t \sigma(\eta_0, Z_\tau w(\tau)) d\tau \quad \text{for every } w \in \mathcal{V}_t. \quad (3.3.5)$$

Substituting the explicit expression of the second variation into (3.3.5), and rearranging the integral terms, gives

$$-\int_0^t [\mathbb{L}_\tau(v(\tau), w(\tau)) + \sigma(\eta_\tau, Z_\tau w(\tau))] d\tau = 0 \quad \text{for every } w \in \mathcal{V}_t.$$

Since the test variation w is arbitrary and can be localized in time, it follows that, for almost every $\tau \in [0, t]$,

$$-\mathbb{L}_\tau(v(\tau), \cdot) - \sigma(\eta_\tau, Z_\tau \cdot) = 0. \quad (3.3.6)$$

The strong Legendre condition implies that $\mathbb{L}_\tau$ is non-degenerate. Hence (3.3.6) uniquely determines the control variation:

$$v(\tau) = \mathbb{L}_\tau^{-1} \sigma(Z_\tau \cdot, \eta_\tau). \quad (3.3.7)$$

Differentiating the definition of η_τ and using (3.3.7), we obtain

$$\dot{\eta}_\tau = Z_\tau v(\tau) = Z_\tau \mathbb{L}_\tau^{-1} \sigma(Z_\tau \cdot, \eta_\tau),$$

so η is a solution of the Jacobi equation with $\eta(0) = \eta_0 \in \Pi_0$.

Conversely, let η solve the Jacobi equation and assume that $\eta(0) \in \Pi_0$. Define

$$v(\tau) := \mathbb{L}_\tau^{-1} \sigma(Z_\tau \cdot, \eta(\tau)).$$

The Jacobi equation gives $\dot{\eta}(\tau) = Z_\tau v(\tau)$, and therefore

$$\eta(\tau) = \eta(0) + \int_0^\tau Z_s v(s) ds.$$

Moreover, the definition of v is equivalent to (3.3.6). Integrating this identity against an arbitrary $w \in \mathcal{V}_t$ recovers (3.3.5). Thus $\eta(t) \in \Lambda_t$, which proves Theorem. $\square$

Now, we prove that the Jacobi curve of an extremal arc is increasing. It is a consequence of the following general proposition.

Proposition 3.3.10. *Let h_t be a time dependent quadratic Hamiltonian function on $T_\lambda(T^*M)$. Define $(B_t)_{t \in \mathbb{R}}$ the flow of h_t . For a given $\Lambda_0 \in \mathfrak{L}(T_\lambda(T^*M))$, let $\Lambda_t = B_t(\Lambda_0)$. Suppose moreover that there are $\Pi, \Delta \in \mathfrak{L}(T_\lambda(T^*M))$, with $\Pi \cap \Delta = \{0\}$, such that $\Lambda_t \cap \Delta = \{0\}$ for every $t \in [0, T]$ and $\Lambda_t = \{(p, S_t p) \mid p \in \Pi\}$, where, as before, we are using the coordinates in $T_\lambda(T^*M)$ such that $\Pi = \{(p, 0) \mid p \in \Pi\}$ and $S_t : \Pi \rightarrow \Delta$ is symmetric. Then*

$$p^\top \dot{S}_t p = 2h_t(p, S_t p), \quad p \in \Pi.$$

In particular, if h_t is positive definite for every $t \in [0, T]$, then also $\dot{S}_t$ is positive definite for every $t \in [0, T]$ and Λ is an increasing curve.

Proof. Using the coordinates introduced in the statement, we have that any integral trajectory of the Hamiltonian vector field $\vec{h}_t$ can be written as $(p(t), q(t)) = (p(t), S_t p(t))$ and in particular

$$\dot{q} = \dot{S}_t p + S_t \dot{p},$$

that is $\vec{h}_t(p, q) = (\dot{p}, \dot{S}_t p + S_t \dot{p})$. Moreover, since the Hamiltonian h_t is quadratic, we have that $\sigma((p, q), \vec{h}_t(p, q)) = 2h_t(p, q)$. Thus, the left hand side of the previous equality reads

$$\sigma((p, S_t p), \vec{h}_t(p, S_t p)) = \sigma((p, S_t p), (\dot{p}, \dot{S}_t p + S_t \dot{p})) = p^T \dot{S}_t p.$$

where we have used the fact that S_t is symmetric. Hence, we see that $\dot{S}_t(p) = 2h_t(p, S_t p)$, as we wanted to show. $\square$

Corollary 3.3.11. *Let $(\tilde{u}, \tilde{\lambda})$ be a regular extremal pair and let Λ be its corresponding Jacobi curve. Then Λ is increasing.*

This is a direct consequence of Proposition 3.3.10, since by the strong Legendre condition b_t is positive definite.

Now, we provide an alternative interpretation of the Jacobi equation, which is more close to the classical case in Riemannian geometry.

In our problem, there are two relevant dynamics: the Hamiltonian dynamics associated with the reference regular control $\tilde{u}$ (encoded by the flow $\tilde{\Phi}_t$), and the dynamics of the maximized Hamiltonian H (encoded by the flow Φ_t). The composition

$$\tilde{\Phi}_t^{-1} \circ \Phi_t$$

compares these two dynamics after transporting both of them back to the reference frame. Since the two flows agree along the reference extremal, this map fixes $\tilde{\lambda}_0$. Its quadratic part at $\tilde{\lambda}_0$ therefore measures the first non-trivial discrepancy between the two dynamics, and, as shown in Corollary 3.3.13 below, this discrepancy is exactly the quadratic Hamiltonian governing the Jacobi equation.

We want to write it as the flow of another Hamiltonian function, i.e.

$$\tilde{\Phi}_{-t} \circ \Phi_t = \overrightarrow{\exp} \int_0^t \vec{\varphi}_s ds, \quad (3.3.8)$$

where here again we follow the notation of the chronological calculus (see Appendix A.1). It follows by the variations formula (see (A.1.1)) that this is

$$\varphi_t(\ell) = (H - h_{\tilde{u}(t)}) \circ \tilde{\Phi}_t(\ell), \quad \ell \in T^*M, t \in [0, T].$$

Indeed, we have

$$\vec{\varphi}_t = \overrightarrow{(H - h_{\tilde{u}(t)}) \circ \tilde{\Phi}_t} = \left(\tilde{\Phi}_{-t} \right)_* \left(\vec{H} - \vec{h}_{\tilde{u}(t)} \right),$$

hence, by the variations formula (see Equation (A.1.1)), we obtain

$$\begin{aligned} \overrightarrow{\exp} \int_0^t \vec{\varphi}_s ds &= \overrightarrow{\exp} \int_0^t \left(\tilde{\Phi}_{-s} \right)_* \vec{H} - \left(\tilde{\Phi}_{-s} \right)_* \vec{h}_{\tilde{u}(s)} ds = \\ &= \overrightarrow{\exp} \int_0^t - \left(\tilde{\Phi}_{-s} \right)_* \vec{h}_{\tilde{u}(s)} ds \circ \overrightarrow{\exp} \int_0^t \left(\tilde{\Phi}_{-s} \right)_* \left(\tilde{\Phi}_s \right)_* \vec{H} ds = \\ &= \left(\overrightarrow{\exp} \int_0^t \vec{h}_{\tilde{u}(s)} ds \right)^{-1} \circ \overrightarrow{\exp} \int_0^t \vec{H} ds. \end{aligned}$$

We have the following relation with the Jacobi equation.

Proposition 3.3.12. *Under the previous notation, we have that*

$$\frac{1}{2} \text{Hess}_{\tilde{\lambda}_0} \varphi_t = b_t.$$

The proof of this proposition is just a standard computation, which is done, for instance, in [13], see Proposition 21.3. We conclude this subsection with a corollary of Proposition 3.3.12.

Corollary 3.3.13. *Under the previous setting, it holds the following identity:*

$$\left(\overrightarrow{\exp} \int_0^t \vec{\varphi}_s ds \right)_{*, \lambda(0)} = \overrightarrow{\exp} \int_0^t \vec{b}_s ds. \quad (3.3.9)$$

Proof. If $h : \mathbb{R}^{2d} \rightarrow \mathbb{R}$ is a quadratic Hamiltonian function, then

$$2h(v) = \sigma(v, \vec{h}[v]), \quad v \in \mathbb{R}^{2d}, \quad (3.3.10)$$

where σ is the standard symplectic form of $\mathbb{R}^{2d}$. To prove this identity, it suffices to show that it holds with respect to any fixed basis. So, we fix any basis of $\mathbb{R}^{2d}$ and we let σ and h be represented by the matrices J and A respectively in this basis, that is $h(v) = v^\top A v$ and $\sigma(v_1, v_2) = -v_1^\top J v_2$. We have that $\vec{h}(v) = 2J A v$ and $\sigma(v, \vec{h}[v]) = -2v^\top J J A v = 2v^\top A v = 2h(v)$, as we claimed.

Since $\varphi_t(\lambda_0) = 0$ and $d_{\lambda_0} \varphi_t = 0$ for every $t \in [0, T]$, then the left-hand side of (3.3.9) is a linear Hamiltonian flow defined on $T_{\lambda_0}(T^*M)$, whose Hamiltonian system is

$$\dot{\eta} = d_{\lambda_0} \vec{\varphi}_t[\eta].$$

Recalling that $d_{\lambda_0} \varphi_t = \sigma_{\lambda_0}(\cdot, \vec{\varphi}_t)$ and differentiating in λ , we obtain

$$d_{\lambda_0}^2 \varphi_t[\eta, \eta] = \sigma_{\lambda_0}(\eta, d_{\lambda_0} \vec{\varphi}_t[\eta]). \quad (3.3.11)$$

By Proposition 3.3.12, the left-hand side of (3.3.11) is equal to the quadratic form $2b_t$. Putting this fact together with Equation (3.3.10) yields

$$\sigma_{\lambda_0}(\eta, \vec{b}_t[\eta]) = 2b_t(\eta) = d_{\lambda_0}^2 \varphi_t[\eta, \eta] = \sigma_{\lambda_0}(\eta, d_{\lambda_0} \vec{\varphi}_t[\eta]), \quad \eta \in T_{\lambda_0}(T^*M).$$

Hence, since σ is non-degenerate, we have $\vec{b}_t = d_{\lambda_0} \vec{\varphi}_t$. Integrating the flow of both sides, the statement of the Corollary follows. $\square$

Remark 3.3.14. The previous Corollary shows that solutions of the Jacobi equation are the transported back solutions of the equation $\dot{\xi} = d_{\tilde{\lambda}_t} \vec{H}[\xi]$. More precisely, let $\tilde{\lambda}_t$ be a regular extremal. Let $\xi(t) \in T_{\tilde{\lambda}_t}(T^*M)$ solve the linearized Hamiltonian system

$$\dot{\xi}(t) = d_{\tilde{\lambda}_t} \vec{H}[\xi(t)]. \quad (3.3.12)$$

Define its pull-back along the reference Hamiltonian flow by

$$\eta(t) := d_{\tilde{\lambda}_t} \tilde{\Phi}_t^{-1} [\xi(t)] \in T_{\lambda_0}(T^*M).$$

By Equation (3.3.8) and Corollary 3.3.13, we obtain

$$\eta(t) = \left(\overrightarrow{\exp} \int_0^t \vec{b}_s ds \right)_* \eta(0),$$

that is, η satisfies the Jacobi equation (3.3.4).

Conversely, if $\eta(t)$ solves the Jacobi equation, then $\xi(t) := (\Phi_t)_{*,\lambda_0} \eta(t)$ solves (3.3.12). Thus, one can re-obtain the Jacobi curve by first solving (3.3.12) with initial conditions in Π , obtaining a curve of subspaces $\Xi_t := (\Phi_t)_* \Pi \in T_{\tilde{\lambda}_t}(T^*M)$ and then transporting back the curve at the initial point applying $(\tilde{\Phi}_t)_*^{-1}$, that is $\Lambda_t = (\tilde{\Phi}_t)_*^{-1} \Xi_t = (\tilde{\Phi}_t)_*^{-1} (\Phi_t)_* \Pi$.

3.4 Morse and Maslov index

In this section, we put in relation the Morse index of the second variation of the endpoint map as introduced in Section 3.2.4 with the Maslov index of the Jacobi curve. First, we recall some definitions and notation needed in the sequel (we mainly refer to [57] for a more extensive introduction to the Maslov index and to [6] for its application in optimal control).

3.4.1 Maslov index for smooth curves

Let (V, σ) be a symplectic linear space. Fix a Lagrangian subspace $L_0 \subset V$. The Maslov train of L_0 is the set of Lagrangian subspaces with non-trivial intersection with L_0 , that is

$$\mathcal{M}_{L_0} := \{W \in \mathfrak{L}(V) : W \cap L_0 \neq \{0\}\}.$$

The Maslov train is a stratified manifold, whose strata are given by

$$\mathcal{M}_{L_0}^k := \{W \in \mathfrak{L}(V) : \dim(W \cap L_0) = k\}, \quad k = 1, \dots, \frac{\dim V}{2}.$$

In this subsection we use the local coordinates of $\mathfrak{L}(V)$ and many definitions introduced in Subsection 2.2.2.

We define the Maslov index of a C^1 curve in the Lagrangian Grassmannian.

Definition 3.4.1 (Maslov index of a curve). Let $(\Lambda_t)_{t \in [0, T]}$ be a C^1 curve in the Lagrangian Grassmannian $\mathfrak{L}(V)$ and take $\Pi \in \mathfrak{L}(V)$. Suppose that the curve Λ has finitely many intersections at times $t_1, \dots, t_\ell$ with the Maslov train $\mathcal{M}_\Pi^1$. Suppose moreover that all intersections are transverse. For $j \in \{1, \dots, \ell\}$, let $v_j \in \Lambda_{t_j} \cap \Pi$ be any vector in the intersection. We define the number $i(j) = 1$ if we have $\dot{\Lambda}_{t_j}(v_j) > 0$ and $i(j) = -1$ if $\dot{\Lambda}_{t_j}(v_j) < 0$. The *Maslov index* of the curve Λ with respect to Π is defined as

$$\mu_\Pi(\Lambda) = \sum_{j=1}^{\ell} i(j) + \frac{1}{2} (\dim(\Lambda_0 \cap \Pi) + \dim(\Lambda_T \cap \Pi) - \dim(\Lambda_0 \cap \Lambda_T \cap \Pi)).$$

Remark 3.4.2. Since the stratum $\mathcal{M}_\Pi^1$ has codimension one in $\mathcal{M}_\Pi$ and $\mathcal{M}_\Pi^2$ has codimension three, we have that the Maslov index defined as in Definition 3.4.1 is invariant under homotopy of the curve Λ with fixed endpoints. Moreover, since $\mathcal{M}_\Pi^1$ is dense in $\mathcal{M}_\Pi$, we can extend the previous definition of Maslov index to any curve Λ such that $\Lambda(0) \cap \Pi = \Lambda(T) \cap \Pi = \{0\}$.

Remark 3.4.3. In the classical references, see [57], the Maslov index μ_Π is defined for curves whose initial and final points are transverse to the subspace Π , hence the terms $\dim(\Lambda_0 \cap \Pi)$, $\dim(\Lambda_T \cap \Pi)$, $\dim(\Lambda_0 \cap \Lambda_T \cap \Pi)$ are all zero. However, in the case of a Jacobi curve associated with an extremal pair, we always have that the initial point has a non-trivial intersection with the space Π . Including these terms in the definition of Maslov index allows us to extend the classical definition to our case.

Remark 3.4.4. If $(\Lambda_t)_{t \in [0, T]}$ is an increasing curve, then the function $t \mapsto \mu_\Pi(\Lambda|_{[0, t]})$ is non-decreasing for every $\Pi \in \mathfrak{L}(V)$, since at every new intersection we must have $\dot{\Lambda}_t > 0$.

We introduce also the positive Maslov index of a triple of Lagrangian subspaces. Let $\Pi, \Lambda_1, \Lambda_2 \in \mathfrak{L}(V)$ be three Lagrangian subspaces. Define the quadratic form $\mathfrak{q} : (\Lambda_1 + \Lambda_2) \cap \Pi / (\Pi \cap \Lambda_1 \cap \Lambda_2) \rightarrow \mathbb{R}$ by

$$\mathfrak{q}(\lambda) = \sigma(\lambda_1, \lambda_2), \quad \lambda_1 \in \Lambda_1, \lambda_2 \in \Lambda_2, \lambda = \lambda_1 + \lambda_2.$$

A direct computation shows that $\mathfrak{q}$ does not depend on the choice of λ_1, λ_2 .

Definition 3.4.5 (Maslov index of a triple). Let $\Pi, \Lambda_1, \Lambda_2 \in \mathfrak{L}(V)$ be three Lagrangian subspaces. The positive Maslov index of the triple $(\Lambda_1, \Pi, \Lambda_2)$ is defined as

$$\text{ind}_\Pi(\Lambda_1, \Lambda_2) := \text{ind}^+ \mathfrak{q} + \frac{1}{2} \ker \mathfrak{q}.$$

We call $\text{ind}_\Pi(\Lambda_1, \Lambda_2)$ the positive Maslov index, stressing the word positive, in contrast with the Maslov index used by some other authors, which is instead the signature of the quadratic form $\mathfrak{q}$. The positive Maslov index of a triple of Lagrangian subspaces enjoys many nice properties, see [6] for more details. We further introduce the notion of increasing and simple curve.

We have the following result, which is proved in [15].

Proposition 3.4.6. *Let $(\Lambda_t)_{t \in [t_0, t_1]}$ be a C^1 curve in the Lagrangian Grassmannian $\mathfrak{L}(V)$. Fix $\Pi \in \mathfrak{L}(V)$. Suppose that the curve Λ is simple and increasing. Then, the Maslov index of the curve Λ with respect to Π is given by*

$$\mu_\Pi(\Lambda) = \text{ind}_\Pi(\Lambda_{t_0}, \Lambda_{t_1}).$$

That is, in the particular case of a simple increasing curve, the Maslov index depends only on its endpoints.

For a proof of this Proposition, see [15], Lemma 2.8 in Chapter 1.

Proposition 3.4.7. *Given any pair $\Lambda_1, \Lambda_2 \in \mathfrak{L}(V)$, there is $\Delta \in \mathfrak{L}(V)$ such that $\Lambda_1 \cap \Delta = \Lambda_2 \cap \Delta = \{0\}$ and satisfying $S_1 \leq S_2$, where $S_1, S_2 : \Lambda_1 \rightarrow \Delta$ denote the matrices corresponding to Λ_1, Λ_2 respectively in the (Λ_1, Δ) -coordinates. In particular, given any pair $\Lambda_1, \Lambda_2 \in \mathfrak{L}(V)$, it is always possible to find a simple increasing curve starting from Λ_1 and ending at Λ_2 .*

Proof. Set $k = \dim(\Lambda_1 \cap \Lambda_2)$ and choose a Darboux basis $e_1, \dots, e_n, f_1, \dots, f_n$ adapted to the pair (Λ_1, Λ_2) , so that

$$\Lambda_1 = \text{span}\{e_1, \dots, e_n\}, \quad \Lambda_2 = \text{span}\{e_1, \dots, e_k, f_{k+1}, \dots, f_n\}.$$

Consider the Lagrangian subspace

$$\Delta = \text{span}\{f_1, \dots, f_k, f_{k+1} - e_{k+1}, \dots, f_n - e_n\}.$$

Then Δ is transverse to both Λ_1 and Λ_2 . In the (Λ_1, Δ) -coordinates, Λ_1 is represented by $S_1 = 0$, whereas Λ_2 is represented by

$$S_2 = \begin{pmatrix} 0_k & 0 \\ 0 & I_{n-k} \end{pmatrix}.$$

Hence the curve $t \mapsto tS_2$, $t \in [0, 1]$, has all the required properties. $\square$

This leads to the following alternative way to compute the Maslov index of an increasing curve in the Lagrangian Grassmannian.

Corollary 3.4.8. *Let $(\Lambda_t)_{t \in [0, T]}$ be a C^1 increasing curve in the Lagrangian Grassmannian $\mathfrak{L}(V)$. Fix $\Pi \in \mathfrak{L}(V)$. Then, the Maslov index of the curve Λ with respect to Π is given by*

$$\mu_\Pi(\Lambda) = \sum_{i=1}^N \text{ind}_\Pi(\Lambda_{t_i}, \Lambda_{t_{i+1}}) \quad (3.4.1)$$

where the partition $0 = t_0 < t_1 < \dots < t_{N+1} = T$ is such that $\Lambda_{|(t_i, t_{i+1})}$ is simple for every $i = 1, \dots, N - 1$.

3.4.2 Maslov index for increasing piecewise smooth curves

We can notice that the formula in (3.4.1) makes sense also for discontinuous curves in the Lagrangian Grassmannian. This allows us to extend the definition of Maslov index to this more general setting.

Definition 3.4.9 (Maslov index of a piecewise C^1 and increasing curve). Let $(\Lambda_t)_{t \in [0, T]}$ be a piecewise C^1 curve in the Lagrangian Grassmannian $\mathfrak{L}(V)$, with finitely many discontinuity times. Suppose that for every interval $I \subset [0, T]$ such that $\Lambda|_I \in C^1$ we have that $\Lambda|_I$ is an increasing curve. Let $0 \leq t_1 < \dots < t_k \leq T$ be the times of discontinuity and denote by

$$\Lambda_j^+ := \lim_{t \rightarrow t_j^+} \Lambda_t, \quad \Lambda_j^- := \lim_{t \rightarrow t_j^-} \Lambda_t, \quad j = 1, \dots, k.$$

Fix $\Pi \in \mathfrak{L}(V)$. For any interval (t_j, t_{j+1}) we further consider the subdivision $t_j = t_{j0} < t_{j1} < \dots < t_{jN_j} = t_{j+1}$ such that Λ is simple on each subinterval $(t_{ji}, t_{j(i+1)})$, for $i = 0, \dots, N_j - 1$. By convention, we set $\Lambda_{j0} = \Lambda_j^+$ and $\Lambda_{jN_j} = \Lambda_{j+1}^-$.

The Maslov index of the curve Λ with respect to Π is defined as

$$\mu_\Pi(\Lambda) = \sum_{j=0}^k \sum_{i=0}^{N_j-1} \text{ind}_\Pi(\Lambda_{t_{ji}}, \Lambda_{t_{j(i+1)}}) + \sum_{j=1}^k \text{ind}_\Pi(\Lambda_j^-, \Lambda_j^+).$$

This definition of Maslov index for an increasing piecewise C^1 curve corresponds to the Maslov index of a (globally) C^1 curve obtained from the discontinuous one by connecting the subspaces Λ_j^- and Λ_j^+ with a simple increasing curve, which is always possible (see Remark 3.4.7). By Proposition 3.4.6, we have that any simple increasing curve connecting Λ_j^- and Λ_j^+ gives the same result.

Remark 3.4.10. In the following, whenever we are dealing with a discontinuous piecewise C^1 curve in the Lagrangian Grassmannian, we can always replace it with a globally C^1 curve having the same Maslov index.

3.4.3 Relation between the Maslov and the Morse index

Now, we go back to the setting of optimal control problems. As before, we consider an extremal pair $(\tilde{u}, \tilde{\lambda})$ and its Jacobi curve Λ_t defined as in 3.3.1. We state a theorem which shows how the Maslov index of the Jacobi curve can be used to compute the negative Morse index of the Hessian of the constrained functional $J|_{F^{-1}(q_1)}$. Recall that we denote by $\Pi = T_{q_0}^* M$ the fiber over q_0 in $T^* M$ and $\dim \Pi = d$.

The following theorem is the second fundamental results on which is based the general strategy for proving necessary optimality conditions of second order. It is proved in a slighter general form in [6], see Theorem 2.4. Here, we recall a statement adapted to our purposes.

Theorem 3.4.11. *Let $(\tilde{u}, \tilde{\lambda})$ be an extremal pair for the OCP (2.3.2), (2.3.3). Suppose that $\text{ind}^- \text{Hess } D_{\tilde{u}} J|_{F^{-1}(q_1)}$ is finite and denote by Λ the Jacobi curve associated with $(\tilde{u}, \tilde{\lambda})$. Suppose that Λ is piecewise C^1 and increasing. Then, it holds that*

$$\text{ind}^- (D_{\tilde{u}}^2 J|_{F^{-1}(q_1)}) = \mu_{\Pi}(\Lambda) - \frac{1}{2} \left(d - \dim \left(\Pi \cap \bigcap_{t \in [0, T]} \Lambda_t \right) \right). \quad (3.4.2)$$

The assumption that Λ is piecewise C^1 is not necessary to prove the theorem, but since we did not define μ_{Π} for an arbitrary discontinuous monotone curve we assume it explicitly. In the reference, the authors lift the Jacobi curve to the universal covering of $\mathfrak{L}(T_{\lambda_0}^*(T^*M))$ and prove Equation (3.4.2) for this lifted curve and the corresponding lifted quantities. However, the convergence in the universal covering implies the convergence in the Lagrangian Grassmannian. Thus, substituting the Leray index in the reference with the Maslov index (see Lemma 2.4 in [6]), the result in the reference implies the one stated above.

Remark 3.4.12. If $\dim \Pi \cap \bigcap_{t \in [0, T]} \Lambda_t = \{0\}$, then the term in the right-hand side inside the parenthesis reduces to d . This takes into account that $\Lambda_0 = \Pi$, but the initial time does not contribute to the Morse index. Indeed, recalling the definition of $\mathfrak{q}$, we have that for any Δ such that $\Delta \cap \Pi = \{0\}$, we have $\text{ind}_{\Pi}(\Pi, \Delta) = \frac{d}{2}$. Hence, if Λ is a simple increasing curve starting from Π and going out from Π without other non-trivial intersections with this subspace, then $\mu_{\Pi}(\Lambda) = \frac{d}{2}$ and we have to cancel this contribution. Finally, the term $\dim \left(\Pi \cap \bigcap_{t \in [0, T]} \Lambda_t \right)$ is a correction that takes into account the fact that Λ_t can have a constant intersection with Π for every $t \in [0, T]$, which already does not contribute to $\mu_{\Pi}(\Lambda)$ and so we do not have to subtract it from $\mu_{\Pi}(\Lambda)$.

We are now in a position to prove Theorem 3.0.1. It is a consequence of Theorem 3.4.11 and Theorem 3.1.7.

Proof of Theorem 3.0.1. Let $(\tilde{u}, \tilde{\lambda})$ be an extremal pair and let Λ be the corresponding Jacobi curve. By Theorem 3.1.7, we know that, under the hypothesis $\text{ind}^- \text{Hess}(J|_{F^{-1}(q_1)}) \geq r$, the extended endpoint map is locally open at $\tilde{u}$. Thus, for every $\varepsilon > 0$ it is possible to find $u_{\varepsilon} \in \mathcal{U}$ close to $\tilde{u}$ in the topology of $\mathcal{U}$ such that $F(u_{\varepsilon}) = F(\tilde{u})$ and $J(u_{\varepsilon}) = J(\tilde{u}) - \varepsilon$. It follows that, $\tilde{u}$ is not a local minimizer in $F^{-1}(q_1)$. Finally, applying Equation (3.4.2) in Theorem 3.4.11, yields the desired result. $\square$

Chapter 4

Second-order sufficient conditions: general principles

In the previous chapter, we developed the general framework for second-order necessary conditions of optimality, based on the constrained Hessian, the Jacobi curve, and the Maslov index. The aim of the present chapter is to describe a complementary approach yielding sufficient conditions for local optimality.

The underlying idea is a generalization of the classical *field of extremals* method from the calculus of variations and Riemannian geometry. Roughly speaking, one embeds the reference extremal into a family of nearby extremals whose projections fill a neighbourhood of the reference trajectory. If this family projects regularly onto the state space, it provides a distinguished extremal passing through each nearby point. Then, the cost of nearby trajectories can be compared by integrating the Poincaré–Cartan form. Thus, the problem of proving local optimality is reduced to constructing a suitable field of extremals and verifying the local invertibility of its projection onto the state space.

In geometric optimal control, this construction applies directly to normal regular extremals, see Chapter 17 of [13]. Indeed, under the regularity assumption, the maximization condition determines uniquely the value of the control in a neighbourhood of the reference extremal and the corresponding maximized Hamiltonian is smooth in this neighbourhood. Its Hamiltonian flow therefore provides the natural family of nearby extremals to which the classical argument can be applied.

This regular setting, however, does not cover many important extremals arising in optimal control. For singular extremals, or for extremals containing switching points, the maximization condition may fail to determine a single smooth Hamiltonian in a neighbourhood of the reference trajectory. The classical construction of the field of extremals therefore cannot be applied directly.

A remarkable feature of the argument is that the maximized Hamiltonian itself is not essential. It can be replaced by a suitable *super-Hamiltonian*, namely a Hamiltonian whose

flow agrees with the reference extremal and which dominates the Pontryagin Hamiltonians in the sense required by the comparison argument; see assumptions **(H5)** and **(H6)** in Subsection 4.1.2. The existence of such a function is substantially weaker than the smooth regularity of the maximized Hamiltonian and makes the field-of-extremals method applicable to much more general classes of extremals. A systematic formulation of sufficient conditions based on this idea is given in [71], and the strategy has proved effective in several different optimal control problems; see, for instance, [70, 29, 61, 30].

In the first part of this chapter, we recall and adapt the general sufficient optimality criterion developed in [71]. The formulation is deliberately flexible: the field of extremals is only required to have the regularity needed for the Poincaré–Cartan argument, while its dynamics is generated by a suitable super-Hamiltonian. The crucial remaining requirement is the local invertibility of the restriction of the canonical projection $\pi : T^*M \rightarrow M$ to the field of extremals.

We then introduce a refinement of this criterion which, to the best of our knowledge, has not been explicitly formulated in the literature. When the field of extremals is a Lipschitz manifold admitting a piecewise smooth Lagrangian structure, we show that the invertibility of the restriction of π can be characterized in terms of the Maslov index of the corresponding curve of tangent Lagrangian subspaces. Although this result is essentially a symplectic reformulation of the invertibility condition, it provides the key link with the second-order theory developed in the previous chapter. In the applications considered later in the thesis, the tangent Lagrangian curve can be obtained as a small perturbation of the initial point of the Jacobi curve of the reference extremal, so that the sufficient optimality criterion can be checked directly through the Maslov index of the Jacobi curve. The final part of the chapter makes the relation with conjugate times and conjugate points explicit.

4.1 Sufficient conditions

The method of fields of extremals provides a geometric sufficient condition for strong local optimality. Its basic idea is to embed the reference extremal into a suitably chosen family of neighbouring extremals and to compare the cost of the reference trajectory with that of any admissible competitor contained in the region covered by this family. The comparison is based on the exactness of the Poincaré–Cartan form on the Hamiltonian flow-out defining the field.

We formulate the construction for the free-final-time problem:

$$J(u) = \int_0^T L(q(t), u(t)) dt \longrightarrow \min,$$

subject to

$$\dot{q}(t) = f(q(t), u(t)), \quad q(0) = q_0, \quad q(T) = q_1,$$

where the final time $T > 0$ is free. For $u \in U$, we denote by

$$h_u(\lambda) := \langle \lambda, f(\pi(\lambda), u) \rangle - L(\pi(\lambda), u)$$

the corresponding normal Pontryagin Hamiltonian. In what follows, we fix $(\tilde{u}, \tilde{\lambda})$ a reference normal extremal pair, and let $\tilde{q} = \pi \circ \tilde{\lambda}$ be its projection. We assume that $\tilde{\lambda}$ is globally Lipschitz and piecewise smooth.

4.1.1 The Poincaré–Cartan form

We first recall the integral invariant underlying the method of fields of extremals. Let $I \subset \mathbb{R}$ be an open interval and let

$$H : I \times T^*M \longrightarrow \mathbb{R}$$

be a possibly time-dependent Hamiltonian satisfying the regularity assumptions described above. We denote its flow by $\Phi_{s,t}^H$.

On the extended phase space $T^*M \times I$, consider the one-form

$$\alpha_H := \vartheta - H dt,$$

where ϑ is the Liouville form on T^*M . The form α_H is called the *Poincaré–Cartan form* associated with H . In canonical coordinates (p, q) on T^*M , it is given by

$$\alpha_H = p dq - H(t, p, q) dt.$$

When H is only locally Lipschitz, the form α_H has locally Lipschitz coefficients. Its pull-back by a Lipschitz map and its exterior differential are therefore understood almost everywhere, or equivalently in the weak sense.

Let $a \in C^\infty(M)$ and consider the exact Lagrangian graph

$$\mathcal{L}_0 := \{d_q a \mid q \in M\} \subset T^*M.$$

The restriction of the Liouville form to $\mathcal{L}_0$ is exact:

$$\vartheta|_{\mathcal{L}_0} = d(a \circ \pi)|_{\mathcal{L}_0}.$$

For every $t \in I$, set

$$\mathcal{L}_t := \Phi_{0,t}^H(\mathcal{L}_0),$$

and define

$$F : \mathcal{L}_0 \times I \longrightarrow T^*M \times I, \quad F(\lambda_0, t) = (\Phi_{0,t}^H(\lambda_0), t).$$

We denote its image by

$$\widehat{\mathcal{L}} := F(\mathcal{L}_0 \times I).$$

Proposition 4.1.1 (Exactness of the Poincaré–Cartan form). *Assume that the Hamiltonian flow $\Phi_{s,t}^H$ is well defined and locally bi-Lipschitz in a neighbourhood of $\widehat{\mathcal{L}}$. Then the pull-back of the Poincaré–Cartan form to $\widehat{\mathcal{L}}$ is exact: there exists a locally Lipschitz function*

$$S : \mathcal{L}_0 \times I \longrightarrow \mathbb{R}$$

such that

$$F^* \alpha_H = dS$$

almost everywhere.

A primitive can be written explicitly. If

$$\lambda_t = \Phi_{0,t}^H(\lambda_0),$$

then one may take

$$S(\lambda_0, t) = a(\pi(\lambda_0)) + \int_0^t \left(\vartheta_{\lambda_s}(\vec{H}(s, \lambda_s)) - H(s, \lambda_s) \right) ds.$$

The proof is the same as the proof of the Poincaré–Cartan integral invariant in [13, Section 17.1.1]. Indeed, the argument is an application of Stokes’ theorem, which holds also in the present Lipschitz setting, and uses only that the sets $\mathcal{L}_t$ are obtained from an exact Lagrangian manifold through a Hamiltonian flow; it does not use that H is the maximized Hamiltonian of an optimal control problem. Consequently, it applies to any Hamiltonian satisfying the preceding regularity assumptions. In the piecewise smooth setting, the argument is applied on each smooth stratum of the flow-out. The corresponding primitives agree on the common strata, and therefore define the locally Lipschitz primitive S appearing in Proposition 4.1.1.

4.1.2 The free-time field of extremals and the super-Hamiltonian

We now use the super-Hamiltonian flow to construct a field of extremals adapted to the free-final-time problem. Let $a \in C^\infty(M)$ and denote by $\mathcal{L}_0$ the graph of its differential. Assume that $\tilde{\lambda}_0 = d_{q_0}a \in \mathcal{L}_0$. Let $H_S : \mathbb{R} \times T^*M \longrightarrow \mathbb{R}$. We assume that:

- (H1) H_S is a Lipschitz continuous time-dependent Hamiltonian.
- (H2) the Hamiltonian system associated with H_S admits a unique flow

$$\Phi_{s,t}^S : T^*M \longrightarrow T^*M,$$

which satisfies the following regularity assumptions:

- (H2a) the map $(s, t, \lambda) \longmapsto \Phi_{s,t}^S(\lambda)$ is Lipschitz continuous;

(H2b) for every $s, t \in \mathbb{R}$, the map $\Phi_{s,t}^S : T^*M \rightarrow T^*M$ is bi-Lipschitz, with inverse $\Phi_{t,s}^S$.

(H3) $H_S(0, \cdot)$ is smooth in a neighbourhood of $\tilde{\lambda}_0$ and that 0 is a regular value of its restriction to $\mathcal{L}_0$, namely

$$d_\lambda (H_S(0, \cdot)|_{\mathcal{L}_0}) \neq 0$$

We fix a relatively open connected subset

$$\mathcal{M}_0 \subset \mathcal{L}_0 \cap H_S(0, \cdot)^{-1}(0)$$

containing $\tilde{\lambda}_0$. By the regular-value assumption, $\mathcal{M}_0$ is a smooth submanifold of dimension $d - 1$. Since it is contained in the Lagrangian manifold $\mathcal{L}_0$, it is isotropic.

For every $t \in I$, define

$$\mathcal{M}_t := \Phi_{0,t}^S(\mathcal{M}_0),$$

We denote the union of the $\mathcal{M}_t$ by

$$\mathcal{M} := \Psi(\mathcal{M}_0 \times I) = \bigcup_{t \in I} \mathcal{M}_t,$$

and we consider the flow-out parametrization

$$\Psi : \mathcal{M}_0 \times I \longrightarrow T^*M, \quad \Psi(\lambda_0, t) = \Phi_{0,t}^S(\lambda_0).$$

In general, by the Lipschitz regularity of $\Phi_{s,t}$, $\mathcal{M}$ is only a Lipschitz immersed manifold. We make the following assumptions on $\mathcal{M}$:

(H4) for $t \in I$ and $\lambda_0 \in \mathcal{M}_0$, we have that $H_S(t, \Psi(\lambda_0, t)) = 0$.

We have two more assumption on H_S .

(H5) For $\lambda_0 \in \mathcal{M}_0$, $t \geq 0$ and denoting $\lambda_t = \Phi_{0,t}^S(\lambda_0)$, the function H_S satisfies

$$H(\lambda_t) \leq H_S(t, \lambda_t).$$

(H6) The function H_S along the reference extremal $\tilde{\lambda}$ satisfies

$$H_S(t, \tilde{\lambda}_t) = h_{\tilde{u}(t)}(\tilde{\lambda}_t) \quad \text{and} \quad \vec{H}_S(t, \tilde{\lambda}_t) = \vec{h}_{\tilde{u}(t)}(\tilde{\lambda}_t)$$

almost everywhere on I .

For a normal extremal couple $(\tilde{u}, \tilde{\lambda})$, if $H_S : I \times T^*M \rightarrow \mathbb{R}$ and $\mathcal{M} \subset T^*M$ satisfy hypotheses from **(H1)** to **(H6)**, we say that H_S is a *super-Hamiltonian* function and $\mathcal{M}$ is a *field of extremals*. The idea of replacing the maximized Hamiltonian by a super-Hamiltonian was first introduced in [70] and successfully reused in many other situations, see [71] and references therein. Strictly speaking, the trajectories $\lambda_t = \Psi(\lambda_0, t)$ are not extremals for the OCP considered above and some authors rather call the set $\mathcal{M}$ a field of *almost*-extremals. By a slight abuse of terminology, we call it simply field of extremals whenever it does not generate confusion.

Remark 4.1.2 (Preservation of the zero level). Since the super-Hamiltonian may depend explicitly on time, the inclusion of $\mathcal{M}_0$ in the zero level of $H_S(0, \cdot)$ does not by itself imply that its flow-out remains in the zero level of H_S . Indeed, if

$$\dot{\lambda}_t = \vec{H}_S(t, \lambda_t),$$

then, at every smooth time,

$$\frac{d}{dt} H_S(t, \lambda_t) = \partial_t H_S(t, \lambda_t) + d_\lambda H_S(t, \lambda_t) \left[\vec{H}_S(t, \lambda_t) \right] = \partial_t H_S(t, \lambda_t),$$

since $d_\lambda H_S \left[\vec{H}_S \right] = \{H_S, H_S\} = 0$. We shall therefore explicitly assume that

$$H_S|_{\mathcal{M}} = 0.$$

A useful way to verify this assumption is to identify an autonomous first integral of the super-Hamiltonian flow. More precisely, suppose that there exist an invariant subset $\Sigma \subset T^*M$ and an autonomous Hamiltonian $K : T^*M \rightarrow \mathbb{R}$ such that

$$H_S(t, \lambda) = K(\lambda), \quad \lambda \in \Sigma,$$

and

$$\{H_S(t, \cdot), K\} = 0.$$

If

$$\mathcal{M}_0 \subset \Sigma \cap K^{-1}(0),$$

then both Σ and the zero level of K are preserved by the super-Hamiltonian flow, and consequently

$$H_S|_{\mathcal{M}} = 0.$$

This mechanism applies to the nice-singular super-Hamiltonian, which will be introduced in later chapters.

4.1.3 Sufficient optimality criterion

We now formulate the sufficient optimality criterion associated with the free-time field constructed in the previous subsection.

Letting $\pi : T^*M \rightarrow M$ be the canonical bundle projection, we assume that

(H7) the restriction $\pi|_{\mathcal{M}}$ is a bi-Lipschitz map onto some open neighbourhood of $\tilde{q}$.

For brevity, we denote $\pi_{\mathcal{M}} := \pi|_{\mathcal{M}}$. The following theorem is a rephrasing of Theorem 3.10 in [71] adapted to our framework.

Theorem 4.1.3 (Sufficient optimality criterion). *Let $(\tilde{u}, \tilde{\lambda})$ be a normal extremal pair, defined on $[0, T]$, for the free-final-time OCP. Suppose that there exist an Hamiltonian function H_S and an initial manifold $\mathcal{M}_0$ satisfying the hypotheses **(H1)** to **(H7)**. Then, the reference trajectory*

$$\tilde{q}(t) = \pi(\tilde{\lambda}_t)$$

is locally strongly optimal. Suppose furthermore that the following condition holds: let (u, q) be an admissible pair defined on the interval I' and such that the image of q is contained in $\pi_{\mathcal{M}}(\mathcal{M})$. Let λ_t be defined by $q(t) = \pi_{\mathcal{M}}(\lambda_t)$. If

$$H(\lambda_t) = H_S(t, \lambda_t), \quad t \in I'$$

implies

$$f(q(t), u(t)) = d_{\lambda_t} \pi_{\mathcal{M}} \left[\vec{H}_S(t, \lambda_t) \right], \quad \text{for a.e. } t \in I' \quad (4.1.1)$$

then $\tilde{q}$ is locally strongly strictly optimal.

Proof. Let u be an admissible control competitor defined on $I' = [0, T_u]$, and let q be its trajectory. Since $\pi_{\mathcal{M}}$ is bi-Lipschitz and q is Lipschitz continuous, the curve

$$\lambda_t := \pi_{\mathcal{M}}^{-1}(q(t)),$$

is also Lipschitz continuous. We denote by $(\tau(t), \lambda_0(t)) \in I \times \mathcal{M}_0$ the pair such that $\lambda_t = \Psi(\tau(t), \lambda_0(t))$. The injectivity of $\pi_{\mathcal{M}}$ and the boundary conditions imply

$$\lambda_0 = \tilde{\lambda}_0, \quad \lambda_{T_u} = \tilde{\lambda}_T.$$

By Proposition 4.1.1, the pull-back of the Liouville form is exact:

$$\Psi^* \vartheta = dS_{\mathcal{M}}$$

almost everywhere on $\mathcal{M}_0 \times I$.

Denoting by $\mathcal{S}$ the concatenation of λ with the curve $[0, T] \ni t \mapsto \tilde{\lambda}(T - t)$, Stokes' theorem for Lipschitz manifolds gives

$$\int_0^{T_u} \vartheta_{\lambda_t}(\dot{\lambda}_t) dt - \int_0^T \vartheta_{\tilde{\lambda}_t}(\dot{\tilde{\lambda}}_t) dt = \int_{\mathcal{S}} dS_{\mathcal{M}} = 0,$$

since the restriction of ϑ to the Lipschitz Lagrangian manifold $\mathcal{M}$ is closed.

Along the reference extremal, **(H4)** and **(H6)** give

$$h_{\tilde{u}(t)}(\tilde{\lambda}_t) = H_S(t, \tilde{\lambda}_t) = 0.$$

Therefore,

$$\vartheta_{\tilde{\lambda}_t}(\dot{\tilde{\lambda}}_t) = \left\langle \tilde{\lambda}_t, \dot{\tilde{q}}(t) \right\rangle = h_{\tilde{u}(t)}(\tilde{\lambda}_t) + L(\tilde{q}(t), \tilde{u}(t)) = L(\tilde{q}(t), \tilde{u}(t)).$$

It follows that

$$\int_0^T \vartheta_{\tilde{\lambda}_t}(\dot{\tilde{\lambda}}_t) dt = J(\tilde{u}). \quad (4.1.2)$$

Along λ_t , the definition of the Liouville form gives

$$\vartheta_{\lambda_t}(\dot{\lambda}_t) = \langle \lambda_t, \dot{q}(t) \rangle.$$

Since

$$\dot{q}(t) = f(q(t), u(t)),$$

one has

$$\vartheta_{\lambda_t}(\dot{\lambda}_t) = h_{u(t)}(\lambda_t) + L(q(t), u(t)) \leq H_S(\tau(t), \lambda_t) + L(q(t), u(t)). \quad (4.1.3)$$

The lifted curve belongs to the zero level of the super-Hamiltonian, and hence

$$H_S(\tau(t), \lambda_t) = 0.$$

Thus,

$$\vartheta_{\lambda_t}(\dot{\lambda}_t) \leq L(q(t), u(t)) \quad (4.1.4)$$

almost everywhere. Integrating and using equations (4.1.4) and (4.1.2), we obtain

$$J(\tilde{u}) = \int_0^T \vartheta_{\tilde{\lambda}_t}(\dot{\tilde{\lambda}}_t) dt = \int_0^{T_u} \vartheta_{\lambda_t}(\dot{\lambda}_t) dt \leq \int_0^{T_u} L(q(t), u(t)) dt = J(u).$$

This proves local strong optimality.

Assume now that the condition (4.1.1) holds and that

$$J(u) = J(\tilde{u}).$$

Then, by (4.1.3), equality must then hold in the pointwise super-Hamiltonian inequality for almost every t :

$$h_{u(t)}(\lambda_t) = H_S(\tau(t), \lambda_t).$$

Indeed, if $h_{u(t)}(\lambda_t) < H_S(\tau(t), \lambda_t)$ for some set of positive measure, then also $J(\tilde{u}) < J(u)$. Then, we must have $H(\lambda_t) = H_S(\tau(t), \lambda_t)$ and Equation (4.1.1) implies

$$\dot{q}(t) = d_{\lambda_t} \pi_{\mathcal{M}} \left[\vec{H}_S(\tau(t), \lambda_t) \right]$$

almost everywhere. Since $\pi_{\mathcal{M}}$ is invertible, this implies

$$\dot{\lambda}_t = \vec{H}_S(\tau(t), \lambda_t)$$

Differentiating $\lambda_t = \Psi(\lambda_0(t), \tau(t))$, yields

$$d_{\lambda_0(t)} \Psi(\cdot, \tau(t))[\dot{\lambda}_0(t)] + (\dot{\tau}(t) - 1) \vec{H}_S(\tau(t), \lambda_t) = 0.$$

Since $d_{\lambda_0(t)} \Psi(\cdot, \tau(t))[\dot{\lambda}_0(t)]$ and $\vec{H}_S(\tau(t), \lambda_t)$ are linearly independent and $\vec{H}_S(\tau(t), \lambda_t) \neq 0$, yields $\tau(t) = t$ and $\lambda_0(t) = \tilde{\lambda}_0$ is constant. Therefore, $\lambda_t = \tilde{\lambda}_t$ and $\tilde{q}$ is strictly optimal. $\square$

4.1.4 Invertibility of $\pi_{\mathcal{M}}$ and the Maslov index

In the previous subsection, we have seen that very little regularity is required on the submanifold $\mathcal{M}$ to prove Theorem 4.1.3. However, without additional structure, it is quite difficult to check that $\mathcal{M}$ verify the hypothesis **(H7)** concerning the local invertibility of $\pi_{\mathcal{M}}$. As we show in later application, if the reference extremal has a piecewise smooth structure, then $\mathcal{M}$ can be found so that it also have a piecewise smooth Lagrangian structure. Namely, in the cases that we study, the field of extremals satisfy the following regularity:

(H8) $\mathcal{M}$ is the union of finitely many smooth d -dimensional Lagrangian submanifolds with boundary

$$\mathcal{M}_1, \dots, \mathcal{M}_{k+1} \subset T^*M,$$

where $d = \dim M$, such that

$$\mathcal{M} = \bigcup_{j=1}^{k+1} \mathcal{M}_j.$$

The interiors of the manifolds $\mathcal{M}_i$ are pairwise disjoint and, for $i = 2, \dots, k$, we have $\mathcal{M}_i \cap \mathcal{M}_j \neq \emptyset$ only if $j \in \{i-1, i, i+1\}$. For every $i \in \{1, \dots, k\}$, the intersection

$$\mathcal{M}_i \cap \mathcal{M}_{i+1} = \partial \mathcal{M}_i \cap \mathcal{M}_{i+1} = \mathcal{M}_i \cap \partial \mathcal{M}_{i+1},$$

is the non-empty common boundary component of $\mathcal{M}_i$ and $\mathcal{M}_{i+1}$. Every boundary component is a smooth $(d-1)$ -dimensional isotropic submanifold of T^*M .

By this piecewise smoothness assumption, we coherently assume that the vector fields $\vec{H}^i := \vec{H}_S|_{\text{int}\mathcal{M}_i}$ and the corresponding flows $\Phi_{s,t}^i := \Phi_{s,t}^S|_{\text{int}\mathcal{M}_i}$ are smooth and have a smooth extension in some open neighbourhood of $\mathcal{M}_i$ in T^*M .

The first consequence of these regularity assumptions is that the generalized tangent family at the points where $\mathcal{M}$ is non-smooth can be explicitly characterized. We use the notion introduced in Definition A.2.8.

Proposition 4.1.4. *Let $\lambda \in \mathcal{M}_i \cap \mathcal{M}_{i+1}$ be a point where $\mathcal{M}$ is non-smooth. Then*

$$\mathcal{T}_{\lambda}\mathcal{M} = \left\{ T_{\lambda}(\mathcal{M}_i \cap \mathcal{M}_{i+1}) + \mathbb{R} \left(\alpha \vec{H}^{i+1}(\lambda) + (1-\alpha) \vec{H}^i(\lambda) \right) \mid \alpha \in [0, 1] \right\}. \quad (4.1.5)$$

This follows directly from Proposition A.2.9, applied to the flow-out parametrization Ψ , since the two one-sided transverse derivatives are given by the corresponding Hamiltonian vector fields. Now, let $\tilde{\lambda}_t$ be an extremal defined on the time interval $[0, T]$ and suppose that there are

$$0 = t_0 < t_1 < \dots < t_k < t_{k+1} = T,$$

such that $\lambda|_{(t_{j-1}, t_j)} \subset \mathcal{M}_j$, for $j = 1, \dots, k+1$. We call the t_j switching times. Then, $\tilde{\lambda}$ is non-smooth at the switching times, and is smooth otherwise. For every t such that $t \in (t_{j-1}, t_j)$, consider the transported-back tangent space

$$\mathfrak{M}_t := (\tilde{\Phi}_{0,t})_*^{-1} T_{\tilde{\lambda}_t} \mathcal{M}_i. \quad (4.1.6)$$

At a non-smooth time t_j , we denote by $\mathfrak{M}_{t_j}^-$ and $\mathfrak{M}_{t_j}^+$ the left and right limit at t_j of $\mathfrak{M}_t$ respectively. In this way, $\mathfrak{M}$ is a piecewise smooth curve in the Lagrangian Grassmannian and, by Proposition 3.3.10, its smooth arcs are increasing. More formally, we have the following result.

Proposition 4.1.5. *Let the curve $\mathfrak{M}$ defined in (4.1.6) be smooth on the time interval (t_j, t_{j+1}) . Define*

$$\varphi_t^j = (H_S(t, \cdot) - h_{\tilde{u}(t)}(\cdot)) \circ \tilde{\Phi}_{t_j, t},$$

and

$$b_t^j = \frac{1}{2} \text{Hess}_{\tilde{\lambda}_{t_j}} \varphi_t^j.$$

Let

$$C_t^j := \overrightarrow{\exp} \int_{t_j}^t \tilde{b}_s^j ds,$$

and define the corresponding linear Hamiltonian flow on the fixed symplectic space $T_{\tilde{\lambda}_0}^*(T^*M)$ by

$$B_t^j := (\tilde{\Phi}_{0, t_j})_*^{-1} \circ C_t^j \circ (\tilde{\Phi}_{0, t_j})_*.$$

Then, letting

$$\mathfrak{M}_{t_j}^+ := \lim_{t \rightarrow t_j^+} \mathfrak{M}_t,$$

we have

$$\mathfrak{M}_t = B_t^j \mathfrak{M}_{t_j}^+.$$

In particular, every smooth arc of the curve $\mathfrak{M}$ is increasing.

Proof. For $t \in (t_j, t_{j+1})$, the field of extremals is generated by the super-Hamiltonian flow. Therefore,

$$T_{\tilde{\lambda}_t} \mathcal{M} = (\Phi_{t_j, t}^S)_* T_{\tilde{\lambda}_{t_j}} \mathcal{M}_{j+1}.$$

By the definition of $\mathfrak{M}_{t_j}^+$,

$$T_{\tilde{\lambda}_{t_j}} \mathcal{M}_{j+1} = (\tilde{\Phi}_{0, t_j})_* \mathfrak{M}_{t_j}^+.$$

Hence

$$\begin{aligned} \mathfrak{M}_t &= (\tilde{\Phi}_{0, t})_*^{-1} T_{\tilde{\lambda}_t} \mathcal{M} \\ &= (\tilde{\Phi}_{0, t_j})_*^{-1} (\tilde{\Phi}_{t_j, t})_*^{-1} (\Phi_{t_j, t}^S)_* (\tilde{\Phi}_{0, t_j})_* \mathfrak{M}_{t_j}^+. \end{aligned}$$

Following the same argument as in Subsection 3.3.2, the map

$$\tilde{\Phi}_{t_j, t}^{-1} \circ \Phi_{t_j, t}^S$$

is the Hamiltonian flow generated by φ_t^j . Since this flow fixes $\tilde{\lambda}_{t_j}$, its linearization at $\tilde{\lambda}_{t_j}$ is generated by the quadratic Hamiltonian b_t^j . Consequently,

$$(\tilde{\Phi}_{t_j,t})_*^{-1}(\Phi_{t_j,t}^S)_* = C_t^j = \overrightarrow{\exp} \int_{t_j}^t \vec{b}_s^j ds.$$

Substituting this identity into the previous expression yields

$$\mathfrak{M}_t = (\tilde{\Phi}_{0,t_j})_*^{-1} C_t^j (\tilde{\Phi}_{0,t_j})_* \mathfrak{M}_{t_j}^+ = B_t^j \mathfrak{M}_{t_j}^+.$$

Moreover, since

$$H_S(t, \lambda) \geq h_{\tilde{u}(t)}(\lambda)$$

and the two Hamiltonians coincide at $\tilde{\lambda}_t$, the function φ_t^j has a minimum at $\tilde{\lambda}_{t_j}$. Therefore,

$$\text{Hess}_{\tilde{\lambda}_{t_j}} \varphi_t^j \geq 0,$$

and hence $b_t^j \geq 0$. Proposition 3.3.10 then implies that the curve

$$t \longmapsto B_t^j \mathfrak{M}_{t_j}^+$$

is increasing. Thus the smooth arc $t \mapsto \mathfrak{M}_t$ is increasing. $\square$

Thanks to this property, we can compute the Maslov index of $\mathfrak{M}$ using Definition 3.4.9. This way, it is possible to link this Maslov index with the inveribility of the function $\pi_{\mathcal{M}}$.

More precisely, let $\Pi := \Pi_{\tilde{\lambda}_0}$. Let t_j be a switching time. To emphasize the difference at switching points, we denote by

$$\mathcal{M}_{t_j} := \left\{ (\tilde{\Phi}_{0,t_j})_*^{-1} \Lambda \mid \Lambda \in \mathcal{T}_{\tilde{\lambda}_{t_j}} \mathcal{M} \right\}$$

the transported-back generalized tangent family at t_j . If V, W are vector spaces, by writing $M_{t_j} \cap V = W$ we mean that for every $\mathfrak{M} \in \mathcal{M}_{t_j}$, we have $\mathfrak{M} \cap V = W$.

We also denote by

$$\Pi_t := T_{\tilde{\lambda}_t} \left(T_{\pi(\tilde{\lambda}_t)}^* M \right)$$

the vertical subspace at $\tilde{\lambda}_t$. For $\Lambda \in \mathcal{T}_{\tilde{\lambda}_t} \mathcal{M}$, we define

$$p_\Lambda : \Lambda \longrightarrow d_{\tilde{\lambda}_t} \pi(\Lambda), \quad [v] \longmapsto d_{\tilde{\lambda}_t} \pi(v).$$

Define the following family of projected generalized tangent maps at $\tilde{\lambda}_{t_j}$ is defined made by the family of linear maps

$$\partial_{\tilde{\lambda}_t} \pi_{\mathcal{M}} := \left\{ p_\Lambda \mid \Lambda \in \mathcal{T}_{\tilde{\lambda}_t} \mathcal{M} \right\} = \left\{ d_{\tilde{\lambda}_t} \pi \circ A \mid A \in \partial_{(\tilde{\lambda}_0, 0)} \Psi \right\}.$$

Theorem 4.1.6 (Maslov criterion for invertibility of the projection). *Let $(\tilde{u}, \tilde{\lambda})$ be a normal extremal pair defined on the time interval $[0, T]$, and assume that there are a super-Hamiltonian H_S and a field of extremals $\mathcal{M}$ satisfying hypotheses from **(H1)** to **(H6)** and **(H8)**. If $\mathfrak{M}_0 \cap \Pi = \{0\}$ or, if zero is a switching point, $\mathcal{M}_0 \cap \Pi = \{0\}$ and*

$$\mu_\Pi(\mathfrak{M}) = 0,$$

then $\pi|_{\mathcal{M}}$ is a bi-Lipschitz map onto its image.

Proof. Since hypothesis **(H7)** is local along the reference extremal, it is enough to prove local invertibility of $\pi_{\mathcal{M}}$ at every point $\tilde{\lambda}_t$, $t \in [0, T]$. Indeed, after possibly restricting $\mathcal{M}$ to a sufficiently small relatively compact neighbourhood of the reference extremal, these local inverses provide the required local bi-Lipschitz projection.

This way, it is enough to prove that the tangent spaces of $\mathcal{M}$ are transverse to the vertical distribution along the reference extremal, including every element of the generalized tangent family at the switching times.

Since the reference Hamiltonian flow $\tilde{\Phi}_{0,t}$ projects onto the flow of the reference controlled dynamics, its differential maps the vertical distribution onto itself. Hence

$$(\tilde{\Phi}_{0,t})_* \Pi = \Pi_t.$$

Suppose first that t is not a switching time and that $\tilde{\lambda}_t \in \mathcal{M}_j$. Then

$$\ker d_{\tilde{\lambda}_t} \pi_{\mathcal{M}} = T_{\tilde{\lambda}_t} \mathcal{M}_j \cap \Pi_t.$$

Transporting this intersection back to $T_{\tilde{\lambda}_0}(T^*M)$ gives

$$(\tilde{\Phi}_{0,t})_*^{-1} (T_{\tilde{\lambda}_t} \mathcal{M}_j \cap \Pi_t) = \mathfrak{M}_t \cap \Pi.$$

Therefore

$$d_{\tilde{\lambda}_t} \pi_{\mathcal{M}} \text{ is an isomorphism} \iff \mathfrak{M}_t \cap \Pi = \{0\}.$$

Now recall that, by Definition 3.4.9, the Maslov index of $\mathfrak{M}$ is the sum of the contributions of its smooth arcs and of the positive Maslov indices associated with its jumps. Every smooth contribution is non-negative because the corresponding arc of $\mathfrak{M}$ is increasing. The jump contributions are non-negative by definition. Hence

$$\mu_\Pi(\mathfrak{M}) = 0$$

implies that every smooth contribution and every jump contribution vanish separately.

Starting from the initial transversality assumption, Proposition 3.4.6 therefore implies that

$$\mathfrak{M}_t \cap \Pi = \{0\}$$

throughout the first smooth arc. Indeed, any non-trivial intersection with Π would give a positive contribution to the Maslov index of this increasing arc.

Consider now a switching time t_j . The corresponding contribution to the Maslov index is

$$\text{ind}_\Pi \left(\mathfrak{M}_{t_j}^-, \mathfrak{M}_{t_j}^+ \right),$$

and, as observed above, it vanishes. Moreover, by the characterization of the generalized tangent family in (4.1.5), the family $\mathcal{M}_{t_j}$ can be parametrized as a simple increasing curve connecting $\mathfrak{M}_{t_j}^-$ and $\mathfrak{M}_{t_j}^+$. Thus, since $\text{ind}_\Pi \left(\mathfrak{M}_{t_j}^-, \mathfrak{M}_{t_j}^+ \right) = 0$ and $\mathfrak{M}_{t_j}^- \cap \Pi = \{0\}$ every element we have that $\mathcal{S}_{\tilde{\lambda}_{t_j}} \mathcal{M} \cap \Pi_{t_j} = \{0\}$ as well. Repeating the same argument successively on the following smooth arcs and at the remaining switching times along the whole curve.

Transporting these relations forward by $(\tilde{\Phi}_{0,t})_*$ shows that

$$T_{\tilde{\lambda}_t} \mathcal{M}_j \cap \Pi_t = \{0\}$$

at every smooth point and that

$$\mathcal{S}_{\tilde{\lambda}_{t_j}} \mathcal{M} \cap \Pi_{t_j} = \{0\}.$$

at every switching point. Thus the classical differential of $\pi_{\mathcal{M}}$ is invertible at smooth points, while every element of the Clarke generalized differential associated with the flow-out parametrization is invertible at switching points. $\square$

4.2 Conjugate times and conjugate points

In this section we introduce conjugate times and conjugate points using the Maslov index of the Jacobi curve. The definition is formulated directly in terms of the Maslov index, while the results below provide equivalent geometric characterizations involving only the intersections of the Jacobi curve with the vertical subspace.

Definition 4.2.1 (Conjugate time, conjugate point). Let $(\tilde{u}, \tilde{\lambda})$ be an extremal pair defined on the time interval $[0, T]$ for the OCP (2.3.2), (2.3.3), and denote by Λ its Jacobi curve. Consider the function

$$m_\Lambda(s) := \mu_\Pi \left(\Lambda|_{[0,s]} \right), \quad s \in (0, T].$$

We say that $t \in (0, T]$ is a time conjugate to 0 if m_Λ is discontinuous at t . Moreover, denoting by $\tilde{q}$ the trajectory corresponding to the extremal $\tilde{\lambda}$, we say that $\tilde{q}(t)$ is conjugate to $\tilde{q}(0)$ along $\tilde{q}$.

Although this definition is concise, the definition of the Maslov index for a discontinuous curve is rather technical. We therefore give alternative characterizations of conjugate times involving only the Jacobi curve and the vertical subspace Π . We first consider a time t at which Λ is continuous.

For $\delta > 0$ sufficiently small, consider the subspace

$$\Pi \cap \bigcap_{t-\delta \leq \tau \leq t} \Lambda_\tau.$$

As δ decreases, these subspaces form an increasing family. Since the ambient space is finite-dimensional, there exists $\delta_0 > 0$ such that this family is constant for every $\delta \in (0, \delta_0]$. We denote its stabilized value by

$$K_t^- := \Pi \cap \bigcap_{t-\delta \leq \tau \leq t} \Lambda_\tau, \quad 0 < \delta \leq \delta_0.$$

Thus, K_t^- consists of the vertical directions which already belong to the Jacobi curve throughout a left neighbourhood of t .

Proposition 4.2.2. *Let $(\tilde{u}, \tilde{\lambda})$ be a normal extremal pair and let Λ be the corresponding Jacobi curve. Let $t \in (0, T]$ be a time such that Λ is continuous at t , and assume that $\mu_\Pi(\Lambda|_{[0, t]})$ is finite. Then t is conjugate to 0 if and only if*

$$K_t^- \subsetneq \Pi \cap \Lambda_t. \quad (4.2.1)$$

Proof. By the monotonicity of the Jacobi curve on every interval of continuity, the function m_Λ is non-decreasing. Moreover, its increments belong to $\frac{1}{2}\mathbb{Z}$. Since $\mu_\Pi(\Lambda|_{[0, t]})$ is finite, positive Maslov contributions cannot accumulate at t . Hence, after possibly decreasing δ_0 , we may assume that no Maslov contribution occurs in $[t - \delta_0, t)$.

By continuity of Λ at t , we may further decrease δ_0 so that $\Lambda|_{[t-\delta, t]}$ is simple for every $\delta \in (0, \delta_0]$. By additivity of the Maslov index and Proposition 3.4.6, we then have

$$m_\Lambda(t) - m_\Lambda(t - \delta) = \mu_\Pi(\Lambda|_{[t-\delta, t]}) = \text{ind}_\Pi(\Lambda_{t-\delta}, \Lambda_t). \quad (4.2.2)$$

Set $K := K_t^-$. By the choice of δ_0 , we have

$$K = \Pi \cap \bigcap_{t-\delta \leq \tau \leq t} \Lambda_\tau$$

for every $\delta \in (0, \delta_0]$. We can therefore reduce the simple increasing curve $\Lambda|_{[t-\delta, t]}$ by the common isotropic subspace K . By Corollary 3.4.8, this reduction does not change its Maslov index. In the reduced symplectic space the common intersection with the reduced vertical subspace is trivial, and therefore the local Maslov contribution at t vanishes if and only if the reduced space Λ_t has trivial intersection with the reduced vertical subspace. Equivalently,

$$\text{ind}_\Pi(\Lambda_{t-\delta}, \Lambda_t) = 0 \iff \Pi \cap \Lambda_t = K_t^-.$$

Combining this equivalence with (4.2.2), we conclude that the Maslov index acquires a positive contribution at t if and only if

$$K_t^- \subsetneq \Pi \cap \Lambda_t.$$

This is precisely the claimed characterization of conjugate times. $\square$

It remains to consider the discontinuity times of the Jacobi curve. Let $t \in (0, T)$ be a discontinuity time of Λ , and denote by Λ^- and Λ^+ its left and right limits at t , respectively. Recall that both spaces are contained in the fixed symplectic space $T_{\tilde{\lambda}_0}(T^*M)$. Set

$$\lambda := \tilde{\lambda}_t,$$

and transport the two limiting spaces to the switching covector by defining

$$\hat{\Lambda}^- := \left(\tilde{\Phi}_{0,t} \right)_* \Lambda^-, \quad \hat{\Lambda}^+ := \left(\tilde{\Phi}_{0,t} \right)_* \Lambda^+.$$

Thus,

$$\hat{\Lambda}^-, \hat{\Lambda}^+ \subset T_\lambda(T^*M).$$

We denote by H^- and H^+ the super-Hamiltonians governing the extremal immediately before and after t , respectively. Under assumption **(H8)**, the structure of the set of Lagrange multipliers at λ gives

$$E := \hat{\Lambda}^- \cap \hat{\Lambda}^+, \quad \hat{\Lambda}^- = E \oplus \mathbb{R}\vec{H}^-(\lambda), \quad \hat{\Lambda}^+ = E \oplus \mathbb{R}\vec{H}^+(\lambda). \quad (4.2.3)$$

Proposition 4.2.3. *Let $(\tilde{u}, \tilde{\lambda})$ be a normal extremal pair and let Λ be the corresponding Jacobi curve. Let $t \in (0, T)$ be a discontinuity time of Λ , and use the notation introduced above. Then t is conjugate to 0 if and only if*

$$d_\lambda \pi(\vec{H}^-(\lambda)) \quad \text{and} \quad d_\lambda \pi(\vec{H}^+(\lambda))$$

belong to different connected components of

$$d_\lambda \pi(\hat{\Lambda}^- + \hat{\Lambda}^+) \setminus d_\lambda \pi(E).$$

Proof. Since the Jacobi curve is left-continuous, the discontinuity of m_Λ at t is determined by the Maslov contribution of the jump from Λ^- to Λ^+ . More precisely,

$$\lim_{s \downarrow t} m_\Lambda(s) - m_\Lambda(t) = \text{ind}_\Pi(\Lambda^-, \Lambda^+).$$

Hence t is conjugate to 0 if and only if

$$\text{ind}_\Pi(\Lambda^-, \Lambda^+) > 0. \quad (4.2.4)$$

Since $\tilde{\Phi}_{0,t}$ is symplectic and maps the initial vertical space Π onto the vertical space

$$\Pi_\lambda := T_\lambda \left(T_{\pi(\lambda)}^* M \right),$$

symplectic invariance gives

$$\text{ind}_\Pi(\Lambda^-, \Lambda^+) = \text{ind}_{\Pi_\lambda}(\hat{\Lambda}^-, \hat{\Lambda}^+). \quad (4.2.5)$$

Set

$$V := d_\lambda \pi \left(\widehat{\Lambda}^- + \widehat{\Lambda}^+ \right), \quad W := d_\lambda \pi(E).$$

By (4.2.3),

$$\dim V - \dim W = 1.$$

Therefore $V \setminus W$ has two connected components. The vectors $d_\lambda \pi(\vec{H}^-(\lambda))$ and $d_\lambda \pi(\vec{H}^+(\lambda))$ belong to different components if and only if there exists $\alpha \in (0, 1)$ such that

$$d_\lambda \pi \left(\alpha \vec{H}^+(\lambda) + (1 - \alpha) \vec{H}^-(\lambda) \right) \in W. \quad (4.2.6)$$

For $\alpha \in [0, 1]$, define

$$\widehat{\Lambda}_\alpha := E + \mathbb{R} \left(\alpha \vec{H}^+(\lambda) + (1 - \alpha) \vec{H}^-(\lambda) \right).$$

Moreover, set

$$K_\lambda := E \cap \Pi_\lambda.$$

Condition (4.2.6) is equivalent to

$$\frac{\widehat{\Lambda}_\alpha \cap \Pi_\lambda}{K_\lambda} \neq \{0\}.$$

By Theorem 4.1.6, such an α exists if and only if

$$\text{ind}_{\Pi_\lambda} \left(\widehat{\Lambda}^-, \widehat{\Lambda}^+ \right) > 0.$$

Combining this equivalence with (4.2.4) and (4.2.5) proves the statement. $\square$

Chapter 5

Singular extremals of optimal control problems with L^1 cost

5.1 Introduction

In this Chapter, we study the optimal control problem of minimizing the L^1 norm of the control for a control affine system.

More precisely, we let M be a smooth manifold of dimension $\dim M = d$ and take $f_0, f_1, \dots, f_m : M \rightarrow TM$, smooth vector fields. Also define U to be the closed Euclidean unit ball, that is $U = \{u \in \mathbb{R}^m \mid |u| \leq 1\}$. We take the set of admissible controls

$$\mathcal{U} = L^\infty([0, T], U) = \{u : [0, T] \rightarrow U \mid u \text{ measurable}\}.$$

The set $\mathcal{U}$ has a natural structure of Banach manifold with boundary. However, since in most of the part of this chapter we consider controls satisfying $|u| < 1$, for our purposes we can simply consider the structure of Banach manifold without boundary modelled on $L^\infty([0, T]; \mathbb{R}^m)$. Given any two points on the manifold $q_0, q_1 \in M$, we want to find $\tilde{u} \in \mathcal{U}$ minimizing

$$J(u) = \int_0^T |u(t)| dt \rightarrow \min.$$

and subject to the control-affine dynamic

$$\begin{aligned} \dot{q} &= f_0(q) + \sum_{j=1}^m u_j(t) f_j(q), \\ q(0) &= q_0, \quad q(T) = q_1. \end{aligned} \tag{5.1.1}$$

In the constrained optimization language, this optimal control problem is

$$\min\{\|u\| \mid u \in F^{-1}(q_1)\},$$

where $F : \mathcal{U} \rightarrow M$ denotes the endpoint map of the control system (5.1.1).

This optimal control problem is relevant for many applications. Indeed, the L^1 norm is a natural choice for every model in which the optimal trajectories are expected to have some *inactivated arcs*, that is arcs along which the control is identically switched off. For instance, this is the case when one wants to minimize the fuel consumption of a vehicle, see [27, 68] and references therein, which has important applications in aerospace engineering.

Another circumstance where the minimization of the L^1 arises is the motion of biological systems. In [22, 37], the movements in animals and human beings were studied and the authors found out that such trajectories have inactivated arcs. They found that these kinds of movements minimize the “absolute work”, which can be modelled using the L^1 norm of the control.

The aim of the present Chapter is to provide effective general tools to study these kinds of problems. As a first step, we apply the Pontryagin Maximum Principle (PMP) to describe the extremals of these problems. As we are going to show in Section 5.2, according to PMP, any optimal solution of Problem 5.1 is the concatenation of one of the following three types of controlled arcs:

- arcs with controls taking values at the boundary of U , i.e. $|u| = 1$;
- inactivated controls, $u = 0$;
- *singular* controls, taking values in the interior of U , that is $|u| < 1$.

We will call the first two kinds of controls *regular*, in opposition to the singular ones. For such extremal trajectories many efficient necessary or sufficient optimality conditions have been developed over the years [14, 62, 18, 55, 60, 31].

In this Chapter, we focus entirely on controls which are purely singular. We prove a generalized Legendre-Clebsch condition, which is a generalization version of the strong Legendre condition discussed in Section 2.4.2. For the L^1 -problem the Legendre form has a non-trivial time-dependent kernel, making it somewhere in between the usual regular trajectories, when the Legendre form is non-degenerate, and classical singular trajectories, when the Legendre form vanishes completely. We determine the second variation that allows us to prove the Legendre condition for this type of extremals and reformulate it in a ready-to-use form.

To formulate this result in a clean way, let us introduce the Hamiltonian functions

$$h_i(\lambda) = \langle \lambda, f_i \rangle, \quad \lambda \in T^*M, \quad i = 0, 1, \dots, m,$$

where f_i are the vector fields appearing in Equation (5.1.1). Moreover, we define $h_c = \frac{1}{2} \sum_{i=1}^m h_i^2$, which is another Hamiltonian function that will be extensively used.

Now we can state one of the main results of the chapter.

Theorem 5.1.1 (Necessary optimality condition). *Let $u \in \mathcal{U}$ be a singular control which is also an optimal solution of Problem 5.1. Let $\tilde{\lambda}(\cdot)$ be its corresponding extremal solving the Hamiltonian system of the PMP. Then*

$$\{h_c, \{h_0, h_c\}\}(\tilde{\lambda}(t)) \geq 0, \quad \text{for } t \in [0, T]. \quad (\text{GLC})$$

In the literature, the conditions like the one in (GLC) are usually called *generalized Legendre-Clebsch conditions*, or sometimes simply generalized Legendre conditions. They play the same role of the strong Legendre but in the context of singular extremals. The strengthening of this condition gives local strong optimality of short singular arcs.

Theorem 5.1.2 (Sufficient conditions, small time). *Let $u \in \mathcal{U}$ be a singular control for Problem 5.1. Denote by $\tilde{q}(\cdot)$ and $\tilde{\lambda}(\cdot)$ its corresponding singular trajectory and extremal respectively. If*

$$\{h_c, \{h_0, h_c\}\}(\tilde{\lambda}(t)) > 0, \quad \text{for } t \in [0, T], \quad (\text{SGLC})$$

then for every $t \in [0, T]$ there is some $\tau > 0$ such that $\tilde{q}|_{[t, t+\tau]}$ is strictly locally strongly optimal.

Conditions like (SGLC) are known in the literature as *strong* generalized Legendre conditions, where the word “strong” emphasizes that the inequality here is strict, while in (GLC) the equal sign is allowed. To prove this theorem, we use the method of overmaximized (or super-) Hamiltonians introduced in Chapter 4.

In order to complete the local optimality study, we need to quantify the time τ in Theorem 5.1.2. This is done via the theory of the Jacobi curve and conjugate points, which were introduced in the previous chapters. We characterize conjugate points by a Jacobi equation adapted to singular extremals, which we formulate in (5.4.3) of Section 5.4. Then, we prove the following two criteria.

Theorem 5.1.3 (Necessary condition, conjugate points). *Let $\tilde{\lambda}(\cdot)$ be a singular extremal for Problem 5.1 satisfying condition (SGLC) and let $\tilde{q}(\cdot)$ be its corresponding trajectory on M . Denote by Λ the Jacobi curve of the extremal pair $(\tilde{u}, \tilde{\lambda})$. Suppose that $\tilde{u}$ is of corank r . Let $N = \mu_\Pi(\Lambda) - \frac{1}{2}(d - \dim \cap_{t \in [0, T]} \Lambda_t)$. If $N \geq r$, then $\tilde{q}(\cdot)$ is not a weak-local minimizer of Problem 5.1 with respect to the L^∞ topology of $\mathcal{U}$.*

Theorem 5.1.4 (Sufficient condition, absence of conjugate points). *Let $\tilde{u} \in \mathcal{U}$ be a singular control for Problem 5.1. Let $\tilde{\lambda}(\cdot)$ be its corresponding extremal and $\tilde{q}(\cdot)$ its corresponding trajectory. Assume furthermore that the strong generalized Legendre condition (SGLC) holds and that there are no conjugate points along the trajectory $\tilde{q}(\cdot)$, for $t \in [0, T]$. Then $\tilde{q}(\cdot)$ is strictly locally strongly optimal.*

It should be stressed that Theorem 5.1.4 can be implemented numerically and used for concrete problems as it was done in [24] for regular extremals.

The structure of the Chapter is as follows. In Section 5.2, we describe the extremal trajectories of our problem. We then study second-order conditions. To simplify the exposition, we invert the order of the proofs. First, we prove the sufficient small-time optimality result of Theorem 5.1.2 in Section 5.3. We then introduce the Jacobi equation, prove Theorem 5.1.4 in Section 5.4, and provide examples in Section 5.5. The rest of the Chapter is dedicated to the study of necessary conditions and the derivation of the Jacobi equation. The reason is that the proofs of sufficient conditions are purely geometric, relying on methods of symplectic geometry and dynamical systems. In contrast, the proofs of necessary conditions are more technical and analytic in nature, and require a careful study of the infinite-dimensional Hessian and the spaces of variations. Therefore, we postpone them to the second half of the Chapter. This does not create any problem for the first part, since only the correct form of the generalized Legendre conditions and the Jacobi equation are needed for the proofs of the sufficient conditions. Finally, to make the main body of the text more readable, we prove in Section 5.7 a number of propositions that are required for the proofs of the three main theorems but involve lengthy technical calculations.

5.2 Description of extremal trajectories

In this section, we fix the notation used in this Chapter. We mainly follow the notation in [5]. We then describe extremal trajectories of Problem 5.1 using PMP and separate them into two categories: regular and singular extremals.

5.2.1 Definitions and notation

To shorten the formulas, we introduce the following notation. First, we use

$$h_{ij} := \{h_i, h_j\}, \quad h_{ijk} := \{h_i, \{h_j, h_k\}\}$$

to indicate nested Poisson brackets. Second, we set $f_I(q) := (f_1, \dots, f_m)$, so that the system (5.1.1) can be abbreviated to

$$\dot{q} = f_0(q) + \langle u, f_I(q) \rangle.$$

We do the same for $h_I(\lambda) := (h_1, \dots, h_m)$, and define $|h_I|(\lambda) := \sqrt{h_1^2(\lambda) + \dots + h_m^2(\lambda)}$ to be its Euclidean norm.

Finally we will often consider differentiation along the flow of a Hamiltonian system

$$\dot{\lambda} = \vec{H}(\lambda).$$

We will often abbreviate a function $h(\lambda)$ just to h when there is no confusion, and abbreviate $h(\lambda(t))$ to $h(t)$, when we consider the values of h along an integral curve $\lambda(t)$ of the Hamiltonian system above.

5.2.2 Description of extremal trajectories

To solve the OCP 5.1, we first apply PMP. By the notation of the previous subsection, the Hamiltonian function $h_u \in C^\infty(T^*M)$ in PMP can be written as:

$$h_u = h_0 + \langle u, h_I \rangle + \nu|u|, \quad h_I = (h_1, \dots, h_m).$$

We study in detail only the normal case, that is, when we normalize $\nu = -1$. The abnormal case, corresponding to $\nu = 0$, coincides with the time-optimal problem, which is already well studied (see [9, 8]). So, henceforth we always assume $\nu = -1$ for every extremal trajectory.

Using spherical coordinates on U , that is, by writing $u = rv$, $r = |u|$, $v \in S^{m-1}$, the Hamiltonian h_u takes the form

$$h_u = h_0 + r(\langle v, h_I \rangle - 1),$$

and the corresponding Hamiltonian vector field becomes

$$\vec{h}_u = \vec{h}_0 + r\langle v, \vec{h}_I \rangle.$$

According to PMP, an optimal control u must maximize h_u for almost every t . This can only happen if v is collinear with h_I . Therefore,

$$v = \frac{h_I}{|h_I|},$$

which yields

$$h_u = h_0 + r(|h_I| - 1).$$

If $|h_I| > 1$, the maximum is achieved when $r = 1$, while if $|h_I| < 1$, the Hamiltonian is maximized for $r = 0$. Therefore, if along an extremal $\lambda(t)$ we have $|h_I(t)| \neq 1$, at each time the value of the maximizing control is uniquely determined by

$$u(t) = \begin{cases} \frac{h_I(t)}{|h_I(t)|}, & \text{if } |h_I(t)| > 1, \\ 0, & \text{if } |h_I(t)| < 1. \end{cases}$$

We will call these controls *regular*.

Now, we want to analyse the case $|h_I(t)| \equiv 1$ on the time interval $[0, T]$, which gives no direct information on r since any $r \in [0, 1]$ realizes same value of the Hamiltonian. We will call these controls *singular*. Notice that, if $|h_I(t)| = 1$ only for a set of zero measure, then the maximizing control is still uniquely determined for a.e. $t \in [0, T]$ (recall that our space of admissible controls is $L^\infty([0, T], U)$). In general we know only that $|h_I|^{-1}(1)$ is a closed set. For simplicity of exposition, we will assume that $|h_I|^{-1}(1) = [0, T]$ is an interval, but

similar conclusions hold if $|h_I|^{-1}(1)$ has positive measure and $t \in |h_I|^{-1}(1)$ is any point of density.

The Hamiltonian vector field reads

$$\vec{h}_u = \vec{h}_0 + r\overrightarrow{h_I}, \quad r(t) \in [0, 1].$$

We define $h_c := \frac{1}{2}|h_I|^2$. Notice that on the hypersurface $|h_I| = 1$ we have $\vec{h}_c = \overrightarrow{\frac{1}{2}|h_I|^2} = |h_I|\overrightarrow{h_I} = \overrightarrow{h_I}$. So, every singular extremal λ can be described as a solution of the Hamiltonian system

$$\dot{\lambda} = \vec{h}_0(\lambda) + r(t)\vec{h}_c(\lambda), \quad r \in L^\infty([0, T]; [0, 1]).$$

Since we are on the surface $|h_I| = 1$ we also have $h_c(t) \equiv \frac{1}{2}|h_I(t)|^2 \equiv \frac{1}{2}$ for all $t \in [0, T]$. We can differentiate this equality in t :

$$0 = \frac{d}{dt}(h_c(t)) = \{h_0 + rh_c, h_c\}(t) = h_{0c}(t).$$

This equation does not give us any information on r . However, it puts an additional constraint on $\lambda(t)$ in order to be a singular extremal. If we differentiate in t the previous equality, we obtain:

$$\frac{d^2}{dt^2}(h_c(t)) = \{h_0 + rh_c, h_{0c}\}(t) = h_{00c}(t) + rh_{c0c}(t) = 0.$$

If $h_{c0c}(\lambda) \neq 0$, we obtain the following feedback equation for the control r :

$$r(\lambda) = -\frac{h_{00c}}{h_{c0c}}(\lambda) = \frac{h_{00c}}{h_{cc0}}(\lambda). \quad (5.2.1)$$

The second equality follows from the antisymmetry of the Poisson bracket, which gives $h_{cc0} = -h_{c0c}$. Hence, the singular control is completely determined, that is

$$u(\lambda) = r(\lambda)h_I(\lambda) \quad (5.2.2)$$

and the Hamiltonian system reads

$$\dot{\lambda} = \vec{h}_0(\lambda) + r(\lambda)\vec{h}_c(\lambda). \quad (5.2.3)$$

It is easy to check that the flow of (5.2.3) preserves the submanifold $\{\lambda \in T^*M \mid |h_I(\lambda)| = 1, h_{0c}(\lambda) = 0\}$. In the following, we always assume that the singular control r is admissible strictly singular, in the sense that $0 < r(t) < 1$ for every $t \in [0, T]$.

In this work, we focus only on the case $h_{c0c}(t) \neq 0$. The case $h_{c0c}(t) \equiv 0$ can happen as well and can be treated similarly although it is a bit more delicate. One can go on and differentiate the equalities $h_{c0c}(t) = 0$ and $h_{00c}(t) = 0$. Then, there are two possibilities: either a non-trivial relation involving r is found after a finite number of differentiations, or you obtain an infinite number of relations. In the first case, a similar analysis to the one that we are going to show can be performed. In the second case, since T^*M is finite-dimensional, there is some $k \in \mathbb{N}$ such that all the relations that we obtain after k differentiations must be linearly dependent. Again, from this fact, one can obtain some information on the optimality of the solutions.

5.3 The super-Hamiltonian function and strong local optimality of short arcs

We now prove the small-time sufficient optimality condition stated in Theorem 5.1.2. Recall that, along a singular extremal, the maximizing condition does not determine the control uniquely, since on the hypersurface $|h_I| = 1$ every control of the form $u = rh_I$, with $r \in [0, 1]$, realizes the maximum of the PMP Hamiltonian. Consequently, contrary to the case of regular extremal arcs, the maximized Hamiltonian does not provide a Hamiltonian flow. On the other hand, Section 5.2 shows that, differentiating the identity $|h_I| = 1$, the singular control is determined by the relation in Equation (5.2.2). The purpose of this section is to extend this dynamics to a smooth Hamiltonian dynamics in a open neighbourhood of the reference extremal.

The construction is based on the method of fields of extremals and on the super-Hamiltonian approach introduced in Chapter 4. For the present problem we only need a smooth fixed-final-time version of the general sufficient criterion, which we restate below in order to make the argument self-contained.

We denote by H the maximized Pontryagin Hamiltonian, namely

$$H(\ell) := \max_{u \in U} h_u(\ell),$$

where

$$h_u(\ell) = \langle \ell, f_0(q) + \langle u, f_I(q) \rangle \rangle - |u|, \quad \ell \in T_q^*M.$$

Let $a \in C^\infty(M)$ and consider the exact Lagrangian graph

$$\mathcal{L}_0 := \{d_q a \mid q \in M\} \subset T^*M.$$

Let $H_S \in C^\infty(\mathbb{R} \times T^*M)$ be a time-dependent Hamiltonian and denote by Φ_t^S the flow of $\vec{H}_S$ from time 0 to time t . For $\ell_0 \in \mathcal{L}_0$, set

$$\ell_t := \Phi_t^S(\ell_0), \quad \mathcal{L}_t := \Phi_t^S(\mathcal{L}_0),$$

and define

$$\mathcal{L} := \{(t, \ell) \in [0, T] \times T^*M \mid \ell \in \mathcal{L}_t\}.$$

Theorem 5.3.1. *Let u be a singular control and let q and λ be its corresponding trajectory and normal extremal. Suppose that there exist a function $a \in C^\infty(M)$ and a time-dependent Hamiltonian $H_S \in C^\infty(\mathbb{R} \times T^*M)$ such that, with the notation introduced above, the following properties are satisfied.*

1. *For every $\ell_0 \in \mathcal{L}_0$ and every $t \in [0, T]$, with $\ell_t = \Phi_t^S(\ell_0)$, it holds that $H_S(t, \ell_t) \geq H(\ell_t)$.*
2. *Along the reference extremal one has $H_S(t, \lambda_t) = H(\lambda_t)$ for almost every $t \in [0, T]$.*

3. Along the reference extremal the previous inequality is a true equality: $\vec{H}_S(t, \lambda_t) = \vec{h}_{u(t)}(\lambda_t)$ for almost every $t \in [0, T]$.
4. There exist open neighbourhoods $\mathcal{V} \subset T^*M$ and $\mathcal{O} \subset M$ of $\lambda([0, T])$ and $q([0, T])$, respectively, such that

$$\text{id} \times \pi : \mathcal{L} \cap ([0, T] \times \mathcal{V}) \longrightarrow [0, T] \times \mathcal{O}, \quad (t, \ell) \longmapsto (t, \pi(\ell)),$$

is a smooth diffeomorphism.

Then q is strictly locally strongly optimal.

This statement is the smooth fixed-final-time counterpart of the more general sufficient criterion of Theorem 4.1.3.

We now apply this criterion to the singular extremal under consideration. The strong generalized Legendre condition

$$h_{c0c}(\lambda_t) > 0, \quad t \in [0, T],$$

is the strict version of the necessary condition stated in Theorem 5.1.1, whose proof will be given in Section 5.6. We show that this condition provides exactly the non-degeneracy needed to construct a super-Hamiltonian satisfying Theorem 5.3.1 on sufficiently short time intervals.

The construction proceeds in three steps. First, we choose an initial submanifold $\mathcal{L}_0$ through the initial covector and contained in the hypersurface $|h_I| = 1$. Second, using the transversality implied by $h_{c0c} > 0$, we construct a smooth Hamiltonian h_S by extending h_0 along the integral curves of $\vec{h}_c$. Finally, we combine h_S with the singular control $r(t)$ to obtain the time-dependent Hamiltonian H_S and verify the four assumptions of Theorem 5.3.1.

Proof of Theorem 5.1.2. We fix a singular extremal trajectory $\lambda(t)$ and $t_0 \in [0, T]$. We consider the restriction of the reference extremal to a sufficiently short interval $[t_0, t_0 + \tau]$. Since the statement is local in time, we may assume that $t_0 = 0$. Set

$$\lambda_0 := \lambda(0), \quad q_0 := q(0).$$

and all the objects constructed below are considered only in sufficiently small neighbourhoods of λ_0 and q_0 .

We apply Theorem 5.3.1. More precisely, we construct a Lagrangian manifold $\mathcal{L}_0 = \text{graph}(da)$, a time-independent Hamiltonian h_S , and a time-dependent Hamiltonian H_S such that the four assumptions of that theorem hold on $[0, \tau]$.

Define

$$\Sigma := \{\ell \in T^*M \mid |h_I(\ell)| = 1\}, \quad \mathcal{S} := \{\ell \in T^*M \mid h_{0c}(\ell) = 0\}.$$

We have $\lambda(t) \in \Sigma \cap \mathcal{S}$. Moreover, the strong generalized Legendre condition gives

$$h_{c0c}(\lambda_0) > 0.$$

Step 1 - Construction of the initial Lagrangian manifold. Consider the Hamilton–Jacobi equation

$$h_c(q, d_q a) = \frac{1}{2}, \quad \text{equivalently} \quad \sum_{i=1}^m |\langle d_q a, f_i(q) \rangle|^2 = 1. \quad (5.3.1)$$

Choose a hypersurface $N \subset M$ through q_0 such that $T_{q_0}N = \ker \lambda_0$, and prescribe $a|_N = 0$. This Cauchy datum is non-characteristic for (5.3.1) at (q_0, λ_0) . Indeed, the differential in the fibre directions of h_c at λ_0 is identified with

$$\sum_{i=1}^m h_i(\lambda_0) f_i(q_0),$$

and

$$\left\langle \lambda_0, \sum_{i=1}^m h_i(\lambda_0) f_i(q_0) \right\rangle = |h_I(\lambda_0)|^2 = 1.$$

Hence this vector is not tangent to N . By the method of characteristics, after possibly shrinking the neighbourhood of q_0 , there exists $a \in C^\infty(M)$ satisfying (5.3.1), $a|_N = 0$, and $d_{q_0} a = \lambda_0$. We set

$$\mathcal{L}_0 := \{(q, d_q a) \mid q \text{ close to } q_0\}.$$

Then $\mathcal{L}_0$ is Lagrangian, $\lambda_0 \in \mathcal{L}_0$, and (5.3.1) implies $\mathcal{L}_0 \subset \Sigma$.

