Singular extremals of optimal control problems with L^1 cost

Andrei Agrachev,[*] Ivan Beschastnyi,[†] Michele Motta [‡]

December 2, 2025

Abstract

We study the optimal control problem for a control-affine system, where we want to minimize the L^1 norm of the control. First, we show how Pontryagin Maximum Principle (PMP) applies to this problem and we divide the extremal trajectories into two categories: regular and singular extremals. Then, we obtain a strong generalized Legendre-Clebsch condition for singular extremals and we show that this condition together with the absence of conjugate points is sufficient to ensure local strong optimality. We provide also some geometric examples where we apply our results. Finally, we prove that generalized Legendre-Clebsch condition is necessary for optimality.

Contents

[*]SISSA, Trieste, Italy
[†]Centre Inria d'Université Côte d'Azur, Sophia-Antipolis, France
[‡]SISSA, Trieste, Italy

1 Introduction

In this paper, we study with the following optimal control problem. Let M be a smooth manifold of dimension $\dim M = n$ and take $f_0, f_1, \ldots, f_m : M \to TM$, smooth vector fields. Define also U to be the radius one closed Euclidean ball, that is $U = \{u \in \mathbb{R}^m \mid |u| \leq 1\}$. We take as set of admissible controls

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

Problem 1 (Optimal control problem (OCP)). *Consider the control-affine system:*

$$\dot{q} = f_0(q) + \sum_{j=1}^{m} u_j(t) f_j(q), \tag{1.1}$$

where $q \in M$ and $u = (u_1, \ldots, u_m) \in \mathcal{U}$ is the control acting on the system. Given any two points on the manifold $q_0, q_T \in M$, we want to find $\tilde{u} \in \mathcal{U}$ such that the corresponding solution q to (1.1) satisfies

$$q(0) = q_0, \quad q(T) = q_T,$$

and the L^1 norm of $\tilde{u}$ is minimal:

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

We will denote by $q_u(t; q_0)$ the solution of (1.1) with control u and initial point q_0 evaluated at time $t \in [0,T]$.

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 you want to minimize the fuel consumption of a vehicle, see [14],[23] 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 [11],[18], the movements in animals and human beings were studied and the authors found out that such trajectories have inactivated arcs. They found that these kind of movements minimize the "absolute work", which can be modelled using the L^1 norm of the control.

The aim of the present paper is to provide effective general tools to study this kind of problems. As a first step we apply the Pontryagin Maximum Principle (PMP) to describe the extremals of this problems. As we are going to show in Section 2, according to PMP, any optimal solution of Problem 1 is the concatenation of one of the following three types of controlled arcs:

- arcs with controls taking values 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 kind 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[8, 22, 9, 19, 20].
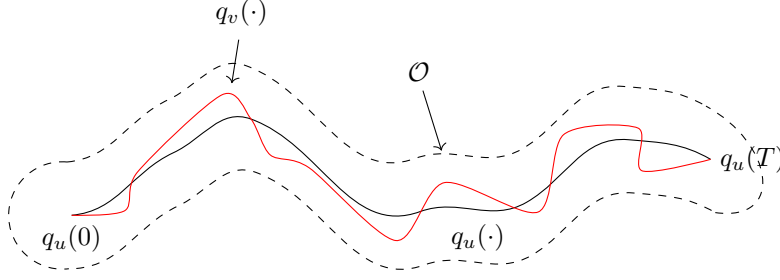
2

Figure 1: Graphical representation of Local Strong Optimality

In this paper, we focus entirely on controls which are purely singular. We prove a generalized Legendre-Clebsch conditions. 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

$$h_i(\lambda) = \langle \lambda, f_i \rangle, \quad \lambda \in T^*M, \ i = 0, 1, \ldots, m, \tag{1.2}$$

where f_i are the vector fields appearing in Equation (1.1). The angled bracket $\langle \cdot, \cdot \rangle$ stand for the duality pairing between tangent and cotangent spaces. Moreover, we denote by $\{\cdot, \cdot\}$ the Poisson bracket between functions $f, g \in C^\infty(T^*M)$ (see Sections 4.1 and 4.2 in [17] for an introduction to the subject).

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

Theorem 1.1 (Necessary optimality condition). *Let $u \in \mathcal{U}$ be a singular control which is also an optimal solution of Problem 1. Let $\lambda(\cdot)$ be its corresponding extremal solving the Hamiltonian system* (2.1) *of the PMP. Define $h_c = \sum_{i=1}^m h_i^2$. Then*

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

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

Definition 1.2 (Local strong optimality). We say that an admissible trajectory $q_u(\cdot)$ and its corresponding control u are *locally strongly optimal* if there exist a neighbourhood $\mathcal{O}$ of $q_u(\cdot)$ such that for any other admissible control $v \in \mathcal{U}$ satisfying

$$q_v(\cdot) \subset \mathcal{O}, \ q_v(0) = q_u(0), \ q_v(T) = q_u(T),$$

we have

$$J(u) \leq J(v). \tag{1.3}$$

We say that $q_u(\cdot)$ is locally strongly *strictly* optimal if the inequality in (1.3) is strict.

Notice that this definition can be equivalently restated saying that $q_u(\cdot)$ is locally strongly optimal if it is a local minimizer among other admissible trajectories, where the topology on the set of the admissible trajectories is the one induced by the uniform topology of $C^0([0, T]; M)$.

3

Theorem 1.3 (Sufficient conditions, small time). *Let $u \in \mathcal{U}$ be a singular control for Problem 1. Denote by $q(\cdot)$ and $\lambda(\cdot)$ its corresponding singular trajectory and extremal respectively. If*

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

then for every $t \in [0, T]$ there is some $\tau > 0$ such that $q_{|[t, t+\tau]}$ 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 (see [25] and references therein for other use cases of this method).

In order to complete the local optimality study, we need to quantify the time τ in Theorem 1.3. This is done via the theory of conjugate points (see Definition 4.1). They are usually defined through solutions of the Jacobi equation, which is a linear non-autonomous ODE associated to a given extremal. We give an explicit form of the Jacobi equation for singular extremals in (4.3) of Section 4. To explain the connection between conjugate points and optimality of a control function, we have to introduce the concept of corank of an extremal trajectory.

Definition 1.4 (Corank of an extremal). Let $q(\cdot)$ be an admissible trajectory for system (1.1) on M satisfying all necessary conditions prescribed by PMP (see Theorem 2.1). The corank of $q(\cdot)$ is the maximum number of linearly independent curves $\lambda(\cdot) \subset T^*M$ which are solution of (2.1) and $q(\cdot) = \pi(\lambda(\cdot))$, where π is the canonical bundle projection.

Theorem 1.5 (Necessary condition, conjugate points). *Let $\lambda(\cdot)$ be a singular extremal for Problem 1 satisfying condition (SGLC) and let $q(\cdot)$ be its corresponding trajectory on M. Suppose that $q(\cdot)$ has corank r. Moreover, assume that there are N isolated conjugate points counted with multiplicity along $q(\cdot)$, for $t \in (0, T)$. If $N \geq r$, then $q(\cdot)$ is not solution of Problem 1.*

It is worth to notice that, if the vector fields $f_0, f_1, \ldots, f_m$ are analytic, then the conjugate points are always isolated. On the other hand, in more singular cases for which condition (SGLC) is violated or the conjugate points are not isolated, it is always possible to apply the theory of Maslov index to obtain similar results, see [3, 4].

After this, one can apply the following sufficient strong local optimality condition.

Theorem 1.6 (Sufficient condition, absence of conjugate points). *Let $u \in \mathcal{U}$ be a singular control for Problem 1. Let $\lambda(\cdot)$ be its corresponding extremal and $q(\cdot)$ its corresponding trajectory. Assume furthermore that strong generalized Legendre condition (SGLC) holds and that there are no conjugate points along the trajectory $q(\cdot)$, for $t \in [0, T]$. Then $q(\cdot)$ is strictly locally strongly optimal.*

It should be stressed that Theorem 1.6 can be implemented numerically and used for concrete problems as it was done in [12] for regular extremals.

The structure of the paper is following. In Section 2 we describe the extremal trajectories of our problem. After that we proceed in studying second order conditions. To simplify the exposition we invert the order of the proofs. First, we prove sufficient small time optimality of Theorem 1.3 in Section 3. After that we introduce the Jacobi equation, prove Theorem 1.6 in Section 4 and provide examples in Section 5. The rest of the paper is dedicated to the study of necessary conditions and derivation of the Jacobi equation. The reason for this, 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, analytic in nature, and requires a careful study of the infinite-dimensional Hessian and the spaces of variations. Therefore, we postpone them to the second half of the paper. This does not create any problem for the first part, since only the correct form of generalized Legendre conditions and the Jacobi equation is needed for the proof of sufficient conditions. Finally, to make the main body

4

of text more readable, in the Appendices A and B we prove a number of propositions which are required for the proofs of the three main theorems, but which constitute long technical calculations.

2   Description of extremal trajectories

In this Section, we fix the notations used in this paper. We mainly follow the book [17]. Then, we describe extremal trajectories of the Problem 1 using PMP and separate them into two categories: regular and singular extremals.

2.1   Definitions and notations

The cotangent bundle T^*M has a canonical symplectic form σ, that allows us to define Hamiltonian vector fields on T^*M as follows. Given a function $h \in C^\infty(T^*M)$, the corresponding Hamiltonian vector field $\vec{h}$ is defined by

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

We recall that we have the following identities, that are used extensively throughout the paper:

$$\{a, b\} = \vec{a}(b) = \langle db, a \rangle = \sigma(\vec{a}, \vec{b}), \quad a, b \in C^\infty(T^*M).$$

To shorten the formulas, we introduce two useful notations. First, we use

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

to indicate nested Poisson bracket. Second, we set $f_I(q) := (f_1, \ldots, f_m)$. so that system (1.1) can be abbreviated to

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

We do the same for $h_I(\lambda) := (h_1, \ldots, h_m)$, and define $|h_I|(\lambda) := \sqrt{h_1^2(\lambda) + \cdots + h_m^2(\lambda)}$ 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.

2.2   Description of extremal trajectories

To solve the Optimal Control Problem 1, a key tool is Pontryagin Maximum Principle (PMP). We recall here a statement adapted to Problem 1 (see [7] for a more detailed discussion about PMP and some of its applications).

Theorem 2.1 (Pontryagin Maximum Principle)**.** *Let $\tilde{u} \in \mathcal{U}$ be a solution of the Optimal Control Problem 1 and denote $\tilde{q}$ the corresponding solution of* (1.1)*. Define the Hamiltonian function*

$$H(\nu, \lambda, u) = \left\langle \lambda, f_0 + \langle u, f_I \rangle \right\rangle + \nu |u|, \quad \lambda \in T^*M, u \in U, \nu \in \mathbb{R}.$$

Then there exists a non-trivial pair

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

such that the following conditions hold:

$$\dot{\lambda}(t) = \vec{H}(\lambda(t), \tilde{u}(t)), \quad \text{for a.e. } t \in [0, T], \tag{2.1}$$

$$H(\nu, \lambda(t), \tilde{u}(t)) = \max_{u \in U} H(\nu, \lambda(t), u) \quad \text{for a.e. } t \in [0, T], \tag{2.2}$$

$$\nu \leq 0.$$

We call *extremal trajectory* any solution λ of the Hamiltonian system (2.1).