Step 2 - Construction and properties of h_S . Since

$$\vec{h}_c(h_{0c}) = h_{c0c},$$

the vector field $\vec{h}_c$ is transverse to $\mathcal{S}$ in a neighbourhood of λ_0 . Thus, after shrinking that neighbourhood, the map

$$(s, \ell_0) \mapsto \exp(s\vec{h}_c)(\ell_0), \quad \ell_0 \in \mathcal{S},$$

is a system of local coordinates. We define h_S in these coordinates by

$$h_S(\exp(s\vec{h}_c)(\ell_0)) := h_0(\ell_0), \quad \ell_0 \in \mathcal{S}. \quad (5.3.2)$$

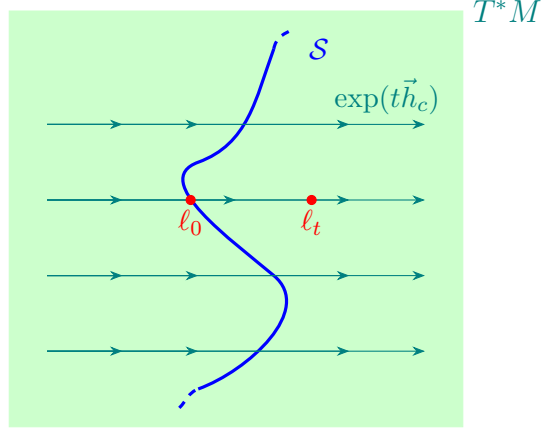
In particular,

$$h_S = h_0 \quad \text{on } \mathcal{S}, \quad \{h_S, h_c\} = 0. \quad (5.3.3)$$

The second identity follows because h_S is constant along the integral curves of $\vec{h}_c$. Notice also that the flow of $\vec{h}_c$ preserves Σ , since $h_c = \frac{1}{2}|h_I|^2$.

We prove the following identity, that will be used below:

$$d_\ell(h_S - h_0) = 0, \quad \forall \ell \in \mathcal{S} \cap \Sigma. \quad (5.3.4)$$

Figure 5.1: Construction of h_S .

To prove this identity, we can consider the splitting

$$T_\ell(T^*M) = T_\ell\mathcal{S} \oplus \mathbb{R}\vec{h}_c(\ell), \quad \ell \in \mathcal{S} \cap \Sigma.$$

By definition of h_S , we have that $h_S(\ell) = h_0(\ell)$ for any $\ell \in \mathcal{S}$, so, if $v \in T_\ell\mathcal{S}$, then $\langle d_\ell(h_S - h_0), v \rangle = 0$. Moreover, we have

$$\langle d_\ell(h_S - h_0), \vec{h}_c \rangle = \langle d_\ell h_S, \vec{h}_c \rangle - \langle d_\ell h_0, \vec{h}_c \rangle = \{h_S, h_c\}(\ell) - \{h_0, h_c\}(\ell) = 0,$$

where the first Poisson bracket vanishes by the definition of h_S and the second by definition of $\mathcal{S}$.

Step 3 - Construction of H_S and verification of the hypotheses of Theorem 5.3.1.

Let

$$r(t) := r(\lambda(t)) = \frac{h_{00c}(\lambda(t))}{h_{cc0}(\lambda(t))},$$

so that

$$u(t) = r(t)h_I(\lambda(t))$$

by (5.2.1) and (5.2.2). Define

$$H_S(t, \ell) := h_S(\ell) + r(t)(|h_I(\ell)| - 1). \quad (5.3.5)$$

After restricting to a neighbourhood of Σ , the function $|h_I|$ is smooth. By (5.3.3) and the identity $\{|h_I|, h_c\} = 0$, we have

$$\{H_S(t, \cdot), h_c\} = 0 \quad \text{for every } t.$$

Hence the flow Φ_t^S of $\vec{H}_S$ preserves Σ . Since $\mathcal{L}_0 \subset \Sigma$, this gives

$$\mathcal{L}_t \subset \Sigma \quad \text{for } t \in [0, \tau], \quad (5.3.6)$$

after possibly reducing τ .

We now verify the four assumptions of Theorem 5.3.1.

Condition (2). For every $t \in [0, \tau]$, the reference extremal belongs to $\Sigma \cap \mathcal{S}$. Therefore, by (5.3.3) and (5.3.5),

$$H_S(t, \lambda(t)) = h_S(\lambda(t)) = h_0(\lambda(t)).$$

Since $H = h_0$ on Σ , condition (2) follows.

Condition (3). At any point of Σ , one has

$$dh_c = d|h_I|, \quad \vec{h}_c = \overrightarrow{|h_I|}.$$

Consequently, by (5.3.4),

$$\vec{H}_S(t, \lambda(t)) = \vec{h}_S(\lambda(t)) + r(t)|\vec{h}_I|(\lambda(t)) = \vec{h}_0(\lambda(t)) + r(t)\vec{h}_c(\lambda(t)).$$

The right-hand side is precisely $\vec{h}_u(\lambda(t))$ by (5.2.3). Thus condition (3) holds.

Condition (1). In view of (5.3.6), it is enough to prove that $h_S \geq h_0$ on Σ near λ_0 . Let $\ell_0 \in \mathcal{S} \cap \Sigma$ be close to λ_0 , and set

$$g_{\ell_0}(s) := h_S(\exp(s\vec{h}_c)(\ell_0)) - h_0(\exp(s\vec{h}_c)(\ell_0)).$$

By (5.3.2) and the fact that $\ell_0 \in \mathcal{S}$,

$$g_{\ell_0}(0) = 0, \quad g'_{\ell_0}(0) = 0, \quad g''_{\ell_0}(0) = -h_{cc0}(\ell_0) = h_{c0c}(\ell_0).$$

By the strong generalized Legendre condition and the continuity of h_{c0c} , we can conclude that there is a neighbourhood $\mathcal{W}_0 \subset T^*M$ of λ_0 such that $h_S(\ell) \geq h_0(\ell)$, for $\ell \in \mathcal{W}_0$. Therefore, for $\ell \in \mathcal{W}_0 \cap \Sigma$, we have $H_S(\ell) = h_S(\ell) \geq h_0(\ell) = H(\ell)$. Recalling that $\mathcal{L}_t = \Phi_t^S(\mathcal{L}_0) \subset \Sigma$ and after possibly shrinking $\mathcal{L}_0$ and τ , we have that condition (1) follows.

Condition (4). Since $\mathcal{L}_0$ is the graph of da , the restriction

$$d_{\lambda_0}\pi|_{T_{\lambda_0}\mathcal{L}_0} : T_{\lambda_0}\mathcal{L}_0 \longrightarrow T_{q_0}M$$

is an isomorphism. Consider

$$(\text{id} \times \pi) : [0, \tau] \times \mathcal{L}_0 \longrightarrow [0, \tau] \times M, \quad (\text{id} \times \pi)(t, \ell_0) := (t, \pi(\Phi_t^S(\ell_0))).$$

Its differential at $(0, \lambda_0)$ is given by

$$D(\text{id} \times \pi)_{(0, \lambda_0)}(s, v) = \left(s, d_{\lambda_0}\pi(v) + s d_{\lambda_0}\pi(\vec{H}_S(0, \lambda_0)) \right), \quad (s, v) \in \mathbb{R} \times T_{\lambda_0}\mathcal{L}_0,$$

and is therefore an isomorphism. The inverse function theorem, together with the smooth dependence of Φ_t^S on (t, ℓ_0) , yields the required local diffeomorphism property in condition (4), after possibly shrinking τ and the neighbourhood defining $\mathcal{L}_0$.

We have verified all the assumptions of Theorem 5.3.1. Therefore, $q|_{[0, \tau]}$ is strictly locally strongly optimal. □

5.4 Jacobi equation and sufficient conditions for optimality

In the previous Section we proved that, under the strong generalized Legendre condition, short arcs of a singular extremal are locally strongly optimal. Now, we want to give a procedure to quantify how long these optimal singular arcs can be. To this aim, we introduce a suitable Jacobi equation and the notion of conjugate points, which are analogous to the one introduced in Section 3.3.2.

5.4.1 Jacobi curve and conjugate points

As before we fix a singular extremal $\tilde{\lambda}(t)$ that satisfies the strong generalized Legendre condition (SGLC). Again, we denote $\tilde{\lambda}_0 = \tilde{\lambda}(0)$.

We recall the definition of the transported-back Hamiltonian function:

$$\mathcal{H}_t(u, \ell) = (h_u - h_{\tilde{u}(t)}) \circ \Phi_t(\ell), \quad u \in U, \ell \in T^*M, t \in [0, T],$$

where $\tilde{u}(t)$ denotes the value of the singular control (5.2.2) and Φ_t is the flow on T^*M of the Hamiltonian system (5.2.3).

Recall also the definition of $Z_t : T_{\tilde{u}(t)}U \rightarrow T_{\lambda_0}(T^*M)$:

$$Z_tv(t) := \left\langle v(t), \frac{\partial \vec{\mathcal{H}}_t}{\partial u}(u, \tilde{\lambda}(0)) \right\rangle_{u=\tilde{u}(t)} = \sum_{j=1}^m v_j(t) [(\Phi_t^{-1})_* \vec{h}_j](\lambda_0), \quad (5.4.1)$$

We can split $T_{\tilde{u}(t)}U = (\mathbb{R}h_I(t)) \oplus (h_I(t))^\perp$ by defining $v(t) = \phi(t)h_I(t) + w(t)$. This splitting separates the two qualitatively different types of control variations: the component parallel to $h_I(t)$ changes the control in the singular direction, whereas the component orthogonal to $h_I(t)$ changes its direction along the regular ones. To shorten the notation we will abbreviate $Z_I(t) := Z_th_I(t)$. Then, using the previously defined splitting, we define

$$\mathcal{Z}_tv(t) := Z_tw(t) - \dot{Z}_I(t)\phi(t).$$

The term $-\dot{Z}_I(t)\phi(t)$ comes from an integration by parts, which is better motivated in Section 5.6.

The next object that we introduce is the generalized Legendre form l_t . Its role is analogous to that of the classical Legendre form in the regular case. It is a quadratic form on $T_{\tilde{u}(t)}U \oplus \mathbb{R}$ defined by

$$l_t(w, \phi) := \begin{pmatrix} w & \phi \end{pmatrix} \begin{pmatrix} \frac{\text{Id}}{r} & \sigma(Z_t \cdot, Z_I(t)) \\ \sigma(Z_t \cdot, Z_I(t))^T & \sigma(Z_I(t), \dot{Z}_I(t)) \end{pmatrix} \begin{pmatrix} w \\ \phi \end{pmatrix}, \quad (5.4.2)$$

where $w \in (h_I(t))^\perp$.

Now, we introduce the Jacobi equation and the notion of conjugate time and conjugate point, which are similar to the one given for regular extremals.

Definition 5.4.1. Let $\tilde{\lambda}(t)$ be an extremal satisfying the strong generalized Legendre condition and denote $\Pi_0 := T_{\lambda_0}^*(T_{q_0}^*M)$, where $\lambda_0 = \tilde{\lambda}(0)$ and $q_0 = \pi(\lambda_0)$. A time T is conjugate to 0 on $\tilde{\lambda}$ if and only if there is a non-constant solution to the following boundary value problem

$$\dot{\eta} = -Z_t l_t^{-1} \sigma(Z_t \cdot, \eta) \quad \text{for a.e. } t \in [0, T], \quad (5.4.3)$$

$$\eta(0) \in (\Pi_0 + \mathbb{R}Z_I(0)) \cap (\mathbb{R}Z_I(0))^\perp = \Pi_0^{\mathbb{R}Z_I(0)}, \quad \eta(T) \in \Pi_0. \quad (5.4.4)$$

Equation (5.4.3) is called the *Jacobi equation* along the extremal $\tilde{\lambda}$. Moreover, if T is conjugate to 0 on $\tilde{\lambda}$, we say that the point $q(T)$ is *conjugate* to $q(0)$ along the trajectory q . The multiplicity of a conjugate point is the number of linearly independent solutions of (5.4.3), (5.4.4).

Theorem 5.4.2. Let $(\tilde{u}, \tilde{\lambda})$ be a singular extremal pair and let Λ denote the corresponding Jacobi curve. Then,

$$\Lambda_t = \{\eta(t) \in T_{\lambda_0}^*(T^*M) \mid \eta \text{ solves (5.4.3), } \eta(0) \in \Pi_0^{\mathbb{R}Z_I(0)}\}$$

We prove this result in 5.6.4, where the Jacobi equation is derived from the second variation.

We have the following property that will be used several times.

Lemma 5.4.3. $\eta_I(t) = Z_I(t)$ is a solution of (5.4.3).

Proof. If we apply the matrix in (5.4.2) to the variation $w(t) = 0$ and $\phi(t) = 1$, we obtain

$$\begin{pmatrix} \frac{\text{Id}}{r} & \sigma(Z_t \cdot, Z_I(t)) \\ \sigma(Z_t \cdot, Z_I(t))^T & \sigma(Z_I(t), \dot{Z}_I(t)) \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \sigma(Z_t \cdot, Z_I(t)) - \sigma(\dot{Z}_I(t), Z_I(t)) = \sigma(Z_t \cdot, Z_I(t)).$$

Hence, identifying l_t^{-1} with the inverse of the matrix in (5.4.2), we have

$$-Z_t l_t^{-1} \sigma(Z_t \cdot, Z_I(t)) = -Z_t(0, 1) = \dot{Z}_I(t).$$

This shows that $\dot{\eta}_I(t) = \dot{Z}_I(t) = -Z_t l_t^{-1} \sigma(Z_t, Z_I(t)) = -Z_t l_t^{-1} \sigma(Z_t, \eta_I(t))$, that is $Z_I(t)$ is a solution of the Jacobi equation. $\square$

The solution $Z_I(t)$ is distinguished because it is generated by the singular direction itself. In general, it does not satisfy the final boundary condition in Definition 5.4.1. Its appearance in the initial condition $\Pi_0^{\mathbb{R}Z_I(0)}$ is intuitively explained by the fact that the singular extremals are contained in a codimension 2 submanifold of T^*M . As a consequence, a field of extremals is also forced to be contained in a smaller submanifold, and then its tangent space is forced to contain the direction $Z_I(t)$.

An explicit example involving the Jacobi equation written in this form is given in Subsection 5.5.3, where we study a concrete example on $SU(2)$. One can also give an equivalent definition of the Jacobi equation using the previously introduced super-Hamiltonian H_S . To do this, one should note that the Jacobi equation is a linear Hamiltonian equation on $T_{\tilde{\lambda}_0}^*(T^*M)$ with a time-dependent quadratic Hamiltonian

$$b_t := \frac{\langle \sigma_{\lambda_0}(Z_t, \eta), l_t^{-1} \sigma_{\lambda_0}(Z_t, \eta) \rangle}{2}. \quad (5.4.5)$$

It should be noted that invertibility of l_t is guaranteed by the strong generalized Legendre condition (see Lemma 5.4.5, which we prove after we introduce the second variation). So it is enough to rewrite this Hamiltonian in terms introduced in the previous sections.

Finally, we can prove Theorem 5.1.3.

Proof of Theorem 5.1.3. Let $\tilde{\lambda}$ be a singular extremal for Problem 5.1 corresponding to the singular control u and let q be its trajectory on M . Let $N = \mu_{\Pi}(\Lambda) - \frac{1}{2}(d - \dim \cap_{t \in [0, T]} \Lambda_t)$. By Theorem 3.0.1, if $\tilde{u}$ is of corank r , we have that if $N \geq r$, then the endpoint map E is locally open at $\tilde{u}$. In particular, u is not weakly-locally optimal. $\square$

In this problem, there are two dynamics: the Hamiltonian dynamics associated with the reference singular control (encoded by the flow Φ_t), and the smooth super-Hamiltonian dynamics (encoded by the flow Φ_t^S) used to construct a field of nearby extremals. Consider the composition

$$\Phi_t^{-1} \circ \Phi_t^S$$

Since the two flows agree along the reference extremal, this map fixes $\tilde{\lambda}_0$. Its quadratic part at $\tilde{\lambda}_0$ therefore measures the first non-trivial discrepancy between the two dynamics, and, as shown in Corollary 5.4.6 below, this discrepancy is exactly the quadratic Hamiltonian governing the Jacobi equation.

We want to write it as the flow of another Hamiltonian function, i.e.

$$\Phi_{-t} \circ \Phi_t^S = \overrightarrow{\exp} \int_0^t \vec{\varphi}_s ds, \quad (5.4.6)$$

where here we follow the notation of the chronological calculus (see Appendix A.1), that is $\overrightarrow{\exp} \int_{\tau_0}^{\tau_1} \vec{h}_t dt$ denotes the (non-autonomous) flow of $\vec{h}_t$ from time τ_1 to time τ_2 . It follows by the variations formula (see (A.1.1)) that this is

$$\varphi_t(\ell) = (H_S(t, \cdot) - H(\tilde{u}(t), \cdot)) \circ \Phi_t(\ell), \quad \ell \in T^*M, t \in [0, T].$$

Indeed, we have

$$\vec{\varphi}_t = \overrightarrow{(H_S(t, \cdot) - H(\tilde{u}(t), \cdot)) \circ \Phi_t} = (\Phi_{-t})_* \left(\vec{H}_S(t, \cdot) - \vec{H}(\tilde{u}(t), \cdot) \right),$$

hence, by the variations formula (see Equation (A.1.1)), we obtain

$$\begin{aligned} \overrightarrow{\exp} \int_0^t \vec{\varphi}_s ds &= \overrightarrow{\exp} \int_0^t (\Phi_{-s})_* \vec{H}_S(s, \cdot) - (\Phi_{-s})_* \vec{H}(u(s), \cdot) ds = \\ &= \overrightarrow{\exp} \int_0^t -(\Phi_{-s})_* \vec{H}(u(s), \cdot) ds \circ \overrightarrow{\exp} \int_0^t (\Phi_{-s})_* (\Phi_s)_* \vec{H}_S(s, \cdot) ds = \\ &= \left(\overrightarrow{\exp} \int_0^t \vec{H}(u(s), \cdot) ds \right)^{-1} \circ \overrightarrow{\exp} \int_0^t \vec{H}_S(s, \cdot) ds. \end{aligned}$$

We have the following relation with the Jacobi equation.

Proposition 5.4.4. *Under the previous notations, we have that for $\tilde{\lambda}_0 \in \Sigma$*

$$\text{Hess}_{\lambda_0} [(H_S(t, \cdot) - H(\tilde{u}(t), \cdot)) \circ \Phi_t] = \langle \sigma_{\lambda_0}(\mathcal{Z}_t, \eta), l_t^{-1} \sigma_{\lambda_0}(\mathcal{Z}_t, \eta) \rangle \quad \eta \in T_{\lambda_0}(T^*M).$$

This proposition is a particular case of Proposition 8.3.2. Its proof is contained in Appendix 8.4.

Finally, we note that there is an important connection between the Legendre condition in Theorem 5.1.2 and the Legendre form l_t .

Proposition 5.4.5. *Given a singular extremal $\tilde{\lambda}(t)$, we have the following equivalence:*

$$l_t \geq 0 \iff h_{c0c}(t) \geq 0.$$

and a similar equivalence holds for strong inequalities.

The proof is a lengthy computation that we postpone to Subsection 5.6.3. An important corollary of this proposition is that the strong Legendre condition ensures that the Hamiltonian b_t in (5.4.5) is non-negative. We will use this fact in the proof of Theorem 7.6.1 presented in the next subsection.

We conclude this subsection with a corollary of Proposition 5.4.4.

Corollary 5.4.6. *Given a singular extremal $\tilde{\lambda}(t)$, define φ_t as in (5.4.6) and let b_t be the quadratic Hamiltonian defined in (5.4.5). Then, it holds the following identity:*

$$\left(\overrightarrow{\exp} \int_0^t \vec{\varphi}_s ds \right)_{*, \lambda(0)} = \overrightarrow{\exp} \int_0^t \vec{b}_s ds.$$

We omit the proof since it is obtained by the very same argument used in Corollary 3.3.13. Finally, the same argument in Remark 3.3.14 yields that solutions to the Jacobi equation (5.4.3) starting from $\Pi^{\mathbb{R}Z_I(0)}$ are in one-to-one correspondence with solutions of $\dot{\xi} = d_{\tilde{\lambda}(t)} \tilde{H}_S[\xi]$ starting from $\Pi^{\mathbb{R}Z_I(0)}$.

5.4.2 Sufficient condition for optimality

This Subsection is entirely devoted to the proof of Theorem 5.1.4. The argument follows closely the idea of the proof of Theorem 21.3 of [13].

We recall the statement of the Theorem that we are going to prove.

Theorem 5.4.7. *Let $\tilde{\lambda}(t)$, $t \in [0, T]$, be a singular extremal for Problem 5.1. Assume furthermore that it satisfies the strong generalized Legendre condition (SGLC) and that the Jacobi equation (5.4.3), (5.4.4) has no non-trivial solution. Then $\tilde{\lambda}(t)$ is strictly locally strongly optimal.*

Proof. We apply Theorem 5.3.1, using the overmaximized Hamiltonian H_S constructed in Section 5.3. As in the proof of Theorem 5.1.2, it is enough to construct a Lagrangian submanifold $\mathcal{L}_0 \subset \Sigma$ through λ_0 such that, if

$$\mathcal{L}_t := \Phi_t^S(\mathcal{L}_0), \quad t \in [0, T],$$

then the restriction of the canonical projection $\pi : T^*M \rightarrow M$ to $\mathcal{L}_t$ is a local diffeomorphism at $\tilde{\lambda}(t)$. We organize the construction in four steps.

Step 1 - Reduction to finding a suitable initial linear subspace.

First, we show how the problem of finding $\mathcal{L}_0$ can be reduced to find a suitable initial tangent space at $\tilde{\lambda}_0$. Set

$$\Pi_t := T_{\tilde{\lambda}(t)}^*(T_{q(t)}^*M).$$

Since $\mathcal{L}_t$ has the same dimension as M , by the inverse function theorem we have that $\pi|_{\mathcal{L}_t}$ is a local diffeomorphism at $\tilde{\lambda}(t)$ if and only if

$$T_{\tilde{\lambda}(t)}\mathcal{L}_t \cap \Pi_t = \{0\}, \quad t \in [0, T].$$

Let $L_0 := T_{\lambda_0}\mathcal{L}_0$. Since $T_{\tilde{\lambda}(t)}\mathcal{L}_t = (\Phi_t^S)_*L_0$ and Φ_t is the cotangent lift of a flow on M , we have $\Pi_t = (\Phi_t)_*\Pi_0$. So, $d_{\tilde{\lambda}(t)}\pi|_{\mathcal{L}_t}$ is invertible if and only if

$$(\Phi_t^S)_*L_0 \cap (\Phi_t)_*\Pi_0 = \{0\}, \quad t \in [0, T]. \quad (5.4.7)$$

Moreover, recall that

$$\Phi_t^S = \Phi_t \circ \overrightarrow{\exp} \int_0^t \vec{\varphi}_s ds,$$

where φ_t is defined in (5.4.6). Therefore, condition (5.4.7) is equivalent to

$$B_t L_0 \cap \Pi_0 = \{0\}, \quad t \in [0, T], \quad (5.4.8)$$

where

$$B_t := \left(\overrightarrow{\exp} \int_0^t \vec{\varphi}_s ds \right)_*.$$

By Corollary 5.4.6, we have that B_t coincides with the flow of the Jacobi equation. Thus the remaining problem is purely linear: we must find a Lagrangian subspace $L_0 \subset T_{\lambda_0} \Sigma$ satisfying (5.4.8).

Step 2 - relation with the Jacobi equation and absence of conjugate points.

Define

$$\Lambda_0 := \Pi_0^{\mathbb{R}Z_I(0)}.$$

This is the Lagrangian space of admissible initial conditions in the Jacobi boundary-value problem (5.4.3), (5.4.4). The absence of conjugate points implies that, for every $t \in (0, T]$,

$$B_t \Lambda_0 \cap \Pi_0 = C_t, \quad (5.4.9)$$

where $C_t := \{v \in \Pi_0 \mid B_s v = v \text{ for every } s \in [0, t]\}$ is the space of constant vertical solutions. We first treat the case $C_t = \{0\}$: under this additional hypothesis, we have $B_t \Lambda_0 \cap \Pi_0 = \{0\}$ for $t \in (0, T]$. Hence, to find a suitable initial space it suffices to find L_0 close to Λ_0 and transverse to Π_0 also for $t = 0$. To achieve this, we use the local coordinates of the Lagrangian Grassmannian introduced in 2.2.2. After that, we reduce the general case $C_t \neq \{0\}$ to the case $C_t = \{0\}$.

Step 3 - construction of L_0 if $C_t = \{0\}$ for every $t \in (0, T]$.

We first explain the reduction imposed by the tangency condition $L_0 \subset T_{\lambda_0} \Sigma$. Since $Z_I(t) \notin \Pi_0$ for every $t \in [0, T]$, choose $\zeta(t) \in \Pi_0$ such that

$$\sigma(\zeta(t), Z_I(t)) \neq 0,$$

and choose a symplectic subspace $\Gamma(t) \subset T_{\lambda_0} \Sigma$ such that

$$T_{\lambda_0}(T^*M) = \Gamma(t) \oplus \mathbb{R}Z_I(t) \oplus \mathbb{R}\zeta(t).$$

The subspace

$$\bar{\Pi}_0(t) := \Gamma(t) \cap \Pi_0 = \Gamma(t) \cap \Lambda_0$$

is Lagrangian in $\Gamma(t)$. This is a concrete symplectic model for the quotient by the characteristic direction $\mathbb{R}Z_I(t)$. In the sequel, the notation B_t also denotes the induced symplectic flow in $\Gamma(t)$.

It is enough to construct a Lagrangian subspace $\bar{L}_0 \subset \Gamma(0)$ such that

$$B_t \bar{L}_0 \cap \bar{\Pi}_0(t) = \{0\}, \quad t \in [0, T]. \quad (5.4.10)$$

Indeed, once (5.4.10) holds, the space

$$L_0 := \bar{L}_0 \oplus \mathbb{R}Z_I(0) \quad (5.4.11)$$

is Lagrangian, belongs to $T_{\lambda_0}\Sigma$, and satisfies (5.4.8).

Choose a Lagrangian subspace $H(0) \subset \Gamma(0)$ transverse to $\bar{\Pi}_0(0)$, and set $H(t) = B_t H(0)$. In the corresponding local Lagrangian charts, $\bar{\Pi}_0(t)$ is represented by the zero symmetric matrix. Since $b_t \geq 0$ by Proposition 5.4.5, Proposition 3.3.10 yields the following monotonicity property: if Ξ_0 is represented by a positive definite symmetric matrix $S_0 > 0$ and $\Xi_t = B_t \Xi_0$ is represented by S_t , then

$$\dot{S}_t \geq 0$$

as long as the curve stays in the chart. Hence $S_t \geq S_0 > 0$. In particular, after choosing $\varepsilon > 0$ sufficiently small, every such curve with $S_0 > 0$ sufficiently close to zero satisfies

$$B_t \bar{L}_0 \cap \bar{\Pi}_0(t) = \{0\}, \quad t \in [0, \varepsilon]. \quad (5.4.12)$$

It remains to show that there is S_0 such that Equation (5.4.12) holds also for $[\varepsilon, T]$. Suppose, by contradiction, that no sufficiently small positive S_0 has this property. Then, there would exist a sequence of positive definite matrices $S_0^{(k)}$, $k \in \mathbb{N}$, such that $S_0^{(k)} \rightarrow 0$ as $k \rightarrow +\infty$ and times $t_k \in [\varepsilon, T]$ for which the corresponding Lagrangian subspaces $\Xi_{t_k}^{(k)}$ meet $\bar{\Pi}_0(t_k)$. Passing to a subsequence, we may assume that $t_k \rightarrow t_\infty \in [\varepsilon, T]$. By continuity of the linear flow and of the Lagrangian Grassmannian, this would imply

$$B_{t_\infty} \Lambda_0 \cap \Pi_0 \neq \{0\},$$

contradicting (5.4.9) and the assumption $C_{t_\infty} = \{0\}$. Thus, in the case $C_t = \{0\}$, one can choose $\bar{L}_0$ satisfying (5.4.10) on the whole interval $[0, T]$.

Step 4 - reduction of the general case to Step 3.

We now allow $C_t \neq \{0\}$. The family $(C_t)_{t>0}$ is monotone non-increasing:

$$C_{t_2} \subset C_{t_1}, \quad 0 < t_1 \leq t_2 \leq T.$$

We set $C_0 := \bar{\Pi}_0(0)$. Since this family is left-continuous and the ambient space is finite-dimensional, it has only finitely many discontinuity times. Let

$$0 = s_0 < s_1 < \cdots < s_k < s_{k+1} = T$$

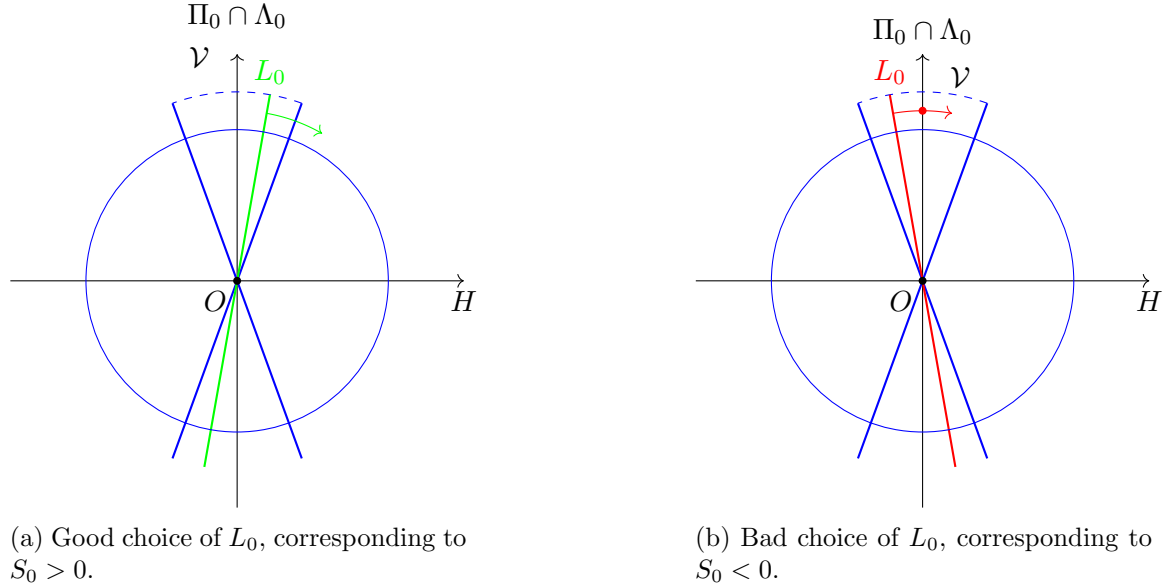


Figure 5.2: The positive side of the local chart gives an initial Lagrangian subspace which does not meet the vertical space for small positive times.

be these times. Choosing complements successively, we may write

$$C_t = E_{i+1} \oplus \cdots \oplus E_k, \quad t \in (s_i, s_{i+1}], \quad i = 0, \dots, k.$$

Thus E_i is the part of the space of constant vertical solutions which disappears at the time s_i .

Choose a dual splitting in the reduced symplectic space,

$$H = F_1 \oplus \cdots \oplus F_k, \quad \sigma_{\lambda_0}(E_i, F_j) = 0 \quad \text{whenever } i \neq j.$$

The next finite-dimensional lemma is the standard iterative version of the construction in Step 3; it is Lemma 21.3 in [13].

Lemma 5.4.8. *Fix any Lagrangian subspace $H_0 \subset \Gamma(0)$ such that $\bar{\Pi}_0(0) \cap H_0 = \{0\}$. Then there exist Lagrangian subspaces $H_1, \dots, H_k$, each transverse to $\bar{\Pi}_0(0)$, and positive numbers*

$$\varepsilon_1 \geq \cdots \geq \varepsilon_k > 0$$

with the following property. Let $\bar{L}_0$ be the Lagrangian subspace represented by εId in the local chart $(\bar{\Pi}_0(0), H_0)$, where $0 < \varepsilon < \varepsilon_i$. Then

1. $B_t \bar{L}_0 \cap \bar{\Pi}_0(t) = \{0\}$ for every $t \in [0, s_i]$;

2. $B_t \bar{L}_0 \cap H_i = \{0\}$ for every $t \in [0, s_i]$, and the corresponding symmetric matrix in the chart $(\bar{\Pi}_0(0), H_i)$ is positive definite.

Applying the lemma with $i = k$ and choosing $0 < \varepsilon < \varepsilon_k$, we obtain a Lagrangian subspace $\bar{L}_0$ satisfying (5.4.10) for every $t \in [0, T]$. Define L_0 by (5.4.11). Then L_0 is Lagrangian, is contained in $T_{\lambda_0} \Sigma$, and satisfies (5.4.8).

Finally, since L_0 is transverse to Π_0 , there exists a local Lagrangian manifold $\mathcal{L}_0 \subset \Sigma$ through λ_0 with $T_{\lambda_0} \mathcal{L}_0 = L_0$; locally, $\mathcal{L}_0$ is the graph of da for a smooth function a defined near q_0 . Condition (5.4.8) gives condition (4) of Theorem 5.3.1, while the construction of H_S in Section 5.3 gives the remaining hypotheses. Therefore Theorem 5.3.1 applies and yields the strict local strong optimality of the reference trajectory. $\square$

Notice that, in the proof of Lemma 5.4.8, the subspaces H_i are necessary because in general there can be a non-trivial intersection $B_t \bar{L}_0 \cap H_0$. But we want to avoid it since we want to use the local coordinates in $H^\natural$, for some Lagrangian subspace H , and this is always possible by means of a small perturbation of the H_i .

5.5 Examples

5.5.1 An example: the Heisenberg group with a drift

We are going now to show an example of Problem 5.1 modelled on the Heisenberg group. We will show that there are some nontrivial singular extremals, but, thanks to the necessary condition of Theorem 5.1.1, one can easily rule out their optimality.

Consider $\mathbb{R}^3$ with coordinates (x, y, z) . Let X, Y, Z be the generators of the Heisenberg algebra, which can be written as

$$\begin{aligned} X &= \partial_x - \frac{y}{2} \partial_z, \\ Y &= \partial_y + \frac{x}{2} \partial_z, \\ Z &= \partial_z. \end{aligned}$$

The three vector fields satisfy

$$[X, Y] = Z, \quad [X, Z] = 0, \quad [Y, Z] = 0.$$

We consider the following control system

$$\begin{aligned} \dot{q} &= \alpha X(q) + \beta Y(q) + \gamma Z(q) + u_1(t)X(q) + u_2(t)Y(q), \\ q(0) &= q_0, \end{aligned}$$

where $\alpha, \beta, \gamma \in \mathbb{R}$, $u_1, u_2 : [0, T] \rightarrow \overline{B_1(0)} \subset \mathbb{R}^2$. As before, we consider the cost functional $J(u) = \|u\|_{L^1}$. The Hamiltonian function for this problem is

$$H(\lambda, u) = h_0(\lambda) + \langle u, h_I(\lambda) \rangle - |u|$$

where

$$\begin{aligned} h_0(\lambda) &= \langle \lambda, \alpha X(q) + \beta Y(q) + \gamma Z(q) \rangle = \alpha h_X(\lambda) + \beta h_Y(\lambda) + \gamma h_Z(\lambda), \\ h_I(\lambda) &= (h_X(\lambda), h_Y(\lambda)), \end{aligned}$$

and $h_X(\lambda) = \langle \lambda, X(q) \rangle$, $h_Y(\lambda) = \langle \lambda, Y(q) \rangle$, $h_Z(\lambda) = \langle \lambda, Z(q) \rangle$.

The condition $u = r \frac{h_I}{|h_I|}$ reads

$$(u_1, u_2) = r(t) \frac{(h_X, h_Y)}{\sqrt{h_X^2 + h_Y^2}},$$

Indeed, the condition $|h_I(t)| = 1$ implies $\frac{1}{2} = h_c = \frac{1}{2}(h_X^2 + h_Y^2)$ along the whole trajectory. If we differentiate in time this equation and simplify, we obtain

$$\begin{aligned} \frac{d}{dt}[h_c(\lambda_t)] &= h_{0c}(\lambda_t) = (\alpha h_Y(\lambda_t) - \beta h_X(\lambda_t))h_Z(\lambda_t) = 0, \\ \frac{d^2}{dt^2}[h_c(\lambda_t)] &= \{h_0 + r h_c, h_{0c}\} = (\alpha^2 + \beta^2 + r(\alpha h_X + \beta h_Y))h_Z^2 = 0. \end{aligned}$$

If $h_Z = 0$ identically, then the two previous equations are trivial. In particular, we have that $h_{c0c}(t) \equiv 0$, so our theorems do not provide any information. We do not analyse in detail this case here. However, a direct computation shows that $h_X(t)$ and $h_Y(t)$ are both constant. This forces $\lambda(t)$ to be constant as well and any control r yields the same singular extremal trajectory.

If instead $h_Z \neq 0$, we obtain that

$$(h_X, h_Y) = -\frac{(\alpha, \beta)}{\sqrt{\alpha^2 + \beta^2}}, \quad r = \sqrt{\alpha^2 + \beta^2},$$

where we have chosen the negative sign for h_X and h_Y in order to have $r \geq 0$. The formula for the singular control u reads

$$u(t) = (-\alpha, -\beta).$$

Thus, the system of equations 5.5.1 for this control is simply

$$\dot{q} = \gamma Z(q).$$

Notice that $h_{c0c}(\lambda_t) = -\sqrt{\alpha^2 + \beta^2} < 0$ if $(\alpha, \beta) \neq (0, 0)$. So, by Theorem 5.1.1, any piece of this extremal is not optimal between its endpoints. On the other hand, if $(\alpha, \beta) = (0, 0)$, then the singular control is $u = 0$, which is also a regular control.

5.5.2 A Martinet-type example

We are now going to show an example, modelled on the flat-Martinet sub-Riemannian structure, for which the singular extremals satisfy the strong generalized Legendre condition. Hence, Theorems 5.1.2 and 5.1.4 apply, meaning that these extremal trajectories are strongly locally optimal.

Let $M = \mathbb{R}^3$, $f_0 = \alpha\partial_x + \beta\partial_y + \gamma\partial_z$, $f_1 = \partial_x$, $f_2 = \partial_y + \frac{x^2}{2}\partial_z$. We consider the control system

$$\dot{q} = f_0(q) + u_1(t)f_1(q) + u_2(t)f_2(q). \quad (5.5.1)$$

The brackets of these vector fields are (we define $f_{ij} = [f_i, f_j]$ and $f_{ijk} = [f_i, [f_j, f_k]]$)

$$\begin{aligned} f_{01} &= 0, & f_{02} &= \alpha x \partial_z, & f_{12} &= x \partial_z, \\ f_{001} &= 0, & f_{002} &= \alpha^2 \partial_z, & f_{101} &= 0, \\ f_{102} &= \alpha \partial_z, & f_{201} &= 0, & f_{202} &= 0. \end{aligned}$$

Let us also denote

$$p_x(\lambda) = \langle \lambda, \partial_x \rangle, \quad p_y(\lambda) = \langle \lambda, \partial_y \rangle, \quad p_z(\lambda) = \langle \lambda, \partial_z \rangle.$$

So, following the notation introduced above, we have:

$$h_{0c} = \frac{1}{2} \{h_0, h_1^2 + h_2^2\} = h_1 h_{01} + h_2 h_{02} = h_2 h_{02},$$

$$h_{00c} = \{h_0, h_2 h_{02}\} = h_{02}^2 + h_2 h_{002},$$

$$h_{c0c} = \frac{1}{2} \{h_1^2 + h_2^2, h_2 h_{02}\} = h_1 h_{12} h_{02} + h_1 h_2 h_{102}.$$

Since for singular extremals we must have $h_{0c} = 0$, we have that a singular extremal must satisfy either 1) $h_2 = 0$ or 2) $h_{02} = 0$ along the trajectory. In case 1), it follows that

$$h_{c0c} = h_1 h_{12} h_{02} = \alpha p_x p_z^2,$$

$$r = -\frac{h_{00c}}{h_{c0c}} = -\frac{h_{02}}{h_1 h_{12}} = -\frac{\alpha}{p_x} = -\alpha p_x \frac{1}{p_x^2}$$

As a result,

$$h_{c0c} > 0 \iff r < 0,$$

so the generalized Legendre condition does not hold in this situation.

In case 2), since $h_{02} = 0$ implies $x = 0$, we have

$$r = -\frac{h_{00c}}{h_{c0c}} = -\frac{h_2 h_{002}}{h_1 h_2 h_{102}} = -\frac{h_{002}}{h_1 h_{102}} = -\frac{\alpha^2 p_z}{p_x(\alpha p_z)} = -\frac{\alpha}{p_x}.$$

Since $h_1 = p_x$ and $h_2 = p_y$ (when restricted to $x = 0$), we obtain

$$0 < h_{c0c} = h_1 h_2 h_{102} = \alpha p_x p_y p_z.$$

and the singular control is

$$u(t) = r(h_1, h_2) = -\frac{\alpha}{p_x}(p_x, p_y) = \left(-\alpha, -\alpha \frac{p_y}{p_x}\right).$$

The Pontryagin Hamiltonian, written in coordinates (x, y, z, p_x, p_y, p_z) for this problem is

$$H(p, q, u) = \alpha p_x + \beta p_y + \gamma p_z + u_1 p_x + u_2 \left(p_y + \frac{x^2}{2} p_z\right) - |u|.$$

Hence, the corresponding Hamiltonian system is

$$\begin{aligned} \dot{x} &= u_1 + \alpha, & \dot{p}_x &= -u_2 x p_z, \\ \dot{y} &= u_2 + \beta, & \dot{p}_y &= 0, \\ \dot{z} &= u_2 \frac{x^2}{2} + \gamma, & \dot{p}_z &= 0. \end{aligned}$$

Therefore, if $x = 0$, the only possibility for p is to be constant.

The trajectory of (5.5.1) corresponding to the control u starting from $(0, y_0, z_0)$ is

$$\begin{aligned} x(t) &= 0, \\ y(t) &= \left(\beta - \frac{\alpha p_y}{p_x}\right)t + y_0, \\ z(t) &= \gamma t + z_0. \end{aligned}$$

We can choose the following normalization for p : assume that $\alpha > 0$ (the case $\alpha < 0$ is similar). Since $(h_1, h_2) = (p_x, p_y)$, we have $p_x^2 + p_y^2 = 1$, so that we can parametrize the singular extremals with $(p_x, p_y) = (\cos \theta, \sin \theta)$. To satisfy the condition $0 < r < 1$, we have to choose $\cos \theta < -\alpha$. Finally, for any such θ , choosing $p_z = \text{sign}(\sin(2\theta))$, we have $h_{c0c} = \frac{\alpha}{2} |\sin(2\theta)| > 0$.

To summarize, for any triple $(\alpha, \beta, \gamma) \in [-1, 1] \times \mathbb{R}^2$, $\alpha \neq 0$, the control system (5.5.1) has an infinite family of singular extremals satisfying the strong generalized Legendre condition. Their corresponding controls are constant and they can be parametrized as

$$u = (-\alpha, -\alpha \tan \theta), \quad \begin{cases} \text{for } \pi - \arccos \alpha < \theta < \pi + \arccos \alpha & \text{if } \alpha > 0, \\ \text{for } -\arccos |\alpha| < \theta < \arccos |\alpha| & \text{if } \alpha < 0. \end{cases}$$

5.5.3 Left-invariant problem in $SU(2)$

In this section we are going to illustrate an example where we use all the optimality conditions developed in this chapter. In particular, we will write explicitly the Jacobi equation and will find all the non-trivial solutions of this system.

Let us consider the following system:

$$\dot{U} = U(\alpha A + \beta B + \gamma C + u_1(t)A + u_2(t)B), \quad U(0) = \text{Id},$$

where $U \in SU(2)$, $\alpha, \beta, \gamma \in \mathbb{R}$, $(u_1, u_2) : [0, T] \rightarrow \overline{B_1(0)} \subset \mathbb{R}^2$ is measurable and

$$A = -\frac{i}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = -\frac{i}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad C = -\frac{i}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

From direct computations, we deduce the following relations:

$$[A, B] = C, \quad [B, C] = A, \quad [C, A] = B. \quad (5.5.2)$$

Therefore our system is strong Lie bracket generating. We want to minimize

$$J(u) = \int_0^T |u(t)| dt.$$

Let $F_0 = \alpha A + \beta B + \gamma C$, $F_I = (A, B)$ and define the left-invariant Hamiltonians as functions living on the dual of the Lie algebra $\mathfrak{su}(2)$

$$\begin{aligned} h_A(\lambda) &= \langle \lambda, A \rangle, \quad h_B(\lambda) = \langle \lambda, B \rangle, \quad h_C(\lambda) = \langle \lambda, C \rangle, \\ h_0(\lambda) &= \langle \lambda, F_0 \rangle, \quad h_I(\lambda) = (h_A(\lambda), h_B(\lambda)), \quad \lambda \in (\mathfrak{su}(2))^*. \end{aligned}$$

The Pontryagin (left-invariant) Hamiltonian function for our optimal control problem is

$$\begin{aligned} H(\lambda, u) &= h_0(\lambda) + \langle u, h_I(\lambda) \rangle - |u| \\ &= \alpha h_A(\lambda) + \beta h_B(\lambda) + \gamma h_C(\lambda) + r \left(\left\langle v, (h_A(\lambda), h_B(\lambda)) \right\rangle - 1 \right). \end{aligned}$$

From the commutation rules (5.5.2) it follows that

$$\{h_A, h_B\} = h_C, \quad \{h_C, h_A\} = h_B, \quad \{h_B, h_C\} = h_A.$$

Singular extremals correspond to

$$|h_I| = 1 \iff 2h_c = h_A^2 + h_B^2 = 1.$$

The first derivative of h_c is

$$\begin{aligned}\dot{h}_c = \{h_0, h_c\} &= \left\{ \alpha h_A + \beta h_B + \gamma h_C, \frac{1}{2}(h_A^2 + h_B^2) \right\} = \\ &= (\alpha h_B - \beta h_A) \{h_A, h_B\} + \gamma (h_A \{h_C, h_A\} + h_B \{h_C, h_B\}) = \\ &= (\alpha h_B - \beta h_A) h_C.\end{aligned}$$

We can further compute

$$\begin{aligned}\dot{h}_A &= \{h_0 + r h_c, h_A\} = -\beta h_C + \gamma h_B - r(t) h_B h_C, \\ \dot{h}_B &= \{h_0 + r h_c, h_B\} = \alpha h_C - \gamma h_A + r(t) h_A h_C, \\ \dot{h}_C &= \{h_0 + r h_c, h_C\} = -\alpha h_B + \beta h_A - r(t) h_A h_B + r(t) h_B h_A = \beta h_A - \alpha h_B.\end{aligned}$$

And the second derivative of h_c is

$$\begin{aligned}\ddot{h}_c &= (\alpha \dot{h}_B - \beta \dot{h}_A) h_C + (\alpha h_B - \beta h_A) \dot{h}_C \\ &= \begin{cases} ((\alpha^2 + \beta^2) h_C + (r h_C - \gamma)(\alpha h_A + \beta h_B)) h_C, & \text{if } \alpha h_B \equiv \beta h_A \\ 0 & \text{if } h_C \equiv 0. \end{cases}\end{aligned}$$

If $h_C \equiv 0$ along the singular extremal, then $h_{c0c}(t) \equiv 0$ along this extremal. Moreover, from the equation $\dot{h}_C = \beta h_A - \alpha h_B$ it follows that either $h_A \equiv h_B \equiv 0$, which forces $\lambda(t) \equiv 0$, or $(\alpha, \beta) = (0, 0)$. In the following, we focus only on the case $h_C \neq 0$. So, from $\ddot{h}_c = 0$, we must have

$$r(t) = \left(\gamma - \frac{(\alpha^2 + \beta^2) h_C}{\alpha h_A + \beta h_B} \right) \frac{1}{h_C}.$$

In particular, if $h_A^2 + h_B^2 = 1$ and $\alpha h_B = \beta h_A$, then

$$(h_A, h_B) = \pm \frac{(\alpha, \beta)}{\sqrt{\alpha^2 + \beta^2}}.$$

So, h_A, h_B are constant. Let us assume $\alpha^2 + \beta^2 = 1$, so that

$$r = \frac{\gamma}{h_C} - (\pm 1)$$

Notice that if $\alpha h_B - \beta h_A = 0$, then $\dot{h}_C(t) = 0$. So, r is constant.

In the end, the value of the singular control must be

$$u_* = r(h_A, h_B) = \left(\frac{\gamma}{h_C} - (\pm 1) \right) (\alpha, \beta). \quad (5.5.3)$$

Let us check when this control satisfies the generalized Legendre condition. A straightforward computation shows that

$$h_{c0c} = h_C^2(\alpha h_A + \beta h_B) = \pm h_C^2 \sqrt{\alpha^2 + \beta^2}.$$

Therefore, in order to have $h_{c0c} > 0$ we need to choose the plus sign in (5.5.3).

Denote $u_* = (u_*^{(1)}, u_*^{(2)})$ and

$$X = \alpha A + \beta B + \gamma C + u_*^{(1)} A + u_*^{(2)} B = \frac{\gamma}{h_C}(\alpha A + \beta B + h_C C).$$

So, to summarize, we have found that the singular control is constant and the solution to the Hamiltonian system is

$$U(t) = e^{tX}, p = \text{const.}$$

Now, we want to find the solutions to the Jacobi equation of this problem.

Let us define $h_I^\perp = (-\beta, \alpha)$. Notice that $P_1 = \alpha A + \beta B$, $P_2 = -\beta A + \alpha B$, and C again form a basis for $\text{SU}(2)$. Since we have assumed $\alpha^2 + \beta^2 = 1$,

$$[P_1, P_2] = C, \quad [P_2, C] = P_1, \quad [C, P_1] = P_2.$$

So, since the commutator relations are the same as (5.5.2), we can assume $\alpha = 1, \beta = 0$. Thus, we have

$$\begin{aligned} h_I &= (1, 0), \quad h_I^\perp = (0, 1), \\ \langle h_I, \vec{h}_I \rangle &= \vec{h}_A, \quad \langle h_I^\perp, \vec{h}_I \rangle = \vec{h}_B. \end{aligned}$$

Given an admissible variation $v \in \ker D_u E_T$, we can represent it as follows:

$$v(t) = \rho(t)h_I + w(t) = \rho(t)(1, 0) + \omega(t)(0, 1),$$

where $\rho(t) := \langle v(t), h_I \rangle$ and $w(t) := \langle v(t), h_I^\perp \rangle$.

In order to write down the Jacobi equation, we have to compute

$$Z_t v = (\Phi_t)_*^{-1} \langle v, \vec{h}_I \rangle.$$

First, we have

$$(\Phi_t)_*^{-1} = e^{t \text{ad} \vec{H}_{u_*}}.$$

Let us compute this exponential in a suitable basis.

As discussed previously, h_C is constant along the singular trajectory. In order to avoid confusion in the following formulas, we introduce a parameter

$$\kappa := h_C(\lambda_0).$$

On $T(T^*M)$, we can take the basis

$$\vec{H}_1 = \vec{h}_B, \quad \vec{H}_2 = \frac{1}{\sqrt{1+\kappa^2}}(-\kappa\vec{h}_A + \vec{h}_C), \quad \vec{H}_3 = \vec{H}_{u_*} = \frac{\gamma}{\kappa}(\vec{h}_A + \kappa\vec{h}_C), \quad \partial_{H_1}, \quad \partial_{H_2}, \quad \partial_{H_3}.$$

Note that $\partial_{H_1}, \partial_{H_2}, \partial_{H_3}$ are well defined, since all of those Hamiltonians are linear functions on the fibres. In this basis, the operator $\text{ad}\vec{H}_3 = \frac{\gamma}{\kappa}\text{ad}(\vec{h}_A + \kappa\vec{h}_C)$ acts as follows:

$$\begin{aligned} [\vec{H}_3, \vec{H}_1] &= k\vec{H}_2, & [\vec{H}_3, \partial_{H_1}] &= 0, \\ [\vec{H}_3, \vec{H}_2] &= -k\vec{H}_1, & [\vec{H}_3, \partial_{H_2}] &= 0, \\ [\vec{H}_3, \vec{H}_3] &= 0, & [\vec{H}_3, \partial_{H_3}] &= 0, \end{aligned}$$

where $k = \frac{\gamma\sqrt{1+\kappa^2}}{\kappa}$. Also

$$[\vec{H}_1, \vec{H}_2] = \frac{1}{k}\vec{H}_3.$$

So, the flow $e^{t\text{ad}\vec{H}_3}$ fixes the vertical directions $\partial_{H_1}, \partial_{H_2}, \partial_{H_3}$ and on the other directions it acts as a rotation generated by the matrix

$$\begin{pmatrix} 0 & -k & 0 \\ k & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

The exponential of this matrix is then

$$e^{t\text{ad}\vec{H}_3} = \begin{pmatrix} \cos(kt) & -\sin(kt) & 0 & 0 & 0 & 0 \\ \sin(kt) & \cos(kt) & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Now, since $Z_t h_I = e^{t\text{ad}\vec{H}_3}(\vec{h}_A)$ and $[\vec{H}_{u_*}, \vec{h}_A] = \gamma\vec{h}_B$, we have

$$\frac{d}{dt}(Z_t h_I) = e^{t\text{ad}\vec{H}_3}(\text{ad}\vec{H}_3(\vec{h}_A)) = \gamma e^{t\text{ad}\vec{H}_3}\vec{h}_B.$$

So that

$$\begin{aligned}
\mathcal{Z}_t v &= -\frac{d}{dt}(Z_t h_I)\phi(t) + (Z_t h_I^\perp)\omega(t) = \\
&= (\omega(t) - \gamma\phi(t))e^{\text{tad}\vec{H}_3}\vec{h}_B = \\
&= (\omega(t) - \gamma\phi(t))e^{\text{tad}\vec{H}_3}\vec{H}_1 = \\
&= (\omega(t) - \gamma\phi(t))(\cos(kt)\vec{H}_1 + \sin(kt)\vec{H}_2).
\end{aligned}$$

Now, to compute the r.h.s. of the Jacobi equation, we need to evaluate $\sigma(\mathcal{Z}_t \cdot, \eta)$, for $\eta \in T_{\lambda_0}(T^*M)$. We can write

$$\eta = \eta_1 \vec{H}_1 + \eta_2 \vec{H}_2 + \eta_3 \vec{H}_3 + \eta_1^* \partial_{H_1} + \eta_2^* \partial_{H_2} + \eta_3^* \partial_{H_3}. \quad (5.5.4)$$

The symplectic form in this basis is

$$\begin{aligned}
\sigma(\partial_{H_\alpha}, \vec{H}_\beta) &= \langle dH_\beta, \partial_{H_\alpha} \rangle = \delta_{\alpha\beta}, \\
\sigma(\partial_{H_\alpha}, \partial_{H_\beta}) &= 0, \\
\sigma(\vec{H}_\alpha, \vec{H}_\beta) &= H_{\alpha\beta}.
\end{aligned}$$

We can calculate the Poisson brackets in a straightforward manner

$$\begin{aligned}
\{H_1, H_2\}(\lambda_0) &= \left\{ h_B, \frac{1}{\sqrt{1+\kappa^2}}(-\kappa h_A + h_C) \right\}(\lambda_0) = \frac{1}{\sqrt{1+\kappa^2}}(\kappa h_C(\lambda_0) + h_A(\lambda_0)) = \sqrt{1+\kappa^2}, \\
\{H_1, H_3\}(\lambda_0) &= \left\{ h_B, \frac{\gamma}{\kappa}(h_A + \kappa h_C) \right\}(\lambda_0) = \gamma(-h_C(\lambda_0)/\kappa + h_A(\lambda_0)) = 0, \\
\{H_2, H_3\}(\lambda_0) &= \left\{ \frac{1}{\sqrt{1+\kappa^2}}(-\kappa h_A + h_C), \frac{\gamma}{\kappa}(h_A + \kappa h_C) \right\}(\lambda_0) = \frac{\gamma(1+\kappa^2)}{\kappa\sqrt{1+\kappa^2}}h_B(\lambda_0) = 0.
\end{aligned}$$

Hence, we have

$$\begin{aligned}
\sigma(\cos(kt)\vec{H}_1 + \sin(kt)\vec{H}_2, \eta) &= \\
&= \sigma(\cos(kt)\vec{H}_1 + \sin(kt)\vec{H}_2, \eta_1 \vec{H}_1 + \eta_2 \vec{H}_2 + \eta_3 \vec{H}_3 + \eta_1^* \partial_{H_1} + \eta_2^* \partial_{H_2} + \eta_3^* \partial_{H_3}) \\
&= \cos(kt)\eta_2 H_{12} - \cos(kt)\eta_1^* - \sin(kt)\eta_1 H_{12} - \sin(kt)\eta_2^* \\
&= \sqrt{1+\kappa^2}(\cos(kt)\eta_2 - \sin(kt)\eta_1) - (\cos(kt)\eta_1^* + \sin(kt)\eta_2^*).
\end{aligned}$$

Let us define $\psi(t, \eta) := \sqrt{1+\kappa^2}(\cos(kt)\eta_2 - \sin(kt)\eta_1) - (\cos(kt)\eta_1^* + \sin(kt)\eta_2^*)$.

The inverse of the Legendre form is

$$l_t = \begin{pmatrix} \frac{1}{r} & -\kappa \\ -\kappa & \gamma\kappa \end{pmatrix} \implies l_t^{-1} = \left(\frac{\gamma\kappa}{r} - \kappa^2 \right)^{-1} \begin{pmatrix} \gamma\kappa & \kappa \\ \kappa & \frac{1}{r} \end{pmatrix},$$

so that

$$l_t^{-1}\sigma(\mathcal{Z}_t, \eta) = \left(\frac{\gamma\kappa}{r} - \kappa^2\right)^{-1} \begin{pmatrix} \gamma\kappa & \kappa \\ \kappa & \frac{1}{r} \end{pmatrix} \begin{pmatrix} 1 \\ -\gamma \end{pmatrix} \psi(t, \eta) = -\frac{\psi(t, \eta)}{\kappa} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Hence

$$\begin{aligned} \dot{\eta} &= -\mathcal{Z}_t l_t^{-1} \sigma(\mathcal{Z}_t, \eta) = \\ &= -\frac{\gamma}{\kappa} \left(\sqrt{1 + \kappa^2} (\eta_2 \cos(kt) - \eta_1 \sin(kt)) - \eta_1^* \cos(kt) - \eta_2^* \sin(kt) \right) (\cos(kt) \vec{H}_1 + \sin(kt) \vec{H}_2). \end{aligned} \quad (5.5.5)$$

First, we are going to find six linearly independent solutions to equation (5.5.5) and then we are going to determine solutions corresponding to conjugate times. If we compare equation (5.5.5) with (5.5.4), we immediately see that $\dot{\eta}_1^* = \dot{\eta}_2^* = \dot{\eta}_3^* = \dot{\eta}_3 = 0$. To determine the solutions of equation (5.5.5), it is sufficient to study the variables η_1 and η_2 keeping the other fixed. Let us denote by $\eta = (\eta_1, \eta_2, \eta_3, \eta_1^*, \eta_2^*, \eta_3^*)$. Elements of the subspace

$$\text{span}\{(0, 0, 1, 0, 0, 0), (0, 0, 0, 0, 0, 1), (0, 1, 0, \sqrt{1 + \kappa^2}, 0, 0), (1, 0, 0, 0, -\sqrt{1 + \kappa^2}, 0)\}$$

are zeros of the r.h.s. of (5.5.5), hence they form the space of constant solutions. So, by linearity of the system, we can reduce to the study of a two-variable ODE, with $(\eta_1, \eta_2) \neq 0$, $\eta^* \equiv 0, \eta_3 \equiv 0$. Let $a = \cos(kt)$, $b = \sin(kt)$ and recall that $k = \gamma(\sqrt{1 + \kappa^2})/\kappa$. The previous system can be written in matrix form:

$$\begin{pmatrix} \dot{\eta}_1 \\ \dot{\eta}_2 \end{pmatrix} = k \begin{pmatrix} ab & -a^2 \\ b^2 & -ab \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}. \quad (5.5.6)$$

Notice that the matrix defining the system of equations is nilpotent:

$$\begin{pmatrix} ab & -a^2 \\ b^2 & -ab \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

So, we can make a (time-dependent) change of variables

$$\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}, \quad (5.5.7)$$

and rewrite system (5.5.6) in these new coordinates:

$$\begin{pmatrix} \dot{\xi}_1 \\ \dot{\xi}_2 \end{pmatrix} = k \begin{pmatrix} -b & a \\ -a & -b \end{pmatrix} \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + k \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}.$$

Simplifying, one obtains

$$\begin{cases} \dot{\xi}_1 = 0, \\ \dot{\xi}_2 = -k\xi_1. \end{cases}$$

The solutions are

$$\xi(t) = \begin{pmatrix} 1 & 0 \\ -kt & 1 \end{pmatrix} \xi(0) = \begin{pmatrix} \xi_1(0) \\ -kt\xi_1(0) + \xi_2(0) \end{pmatrix}.$$

Moreover, since the change of basis in (5.5.7) is the identity for $t = 0$, we have $\xi(0) = (\eta_1(0), \eta_2(0))$ and

$$\begin{pmatrix} \eta_1(t) \\ \eta_2(t) \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} \eta_1(0) \\ -kt\eta_1(0) + \eta_2(0) \end{pmatrix} = \begin{pmatrix} (\cos(kt) + kt \sin(kt))\eta_1(0) - \sin(kt)\eta_2(0) \\ (\sin(kt) - kt \cos(kt))\eta_1(0) + \cos(kt)\eta_2(0) \end{pmatrix}.$$

So, we have obtained a family of six linearly independent solutions for the equation (5.5.5). We still have to determine the solutions with the appropriate boundary conditions, which are

$$\begin{aligned} \eta(0) &\in (T_{\lambda_0}(T_{q_0}^* M) + \mathbb{R}Z_I(0)) \cap \mathbb{R}Z_I(0)^\perp, \\ \eta(T) &\in T_{\lambda_0}(T_{q_0} M). \end{aligned} \tag{5.5.8}$$

Recall that $Z_I(0) = \vec{h}_A = -\frac{\kappa}{\sqrt{1+\kappa^2}}\vec{H}_2 + \frac{\kappa}{\gamma(1+\kappa^2)}\vec{H}_3 =: c_2\vec{H}_2 - c_3\vec{H}_3$. A direct verification shows that

$$\begin{aligned} (T_{\lambda_0}(T_{q_0}^* M) + \mathbb{R}Z_I(0)) &= \text{span}\{\partial_{H_1}, \partial_{H_2}, \partial_{H_3}, \vec{h}_A\}, \\ \mathbb{R}Z_I(0)^\perp &= \text{span}\{\vec{h}_A, \vec{H}_3, \partial_{H_1}, c_3\partial_{H_2} + c_2\partial_{H_3}, \vec{H}_1 - \sqrt{1+\kappa^2}\partial_{H_2}\}. \end{aligned}$$

Therefore,

$$(T_{\lambda_0}(T_{q_0}^* M) + \mathbb{R}Z_I(0)) \cap \mathbb{R}Z_I(0)^\perp = \text{span}\{\vec{h}_A, \partial_{H_1}, c_3\partial_{H_2} + c_2\partial_{H_3}\}.$$

and $T_{\lambda_0}(T_{q_0}^* M) = \text{span}\{\partial_{H_1}, \partial_{H_2}, \partial_{H_3}\}$. Here all the vector fields are intended to be evaluated at λ_0 .

We have to find the solutions for $\eta(0) \in \text{span}\{\vec{h}_A, \partial_{H_1}, c_3\partial_{H_2} + c_2\partial_{H_3}\}$ and see for which $T > 0$ we realize (5.5.8). In order to determine these solutions, for each of the three vectors $\vec{h}_A, \partial_{H_1}, c_3\partial_{H_2} + c_2\partial_{H_3}$, we write it as the sum of a constant solution of (5.5.5) and a time-varying solution.

- If $\eta(0) = \vec{h}_A = c_2\vec{H}_2 - c_3\vec{H}_3$, then

$$\eta(t) = -\sin(kt)c_2\vec{H}_1 + \cos(kt)c_2\vec{H}_2 - c_3\vec{H}_3.$$

- If $\eta(0) = \sqrt{1 + \kappa^2} \partial_{H_1} = (\sqrt{1 + \kappa^2} \partial_{H_1} + \vec{H}_2) - \vec{H}_2$, then

$$\begin{aligned} \eta(t) &= (\sqrt{1 + \kappa^2} \partial_{H_1} + \vec{H}_2) + \sin(kt) \vec{H}_1 - \cos(kt) \vec{H}_2 = \\ &= \sqrt{1 + \kappa^2} \partial_{H_1} + \sin(kt) \vec{H}_1 + (1 - \cos(kt)) \vec{H}_2. \end{aligned}$$

- If $\eta(0) = -\sqrt{1 + \kappa^2} (c_3 \partial_{H_2} + c_2 \partial_{H_3}) = (-\sqrt{1 + \kappa^2} (c_3 \partial_{H_2} + c_2 \partial_{H_3}) + c_3 \vec{H}_1) - c_3 \vec{H}_1$, then

$$\begin{aligned} \eta(t) &= -\sqrt{1 + \kappa^2} c_2 \partial_{H_3} + c_3 (\vec{H}_1 - \sqrt{1 + \kappa^2} \partial_{H_2}) - c_3 ((\cos(kt) + kt \sin(kt)) \vec{H}_1 + (\sin(kt) - kt \cos(kt)) \vec{H}_2) = \\ &= -\sqrt{1 + \kappa^2} c_2 \partial_{H_3} - \sqrt{1 + \kappa^2} c_3 \partial_{H_2} + c_3 (1 - \cos(kt) - kt \sin(kt)) \vec{H}_1 - c_3 (\sin(kt) - kt \cos(kt)) \vec{H}_2. \end{aligned}$$

We see that any initial condition of the form

$$a_1 \vec{h}_A + a_2 \sqrt{1 + \kappa^2} \partial_{H_1} - a_3 \sqrt{1 + \kappa^2} (c_3 \partial_{H_2} + c_2 \partial_{H_3}), \quad a_i \in \mathbb{R},$$

results in a solution that has a nonzero component along $\vec{H}_3$, unless $a_1 = 0$. So, since we want to satisfy (5.5.8), we can restrict to this case. Hence, we have to find $a_2, a_3 \in \mathbb{R}$ such that

$$\begin{aligned} a_3 c_3 (1 - \cos(kt) - kt \sin(kt)) + a_2 \sin(kt) &= 0, \\ -a_3 c_3 (\sin(kt) - kt \cos(kt)) + a_2 (1 - \cos(kt)) &= 0. \end{aligned}$$

So, in order to have a non-zero solution, we must require that the determinant of the matrix defining the system of equations with respect to a_2, a_3 is zero:

$$(1 - \cos(kt) - kt \sin(kt))(1 - \cos(kt)) + (\sin(kt) - kt \cos(kt)) \sin(kt) = 0,$$

which simplifies to

$$2 - 2 \cos(kt) - kt \sin(kt) = 0,$$

and applying the half-angle formula, we obtain

$$2 \sin\left(\frac{kt}{2}\right)^2 - kt \sin\left(\frac{kt}{2}\right) \cos\left(\frac{kt}{2}\right) = 0.$$

Therefore, we have a conjugate point at $t = 2\pi/k$ and for $\bar{t}$ that satisfies

$$\tan\left(\frac{k\bar{t}}{2}\right) = \frac{k\bar{t}}{2}.$$

5.6 Second variation and the generalized Legendre condition

5.6.1 Definition of the endpoint map and the Hessian

In this section, we compute the Hessian of the extended endpoint map for the problem 5.1. We recall that the endpoint map is defined by

$$F : \mathcal{U} \rightarrow M, \quad F(u) = q(T; q_0, u),$$

while the extended endpoint-map is

$$E : \mathcal{U} \rightarrow M \times \mathbb{R}, \quad E(u) = (q_u(T; q_0), J(u)).$$

Recall that, from Theorem 3.1.7, that if $\text{ind}^-(\lambda, \nu) \text{Hess}_u E = +\infty$ for a critical point u , then we immediately know that E is locally open at u , hence u is not a weakly locally optimal. So, our first goal is to compute the Hessian of E and then to find a condition that ensures $\text{ind}^-(\lambda, \nu) \text{Hess}_u E < +\infty$. Moreover, in Subsection 5.6.4, we show that it is possible to compute $\text{ind}^-(\lambda, \nu) \text{Hess}_u E$ finding the solutions of the Jacobi equation, which together with Theorem 3.1.7 will prove Theorem 5.1.3.

From now on we fix a critical point $\tilde{u} \in \mathcal{U}$ of E and we assume that it is a singular control of the form (5.2.2). Moreover, we denote by $\tilde{\lambda}$ its corresponding singular extremal, which is a solution of Equation (5.2.3). For the sake of brevity, we will also use the notation

$$Q := -(\lambda, \nu) \text{Hess}_u E.$$

5.6.2 Explicit expression for the second variation form

The formula of the Hessian of the endpoint map E for a general control system is computed in [13], Section 20.3. It reads as:

$$Q(v) = \int_0^T \left(-\frac{d^2 H(\tilde{\lambda}_0, \tilde{u}(t))}{du^2} (v(t), v(t)) + \sigma_{\tilde{\lambda}_0} \left(Z_t v(t), \int_0^t Z_s v(s) ds \right) \right) dt,$$

for $v \in \ker D_u E$, where the first summand is the Hessian of the PMP Hamiltonian with respect to the control variables. The map $Z_t : T_{\tilde{u}(t)} U \rightarrow T_{\tilde{\lambda}_0}^*(T^*M)$ was defined in (5.4.1), but we recall the definition here:

$$Z_t v(t) := \left\langle v(t), [(\Phi_t^{-1})_* \vec{h}_I](\tilde{\lambda}_0) \right\rangle,$$

where Φ_t is the flow on T^*M of the Hamiltonian system (5.2.3) defined by the PMP Hamiltonian. We also recall that this quadratic form must be evaluated on the set of admissible variations given by the kernel of the first differential. By definition, a variation

$v \in T_u\mathcal{M} = L^\infty([0, T], \mathbb{R}^m)$ is admissible if $v \in \ker D_u E_T$. In symplectic language, this can be equivalently restated by the following equation (see [13], Equation (21.4)):

$$\int_0^T Z_t v(t) dt \in \Pi_0, \quad (5.6.1)$$

where $\Pi_0 = T_{\tilde{\lambda}_0}^*(T_{q_0}^* M)$ as before is the vertical fibre.

It is important to note that both the form Q and the first differential in (5.6.1) are continuous in a weaker norm of $L^2([0, T], \mathbb{R}^m)$. As a first step we extend them to this weaker topology.

As before we separate the component of v in the direction h_I , which is the singular control, from all the others. Namely, we write $v = \rho h_I + w$, where $v \in \ker D_u E_T$ and $\langle h_I, w \rangle = 0$.

We compute explicitly the term $\frac{d^2 H(\tilde{\lambda}_0, \tilde{u}(t))}{du^2}(v, v)$. Since the only non-linear part is the Euclidean norm of the control, we differentiate just this part and neglect the linear term. Let $g(u) = |u| = \left(\sum_{j=1}^m u_j^2\right)^{\frac{1}{2}}$. We have $\frac{\partial g}{\partial u_j} = u_j \left(\sum_{j=1}^m u_j^2\right)^{-\frac{1}{2}}$, and

$$\frac{\partial^2 g}{\partial u_i \partial u_j} = -u_i u_j \left(\sum_{j=1}^m u_j^2\right)^{-\frac{3}{2}}, \text{ if } i \neq j, \quad \frac{\partial^2 g}{\partial u_j^2} = -u_j^2 \left(\sum_{j=1}^m u_j^2\right)^{-\frac{3}{2}} + \left(\sum_{j=1}^m u_j^2\right)^{-\frac{1}{2}}.$$

We can represent the matrix associated with the Hessian of g as

$$\text{Hess}_u g = |u|^{-1} \text{Id} - |u|^{-3} u u^\top.$$

Let $v \in \ker D_u E_T$, $v = \rho h_I + w$, where $\langle h_I, w \rangle = 0$. Then, recalling the notation $u = r h_I$, we obtain

$$\text{Hess}_u g[v, v] = \frac{1}{r} |v|^2 - \frac{1}{r^3} (u^\top v)^2 = \frac{1}{r} (\rho^2 + |w|^2) - \frac{1}{r^3} (\rho r)^2 = \frac{|w|^2}{r},$$

and in particular

$$\frac{d^2 H(\tilde{\lambda}_0, \tilde{u}(t))}{du^2}(v, v) = -\frac{|w|^2}{r}.$$

Then we obtain

$$Q(v) = \int_0^T \left(\frac{|w(t)|^2}{r} + \sigma_{\tilde{\lambda}_0} \left(Z_t(\rho(t)h_I(t) + w(t)), \int_0^t Z_s(\rho(s)h_I(s) + w(s)) ds \right) \right) dt \quad (5.6.2)$$

$$= \int_0^T \frac{|w(t)|^2}{r} dt + \int_0^T \sigma_{\tilde{\lambda}_0} \left(\rho(t)Z_I(t), \int_0^t \rho(s)Z_I(s)ds \right) dt + \quad (5.6.3)$$

$$+ \int_0^T \sigma_{\tilde{\lambda}_0} \left(\rho(t)Z_I(t), \int_0^t Z_s w(s)ds \right) dt + \quad (5.6.4)$$

$$+ \int_0^T \sigma_{\tilde{\lambda}_0} \left(Z_t w(t), \int_0^t \rho(s)Z_I(s)ds \right) dt + \quad (5.6.5)$$

$$+ \int_0^T \sigma_{\tilde{\lambda}_0} \left(Z_t w(t), \int_0^t Z_s w(s)ds \right) dt.$$

The form Q is not coercive in the L^2 -topology. To see this, we can consider the family variations of the form $w \equiv 0$ and $\rho_\varepsilon(t) = \rho_1(t/\varepsilon)$, $0 < \varepsilon \leq 1$ and for any admissible ρ_1 . The same argument used in Proposition 20.1 in [13] shows that, since $v = \rho h_I$ is in the kernel of $\partial_u^2 H$, it holds that $Q = O(\varepsilon^2)$ for $\varepsilon \rightarrow 0$, while $\|\rho_\varepsilon\| = O(\varepsilon)$. To obtain coercivity we introduce as a new variable

$$\phi(t) = \int_0^t \rho(s) ds$$

and integrate by parts to replace the ρ variable with ϕ variable everywhere. For example, if we integrate by parts the first order condition (5.6.1), we get

$$\begin{aligned} \int_0^T Z_t v(t) dt &= \int_0^T Z_t \left(\rho(t)h_I(\tilde{\lambda}_t) + w(t) \right) dt = \\ &= \phi(T)Z_I(T) - \int_0^T \phi(t)\dot{Z}_I(t)dt + \int_0^T Z_t w(t) dt = \\ &= \phi(T)Z_I(T) + \int_0^T \mathcal{Z}_t v(t)dt \in \Pi_0, \end{aligned} \quad (5.6.6)$$

where we have shortened

$$\mathcal{Z}_t v(t) := -\phi(t)\dot{Z}_I(t) + Z_t w(t).$$

Similarly, integrating by parts in the formula for Q (see Appendix 5.7) we obtain an expression:

$$\begin{aligned} Q(v_1, v_2) &= \phi_1(T)\sigma \left(Z_I(T), \int_0^T \mathcal{Z}_t v_2(t)dt \right) + \int_0^T \langle v_1(t), l_t v_2(t) \rangle dt \\ &+ \int_0^T \sigma \left(\mathcal{Z}_t v_1(t), \int_0^t \mathcal{Z}_s v_2(s) ds \right) dt, \end{aligned} \quad (5.6.7)$$

where

$$l_t := \begin{pmatrix} \frac{\text{Id}}{r} & \sigma(Z_t \cdot, Z_I(t)) \\ \sigma(Z_t \cdot, Z_I(t))^T & \sigma(Z_I(t), \dot{Z}_I(t)) \end{pmatrix}.$$

and $l_t \in \mathbb{R}^{(m+1) \times (m+1)}$ with $\text{Id} \in \mathbb{R}^{m \times m}$.

Proposition 5.6.1. *The operator in (5.6.6) and the quadratic form (5.6.7) are continuous on the space of variations*

$$(\phi(T), \phi, w) \in \mathbb{R} \times L^2([0, T]) \times L^2([0, T], \mathbb{R}^{m-1}).$$

Here $\phi(T)$ and ϕ are considered as independent variations. It is not a problem, since the image of the map

$$\rho \mapsto \left(\int_0^T \rho(t) dt, \int_0^\cdot \rho(t) dt \right),$$

is dense inside $\mathbb{R} \times L^2([0, T], \mathbb{R})$. We will slightly abuse the notation from now on and write $v = (w, \phi)$ for the admissible variations.

Proof. Since the reference extremal is fixed, the maps $t \mapsto Z_t$, $t \mapsto \dot{Z}_I(t)$, and $t \mapsto l_t$ are continuous on the compact interval $[0, T]$; in particular they are uniformly bounded. Hence, for $v = (\phi, w) \in L^2([0, T]) \times L^2([0, T], \mathbb{R}^{m-1})$, the map

$$t \mapsto Z_t v(t) = -\phi(t) \dot{Z}_I(t) + Z_t w(t)$$

belongs to $L^2([0, T], T_{\lambda_0}^*(T^*M))$. Since $[0, T]$ has finite measure, it also belongs to L^1 , and therefore

$$v \mapsto \phi(T) Z_I(T) + \int_0^T Z_t v(t) dt$$

is a continuous linear map from $\mathbb{R} \times L^2([0, T]) \times L^2([0, T], \mathbb{R}^{m-1})$ to $T_{\lambda_0}^*(T^*M)$. This proves the continuity of the first-order operator in (5.6.6).

The same boundedness argument gives the continuity of the quadratic form in (5.6.7). Indeed, the term

$$\int_0^T \langle v_1(t), l_t v_2(t) \rangle dt$$

is continuous by the Cauchy-Schwarz inequality and the uniform boundedness of l_t . Moreover, the operator

$$v \mapsto \left(t \mapsto \int_0^t Z_s v(s) ds \right)$$

is continuous from L^2 to L^∞ , but $L^\infty[0, T]$ is continuously embedded into $L^2[0, T]$. Hence, the map above is a continuous map from L^2 to L^2 . $\square$

5.6.3 Proof of Theorem 5.1.1

To prove Theorem 5.1.1 we restrict the form Q in (5.6.7) to the subspace of finite codimension with $\phi(T) = 0$. Then, essentially we have (see Appendix 5.7 for a precise formula)

$$Q(v) = \int_0^T \langle v(t), l_t v(t) \rangle dt + \text{other terms.} \quad (5.6.8)$$

Following the approach in [13, Chapter 20], we have to check that the form Q has a finite negative index.

Proposition 5.6.2. *If the quadratic form Q has finite negative index, then the Legendre form l_t is non-negative for almost all $t \in [0, T]$. Moreover, if the Legendre form l_t is strictly positive definite for almost every $t \in [0, T]$, then for every $\tau \in [0, T)$ there is $\varepsilon > 0$ such that the form Q restricted to variations supported in $[\tau, \tau + \varepsilon]$ is strictly positive definite. In particular, Q has finite negative index.*

The proof is analogous to the proof of Proposition 20.1 of [13] and we do not repeat it, but only give the main idea. The proof is by contradiction. We assume, that there is a small open interval $I \subset [0, T]$ on which the Legendre form is not non-negative definite. Then choose $t \in I$ and consider variations concentrated on a small interval $[t, t + \varepsilon]$. This will give us a family of quadratic forms parametrized by the parameter $\varepsilon > 0$. A suitable rescaling can be used to transform it into a continuous family of forms on L^2 functions on the interval $[0, 1]$. By writing down the asymptotic expansion in ε , we find that the dominating term is given exactly by the integral of the Legendre form (it is already written as highest order term in (5.6.8)). By assumption, at every moment of time the Legendre form has a non-trivial subspace on which it is negative-definite. By taking any bounded variation v , such that for a.e. moment of time t it belongs to the negative subspace of the Legendre form, we, hence, obtain a variation that for $\varepsilon > 0$ small enough will be on the negative subspace of the Hessian. We can multiply this variation v by any smooth function and it will still belong to the negative subspace of the Hessian. Therefore, its Morse index is infinite and the extremal trajectory is not optimal.

Notice that technically we need to consider the form $\langle l_t v, v \rangle$ restricted to the subspace $\langle h_I(t), w \rangle = 0$. Keeping an extra variable simplifies the exposition, and we will show that this additional direction does not change the sign of the matrix l_t . Let us now see when $l_t \geq 0$. Since $r > 0$ by assumption, by Sylvester's criterion it suffices to check that the determinant of the matrix is positive. Notice that

$$\begin{pmatrix} \frac{\text{Id}}{r} & a \\ a^T & \alpha \end{pmatrix} = \begin{pmatrix} \text{Id} & 0 \\ ra^T & 1 \end{pmatrix} \begin{pmatrix} \frac{\text{Id}}{r} & 0 \\ 0 & \alpha - r|a|^2 \end{pmatrix} \begin{pmatrix} \text{Id} & ra \\ 0 & 1 \end{pmatrix}$$

where $a = \sigma(Z_t \cdot, Z_I(t))$ and $\alpha = \sigma(Z_I(t), \dot{Z}_I(t))$. Hence, the previous matrix is non-negative if and only if

$$\sigma(Z_I(t), \dot{Z}_I(t)) - r \|\sigma(Z_t \cdot, Z_I(t))\|^2 \geq 0, \quad (5.6.9)$$

We point out that the form l_t is non-negative definite for $(w, \phi) \in (T_{\tilde{u}(t)}U \cap h_I^\perp) \times \mathbb{R}$ if and only if it is non-negative definite for $(w, \phi) \in T_{\tilde{u}(t)}U \times \mathbb{R}$. One implication is obvious. Concerning the other one, if you take $(w, \phi) \in T_{\tilde{u}(t)}U \times \mathbb{R}$, then, writing $w = \tilde{w} + \rho h_I$, with $\tilde{w} \in h_I^\perp$, we have

$$(\tilde{w} + \rho h_I, 0)l_t(0, \phi)^T = \phi\sigma(Z_t(\tilde{w} + \rho h_I), Z_t h_I) = \phi\sigma(Z_t(\tilde{w}), Z_t h_I) = (\tilde{w}, 0)l_t(0, \phi)^T.$$

Hence, assuming $(\tilde{w}, \phi)l_t(\tilde{w}, \phi)^T \geq 0$, we obtain that also $(w, \phi)l_t(w, \phi)^T = (\tilde{w}, \phi)l_t(\tilde{w}, \phi)^T \geq 0$, as we claimed.

Let us rewrite condition (5.6.9) explicitly in terms of the Hamiltonians h_i . Recall, that in Proposition 5.4.5 we claimed that $l_t \geq 0$ is equivalent to $h_{c0c}(t) \geq 0$. So, to prove the Proposition, we only need to show that (5.6.9) is equivalent to $h_{c0c}(t) \geq 0$.

Proof of Proposition 5.4.5. From the definition of Z_t , we have

$$Z_I(t) = (\Phi_t^{-1})_*(\langle h_I(\tilde{\lambda}_t), \vec{h}_I \rangle).$$

If we differentiate in t , we obtain

$$\begin{aligned} \dot{Z}_I(t) &= (\Phi_t^{-1})_* \left\langle \frac{d}{dt}[h_I(\tilde{\lambda}_t)], \vec{h}_I \right\rangle + (\Phi_t^{-1})_* \left[\vec{h}_0 + \langle \tilde{u}(t), \vec{h}_I \rangle, \langle h_I(\tilde{\lambda}_t), \vec{h}_I \rangle \right] \\ &= (\Phi_t^{-1})_* \left\langle \{h_0 + \langle \tilde{u}(t), h_I \rangle, h_I\}(\tilde{\lambda}_t), \vec{h}_I \right\rangle + (\Phi_t^{-1})_* \left[\vec{h}_0, \langle h_I(\tilde{\lambda}_t), \vec{h}_I \rangle \right] \\ &= (\Phi_t^{-1})_* \left(\sum_{i=1}^m \left(h_{0i} + r(t) \sum_{j=1}^m h_j(\tilde{\lambda}_t) h_{ji} \right) \vec{h}_i + \sum_{i=1}^m h_i(\tilde{\lambda}_t) \vec{h}_{0i} \right) = \\ &= (\Phi_t^{-1})_* \left(\vec{h}_{0c} + r(t) \sum_{i,j=1}^m h_j(\tilde{\lambda}_t) h_{ji} \vec{h}_i \right). \end{aligned}$$

So, we can compute $\sigma_{\tilde{\lambda}_0}(Z_I(t), \dot{Z}_I(t))$:

$$\begin{aligned} \sigma_{\tilde{\lambda}_0}(Z_I(t), \dot{Z}_I(t)) &= \sigma_{\tilde{\lambda}(t)} \left(\sum_{i=1}^m h_i(\tilde{\lambda}_t) \vec{h}_i, \vec{h}_{0c} + r(t) \sum_{i,j=1}^m h_j(\tilde{\lambda}_t) h_{ji} \vec{h}_i \right) = \\ &= h_{c0c} + r \sum_{i,j,l=1}^m h_j h_l h_{ji} h_{li}. \end{aligned}$$

The second sum is (every function now is evaluated in $\tilde{\lambda}(t)$)

$$\begin{aligned} \sum_{i,l,j=1}^m h_i h_{jl} h_j h_{il} &= \sum_{i,l=1}^m h_i h_{il} \sum_j h_{jl} h_j = \sum_{l=1}^m \left[\left(\sum_j h_{jl} h_j \right) \sum_i h_i h_{il} \right] = \sum_{l=1}^m \left(\sum_j h_{jl} h_j \right)^2 \\ &= \sum_l \{h_l, h_c\}^2 = |\{h_c, h_I\}|^2 = |h_{cI}|^2, \end{aligned}$$

where h_{cI} is the vector whose i -th component is h_{ci} . Thus, putting everything together, we obtain

$$\sigma_{\tilde{\lambda}_0} \left(Z_I(t), \dot{Z}_I(t) \right) = h_{c0c}(\tilde{\lambda}_t) + r(\tilde{\lambda}_t) |h_{cI}(\tilde{\lambda}_t)|^2. \quad (5.6.10)$$

On the other hand, for any admissible variation v the second addendum in (5.6.9) is

$$\begin{aligned} \sigma_{\tilde{\lambda}_0} (Z_t v, Z_I(t)) &= \sigma_{\tilde{\lambda}_0} (\langle v(t), (\Phi_t^{-1})_* \vec{h}_I \rangle, \langle h_I(\tilde{\lambda}_t), (\Phi_t^{-1})_* \vec{h}_I \rangle) = \sum_{i,j} v_i(t) h_j(\tilde{\lambda}_t) \sigma_{\lambda(t)}(\vec{h}_i, \vec{h}_j) = \\ &= \sum_{i,j} v_i(t) h_j(\lambda_t) h_{ij}. \end{aligned}$$

Thus

$$|\sigma_{\tilde{\lambda}_0} (Z_t \cdot, Z_I(t))|^2 = \sum_i \left(\sum_j h_j h_{ij} \right)^2 = |h_{cI}|^2. \quad (5.6.11)$$

Finally, using (5.6.10) and (5.6.11), the left hand side of Equation (5.6.9) reads

$$\sigma_{\tilde{\lambda}_0} \left(Z_I(t), \dot{Z}_I(t) \right) - r \|\sigma_{\tilde{\lambda}_0} (Z_t \cdot, Z_I(t))\|^2 = h_{c0c} + r |h_{cI}|^2 - r |h_{cI}|^2 = h_{c0c},$$

which proves the result. $\square$

5.6.4 Derivation of the Jacobi equation and proof of Theorem 5.1.3

We consider a singular extremal for which the strong Legendre condition is satisfied. In this subsection we shorten $\sigma = \sigma_{\tilde{\lambda}_0}$.

In Proposition 5.6.2, we saw that the second variation is positive definite on short intervals if strong generalized Legendre condition holds. Now, we describe how we can quantify the length of these intervals. We define

$$\begin{aligned} \mathcal{V}_t &:= \{(\phi(t), \phi, w) \in \mathbb{R} \times L^2([0, T]) \times L^2([0, T], \mathbb{R}^{m-1}) \mid \text{supp}(\phi, w) \subseteq [0, t]\} \cap \ker D_u E \\ t_* &:= \sup\{t \in [0, T] \mid Q|_{\mathcal{V}(t)} \text{ is positive definite}\}. \end{aligned}$$

If $t_* < T$, we say t_* is conjugate to 0 along the trajectory $\tilde{\lambda}$. The remaining part of this subsection is devoted to proving that this notion of conjugate time is equivalent to the one given in Definition 5.4.1.

We have the following result.

Proposition 5.6.3. *Let $\tilde{\lambda}$ be a singular extremal of Problem 5.1 satisfying strong generalized Legendre condition. Suppose that $t_* \in (0, T]$. Then $Q|_{\mathcal{V}(t_*)}$ is singular.*

The proof of this fact follows very closely the proof of Proposition 21.1 in [13], so we omit it. Now, we show how this result can be translated in a more geometric framework and obtain the so-called Jacobi equation. Since now we are interested only on the dynamics up to time t_* , to simplify the notation we simply rename t_* to T .

To derive the Jacobi equation, we look for admissible variations $\bar{v}$ in the kernel of Q :

$$Q(\cdot, \bar{v}) = 0.$$

On the other hand, an admissible variation must satisfy (5.6.6) or, equivalently,

$$\sigma\left(\cdot_{\phi(T)}(Z_I(T)) + \int_0^T \mathcal{Z}_t \cdot_v dt, \nu\right) = 0, \quad \forall \nu \in \Pi_0.$$

Therefore, $\ker D_u E_T$ can be identified with a common kernel of a finite number of linear forms. Since, by assumption $Q(\cdot, \bar{v}) = 0$ on $\ker D_u E_T$, it follows that there is $\nu \in \Pi_0$ such that

$$Q(\cdot, \bar{v}) = \sigma\left(\cdot_{\phi(T)}(Z_I(T)) + \int_0^T \mathcal{Z}_t \cdot_v dt, \nu\right), \quad \text{on } T_u \mathcal{U}. \quad (5.6.12)$$

The equation (5.6.12) splits into two equations:

$$\sigma\left(Z_I(T), \int_0^T \mathcal{Z}_t \bar{v} dt\right) = \sigma\left(Z_I(T), \nu\right) \quad (5.6.13)$$

$$\int_0^T l_t(\cdot_v, \bar{v}) dt + \int_0^T \sigma\left(\mathcal{Z}_t \cdot_v, \int_0^t \mathcal{Z}_s \bar{v} ds\right) dt = \int_0^T \sigma(\mathcal{Z}_t \cdot_v, \nu) dt \quad (5.6.14)$$

We denote

$$\eta(t) := \nu - \int_0^t \mathcal{Z}_s \bar{v} ds. \quad (5.6.15)$$

Since we have the equalities for the forms in (5.6.14), we can pass to equality of the integrands:

$$l_t(\cdot, \bar{v}) = \sigma(\mathcal{Z}_t \cdot_v, \eta(t)) \quad \text{for a.e. } t \in [0, T]$$

Then, since by assumption $l_t(\cdot, \bar{v})$ is an invertible linear map, we can write

$$\bar{v}(t) = l_t^{-1} \sigma\left(\mathcal{Z}_t \cdot_v, \eta(t)\right) \quad \text{for a.e. } t \in [0, T].$$

Finally, using (5.6.15) we find

$$\dot{\eta}(t) = -Z_t \bar{v}(t) = -Z_t l_t^{-1} \sigma(Z_t \cdot, \eta(t)) \quad \text{for a.e. } t \in [0, T].$$

which gives us a non-autonomous linear ODE for η . Therefore, we have derived the *Jacobi equation*.

Concerning (5.6.13), we have

$$\sigma(\eta(T), Z_I(T)) = 0. \quad (5.6.16)$$

It is common to rewrite this condition using skew-orthogonal complement. Using this notation, we can take into account both (5.6.6) and (5.6.16) as

$$\eta(T) \in (\Pi_0 + \mathbb{R}Z_I(T)) \cap (\mathbb{R}Z_I(T))^\perp = \Pi_0^{\mathbb{R}Z_I(T)}.$$

This motivates the following definition.

Definition 5.6.4. A time T is a conjugate time for 0 if and only if there is a non-trivial solution to the following boundary value problem:

$$\dot{\eta}(t) = -Z_t l_t^{-1} \sigma(Z_t \cdot, \eta(t)) \quad \text{for a.e. } t \in [0, T], \quad (5.6.17)$$

$$\eta(0) \in \Pi_0, \eta(T) \in \Pi_0^{\mathbb{R}Z_I(T)}. \quad (5.6.18)$$

The corresponding point $\pi(\tilde{\lambda}(T))$ is said to be conjugate to $\pi(\tilde{\lambda}(0))$.