By definition (1.2) and notations of the previous subsection, the Hamiltonian function $H \in C^\infty(T^*M)$ in PMP can be written as:

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

We will 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 [6, 5]).

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

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

and the corresponding Hamiltonian vector field becomes

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

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

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

which yields

$$H = h_0 + r(|h_I| - 1).$$

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 moment of time the maximum in (2.2) is attained at a unique point in U and the control u is completely 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 $|h_I(t)| \equiv 1$, which gives no direct information on r since any $r \in [0, 1]$ realizes the maximum in (2.2). We will call these controls *singular*. The Hamiltonian vector field reads

$$\vec{H} = \vec{h}_0 + r\overrightarrow{|h_I|}, \quad r(t) \in [0, 1].$$

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

$$\dot{\lambda} = \vec{h}_0(\lambda) + r(t)\vec{h}_c(\lambda), \quad r \in L^\infty([0, T]; [0, 1]).$$

6

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:

$$0 = \frac{d}{dt}\left(h_c(t)\right) = \{h_0 + rh_c, h_c\}(t) = h_{0c}(t).$$

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:

$$\frac{d^2}{dt^2}\left(h_c(t)\right) = \{h_0 + rh_c, h_{0c}\}(t) = h_{00c}(t) + rh_{c0c}(t) = 0.$$

If $h_{c0c}(\lambda) \neq 0$, we obtain the following equation for the control r:

$$r(\lambda) = -\frac{h_{00c}}{h_{c0c}}(\lambda) = \frac{h_{00c}}{h_{cc0}}(\lambda).$$

Hence, the singular control is completely determined, that is

$$u(\lambda) = r(\lambda)h_I(\lambda) \tag{2.3}$$

and the Hamiltonian system reads

$$\dot{\lambda} = \vec{h}_0(\lambda) + r(\lambda)\vec{h}_c(\lambda). \tag{2.4}$$

It is easy to check that the flow of (2.4) preserves the submanifold $\{\lambda \in T^*M \mid |h_I(\lambda)| = 1, \, h_{0c}(\lambda) = 0\}$.

In this work, we focus only 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.

3 The super-Hamiltonian function and sufficient conditions for strong local optimality of short arcs

We now describe sufficient conditions for local strong optimality (see Definition 1.2) for singular extremals of Problem 1. In the previous section we assumed that $h_{c0c}(t) \neq 0$ to derive the equations for the singular extremals. The goal of this section is to prove that $h_{c0c} > 0$ implies strong local optimality of short arcs (see Theorem 1.3 and explanations before). This condition can be derived as a necessary condition for optimality, and this will be done in Section 4. For now, we will consider it as an educated guess.

A general strategy for determining sufficient conditions is the so-called method of *field of extremals* (see, for example, [7], Chapter 17). It is mostly used when the control determined by PMP is unique. More precisely, one must have that

$$\forall t, \, \exists! u \in U \text{ s.t. } H(\lambda(t), u) = \max_{v \in U} H(\lambda(t), v). \tag{3.1}$$

This way, the maximized Hamiltonian function and the flow of the corresponding Hamiltonian vector fields are well defined on the whole T^*M. This is not true any more for singular extremals, for which the maximum in (3.1) is attained for infinite many values of $u \in U$. However, it is still possible to adapt the technique of the field of

extremals to the singular case. It was first carried out in [24], and then the technique was extended also to other cases (see, for instance [15, 21, 16]).

Now, we recall a particular version of this principle, which is well suited to our case. Let $a \in C^\infty(M)$, $H_S \in C^\infty(\mathbb{R} \times T^*M)$ be a time-dependent Hamiltonian, Φ_t^S the flow of $\vec{H}_S$ on T^*M. We define the following submanifolds:

$$\mathcal{L}_0 = \{(q, d_q a) \in T^*M \,|\, q \in M\},$$

$$\mathcal{L}_t = \Phi_t^S(\mathcal{L}_0),$$

$$\mathcal{L} = \{(t, \lambda) \,|\, t \in [0, T], \, \lambda \in \mathcal{L}_t\} \subset T^*M \times \mathbb{R}.$$

We recall the definitions of the Hamiltonian function from PMP and the relative maximized Hamiltonian

$$H(\lambda, u) = \langle \lambda, f_0(q) + \langle u, f_I(q) \rangle \rangle - |u|, \qquad H_M(\lambda) = \max_{u \in U} H(\lambda, u).$$

Theorem 3.1 (Stefani, Zezza). *Let $u(t)$ be a singular control and let $q(t)$ and $\lambda(t)$ be its corresponding solution of (1.1) and (2.1). Suppose that $H_S \in C^\infty(\mathbb{R} \times T^*M)$ is a time-dependent Hamiltonian satisfying*

1. *$H_S(t, \ell_t) \geq H_M(\ell_t)$, where $\ell_t = \Phi_t^S(\ell_0)$, Φ_t^S the flow of $\vec{H}_S$, $\ell_0 \in \mathcal{L}_0$;*

2. *$H_S(t, \lambda_t) = H_M(\lambda_t)$ for a.e. $t \in [0, T]$;*

3. *$\vec{H}_S(t, \lambda_t) = \vec{H}(\lambda_t, u(t))$ for a.e. $t \in [0, T]$;*

4. *there are open neighbourhoods $\mathcal{V} \subset T^*M$ and $\mathcal{O} \subset M$ of the singular extremal $\lambda(\cdot)$ and the trajectory $q(\cdot)$ respectively such that the function $\mathrm{id} \times \pi : \mathcal{L} \to \mathbb{R} \times M$*

$$(\mathrm{id} \times \pi)(t, \ell_0) = (t, \pi(\ell_t)),$$

   is a smooth diffeomorphism from $\big([0, T] \times \mathcal{V}\big) \cap \mathcal{L}$ onto $[0, T] \times \mathcal{O}$. Here $\pi : T^*M \to M$ denotes the standard bundle projection.

Then, q is strictly locally strongly optimal.

The proof of this statement is a straightforward adaptation of Theorem 17.1 in [7]. A more general form of this result, which applies also when Φ_t^S is not C^∞, was proven in [25].

The geometric interpretation of this method is the following. The submanifold $\mathcal{L}_0$ represents the a set of initial conditions for extremals of PMP. If there would have been a unique maximizing control for the Hamiltonian, there would have been no need for H_S. We could have used H_M to propagate $\mathcal{L}_0$ along its flow to obtain a field of extremals. In essence, if the projection $\mathrm{id} \times \pi$ restricted to $\mathbb{R} \times \mathcal{L}$ is not one-to-one, then there would be several extremal trajectories arriving to the same point. Therefore, we can not guarantee optimality. If the projection is a local diffeomorphism, then one can prove local optimality using the integral of the Poincaré-Cartan form (see [7, Chapter 17]).

In the singular case there is no nice smooth Hamiltonian in the neighbourhood of a singular trajectory that would also describe the extremal trajectories in its small neighbourhood. Nevertheless, we saw in Section 2 that one can define a smooth Hamiltonian if we restricted to the level set $|h_I| = 1$. The idea is to extend it to the Hamiltonian H_S, which under the conditions of Theorem 3.1, allows one to prove minimality.

So what we claim is that the condition $h_{c0c}(t) > 0$ is sufficient to construct such a Hamiltonian H_S. This will give the proof of Theorem 1.3 that comes next.

Proof of Theorem 1.3. The proof is separated into two parts. First, we find the Lagrangian manifold $\mathcal{L}_0$ and a time-independent Hamiltonian h_S which satisfy points 2 and 4 of Theorem (3.1). After that, we show how to obtain from h_S a time-dependent Hamiltonian H_S that satisfies all four points. We fix a singular extremal trajectory $\lambda(t)$ that satisfies the assumptions of the theorem. Since the statement is local in time, we can prove it at $t = 0$. Then the same argument applies for any $t \in [0, T]$.

We begin with construction of the Lagrangian manifold $\mathcal{L}_0$. Define

$$\Sigma = \{\ell \in T^*M \mid |h_I(\ell)| = 1\}$$
$$\mathcal{S} = \{\ell \in T^*M \mid h_{0c}(\ell) = 0\}.$$

First, we find $a \in C^\infty(M)$ such that $\lambda(0) \in \mathcal{L}_0$ and $\mathcal{L}_0 \subseteq \Sigma$. Notice that this is equivalent to the equation

$$d_{q(0)}a = \lambda(0), \qquad \sum_{i=1}^{m} |\langle d_q a, f_i(q)\rangle|^2 = 1.$$

We can solve this equation locally using the method of the characteristics. We have to find a non-characteristic hypersurface $N \subset M$ and take $a_{|N} = 0$. We can choose in particular any N such that $T_{q_0}N \subset \ker \lambda_0$, so that N is transversal to the distribution $\mathrm{span}\{f_1, \ldots f_m\}$ at q_0 since $|\langle \lambda_0, f_I(q_0)\rangle|^2 = 1$ (and by continuity is transversal also in a neighborhood of q_0), so our equation satisfy the non-characteristic assumption. This will ensure that point 4 of Theorem 3.1 holds for sufficiently small time if Σ is preserved by the super-Hamiltonian, since

$$d_{(0,\lambda(0))}\Psi = \begin{pmatrix} 1 & * \\ 0 & d_{\lambda(0)}\pi \end{pmatrix}.$$

Now, we want to find an Hamiltonian function h_S such that point 2 and 4 of Theorem 3.1 holds. Our reference control is

$$u(t) = \frac{h_{00c}(t)}{h_{c0c}(t)}h_I(t),$$

and h_S must satisfy along the trajectory $\lambda(t)$

$$h_S(t) = H(\lambda(t), u(t)) = h_0(t),$$

which is satisfied if, for instance, we set

$$h_S(\ell) = h_0(\ell) \quad \text{for } \ell \in \mathcal{S} \cap \Sigma.$$

In order to have $\vec{h}_S$ tangent to Σ (which implies $\mathcal{L}_t \subset \Sigma$ for $t > 0$) we can solve

$$\langle d_\ell h_c, \vec{h}_S\rangle = 0, \quad \ell \in T^*M \tag{3.2}$$
$$h_S(\ell) = h_0(\ell), \qquad \ell \in \mathcal{S}.$$

We have from (3.2)

$$\langle d_\ell h_S, \vec{h}_c\rangle = \{h_S, h_c\}(\ell) = -\langle d_\ell h_c, \vec{h}_S\rangle = 0.$$