Recall that previously we have given a slightly different definition of a conjugate point in Definition 5.4.1. But the two definitions are actually equivalent, even though Definition 5.6.4 is more convenient in practice.

Proposition 5.6.5. *Definitions 5.4.1 and 5.6.4 are equivalent.*

Proof. We show that solutions of Equation (5.6.17) satisfy boundary conditions (5.6.18) if and only if they satisfy also the boundary conditions (5.4.4). By Lemma (5.4.3) $\eta(t) = Z_I(t)$ is a solution of the Jacobi equation. However, $Z_I(0) \notin \Pi_0$ and $Z_I(T) \notin \Pi_0$. Therefore, this solution never satisfies (5.6.18) or (5.4.4). So, if T is conjugate to 0 there must be some vector $\nu \in \Pi_0$ such that the corresponding solution to (5.6.17) with $\eta(0) = \nu$ is such that $\eta(T) \in \Pi_0 \cap Z_I(T)^\perp$. Hence, we must have $\sigma(\eta(T), Z_I(T)) = 0$. Since the Equation (5.6.17) is Hamiltonian, we will have that if and only if $\sigma(\eta(0), Z_I(0)) = 0$. $\square$

5.7 Proof of formula (5.6.8)

Here we are going to show how to obtain (5.6.8) from (5.6.2). We have just to integrate by parts with respect to ρ in every integral where it appears.

For the first integral (5.6.3), we have

$$\begin{aligned}
& \int_0^T \rho_1(t) \sigma \left(Z_I(t), \int_0^t \rho_2(s) Z_I(s) ds \right) dt = \\
& \stackrel{\text{int.by parts}}{=} \int_0^T \rho_1(t) \sigma \left(Z_I(t), \phi_2(t) Z_I(t) - \int_0^t \phi_2(s) \dot{Z}_I(s) ds \right) dt = \\
& = - \int_0^T \rho_1(t) \sigma \left(Z_I(t), \int_0^t \phi_2(s) \dot{Z}_I(s) ds \right) dt \stackrel{\text{Fubini}}{=} - \int_0^T \left(\int_s^T \rho_1(t) \sigma \left(Z_I(t), \phi_2(s) \dot{Z}_I(s) \right) dt \right) ds \\
& = - \int_0^T \phi_2(s) \sigma \left(\int_s^T \rho_1(t) Z_I(t) dt, \dot{Z}_I(s) \right) ds = \\
& \stackrel{\text{int.by parts}}{=} - \int_0^T \phi_2(s) \sigma \left(\phi_1(T) Z_I(T) - \phi_1(s) Z_I(s) - \int_s^T \phi_1(t) \dot{Z}_I(t) dt, \dot{Z}_I(s) \right) ds = \\
& = - \sigma \left(\phi_1(T) Z_I(T), \int_0^T \phi_2(s) \dot{Z}_I(s) ds \right) + \\
& + \int_0^T \phi_1(s) \phi_2(s) \sigma \left(Z_I(s), \dot{Z}_I(s) \right) ds + \int_0^T \sigma \left(\phi_1(t) \dot{Z}_I(t), \int_0^t \phi_2(s) \dot{Z}_I(s) ds \right) dt.
\end{aligned}$$

A similar computation for (5.6.4) yields

$$\begin{aligned}
& \int_0^T \sigma \left(\rho_1(t) Z_I(t), \int_0^t Z_s w_2(s) ds \right) dt \stackrel{\text{Fubini}}{=} \int_0^T \sigma \left(\int_s^T \rho_1(t) Z_I(t) dt, Z_s w_2(s) \right) ds \\
& = \int_0^T \phi_1(T) \sigma \left(Z_I(T), Z_s w_2(s) \right) ds - \int_0^T \phi_1(s) \sigma \left(Z_I(s), Z_s w_2(s) \right) ds - \\
& - \int_0^T \sigma \left(\int_s^T \phi_1(t) \dot{Z}_I(t) dt, Z_s w_2(s) \right) ds = \\
& = \sigma \left(\phi_1(T) Z_I(T), \int_0^T Z_s w_2(s) ds \right) + \int_0^T \phi_1(s) \sigma \left(Z_s w_2(s), Z_I(s) \right) ds - \\
& - \int_0^T \sigma \left(\phi_1(t) \dot{Z}_I(t), \int_0^t Z_s w_2(s) ds \right) dt
\end{aligned}$$

And similarly for (5.6.5)

$$\begin{aligned} & \int_0^T \sigma \left(Z_t w_1(t), \int_0^t \rho_2(s) Z_I(s) ds \right) dt = \\ &= \int_0^T \phi_2(t) \sigma (Z_t w_1(t), Z_I(t)) dt - \int_0^T \sigma \left(Z_t w_1(t), \int_0^t \phi_2(s) \dot{Z}_I(s) ds \right) dt \end{aligned}$$

So, in the end, we can write $Q_T(v_1, v_2) = Q_1(v_1, v_2) + Q_2(v_1, v_2) + Q_3(v_1, v_2)$, where

$$\begin{aligned} Q_1(v_1, v_2) = & \int_0^T \left(\frac{\langle w_1(t), w_2(t) \rangle}{r(t)} + \phi_1(t) \phi_2(t) \sigma (Z_I(t), \dot{Z}_I(t)) + \right. \\ & \left. + \phi_1(t) \sigma (Z_t w_2(t), Z_I(t)) + \phi_2(t) \sigma (Z_t w_1(t), Z_I(t)) \right) dt, \end{aligned}$$

$$\begin{aligned} Q_2(v_1, v_2) = & \int_0^T \left(\sigma \left(\phi_1(t) \dot{Z}_I(t), \int_0^t \phi_2(s) \dot{Z}_I(s) ds \right) - \sigma_{\lambda_0} \left(\phi_1(t) \dot{Z}_I(t), \int_0^t Z_s w_2(s) ds \right) - \right. \\ & \left. - \sigma \left(Z_t w_1(t), \int_0^t \phi_2(s) \dot{Z}_I(s) ds \right) \right) + \sigma \left(Z_t w_1(t), \int_0^t Z_s w_2(s) ds \right) dt, \end{aligned}$$

$$Q_3(v_1, v_2) = \sigma \left(\phi_1(T) Z_I(T), \int_0^T (-\phi_2(s) \dot{Z}_I(s) + Z_s w_2(s)) ds \right)$$

One can recognize in Q_1 the principal part of formula (5.6.8).

Chapter 6

Piecewise regular extremals - part 1: main results and a non-convex one-dimensional example

6.1 Introduction

6.1.1 Formulation of the problem and motivations

In this chapter we study the following class of optimal control problems. We first give the precise formulation and the problems and we state the related optimality conditions. Then, we discuss in detail a class of examples to which we apply these optimality conditions and finally, in Chapter 7 we give the complete proofs of our theorems.

Let $U \subset \mathbb{R}^m$ be a subset having the following structure: it is the union of finitely many smooth submanifolds U_i without boundary

$$U = \bigcup_{i=1}^n U_i,$$

with $U_i \cap U_j = \emptyset$ if $i \neq j$. We denote by $m_i = \dim U_i$. Moreover, let M be a d -dimensional smooth manifold and $f : M \times U \rightarrow TM$ such that $f(\cdot, u)$ is a smooth vector field for all $u \in U$ and $f|_{M \times U_i}$ is smooth for every $i \in \{1, \dots, n\}$. For $q_0, q_1 \in M$, we consider the control system

$$\dot{q} = f(q, u(t)), \tag{6.1.1}$$

$$q(0) = q_0, q(T) = q_1, \tag{6.1.2}$$

where $u \in L^\infty([0, T]; U)$. In this Chapter, we mainly focus on the case with free final time $T > 0$. However, in Section 7.2 we show how the problem with fixed final time can be

reduced to the free final time. For this reason, Theorems 6.1.7 and 6.1.8 are formulated for both cases.

Let us denote by $\mathfrak{U}_T := L^\infty([0, T]; U)$ the class of admissible controls on the interval $[0, T]$ and by $\mathfrak{U} = \cup_{T>0} \mathfrak{U}_T$. Moreover, given any control $u \in \mathfrak{U}$, we denote by $q(t; q_0, u)$ the solution of Equation (6.1.1) with initial point q_0 and control equal to u . If there is no ambiguity on the initial point, we simply write $q(t; u)$ instead of $q(t; q_0, u)$.

Let $L : M \times U \rightarrow \mathbb{R}$ be a continuous function such that for every U_i the restriction $L|_{M \times U_i}$ is smooth. We consider the cost of the form $J : \mathfrak{U} \rightarrow \mathbb{R}$ defined as

$$J(u) = \int_0^T L(q(t; u), u(t)) dt, \quad u \in \mathfrak{U}_T. \quad (6.1.3)$$

We want to study the following optimal control problem (OCP) with free final time

$$\min\{J(u) \mid u \in \mathfrak{U}, q(\cdot; u) \text{ satisfies (6.1.1), (6.1.2)}\}, \quad (6.1.4)$$

and the optimal control problem with fixed final time

$$\min\{J(u) \mid u \in \mathfrak{U}_T, q(\cdot; u) \text{ satisfies (6.1.1), (6.1.2), with } T > 0 \text{ fixed}\}. \quad (6.1.5)$$

To motivate our framework, we provide some examples of control systems which are in the form described before.

Example 6.1.1 (Bang-bang control systems). Let us consider a control system of the form

$$\dot{q} = \sum_{i=1}^n u_i(t) f_i(q), \quad u \in U = \bigcup_{i=1}^n U_i.$$

where $f_1, \dots, f_n$ are smooth vector fields on M and $U_i = \{e_i\}$, e_i being the i^{th} -vector of the canonical basis of $\mathbb{R}^n$. We consider the time-optimal control problem, that is, we want to minimize the final time T to reach a point q_1 from a point q_0 . This corresponds to taking $L(q, u) = 1$ in (6.1.3). This includes the well known case of *bang-bang* trajectories, whose corresponding control are piecewise constant functions with values in U .

Example 6.1.2 (OCP with L^1 cost). Let us consider a control system of the form

$$\dot{q} = f_0(q) + \sum_{i=1}^m u_i(t) f_i(q), \quad u \in U = \{u \in \mathbb{R}^m \mid |u| \leq 1\}.$$

where, again, $f_0, f_1, \dots, f_m$ are smooth vector fields on M . Given two points $q_0, q_1 \in M$, we consider the optimal control problem given by the minimization of the functional

$$J(u) = \int_0^T |u(t)| dt,$$

among all controls steering the system from q_0 to q_1 . Even though the integrand of J is not smooth, the set of control values U can be conveniently divided into three submanifolds without boundary: the origin $U_0 = \{0\}$, the interior of the ball $U_1 = \{u \in \mathbb{R}^m \mid |u| < 1\}$ and the boundary of the ball $U_2 = \{u \in \mathbb{R}^m \mid |u| = 1\}$. It is straightforward to check that this optimal control problem is in the form considered above.

Example 6.1.3 (Problem of Calculus of Variations with a non-convex Lagrangian). Let $M = \mathbb{R}^d$ and $U = \mathbb{R}^d$. We consider the free control system

$$\dot{q} = u, \quad q(0) = q_0, \quad q(T) = q_1.$$

The set of admissible control over the time interval $[0, T]$ is $\mathcal{U}_T = L^\infty([0, T]; \mathbb{R}^d)$. Let $L \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R})$ be a positive and super-linear function, that is, for every $q \in \mathbb{R}^d$ it satisfies

$$\lim_{|u| \rightarrow +\infty} \frac{L(q, u)}{|u|} = +\infty.$$

In the literature of Calculus of Variations, a function L satisfying the above assumptions together with strict convexity is called a Tonelli Lagrangian. In our framework, we do not assume convexity: we suppose that L is a *double-well* function, that is it has exactly three distinct critical points, two of which are local minimum points and one is a local maximum point. We want to minimize the functional

$$J(u) = \int_0^T L(q(t), u(t)) dt.$$

Dividing $U = \mathbb{R}^d$ into domains of convexity for L , one can obtain this non-convex problem as a problem in the form above. In Section 6.2, we discuss more details of this example in the one-dimensional case.

In the general framework of optimal control, Pontryagin maximum principle (PMP) provides a powerful tool for solving OCPs by characterizing necessary conditions for optimality. One possible statement of PMP is given in Theorem 2.4.1. We refer to [13] for a wider introduction to PMP and some of its applications. However, in general admissible controls satisfying PMP need not to be optimal, not even locally. Furthermore, even if PMP determines uniquely the control, it does not provide any information about where the control ceases to be the optimal solution, similarly to what happens to the solution of the geodesics equation on Riemannian manifolds.

For some OCPs, PMP alone can be enough to find the optimal solution, but there are many instances for which this task can be quite difficult without additional information. For this reason, in the Sixties some first examples of necessary conditions of second order were introduced, see for instance [51, 46, 52]. These ideas were subsequently recast in a geometric and symplectic framework, see for instance [1, 4, 3], and further developed in the monograph [13]. In [24], a numerical implementation to check second order conditions in

Name	Symbol	Definition	Comments
PMP Hamiltonian	$h_u(\lambda)$	$\langle \lambda, f_u(q) \rangle - L(q, u)$	-
PMP Hamiltonian on the U_j	$h_u^j(\lambda)$	$\langle \lambda, f_u(q) \rangle - L(q, u),$ $u \in U_j$	-
Maximized Hamiltonian over U_j	$H^j(\lambda)$	$\max_{v \in U_j} h_v(\lambda) = h_{u_j(\lambda)}(\lambda)$	-
Maximized Hamiltonian	$H(\lambda)$	$\max_{v \in U} h_v(\lambda) = h_{u_{\max}(\lambda)}(\lambda)$	Along $\tilde{\lambda}_j$, $H(\tilde{\lambda}_t) = H^j(\tilde{\lambda}_t)$ for $t \in (t_{j-1}, t_j)$
Transported-back Hamiltonian vector fields at $\tilde{\lambda}_0$	$\vec{\mathcal{H}}^j$	$(\tilde{\Phi}_{0,t}^{-1})_* \vec{H}^j \Big _{\tilde{\lambda}_0}$	Constant for $t \in [t_{j-1}, t_j]$
Transported-back Hamiltonian	$h_t(\lambda, u)$	$(h_u(\cdot) - h_{\tilde{u}(t)}) \circ \tilde{\Phi}_t$	At $\lambda = \tilde{\lambda}_0$, $h_u(\tilde{\lambda}_0) = h_u(\tilde{\lambda}_t) - H^j(\tilde{\lambda}_t)$ for $t \in (t_{j-1}, t_j)$

Table 6.1: Notation for some Hamiltonian functions introduced throughout the chapter.

an automated way is studied. More recent contributions include [14, 62], where the case of *bang-bang* extremals is studied, [59], where second order conditions for weak-optimality of broken-extremals are considered, and [16, 71, 61, 29, 34, 33, 17], where significant advances in second-order theory were made in many different directions.

The aim of this part of the thesis is to prove both a necessary and a sufficient condition for optimality for the class of problems introduced before. The common feature of the previous examples is that the strong Legendre condition is satisfied on portions of the trajectory, but not on the whole time interval. Classical second-order theory applies naturally to smooth regular extremals, see [13], Chapter 20 and 21. The purpose of this part of the thesis is to provide a unified geometric framework for the wide class of piecewise regular extremals, which includes as particular instances regular extremals and *bang-bang* extremals, and to derive both necessary and sufficient second-order optimality conditions.

6.1.2 Second order optimality conditions for piecewise regular extremals

In this subsection we state the main results about piecewise regular extremals, which are Theorems 6.1.7 and 6.1.8. First, we have to set some more notation. We recall the strong Legendre condition, both for regular arcs and at and at switching points, in a form suited to our formulation. As before, in the present chapter we consider only normal extremal

trajectories. We denote by

$$h_u(\lambda) = \langle \lambda, f(q, u) \rangle - L(q, u), \quad \lambda \in T^*M, u \in U.$$

the normal PMP Hamiltonian function. For $i \in \{1, \dots, n\}$, we define h_u^i to be the restriction of the Hamiltonian function h_u to the submanifold U_i , i.e. $h_u^i(\lambda) = h_u(\lambda)$ for $u \in U_i$ and $\lambda \in T^*M$. Moreover, we introduce the maximized Hamiltonian functions $H, H^i : T^*M \rightarrow \mathbb{R}$ defined as

$$H^i(\lambda) = \max_{u \in U_i} h_u^i(\lambda), \quad H(\lambda) = \max_{u \in U} h_u(\lambda).$$

In particular, we have that $H(\lambda) = \max_{i=1, \dots, n} H^i(\lambda)$. From PMP, we immediately obtain the following Corollary (see Subsections 2.4.2 and 2.4.3).

Corollary 6.1.4 (Legendre condition). *Let $\tilde{u} \in \mathfrak{U}_T$ be an optimal solution either for Problem (6.1.4) or for (6.1.5). Suppose that $\tilde{\lambda}$ is an extremal trajectory of $\tilde{u}$. If for $\ell \in \{1, \dots, n\}$ we have*

$$H(\tilde{\lambda}_t) = H^\ell(\tilde{\lambda}_t),$$

for $t \in I \subset [0, T]$, where I is an interval, then

$$\left. \frac{\partial^2 h_u^\ell}{\partial u^2} \right|_{u=\tilde{u}(t), \lambda=\tilde{\lambda}_t} \leq 0, \quad \text{for a.e. } t \in I. \quad (\text{LC1})$$

Moreover, if for $\tau \in [0, T]$ and $\varepsilon > 0$ there are $\ell_1, \ell_2 \in \{1, \dots, n\}$ such that it holds that $H^{\ell_1}(\tilde{\lambda}_\tau) = H^{\ell_2}(\tilde{\lambda}_\tau) > \max_{i \neq \ell_1, \ell_2} \{H^i(\tilde{\lambda}_t)\}$ and

$$\begin{aligned} H(\tilde{\lambda}_t) &= H^{\ell_1}(\tilde{\lambda}_t) > \max_{i \neq \ell_1} \{H^i(\tilde{\lambda}_t)\}, \quad t \in (\tau - \varepsilon, \tau), \\ H(\tilde{\lambda}_t) &= H^{\ell_2}(\tilde{\lambda}_t) > \max_{i \neq \ell_2} \{H^i(\tilde{\lambda}_t)\}, \quad t \in (\tau, \tau + \varepsilon), \end{aligned}$$

then

$$\left\{ H^{\ell_1}, H^{\ell_2} \right\}(\tilde{\lambda}_\tau) \geq 0. \quad (\text{LC2})$$

We recall that the bilinear form in the left-hand side of (LC1) is called *Legendre form* and the inequalities in (LC1) and (LC2) are called *Legendre conditions* in the literature. If the inequalities are strict, they are referred to as *strong Legendre conditions* and the corresponding extremal arc is called *regular*.

In the book [13], extremal trajectories satisfying strong Legendre condition on the whole time interval are studied in detail (in particular, see Chapter 20 and 21). However, in many OCPs, see for instance Section 6.2, extremal trajectories are obtained by concatenating many regular arcs, while (LC1) is not satisfied on the whole time interval. For this reason, we introduce following class of extremal trajectories.

Definition 6.1.5 (Piecewise regular extremal). Let $\tilde{\lambda} : [0, T] \rightarrow T^*M$ be an extremal trajectory either for Problem (6.1.4) or for (6.1.5), corresponding to a control $\tilde{u} : [0, T] \rightarrow U$. We say that $\tilde{\lambda}$ is *piecewise regular* if it holds that:

1. there are $0 = t_0 < t_1 < \dots < t_k < t_{k+1} = T$ such that for $t \in (t_{j-1}, t_j)$ there is a unique $\ell(j) \in \{1, \dots, n\}$ for which $H(\tilde{\lambda}_t) = H^{\ell(j)}(\tilde{\lambda}_t)$ and it holds that

$$\left. \frac{\partial^2 h_u^{\ell(j)}}{\partial u^2} \right|_{u=\tilde{u}(t), \lambda=\tilde{\lambda}_t} < 0, \quad \text{for a.e. } t \in (t_{j-1}, t_j). \quad (\text{SLC1})$$

2. For every $j = 1, \dots, k$, we have $\{H^{\ell(j)}, H^{\ell(j+1)}\}(\tilde{\lambda}_{t_j}) > 0$.

We call $t_1, \dots, t_k$ the *switching times* of the control $\tilde{u}$. Finally, we say that the pair $(\tilde{u}, \tilde{\lambda})$ is a *piecewise regular extremal pair*.

To avoid using too many indices, we denote by U_j the submanifold $U_{\ell(j)}$ for which Equation (SLC1) holds for $t \in (t_{j-1}, t_j)$, by u_j the corresponding maximizing control on U_j and H^j the maximized Hamiltonian over U_j .

If $\tilde{\lambda}$ is a piecewise regular extremal, there is a neighbourhood of $\tilde{\lambda}$ in T^*M such that the maximizing control

$$u_{\max}(\lambda) := \arg \max_{v \in U} h_v(\lambda), \quad \lambda \in \mathcal{N}$$

is a well defined piecewise smooth function (see Proposition 7.5.1). In particular, the flow of $\vec{H}$ is well defined in $\mathcal{N}$ and is piecewise smooth and Lipschitz continuous.

Consequently, the set of Lagrange multipliers carries naturally a structure of piecewise smooth Lagrangian manifold and, therefore, the corresponding Jacobi curve is a piecewise smooth curve of Lagrangian subspaces. The purpose of our work is to characterize precisely the Jacobi curve of a piecewise regular extremal pair and, as a consequence, obtain the second-order optimality condition following the spirit of Chapters 3 and 4. To this aim, we introduce the following objects:

1. the flow on T^*M of $\vec{h}_{\tilde{u}}$, which we denote by $(\tilde{\Phi}_{\tau_1, \tau_2})_{\tau_1, \tau_2 \in \mathbb{R}}$;
2. the autonomous flow on T^*M of $\vec{H}$, denoted by Φ_t .
3. the transported-back flow ϕ_t is defined by

$$\phi_t = \tilde{\Phi}_{0,t}^{-1} \circ \Phi_{t-t_{j-1}} \circ \tilde{\Phi}_{0,t_{j-1}} \quad t \in (t_{j-1}, t_j), \quad j = 1, \dots, k+1; \quad (6.1.6)$$

4. the transported-back Hamiltonian vector fields at $\tilde{\lambda}_0$, $\vec{\mathcal{H}}^j := (\tilde{\Phi}_{0,t}^{-1})_* \vec{H}^j|_{\tilde{\lambda}_0}$ for $j = 1, \dots, k+1$, $t \in (t_{j-1}, t_j)$;

Notice that $(\tilde{\Phi}_{0,t}^{-1})_* \vec{H}^j|_{\tilde{\lambda}_0} = (\tilde{\Phi}_{0,t_{j-1}}^{-1})_* \vec{H}^j|_{\tilde{\lambda}_0}$ for $t \in [t_{j-1}, t_j]$ since $\vec{H}^j$ generates the flow $\tilde{\Phi}_{\tau_1, \tau_2}$ for $\tau_1, \tau_2 \in [t_{j-1}, t_j]$. Hence, the vector $\vec{H}^j$ does not depend on time.

Theorem 6.1.6 (Structure of the Jacobi curve). *Let $(\tilde{u}, \tilde{\lambda})$ be a piecewise regular extremal pair for Problem (6.1.4). Suppose that $(\tilde{u}, \tilde{\lambda})$ are defined on the time interval $[0, T]$. Then, the Jacobi curve $\Lambda : [0, T] \rightarrow \mathfrak{L}(T_{\tilde{\lambda}_0}(T^*M))$ of the pair $(\tilde{u}, \tilde{\lambda})$ is constructed via the following algorithm: we set $\Lambda_0 = \Pi$; then, for $j = 0, 1, \dots, k$:*

1. *we have that $\Lambda_{t_j}^+ = \Lambda_{t_j}^{\vec{H}^{j+1}}$;*
2. *for $t \in (t_j, t_{j+1})$, $\Lambda_t = B_t \Lambda_{t_j}^+$, where $B_t := (\phi_t)_*$ is the linearization flow of the flow ϕ_t ;*
3. *for $t = t_{j+1}$, $\Lambda_{t_{j+1}} = \lim_{t \rightarrow t_{j+1}-} \Lambda_t$.*

As an immediate corollary of Theorems 3.1.7 and 6.1.6, we obtain the following necessary second-order condition of optimality for piecewise regular extremals.

Theorem 6.1.7 (Necessary condition for optimality). *Let $(\tilde{u}, \tilde{\lambda})$ be a piecewise regular extremal pair, either for Problem (6.1.4) or for (6.1.5). Suppose that $\tilde{u}$ is of corank r . Then, if $\mu_\Pi(\Lambda) - \left(\frac{d}{2} - \dim\left(\Pi \cap \bigcap_{t \in [0, T]} \Lambda_t\right)\right) \geq r$, $\tilde{u}$ is not a weak-local minimizer with respect to the L^1 -topology of $\mathfrak{U}$.*

Finally, the absence of conjugate points is a sufficient condition for optimality.

Theorem 6.1.8 (Sufficient condition, absence of conjugate times). *Let $(\tilde{u}, \tilde{\lambda})$ be a piecewise regular extremal pair, either for Problem (6.1.4) or for (6.1.5). Suppose that there are no conjugate times along the trajectory $\tilde{\lambda}$. Then, $\tilde{q}$ is locally strongly strictly optimal.*

6.2 A one-dimensional non-convex Lagrangian example

Before going into the details of the proofs of our main results, we discuss the class of examples introduced in Example 6.1.3. This section has two purposes. First, it shows that piecewise regular extremals are natural optimal candidates in a simple non-convex variational problem. Second, it illustrates that, despite the abstract definition, the Jacobi curve and conjugate points can be computed concretely and, in dimension one, this reduces to a simple sign condition.

Consider the free control system

$$\dot{q} = u, \tag{6.2.1}$$

where $q \in \mathbb{R}$ and $u \in \mathbb{R}$, i.e., using the notation introduced before, we have $U = \mathbb{R}$ and $\mathcal{U} = L^\infty([0, T]; \mathbb{R})$. The final time $T > 0$ can be either free or fixed. We fix two values $q_0, q_1 \in \mathbb{R}$ and we want to minimize the functional

$$J(u) = \int_0^T L(q(t), u(t)) dt, \tag{6.2.2}$$

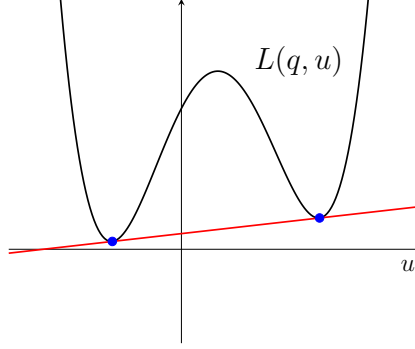


Figure 6.1: Example of double-well functional for fixed q , highlighting the bi-tangent line.

subject to (6.2.1) and $q(0) = q_0$ and $q(T) = q_1$, where the function $L \in C^\infty(\mathbb{R} \times \mathbb{R}; \mathbb{R})$ satisfies the following hypothesis:

1. for every $q \in \mathbb{R}$, the function $u \mapsto L(q, u)$ is super-linear, that is $\lim_{|u| \rightarrow +\infty} \frac{L(q, u)}{|u|} = +\infty$;
2. $L(q, u) > 0$ for every $q, u \in \mathbb{R}$;
3. for every $q \in \mathbb{R}$, $\partial_u^2 L(q, \cdot)$ has only simple zeros;
4. for every $q \in \mathbb{R}$, $L(q, \cdot)$ is a *double-well* function, that is it has exactly three distinct critical points, two of which are local minimum points and one is a local maximum point.

In the Calculus of Variations language, a function L satisfying assumptions 1. together with strict convexity is said to be a Tonelli Lagrangian function. Here, we study this minimization problem in dimension 1 but without the usual convexity assumption. We refer to [56] for an introduction to Tonelli Lagrangian functions and their role in the Calculus of Variations. We point out that Theorems 6.1.7 and 6.1.8 apply also to the case $d > 1$, but a precise analysis of this case goes beyond the purpose of this thesis. Assumption 2. in the case of free final time is necessary to guarantee that $\inf_u J(u) > -\infty$. In the case of fixed time, if L is negative somewhere and $\inf_{q,u} L(q, u) > -\infty$, then by adding a constant to L assumption 2. is always satisfied and this modification of L does not change the minimizing controls. Finally, assumptions 3. and 4. are present just to simplify the exposition: our arguments generalize directly to the case of integrands admitting a many local minima and maxima.

First, we state our results about this class of examples: our first theorem says that minimizers of this OCPs exist and, under some natural conditions, the extremal trajectories are piecewise regular; then, we give a practical characterization of conjugate points for both

the free and fixed final time problem. After this first part, we discuss in more detail the case of a particular Lagrangian function of the form $L(q, u) = \phi(u) + V(q)$, with $V(q) = \frac{q^2}{2}$, which corresponds to a “generalized” harmonic oscillator. This way, we illustrate some intuitive ideas behind our analysis in a simpler case. Finally, we give a full proof of these results.

6.2.1 Statements of the results

Our first result is about the structure of the extremal trajectories for this class of problems. By PMP, we define the Hamiltonian function

$$h_u(p, q) = pu - L(q, u), \quad (6.2.3)$$

where $p \in \mathbb{R}^*$ and $q, u \in \mathbb{R}$ as before. In the following, we identify $\mathbb{R}^*$ with $\mathbb{R}$ by fixing a global coordinate, so that $T^*\mathbb{R} \simeq \mathbb{R}_p \times \mathbb{R}_q$. The corresponding Hamiltonian system reads

$$\begin{cases} \dot{q} = u, \\ \dot{p} = \partial_q L(q, u), \end{cases} \quad (6.2.4)$$

where control u has to satisfy the maximization principle $u(t) = \arg \max_{v \in \mathbb{R}} h_v(\lambda_t)$, where $\lambda_t = (p(t), q(t))$ denotes the extremal trajectory. In this case, we show that there are two competing controls, each being the maximizing one on some specific region of the plane $\mathbb{R}_p \times \mathbb{R}_q$. Each of these two regions corresponds, in some sense, to a convexity region of the function L . Thus, in the terminology of Section 6.1.1, there are two submanifolds U_1, U_2 and the maximized Hamiltonian function is $H(\lambda) = \max\{H^1(\lambda), H^2(\lambda)\}$, where the two functions H^1, H^2 are smooth.

Theorem 6.2.1. *Assume that the function L satisfies the assumptions 1. to 4. listed above. Moreover, let $H^1, H^2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ be defined as in Equation 6.2.3 and define $\Sigma = \{\lambda \in \mathbb{R}^2 \mid H^1(\lambda) = H^2(\lambda)\}$ and $\mathcal{S} := \{\lambda \in \mathbb{R}^2 \mid \{H^1, H^2\}(\lambda) = 0\}$. Suppose that the set $\mathcal{S} \cap \Sigma =: \mathcal{Z}$ is discrete. Then, the following conclusions hold.*

1. *The set of initial values of $p_0 \in \mathbb{R}$ for which the corresponding extremal trajectory is not piecewise regular is discrete.*
2. *If in addition to the previous assumptions we have that $\alpha \vec{H}^1(\lambda) + (1 - \alpha) \vec{H}^2(\lambda) \neq 0$ for all $\alpha \in [0, 1]$ and all $\lambda \in \mathcal{Z}$, then the minimum of the functional (6.2.2) exists for both the free and the fixed final time problem.*

The first assertion shows that, outside a discrete exceptional set of initial covectors, these candidates fall within the piecewise regular framework developed in the Chapter. The second one guarantees the existence of optimal candidates in this class of problems. After the proof of these results, we also discuss briefly what happens when these hypotheses

fail, see Remark 6.2.5. Notice that, for generic smooth functions H^1, H^2 , the two sets Σ and $\mathcal{S}$ intersect transversally. Hence, the hypothesis on $\mathcal{Z}$ is satisfied generically.

The next two propositions are not specific to the double-well Lagrangians considered in this section. They apply to every one-dimensional optimal control problem for which the reference extremal is piecewise regular. Their role is to give a concrete characterization of conjugate times in dimension one, where the Jacobi curve has a rather easy and explicit description.

First of all, since we have a global chart for the cotangent space, that is $T^*M = T^*\mathbb{R}_q \simeq \mathbb{R}_p \times \mathbb{R}_q$, we do not need to transport back the dynamics to the initial point. More precisely, the Jacobi curve in the case of free final time, since $\dim \Lambda_t = 1$, is completely determined by the condition $\vec{H}(\tilde{\lambda}_t) \in \Lambda_t$ and the discontinuities of Λ at the switching times corresponds exactly to the discontinuities of $\vec{H}(\tilde{\lambda}_t)$. Conjugate times can be characterized by the following property.

Proposition 6.2.2 (Conjugate times in one-dimensional problems with free final time). *Let $(\tilde{u}, \tilde{\lambda})$ be a piecewise regular extremal pair for problem (6.1.4) with $M = \mathbb{R}$. Then, denoting by $\tilde{q} = \pi(\tilde{\lambda})$, we have that a time $t \in [0, T]$ is conjugate to 0 along $\tilde{\lambda}$ if and only if*

- *either $\dot{\tilde{q}}(t) = 0$, if t is not a switching time;*
- *or $\lim_{\tau \rightarrow t+} \text{sgn}(\dot{\tilde{q}}(\tau)) \neq \lim_{\tau \rightarrow t-} \text{sgn}(\dot{\tilde{q}}(\tau))$, if t is a switching time.*

For the problem with fixed final time, as explained in Section 7.2, we need to introduce an additional time variable to reduce to the free final time case (see Equation (7.2.3)). So, as in that Section, we let $t = t(s)$, where s is the new time variable, so that the state manifold is $\mathbb{R}_t \times \mathbb{R}_q$ and its cotangent space is identified with $(\mathbb{R}_\tau \times \mathbb{R}_p) \times (\mathbb{R}_t \times \mathbb{R}_q)$. Recall that, along the reference extremal, we take $\nu \equiv 0$, hence $t(s) = s$; we therefore use the same symbols t_j for the switching times in the original and reparametrized time. Thus, the Jacobi curve is a curve of two-dimensional subspaces, which can be represented in a matrix form as

$$\Lambda = \begin{bmatrix} \tau_1 & \tau_2 \\ t_1 & t_2 \\ y_1 & y_2 \\ x_1 & x_2 \end{bmatrix}, \quad (6.2.5)$$

where (τ_i, t_i, y_i, x_i) , $i = 1, 2$ are any two vectors generating the subspace Λ . Here, with a slight abuse of notation, we use again the letters t, τ to denote both the variables in $T^*(\mathbb{R}_t \times M)$ and the corresponding variables in the tangent space $T_{(t, \lambda)}(T^*(\mathbb{R}_t \times M))$. Clearly, replacing the one of the two column by any linear combination of these vectors

yields the same subspace Λ . In this notation, the initial subspace Λ_0 is given by

$$\Lambda_0 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad (6.2.6)$$

and the algorithm to obtain the Jacobi curve goes as follows:

- $\Lambda_0^+ = (\Lambda_0 \cap (\partial_t + \vec{H}(\lambda_0))^\perp) + \mathbb{R}(\partial_t + \vec{H}(\lambda_0))$;
- $\Lambda_s = (\Phi_s)_*(\Lambda_0^+)$, $s \in (0, t_1]$. Here $\Phi_s = e^{s\vec{H}}$ and $(\Phi_s)_*$ denotes its linearization in the variable λ at the point $\tilde{\lambda}_s$. Moreover, $(\Phi_s)_*(\partial_t|_{\lambda_0}) = \partial_t|_{\tilde{\lambda}_s}$ and $(\Phi_s)_*(\partial_\tau|_{\lambda_0}) = \partial_\tau|_{\tilde{\lambda}_s}$, so that the first two rows of Λ are constant for $s \in (0, t_1]$. Notice also that $\vec{H}(\tilde{\lambda}_s)$ coincides either with $\vec{H}^1(\tilde{\lambda}_s)$ or with $\vec{H}^2(\tilde{\lambda}_s)$ in this time interval.
- $\Lambda_{t_1}^+ = (\Lambda_{t_1} \cap (\partial_t + \vec{H}(\tilde{\lambda}_{t_1+}))^\perp) + \mathbb{R}(\partial_t + \vec{H}(\tilde{\lambda}_{t_1+}))$, where $\vec{H}(\tilde{\lambda}_{t_1+}) = \lim_{s \rightarrow t_1+} \vec{H}(\tilde{\lambda}_s)$;
- then repeat the same procedure on the remaining part of the interval $[0, T]$.

In this case, the system of equations governing the evolution of Λ_t in the interval of smoothness are obtained by linearizing with respect to p, q the Equation (6.2.4) with u equal to the maximizing control. If T is fixed, the conjugate times can also be characterized in a similar fashion to what happens in the case of T free.

Proposition 6.2.3 (Conjugate times in one-dimensional problems with fixed final time). *Let $(\tilde{u}, \tilde{\lambda})$ be a piecewise regular extremal pair for the problem (6.2.1), (6.2.2). Let Λ be represented in matrix form as in Equation (6.2.5) and choose the representation such that $(\tau_1(s), t_1(s), y_1(s), x_1(s)) = \partial_t + \vec{H}(\tilde{\lambda}_s)$ and $\tau_2(s) = 1, t_2(s) = 0$ for all $s \in [0, T]$. Then, a time $s \in [0, T]$ is conjugate to 0 along $\tilde{\lambda}$ if and only if*

1. *either $x_2(s) = 0$, if s is not a switching time;*
2. *or $\lim_{\varsigma \rightarrow s+} \text{sgn}(x_2(\varsigma)) \neq \lim_{\varsigma \rightarrow s-} \text{sgn}(x_2(\varsigma))$, if s is a switching time.*

Moreover, using again these coordinates, the right limit of Λ at a switching time t_j can be computed as follows: the first column can be chosen equal to $\partial_t + \vec{H}(\tilde{\lambda}_{t_j+})$ and, letting $\vec{H}(\tilde{\lambda}_{t_j-}) = \lim_{s \rightarrow t_j-} \vec{H}(\tilde{\lambda}_s)$, the second one is equal to

$$\partial_\tau + \alpha(\vec{H}(\tilde{\lambda}_{t_j-}) - \vec{H}(\tilde{\lambda}_{t_j+})) + \begin{pmatrix} y_2(t_j) \\ x_2(t_j) \end{pmatrix}, \quad \text{where } \alpha = \frac{\sigma(\vec{H}(\tilde{\lambda}_{t_j+}), (y_2(t_j), x_2(t_j))^\top) - 1}{\sigma(\vec{H}(\tilde{\lambda}_{t_j-}), \vec{H}(\tilde{\lambda}_{t_j+}))}.$$

For every piecewise regular extremal in one-dimensional OCP, Propositions 6.2.2 and 6.2.3 provide an explicit test for the absence of conjugate points. Therefore, whenever the corresponding sign condition rules out the presence of conjugate times on $(0, T]$, Theorem 6.1.8 yields strong local optimality of the trajectory. Conversely, if the Jacobi curve acquires sufficient Maslov index, for instance, in the corank-one case after the first conjugate time, Theorem 6.1.7 excludes local optimality.

6.2.2 Generalized harmonic oscillator

Before going into the details of the proof of the results exposed in the previous Subsection, to give some intuition behind our arguments we discuss a particular example modelled on the classical harmonic oscillator.

For this model, we can carry out exact computations, which also allows us to compare the free final time and the fixed final time cases. More precisely, we find that switching times are never conjugate to 0 in the free final time problem, while conjugate times may occur along regular arcs. In the fixed final time problem, the additional time variable changes the Jacobi curve and switching times may contribute to conjugacy according to the sign condition in Proposition 6.2.3. We point out that most of the analysis in the remaining part of this Subsection can be extended without substantial changes to any Lagrangian function of the form $L(q, u) = \phi(u) - V(q) + C$. Our particular choice is mainly due to expository reasons: in this case computations are particularly simple and can be carried out exactly, simplifying greatly the presentation although retaining the main ideas behind the arguments for the general case.

We consider $L(q, u) = \phi(u) - V(q) + C$, with $V(q) = \frac{q^2}{2}$, $C > 0$ and

$$\phi(u) = \begin{cases} \frac{1}{2}(u + \frac{1}{2})^2 - \frac{1}{2}, & u \leq \frac{1}{2}, \\ \frac{1}{2}(u - \frac{1}{2})^2 + \frac{1}{2}, & u \geq \frac{3}{2}, \end{cases}$$

is such that ϕ is globally C^2 and $\phi(u) > u - \frac{1}{2}$ for $u \in (\frac{1}{2}, \frac{3}{2})$. We notice that L does not satisfy the assumption $L > 0$. In order to retain the existence discussion within the global framework of Theorem 4.1, we modify the potential V outside a sufficiently large compact set. More precisely, fix $R > 0$ and choose a function $V_R \in C^\infty(\mathbb{R})$ such that

$$V_R(q) = \frac{q^2}{2} \quad \text{for } |q| \leq R,$$

and such that V_R is bounded from above on $\mathbb{R}$. Since ϕ is bounded from below, one can choose a constant $C_R > 0$ such that

$$L_R(q, u) := \phi(u) - V_R(q) + C_R > 0 \quad \text{for every } (q, u) \in \mathbb{R}^2.$$

The Lagrangian L_R is super-linear in the control variable and has the same double-well structure with respect to u as ϕ . Hence it satisfies the hypotheses of Theorem 4.1.

In the following, we carry out the computations using the explicit expression

$$V(q) = \frac{q^2}{2}.$$

If the reference extremal and the portion of its phase portrait considered below are contained in $\{|q| < R\}$, then the Hamiltonian vector field associated with L_R agrees there with the Hamiltonian vector field of the oscillator model. Consequently, all the computations of the extremal flow, the Jacobi curve, and the conjugate times remain unchanged along the reference extremal. The globally modified Lagrangian L_R is used only to ensure that the existence statement follows from Theorem 4.1.

The PMP Hamiltonian function of this problem is $h_u(p, q) = pu - L(q, u) = pu - \phi(u) + V(q) - C$, and, letting $\psi(p) := \max_{u \in \mathbb{R}} pu - \phi(u)$, the maximized Hamiltonian function $H = \max_{u \in \mathbb{R}} h_u$ is

$$H(p, q) = \psi(p) + V(q) = \frac{p^2}{2} + \frac{|p-1|}{2} + \frac{q^2}{2} - C.$$

Indeed, by the super-linearity assumption on L , we have that the maximum of h_u is attained at a critical point and $\partial_u h_u(p, q) = 0$ if and only if $p = \phi'(u)$. Hence, a direct computation shows that $h_u(p, q)$ reaches its global maximum for $u = p - \frac{1}{2}$ if $p < 1$ and $u = p + \frac{1}{2}$ if $p > 1$. For $p = 1$, the line $y(u) = u - \frac{1}{2}$ is tangent to the graph of ϕ both at $u = \frac{1}{2}$ and at $u = \frac{3}{2}$. So, defining $u_1(p) = p - \frac{1}{2}$ and $u_2(p) = p + \frac{1}{2}$, the maximized Hamiltonian is the maximum of the two Hamiltonians $H^1(p, q) := h_{u_1(p)}(p, q)$ and $H^2(p, q) := h_{u_2(p)}(p, q)$. First, we focus on the structure of the extremal trajectories, showing that most of them

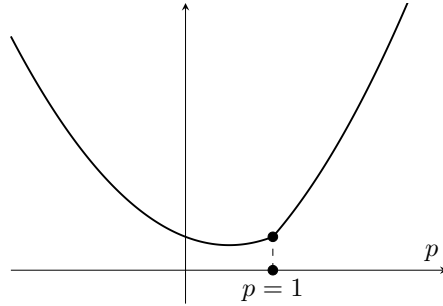


Figure 6.2: Graph of $H(p, q) = \frac{p^2}{2} + \frac{|p-1|}{2} + \frac{q^2}{2} - 1$ for fixed q .

are piecewise regular, and we find the conjugate points. Then, in the last part of the subsection, we prove existence of the minimum for the OCP.

By PMP, an extremal trajectory satisfies the Hamiltonian equations $\dot{p} = -\partial_q V(q)$, $\dot{q} = \partial_p \psi(p)$, which reads

$$\begin{cases} \dot{p} = -q, \\ \dot{q} = p - \frac{1}{2}, \end{cases} \quad \text{if } p(t) < 1, \text{ or } \begin{cases} \dot{p} = -q, \\ \dot{q} = p + \frac{1}{2}, \end{cases} \quad \text{if } p(t) > 1, \quad (6.2.7)$$

and the set of switching points is $\Sigma = \{(1, q) \mid q \in \mathbb{R}\}$. The Poisson bracket $\{H^1, H^2\}$ is easily computed:

$$\{H^1, H^2\}(p, q) = \frac{\partial H^1}{\partial p} \frac{\partial H^2}{\partial q} - \frac{\partial H^1}{\partial q} \frac{\partial H^2}{\partial p} = \left(p - \frac{1}{2}\right)q - \left(p + \frac{1}{2}\right)q = -q.$$

Thus, the set $\mathcal{S} = \{(p, 0) \mid p \in \mathbb{R}\}$ and the set $\mathcal{Z} = \Sigma \cap \mathcal{S} = \{(1, 0)\}$ is discrete.

Since the system (6.2.1) is one-dimensional and by PMP the maximized Hamiltonian is constant along extremal trajectories, the solutions of $\dot{\lambda} = \vec{H}(\lambda)$ are contained in the level sets of H (see Figure 6.3), that is, the system (6.2.4) is integrable. In particular, switching points cannot accumulate and an extremal trajectory fails to be piecewise regular only if the Poisson bracket $\{H^1, H^2\} = 0$ at the switching point. This is possible only at the point $(p, q) = (1, 0)$, which is contained in the level set $H(\lambda) = \frac{1}{2} - C$. Explicit solutions of

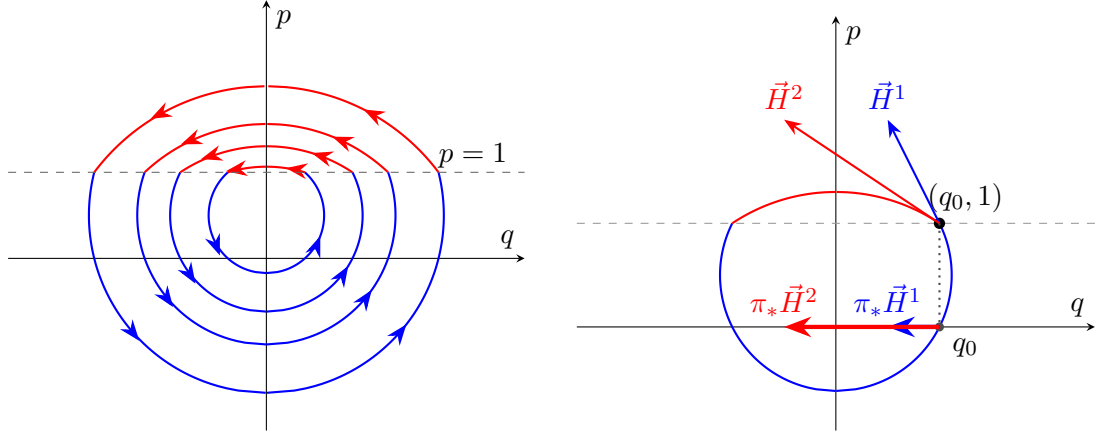


Figure 6.3: Phase portrait of $H(p, q) = \frac{p^2}{2} + \frac{|p-1|}{2} + \frac{q^2}{2} - 1$. In this case, the vector fields $\pi_* \vec{H}^1$ and $\pi_* \vec{H}^2$ have the same sign at switching points, thus switching points are never conjugate.

these equations are easily computed: if $p(0) < 1$, then the first smooth arc of the solution satisfies $\ddot{q} = -q$, $\ddot{p} = -p + \frac{1}{2}$ having solution

$$q(t) = q(0) \cos t + \left(p(0) - \frac{1}{2}\right) \sin t, \quad p(t) = -q(0) \sin t + \left(p(0) - \frac{1}{2}\right) \cos t + \frac{1}{2}. \quad (6.2.8)$$

If instead $p(0) > 1$, then

$$q(t) = q(0) \cos t + \left(p(0) + \frac{1}{2}\right) \sin t, \quad p(t) = -q(0) \sin t + \left(p(0) + \frac{1}{2}\right) \cos t - \frac{1}{2}.$$

Then, when passing from one regular arc to the next one, one can find an analogous formula by replacing the initial conditions $(p(0), q(0))$ with the initial point of the new regular arc.

Now, we fix any piecewise regular extremal $\tilde{\lambda}$ and we want to determine its Jacobi curve. We begin with the case of free final time. As already pointed out, the Jacobi curve is $\Lambda_0 = (1, 0)^\top$ and for $t > 0$

$$\Lambda_t = \mathbb{R}\vec{H}(p(t), q(t)) = \begin{cases} \mathbb{R}\vec{H}^1(p(t), q(t)) = \mathbb{R}(-q(t), p(t) - \frac{1}{2}) & \text{if } p(t) < 1, \\ \mathbb{R}\vec{H}^2(p(t), q(t)) = \mathbb{R}(-q(t), p(t) + \frac{1}{2}) & \text{if } p(t) > 1, \end{cases}$$

and extended such that Λ_t is left-continuous at the switching times. If we apply Proposition 6.2.2, we see that for $p(t) \approx 1$ we have that $\dot{q}$ is positive and it follows that switching times are never conjugate times. This is also seen clearly from the picture (see Figure 6.3): at the switching point, we have $\Lambda_{t_j} + \Lambda_{t_j}^+ = \mathbb{R}^2$ and $\Lambda_{t_j} \cap \Lambda_{t_j}^+ = \{0\}$, so $\pi_*(\Lambda_{t_j} + \Lambda_{t_j}^+) \setminus \pi_*(\Lambda_{t_j} \cap \Lambda_{t_j}^+) = \mathbb{R} \setminus \{0\}$ and t_j is conjugate if and only if $\pi_*\vec{H}^1(\tilde{\lambda}_{t_j})$ and $\pi_*\vec{H}^2(\tilde{\lambda}_{t_j})$ have different signs, which is never the case in our example. Thus, conjugate times can occur only along regular arcs on which $p(t) < 1$, since on an arc with $p(t) > 1$ one has

$$\dot{q}(t) = p(t) + \frac{1}{2} > \frac{3}{2},$$

and hence $\dot{q}$ cannot vanish.

Let (a, b) be a time interval such that $\tilde{\lambda}|_{(a, b)}$ is a regular arc on which $p(t) < 1$, and write

$$\lambda(a) = (p_a, q_a).$$

Setting $s = t - a$, the solution on this arc is obtained from (6.2.8) by taking $(p(0), q(0)) = (p_a, q_a)$. Therefore,

$$\dot{q}(t) = p(t) - \frac{1}{2} = -q_a \sin s + \left(p_a - \frac{1}{2}\right) \cos s.$$

By Proposition 4.2, the conjugate times in the interval (a, b) are precisely the times $t = a + s$ such that

$$-q_a \sin s + \left(p_a - \frac{1}{2}\right) \cos s = 0, \quad 0 < s < b - a.$$

If $(q_a, p_a - \frac{1}{2}) \neq (0, 0)$, let $s_* \in (0, \pi]$ be the first positive solution of this equation. Then all the conjugate times on the same regular arc are of the form

$$t = a + s_* + k\pi, \quad k \in \mathbb{N} \cup \{0\},$$

provided that $t \in (a, b)$. After a switching time, the same computation applies by taking as (p_a, q_a) the initial point of the next regular arc. This completes the discussion about conjugate points in the case of free final time.

Let us analyse conjugate points for the fixed final time problem. As described before, in this case we have to add a variable to the state space (and, consequently, another one

to the cotangent space), so that the Jacobi curve is a curve of two-dimensional subspaces. The initial space is given by (6.2.6). We assume, for simplicity, that on the time interval $[0, t_1]$ the maximized Hamiltonian is equal to H^1 , on $[t_1, t_2]$ is equal to H^2 , then on $[t_2, t_3]$ is again H^1 and so on. The case with the reversed order can be dealt with similarly. Then, following the convention introduced before, the right limit of Λ at $t = 0$ can be represented as

$$\Lambda_0^+ = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ -q_0 & -\frac{1}{p_0-1/2} \\ p_0 - \frac{1}{2} & 0 \end{bmatrix}.$$

where $\tilde{\lambda}_0 = (p_0, q_0)$ is the initial point of the extremal trajectory. For $t \in (0, t_1)$, the first two rows are constant, while the last two rows evolve according to the equation obtained by linearizing (6.2.7): we have that $\dot{y}_i = -\partial_q^2 V(q(s))x_i$, $\dot{x}_i = \partial_p^2 \psi(p(s))y_i$, which reads

$$\begin{cases} \dot{y}_i = -x_i, \\ \dot{x}_i = y_i, \end{cases} \quad i = 1, 2, \quad \text{with} \quad \begin{pmatrix} y_1(0) \\ x_1(0) \end{pmatrix} = \begin{pmatrix} -q_0 \\ p_0 - \frac{1}{2} \end{pmatrix}, \quad \begin{pmatrix} y_2(0) \\ x_2(0) \end{pmatrix} = \begin{pmatrix} -\frac{1}{p_0-1/2} \\ 0 \end{pmatrix}.$$

For $p_0 \neq \frac{1}{2}$, the solutions are

$$x_2(s) = -\frac{1}{p_0 - 1/2} \sin s, \quad y_2(s) = -\frac{1}{p_0 - 1/2} \cos s.$$

and $(y_1(t_1), x_1(t_1)) = \vec{H}^1(p(t_1), q(t_1))$. For $p_0 = \frac{1}{2}$, the resulting Jacobi curve is continuous at $t = 0$ and the solutions are $x_2(t) = 0$, $y_2(t) = 0$ constant. We see immediately that, if $p_0 \neq \frac{1}{2}$, $s = k\pi$ are conjugate times.

We have now to characterize switching times that are also conjugate times. The Jacobi curve is determined up to time t_1 :

$$\Lambda_{t_1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ y_1(t_1) & y_2(t_1) \\ x_1(t_1) & x_2(t_1) \end{bmatrix}.$$

Denote by $\eta_2(t) = (y_2(t), x_2(t))^\top$. By Proposition 6.2.3, the right limit of Λ at t_1 is

$$\Lambda_{t_1}^+ = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ \vec{H}^2 & \alpha(\vec{H}^1 - \vec{H}^2) + \eta_2 \end{bmatrix}, \quad \text{where } \alpha = \frac{\sigma_{\lambda_{t_1}}(\vec{H}^2, \eta_2) - 1}{\sigma_{\lambda_{t_1}}(\vec{H}^1, \vec{H}^2)},$$

where $\vec{H}^1, \vec{H}^2$ are evaluated at λ_{t_1} and η_2 is evaluated at t_1 . We have that

$$\vec{H}^1 - \vec{H}^2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \sigma_{\lambda_{t_1}}(\vec{H}^2(\lambda_{t_1}), \eta(t_1)) = -q(t_1)x_2(t_1) - \frac{3}{2}y_2(t_1).$$

After some simplifications, the formula for $\Lambda_{t_1}^+$ reads

$$\Lambda_{t_1}^+ = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ -q(t_1) & y_2(t_1) \\ \frac{3}{2} & -\frac{3}{2} \frac{y_2(t_1)}{q(t_1)} - \frac{1}{q(t_1)} \end{bmatrix}.$$

In particular, by Proposition 6.2.3, t_1 is conjugate to 0 if

$$x_2(t_1) \left(-\frac{3}{2} \frac{y_2(t_1)}{q(t_1)} - \frac{1}{q(t_1)} \right) \leq 0.$$

If $p_0 = \frac{1}{2}$, then $x_2(t_1) = y_2(t_1) = 0$ and the previous inequality holds true. Otherwise, since $-q = \sigma(\vec{H}^1, \vec{H}^2) > 0$, it reduces to $x_2(t_1)(3y_2(t_1) + 2) \leq 0$, and, replacing the formulas for x_2 and y_2 , this condition reads

$$3P_0^2 \sin t_1 \cos t_1 + 2P_0 \sin t_1 \leq 0,$$

where $P_0 = -\frac{1}{p_0 - 1/2}$. Letting $\theta = \arccos \frac{2p_0 - 1}{3}$, a direct computation shows that t_1 satisfies this inequality if

1. $t_1 \in [\pi, 2\pi]$ if $p_0 \leq -1$;
2. $t_1 \in [\theta, \pi] \cup [2\pi - \theta, 2\pi]$ if $-1 \leq p_0 < 1$.

Finally, since $p(t_1) = 1$, from Equation (6.2.8) t_1 must satisfy

$$-q(0) \sin t_1 + \left(p(0) - \frac{1}{2} \right) \cos t_1 = \frac{1}{2}.$$

Thus, putting together these conditions one can find whether t_1 is conjugate to zero depending on $p(0)$ and $q(0)$.

We are left to prove that the minimum for the OCP exists. To this aim, one can show that the set of extremal trajectories for the OCP with ϕ coincides with the set of extremal trajectories for the OCP with the convexification of ϕ . More precisely, we define $\phi^C(u) = u - \frac{1}{2}$ for $\frac{1}{2} \leq u \leq \frac{3}{2}$ and $\phi^C(u) = \phi(u)$ otherwise.

Proposition 6.2.4. *Under the assumptions of Theorem 6.2.1, a curve $\lambda : [0, T] \rightarrow \mathbb{R}_p \times \mathbb{R}_q$ is an extremal trajectory for the OCP (6.2.1), (6.2.2) with $L(q, u) = \phi(u) - V(q)$ if and only if λ is also an extremal trajectory for the OCP with $L(q, u) = \phi^C(u) - V(q)$.*

Sketch of the proof. Since the maximizing control for the OCP with $L(q, u) = \phi(u) - V(q)$ does not take values in the interval where the convexity of ϕ fails, any extremal trajectory for the OCP with $L(q, u) = \phi(u) - V(q)$ is also an extremal trajectory for the OCP with $L(q, u) = \phi^C(u) - V(q)$. On the other hand, an extremal trajectory $(p(t), q(t))$ for the OCP with $L(q, u) = \phi^C(u) - V(q)$ is not an extremal trajectory for the problem with $L(q, u) = \phi(u) - V(q)$ only if $p(t) = 1$ for some non-trivial time interval. But this is not possible since $p(t) = 1$ forces $q(t) = 0$ in the same time interval and we have $\vec{H}^1(1, 0) \neq 0$, $\vec{H}^2(1, 0) \neq 0$. $\square$

For ϕ^C , the existence of a minimizing trajectory follows by applying the direct method of the Calculus of Variations. By the previous Proposition, a minimizing extremal for the problem with $L = \phi^C - V$ is also an extremal for the problem with $L = \phi - V$. Since the minimum of the problem with $L = \phi^C - V$ is less than or equal to the infimum for the problem $L = \phi - V$, we obtain that this infimum is indeed a minimum coinciding with the minimum of the problem with $L = \phi^C - V$.

6.2.3 Proof of Theorem 6.2.1

The remaining part of this Section is devoted to the proof of the results stated in Subsection 6.2.1.

We apply PMP to the OCP defined by Equations (6.2.1), (6.2.2). That is, letting $\lambda = (p, q) \in T^*\mathbb{R} \simeq \mathbb{R}^2$, we define the Hamiltonian function $h_u(p, q) = pu - L(q, u)$ as in (6.2.3) and we want to determine the value of u maximizing $h_u(p, q)$ for fixed p and q . By the super-linearity of the function L , for fixed (p, q) , we have

$$h_u(p, q) = |u| \left(\left\langle p, \frac{u}{|u|} \right\rangle - \frac{L(q, u)}{|u|} \right) \leq |u| \left(|p| - \frac{L(q, u)}{|u|} \right) \xrightarrow{|u| \rightarrow +\infty} -\infty.$$

Thus, the maximum point of the Hamiltonian $h_u(p, q)$ with respect to u is a critical point:

$$\partial_u h_u(p, q) = 0, \text{ equivalently } p = \partial_u L(q, u).$$

Geometrically, if u_0 is a critical point of $u \mapsto h_u(p, q)$, then the affine function

$$y_{p, u_0}(u) := p(u - u_0) + L(q, u_0)$$

is tangent to the graph of $L(q, \cdot)$ at u_0 . Thus, for fixed $p, q \in \mathbb{R}$, to find the value of the control maximizing $h_u(p, q)$ it suffices to find u_0 such that the line y_{p, u_0} is tangent to $L(q, \cdot)$ at the point u_0 .

Step 1 - The two maximizing branches. By assumptions 3 and 4, for fixed (p, q) the equation $p = \partial_u L(q, u)$ has either one or three solutions, apart from the limiting case in which one of the solutions is double. When three distinct solutions are present, the two outer ones correspond to strict local maxima of $u \mapsto h_u(p, q)$, whereas the middle one corresponds to a local minimum.

Therefore, by the implicit function theorem, there exist maximal open sets $\Omega_1, \Omega_2 \subset \mathbb{R}^2$ and smooth functions

$$u_i : \Omega_i \longrightarrow \mathbb{R}, \quad i = 1, 2,$$

such that $u_1(p, q) < u_2(p, q)$ whenever both are defined, and u_1, u_2 are precisely the two local maximizing controls.

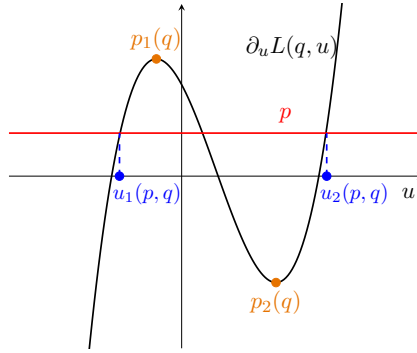


Figure 6.4: The graph of $\partial_u L(q, \cdot)$ and the value of the slope of the bi-tangent line.

The domains of these two branches have a simple form. Indeed, there exist smooth functions $p_1, p_2 : \mathbb{R} \rightarrow \mathbb{R}$, with $p_2(q) < p_1(q)$, such that

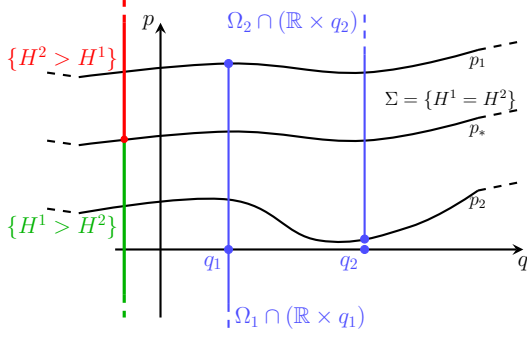
$$\Omega_1 \cap (\mathbb{R} \times \{q\}) = (-\infty, p_1(q)) \times \{q\}, \text{ and } \Omega_2 \cap (\mathbb{R} \times \{q\}) = (p_2(q), +\infty) \times \{q\}.$$

To prove this, fix q and consider the values $p_1(q)$ and $p_2(q)$ corresponding respectively to the local maximum and the local minimum of the function $u \mapsto \partial_u L(q, u)$, see Figure 6.4. Their smooth dependence on q follows from the implicit function theorem and assumption 3. Define

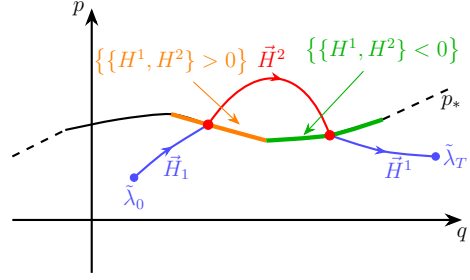
$$H^i(\lambda) := h_{u_i(\lambda)}(\lambda), \quad \lambda \in \Omega_i, \quad i = 1, 2,$$

and take any smooth extension of H^i for $\lambda \notin \Omega_i$. For $p < p_2(q)$, the only local maximum is $u_1(p, q)$, and it is necessarily the global maximum. Similarly, for $p > p_1(q)$, the only local maximum is $u_2(p, q)$, and it is the global maximum. Then, we have $H(\lambda) = \max\{H^1(\lambda), H^2(\lambda)\}$. Define $\Sigma = \{\lambda \in T^*\mathbb{R} \mid H^1(\lambda) = H^2(\lambda)\}$.

Step 2 - The set Σ is a graph. We want to prove the following: for every $q \in \mathbb{R}$, there is a unique value $p \in (p_2(q), p_1(q))$ such that $H^1(p, q) = H^2(p, q)$, which depends smoothly on q .



(a) Graphical illustration of Step 1.



(b) Structure of the extremal trajectories in the phase space.

Writing explicitly $H^1(p, q) = H^2(p, q)$ and taking into account that, by our assumption, we have $u_1 < u_2$, this equality is equivalent to

$$p = \frac{L(q, u_2(p, q)) - L(q, u_1(p, q))}{u_2(p, q) - u_1(p, q)},$$

that is, the two straight affine lines

$$u \mapsto p(u - u_1(p, q)) + L(q, u_1(p, q)) \quad \text{and} \quad u \mapsto p(u - u_2(p, q)) + L(q, u_2(p, q)),$$

coincide. Notice that this is equivalent to saying that the line tangent to the graph of $L(q, \cdot)$ at $u_1(p, q)$ coincides with the tangent at $u_2(p, q)$. By the double-well shape of the graph of $L(q, \cdot)$, this can happen only for a single value of p , which depends smoothly on q . This proves that Σ is a graph.

Step 3 - The maximizing Hamiltonian. As a corollary of the previous step, we determine precisely the maximized Hamiltonian. For fixed $q \in \mathbb{R}$ and denoting by $p_*(q)$ the unique $p \in \mathbb{R}$ such that $H^1(p, q) = H^2(p, q)$, by continuity of the functions H^1, H^2 we have that

$$H(p, q) = \begin{cases} H^1(p, q) & \text{if } p \leq p_*(q), \\ H^2(p, q) & \text{if } p \geq p_*(q). \end{cases}$$

This gives the PMP description of the extremals. Away from Σ , an extremal satisfies

$$\dot{\lambda} = \begin{cases} \vec{H}^1(\lambda) & \text{in } \{H^1 > H^2\}, \\ \vec{H}^2(\lambda) & \text{in } \{H^2 > H^1\}. \end{cases} \quad (6.2.9)$$

At points of Σ , both controls are maximizing. In particular, every extremal trajectory is contained in a level set of the maximized Hamiltonian H . More precisely, H is constant along each regular arc and its value is preserved when the extremal switches from one branch to the other, that is, system (6.2.9) is integrable.

Step 4 - Transversality at non-degenerate switching points. Let $\bar{\lambda} \in \Sigma$. Along an arc generated by either $\vec{H}^1$ or $\vec{H}^2$, one has

$$\frac{d}{dt}(H_2(\lambda_t) - H_1(\lambda_t)) = \{H_1, H_2\}(\lambda_t). \quad (6.2.10)$$

Consequently, if $\{H_1, H_2\}(\bar{\lambda}) \neq 0$, then both Hamiltonian vector fields $\vec{H}_1$ and $\vec{H}_2$ are transverse to Σ at $\bar{\lambda}$. Thus an extremal reaching $\bar{\lambda}$ crosses the switching set transversally, and the corresponding switching time is isolated.

Furthermore, the maximizing branch changes when $H_2 - H_1$ changes sign. Therefore, at every switching point outside

$$\mathcal{Z} := \{\lambda \in \Sigma : \{H_1, H_2\}(\lambda) = 0\},$$

the strong switching Legendre condition is satisfied.

Step 5 - Finiteness of switching times and piecewise regularity. We can now prove point 1. in Theorem 6.2.1. First, we prove that switching times cannot accumulate when $\mathcal{Z}$ is discrete. Assume by contradiction that a non-constant extremal trajectory has infinitely many switching times t_j converging to some t_∞ . Since λ is continuous, the switching points $\lambda(t_j)$ converge to $\lambda(t_\infty) \in \Sigma$. Moreover, consecutive switching arcs alternate between the regions $\{H_1 > H_2\}$ and $\{H_2 > H_1\}$. By (6.2.10), the corresponding values of $\{H_1, H_2\}$ have alternating signs. Hence,

$$\{H_1, H_2\}(\lambda(t_\infty)) = 0,$$

and therefore

$$\lambda(t_\infty) \in \mathcal{Z}.$$

Since $\mathcal{Z}$ is discrete, the only possible accumulation configuration would be a piecewise smooth spiral around the point $\lambda(t_\infty)$. This is impossible because the dynamics admit the first integral H : every extremal remains in a single level set of H , whereas a spiral converging to an isolated attracting equilibrium would force H to be locally constant near that equilibrium. Thus switching times cannot accumulate. It follows that every non-constant extremal has only finitely many switching times.

Moreover, on every open interval between two consecutive switching times, the maximizing control is unique and satisfies

$$\partial_{uu}^2 h_u(p, q) = -\partial_{uu}^2 L(q, u) < 0.$$

At a switching point outside $\mathcal{Z}$, the strong switching Legendre condition follows from the transversality argument above. Hence an extremal can fail to be piecewise regular only if it meets $\mathcal{Z}$.

Fix $q_0 \in \mathbb{R}$. Since $\mathcal{Z}$ is discrete and the system is integrable, the set of initial covectors $p_0 \in \mathbb{R}$ for which the extremal with initial condition

$$\lambda(0) = (p_0, q_0)$$

meets $\mathcal{Z}$ is discrete. Therefore, outside a discrete set of initial covectors, the corresponding extremal trajectory is piecewise regular.

It remains to prove the existence assertion in Theorem 4.1.

Step 6 - Existence of minimizers. Let $L^C(q, \cdot)$ denote the convexification of $L(q, \cdot)$ with respect to the control variable. The construction above shows that the maximized Hamiltonian associated with L^C coincides with H . Indeed, the common tangent line of slope $p^*(q)$ determines the affine portion introduced by the convexification, and the maximizers remain on the two outer branches where L^C agrees with L .

The direct method in the Calculus of Variations yields the existence of a minimizing control u^C for the problem associated with L^C .

If u^C is such that $L^C(q(t), u^C(t)) = L(q(t), u^C(t))$ for every $t \in [0, T]$, then, since $L^C \leq L$, u^C is a minimizer also for the original functional L , and so point 2. of Theorem 6.2.1 follows.

Suppose, by contradiction, that $L^C(q(t), u^C(t)) < L(q(t), u^C(t))$. Then, denoting by λ^C the extremal for L^C corresponding to the control u^C , we must have that there is some non-trivial time interval for which $\lambda_t^C \in \Sigma$, since outside this set the Hamiltonian systems for the extremals of L^C and L are the same. This implies that $\dot{\lambda}_t = \alpha(t)\vec{H}^1(\lambda_t) + (1 - \alpha(t))\vec{H}^2(\lambda_t)$, for some function α satisfying $\alpha(t) \in [0, 1]$. But then the same computation as the one in Equation (6.2.10) forces $\{H^1, H^2\}(\lambda_t^C) = 0$ on this time interval, i.e. $\lambda_t \in \mathcal{Z}$. So, since $\mathcal{Z}$ is discrete, λ_t is constant. But this is not possible since, by assumption $\alpha\vec{H}^1(\lambda) + (1 - \alpha)\vec{H}^2(\lambda) \neq 0$ for all $\alpha \in [0, 1]$ and all $\lambda \in \mathcal{Z}$.

This completes the proof of Theorem 6.2.1. $\square$

Remark 6.2.5. Notice that if $\{H^1, H^2\}(\bar{\lambda}) = 0$ at a switching point, i.e. $\bar{\lambda} \in \mathcal{Z}$, then we have three cases:

- either $\vec{H}^1(\bar{\lambda}) = \vec{H}^2(\bar{\lambda}) = 0$, in which case $\bar{\lambda}$ is an equilibrium point and the only possible extremal trajectory passing through $\bar{\lambda}$ is the constant trajectory;
- either one among $\vec{H}^1(\bar{\lambda})$ or $\vec{H}^2(\bar{\lambda})$ is zero and the other one is non-zero, so that there is still one unique way to extend any extremal trajectory passing through the point $\bar{\lambda}$;
- or both $\vec{H}^1(\bar{\lambda}) \neq 0$ and $\vec{H}^2(\bar{\lambda}) \neq 0$: in this case the extremal trajectory passing through the point $\bar{\lambda}$ is not uniquely determined after this point, because one can choose both the vectors $\vec{H}^1(\bar{\lambda})$ and $\vec{H}^2(\bar{\lambda})$ to extend the trajectory, that is $\bar{\lambda}$ is a branching point;

Even though the case $\{H^1, H^2\}(\bar{\lambda}) = 0$ could be studied with our approach, a detailed analysis of these degenerate cases goes beyond the purposes of this thesis.

6.2.4 Proof of Proposition 6.2.2

In this Subsection, we prove Proposition 6.2.2. Fix a reference piecewise regular normal extremal pair $(\tilde{u}, \tilde{\lambda})$.

We point out that in the following computations we only need that the state manifold M is one-dimensional and we do not require the structure introduced in the problem (6.2.1),(6.2.2). So, the results contained in this subsection apply to any one-dimensional control system of the form (6.1.4) or (7.2.2).

This characterization in the free final time case could be obtained as well by an elementary argument. Nonetheless, in our opinion it is a good illustration of our theory.

Proof of Proposition 6.2.2. As already pointed out in Section 6.2.1, the cotangent space is trivial, $T^*M \simeq \mathbb{R}^* \times M$, so we have a global chart and we do not have to transport all the dynamics back to the initial point. Consequently, the tangent space $T_{\tilde{\lambda}_0}(T^*M)$ is also identified with $\mathbb{R}^* \times \mathbb{R}$. The Lagrangian Grassmannian $\mathfrak{L}(T_{\tilde{\lambda}_0}(T^*M)) \simeq \mathfrak{L}(\mathbb{R}^* \times \mathbb{R})$ identifies with one-dimensional linear subspaces of $\mathbb{R}^* \times \mathbb{R}$, i.e. $\mathbb{RP}^1 \simeq S^1$. As a convenient set of coordinates, we can choose to identify $T_{\tilde{\lambda}_0}(T_{q_0}^*M) \simeq [0, 1]^\top$ and $T_{\tilde{\lambda}_0}(T_{q_0}M) \simeq [1, 0]^\top$. Here we denote with the squared brackets $[\cdot, \cdot]$ points on the projective space $\mathbb{RP}^1$ and with the round brackets $(\cdot, \cdot)$ points in the space $\mathbb{R}^* \times \mathbb{R}$. In these coordinates, the canonical symplectic form is given by

$$\sigma((0, 1)^\top, (1, 0)^\top) = 1,$$

and vectors rotating in the clockwise direction are given a positive orientation: to show this, we parametrize points in S^1 in a slightly unusual way, that is we set

$$S^1 = \{(\sin \theta, \cos \theta)^\top \mid \theta \in [0, 2\pi)\},$$

so that the point corresponding to $\theta = 0$ is the vertical subspace $[0, 1]^\top$ and as θ increases the point $(\sin \theta, \cos \theta)$ moves clockwise. We have to show that vectors moving clockwise are given a positive orientation. Take any function $\theta : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ such that $\dot{\theta} > 0$, we have

$$\begin{aligned} & \sigma \left((\sin \theta_0, \cos \theta_0)^\top, \left. \frac{d}{dt} \right|_{t=0} (\sin \theta(t), \cos \theta(t))^\top \right) = \\ & = \dot{\theta}(0) \sigma \left((\sin \theta_0, \cos \theta_0)^\top, (\cos \theta_0, -\sin \theta_0)^\top \right) = \dot{\theta}(0) > 0, \end{aligned}$$

as we wanted to show.

The Jacobi curve is the curve $\Lambda : [0, T] \rightarrow \mathbb{RP}^1$ given by $\Lambda_0 = [0, 1]^\top$ and $\Lambda_t = \mathbb{R}\vec{h}_{\tilde{u}(t)}$ for $t > 0$. Indeed, since $\dim M = 1$, the condition $\vec{h}_{\tilde{u}(t)} \in \Lambda(t)$ completely determines the Jacobi curve.

In particular, if t is a switching time and we denote by $\vec{h}_-$ and $\vec{h}_+$ the Hamiltonian vector fields generating the maximized flow before and after t respectively and supposing $\vec{h}_- = r_-(\sin \theta_-, \cos \theta_-)^\top$ and $\vec{h}_+ = r_+(\sin \theta_+, \cos \theta_+)^\top$, with $r_\pm > 0$ and $\theta_\pm \in [0, 2\pi)$, we have that

$$\sigma(\vec{h}_-, \vec{h}_+) = r_- r_+ \sin(\theta_+ - \theta_-).$$

Hence, we have $\sigma(\vec{h}_-, \vec{h}_+) > 0$ if and only if $\sin(\theta_+ - \theta_-) > 0$, that is $\theta_+ \in (\theta_-, \theta_- + \pi) \bmod 2\pi$.

Now, we want to characterize conjugate times. Suppose first that the curve Λ is continuous at $t \in [0, T]$. Recall that, in this case, t is a conjugate time if and only if Equation (4.2.1) holds. If $\vec{h}_{\tilde{u}(t)}$ is not constant in time, then the left-hand side is $\{0\}$ and, since both Λ and Π have dimension 1, this condition reduces simply to $\mathbb{R}\vec{h}_{\tilde{u}(t)} = [0, 1]^\top$. This happens if and only if $\pi_* \vec{h}_{\tilde{u}(t)} = 0$, that is $\dot{\tilde{q}}(t) = 0$.

If instead t is a switching time and using the same notation introduced above, we have that $\pi_*(\Lambda_t + \Lambda_t^+) = \mathbb{R}$ and $\pi_*(\Lambda_t \cap \Lambda_t^+) = \{0\}$. Thus, t is a conjugate time if and only if $\pi_* \vec{h}_-$ and $\pi_* \vec{h}_+$ belong to different connected components of $\mathbb{R} \setminus \{0\}$, that is they have opposite sign.

This concludes the proof. $\square$

6.2.5 Proof of Proposition 6.2.3

In this Subsection we prove the characterization of conjugate points in the case of fixed final time, see Proposition 6.2.3. As explained in Section 7.2, we have to consider the extended state space $\mathbb{R} \times M$, where the first $\mathbb{R}$ corresponds to the added time variable. The purpose of the following calculation is only to track the second coordinate x_2 of the Jacobi curve (see notation introduced in Proposition 6.2.3): away from switching times, $x_2 = 0$ detects conjugacy, while at a switching time the relevant information is whether x_2 changes sign.

Thus, the cotangent space is $T^*(\mathbb{R} \times M) \simeq (\mathbb{R}^* \times \mathbb{R}) \times T^*M$. First of all, we have to write the Hamiltonian system with the time as additional variable:

$$\begin{cases} \dot{t}(s) = 1 + \nu(s), \\ \dot{q}(s) = (1 + \nu(s))u(s). \end{cases} \quad \begin{cases} \dot{\tau}(s) = 0, \\ \dot{p}(s) = (1 + \nu(s))\partial_q L(q(s), u(s)). \end{cases}$$

We immediately notice that the right-hand sides of the equations for t and τ depend on neither p, q nor u . Thus, in the linearized system, i.e. the Jacobi equation, the corresponding linearized variables are constant for $s \in (t_{j-1}, t_j)$.

Proof of Proposition 6.2.3. Again, for simplicity of exposition, we deal with the simpler case of one single switching time, but repeating the same argument one obtains the same result for any finite number of switching times.

We recall briefly the structure of the Jacobi curve: we have that $\Lambda_0 = \partial_\tau \oplus \Pi$. Then, the discontinuity at 0 is computed as $\Lambda_0^+ = \Lambda_0^{\partial_t + \vec{H}^1(\tilde{\lambda}_0)}$. If $\vec{H}^1(\tilde{\lambda}_0)$ is vertical, that is $x_1(0) = 0$, then $\Lambda_0^+ = \Lambda_0$. Otherwise, denoting by $\vec{H}^1(s) = \vec{H}^1(\tilde{\lambda}_s) = [0, 0, y_1(s), x_1(s)]^\top$, we have

$$\Lambda_0^+ = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ y_1(0) & -1/x_1(0) \\ x_1(0) & 0 \end{bmatrix}.$$

Then, for $s \in (0, t_1]$, we have that the τ, t components are constant and the p, q components evolve according to the linearized flow $(\Phi_s^1)_*$ (which was defined in (6.1.6)). Thus, we have

$$\Lambda_s = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ y_1(s) & y_2(s) \\ x_1(s) & x_2(s) \end{bmatrix} = \begin{bmatrix} \partial_t + \vec{H}^1(s) & \partial_\tau + \eta(s) \end{bmatrix}, \quad s \in (0, t_1].$$

where $\eta(s) = (\Phi_s^1)_*[-\frac{1}{x_1(0)}, 0]^\top$. Then, using a similar notation, we define $\Lambda_{t_1}^+ = \Lambda_{t_1}^{\partial_t + \vec{H}^2(t_1)}$ and $\Lambda_s = (\Phi_s^2)_* \Lambda_{t_1}^+$ for $s \in (t_1, t_2]$, and then repeat the same algorithm for all the subsequent regular subintervals.

If the vector field $\vec{H}^1(s)$ is non-constant in s , from $\partial_t + \vec{H}^1(s) \in \Lambda_s$ it follows that there are no constant vertical vectors contained in Λ_s .

First, suppose that the Jacobi curve is continuous at $s \in [0, T]$. We can easily see that the vector $\partial_t + \vec{h}_{\tilde{u}(s)}$ is never contained in $\mathbb{R}\partial_\tau \oplus \Pi$, since its projection on the ∂_t component is always non-zero. Moreover, every linear combination of $\partial_t + \vec{h}_{\tilde{u}(t)}$ and $\partial_\tau + \eta(t)$ with non-zero coefficients cannot be in $\mathbb{R}\partial_\tau \oplus \Pi$ as well. Then, following our notation, the only possibility for s to be a conjugate time is if $x_2(s) = 0$.

If the Jacobi curve is not continuous at s , that is $s = t_1$, then we can compute the Maslov index $\text{ind}_\Pi(\Lambda_{t_1}, \Lambda_{t_1}^+)$ of the curve (see Section 3.4).

The following computation shows that the Maslov index is non-zero at a switching time if and only if the fourth component x_2 changes sign. We know that $\Lambda_{t_1}^+ = \Lambda_{t_1}^{\partial_t + \vec{H}^2(t_1)}$. Denoting by $\vec{H}^2 = [0, 0, w, z]^\top$, we have to find $\eta^+ = \alpha(\partial_t + \vec{H}^1(t_1)) + \partial_\tau + \eta(t_1) \in \Lambda_{t_1}$ such that $\sigma(\eta^+, \partial_t + \vec{H}^2(t_1)) = 0$. Writing explicitly, we obtain (we omit the dependence on t_1 , since it is fixed):

$$\sigma(\eta^+, \partial_t + \vec{H}^2) = \sigma(\alpha(\partial_t + \vec{H}^1) + \partial_\tau + \eta, \partial_t + \vec{H}^2) = \alpha\sigma(\vec{H}^1, \vec{H}^2) + 1 + \sigma(\eta, \vec{H}^2) = 0.$$

Solving for α yields

$$\alpha = \frac{\sigma(\vec{H}^2, \eta) - 1}{\sigma(\vec{H}^1, \vec{H}^2)}. \quad (6.2.11)$$

Notice that α can be computed using only the projection of Λ_{t_1} in $T_{\lambda_{t_1}}(T^*M)$. So, the right limit of the Jacobi curve is

$$\Lambda_{t_1}^+ = \begin{bmatrix} \partial_t + \vec{H}^2(t_1) & \eta^+(t_1) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & \alpha \\ w(t_1) & \alpha y_1(t_1) + y_2(t_1) \\ z(t_1) & \alpha u_1(t_1) + x_2(t_1) \end{bmatrix}.$$

We can simplify the expression of the second column by subtracting $\alpha(\partial_t + \vec{H}^2(t_1))$:

$$\begin{aligned} \Lambda_{t_1}^+ &= \begin{bmatrix} \partial_t + \vec{H}^2(t_1) & \eta^+(t_1) - \alpha(\partial_t + \vec{H}^2(t_1)) \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ w(t_1) & \alpha(y_1(t_1) - w(t_1)) + y_2(t_1) \\ z(t_1) & \alpha(x_1(t_1) - z(t_1)) + x_2(t_1) \end{bmatrix}. \end{aligned}$$

We compute now the Maslov index. Recall that $\text{ind}_\Pi(\Lambda_{t_1}, \Lambda_{t_1}^+) = \text{ind}^+ \mathfrak{q} + \frac{1}{2} \ker \mathfrak{q}$, where $\mathfrak{q} : ((\Lambda_{t_1} + \Lambda_{t_1}^+) \cap \Pi) / (\Lambda_{t_1} \cap \Lambda_{t_1}^+ \cap \Pi) \rightarrow \mathbb{R}$ is defined by $\mathfrak{q}(\nu) = \sigma(v_1, v_2)$, with $v_1 + v_2 = \nu$. We have that $\mathfrak{q}$ is a quadratic form defined on the one-dimensional space $(\Lambda_{t_1} + \Lambda_{t_1}^+) \cap \Pi$. The space $\Lambda_{t_1} + \Lambda_{t_1}^+$ is given by

$$\Lambda_{t_1} + \Lambda_{t_1}^+ = \begin{bmatrix} \partial_t + \vec{H}^1(t_1) & \partial_t + \vec{H}^2(t_1) & \partial_\tau + \eta(t_1) \end{bmatrix}.$$

We have to find $\nu \in \mathbb{R}\partial_\tau \oplus \Pi$ such that

$$\nu = \beta_1 \left(\partial_t + \vec{H}^1(t_1) \right) + \beta_2 \left(\partial_t + \vec{H}^2(t_1) \right) + \beta_3 (\partial_\tau + \eta(t_1)),$$

for $(\beta_1, \beta_2, \beta_3) \neq (0, 0, 0)$. Clearly, in order to eliminate the component in the direction ∂_t , we have to choose

$$\beta_1 = -\beta_2.$$

Since at a genuine switch we have $x_1(t_1) - z(t_1) \neq 0$, it follows that we must choose $\beta_3 \neq 0$. In particular, up to rescale ν , we can always suppose $\beta_3 = 1$. Hence

$$\nu = \beta_1 \left(\vec{H}^1(t_1) - \vec{H}^2(t_1) \right) + \partial_\tau + \eta(t_1).$$

The condition $\nu \in \partial_\tau \oplus \Pi$ is equivalent to $\beta_1 \left(\vec{H}^1(t_1) - \vec{H}^2(t_1) \right) + \eta(t_1) \in \Pi = \text{span}\{(0, 1)^\top\}$. This reads,

$$\beta_1(x_1(t_1) - z(t_1)) + x_2(t_1) = 0 \iff \beta_1 = \frac{x_2(t_1)}{z(t_1) - x_1(t_1)}.$$

Finally, we can compute $\mathfrak{q}(\nu)$

$$\begin{aligned}\mathfrak{q}(\nu) &= \sigma(\beta_1(\partial_t + \vec{H}^1) + \partial_\tau + \eta, -\beta_1(\partial_t + \vec{H}^2)) = \\ &= -\beta_1^2 \sigma(\vec{H}^1, \vec{H}^2) - \beta_1 - \beta_1 \sigma(\eta, \vec{H}^2) = \\ &= -\beta_1^2 \sigma(\vec{H}^1, \vec{H}^2) + \beta_1(\sigma(\vec{H}^2, \eta) - 1).\end{aligned}$$

Finally, the time t_1 is conjugate to 0 if and only if $\text{ind}_\Pi(\Lambda_{t_1}, \Lambda_{t_1}^+) > 0$. So, t_1 is not conjugate to 0 if $\mathfrak{q}(\nu) < 0$, which is equivalent to

$$\beta_1^2 - \beta_1 \frac{(\sigma(\vec{H}^2, \eta) - 1)}{\sigma(\vec{H}^1, \vec{H}^2)} = \beta_1^2 - \beta_1 \alpha = \beta_1(\beta_1 - \alpha) > 0,$$

where α is the same as in (6.2.11) and we have used the fact that $\sigma(\vec{H}^1, \vec{H}^2) > 0$. We compute the product $x_2(t_1)x_2(t_1+)$ (for the sake of readability, we omit the s and denote $x_2(t_1+)$ by x_{2+}):

$$\begin{aligned}x_2 x_{2+} &= x_2^2 + \alpha x_2(x_1 - z) = \\ &= \beta_1^2(x_1 - z)^2 - \alpha \beta_1(x_1 - z)^2 = \\ &= \beta_1(\beta_1 - \alpha)(x_1 - z)^2.\end{aligned}$$

Thus, t_1 is not a conjugate time if and only if $x_2(t_1)$ and $x_2(t_1+)$ have the same sign, as we wanted to show. $\square$

Chapter 7

Piecewise regular extremals - part 2: proofs of main results

In this chapter, we apply the general principles developed in Chapters 3 and 4 to piecewise-regular extremals. Our main purpose is to prove the second-order necessary and sufficient optimality conditions stated in Chapter 6.

The chapter is divided into two parts. The first part is devoted to the necessary optimality condition, whereas the second concerns the sufficient optimality condition.

We begin by introducing a Banach manifold structure on a suitable class of admissible controls that are close to a fixed reference control in an appropriate topology. This construction provides the functional-analytic framework needed to define the extended endpoint map, both for problems with free final time and for problems with fixed final time, which are considered in Sections 7.1 and 7.2, respectively. We then focus on the free-final-time problem. In Section 7.3, we compute the first and second differentials of the extended endpoint map. The fixed-final-time case can be treated by analogous arguments and is therefore not developed in detail. Building on these computations, in Section 7.4 we derive the characterization of the Jacobi curve stated in Theorem 6.1.6. The necessary optimality condition of Theorem 6.1.7 then follows by applying the general criterion given in Theorem 3.0.1.

The second part is devoted to sufficient optimality. In Section 7.5, we study the regularity of the Hamiltonian flow generated by the maximized Hamiltonian. This regularity allows us to construct a field of extremals satisfying the hypotheses introduced in Chapter 4. Combining this construction with the absence of conjugate points, in Section 7.6 we prove the sufficient optimality condition stated in Theorem 6.1.8.

7.1 The endpoint map

The aim of this Section is to introduce the endpoint map associated with the problem (6.1.4). To this aim, we build a structure of Banach manifold on control competitors of a given reference control. Relying on this Banach manifold structure, in Section 7.3.2 we compute the derivatives of the endpoint map.

The purpose of the next construction is to obtain a local parametrization of a class of admissible control competitors whose corresponding extremal is in a neighbourhood of the reference one. Two kinds of variations have to be considered. First, the control may vary inside the same smooth submanifold of U . Second, the switching times themselves must be allowed to change. The first type of variation is described by local coordinates on the sets U_j , whereas the second is encoded by a time reparametrization. Combining these two ingredients yields a Banach manifold of controls on which the endpoint map becomes smooth.

Now, we make this construction rigorous. We take (time-dependent) local coordinates on each U_j in a neighbourhood of $\tilde{u}(t)$, depending on $\tilde{\lambda}_t$, such that $\tilde{u}(t) = 0$. This way, we can identify an open neighbourhood of $\tilde{u}(t)$ in U_j with an open neighbourhood $V_j \subset \mathbb{R}^{\dim U_j}$ of $0 \in \mathbb{R}^{\dim U_j}$ for every $j = 1, \dots, k+1$. Hence, any other control $u \in L^\infty((t_{j-1}, t_j); U_j)$ close enough to $\tilde{u}|_{(t_{j-1}, t_j)}$ can be identified with a control function with values V_j for $t \in (t_{j-1}, t_j)$. Using these coordinates, we can rewrite the Hamiltonian system (2.4.1) as

$$\dot{\tilde{\lambda}} = \vec{h}_0^j(\tilde{\lambda}), \quad t \in (t_{j-1}, t_j), \quad j = 1, \dots, k+1.$$

We now introduce the time-reparametrization variable, whose purpose is to encode variations of the switching times. Let $\nu \in L^\infty(\mathbb{R})$ be a function such that $\nu(t) > -1$ and we denote by $\nu_j := \nu|_{(t_{j-1}, t_j)}$. The control system with time reparametrization ν is

$$\dot{\lambda} = (1 + \nu_j(t)) \vec{h}_0^j(\lambda), \quad t \in (t_{j-1}, t_j), \quad j = 1, \dots, k+1.$$

The Hamiltonian lift of system (6.1.1) with time reparametrization ν is

$$\dot{\lambda} = (1 + \nu_j(t)) \vec{h}_{u(t)}^j(\lambda), \quad u(t) \in V_j, \quad t \in (t_{j-1}, t_j), \quad j = 1, \dots, k+1, \quad (7.1.1)$$

where here we are using the identification of a neighbourhood of $\tilde{u}(t)$ in U_j with V_j . The cost functional with time reparametrizations is given by

$$J(\nu, u) = \int_0^T (1 + \nu(t)) L(q(t; \nu, u), u(t)) dt, \quad (\nu, u) \in \mathcal{U},$$

where we are omitting the dependence on the initial point q_0 since it is fixed.