This implies that h_S is constant along the flow lines of $\vec{h}_c$. In addition, since $\vec{h}_c(|h_I|) = \{h_c, |h_I|\} = 0$, the flow $\vec{h}_c$ preserves Σ. Also, since $\vec{h}_c(h_{0c})(\ell) = h_{c0c}(\ell) > 0$ for all $\ell \in \mathcal{S}$ sufficiently close to $\lambda(0)$, the flow of $\vec{h}_c$ is locally transversal to $\mathcal{S}$. So, for $\ell \in T^*M$ close enough to $\mathcal{S}$, we have that $\ell = \ell_t := \exp(t\vec{h}_c)(\ell_0)$, for a unique couple $(\ell_0, t) \in \mathcal{S} \times \mathbb{R}$, and the solution of Equation (3.2) is
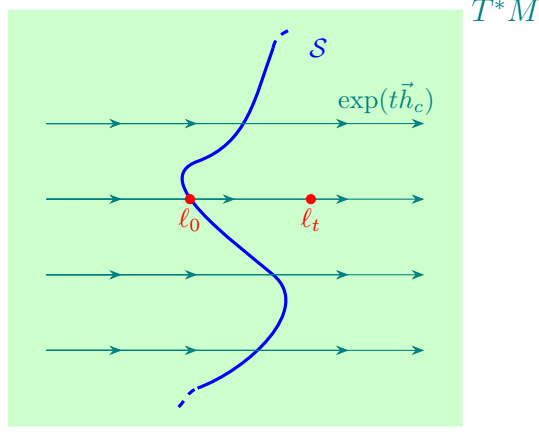
$$h_S(\ell_t) = h_0(\ell_0),$$

Figure 2: Construction of h_S.

as depicted in Figure 2. This choice satisfies point 2 of Theorem 3.1.

Now we start constructing the Hamiltonian H_S and focus on point 3 and then on point 1. First, we show that

$$d_\ell(h_S - h_0) = 0, \tag{3.3}$$

for $\ell \in \mathcal{S} \cap \Sigma$. To prove the claim, 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 $\mathcal{S}$.

Next notice that in all of the previous discussion we only used that $\{h_S, h_c\} = 0$. It means that we can freely add to h_S any function that Poisson commutes with h_c and is zero on Σ. We define the final H_S as the time-dependent Hamiltonian

$$H_S(t, \ell) = h_S(\ell) + r(t)(|h_I(\ell)| - 1),$$

where $r(t) = \frac{h_{00c}(t)}{h_{cc0}(t)}$ and the hamiltonians are evaluated along the reference extremal $\lambda(t)$. From (3.3) we obtain point 3, i.e. $\vec{H}_S(t, \lambda(t)) = \vec{H}(\lambda(t), u(t))$.

It only remains to prove that H_S satisfies point 1. For $\ell \in \Sigma$, we have

$$H_S(t, \ell) = h_S(\ell),$$
$$H_M(\ell) = h_0(\ell).$$

Since by construction $\mathcal{L}_t$ is contained in Σ, it is enough to prove that the difference $h_S - h_0$ is positive in a small neighbourhood of $\lambda(0) \in \Sigma \cap \mathcal{S}$ inside Σ.

10

We have $(h_S - h_0)|_{\Sigma \cap S} = 0$ by construction. Therefore, it is enough to study the local behaviour of $h_S - h_0$ along directions inside Σ transversal to $\Sigma \cap S$. Previously we have seen that $\vec{h}_c$ is transversal to $\Sigma \cap S$ and preserves Σ. Hence, given $l \in \Sigma \cap S$, it is enough to check that $\tau = 0$ is a local minimum of the function

$$h_S(\exp(\tau \vec{h}_c)(\ell)) - h_0(\exp(\tau \vec{h}_c)(\ell)).$$

Using the definition of h_S and expanding into Taylor polynomial the second summand we find

$$h_S(\exp(\tau \vec{h}_c)(\ell)) - h_0(\exp(\tau \vec{h}_c)(\ell)) = h_0(\ell) - \left(h_0(\ell) + \tau h_{c0}(\ell) + \frac{\tau^2}{2} h_{cc0}(\ell) \right) + o(\tau^2) = \frac{\tau^2}{2} h_{c0c}(\ell) + o(\tau^2).$$

Since $h_{c0c}(\lambda_0) > 0$ for ℓ in a suitable compact neighbourhood of λ_0 contained in Σ, we have the desired inequality of point 1 of the Theorem. $\qquad \square$

4   Jacobi equation and sufficient conditions for optimality

In the last 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 arc can be. To this aim, we introduce a suitable Jacobi equation and the notion of conjugate points. These concepts are classical in Riemannian geometry and they have been successfully generalized to some instances in optimal control theory, see [7, Chapter 21].

4.1   Linear symplectic geometry and conjugate points

In the classical calculus of variations the Jacobi equation is essentially the linearization of the Hamiltonian equation along a given extremal $\lambda(t)$, and it is used for studying the behaviour of extremals in a neighbourhood of $\lambda(t)$. The Morse index theorem states that under the strong Legendre condition, the number of negative eigenvalues of the Hessian of the second variation can be computed by counting solutions to a certain boundary value problem involving the Jacobi equation. In the singular case, that we study, this is also true. However, definition of the Jacobi equation is more involved.

As before we fix a singular extremal $\lambda(t)$ with $\lambda(0) = \lambda_0$ that satisfies the strong generalized Legendre condition (SGLC). For this section we will shorten the notation value of the symplectic form at λ_0 as $\sigma := \sigma_{\lambda_0}$. To define conjugate points we need to introduce some standard tools from symplectic geometry. For more details about this topic, we refer to [13].

Let V be a $2n$-dimensional vector space and σ a symplectic form on V. We say that a subspace $W \subset V$ is *symplectic* if $\sigma|_W$ is non-degenerate. Contrary to that, we say that a subspace $\Lambda \subset V$ is *isotropic* if $\sigma_{|\Lambda} = 0$. A linear subspace Λ is called *Lagrangian* if it is isotropic and maximal, which is equivalent to $\dim \Lambda = n$. Alternatively, we can define the *skew-orthogonal compliment* of Λ as

$$\Lambda^{\angle} = \{\lambda \in V : \sigma(\lambda, \mu) = 0, \forall \mu \in \Lambda\}.$$

A subspace Λ is isotropic if and only if $\Lambda \subset \Lambda^{\angle}$ and Lagrangian if and only if $\Lambda^{\angle} = \Lambda$. Given a Lagrangian subspace Λ and a isotropic subspace Γ we can construct a new Lagrangian subspace such that it contains Γ and has a maximal intersection with Λ:

$$\Lambda^{\Gamma} := (\Lambda + \Gamma) \cap \Gamma^{\angle}. \tag{4.1}$$

For the proofs of sufficient conditions we will need to understand the structure of the set of all Lagrangian subspaces of V, denoted by $L(V)$. It is an $(n+1)n/2$-dimensional compact submanifold of the Grassmannian,

called the Lagrangian Grassmannian. A convenient choice of local coordinate on $L(V)$ is constructed as follows. Take $\Pi \in L(V)$ and fix any $\Delta \in L(V)$ such that $\Delta \oplus \Pi = V$. Then, there is a one-to-one correspondence between Lagrangian subspaces transversal to Δ, which we will denote by $\Delta^{\pitchfork}$, and the set of symmetric matrices of dimension n, which we will denote by $\mathrm{Sym}(n)$. More precisely, if we choose Darboux coordinates for (V, σ) such that

$$\Pi = \{(p, 0)\}, \quad \Delta = \{(0, q)\}, \quad p \in \mathbb{R}^{n*}, q \in \mathbb{R}^n$$

then, for any Lagrangian subspace Λ transversal to Δ there is a unique $S \in \mathrm{Sym}(n)$ such that

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

Now we can define the Jacobi equation and the conjugate points. Let $Z_t : T_{u(t)}U \to T_{\lambda_0}(T^*M)$ be the time-dependent map defined by

$$Z_t v(t) := \left\langle v(t), \left[(\Phi_t^{-1})_* \vec{h}_I\right](\lambda_0) \right\rangle, \tag{4.2}$$

where Φ_t is the flow on T^*M of the Hamiltonian system (2.4). We can split $T_{u(t)}U = (\mathbb{R}h_I(t)) \oplus (h_I(t))^\perp$ by defining $v(t) = \phi(t)h_I(t) + w(t)$. To shorten the notation we will abbreviate $Z_I(t) := Z_t h_I(t)$. Then, using the previously defined splitting, we define

$$\mathcal{Z}_t v(t) := Z_t w(t) - \dot{Z}_I(t)\phi(t).$$

In addition, we define the Legendre form on $T_{u(t)}U \oplus \mathbb{R}$ as

$$l_t(w, \phi) := \begin{pmatrix} w & \phi \end{pmatrix} \begin{pmatrix} \frac{\mathrm{Id}}{r} & \sigma\big(Z_t \cdot, Z_I(t)\big) \\ \sigma\big(Z_t \cdot, Z_I(t)\big)^T & \sigma\big(Z_I(t), \dot{Z}_I(t)\big) \end{pmatrix} \begin{pmatrix} w \\ \phi \end{pmatrix},$$

where $w \in (h_I(t))^\perp$.

Definition 4.1. Let $\lambda(t)$ be an extremal satisfying the strong generalized Legendre condition and denote by $\Pi_0 := T_{\lambda_0}(T_{q_0}^* M)$. A moment of time T is conjugate to 0 on λ if and only if there is a non-constant solution to the following boundary value problem

$$\dot{\eta} = -\mathcal{Z}_t l_t^{-1} \sigma\big(\mathcal{Z}_t \cdot, \eta\big) \quad \text{for a.e. } t \in [0, T], \tag{4.3}$$

$$\eta(0) \in \Big(\Pi_0 + \mathbb{R}Z_I(0)\Big) \cap \big(\mathbb{R}Z_I(0)\big)^{\angle} = \Pi_0^{\mathbb{R}Z_I(0)}, \quad \eta(T) \in \Pi_0. \tag{4.4}$$

Equation (4.3) is called the *Jacobi equation* along the extremal λ. Moreover, if T is conjugate to 0 on λ, 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 (4.3),(4.4).

We have the following property that will be used several times.

Lemma 4.2. $\eta_I(t) = Z_I(t)$ is a solution of (4.3).

Proof. If we take the variation $w(t) = 0$ and $\phi(t) = 1$, by definition of l_t we have $l_t(0, 1) = \sigma(Z_t \cdot w, Z_I(t)) - \sigma(\dot{Z}_I(t), Z_I(t)) = \sigma(\mathcal{Z}_t \cdot, Z_I(t))$. So

$$-\mathcal{Z}_t l_t^{-1} \sigma\Big(\mathcal{Z}_t \cdot, Z_I(t)\Big) = -\mathcal{Z}_t(0, 1) = \dot{Z}_I(t).$$

$\square$

12

The Jacobi equation (4.3) is obtained by pulling back all of the information along an extremal. An explicit example involving Jacobi equation written in this form is given in Subsection 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 $T_{\lambda_0}(T^*M)$ with a time-dependent quadratic Hamiltonian

$$b_t := \frac{\left\langle \sigma_{\lambda_0}(\mathcal{Z}_t \cdot, \eta), l_t^{-1} \sigma_{\lambda_0}(\mathcal{Z}_t \cdot, \eta) \right\rangle}{2}. \tag{4.5}$$

It should be noted that invertibility of l_t is guaranteed by the strong generalized Legendre condition (see Lemma 4.4, that we prove after we introduce the second variation). So it is enough to rewrite this Hamiltonian in terms introduced in the previous sections.