We define the space of admissible controls with time reparametrizations on the interval $[0, T]$ as

$$\mathcal{U} := \{(\nu, u) \mid \nu \in L^\infty([0, T]; (-1, +\infty)), u \in L^\infty([0, T]; V_j) \text{ for } t \in (t_{j-1}, t_j)\}.$$

Moreover, for $t \in [0, T]$, we denote by $\mathcal{U}_t := \{(\nu, u) \in \mathcal{U} \mid \text{supp}(\nu, u) \subset [0, t]\}$. Notice that, $\mathcal{U}$ has a structure of Banach manifold modelled on the Banach space

$$L^\infty([0, T]; \mathbb{R}) \times \left(L^\infty([0, t_1]; \mathbb{R}^{\dim U_1}) \times \cdots \times L^\infty([t_k, T]; \mathbb{R}^{\dim U_{k+1}}) \right).$$

As this choice of control variations plays a central role in all the rest of the Chapter, it is worth spending a few more words to clarify the structure of this space:

- if $\nu(t) = 0$ for $t \in [0, T]$, then the control u in the right-hand side of (7.1.1) corresponds to a control whose value $u(t)$ is in the same submanifold U_j where also $\tilde{u}(t)$ is. In particular, if $\nu(t) = 0$ for $t \in [0, T]$, u is close to $\tilde{u}$ in L^∞ ;
- if $\nu(t) \neq 0$, then the control (ν, u) in (7.1.1) corresponds to the control $v \in \mathfrak{U}_{T+\int_0^T \nu(s)ds}$ admissible for the system (6.1.1) defined by $v(t + \int_0^t \nu(s)ds) = u(t)$. In particular, $v(t)$ and $\tilde{u}(t)$ can be in different submanifolds of U . This fact allows us to change the switching times and the resulting control v is close to $\tilde{u}$ in L^1 but not in L^∞ .

Introducing time reparametrizations is crucial for several reasons: first of all, the manifolds U_j may be reduced to a single point (see Example 6.1.1) and in this case there is no variation of the control which is close to the reference one in the L^∞ norm. Second, we can keep $T > 0$ fixed, while still considering trajectories with different ending time. Furthermore, time reparametrizations allow us to consider controls with values in a submanifold different from U_j , but having at the same time a nice parametrization.

Indeed, the space $\mathcal{U}$ is a much smaller space than $\mathfrak{U}$. Of course, to obtain necessary conditions for optimality we can restrict to consider control competitors in a smaller set. The remarkable fact is that, as we are going to show, this class of variations is rich enough to obtain sufficient conditions for local optimality as well.

We define the endpoint map on the cotangent bundle to be the map $\mathcal{E} : \mathcal{U} \rightarrow T^*M$ defined as

$$\mathcal{E}(\nu, u) = \overrightarrow{\exp} \int_0^T (1 + \nu(t)) \vec{h}_{u(t)} dt \circ \tilde{\lambda}_0,$$

and we denote by $\lambda(t; \tilde{\lambda}_0, \nu, u)$ the solution of (7.1.1), with the convention that if $(\nu, u) \in \mathcal{U}_t$ we stop the evolution of the system at time t . Finally, we introduce the endpoint map on the base manifold M to be the map $E : \mathcal{U} \rightarrow M$, $E(\nu, u) = \pi(\mathcal{E}(\nu, u))$:

$$E(\nu, u) = \pi \circ \overrightarrow{\exp} \int_0^T (1 + \nu(t)) \vec{h}_{u(t)} dt \circ \tilde{\lambda}_0.$$

7.2 Optimal control problems with fixed final time

In this section, we show how the problem with fixed final time can be rephrased as a problem with free final time. In particular, we state the Theorem about the structure of the Jacobi curve for the fixed final time problem, which is not given in the previous chapter.

The idea of reducing the fixed final time problem to the free final time one is classical. Instead of prescribing the terminal time, we introduce the physical time as an additional state variable. The resulting problem has free final time with respect to the new time parameter s , while the original fixed-final-time constraint is recovered by imposing the terminal condition $t(S) = T$. This reformulation allows us to apply directly the free-final-time theory.

Let U be as in Section 6.1.1. We consider the control system on the manifold M

$$\dot{q}(t) = f(q(t), u(t)), \quad q(0) = q_0, q(T) = q_1, \quad t \in [0, T], \quad (7.2.1)$$

where f, u, q_0, q_1 are defined as before, and $T > 0$ is fixed. Similarly, we take the cost functional

$$J(u) = \int_0^T L(q(t), u(t)) dt.$$

We want to study the problem

$$\min\{J(u) \mid q(\cdot; u) \text{ satisfies (7.2.1)}\}. \quad (7.2.2)$$

Again, we can apply PMP, obtaining the Hamiltonian system

$$\dot{\lambda} = \vec{h}_{u(t)}(\lambda),$$

where $h_u(\lambda) = \langle \lambda, f(q, u) \rangle - L(q, u)$, for $\lambda \in T^*M$, $u \in U$, $q = \pi(\lambda)$. We want to define an endpoint map similar to the one introduced in Section 7.1. To this aim, we introduce an extended control system on $M \times \mathbb{R}$, where we add the time t as a new state-variable. In addition to the control u , we consider also time reparametrizations as admissible trajectories. More precisely, on the extended state manifold $\widehat{M} = M \times \mathbb{R}$ we consider

$$\begin{cases} \dot{q}(s) = (1 + \nu(s))f(q(s), u(s)), \\ \dot{t} = 1 + \nu(s), \end{cases}$$

$$(q(0), t(0)) = (q_0, 0), (q(S), t(S)) = (q_T, T), \quad s \in [0, S].$$

Denoting by $\tau = \langle dt, v \rangle$, for $v \in T_{(q,t)}\widehat{M}$, the new Hamiltonian is $\mathcal{H}_{(\nu, u)}(\tau, \lambda) = (1 + \nu)(h_u(\lambda) + \tau)$. We can introduce a Hamiltonian system in $T^*\widehat{M}$ analogous to the one in

Equation (7.1.1), but where we keep this new time variable:

$$\begin{cases} \dot{\lambda}_s = (1 + \nu(s)) \vec{h}_{u(s)}(\lambda_s), \\ \dot{\tau} = 0, \\ \dot{t} = 1 + \nu(s), \end{cases} \quad (7.2.3)$$

As in the case of free final time, we consider a piecewise regular extremal pair $(\tilde{u}, \tilde{\lambda})$ and we study optimality of the corresponding trajectory $\tilde{q} = \pi \circ \tilde{\lambda}$ on M .

Along the reference extremal we have $\nu \equiv 0$, so that $t(s) = s$ and $S = T$. Therefore the switching times are the same in both parametrizations. The only substantial difference with the free-final-time case is that the Jacobi curve now lives in the tangent space of the extended cotangent bundle $T^*\widehat{M}$, and the initial vertical space and jump transformations must be modified accordingly. Notice that, by PMP, $h_{\tilde{u}(s)}(\tilde{\lambda}_s)$ is constant, hence for the reference extremal with $\nu(s) = 0$ we have that $\mathcal{H}_{\tilde{u}(s),0}(\tilde{\lambda}_s, \tau_s) = 0$, which reads $\tau_s = -h_{\tilde{u}(s)}(\tilde{\lambda}_s)$. As before, we can fix time-dependent coordinates in U such that $\tilde{u}(s) = 0$ for all $s \in [0, T]$ and parametrize the value of controls close to $\tilde{u}(s)$ by means of these coordinates. Therefore, the endpoint map of this system is

$$E(\nu, u) = \widehat{\pi} \circ \overrightarrow{\text{exp}} \int_0^T (1 + \nu(s)) (\partial_t + \vec{h}_{u(s)}) ds \circ (\tilde{\lambda}_0, \tau_0, 0),$$

where here $\widehat{\pi} : T^*\widehat{M} \rightarrow \widehat{M}$ is the canonical bundle projection. In particular, we see that the partial derivative of $\mathcal{H}$ with respect to ν is $\partial_t + \vec{h}_{u(s)}$, while in the case of free final time was simply $\vec{h}_{u(t)}$. Thus, the formula for the jumps at the switching points is simply determined by replacing $\vec{\mathcal{H}}^{j+1}$ with $\partial_t + \vec{\mathcal{H}}^{j+1}$.

Theorem 7.2.1 (Structure of Jacobi curve, fixed final time). *Let $(\tilde{u}, \tilde{\lambda})$ be a piecewise regular extremal pair for the Problem with fixed final time. The Jacobi curve Λ for this problem is a curve of Lagrangian subspaces in $\mathfrak{L}(T_{(\tilde{\lambda}_0, \tau_0, 0)}(T^*\widehat{M}))$ is obtained as follows: we set $\Lambda_0 = \Pi \oplus \mathbb{R}\partial_\tau$, then, for $j = 0, 1, \dots, k$ we set*

1. $\Lambda_{t_j}^+ = \Lambda_{t_j}^{\partial_t + \vec{\mathcal{H}}^{j+1}};$
2. $\Lambda_s = (\phi_s^{j+1})_* \Lambda_{t_j}^+$ for $s \in (t_j, t_{j+1}]$.

Here ϕ_s^j were introduced in Equation 6.1.6. Notice that the right-hand side of two equations for τ and t in (7.2.3) does not depend on the state variables λ, τ, t , hence the corresponding linearized variables are constant of the linearized dynamic. In particular, the linearized flow $(\phi_s)_*$ extends naturally to this case.

7.3 Differential and Hessian of the endpoint map

In this section, we derive a formula for the differential and the Hessian of the endpoint map introduced in 7.1. This formula is then used in Section 7.4 to prove Theorem 7.4.1 about the structure of the Jacobi curve corresponding to a piecewise regular extremal. First, we recall some formula of symplectic geometry which is needed in our computations. Then, we deduce the formula for the Hessian from the general one recalled in (3.2.4) in which the contributions of the time variations is separated by the one coming from control variations.

7.3.1 Useful formulas

In this subsection, we recall two elementary identities that will be used in the computation of the Hessian. Let M, U_1, U_2 be smooth manifolds and let $f^i : M \times U_i \rightarrow TM$ be smooth maps such that $f^i(\cdot, u) \in \text{Vec}(M)$ for every $u \in U_i$, $i = 1, 2$. For $u \in U_i$, define the fiber-linear Hamiltonian

$$h_u^i(\lambda) := \langle \lambda, f^i(\pi(\lambda), u) \rangle, \quad \lambda \in T^*M. \quad (7.3.1)$$

We denote by ϑ the tautological, or Liouville, one-form on T^*M .

Lemma 7.3.1. *Let $u_i : T^*M \rightarrow U_i$, $i = 1, 2$, be smooth maps satisfying*

$$\left. \frac{\partial h_u^i}{\partial u}(\lambda) \right|_{u=u_i(\lambda)} = 0, \quad \lambda \in T^*M.$$

Define $h^i(\lambda) := h_{u_i(\lambda)}^i(\lambda)$, $i = 1, 2$. Then

$$\vartheta_\lambda(\vec{h}^i(\lambda)) = h^i(\lambda), \quad i = 1, 2,$$

and

$$\vartheta_\lambda([\vec{h}^1, \vec{h}^2](\lambda)) = \sigma_\lambda(\vec{h}^1(\lambda), \vec{h}^2(\lambda)).$$

Proof. By definition of the Hamiltonian vector field, for every fixed $u \in U_i$ we have

$$d_\lambda h_u^i = \sigma_\lambda(\cdot, \vec{h}_u^i).$$

The chain rule gives

$$d_\lambda h^i = \left. \frac{\partial h_u^i}{\partial \lambda} \right|_{u=u_i(\lambda)} + \left. \frac{\partial h_u^i}{\partial u} \right|_{u=u_i(\lambda)} \circ d_\lambda u_i.$$

The second term vanishes by (7.3.1), and therefore

$$d_\lambda h^i = \sigma_\lambda(\cdot, \vec{h}_{u_i(\lambda)}^i).$$

Hence

$$\vec{h}^i(\lambda) = \vec{h}_{u_i(\lambda)}^i(\lambda).$$

Since h_u^i is fiber-linear, it follows that

$$\vartheta_\lambda(\vec{h}_u^i(\lambda)) = h_u^i(\lambda),$$

which proves the first identity.

Concerning the second identity, the conventions of Section 2.2 give

$$[\vec{h}^1, \vec{h}^2] = \overrightarrow{\{h^1, h^2\}} = \overrightarrow{\sigma(\vec{h}^1, \vec{h}^2)}.$$

The Hamiltonian

$$\lambda \longmapsto \sigma_\lambda(\vec{h}^1(\lambda), \vec{h}^2(\lambda))$$

is fiber-linear. Applying the first identity to this Hamiltonian yields the conclusion. $\square$

7.3.2 Differential and Hessian of the endpoint map

Let

$$0 = t_0 < t_1 < \cdots < t_k < t_{k+1} = T$$

be the partition associated with the piecewise regular extremal pair $(\tilde{u}, \tilde{\lambda})$, and set

$$I_j := (t_{j-1}, t_j), \quad j = 1, \dots, k+1.$$

We use the local control coordinates introduced in Section 7.1, in which the reference control is represented by $\tilde{u}(t) = 0$ on every interval I_j . For brevity, we denote the tangent space to the control manifold at the reference control by

$$\mathcal{V} := T_{(0,0)}\mathcal{U}.$$

Every variation $w \in \mathcal{V}$ is written as

$$w = (\theta, v),$$

where $(\theta, 0)$ is called a *time variation* and $(0, v)$ is called a *control variation*. For later use, we introduce the subspaces

$$\mathcal{V}_\tau := \{w \in \mathcal{V} \mid \text{supp } w \subset [0, \tau]\}, \quad 0 \leq \tau \leq T,$$

and

$$\mathcal{V}_{\tau_1, \tau_2} := \{w \in \mathcal{V} \mid \text{supp } w \subset [\tau_1, \tau_2]\}, \quad 0 \leq \tau_1 < \tau_2 \leq T.$$

Let $\tilde{\Phi}_{0,t}$ denote the Hamiltonian flow of the reference extremal, so that

$$\tilde{\lambda}_t = \tilde{\Phi}_{0,t}(\tilde{\lambda}_0).$$

We transport the lifted endpoint map back to the initial cotangent space by setting

$$\mathcal{G}(\nu, u) := \tilde{\Phi}_{0,T}^{-1}(\mathcal{E}(\nu, u)),$$

and define the corresponding map on the base manifold by

$$G := \pi \circ \mathcal{G}.$$

In particular,

$$\mathcal{G}(0, 0) = \tilde{\lambda}_0, \quad G(0, 0) = q_0.$$

For $t \in I_j$, define the transported-back Hamiltonian

$$\mathfrak{h}_t(\lambda, \nu, u) := ((1 + \nu)h_u^j - h_0^j) \circ \tilde{\Phi}_{0,t}(\lambda).$$

Its Hamiltonian vector field is

$$\vec{\mathfrak{h}}_t(\lambda, \nu, u) = (\tilde{\Phi}_{0,t}^{-1})_*((1 + \nu)\vec{h}_u^j - \vec{h}_0^j) \Big|_{\tilde{\Phi}_{0,t}(\lambda)}.$$

The variation formula of chronological calculus gives

$$\mathcal{G}(\nu, u) = \overrightarrow{\exp} \int_0^T \vec{\mathfrak{h}}_t(\cdot, \nu(t), u(t)) dt \circ \tilde{\lambda}_0.$$

For $w(t) = (\theta(t), v(t))$, define

$$X_t[w(t)] := \frac{\partial \vec{\mathfrak{h}}_t}{\partial(\nu, u)}(\tilde{\lambda}_0, \nu, u)[\theta(t), v(t)] \Big|_{(\nu, u)=(0,0)},$$

and

$$Z_t v(t) := \frac{\partial \vec{\mathfrak{h}}_t}{\partial u}(\tilde{\lambda}_0, 0, u)[v(t)] \Big|_{u=0}.$$

Recalling the transported-back Hamiltonian vector fields $\vec{\mathcal{H}}^j$ introduced in Chapter 6, we have

$$X_t[\theta(t), v(t)] = \theta(t)\vec{\mathcal{H}}^j + Z_t v(t), \quad t \in I_j.$$

We introduce the notation

$$\Theta_j := \int_{t_{j-1}}^{t_j} \theta(t) dt, \quad \eta_v^j := \int_{t_{j-1}}^{t_j} Z_t v(t) dt.$$

The first-order formula of Chapter 3 now gives

$$DG[w] = d_{\tilde{\lambda}_0} \pi \left[\int_0^T X_t[w(t)] dt \right] = d_{\tilde{\lambda}_0} \pi \left[\sum_{j=1}^{k+1} \Theta_j \vec{\mathcal{H}}^j + \int_0^T Z_t v(t) dt \right]. \quad (7.3.2)$$

Pairing this identity with $\tilde{\lambda}_0$ and using Lemma 7.3.1, we obtain

$$\left\langle \tilde{\lambda}_0, DG[w] \right\rangle = \sum_{j=1}^{k+1} \int_{t_{j-1}}^{t_j} \theta(t) h_0^j(\tilde{\lambda}_t) dt + \sum_{j=1}^{k+1} \int_{t_{j-1}}^{t_j} \frac{\partial h_u^j}{\partial u}(\tilde{\lambda}_t)[v(t)] \Big|_{u=0} dt.$$

The first integrand is the value of the Hamiltonian along the reference extremal, while the second one contains the term with the derivative of the control variables. Thus, arbitrary localized control variations recover the fact that optimal controls must be critical points of the PMP Hamiltonian function, and arbitrary time variations recover the zero-Hamiltonian condition for the free-final-time problem. Since $(\tilde{u}, \tilde{\lambda})$ satisfies PMP, the covector $\tilde{\lambda}_0$ annihilates the image of DG .

It follows from (7.3.2) that

$$\ker DG = \left\{ (\theta, v) \in \mathcal{V} \left| \sum_{j=1}^{k+1} \Theta_j \vec{\mathcal{H}}^j + \int_0^T Z_t v(t) dt \in \Pi_0 \right. \right\},$$

where

$$\Pi_0 := T_{\tilde{\lambda}_0} (T_{q_0}^* M).$$

For brevity, we set

$$\mathcal{K} := \ker DG, \quad \mathcal{K}_\tau := \mathcal{K} \cap \mathcal{V}_\tau, \quad \mathcal{K}_{\tau_1, \tau_2} := \mathcal{K} \cap \mathcal{V}_{\tau_1, \tau_2}.$$

We now turn to the second differential. Unless otherwise specified, all derivatives of $\mathbf{h}_t$ and $\vec{\mathbf{h}}_t$ appearing below are taken with respect to (ν, u) and evaluated at $(\tilde{\lambda}_0, 0, 0)$. The general formula proved in Chapter 3 yields the quadratic form

$$Q(w) := - \int_0^T D^2 \mathbf{h}_t[w(t)]^2 dt - \int_0^T \int_0^t \sigma_{\tilde{\lambda}_0}(X_s[w(s)], X_t[w(t)]) ds dt, \quad w \in \mathcal{V}. \quad (7.3.3)$$

Its restriction to $\mathcal{K}$ is the second differential of the extended endpoint map:

$$Q|_{\mathcal{K}} = D^2 J - \tilde{\lambda}_0 D^2 G.$$

The Legendre form $D^2 \mathbf{h}_t$ in (7.3.3) depends only on control-control contributions. Indeed, the transported Hamiltonian is affine in ν , and therefore

$$D^2 \mathbf{h}_t[\theta(t), 0]^2 = 0.$$

Moreover,

$$D^2 \mathbf{h}_t[(\theta(t), 0), (0, v(t))] = \theta(t) \frac{\partial h_u^j}{\partial u}(\tilde{\lambda}_t)[v(t)] \Big|_{u=0} = 0, \quad t \in I_j.$$

Consequently,

$$D^2\mathbf{h}_t[w(t)]^2 = D^2\mathbf{h}_t[v(t)]^2. \quad (7.3.4)$$

The following identity allows us to separate the mixed time-control terms interval by interval.

Proposition 7.3.2. *For every $j = 1, \dots, k+1$, every $t \in I_j$, and every $v \in T_0U_j$, one has*

$$\left\{ h_0^j, \frac{\partial h_u^j}{\partial u}[v] \right\} \Big|_{u=0} (\tilde{\lambda}_t) = 0. \quad (7.3.5)$$

Equivalently, in transported-back coordinates,

$$\sigma_{\tilde{\lambda}_0}(\vec{\mathcal{H}}^j, Z_tv) = 0. \quad (7.3.6)$$

Proof. In the local coordinates fixed in Section 7.1, the reference control is identically zero on I_j . By the first order condition, we have that $\tilde{u}(t)$ is a critical point of $h_u^j(\tilde{\lambda}_t)$:

$$\frac{\partial h_u^j}{\partial u}(\tilde{\lambda}_t) \Big|_{u=0} = 0, \quad t \in I_j.$$

Differentiating this identity along the reference extremal and using

$$\dot{\tilde{\lambda}}_t = \vec{h}_0^j(\tilde{\lambda}_t)$$

gives (7.3.5). Since $\tilde{\Phi}_{0,t}$ is symplectic, transporting the two Hamiltonian vector fields back to $\tilde{\lambda}_0$ gives (7.3.6). $\square$

We first compute the term in (7.3.3) with the double integral in the case of a single switching time. Assume therefore that $k = 1$, so that

$$I_1 = (0, t_1), \quad I_2 = (t_1, T).$$

For pure time variations, the contributions for which s and t belong to the same interval vanish by skew-symmetry: for $t, s \in I_1$,

$$\sigma_{\tilde{\lambda}_0}(X_s[\theta(s), 0], X_t[\theta(t), 0]) = \theta(s)\theta(t)\sigma_{\tilde{\lambda}_0}(\vec{\mathcal{H}}^1, \vec{\mathcal{H}}^1) = 0,$$

and similarly if $t, s \in I_2$ we have $\sigma_{\tilde{\lambda}_0}(X_s[\theta(s), 0], X_t[\theta(t), 0]) = 0$. Hence,

$$\begin{aligned} \int_0^T \int_0^t \sigma_{\tilde{\lambda}_0}(X_s[\theta(s), 0], X_t[\theta(t), 0]) ds dt &= \int_{t_1}^T \int_0^{t_1} \theta(s)\theta(t)\sigma_{\tilde{\lambda}_0}(\vec{\mathcal{H}}^1, \vec{\mathcal{H}}^2) ds dt \\ &= \Theta_1\Theta_2\sigma_{\tilde{\lambda}_0}(\vec{\mathcal{H}}^1, \vec{\mathcal{H}}^2). \end{aligned} \quad (7.3.7)$$

The control-control contribution is simply

$$\int_0^T \int_0^t \sigma_{\tilde{\lambda}_0}(Z_s v(s), Z_t v(t)) ds dt.$$

For the first mixed contribution, Proposition 7.3.2 eliminates the regions contained in $I_1 \times I_1$ and $I_2 \times I_2$. Thus,

$$\begin{aligned} \int_0^T \int_0^t \sigma_{\tilde{\lambda}_0}(X_s[\theta(s), 0], X_t[0, v(t)]) ds dt &= \int_{t_1}^T \int_0^{t_1} \theta(s) \sigma_{\tilde{\lambda}_0}(\vec{\mathcal{H}}^1, Z_t v(t)) ds dt \\ &= \Theta_1 \sigma_{\tilde{\lambda}_0}(\vec{\mathcal{H}}^1, \eta_v^2). \end{aligned}$$

Similarly,

$$\int_0^T \int_0^t \sigma_{\tilde{\lambda}_0}(X_s[0, v(s)], X_t[\theta(t), 0]) ds dt = \Theta_2 \sigma_{\tilde{\lambda}_0}(\eta_v^1, \vec{\mathcal{H}}^2). \quad (7.3.8)$$

Combining (7.3.4) and (7.3.7)–(7.3.8), we obtain

$$\begin{aligned} Q(\theta, v) &= - \int_0^T D^2 \mathbf{h}_t[v(t)]^2 dt - \int_0^T \int_0^t \sigma_{\tilde{\lambda}_0}(Z_s v(s), Z_t v(t)) ds dt \\ &\quad - \Theta_1 \Theta_2 \sigma_{\tilde{\lambda}_0}(\vec{\mathcal{H}}^1, \vec{\mathcal{H}}^2) - \Theta_1 \sigma_{\tilde{\lambda}_0}(\vec{\mathcal{H}}^1, \eta_v^2) - \Theta_2 \sigma_{\tilde{\lambda}_0}(\eta_v^1, \vec{\mathcal{H}}^2). \end{aligned}$$

The formula for finitely many switching times follows by decomposing the triangle

$$\{(s, t) \in [0, T]^2 \mid s \leq t\}$$

into the regions $I_i \times I_j$, with $i \leq j$. On each diagonal region $I_j \times I_j$, the time-time contribution vanishes by skew-symmetry and the mixed contributions vanish by Proposition 7.3.2. Every region $I_i \times I_j$ with $i < j$ gives the same four cross terms as in the one-switch computation. Consequently,

$$\begin{aligned} Q(\theta, v) &= - \int_0^T D^2 \mathbf{h}_t[v(t)]^2 dt - \int_0^T \int_0^t \sigma_{\tilde{\lambda}_0}(Z_s v(s), Z_t v(t)) ds dt \\ &\quad - \sum_{1 \leq i < j \leq k+1} \Theta_i \Theta_j \sigma_{\tilde{\lambda}_0}(\vec{\mathcal{H}}^i, \vec{\mathcal{H}}^j) \\ &\quad - \sum_{1 \leq i < j \leq k+1} \left[\Theta_j \sigma_{\tilde{\lambda}_0}(\eta_v^i, \vec{\mathcal{H}}^j) + \Theta_i \sigma_{\tilde{\lambda}_0}(\vec{\mathcal{H}}^i, \eta_v^j) \right]. \end{aligned} \quad (7.3.9)$$

This is the required decomposition of the Hessian into the pointwise control term, the non-local control-control interaction, the interaction between time variations on distinct arcs, and the two mixed time-control interactions.

Finally, for $0 \leq \tau_1 < \tau_2 \leq T$, we define the quadratic form

$$Q_{\tau_1, \tau_2}(w) := - \int_{\tau_1}^{\tau_2} D^2 \mathbf{h}_t[w(t)]^2 dt - \int_{\tau_1}^{\tau_2} \int_{\tau_1}^t \sigma_{\tilde{\lambda}_0}(X_s[w(s)], X_t[w(t)]) ds dt, \quad w \in \mathcal{V}_{\tau_1, \tau_2}.$$

The quadratic forms Q and Q_{τ_1, τ_2} will be used in the next section to construct the $\mathcal{L}$ -derivative and the Jacobi curve of the piecewise regular extremal.

7.4 Jacobi curve of a piecewise regular extremal

In the previous sections we introduced the endpoint map and computed its differential and Hessian. Building on these results, in this section we prove the characterization of the Jacobi curve of a piecewise regular extremal pair $(\tilde{u}, \tilde{\lambda})$ stated in Theorem 6.1.6. We then give an interpretation of the regular arcs of the Jacobi curve analogous to the one developed in Subsection 3.3.2, with the appropriate initial conditions and jumps produced by the time variations at the switching points.

7.4.1 Proof of Theorem 6.1.6

The proof combines the working characterization of the Jacobi curve in Proposition 3.3.5 with the recursive construction of Lemma 3.3.6. Proposition 7.3.2 allows us to separate the contribution of the time variations from that of the control variations on each interval of regularity. More precisely, the time variations determine the right limit at the initial point of each regular arc, while the subsequent evolution is governed by the Jacobi equation for regular extremals.

For the reader convenience, we recall some notation and restate precisely the result. For $t \in (t_{j-1}, t_j)$, let $b_t : T_{\tilde{\lambda}_0}(T^*M) \rightarrow \mathbb{R}$ be the quadratic Hamiltonian

$$b_t(\eta) := -\frac{1}{2} \left\langle \sigma_{\tilde{\lambda}_0}(Z_t \cdot, \eta), \mathbb{L}_t^{-1} \sigma_{\tilde{\lambda}_0}(Z_t \cdot, \eta) \right\rangle, \quad (7.4.1)$$

where Z_t was introduced in (3.2.1), $\mathbb{L}_t$ is the Legendre form introduced in (2.4.2) and we use the convention that both Z_t and $\mathbb{L}_t$ for $t \in (t_{j-1}, t_j)$ are defined on $T_{\tilde{u}(t)}U_j$. Moreover we define B_t to be the flow

$$B_t = \overrightarrow{\exp} \int_{t_{j-1}}^t \vec{b}_s ds, \quad t \in (t_{j-1}, t_j), \quad j = 1, \dots, k+1.$$

Theorem 7.4.1. *Let $(\tilde{u}, \tilde{\lambda})$ be a piecewise regular extremal pair for problem (6.1.4). Denote by $(\Lambda_t)_{t \in [0, T]}$ the Jacobi curve associated with $(\tilde{u}, \tilde{\lambda})$. Then, Λ is a left-continuous piecewise smooth curve in the Lagrangian Grassmannian of $T_{\tilde{\lambda}_0}(T^*M)$, with discontinuities at*

$0, t_1, \dots, t_k$. It is determined recursively by the following algorithm. We set $\Lambda_0 = \Pi_{\tilde{\lambda}_0}$ and, for every $j = 1, \dots, k+1$, the following properties hold.

1. For every $t \in (t_{j-1}, t_j]$, one has $\vec{\mathcal{H}}^j \in \Lambda_t$. In particular, the right limit at t_{j-1} is

$$\Lambda_{t_{j-1}}^+ = \Lambda_{t_{j-1}}^{\vec{\mathcal{H}}^j} = \left(\Lambda_{t_{j-1}} \cap (\vec{\mathcal{H}}^j)^\perp \right) + \mathbb{R}\vec{\mathcal{H}}^j.$$

2. For every $t \in (t_{j-1}, t_j]$, the space Λ_t is generated by the solutions of the Cauchy problem on $T_{\tilde{\lambda}_0}^*(T^*M)$

$$\dot{\eta}(t) = Z_t \mathbb{L}_t^{-1} \sigma_{\tilde{\lambda}_0}(Z_t \cdot, \eta(t)), \quad \eta(t_{j-1}) \in \Lambda_{t_{j-1}}^+. \quad (7.4.2)$$

If B_t denotes the flow of the right-hand side of (7.4.2), then $\Lambda_t = B_t \Lambda_{t_{j-1}}^+$.

Proof. We compute the Jacobi curve recursively by means of the following algorithm. For each time interval (t_{j-1}, t_j) , $j = 1, \dots, k+1$:

1. we compute $\Lambda_t(V_j)$, $t \in (t_{j-1}, t_j]$, where $V_j = \mathcal{V}_{t_{j-1}} \oplus V_\theta$ is the subspace of variations generated by $\mathcal{V}_{t_{j-1}}$, that is, all admissible variations supported in $[0, t_{j-1}]$, and $V_\theta = \text{span}\{\mathbb{1}_{[t_{j-1}, t]}\}$, the time variations supported in (t_{j-1}, t) . This is done applying Lemma 3.3.6, with $\mathcal{V}_1 = \mathcal{V}_{t_{j-1}}$ and $V_2 = V_\theta$. In particular, we will find that $\Lambda_t(V_j) = \Lambda_{t_{j-1}}^{\vec{\mathcal{H}}^j}$, that is, $\Lambda_t(V_j)$ is constant for $t \in (t_{j-1}, t_j)$. Notice that, since Q_τ depends only on the integral of the variations of time on the intervals (t_{j-1}, t_j) , we can restrict to consider θ constant on some subinterval of (t_{j-1}, t_j) .
2. We compute Λ_t , $t \in (t_{j-1}, t_j]$, from the previous approximation $\Lambda_t(V_j)$. This is done by using the splitting $\mathcal{V}_t = V_j \oplus W_j$, W_j being the set of control variations supported in (t_{j-1}, t_j) . Since $\Lambda_t(V_j)$ is constant for $t \in (t_{j-1}, t_j)$ and time variations commutes with control variations, it is possible to characterize solutions of Equation (3.3.1) as solutions of a suitable Jacobi type equation, following the same strategy that is applied to the case of a regular arc.

We first prove points 1. and 2. in the Theorem's statement on the interval $(0, t_1]$ and then explain how the same construction is iterated on the subsequent intervals of regularity.

We begin with point 1. By definition,

$$\Lambda_0 = \Pi = T_{\tilde{\lambda}_0}^*(T_{q_0}^*M) \subset T_{\tilde{\lambda}_0}^*(T^*M).$$

Fix $t \in (0, t_1)$ and consider the one-dimensional space of time variations

$$V_\theta := \text{span} \left\{ (\mathbb{1}_{(0,t)}, 0) \right\}.$$

We compute $\Lambda_t(V_\theta)$ from Λ_0 by applying Lemma 3.3.6. Since the quadratic form Q_t depends only on the integral of the time variations, as follows from Equation (7.3.9), it is enough

to consider time variations that are constant on a subinterval of $(0, t_1)$. For $w_1, w_2 \in V_\theta$, the pointwise second-order term vanishes because $D^2\mathbf{h}_s|_{V_\theta} = 0$, while the term inside the double integral vanishes because

$$\sigma_{\tilde{\lambda}_0}(X_s w_1(s), X_r w_2(r)) = \sigma_{\tilde{\lambda}_0}(\vec{\mathcal{H}}^1, \vec{\mathcal{H}}^1) = 0.$$

Consequently,

$$Q_t(w_1, w_2) = 0, \quad w_1, w_2 \in V_\theta.$$

Moreover, the right-hand side of Equation (3.3.3) is

$$\int_0^t \sigma_{\tilde{\lambda}_0}(\eta, X_s w_2(s)) ds = \int_0^t \sigma_{\tilde{\lambda}_0}(\eta, \vec{\mathcal{H}}^1) ds = t \sigma_{\tilde{\lambda}_0}(\eta, \vec{\mathcal{H}}^1), \quad \eta \in \Lambda_0. \quad (7.4.3)$$

Since $(\vec{\mathcal{H}}^1)^\perp$ has codimension one in $T_{\tilde{\lambda}_0}(T^*M)$ and Λ_0 has dimension d , Grassmann's identity gives

$$\dim(\Lambda_0 \cap (\vec{\mathcal{H}}^1)^\perp) \in \{d-1, d\}.$$

Suppose first that this dimension is d . Then $\Lambda_0 \subset (\vec{\mathcal{H}}^1)^\perp$ and, since Λ_0 is Lagrangian, $\vec{\mathcal{H}}^1 \in \Lambda_0$. This implies that (7.4.3) also vanishes and Therefore,

$$\Lambda_0^{\vec{\mathcal{H}}^1} = \Lambda_0 = \Pi.$$

Hence, Lemma 3.3.6 yields $\Lambda_t(V_\theta) = \Pi$. Suppose instead that

$$\dim(\Lambda_0 \cap (\vec{\mathcal{H}}^1)^\perp) = d-1.$$

Then there exists $\eta_1 \in \Lambda_0$ such that $\sigma_{\tilde{\lambda}_0}(\vec{\mathcal{H}}^1, \eta_1) \neq 0$, and there exist $d-1$ linearly independent vectors $\eta_2, \dots, \eta_d \in \Lambda_0$ such that

$$\sigma_{\tilde{\lambda}_0}(\vec{\mathcal{H}}^1, \eta_i) = 0, \quad i = 2, \dots, d.$$

Thus $\eta_2, \dots, \eta_d$ annihilate the right-hand side of (7.4.3). In addition, the choice $\eta = 0$ and $w_1 = \mathbf{1}_{(0,t)}$ in Equation (3.3.2) satisfies Equation (3.3.3), because both sides vanish on V_θ . Lemma 3.3.6 therefore yields

$$\Lambda_t(V_\theta) = \text{span} \left\{ \vec{\mathcal{H}}^1, \eta_2, \dots, \eta_d \right\}.$$

Combining the two possible dimension cases, we obtain

$$\Lambda_t(V_\theta) = \left(\Pi \cap (\vec{\mathcal{H}}^1)^\perp \right) + \mathbb{R}\vec{\mathcal{H}}^1 = \Lambda_0^{\vec{\mathcal{H}}^1}, \quad t \in (0, t_1].$$

We now add the remaining admissible variations. Let $W_t \subset \mathcal{V}_t$ be any space of variations containing V_θ . Proposition 7.3.2 implies that the time variation commutes symplectically with every additional variation, namely

$$\sigma_{\tilde{\lambda}_0}(X_s w(s), X_s w_\theta(s)) = \sigma_{\tilde{\lambda}_0}(X_s w(s), \vec{\mathcal{H}}^1) = 0, \quad s \in (0, t),$$

for every $w \in W_t$ and $w_\theta \in V_\theta$. Hence the vector $\vec{\mathcal{H}}^1$ satisfies Equation (3.3.1) also when the test variations range over W_t . It follows that $\vec{\mathcal{H}}^1 \in \Lambda_t(W_t)$. Passing to the generalized limit over the finite-dimensional approximations gives

$$\vec{\mathcal{H}}^1 \in \Lambda_t, \quad t \in (0, t_1].$$

In particular, $\vec{\mathcal{H}}^1 \in \Lambda_0^+$, and point 1 is proved on the first interval.

We now prove point 2 on $(0, t_1]$. Fix $t \in (0, t_1]$ and denote by W_t the space of control variations supported in $(0, t)$. We have already shown that the time variations generate the constant Lagrangian subspace

$$\Lambda_t(V_\theta) = \Lambda_0^+ = \Lambda_0^{\vec{\mathcal{H}}^1} = \Pi^{\vec{\mathcal{H}}^1}.$$

The crucial point is that the contribution of the time variations can be incorporated entirely into this new initial condition before adding the control variations. Indeed, Equation (7.3.4) shows that the pointwise part of the second variation has no mixed time-control terms, while Proposition 7.3.2 gives

$$\sigma_{\tilde{\lambda}_0}(\vec{\mathcal{H}}^1, Z_s v) = 0, \quad v \in T_0 U_1.$$

Hence, also the mixed time-control contributions in the double integral defining Q_t vanish. Since Q_t vanishes on pairs of time variations, for $\theta_i \in V_\theta$ and $v_i \in W_t$ we obtain

$$Q_t(\theta_1 + v_1, \theta_2 + v_2) = Q_t(v_1, v_2).$$

Therefore, once the equations corresponding to the time variations have been used to replace Π with Λ_0^+ , the remaining equations involve only control variations and are identical to those arising for a regular extremal arc. This allows us to repeat the proof of Theorem 3.3.9 with the initial vertical space replaced by Λ_0^+ .

More explicitly, we have to determine $\eta_0 \in \Lambda_0^+$ and $v \in W_t$ satisfying

$$Q_t(v, w) = \int_0^t \sigma_{\tilde{\lambda}_0}(\eta_0, Z_s w(s)) ds, \quad w \in W_t.$$

Setting

$$\eta(s) = \eta_0 + \int_0^s Z_r v(r) dr,$$

the computation from Equation (3.3.5) to Equation (3.3.6) applies without modification and gives

$$-\mathbb{L}_s(v(s), \cdot) - \sigma_{\tilde{\lambda}_0}(\eta(s), Z_s \cdot) = 0.$$

The strong Legendre condition then yields

$$\dot{\eta}(s) = Z_s \mathbb{L}_s^{-1} \sigma_{\tilde{\lambda}_0}(Z_s \cdot, \eta(s)), \quad \eta(0) \in \Lambda_0^+.$$

The converse implication follows exactly as in the final part of the proof of Theorem 3.3.9: every solution of this Cauchy problem determines a control variation satisfying the preceding equation. Consequently,

$$\Lambda_t = B_t^1 \Lambda_0^+ = B_t^1 \Pi^{\vec{\mathcal{H}}^1}, \quad t \in (0, t_1].$$

Finally, Proposition 7.3.2 also shows that the constant curve

$$\eta(s) = \vec{\mathcal{H}}^1$$

solves the Jacobi equation, since its right-hand side vanishes. This is consistent with point 1 and with the inclusion $\vec{\mathcal{H}}^1 \in \Lambda_t$ for every $t \in (0, t_1]$.

It remains to iterate the construction. Consider first the interval $(t_1, t_2]$ and let V_θ be the space of time variations supported in (t_1, t_2) . These variations determine the new initial condition at t_1 for the Jacobi equation on the second regular arc. Since $\vec{\mathcal{H}}^1 \in \Lambda_{t_1}$ and

$$\sigma_{\tilde{\lambda}_0}(\vec{\mathcal{H}}^1, \vec{\mathcal{H}}^2) \neq 0,$$

the inclusion $\Lambda_{t_1} \subset (\vec{\mathcal{H}}^2)^\perp$ is impossible. Since $(\vec{\mathcal{H}}^2)^\perp$ has codimension one, it follows that

$$\dim(\Lambda_{t_1} \cap (\vec{\mathcal{H}}^2)^\perp) = d - 1.$$

We may therefore choose $d - 1$ linearly independent vectors $\eta_2, \dots, \eta_d$ in this intersection. As on the first interval, $Q_t(w_1, w_2) = 0$ for all $w_1, w_2 \in V_\theta$, and the pair $\eta = 0, w_1 = \mathbb{1}_{(t_1, t_2)}$ satisfies Equation (3.3.3). Lemma 3.3.6 then gives

$$\Lambda_t(V_\theta) = \Lambda_{t_1}^{\vec{\mathcal{H}}^2}, \quad t \in (t_1, t_2].$$

In particular,

$$\Lambda_{t_1}^+ = \Lambda_{t_1}^{\vec{\mathcal{H}}^2}.$$

The same commutation argument used on the first interval shows that

$$\vec{\mathcal{H}}^2 \in \Lambda_t, \quad t \in (t_1, t_2].$$

Finally, applying again Proposition 3.3.5 again to add the control variations, and repeating the derivation above, shows that Λ_t on $(t_1, t_2]$ is generated by the solutions of (7.4.2) with initial condition in $\Lambda_{t_1}^+$.

Repeating the same two steps on every interval $(t_{j-1}, t_j]$ proves both assertions of the theorem for $j = 1, \dots, k+1$. On each interval the Jacobi curve is generated by the flow of a linear differential equation and is therefore smooth. Its value at t_j is the endpoint of the smooth arc defined on $(t_{j-1}, t_j]$, while the time variations determine the right limit $\Lambda_{t_j}^+$. Hence the curve is left continuous and its only possible discontinuities occur at $0, t_1, \dots, t_k$. $\square$

7.4.2 Hamiltonian characterization of the regular arcs

We now give a characterization of the smooth pieces of the Jacobi curve parallel to the one obtained for regular extremals in Subsection 3.3.2. First, we prove that regular arcs of the Jacobi curve of a piecewise regular extremal are increasing.

Corollary 7.4.2. *Let $(\tilde{u}, \tilde{\lambda})$ be a piecewise regular extremal pair and let Λ be its Jacobi curve. Then the restriction of Λ to every interval (t_{j-1}, t_j) is increasing.*

Proof. The strong Legendre condition implies that $\mathbb{L}_t$ is negative definite on each interval (t_{j-1}, t_j) . Hence the quadratic Hamiltonian b_t in (7.4.1) is positive definite on each interval (t_{j-1}, t_j) . The conclusion follows from Proposition 3.3.10 applied to

$$\Lambda_t = B_t \Lambda_{t_{j-1}}^+.$$

$\square$

To compare the Jacobi equation with the maximized Hamiltonian dynamics, define

$$H_t := (H - h_{\tilde{u}(t)}) \circ \tilde{\Phi}_{0,t}, \quad t \in [0, T]$$

The corresponding Hamiltonian flow compares the maximized flow with the flow generated by the reference control after both have been transported back to $\tilde{\lambda}_0$.

Proposition 7.4.3. *For every $j = 1, \dots, k+1$, the following identities hold on (t_{j-1}, t_j) .*

1. *The transported-back flow ϕ_t is generated by H_t , namely*

$$\phi_t = \overrightarrow{\exp} \int_{t_{j-1}}^t \vec{H}_s ds.$$

2. *The function H_t vanishes to first order at $\tilde{\lambda}_0$ and its Hessian is the quadratic Hamiltonian of the Jacobi equation:*

$$H_t(\tilde{\lambda}_0) = 0, \quad d_{\tilde{\lambda}_0} H_t = 0, \quad b_t = \frac{1}{2} \text{Hess}_{\tilde{\lambda}_0} H_t.$$

3. The linearization of ϕ_t at $\tilde{\lambda}_0$ is the flow of the Jacobi equation:

$$(\phi_t)_* = B_t.$$

Proof. The definition (6.1.6) and the variation formula for non-autonomous flows give

$$\tilde{\Phi}_{0,t}^{-1} \circ \Phi_{t-t_{j-1}} \circ \tilde{\Phi}_{0,t_{j-1}} = \overrightarrow{\exp} \int_{t_{j-1}}^t \overrightarrow{(H - h_{\tilde{u}(s)}) \circ \tilde{\Phi}_{0,s}} ds,$$

which proves the first identity. Along the reference extremal, the control $\tilde{u}(t)$ maximizes h_u , so H and $h_{\tilde{u}(t)}$ have the same value and the same first differential at $\tilde{\lambda}_t$. After composition with $\tilde{\Phi}_{0,t}$, this gives

$$H_t(\tilde{\lambda}_0) = 0, \quad d_{\tilde{\lambda}_0} H_t = 0.$$

The computation of the second differential is exactly the regular-arc computation of Proposition 3.3.12, applied on (t_{j-1}, t_j) after translating the initial time and transporting the dynamics to $T_{\tilde{\lambda}_0}(T^*M)$.

Finally, a similar strategy combined with the argument used in Corollary 3.3.13 yields also point 3. in the statement. $\square$

Remark 7.4.4. Equation (7.4.2) is written on the fixed symplectic space $T_{\tilde{\lambda}_0}(T^*M)$ because the entire dynamics have been transported back along the reference flow. An equivalent formulation is obtained in the moving spaces $T_{\tilde{\lambda}_t}(T^*M)$. Define

$$\Xi_t := (\tilde{\Phi}_{0,t})_* \Lambda_t.$$

On the interval (t_{j-1}, t_j) , the space Ξ_t is generated by the solutions of

$$\dot{\xi}(t) = d_{\tilde{\lambda}_t} \vec{H}[\xi(t)], \quad \xi(t_{j-1}) \in \Xi_{t_{j-1}}^+,$$

where

$$\Xi_{t_{j-1}}^+ = \Xi_{t_{j-1}}^{\vec{H}^j(\lambda_{t_{j-1}})} = \left(\Xi_{t_{j-1}} \cap \vec{H}^j(\lambda_{t_{j-1}})^\perp \right) + \mathbb{R} \vec{H}^j(\lambda_{t_{j-1}}).$$

Thus, the Jacobi curve is obtained by alternating the linearized Hamiltonian evolution on each regular arc with the jump prescribed at each switching point.

7.5 The maximized Hamiltonian flow and the field of extremals

In the previous sections of this Chapter we found the structure of the Jacobi curve associated with a piecewise regular extremal and, as a direct consequence, we obtained a second-order necessary condition for optimality. The purpose of the remaining part of this Chapter is to prove a second-order sufficient optimality condition.

In this section, we show that, in the case of a piecewise regular extremal, the maximized Hamiltonian produces a piecewise smooth flow in a neighbourhood of the reference extremal, and then it can be used naturally as a super-Hamiltonian function. This allows us to prove that the set of Lagrange multipliers in this case has a structure of Lipschitz (piecewise-smooth) manifold and also to produce a suitable field of extremals, which will be used to prove the sufficient optimality condition.

7.5.1 The maximized Hamiltonian and its flow

Let $(\tilde{u}, \tilde{\lambda})$ be a piecewise regular extremal pair for the optimal control problem (6.1.4), and let

$$h_u(\lambda) = \langle \lambda, f(\pi(\lambda), u) \rangle - L(\pi(\lambda), u)$$

be the corresponding normal PMP Hamiltonian. We retain the notation introduced in Definition 6.1.5: the switching times are

$$0 = t_0 < t_1 < \cdots < t_k < t_{k+1} = T,$$

and, on the interval (t_{j-1}, t_j) , the reference control belongs to the smooth submanifold U_j . In the following Proposition, we prove that the strong Legendre condition determines, locally around the reference extremal, a smooth maximizing control u_j , that is a smooth function from a neighbourhood in T^*M of a smooth arc of $\tilde{\lambda}$ to U_j such that

$$H^j(\lambda) = h_{u_j(\lambda)}(\lambda).$$

Given a piecewise regular extremal $\tilde{\lambda}$ and a neighbourhood of its image $\mathcal{N} \subset T^*M$, we say that the set

$$\mathcal{N}_j := \{\lambda \in \mathcal{N} \mid H^j(\lambda) > \max_{i \neq j} H^i(\lambda)\}$$

is a regular region for $\tilde{\lambda}$ inside $\mathcal{N}$, while we say that the set

$$\Sigma_j := \{\lambda \in \mathcal{N} \mid H^j(\lambda) = H^{j+1}(\lambda)\}$$

is a switching set of $\tilde{\lambda}$ inside $\mathcal{N}$. We recall that, since all the relevant dynamics are confined in a relatively compact neighbourhood of $\tilde{\lambda}$, we assume, without loss of generality, that all the vector fields under consideration are complete.

Proposition 7.5.1. *Let $(\tilde{u}, \tilde{\lambda})$ be a piecewise regular extremal pair. There exist a neighbourhood $\mathcal{N} \subset T^*M$ of the image of $\tilde{\lambda}$ and an autonomous flow Φ on $\mathcal{N}$ with the following properties:*

1. *on every regular region $\mathcal{N}_j$ for $\tilde{\lambda}$, there is a unique smooth maximizing control $u_j : \mathcal{N}_j \rightarrow U_j$, that is a smooth function realizing*

$$H(\lambda) = H^j(\lambda) = h_{u_j(\lambda)}(\lambda), \quad \lambda \in \mathcal{N}_j;$$

2. the switching sets Σ_j are smooth hypersurfaces containing $\tilde{\lambda}_{t_j}$, and both $\vec{H}^j$ and $\vec{H}^{j+1}$ are transverse to Σ_j at every point;
3. the maximized Hamiltonian H is locally Lipschitz on $\mathcal{N}$ and smooth outside the union of the switching hypersurfaces;
4. in $\mathcal{N} \setminus (\cup_{j=1}^k \Sigma_j)$, the maximizing control

$$u_{\max}(\lambda) = \operatorname{argmax}_{u \in U} h_u(\lambda)$$

is uniquely defined and smooth, while on Σ_j the two adjacent controls $u_j(\lambda)$ and $u_{j+1}(\lambda)$ realize the same maximum;

5. the Hamiltonian flow of H , denoted by Φ , is obtained by concatenating the smooth Hamiltonian flows of the H^j at their first intersections with the switching hypersurfaces. The map $(t, \lambda) \mapsto \Phi_t(\lambda)$ is Lipschitz on its domain, and every map Φ_t is locally bi-Lipschitz, with inverse Φ_{-t} . Every extremal trajectory sufficiently close to $\tilde{\lambda}$ is an orbit of this flow. In particular,

$$H(\Phi_t(\lambda)) = H(\lambda). \quad (7.5.1)$$

Proof. We first work on a compact subinterval I of (t_{j-1}, t_j) . Along the reference extremal, PMP and the strong Legendre condition give

$$\left. \frac{\partial h_u}{\partial u} \right|_{u=\tilde{u}(t), \lambda=\tilde{\lambda}_t} = 0, \quad \left. \frac{\partial^2 h_u}{\partial u^2} \right|_{u=\tilde{u}(t), \lambda=\tilde{\lambda}_t} < 0.$$

The implicit function theorem therefore gives a smooth feedback $u_j(\lambda)$, defined on a sufficiently small neighbourhood in T^*M of $\tilde{\lambda}|_I$ satisfying

$$\left. \frac{\partial h_u}{\partial u} \right|_{u=u_j(\lambda)} = 0.$$

After reducing the neighbourhood, the Legendre form remains negative definite, and hence $u_j(\lambda)$ is the unique maximizer on the submanifold U_j . The strict separation of H^j from the other H^i along every compact subinterval, together with continuity, shows that H^j also realizes the maximum over the whole control set in a neighbourhood of the corresponding reference arc. Moreover,

$$dH^j(\lambda) = d_\lambda h_{u_j(\lambda)}(\lambda),$$

because the derivative with respect to the control vanishes at $u_j(\lambda)$. Consequently,

$$\vec{H}^j(\lambda) = \vec{h}_{u_j(\lambda)}(\lambda).$$

We now consider a switching point $\tilde{\lambda}_{t_j}$ and set

$$G_j := H^{j+1} - H^j.$$

By the strong switching condition, we have $\{H^j, H^{j+1}\}(\tilde{\lambda}_{t_j}) \neq 0$. In particular,

$$d_{\tilde{\lambda}_{t_j}} G_j \neq 0,$$

and the implicit function theorem shows that $\Sigma_j = G_j^{-1}(0)$ is a smooth hypersurface near $\tilde{\lambda}_{t_j}$. Furthermore,

$$dG_j [\vec{H}^j] = \{H^j, H^{j+1}\} \neq 0, \quad dG_j [\vec{H}^{j+1}] = \{H^{j+1}, H^j\} \neq 0,$$

so both adjacent Hamiltonian vector fields are transverse to Σ_j at $\tilde{\lambda}_{t_j}$. By continuity, the transversality persists in a sufficiently small neighbourhood of the switching point.

Let Φ^j denote the smooth Hamiltonian flow of H^j . For λ close to $\tilde{\lambda}_{t_{j-1}}$, the first intersection with Σ_j is determined by the equation

$$G_j(\Phi_s^j(\lambda)) = 0.$$

Let $g_j(s, \lambda) := G_j(\Phi_s^j(\lambda))$. Since $\partial_s g(t_j - t_{j-1}, \tilde{\lambda}_{t_{j-1}}) = \{H^j, H^{j+1}\}(\tilde{\lambda}_{t_j}) \neq 0$, the implicit function theorem provides a smooth hitting-time function $\tau_j(\lambda)$. The local trajectory is obtained by following Φ^j up to time $\tau_j(\lambda)$ and then following Φ^{j+1} from the hitting point. Repeating the construction at the finitely many switching hypersurfaces defines the local flow Φ in a neighbourhood of the whole reference trajectory.

Thus, by this construction, $(t, \lambda) \mapsto \Phi_t(\lambda)$ is Lipschitz continuous. The same construction performed backwards gives Φ_{-t} , and therefore every Φ_t is locally bi-Lipschitz. Since a nearby extremal must satisfy the maximizing condition on each regular region, uniqueness of the maximizing feedback and of the crossing times shows that it coincides with an orbit of Φ . Finally, on an arc governed by H^j one has

$$\frac{d}{dt} H^j(\Phi_t(\lambda)) = \{H^j, H^j\}(\Phi_t(\lambda)) = 0.$$

At a switching point the two adjacent Hamiltonians have the same value. Therefore the conserved values on consecutive arcs agree, which proves (7.5.1). $\square$

Corollary 7.5.2. *The maximized Hamiltonian H is a super-Hamiltonian for $(\tilde{u}, \tilde{\lambda})$.*

Proof. Let $a \in C^\infty(M)$ satisfy

$$\tilde{\lambda}_0 = d_{q_0} a,$$

and suppose that 0 is a regular value of the restriction of H to the exact Lagrangian graph

$$\mathcal{L}_0 = \{d_q a | q \in M\}.$$

If $\mathcal{M}_0$ is a relatively open connected subset of

$$\mathcal{L}_0 \cap H^{-1}(0)$$

containing $\tilde{\lambda}_0$, then the flow-out

$$\mathcal{M} = \{\Phi_t(\lambda_0) | \lambda_0 \in \mathcal{M}_0, t \in I\}$$

satisfies hypotheses **(H1)**–**(H6)** of Chapter 4. Indeed, the regularity and local bi-Lipschitz properties follow from Proposition 7.5.1, the super-Hamiltonian inequality holds with equality, and (7.5.1) implies that H vanishes on the whole flow-out. Along the reference extremal, PMP gives

$$H(\tilde{\lambda}_t) = h_{\tilde{u}(t)}(\tilde{\lambda}_t), \quad \vec{H}(\tilde{\lambda}_t) = \vec{h}_{\tilde{u}(t)}(\tilde{\lambda}_t)$$

for almost every t . □

Thus, the only hypothesis of the sufficient criterion which remains to be verified is the local invertibility of the projection of $\mathcal{M}$ onto M .

7.5.2 Structure of the set of Lagrange multipliers

We now apply the preceding construction to the set $\mathcal{L}$ of normal Lagrange multipliers introduced in Definition 3.1.4. Our aim is to show that in the case of a piecewise regular extremal, the set of Lagrange multipliers satisfies the hypothesis **(H8)** in Chapter 4. In particular, it is possible to find a field of extremals satisfying that assumption. Let

$$\Pi = T_{q_0}^* M$$

be the initial vertical fiber and set

$$\mathcal{L}_0 = \Pi \cap H^{-1}(0).$$

Locally around the reference extremal, PMP and the uniqueness statement in Proposition 7.5.1 identify $\mathcal{L}$ with the flow-out of $\mathcal{L}_0$:

$$\mathcal{L} = \{\Phi_t(\lambda_0) | \lambda_0 \in \mathcal{L}_0, t \in I\}. \tag{7.5.2}$$

We first consider the regular case.

Proposition 7.5.3 (Structure of $\mathcal{L}$ near a regular extremal). *Let $(\tilde{u}, \tilde{\lambda})$ be a regular extremal pair. Assume that*

$$\vec{H}(\tilde{\lambda}_0) \notin T_{\tilde{\lambda}_0} \mathcal{L}_0.$$

Then there exists a neighbourhood $\mathcal{O}_{\tilde{\lambda}} \subset T^*M$ of the image of $\tilde{\lambda}$ such that $\mathcal{L} \cap \mathcal{O}_{\tilde{\lambda}}$ is a smooth immersed Lagrangian submanifold of dimension d . If

$$\lambda = \Phi_t(\lambda_0), \quad \lambda_0 \in \mathcal{L}_0,$$

then

$$T_{\lambda} \mathcal{L} = d_{\lambda_0} \Phi_t (T_{\lambda_0} \mathcal{L}_0) + \mathbb{R} \vec{H}(\lambda). = d_{\lambda_0} \Phi_t \left(T_{\lambda_0} \mathcal{L}_0 + \mathbb{R} \vec{H}(\lambda_0) \right). \quad (7.5.3)$$

Proof. The transversality assumption implies that 0 is a regular value of $H|_{\Pi}$ near $\tilde{\lambda}_0$, hence $\mathcal{L}_0$ is a smooth $(d-1)$ -dimensional submanifold of the vertical fibre Π . Since Π is Lagrangian, $\mathcal{L}_0$ is isotropic. Consider the flow-out map

$$\Xi : \mathcal{L}_0 \times I \longrightarrow T^*M, \quad \Xi(\lambda_0, t) = \Phi_t(\lambda_0).$$

Its differential is

$$d_{(\lambda_0, t)} \Xi(v, s) = d_{\lambda_0} \Phi_t(v) + s \vec{H}(\Phi_t(\lambda_0)).$$

At $(\tilde{\lambda}_0, 0)$, the transversality assumption implies that this map has rank d . After reducing $\mathcal{L}_0$ and the time interval, the same holds in a neighbourhood of the reference flow-out, and hence Ξ is an immersion. Equation (7.5.2) then gives the smooth immersed structure of $\mathcal{L}$ and the tangent-space formula (7.5.3).

It remains to prove that the tangent spaces are Lagrangian. If $v \in T_{\lambda_0} \mathcal{L}_0$, then

$$\sigma_{\lambda_0} \left(v, \vec{H}(\lambda_0) \right) = dH(\lambda_0)[v] = 0,$$

because $\mathcal{L}_0 \subset H^{-1}(0)$. Thus

$$T_{\lambda_0} \mathcal{L}_0 + \mathbb{R} \vec{H}(\lambda_0)$$

is a d -dimensional isotropic subspace, and therefore it is Lagrangian. Since every smooth Hamiltonian flow preserves the symplectic form, the image in (7.5.3) is Lagrangian as well. $\square$

We now turn to the piecewise regular case. The local description of the switching flow obtained in the proof of Proposition 7.5.1 already contains the coordinate construction needed below. We shall therefore refer to that proof whenever the same hitting-time argument is involved, and only record the additional geometric information specific to the flow-out of $\mathcal{L}_0$.

Proposition 7.5.4 (Structure of $\mathcal{L}$ near a piecewise regular extremal). *Let $(\tilde{u}, \tilde{\lambda})$ be a piecewise regular extremal pair, with switching times $t_1, \dots, t_k$. Assume that*

$$\vec{H}^1(\tilde{\lambda}_0) \notin T_{\tilde{\lambda}_0} \mathcal{L}_0.$$

Then there exists a neighbourhood $\mathcal{O}_{\tilde{\lambda}} \subset T^*M$ of the image of $\tilde{\lambda}$ such that $\mathcal{L} \cap \mathcal{O}_{\tilde{\lambda}}$ is a d -dimensional Lipschitz immersed manifold. It admits a decomposition

$$\mathcal{L} \cap \mathcal{O}_{\tilde{\lambda}} = \bigcup_{j=1}^{k+1} \mathcal{L}_j,$$

where every $\mathcal{L}_j$ is a smooth Lagrangian submanifold with boundary and the interiors of the $\mathcal{L}_j$ are pairwise disjoint. Consecutive pieces meet along the smooth $(d-1)$ -dimensional isotropic submanifolds

$$\Gamma_j := \mathcal{L}_j \cap \mathcal{L}_{j+1} \subset \Sigma_j, \quad j = 1, \dots, k.$$

In the relative topology of $\mathcal{O}_{\tilde{\lambda}}$, their boundaries are

$$\begin{aligned} \partial \mathcal{L}_1 &= \Gamma_1, \\ \partial \mathcal{L}_j &= \Gamma_{j-1} \cup \Gamma_j, \quad j = 2, \dots, k, \\ \partial \mathcal{L}_{k+1} &= \Gamma_k. \end{aligned}$$

At every smooth point $\lambda \in \mathcal{L}_j$, one has

$$\vec{H}^j(\lambda) \in T_{\lambda} \mathcal{L}_j.$$

If $\lambda \in \Gamma_j$, then

$$\begin{aligned} T_{\lambda} \mathcal{L}_j &= T_{\lambda} \Gamma_j \oplus \mathbb{R} \vec{H}^j(\lambda), \\ T_{\lambda} \mathcal{L}_{j+1} &= T_{\lambda} \Gamma_j \oplus \mathbb{R} \vec{H}^{j+1}(\lambda), \end{aligned}$$

and consequently

$$T_{\lambda} \mathcal{L}_j \cap T_{\lambda} \mathcal{L}_{j+1} = T_{\lambda} \Gamma_j.$$

Finally, the generalized tangent family of $\mathcal{L}$ at $\lambda \in \Gamma_j$ is

$$T_{\lambda} \mathcal{L} = \bigcup_{\alpha \in [0,1]} \left(T_{\lambda} \Gamma_j + \mathbb{R} \left(\alpha \vec{H}^j(\lambda) + (1-\alpha) \vec{H}^{j+1}(\lambda) \right) \right). \quad (7.5.4)$$

In particular, both $\vec{H}^j(\lambda)$ and $\vec{H}^{j+1}(\lambda)$ are transverse to Γ_j .

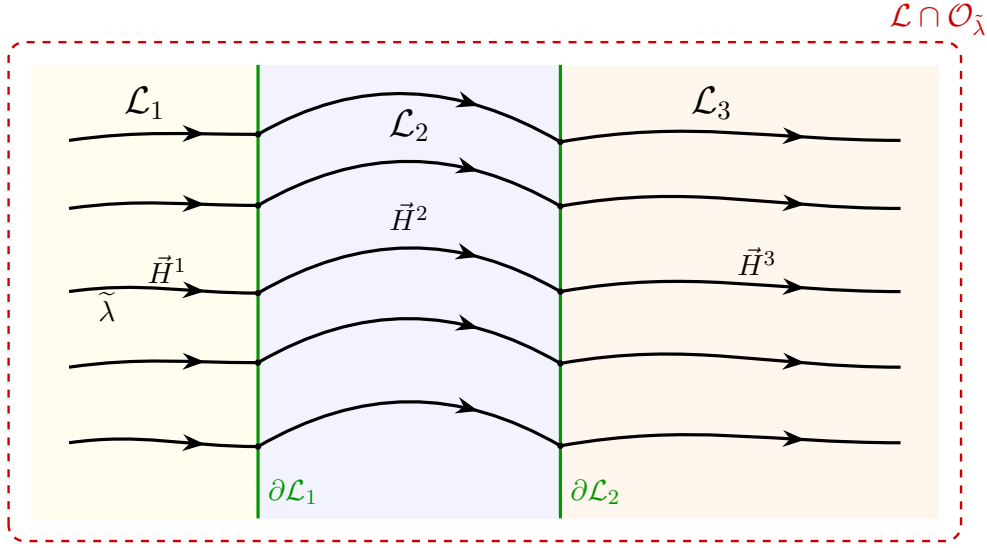


Figure 7.1: Graphical representation of Proposition 7.5.4.

Proof. Away from the switching hypersurfaces, the statement follows from Proposition 7.5.3, applied on the regular arcs of the reference extremal. It is therefore sufficient to describe the gluing near each $\tilde{\lambda}_{t_j}$.

Let $\mathcal{L}_j$ denote the smooth incoming Lagrangian piece. Since $\vec{H}^j$ is tangent to $\mathcal{L}_j$ and, by Proposition 7.5.1, transverse to Σ_j , the intersection

$$\Gamma_j = \mathcal{L}_j \cap \Sigma_j$$

is a smooth submanifold of $\mathcal{L}_j$ of codimension one. Hence $\dim \Gamma_j = d-1$, and Γ_j is isotropic because it is contained in the Lagrangian manifold $\mathcal{L}_j$.

The hitting-time construction in the proof of Proposition 7.5.1, restricted to the flow-out of $\mathcal{L}_0$, gives the local parametrization

$$\Theta_j : \Gamma_j \times (-\delta, \delta) \longrightarrow \mathcal{L}, \quad \Theta_j(\eta, s) = \begin{cases} \Phi_s^j(\eta), & s \leq 0, \\ \Phi_s^{j+1}(\eta), & s \geq 0. \end{cases} \quad (7.5.5)$$

As proved there, after reducing δ and the neighbourhood, Θ_j is bi-Lipschitz onto its image and is smooth on each half-cylinder. This proves the local Lipschitz manifold structure.

The transversality of $\vec{H}^j$ and $\vec{H}^{j+1}$ also yields

$$T_\lambda \Gamma_j = T_\lambda \mathcal{L}_j \cap T_\lambda \mathcal{L}_{j+1}$$

Finally, the same piecewise smooth parametrization (7.5.5) gives the the generalized tangent. Its generalized differential consists of the convex combinations of the two one-sided differentials in the transverse variable, while the differential in the Γ_j -directions is the

same on both sides. Definition A.2.8 therefore gives (7.5.4). Since there are finitely many switching times, the local constructions can be patched after choosing disjoint switching neighbourhoods, which completes the proof. $\square$

Remark 7.5.5 (Structure of fields of extremals). The proof of Proposition 7.5.4 uses only the piecewise Hamiltonian structure established in Proposition 7.5.1, the transversality of the switching hypersurfaces, and the isotropicity of the initial section contained in $H^{-1}(0)$. It does not use the fact that the initial section is contained in the vertical fiber $T_{q_0}^*M$. Consequently, if $\mathcal{M}_0$ is the codimension-one isotropic section of an exact Lagrangian graph introduced after Proposition 7.5.1, then its flow-out $\mathcal{M}$ has the same piecewise smooth Lagrangian structure and the same tangent-cone description as $\mathcal{L}$. This is the field of extremals to which the sufficient criterion will be applied in Section 7.6.

7.6 Sufficient condition for optimality

The aim of this section is to prove Theorem 6.1.8. We recall its statement for the reader's convenience.

Theorem 7.6.1. *Let $(\tilde{u}, \tilde{\lambda})$ be a piecewise regular normal extremal pair. Assume that there are no conjugate points along the trajectory $q(\cdot, \tilde{u})$. Then, the trajectory $q(\cdot, \tilde{u})$ is locally strongly strictly optimal.*

The proof is based on the sufficient optimality criterion developed in Chapter 4, see Theorem 4.1.3. According to this criterion, it is enough to embed the reference extremal in a field of extremals generated by a suitable super-Hamiltonian and whose projection onto the state space is locally invertible.

Most of this construction has already been carried out in Section 7.5. More precisely, Proposition 7.5.1 shows that the maximized Hamiltonian H generates a locally bi-Lipschitz piecewise smooth Hamiltonian flow Φ in a neighbourhood of the reference extremal. Moreover, Remark 7.5.5 shows that the flow-out of any admissible initial exact Lagrangian graph has the piecewise smooth Lagrangian structure required in Chapter 4. Consequently, the maximized Hamiltonian and the corresponding flow-out satisfy hypotheses **(H1)**–**(H6)** and **(H8)** of Theorem 4.1.3.

The only remaining point is therefore to choose the initial exact Lagrangian graph so that the projection of its flow-out satisfies hypothesis **(H7)**. We first show that this choice can be reduced to the choice of the tangent space of the initial exact graph. We then study the evolution of this tangent space and, by Theorem 4.1.6, we relate Maslov index of this curve to the invertibility of π restricted to the corresponding field of extremals.

The argument is presented for the free-final-time problem. The fixed-final-time case follows by applying the same construction to the extended free-time problem on $M \times \mathbb{R}$ introduced in Section 7.2.

Step 1: reduction to the choice of the initial tangent space.

As before, set

$$\Pi := T_{\tilde{\lambda}_0} (T_{q_0}^* M).$$

We have to find $a \in C^\infty(M)$ such that the corresponding fields of extremals satisfies the hypothesis **(H7)**. Again, to check the invertibility of the projection, we can apply the inverse function theorem, so that it suffices to find the good initial tangent space to the manifold.

More precisely, for fixed $\tilde{\lambda}_0 \in T_{q_0}^* M$ and $L_0 \in \mathfrak{L}(T_{\tilde{\lambda}_0}(T^*M))$, a function $a \in C^\infty(M)$ such that $a(q_0) = 0$ and

$$d_{q_0} a = \tilde{\lambda}_0, \quad \mathcal{L}_0 := \{d_q a \mid q \in M\}, \quad T_{\tilde{\lambda}_0} \mathcal{L}_0 = L_0.$$

is determined uniquely as an equivalence class of the 2-jet at q_0 . Thus, the problem of finding a and $\mathcal{L}_0$ is reduced to finding L_0 .

Since the reference trajectory is non-constant on its first regular arc, the Hamiltonian vector field generating this arc is not vertical:

$$\vec{\mathcal{H}}^1 \notin \Pi.$$

Hence, after taking L_0 sufficiently close to Π , we also have

$$\vec{\mathcal{H}}^1 \notin L_0. \tag{7.6.1}$$

Since L_0 is Lagrangian, Equation (7.6.1) implies that 0 is a regular value of the restriction of H to $\mathcal{L}_0$. Indeed, if

$$d_{\tilde{\lambda}_0} (H|_{\mathcal{L}_0}) = 0,$$

then

$$\sigma_{\tilde{\lambda}_0} (v, \vec{\mathcal{H}}^1) = 0, \quad v \in L_0,$$

and therefore

$$\vec{\mathcal{H}}^1 \in L_0^\angle = L_0,$$

which contradicts Equation (7.6.1).

Let $\mathcal{M}_0$ be a relatively open connected neighbourhood of $\tilde{\lambda}_0$ in

$$\mathcal{L}_0 \cap H^{-1}(0),$$

and let $\mathcal{M}$ be its flow-out under the maximized Hamiltonian flow. The tangent space of the initial section is

$$T_{\tilde{\lambda}_0} \mathcal{M}_0 = L_0 \cap (\vec{\mathcal{H}}^1)^\angle.$$

The tangent space of the full flow-out at the initial point is consequently

$$\mathfrak{M}_0 := T_{\tilde{\lambda}_0} \mathcal{M} = \left(L_0 \cap (\vec{\mathcal{H}}^1)^\angle \right) + \mathbb{R} \vec{\mathcal{H}}^1.$$

Thus, L_0 denotes the tangent space of the initial exact graph, whereas $\mathfrak{M}_0$ denotes the tangent space of the field of extremals obtained after imposing the zero-Hamiltonian constraint and flowing in the Hamiltonian direction.

By the results of Section 7.5, the field $\mathcal{M}$ constructed in this way satisfies all the assumptions of Theorem 4.1.3, except possibly hypothesis **(H7)**. It remains to choose L_0 so that the projection

$$\pi_{\mathcal{M}} := \pi|_{\mathcal{M}}$$

is locally invertible along the reference extremal.

Step 2: evolution of the tangent spaces.

We now describe the linear evolution of $\mathfrak{M}_0$. Recall that the switching times are

$$0 = t_0 < t_1 < \cdots < t_k < t_{k+1} = T.$$

For $t \in (t_{j-1}, t_j]$, let

$$B_t = \overrightarrow{\exp} \int_{t_{j-1}}^t \vec{b}_s ds$$

be the linear symplectic flow introduced in Section 7.4. The dependence of B_t on the regularity interval containing t will always be understood.

Starting from

$$\mathfrak{M}_{t_0}^+ := \mathfrak{M}_0 = L_0^{\vec{\mathcal{H}}^1},$$

we define recursively

$$\mathfrak{M}_t := B_t \mathfrak{M}_{t_{j-1}}^+, \quad t \in (t_{j-1}, t_j], \quad j = 1, \dots, k+1.$$

At every switching time t_j , the outgoing space is

$$\mathfrak{M}_{t_j}^+ := \mathfrak{M}_{t_j}^{\vec{\mathcal{H}}^{j+1}} = \left(\mathfrak{M}_{t_j} \cap \left(\vec{\mathcal{H}}^{j+1} \right)^{\angle} \right) + \mathbb{R} \vec{\mathcal{H}}^{j+1}.$$

The following regular arc starts from the outgoing space $\mathfrak{M}_{t_j}^+$.

At a regular time, we set

$$\mathcal{B}_t(L_0) := \{\mathfrak{M}_t\}.$$

At a switching time t_j , define

$$E_j := \mathfrak{M}_{t_j} \cap \mathfrak{M}_{t_j}^+.$$

The generalized tangent family at t_j is

$$\mathcal{B}_{t_j}(L_0) := \left\{ E_j + \mathbb{R} \left(\alpha \vec{\mathcal{H}}^j + (1 - \alpha) \vec{\mathcal{H}}^{j+1} \right) \mid \alpha \in [0, 1] \right\}.$$

Proposition 7.5.4 and Proposition 7.4.3 show that this recursion describes the transported-back tangent spaces of the field. At every regular time,

$$\left(\tilde{\Phi}_{0,t}^{-1}\right)_* T_{\tilde{\lambda}_t} \mathcal{M} = \mathfrak{M}_t.$$

At a switching time, the transported-back generalized tangent family is precisely $\mathcal{B}_{t_j}(L_0)$.

The same construction starting from $L_0 = \Pi$ gives the Jacobi curve. Indeed,

$$\mathfrak{M}_0 = \Pi^{\vec{\mathcal{H}}^1} = \Lambda_0^+,$$

and Theorem 7.4.1 yields

$$\mathcal{B}_t(\Pi) = \{\Lambda_t\}, \quad t \in (0, T] \setminus \{t_1, \dots, t_k\}.$$

At a switching time,

$$\mathcal{B}_{t_j}(\Pi) = \left\{ \left(\Lambda_{t_j} \cap \Lambda_{t_j}^+ \right) + \mathbb{R} \left(\alpha \vec{\mathcal{H}}^j + (1 - \alpha) \vec{\mathcal{H}}^{j+1} \right) \mid \alpha \in [0, 1] \right\}.$$

The recursive construction involves only finitely many continuous linear symplectic flows, extension operations, and compact families parametrized by $\alpha \in [0, 1]$. Consequently, the set-valued evolution $(t, L_0) \mapsto \mathcal{B}_t(L_0)$ has a closed graph in the region under consideration.

At a smooth point of $\mathcal{M}$, the differential of $\pi_{\mathcal{M}}$ is invertible if and only if the corresponding tangent space is transverse to the vertical space. At a switching point, the same condition must hold for every member of the generalized tangent family. After transporting everything back to $T_{\tilde{\lambda}_0}(T^*M)$, hypothesis **(H7)** is therefore reduced to

$$L \cap \Pi = \{0\}, \quad t \in [0, T], \quad L \in \mathcal{B}_t(L_0). \quad (7.6.2)$$

Thus, the construction of a suitable field of extremals is reduced to the finite-dimensional problem of choosing an initial Lagrangian subspace L_0 for which Equation (7.6.2) holds.

Step 3: the case without constant vertical solutions.

For $t \in (0, T]$, define the space of constant vertical directions by

$$C_t := \Pi \cap \bigcap_{0 < \tau \leq t} \Lambda_\tau.$$

The characterization of conjugate times given in Section 4.2 implies that the absence of conjugate points is equivalent to

$$L \cap \Pi = C_t, \quad t \in (0, T], \quad L \in \mathcal{B}_t(\Pi). \quad (7.6.3)$$

We first consider the case

$$C_t = \{0\}, \quad t \in (0, T].$$

In this case, Equation (7.6.3) shows that the Jacobi curve and all the intermediate spaces at the switching times are transverse to Π for every positive time. The obstruction occurs only at the initial time, since $\Lambda_0 = \Pi$.

We make the following claim, which is proved afterwards.

Claim 7.6.2. *There is $\varepsilon > 0$ such that for any neighbourhood $\mathcal{N} \subset \mathfrak{L}(T_{\lambda_0}^*(T^*M))$ of Π there is $L_0 \in \mathcal{N} \setminus \{\Pi\}$ such that*

$$L \cap \Pi = \{0\}, \quad t \in [0, \varepsilon), \quad L \in \mathcal{B}_t(L_0).$$

In particular, the claim implies that there exists $\varepsilon > 0$ and a sequence of Lagrangian subspaces

$$L_0^{(n)} \longrightarrow \Pi, \quad \text{as } n \rightarrow +\infty,$$

such that

$$L \cap \Pi = \{0\}, \quad t \in [0, \varepsilon), \quad L \in \mathcal{B}_t(L_0^{(n)}),$$

for every $n \in \mathbb{N}$.

We now show that, from the Claim and the assumption that there are no conjugate times and no non-zero constant vertical directions, the same transversality property holds on the whole interval for all sufficiently large n . Suppose, by contradiction, that this is false. After passing to a subsequence, there exist times

$$t_n \in [\varepsilon, T]$$

and Lagrangian subspaces

$$L_n \in \mathcal{B}_{t_n}(L_0^{(n)})$$

such that

$$L_n \cap \Pi \neq \{0\}. \tag{7.6.4}$$

By compactness of $[\varepsilon, T]$, after extracting a further subsequence we may assume that

$$t_n \longrightarrow t_\infty \in [\varepsilon, T].$$

The Lagrangian Grassmannian is compact, so we may also assume that

$$L_n \longrightarrow L_\infty.$$

Thus, by (7.6.4) we have that $L_\infty \cap \Pi \neq \{0\}$. The closed-graph property of the evolution $\mathcal{B}$ gives

$$L_\infty \in \mathcal{B}_{t_\infty}(\Pi).$$

This implies that t_∞ is a conjugate time, which contradicts our assumptions.

Theorem 4.1.6 therefore shows that every element of the generalized differential of $\pi_{\mathcal{M}}$ is an isomorphism along the reference extremal. Equivalently, $\pi_{\mathcal{M}}$ is a local diffeomorphism in a neighbourhood of the reference extremal.

Thus, in the case of no constant vertical solutions, the proof is reduced to the previous Claim.

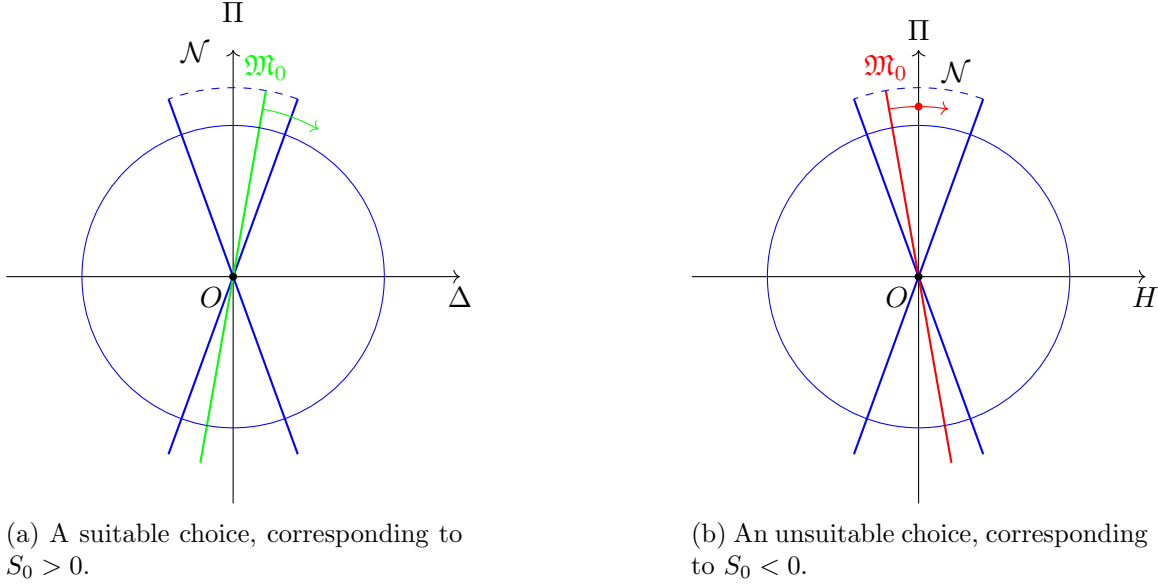


Figure 7.2: In dimension one, the positive side of the local chart gives an initial Lagrangian subspace whose evolution does not meet the vertical space for small positive times.

Proof of the Claim. To prove the claim, we use the local coordinate of $\mathfrak{L}(T_{\tilde{\lambda}_0}(T^*M))$. Fix a Lagrangian space Δ such that $\Pi \cap \Delta = \{0\}$ and use the corresponding local (Π, Δ) coordinates in $\Delta^\natural$, in which Π is represented by the zero symmetric matrix. We can choose Δ so that $\Pi^{\vec{H}^1} \in \Delta^\natural$. Since $\Delta^\natural$ is open, there is some ε such that also $B_t \Pi^{\vec{H}^1} \in \Delta^\natural$ for $t \in [0, \varepsilon]$.

Fix any neighbourhood $\mathcal{N}$ of Π . If $L \in \mathcal{N}$ and S is the symmetric matrix representing L in these coordinates, then, without loss of generality, we can assume that $\mathcal{N}$ is represented by the set of matrices $\{-\delta \text{Id} < S < \delta \text{Id}\}$.

Choose L_0 represented in this chart by a positive definite symmetric matrix

$$S_0 = \delta_0 \text{Id} > 0.$$

By Proposition 3.3.10, the curve $\mathcal{B}_t(L_0)$ is increasing. Thus, $\mathcal{B}_t(L_0)$ belongs to $\Delta^\natural$ for $t \in [0, \varepsilon)$, provided that δ_0 is small enough. Denoting by S_t its representing symmetric matrix, monotonicity gives

$$S_t \geq S_0 > 0. \quad (7.6.5)$$

Equation (7.6.5) implies that no member of $\mathcal{B}_t(L_0)$ intersects Π on this interval. This completes the proof of the Claim. $\square$

This completely proves Theorem 7.6.1 in the case $C_t = \{0\}$ for $t > 0$.

Step 4: reduction by the constant vertical directions.

We now consider the general case $C_t \neq \{0\}$ for $t > 0$. By definition, the family $(C_t)_{t>0}$ is monotone non-increasing:

$$C_{t_2} \subset C_{t_1}, \quad 0 < t_1 \leq t_2 \leq T.$$

Every strict inclusion decreases the dimension. Since the ambient space is finite-dimensional, the family $(C_t)_{t>0}$ takes only finitely many distinct values.

Choose times

$$0 = s_0 < s_1 < \cdots < s_m = T$$

such that C_t is constant on every interval $(s_{i-1}, s_i]$. We can decompose the vertical space as

$$\Pi = E_0 \oplus E_1 \oplus \cdots \oplus E_m$$

in such a way that

$$C_t = E_i \oplus \cdots \oplus E_m, \quad t \in (s_{i-1}, s_i], \quad i = 1, \dots, m.$$

Set

$$\Sigma_0 := T_{\tilde{\lambda}_0}(T^*M).$$

Choose a Lagrangian subspace

$$\Delta_0 \in \mathfrak{L}(\Sigma_0)$$

such that

$$\Delta_0 \cap \Pi = \{0\}.$$

We can decompose Δ_0 as

$$\Delta_0 = F_0 \oplus F_1 \oplus \cdots \oplus F_m$$

so that

$$\sigma_{\tilde{\lambda}_0}(E_i, F_j) = 0, \quad i \neq j,$$

and every $E_i \oplus F_i$ is a symplectic subspace.

For each i , choose an isomorphism

$$\mathfrak{I}_i: E_i \longrightarrow F_i$$

such that the bilinear form

$$(v, w) \longmapsto \sigma_{\tilde{\lambda}_0}(v, \mathfrak{I}_i w)$$

is symmetric and positive definite. We then set

$$\mathfrak{I} := \mathfrak{I}_0 \oplus \cdots \oplus \mathfrak{I}_m: \Pi \longrightarrow \Delta_0.$$

In particular, the graph of $\varepsilon \mathfrak{I}$ is a Lagrangian subspace for every $\varepsilon \in \mathbb{R}$ and lies on the positive side of the chart (Π, Δ_0) when $\varepsilon > 0$.

The following lemma constructs the initial subspace iteratively. At each step, the constant vertical directions are removed by symplectic reduction, so that the argument of Step 3 can be applied in the quotient.

Lemma 7.6.3. *There exist Lagrangian subspaces*

$$\Delta_1, \dots, \Delta_m \in \mathfrak{L}(\Sigma_0)$$

transverse to Π and positive numbers

$$\varepsilon_1 \geq \dots \geq \varepsilon_m > 0$$

with the following property. For $0 < \varepsilon < \varepsilon_i$, let $L_0(\varepsilon)$ be the Lagrangian subspace represented by $\varepsilon\mathfrak{J}$ in the chart (Π, Δ_0) . Then:

1. *every $L \in \mathcal{B}_t(L_0(\varepsilon))$ satisfies*

$$L \cap \Pi = \{0\}, \quad t \in [0, s_i];$$

2. *every $L \in \mathcal{B}_t(L_0(\varepsilon))$ is transverse to Δ_i for $t \in [0, s_i]$, and its representing symmetric form in the chart (Π, Δ_i) is positive definite.*

Proof. We argue by induction on i .

In the computations below, B_t denotes the linear symplectic propagator on a regular arc or on one of the smooth monotone curves inserted at a switching time. Every member of the generalized family $\mathcal{B}_t(L_0(\varepsilon))$ is reached by one of these propagators. Consequently, the argument applies both to the regular tangent spaces and to all the intermediate spaces at the switching times.

The case $i = 0$ corresponds to the initial choice made in the proof of the Claim. After reducing the upper bound for ε if necessary, the initial spaces are transverse to both Π and Δ_0 , and their representing forms in the chart (Π, Δ_0) are positive definite.

Assume now that the conclusion holds up to time s_i for some $i \in \{0, \dots, m-1\}$. We prove it on the larger interval $[0, s_{i+1}]$.

For $t \in (s_i, s_{i+1}]$, the space of constant vertical directions is

$$C_t = E_{i+1} \oplus \dots \oplus E_m.$$

Set

$$\begin{aligned} E'_1 &:= E_0 \oplus \dots \oplus E_i, \\ E'_2 &:= C_t = E_{i+1} \oplus \dots \oplus E_m. \end{aligned}$$

Choose a decomposition

$$\Delta_i = F'_1 \oplus F'_2$$

such that $E'_1 \oplus F'_1$ and $E'_2 \oplus F'_2$ are symplectic and mutually symplectically orthogonal.

Fix $0 < \varepsilon < \varepsilon_i$ and denote by

$$\mathfrak{M}_0(\varepsilon)$$

the unique member of $\mathcal{B}_0(L_0(\varepsilon))$. Since the vectors in E'_2 are constant vertical solutions on $(s_i, s_{i+1}]$, every propagator occurring on this interval fixes E'_2 :

$$B_t E'_2 = E'_2.$$

It follows that its symplectic orthogonal is invariant as well:

$$B_t (E'_2)^\perp = (E'_2)^\perp.$$

With the preceding decomposition,

$$(E'_2)^\perp = E'_1 \oplus E'_2 \oplus F'_1.$$

Consider the symplectic quotient

$$\tilde{\Sigma}_i := (E'_2)^\perp / E'_2$$

and the reduced vertical space

$$\tilde{\Pi}_i := \Pi / E'_2.$$

The invariance relations above induce a symplectic evolution on $\tilde{\Sigma}_i$.

The reduction of the initial tangent space is

$$\tilde{\mathfrak{M}}_0(\varepsilon) := (\mathfrak{M}_0(\varepsilon) \cap (E'_2)^\perp) / E'_2.$$

More generally, if $L \in \mathcal{B}_t(L_0(\varepsilon))$, its reduction is

$$\tilde{L} := (L \cap (E'_2)^\perp) / E'_2.$$

By Equation (7.6.3), every member of the reference evolution satisfies

$$L \cap \Pi = E'_2, \quad t \in (s_i, s_{i+1}], \quad L \in \mathcal{B}_t(\Pi).$$

Therefore, after passing to the quotient,

$$\tilde{L} \cap \tilde{\Pi}_i = \{0\}, \quad t \in (s_i, s_{i+1}].$$

Thus, the reduced evolution has no non-zero constant vertical directions on this interval.

We can now apply in $\tilde{\Sigma}_i$ the sequential argument used in Step 3. After decreasing ε_i to a number

$$0 < \varepsilon_{i+1} \leq \varepsilon_i,$$

we obtain

$$\tilde{L} \cap \tilde{\Pi}_i = \{0\}, \quad t \in (s_i, s_{i+1}], \quad L \in \mathcal{B}_t(L_0(\varepsilon)), \quad 0 < \varepsilon < \varepsilon_{i+1}. \quad (7.6.6)$$

We lift Equation (7.6.6) to Σ_0 . Fix a member

$$L = B_t \mathfrak{M}_0(\varepsilon) \in \mathcal{B}_t(L_0(\varepsilon)).$$

Since $\Pi \subset (E'_2)^\perp$, we have

$$\begin{aligned} L \cap \Pi &= B_t \mathfrak{M}_0(\varepsilon) \cap B_t (E'_2)^\perp \cap \Pi \\ &= B_t \left(\mathfrak{M}_0(\varepsilon) \cap (E'_2)^\perp \right) \cap \Pi. \end{aligned}$$

Equation (7.6.6) implies that

$$L \cap \Pi \subset E'_2. \quad (7.6.7)$$

On the other hand, the induction hypothesis gives

$$\mathfrak{M}_0(\varepsilon) \cap E'_2 = \{0\},$$

because $E'_2 \subset \Pi$. Using the invariance of E'_2 , we obtain

$$\begin{aligned} L \cap E'_2 &= B_t \mathfrak{M}_0(\varepsilon) \cap B_t E'_2 \\ &= B_t (\mathfrak{M}_0(\varepsilon) \cap E'_2) \\ &= \{0\}. \end{aligned} \quad (7.6.8)$$

Combining Equations (7.6.7) and (7.6.8), we conclude that

$$L \cap \Pi = \{0\}, \quad t \in (s_i, s_{i+1}], \quad L \in \mathcal{B}_t(L_0(\varepsilon)).$$

Together with the induction hypothesis, this proves the first part of the lemma on $[0, s_{i+1}]$.

It remains to prove the second part. The family

$$\{L \mid L \in \mathcal{B}_t(L_0(\varepsilon)), \quad t \in [0, s_{i+1}], \quad 0 \leq \varepsilon \leq \varepsilon_{i+1}\}$$

is compact and every one of its members is transverse to Π . Consequently, we can choose a Lagrangian subspace Δ_{i+1} sufficiently close to Π which is transverse to every member of this family.