Let Φ_t the flow of the hamiltonian vector field $\vec{H}(u(t), \cdot)$ and Φ_t^S the flow of $\vec{H}_S(t, \cdot)$. Consider the composition

$$\Phi_{-t} \circ \Phi_t^S.$$

This is diffeomorphism of the cotangent bundle which fixes the point λ_0. We want to write it as the flow of another Hamiltonian function, i.e.

$$\Phi_{-t} \circ \Phi_t^S = \overrightarrow{\exp} \int_0^t \vec{\varphi}_s \, ds. \tag{4.6}$$

It follows by the variation formula (see [7]) that this is

$$\varphi_t(\ell) = \big(H_S(t, \cdot) - H(u(t), \cdot) \big) \circ \Phi_t(\ell), \quad \ell \in T^*M, t \in [0, T].$$

Indeed, we have

$$\vec{\varphi}_t = \overrightarrow{\big(H_S(t, \cdot) - H(u(t), \cdot) \big) \circ \Phi_t} = (\Phi_{-t})_* \left(\vec{H}_S(t, \cdot) - \vec{H}(u(t), \cdot) \right),$$

hence, by the twisted variation formula (see Equation (2.28) in [7]), we obtain

$$\overrightarrow{\exp} \int_0^t \vec{\varphi}_s \, ds = \overrightarrow{\exp} \int_0^t (\Phi_{-s})_* \vec{H}_S(s, \cdot) - (\Phi_{-s})_* \vec{H}(u(s), \cdot) \, ds =$$

$$= \overrightarrow{\exp} \int_0^t - (\Phi_{-s})_* \vec{H}(u(s), \cdot) \, ds \circ \overrightarrow{\exp} \int_0^t (\Phi_{-s})_* (\Phi_s)_* \vec{H}_S(s, \cdot) \, ds =$$

$$= \left(\overrightarrow{\exp} \int_0^t \vec{H}(u(s), \cdot) \, ds \right)^{-1} \circ \overrightarrow{\exp} \int_0^t \vec{H}_S(s, \cdot) \, ds.$$

We have the following relation with the Jacobi equation.

Proposition 4.3. *Under the previous notations, we have that for $\lambda_0 \in \Sigma$*

$$\operatorname{Hess}_{\lambda_0} \left[\big(H_S(t, \cdot) - H(u(t), \cdot) \big) \circ \Phi_t \right] = \left\langle \sigma_{\lambda_0}(\mathcal{Z}_t \cdot, \eta), l_t^{-1} \sigma_{\lambda_0}(\mathcal{Z}_t \cdot, \eta) \right\rangle \quad \eta \in T_{\lambda_0}(T^*M). \tag{4.7}$$

The proof of this proposition is given in Appendix B.

Finally, we note that there is an important connection between the Legendre condition in Theorem 1.3 and the Legendre form l_t.

Proposition 4.4. *Given a singular extremal $\lambda(t)$, 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 lengthy computation that we postpone to Subsection 6.3. An important corollary of this proposition is that the strong Legendre condition ensures that the Hamiltonian b_t in (4.5) is non-negative. We will use this fact in the proof of Theorem 1.6 presented in the next subsection.

4.2   Sufficient condition for optimality

This Subsection is entirely devoted to the proof of Theorem 1.6. The argument follows closely the idea of the proof of Theorem 21.3 of [7].

We recall the following Lemma, which is proved in [7]. It will play a crucial role in the proof of Theorem 1.6.

Lemma 4.5. *Let V be a symplectic space and let h_t be a time dependent quadratic Hamiltonian function on V. Define $(\phi_t)_{t \in \mathbb{R}}$ the flow of h_t. For a given $\Lambda_0 \in L(V)$, let $\Lambda_t = \phi_t(\Lambda_0)$. Suppose moreover that $\Lambda_t \cap \Delta = \{0\}$ for every $t \in [0, T]$ and let $\Lambda_t = \{(p, S_t p) \mid p \in \Pi\}$, $S_t \in \mathrm{Sym}(n)$. Then*

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

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

We recall the statement of the Theorem that we are going to prove.

Theorem 4.6. *Let $\lambda(t)$, $t \in [0, T]$, be a singular extremal for Problem 1. Assume furthermore that it satisfies the strong generalized Legendre condition* (SGLC) *and that the Jacobi equation* (4.3),(4.4) *has no non-trivial solution. Then $\lambda(t)$ is strictly locally strongly optimal.*

Proof. We want to apply Theorem 3.1. We already saw in Section 3 how to define the overmaximized-Hamiltonian H_S. So, to prove the Theorem, we have to find a suitable function $a \in C^\infty(M)$ such that, denoting $\mathcal{L}_0 = \{(q, d_q a) \in T^*M\}$ the graph of da and $\mathcal{L}_t := \Phi_t^S(\mathcal{L}_0)$, the canonical projection

$$\pi_t : \mathcal{L}_t \to M$$

is a smooth local diffeomorphism from a neighbourhood of λ_t onto its image for any $t \in [0, T]$ and $T_{\lambda_0}\mathcal{L}_0 \subset T_{\lambda_0}\Sigma$. This is equivalent to require that the tangent space $T_{\lambda_t}\mathcal{L}_t$ has zero intersection with the vertical space Π_0:

$$T_{\lambda_t}\mathcal{L}_t \cap T_{\lambda_t}\left(T_{q(t)}^*M\right) = \{0\}, \quad t \in [0, T]. \tag{4.8}$$