In the chart (Π, Δ_i) , we choose Δ_{i+1} to be represented by the negative definite form

$$-\delta \text{Id}, \quad \delta > 0.$$

Let S be the positive definite symmetric matrix representing one of the spaces under consideration in the chart (Π, Δ_i) . The same space is represented in the chart (Π, Δ_{i+1}) by

$$S' = (S^{-1} + \delta^{-1} \text{Id})^{-1} = \delta S (\delta \text{Id} + S)^{-1}.$$

In particular,

$$S' > 0.$$

In the special case $S = \varepsilon \text{Id}$, this formula becomes

$$S' = \frac{\varepsilon}{1 + \varepsilon/\delta} \text{Id} > 0.$$

Every regular arc and every connecting curve is increasing. Hence, as long as the curve remains in the chart (Π, Δ_{i+1}) , its representing matrix satisfies

$$\dot{S}_t \geq 0.$$

Since Δ_{i+1} is transverse to the whole family, the chart is valid on $[0, s_{i+1}]$. Positivity at the beginning of each new arc therefore implies

$$S_t > 0, \quad t \in [0, s_{i+1}].$$

This proves the second part of the lemma and completes the induction. $\square$

Apply Lemma 7.6.3 with $i = m$ and choose

$$0 < \varepsilon < \varepsilon_m.$$

The corresponding initial Lagrangian subspace $L_0(\varepsilon)$ satisfies

$$L \cap \Pi = \{0\}, \quad t \in [0, T], \quad L \in \mathcal{B}_t(L_0(\varepsilon)).$$

Since ε can be chosen arbitrarily small, $L_0(\varepsilon)$ can be taken in any prescribed sufficiently small neighbourhood of Π .

The field of extremals constructed from $L_0(\varepsilon)$ therefore has no non-zero common vertical directions and its curve of transported-back generalized tangent spaces has zero Maslov index. By Theorem 4.1.6, the projection $\pi_{\mathcal{M}}$ is locally bi-Lipschitz in a neighbourhood of the reference extremal and hypothesis **(H7)** is satisfied.

All the remaining hypotheses of Theorem 4.1.3 were verified in Section 7.5. The transversality of the Hamiltonian vector fields to the switching hypersurfaces gives the strictness condition at the switching points. It follows that $\tilde{q}$ is locally strongly strictly optimal.

The fixed-final-time case follows by applying the same construction to the extended free-time problem introduced in Section 7.2. This fully completes the proof of Theorem 7.6.1.

Chapter 8

Second-order optimality conditions for nice extremals

This chapter is a further step in the second-order optimality analysis developed in the preceding chapters. Our aim is to extend the theory of piecewise regular extremals introduced in Chapters 6 and 7 to trajectories containing singular arcs, combining the regular and singular constructions developed separately in the previous parts of the thesis.

We keep the general setting and notation introduced in Chapter 6. Recall that a piecewise regular extremal $\tilde{\lambda} : [0, T] \rightarrow T^*M$ admits a finite partition

$$0 = t_0 < t_1 < \cdots < t_k < t_{k+1} = T$$

such that, on every interval (t_j, t_{j+1}) , the Pontryagin maximization condition determines a unique smooth maximizing submanifold and the strong Legendre condition holds. The corresponding maximized Hamiltonian vector field is therefore smooth along each interval, while having jumps at isolated switching times. In Chapters 6 and 7, we described the Jacobi curve by alternating its smooth Hamiltonian evolution with explicit discrete dynamics at the switching times and derived both necessary and sufficient optimality conditions.

Important classes of optimal trajectories fall outside this framework because the degeneracy of the maximization condition may occur on a whole time interval rather than at isolated switching times. This is not merely a technical possibility. Already in simple problems, such as the Dubins car, singular arcs are unavoidable for certain endpoint configurations: in particular, connecting points that are sufficiently far apart requires optimal trajectories containing a singular segment. Therefore, any general theory must be able to accommodate such trajectories. Second-order analysis of trajectories containing singular arcs has been the subject of a substantial literature; see, among others, [70, 64, 61, 63, 29, 30, 34, 19].

The singularities considered in this chapter are the simplest possible degeneracies arising within the framework of Chapter 6. Roughly speaking, along a *first-order singular arc*

exactly two smooth Hamiltonians among the H^j that are simultaneously maximizing. The Pontryagin Maximum Principle therefore no longer determines the maximizing dynamics completely. This is the main difference with the piecewise regular setting, where the maximizer is uniquely determined away from isolated switching times.

The chapter is conceptually divided into two parts. The first part is devoted to the second-order analysis of a single first-order singular arc. It extends the approach developed in Chapter 5 for the singular L^1 problem. The main novelty is that we no longer rely on the particular control-affine and L^1 structure considered there. Within this setting, we derive the corresponding second-order necessary conditions, derive the Jacobi equation, and introduce a suitable super-Hamiltonian whose linearization reproduces this dynamics. Thus, this first part provides the singular analogue of the constructions used for regular arcs in the preceding chapters.

The second part combines the analysis of singular extremal arcs with the piecewise regular theory of Chapters 6 and 7. We call *nice extremals* the extremals obtained by concatenating finitely many regular arcs and first-order singular arcs.

We first characterize the Jacobi curve of a nice extremal by combining the piecewise regular evolution with the evolution along first-order singular arcs. Then, by applying Theorem 3.0.1, we obtain the corresponding second-order necessary optimality condition. We finally turn to sufficient optimality, where the new difficulty is the construction of a field of extremals through regular–singular junctions. Following the approach used in [61], we construct an appropriate super-Hamiltonian near these junctions and combine it with the regular and singular constructions developed earlier. For nice extremals containing at most one singular arc, this allows us to prove that the absence of conjugate points implies strong local optimality. We believe that the restriction to one singular arc is technical, and we expect the same result to hold for nice extremals containing an arbitrary finite number of singular arcs. This final part is still a work in progress.

8.1 Analysis of first-order singular extremals

In this section, we introduce the notions of first-order singular extremal arcs and of nice extremals, which are the main focus of this chapter. In particular, we derive a feedback formula for controls corresponding to first-order singular extremal arcs.

The construction developed here combines the two classes of extremals studied in the preceding chapters. In Chapter 5, we considered singular extremal arcs for which the Pontryagin Maximum Principle does not determine the control and a feedback formula is recovered by differentiating the singularity relations. In Chapters 6 and 7, instead, we considered piecewise regular extremals for which a unique maximizing Hamiltonian governs each open arc, while changes of maximizing Hamiltonian occur at isolated switching times. The extremals studied in the present chapter may contain both regular arcs and intervals along which two different Hamiltonians are simultaneously maximizing. On these latter

intervals, the relative weight of the two maximizing Hamiltonians plays the role of the singular control and is determined through a differentiated switching relation, similarly to the singular feedback of Chapter 5.

First, we recall our setting. Let $U \subset \mathbb{R}^m$ be the disjoint union of finitely many smooth submanifolds without boundary,

$$U = \bigcup_{i=1}^n U_i, \quad U_i \cap U_j = \emptyset \quad \text{for } i \neq j.$$

On a smooth manifold M , we consider the control system

$$\dot{q}(t) = f(q(t), u(t)), \quad (8.1.1)$$

with boundary conditions

$$q(0) = q_0, \quad q(T) = q_1, \quad (8.1.2)$$

where $q_0, q_1 \in M$, $u \in \mathfrak{U}_T := L^\infty([0, T]; U)$, and $f : M \times U \rightarrow TM$ satisfies the regularity assumptions of Section 6.1.1. As in Chapters 6 and 7, the final time may be either fixed or free. We write $\mathfrak{U} = \bigcup_{T>0} \mathfrak{U}_T$ and denote by $q(\cdot; q_0, u)$, or simply by $q(\cdot; u)$, the trajectory corresponding to u .

Let $L : M \times U \rightarrow \mathbb{R}$ satisfy the assumptions of Section 6.1.1 on every $M \times U_i$. The cost functional is

$$J(u) = \int_0^T L(q(t; u), u(t)) dt.$$

We consider both the free- and the fixed-final-time formulations. The optimal control problem with free final time is

$$\min \{J(u) \mid u \in \mathfrak{U}, \quad q(\cdot; u) \text{ satisfies (8.1.1) and (8.1.2)}\}. \quad (8.1.3)$$

For a fixed final time $T > 0$, the corresponding problem is

$$\min \{J(u) \mid u \in \mathfrak{U}_T, \quad q(\cdot; u) \text{ satisfies (8.1.1) and (8.1.2)}\}. \quad (8.1.4)$$

Under the preceding assumptions, the set of admissible velocity–cost pairs need not be convex. As a consequence, the minimum in the OCPs (8.1.3) and (8.1.4) does not exist in general. For this reason, we now introduce a relaxation of the previous problem, for which existence of the minimum is guaranteed by Filippov’s theorem.

For every $q \in M$, define the extended set of admissible velocities at q :

$$\mathcal{A}(q) := \{(f(q, u), L(q, u)) \mid u \in U\} \subset T_q M \times \mathbb{R}.$$

The appropriate relaxation of the integral-cost problem is obtained by replacing $\mathcal{A}(q)$ with its convex hull.

Set

$$d := \dim M, \quad N := d + 2,$$

and introduce the simplex

$$\Delta_N := \left\{ \mathbf{v} = (v_1, \dots, v_N) \in [0, 1]^N \mid \sum_{l=1}^N v_l = 1 \right\}.$$

By Carathéodory's theorem, every element of $\text{co } \mathcal{A}(q)$ can be represented as

$$\sum_{l=1}^N v_l (f(q, u_l), L(q, u_l)), \quad \mathbf{v} \in \Delta_N, \quad u_l \in U.$$

For $T > 0$, define the class of relaxed controls

$$\mathfrak{U}_T^{\text{rel}} := L^\infty([0, T]; U^N \times \Delta_N), \quad \mathfrak{U}^{\text{rel}} := \bigcup_{T>0} \mathfrak{U}_T^{\text{rel}}.$$

We write a relaxed control as

$$\mathbf{u} = (u_1, \dots, u_N), \quad \mathbf{v} = (v_1, \dots, v_N).$$

The relaxed dynamics is

$$\dot{q}(t) = \sum_{l=1}^N v_l(t) f(q(t), u_l(t)), \quad (8.1.5)$$

with boundary conditions

$$q(0) = q_0, \quad q(T) = q_1. \quad (8.1.6)$$

The corresponding relaxed cost is

$$\bar{J}(\mathbf{u}, \mathbf{v}) := \int_0^T \sum_{l=1}^N v_l(t) L(q(t), u_l(t)) dt. \quad (8.1.7)$$

Then, the relaxed problem with free final time is

$$\min \left\{ \bar{J}(\mathbf{u}, \mathbf{v}) \mid (\mathbf{u}, \mathbf{v}) \in \mathfrak{U}^{\text{rel}}, \quad q \text{ satisfies (8.1.5) and (8.1.6)} \right\}. \quad (8.1.8)$$

For a fixed final time $T > 0$, we consider

$$\min \left\{ \bar{J}(\mathbf{u}, \mathbf{v}) \mid (\mathbf{u}, \mathbf{v}) \in \mathfrak{U}_T^{\text{rel}}, \quad q \text{ satisfies (8.1.5) and (8.1.6)} \right\}. \quad (8.1.9)$$

Admissible trajectories for the original problems are contained in the relaxed ones by choosing one of the v_i equal to 1 and all the other equal to 0. Under the standard

compactness, regularity, and completeness assumptions ensuring existence for convex optimal control problems, the relaxed problems admit minimizers. Moreover, by a standard argument, see for instance Theorem 8.2 in [13], any admissible trajectory for the relaxed control system (8.1.5) can be uniformly approximated by an admissible trajectory of the original control system (8.1.1).

Recall the definition of the normal Pontryagin Hamiltonian

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

Recall also the notation

$$H^i(\lambda) = \max_{u \in U_i} h_u(\lambda), \quad H(\lambda) = \max_{u \in U} h_u(\lambda).$$

As in the preceding chapters, we restrict the discussion to normal extremals. The normal Hamiltonian of the relaxed problem is

$$\bar{h}_{\mathbf{u}, \mathbf{v}}(\lambda) = \sum_{l=1}^N v_l h_{u_l}(\lambda).$$

Since this Hamiltonian is affine in the $\mathbf{v}$, we have

$$\max_{\substack{\mathbf{u} \in U^N \\ \mathbf{v} \in \Delta_N}} \bar{h}_{\mathbf{u}, \mathbf{v}}(\lambda) = \max_{u \in U} h_u(\lambda) = H(\lambda).$$

Consequently, if $(\mathbf{u}, \mathbf{v})$ is maximizing at λ , then

$$v_l > 0 \implies h_{u_l}(\lambda) = H(\lambda).$$

In other words, every control appearing with positive weight must be a maximizing control.

We now restrict the analysis to relaxed extremals for which exactly two different Hamiltonians among the H^i are simultaneously maximizing. More precisely, assume that along a reference extremal trajectory $\tilde{\lambda}$ there exist two distinct indices $j_0, j_1 \in \{1, \dots, n\}$ such that $H(\tilde{\lambda}_t) = H^{j_0}(\tilde{\lambda}_t) = H^{j_1}(\tilde{\lambda}_t)$. After possibly a relabelling, the relaxed dynamics and cost take the form

$$\dot{q}(t) = v(t)f(q(t), u_0(t)) + (1 - v(t))f(q(t), u_1(t)), \quad (8.1.10)$$

and

$$\bar{L}(q, u_0, v, u_1) = vL(q, u_0) + (1 - v)L(q, u_1), \quad (8.1.11)$$

where $v(t) \in [0, 1]$.

Whenever the strong Legendre condition holds

$$\frac{\partial h_u^{j_i}}{\partial u^2}(\tilde{\lambda}_t) < 0,$$

for $i = 0, 1$, the implicit function theorem locally determines the maximizing control as a smooth function $\hat{u}_i(\lambda)$, and the corresponding function $H^i(\lambda) = h_{\hat{u}_i(\lambda)}^i(\lambda)$ is smooth. We introduce the notation

$$H^{ij} := \{H^i, H^j\}, \quad H^{ijk} := \{H^i, H^{jk}\}.$$

Definition 8.1.1 (First-order singular extremal arc). Let $(\tilde{u}_0, \tilde{v}, \tilde{u}_1, \tilde{\lambda})$ be a normal extremal of either the relaxed free-final-time problem (8.1.8) or the relaxed fixed-final-time problem (8.1.9), written in the reduced form (8.1.10). We say that $\tilde{\lambda}|_{[\tau_1, \tau_2]}$ is a first-order singular extremal arc if there exist two distinct indices $j_0, j_1 \in \{1, \dots, n\}$ such that

$$\tilde{u}_0(t) \in U_{j_0}, \quad \tilde{u}_1(t) \in U_{j_1},$$

for almost every $t \in [\tau_1, \tau_2]$, and the following conditions hold.

1. The Hamiltonians H^{j_0} and H^{j_1} are simultaneously maximizing along the arc:

$$H^{j_0}(\tilde{\lambda}_t) = H^{j_1}(\tilde{\lambda}_t) = H(\tilde{\lambda}_t), \quad t \in [\tau_1, \tau_2].$$

Moreover, they are the only maximizing Hamiltonians for $t \in [\tau_1, \tau_2]$:

$$H^{j_0}(\tilde{\lambda}_t) = H^{j_1}(\tilde{\lambda}_t) > \max_{k \notin \{j_0, j_1\}} H^k(\tilde{\lambda}_t).$$

2. The strong Legendre condition holds on both U_{j_0} and U_{j_1} :

$$\left. \frac{\partial^2 h_{\tilde{u}}^{j_i}}{\partial u^2} \right|_{u=\tilde{u}_i(t), \lambda=\tilde{\lambda}_t} < 0, \quad i = 0, 1,$$

for almost every $t \in (\tau_1, \tau_2)$, where

$$\tilde{u}_i(t) = \hat{u}_{j_i}(\tilde{\lambda}_t).$$

3. The first-order non-degeneracy conditions

$$H^{j_0 j_0 j_1}(\tilde{\lambda}_t) < 0, \quad H^{j_1 j_1 j_0}(\tilde{\lambda}_t) < 0 \tag{8.1.12}$$

hold for every $t \in [\tau_1, \tau_2]$.

The first condition describes the failure of the maximization condition to select a unique Hamiltonian evolution. We say that the two functions H^{j_0} and H^{j_1} are the *active* along the extremal $\tilde{\lambda}$. Moreover, we call the points $\tilde{\lambda}_{\tau_0}$ and $\tilde{\lambda}_{\tau_1}$ the incoming and, respectively, outgoing regular-singular junctions of the extremal $\tilde{\lambda}$. The strong Legendre condition nevertheless determines uniquely the maximizing control inside each of the two submanifold

U_{j_0} and U_{j_1} . The only variable that remains undetermined by the Pontryagin Maximum Principle is therefore $\tilde{v}$. The last condition allows us to recover the value of $\tilde{v}$ by differentiating the equality of the two maximizing Hamiltonians. We now derive this feedback and the Hamiltonian dynamics generated by it. To simplify the notation, we consider a purely first-order singular arc on $[0, T]$ and relabel the two maximizing Hamiltonians as H^0 and H^1 . Thus,

$$H^0(\tilde{\lambda}_t) = H^1(\tilde{\lambda}_t) = H(\tilde{\lambda}_t), \quad t \in [0, T].$$

Proposition 8.1.2 (Singular feedback). *Let $\tilde{\lambda}$ be a first-order singular extremal arc associated with the two maximizing Hamiltonians H^0 and H^1 . Then*

$$H^{01}(\tilde{\lambda}_t) = 0, \quad t \in [0, T]. \quad (8.1.13)$$

Moreover, $\tilde{v}$ is uniquely determined by

$$\tilde{v}(t) = \frac{H^{110}(\tilde{\lambda}_t)}{H^{001}(\tilde{\lambda}_t) + H^{110}(\tilde{\lambda}_t)}, \quad t \in [0, T], \quad (8.1.14)$$

and satisfies $\tilde{v}(t) \in (0, 1)$.

Proof. Along the relaxed extremal, the maximization with respect to u_0 and u_1 gives

$$\dot{\tilde{\lambda}}_t = \tilde{v}(t)\tilde{H}^0(\tilde{\lambda}_t) + (1 - \tilde{v}(t))\tilde{H}^1(\tilde{\lambda}_t).$$

Since $H^0(\tilde{\lambda}_t) - H^1(\tilde{\lambda}_t) = 0$, differentiating with respect to t yields

$$\begin{aligned} 0 &= \frac{d}{dt} (H^0 - H^1)(\tilde{\lambda}_t) \\ &= \tilde{v}(t)\{H^0, H^0 - H^1\}(\tilde{\lambda}_t) + (1 - \tilde{v}(t))\{H^1, H^0 - H^1\}(\tilde{\lambda}_t) \\ &= -H^{01}(\tilde{\lambda}_t). \end{aligned}$$

This proves (8.1.13). The first differentiated relation is again at least Lipschitz in t since it does not contain $\tilde{v}(t)$, so we differentiate it once more and obtain

$$\begin{aligned} 0 &= \frac{d}{dt} H^{01}(\tilde{\lambda}_t) \\ &= \tilde{v}(t)H^{001}(\tilde{\lambda}_t) - (1 - \tilde{v}(t))H^{110}(\tilde{\lambda}_t) \\ &= \tilde{v}(t) \left(H^{001}(\tilde{\lambda}_t) + H^{110}(\tilde{\lambda}_t) \right) - H^{110}(\tilde{\lambda}_t). \end{aligned}$$

By (8.1.12), the coefficient $\left(H^{001}(\tilde{\lambda}_t) + H^{110}(\tilde{\lambda}_t) \right)$ is strictly negative, and (8.1.14) follows. The same inequalities also give

$$\tilde{v}(t) = \frac{H^{110}(\tilde{\lambda}_t)}{H^{001}(\tilde{\lambda}_t) + H^{110}(\tilde{\lambda}_t)} > 0, \quad 1 - \tilde{v}(t) = \frac{H^{001}(\tilde{\lambda}_t)}{H^{001}(\tilde{\lambda}_t) + H^{110}(\tilde{\lambda}_t)} > 0.$$

□

The preceding calculation explains the terminology first-order singular. The equality $H^0 = H^1$ expresses the fact that the PMP Hamiltonian is constant with respect to the control v , its first derivative gives the additional control-independent relation $H^{01} = 0$, and its second derivative determines the control v . Geometrically, Equation (8.1.13) implies that both $\vec{H}^0$ and $\vec{H}^1$ are tangent, along the reference arc, to the level set $\{H^0 = H^1\}$. Moreover, the strict inequalities in (8.1.12) imply that the two vector fields are transverse to the level set $\{H^{01} = 0\}$ in opposite directions.

The Hamiltonian flow used in the subsequent second-variation computations is the flow generated by the reference control. More precisely, we consider the Hamiltonian function (denote $\tilde{u} = (\tilde{u}_0, \tilde{u}_1)$ for brevity)

$$h_{(\tilde{u}(t), \tilde{v}(t))}(\lambda) := \tilde{v}(t)h_{\tilde{u}_0(t)}^0(\lambda) + (1 - \tilde{v}(t))h_{\tilde{u}_1(t)}^1(\lambda).$$

Then, a reference first-order singular extremal satisfies

$$\dot{\tilde{\lambda}}_t = \vec{h}_{(\tilde{u}(t), \tilde{v}(t))}(\tilde{\lambda}_t), \quad \tilde{v}(t) = \frac{H^{110}(\tilde{\lambda}_t)}{H^{001}(\tilde{\lambda}_t) + H^{110}(\tilde{\lambda}_t)}. \quad (8.1.15)$$

We denote by $\tilde{\Phi}_{s,t}$ the Hamiltonian flow generated by Equation (8.1.15). This is the reference Hamiltonian flow that will be used to transport the first and second variations back to $T_{\tilde{\lambda}_0}(T^*M)$.

The feedback formula also defines an autonomous Hamiltonian dynamics that preserves the singular locus. By continuity, after shrinking a neighbourhood $\mathcal{O}$ of the reference arc, we may assume that

$$H^{001}(\lambda) < 0, \quad H^{110}(\lambda) < 0, \quad \lambda \in \mathcal{O}. \quad (8.1.16)$$

On $\mathcal{O}$, define the smooth singular feedback

$$v_s(\lambda) := \frac{H^{110}(\lambda)}{H^{001}(\lambda) + H^{110}(\lambda)}$$

and the singular feedback Hamiltonian

$$H_s(\lambda) := v_s(\lambda)H^0(\lambda) + (1 - v_s(\lambda))H^1(\lambda).$$

Let

$$\Sigma := \{\lambda \in \mathcal{O} \mid H^0(\lambda) = H^1(\lambda)\}, \quad \mathcal{S} := \{\lambda \in \mathcal{O} \mid H^{01}(\lambda) = 0\}.$$

Since

$$H_s = H^1 + v_s(H^0 - H^1),$$

its Hamiltonian vector field satisfies

$$\vec{H}_s(\lambda) = v_s(\lambda)\vec{H}^0(\lambda) + (1 - v_s(\lambda))\vec{H}^1(\lambda), \quad \lambda \in \Sigma.$$

In particular, the reference extremal is also an integral curve of $\vec{H}_s$. A direct computation gives

$$\vec{H}_s(H^0 - H^1) = -H^{01} \quad \text{on } \Sigma,$$

and

$$\vec{H}_s(H^{01}) = v_s H^{001} - (1 - v_s) H^{110} = 0 \quad \text{on } \mathcal{S} \cap \Sigma.$$

Therefore, $\Sigma \cap \mathcal{S}$ is invariant under the Hamiltonian flow of H_s . Moreover, by (8.1.16), the differentials of $H^0 - H^1$ and H^{01} are independent along the reference arc, then $\Sigma \cap \mathcal{S}$ is a codimension-two submanifold of T^*M near that arc.

Example 8.1.3 (The Dubins car). A classical example illustrating the preceding construction is provided by the Dubins car; see, for instance, [13, Section 13.5]. On the state space $M = \mathbb{R}^2 \times S^1$, $q = (x, y, \theta)$, consider the time-optimal control system

$$\begin{aligned} \dot{x} &= \cos \theta, \\ \dot{y} &= \sin \theta, \quad u \in U_0 \cup U_1, \quad U_0 = \{-1\}, \quad U_1 = \{1\}. \\ \dot{\theta} &= u, \end{aligned}$$

The relaxation of this bang-bang control system gives the model with angular velocity in $[-1, 1]$. In the reference, it is shown that every trajectory connecting any two points sufficiently far apart from each other is obtained by concatenating a circular arc, a straight segment, and another circular arc is a concrete example of a nice extremal:

$$\text{regular arc} - \text{first-order singular arc} - \text{regular arc}.$$

Thus, this shows that even in very simple control problems first-order singular trajectories arises naturally.

Definition 8.1.4 (Nice extremal). A normal extremal $\tilde{\lambda} : [0, T] \rightarrow T^*M$ of either the relaxed free-final-time problem (8.1.8) or the relaxed fixed-final-time problem (8.1.9) is called *nice* if there exists a finite partition

$$0 = t_0 < t_1 < \cdots < t_k < t_{k+1} = T$$

such that every restriction

$$\tilde{\lambda}|_{(t_j, t_{j+1})}$$

is either a regular extremal arc satisfying the strong Legendre condition or a first-order singular extremal arc in the sense of Definition 8.1.1.

Remark 8.1.5. We make explicit a consequence of the preceding definition. Suppose, for instance, that a regular arc associated with a maximized Hamiltonian H^0 is followed at time t_j by a first-order singular arc, and relabel the second active Hamiltonian on the

singular arc as H^1 . The switching from the regular to the singular arc occurs when the regular extremal reaches the equality set

$$\Sigma = \{H^0 = H^1\}.$$

Moreover, by Proposition 8.1.2, at the initial point of the singular arc one also has

$$H^{01}(\tilde{\lambda}_{t_j}) = 0.$$

Hence the regular Hamiltonian H^0 is one of the two Hamiltonian that generates the singular arc. In particular, instead of crossing the hypersurface Σ through an ordinary regular switching, the extremal can remain tangent to Σ and continue according to the singular control determined above, which combines the Hamiltonian vector fields $\vec{H}^0$ and $\vec{H}^1$. The analogous statement holds when a first-order singular arc is followed by a regular one: the outgoing regular Hamiltonian is one of the two Hamiltonians active on the preceding singular arc.

This definition contains the piecewise regular extremals of Chapters 6 and 7 as the special case in which every open arc is regular. When the extremal consists of a single first-order singular arc, its analysis follows the same general scheme used for the singular extremals of Chapter 5: singular feedback, generalized Legendre condition, super-Hamiltonian, Jacobi equation, and conjugate times. The genuinely new situation considered in this chapter is the concatenation of regular and first-order singular arcs. Its second-order analysis combines the singular constructions of Chapter 5 with the switching and concatenation techniques developed in Chapters 6 and 7.

The remainder of the chapter studies first the second variation and the Hamiltonian geometry of a purely first-order singular arc and then extends these constructions to nice extremals by concatenating the corresponding regular and singular contributions.

8.2 Second variation of first-order singular extremals

In this section, we develop the second-order analysis of a first-order singular arc. The construction closely parallels the one carried out in Section 5.6.1 for the singular L^1 problem: starting from the endpoint map, we compute the second variation and, by applying an integration by parts, we transform it to obtain a coercive form. We then introduce the corresponding augmented space of variations and identify the generalized Legendre form, corresponding to the leading term of the second variation. Finally, under the strong generalized Legendre condition, we derive the associated Jacobi equation and characterize the Jacobi curve along the singular arc. The main difference with Chapter 5 is that the present construction must simultaneously account for variations of the controls on the two active submanifolds and for the singular direction.

8.2.1 Endpoint map and admissible variations

As in the previous section, we fix a first-order singular extremal arc on $[0, T]$, relabel the two active submanifolds as U_0 and U_1 , and write the reference reduced control as

$$\tilde{u} = (\tilde{u}_0, \tilde{u}_1), \quad \tilde{u}_i(t) \in U_i, \quad \tilde{v}(t) \in (0, 1).$$

In the following discussion, we keep the final time fixed. This is sufficient for the second-order necessary conditions, since, by Proposition 7.3.2, time variations commute with control variation if there are no switching points.

As in Section 7.1, we use time-dependent local coordinates on the submanifolds U_i . For $i = 0, 1$, let $d_i = \dim U_i$ and choose smooth families of charts

$$\kappa_i(t, \cdot) : V_i \subset \mathbb{R}^{d_i} \longrightarrow U_i, \quad \kappa_i(t, 0) = \tilde{u}_i(t).$$

After shrinking the sets V_i and a neighbourhood of $\tilde{v}$ in $L^\infty([0, T])$, we obtain a Banach manifold $\mathcal{U}_{\text{sing}}$ of controls

$$u = (u_0, u_1), \quad u_i(t) = \kappa_i(t, \xi_i(t)), \quad v(t) \in (0, 1),$$

modelled on

$$L^\infty([0, T]; \mathbb{R}^{d_0}) \times L^\infty([0, T]; \mathbb{R}) \times L^\infty([0, T]; \mathbb{R}^{d_1}).$$

We identify u_i with its coordinate representative ξ_i whenever no confusion is possible. Unlike the construction for piecewise regular extremals, no time-reparametrization variable is introduced: the controls vary only inside the two maximizing submanifolds and in the direction of v .

Now, we introduce the endpoint maps, the extended endpoint maps and their lifts, similarly to what is done in Section 7.1. For $(u, v) \in \mathcal{U}_{\text{sing}}$, let $q(\cdot; u, v)$ be the trajectory with $q(0; u, v) = q_0$ satisfying (8.1.10). Recall also the definition of $\bar{J}$ in (8.1.7). Define

$$\begin{aligned} F(u, v) &:= q(T; u, v), \\ E(u, v) &:= (F(u, v), \bar{J}(u, v)). \end{aligned} \tag{8.2.1}$$

The fixed-endpoint competitors form the level set $F^{-1}(q_1)$, and the maps in (8.2.1) by their definition are smooth.

Recall also the Hamiltonian dynamics in (8.1.15) and the corresponding flow $\tilde{\Phi}_{s,t}$. Fixing the initial covector $\tilde{\lambda}_0$, define the corresponding Hamiltonian lift and its transported-back version by

$$\begin{aligned} \mathcal{E}(u, v) &:= \overrightarrow{\text{exp}} \int_0^T \vec{h}_{(u(s), v(s))} ds \circ \tilde{\lambda}_0, \\ \mathcal{G}(u, v) &:= \tilde{\Phi}_{0,T}^{-1} \circ \mathcal{E}(u, v), \\ G(u, v) &:= \pi \circ \mathcal{G}(u, v), \end{aligned}$$

These maps satisfy

$$\pi \circ \mathcal{E} = F, \quad \mathcal{E}(\tilde{u}, \tilde{v}) = \tilde{\lambda}_T, \quad \mathcal{G}(\tilde{u}, \tilde{v}) = \tilde{\lambda}_0, \quad G(\tilde{u}, \tilde{v}) = q_0.$$

Finally, the tangent space of control variations is

$$\mathcal{V} := T_{(\tilde{u}, \tilde{v})} \mathcal{U}_{\text{sing}} \simeq L^\infty([0, T]; \mathcal{W}_t), \quad \mathcal{W}_t := T_{\tilde{u}_0(t)} U_0 \oplus \mathbb{R} \oplus T_{\tilde{u}_1(t)} U_1.$$

By our choice of coordinates, we can identify $\mathcal{W}_t \simeq \mathbb{R}^{d_0+1+d_1}$. We write

$$w = (w_0, \omega, w_1) \in \mathcal{V},$$

where $w_i(t) \in T_{\tilde{u}_i(t)} U_i$ and ω is the infinitesimal variation of v . For later use, set

$$\begin{aligned} \mathcal{V}_t &:= \{w \in \mathcal{V} \mid \text{supp } w \subset [0, t]\}, \\ \mathcal{V}_{a,b} &:= \{w \in \mathcal{V} \mid \text{supp } w \subset [a, b]\}, \quad 0 \leq a < b \leq T. \end{aligned}$$

8.2.2 First and second variation of the endpoint map

We now compute the first and second variations of the endpoint map introduced in the previous subsection. As in the previous chapters, the entire dynamics is transported back along the reference Hamiltonian flow and regarded within the fixed symplectic space $T_{\tilde{\lambda}_0}^*(T^*M)$.

Recall that the reduced relaxed Hamiltonian is

$$h_{u_0, v, u_1} = v h_{u_0}^0 + (1 - v) h_{u_1}^1.$$

For every $t \in [0, T]$, define the transported-back perturbation Hamiltonian by

$$\mathbf{h}_t(\lambda, u_0, v, u_1) := (h_{(u_0, v, u_1)} - h_{(\tilde{u}, \tilde{v})}) \circ \tilde{\Phi}_{0,t}(\lambda).$$

By construction,

$$\mathbf{h}_t(\tilde{\lambda}_0, \tilde{u}_0(t), \tilde{v}(t), \tilde{u}_1(t)) = 0.$$

Moreover, the variation formula of chronological calculus gives

$$\mathcal{G}(u, v) = \overrightarrow{\exp} \int_0^T \vec{\mathbf{h}}_s(\cdot, u_0(s), v(s), u_1(s)) ds \circ \tilde{\lambda}_0.$$

Let

$$\sigma := \sigma_{\tilde{\lambda}_0}, \quad \Pi := T_{\tilde{\lambda}_0}^*(T_{q_0}^*M).$$

For $i = 0, 1$, we introduce the transported-back reference Hamiltonian vector fields

$$\vec{\mathcal{H}}_t^i := (\tilde{\Phi}_{0,t}^{-1})_* \vec{h}_{\tilde{u}_i(t)}^i \Big|_{\tilde{\lambda}_t} \in T_{\tilde{\lambda}_0}^*(T^*M),$$

whose difference is the distinguished singular direction

$$\Delta_t := \vec{\mathcal{H}}_t^0 - \vec{\mathcal{H}}_t^1.$$

The variation-to-tangent maps are defined by

$$Z_t^0 w_0 := \tilde{v}(t) \left(\tilde{\Phi}_{0,t}^{-1} \right)_* \frac{\partial \vec{h}_u^0}{\partial u} [w_0] \Big|_{u=\tilde{u}_0(t), \lambda=\tilde{\lambda}_t}$$

and

$$Z_t^1 w_1 := (1 - \tilde{v}(t)) \left(\tilde{\Phi}_{0,t}^{-1} \right)_* \frac{\partial \vec{h}_u^1}{\partial u} [w_1] \Big|_{u=\tilde{u}_1(t), \lambda=\tilde{\lambda}_t}.$$

Thus, for a variation

$$w = (w_0, \omega, w_1) \in \mathcal{V},$$

we set

$$X_t[w(t)] := Z_t^0 w_0(t) + \omega(t) \Delta_t + Z_t^1 w_1(t). \quad (8.2.2)$$

The first-order term in the chronological expansion gives

$$D_{(\tilde{u}, \tilde{v})} \mathcal{G}[w] = \int_0^T X_s[w(s)] ds.$$

Consequently,

$$D_{(\tilde{u}, \tilde{v})} G[w] = d_{\tilde{\lambda}_0} \pi \left[\int_0^T X_s[w(s)] ds \right]. \quad (8.2.3)$$

In particular, we obtain

$$\mathcal{K} := \ker (D_{(\tilde{u}, \tilde{v})} G|_{\mathcal{V}}) = \left\{ w \in \mathcal{V} \mid \int_0^T X_s[w(s)] ds \in \Pi \right\}. \quad (8.2.4)$$

The first differential also recovers the fact that the control is a critical point of the PMP Hamiltonian along the reference extremal. Indeed, pairing (8.2.3) with $\tilde{\lambda}_0$ and using the transported Hamiltonian representation yields

$$\begin{aligned} \langle \tilde{\lambda}_0, D_{(\tilde{u}, \tilde{v})} G[w] \rangle &= \int_0^T \left(\tilde{v}(t) \frac{\partial h_u^0}{\partial u} (\tilde{\lambda}_t) [w_0(t)] \Big|_{u=\tilde{u}_0(t)} \right. \\ &\quad + \omega(t) \left(h_{\tilde{u}_0(t)}^0 - h_{\tilde{u}_1(t)}^1 \right) (\tilde{\lambda}_t) \\ &\quad \left. + (1 - \tilde{v}(t)) \frac{\partial h_u^1}{\partial u} (\tilde{\lambda}_t) [w_1(t)] \Big|_{u=\tilde{u}_1(t)} \right) dt. \end{aligned}$$

Since $((\tilde{u}, \tilde{v}), \tilde{\lambda})$ is a normal extremal pair, $\tilde{\lambda}_0$ annihilates the image of $D_{(\tilde{u}, \tilde{v})}G$. Arbitrary localized variations in the three control components therefore give

$$\left. \frac{\partial h_u^i}{\partial u}(\tilde{\lambda}_t) \right|_{u=\tilde{u}_i(t)} = 0, \quad i = 0, 1, \quad (8.2.5)$$

and

$$h_{\tilde{u}_0(t)}^0(\tilde{\lambda}_t) = h_{\tilde{u}_1(t)}^1(\tilde{\lambda}_t),$$

for almost every $t \in [0, T]$. These identities coincide with the maximization conditions on the submanifolds U_i and with the simultaneous maximizing condition $H^0(\tilde{\lambda}_t) = H^1(\tilde{\lambda}_t) = H(\tilde{\lambda}_t)$.

We now turn to the second variation. For $w, w' \in \mathcal{K}$, define the Hessian by

$$Q(w, w') := \left\langle \tilde{\lambda}_0, \text{Hess}_{(\tilde{u}, \tilde{v})} G(w, w') \right\rangle = \left\langle (\tilde{\lambda}_T, -1), \text{Hess}_{(\tilde{u}, \tilde{v})} E(w, w') \right\rangle. \quad (8.2.6)$$

As usual, we denote by $Q(w)$ the associated quadratic form. The general second-variation formula of Chapter 3 gives the following extension of Q to the whole space $\mathcal{V}$:

$$Q(w) = - \int_0^T D^2 \mathbf{h}_t[w(t)]^2 dt - \int_0^T \int_0^t \sigma(X_s[w(s)], X_t[w(t)]) ds dt. \quad (8.2.7)$$

Its restriction to $\mathcal{K}$ is precisely the scalar second variation defined in (8.2.6).

It remains to describe the Hessian $D^2 \mathbf{h}_t$ in the first integral of (8.2.7). Let $w = (w_0, \omega, w_1)$ and $w' = (w'_0, \omega', w'_1)$. Differentiating the affine relaxed Hamiltonian and evaluating along the reference extremal gives

$$\begin{aligned} D^2 \mathbf{h}_t[w, w'] &= \tilde{v}(t) \left. \frac{\partial^2 h_u^0}{\partial u^2}(\tilde{\lambda}_t) [w_0, w'_0] \right|_{u=\tilde{u}_0(t)} \\ &\quad + (1 - \tilde{v}(t)) \left. \frac{\partial^2 h_u^1}{\partial u^2}(\tilde{\lambda}_t) [w_1, w'_1] \right|_{u=\tilde{u}_1(t)} \\ &\quad + \omega \left(\left. \frac{\partial h_u^0}{\partial u}(\tilde{\lambda}_t) [w'_0] \right|_{u=\tilde{u}_0(t)} - \left. \frac{\partial h_u^1}{\partial u}(\tilde{\lambda}_t) [w'_1] \right|_{u=\tilde{u}_1(t)} \right) \\ &\quad + \omega' \left(\left. \frac{\partial h_u^0}{\partial u}(\tilde{\lambda}_t) [w_0] \right|_{u=\tilde{u}_0(t)} - \left. \frac{\partial h_u^1}{\partial u}(\tilde{\lambda}_t) [w_1] \right|_{u=\tilde{u}_1(t)} \right). \end{aligned}$$

By (8.2.5), the last two lines vanish. Therefore,

$$\begin{aligned} D^2 \mathbf{h}_t[w(t)]^2 &= \tilde{v}(t) \left. \frac{\partial^2 h_u^0}{\partial u^2}(\tilde{\lambda}_t) [w_0(t)]^2 \right|_{u=\tilde{u}_0(t)} \\ &\quad + (1 - \tilde{v}(t)) \left. \frac{\partial^2 h_u^1}{\partial u^2}(\tilde{\lambda}_t) [w_1(t)]^2 \right|_{u=\tilde{u}_1(t)}. \end{aligned}$$

Equivalently, introducing the ordinary Legendre forms

$$\mathbb{L}_t^0 := -\tilde{v}(t) \frac{\partial^2 h_u^0}{\partial u^2}(\tilde{\lambda}_t) \Big|_{u=\tilde{u}_0(t)}, \quad \mathbb{L}_t^1 := -(1 - \tilde{v}(t)) \frac{\partial^2 h_u^1}{\partial u^2}(\tilde{\lambda}_t) \Big|_{u=\tilde{u}_1(t)}, \quad (8.2.8)$$

we have

$$-D^2 \mathbf{h}_t[w(t)]^2 = \mathbb{L}_t^0[w_0(t)]^2 + \mathbb{L}_t^1[w_1(t)]^2. \quad (8.2.9)$$

The forms $\mathbb{L}_t^0$ and $\mathbb{L}_t^1$ are positive definite by the strong Legendre condition on the two active submanifolds. However, the pointwise term contains no contribution in the direction ω , because both the relaxed dynamics and the relaxed cost depend affinely on v . The Goh transformation introduced in the next subsection is precisely aimed at recovering the missing coercive term in this singular direction.

8.2.3 Goh transformation

The expression in (8.2.7) is degenerate in the variation ω corresponding to the control v , because the pointwise Hessian does not contain any quadratic term in this direction. We remove this degeneracy by performing a Goh transformation.

Let

$$w = (w_0, \omega, w_1) \in \mathcal{V},$$

and introduce the primitive

$$\vartheta(t) := \int_0^t \omega(s) ds. \quad (8.2.10)$$

The transformed variation is

$$\zeta(t) := (w_0(t), \vartheta(t), w_1(t)) \in \mathcal{W}_t, \quad (8.2.11)$$

where

$$\mathcal{W}_t := T_{\tilde{u}_0(t)}U_0 \oplus \mathbb{R} \oplus T_{\tilde{u}_1(t)}U_1.$$

We define the desingularized variation-to-tangent map

$$\mathcal{Z}_t(w_0, \vartheta, w_1) := \mathcal{Z}_t^0 w_0 - \vartheta \dot{\Delta}_t + \mathcal{Z}_t^1 w_1. \quad (8.2.12)$$

Here and below, $\dot{\Delta}_t$ denotes the derivative of the Lipschitz curve $t \mapsto \Delta_t$, defined for almost every $t \in [0, T]$.

Since $\omega = \dot{\vartheta}$ and $\vartheta(0) = 0$, integration by parts gives

$$\int_0^t X_s[w(s)] ds = \vartheta(t)\Delta_t + \int_0^t \mathcal{Z}_s \zeta(s) ds. \quad (8.2.13)$$

Consequently, the endpoint-preserving condition in (8.2.4) becomes

$$\mathcal{K}_t = \left\{ w \in \mathcal{V}_t \mid \vartheta(t)\Delta_t + \int_0^t \mathcal{Z}_s \zeta(s) ds \in \Pi \right\}.$$

The first term in (8.2.13) is the boundary contribution generated by the transformation, while the integral contains only the distributed variables ζ .

We now transform the second variation. For every $t \in [0, T]$, define the linear forms

$$a_t^0(w_0) := \sigma(Z_t^0 w_0, \Delta_t), \quad a_t^1(w_1) := \sigma(Z_t^1 w_1, \Delta_t), \quad (8.2.14)$$

and the scalar coefficient

$$\alpha_t := \sigma(\Delta_t, \dot{\Delta}_t). \quad (8.2.15)$$

The symmetric bilinear form l_t on $\mathcal{W}_t$ is determined by

$$\begin{aligned} l_t[(w_0, \vartheta, w_1), (w'_0, \vartheta', w'_1)] &:= \mathbb{L}_t^0[w_0, w'_0] + \mathbb{L}_t^1[w_1, w'_1] \\ &\quad + \vartheta' a_t^0(w_0) + \vartheta a_t^0(w'_0) \\ &\quad + \vartheta' a_t^1(w_1) + \vartheta a_t^1(w'_1) \\ &\quad + \alpha_t \vartheta \vartheta'. \end{aligned} \quad (8.2.16)$$

In particular,

$$\begin{aligned} l_t[\zeta(t), \zeta(t)] &= \mathbb{L}_t^0[w_0(t)]^2 + \mathbb{L}_t^1[w_1(t)]^2 \\ &\quad + 2\vartheta(t) a_t^0(w_0(t)) + 2\vartheta(t) a_t^1(w_1(t)) \\ &\quad + \alpha_t \vartheta(t)^2. \end{aligned}$$

Proposition 8.2.1 (Goh-transformed second variation). *For every $w = (w_0, \omega, w_1) \in \mathcal{V}$, with ϑ and ζ defined by (8.2.10) and (8.2.11), one has*

$$\begin{aligned} Q(w) &= \vartheta(T) \sigma \left(\Delta_T, \int_0^T \mathcal{Z}_t \zeta(t) dt \right) + \int_0^T l_t[\zeta(t), \zeta(t)] dt + \\ &\quad + \int_0^T \sigma \left(\mathcal{Z}_t \zeta(t), \int_0^t \mathcal{Z}_s \zeta(s) ds \right) dt. \end{aligned} \quad (8.2.17)$$

Proof. Set

$$Y_t := Z_t^0 w_0(t) + Z_t^1 w_1(t), \quad B_t := \vartheta(t) \Delta_t.$$

By (8.2.2) and (8.2.12),

$$X_t[w(t)] = \dot{B}_t + \mathcal{Z}_t \zeta(t).$$

Moreover, (8.2.13) yields

$$\int_0^t X_s[w(s)] ds = B_t + \int_0^t \mathcal{Z}_s \zeta(s) ds. \quad (8.2.18)$$

Using the skew-symmetry of σ , the double integral in (8.2.7) can be rewritten as

$$- \int_0^T \int_0^t \sigma(X_s[w(s)], X_t[w(t)]) ds dt = \int_0^T \sigma \left(X_t[w(t)], \int_0^t X_s[w(s)] ds \right) dt.$$

Substituting (8.2.18) in the formula (8.2.7) and integrating the term containing B_t by parts, we obtain

$$\begin{aligned}
& - \int_0^T \int_0^t \sigma(X_s[w(s)], X_t[w(t)]) ds dt \\
& = \vartheta(T) \sigma \left(\Delta_T, \int_0^T \mathcal{Z}_t \zeta(t) dt \right) \\
& \quad + \int_0^T [\alpha_t \vartheta(t)^2 + 2\vartheta(t) a_t^0(w_0(t)) + 2\vartheta(t) a_t^1(w_1(t))] dt \\
& \quad + \int_0^T \sigma \left(\mathcal{Z}_t \zeta(t), \int_0^t \mathcal{Z}_s \zeta(s) ds \right) dt.
\end{aligned} \tag{8.2.19}$$

Combining (8.2.19) with (8.2.9) gives (8.2.17). $\square$

Formula (8.2.17) separates the transformed second variation into a boundary term, the local form l_t , and a Volterra-type symplectic term involving only the desingularized map $\mathcal{Z}_t$. For variations in $\mathcal{V}$, the terminal value $\vartheta(T)$ is the trace of the primitive ϑ . In the next subsection, these two quantities will be regarded as independent variables in the natural completion of the transformed variation space. The algebraic structure and positivity of the local form l_t will be analyzed afterwards.

8.2.4 Augmented second variation

For a variation in $\mathcal{V}$, the primitive introduced in (8.2.10) belongs to $W^{1,\infty}([0, T])$, satisfies $\vartheta(0) = 0$, and its terminal value $\vartheta(T)$ is determined by the function ϑ . Formula (8.2.17), however, depends continuously on ϑ and on its terminal value when these two quantities are regarded as independent variables. We now introduce the corresponding completion, which is the natural space for the construction of the $\mathcal{L}$ -derivative and of the Jacobi curve.

Using the time-dependent coordinates fixed in Subsection 8.2.1, we identify the spaces $\mathcal{W}_t$ with $\mathbb{R}^{d_0+1+d_1}$ and set

$$\widehat{\mathcal{V}} := \mathbb{R} \times L^2([0, T]; \mathcal{W}_t). \tag{8.2.20}$$

An element of $\widehat{\mathcal{V}}$ is denoted by

$$\widehat{w} = (\vartheta_T, \zeta), \quad \zeta = (w_0, \vartheta, w_1).$$

The scalar ϑ_T and the middle component ϑ of ζ are independent in (8.2.20). In particular, no trace of the L^2 -function ϑ is prescribed at T . We equip $\widehat{\mathcal{V}}$ with the Hilbert norm

$$\|(\vartheta_T, \zeta)\|_{\widehat{\mathcal{V}}}^2 := |\vartheta_T|^2 + \int_0^T |\zeta(t)|^2 dt. \tag{8.2.21}$$

The Goh transformation defines a linear map

$$\begin{aligned}
& \mathcal{T}_{\text{Goh}} : \mathcal{V} \longrightarrow \widehat{\mathcal{V}}, \\
& w = (w_0, \omega, w_1) \longmapsto (\vartheta(T), (w_0, \vartheta, w_1)),
\end{aligned}$$

where $\vartheta(t) = \int_0^t \omega(s) ds$. We denote its image by $\mathcal{V}_{\text{Goh}}$.

Lemma 8.2.2. *The subspace $\mathcal{V}_{\text{Goh}}$ is dense in $\widehat{\mathcal{V}}$.*

Proof. It is sufficient to show that any pair $(\vartheta_T, \vartheta) \in \mathbb{R} \times L^2([0, T]; \mathbb{R})$ can be approximated, with respect to the norm in (8.2.21), by a pair of the type $(\vartheta(T), \vartheta)$, where $\vartheta(t) = \int_0^t \omega(s) ds$, for some $\omega \in L^2([0, T]; \mathbb{R})$. Choose $\psi_n \in C_c^\infty((0, T))$ such that $\psi_n \rightarrow \vartheta$ in $L^2([0, T])$. Let $\beta_n \in W^{1,\infty}([0, T])$ be supported in $[T - 1/n, T]$ and satisfy

$$\beta_n(0) = 0, \quad \beta_n(T) = \vartheta_T, \quad \|\beta_n\|_{L^2} \rightarrow 0.$$

Then

$$\vartheta_n := \psi_n + \beta_n$$

belongs to $W^{1,\infty}([0, T])$, satisfies $\vartheta_n(0) = 0$ and $\vartheta_n(T) = \vartheta_T$, and converges to ϑ in $L^2([0, T])$. Setting $\omega_n = \dot{\vartheta}_n$, we obtain the required approximation. $\square$

For every $t \in [0, T]$, define the extended first-variation operator

$$\begin{aligned} \widehat{A}_t : \mathbb{R} \times L^2([0, t]; \mathcal{W}_s) &\longrightarrow T_{\tilde{\lambda}_0}(T^*M), \\ \widehat{A}_t(\vartheta_t, \zeta) &:= \vartheta_t \Delta_t + \int_0^t \mathcal{Z}_s \zeta(s) ds. \end{aligned} \tag{8.2.22}$$

For a variation $w = (w_0, \omega, w_1) \in \mathcal{V}$, formula (8.2.13) gives

$$\widehat{A}_t(\vartheta(t), \zeta|_{[0,t]}) = \int_0^t X_s[w(s)] ds. \tag{8.2.23}$$

The augmented endpoint-preserving space up to time t is therefore

$$\widehat{\mathcal{K}} := \left\{ (\vartheta_T, \zeta) \in \mathbb{R} \times L^2([0, T]; \mathcal{W}) \mid \widehat{A}_T(\vartheta_T, \zeta) \in \Pi \right\}.$$

We write $\widehat{A} := \widehat{A}_T$. In particular, (8.2.23) implies

$$\mathcal{T}_{\text{Goh}}(\mathcal{K}) = \mathcal{V}_{\text{Goh}} \cap \widehat{\mathcal{K}}.$$

We next extend the transformed quadratic form. For $(\vartheta_T, \zeta) \in \mathbb{R} \times L^2([0, T]; \mathcal{W}_s)$, set

$$\begin{aligned} \widehat{Q}(\vartheta_T, \zeta) &:= \vartheta_T \sigma \left(\Delta_T, \int_0^T \mathcal{Z}_s \zeta(s) ds \right) + \int_0^T l_s[\zeta(s), \zeta(s)] ds \\ &\quad + \int_0^T \sigma \left(\mathcal{Z}_s \zeta(s), \int_0^s \mathcal{Z}_r \zeta(r) dr \right) ds. \end{aligned} \tag{8.2.24}$$

For every variation $w \in \mathcal{V}$, Proposition 8.2.1 yields

$$Q(w) = \widehat{Q}(\mathcal{T}_{\text{Goh}} w).$$

The restriction of $\widehat{Q}$ to $\widehat{\mathcal{K}}$ will be called the augmented second variation.

Proposition 8.2.3. *For every $t \in [0, T]$, the operator $\hat{A}_t$ and the quadratic form $\hat{Q}$ are continuous on their respective domains. Consequently, $\hat{\mathcal{K}}$ is a closed subspace of finite codimension.*

Proof. The curves Δ_t , Z_t^0 , Z_t^1 , $\dot{\Delta}_t$, and l_t are bounded on the compact interval under consideration. Hence there is a constant $C > 0$ such that

$$\|\mathcal{Z}_s \zeta(s)\| \leq C |\zeta(s)|$$

for almost every s . It follows from the Cauchy–Schwarz inequality that

$$\left\| \int_0^r \mathcal{Z}_s \zeta(s) ds \right\| \leq C \sqrt{t} \|\zeta\|_{L^2([0, t])}, \quad r \in [0, t].$$

This estimate proves the continuity of $\hat{A}_t$. The local term in (8.2.24) is continuous by the uniform boundedness of l_t . The same estimate gives

$$\left| \int_0^T \sigma \left(\mathcal{Z}_s \zeta(s), \int_0^s \mathcal{Z}_r \zeta(r) dr \right) ds \right| \leq C \|\zeta\|_{L^2([0, t])}^2.$$

The boundary term is bounded by $C |\vartheta_T| \|\zeta\|_{L^2([0, T])}$. Therefore,

$$\left| \hat{Q}(\vartheta_T, \zeta) \right| \leq C \left(|\vartheta_T|^2 + \|\zeta\|_{L^2([0, T])}^2 \right),$$

which proves the continuity of $\hat{Q}$. Finally, $\hat{\mathcal{K}}$ is the inverse image of the finite-dimensional subspace Π under the continuous map $\hat{A}_t$, so it is closed and has finite codimension. $\square$

Proposition 8.2.4 (Invariance of the $\mathcal{L}$ -derivative). *Let $((\tilde{u}, \tilde{v}), \tilde{\lambda})$ be a first-order singular extremal pair. If their Jacobi curve over the space of variation $\mathcal{V}$ is well defined, then it coincide with the Jacobi curve obtained over the space of variations $\hat{\mathcal{V}}$.*

This is a direct consequence of Lemma 8.2.2 and point 4. in Remark 3.1.10.

The augmented construction separates the terminal contribution $\vartheta_T \Delta_T$ from the distributed contribution generated by $\mathcal{Z}_t$. This separation is essential in the subsequent construction of the Jacobi curve, because the terminal variable produces the one-dimensional extension in the direction Δ_t , while the distributed variables determine the continuous Jacobi evolution.

8.2.5 Generalized Legendre condition

The Goh transformation isolates in (8.2.17) the pointwise symmetric form l_t on $\mathcal{W}_t$. This form contains the ordinary Legendre terms associated with the two maximizing submanifolds and the additional direction generated by the primitive ϑ of ω . We now analyze its algebraic structure and relate its positivity to the first-order non-degeneracy conditions introduced in Definition 8.1.1.

Definition 8.2.5 (Generalized Legendre condition). The symmetric form l_t introduced in (8.2.16) is called the *generalized Legendre form* of the first-order singular arc. We say that the *generalized Legendre condition* holds if

$$l_t \geq 0 \quad (8.2.25)$$

for almost every $t \in [0, T]$. We say that the strong generalized Legendre condition holds if

$$l_t[\zeta, \zeta] > 0 \quad \zeta \in \mathcal{W}_t, \quad (8.2.26)$$

for almost every $t \in [0, T]$.

With respect to the order (w_0, ϑ, w_1) fixed above, the generalized Legendre form has the block representation

$$l_t = \begin{pmatrix} \mathbb{L}_t^0 & a_t^0 & 0 \\ (a_t^0)^* & \alpha_t & (a_t^1)^* \\ 0 & a_t^1 & \mathbb{L}_t^1 \end{pmatrix}.$$

Here the linear forms a_t^0, a_t^1 and the scalar α_t are those defined in (8.2.14) and (8.2.15), while $\mathbb{L}_t^0, \mathbb{L}_t^1$ are the ordinary Legendre forms in (8.2.8). We recall the following matrix decomposition:

$$\begin{pmatrix} A & a & 0 \\ a^\top & \alpha & b^\top \\ 0 & b & B \end{pmatrix} = \begin{pmatrix} \text{Id} & A^{-1}a & 0 \\ 0 & 1 & 0 \\ 0 & B^{-1}b & \text{Id} \end{pmatrix}^\top \begin{pmatrix} A & 0 & 0 \\ 0 & \tilde{\alpha} & 0 \\ 0 & 0 & B \end{pmatrix} \begin{pmatrix} \text{Id} & A^{-1}a & 0 \\ 0 & 1 & 0 \\ 0 & B^{-1}b & \text{Id} \end{pmatrix},$$

where A, B are matrices and a, b vectors of appropriate dimensions, $\alpha \in \mathbb{R}$ and $\tilde{\alpha} = \alpha - a^\top A^{-1}a - b^\top B^{-1}b$. By the strong Legendre condition on the submanifold U_i , $\mathbb{L}_t^0$ and $\mathbb{L}_t^1$ are invertible, so we apply this decomposition to the generalized Legendre form l_t and define

$$\gamma_t := \alpha_t - \langle (\mathbb{L}_t^0)^{-1}a_t^0, a_t^0 \rangle - \langle (\mathbb{L}_t^1)^{-1}a_t^1, a_t^1 \rangle. \quad (8.2.27)$$

Lemma 8.2.6 (Block reduction). Assume that $\mathbb{L}_t^0$ and $\mathbb{L}_t^1$ satisfy the strong Legendre condition. Then, for almost every $t \in [0, T]$, the generalized Legendre form is congruent to

$$\mathbb{L}_t^0 \oplus \gamma_t \oplus \mathbb{L}_t^1.$$

More precisely, for every $(w_0, \vartheta, w_1) \in \mathcal{W}_t$ one has

$$\begin{aligned} l_t[(w_0, \vartheta, w_1), (w_0, \vartheta, w_1)] &= \mathbb{L}_t^0 [w_0 + \vartheta(\mathbb{L}_t^0)^{-1}a_t^0]^2 \\ &\quad + \gamma_t \vartheta^2 \\ &\quad + \mathbb{L}_t^1 [w_1 + \vartheta(\mathbb{L}_t^1)^{-1}a_t^1]^2. \end{aligned} \quad (8.2.28)$$

Consequently, l_t is non-negative, respectively positive definite, if and only if γ_t is non-negative, respectively positive, at that time.

Proof. Formula (8.2.28) follows by expanding the two squares and using (8.2.27). The remaining assertions follow from the positive definiteness of $\mathbb{L}_t^0$ and $\mathbb{L}_t^1$. $\square$

It remains to compute an explicit formula for the scalar term γ_t . For this purpose, the distinction between the brackets of the maximized Hamiltonians and the brackets obtained by freezing the reference controls must be kept explicit. Set

$$\begin{aligned} h_t^{001} &:= \left\{ h_{\tilde{u}_0(t)}^0, \left\{ h_{\tilde{u}_0(t)}^0, h_{\tilde{u}_1(t)}^1 \right\} \right\}, \\ h_t^{110} &:= \left\{ h_{\tilde{u}_1(t)}^1, \left\{ h_{\tilde{u}_1(t)}^1, h_{\tilde{u}_0(t)}^0 \right\} \right\}. \end{aligned} \quad (8.2.29)$$

These Hamiltonians are obtained by keeping $\tilde{u}_0(t)$ and $\tilde{u}_1(t)$ fixed when the Poisson brackets are computed. They must not be confused with H^{001} and H^{110} , for which the maximizing feedbacks vary with the covector.

For a fixed t , abbreviate

$$\vec{\delta}_t := \vec{h}_{\tilde{u}_0(t)}^0(\tilde{\lambda}_t) - \vec{h}_{\tilde{u}_1(t)}^1(\tilde{\lambda}_t),$$

and define

$$\rho_t^i := d_{\tilde{\lambda}_t} \widehat{u}_i[\vec{\delta}_t], \quad i = 0, 1.$$

Lemma 8.2.7 (Identification of the scalar block). *Along the reference first-order singular arc, one has*

$$a_t^i(w_i) = -\mathbb{L}_t^i[w_i, \rho_t^i], \quad i = 0, 1,$$

and

$$\gamma_t = - (H^{001} + H^{110})(\tilde{\lambda}_t). \quad (8.2.30)$$

Proof. In the following computation, all derivatives of h_u^i are evaluated at $(\tilde{\lambda}_t, \tilde{u}_i(t))$ unless a different base point is explicitly indicated. Differentiating the identity

$$\left. \frac{\partial h_u^i}{\partial u} \right|_{u=\widehat{u}_i(\lambda)} = 0$$

with respect to λ gives

$$d_\lambda \left(\frac{\partial h_u^i}{\partial u} [w_i] \right) [\xi] + \frac{\partial^2 h_u^i}{\partial u^2} [w_i, d_\lambda \widehat{u}_i[\xi]] = 0. \quad (8.2.31)$$

By the symplectic invariance of the reference flow and the definition of Z_t^i , we have

$$a_t^0(w_0) = \tilde{v}(t) \sigma_{\tilde{\lambda}_t} \left(\frac{\partial \vec{h}_u^0}{\partial u} [w_0], \vec{\delta}_t \right) = \tilde{v}(t) \frac{\partial^2 h_u^0}{\partial u^2} [w_0, \rho_t^0] = -\mathbb{L}_t^0[w_0, \rho_t^0].$$

The same computation gives a similar identity for a_t^1 . Hence,

$$\langle (\mathbb{L}_t^i)^{-1} a_t^i, a_t^i \rangle = \mathbb{L}_t^i[\rho_t^i, \rho_t^i], \quad i = 0, 1. \quad (8.2.32)$$

We now compute α_t . Differentiating the transported-back singular direction gives

$$\dot{\Delta}_t = (\tilde{\Phi}_{0,t}^{-1})_* \left(\vec{h}_t^{10} + \frac{\partial \vec{h}_u^0}{\partial u} [\dot{u}_0(t)] - \frac{\partial \vec{h}_u^1}{\partial u} [\dot{u}_1(t)] \right),$$

where

$$h_t^{10} := \left\{ h_{\tilde{u}_1(t)}^1, h_{\tilde{u}_0(t)}^0 \right\}.$$

Using (8.2.15), (8.2.31), and the identities in (8.2.29), we obtain

$$\alpha_t = - (h_t^{001} + h_t^{110}) (\tilde{\lambda}_t) - \frac{\partial^2 h_u^0}{\partial u^2} [\dot{u}_0(t), \rho_t^0] + \frac{\partial^2 h_u^1}{\partial u^2} [\dot{u}_1(t), \rho_t^1]. \quad (8.2.33)$$

Introduce

$$\eta_t^0 := d_{\tilde{\lambda}_t} \widehat{u}_0 [\vec{h}_{\tilde{u}_1(t)}^1], \quad \eta_t^1 := d_{\tilde{\lambda}_t} \widehat{u}_1 [\vec{h}_{\tilde{u}_0(t)}^0].$$

Since the reference Hamiltonian vector field is

$$\tilde{v}(t) \vec{h}_{\tilde{u}_0(t)}^0 + (1 - \tilde{v}(t)) \vec{h}_{\tilde{u}_1(t)}^1,$$

the derivatives of $\tilde{u}_0$ and $\tilde{u}_1$ satisfy

$$\dot{\tilde{u}}_0(t) = \eta_t^0 + \tilde{v}(t) \rho_t^0, \quad \dot{\tilde{u}}_1(t) = \eta_t^1 - (1 - \tilde{v}(t)) \rho_t^1. \quad (8.2.34)$$

Substituting (8.2.32), (8.2.33), and (8.2.34) into (8.2.27) yields

$$\gamma_t = - (h_t^{001} + h_t^{110}) (\tilde{\lambda}_t) - \frac{\partial^2 h_u^0}{\partial u^2} [\eta_t^0, \rho_t^0] + \frac{\partial^2 h_u^1}{\partial u^2} [\eta_t^1, \rho_t^1]. \quad (8.2.35)$$

Finally, set

$$\mathcal{B}(\lambda, u_0, u_1) := \{ h_{u_0}^0, h_{u_1}^1 \} (\lambda).$$

Since

$$H^{01}(\lambda) = \mathcal{B}(\lambda, \widehat{u}_0(\lambda), \widehat{u}_1(\lambda)),$$

the chain rule in the direction $\vec{\delta}_t$ gives

$$(H^{001} + H^{110}) (\tilde{\lambda}_t) = (h_t^{001} + h_t^{110}) (\tilde{\lambda}_t) + \frac{\partial^2 h_u^0}{\partial u^2} [\eta_t^0, \rho_t^0] - \frac{\partial^2 h_u^1}{\partial u^2} [\eta_t^1, \rho_t^1]. \quad (8.2.36)$$

Comparing (8.2.35) and (8.2.36) proves (8.2.30). $\square$

Proposition 8.2.8 (Positivity of the generalized Legendre form). *Let $\tilde{\lambda}$ be a first-order singular extremal arc in the sense of Definition 8.1.1. Then the strong generalized Legendre condition (8.2.26) holds, that is l_t is positive definite for almost every $t \in [0, T]$.*

Proof. By the strong Legendre condition, the forms $\mathbb{L}_t^0$ and $\mathbb{L}_t^1$ are positive definite for almost every t . By Lemma 8.2.7 and (8.1.12),

$$\gamma_t = -H^{001}(\tilde{\lambda}_t) - H^{110}(\tilde{\lambda}_t) > 0.$$

Then, the statement follows from Lemma 8.2.6. $\square$

We conclude by recording the variational consequences of the generalized Legendre condition. For $0 \leq a < b \leq T$, let $\widehat{Q}_{a,b}$ and $\widehat{\mathcal{K}}_{a,b}$ denote the augmented second variation and the augmented endpoint-preserving space obtained by repeating the construction of Subsection 8.2.4 on the restricted arc $[a, b]$, with all tangent vectors transported back to $T_{\tilde{\lambda}_a}(T^*M)$.

Proposition 8.2.9 (Variational consequences). *The following statements hold.*

1. *If the negative index of $\widehat{Q}|_{\widehat{\mathcal{K}}}$ is finite, then the generalized Legendre condition (8.2.25) holds.*
2. *If the strong generalized Legendre condition holds, then, for every $\tau \in [0, T)$, there exists $\varepsilon > 0$ such that*

$$\widehat{Q}_{\tau, \tau+\varepsilon}|_{\widehat{\mathcal{K}}_{\tau, \tau+\varepsilon}} > 0. \quad (8.2.37)$$

In particular, $\widehat{Q}|_{\widehat{\mathcal{K}}}$ has finite negative index.

Proof. This is the same localization and rescaling argument used in Proposition 5.6.2. If l_t has a negative direction on a set of positive measure, one can construct infinitely many variations with disjoint sufficiently small supports for which the local integral in (8.2.24) is negative, while the term with the double integral is of lower order. After imposing the finite-dimensional endpoint condition, this gives an infinite-dimensional negative subspace of the augmented second variation, proving the first assertion by contradiction. Under (8.2.26), the local integral is uniformly coercive. On a sufficiently short interval, the boundary term and the term with the double integral are lower-order perturbations after rescaling, and the finite-dimensional endpoint constraint does not destroy coercivity. This proves (8.2.37). A finite covering of $[0, T]$ by such intervals gives the finite-index conclusion. $\square$

Proposition 8.2.8 shows that the first-order singularity assumptions recover precisely the coercive direction that is absent from the original pointwise Hessian. The form l_t and the de-singularized map $\mathcal{Z}_t$ will be used in the subsequent construction of the Jacobi equation and of the Jacobi curve along a nice-singular arc.

8.2.6 Jacobi equation and Jacobi curve of a first-order singular extremal

In this subsection, we derive a characterization of the Jacobi curve associated with a first-order singular extremal similar to the one that was obtained in Subsection 5.4.1. That is, we show that the evolution of the Jacobi curve is governed by a suitable Jacobi equation. Since some of these arguments are a very close repetition of the ones already described previously, the exposition is intentionally concise.

Let $((\tilde{u}, \tilde{v}), \tilde{\lambda})$ be the first-order singular extremal pair fixed throughout this section, and let $(\Lambda_t)_{t \in [0, T]}$ be its Jacobi curve, defined as the $\mathcal{L}$ -derivative associated with the filtration of control variations introduced in Subsection 8.2.1. For every $t \in (0, T]$, denote by $\hat{\mathcal{V}}_t$ and $\hat{\mathcal{Q}}_t$ the augmented variation space and the augmented second variation obtained by repeating the construction of Subsection 8.2.4 on the interval $[0, t]$. By Lemma 8.2.2 and Proposition 8.2.4, the space Λ_t can be computed equivalently using the original variations or the augmented Goh variations.

We recall the following notation. If Λ is a Lagrangian subspace and Γ is an isotropic subspace of the same symplectic vector space, we set

$$\Lambda^\Gamma := (\Lambda + \Gamma) \cap \Gamma^\perp = (\Lambda \cap \Gamma^\perp) + \Gamma.$$

By Proposition 8.2.8, the generalized Legendre form l_t is positive definite for almost every $t \in [0, T]$. We identify it with the associated linear map

$$l_t : \mathcal{W}_t \longrightarrow \mathcal{W}_t^*,$$

and denote its inverse by l_t^{-1} .

We now derive the equation governing the continuous evolution of the Jacobi curve. Applying the same procedure described in 5.6.4 to the augmented second variation for a first-order singular extremal and testing independently the ϑ component of the augmented variations, we obtain the following relation

$$l_t[\zeta(t), \cdot] = \sigma(\mathcal{Z}_t \cdot, \eta(t)). \quad (8.2.38)$$

Hence,

$$\zeta(t) = l_t^{-1} \sigma(\mathcal{Z}_t \cdot, \eta(t)).$$

Defining

$$\eta(t) = \eta_0 - \int_0^t \mathcal{Z}_s \zeta(s) ds, \quad t \in [0, T], \eta \in \Pi,$$

and combining the last two identities, we obtain the Jacobi equation

$$\dot{\eta}(t) = -\mathcal{Z}_t l_t^{-1} \sigma(\mathcal{Z}_t \cdot, \eta(t)), \quad \text{for a.e. } t \in [0, T]. \quad (8.2.39)$$

Equation (8.2.39) is a linear Hamiltonian equation on $T_{\tilde{\lambda}_0}^*(T^*M)$. Its time-dependent quadratic Hamiltonian is

$$b_t(\eta) := \frac{1}{2} \langle \sigma(\mathcal{Z}_t \cdot, \eta), l_t^{-1} \sigma(\mathcal{Z}_t \cdot, \eta) \rangle, \quad \eta \in T_{\tilde{\lambda}_0}^*(T^*M).$$

Indeed, with the symplectic convention fixed in Section 2.2, one has

$$\vec{b}_t(\eta) = -\mathcal{Z}_t l_t^{-1} \sigma(\mathcal{Z}_t \cdot, \eta).$$

We denote by

$$B_{s,t} := \overrightarrow{\exp} \int_s^t \vec{b}_\tau d\tau, \quad B_t := B_{0,t}, \quad (8.2.40)$$

the corresponding linear symplectic flow.

The distinguished singular direction is itself a solution of the Jacobi equation.

Lemma 8.2.10. *The curve $t \mapsto \Delta_t$ satisfies Equation (8.2.39). Equivalently,*

$$B_{s,t} \Delta_s = \Delta_t, \quad 0 \leq s \leq t \leq T. \quad (8.2.41)$$

Proof. Let $e_v := (0, 1, 0) \in \mathcal{W}_t$. By the definitions of l_t and $\mathcal{Z}_t$, one has

$$l_t[\cdot, e_v] = \sigma(\mathcal{Z}_t \cdot, \Delta_t), \quad \mathcal{Z}_t e_v = -\dot{\Delta}_t.$$

Therefore,

$$-\mathcal{Z}_t l_t^{-1} \sigma(\mathcal{Z}_t \cdot, \Delta_t) = -\mathcal{Z}_t e_v = \dot{\Delta}_t,$$

which proves the first assertion. The identity (8.2.41) follows from uniqueness for the linear Cauchy problem. $\square$

We can now identify the Jacobi curve explicitly.

Theorem 8.2.11 (Jacobi curve of a first-order singular extremal). *The Jacobi curve associated with $((\tilde{u}, \tilde{v}), \tilde{\lambda})$ satisfies*

$$\Lambda_0 = \Pi, \quad \Lambda_0^+ = \Pi^{\mathbb{R}\Delta_0}, \quad \Lambda_t = B_{0,t} \left(\Pi^{\mathbb{R}\Delta_0} \right), \quad t \in (0, T]. \quad (8.2.42)$$

Here Λ_0^+ denotes the right limit of the Jacobi curve at $t = 0$.

Proof. Fix $t \in (0, T]$. By Proposition 8.2.4, the space Λ_t can be computed on $\hat{\mathcal{V}}_t$. The distributed component ζ of an augmented variation is governed by the algebraic relation (8.2.38); hence its contribution to the $\mathcal{L}$ -derivative evolves according to the Jacobi flow $B_{0,t}$.

The independent terminal variable contributes the vector $\vartheta_t \Delta_t$ to the first variation through (8.2.22), while the restriction of the augmented quadratic form to this one-dimensional

variation space vanishes. The finite-dimensional extension argument of Lemma 3.3.6, applied after the distributed variations, therefore replaces the propagated Lagrangian subspace $B_{0,t}\Pi$ with its Lagrangian extension in the direction $\mathbb{R}\Delta_t$. Consequently,

$$\Lambda_t = (B_{0,t}\Pi)^{\mathbb{R}\Delta_t}.$$

Since $B_{0,t}$ is symplectic and, by Lemma 8.2.10, $B_{0,t}\Delta_0 = \Delta_t$, Lagrangian extension commutes with the flow:

$$(B_{0,t}\Pi)^{\mathbb{R}\Delta_t} = B_{0,t}(\Pi^{\mathbb{R}\Delta_0}).$$

The general definition of the Jacobi curve gives $\Lambda_0 = \Pi$, whereas the preceding formula yields the right limit $\Lambda_0^+ = \Pi^{\mathbb{R}\Delta_0}$. This proves (8.2.42). $\square$

8.3 Super-Hamiltonian and Jacobi equation

In this section, we construct a smooth super-Hamiltonian associated with a reference first-order singular extremal pair $((\tilde{u}, \tilde{v}), \tilde{\lambda})$. We then show that, in the reference transported-back frame, the linearization of its Hamiltonian flow coincides with the Jacobi flow $B_{s,t}$ introduced in Subsection 8.2.6. Throughout the section, H^0 and H^1 denote the two maximizing Hamiltonians, and $\tilde{\Phi}_{s,t}$ denotes the reference Hamiltonian flow generated by $h_{(\tilde{u}(t), \tilde{v}(t))}$.

Recall the submanifolds introduced in Section 8.1,

$$\Sigma = \{\lambda \in \mathcal{O} \mid H^0(\lambda) = H^1(\lambda)\}, \quad \mathcal{S} = \{\lambda \in \mathcal{O} \mid H^{01}(\lambda) = 0\},$$

where $\mathcal{O} \subset T^*M$ is a neighbourhood of the reference arc on which

$$H^{001} < 0, \quad H^{110} < 0.$$

In particular,

$$dH^{01}[\vec{H}^1 - \vec{H}^0] = \{H^1 - H^0, H^{01}\} = -H^{001} - H^{110} > 0 \quad (8.3.1)$$

along the reference arc. After possibly shrinking $\mathcal{O}$, the set $\mathcal{S}$ is therefore a smooth hypersurface and the vector field $\vec{H}^1 - \vec{H}^0$ is transverse to it.

We define H_S^0 as the solution of the characteristic Cauchy problem

$$\begin{cases} \{H_S^0, H^1 - H^0\} = 0 & \text{in } \mathcal{O}, \\ H_S^0 = H^0 & \text{on } \mathcal{S}. \end{cases} \quad (8.3.2)$$

By (8.3.1), the method of characteristics gives a unique smooth solution after a further restriction of $\mathcal{O}$. We then set

$$H_S^1 := H_S^0 + H^1 - H^0.$$

This definition immediately gives

$$H_S^1 = H^1 \quad \text{on } \mathcal{S}, \quad \{H_S^1, H^0 - H^1\} = 0,$$

and

$$H_S^0 - H_S^1 = H^0 - H^1.$$

In particular,

$$\vec{H}_S^0 - \vec{H}_S^1 = \vec{H}^0 - \vec{H}^1.$$

We can now introduce the time-dependent Hamiltonian

$$\begin{aligned} H_S(t, \lambda) &:= H_S^0(\lambda) + (1 - \tilde{v}(t))(H^1 - H^0)(\lambda) \\ &= H_S^1(\lambda) + \tilde{v}(t)(H^0 - H^1)(\lambda). \end{aligned} \tag{8.3.3}$$

We denote by $\Phi_{s,t}^S$ its Hamiltonian flow.

Proposition 8.3.1 (Super-Hamiltonian properties). *After possibly shrinking $\mathcal{O}$, the Hamiltonian H_S satisfies the following properties.*

1. For every $t \in [0, T]$ and every $\lambda \in \mathcal{O} \cap \Sigma$,

$$H_S(t, \lambda) \geq H(\lambda).$$

2. Along the reference extremal,

$$H_S(t, \tilde{\lambda}_t) = H^0(\tilde{\lambda}_t) = H^1(\tilde{\lambda}_t) = H(\tilde{\lambda}_t).$$

3. The reference extremal is an integral curve of $\vec{H}_S$, namely

$$\vec{H}_S(t, \tilde{\lambda}_t) = \tilde{v}(t)\vec{H}^0(\tilde{\lambda}_t) + (1 - \tilde{v}(t))\vec{H}^1(\tilde{\lambda}_t) = \dot{\tilde{\lambda}}_t.$$

Consequently,

$$\Phi_{0,t}^S(\tilde{\lambda}_0) = \tilde{\lambda}_t. \tag{8.3.4}$$

The proof of this Proposition is a repetition of the same argument used in the proof of Theorem 5.1.2, just changing accordingly the notation. For this reason, we omit it.

We now compare the super-Hamiltonian dynamics with the reference dynamics. Define

$$C_{0,t} := \tilde{\Phi}_{0,t}^{-1} \circ \Phi_{0,t}^S.$$

By (8.3.4),

$$C_{0,t}(\tilde{\lambda}_0) = \tilde{\lambda}_0. \tag{8.3.5}$$

The variations formula yields

$$C_{0,t} = \overrightarrow{\exp} \int_0^t \vec{\varphi}_s ds,$$

where

$$\varphi_t := (H_S(t, \cdot) - h_{(\tilde{u}(t), \tilde{v}(t))}) \circ \tilde{\Phi}_{0,t}. \quad (8.3.6)$$

The value and the first differential of the two Hamiltonians agree along the reference extremal, and therefore

$$\varphi_t(\tilde{\lambda}_0) = 0, \quad d_{\tilde{\lambda}_0} \varphi_t = 0.$$

Thus, the first non-trivial derivative of φ_t at $\tilde{\lambda}_0$ is its Hessian.

Proposition 8.3.2. *For every $\eta_1, \eta_2 \in T_{\tilde{\lambda}_0}(T^*M)$,*

$$\text{Hess}_{\tilde{\lambda}_0} \varphi_t[\eta_1, \eta_2] = \langle \sigma(\mathcal{Z}_t \cdot, \eta_1), l_t^{-1} \sigma(\mathcal{Z}_t \cdot, \eta_2) \rangle. \quad (8.3.7)$$

In particular,

$$\frac{1}{2} \text{Hess}_{\tilde{\lambda}_0} \varphi_t[\eta, \eta] = b_t(\eta).$$

The proof of Proposition 8.3.2 is given in Section 8.4. It is precisely at this point that the two ordinary Legendre blocks and the Goh direction, collected in the generalized Legendre form l_t , enter the linearization of the super-Hamiltonian dynamics.

Corollary 8.3.3. *The linearization of φ_t coincides with the Jacobi flow introduced in (8.2.40):*

$$d_{\tilde{\lambda}_0} C_{0,t} = d_{\tilde{\lambda}_0} \left(\tilde{\Phi}_{0,t}^{-1} \circ \Phi_{0,t}^S \right) = B_{0,t}. \quad (8.3.8)$$

Consequently, the Jacobi curve of Theorem 8.2.11 admits the representation

$$\Lambda_t = d_{\tilde{\lambda}_0} \left(\tilde{\Phi}_{0,t}^{-1} \circ \Phi_{0,t}^S \right) \left(\Pi^{\mathbb{R}\Delta_0} \right), \quad t \in (0, T]. \quad (8.3.9)$$

Proof. By (8.3.5), the curve $\tilde{\lambda}_0$ is a fixed solution of the Hamiltonian system generated by φ_t . Its linearization is therefore generated by the quadratic Hamiltonian

$$\eta \longmapsto \frac{1}{2} \text{Hess}_{\tilde{\lambda}_0} \varphi_t[\eta, \eta].$$

By Proposition 8.3.2, this quadratic Hamiltonian is b_t . Hence its linearized flow is $B_{0,t}$, which proves (8.3.8). Equation (8.3.9) then follows from (8.2.42). $\square$

Remark 8.3.4. The preceding result gives an equivalent formulation of the Jacobi equation in the moving symplectic space $T_{\tilde{\lambda}_t}(T^*M)$. A curve $\eta(t) \in T_{\tilde{\lambda}_0}(T^*M)$ solves (8.2.39) if and only if

$$\xi(t) := d_{\tilde{\lambda}_0} \tilde{\Phi}_{0,t}[\eta(t)] \in T_{\tilde{\lambda}_t}(T^*M)$$

solves the linearized super-Hamiltonian equation

$$\dot{\xi}(t) = d_{\tilde{\lambda}_t} \vec{H}_S(t, \cdot)[\xi(t)].$$

Thus, Equation (8.2.39) is exactly the linearized super-Hamiltonian dynamics written in the transported-back reference frame.

8.4 Proof of Proposition 8.3.2

In this section, we prove the identity in Equation (8.3.7). First, we compute the Hessian of φ_t and obtain an explicit expression. Then, show that the quadratic Hamiltonian b_t reduces to the same expression.

Fix $t \in [0, T]$ and set

$$\lambda := \tilde{\lambda}_t, \quad v := \tilde{v}(t), \quad u_i := \tilde{u}_i(t), \quad h_i := h_{u_i}^i, \quad i = 0, 1.$$

We denote by $\hat{u}_i$ the maximizing feedback associated with H^i and set

$$D_i := d_{\lambda} \hat{u}_i, \quad i = 0, 1.$$

Moreover, we introduce the notation

$$\vec{\delta} := \vec{H}^0(\lambda) - \vec{H}^1(\lambda) = \vec{h}_0(\lambda) - \vec{h}_1(\lambda),$$

and we denote by

$$\vec{\delta}^{(1)} := d_{\tilde{\lambda}_0} \tilde{\Phi}_{0,t}[\dot{\Delta}_t].$$

Finally, set

$$K_t := H_S(t, \cdot) - h_{(\tilde{u}(t), \tilde{v}(t))}.$$

By (8.3.6),

$$\varphi_t = K_t \circ \tilde{\Phi}_{0,t}.$$

Since $K_t(\lambda) = 0$ and $d_{\lambda} K_t = 0$, for every $\eta_1, \eta_2 \in T_{\tilde{\lambda}_0}(T^*M)$, setting

$$\xi_i := d_{\tilde{\lambda}_0} \tilde{\Phi}_{0,t}[\eta_i], \quad i = 1, 2,$$

we have

$$\text{Hess}_{\tilde{\lambda}_0} \varphi_t[\eta_1, \eta_2] = \text{Hess}_{\lambda} K_t[\xi_1, \xi_2].$$

We also move the desingularized map to $T_{\lambda}(T^*M)$ by setting

$$\hat{\mathcal{Z}} := d_{\tilde{\lambda}_0} \tilde{\Phi}_{0,t} \circ \mathcal{Z}_t.$$

We also rewrite the generalized Legendre form in the moving symplectic space. To this aim, define

$$\begin{aligned}\widehat{Z}^0 w_0 &:= d_{\tilde{\lambda}_0} \tilde{\Phi}_{0,t} [Z_t^0 w_0] = v \frac{\partial \vec{h}_u^0}{\partial u} \Big|_{u=u_0} [w_0], \\ \widehat{Z}^1 w_1 &:= d_{\tilde{\lambda}_0} \tilde{\Phi}_{0,t} [Z_t^1 w_1] = (1-v) \frac{\partial \vec{h}_u^1}{\partial u} \Big|_{u=u_1} [w_1].\end{aligned}$$

Thus,

$$\widehat{\mathcal{Z}}(w_0, \vartheta, w_1) = \widehat{Z}^0 w_0 - \vartheta \vec{\delta}^{(1)} + \widehat{Z}^1 w_1.$$

Correspondingly, define the moved-forward coefficients

$$\widehat{a}_t^0(w_0) := \sigma_\lambda(\widehat{Z}^0 w_0, \vec{\delta}), \quad \widehat{a}_t^1(w_1) := \sigma_\lambda(\widehat{Z}^1 w_1, \vec{\delta}),$$

and

$$\widehat{\alpha}_t := \sigma_\lambda(\vec{\delta}, \vec{\delta}^{(1)}).$$

The moved-forward generalized Legendre form $\widehat{l}_t$ is then defined, for $(w_0, \vartheta, w_1), (\bar{w}_0, \bar{\vartheta}, \bar{w}_1) \in \mathcal{W}_t$, by

$$\begin{aligned}\widehat{l}_t[(w_0, \vartheta, w_1), (\bar{w}_0, \bar{\vartheta}, \bar{w}_1)] &= \mathbb{L}_t^0[w_0, \bar{w}_0] + \mathbb{L}_t^1[w_1, \bar{w}_1] \\ &\quad + \bar{\vartheta} \widehat{a}_t^0(w_0) + \vartheta \widehat{a}_t^0(\bar{w}_0) \\ &\quad + \bar{\vartheta} \widehat{a}_t^1(w_1) + \vartheta \widehat{a}_t^1(\bar{w}_1) + \widehat{\alpha}_t \vartheta \bar{\vartheta}.\end{aligned}$$

Notice that the ordinary Legendre forms $\mathbb{L}_t^0$ and $\mathbb{L}_t^1$ act on the control spaces and therefore are not affected by the change of symplectic frame. Since $d_{\tilde{\lambda}_0} \tilde{\Phi}_{0,t}$ is symplectic, the definitions in (8.2.41)–(8.2.42) give

$$\widehat{a}_t^0 = a_t^0, \quad \widehat{a}_t^1 = a_t^1, \quad \widehat{\alpha}_t = \alpha_t.$$

Consequently,

$$\widehat{l}_t = l_t.$$

Thus, in what follows, we continue to denote this common bilinear form by l_t .

Since $\tilde{\Phi}_{0,t}$ is Hamiltonian, it is enough to prove

$$\text{Hess}_\lambda K_t[\xi_1, \xi_2] = \left\langle \sigma_\lambda(\widehat{\mathcal{Z}} \cdot, \xi_1), l_t^{-1} \sigma_\lambda(\widehat{\mathcal{Z}} \cdot, \xi_2) \right\rangle. \quad (8.4.1)$$

We first record a preliminary identity that will be used repeatedly below. More precisely, we make the following claim.

Claim 8.4.1. *For every $\xi \in T_\lambda(T^*M)$, we have that*

$$\sigma_\lambda(\xi, \vec{\delta}^{(1)}) = -d_\lambda H^{01}[\xi] + \mathbb{L}_t^0[D_0 \vec{\delta}, D_0 \xi] + \mathbb{L}_t^1[D_1 \vec{\delta}, D_1 \xi]. \quad (8.4.2)$$

We derive this identity explicitly. The ingredients are the differentiated stationarity relation (8.2.80), the formula for $\dot{\Delta}_t$ in (8.2.83), the representation of the maximized bracket in (8.2.90)–(8.2.91), and the definition of the ordinary Legendre forms in (8.2.32).

Set

$$Y := v\vec{h}_0 + (1-v)\vec{h}_1 = \dot{\vec{\lambda}}_t.$$

Since

$$u_i = \widehat{u}_i(\lambda), \quad i = 0, 1,$$

differentiating the feedbacks along the reference extremal gives

$$\dot{u}_i(t) = D_i Y, \quad i = 0, 1.$$

Pushing forward (8.2.83) to $T_\lambda(T^*M)$ therefore yields

$$\vec{\delta}^{(1)} = \vec{h}_t^{10} + \frac{\partial \vec{h}_u^0}{\partial u} \Big|_{u=u_0} [D_0 Y] - \frac{\partial \vec{h}_u^1}{\partial u} \Big|_{u=u_1} [D_1 Y]. \quad (8.4.3)$$

We shall use differentiated stationarity in the following form. For $i = 0, 1$, $w_i \in T_{u_i} U_i$, and $\xi \in T_\lambda(T^*M)$, one has

$$\sigma_\lambda \left(\xi, \frac{\partial \vec{h}_u^i}{\partial u} \Big|_{u=u_i} [w_i] \right) = - \frac{\partial^2 h_u^i}{\partial u^2} \Big|_{(\lambda, u_i)} [w_i, D_i \xi]. \quad (8.4.4)$$

Indeed, the left-hand side is

$$d_\lambda \left(\frac{\partial h_u^i}{\partial u} \Big|_{u=u_i} [w_i] \right) [\xi],$$

and differentiating the stationarity identity

$$\frac{\partial h_u^i}{\partial u} \Big|_{u=\widehat{u}_i(\ell)} = 0$$

at $\ell = \lambda$ in the direction ξ gives precisely (8.4.4).

Pairing (8.4.3) with ξ and using $h_t^{10} = -h_t^{01}$ together with (8.4.4), we obtain

$$\sigma_\lambda(\xi, \vec{\delta}^{(1)}) = -d_\lambda h_t^{01}[\xi] - \frac{\partial^2 h_u^0}{\partial u^2} \Big|_{(\lambda, u_0)} [D_0 Y, D_0 \xi] + \frac{\partial^2 h_u^1}{\partial u^2} \Big|_{(\lambda, u_1)} [D_1 Y, D_1 \xi]. \quad (8.4.5)$$

It remains to replace the frozen-control bracket h_t^{01} by the maximized bracket H^{01} . Recall from (8.2.90)–(8.2.91) that

$$B(\ell, u_0, u_1) = \{h_{u_0}^0, h_{u_1}^1\}(\ell), \quad H^{01}(\ell) = B(\ell, \widehat{u}_0(\ell), \widehat{u}_1(\ell)).$$

The chain rule gives

$$d_\lambda H^{01}[\xi] = d_\lambda h_t^{01}[\xi] + \frac{\partial B}{\partial u_0}[D_0\xi] + \frac{\partial B}{\partial u_1}[D_1\xi]. \quad (8.4.6)$$

We compute the two last terms separately. For $w_0 \in T_{u_0}U_0$, using (8.4.4), we have

$$\begin{aligned} \frac{\partial B}{\partial u_0}[w_0] &= \left\{ \frac{\partial h_u^0}{\partial u} \Big|_{u=u_0} [w_0], h_1 \right\} \\ &= -d_\lambda \left(\frac{\partial h_u^0}{\partial u} \Big|_{u=u_0} [w_0] \right) [\vec{h}_1] = \frac{\partial^2 h_u^0}{\partial u^2} \Big|_{(\lambda, u_0)} [w_0, D_0 \vec{h}_1]. \end{aligned}$$

Similarly, for $w_1 \in T_{u_1}U_1$,

$$\begin{aligned} \frac{\partial B}{\partial u_1}[w_1] &= \left\{ h_0, \frac{\partial h_u^1}{\partial u} \Big|_{u=u_1} [w_1] \right\} \\ &= d_\lambda \left(\frac{\partial h_u^1}{\partial u} \Big|_{u=u_1} [w_1] \right) [\vec{h}_0] = -\frac{\partial^2 h_u^1}{\partial u^2} \Big|_{(\lambda, u_1)} [w_1, D_1 \vec{h}_0]. \end{aligned}$$

Substituting these identities into (8.4.6) yields

$$d_\lambda H^{01}[\xi] = d_\lambda h_t^{01}[\xi] + \frac{\partial^2 h_u^0}{\partial u^2} \Big|_{(\lambda, u_0)} [D_0\xi, D_0\vec{h}_1] - \frac{\partial^2 h_u^1}{\partial u^2} \Big|_{(\lambda, u_1)} [D_1\xi, D_1\vec{h}_0].$$

Using this formula in (8.4.5), we obtain

$$\begin{aligned} \sigma_\lambda(\xi, \vec{\delta}^{(1)}) &= -d_\lambda H^{01}[\xi] + \frac{\partial^2 h_u^0}{\partial u^2} \Big|_{(\lambda, u_0)} [D_0\xi, D_0(\vec{h}_1 - Y)] \\ &\quad + \frac{\partial^2 h_u^1}{\partial u^2} \Big|_{(\lambda, u_1)} [D_1\xi, D_1(Y - \vec{h}_0)]. \end{aligned}$$

Since

$$\vec{h}_1 - Y = -v\vec{\delta}, \quad Y - \vec{h}_0 = -(1-v)\vec{\delta},$$

the preceding identity becomes

$$\begin{aligned} \sigma_\lambda(\xi, \vec{\delta}^{(1)}) &= -d_\lambda H^{01}[\xi] - v \frac{\partial^2 h_u^0}{\partial u^2} \Big|_{(\lambda, u_0)} [D_0\vec{\delta}, D_0\xi] \\ &\quad - (1-v) \frac{\partial^2 h_u^1}{\partial u^2} \Big|_{(\lambda, u_1)} [D_1\vec{\delta}, D_1\xi]. \end{aligned}$$

By the definition of $\mathbb{L}_t^0$ and $\mathbb{L}_t^1$ in (8.2.32), this is exactly (8.4.2). Thus, the claim is proved.

Now, recalling that

$$\mathcal{S} = \{\ell \in \mathcal{O} \mid H^{01}(\ell) = 0\},$$

we have

$$d_\lambda H^{01}[\vec{\delta}] = H^{001}(\lambda) + H^{110}(\lambda) \neq 0.$$

Hence $\vec{\delta}$ is transverse to $\mathcal{S}$, and therefore

$$T_\lambda(T^*M) = \mathbb{R}\vec{\delta} \oplus T_\lambda\mathcal{S}. \quad (8.4.7)$$

Computation of the left-hand side. Using (8.3.3), K_t can be written as

$$K_t = (H_S^0 - H^0) + v(H^0 - h_0) + (1 - v)(H^1 - h_1).$$

At λ , each of the three terms on the right-hand side, as well as its first differential, vanishes.

We first compute the Hessians of the last two terms. The point is that H^i is obtained by substituting the maximizing feedback into h_u^i ,

$$H^i(\ell) = h_{\hat{u}_i(\ell)}^i(\ell),$$

whereas $h_i = h_{u_i}^i$ is obtained by freezing the maximizing control at its reference value $u_i = \hat{u}_i(\lambda)$. The two functions therefore have the same value and the same first differential at λ , while their second-order difference is determined by the variation of the maximizing feedback.

Indeed, we have

$$\left. \frac{\partial h_u^i}{\partial u} \right|_{(\lambda, u_i)} = 0,$$

and hence, in a neighbourhood of λ ,

$$d_\ell H^i[\xi_1] = d_\ell h_{\hat{u}_i(\ell)}^i[\xi_1].$$

Differentiating this identity at $\ell = \lambda$ in the direction ξ_2 gives

$$\begin{aligned} \text{Hess}_\lambda H^i[\xi_1, \xi_2] &= \text{Hess}_\lambda h_i[\xi_1, \xi_2] \\ &\quad + d_\lambda \left(\left. \frac{\partial h_u^i}{\partial u} \right|_{u=u_i} [D_i \xi_2] \right) [\xi_1]. \end{aligned}$$

Applying (8.4.4) with $w_i = D_i \xi_2$, we conclude that

$$\text{Hess}_\lambda (H^i - h_i)[\xi_1, \xi_2] = - \left. \frac{\partial^2 h_u^i}{\partial u^2} \right|_{(\lambda, u_i)} [D_i \xi_1, D_i \xi_2].$$

Multiplying by the corresponding coefficients and using again (8.2.8), we obtain

$$\begin{aligned} v \operatorname{Hess}_\lambda(H^0 - h_0)[\xi_1, \xi_2] &= \mathbb{L}_t^0[D_0\xi_1, D_0\xi_2], \\ (1 - v) \operatorname{Hess}_\lambda(H^1 - h_1)[\xi_1, \xi_2] &= \mathbb{L}_t^1[D_1\xi_1, D_1\xi_2]. \end{aligned} \quad (8.4.8)$$

For the first term, $H_S^0 - H^0$ vanishes on $\mathcal{S}$, while the characteristic equation (8.3.2) gives

$$d(H_S^0 - H^0)[\vec{\delta}] = -H^{01}$$

along $\mathcal{S}$. Since

$$d_\lambda(H_S^0 - H^0) = 0,$$

we obtain

$$\operatorname{Hess}_\lambda(H_S^0 - H^0)[X_1, X_2] = 0, \quad X_1, X_2 \in T_\lambda\mathcal{S},$$

and

$$\operatorname{Hess}_\lambda(H_S^0 - H^0)[\vec{\delta}, \xi] = -d_\lambda H^{01}[\xi].$$

Combining these identities with (8.4.2) and (8.4.8), we find

$$\begin{aligned} \operatorname{Hess}_\lambda K_t[\vec{\delta}, \xi] &= \sigma_\lambda(\xi, \vec{\delta}^{(1)}), \\ \operatorname{Hess}_\lambda K_t[X_1, X_2] &= \mathbb{L}_t^0[D_0X_1, D_0X_2] + \mathbb{L}_t^1[D_1X_1, D_1X_2], \end{aligned} \quad (8.4.9)$$

for every $\xi \in T_\lambda(T^*M)$ and $X_1, X_2 \in T_\lambda\mathcal{S}$.

By (8.4.7), we can write

$$\xi_i = a_i \vec{\delta} + X_i, \quad a_i \in \mathbb{R}, \quad X_i \in T_\lambda\mathcal{S}, \quad i = 1, 2.$$

By bilinearity and symmetry of the Hessian, Equation (8.4.9) yields

$$\begin{aligned} \operatorname{Hess}_\lambda K_t[\xi_1, \xi_2] &= a_1 a_2 \sigma_\lambda(\vec{\delta}, \vec{\delta}^{(1)}) \\ &\quad + a_1 \sigma_\lambda(X_2, \vec{\delta}^{(1)}) + a_2 \sigma_\lambda(X_1, \vec{\delta}^{(1)}) \\ &\quad + \mathbb{L}_t^0[D_0X_1, D_0X_2] + \mathbb{L}_t^1[D_1X_1, D_1X_2]. \end{aligned} \quad (8.4.10)$$

We now show that the right-hand side of (8.4.1) gives exactly the same expression.

Computation of the right-hand side. By the definition of $\mathcal{Z}_t$, the moving desingularized map satisfies

$$\begin{aligned} \widehat{\mathcal{Z}}(w_0, \vartheta, w_1) &= v \frac{\partial \vec{h}_u^0}{\partial u} \Big|_{u=u_0} [w_0] - \vartheta \vec{\delta}^{(1)} \\ &\quad + (1 - v) \frac{\partial \vec{h}_u^1}{\partial u} \Big|_{u=u_1} [w_1]. \end{aligned}$$

Let

$$e := (0, 1, 0) \in \mathcal{W}_t,$$

and, for $X \in T_\lambda \mathcal{S}$, set

$$\omega_X := (-D_0 X, 0, -D_1 X) \in \mathcal{W}_t.$$

We first identify the l_t -representer associated with $\vec{\delta}$. The definitions of a_t^0 , a_t^1 , α_t , and l_t give

$$l_t(\cdot, e) = \sigma_\lambda(\widehat{\mathcal{Z}}\cdot, \vec{\delta}). \quad (8.4.11)$$

We next identify the representer associated with a vector $X \in T_\lambda \mathcal{S}$. Equation (8.4.4) gives, for every $(w_0, 0, w_1) \in \mathcal{W}_t$,

$$\begin{aligned} l_t((w_0, 0, w_1), \omega_X) &= -\mathbb{L}_t^0[w_0, D_0 X] - \mathbb{L}_t^1[w_1, D_1 X] \\ &= \sigma_\lambda(\widehat{\mathcal{Z}}(w_0, 0, w_1), X). \end{aligned}$$

Since $X \in T_\lambda \mathcal{S}$, we have $d_\lambda H^{01}[X] = 0$, and therefore (8.4.2) gives

$$\begin{aligned} l_t(e, \omega_X) &= \mathbb{L}_t^0[D_0 \vec{\delta}, D_0 X] + \mathbb{L}_t^1[D_1 \vec{\delta}, D_1 X] \\ &= \sigma_\lambda(X, \vec{\delta}^{(1)}) \\ &= \sigma_\lambda(\widehat{\mathcal{Z}}e, X). \end{aligned}$$

By bilinearity, it follows that

$$l_t(\cdot, \omega_X) = \sigma_\lambda(\widehat{\mathcal{Z}}\cdot, X), \quad X \in T_\lambda \mathcal{S}. \quad (8.4.12)$$

Consequently, if

$$\xi = a\vec{\delta} + X, \quad X \in T_\lambda \mathcal{S},$$

then (8.4.11) and (8.4.12) imply

$$l_t(\cdot, ae + \omega_X) = \sigma_\lambda(\widehat{\mathcal{Z}}\cdot, \xi).$$

By Proposition 8.2.8, the form l_t is non-degenerate, and hence

$$l_t^{-1}\sigma_\lambda(\widehat{\mathcal{Z}}\cdot, \xi) = ae + \omega_X.$$

It follows that the right-hand side of (8.4.1) is

$$\begin{aligned} &\left\langle \sigma_\lambda(\widehat{\mathcal{Z}}\cdot, \xi_1), l_t^{-1}\sigma_\lambda(\widehat{\mathcal{Z}}\cdot, \xi_2) \right\rangle \\ &= l_t(a_1e + \omega_{X_1}, a_2e + \omega_{X_2}). \end{aligned}$$

It remains to expand the last expression. From (8.4.11) and the fact that

$$\widehat{\mathcal{Z}}e = -\vec{\delta}^{(1)},$$

we obtain

$$l_t(e, e) = \sigma_\lambda(\vec{\delta}, \vec{\delta}^{(1)}).$$

Similarly, (8.4.12) gives

$$l_t(e, \omega_X) = \sigma_\lambda(X, \vec{\delta}^{(1)}).$$

Finally, applying (8.4.4) with $(w_0, w_1) = (-D_0X_1, -D_1X_1)$ gives

$$l_t(\omega_{X_1}, \omega_{X_2}) = \mathbb{L}_t^0[D_0X_1, D_0X_2] + \mathbb{L}_t^1[D_1X_1, D_1X_2].$$

Therefore,

$$\begin{aligned} & \left\langle \sigma_\lambda(\widehat{\mathcal{Z}} \cdot, \xi_1), l_t^{-1} \sigma_\lambda(\widehat{\mathcal{Z}} \cdot, \xi_2) \right\rangle \\ &= a_1 a_2 \sigma_\lambda(\vec{\delta}, \vec{\delta}^{(1)}) \\ & \quad + a_1 \sigma_\lambda(X_2, \vec{\delta}^{(1)}) + a_2 \sigma_\lambda(X_1, \vec{\delta}^{(1)}) \\ & \quad + \mathbb{L}_t^0[D_0X_1, D_0X_2] + \mathbb{L}_t^1[D_1X_1, D_1X_2]. \end{aligned}$$

This is exactly the expression obtained for the left-hand side in (8.4.10), and hence (8.4.1) follows.

Finally, recalling that

$$\xi_i = d_{\tilde{\lambda}_0} \tilde{\Phi}_{0,t}[\eta_i], \quad \widehat{\mathcal{Z}} = d_{\tilde{\lambda}_0} \tilde{\Phi}_{0,t} \circ \mathcal{Z}_t,$$

the symplecticity of $\tilde{\Phi}_{0,t}$ gives

$$\sigma_\lambda(\widehat{\mathcal{Z}} \cdot, \xi_i) = \sigma(\mathcal{Z}_t \cdot, \eta_i), \quad i = 1, 2.$$

Substituting this identity into (8.4.1), we conclude that

$$\text{Hess}_{\tilde{\lambda}_0} \varphi_t[\eta_1, \eta_2] = \left\langle \sigma(\mathcal{Z}_t \cdot, \eta_1), l_t^{-1} \sigma(\mathcal{Z}_t \cdot, \eta_2) \right\rangle.$$

This is (8.3.7) and completes the proof.

8.5 Jacobi curve for a nice extremal

The aim of this section is to establish a characterization of the Jacobi curve associated with a nice extremal pair $(\tilde{u}, \tilde{\lambda})$ via an appropriate algorithm, similarly to what is done in Theorem 6.1.6. We briefly state the definition of endpoint map and derive the first and second differential, which are needed for the proof of the main theorem of the section. Since this first part is very close to the corresponding part for the piecewise regular extremal, we treat this part schematically just underlining the principal differences.

8.5.1 Endpoint map near a nice extremal

Let $(\tilde{u}, \tilde{\lambda})$ be a nice extremal pair and let

$$0 = t_0 < t_1 < \cdots < t_k < t_{k+1} = T$$

be a partition as in Definition 8.1.4. We set

$$I_j := (t_{j-1}, t_j), \quad j = 1, \dots, k+1,$$

and denote by $\mathcal{R}$ and $\mathcal{S}$ the sets of indices corresponding respectively to regular and first-order singular arcs.

We construct a local space of competitors by combining the constructions of Sections 7.1 and 8.2.1. If $j \in \mathcal{R}$, we use time-dependent local coordinates on the unique maximizing submanifold U_j containing $\tilde{u}|_{I_j}$, exactly as in Section 7.1, and denote the corresponding local control by u^j . If $j \in \mathcal{S}$, we denote by $U_{j,0}$ and $U_{j,1}$ the two active submanifolds and use the reduced variables

$$u^j = (u_0^j, u_1^j), \quad v^j \in (0, 1),$$

introduced in Subsection 8.2.1.

As in Section 7.1, we also introduce a time-reparametrization variable $\nu \in L^\infty([0, T])$, with $1 + \nu > 0$, in order to encode variations of the switching times. We denote by $\mathbf{u}$ the collection of all the local control variables just introduced and by $\tilde{\mathbf{u}}$ its reference value, that is

$$\mathbf{u}(t) = \begin{cases} u(t) & \text{if } t \in I_j \text{ and } j \in \mathcal{R}, \\ (u_0(t), v(t), u_1(t)) & \text{if } t \in I_j \text{ and } j \in \mathcal{S}. \end{cases}$$

For $t \in I_j$, set

$$h_{\mathbf{u}(t)} := \begin{cases} h_{u^j(t)}^j, & j \in \mathcal{R}, \\ h_{(u^j(t), v^j(t))}^j, & j \in \mathcal{S}, \end{cases}$$

where, on a singular interval,

$$h_{(u^j, v^j)} = v^j h_{u_0^j}^{j,0} + (1 - v^j) h_{u_1^j}^{j,1}.$$

We denote by $\mathcal{U}_{\text{nice}}$ the resulting Banach manifold of local competitors $(\nu, \mathbf{u})$. The corresponding state endpoint map and extended endpoint map are defined as in Sections 7.1 and 8.2.1; on a regular interval the running cost is L , whereas on a singular interval it is the reduced relaxed cost introduced in (8.1.11). We denote them by

$$F : \mathcal{U}_{\text{nice}} \longrightarrow M, \quad E = (F, \overline{J}).$$

Fixing the initial covector $\tilde{\lambda}_0$, the Hamiltonian lift of the endpoint map is

$$\mathcal{E}(\nu, \mathbf{u}) = \overrightarrow{\text{exp}} \int_0^T (1 + \nu(t)) \vec{h}_{\mathbf{u}(t)} dt \circ \tilde{\lambda}_0.$$

Let $\tilde{\Phi}_{s,t}$ denote the Hamiltonian flow of the control corresponding to the reference extremal. As before, we transport the endpoint map back to the fixed initial symplectic space and set

$$\mathcal{G}(\nu, \mathbf{u}) := \tilde{\Phi}_{0,T}^{-1} \circ \mathcal{E}(\nu, \mathbf{u}), \quad G := \pi \circ \mathcal{G}.$$

In particular,

$$\mathcal{G}(0, \tilde{\mathbf{u}}) = \tilde{\lambda}_0, \quad G(0, \tilde{\mathbf{u}}) = q_0.$$

Finally, we denote the tangent space of variations by

$$\mathcal{V} := T_{(0, \tilde{\mathbf{u}})} \mathcal{U}_{\text{nice}}.$$

A variation is written on a regular interval as

$$w|_{I_j} = (\theta, w^j), \quad j \in \mathcal{R},$$

and on a singular interval as

$$w|_{I_j} = (\theta, w_0^j, \omega^j, w_1^j), \quad j \in \mathcal{S}.$$

For $t \in [0, T]$, we use the same support filtration as in Section 7.3.2,

$$\mathcal{V}_t := \{w \in \mathcal{V} \mid \text{supp } w \subset [0, t]\}.$$

The Jacobi curve considered below is the $\mathcal{L}$ -derivative associated with this filtration.

8.5.2 First and second differential of the endpoint map

We now record only the formulas needed to determine the Jacobi curve. The computations are the same as in Section 7.3.2, with the singular contribution on each interval I_j , $j \in \mathcal{S}$, replaced by the one derived in Section 8.2.

Define the transported perturbation Hamiltonian by

$$\mathbf{h}_t(\lambda, \nu, \mathbf{u}) := \left((1 + \nu) h_{\mathbf{u}(t)} - h_{\tilde{\mathbf{u}}(t)} \right) \circ \tilde{\Phi}_{0,t}(\lambda).$$

For $w \in \mathcal{V}$, let

$$X_t[w(t)] := D_{(\nu, \mathbf{u})} \vec{\mathbf{h}}_t(\tilde{\lambda}_0, 0, \tilde{\mathbf{u}})[w(t)].$$

On a regular interval I_j , the notation of Section 7.3.2 gives

$$X_t[w(t)] = \theta(t) \vec{\mathcal{H}}_t^j + Z_t^j w^j(t), \quad j \in \mathcal{R}.$$

On a singular interval, denote by

$$\Delta_t^j := \vec{\mathcal{H}}_t^{j,0} - \vec{\mathcal{H}}_t^{j,1}$$

the distinguished singular direction and by

$$\vec{\mathcal{H}}_t^{j,s} := \tilde{v}^j(t) \vec{\mathcal{H}}_t^{j,0} + (1 - \tilde{v}^j(t)) \vec{\mathcal{H}}_t^{j,1}$$

the transported vector field of the reference singular dynamics. Then the first variation has the decomposition

$$X_t[w(t)] = \theta(t) \vec{\mathcal{H}}_t^{j,s} + Z_t^{j,0} w_0^j(t) + \omega^j(t) \Delta_t^j + Z_t^{j,1} w_1^j(t), \quad j \in \mathcal{S}.$$

The general first-order variation formula therefore gives

$$DG[w] = d_{\tilde{\lambda}_0} \pi \left[\int_0^T X_t[w(t)] dt \right]. \quad (8.5.1)$$

Consequently, denoting by $\Pi = T_{\tilde{\lambda}_0}^*(T_{q_0}^* M)$, the endpoint-preserving kernel is

$$\mathcal{K} = \left\{ w \in \mathcal{V} \mid \int_0^T X_t[w(t)] dt \in \Pi \right\}.$$

Similarly, the second differential is obtained directly from the general formula of Section 7.3.2. The associated quadratic form on $\mathcal{V}$ is

$$Q(w) = - \int_0^T D^2 \mathbf{h}_t[w(t)]^2 dt - \int_0^T \int_0^t \sigma(X_s[w(s)], X_t[w(t)]) ds dt. \quad (8.5.2)$$

Its restriction to $\mathcal{K}$ coincides with the constrained second differential of the endpoint map.

For later use, it remains only to identify the form taken by these expressions when the variations are localized inside one of the intervals I_j . If $j \in \mathcal{R}$, the restriction of (8.5.1)–(8.5.2) to I_j is exactly the one studied in Sections 7.3.2 and 7.4.1. In particular, after the contribution of the time variations has been incorporated into the initial Lagrangian subspace, the remaining control variations are described by Z_t^j and by the ordinary Legendre form $\mathbb{L}_t^j$, and generate the regular Jacobi equation.

Suppose instead that $j \in \mathcal{S}$ and set $a = t_{j-1}$. For a variation supported in $[a, t] \subset [t_{j-1}, t_j]$, introduce the Goh primitive

$$\vartheta^j(s) := \int_a^s \omega^j(r) dr$$

and the transformed distributed variation

$$\zeta^j = (w_0^j, \vartheta^j, w_1^j).$$

Repeating the construction of Subsections 8.2.3 and 8.2.4 on the interval $[a, t]$, the contribution of the singular control variations to the first differential extends continuously to

$$\widehat{A}_{a,t}^j(\vartheta_t^j, \zeta^j) = \vartheta_t^j \Delta_t^j + \int_a^t \mathcal{Z}_s^j \zeta^j(s) ds.$$

The corresponding quadratic form is the translated version of the augmented second variation of Subsection 8.2.4, with generalized Legendre form l_s^j and desingularized map $\mathcal{Z}_s^j$. By Proposition 8.2.4, replacing the original singular variations with these augmented variations does not change the associated $\mathcal{L}$ -derivative.

Therefore, on each regular interval the localized variational problem is exactly the one used in the proof of Theorem 7.4.1, whereas on each singular interval it is exactly the translated version of the variational problem studied in Subsection 8.2.6. The time variations can be separated from the local control variations by the same argument used in Proposition 7.3.2. These observations are the only ingredients from the endpoint-map computation that will be needed in the proof of Theorem 8.5.1.

8.5.3 Structure of the Jacobi curve for a nice extremal

We can now characterize the Jacobi curve of a nice extremal by combining the recursive construction of Theorem 7.4.1 with the analysis of a first-order singular arc developed in Section 8.2. The preceding subsections show that, after the contribution of the variations supported before t_{j-1} has been incorporated into $\Lambda_{t_{j-1}}$, the variational problem on (t_{j-1}, t_j) is precisely one of the two problems already considered: the regular one of Section 7.4, or the first-order singular one of Subsection 8.2.6.

For every $j = 1, \dots, k+1$, let $B_{s,t}^j$ denote the linear symplectic flow associated with the corresponding Jacobi equation on (t_{j-1}, t_j) . If the j -th arc is regular, this is the regular Jacobi flow of Theorem 7.4.1. If the j -th arc is first-order singular, it is the flow generated by

$$\dot{\eta}(t) = -\mathcal{Z}_t^j(l_t^j)^{-1} \sigma(\mathcal{Z}_t^j \cdot, \eta(t)). \quad (8.5.3)$$

In the latter case, by Corollary 8.3.3, this flow is equivalently the linearization of the flow φ_t introduced before, after identifying the tangent spaces through the reference Hamiltonian flow.

Theorem 8.5.1 (Structure of the Jacobi curve). *Let $(\tilde{u}, \tilde{\lambda})$ be a nice extremal pair associated with the partition*

$$0 = t_0 < t_1 < \dots < t_k < t_{k+1} = T.$$

Denote by

$$\Lambda : [0, T] \longrightarrow \mathfrak{L}\left(T_{\tilde{\lambda}_0}(T^*M)\right)$$

the corresponding Jacobi curve. Then Λ is a left-continuous piecewise smooth curve whose possible discontinuities are $t_0, \dots, t_k$. It is determined recursively by setting

$$\Lambda_0 = \Pi$$

and applying the following rules on every interval $(t_{j-1}, t_j]$.

1. Suppose that the j -th arc is regular and denote by $\vec{\mathcal{H}}_t^j$ its transported Hamiltonian vector field. Then

$$\Lambda_{t_{j-1}}^+ = \Lambda_{t_{j-1}}^{\mathbb{R}\vec{\mathcal{H}}_{t_{j-1}}^j}, \quad (8.5.4)$$

and

$$\Lambda_t = B_{t_{j-1}, t}^j \Lambda_{t_{j-1}}^+, \quad t \in (t_{j-1}, t_j]. \quad (8.5.5)$$

In particular,

$$\vec{\mathcal{H}}_t^j \in \Lambda_t, \quad t \in (t_{j-1}, t_j].$$

2. Suppose that the j -th arc is first-order singular, with active Hamiltonians $H^{j,0}$ and $H^{j,1}$, and set

$$\Delta_t^j = \vec{\mathcal{H}}_t^{j,0} - \vec{\mathcal{H}}_t^{j,1}.$$

Then

$$\Lambda_{t_{j-1}}^+ = \Lambda_{t_{j-1}}^{\mathbb{R}\Delta_{t_{j-1}}^j}, \quad (8.5.6)$$

and

$$\Lambda_t = B_{t_{j-1}, t}^j \Lambda_{t_{j-1}}^+, \quad t \in (t_{j-1}, t_j]. \quad (8.5.7)$$

In particular,

$$\Delta_t^j \in \Lambda_t, \quad t \in (t_{j-1}, t_j]. \quad (8.5.8)$$

Proof. The proof follows recursively the same argument used for Theorem 7.4.1. Assume that $\Lambda_{t_{j-1}}$ has already been determined and fix $t \in (t_{j-1}, t_j]$. By the causality property of the $\mathcal{L}$ -derivative, all variations supported in $[0, t_{j-1}]$ are completely encoded by the initial Lagrangian subspace $\Lambda_{t_{j-1}}$, so that it remains only to add the variations supported in (t_{j-1}, t) .

If the j -th arc is regular, there is nothing new to prove. The argument of Theorem 7.4.1 applies verbatim after translating the initial time from 0 to t_{j-1} and replacing the initial vertical subspace by $\Lambda_{t_{j-1}}$. The time variations give the new initial subspace (8.5.4), while the ordinary control variations generate the regular Jacobi flow. This gives (8.5.5) and the inclusion $\vec{\mathcal{H}}_t^j \in \Lambda_t$.

Suppose now that the j -th arc is first-order singular. We apply on $[t_{j-1}, t]$ the construction of Section 8.2, again replacing the initial vertical subspace Π by the already constructed space $\Lambda_{t_{j-1}}$. By Proposition 8.2.4, passing from the original variations to the augmented space does not modify the $\mathcal{L}$ -derivative. The independent terminal variable

ϑ_T produces precisely the symplectic transformation yielding the new initial subspace, as in Theorem 8.2.11, and therefore gives (8.5.6). The distributed augmented variations are governed by the generalized Legendre form l_t^j and the map $\mathcal{Z}_t^j$, hence their contribution evolves according to (8.5.3). This proves (8.5.7). Moreover, Lemma 8.2.10, translated to the interval $[t_{j-1}, t_j]$, gives (8.5.8). By Corollary 8.3.3, the same evolution can equivalently be described as the linearization of the flow φ_t of Section 8.3.

Iterating these two constructions for $j = 1, \dots, k + 1$ determines the Jacobi curve on the whole interval. Each of its restrictions to (t_{j-1}, t_j) is smooth, while Λ_{t_j} is obtained by left continuation from the preceding arc. Hence Λ is left-continuous and piecewise smooth, with possible right jumps only at the initial time and at the junction times. $\square$

Remark 8.5.2. Along a first-order singular arc, both active Hamiltonian directions belong to the Jacobi curve. Indeed, ordinary time variations give

$$\vec{\mathcal{H}}_t^{j,s} = \tilde{v}^j(t) \vec{\mathcal{H}}_t^{j,0} + (1 - \tilde{v}^j(t)) \vec{\mathcal{H}}_t^{j,1} \in \Lambda_t,$$

whereas (8.5.8) gives

$$\vec{\mathcal{H}}_t^{j,0} - \vec{\mathcal{H}}_t^{j,1} \in \Lambda_t.$$

Since $\tilde{v}^j(t) \in (0, 1)$, it follows that

$$\vec{\mathcal{H}}_t^{j,0}, \vec{\mathcal{H}}_t^{j,1} \in \Lambda_t. \quad (8.5.9)$$

In particular, if a first-order singular arc is followed at t_j by a regular arc, the remark following Definition 8.1.4 shows that the Hamiltonian of the outgoing regular arc is one of the two active Hamiltonians of the singular arc. Hence its transported Hamiltonian vector field already belongs to Λ_{t_j} by (8.5.9), and the extension prescribed by (8.5.4) is trivial:

$$\Lambda_{t_j}^+ = \Lambda_{t_j}.$$

Thus the Jacobi curve is continuous at every junction from a first-order singular arc to a regular arc. At a regular–singular junction, instead, the possible right jump is the extension in the singular direction $\mathbb{R}\Delta_{t_j}^{j+1}$ prescribed by (8.5.6).

8.6 The super-Hamiltonian near regular–singular junctions

In this section, we extend the construction exposed in Section 8.3 of the super-Hamiltonian function for a purely first-order singular arc to a nice extremal containing a first-order singular arc. The only new point is the behaviour of the field of extremals at the junctions between regular and singular arcs. We follow the construction introduced for bang–singular extremals in [64] and subsequently developed in [61], adapting it to the super-Hamiltonian of Section 8.3.

This part is still a work in progress. For this reason, the arguments leading to some of the results of this section are not presented in full detail. In particular, in the end of the section we state what we think is the natural conjecture about a sufficient optimality condition for nice extremal without conjugate points. However, it does not follow directly from the arguments presented here or present in the literature, but we think that it is achievable by some refinement of these techniques.

For simplicity, we consider a nice extremal with the following regular–singular–regular structure:

- the first arc, defined on $[0, t_1]$, is regular and is generated by the maximized Hamiltonian H^0 ;
- the second arc, defined on $[t_1, t_2]$, is a first-order singular arc with simultaneous maximized Hamiltonians H^0 and H^1 ;
- the final arc, defined on $[t_2, T]$, is regular and generated by H^1 .

We denote by H_S^0 , H_S^1 , and H_S the Hamiltonians constructed in Section 8.3. Recall in particular that

$$\{H_S^0, H^1 - H^0\}(\lambda) = 0, \quad \{H_S^1, H^1 - H^0\}(\lambda) = 0, \quad \lambda \in T^*M,$$

so that the Hamiltonian flows of both H_S^0 and H_S^1 leave Σ invariant, while

$$H_S(t, \lambda) = H_S^0(\lambda) + (1 - \tilde{v}(t))(H^1 - H^0)(\lambda)$$

is the super-Hamiltonian associated with the singular arc.

In Section 7.5, we proved that, for a piecewise regular extremal, the maximized Hamiltonian generates an Hamiltonian flow which can be used to construct a field of extremals. For a first-order singular arc, the maximized Hamiltonian do not produce a well defined flow in a neighbourhood of the reference extremal. However, in Section 8.3 we showed that a smooth super-Hamiltonian near the singular extremal can be constructed, so that Theorem 4.1.3 can be still applied to prove local optimality. For a nice extremal, however, simply concatenating these two constructions is not enough to obtain a field of extremals.

The obstruction occurs at the regular–singular junctions. Around these points, the maximized Hamiltonian flow divides any neighbourhood of the preceding regular arc into two families (see Figure 8.1d): trajectories on one side reach and cross the switching hypersurface Σ , whereas trajectories on the other side do not reach Σ . The trajectories reaching Σ tangentially form the separating family, and the reference extremal belongs precisely to this boundary. This is the local geometry underlying the bang–singular construction in [64, 61]. Consequently, the field obtained from the maximized Hamiltonian alone is one-sided at the reference regular–singular extremal, and its projection cannot cover a full neighbourhood of the reference trajectory in M . To overcome this problem, one can introduce an additional segment of the dynamics generated by the Hamiltonians H_S^0 and H_S^1 .

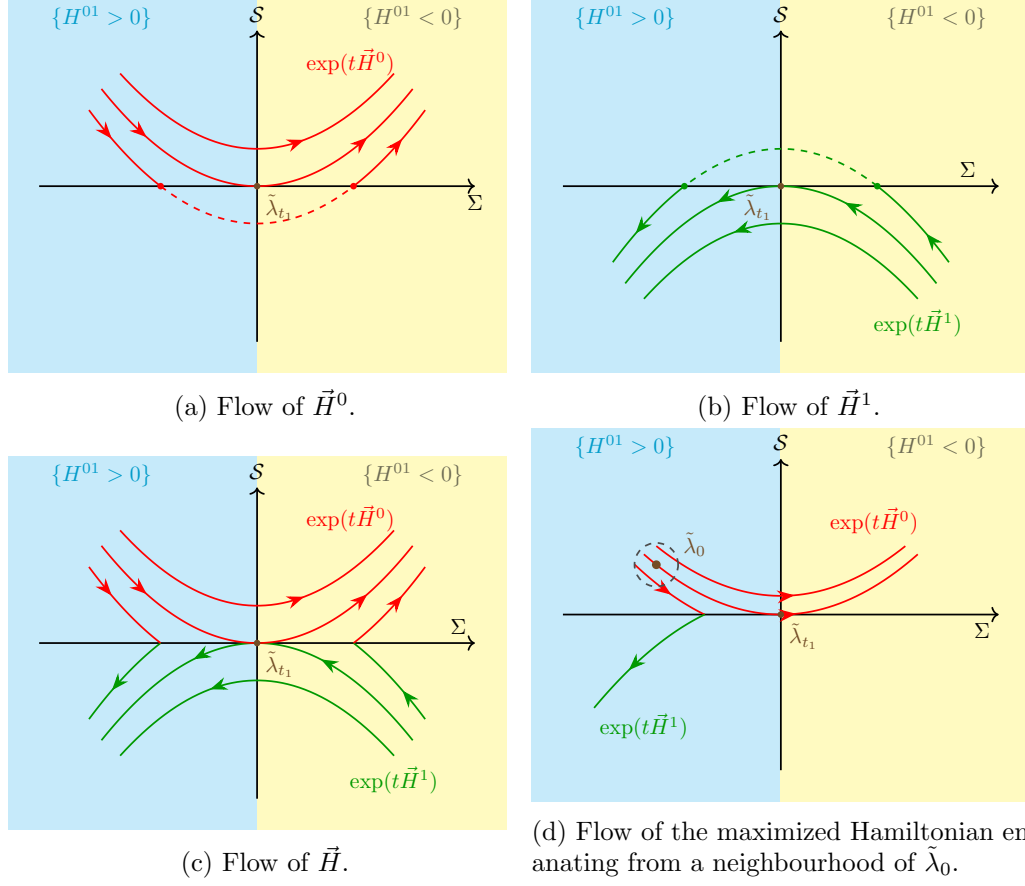


Figure 8.1: Flow of the maximized Hamiltonians in a neighbourhood of $\tilde{\lambda}_{t_1}$. The horizontal axis correspond to $\Sigma = \{H^0 = H^1\}$, while the vertical axis is $S = \{H^{01} = 0\}$. Above Σ we have $H^0 > H^1$, below Σ we have $H^0 < H^1$.

First of all, we make this local geometry precise, that is we show that either trajectories of $\vec{H}^0$ nearby $\tilde{\lambda}$ crosses Σ , either they meet Σ transversally or they do not reach it.

Proposition 8.6.1 (Local geometry of a regular–singular junction). *Let $(\tilde{u}, \tilde{\lambda})$ have the regular–singular–regular structure described above, and let $t \in (0, t_1)$. Then, there is a sufficiently small neighbourhood $\mathcal{N}_t$ of $\tilde{\lambda}_t$ and $\delta > 0$ such that $\mathcal{N}_t$ is partitioned into three non-empty sets: letting $\lambda_s = \exp(s\vec{H}^0)(\lambda)$, the three sets are*

$$\begin{aligned} \mathcal{N}_t^1 &:= \{\lambda \in \mathcal{N}_t \mid \exists! s \in [0, t_1 - t + \delta], \lambda_s \in \Sigma, H^{01}(\lambda_s) \neq 0\}, \\ \mathcal{N}_t^2 &:= \{\lambda \in \mathcal{N}_t \mid \exists! s \in [0, t_1 - t + \delta], \lambda_s \in \Sigma, H^{01}(\lambda_s) = 0\}, \\ \mathcal{N}_t^3 &:= \{\lambda \in \mathcal{N}_t \mid H^0(\lambda_s) > H^1(\lambda_s), \forall s \in [0, t_1 - t + \delta]\}. \end{aligned}$$

Moreover, $\mathcal{N}_t^2$ is a smooth hypersurface through $\tilde{\lambda}_t$.

Proof. Since $H^{01}(\tilde{\lambda}_{t_1}) = 0$ and $\vec{H}^0(H^{01})(\tilde{\lambda}_{t_1}) = H^{001}(\tilde{\lambda}_{t_1}) \neq 0$, by implicit function theorem there is a neighbourhood $\mathcal{N}_t \subset T^*M$ of $\tilde{\lambda}_t$ and a function $\tau : \mathcal{N}_t \rightarrow \mathbb{R}$ such that, $H^{01}(\lambda_{\tau(\lambda)}) = 0$ for every $\lambda \in \mathcal{N}_t$. Moreover, for every $\lambda \in \mathcal{N}_t$, the function

$$s \mapsto H^0(\lambda_s) - H^1(\lambda_s),$$

has a critical point at $s = \tau(\lambda)$, since $\frac{d}{ds}[H^0(\lambda_s) - H^1(\lambda_s)] = H^{10}(\lambda_s)$. Moreover

$$\frac{d^2}{ds^2}[H^0(\lambda_s) - H^1(\lambda_s)] = -H^{001}(\lambda_s),$$

and $-H^{001}(\tilde{\lambda}_{t_1}) > 0$, thus the previous critical points are local minima. The value of $H^1 - H^0$ at this local minima is either negative, zero, or positive according as the corresponding trajectory reaches Σ transversally, reaches it tangentially, or does not reach it (see Figure 8.1a). Since Σ is a smooth hypersurface, the zero level of this maximal value is itself a smooth hypersurface after reducing $\mathcal{N}_t$. This gives the stated decomposition. $\square$

In view of Proposition 8.6.1, the maximized Hamiltonian flow has to be modified on the side which does not provide a continuation compatible with the singular field.

Recall that, by the definitions given in Section 8.3, we have

$$H_S^0(\lambda) = H^0(\lambda), \quad d_\lambda H_S^0 = d_\lambda H^0 \quad \lambda \in \mathcal{S}, \quad (8.6.1)$$

$$H_S^1(\lambda) = H^1(\lambda), \quad d_\lambda H_S^1 = d_\lambda H^1 \quad \lambda \in \mathcal{S}. \quad (8.6.2)$$

In particular, $\{H_S^0, H^{01}\}(\lambda_{t_1}) = H^{001}(\lambda_{t_1}) \neq 0$, and similarly $\{H_S^0, H^{01}\}(\lambda_{t_2}) \neq 0$, so that $\vec{H}_S^0(\tilde{\lambda}_{t_1})$ and $\vec{H}_S^1(\tilde{\lambda}_{t_2})$ are transversal to $\mathcal{S}$. This fact allow us to define the following correcting dynamics.

Proposition 8.6.2 (Hitting times at the junctions). *There exist relatively open neighbourhoods $\mathcal{N}_1$ and $\mathcal{N}_2$ in Σ of $\tilde{\lambda}_{t_1}$ and $\tilde{\lambda}_{t_2}$ respectively and smooth functions τ_0 and τ_1 , defined respectively on these neighbourhoods, such that*

$$H^{01}\left(\exp((\tau_0(\lambda) - t_1)\vec{H}_S^0)(\lambda)\right) = 0, \quad \lambda \in \mathcal{N}_1$$

and

$$H^{01}\left(\exp((\tau_1(\lambda) - t_2)\vec{H}_S^1)(\lambda)\right) = 0 \quad \lambda \in \mathcal{N}_2.$$

In particular,

$$\tau_0(\tilde{\lambda}_{t_1}) = t_1, \quad \tau_1(\tilde{\lambda}_{t_2}) = t_2.$$

For points sufficiently close to the reference junctions satisfying $H^{01}(\lambda) < 0$, one has

$$\tau_0(\lambda) < t_1, \quad \tau_1(\lambda) > t_2.$$

Moreover, for $i = 1, 2$, $\tau_i(\lambda) = t_i$ if and only if $\lambda \in \Sigma \cap \mathcal{S}$.

Proof. For the incoming junction, apply the implicit function theorem to

$$H^{01} \left(\exp((s - t_1)\vec{H}_S^0)(\lambda) \right) = 0$$

near $(s, \lambda) = (t_1, \tilde{\lambda}_{t_1})$. The derivative with respect to s at this point is, by (8.6.1),

$$\{H_S^0, H^{01}\}(\tilde{\lambda}_{t_1}) = H^{001}(\tilde{\lambda}_{t_1}) < 0,$$

so that a unique smooth solution $s = \tau_0(\lambda)$ exists. Since the H_S^0 -flow preserves Σ , the corresponding hitting point belongs to $\Sigma \cap \mathcal{S}$, and the sign of the preceding derivative gives $\tau_0(\lambda) < t_1$ when $H^{01}(\lambda) < 0$. The outgoing junction is analogous, since

$$\{H_S^1, H^{01}\} = -H^{110} > 0$$

on $\Sigma \cap \mathcal{S}$, which gives the function τ_1 and the inequality $\tau_1(\lambda) > t_2$. $\square$

We now use the hitting times introduced above to construct a local field of extremals through the regular–singular–regular block.

Definition 8.6.3 (Regular-singular-regular field of extremals). Consider a sufficiently small relatively open neighbourhood $\mathcal{N}_1$ of $\tilde{\lambda}_{t_1}$ in Σ and define $\mathcal{N}_2 := \Phi_{t_1, t_2}^S(\mathcal{N}_1)$. Since the flow of H_S preserves Σ , after reducing the neighbourhood if necessary this produces a neighbourhood of the reference singular arc contained in Σ , whose trajectories have their values at t_1 and t_2 in $\mathcal{N}_1$ and $\mathcal{N}_2$, respectively.

We extend each of these trajectories through the two junctions with the regular arcs as follows. More precisely, for each point $\lambda \in \mathcal{N}_1$, we construct a trajectory $[0, T] \ni t \mapsto \lambda_t$ as follows:

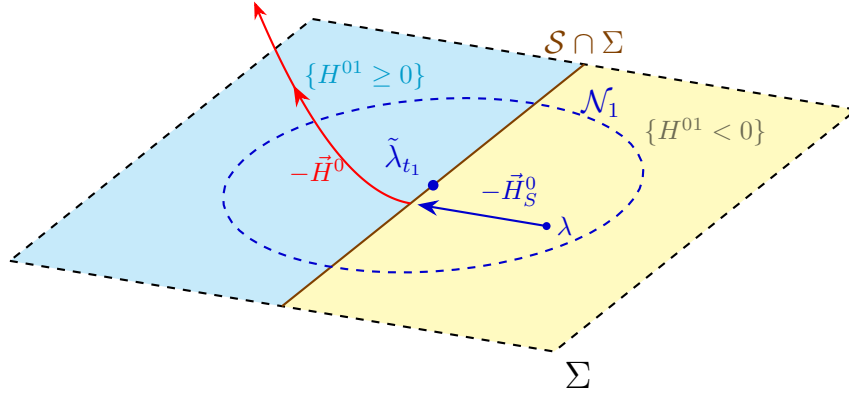
1. For $\lambda \in \mathcal{N}_1$, we set $\lambda = \lambda_{t_1}$. Then, if $H^{01}(\lambda) \geq 0$, the trajectory is extended backwards directly by the flow of H^0 , that is

$$\lambda_t = \exp(-(t_1 - t)\vec{H}^0)(\lambda), \text{ for } t \in [0, t_1].$$

If instead $H^{01}(\lambda) < 0$, it is first extended backwards by the flow of H_S^0 , for a time equal to $\tau_0(\lambda)$, so that it meets $\mathcal{S}$, and then continued backwards by the flow of H^0 . That is

$$\lambda_t = \begin{cases} \exp(-(t_1 - t)\vec{H}_S^0)(\lambda), & t \in [\tau_0(\lambda), t_1], \\ \exp(-(\tau_0(\lambda) - t)\vec{H}^0)(\lambda_{\tau_0(\lambda)}), & t \in [0, \tau_0(\lambda)]. \end{cases}$$

2. On the singular interval $[t_1, t_2]$, the trajectory is generated by the time-dependent Hamiltonian H_S .

Figure 8.2: Graphical representation of the construction on the side $H^{01} < 0$.

3. For $\lambda \in \mathcal{N}_2$ if $H^{01}(\lambda) \geq 0$, the trajectory is extended forward directly by the flow of H^1 , that is

$$\lambda_t = \exp((t - t_2)\vec{H}^1)(\lambda), \text{ for } t \in [t_2, T].$$

If $H^{01}(\lambda) < 0$, it is first propagated forward by the flow of H_S^1 for a time equal to $\tau_1(\lambda)$, so that it meets $\mathcal{S}$, and then continued forward by the flow of H^1 . That is

$$\lambda_t = \begin{cases} \exp((t - t_2)\vec{H}_S^0)(\lambda), & t \in [t_2, \tau_1(\lambda)], \\ \exp((t - \tau_1(\lambda))\vec{H}^0)(\lambda_{\tau_1(\lambda)}), & t \in [\tau_1(\lambda), T]. \end{cases}$$

We call the resulting family a regular-singular-regular field of extremals.

The corrected portions generated by H_S^0 and H_S^1 remain in Σ . At their intersections with $\mathcal{S}$, Equations (8.6.1) and (8.6.2) imply that the corresponding Hamiltonian vector fields agree with $\vec{H}^0$ and $\vec{H}^1$, respectively. Hence the corrected and ordinary portions join with the same tangent vector.

After reducing the neighbourhoods if necessary, we can choose $\varepsilon > 0$ such that all the relevant incoming and outgoing hitting times satisfy

$$t_1 - \frac{\varepsilon}{2} < \tau_0(\lambda) \leq t_1, \quad t_2 \leq \tau_1(\lambda) < t_2 + \frac{\varepsilon}{2}.$$

We can therefore prescribe the Hamiltonian generating this family of trajectories by

$$H^S(t, \lambda) := \begin{cases} H^0(\lambda), & 0 \leq t \leq t_1 - \varepsilon, \\ H^0(\lambda), & t_1 - \varepsilon < t < t_1, \quad H^{01}(\lambda) \geq 0, \\ H_S^0(\lambda), & t_1 - \varepsilon < t < t_1, \quad H^{01}(\lambda) < 0, \\ H_S(t, \lambda), & t_1 \leq t \leq t_2, \\ H_S^1(\lambda), & t_2 < t < t_2 + \varepsilon, \quad H^{01}(\lambda) < 0, \\ H^1(\lambda), & t_2 < t < t_2 + \varepsilon, \quad H^{01}(\lambda) \geq 0, \\ H^1(\lambda), & t_2 + \varepsilon \leq t \leq T. \end{cases} \quad (8.6.3)$$

This definition describes the Hamiltonian dynamics along the corrected field constructed above. To obtain a Hamiltonian defined on a full neighbourhood of this field, we can follow the local matching procedure introduced in [64, 61].

The ingredients required for the same construction are available in the present setting. Indeed, the flows of H_S^0 and H_S^1 preserve Σ , Proposition 8.6.2 provides smooth hitting times of $\mathcal{S}$, and the first-order contact relations (8.6.1)–(8.6.2) guarantee the compatibility of the correcting and regular dynamics at the corresponding hitting points. Consequently, after possibly reducing the neighbourhoods of the two junctions, the matching argument of [64, 61] applies with H^0 , H_S^0 at the incoming junction and with H_S^1 , H^1 at the outgoing junction.

In this way, the definition in Equation (8.6.3) can be completed to a time-dependent Hamiltonian, defined on a neighbourhood of the whole regular–singular–regular field and having the regularity required in Theorem 4.1.3. Its Hamiltonian flow coincides, on the relevant trajectories, with the regular-singular-regular field of extremals described above. Since the detailed argument realizing this part would be quite long and technical, we do not repeat it here.

With a slight abuse of notation, we continue to denote this local extension by H^S .

Proposition 8.6.4 (Super-Hamiltonian at a regular–singular–regular junction). *After possibly reducing the neighbourhood of the reference extremal, H^S generates the corrected local field described above and satisfies the super-Hamiltonian inequality*

$$H(\lambda_t) \leq H^S(t, \lambda_t) \quad (8.6.4)$$

along every trajectory λ_t of this field. Moreover, the reference extremal is an integral curve of $\vec{H}^S$ and

$$H^S(t, \tilde{\lambda}_t) = H(\tilde{\lambda}_t)$$

for every $t \in [0, T]$.

Proof. Outside of Σ , H^S coincides with H^0 or H^1 in the regions where the corresponding Hamiltonian is the maximizing one, so (8.6.4) holds with equality. For $\lambda \in \Sigma \cap \{H^{01} < 0\}$ close to $\tilde{\lambda}_{t_1}$, we have that $\exp(s\vec{H}_S^0)(\lambda) \in \Sigma$ for $s > 0$ and similarly for $\lambda \in \Sigma \cap \{H^{01} < 0\}$ close to $\tilde{\lambda}_{t_2}$. Then, Proposition 8.3.1 gives

$$H_S^0 = H_S^1 = H_S(t, \cdot) \geq H \quad \text{on } \Sigma.$$

The same Proposition gives the required domination and agreement with the reference extremal on the singular interval $[t_1, t_2]$. Finally, the localization described above is performed away from the relevant field and therefore does not modify any of these properties. $\square$

This completes the construction of a super-Hamiltonian through a regular–first-order singular–regular extremal. With minor modifications, the same geometric construction applies to the case of a nice extremal containing one singular arc.

We can now combine the preceding construction with the analysis of the Jacobi curve developed in Section 8.5.3. The essential point is that the super-Hamiltonian constructed above provides the nonlinear field of extremals through the regular–singular and singular–regular junctions, while its linearized evolution along the reference extremal agrees with the evolution of the Jacobi curve described in the previous sections. Therefore, in the absence of conjugate points, the same perturbation and local-invertibility argument used for piecewise regular extremals in Section 7.6 applies without any further construction.

Theorem 8.6.5 (Sufficient optimality for nice extremals). *Let $(\tilde{u}, \tilde{\lambda})$ be a nice normal extremal pair containing at most one first-order singular arc. Suppose that there are no points conjugate to q_0 along the trajectory $\tilde{q} = \pi \circ \tilde{\lambda}$. Then $\tilde{q}$ is locally strongly optimal.*

The restriction to a single singular arc appears to be only technical. Indeed, Theorem 8.5.1 already describes the Jacobi curve for an arbitrary finite concatenation of regular and first-order singular arcs, while the super-Hamiltonian construction developed above is local at each regular–singular or singular–regular junction. This motivates the following conjecture.

Conjecture 8.6.6. *Let $(\tilde{u}, \tilde{\lambda})$ be a nice normal extremal pair containing a finite number of first-order singular arcs. If there are no points conjugate to q_0 along $\tilde{q} = \pi \circ \tilde{\lambda}$, then $\tilde{q}$ is locally strongly optimal.*

The remaining difficulty is to iterate the local super-Hamiltonian modifications while preserving a single globally consistent Lipschitz field and the compatibility of its generalized tangent dynamics with the recursively defined Jacobi curve. No new local obstruction is expected to arise from the presence of additional singular arcs.

Chapter 9

Asymptotics of motion planning complexity for control-affine systems

9.1 Introduction

Given a control system on a manifold M of the form

$$\dot{q} = f(q, u), \quad q \in M, u \in \mathbb{R}^m,$$

the motion planning problem consists in finding an admissible trajectory q_u steering the system between a given initial position and a final one, usually under further requirements, such as obstacle avoidance. One possible approach to solve such a problem is to split it into two different tasks: (i) find a (usually non-admissible) trajectory Γ solving the problem (i.e., connecting the two given points without touching the obstacles), and (ii) track Γ by means of admissible trajectories. Since the first step depends only on the topology of the manifold and of the obstacles, we focus on the second one, which depends only on the control system near the path Γ .

To measure tracking accuracy, we introduce a distance $d : M \times M \rightarrow [0, +\infty)$ and say that a trajectory q_u tracks Γ with precision ε if $q_u([0, T]) \subset \text{Tube}_\varepsilon(\Gamma)$, where

$$\text{Tube}_\varepsilon(\Gamma) = \{p \in M \mid d(\Gamma, p) \leq \varepsilon\}.$$

Moreover, we assume that a cost function $J : L^\infty([0, T]; \mathbb{R}^m) \rightarrow \mathbb{R}$ is provided, and look for trajectories solving step (ii) in time $T > 0$ while minimizing J . This is quantified by the ε -complexity associated with a compact curve $\Gamma \subset M$ connecting $q_1, q_2 \in M$, which for $\varepsilon > 0$ is the quantity

$$\text{MC}_\varepsilon(\Gamma, T) = \frac{1}{\varepsilon} \inf \left\{ J(u) \mid q_u([0, T]) \subset \text{Tube}_\varepsilon(\Gamma), q_u(0) = q_1, q_u(T) = q_2 \right\}.$$

When Γ is not an admissible trajectory of the control system, the complexity diverges as ε tends to 0 and the aim of this Chapter is to precisely characterize its asymptotic behaviour in a specific class of control-affine systems.

In the case of a sub-Riemannian (or nonholonomic) control system, i.e., of a control system linear w.r.t. the control, a natural choice for the cost is the L^1 -norm of the controls. Due to time reversibility of such systems, the associated value function is in fact a distance, called the Carnot–Carathéodory distance, that endows M with a metric space structure. In this context, various definitions of complexity have been proposed (see [21, 47]), and asymptotic estimates up to multiplicative constants are known (see [48]). We also mention that for a restricted class of sub-Riemannian systems, i.e., one-step bracket-generating systems or systems with two controls in dimension ≤ 6 , precise asymptotic estimates and explicit asymptotically optimal syntheses are obtained in the series of papers [23, 44, 38, 40, 42, 43, 41, 39, 69] (see [36] for a review). We also mention [45], where approximate optimal syntheses are derived.

In [49], the results of [48] have been partially extended to the control-affine case. Namely, the authors define different notions of complexities for such systems and find the corresponding rate of divergence under the strong Hörmander condition. The aim of the current chapter is to refine the results of [49] in the case of co-rank one distributions, providing precise asymptotics with explicit constants for all generic situations.

9.1.1 Setting and main results

Let M be a smooth and connected manifold of dimension $n \geq 3$, on which is given a family of smooth vector fields $X_0, X_1, \dots, X_m \in \text{Vec}(M)$, $m \in \mathbb{N}^*$, defining a control system of the form

$$\dot{q} = X_0(q) + \sum_{j=1}^m u_j X_j(q), \quad u \in \mathbb{R}^m. \quad (9.1.1)$$

Throughout the Chapter, we assume that the distribution $\Delta = \text{span}\{X_1, \dots, X_m\}$ satisfies the following assumptions:

(H0) Strong Hörmander condition: for any $q \in M$, it holds $\text{Lie } \Delta(q) = T_q M$.

(H1) Co-rank one: $\dim \Delta(q) = n - 1$ for any $q \in M$.

Assuming **(H0)**, the control-linear part of system (9.1.1) defines a sub-Riemannian structure on M . We denote by g the associated metric and $|\cdot|_g$ the corresponding norm, which is defined by

$$|v|_g = \min \left\{ |u| \mid \sum_{j=1}^m u_j X_j(q) = v \right\}, \quad \forall q \in M, v \in \Delta(q). \quad (9.1.2)$$

For simplicity of notation, we set $|v|_g = +\infty$ whenever $v \notin \Delta(q)$. It is standard to observe that the motion complexity problem with cost $J(u) = \|u\|_{L^1}$ for system (9.1.1) depends only on the pair (Δ, g) and not on the precise choice of frame $\{X_1, \dots, X_m\}$.

In this situation, [49, Theorem 1.3] states that, for $\varepsilon \rightarrow 0$, the complexity $\text{MC}_\varepsilon(\Gamma, T)$ diverges as ε^{-2} whenever Γ is not an admissible trajectory of the control system and T is sufficiently small. In this work, following [69], we derive precise asymptotics. More precisely, we consider curves Γ satisfying the following assumption:

(H2) The smooth curve $\Gamma \subset M$ is oriented, compact with non-empty boundary, and such that¹ $T\Gamma \pitchfork \Delta$. We label the boundary points as $\partial\Gamma = \{q_1, q_2\}$ consistently with this orientation.

The choice of orientation is crucial because system (9.1.1) is not time-reversible: the drift X_0 may either assist or hinder our motion along Γ . To formalize this distinction, we introduce the following definition.

Definition 9.1.1. Let $T^+\Gamma$ be the set of positively oriented tangent vectors to Γ . Then, a one-form ω is associated with (Δ, Γ) if $\Delta = \ker \omega$ and $\omega(v) > 0$ for any $v \in T^+\Gamma$.

Observe that the existence of forms associated with (Δ, Γ) is guaranteed by assumption **(H1)**. One can check that the set of all forms associated with (Δ, Γ) is given by $\{\varphi\omega \mid \varphi \in C^\infty(M), \varphi > 0\}$, where ω is any form associated with (Δ, Γ) .

Since $T_q M = \Delta(q) \oplus T_q \Gamma$ for any $q \in \Gamma$, we have the decomposition

$$X_0(q) = X_0^\Delta(q) + X_0^\Gamma(q), \quad \text{with } X_0^\Delta(q) \in \Delta, X_0^\Gamma(q) \in T\Gamma, \quad \forall q \in \Gamma. \quad (9.1.3)$$

Then, we define

$$T_\Gamma := \inf\{t > 0 \mid e^{tX_0^\Gamma} q_1 = q_2\}. \quad (9.1.4)$$

We say that X_0 and Γ are co-oriented if $T_\Gamma < +\infty$. Observe that this is equivalent to the fact that $\omega(X_0) > 0$ on Γ for any form ω associated with (Δ, Γ) .

Our first result concerns step 2 distributions. Here and henceforth, we let $f(\varepsilon) \simeq g(\varepsilon)$ if $\lim_{\varepsilon \rightarrow 0} f(\varepsilon)/g(\varepsilon) = 1$.

Theorem 9.1.2 (Step 2 case). *Assume that Δ and Γ are respectively a distribution and a curve on a smooth manifold M satisfying **(H1)**, **(H2)**. Assume moreover that Δ is of step 2 along Γ , i.e., that $\dim \text{span}\{\Delta(q) + [\Delta, \Delta](q)\} = n$ for any $q \in \Gamma$.*

Let ω be a form associated with (Δ, Γ) and define

$$\alpha(q) := \max \omega_q([X_1, X_2]), \quad (9.1.5)$$

where the maximum is taken among all vector fields such that $X_1(q'), X_2(q') \in \Delta$ and $|X_1(q')|_g = |X_2(q')|_g = 1$ for any q' in a neighborhood of q .

Then, recalling the definition of T_Γ in (9.1.4), we have:

¹The assumption $T\Gamma \pitchfork \Delta$ is present just to simplify the exposition. See [69].

- if $T_\Gamma < T$:

$$\text{MC}_\varepsilon(\Gamma, T) \simeq \frac{2(T - T_\Gamma)}{\varepsilon^2} \min_{q \in \Gamma} \frac{\omega(X_0(q))}{\alpha(q)}; \quad (9.1.6)$$

- if $T_\Gamma = T$:

$$\text{MC}_\varepsilon(\Gamma, T) \simeq \frac{1}{\varepsilon} \int_\Gamma \frac{|X_0^\Delta|}{\omega(X_0)} \omega, \quad (9.1.7)$$

where X_0^Δ is defined in (9.1.3);

- if $T_\Gamma > T$: there exists $\Omega \subset \Gamma$ defined in Lemma 9.4.3, such that $|\Gamma \setminus \Omega| > 0$, encoding the portion of Γ where the drift helps reaching the endpoint. If $|\Omega| > 0$, then

$$\text{MC}_\varepsilon(\Gamma, T) \simeq \frac{2}{\varepsilon^2} \int_{\Gamma \setminus \Omega} \frac{1}{\alpha} \omega. \quad (9.1.8)$$

If, instead, $|\Omega| = 0$, then

$$\text{MC}_\varepsilon(\Gamma, T) \simeq \frac{2}{\varepsilon^2} \left(\int_\Gamma \frac{1}{\alpha} \omega - T \min_{q \in \Gamma} \frac{|\omega(X_0)|}{\alpha(q)} \right). \quad (9.1.9)$$

The three cases of the theorem can be interpreted as follows:

1. The drift is too strong along $T^+\Gamma$, and thus we have to “slow down” the motion. The optimal strategy is to do this when the ratio between the component of the drift along $T\Gamma$ and the maximal velocity at which the system can travel in the $T\Gamma$ direction is minimal.
2. The component of the drift along $T\Gamma$ steers the system exactly to the desired endpoint, so it is sufficient to neutralize the components of the drift along Δ .
3. The component of the drift along $T^+\Gamma$ is either too weak or directed in the opposite direction, so we also have to “push” the system towards the desired endpoint. In this case, the set Ω corresponds to the parts of Γ where the drift is “helping” to reach the endpoint. Consequently, on Ω we can simply follow the flow of the drift, while on $\Gamma \setminus \Omega$ we have to push the system thus contributing to the overall complexity. If Ω has zero measure, we have to push the system along the whole curve Γ , and the additional term in Equation (9.1.9) takes into account the final time constraint.

Remark 9.1.3. Notice that, if Δ is of step 2 along Γ , then at every point $q \in \Gamma$ there are X_1, X_2 tangent to Δ such that $[X_1, X_2](q) \lrcorner \Delta(q)$, hence $[X_1, X_2](q) \neq 0$ and in particular $\alpha(q) \neq 0$.

Remark 9.1.4. The order of the above asymptotics is drastically different when $T \neq T_\Gamma$ or $T = T_\Gamma$, passing from $O(\varepsilon^{-2})$ to $O(\varepsilon^{-1})$. A natural question, outside the scope of the present Chapter, is whether the two regimes merge smoothly as $T \rightarrow T_\Gamma$. Notice that this limit cannot be computed by elementary means from Theorem 9.1.2 by inverting the order of the limit in T and in ε , since both the quantities $\text{MC}_\varepsilon(\Gamma, T)$ and $\frac{2}{\varepsilon^2}C_0(\Gamma, T)$ are unbounded for every T as $\varepsilon \rightarrow 0$. Indeed, such a result requires the derivation of finer asymptotic expansions in the case $T \neq T_\Gamma$, of the type

$$\text{MC}_\varepsilon(\Gamma, T) \simeq \frac{2}{\varepsilon^2}C_0(\Gamma, T) + \frac{1}{\varepsilon}C_1(\Gamma, T) + o(\varepsilon^{-1}),$$

where the small-o term is uniform as $T \rightarrow T_\Gamma$ and $C_0(\Gamma, T), C_1(\Gamma, T)$ are functions such that

$$C_0(\Gamma, T) \rightarrow 0, \quad \text{and} \quad C_1(\Gamma, T) \rightarrow \int_\Gamma \frac{|X_0^\Delta|}{\omega(X_0)} \omega, \quad \text{as } T \rightarrow T_\Gamma. \quad (9.1.10)$$

Although the first condition in (9.1.10) is already contained in Theorem 9.1.2, the second one is not, and it cannot be derived directly from our computations. Indeed, the asymptotic optimal synthesis we exhibit is derived up to a $O(\varepsilon^{-1})$ term, which cannot be easily tracked (see, e.g., the proof of Theorem 9.3.3).

Remark 9.1.5. Letting $X_0 = 0$ we recover the result of [69]. Indeed, in this case $T_\Gamma = +\infty$, and $\Omega = \emptyset$.

The technique of proof for Theorem 9.1.2 consists in explicitly exhibiting an asymptotic optimal synthesis. Although similar in spirit to the synthesis considered in [69] for the control-linear case, the presence of a drift introduces some non-trivial technical difficulties, for example in showing that controls driving the system “against the drift” cannot be asymptotically optimal when the drift pushes the system in the direction of Γ (Theorem 9.3.4). The argument proceeds by first treating the balanced case $T = T_\Gamma$ and introducing normal coordinates adapted to (Δ, Γ) in Section 9.2. Then, in Section 9.3 we reduce the motion planning problem to a one-dimensional control problem for the variable z , encoded by the reduced motion complexity rmc_ε . The core of the proof is the analysis of this reduced problem, carried out in Section 9.3.2 and Section 9.4, where we determine the sign of optimal controls according to the relation between T and T_Γ and derive the corresponding asymptotics.

Since the step 2 assumption holds generically in dimension $n \geq 4$ (see the discussion preceding Theorem 9.2.4), we henceforth focus on the case $n = 3$, and in particular on the following situation, which is generic, as shown in [69].

Definition 9.1.6. For $n = 3$, we say that the couple (Δ, Γ) is of *generic type* if Γ satisfies **(H2)** and there exists a *finite* set of points $\Sigma \subset \Gamma$ (possibly empty), called *Martinet points*,

on which Δ is step 3, while it is step 2 on $\Gamma \setminus \Sigma$. More precisely, for any $q \in \Sigma$,

$$\begin{aligned} \dim \text{span}\{\Delta(q) + [\Delta, \Delta](q)\} &= 2, \\ \dim \text{span}\{\Delta(q) + [\Delta, \Delta](q) + [\Delta, [\Delta, \Delta]](q)\} &= 3. \end{aligned}$$

Moreover, we require that, at every Martinet point $q \in \Sigma$, it holds that $d_q \alpha|_{\Delta(q)} \neq 0$ and $d_q \alpha|_{T_q \Gamma} \neq 0$.

Remark 9.1.7. The assumption $d_q \alpha|_{T_q \Gamma} \neq 0$ in the definition of generic type means that the curve Γ intersects transversally the Martinet surface $\mathcal{M} = \{q \in M \mid \alpha(q) = 0\}$, which contains all the points where Δ is of step 3. The case where Γ can become tangent to $\mathcal{M}$ is quite delicate, due among other things to the existence of abnormal extremals contained in $\mathcal{M}$, and is not treated in this Chapter.

Near Martinet points, tracking Γ can become more expensive, and indeed in the sub-Riemannian case they dominate the asymptotic [69]. In the presence of a drift, however, this is not always true: if the drift is co-oriented with Γ at Martinet points, we can exploit it to bypass these points. We then have the following result.

Theorem 9.1.8. *Assume that $n = 3$ and that (Δ, Γ) is of generic type. Let ω be a form associated with (Δ, Γ) and let $\Sigma^- = \{q_i \in \Sigma \mid \omega(X_0(q_i)) \leq 0\}$. We have the following:*

- *If $\Sigma^- = \emptyset$, the same conclusions as in Theorem 9.1.2 hold.*
- *If $\Sigma^- = \{q_1, \dots, q_r\}$, then*

$$\text{MC}_\varepsilon(\Gamma, T) \simeq -\frac{\log \varepsilon}{\varepsilon^2} \sum_{i=1}^r \frac{4}{\kappa_i}.$$

Here, $\kappa_1, \dots, \kappa_r$ are defined by

$$\kappa_i = |\omega([W_\Gamma, [X_1, X_2]](q_i)) - d\omega(W_\Gamma, [X_1, X_2](q_i))|, \quad i = 1, \dots, r, \quad (9.1.11)$$

where $\{X_1, X_2\} \subset \text{Vec}(M)$ is a local orthonormal frame for the metric g near q_i , and W_Γ is any vector field such that $\omega(W_\Gamma) = 1$ and $W_\Gamma|_\Gamma \subset T\Gamma$.

This Theorem generalizes the result in [40] in the sub-Riemannian setting, which corresponds to $X_0 = 0$ in our notations.

Remark 9.1.9. Replacing ω with another form associated with (Δ, Γ) does not change the value of κ_i , as can be checked by direct computations. Moreover, in Proposition 9.6.1 we show that the above definition of κ_i coincides with the one provided in [69], that is $\kappa_i = d\alpha(W_\Gamma(q_i))$. In particular, the above result generalises the result of that reference.

Structure of the chapter

The remainder of the chapter is organized as follows: in Section 9.2 we prove some preliminary results and establish the asymptotic in Equation (9.1.7); in Section 9.3 we introduce the notion of *reduced motion complexity* and we show how the problem of finding an asymptotic for MC_ε can be reduced to the study of this new simplified quantity, see Section 9.3.1; then, in Section 9.3.2 we prove some important properties of the reduced motion complexity which will be used in the proof of Theorem 9.1.2; in Section 9.4 we prove Theorem 9.1.2; finally, in Section 9.5 we prove Theorem 9.1.8: we first deal with the case of a single Martinet point (see Section 9.5.1) and then we show how the general case can be reduced to this one.

9.2 Preliminaries and normal forms

In this Section we first prove a continuity result for the quantity $\text{MC}_\varepsilon(\Gamma, T)$. In particular, this leads to a simple proof of Equation (9.1.7). Then in Section 9.2.3, we introduce normal coordinates in $\text{Tube}_\varepsilon(\Gamma)$.

9.2.1 Continuity of motion complexity

Since we are going to consider motion complexities associated to different control systems, in this section we specify the dependence of the motion complexity on the vector fields $X_0, X_1, \dots, X_m$: using the multi-index notation $X_I = (X_0, X_1, \dots, X_m)$, we write $\text{MC}_\varepsilon(\Gamma, T, X_I)$.

Lemma 9.2.1. *Let $(X_I^j)_{j \in \mathbb{N}}$ be a sequence of $(m+1)$ -tuples of smooth vector fields on M converging to X_I^∞ in the product topology of $\text{Vec}(M)^{m+1}$ induced by the C^0 topology of $\text{Vec}(M)$. Then, for every $\varepsilon > 0$, we have*

$$\lim_{j \rightarrow +\infty} \text{MC}_\varepsilon(\Gamma, T, X_I^j) = \text{MC}_\varepsilon(\Gamma, T, X_I^\infty)$$

Proof. We begin by showing the following inequality:

$$\liminf_{j \rightarrow +\infty} \text{MC}_\varepsilon(\Gamma, T, X_I^j) \geq \text{MC}_\varepsilon(\Gamma, T, X_I^\infty). \quad (9.2.1)$$

We fix $\delta > 0$ and for every $j \in \mathbb{N}$ we take a control u_j admissible for $\text{MC}_\varepsilon(\Gamma, T, X_I^j)$ such that

$$\frac{1}{\varepsilon} \|u_j\|_1 \leq \text{MC}_\varepsilon(\Gamma, T, X_I^j) + \delta. \quad (9.2.2)$$

Let us denote by q_j the trajectory associated to the control u_j . By (9.2.1) and (9.2.2), we have that $\sup_j \|u_j\|_1 < +\infty$. We claim that, up to a subsequence, the q_j converges

uniformly to a limit trajectory q_∞ . Indeed, all the q_j are contained in $\text{Tube}_{\varepsilon+\eta}^\infty(\Gamma)$ for some $\eta > 0$, hence they are equibounded. Moreover, they are equi-Lipschitz continuous:

$$|q_j(t_1) - q_j(t_2)| = \left| \int_{t_1}^{t_2} X_0^j(q_j) + \sum_{i=1}^m u_j(s) X_i(q_j) ds \right| \leq (C_1 + \|u_j\|_1 C_2) |t_1 - t_2|$$

where $C_1 = \sup_{j \in \mathbb{N}, q \in \text{Tube}_\varepsilon(\Gamma)} |X_0^j(q)|$ and $C_2 = \max_{i \in 1, \dots, m, q \in \text{Tube}_\varepsilon(\Gamma)} |X_i(q)|$. Thus, by Arzelà-Ascoli Theorem, the trajectories q_j admit a converging subsequence and we denote by q_∞ its limit. Now, the same argument used in Theorem 3.41 in [5] in our case yields that q_∞ is admissible for the problem $\text{MC}_\varepsilon(\Gamma, T, X_I^\infty)$. We point out two differences between our case and the one treated in the reference: first, they take into consideration only control-linear systems, but their proof does not require any particular structure of the control system; second, in the reference the set of admissible velocities is the same for every trajectory of the sequence, while in our case it changes for every j . However, this fact does not affect the argument thanks to the assumption $X_I^j \rightarrow X_I^\infty$ uniformly as $j \rightarrow +\infty$.

Thus, denoting by u_∞ the control corresponding to q_∞ , again by Theorem 3.41 in [5] we have $\|u_\infty\|_1 \leq \liminf_{j \rightarrow +\infty} \|u_j\|_1$. Hence, we obtain

$$\text{MC}_\varepsilon(\Gamma, T, X_I^\infty) \leq \frac{1}{\varepsilon} \|u_\infty\| \leq \frac{1}{\varepsilon} \liminf_{j \rightarrow +\infty} \|u_j\|_1 \leq \liminf_{j \rightarrow +\infty} \text{MC}_\varepsilon(\Gamma, T, X_I^j) + \delta,$$

which, thanks to the arbitrariness of δ , yields the desired inequality.

Now, we want to prove the opposite inequality with $\limsup$ in place of $\liminf$. By definition of $\text{MC}_\varepsilon(\Gamma, T, X_I^\infty)$, for every $\delta > 0$ there is a control $u = (u_1, \dots, u_m)$ admissible for $\text{MC}_\varepsilon(\Gamma, T, X_I^\infty)$ such that

$$\frac{1}{\varepsilon} \|u\|_1 \leq \text{MC}_\varepsilon(\Gamma, T, X_I^\infty) + \delta.$$

Let us denote by q_∞ the trajectory associated to u . Now, for each $j \in \mathbb{N}$, consider the trajectory q_j of the control system (9.1.1) with drift X_0^j and control u . Then, standard arguments show that the C^0 convergence of X_I^j to X_I^∞ yields the uniform convergence of q_j to q_∞ on $[0, T]$.

Since $q_j \rightarrow q_\infty$ uniformly, we have that, for every j big enough, there are some $\eta_j > 0$ and $\tau_j > 0$ such that $q_j(t) \in \text{Tube}_{\varepsilon+\eta_j}^j(\Gamma)$, and $|q_j(T - \tau_j) - q_\infty(T)| \leq \delta$, where, $\text{Tube}_\varepsilon^j(\Gamma)$ denotes the sub-Riemannian tube associated with the family X_I^j , and $|\cdot|$ denotes the euclidean distance in a suitable local chart. Then, since for the driftless system the L^1 cost to join $q_j(T - \tau_j)$ to $q_\infty(T)$ is bounded by $C\delta^{1/2}$ (this is the Ball-Box Theorem, see [21] or [50, Corollary 2.1]), a standard time-rescaling argument (see also [65, Theorem 1.1]) yields that we can reach $q_\infty(T)$ from $q_j(T - \tau_j)$ in time τ_j with a control v_j , whose L^1 -norm is bounded by $\frac{1}{\varepsilon} \|v_j\|_1 \leq C \left(\delta^{1/2} + \tau_j^{1/2} \right)$. Since the concatenation of the trajectory

$q_j|_{[0, T-\tau_j]}$ and the trajectory given by v_j is admissible for $\text{MC}_{\varepsilon+\eta_j}(\Gamma, T, X_I^j)$, we have that $\text{MC}_{\varepsilon+\eta_j}(\Gamma, T, X_I^j)$ is bounded by

$$\frac{1}{\varepsilon} \left(\int_0^{T-\tau_j} |u(t)| dt + \int_{T-\tau_j}^T |v_j(t)| dt \right) \leq \text{MC}_{\varepsilon}(\Gamma, T, X_I^{\infty}) + \delta + C \left(\delta^{1/2} + \tau_j^{1/2} \right).$$

Notice that, as $j \rightarrow +\infty$, we can take $\eta_j \rightarrow 0$ and $\tau_j \rightarrow 0$. Thus, taking the limsup in j , we obtain the inequality in (9.2.1). Here, we use the fact that $\varepsilon \mapsto \text{MC}_{\varepsilon}(\Gamma, T, X_I^j)$ is continuous for $\varepsilon > 0$, which can be proved by following the same arguments as in the first part of this proof. $\square$

9.2.2 Decomposition of the drift and proof of (9.1.7)

In this section, we first prove that replacing the drift X_0 with its vertical part X_0^{Γ} affects the motion complexity by a factor of order $O(\varepsilon^{-1})$. This allows us to prove (9.1.7) of Theorem (9.1.2).

Proposition 9.2.2. *Let $X_0 = X_0^{\Delta} + X_0^{\Gamma}$, where X_0^{Δ} and X_0^{Γ} are vector fields defined in a neighbourhood of Γ that satisfy (9.1.3). Then, denoting by $X_I = (X_0, X_1, \dots, X_m)$ and $X_I^{\Gamma} = (X_0^{\Gamma}, X_1, \dots, X_m)$, there are constants $0 \leq C < +\infty$ and $\varepsilon_0 > 0$ such that for $0 < \varepsilon < \varepsilon_0$ it holds that*

$$|\text{MC}_{\varepsilon}(\Gamma, T, X_I) - \text{MC}_{\varepsilon}(\Gamma, T, X_I^{\Gamma})| \leq C\varepsilon^{-1}.$$

Proof. Let v be an admissible control for the problem $\text{MC}_{\varepsilon}(\Gamma, T, X_0^{\Gamma})$. Then

$$\dot{q}_v = X_0^{\Gamma}(q_v) + \sum_{j=1}^m v_j(t) X_j(q_v)$$

Since, by definition, $X_0^{\Delta}(q) = \sum_{j=1}^m a_j(q) X_j(q)$, $q \in M$, for some suitable choice of $a = (a_1, \dots, a_m) \in C^{\infty}(M)^m$, we have that $v + a \circ q_v$ is an admissible control for $\text{MC}_{\varepsilon}(\Gamma, T, X_0)$ corresponding to the same trajectory q_v . Notice that we can choose a such that $|a| = |X_0^{\Delta}|_g$, where $|X_0^{\Delta}|_g$ is the sub-Riemannian norm, which was introduced in (9.1.2), of the vector field X_0^{Δ} . So, for any admissible control v for $\text{MC}_{\varepsilon}(\Gamma, X_0^{\Gamma})$ there is a control u admissible for $\text{MC}_{\varepsilon}(\Gamma, X_0)$ with cost

$$J(u) = \int_0^T |u(t)| dt \leq J(v) + \int_0^T |a(q_v(t))| dt \leq J(v) + C_1,$$

where $0 < C_1 \leq T \max_{q \in \text{SR-Tube}_{\varepsilon}(\Gamma)} |a(q)|$. Similarly, for any control u admissible for $\text{MC}_{\varepsilon}(\Gamma, T, X_I)$, defining $v = u - a \circ q_u$ we obtain an admissible control for the problem $\text{MC}_{\varepsilon}(\Gamma, T, X_I^{\Gamma})$ corresponding to the same trajectory q_u and

$$J(v) = \int_0^T |v(t)| dt \leq J(u) + \int_0^T |a(q_u(t))| dt \leq J(u) + C_2,$$

where again $0 < C_2 \leq T \max_{q \in \text{SR-Tube}_\varepsilon(\Gamma)} |a(q)|$. We conclude that there is some constant $C > 0$ such that the following holds: for all u admissible for $\text{MC}_\varepsilon(\Gamma, T, X_I)$ (resp. v admissible for $\text{MC}_\varepsilon(\Gamma, T, X_I^\Gamma)$) there is v admissible for $\text{MC}_\varepsilon(\Gamma, T, X_I^\Gamma)$ (resp. u admissible for $\text{MC}_\varepsilon(\Gamma, T, X_I)$) such that $|J(u) - J(v)| \leq C$. Now, fix $\varepsilon > 0$ and take a minimizing sequence $(v_k)_{k \in \mathbb{N}}$ for $\text{MC}_\varepsilon(\Gamma, T, X_0^\Gamma)$ and let $u_k = v_k + a_k \circ q_{u_k}$, as above. We obtain $\frac{1}{\varepsilon} J(v_k) + \frac{C}{\varepsilon} \geq \frac{1}{\varepsilon} J(u_k) \geq \text{MC}_\varepsilon(\Gamma, T, X_0)$, and passing to the limit for $k \rightarrow +\infty$, we obtain $\text{MC}_\varepsilon(\Gamma, X_0^\Gamma) + \frac{C}{\varepsilon} \geq \text{MC}_\varepsilon(\Gamma, X_0)$. Exchanging X_0 with X_0^Γ , the very same argument yields the opposite inequality, proving the statement. $\square$

Now, we can prove directly the asymptotic for $T = T_\Gamma$ in Theorem 9.1.2.

Proposition 9.2.3. *Suppose that $T = T_\Gamma$. Then (9.1.7) in Theorem 9.1.2 holds.*

Proof. If $T = T_\Gamma$, then the curve Γ is admissible for the control system

$$\dot{q}_u = X_0^\Gamma(q_u) + \sum_{j=1}^m u_j X_j(q_u),$$

with control $v_0 = 0$, which is of course the unique minimizer for $\text{MC}_\varepsilon(\Gamma, T, X_0^\Gamma)$ and in particular $\text{MC}_\varepsilon(\Gamma, T, X_0^\Gamma) = 0$, for every $\varepsilon \geq 0$. Moreover, we know that Γ is an admissible trajectory also for $\text{MC}_\varepsilon(\Gamma, T, X_0)$, hence there is a constant $C > 0$ such that $\text{MC}_\varepsilon(\Gamma, T, X_0) \simeq C\varepsilon^{-1}$, $\varepsilon \rightarrow 0$. In particular, the control $u_0(t) = -a(e^{tX_0^\Gamma}(q_1))$ is admissible for $\text{MC}_\varepsilon(\Gamma, T, X_0)$ for every $\varepsilon \geq 0$ and realizes the trajectory Γ . Hence, we have $\lim_{\varepsilon \rightarrow 0} \varepsilon \text{MC}_\varepsilon(\Gamma, T, X_0) = J(u_0)$. Choosing any form ω associated with (Δ, Γ) , the parametrization $\Gamma(t) = e^{tX_0^\Gamma}(q_1)$ yields the following intrinsic formula for $J(u_0)$:

$$\int_0^T |u_0(t)| dt = \int_0^T |a(e^{tX_0^\Gamma}(q_1))| dt = \int_0^T \frac{|X_0^\Delta(\Gamma)|}{\omega_\Gamma(X_0(\Gamma))} \omega_\Gamma(\dot{\Gamma}) dt = \int_\Gamma \frac{|X_0^\Delta|}{\omega(X_0)} \omega.$$

Here, we used that $\omega(\dot{\Gamma})|_{\Gamma(t)} = \omega(X_0)|_{\Gamma(t)}$. This completes the proof. $\square$

9.2.3 Normal forms

Henceforth, we fix a form ω associated with (Δ, Γ) . By requiring that $\omega(\dot{\Gamma}) \equiv 1$, this singles out an associated parametrisation $\Gamma : [0, S] \rightarrow M$, that we will exploit to construct normal coordinates in $\text{Tube}_\varepsilon(\Gamma)$.

In view of Proposition 9.2.3, from now on we always assume $T \neq T_\Gamma$. Since we are going to show that $\text{MC}_\varepsilon(\Gamma, T) = O(\varepsilon^{-2})$, we can neglect every term in the asymptotic of smaller order. This allows to introduce appropriate normal forms.

The choice of ω determines a skew-symmetric endomorphism $A_q : \Delta(q) \rightarrow \Delta(q)$ as follows:

$$g_q(A_q v, w) = d\omega(v, w) \quad \forall v, w \in \Delta(q).$$

By skew-symmetry, A_q has purely imaginary spectrum, and we denote its non-zero eigenvalues by $\pm i\alpha_1(q), \dots, \pm i\alpha_k(q)$, $k = n/2$ if n is even or $k = (n-1)/2$ otherwise. Then, there exists an orthonormal frame $X_1, \dots, X_{n-1}$ such that A_q is block diagonal with k two-dimensional blocks of the form

$$\begin{pmatrix} 0 & \alpha_i(q) \\ -\alpha_i(q) & 0 \end{pmatrix}.$$

Notice that the step 2 assumption implies that $A_q \neq 0$ and hence has at least a non-zero eigenvalue. Moreover, by Lemma 9.2.1, it is not restrictive to suppose that the greatest eigenvalue in absolute value is simple. That is, we can order the eigenvalues of A_q in such a way that $\alpha_1(q) > \max\{\alpha_2(q), \dots, \alpha_k(q)\} \geq 0$. Notice that $\alpha_1 = \alpha$, where α was defined in (9.1.5).

Let $\mathfrak{X}$ denote the set of those (Δ, g, Γ) satisfying **(H0)**, **(H1)**, **(H2)**. This is the set of relevant motion planning problems, in the language of [69]. Then, adapting the argument of [69, Theorem 2.2] one obtains that there exists an open and dense subset $\mathfrak{X}_* \subset \mathfrak{X}$ such that for any $(\Delta, g, \Gamma) \in \mathfrak{X}_*$ it holds that: (i) if $n \geq 4$, then Δ is step 2; (ii) if $n = 3$, then (Δ, Γ) is of generic type according to Definition 9.1.6. We have the following, which can be proved as [69, Theorem 3.4] and references therein.

Theorem 9.2.4 (Normal coordinates). *Let $(\Delta, g, \Gamma) \in \mathfrak{X}_*$. Then, there exists $\varepsilon_0 > 0$ and a coordinate system $(\xi, z) : \mathbb{R}^{n-1} \times \mathbb{R} \rightarrow \text{Tube}_{\varepsilon_0}(\Gamma)$ such that*

- $\Gamma(s) = (0, s)$ for $s \in [0, S]$. In particular, $q_1 = (0, 0)$ and $q_2 = (0, S)$.
- There exists a generating frame $\{X_1, \dots, X_{n-1}\}$ for Δ , which is orthonormal for the metric g , and that reads

$$X_i(\xi, z) = \sum_{j=1}^{n-1} Q_{ij}(\xi, z) \partial_{\xi_j} + L_j(\xi, z) \partial_z,$$

where the matrix $Q(\xi, z)$ is symmetric, $Q(\xi, z)\xi = 0$, $L(\xi, z) \cdot \xi = 0$, and

$$Q(\xi, z) = \text{Id} + O(|\xi|^2) \quad \text{and} \quad L(\xi, z) = A_{(0,z)}\xi + O(|\xi|^2).$$

- If $n = 3$, letting $\xi = (x, y) \in \mathbb{R}^2$, the above orthonormal frame reads

$$\begin{aligned} X_1(x, y, z) &= (1 + y^2 \beta(x, y, z)) \partial_x - xy \beta(x, y, z) \partial_y + \frac{y}{2} \gamma(x, y, z) \partial_z, \\ X_2(x, y, z) &= -xy \beta(x, y, z) \partial_x + (1 + x^2 \beta(x, y, z)) \partial_y - \frac{x}{2} \gamma(x, y, z) \partial_z, \end{aligned}$$

where β and γ are smooth functions and $|\gamma(0, 0, z)| = |\omega([X_1, X_2])|_{(0,0,z)} = \alpha(0, 0, z)$, where α is defined in (9.1.5).

Remark 9.2.5. In the case $n \geq 4$, [69, Theorem 3.4] holds under a stronger assumption than the fact that Δ is step 2. Namely, the authors assume that all non-zero eigenvalues of A_q are simple, and that $\dim \ker A_q$ is the smallest possible while retaining genericity. This requirement is due to the fact that in [69] the authors require a precise ordering of the blocks in $A_{(0,z)}$, while we only need to identify the block corresponding to the eigenvalue having largest modulus.

9.3 Step 2 case: the reduced motion complexity

In this Section we collect a series of technical results needed for the proof of the step 2 case. Throughout this Section, we assume that the distribution Δ satisfies the assumptions of Theorem 9.1.2.

First, we introduce the notion of *reduced motion complexity* and we show that the problem of finding an asymptotic as $\varepsilon \rightarrow 0$ for $\text{MC}_\varepsilon(\Gamma, T)$ can be reduced to finding an asymptotic of this reduced motion complexity (see Theorem 9.3.3 and Section 9.3.1). Finally, in Section 9.3.2 we prove some properties of the reduced motion complexity which are instrumental to the proof of Theorem 9.1.2.

We set $a_0 := \omega(X_0) \in C^\infty(M)$. Also, with a small abuse of notation, for $z \in \mathbb{R}$, we write $\alpha(z) := \alpha(0, z)$ and $a_0(z) = a_0(0, z)$, where in the r.h.s. we are using the coordinates (ξ, z) introduced in Theorem 9.2.4.

Definition 9.3.1. Given $z_f > 0$, for any $\varepsilon > 0$ and $T > 0$, the *reduced motion complexity* is

$$\text{rmc}_\varepsilon(z_f, T) = \inf \frac{2}{\varepsilon^2} J(v), \quad J(v) = \int_0^T |v| dt, \quad (9.3.1)$$

where the infimum is taken over all controls $v \in L^\infty([0, T]; \mathbb{R})$ such that there exists a solution $z : [0, T] \rightarrow \mathbb{R}$ of

$$\dot{z} = a_0(z) - \alpha(z)v, \quad \text{with} \quad z(0) = 0 \text{ and } z(T) = z_f. \quad (9.3.2)$$

Remark 9.3.2. Under the step 2 assumption, the control system (9.3.2) is feedback equivalent to the free system $\dot{z} = \tilde{v}(t)$, $\tilde{v} : [0, T] \rightarrow \mathbb{R}$. Indeed, $\alpha > 0$ and thus the transformation $v \in \mathbb{R} \mapsto a_0(z) - \frac{\varepsilon}{2}\alpha(z)v$ is a diffeomorphism onto $\mathbb{R}$. Hence, any function in $z \in W^{1,\infty}([0, T])$ is an admissible trajectory.

The main result of this Section is the following.

Theorem 9.3.3. *Let $T \neq T_\Gamma$. Then,*

$$\text{MC}_\varepsilon(\Gamma, T) = \text{rmc}_\varepsilon \left(\int_\Gamma \omega, T \right) + O(\varepsilon^{-1}), \quad \text{as } \varepsilon \rightarrow 0.$$

As a consequence of the above, if there exists $C > 0$ such that $\text{rmc}_\varepsilon(\int_\Gamma \omega, T) \simeq \frac{C}{\varepsilon^2}$, then, $\text{MC}_\varepsilon(\Gamma, T) \simeq \frac{C}{\varepsilon^2}$. Obtaining such a constant C is the aim of Section 9.4, and requires the following, proved in Section 9.3.2.

Theorem 9.3.4. *The infimum in the definition of $\text{rmc}_\varepsilon(z_f, T)$ can be taken over controls v such that the corresponding z coordinate is non-decreasing. Moreover, we can assume that $v \leq 0$ if $T < T_\Gamma$ and that $v \geq 0$ if $T > T_\Gamma$.*

9.3.1 Proof of Theorem 9.3.3

The proof is composed of two steps: first, we show how to use the normal coordinates introduced in Theorem 9.2.4 to rewrite the control system in a convenient way, which allows us to perform precise computations. Then, we show how to reduce to considering only the vertical variable.

From now on we denote $d = \lfloor \frac{n-1}{2} \rfloor$. From Theorem 9.2.4, in the step 2 case, we have two alternatives. If $\dim M$ is odd, then $n = 2d + 1$ and using the local coordinates $(\xi, z) = (x, y, z) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}$ introduced above, the vector fields of the control system (9.1.1) can be chosen in the form

$$X_{2i-1}(x, y, z) = \partial_{x_i} + \alpha_i(z) \frac{y_i}{2} \partial_z + O(|\xi|^2), \quad (9.3.3)$$

$$X_{2i}(x, y, z) = \partial_{y_i} - \alpha_i(z) \frac{x_i}{2} \partial_z + O(|\xi|^2), \quad (9.3.4)$$

for $i = 1, \dots, d$. If, instead, $n = \dim M$ is even, then $n = 2d + 2$ and the normal forms for the X_i are the same as (9.3.3), (9.3.4) for $i = 1, \dots, d$, and $X_{n-1} = \partial_{x_{n-1}} + O(|\xi|^2)$.

In both cases, one can introduce polar coordinates $(r_i, \theta_i) \in \mathbb{R}_{>0} \times S^1$ in every (x_i, y_i) -plane by letting $(x_i, y_i) = r_i(\cos \theta_i, \sin \theta_i)$. Moreover, we denote by $\mathbf{v} = (v_1, \dots, v_d)$ and $\mathbf{w} = (w_1, \dots, w_d)$, where $(w_i, v_i)^\top := R_{\theta_i}(u_{2i-1}, u_{2i})^\top$, the control system (9.1.1) reads

$$\begin{cases} \dot{r}_i = w_i + O(|r|), \\ \dot{\theta}_i = -\frac{v_i}{r_i} + O(|r|), \\ \dot{z} = a_0(r, \theta, z) - \sum_{i=1}^d \alpha_i(r, \theta, z) \frac{r_i}{2} v_i + O(|r|), \end{cases} \quad \text{for } i = 1, \dots, d, \quad (9.3.5)$$

with the extra equation $\dot{x}_{n-1} = u_{n-1}(t) + O(|r|^2)$ if $\dim M$ is even. Here, $O(|r|)$ stands for a function ϕ such that $|\phi(r, \theta, z)| \leq C|r|$ for every $(\theta, z) \in \mathbb{T}^d \times \mathbb{R}$, provided that $r \in \mathbb{R}_{>0}^d$ is in a sufficiently small neighbourhood of 0.

The main idea is that, up to an error of order $O(\varepsilon^{-1})$, we can ignore the θ variable in (9.3.5), replace $a_0(r, \theta, z)$ and $\alpha_i(r, \theta, z)$ with $a_0(0, 0, z)$ and $\alpha_i(0, 0, z)$, respectively, and ignore the $O(|r|)$ terms. To make this precise, we introduce the following intermediate notion of reduced motion complexity.

Definition 9.3.5. Consider the control system

$$\begin{cases} \dot{r} = \mathbf{w}, \\ \dot{z} = a_0(z) - \sum_{i=1}^d \alpha_i(z) \frac{r_i}{2} v_i. \end{cases} \quad (9.3.6)$$

where $r_i \geq 0$, $z \in \mathbb{R}$ and $\mathbf{v}, \mathbf{w} \in L^\infty([0, T]; \mathbb{R}^d)$ are the controls. The *auxiliary reduced motion complexity* is

$$\text{a-rmc}_\varepsilon(z_f, T) := \inf \|(\mathbf{v}, \mathbf{w})\|_{L^1},$$

where the infimum is taken over all controls $\mathbf{v}, \mathbf{w} \in L^\infty([0, T]; \mathbb{R}^d)$ such that the corresponding solution of (9.3.6) satisfies the boundary conditions $r(0) = 0$, $z(0) = 0$, $r(T) = 0$, $z(T) = z_f$, and is subject to the constraint $|r(t)| \leq \varepsilon$.

Lemma 9.3.6. *We have the following asymptotic equivalence*

$$\text{MC}_\varepsilon(\Gamma, T) = \text{a-rmc}_\varepsilon\left(\int_\Gamma \omega, T\right) + O(\varepsilon^{-1}), \quad \text{as } \varepsilon \rightarrow 0. \quad (9.3.7)$$

Proof. Let $\phi_{r_i}, \phi_{\theta_i}, \phi_z$ be the $O(|r|)$ terms in the Equations 9.3.5 of r_i, θ_i, z respectively, $i = 1, \dots, d$. Replacing w_i, v_i in (9.3.5) by $\tilde{w}_i = w_i - \phi_{r_i}$ and $\tilde{v}_i = v_i + \frac{2\phi_z}{mr_i\alpha_i}$, there exists a constant $C > 0$ such that $\|w_i - \tilde{w}_i\|_\infty \leq C$ and $\|v_i - \tilde{v}_i\|_\infty \leq C$. This yields $\frac{1}{\varepsilon}|J(\tilde{\mathbf{v}}, \tilde{\mathbf{w}}) - J(\mathbf{v}, \mathbf{w})| \leq \frac{2dCT}{\varepsilon}$. Then, using an argument similar to the one used in Proposition 9.2.2, one obtains that this modification impacts the expansion of the motion complexity starting at order ε^{-1} . The equivalence in (9.3.7) follows. Finally, since the variables θ_i do not appear in the equations of r, z , nor in the boundary conditions, we can neglect them. $\square$

It is quite reasonable to expect that, in order to minimize the cost in (9.3.6), one should aim to obtain $r_1 = \varepsilon$ and then to only use the control v_1 to move in the z direction, since α_1 is the biggest eigenvalue of $A_{(0,z)}$. This effectively reduces the problem to a one-dimensional control system for z , which is exactly the one appearing in the definition of rmc_ε . The following makes this intuition precise.

Proposition 9.3.7. *Let $T \neq T_\Gamma$. Then, for any $z_f > 0$ the following asymptotic equivalence holds:*

$$\text{a-rmc}_\varepsilon(z_f, T) \simeq \text{rmc}_\varepsilon(z_f, T) + O(1) \quad \text{as } \varepsilon \rightarrow 0.$$

In the proof of Proposition 9.3.7, we need the following technical Lemma.

Lemma 9.3.8. *Fix $\varepsilon > 0$ and let $(\delta_k)_{k \in \mathbb{N}}$ be a sequence such that $\delta_k \rightarrow 0$ as $k \rightarrow +\infty$. Then $\lim_{k \rightarrow +\infty} \text{a-rmc}_\varepsilon(\Gamma, T + \delta_k) = \text{a-rmc}_\varepsilon(\Gamma, T)$*

Since the proof of this Lemma is very similar to the first part of the proof of Lemma 9.2.1, we omit it. Now, we prove Proposition 9.3.7.

Proof. Letting $u = 2v/\varepsilon$ in the definition of rmc_ε , we have

$$\text{rmc}_\varepsilon(z_f, T) = \inf \frac{1}{\varepsilon} \|u\|_1, \quad \text{where} \quad \dot{z} = a_0(z) - \alpha_1(z) \frac{\varepsilon}{2} u, \quad (9.3.8)$$

We remark that this dynamic coincides with that of the z variable in the dynamics (9.3.6) of a-rmc_ε , if $r = \varepsilon e_1$, where $e_1 = (1, 0, \dots, 0)$.

We start by showing that $\text{a-rmc}_\varepsilon(z_f, T) \geq \text{rmc}_\varepsilon(z_f, T)$. Fix a control $(\mathbf{v}, \mathbf{w})$ that is admissible for $\text{a-rmc}_\varepsilon(z_f, T)$. Considering the corresponding z coordinates, given by (9.3.6), we then let $u = \sum_{i=1}^d \frac{r_i}{\varepsilon} \frac{\alpha_i(z)}{\alpha_1(z)} v_i$, on $[0, T]$. With this choice of control, the z coordinate given by (9.3.6) is the same as the one appearing in the dynamic of (9.3.8). Recall that $\alpha_1 \geq \alpha_i$ for any $i = 1, \dots, d$, and that, by definition of $\text{a-rmc}_\varepsilon(z_f, T)$ we have $|r| = |(r_1, \dots, r_d)| \leq \varepsilon$. Thus, on $[0, T]$, $|u| \leq \sum_{i=1}^d \left| \frac{r_i}{\varepsilon} \right| |v_i| \leq \frac{|r|}{\varepsilon} |\mathbf{v}| \leq |\mathbf{v}|$. In particular, $\|u\|_1 \leq \|(\mathbf{v}, \mathbf{w})\|_1$, completing the proof of the first inequality.

To complete the proof of the statement, we are left to show the opposite inequality. To this aim, we prove that there exists $C > 0$, independent of ε , such that for any $\delta > 0$ sufficiently small it holds $\text{a-rmc}_\varepsilon(z_f, T + \delta) \leq \text{rmc}_\varepsilon(z_f, T) + C \left(1 + \frac{\delta^{1/2}}{\varepsilon}\right)$. By Lemma 9.3.8, taking the limit as $\delta \rightarrow 0$ yields that $\text{a-rmc}_\varepsilon(z_f, T) \leq \text{rmc}_\varepsilon(z_f, T) + C$.

Consider a control u admissible for $\text{rmc}_\varepsilon(z_f, T)$, in the formulation (9.3.8). We have to construct a control $(\mathbf{v}, \mathbf{w})$ that is admissible for $\text{a-rmc}_\varepsilon(z_f, T + \delta)$ and such that $\|(\mathbf{v}, \mathbf{w})\|_1 \leq \|u\|_1 + C(\varepsilon + \delta)$. The control $(\mathbf{v}, \mathbf{w})$ can be obtained by concatenating three control arcs:

1. A control driving the system (9.3.6) from $(0, 0)$ to $(\varepsilon e_1, 0)$ on the time interval $[0, \delta/2]$;
2. The control $\mathbf{w} = 0$ and $\mathbf{v} = (u(\cdot - \delta/2), 0, \dots, 0)$ on the time interval $[\delta/2, T + \delta/2]$, which drives the system to $(\varepsilon e_1, z_f)$;
3. A control driving the system to $(0, z_f)$ on the time interval $[T + \delta/2, T + \delta]$.

Using the Ball-Box Theorem and a time rescaling as in the proof of Lemma 9.3.8 (see also [65, Theorem 1.1]), one proves that the cost for the first and the last step is bounded by $C(\varepsilon + \delta^{1/2})$, proving the second inequality and the statement. $\square$

Finally, putting together Lemma 9.3.6 and Proposition 9.3.7, Theorem 9.3.3 follows.

9.3.2 Proof of Theorem 9.3.4

We split the proof of Theorem 9.3.4 in two parts: first we prove that z is non-decreasing (Lemmas 9.3.9 and 9.3.10), then we prove that the control v has a prescribed sign depending on the difference $T - T_\Gamma$ (Lemma 9.3.11).

Lemma 9.3.9. *Let v be an admissible control for $\text{rmc}_\varepsilon(z_f, T)$. Then, there exists $C > 0$ such that for any $\eta > 0$, there exists an admissible control v_η such that z_{v_η} is piecewise affine and $|\|v\|_{L^1} - \|v_\eta\|_{L^1}| \leq C\eta$.*

Proof. Thanks to Remark 9.3.2, the set of admissible trajectories is $W^{1,\infty}([0, T])$, where affine functions are dense in the $W^{1,1}$ topology. So, for any admissible trajectory z_v with control v , we can find another admissible trajectory z_η with some other control v_η , such that z_η is piecewise affine and $\|z_v - z_\eta\|_{W^{1,1}} \leq \eta$. Recall that $\dot{z}_v = a_0(z) - \alpha(z)v$. Hence, we can obtain a formula of v in terms of z and $\dot{z}$, which guarantees the existence of a constant $C > 0$ such that

$$\left| \int_0^T |v(t)| - |v_\eta(t)| dt \right| \leq \tilde{C} \left(\int_0^T |a_0(z_v(t)) - a_0(z_\eta(t))| dt + \int_0^T |\dot{z}_v(t) - \dot{z}_\eta(t)| dt \right),$$

Since a_0 is Lipschitz continuous, the statement follows by the approximation in $W^{1,1}$. $\square$

The following result proves the first part of Theorem 9.3.4.

Lemma 9.3.10. *The infimum in (9.3.1) can be taken on controls v such that the corresponding z_v is non-decreasing.*

Proof. Let v be a control admissible for (9.3.1), (9.3.2). The statement will be proved by showing that then there exists an admissible control $\tilde{v}$ such that $\dot{z}_{\tilde{v}} \geq 0$ and $\|\tilde{v}\|_1 \leq \|v\|_1$. By Lemma 9.3.9, we can suppose that z_v is piecewise affine, that is $z_v(t) = \sum_{j=1}^k f_j(t) \mathbf{1}_{[t_j, t_{j+1}]}(t)$, $f_j(t) = A_j t + B_j$, where $A_j, B_j \in \mathbb{R}$, $0 < t_1 < \dots < t_k < T$. We have to show that $A_j > 0$ for every $j = 1, \dots, k$. By contradiction, suppose that $A_{j_0} < 0$ for some $j_0 \in \{1, \dots, k\}$. Then there is some $s_1 > t_{j_0}$ such that $z_v(t_{j_0}) = z_v(s_1) =: z_1$. Then, the function z_v attains a local minimum for some $t_{j_1} < s_1$, $j_1 > j_0$. Define $z_0 := z_v(t_{j_1})$. Take any $s_0 \in [0, t_{j_0}]$ such that $z_v(s_0) = z_0$. Since the function $|a_0|/\alpha$ is continuous, it must have a minimum point $z_{\min}$ in the interval $[z_0, z_1]$. Moreover, we can find $\tau_0 \in [s_0, t_{j_0}]$ and $\tau_1 \in [t_{j_1}, s_1]$ such that $z_v(\tau_0) = z_v(\tau_1) = z_{\min}$.

As a consequence, there is some control $\tilde{v}$ such that $z_{\tilde{v}}$ coincides with z_v in $[0, \tau_0]$ and $[\tau_1, T]$ and $z_{\tilde{v}}(t) = z_{\min}$ for $t \in (\tau_0, \tau_1)$. (See Figure 9.1.) Notice that $\tilde{v}(t) = 2a_0(z_{\min})/(\varepsilon\alpha(z_{\min}))$ for $t \in (\tau_0, \tau_1)$.

In order to complete the proof, we are left to show that $\|\tilde{v}\|_1 \leq \|v\|_1$. By definition of $\tilde{v}$, we have that

$$\int_0^T |\tilde{v}(t)| dt = \int_0^{\tau_0} |v(t)| dt + \int_{\tau_0}^{\tau_1} \frac{|a_0(z_{\min})|}{\alpha(z_{\min})} dt + \int_{\tau_1}^T |v(t)| dt.$$

Now, if $|a_0(z_{\min})| = 0$, then we have $\|\tilde{v}\|_1 \leq \|v\|_1$ trivially. Otherwise, on the interval $[\tau_0, \tau_1]$ we have either $a_0 > 0$ or $a_0 < 0$. In both cases, we have

$$\begin{aligned} \int_{\tau_0}^{\tau_1} \frac{|a_0(z_{\min})|}{\alpha(z_{\min})} dt &\leq \int_{\tau_0}^{\tau_1} \frac{|a_0(z_v(t))|}{\alpha(z_v(t))} dt = \pm \int_{\tau_0}^{\tau_1} \frac{a_0(z_v(t))}{\alpha(z_v(t))} dt \\ &= \pm \int_{\tau_0}^{\tau_1} \left(\frac{\dot{z}_v}{\alpha(z_v(t))} + v(t) \right) dt \leq \pm \int_{z_v(\tau_0)}^{z_v(\tau_1)} \frac{1}{\alpha(z)} dz + \int_{\tau_0}^{\tau_1} |v(t)| dt = \int_{\tau_0}^{\tau_1} |v(t)| dt. \end{aligned}$$

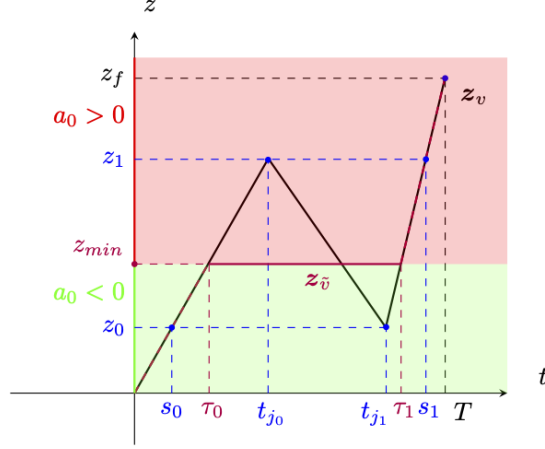


Figure 9.1: Construction for the proof of Lemma 9.3.10 in case $a_0(z_{\min}) = 0$.

The last equality follows since $z_v(\tau_0) = z_v(\tau_1)$. Hence, we obtain that $\|\tilde{v}\|_1 \leq \|v\|_1$, thus completing the proof. $\square$

We conclude the proof of Theorem 9.3.4 thanks to the following result.

Lemma 9.3.11. *The infimum in (9.3.1) can be taken on controls v such that*

1. $v \leq 0$ if $T < T_\Gamma$;
2. $v \geq 0$ if $T > T_\Gamma$.

Proof. Assume that $T > T_\Gamma$. In this case, the drift drives the system beyond the target. Namely, letting z_0 be the solution of (9.3.2) with $v \equiv 0$, we have $z_0(T) > z_f$ and thus, for any control v admissible for $\text{rmc}_\varepsilon(z_f, T)$, it holds that $v(t) > 0$ for some $t \in [0, T]$. Let us show that for any control v that is admissible for $\text{rmc}_\varepsilon(z_f, T)$ and takes negative values, there exists a non-negative admissible control with smaller L^1 -norm. By Lemma 9.3.9, we can restrict to consider the case where v is affine. Letting $t_* := \inf\{t \in [0, T] \mid v(t) < 0\} \in [0, T]$, define

$$\tilde{v}(t) = \begin{cases} v(t) & \text{if } t \in [0, t_*], \\ 0 & \text{if } t \in [t_*, T]. \end{cases}$$

Observe that if $v(0) = 0$, then $t_* = 0$, and hence $\tilde{v}(t) = 0$ for all $t \in [0, T]$.

There are three possibilities: 1. $z_{\tilde{v}}(T) < z_f$, 2. $z_{\tilde{v}}(T) = z_f$, 3. $z_{\tilde{v}}(T) > z_f$. Clearly, if $z_{\tilde{v}}(T) = z_f$, then $\tilde{v}$ is the control we were looking for.

Assume that we are in the first case. Let us consider the function $\phi(t) := z_0(T - t; z_v(t)) - z_f$, where $z_0(s; \zeta)$ denotes the solution of the Cauchy problem $\dot{z} = a_0(z)$, $z(0) = \zeta$,

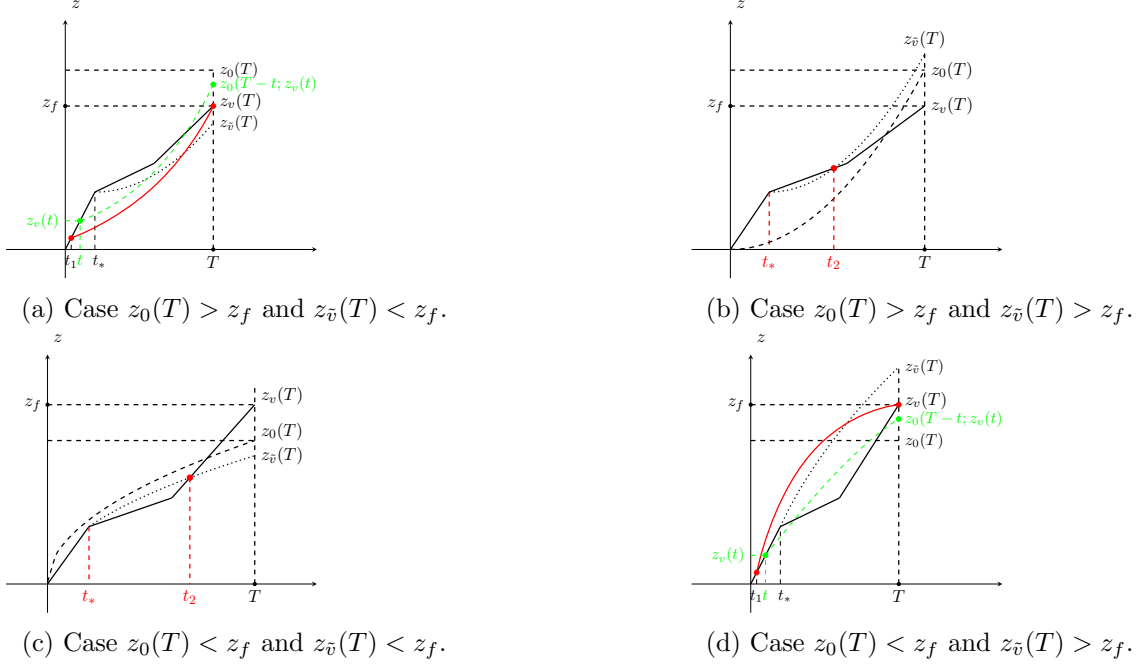


Figure 9.2: Construction for Lemma 9.3.11.

evaluated at time s . By the continuous dependence on the initial value of a solution of a differential equation, we have that ϕ is continuous. Moreover, it is positive at $t = 0$, and we have $\phi(t_*) = z_{\tilde{v}}(T) - z_f < 0$. So, there exists a time $t_1 \in [0, t_*]$ such that the control equal to v up to time t_1 and 0 for $t \geq t_1$ steers the system from 0 to z_f . By construction, this control is never negative and its L^1 -norm is smaller or equal than the norm of v .

In the third case, it holds that $z_{\tilde{v}}(T) > z_v(T)$. Moreover, $T > T_\Gamma$ implies $a_0 > 0$ on the whole curve Γ , and hence $\dot{z}_v(t_*) > \dot{z}_{\tilde{v}}(t_*)$. Then, there exists $t_2 \in [t_*, T]$ such that $z_v(t_2) = z_{\tilde{v}}(t_2)$. This implies that the control

$$\tilde{v}(t) = \begin{cases} v(t) & \text{if } t \in [0, t_*] \cup [t_2, T], \\ 0 & \text{otherwise,} \end{cases}$$

steers the system from 0 to z_f and $\|\tilde{v}\|_1 \leq \|v\|_1$. This ends the case $T > T_\Gamma$. Assume now $T < T_\Gamma$. By Lemma 9.3.10 it holds that $0 < \dot{z}_v(t) = a_0(z(t)) - \alpha(z(t))v(t)$, which implies that $v(t) < a_0(z(t))/\alpha(z(t))$. Hence, $v(t) \leq 0$ for any t such that $a_0(z(t)) \leq 0$. This allows us to restrict to the case of $a_0 > 0$ on the whole Γ . In this case, one can adapt the argument for the case $T > T_\Gamma$ to show that $v \leq 0$ (see Figure 9.2). We have that $\phi(0) < 0$, where ϕ is defined as before. In case $z_{\tilde{v}}(T) < z_f$, we have that the trajectory $z_{\tilde{v}}$ intersects the trajectory z_0 at some $t_2 > t_* := \inf\{t \in [0, T] \mid v(t) > 0\}$, so that one can argue as for

point 3 in the case $T > T_\Gamma$. On the other hand, if $z_{\bar{v}}(T) > z_f$ we have that $\phi(t_*) > 0$, so one can argue as for point 1 in the case $T > T_\Gamma$. $\square$

9.4 Step 2 case: proof of Theorem 9.1.2

In this section, we assume that $T \neq T_\Gamma$. First, we define the following quantities

$$I^*(\Gamma, T) = \sup \left\{ \int_0^T \frac{a_0(z(t))}{\alpha(z(t))} dt \mid z \text{ solution of (9.3.2) with } v \leq 0 \right\},$$

$$I_*(\Gamma, T) = \inf \left\{ \int_0^T \frac{a_0(z(t))}{\alpha(z(t))} dt \mid z \text{ solution of (9.3.2) with } v \geq 0 \right\}.$$

Lemma 9.4.1. *Under the assumptions of Theorem 9.1.2, we have*

1. If $T < T_\Gamma$: $\text{MC}_\varepsilon(\Gamma, T) \simeq \frac{2}{\varepsilon^2} \left[\int_0^S \frac{dz}{\alpha(z)} - I^*(\Gamma, T) \right].$
2. If $T > T_\Gamma$: $\text{MC}_\varepsilon(\Gamma, T) \simeq \frac{2}{\varepsilon^2} \left[I_*(\Gamma, T) - \int_0^S \frac{dz}{\alpha(z)} \right].$

Remark 9.4.2. It is possible to give a direct proof that the quantities in the squared brackets are indeed positive. We omit this, since it will follow directly from the explicit formulas for $I^*(\Gamma, T)$ and $I_*(\Gamma, T)$ given in Lemma 9.4.3 and Lemma 9.4.4.

Proof. We only present an argument for the case $T < T_\Gamma$, since the other follows by a straightforward adaptation of the same argument. By Theorem 9.3.3, it suffices to prove the statement replacing $\text{MC}_\varepsilon(\Gamma, T)$ by $\text{rmc}_\varepsilon(z_f, T)$, where $z_f = \int_\Gamma \omega$. Let v be an admissible control for $\text{rmc}_\varepsilon(z_f, T)$. By Lemma 9.3.11, we can assume $v \leq 0$. By (9.3.2), we can express v as a function of a_0, α . Then, for every such control, we have that

$$\|v\|_1 = - \int_0^T v dt = \int_0^T \frac{\dot{z} - a_0(z)}{\alpha(z)} dt = \int_0^{T_\Gamma} \frac{dz}{\alpha(z)} - \int_0^T \frac{a_0(z)}{\alpha(z)} dt.$$

Thus, taking the infimum w.r.t. non-positive admissible controls for $\text{rmc}_\varepsilon(\Gamma, T)$, we obtain

$$\text{rmc}_\varepsilon(\Gamma, T) = \frac{2}{\varepsilon^2} \left[\int_0^{T_\Gamma} \frac{dz}{\alpha(z)} - \sup_{v \leq 0} \int_0^T \frac{a_0(z)}{\alpha(z)} dt \right] = \frac{2}{\varepsilon^2} \left[\int_0^{T_\Gamma} \frac{dz}{\alpha(z)} - I^*(\Gamma, T) \right].$$

$\square$

In the following two lemmata, proved in Section 9.4.1, we present closed formulas for $I^*(\Gamma, T)$ and $I_*(\Gamma, T)$. Henceforth, for any $H \in \mathbb{R}$ and $c > 0$, we denote by $z(\cdot; H, c)$ the

solution of the ODE in (9.3.2), with initial condition $z(0) = 0$, and associated with the piece-wise constant feedback control given by

$$v(z; H, c) = \begin{cases} -c & \text{if } H > \frac{a_0(z)}{\alpha(z)}, \\ 0 & \text{if } H < \frac{a_0(z)}{\alpha(z)}. \end{cases} \quad (9.4.1)$$

Lemma 9.4.3. *If $T < T_\Gamma$, there exists a unique value of $H_\infty > 0$ for which it holds that $\lim_{c \rightarrow +\infty} z(T; H_\infty, c) = z_f$. Moreover, letting $\Omega = \{z \in [0, z_f] \mid \frac{a_0(z)}{\alpha(z)} > H_\infty\}$, we have that*

$$I^*(\Gamma, T) = \begin{cases} \int_\Omega \frac{dz}{\alpha(z)} & \text{if } |\Omega| > 0 \\ T \min_{q \in \Gamma} \frac{|\omega(X_0)|}{\alpha(q)} & \text{if } |\Omega| = 0. \end{cases}$$

Lemma 9.4.4. *If $T > T_\Gamma$, it holds that*

$$I_*(\Gamma, T) = \int_0^S \frac{dz}{\alpha(z)} + (T - T_\Gamma) \min_\Gamma \frac{a_0}{\alpha}.$$

We are finally in a position to complete the proof of the theorem.

Proof. The case $T = T_\Gamma$ follows from Proposition 9.2.3. The case $T_\Gamma > T$ is an immediate consequence of Lemma 9.4.1 and Lemma 9.4.4, recalling that $a_0(z) = \omega(X_0)|_{(0,0,z)}$. The case $T_\Gamma < T$ follows similarly, using Lemma 9.4.3 and observing that

$$\int_0^S \frac{dz}{\alpha(z)} = \int_\Gamma \frac{1}{\alpha}.$$

□

9.4.1 Proofs of the auxiliary lemmata

Proof of Lemma 9.4.3 First, observe that we have

$$I^*(\Gamma, T) = \sup \left\{ \int_0^T \frac{a_0(z)}{\alpha(z)} dt \mid \dot{z} = a_0(z) + \alpha(z)v, z(0) = 0, z(T) = z_f, v \geq 0 \right\}.$$

Here, we changed the sign of the control v to reduce possible confusion. For $c \in (0, +\infty]$, we let

$$V(c) := \inf \{ \|v\|_1 \mid 0 \leq v \leq c, z_v(T) = z_f \} \quad (9.4.2)$$

Then, since $\dot{z} = a_0(z) + \alpha(z)v$, we have $\int_0^T \frac{a_0(z(t))}{\alpha(z(t))} dt = \int_0^T \frac{\dot{z}(t)}{\alpha(z(t))} - v(t) dt = \int_0^{z_f} \frac{dz}{\alpha(z)} - \|v\|_{L^1}$, which implies

$$I^*(\Gamma, T) = \int_0^{z_f} \frac{dz}{\alpha(z)} - V(+\infty). \quad (9.4.3)$$

Hence, in the following we focus on computing $V(+\infty)$.

First, we notice that we can apply the Pontryagin Maximum Principle (PMP) to compute $V(c)$ for any $c < +\infty$. This will be enough thanks to the following result.

Proposition 9.4.5. *The function $V : (0, +\infty] \rightarrow \mathbb{R}_{\geq 0}$ is non-increasing. Moreover, $\lim_{c \rightarrow +\infty} V(c) = V(+\infty)$.*

Proof. The monotonicity of V is immediate from the definition. This implies the existence of the limit and the fact that $\lim_{c \rightarrow +\infty} V(c) \geq V(+\infty)$. To see the opposite inequality, we consider a maximizing sequence $(v_j)_{j \in \mathbb{N}}$ for $V(+\infty)$. Since any v_j is in L^∞ , then it is admissible for $V(\|v_j\|_\infty)$. This implies $V(\|v_j\|_\infty) \leq \|v_j\|_1$ for every $j \in \mathbb{N}$ and, in particular, $\lim_{j \rightarrow +\infty} V(\|v_j\|_\infty) \leq \lim_{j \rightarrow +\infty} \|v_j\|_1 = V(+\infty)$. $\square$

Now, we distinguish two cases: first, we consider the case of $\max_{z \in [0, z_f]} a_0(z) \geq 0$; then, we study the case $\max_{z \in [0, z_f]} a_0(z) < 0$.

Applying the PMP (see [2]) to (9.4.2), we obtain the following.

Lemma 9.4.6. *Assume that $\max_{z \in [0, z_f]} a_0(z) \geq 0$. Let $\tilde{v}$ be an optimal solution to (9.4.2) for c sufficiently big. Then there exists $H \geq 0$ such that $\tilde{v}$ is the feedback control $\tilde{v}(t) = -v(z(t); H, c)$, where $v(z; H, c)$ is defined in (9.4.1).*

Proof. In order to apply the PMP, we define the Hamiltonian function

$$h(\xi, \xi_0, z, v) = \xi(a_0(z) + \alpha(z)v) + \xi_0 v, \quad v \in [0, c], \quad \xi, z \in \mathbb{R}, \quad \xi_0 \in \{0, -1\}. \quad (9.4.4)$$

Here, the choice $\xi_0 = -1$ corresponds to normal extremals, while $\xi_0 = 0$ corresponds to abnormal extremals. By PMP, an extremal is a curve $(\xi(t), z(t))$, where ξ satisfies the adjoint equation

$$\dot{\xi} = -\xi(a'_0(z) + \tilde{v}\alpha'(z)). \quad (9.4.5)$$

Here we denoted by a'_0 (resp. α') the derivative of a_0 (resp. α) w.r.t. z . Moreover, the value of the control $\tilde{v}$ maximizes the value of h , and hence the control satisfies

$$\tilde{v}(t) = \begin{cases} c & \text{if } \xi\alpha(z) + \xi_0 > 0, \\ 0 & \text{if } \xi\alpha(z) + \xi_0 < 0. \end{cases} \quad (9.4.6)$$

If $\xi\alpha(z) + \xi_0 = 0$, $\tilde{v}$ is not determined directly, i.e., the control is singular.

We start by showing that there are no abnormal extremals. These are associated with $\xi_0 = 0$, but $\xi \neq 0$ for all times. As a consequence, (9.4.6) reads

$$\tilde{v}(t) = \begin{cases} c & \text{if } \xi > 0, \\ 0 & \text{if } \xi < 0, \end{cases}$$

and, thus, either $v \equiv 0$ or $v \equiv c$. But, since $T \neq T_\Gamma$, up to considering c sufficiently big, these controls do not steer the system from 0 to z_f .

We are then left with the normal case ($\xi_0 = -1$). We now claim that there are no singular controls, which by (9.4.6) amounts to asking that $\xi\alpha(z) \neq 1$ a.e. on $[0, T]$. We can differentiate the identity $\xi = 1/\alpha$ to obtain:

$$\dot{\xi} = -\frac{\alpha'(z)\dot{z}}{\alpha(z)^2} = -\frac{\alpha'(z)(a_0(z) + \alpha(z)\tilde{v})}{\alpha(z)^2} \quad (9.4.7)$$

Putting together (9.4.7) and (9.4.5), and using again $\xi = 1/\alpha$, yields $(\log(a_0/\alpha))' = 0$, which is equivalent to $(a_0/\alpha)' = 0$. By Lemma 9.2.1, up to perturbing the drift X_0 we can assume that this identity cannot hold for z in an interval $I \subset [0, z_1]$. So, this implies that the trajectory z is constant. However, one can show that, under the hypothesis of (9.4.2), trajectories for z which are constant on some time interval are not optimal.

From (9.4.6), we have thus shown that the control $\tilde{v}$ stands defined by

$$\tilde{v}(t) = \begin{cases} c & \text{if } \xi(t) > \frac{1}{\alpha(z(t))}, \\ 0 & \text{if } \xi(t) < \frac{1}{\alpha(z(t))}. \end{cases} \quad (9.4.8)$$

Let H be the constant value of the Hamiltonian h along the extremal trajectory. Then, solving for ξ the definition (9.4.4) of h we have $\xi(t) = \frac{H + \tilde{v}(t)}{a_0(z(t)) + \alpha(z(t))\tilde{v}(t)}$. Since a_0 and α are bounded, up to considering c sufficiently big, we have $a_0 + \alpha c > 0$. So, the statement follows from the fact that $\xi = 1/\alpha(z)$ is equivalent to $H > a_0(z)/\alpha(z)$, which is easily checked using (9.4.8) and the definition of H . $\square$

From Lemma 9.4.6, we have simply to find for which value of $H \in \mathbb{R}$ the control $\tilde{v}$ defined in (9.4.1) steers the system from $z(0) = 0$ to $z(T) = z_f$.

Claim 9.4.7. *Assume that $\max_{z \in [0, z_f]} a_0(z) \geq 0$. For any $c \in \mathbb{R}$ sufficiently large, there is a unique value $H(c) > 0$ of H such that the control given in (9.4.8) is the optimal control.*

For $H > 0$, let $E_c(H) := z(T; H, c)$ be the solution of $\dot{z} = a_0(z) + \alpha(z)v$, $z(0) = 0$, evaluated at time T , where v is defined as in (9.4.1). Then, E_c is monotone non-decreasing. Indeed, if $H_1 < H_2$, then $\dot{z}(t; H_1, c) \leq \dot{z}(t; H_2, c)$. Moreover, since $T < T_\Gamma$, we have that $\lim_{H \rightarrow 0+} E_c(H) < z_f$. On the other hand, for any fixed $H > \max_{z \in [0, z_f]} a_0(z)/\alpha(z)$, we have that $z(T; H, c) \rightarrow +\infty$ if $c \rightarrow +\infty$. So, if c is large enough, we know that $E_c(H) > z_f$ for $H > \max_z a_0(z)/\alpha(z)$. Since E_c is continuous, this proves the Claim. $\square$

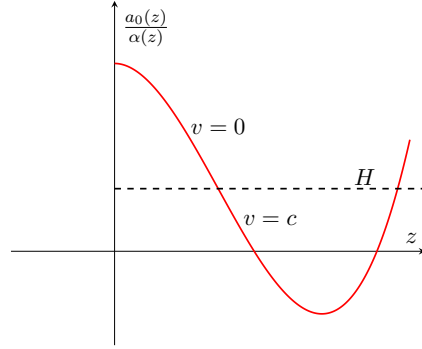


Figure 9.3: Graphical representation of the extremal controls. If the value of a_0/α is above H , then $v = 0$ and $\dot{z} = a_0$. When the value of a_0/α is below H , then $v = c$. As c grows, the extremal trajectories move faster and faster through the region where $a_0/\alpha < H$.

Claim 9.4.8. *The function $c \mapsto H(c)$ is monotone non-increasing. In particular, the following limit exists $H_\infty := \lim_{c \rightarrow +\infty} H(c)$.*

Suppose, by contradiction, that there are $c_1 < c_2$ such that $H(c_1) < H(c_2)$. Then, denoting by z_1, z_2 the trajectories of the control system corresponding to the resulting controls, we have that $\dot{z}_2 \geq \dot{z}_1$, with strict inequality on a set of positive measure. So, $z_2(T) > z_1(T) = z_f$, which is impossible since z_2 must satisfy the same boundary conditions of z_1 . This proves the Claim. $\square$

Observe that H_∞ can be also characterized by the property

$$\lim_{c \rightarrow +\infty} E_c(H_\infty) = \lim_{c \rightarrow +\infty} z(T; H_\infty, c) = z_f. \quad (9.4.9)$$

Indeed, by the same argument above, the function $c \mapsto z(T; H, c)$ is strictly increasing for any $H > 0$, so the limit in (9.4.9) exists and the limit function is again a monotone function of H . So, (9.4.9) uniquely determines the value of H_∞ . Concerning the limit trajectory z_∞ as $c \rightarrow +\infty$, the optimal control strategy is characterized as follows (see also Figure 9.3): whenever $a_0(z(t))/\alpha(z(t)) > H_\infty$, we apply zero control and the trajectory evolves according to $\dot{z} = a_0(z)$; whenever $a_0(z(t))/\alpha(z(t)) \leq H_\infty$, we apply maximum control to reach the next point $z' \in (z(t), z_f]$ where $a_0(z')/\alpha(z') = H_\infty$.

It remains to compute the value of $I^*(\Gamma, T)$, which can be done exploiting (9.4.3). To this aim, we let

$$\Omega(c) = \{z \in [0, z_f] \mid a_0(z)/\alpha(z) > H(c)\}. \quad (9.4.10)$$

Then, denoting by z_c the optimal solution of $I^*(c)$, and omitting its dependence on t , we

have

$$\begin{aligned} I^*(c) &= \int_0^T \frac{a_0(z_c)}{\alpha(z_c)} dt = \int_{z_c^{-1}(\Omega(c))} \frac{\dot{z}_c}{\alpha(z_c)} dt + \int_{[0,T] \setminus z_c^{-1}(\Omega(c))} \frac{\dot{z}_c - \alpha(z_c)c}{\alpha(z_c)} dt = \\ &= \int_{\Omega(c)} \frac{dz}{\alpha(z)} + \int_{[0,z_f] \setminus \Omega(c)} \frac{dz}{\alpha(z)} - c \int_{[0,z_f] \setminus \Omega(c)} \frac{dz}{a_0(z) + \alpha(z)c} \xrightarrow{c \rightarrow +\infty} \int_{\Omega} \frac{dz}{\alpha(z)}, \end{aligned}$$

Since $\lim_{c \rightarrow +\infty} I^*(c) = I^*$, this proves Lemma 9.4.3 in the case $\max_{z \in [0, z_f]} a_0(z) \geq 0$.

In the case $\max_{z \in [0, z_f]} a_0(z) < 0$, we have the following analogue of Lemma 9.4.6

Lemma 9.4.9. *Assume that $\max_{z \in [0, z_f]} a_0(z) < 0$. Let $\tilde{v}$ be an optimal solution to (9.4.2). Then, for c sufficiently big, there are $t_1, t_2 \in [0, T]$, $t_1 < t_2$ such that $\tilde{v}(t) = c$ for $t \in [0, t_1] \cup [t_2, T]$ and $\tilde{v}(t) = |a_0(z_{\min})|/\alpha(z_{\min})$ for $t \in (t_1, t_2)$, where $z_{\min} := \arg \min |a_0(z)|/\alpha(z)$.*

The proof of this Lemma closely follows the one of Lemma 9.4.6. The difference here is that, since $a_0(z) < 0$ for every $z \in [0, z_f]$, for any $H > 0$ we have that $v(t; H, c) = c$ for every $t \in [0, T]$. For such control it clearly holds that $z(T; H, c) \rightarrow +\infty$ as $c \rightarrow +\infty$, hence it is not admissible for our control problem. In order to arrive at the final point z_f at time T , we can use a singular control that keep z constant for the time needed, which corresponds to the time interval (t_1, t_2) in the statement. In this case, I_* is

$$I^* = T \frac{|a_0(z_{\min})|}{\alpha(z_{\min})} = T \min_{q \in \Gamma} \frac{|\omega(X_0)|}{\alpha(q)}.$$

and hence $\text{MC}_\varepsilon(\Gamma, T) \simeq \frac{2}{\varepsilon^2} \left(\int_{\Gamma} \frac{1}{\alpha} \omega - T \min_{q \in \Gamma} \frac{|\omega(X_0)|}{\alpha(q)} \right)$. This finally concludes the proof of Lemma 9.4.3.

Proof of Lemma 9.4.4 Similar computations as those leading to (9.4.3) yield

$$I_*(\Gamma, T) = \int_0^{z_f} \frac{dz}{\alpha(z)} + V(+\infty).$$

We can then proceed as in the case of $I^*(\Gamma, T)$: we reduce to the optimal control problem with controls $v \in [0, c]$, then apply the PMP and take the limit as $c \rightarrow +\infty$.

We have the same extremals as before. However, since by Lemma 9.3.10 we can restrict to consider trajectories with $\dot{z} \geq 0$, for c large enough it cannot happen that $v(t) = c$. Thus, the only possible trajectories are concatenations of arcs that either correspond to the control $v = 0$ or that are such that z remains constant.

By assumption, we have that $\min_z a_0(z)/\alpha(z) > 0$. We know that if we simply apply the control $v = 0$ for every $t > 0$ we get to the point z_f in time $T_\Gamma < T$. So, in particular, with this control strategy, we arrive at a point $z_{\min}$ realizing the minimum of a_0/α at some

time $\tau_0 < T_\Gamma$. Consider the control

$$\bar{v}(t) = \begin{cases} 0 & \text{if } t \in [0, \tau_0], \\ \frac{a_0(z_{\min})}{\alpha(z_{\min})} & \text{if } t \in [\tau_0, \tau_0 + T - T_\Gamma], \\ 0 & \text{if } t \in [\tau_0 + T - T_\Gamma, T]. \end{cases}$$

The control clearly drives the system to the desired endpoint, and, since for $v = 0$ we have $\dot{z} = a_0(z)$, we have $\int_0^T \frac{a_0(z(t))}{\alpha(z(t))} dt = \int_\Gamma \frac{1}{\alpha} + (T - T_\Gamma) \frac{a_0(z_{\min})}{\alpha(z_{\min})}$. Observe that $\bar{v}$ is an extremal control. It is easily checked that any other extremal control has bigger or equal cost and thus that $\bar{v}$ is the optimal control.

9.5 Distributions of generic type

The aim of this Section is to prove Theorem 9.1.8. As discussed in the Introduction, if $\dim M \geq 4$, then the step 2 assumption in Theorem 9.1.2 is generic. If $\dim M = 3$, this is no longer true. For this reason, in this section we deal with the case $\dim M = 3$, where the pair (Δ, Γ) is of generic type (see Definition 9.1.6).

We split the proof of Theorem 9.1.8 in two parts. First, in Section 9.5.1 we deal with the case of a curve Γ with a single Martinet point, see Theorem 9.5.3. After that, in Section 9.5.2, we show how to reduce the general case to the first one, using some properties of the motion complexity.

To simplify the statements, throughout all the remaining part of this section, Δ is a corank 1 distribution on M , Γ is a smooth simple curve transverse to Δ , ω is a 1-form associated with (Δ, Γ) , which we suppose to be a pair of generic type. As before, we fix a parametrization of Γ such that $\omega(\dot{\Gamma}) = 1$. Moreover, g is a sub-Riemannian metric over (M, Δ) and the drift X_0 is tangent to Γ .

9.5.1 Case of a single Martinet point

In this section we assume that on the curve Γ there is only one Martinet point, which we denote by $\bar{q}$.

We start by recalling, from Theorem 9.2.4, a convenient coordinate frame to carry out our computations. First, thanks to Theorem 9.2.4, we fix coordinates $(x, y, z) : \mathbb{R}^3 \rightarrow \text{Tube}_{\varepsilon_0}(\Gamma)$ such that $\bar{q} = (0, 0, 0)$, $\Gamma(s) = (0, 0, s)$ for $s \in [-S, S]$, and the expression of the drift X_0 and of an orthonormal frame $\{X_1, X_2\}$ for the metric g is

$$\begin{aligned} X_0(x, y, z) &= a_0(x, y, z)\partial_z + O(|x| + |y|) \\ X_1(x, y, z) &= \partial_x - \frac{y}{2}\gamma(x, y, z)\partial_z + O(x^2 + y^2), \\ X_2(x, y, z) &= \partial_y + \frac{x}{2}\gamma(x, y, z)\partial_z + O(x^2 + y^2). \end{aligned} \tag{9.5.1}$$

Here, a_0, γ are smooth functions and $|\gamma(x, y, z)| = \alpha(x, y, z)$, where $\alpha = |\omega([X_1, X_2])|$ is the function defined in (9.1.5).

Observe that the function γ appearing above differs from the one appearing in Theorem 9.2.4 if $(x, y) \neq (0, 0)$, but it follows from the following result that the two functions differ by a $O(x^2 + y^2)$ term.

Lemma 9.5.1. *Let $\bar{q}$ be a Martinet point of Γ , i.e., such that $\alpha(\bar{q}) = 0$. Then, α is Lipschitz continuous and smooth on the set $\{\alpha > 0\}$. Moreover, there exists a neighborhood U of $\bar{q}$ such that the set $\{\alpha > 0\} \cap U$ has exactly 2 connected components, A_+ and A_- , which are such that $(0, 0, z) \in A_+$ if $z > 0$ and $(0, 0, z) \in A_-$ if $z < 0$. Then, there exists $(\beta_1, \beta_2) \neq (0, 0)$ such that, in coordinates, we have*

$$\alpha(x, y, z) = \begin{cases} \kappa z + \beta_1 x + \beta_2 y + O(x^2 + y^2 + z^2), & \text{on } A_+, \\ -\kappa z - \beta_1 x - \beta_2 y + O(x^2 + y^2 + z^2), & \text{on } A_-. \end{cases} \quad (9.5.2)$$

Here, $\kappa > 0$ is defined by (9.1.11).

Proof. The first part of the statement is a direct consequence of the definition of α and of the generic type assumption. Formula (9.5.2) follows directly from a Taylor expansion of α at the point $(0, 0, 0)$. Indeed, by Lemma 9.6.1, we have that $\kappa = \lim_{z \rightarrow 0} |\partial_z \alpha(0, 0, z)|$. The last requirement in the generic type assumption implies that $\kappa > 0$. Finally,

$$(\beta_1, \beta_2, \kappa)^T = \lim_{\substack{(x, y, z) \rightarrow 0 \\ (x, y, z) \in A_+}} \nabla \alpha(x, y, z) = \nabla \gamma(0, 0, 0) = - \lim_{\substack{(x, y, z) \rightarrow 0 \\ (x, y, z) \in A_-}} \nabla \alpha(x, y, z),$$

due to the fact that $\alpha = \gamma$ on A_+ and $\alpha = -\gamma$ on A_- . $\square$

Now, using the notation of Definition 9.1.6, we have two possibilities: either $\Sigma = \Sigma^+ = \{\bar{q}\}$ and $\Sigma^- = \emptyset$ or, vice-versa, $\Sigma = \Sigma^- = \{\bar{q}\}$ and $\Sigma^+ = \emptyset$. In both cases, the definition of T_Γ that we gave before is still meaningful (see (9.1.4)). Hence, for the first case, we have the following result, which is a direct adaptation of Theorem 9.1.2.

Theorem 9.5.2. *Assume that $\Sigma = \Sigma^+ = \{\bar{q}\}$ and $\Sigma^- = \emptyset$. Then, the same asymptotic equivalences as in Theorem 9.1.2 hold. Moreover, in the case $T_\Gamma > T$, we have that $\bar{q} \in \text{int } \Omega$. In particular, the integral in (9.1.8) is finite.*

Proof. One can prove the first two points by proceeding exactly in the same way as in Theorem 9.1.2. The only thing worth noticing is that, since $\lim_{q \rightarrow \bar{q}} a_0(q)/\alpha(q) = +\infty$, the minimum in (9.1.6) is not attained at $\bar{q}$.

Also the third point follows as in Theorem 9.1.2, with minor adaptations. More precisely, the only fact one has to check is that the integral in (9.1.8) is finite. To show this, we prove that the point $\bar{q} \in \text{int } \Omega$, so that $1/\alpha$ is bounded on $\Gamma \setminus \Omega$. To prove this, recall the definitions of Ω and the value H_∞ , see Equation (9.4.10) and Claim 9.4.8. From Claim

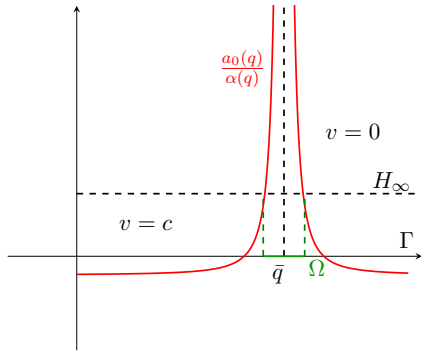


Figure 9.4: Graphical representation of for Theorem 9.5.2: The point $\bar{q}$ lies in the interior of Ω .

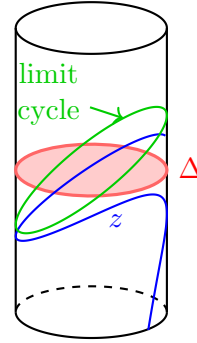


Figure 9.5: Graphical representation of the trajectory z of point 1. in Claim 9.5.5.

9.4.8, we know that $H_\infty \leq H(0) < +\infty$. Hence, H_∞ must be finite. As a consequence, since $\lim_{q \rightarrow \bar{q}} a_0(q)/\alpha(q) = +\infty$, the set $V := \{q \in \Gamma \mid a(q)/\alpha(q) > 2H_\infty\}$ is an open neighbourhood of $\bar{q}$ and $\bar{V} \subset \Omega$. Thus, the point $\bar{q} \in \text{int } \Omega$ (see Figure 9.4). $\square$

On the other hand, the second case $\Sigma = \Sigma^- = \{\bar{q}\}$ is more delicate.

Theorem 9.5.3. *Suppose that $\Sigma = \Sigma^- = \{\bar{q}\}$, that is, there is a single Martinet point on the curve Γ and we have $\omega_{\bar{q}}(X_0(\bar{q})) < 0$. Then, letting κ be defined as in (9.1.11), it holds the following*

$$\text{MC}_\varepsilon(\Gamma, T) \simeq -\frac{4}{\kappa \varepsilon^2} \ln \varepsilon, \quad \varepsilon \rightarrow 0. \quad (9.5.3)$$

Remark 9.5.4. Notice that the above asymptotic depends only on κ , hence only on the local structure of Γ near $\bar{q}$: it involves neither the whole curve nor the time T .

Proof. Let us consider the coordinates and the orthonormal frame introduced in equation (9.5.1). In these coordinates, after some simplifications very similar to those carried out in Section 9.3.1, the control system (9.1.1) can be expressed in cylindrical coordinates as

$$\begin{cases} \dot{r} = v_1, \\ \dot{\theta} = \frac{v_2}{r}, \\ \dot{z} = a_0(z) + \frac{r}{2} \gamma(z) v_2. \end{cases} \quad (9.5.4)$$

Here, with again a slight abuse of notation, we write $a_0(z)$ and $\gamma(z)$ in place of $a_0(0, \cdot, z)$ and $\gamma(0, \cdot, z)$. As we showed in the case of a step 2 distribution, these simplifications in the control system affect the asymptotic of MC_ε starting from the order $O(\varepsilon^{-1})$. Hence, they will not change the final result that we are going to obtain. We omit the details about these

reductions, since it would be a very close repetition of all the arguments already exposed in Section 9.3.1.

Fix $\varepsilon > 0$. As in the step 2 case (see Proposition 9.4.5), we can fix a uniform bound $c > 0$ on the norm of the controls and then take the limit as $c \rightarrow +\infty$. Again as it was done in the proof of the step 2 case, applying Pontryagin Maximum Principle one can prove that an optimal control is the concatenation of *bang* arcs, and for each arc we have either $v_2(t) = c$ or $v_2(t) = 0$. Now, we restrict to a sufficiently small neighbourhood of the Martinet point $\bar{q}$, so that the Taylor expansion in Lemma 9.5.1 holds, and we want to find a lower bound for the L^1 -norm of a bang arc with $v_2(t) = c$ in this neighbourhood of $\bar{q}$. As a consequence, from now on, we restrict to consider just controls of maximal norm steering the system from some initial point $z(t_1) < 0$ to a final point $z(t_2) > 0$ not depending on ε .

By (9.5.4), we have that there exists $C, K > 0$ independent of ε such that

$$|\dot{z}(t)| \leq C + \frac{c\varepsilon}{2}(\kappa|z(t)| + \varepsilon K), \quad \forall t \in [t_1, t_2].$$

So, applying Gronwall's Lemma, we obtain

$$|z(t_2)| \leq \left(\frac{\tilde{C} + \frac{c\varepsilon^2}{2}K}{\frac{c\varepsilon}{2}\kappa} + |z(t_1)| \right) e^{\frac{c\varepsilon}{2}\kappa(t_2-t_1)} = \left(\frac{2\tilde{C}}{c\varepsilon\kappa} + \frac{\varepsilon K}{\kappa} + |z(t_1)| \right) e^{\frac{\varepsilon}{2}\kappa\|v\|_1},$$

where we have used that $c(t_2 - t_1) = \|v\|_1$. Since $c \rightarrow +\infty$ and $\varepsilon > 0$ is fixed, we can assume $c\varepsilon^2 > 1$, so that $2\tilde{C}/c\varepsilon\kappa \leq 2\tilde{C}\varepsilon/\kappa$. Taking the logarithm on both sides and neglecting lower-order terms, we obtain the desired estimate:

$$\|v\|_1 \geq \frac{2}{\kappa\varepsilon} \ln \left(\frac{\kappa}{\varepsilon\tilde{K}} |z(t_2)| \right) = -\frac{2}{\kappa\varepsilon} \ln \varepsilon + O\left(\frac{1}{\varepsilon}\right).$$

By this argument, it already follows that the complexity of tracking a curve containing a Martinet point is at least of order $O(-\varepsilon^{-2} \ln \varepsilon)$, independent of the boundary conditions and final time $T > 0$. Therefore, given a neighbourhood $V_{\bar{q}}$ of $\bar{q}$, it suffices to compute the asymptotic of $\text{MC}_\varepsilon(\Gamma \cap V_{\bar{q}}, T)$ without imposing exact boundary conditions or fixing T precisely. This is justified as follows: by Theorem 9.1.2, one can steer the system from $\Gamma(0)$ to a point near the Martinet point with cost $O(\varepsilon^{-2})$, which is asymptotically negligible compared to $O(-\varepsilon^{-2} \ln \varepsilon)$. Similarly, after crossing the region near $\bar{q}$, one can reach the final point $\Gamma(S)$ with an additional cost of $O(\varepsilon^{-2})$. Thus, the Martinet point dominates the total complexity.

The following claim completes the proof:

Claim 9.5.5. *The asymptotically optimal synthesis for the motion planning problem restricted to a neighbourhood of the Martinet point consists of three arcs:*

1. a first bang arc, with $v_1(t) = 0$ and $v_2(t) = c$, starting from a point $q_0 = (\varepsilon, \theta_0, z_0)$, with $z_0 < 0$, reaching a point $q_1 = (\varepsilon, \theta_1, z_1)$, with $z_1 > 0$ and $z_1 \leq \tilde{C}\varepsilon^3$, for some constant $\tilde{C} > 0$.

2. a second bang arc with $v_1(t) = c$ and $v_2(t) = 0$ on a short time interval, steering the system from $q_1 = (\varepsilon, \theta_1, z_1)$ to $q_2 = (\varepsilon, -\theta_1, z_2)$, where $z_1 > z_2 > 0$.
3. a third bang arc, with $v_1(t) = 0$ and $v_2(t) = c$, starting from a point $q_3 = (\varepsilon, -\theta_1, z_2)$ and reaching a point $q_4 = (\varepsilon, \theta_3, z_3)$, where $z_3 > 0$ is independent of ε .

Here, $c > C\varepsilon^{-4}$ for some constant $C > 0$. The L^1 -norm of the control of the first and the third bang arc can be bounded above by $-\frac{2}{\kappa\varepsilon} \ln \varepsilon$, while the norm of the second arc is of lower order.

Since optimal trajectories are made by concatenations of *bang* arcs and that, for $\varepsilon \ll 1$, the *bang* arcs with $v_2(t) = 0$ do not yield trajectories crossing the Martinet surface, we have that indeed the optimal synthesis must begin, as in point 1, with a bang arc with $v_2(t) = c$. Let us show that the maximal z coordinate obtained with such a bang control starting from q_0 is positive and of the order $O(\varepsilon^3)$. Moreover, we show that this can be done with L^1 -norm of the control bounded above by $-\frac{2}{\kappa\varepsilon} \ln \varepsilon$.

Since we are localized near a Martinet point, up to modify the asymptotic of the motion complexity by a $O(\varepsilon^{-1})$ term, we can assume that the drift is constant $a_0(z) = -a$, for some $a > 0$. By (9.5.2), we have that $\alpha(r, \theta, z) = -\kappa z + r\tilde{\alpha}_1 \cos(\theta + \tilde{\theta}) + o(|r| + |z|)$, $(r, \theta, z) \in A^-$, and for some constants $\tilde{\alpha}_1 \in \mathbb{R}$, $\tilde{\theta} \in [0, 2\pi)$ and $(r, z) \rightarrow 0$. Notice that, under the generic type assumption, we have that $\tilde{\alpha}_1 \neq 0$. Now, we can consider controls for which $r = \varepsilon$ and $v_2 = c$ constant. So, the solution for θ reads $\theta(t) = \frac{c}{\varepsilon}t + \theta_0$. Moreover, the equation for z for such controls is

$$\dot{z} = -a - \frac{\varepsilon c}{2} \kappa z + \frac{\varepsilon^2}{2} c \tilde{\alpha}_1 \cos(\theta(t) + \tilde{\theta}). \quad (9.5.5)$$

This equation can be integrated explicitly:

$$z(t) = e^{-\frac{\varepsilon \kappa c t}{2}} \left[z_0 - 2a \frac{e^{\frac{\varepsilon \kappa c t}{2}} - 1}{\varepsilon \kappa c} + \frac{\tilde{\alpha}_1 \varepsilon^3}{4 + \kappa^2 \varepsilon^4} \left(\kappa \varepsilon^2 \left(e^{\frac{\varepsilon}{2} \kappa c t} \cos \theta(t) - \cos \theta_0 \right) + 2 \left(e^{\frac{\varepsilon}{2} \kappa c t} \sin \theta(t) - \sin \theta_0 \right) \right) \right].$$

For fixed c and $t \rightarrow +\infty$, the trajectory has a limit cycle, whose equations are:

$$r = \varepsilon, \quad z = -\frac{2a}{\varepsilon \kappa c} + \frac{\tilde{\alpha}_1 \varepsilon^3}{4 + \kappa^2 \varepsilon^4} (\kappa \varepsilon^2 \cos \theta + 2 \sin \theta). \quad (9.5.6)$$

Notice that, since $\tilde{\alpha}_1 \neq 0$, the limit cycle is transverse to the Martinet surface $\{\alpha = 0\}$ (see Figure 9.5). The maximum value of z is attained at $\sin \theta \simeq 1$ and this is positive for every $\varepsilon > 0$ if and only if $-\frac{2a}{\varepsilon \kappa c} + \frac{\tilde{\alpha}_1 \varepsilon^3}{2} \geq 0$. Hence, we must choose $c \geq C\varepsilon^{-4}$ for $C = 8a/(\tilde{\alpha}_1 \kappa)$, otherwise the corresponding trajectories cannot have $z > 0$. So, for fixed $\varepsilon, t_1 > 0$ and

c sufficiently large, starting from $z_0 < 0$, we obtain that there is a point of the resulting trajectory arbitrarily close to the point of the limit cycle above the Martinet point, that is, we have $z(t_1) > 0$. Notice that, by (9.5.6), $z(t_1) \leq \tilde{\alpha}_1 \varepsilon^3 / 2$.

Now, we have to estimate the L^1 -norm of this control up to t_1 . Recall that the solution to (9.5.5) with constant control is

$$\begin{aligned} z(t) &= e^{\frac{-\varepsilon \kappa c t}{2}} z_0 - a \int_0^t e^{\frac{\varepsilon \kappa c (t-s)}{2}} ds + \frac{\tilde{\alpha}_1 \varepsilon^2}{2} \int_0^t e^{\frac{\varepsilon \kappa c (t-s)}{2}} \cos(\theta(s) - \tilde{\theta}) ds \\ &\geq e^{\frac{-\varepsilon \kappa c t}{2}} z_0 + \left(e^{\frac{-\varepsilon \kappa c t}{2}} - 1 \right) \left(\frac{2a}{\kappa c \varepsilon} + \frac{\tilde{\alpha}_1 \varepsilon}{\kappa c} \right), \end{aligned}$$

where we have used $\cos(\theta(s) - \tilde{\theta}) \geq -1$. So, choosing $z_0 = -\frac{1}{2} \left(\frac{2a}{\kappa c \varepsilon} + \frac{\tilde{\alpha}_1 \varepsilon}{\kappa c} \right)$, we obtain

$$\left(z(t) + \frac{2a + \tilde{\alpha}_1 \varepsilon^2}{\kappa c \varepsilon} \right) \frac{\kappa c \varepsilon}{2a + \tilde{\alpha}_1 \varepsilon^2} \geq \exp \left(-\frac{\kappa c \varepsilon t}{2} \right),$$

and taking the logarithm on both sides, for $t = t_1$ we have $\ln \left(z(t_1) \frac{\kappa c \varepsilon}{2a} + 1 \right) \geq -\frac{\kappa c \varepsilon t_1}{2}$. Neglecting lower order terms, we obtain

$$\|v\|_1 = ct_1 \leq -\frac{2}{\kappa \varepsilon} \ln \varepsilon. \quad (9.5.7)$$

Notice that this is the L^1 -norm of the control v required to reach a point $z(t) = O(\varepsilon^3)$, $z(t) > 0$. In particular, this depends only on the product ct_1 . Hence, up to enlarge c , we can take t_1 as small as we want. This ends the proof of point 1 in the Claim.

We have to prove point 2 and 3. By means of a constant control of the form $v_1(t) = c$, $v_2(t) = 0$ on a time interval of length $2\varepsilon/c$, we can go from a point $(\varepsilon, \theta, z_1)$ to $(\varepsilon, -\theta, z_1 - a\frac{2\varepsilon}{c})$. The L^1 -norm of this control is 2ε , so it is negligible in our asymptotic. If c is big enough, then $z_1 - a\frac{2\varepsilon}{c} > 0$, hence this piece of trajectory crosses the Martinet surface. Thus, also point 2 follows. Finally, reversing the strategy applied to point 1, one can reach a point whose z -coordinate is positive and independent of ε , and the required L^1 -norm of the control is the same cost as in Equation (9.5.7). Thus, the total cost to cross the region near the Martinet point is the right-hand side of (9.5.7) multiplied by 2, which yields the asymptotic (9.5.3) in the Theorem.

The fact that the switching from point 2 to point 3 yields an asymptotically optimal synthesis follows by the same argument as the one used to justify the switch from point 1 to point 2, applied to the backward problem obtained by considering the drift $-X_0$ and the curve Γ with reverse orientation. $\square$

9.5.2 Case of many Martinet Points

In this section we generalize Theorem 9.5.3 to the case of finitely many Martinet points, completing the proof of Theorem 9.1.8.

As before, let (Δ, Γ) be a pair of generic type, and denote by $q_1 := \Gamma(s_1), \dots, q_r := \Gamma(s_r)$ the points in Σ^- . We have that $\alpha(q_i) = 0$, $i = 1, \dots, r$, where α was defined in (9.1.5). Recall also the definition of $\kappa_1, \dots, \kappa_r$, see Equation (9.1.11). We have to prove that

$$\text{MC}_\varepsilon(\Gamma, T) \simeq \sum_{i=1}^r -\frac{4}{\kappa_i \varepsilon^2} \ln \varepsilon, \quad \varepsilon \rightarrow 0.$$

First, we show that the r.h.s. is a lower bound for the motion complexity and then we find a sequence of controls realizing the asymptotic.

Take any control u admissible for $\text{MC}_\varepsilon(\Gamma, T)$ and denote by q_u the corresponding trajectory. Let V_i be an open neighbourhood of q_i , $\bar{V}_i \subset \text{Tube}_\varepsilon(\Gamma)$, for every $i = 1, \dots, r$. Let $I_i := q_u^{-1}(V_i)$. Then, since the asymptotic in (9.5.3) does not depend on the exact initial and final points of the curve that we are approximating, the curve $q_u|_{I_i}$ is an admissible competitor of the motion complexity of $\Gamma \cap V_i$. Then, by Theorem 9.5.3, we must have $\int_{I_i} |u(t)| dt \geq -\frac{4}{\kappa_i \varepsilon^2} \ln \varepsilon$. On the other hand, we have that the motion complexity of Γ away from Martinet points is of order $O(\frac{1}{\varepsilon^2})$. Thus, putting everything together, we obtain

$$\frac{1}{\varepsilon} \int_0^T |u(t)| dt = \frac{1}{\varepsilon} \sum_{i=1}^r \int_{I_i} |u(t)| dt + O\left(\frac{1}{\varepsilon^2}\right) \geq \sum_{i=1}^r -\frac{4}{\kappa_i \varepsilon^2} \ln \varepsilon + O\left(\frac{1}{\varepsilon^2}\right), \quad (9.5.8)$$

and taking the infimum on the left hand side, we obtain the desired inequality.

To conclude, we notice that, applying the procedure already used in the proof of Theorem 9.5.3, it is possible to find a sequence of controls $(u_j)_{j \in \mathbb{N}}$ admissible for $\text{MC}_\varepsilon(\Gamma, T)$ and such that they realize the asymptotic as in Theorem 9.5.3 near each Martinet point q_i , that is, for every $i = 1, \dots, r$ we have $\lim_{j \rightarrow +\infty} \int_{I_i} |u_j(t)| dt = -\frac{4}{\kappa_i \varepsilon^2} \ln \varepsilon$. Thus, replacing u with such u_j in (9.5.8) and taking the limit as $j \rightarrow +\infty$, we obtain the desired result. This concludes the proof of Theorem 9.1.8. $\square$

9.6 Intrinsic definition of κ

Let ω be a form associated with (Δ, Γ) . This form defines a vector field W_Γ on Γ such that $W_\Gamma(q) \in T_q \Gamma$ and $\omega(W_\Gamma) = 1$. Pick any point $q_i \in \Sigma$. Then the quantities κ_i are defined in [69] as

$$\kappa_i = |\varphi'(0)|, \quad \text{where} \quad \varphi(t) := \frac{d}{dt} \Big|_{t=0} \omega([X_1, X_2](e^{tW_\Gamma}(q_i))).$$

Let us show that these coincide with the expression in (9.1.11).

Lemma 9.6.1. *Consider any extension of W_Γ in a neighbourhood of Γ such that $\omega(W_\Gamma) \equiv 1$. Then, for any $q_i \in \Sigma$ we have*

$$\varphi'(0) = \omega([W_\Gamma, [X_1, X_2]])(q_i) - d\omega(W_\Gamma, [X_1, X_2])(q_i). \quad (9.6.1)$$

In particular, this value is independent of the choice of the extension.

Proof. By Cartan formula, we have

$$\left. \frac{d}{dt} \right|_{t=0} \omega([X_1, X_2](e^{tW_\Gamma}(q_i))) = (\mathcal{L}_{[X_1, X_2]}\omega(W_\Gamma) - i_{[X_1, X_2]}d\omega(W_\Gamma))|_{q_i}$$

By definition of Lie derivative, we get

$$\mathcal{L}_{[X_1, X_2]}\omega(W_\Gamma) = [X_1, X_2](\omega(W_\Gamma)) - \omega([X_1, X_2], W_\Gamma) = \omega([W_\Gamma, [X_1, X_2]]).$$

Here, we used the fact that $\omega(W_\Gamma) \equiv 1$ and thus $[X_1, X_2](\omega(W_\Gamma)) = 0$. Finally, (9.6.1) follows from the fact that $i_{[X_1, X_2]}d\omega(W_\Gamma) = -d\omega(W_\Gamma, [X_1, X_2])$.

The fact that the resulting value is independent of the choice of the extension is a consequence of the definition of κ_i . Let us anyway check it directly by considering a different extension $\tilde{W}$. Since $\omega(\tilde{W}_\Gamma) = \omega(W_\Gamma)$ we have that $\tilde{W}_\Gamma = W_\Gamma + Y$ for some $Y \in \ker \omega = \Delta$. Then, the statement follows by linearity of the objects under consideration and the fact that $\omega([Y, [X_1, X_2]]) = d\omega(Y, [X_1, X_2])$ since both Y and $[X_1, X_2]$ belong to Δ at q . $\square$

Remark 9.6.2. If instead of ω we consider $\tilde{\omega} = \varphi\omega$, for any $\varphi > 0$, we have

$$\tilde{\kappa} = \tilde{W}_\Gamma(\varphi)\omega([X_1, X_2]) + \varphi \left. \frac{d}{dt} \right|_{t=0} \omega([X_1, X_2](e^{t\tilde{W}_\Gamma})) = W_\Gamma(\omega([X_1, X_2])) = \kappa,$$

where every function is evaluated at q_i and we used the identities $\omega([X_1, X_2])|_{q_i} = 0$ at the Martinet point q_i and $\tilde{W}_\Gamma = \frac{1}{\varphi}W_\Gamma$. So, κ does not depend on the choice of ω .

Chapter 10

Lyapunov exponents of linear switched systems

10.1 Introduction

Let

$$\dot{y} = A(t)y, \quad y \in \mathbb{R}^n,$$

be a linear system of ordinary differential equations; here $t \mapsto A(t)$, $t \geq 0$, is a measurable bounded family of $n \times n$ -matrices. Let $X(t)$ be the fundamental matrix of system (1.1), the solutions of (1.1) have a form: $y(t) = X(t)y(0)$.

Principal Lyapunov exponent of system (1.1) is defined as follows:

$$\ell(A(\cdot)) = \limsup_{t \rightarrow \infty} \frac{1}{t} \ln \|X(t)\|. \quad (10.1.1)$$

This quantity does not depend on the choice of norm in the space of matrices and is a natural measure of instability of system (1.1). Indeed, system (1.1) is asymptotically stable with an exponential convergence rate if and only if $\ell(A(\cdot)) < 0$.

Principal Lyapunov exponents play a key role in the general theory of Dynamical Systems, see [72], [20].

If $A(t) \equiv A$ is a constant matrix, then $\ell(A)$ is maximum of the real parts of the eigenvalues of A . Moreover, for any $A(\cdot)$ and any $c \in \mathbb{R}$, we have $\ell(A(\cdot) + cI) = \ell(A(\cdot)) + c$.

In what follows, we use standard notations: $\mathfrak{gl}_n(\mathbb{R})$ is Lie algebra of all real $n \times n$ -matrices and $GL_n(\mathbb{R})$ is Lie group of all nondegenerate real $n \times n$ -matrices. Similarly $\mathfrak{sl}_n(\mathbb{R})$ is Lie algebra of all real $n \times n$ -matrices with zero trace and $SL_n(\mathbb{R})$ is Lie group of all real $n \times n$ -matrices whose determinant equals 1.

Let $S \subset \mathfrak{sl}_n(\mathbb{R})$ be a compact subset. A switched system induced by S is the set of systems of ordinary differential equations

$$\dot{y} = A(t)y, \quad A(t) \in S, \quad 0 \leq t < \infty,$$

where $t \mapsto A(t)$, $t \geq 0$, is any measurable map with values in S .

Let $\mathcal{A}_S \subset L_\infty([0, \infty); S)$ be the set of all such maps. Principal Lyapunov exponent of the switched system is defined as follows:

$$\ell(S) = \sup_{A(\cdot) \in \mathcal{A}_S} \ell(A(\cdot)).$$

Clearly, $\ell(S) \geq \sup_{A \in S} \ell(A)$.

We say that the switched system is exponentially stable if $\ell(S) < 0$. Stability of switched dynamical systems is really important for applications and there is a big literature devoted to this subject, see for instance [10, 25, 28, 54, 53, 66, 67] and references there.

A more natural mathematical problem is to compute or at least to estimate principal Lyapunov exponent of any switched system; many results about stability can be naturally generalized in this direction. This chapter is devoted to the detailed analysis of the most simple but fundamental case of two matrices generating Lie algebra $\mathfrak{sl}_2(\mathbb{R})$. We give a simple explicit formula of the maximal Lyapunov exponent as a function of the matrices and explicitly describe switching functions which realize maximal Lyapunov exponents.

This work is mainly motivated by paper [10]. The paper [10] is about stability but if you analyse proofs of the main results you will find the characterization of Lie subalgebras $\mathfrak{g} \subset \mathfrak{gl}_n(\mathbb{R})$ such that

$$\ell(S) = \sup_{A \in S} \ell(A), \quad \forall S \subset \mathfrak{g}. \quad (10.1.2)$$

Namely, a Lie subalgebra $\mathfrak{g} \subset \mathfrak{gl}_n(\mathbb{R})$ has property (10.1.2) if and only if $\mathfrak{g}$ does not contain a subalgebra isomorphic to $\mathfrak{sl}_2(\mathbb{R})$.

In the current , we explicitly compute $\ell(\{A, B\})^1$ for any pair of matrices A, B which generates Lie algebra $\mathfrak{sl}_2(\mathbb{R})$. Such a formula automatically provides us with an explicit expression of $\ell(A, B)$ for matrices A, B of any size if these matrices generate a Lie algebra isomorphic to $\mathfrak{sl}_2(\mathbb{R})$.

Indeed, let $\Psi : \mathfrak{sl}_2(\mathbb{R}) \rightarrow Lie(A, B)$ be such isomorphism; then Ψ is a representation of $\mathfrak{sl}_2(\mathbb{R})$. The representation Ψ is a direct sum of irreducible representations, $\Psi = \bigoplus_{i=1}^k \Psi_i$, where $\Psi_i : \mathfrak{sl}_2(\mathbb{R}) \rightarrow \mathfrak{sl}_{n_i}(\mathbb{R})$, $i = 1, \dots, k$, $n_1 \geq \dots \geq n_k \geq 2$. Then

$$\ell(A, B) = (n_1 - 1)\ell(\Psi^{-1}(A), \Psi^{-1}(B)).$$

Indeed, for any $n \geq 2$ there is a unique up-to an isomorphism irreducible n -dimensional representation of $\mathfrak{sl}_2(\mathbb{R})$. The representation of the group $SL_2(\mathbb{R})$ generated by this representation of $\mathfrak{sl}_2(\mathbb{R})$ is the $(n - 1)$ th symmetric power of the basic two-dimensional representation. If $X \in SL_2(\mathbb{R})$ has eigenvalues λ, λ^{-1} , then the $(n - 1)$ th symmetric power of X has the eigenvalues $\lambda^{n-1}, \lambda^{n-3}, \dots, \lambda^{1-n}$. We hope that our computation will serve as

¹In what follows we use shortened notation $\ell(A, B) \doteq \ell(\{A, B\})$

a building block for eventual estimates of the Lyapunov exponents in much more general cases.

Our final result can be stated as follows. To simplify the notation, set $a = \text{tr}(A^2)$, $b = \text{tr}(B^2)$, $c = \text{tr}(AB)$.

Theorem 10.1.1 (Main Theorem). *Suppose that $A, B \in \mathfrak{sl}_2(\mathbb{R})$, with $a \geq b$.*

1. *If $a \geq 0$ and $c > a$ or $a < 0$ and $c \geq \sqrt{ab}$, then*

$$\ell(A, B) = \frac{1}{2} \sqrt{\frac{c^2 - ab}{c - \frac{a+b}{2}}};$$

2. *If $a \geq 0$ and $c \leq a$, then*

$$\ell(A, B) = \sqrt{\frac{a}{2}};$$

3. *If $a < 0$ and $c \leq -\sqrt{ab}$, then*

$$\ell(A, B) = \frac{2}{\pi} \frac{1}{\sqrt{\frac{-2}{a}} + 3\sqrt{\frac{-2}{b}}} \text{arcosh} \left(\frac{-c}{\sqrt{ab}} \right).$$

We point out that for every $A, B \in \mathfrak{sl}_2(\mathbb{R})$ such that $a, b \leq 0$, then $c^2 \geq ab$ (see Lemma 10.6.1).

As we will show in the proof of this result, we also find the trajectories realizing these exponents (see also Table 10.1). More specifically, in point 1. the value of the Lyapunov exponent is not attained exactly, but it is the supremum among a family of periodic piecewise constant right hand side switching faster and faster. We compute exactly this trajectories and its limit as the switching time goes to zero. In point 2. the Lyapunov exponent is exactly attained by the trajectory with a constant right hand side. In point 3. the Lyapunov exponent is attained by a fixed periodic piecewise constant right hand side. The whole is devoted to prove our Main Theorem and it is structured as follows: in Section 10.2 we show how computing the Principal Lyapunov exponent $\ell(A, B)$ can be reduced to an optimal control problem and we apply Pontryagin Maximum Principle (PMP) to this problem. In Section 10.3 we deduce from PMP some general properties of Pontryagin extremals. In particular, we divide them into two categories: *bang-bang* extremals and singular extremals. In Section 10.4 and 10.5 we study *bang-bang* and singular extremals respectively. In Section 10.6 we show which extremal is the optimal one, depending on $\text{tr}(A^2)$, $\text{tr}(B^2)$ and $\text{tr}(AB)$.

10.2 Reduction to OC Problem and application of PMP

As explained in the Introduction, we will study the switched system induced by two matrices $A, B \in \mathfrak{sl}_2(\mathbb{R})$:

$$\begin{cases} \dot{X}(t) = X(t)(u(t)A + (1 - u(t))B), \\ X(0) = \text{Id}, \end{cases} \quad (10.2.1)$$

where $X : [0, T] \rightarrow \text{SL}_2(\mathbb{R})$ and $u \in \mathcal{U} := \{v : [0, T] \rightarrow \{0, 1\} \mid v \text{ measurable}\}$. We suppose moreover that $A, B, [A, B]$ are linearly independent and, without loss of generality, $\text{tr}(A^2) \geq \text{tr}(B^2)$.

We remark that taking $A, B, [A, B]$ linearly independent is not restrictive, because otherwise $\ell(A, B)$ can be computed as in [10].

Notice that in (10.2.1) we write $\dot{X} = XA(t)$, whereas in the Introduction, the expression was $\dot{X} = A(t)X$. Using the Introduction's notation, system (10.2.1) should appear as $\dot{X}^T = X^T A(t)^T$, where the superscript “ T ” represents the transpose. For simplification, we will omit this superscript. We prefer to solve the system with A and X in reversed order because system (10.2.1) is left-invariant.

Now, our aim is to find the principal Lyapunov exponent $\ell(A, B)$. Fixed $u \in \mathcal{U}$, we will denote by $X(\cdot; u)$ the solution of (10.2.1) with control equal to u .

A straightforward way to relax our problem (10.2.1) is to let $u : [0, T] \rightarrow [0, 1]$ be a measurable function rather than restricting to $u(t) \in \{0, 1\}$. This choice is quite natural, since the trajectories with $u(t) \in [0, 1]$ can be uniformly approximated by trajectories with controls $u(t) \in \{0, 1\}$ (see, for instance, Theorem 8.2 in [13]).

This relaxation will allow us to apply classical results from Control Theory, such as Filippov theorem, which requires the convexity of the set of admissible velocities.

Given $X \in \text{SL}_2(\mathbb{R})$, we denote by $\lambda(X)$ the eigenvalue of X bigger than 1 if both eigenvalues are real. If X has complex eigenvalues, we denote by $\lambda(X)$ the real part of these eigenvalues. Another crucial observation is that if the norm in (10.1.1) is the norm induced by the Euclidean metric on $\mathbb{R}^2$, then its value coincides with the absolute value of $\lambda(X(T))$. Hence, if for every $T > 0$ we can compute

$$c(T) := \sup_{u \in \mathcal{U}} \lambda(X(T; u)),$$

then, we obtain that

$$\ell(A, B) = \limsup_{T \rightarrow +\infty} \frac{1}{T} \ln c(T). \quad (10.2.2)$$

Strictly speaking, the previous equality should be a “less than or equal to” a priori, since the optimal strategy to maximize λ up to time T may depend on T itself. However, a posteriori, we found that, once A and B are fixed, for any $T_1 < T_2$ the optimal control for the problem with final time T_1 is the restriction to $[0, T_1]$ of the optimal control for the problem with final time T_2 . For this reason, we obtain the equality in (10.2.2).

In terms of Control Theory, the optimal strategy will be one of the following: 1. Constant singular control (case 1 of the Main Theorem). 2. Constant control equal to 1 (case 2). 3. Bang-bang periodic control (case 3). These concepts will be defined later.

Now, instead of finding $\sup_u \lambda(X(T; u))$, one can find $\sup_u f(\lambda(X(T; u)))$, for a suitable f monotone in λ . A convenient choice of f is the trace $\text{tr}(X(T; u))$: besides being a monotone function of λ , it is a linear function of X and it is invariant after a change of basis.

Summing up all these observations, we can reduce the original problem of computing $\ell(A, B)$ into the following Optimal Control problem:

Problem 1. Let $A, B \in \mathfrak{sl}_2(\mathbb{R})$ such that $A, B, [A, B]$ are linearly independent and fix $T \geq 0$. Let $\mathcal{U} = \{u : [0, T] \rightarrow [0, 1] \mid u \text{ measurable}\}$ and for each $u \in \mathcal{U}$ denote by $X(\cdot; u)$ the solution of the following Cauchy problem on $\text{SL}_2(\mathbb{R})$:

$$\begin{cases} \dot{X}(t) = X(t)(u(t)A + (1 - u(t))B), \\ X(0) = \text{Id} \in \text{SL}_2(\mathbb{R}), \end{cases} \quad t \in [0, T]. \quad (10.2.3)$$

We want to find $\tilde{u} \in \mathcal{U}$ which maximizes $\text{tr}X(T)$:

$$\text{tr}X(T; \tilde{u}) = \max_{u \in \mathcal{U}} \text{tr}X(T; u). \quad (10.2.4)$$

If we are able to solve this problem for each fixed T , then from the estimate of $\text{tr}X(T; u)$ we can deduce an estimate for $\|X(T; u)\|$ which is uniform with respect to $u \in \mathcal{U}$ and depends only on T . Then, the desired value of the principal Lyapunov exponent follows.

Existence of maximum in (10.2.4) follows by Filippov Theorem (see, for instance, Theorem 10.1 in [13]).

So, we can apply Pontryagin Maximum Principle (PMP) to our problem. For a general reference, see Theorem 12.3 in [13]. We recall here the precise statement of PMP suited to our setting.

10.2.1 Pontryagin Maximum Principle

Let M be a smooth manifold of dimension n , $U \subset \mathbb{R}^m$, and $(f_u)_{u \in U}$, a family of smooth vector field on M . Suppose furthermore that $M \times \overline{U} \ni (q, u) \mapsto f_u(q)$ and $M \times \overline{U} \ni (q, u) \mapsto \frac{\partial f_u}{\partial q}(q)$ are continuous.

Fix $T > 0$ and define $\mathcal{U} = \{u : [0, T] \rightarrow U, u \in L_{loc}^\infty([0, T], U)\}$ the space of admissible controls. Fix $q_0 \in M$ and for all $u \in \mathcal{U}$ denote by $q_u : [0, T] \rightarrow M$ the solution to the Cauchy Problem

$$\begin{cases} \dot{q}(t) = f(q(t), u(t)), \\ q(0) = q_0. \end{cases} \quad (10.2.5)$$

Existence and uniqueness of solution to (10.2.5) are guaranteed by Carathéodory Theorem, see, for instance, [26].

Let $a : M \rightarrow \mathbb{R}$ be a smooth function. We define the cost functional

$$J(u) = a(q(T; u)), \quad u \in \mathcal{U},$$

where $q(\cdot; u)$ denotes the solution to (10.2.5) corresponding to the choice of the control $u \in \mathcal{U}$. In the following, given any $h \in T^*M$, we will denote by $\vec{h}$ the hamiltonian vector field of h , that is $dh = \sigma(\cdot, \vec{h})$, where σ is the standard symplectic form of T^*M .

Theorem 10.2.1 (Pontryagin Maximum Principle). *Let $\tilde{u} \in \mathcal{U}$ be a local maximum point for J . Define $h_u(\lambda) = \langle \lambda, f(q, u) \rangle$, for $q \in M$, $u \in U$ and $\lambda \in T_q^*M$. Then, there exists a Lipschitz curve $\lambda : [0, T] \rightarrow T^*M$ such that*

$$\begin{aligned} \lambda(t) &\neq 0, \\ \dot{\lambda}(t) &= \vec{h}_{\tilde{u}(t)}(\lambda(t)), \end{aligned} \tag{10.2.6}$$

$$h_{\tilde{u}(t)}(\lambda(t)) = \max_{u \in U} h_u(\lambda_t), \tag{10.2.7}$$

for almost all $t \in [0, T]$, and moreover

$$\lambda_T = d_{q_{\tilde{u}}(T)} a. \tag{10.2.8}$$

We will call (10.2.6) the Hamiltonian system associated to (10.2.5), (10.2.7) is the maximal condition and (10.2.8) is the transversality condition. We will call *extremal* any curve λ satisfying all the necessary condition given by PMP.

We recall now some results about left-invariant Hamiltonian systems on Lie groups. For more detailed reference see Chapter 18 in [13] and Section 7.6 in [5].

Let $M \subset GL(N)$ be a Lie Group and $\mathcal{M}$ its Lie algebra. Given $q \in M$, define $L_q(p) = qp$, $p \in M$, the multiplication on the left by q .

Take a Hamiltonian function $h \in C^\infty(T^*M)$ and define

$$\mathcal{H}(\xi, q) = h(L_{q^{-1}}^* \xi, q), \quad \xi \in T_q^*M, q \in M.$$

We will call $\mathcal{H}$ the trivialized Hamiltonian function of h . We say that h is left-invariant if $\mathcal{H}$ does not depend on q . In this case, the function $\mathcal{H}$ is smooth on $\mathcal{M}^*$.

Recall the definition of the adjoint representation: given $v \in \mathcal{M}$, $\text{adv}(w) = [v, w]$, where $w \in \mathcal{M}$. The Killing form K of $\mathcal{M}$ is defined by $K(v_1, v_2) := \text{tr}(\text{adv}_1 \circ \text{adv}_2)$, for $v_1, v_2 \in \mathcal{M}$.

Proposition 10.2.2. *Let M be a Lie group and take $h \in C^\infty(T^*M)$. Let $\mathcal{M}$ be the Lie algebra of M and K the Killing form of $\mathcal{M}$. Suppose that K is non-degenerate. If h is left-invariant, then its associated Hamiltonian system can be rewritten as*

$$\begin{cases} \dot{q} = L_{q*} d\mathcal{H}, \\ \dot{\xi} = (\text{ad } d\mathcal{H})^* \xi, \end{cases} \quad \text{or} \quad \begin{cases} \dot{q} = L_{q*} d\mathcal{H}, \\ \dot{\eta} = [\eta, d\mathcal{H}], \end{cases} \tag{10.2.9}$$

where $q : [0, T] \rightarrow M$, $\xi : [0, T] \rightarrow \mathcal{M}^*$, $\eta : [0, T] \rightarrow \mathcal{M}$ and ξ and η are dual with respect to K :

$$K(\eta, v) = \langle \xi, v \rangle \quad \forall v \in \mathcal{M}.$$

Now, we go back to the discussion about Problem 1.

Corollary 10.2.3 (PMP applied to Problem 1). *Let $\tilde{u}$ be an optimal solution to Problem 1. Define*

$$H(\eta, u) = \text{tr}(\eta(uA + (1 - u)B)) \quad \eta \in \mathfrak{sl}_2(\mathbb{R}), u \in [0, 1]. \quad (10.2.10)$$

Then there is $\eta : [0, T] \rightarrow \mathfrak{sl}_2(\mathbb{R})$ such that:

1. *there is some $\eta_0 \in \mathfrak{sl}_2(\mathbb{R})$ such that η satisfies the equation*

$$\begin{cases} \dot{\eta}(t) = [\eta(t), \tilde{u}(t)A + (1 - \tilde{u}(t))B], \\ \eta(0) = \eta_0, \end{cases} \quad (10.2.11)$$

for a.e. $t \in [0, T]$.

2. *the following two conditions hold*

$$H(\eta(t), \tilde{u}(t)) = \max_{u \in [0, 1]} H(\eta(t), u), \quad \text{for a.e. } t \in [0, T]. \quad (10.2.12)$$

$$\eta(T) = X_{\tilde{u}}(T) - \frac{\text{tr}(X_{\tilde{u}}(T))}{2} I. \quad (10.2.13)$$

We will call (10.2.10) the *adjoint equation* and its solution η *adjoint trajectory*.

Proof. We apply PMP as formulated in Theorem 10.2.1 to Problem 1: (10.2.10) is the trivialized Hamiltonian function corresponding to the one in the statement of PMP, identifying tangent and cotangent space of $\text{SL}_2(\mathbb{R})$ through Killing form of $\mathfrak{sl}_2(\mathbb{R})$, which is non-degenerate since $\mathfrak{sl}_2(\mathbb{R})$ is simple. Recall that Killing form K on $\mathfrak{sl}_2(\mathbb{R})$ can be computed as

$$K(M_1, M_2) = 4\text{tr}(M_1 M_2), \quad M_1, M_2 \in \mathfrak{sl}_2(\mathbb{R}).$$

The system (10.2.11) is simply the equation in the fibers of (10.2.6) (see (10.2.9)).

Condition (10.2.12) is a straightforward consequence of (10.2.7) and (10.2.10).

It remains to prove (10.2.13). We have to apply (10.2.8). Let $\text{tr} : \text{SL}_2(\mathbb{R}) \rightarrow \mathbb{R}$ be the trace operator on $\text{SL}_2(\mathbb{R})$ and set $X = X_{\tilde{u}}(T)$. In our setting (10.2.8) reads

$$\text{tr}(\eta(T)U) = \langle d_X \text{tr}, XU \rangle, \quad \forall U \in \mathfrak{sl}_2(\mathbb{R}).$$

On the other hand, we have

$$\langle d_X \text{tr}, XU \rangle = \frac{d}{dt} (\text{tr}(X e^{tU}))|_{t=0} = \text{tr}(XU).$$

Moreover $\text{tr}(XU) = \text{tr}((X - \frac{\text{tr}(X)}{2}I)U)$ and $X - \frac{\text{tr}(X)}{2}I \in \mathfrak{sl}_2(\mathbb{R})$. So, we obtained $K(\eta(T), U) = K((X - \frac{\text{tr}(X)}{2}I), U)$ for every $U \in \mathfrak{sl}_2(\mathbb{R})$. Since the Killing form of $\mathfrak{sl}_2(\mathbb{R})$ is a non-degenerate bilinear form, (10.2.13) follows. $\square$

In the remaining part of the chapter, we are going to find *all* the possible extremal trajectories that satisfy Corollary 10.2.3. We will see that there are four possibilities: constant control $u = 1$, two different kinds of periodic piecewise constant controls and singular controls (which will be defined later, Definition 10.3.5). Depending on the values of $\text{tr}(A^2)$, $\text{tr}(B^2)$, $\text{tr}(AB)$, some of them will be extremals and some of them will not.

	$a > 0$ and $c \geq a$	$a \geq 0$ and $b \leq c < a$	$a \geq 0$ and $c < b$	$a \leq 0$ and $c \geq a$	$a \leq 0$ and $c < b$
Regular constant control $u = 1$	Yes	Yes	Yes	No	No
Periodic control of 1 st kind	No	No	No	Yes	Yes
Periodic control of 2 nd kind	Yes	No	Yes	Yes	Yes
Singular constant controls	Yes	No	Yes	Yes	Yes

Table 10.1: Classification of the extremals, depending on the values of $a = \text{tr}(A^2)$, $b = \text{tr}(B^2)$, $c = \text{tr}(AB)$. In this table, for all the possible instances of a, b, c we wrote which are the possible kind of extremals among all the families that we found. The green cells are the optimal one.

10.3 General properties of extremals

In this Section we discuss some general properties of solutions of (10.2.11). In particular, we are going to show that this Equation admits two first integrals. So, we can reduce the number of the degrees of freedom from 3 to 1. So, we will determine geometrically the trajectories of the solutions, without solving explicitly the equations.

In the following, if it is not important to specify the control we will denote by X the solution of (10.2.3) instead of $X(\cdot; u)$.

Proposition 10.3.1. *Let η be a solution of (10.2.11). Then*

$$\eta(t) = X(t)^{-1} \eta_0 X(t). \quad (10.3.1)$$

This is an immediate consequence of

$$\frac{d}{dt}[X(t)^{-1}] = -X(t)^{-1} \dot{X}(t) X(t)^{-1}.$$

In particular, we obtain the following Corollary.

Corollary 10.3.2. *Let η be a solution of (10.2.11). Then $\det \eta(t) = \eta_0 = \text{const.}$ for every $t \geq 0$.*

For $M \in \mathfrak{sl}_2(\mathbb{R})$, we have the formula $2 \det(M) = -\text{tr}(M^2)$. In what follows, it will be more convenient to use the expression $\text{tr}(M^2)$ instead of $\det M$.

Proposition 10.3.3. *Let $\tilde{u}$ be an optimal solution of (OCP) and let (X, η) be an extremal corresponding to $\tilde{u}$. Then, if $T > 0$ is large enough, it holds*

$$\text{tr}(\eta(t)^2) > 0 \quad \forall t \in [0, T].$$

In particular, $\eta(t)$ has real eigenvalues for any $t \in [0, T]$.

Proof. From transversality condition (10.2.13), it follows that

$$\text{tr}(\eta(T)^2) = \frac{1}{2}(\text{tr} X(T))^2 - 2.$$

So, since we want to maximize $\text{tr}(X(T))$, for large T we will have $\text{tr}(X(T))^2 > 4$. Moreover, since $\text{tr}(\eta^2)$ is constant, we can deduce that $\text{tr}(\eta(t)^2) > 0$ for all $t \in [0, T]$. $\square$

10.3.1 Invariant Hyperboloid

We now want to give a geometric interpretation of Corollary 10.3.2. Equation (10.2.11) is invariant after a dilation, in the sense that for any $c \in \mathbb{R}$, if $\eta(\cdot)$ is a solution with initial value $\eta(0) = \eta_0$, then $\tilde{\eta}(t) = c\eta(t)$ is a solution with initial value $\tilde{\eta}(0) = c\eta_0$.

So, it is not restrictive to suppose that $\text{tr}(\eta(t)^2) = \text{tr}(\eta_0^2) = 1$. Hence, the trajectory η is contained in the set

$$H := \{M \in \mathfrak{sl}_2(\mathbb{R}) \mid \text{tr}(M^2) = 1\},$$

which is a (connected) hyperboloid in the three-dimensional vector space $\mathfrak{sl}_2(\mathbb{R})$. To see this, just take

$$M = \begin{pmatrix} m_1 & m_2 \\ m_3 & -m_1 \end{pmatrix}.$$

Then

$$\text{tr}(M^2) = 2(m_1^2 + m_2 m_3) = 2 \left(m_1^2 + \left(\frac{m_2 + m_3}{2} \right)^2 - \left(\frac{m_2 - m_3}{2} \right)^2 \right).$$

Since $A, B, [A, B]$ are linearly independent, we can write

$$M = xA + yB + z[A, B],$$

and the equation defining H reads

$$x^2 \text{tr}(A^2) + y^2 \text{tr}(B^2) + z^2 \text{tr}([A, B]^2) + 2xy \text{tr}(AB) = 1.$$

We recall an identity that we will use many times in this chapter.

Lemma 10.3.4. *Let $A, B \in \mathfrak{sl}_2(\mathbb{R})$. Then*

$$\mathrm{tr}([A, B]^2) = 2\mathrm{tr}(AB)^2 - 2\mathrm{tr}(A^2)\mathrm{tr}(B^2).$$

The proof is an elementary exercise of Linear Algebra and we omit it.

10.3.2 Switching function, *bang-bang* and singular extremals

Now we want to analyze the condition on the Hamiltonian (10.2.12). It can be rewritten more explicitly as

$$\begin{aligned} \mathrm{tr}(\eta(t)B) + \tilde{u}(t)\mathrm{tr}(\eta(t)(A - B)) &= \max_{u \in [0,1]} [\mathrm{tr}(\eta(t)B) + u\mathrm{tr}(\eta(t)(A - B))] \\ &= \mathrm{tr}(\eta(t)B) + \max_{u \in [0,1]} u\mathrm{tr}(\eta(t)(A - B)), \end{aligned}$$

for almost every $t \in [0, T]$. Hence, we are lead to define $\varphi : [0, T] \rightarrow \mathbb{R}$,

$$\varphi(t) = \mathrm{tr}(\eta(t)(A - B)). \quad (10.3.2)$$

We will call φ *switching function*. Notice that φ is at least Lipschitz. If φ has only isolated zeros, then a control u , in order to satisfy Pontryagin Maximum Principle, must be of the form

$$u(t) = \begin{cases} 1 & \text{if } \varphi(t) > 0, \\ 0 & \text{if } \varphi(t) < 0, \end{cases} \quad (10.3.3)$$

that is, the isolated zeros of φ are exactly the times when there is a *switch* in the dynamics. If $\varphi(t_0) = 0$, we will call t_0 a *switching time* and the corresponding $\eta(t_0)$ *switching point*. If φ has any nonisolated zero, then Pontryagin Maximum Principle does not determine the control directly, but, as we will see in Section 10.5, in our case it is still possible to bypass this obstacle and determine the control.

Definition 10.3.5 (*bang-bang control, singular control*). We will call *bang-bang* a control u and its corresponding extremal if φ has only isolated zeros.

If instead $\varphi \equiv 0$, we will call *singular* the corresponding control u and its extremal.

A control u which takes only extremal values, i.e. $u(t) \in \{0, 1\}$ for a.e. $t \in [0, T]$, is called *bang-bang*. We will first focus on *bang-bang* extremals and then we will discuss separately the singular case.

10.4 *Bang-bang* extremals

10.4.1 Properties of *bang-bang* extremals

The goal of this Section is to prove that extremal controls must be either constant or periodic piecewise constant.

Theorem 10.4.1. *Any admissible control corresponding to a bang-bang extremal must belong to one of the following two families of controls:*

- constant controls, that is either $u(t) \equiv 1$ or $u(t) \equiv 0$ for all $t \in [0, T]$;
- bang-bang periodic controls, that is controls in the form

$$u(t) = \begin{cases} 1 & \text{if } t \in [0, \tau), \\ 0 & \text{if } t \in [\tau, T/k), \end{cases} \quad (10.4.1)$$

for $k \in \mathbb{N}$ and $\tau \in (0, T/k)$ and then extended by periodicity on the whole interval $[0, T]$.

We explain briefly the idea behind this result. We have two first integrals for equation (10.2.11): the quantity $\det \eta(t)$ and the Hamiltonian $H(\eta(t), u(t))$. In the 3D linear space $\mathfrak{sl}_2(\mathbb{R})$, the set $\{\det \eta = \text{const}\}$ is a connected hyperboloid, while $\{H(\eta, 0) = \text{const}, \text{tr}(\eta(A - B)) < 0\}$ and $\{H(\eta, 1) = \text{const}, \text{tr}(\eta(A - B)) > 0\}$ form half-planes. Thus, any trajectory η solving (10.2.11) must be in the intersections of these surfaces, resulting in a concatenation of conic paths, which can be either unbounded or a loop. If unbounded, after a finite number of switches, the control remains constant. If instead η is contained in a loop, then the solution η and the switching function are periodic, making the control periodic as well.

Before proving Theorem 10.4.1, we prove the following useful remark: the solution η of (10.2.11) satisfies transversality condition (10.2.13) if and only if it is periodic.

Proposition 10.4.2. *A couple (X, η) solution of (10.2.3) and (10.2.11) respectively satisfies transversality condition (10.2.13) at time $T > 0$ if and only if η is periodic of period T .*

Proof. If (10.2.13) is satisfied, then

$$X(T)^{-1}\eta(0)X(T) = \eta(T) = X(T) - \frac{\text{tr}(X(T))}{2}\text{Id},$$

and simply multiplying both sides by $X(T)$ on the left and by $X(T)^{-1}$ on the left one obtains that also

$$\eta(0) = X(T) - \frac{\text{tr}(X(T))}{2}\text{Id} = \eta(T),$$

that is, η is periodic of period T .

Viceversa, if $\eta(0) = \eta(T)$, then, from $\eta(T) = X(T)^{-1}\eta(0)X(T)$, it follows that

$$X(T)\eta(0) = \eta(0)X(T),$$

that is, $\eta(0), \eta(T)$ and $X(T)$ commutes. But then also $X(T) - \frac{\text{tr}(X(T))}{2}\text{Id}$ and $\eta(T)$ commutes, and since they are both matrices in $\mathfrak{sl}_2(\mathbb{R})$, which is semisimple, they must be proportional, i.e. there is $c \in \mathbb{R}$ such that :

$$c\eta(T) = X(T) - \frac{\text{tr}(X(T))}{2}\text{Id}.$$

Hence, since we can always rescale η , transversality condition (10.2.13) is satisfied. $\square$

The remaining part of this section and Subection 10.4.2 are devoted to the proof of Theorem 10.4.1.

In order to find the switches of a trajectory η , we define the plane of switching points, that is the subset of $\mathfrak{sl}_2(\mathbb{R})$ where the switching function ϕ is zero:

$$\Pi = \{M \in \mathfrak{sl}_2(\mathbb{R}) \mid \text{tr}(MA) = \text{tr}(MB)\}.$$

Since the matrices $A, B, [A, B]$ form a basis of $\mathfrak{sl}_2(\mathbb{R})$ by hypothesis, we can write the equation $\text{tr}(MA) = \text{tr}(MB)$ as

$$xa + yc = xc + yb,$$

where, according to the notation introduced in the Main Theorem, we set $a = \text{tr}(A^2)$, $b = \text{tr}(B^2)$, $c = \text{tr}(AB)$. So, if $c \neq b$, we have

$$y = \frac{c - a}{c - b}x. \quad (10.4.2)$$

So, if we impose $\text{tr}(M^2) = 1$, that is if we intersect the hyperboloid H with the plane Π we obtain the equation of a conic

$$\frac{(2c - a - b)(c^2 - ab)}{(c - b)^2}x^2 + 2(c^2 - ab)z^2 = 1. \quad (10.4.3)$$

Definition 10.4.3 (Switching curve). We will call *switching curve* S the intersection curve $\Pi \cap H$, whose equation is given by (10.4.2), (10.4.3).

So, the coefficient in front of x is positive if and only if

$$(2c - a - b)(c^2 - ab) > 0.$$

Hence, if $c^2 - ab > 0$, then the switching curve is an ellipse if $c > \frac{a+b}{2}$, and is an hyperbola if $c^2 - ab > 0$ and $c < \frac{a+b}{2}$.

If instead $c^2 - ab < 0$, the switching curve is an hyperbola if $c < \frac{a+b}{2}$. Notice that $c > \frac{a+b}{2}$ implies $c^2 - ab > 0$, so we can neglect the case $c > \frac{a+b}{2}$ and $c^2 - ab < 0$.

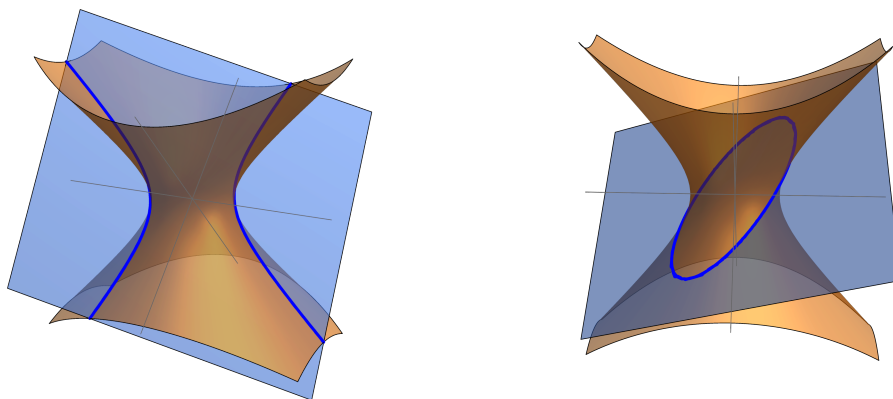


Figure 10.1: Examples of switching curves. In both pictures you can see the orange hyperboloid H , the blue plane Π and their intersection, the switching curve S .

10.4.2 Hyperboloid foliation

Let us now fix an initial point on the hyperboloid:

$$\eta_0 \in H.$$

Define the half spaces and the half hyperboloids

$$\begin{aligned} \Pi^+ &= \{M \in \mathfrak{sl}_2(\mathbb{R}) \mid \text{tr}(MA) > \text{tr}(MB)\}, & H^+ &= H \cap \Pi^+, \\ \Pi^- &= \{M \in \mathfrak{sl}_2(\mathbb{R}) \mid \text{tr}(MA) < \text{tr}(MB)\}, & H^- &= H \cap \Pi^-. \end{aligned}$$

From condition (10.3.3), we can see that if $\eta_0 \in H^+$, the control u , in order to satisfy PMP, must be equal to 1 until the trajectory η meets the switching curve S . Notice that, as far as $u(t) = 1$, the Hamiltonian function of the problem is $H(\eta(t), 1) = \text{tr}(\eta(t)A)$, hence the quantity $\text{tr}(\eta(t)A)$ is constant.

Similarly, if $\eta(t) \in H^-$ for $t \in [\tau_1, \tau_2]$, for some $0 < \tau_1 < \tau_2$, then $H(\eta(t), 0) = \text{tr}(\eta(t)B)$ is constant.

Define the planes

$$\begin{aligned} \Pi_{A,\gamma} &= \{M \in \mathfrak{sl}_2(\mathbb{R}) \mid \text{tr}(MA) = \gamma\} \\ \Pi_{B,\gamma} &= \{M \in \mathfrak{sl}_2(\mathbb{R}) \mid \text{tr}(MB) = \gamma\}, \end{aligned}$$

for $\gamma \in \mathbb{R}$.

So, if $\eta_0 \in H^+$ is such that $\text{tr}(\eta_0 A) = \gamma$, then the trajectory η is contained in the curve

$H^+ \cap \Pi_{A,\gamma}$ (see Figure 10.3). In particular, we can suppose that geometrically the trajectory η coincides with the whole curve $H^+ \cap \Pi_{A,\gamma}$. Indeed, if η accumulates at some point, then the control $u(t)$ is identically equal to 1 or to 0 for every t sufficiently large and, as we will see, the corresponding Lyapunov exponents are easy to compute.

Now, we want to study all the possible configurations of H, S and the planes $\Pi_{A,\gamma}, \Pi_{B,\gamma}$. Consider the straight line given by the intersection of the three planes $\Pi, \Pi_{A,\gamma}, \Pi_{B,\gamma}$:

$$r_\gamma : \begin{cases} \text{tr}(MA) = \gamma, \\ \text{tr}(MB) = \gamma, \\ \text{tr}(MA) = \text{tr}(MB). \end{cases} \quad (10.4.4)$$

There are three cases: either this line does not intersect the hyperboloid H , either it is tangent to the hyperboloid, or it has two distinct intersection with the hyperboloid (See Figure 10.2).

We will see later that if r_γ is tangent, then the corresponding trajectory η is constant and in particular it is a singular trajectory. So, we will consider first the non-tangent cases and we postpone the tangent case to the Section of singular extremals.

If r_γ does not intersect the hyperboloid, then the trajectory η starting from η_0 does not meet the switching curve S . This means that η is confined either in the region H^+ or H^- , so the control u is identically equal to 1 or 0 for all times. If instead there are two distinct

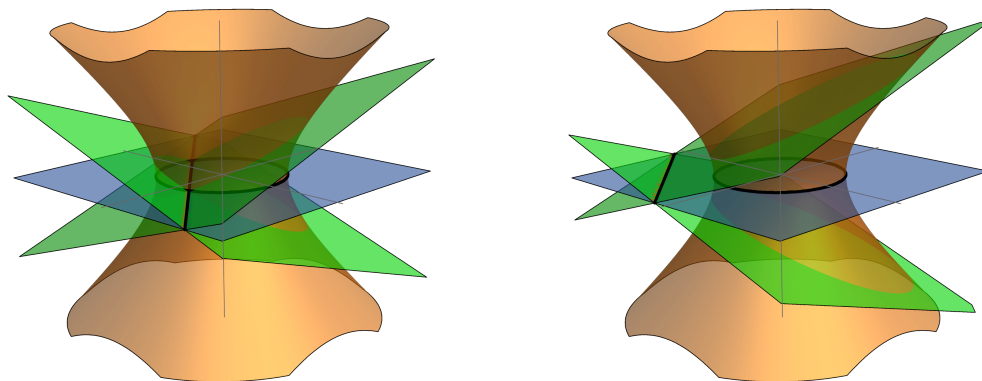


Figure 10.2: Example of possible position of the line r_γ .

intersection between r_γ and H , then there are three possibilities (see Figure 10.3):

1. there is exactly one switch;

2. there are two switches and after the second switch the adjoint trajectory does not meet anymore the switching curve;
3. the adjoint trajectory is periodic with infinite switches.

Since the trajectories with no switches give the same Lyapunov exponents as the trajectories with a finite number of switches, we will not consider them. This ends the proof of Theorem 10.4.1.

Before ending this Subsection, we provide explicit formulas for the coordinates of the switching points.

Using the same notation introduced before, we can write $\eta_0 = xA + yB + z[A, B]$. Then, from (10.4.2), (10.4.4) we obtain

$$x = \frac{b-c}{c^2-ab}\gamma, \quad y = \frac{a-c}{c^2-ab}\gamma, \quad z = \pm \frac{\sqrt{(a+b-2c)\gamma^2 + c^2 - ab}}{\sqrt{2}|c^2 - ab|}. \quad (10.4.5)$$

Thus, there are two distinct intersection if γ is such that

$$(a+b-2c)\gamma^2 + (c^2 - ab) > 0.$$

If $a+b-2c > 0$ and $c^2 > ab$ then the previous condition holds for every $\gamma \in \mathbb{R}$. If instead $a+b-2c > 0$ and $c^2 < ab$ this condition corresponds to

$$\gamma > \frac{1}{\sqrt{2}}\sqrt{\frac{2(c^2 - ab)}{2c - a - b}} \quad \text{or} \quad \gamma < -\frac{1}{\sqrt{2}}\sqrt{\frac{2(c^2 - ab)}{2c - a - b}}.$$

If instead $a+b-2c < 0$, this implies $c^2 > ab$, hence γ can assume the values

$$-\frac{1}{\sqrt{2}}\sqrt{\frac{2(c^2 - ab)}{2c - a - b}} < \gamma < \frac{1}{\sqrt{2}}\sqrt{\frac{2(c^2 - ab)}{2c - a - b}}.$$

Notice that if $\gamma = 0$, then the line r_γ is tangent to the hyperboloid.

As already mentioned in the Introduction, if the control u in equation (10.2.3) is constant equal to 1, then the corresponding principal Lyapunov exponent is equal to the positive eigenvalue of A , i.e.

$$\ell_{u=1}(A, B) = \sqrt{\frac{a}{2}}.$$

Analogously, if the control u is constant equal to 0, then

$$\ell_{u=0}(A, B) = \sqrt{\frac{b}{2}}.$$

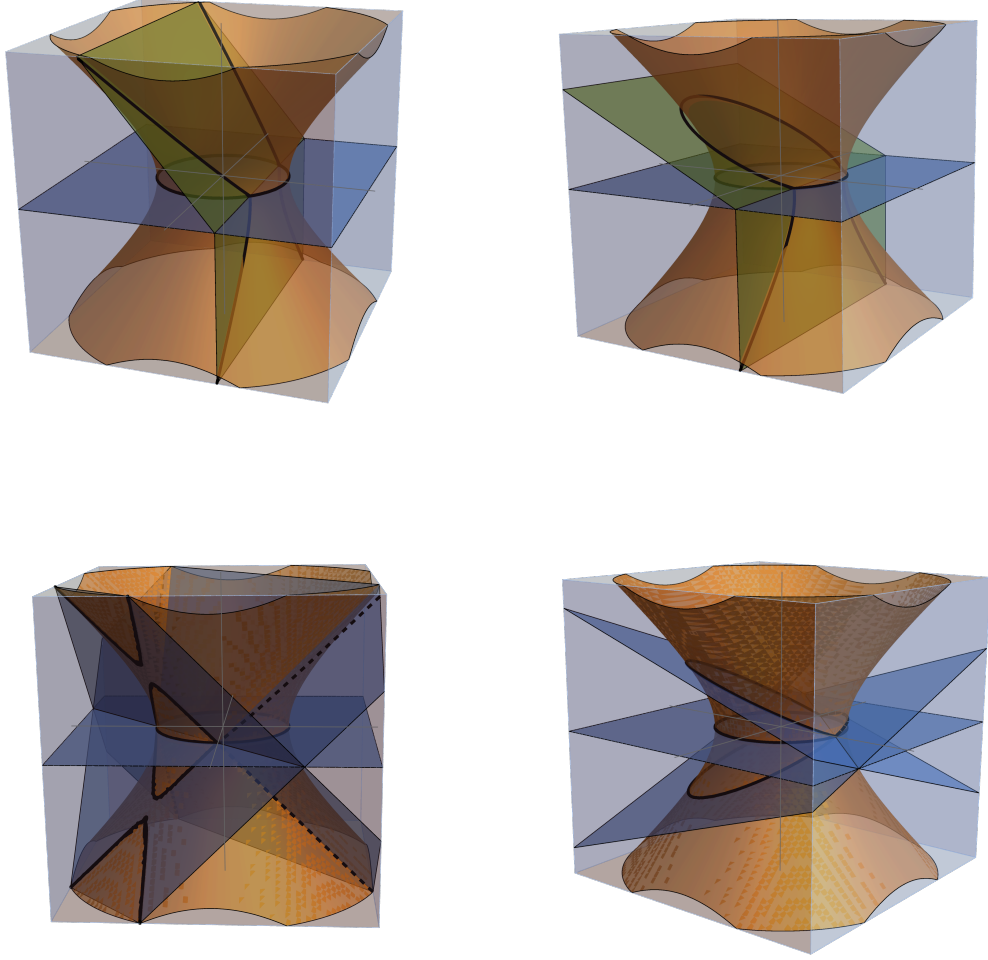


Figure 10.3: The figure illustrates the possible cases when the switching curve is a horizontal circle: in the first figure the intersection between the hyperboloid H and the planes $\Pi_{A,\gamma}$ and $\Pi_{B,\gamma}$ are two hyperbolas and the corresponding adjoint trajectory has only one switch. In the second figure the intersection with one plane is an ellipse and the other is a hyperbola, and the resulting adjoint trajectory has two switch. In the third case, the two intersections between the hyperboloid and the planes are hyperbolas, but in this case there are three connected components and one of them is closed, which corresponds to a periodic adjoint trajectory. The last figure is similar to the third one but with ellipses instead of hyperbolas.

Clearly, since we are assuming $a \geq b$, we have $\ell_{u=1} \geq \ell_{u=0}$.

Concerning of a periodic control, the principal Lyapunov exponent can be computed as follows: let T be the period of the control, then

$$\begin{aligned}\ell_{\text{per}}(T) &= \limsup_{n \rightarrow +\infty} \frac{1}{nT} \log \|X(nT)\| = \\ &= \limsup_{n \rightarrow +\infty} \frac{1}{nT} \log \|X(T)^n\| = \\ &= \limsup_{n \rightarrow +\infty} \frac{1}{nT} \log \lambda(T)^n = \frac{1}{T} \log \lambda(T),\end{aligned}$$

where $\lambda(T)$ is the greatest eigenvalue of $X(T)$. We can further simplify the expression for $\ell_{\text{per}}(T)$. Indeed, the functional of our Optimal Control Problem does not appear explicitly in the formula for $\ell_{\text{per}}(T)$. However, we can deduce the value the eigenvalue from the trace solving

$$\lambda(T)^2 - \text{tr}(X(T))\lambda(T) + 1 = 0,$$

which gives

$$\lambda(T) = \frac{\text{tr}(X(T)) + \sqrt{\text{tr}(X(T))^2 - 4}}{2}.$$

Using the identity $\text{arcosh } x = \log(x + \sqrt{x^2 - 1})$, we obtain

$$\ell_{\text{per}}(T) = \frac{\text{arcosh}\left(\frac{\text{tr}(X(T))}{2}\right)}{T}.$$

10.4.3 *Bang-bang* periodic controls

Up to this point, we have demonstrated that adjoint trajectories can be classified into two distinct categories: those without switches and those that are piecewise constant and periodic.

For periodic adjoint trajectories, the switching function (10.3.2) is also periodic, thereby the associated control is periodic too. Nevertheless, in order to satisfy PMP, we have to choose carefully the switching time τ in (10.4.1). In this Section, we will see how we can determine this τ .

From (10.3.1), finding all periodic solution to equation adjoint equation (10.2.11) amounts to finding $X(T)$ in the form

$$X(T) = e^{\tau A} e^{\sigma B},$$

with $\tau + \sigma = T$ and $0, \tau, T$ switching times.

We will proceed as follows: for every t we will find $\sigma = \sigma(\tau)$ such that $0, \tau, T = \tau + \sigma(\tau)$ are switching points. The following result simplifies this task.

Proposition 10.4.4. *Define the bang-bang control*

$$u(t) = \begin{cases} 1 & \text{if } t \in [0, \tau], \\ 0 & \text{if } t \in [\tau, T]. \end{cases}$$

Let X be the corresponding solution of (10.2.3) and η be the solution of (10.2.11). Suppose that η satisfies the condition of transversality (10.2.13) at time T . If τ is a switching time, then also 0 and T are switching times.

So, if we know that η satisfies transversality condition at T and τ is a switching time, we know that automatically T is a switching time, so we do not have to check that the switching condition is satisfied also at time T .

Proof. We have to check that $\text{tr}(\eta(T)A) = \text{tr}(\eta(T)B)$. From transversality condition (10.2.13), this is equivalent to $\text{tr}(X(T)A) = \text{tr}(X(T)B)$. transversality condition also implies, by Proposition 10.4.2, that η is periodic of period T , so it must be

$$\eta_0 = X(T) - \frac{\text{tr}(X(T))}{2} \text{Id}.$$

So, putting together this identity with the formula (10.3.1), we obtain that

$$\eta(\tau) = e^{\sigma B} e^{\tau A} - \frac{\text{tr}(X(T))}{2} \text{Id}.$$

We are assuming that τ to be a switching time, so it must satisfy

$$\text{tr}(\eta(\tau)A) = \text{tr}(\eta(\tau)B).$$

Using the formula for η that we have just obtained, we get

$$\text{tr}(e^{\sigma B} e^{\tau A} A) = \text{tr}(e^{\sigma B} e^{\tau A} B). \quad (10.4.7)$$

The left hand side can be rewritten as

$$\text{tr}(e^{\sigma B} e^{\tau A} A) = \text{tr}(e^{\sigma B} A e^{\tau A}) = \text{tr}(e^{\tau A} e^{\sigma B} A) = \text{tr}(X(T)A).$$

Similarly, the right hand side of (10.4.7) is equal to $\text{tr}(X(T)B)$. Hence, it follows that also

$$\text{tr}(X(T)A) = \text{tr}(X(T)B), \quad (10.4.8)$$

which implies that T (and thus, since η is T -periodic, also 0) is a switching time. $\square$

Remark 10.4.5. If we consider a solution of (10.2.3) with one switch, the final point is in the form

$$X(T) = X(\tau, \sigma) = e^{\tau A} e^{\sigma B}$$

where $T = \tau + \sigma$. So, we can see the cost functional as a function of t and s :

$$\Phi(\tau, \sigma) = \text{tr}(e^{\tau A} e^{\sigma B}).$$

With this notation, if τ is a switching point, from the previous proof follows that $X(T)$ must satisfy Equation (10.4.8), which can be rewritten as

$$\frac{\partial \Phi}{\partial \tau}(\tau, \sigma) = \frac{\partial \Phi}{\partial \sigma}(\tau, \sigma). \quad (10.4.9)$$

To summarize, given any $\tau > 0$, equation (10.4.9) give us a condition to find $\sigma > 0$ as a function of τ such that the control

$$u(t) = \begin{cases} 1 & \text{if } t \in [0, \tau], \\ 0 & \text{if } t \in [\tau, \tau + \sigma], \end{cases}$$

satisfies Pontryagin Maximum Principle on the time interval $[0, \tau + \sigma]$. The corresponding solution to the adjoint system η satisfies transversality condition on the same time interval, so one can extend this control by periodicity on the whole line $[0, +\infty)$.

10.4.4 Principal Lyapunov exponent for periodic controls

Now, we are going to compute the upper bound for Principal Lyapunov exponent of periodic controls. As we saw previously, in order to find the extremals of our problem we have to study the functional $\Phi : [0, +\infty) \times [0, +\infty) \rightarrow \mathbb{R}$

$$\Phi(\tau, \sigma) = \text{tr}(e^{\tau A} e^{\sigma B}).$$

Since equations (10.2.3) and (10.2.11) are invariant after a change of basis, if A and B are not nilpotent (see Remark 10.4.8 for the nilpotent case), without loss of generality we can assume $\text{spec}(A) = \{\lambda, -\lambda\}$, $\text{spec}(B) = \{\mu, -\mu\}$, with $\lambda, \mu \in \mathbb{R} \cup i\mathbb{R}$. Up to exchange the role of A and B , we can assume $\text{tr}(A^2) \geq \text{tr}(B^2)$.

Proposition 10.4.6. *With this notations, the functional in (10.4.4) is*

$$\Phi(\tau, \sigma) = 2 \cosh(\lambda\tau) \cosh(\mu\sigma) + \frac{\text{tr}(AB)}{\lambda\mu} \sinh(\lambda\tau) \sinh(\mu\sigma). \quad (10.4.10)$$

This formula is a simple consequence of the following Lemma.

Lemma 10.4.7. *If $M \in \mathfrak{sl}_2(\mathbb{R})$, with positive eigenvalue $\alpha > 0$, then*

$$e^{tM} = \cosh(\alpha t) \text{Id} + \frac{\sinh(\alpha t)}{\alpha} M. \quad (10.4.11)$$

Proof. From the characteristic polynomial of M , we know that $M^2 = \alpha^2 \text{Id}$, and so $M^{2k} = \alpha^{2k} \text{Id}$, $M^{2k+1} = \alpha^{2k} M$. Using these facts in the series expansion of e^{tM} one obtains the formula above. $\square$

Remark 10.4.8. If $M \in \mathfrak{sl}_2(\mathbb{R})$ is nilpotent, i.e. $M^2 = 0$, then

$$e^{tM} = \text{Id} + tM. \quad (10.4.12)$$

So, if either A or B is nilpotent, one can just replace formula (10.4.10) with a formula for Φ obtained using (10.4.12).

Imaginary eigenvalues

We are now going to examine a special case. Suppose that $\text{tr}(A^2), \text{tr}(B^2) < 0$, i.e. A, B have imaginary eigenvalues.

Proposition 10.4.9. *If both A, B have imaginary eigenvalues, the functional Φ is*

$$\Phi(\tau, \sigma) = \text{tr}(\exp(\tau A) \exp(\sigma B)) = 2 \cos(\lambda \tau) \cos(\mu \sigma) - 2\gamma \sin(\lambda \tau) \sin(\mu \sigma),$$

$$\text{Where } \gamma = \frac{|\text{tr}(AB)|}{\sqrt{\text{tr}(A^2)\text{tr}(B^2)}}.$$

This is simply a consequence of formula (10.4.10) and $\cosh(i\lambda\tau) = \cos(\lambda\tau)$, $\sinh(i\mu\sigma) = i \sin(\mu\sigma)$. With this formula, the switching equation (10.4.9) reads

$$(\gamma\lambda - \mu) \cos(\lambda\tau) \sin(\mu\sigma) = (\gamma\mu - \lambda) \cos(\mu\sigma) \sin(\lambda\tau),$$

from which we obtain immediately that there are always solutions for $\cos(\lambda\tau), \cos(\mu\sigma) = 0$, that is

$$(\tau, \sigma) \in \left\{ \left(\frac{\pi}{2\lambda}, \frac{3\pi}{2\mu} \right), \left(\frac{\pi}{2\lambda}, \frac{\pi}{2\mu} \right), \left(\frac{3\pi}{2\lambda}, \frac{\pi}{2\mu} \right) \right\}.$$

From these choices the value of this functional is $\Phi(\tau, \sigma) = \pm 2\gamma$. Using formula (10.4.2) one obtains

$$\ell_{\text{per,im}}(A, B) = \frac{1}{\tau + \sigma} \text{arcosh}(\gamma).$$

If $\text{tr}(AB) < 0$, then we have

$$\ell_{\text{per,im}}(A, B) = \frac{2}{\pi \left(\sqrt{\frac{-2}{\text{tr}(A^2)}} + 3 \sqrt{\frac{-2}{\text{tr}(B^2)}} \right)} \text{arcosh} \left(\frac{-\text{tr}(AB)}{\sqrt{\text{tr}(A^2)\text{tr}(B^2)}} \right).$$

Notice that we have chosen the couple (τ, σ) so to maximize $\frac{1}{\tau+\sigma}$ under the hypothesis $\text{tr}(B^2) < \text{tr}(A^2) < 0$.

If instead $\text{tr}(AB) > 0$, then

$$\ell_{\text{per,im}}(A, B) = \frac{2}{\pi \left(\sqrt{\frac{-2}{\text{tr}(A^2)}} + \sqrt{\frac{-2}{\text{tr}(B^2)}} \right)} \text{arcosh} \left(\frac{\text{tr}(AB)}{\sqrt{\text{tr}(A^2)\text{tr}(B^2)}} \right).$$

We will see in the next subsection how to deal with the extremals with $\cos(\lambda\tau), \cos(\mu\sigma) \neq 0$.

General case

We suppose now $\cosh(\lambda\tau), \cosh(\mu\sigma) \neq 0$ (we already saw the case $\cosh(\lambda\tau), \cosh(\mu\sigma) = 0$, which can happen if $\lambda, \mu \in i\mathbb{R}$). With this assumption, equation (10.4.9) becomes

$$\tanh(\mu\sigma) = \frac{\mu \text{tr}(AB) - \text{tr}(A^2)}{\lambda \text{tr}(AB) - \text{tr}(B^2)} \tanh(\lambda\tau). \quad (10.4.13)$$

If $\frac{\text{tr}(AB) - \text{tr}(A^2)}{\text{tr}(AB) - \text{tr}(B^2)} > 0$, then, for every $\tau > 0$ we can find $\sigma = \sigma(\tau)$ for which (10.4.13) holds. In particular

$$\sigma(\tau) = \frac{\text{tr}(AB) - \text{tr}(A^2)}{\text{tr}(AB) - \text{tr}(B^2)} \tau + o(\tau) \quad \text{as } \tau \rightarrow 0. \quad (10.4.14)$$

So, for every $\tau > 0$ we obtained an $\sigma(\tau)$ such that the *bang-bang* control defined by

$$u(t) = \begin{cases} 1 & \text{if } t \in [0, \tau), \\ 0 & \text{if } t \in [\tau, \tau + \sigma(\tau)), \end{cases}$$

and extended by periodicity for $t \in [\tau + \sigma(\tau), +\infty)$, satisfies all conditions of Pontryagin Maximum Principle.

Now, it remains just to determine for which τ the Principal Lyapunov exponent of the system is maximal. Define the family of controls

$$u_{\tau,\sigma}(t) = \begin{cases} 1 & \text{if } t \in [0, \tau), \\ 0 & \text{if } t \in [\tau, \tau + \sigma), \end{cases} \quad \tau, \sigma > 0,$$

and denote by $X_{\tau,\sigma}$ the solution of (10.2.3) with control equal to $u_{\tau,\sigma}$. Let $\alpha(\tau, \sigma)$ be the greatest eigenvalue of $X_{\tau,\sigma}(\tau + \sigma)$, so that the Principal Lyapunov exponent corresponding to the control $u_{\tau,\sigma}$ is

$$\ell(\tau, \sigma) = \frac{1}{\tau + \sigma} \log(\alpha(\tau, \sigma)).$$

The following result helps us in this task.

Lemma 10.4.10. *With the notation introduced above:*

1. *if $\text{tr}([A, B]^2) > 0$, then $\ell(\frac{\tau}{2}, \frac{\sigma}{2}) > \ell(\tau, \sigma)$;*
2. *if $\text{tr}([A, B]^2) < 0$, then $\ell(\frac{\tau}{2}, \frac{\sigma}{2}) < \ell(\tau, \sigma)$;*

We first show how to obtain the estimate for the Principal Lyapunov exponent, then we prove the Lemma.

Theorem 10.4.11. *If $\text{tr}([A, B]^2) > 0$, then for every $\tau > 0$ it holds*

$$\ell(\tau, \sigma(\tau)) < \lim_{\tau_1 \rightarrow 0} \ell(\tau_1, \sigma(\tau_1)) = \frac{1}{2} \sqrt{\frac{\text{tr}([A, B]^2)}{2\text{tr}(AB) - \text{tr}(A^2) - \text{tr}(B^2)}} \quad (10.4.15)$$

If instead $\text{tr}([A, B]^2) < 0$, then control with switches are not optimal.

Proof. First, suppose that $\text{tr}([A, B]^2) > 0$. One can check that this condition implies that the coefficient in (10.4.13) is smaller than 1:

$$\frac{\mu \text{tr}(AB) - \text{tr}(A^2)}{\lambda \text{tr}(AB) - \text{tr}(B^2)} < 1.$$

So, for every $\tau > 0$ it makes sense to take the inverse hyperbolic tangent in (10.4.13) on both sides. Moreover, from point (1) of Lemma 10.4.10, for any $\tau > 0$ we have

$$\ell\left(\frac{\tau}{2}, \frac{\sigma(\tau)}{2}\right) > \ell(\tau, \sigma(\tau))$$

In particular, we can find $\tau_1 < \tau$ such that $\tau_1 + s(\tau_1) = \frac{1}{2}(\tau + \sigma(\tau))$, so that

$$\ell(\tau_1, s(\tau_1)) > \ell\left(\frac{\tau}{2}, \frac{\sigma(\tau)}{2}\right) > \ell(\tau, \sigma(\tau)).$$

So, the inequality in (10.4.15) follows. To evaluate the limit, one can use Taylor expansions for $\sigma(\tau)$ as $\tau \rightarrow 0$.

Suppose now that $\text{tr}([A, B]^2) < 0$. Notice that this can happen only if $\text{tr}(A^2), \text{tr}(B^2) > 0$. Similarly to the previous case, this inequality implies

$$\frac{\mu \text{tr}(AB) - \text{tr}(A^2)}{\lambda \text{tr}(AB) - \text{tr}(B^2)} > 1.$$

So, we can find $\sigma = \sigma(\tau)$ solving (10.4.13) only for $\tau < \bar{\tau}$, where

$$\bar{\tau} = \frac{1}{\lambda} \text{arctanh} \left(\frac{\lambda \text{tr}(AB) - \text{tr}(A^2)}{\mu \text{tr}(AB) - \text{tr}(B^2)} \right).$$

Moreover, $\lim_{\tau \rightarrow \bar{\tau}} \sigma(\tau) = +\infty$. Then, from point (2) of Lemma 10.4.10, we have $\ell(2\tau, 2\sigma(\tau)) > \ell(\tau, \sigma(\tau))$. So, similarly to the previous case, this implies that we can find τ_1 such that

$$\ell(\tau_1, \sigma(\tau_1)) > \ell(\tau, \sigma(\tau)).$$

So, we obtain that

$$\ell(\tau, \sigma(\tau)) < \lim_{\tau_1 \rightarrow \bar{\tau}} \ell(\tau_1, \sigma(\tau_1)).$$

But as $\tau \rightarrow \bar{\tau}$, the trajectory spends more and more time with control equal to 0. So, this strategy is for sure worse than spending all the time with control equal to 1, which is always admissible if $\text{tr}([A, B]^2) < 0$. Hence, in this case switches are not optimal. $\square$

We now prove Lemma 10.4.10.

Proof of Lemma 10.4.10. Recall that

$$l(\tau, \sigma) = \frac{1}{\tau + \sigma} \log(\alpha(\tau, \sigma)).$$

Hence, we can write

$$\ell\left(\frac{\tau}{2}, \frac{\sigma}{2}\right) - \ell(\tau, \sigma) = \frac{2}{\tau + \sigma} \log\left(\alpha\left(\frac{\tau}{2}, \frac{\sigma}{2}\right)\right) - \frac{1}{\tau + \sigma} \log(\alpha(\tau, \sigma)) = \frac{1}{\tau + \sigma} \log\left(\frac{\alpha\left(\frac{\tau}{2}, \frac{\sigma}{2}\right)^2}{\alpha(\tau, \sigma)}\right).$$

So, we can see that this difference is positive if and only if

$$\frac{\alpha\left(\frac{\tau}{2}, \frac{\sigma}{2}\right)^2}{\alpha(\tau, \sigma)} > 1.$$

Since $\alpha(\tau, \sigma) > 1$ for any $\tau, \sigma > 0$, we have that

$$\alpha\left(\frac{\tau}{2}, \frac{\sigma}{2}\right)^2 > \alpha(\tau, \sigma) \quad \text{iff} \quad \alpha\left(\frac{\tau}{2}, \frac{\sigma}{2}\right)^2 + \frac{1}{\alpha\left(\frac{\tau}{2}, \frac{\sigma}{2}\right)^2} > \alpha(\tau, \sigma) + \frac{1}{\alpha(\tau, \sigma)},$$

which is the same as

$$\text{tr}\left(X\left(\frac{\tau}{2}, \frac{\sigma}{2}\right)^2\right) > \text{tr}(X(\tau, \sigma)).$$

The term in the left hand side can be rewritten as

$$X\left(\frac{\tau}{2}, \frac{\sigma}{2}\right)^2 = X(\tau, \sigma) + e^{\frac{\tau}{2}A} [e^{\frac{\sigma}{2}B}, e^{\frac{\tau}{2}A}] e^{\frac{\sigma}{2}B}.$$

So, in the end, it amounts just to compute $\text{tr}\left(e^{\frac{\tau}{2}A} [e^{\frac{\sigma}{2}B}, e^{\frac{\tau}{2}A}] e^{\frac{\sigma}{2}B}\right)$, which can be done using formula (10.4.11). One obtains

$$\text{tr}\left(e^{\frac{\tau}{2}A} [e^{\frac{\sigma}{2}B}, e^{\frac{\tau}{2}A}] e^{\frac{\sigma}{2}B}\right) = \frac{\sinh(\lambda \frac{\tau}{2})^2 \sinh(\mu \frac{\sigma}{2})^2}{2\lambda^2 \mu^2} \text{tr}([A, B]^2).$$

So, if $\text{tr}([A, B]^2) > 0$, then we obtained point (1) of the Lemma. If instead $\text{tr}([A, B]^2) < 0$, then all inequalities are reversed and we obtain point (2). $\square$

Remark 10.4.12. Notice that this is a completely general result about product of matrix exponentials. Indeed, it holds for a generic control $u_{\tau,\sigma}$, even if it does not correspond to a Pontryagin extremal.

Limit trajectory as switching time tends to zero

In the previous Subsection we saw that if $\text{tr}([A, B]^2) > 0$, then the upper bound for the principal Lyapunov is obtained as the switching time tends to zero. Now, we want to discuss some more details about the limit trajectory as the switching time goes to zero.

Proposition 10.4.13. *The limit trajectory as the switching time goes to zero is the singular trajectory.*

Proof. First, recall that from (10.4.14)

$$\sigma(\tau) = \frac{\text{tr}(AB) - \text{tr}(A^2)}{\text{tr}(AB) - \text{tr}(B^2)}\tau + o(\tau) \quad \text{as } \tau \rightarrow 0.$$

So, our trajectory spends τ time with control equal to 1, which corresponds to matrix A , and approximately $\frac{\text{tr}(AB) - \text{tr}(A^2)}{\text{tr}(AB) - \text{tr}(B^2)}\tau$ time with control equal to 0, corresponding to matrix B . For simplicity, let us call

$$c = \frac{\text{tr}(AB) - \text{tr}(A^2)}{\text{tr}(AB) - \text{tr}(B^2)},$$

the coefficient appearing in the expansion of $\sigma(\tau)$.

So, by convex approximation (see Theorem 8.2 in [13]), we know that as the number of switches goes to infinity (i.e., τ goes to zero), the flow tends uniformly (in the $C^\infty(\text{SL}_2(\mathbb{R}))$ -topology) to

$$T \mapsto \exp(T(vA + (1-v)B)),$$

where

$$v = \frac{1}{1+c} = \frac{\text{tr}(AB) - \text{tr}(B^2)}{2\text{tr}(AB) - \text{tr}(A^2) - \text{tr}(B^2)},$$

which corresponds to the value of the singular control (see Section 10.5). $\square$

10.5 Singular extremals

In this Section we are going to deal with the case of singular extremals (see Definition 10.3.5). Recall the definition of the Hamiltonian function H and the switching function φ

$$H(\eta, u) = \text{tr}(\eta(uA + (1-u)B)), \quad \varphi(t) = \text{tr}(\eta(t)(A - B)).$$

Recall also that the singular extremals correspond to the case of $\varphi(t) = 0$ for $t \in [\tau_1, \tau_2]$, and $\tau_1, \tau_2 > 0$. In this case, since φ is Lipschitz, we can differentiate it:

$$0 = \dot{\varphi}(t) = \text{tr}(\dot{\eta}(t)(A - B)) = \text{tr}\left([\eta(t), uA + (1 - u)B](A - B)\right) = \text{tr}(\eta(t)[A, B]),$$

where in the last equality we used the identity $\text{tr}([M_1, M_2]M_3) = \text{tr}(M_1[M_2, M_3])$. Notice that $\dot{\varphi}$ does not depend on u and in particular it is again a Lipschitz function. So, if η_* is a point of a singular trajectory, it must satisfy the two following conditions:

$$\begin{cases} \text{tr}(\eta_*(A - B)) = 0, \\ \text{tr}(\eta_*[A, B]) = 0 \end{cases}. \quad (10.5.1)$$

These two conditions are linear in η_* , so, since A, B and $[A, B]$ are linearly independent, they determine a straight line in $\mathfrak{sl}_2(\mathbb{R})$.

If we write $\eta_* = xA + yB + z[A, B]$, then the second equation implies $z = 0$. Hence, from (10.4.5), we see that if we take $\gamma_* = \text{tr}(\eta_*A)$, the line r_{γ_*} is tangent to the hyperboloid H (see discussion in Subsection 10.4.2).

There are two intersection between the line (10.5.1) and the hyperboloid H described in Section 10.2 if and only if

$$(2\text{tr}(AB) - \text{tr}(A^2) - \text{tr}(B^2))\text{tr}([A, B]^2) > 0.$$

If there are no intersection, then there are no singular extremals. So, from now on we will assume that this inequality hold.

We can compute explicitly these two intersection: let us call η_* one of the two intersection point. Then, from (10.4.2) and $\gamma = 0$ we obtain

$$\eta_* = \alpha_* A + \alpha_* \frac{\text{tr}(AB) - \text{tr}(A^2)}{\text{tr}(AB) - \text{tr}(B^2)} B,$$

where $\alpha_* \in \mathbb{R}$ is chosen such that $\text{tr}(\eta_*^2) = 1$. Since the intersection between (10.5.1) and H is a discrete set and since the solution η_* is Lipschitz, the only possibility for η_* is to be constant. Hence

$$0 = \dot{\eta}_*(t) = [\eta_*, u_* A + (1 - u_*) B],$$

from which we can determine the value of the singular control u_*

$$u_* = \frac{\text{tr}(AB) - \text{tr}(B^2)}{2\text{tr}(AB) - \text{tr}(A^2) - \text{tr}(B^2)}.$$

In order to be an admissible control, it must hold $u_* \in [0, 1]$.

Proposition 10.5.1. *The value u_* is an admissible control if and only if*

$$\text{tr}(AB) \geq \text{tr}(A^2) \quad \text{or} \quad \text{tr}(AB) \leq \text{tr}(B^2).$$

So, if we take the constant singular control equal to u_* , then the associated Lyapunov exponent is simply the greatest eigenvalue of $u_*A + (1-u_*)B$, which can be easily computed:

$$l = \frac{1}{2} \sqrt{\frac{\text{tr}([A, B]^2)}{2\text{tr}(AB) - \text{tr}(A^2) - \text{tr}(B^2)}}. \quad (10.5.2)$$

A direct computation shows that

$$\begin{aligned} \ell_{\text{sing}} > \ell_{u=1} &\iff \text{tr}\left((u_*A + (1-u_*)B)^2\right) \geq \text{tr}(A^2) \\ &\iff \text{tr}(AB) \geq \frac{\text{tr}(A^2) + \text{tr}(B^2)}{2}. \end{aligned}$$

So, taking into account that singular control is admissible if $\text{tr}(AB) \geq \text{tr}(A^2)$, we obtain that singular constant control are better than constant control if $\text{tr}(AB) \geq \text{tr}(A^2)$. On the other hand, if $\text{tr}(AB) \leq \text{tr}(B^2)$ then constant control is better than constant singular control.

The following result ends the discussion about the singular case.

Proposition 10.5.2. *There are no solutions to the adjoint system (10.2.11) containing a bang-bang piece and a singular piece.*

Proof. We already noticed that for η_* we have $\gamma = 0$ in (10.4.5) and the intersection line between the planes Π_{A,c_*} and Π_{B,c_*} is tangent to the hyperboloid H , where $c_* = \text{tr}(A\eta_*)$. Recall from Section 10.4 that as the switching time in a *bang-bang* trajectory in $\text{SL}_2(\mathbb{R})$ tends to zero, the resulting trajectory tends to a singular trajectory. The convergence in $\text{SL}_2(\mathbb{R})$ implies that also the sequence of the corresponding adjoint trajectories tends to the singular adjoint trajectory, which is constant. Thus, in a neighbourhood of η_* sufficiently small and contained in the hyperboloid H , the phase portrait is as in Figure 10.4.

Thus, any possible trajectory starting from η_* and escaping the switching curve S would inevitably cross some of the periodic trajectories tending to η_* , contradicting the local uniqueness of the solution of an ODE. $\square$

So, to resume, we found that if an extremal contains a singular piece, then the whole extremal is singular, and the corresponding Lyapunov exponent is given by (10.5.2). This is the optimal solution if $\text{tr}(AB) \geq \text{tr}(A^2)$.

10.6 Final remarks

Since the nature of the extremals changes a lot with the values of $\text{tr}(A^2), \text{tr}(B^2), \text{tr}(AB)$, we end the discussion with some final remarks about which is the optimal strategy depending on these values. As before, we always assume $\text{tr}(A^2) \geq \text{tr}(B^2)$.

So far, we have obtained that all possible extremals satisfying PMP are in one of the following form:

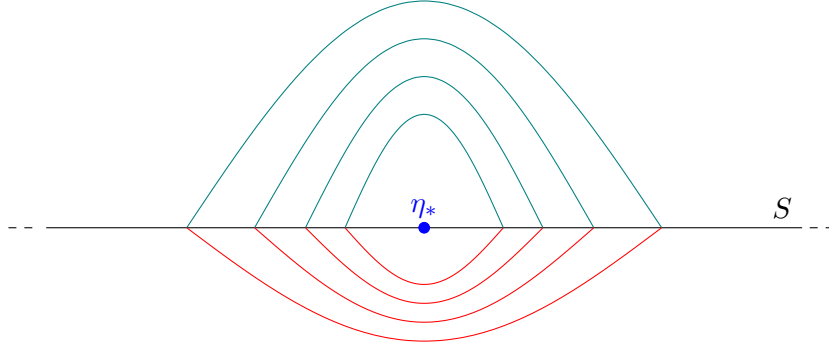


Figure 10.4: Phase portrait on the hyperboloid H for the adjoint system (10.2.11) near η_* . The black line S is the switching curve, η_* is the singular point. Above and below S we have the two *bang* pieces, and as the switching time goes to zero the adjoint trajectory tends to η_* .

- constant control;
- constant singular control;
- if $\text{tr}(A^2) < 0$, *bang-bang* periodic control.

As already pointed out, principal Lyapunov exponent associated with the periodic bang-bang extremals accumulating at singular extremals are bounded above by the principal Lyapunov exponent associated with singular extremal, so we can omit them from the discussion.

First, assume that $\text{tr}(A^2) \geq 0$. In this case we already saw in Section 10.5 that if $\text{tr}(AB) \geq \text{tr}(A^2)$, then the value of the Principal Lyapunov exponent given by the constant singular control is bigger than the one given by the constant control. Viceversa, if $\text{tr}(AB) \leq \text{tr}(A^2)$, then the optimal solution is obtained for $u = 1$ constant. So, the first two point of the Main Theorem are proved.

We turn now to the $\text{tr}(A^2) < 0$ case. We have the following restriction on the possible values of $\text{tr}(AB)$.

Lemma 10.6.1. *If $\text{tr}(B^2) \leq \text{tr}(A^2) < 0$, then $|\text{tr}(AB)| \geq \sqrt{\text{tr}(A^2)\text{tr}(B^2)}$. In particular $\text{tr}([A, B]^2) \geq 0$.*

Proof. Up to a change of base, we can suppose

$$A = \begin{pmatrix} 0 & -\lambda \\ \lambda & 0 \end{pmatrix}, \quad B = P \begin{pmatrix} 0 & -\mu \\ \mu & 0 \end{pmatrix} P^{-1}, \quad P = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}_2(\mathbb{R}), \quad \lambda, \mu \in \mathbb{R}.$$

Since A in this form commutes with rotations, that is matrix of the form

$$R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad \theta \in [0, 2\pi],$$

we can use another change of base given by a rotation matrix to kill one element of P . That is, we can further assume that P is in the form

$$P = \begin{pmatrix} \alpha & \beta \\ 0 & \frac{1}{\alpha} \end{pmatrix}, \quad \alpha, \beta \in \mathbb{R}, \quad \alpha \neq 0.$$

With A, B in this form, we can compute directly $\text{tr}(AB)$:

$$\text{tr}(AB) = -\lambda\mu \left(\frac{1}{\alpha^2} + \alpha^2 + \beta^2 \right).$$

Since $\frac{1}{\alpha^2} + \alpha^2 \geq 2$, we obtain

$$|\text{tr}(AB)| \geq 2|\lambda\mu| = \sqrt{\text{tr}(A^2)\text{tr}(B^2)}.$$

□

Remark 10.6.2. From the previous proof, we can give a geometric interpretation of the sign of $\text{tr}(AB)$ in the case $\text{tr}(B^2) \leq \text{tr}(A^2) < 0$. Define the cone of imaginary matrices in $\mathfrak{sl}_2(\mathbb{R})$:

$$\mathcal{C} = \{M \in \mathfrak{sl}_2(\mathbb{R}) \mid \text{tr}(M^2) < 0\}.$$

Using the form for A, B introduced in the previous proof, if $\text{tr}(AB) < 0$, then $\lambda\mu > 0$, that is A, B are in the same connected component. If instead $\text{tr}(AB) > 0$, then $\lambda\mu < 0$, hence A, B are in two different connected component.

In particular, if $\text{tr}(AB) < 0$, then the whole segment of velocities

$$\{uA + (1 - u)B \mid u \in [0, 1]\}$$

is inside the cone $\mathcal{C}$. In this case, if the singular control is admissible, then it has imaginary eigenvalues, hence singular constant control cannot be optimal (the resulting trajectory in $\text{SL}_2(\mathbb{R})$ would be periodic, hence bounded). So, in this case the best admissible strategy is the one given in Subsubsection 10.4.4 “Imaginary eigenvalues”, and we obtain the third point of Main Theorem.

If instead $\text{tr}(AB) > 0$, then part of the segment of velocities is outside the cone, and in particular singular control is admissible since $\text{tr}(A^2) < 0$.

The following Lemma shows that in this case constant singular control are again optimal.

Lemma 10.6.3. *If $\text{tr}(B^2) \leq \text{tr}(A^2) < 0$ and $\text{tr}(AB) \geq \sqrt{\text{tr}(A^2)\text{tr}(B^2)}$, then*

$$\frac{2}{\pi} \frac{1}{\sqrt{\frac{-2}{\text{tr}(A^2)}} + \sqrt{\frac{-2}{\text{tr}(B^2)}}} \text{arcosh} \left(\frac{\text{tr}(AB)}{\sqrt{\text{tr}(A^2)\text{tr}(B^2)}} \right) \leq \sqrt{\frac{1}{2} \frac{\text{tr}(AB)^2 - \text{tr}(A^2)\text{tr}(B^2)}{2\text{tr}(AB) - \text{tr}(A^2) - \text{tr}(B^2)}}.$$

Proof. First, notice that the two quantities are homogeneous in $\text{tr}(A^2), \text{tr}(B^2), \text{tr}(AB)$. So, left hand side can be rewritten as

$$\frac{2}{\pi} a \frac{\sqrt{b}}{1 + \sqrt{b}} \text{arcosh} \left(\frac{-c}{\sqrt{b}} \right),$$

where $a = \sqrt{\frac{\text{tr}(A^2)}{-2}}$, $b = \frac{\text{tr}(B^2)}{\text{tr}(A^2)}$, $c = \frac{\text{tr}(AB)}{\text{tr}(A^2)}$. From the hypothesis of the Lemma, we deduce $a > 0$, $b \geq 1$, $c \leq -\sqrt{b} \leq -1$. Using similar notation, the right hand side of the inequality in the thesis is

$$a \sqrt{\frac{c^2 - b}{b + 1 - 2c}}.$$

Define

$$f(b, c) = \frac{2}{\pi} \frac{\sqrt{b}}{1 + \sqrt{b}} \text{arcosh} \left(\frac{-c}{\sqrt{b}} \right),$$

$$g(b, c) = \sqrt{\frac{c^2 - b}{b + 1 - 2c}}.$$

Notice that for every b , if $c \rightarrow -\sqrt{b}$ then both functions tend to zero, i.e.:

$$\forall b \geq 1 \quad \lim_{c \rightarrow -\sqrt{b}} f(b, c) = 0 = \lim_{c \rightarrow -\sqrt{b}} g(b, c).$$

Hence, if for every $b \geq 1$ and every $c \leq -\sqrt{b}$ we have

$$\frac{\partial f}{\partial c}(b, c) \geq \frac{\partial g}{\partial c}(b, c), \tag{10.6.1}$$

then we obtain $f(b, c) \leq g(b, c)$ for every $b \geq 1$ and $c \leq -\sqrt{b}$. These derivatives are

$$\frac{\partial g}{\partial c}(b, c) = -\frac{2}{\pi} \frac{\sqrt{b}}{\sqrt{b} + 1} \frac{1}{\sqrt{c^2 - b}}$$

$$\frac{\partial g}{\partial c}(b, c) = \frac{1}{\sqrt{c^2 - b}} \frac{(c - 1)(b - c)}{(b + 1 - 2c)^{\frac{3}{2}}}.$$

So (10.6.1) reduces to

$$\frac{2}{\pi} \frac{\sqrt{b}}{\sqrt{b} + 1} \leq \frac{(1 - c)(b - c)}{(b + 1 - 2c)^{\frac{3}{2}}}.$$

Since we want this inequality to hold for every $c \leq -\sqrt{b}$ and since the right hand side is decreasing in c , it is sufficient to prove

$$\frac{2}{\pi} \frac{\sqrt{b}}{\sqrt{b} + 1} \leq \frac{(1 + \sqrt{b})(b + \sqrt{b})}{(b + 1 + 2\sqrt{b})^{\frac{3}{2}}}.$$

After a few simplifications, this reduces to

$$(\sqrt{b} + 1)^{\frac{3}{2}} \geq \frac{2}{\pi},$$

which is true for every $b \geq 1$. □

This concludes the proof of the fourth case of Main Theorem.

Appendix A

Appendix

A.1 Chronological calculus and the variation formula

In this section, we briefly recall the notation and the main formulas of chronological calculus that are used throughout the thesis. In particular, we recall the variation formula for non-autonomous flows and the Volterra expansion up to second order. For a more complete treatment, we refer to [13, Chapter 2] and [5, Chapter 6].

Let X_t be a complete non-autonomous smooth vector field on the manifold M . More generally, the same formulas remain valid under the standard measurable-in-time assumptions for which the corresponding non-autonomous differential equation is well posed. For $s, t \in \mathbb{R}$, we denote by $\Phi_{s,t}$ its flow, so that, for every $q \in M$, $\Phi_{s,t}(q)$ is the value at time t of the solution of

$$\dot{q}(\tau) = X_\tau(q(\tau)), \quad q(s) = q.$$

The flow satisfies

$$\Phi_{s,s} = \text{Id}, \quad \Phi_{r,t} = \Phi_{s,t} \circ \Phi_{r,s}, \quad \Phi_{s,t}^{-1} = \Phi_{t,s}.$$

We denote the flow from time s to time t by the right chronological exponential

$$\Phi_{s,t} = \overrightarrow{\exp} \int_s^t X_\tau d\tau.$$

When the initial time is 0, we use the abbreviated notation

$$\Phi_t := \Phi_{0,t}.$$

If $F : M \rightarrow M$ is a diffeomorphism and Y is a smooth vector field, we denote by F_*Y its push-forward, defined by

$$(F_*Y)(q) = d_{F^{-1}(q)}F[Y(F^{-1}(q))].$$

We now recall the variation formula. Let X_t and Y_t be two non-autonomous vector fields and denote by $\Phi_{0,t}$ and $\Psi_{0,t}$ the flows generated by X_t and $X_t + Y_t$, respectively. The flow of the perturbed vector field can be decomposed into the reference flow and a correction transported back by the reference dynamics.

Lemma A.1.1 (Variation formula). *With the notation above, one has*

$$\Psi_{0,t} = \Phi_{0,t} \circ \overrightarrow{\exp} \int_0^t \left(\Phi_{0,s}^{-1} \right)_* Y_s ds. \quad (\text{A.1.1})$$

We omit the proof of this Lemma, which can be found in the references cited above.

The variation formula is particularly useful when the perturbation depends smoothly on a parameter. In that case, the first and second derivatives of the perturbed flow can be read directly from the first terms of the Volterra expansion. We recall only the order needed in the subsequent applications; estimates for the remainder and the general expansion can be found in [5, Sections 6.4 and 6.6].

Proposition A.1.2 (Volterra expansion up to second order). *Let Y_t^ε be a smooth one-parameter family of non-autonomous vector fields such that*

$$Y_t^0 = 0, \quad \left. \frac{\partial Y_t^\varepsilon}{\partial \varepsilon} \right|_{\varepsilon=0} = Y_t, \quad \left. \frac{\partial^2 Y_t^\varepsilon}{\partial \varepsilon^2} \right|_{\varepsilon=0} = Z_t. \quad (\text{A.1.2})$$

Denote by

$$P_t^\varepsilon = \overrightarrow{\exp} \int_0^t Y_s^\varepsilon ds$$

the corresponding flow. Then, for every $a \in C^\infty(M)$,

$$\begin{aligned} a \circ P_t^\varepsilon &= a + \varepsilon \int_0^t Y_s a ds \\ &\quad + \frac{\varepsilon^2}{2} \left(\int_0^t Z_s a ds + 2 \int_{0 \leq s_2 \leq s_1 \leq t} Y_{s_2} (Y_{s_1} a) ds_1 ds_2 \right) + o(\varepsilon^2). \end{aligned} \quad (\text{A.1.3})$$

Proof. Vector fields act as derivations on $C^\infty(M)$. The Volterra expansion of the chronological exponential gives

$$a \circ P_t^\varepsilon = a + \int_0^t Y_{s_1}^\varepsilon a ds_1 + \int_{0 \leq s_2 \leq s_1 \leq t} Y_{s_2}^\varepsilon (Y_{s_1}^\varepsilon a) ds_1 ds_2 + O(\varepsilon^3).$$

Expanding Y_t^ε at $\varepsilon = 0$ by (A.1.2) and collecting the terms of order one and two gives (A.1.3). $\square$

In particular, after applying the variation formula, the first differential of an endpoint map is obtained from the first integral in (A.1.3), while its second differential contains both the pointwise second-order variation of the infinitesimal generator and the time-ordered double integral. This is the form of the chronological expansion used in Section 7.3.

A.2 Lipschitz maps and piecewise smooth manifolds

In this section, we recall the elementary nonsmooth tools used in the construction of piecewise smooth fields of extremals and in the study of their projections. Throughout the first part of the section, $\Omega \subset \mathbb{R}^n$ is an open set and $f : \Omega \rightarrow \mathbb{R}^m$ is a locally Lipschitz continuous map. We write $\text{co } S$ for the convex hull of a set S and $\bar{S}$ for its closure.

Definition A.2.1 (Clarke generalized differential). Let $x \in \Omega$. The *Clarke generalized differential*, or *generalized Jacobian*, of f at x is

$$\partial f(x) := \overline{\text{co} \left\{ \lim_{k \rightarrow \infty} Df(x_k) \mid x_k \rightarrow x, f \text{ is differentiable at } x_k \right\}}.$$

The following standard properties follow from Rademacher's theorem and the definition of the generalized differential.

Proposition A.2.2 (Basic properties). *Let $f : \Omega \rightarrow \mathbb{R}^m$ be locally Lipschitz and let $x \in \Omega$. Then the following statements hold.*

1. *The set $\partial f(x)$ is a non-empty compact convex subset of $\mathbb{R}^{m \times n}$.*
2. *If f is C^1 in a neighbourhood of x , then*

$$\partial f(x) = \{Df(x)\}.$$

3. *The set-valued map $x \mapsto \partial f(x)$ is locally bounded and upper semicontinuous.*

Here upper semicontinuity means that, for every open set $\mathcal{O} \subset \mathbb{R}^{m \times n}$ containing $\partial f(x)$, there exists a neighbourhood V of x such that $\partial f(y) \subset \mathcal{O}$ for every $y \in V$.

Definition A.2.3 (Clarke regular point). Assume that $m \leq n$. A point $x \in \Omega$ is called a *Clarke regular point* of f if

$$\text{rank } A = m \quad \text{for every } A \in \partial f(x).$$

If $m = n$, this is equivalent to requiring every $A \in \partial f(x)$ to be invertible.

The nonsmooth analogue of the classical inverse function theorem is the following result of Clarke [32].

Theorem A.2.4 (Clarke's local inverse function theorem). *Let $\Omega \subset \mathbb{R}^n$ be open and let $f : \Omega \rightarrow \mathbb{R}^n$ be locally Lipschitz. If x_0 is a Clarke regular point of f , then there exist open neighbourhoods O_{x_0} of x_0 and $O_{f(x_0)}$ of $f(x_0)$ such that*

$$f : O_{x_0} \longrightarrow O_{f(x_0)}$$

is a bi-Lipschitz map.

We will also use the following elementary chain rule.

Proposition A.2.5 (Chain rule). *Let $f : \Omega \rightarrow \mathbb{R}^m$ be locally Lipschitz and let $g : O \subset \mathbb{R}^m \rightarrow \mathbb{R}^k$ be a C^1 map defined on a neighbourhood of $f(x)$. Then*

$$\partial(g \circ f)(x) = Dg(f(x)) \circ \partial f(x),$$

where

$$Dg(f(x)) \circ \partial f(x) := \{Dg(f(x)) \circ A \mid A \in \partial f(x)\}.$$

A particularly simple situation, which is the one occurring at the switching hypersurfaces considered in this thesis, is that of a continuous map which is smooth on the two sides of a smooth hypersurface.

Proposition A.2.6 (Generalized differential of a two-sided smooth map). *Let $S \subset \Omega$ be a smooth hypersurface and assume that, locally near $x \in S$, the complement $\Omega \setminus S$ has two components Ω^- and Ω^+ . Suppose that f is locally Lipschitz and that the restrictions of f to Ω^- and Ω^+ admit C^1 extensions up to S . Denote their differentials at x by $D^-f(x)$ and $D^+f(x)$. Then*

$$\partial f(x) = \text{co} \{D^-f(x), D^+f(x)\}.$$

We now introduce the class of Lipschitz manifolds that will be used below.

Definition A.2.7 (Lipschitz manifold). A subset $\mathcal{M} \subset \mathbb{R}^n$ is called a d -dimensional *Lipschitz manifold* if, for every $x \in \mathcal{M}$, there exist an open neighbourhood $O_x \subset \mathbb{R}^n$ and a locally Lipschitz map

$$f : O_x \longrightarrow \mathbb{R}^{n-d}$$

such that

$$\mathcal{M} \cap O_x = f^{-1}(0),$$

and every point of $\mathcal{M} \cap O_x$ is a Clarke regular point of f . Equivalently, 0 is a Clarke regular value of f along $\mathcal{M} \cap O_x$.

In the applications considered in this thesis, the local Lipschitz structure has a more specific form. Near every nonsmooth point, the manifold is given by a bi-Lipschitz parametrization which is smooth on the two sides of a smooth hypersurface. More precisely, locally we consider maps of the form

$$\Theta : V \times (-\delta, \delta) \longrightarrow \mathcal{M}, \tag{A.2.1}$$

where $V \subset \mathbb{R}^{d-1}$ is open, Θ is bi-Lipschitz onto its image, and its restrictions to $V \times (-\delta, 0]$ and $V \times [0, \delta)$ admit smooth extensions to the corresponding closed half-neighbourhoods. This is precisely the local structure arising from the flow-box parametrizations at the switching hypersurfaces of a piecewise regular extremal.

For such parametrizations, it is useful to retain the whole family of limiting tangent spaces rather than combining them into a single tangent cone.

Definition A.2.8 (Generalized tangent family). Let $\Theta : W \subset \mathbb{R}^d \rightarrow \mathcal{M}$ be a local bi-Lipschitz parametrization, let $y \in W$, and set $x = \Theta(y)$. The *generalized tangent family* associated with Θ at x is

$$\mathcal{T}_x \mathcal{M} := \{\text{Im } L \mid L \in \partial\Theta(y)\}.$$

The generalized differential in Definition A.2.8 is computed in any smooth ambient coordinate system. Changes of smooth ambient coordinates act on its elements by composition with an invertible differential and therefore do not affect the corresponding rank and transversality conditions. In the applications below, the parametrization Θ is fixed by the geometric construction, so we suppress it from the notation. At a smooth point, the generalized tangent family reduces to the ordinary tangent space:

$$\mathcal{T}_x \mathcal{M} = \{T_x \mathcal{M}\}.$$

The two-sided structure described above yields an explicit formula for this family.

Proposition A.2.9 (Generalized tangent family at a switching hypersurface). *Let Θ be as in (A.2.1) and let*

$$x = \Theta(y_0, 0).$$

Assume that the two one-sided differentials agree on the directions tangent to V , and denote by

$$E_x := d_{y_0} \Theta(\mathbb{R}^{d-1} \times \{0\})$$

the corresponding $(d-1)$ -dimensional subspace. Let

$$v^- := \partial_s^- \Theta(y_0, 0), \quad v^+ := \partial_s^+ \Theta(y_0, 0)$$

be the two one-sided transverse derivatives. Then

$$\mathcal{T}_x \mathcal{M} = \{E_x + \mathbb{R}(\alpha v^- + (1-\alpha)v^+) \mid \alpha \in [0, 1]\}. \quad (\text{A.2.2})$$

Proof. Proposition A.2.6 gives

$$\partial\Theta(y_0, 0) = \{\alpha D^- \Theta(y_0, 0) + (1-\alpha) D^+ \Theta(y_0, 0) \mid \alpha \in [0, 1]\}.$$

Since the two differentials coincide on the directions tangent to V , only the transverse direction is affected by the convex combination. Taking the images of these maps gives (A.2.2). $\square$

We finally record the local invertibility criterion in the form used in the thesis. Since Theorem A.2.4 is local and invariant under smooth changes of coordinates, it immediately extends to maps between finite-dimensional Lipschitz manifolds.

Corollary A.2.10 (Local invertibility on a Lipschitz manifold). *Let $\mathcal{M}$ be a d -dimensional Lipschitz manifold, let N be a d -dimensional smooth manifold, and let $F : \mathcal{M} \rightarrow N$ be locally Lipschitz. Fix $x \in \mathcal{M}$ and a local bi-Lipschitz parametrization*

$$\Theta : W \subset \mathbb{R}^d \longrightarrow \mathcal{M}, \quad \Theta(y) = x.$$

If every element of the Clarke generalized differential of $F \circ \Theta$ at y is an isomorphism, then F is a local bi-Lipschitz map near x . In particular, if $\mathcal{M}$ is contained in a smooth manifold P and F is the restriction to $\mathcal{M}$ of a smooth map $\widehat{F} : P \rightarrow N$, it is sufficient that

$$d_x \widehat{F} \circ L : \mathbb{R}^d \longrightarrow T_{F(x)} N$$

be an isomorphism for every $L \in \partial\Theta(y)$.

The case needed for fields of extremals is obtained by taking $P = T^*M$, $\widehat{F} = \pi$, and $\dim \mathcal{M} = \dim M = d$. Let $\lambda \in \mathcal{M}$ and let Θ be a local parametrization as above. Then the restriction

$$\pi_{\mathcal{M}} := \pi|_{\mathcal{M}}$$

is locally bi-Lipschitz near λ provided

$$d_{\lambda} \pi \circ L : \mathbb{R}^d \longrightarrow T_{\pi(\lambda)} M$$

is an isomorphism for every $L \in \partial\Theta(y)$. Equivalently, in terms of the generalized tangent family, this condition is

$$\Lambda \cap T_{\lambda} \left(T_{\pi(\lambda)}^* M \right) = \{0\} \quad \text{for every } \Lambda \in \mathcal{T}_{\lambda} \mathcal{M}. \quad (\text{A.2.3})$$

Thus, at a smooth point (A.2.3) reduces to the usual transversality of $T_{\lambda} \mathcal{M}$ with the vertical subspace, while at a switching point it must hold for every member of the connecting family (A.2.2).

We conclude with a version of Stokes' theorem for Lipschitz manifolds.

Theorem A.2.11 (Stokes' theorem for Lipschitz manifolds). *Let $\mathcal{S}$ be a compact oriented d -dimensional Lipschitz manifold with boundary, contained in a smooth manifold P , and let ω be a smooth $(d-1)$ -form defined on a neighbourhood of $\mathcal{S}$ in P . Then*

$$\int_{\mathcal{S}} d\omega = \int_{\partial\mathcal{S}} \omega,$$

where $\partial\mathcal{S}$ is endowed with the orientation induced by $\mathcal{S}$, and the integrals are understood through oriented Lipschitz parametrizations.

Since Lipschitz manifolds are sets of finite perimeter, this formulation of the Stokes' theorem is obtained, for instance, as a particular case of the Gauss–Green theorem for BV functions. A standard reference for this result is, for instance, [35, Section 5.8].

Bibliography

- [1] A. A. Agrachev. “A necessary condition for second order optimality in the general nonlinear case”. In: *Mat. Sb. (N.S.)* 102(144).4 (1977), pp. 551–568, 632. ISSN: 0368-8666.
- [2] A. Agrachev, U. Boscain, and M. Sigalotti. “A Gauss-Bonnet-like formula on two-dimensional almost-Riemannian manifolds”. In: *Discrete Contin. Dyn. Syst.* 20.4 (2008), pp. 801–822. ISSN: 1078-0947.
- [3] A. A. Agrachev. “Feedback-invariant optimal control theory and differential geometry. II. Jacobi curves for singular extremals”. In: *J. Dynam. Control Systems* 4.4 (1998), pp. 583–604. ISSN: 1079-2724,1573-8698.
- [4] A. A. Agrachev and R. V. Gamkrelidze. “Feedback-invariant optimal control theory and differential geometry. I. Regular extremals”. In: *J. Dynam. Control Systems* 3.3 (1997), pp. 343–389. ISSN: 1079-2724,1573-8698.
- [5] Andrei Agrachev, Davide Barilari, and Ugo Boscain. *A comprehensive introduction to sub-Riemannian geometry*. Vol. 181. Cambridge Studies in Advanced Mathematics. From the Hamiltonian viewpoint, With an appendix by Igor Zelenko. Cambridge University Press, Cambridge, 2020, pp. xviii+745. ISBN: 978-1-108-47635-5.
- [6] Andrei Agrachev and Ivan Beschastnyi. “Jacobi fields in optimal control: Morse and Maslov indices”. In: *Nonlinear Analysis* 214 (2022), p. 112608. ISSN: 0362-546X.
- [7] Andrei Agrachev, Ivan Beschastnyi, and Michele Motta. *Singular extremals of optimal control problems with L^1 cost*. 2025. arXiv: 2511.21527 [math.OC].
- [8] Andrei A. Agrachev and Carolina Biolo. “Optimality of Broken Extremals”. In: *Journal of Dynamical and Control Systems* 25.2 (2018), pp. 289–307. ISSN: 1573-8698.
- [9] Andrei A. Agrachev and Carolina Biolo. “Switching in Time-Optimal Problem with Control in a Ball”. In: *SIAM Journal on Control and Optimization* 56.1 (2018), pp. 183–200. eprint: <https://doi.org/10.1137/16M110304X>.
- [10] Andrei A. Agrachev and Daniel Liberzon. “Lie-Algebraic Stability Criteria for Switched Systems”. In: *SIAM Journal on Control and Optimization* 40.1 (2001), pp. 253–269.

- [11] Andrei A. Agrachev and Michele Motta. “Lyapunov exponents of linear switched systems”. In: *Regul. Chaotic Dyn.* 30.4 (2025), pp. 481–503. ISSN: 1560-3547,1468-4845.
- [12] Andrei A. Agrachev and Michele Motta. *Second order optimality conditions for piecewise regular extremals in Optimal Control*. 2026. arXiv: 2607.10434 [math.OC].
- [13] Andrei A. Agrachev and Yuri L. Sachkov. *Control theory from the geometric viewpoint*. Vol. 87. Encyclopaedia of Mathematical Sciences. Control Theory and Optimization, II. Springer-Verlag, Berlin, 2004, pp. xiv+412. ISBN: 3-540-21019-9.
- [14] Andrei A. Agrachev, Gianna Stefani, and Pierluigi Zezza. “Strong optimality for a bang-bang trajectory”. In: *SIAM J. Control Optim.* 41.4 (2002), pp. 991–1014. ISSN: 0363-0129,1095-7138.
- [15] Andrei A. Agrachev et al. *Nonlinear and optimal control theory*. Ed. by Paolo Nistri and Gianna Stefani. Vol. 1932. Lecture Notes in Mathematics. Lectures given at the C.I.M.E. Summer School held in Cetraro, June 19–29, 2004. Springer-Verlag, Berlin; Fondazione C.I.M.E., Florence, 2008, pp. xiv+351. ISBN: 978-3-540-77644-4.
- [16] M. S. Aronna, J. F. Bonnans, and B. S. Goh. “Second order analysis of control-affine problems with scalar state constraint”. In: *Math. Program.* 160.1-2 (2016), pp. 115–147. ISSN: 0025-5610,1436-4646.
- [17] M. Soledad Aronna, Monica Motta, and Franco Rampazzo. “A higher-order maximum principle for impulsive optimal control problems”. In: *SIAM J. Control Optim.* 58.2 (2020), pp. 814–844. ISSN: 0363-0129,1095-7138.
- [18] M. Soledad Aronna et al. “Quadratic order conditions for bang-singular extremals”. In: *Numer. Algebra Control Optim.* 2.3 (2012), pp. 511–546. ISSN: 2155-3289,2155-3297.
- [19] M. Soledad Aronna et al. “Quadratic order conditions for bang-singular extremals”. In: *Numer. Algebra Control Optim.* 2.3 (2012), pp. 511–546. ISSN: 2155-3289,2155-3297.
- [20] L. Barreira and Y.B. Pesin. *Lyapunov Exponents and Smooth Ergodic Theory*. Student Mathematical Library. American Mathematical Soc., 2002.
- [21] A. Bellaïche. “The tangent space in sub-Riemannian geometry”. In: *Sub-Riemannian geometry*. Vol. 144. Progr. Math. Basel: Birkhäuser, 1996, pp. 1–78.
- [22] Bastien Berret et al. “The inactivation principle: mathematical solutions minimizing the absolute work and biological implications for the planning of arm movements”. In: *PLoS Comput. Biol.* 4.10 (2008), e1000194, 25. ISSN: 1553-734X,1553-7358.
- [23] Nicolas Boizot and Jean-Paul Gauthier. “On the motion planning of the ball with a trailer”. In: *Mathematical Control and Related Fields* 3.3 (2013), pp. 269–286. ISSN: 2156-8472.

- [24] Bernard Bonnard, Jean-Baptiste Caillaud, and Emmanuel Trélat. “Second order optimality conditions in the smooth case and applications in optimal control”. In: *ESAIM: Control, Optimisation and Calculus of Variations* 13.2 (2007), pp. 207–236.
- [25] Ugo Boscaiu. “Stability of Planar Switched Systems: The Linear Single Input Case”. In: *SIAM Journal on Control and Optimization* 41.1 (2002), pp. 89–112.
- [26] Alberto Bressan and Benedetto Piccoli. *Introduction to the mathematical theory of control*. Vol. 2. AIMS Series on Applied Mathematics. American Institute of Mathematical Sciences (AIMS), Springfield, MO, 2007, pp. xiv+312. ISBN: 978-1-60133-002-4.
- [27] Z. Chen, J.-B. Caillaud, and Y. Chitour. “ L^1 -minimization for mechanical systems”. In: *SIAM J. Control Optim.* 54.3 (2016), pp. 1245–1265. ISSN: 0363-0129,1095-7138.
- [28] Y. Chitour, M. Gaye, and P. Mason. “Geometric and asymptotic properties associated with linear switched systems”. In: *Journal of Differential Equations* 259.11 (2015), pp. 5582–5616. ISSN: 0022-0396.
- [29] Francesca Chittaro and Gianna Stefani. “Minimum-time strong optimality of a singular arc: the multi-input non involutive case”. In: *ESAIM Control Optim. Calc. Var.* 22.3 (2016), pp. 786–810. ISSN: 1292-8119,1262-3377.
- [30] Francesca C. Chittaro and Laura Poggiolini. “Singular extremals in L^1 -optimal control problems: sufficient optimality conditions”. In: *ESAIM Control Optim. Calc. Var.* 26 (2020), Paper No. 99, 33. ISSN: 1292-8119,1262-3377.
- [31] Chittaro, Francesca C. and Poggiolini, Laura. “Singular extremals in L^1 -optimal control problems: sufficient optimality conditions”. In: *ESAIM: COCV* 26 (2020), p. 99.
- [32] F. H. Clarke. “On the inverse function theorem”. In: *Pacific J. Math.* 64.1 (1976), pp. 97–102. ISSN: 0030-8730,1945-5844.
- [33] Andrei Dmitruk and Nikolai Osmolovskii. “Local minimum principle for an optimal control problem with a nonregular mixed constraint”. In: *SIAM J. Control Optim.* 60.4 (2022), pp. 1919–1941. ISSN: 0363-0129,1095-7138.
- [34] Andrei V. Dmitruk. “Jacobi type conditions for singular extremals”. In: *Control Cybernet.* 37.2 (2008), pp. 285–306. ISSN: 0324-8569,2720-4278.
- [35] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992, pp. viii+268. ISBN: 0-8493-7157-0.
- [36] J.-P. Gauthier. *Motion planning for kinematic systems*. Preprint, in honor of Bernard Bonnard. 2012.
- [37] J.-P. Gauthier, B. Berret, and F. Jean. “A biomechanical inactivation principle”. In: *Tr. Mat. Inst. Steklova* 268 (2010), pp. 100–123. ISSN: 0371-9685,3034-1809.

- [38] J.-P. Gauthier, B. Jakubczyk, and V. Zakalyukin. “Motion planning and fastly oscillating controls”. In: *SIAM J. Control Optim.* 48.5 (2009), pp. 3433–3448. ISSN: 0363-0129.
- [39] J.-P. Gauthier and V. Zakalyukin. “Entropy estimations for motion planning problems in robotics”. In: *Volume In honor of Dmitry Victorovich Anosov, Proceedings of the Steklov Mathematical Institute*. Vol. 256. 2007, pp. 62–79.
- [40] J.-P. Gauthier and V. Zakalyukin. “On the codimension one motion planning problem”. In: *J. Dyn. Control Syst.* 11.1 (2005), pp. 73–89. ISSN: 1079-2724.
- [41] J.-P. Gauthier and V. Zakalyukin. “On the motion planning problem, complexity, entropy, and nonholonomic interpolation”. In: *J. Dyn. Control Syst.* 12.3 (2006), pp. 371–404. ISSN: 1079-2724.
- [42] J.-P. Gauthier and V. Zakalyukin. “On the one-step-bracket-generating motion planning problem”. In: *J. Dyn. Control Syst.* 11.2 (2005), pp. 215–235. ISSN: 1079-2724.
- [43] J.-P. Gauthier and V. Zakalyukin. “Robot motion planning, a wild case”. In: *In Proceedings of the 2004 Suszda conference on dynamical systems, volume 250 of Proceedings of the Steklov Mathematical Institute*. 2005.
- [44] Jean-Paul Gauthier, Bronisław Jakubczyk, and Vladimir Zakalyukin. “Motion Planning and Fastly Oscillating Controls”. In: *SIAM Journal on Control and Optimization* 48.5 (Jan. 2010), pp. 3433–3448. ISSN: 0363-0129, 1095-7138.
- [45] Jean-Paul Gauthier and Matthias Kawschitz. “Minimal complexity sinusoidal controls for path planning”. In: *53rd IEEE Conference on Decision and Control*. 2014, pp. 3731–3736.
- [46] B. S. Goh. “Necessary conditions for singular extremals involving multiple control variables”. In: *SIAM J. Control* 4 (1966), pp. 716–731.
- [47] F. Jean. “Complexity of nonholonomic motion planning”. In: *Internat. J. Control* 74.8 (2001), pp. 776–782. ISSN: 0020-7179.
- [48] F. Jean. “Entropy and complexity of a path in sub-Riemannian geometry”. In: *ESAIM Control Optim. Calc. Var.* 9 (2003), 485–508 (electronic). ISSN: 1292-8119.
- [49] F. Jean and D. Prandi. “Complexity of Control-Affine Motion Planning”. In: *SIAM Journal on Control and Optimization* 53.2 (2015), pp. 816–844.
- [50] Frédéric Jean. *Control of Nonholonomic Systems: From Sub-Riemannian Geometry to Motion Planning*. SpringerBriefs in Mathematics. Cham: Springer International Publishing, 2014. ISBN: 978-3-319-08689-7.
- [51] Henry J. Kelley. “A second variation test for singular extremals”. In: *AIAA J.* 2 (1964), pp. 1380–1382. ISSN: 0001-1452.
- [52] Arthur J. Krener. “The high order maximal principle and its application to singular extremals”. In: *SIAM J. Control Optim.* 15.2 (1977), pp. 256–293. ISSN: 0363-0129.

- [53] Daniel Liberzon. *Commutation relations and stability of switched systems: a personal history*. 2023. arXiv: 2304.11155 [math.OC].
- [54] Daniel Liberzon. *Switching in systems and control*. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 2003, pp. xiv+233. ISBN: 0-8176-4297-8.
- [55] Helmut Maurer and Nikolai P. Osmolovskii. “Second order sufficient conditions for time-optimal bang-bang control”. In: *SIAM J. Control Optim.* 42.6 (2004), pp. 2239–2263. ISSN: 0363-0129,1095-7138.
- [56] Marco Mazzucchelli. *Critical point theory for Lagrangian systems*. Vol. 293. Progress in Mathematics. Birkhäuser/Springer Basel AG, Basel, 2012, pp. xii+187. ISBN: 978-3-0348-0162-1.
- [57] Dusa McDuff and Dietmar Salamon. *Introduction to symplectic topology*. Third. Oxford Graduate Texts in Mathematics. Oxford University Press, Oxford, 2017, pp. xi+623. ISBN: 978-0-19-879490-5.
- [58] Michele Motta and Dario Prandi. *Asymptotics of motion planning complexity for control-affine systems*. 2026. arXiv: 2511.17130 [math.DS].
- [59] Nikolai P. Osmolovskii and Helmut Maurer. “Second-order optimality conditions for broken extremals and bang-bang controls: theory and applications”. In: *Advances in mathematical modeling, optimization and optimal control*. Vol. 109. Springer Optim. Appl. Springer, [Cham], 2016, pp. 147–201. ISBN: 978-3-319-30784-8.
- [60] Benedetto Piccoli and Héctor J. Sussmann. “Regular Synthesis and Sufficiency Conditions for Optimality”. In: *SIAM Journal on Control and Optimization* 39.2 (2000), pp. 359–410.
- [61] L. Poggiolini and G. Stefani. “Bang-singular-bang extremals: sufficient optimality conditions”. In: *J. Dyn. Control Syst.* 17.4 (2011), pp. 469–514. ISSN: 1079-2724,1573-8698.
- [62] Laura Poggiolini and Gianna Stefani. “State-local optimality of a bang-bang trajectory: a Hamiltonian approach”. In: *Systems Control Lett.* 53.3-4 (2004), pp. 269–279. ISSN: 0167-6911,1872-7956.
- [63] Laura Poggiolini and Gianna Stefani. “Strong local optimality for a bang-bang-singular extremal: the fixed-free case”. In: *SIAM Journal on Control and Optimization* 56.3 (2018), pp. 2274–2294.
- [64] Laura Poggiolini and Gianna Stefani. “Sufficient optimality conditions for a bang-singular extremal in the minimum time problem”. In: *Control Cybernet.* 37.2 (2008), pp. 469–490. ISSN: 0324-8569,2720-4278.
- [65] Dario Prandi. “Hölder equivalence of the value function for control-affine systems”. In: *ESAIM Control Optim. Calc. Var.* 20.4 (2014), pp. 1224–1248. ISSN: 1292-8119,1262-3377.

- [66] Vladimir Yu. Protasov and Asiiat Musaeva. “Second-order linear switching systems with arbitrary control sets: stability and invariant norms”. In: *Eur. J. Control* 85 (2025), Paper No. 101346, 13. issn: 0947-3580,1435-5671.
- [67] E.S. Pyatnitskiy and L.B. Rapoport. “Criteria of asymptotic stability of differential inclusions and periodic motions of time-varying nonlinear control systems”. In: *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* 43.3 (1996), pp. 219–229.
- [68] V. Romano and P. Corallo F. Pontani. M. Teofilatto. “Minimum-Fuel Orbit Transfers Using Modified Equinoctial Elements via Indirect Heuristic Method”. In: *9th Eur. Conf. Aeronaut. Space Sci. (EUCASS)* (2022).
- [69] C. Romero-Meléndez, J. P. Gauthier, and F. Monroy-Pérez. “On complexity and motion planning for co-rank one sub-Riemannian metrics”. In: *ESAIM Control Optim. Calc. Var.* 10.4 (2004), pp. 634–655. issn: 1292-8119.
- [70] G. Stefani. “Strong optimality of singular trajectories”. In: *Geometric control and nonsmooth analysis*. Vol. 76. Ser. Adv. Math. Appl. Sci. World Sci. Publ., Hackensack, NJ, 2008, pp. 300–326. isbn: 978-981-277-606-8.
- [71] Gianna Stefani and Pierluigi Zezza. “A Hamiltonian approach to sufficiency in optimal control with minimal regularity conditions: Part I”. In: *Variational methods*. Vol. 18. Radon Ser. Comput. Appl. Math. De Gruyter, Berlin, 2017, pp. 496–515. isbn: 978-3-11-043923-6.
- [72] Marcelo Viana. *Lectures on Lyapunov exponents*. Vol. 145. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2014, pp. xiv+202. isbn: 978-1-107-08173-4.

Ringraziamenti

Il primo ringraziamento che voglio esprimere non può non andare al mio relatore, Prof. Andrei Agrachev, per avermi seguito con grande cura e per tutto quello che mi ha insegnato in questi anni di dottorato e, in precedenza, durante la mia tesi di laurea magistrale. Per me è un grandissimo onore essere un suo studente e continuerò sempre a considerarlo un esempio a cui ispirarmi.

Un secondo sentito ringraziamento va ai miei coautori Ivan Beschastnyi e Dario Prandi, dai quali ho potuto imparare moltissimo. Senza il loro aiuto, questo percorso sarebbe stato decisamente più difficile.

Vorrei anche ringraziare il mio cosupervisore, Prof. Luca Rizzi, per il suo supporto durante questi anni.

Infine, un grazie di cuore va a tutta la mia famiglia e a tutti i miei amici che mi hanno accompagnato durante questa avventura, da quelli più recenti che ho conosciuto qui a Trieste e alla SISSA, a quelli degli anni dell'università a Bologna, a quelli di più vecchia data, diventati ormai compagni inseparabili di vita. Qualche volta la distanza geografica rende più rare le occasioni per stare insieme, ma nonostante questo sento che i nostri legami sono rimasti ben saldi e profondi. Vorrei nominarvi uno per uno, ma l'elenco diventerebbe oltremodo lungo e trovare le parole giuste per ringraziarvi tutti non è facile. Sappiate solo che, anche se siamo distanti, vi porto sempre nei miei pensieri.