Define $T_{\lambda_0}\mathcal{L}_0 = L_0$. Recall that $T_{\lambda_t}\mathcal{L}_t = (\Phi_t^S)_* L_0$. Also, since Φ_t is the lift of a flow on M, we have $T_{\lambda_t}\left(T_{q(t)}^*M\right) = (\Phi_t)_* T_{\lambda_0}\left(T_{q_0}^*M\right)$. Moreover, we know that $\Phi_t^S = \Phi_t \circ \overrightarrow{\exp} \int_0^t \vec{\varphi}_s ds$, where φ_t was defined in Equation (4.6). So, putting everything together, from (4.8) we obtain

$$(\Phi_t^S)_* L_0 \cap (\Phi_t)_* \Pi_0 = \{0\}, \quad t \in [0, T],$$

$$(\Phi_t)_* B_t L_0 \cap (\Phi_t)_* \Pi_0 = \{0\}, \quad t \in [0, T],$$

$$B_t L_0 \cap \Pi_0 = \{0\}, \quad t \in [0, T],$$

14

where

$$B_t = \left(\overrightarrow{\exp} \int_0^t \vec{\varphi}_s ds \right)_* .$$

Recall also that $d_{\lambda_0}^2 \varphi_t$ is the Hamiltonian function b_t of Jacobi equation. So, we have

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

Now, we know, by assumption, that there are no conjugate point along the trajectory λ_t, that is, there are no non-trivial solution

$$B_t \Lambda_0 \cap \Pi_0 = C_t, \quad t \in (0, T],$$

where C_t is the space of constant vertical solutions and Λ_0 is the linear subspace of initial condition for Jacobi equation (4.4).

We first consider the case of $C_t = \{0\}$. Then, we will see how we can reduce the general case to this simpler one. If $C_t = \{0\}$, the idea is the following: once we fix ε small enough, in every neighbourhood of Π_0 it is possible to find a suitable Lagrange subspace L_0 for which there are no conjugate points for $t < \varepsilon$:

$$B_t L_0 \cap \Pi_0 = \{0\} \quad \text{for } t \in [0, \varepsilon]. \tag{4.9}$$

Notice that, for every $\varepsilon > 0$ and any neighbourhood $\mathcal{V}$ of Λ_0, there are Lagrangian subspaces $L \in \mathcal{V}$ such that $B_t L \cap \Pi_0 \neq \{0\}$ with $t < \varepsilon$ (see Figure 3), so L_0 must be chosen carefully. Once we have chosen L_0 in a sufficiently small neighbourhood of Λ_0 and satisfying (4.9), the continuity of the flow B_t ensures that there are no conjugate points also for $t \in [\varepsilon, T]$. In order to find a suitable L_0 tangent to Σ, we can use the local coordinates introduced

in the previous subsection. Since we want L_0 to be tangent to Σ, first of all we can restrict to a symplectic subspace of $T_{\lambda_0}(T^*M)$ of dimension $2n - 2$. More precisely, since $Z_I(t) \notin \Pi_0$ for every $t \in [0, T]$, there is some $\zeta(t) \in \Pi_0$ such that $\sigma(\zeta(t), Z_I(t)) \neq 0$ and then we can fix any subspace $\Gamma(t) \subset T_{\lambda_0}\Sigma$ symplectic such that $\Gamma(t) \oplus \mathbb{R}Z_I(t) \oplus \mathbb{R}\zeta(t) = T_{\lambda_0}(T^*M)$. Moreover, since B_t is symplectic on $T_{\lambda_0}(T^*M)$, B_t restricted to $\Gamma(t)$ is also symplectic.

Now, we must have $\dim(\Gamma(t) \cap \Pi_0) = n - 1$. In particular $\Gamma(t) \cap \Pi_0 = \Gamma(t) \cap \Lambda_0 =: \bar{\Pi}_0(t)$ is a Lagrangian subspace in $\Gamma(t)$. So, if we find some $\bar{L}_0$ Lagrangian in $\Gamma(0)$ and satisfying

$$B_t \bar{L}_0 \cap \bar{\Pi}_0(t) = \{0\} \quad \text{for } t \in [0, \varepsilon],$$

then $L_0 := \bar{L}_0 \oplus \mathbb{R}Z_I(0)$ satisfies (4.9). Moreover, L_0 is a Lagrangian subspace since $\Gamma(0) \subset Z_I(0)^{\angle}$ implies

$$\sigma(Z_I(0), v) = 0, \quad \forall v \in \bar{L}_0 \subset \Gamma(0).$$

So, from now on, our ambient space will be this subspace $\Gamma(t)$. Notice that there are always non-trivial choices of $\Gamma(t)$ if $\dim M = n \geq 2$.

Now, we can choose any Lagrangian subspace $H(0)$ such that $\bar{\Pi}_0 \cap H(0) = \{0\}$ and, denoting by $H(t) = B_t H(0)$, consider the local chart $H(t)^{\pitchfork}$ in which $\bar{\Pi}_0(t)$ corresponds to the null symmetric matrix.

Recall that as a corollary of Proposition 4.4 we have $b_t \geq 0$. This means that, by Lemma 4.5, if we choose any initial Lagrangian subspace Ξ_0 corresponding to a matrix $S_0 > 0$ and we let it evolve by the flow of B_t, then, denoting by S_t the matrix corresponding to $\Xi_t := B_t \Xi_0$, we have

$$\dot{S}_t \geq 0,$$

15
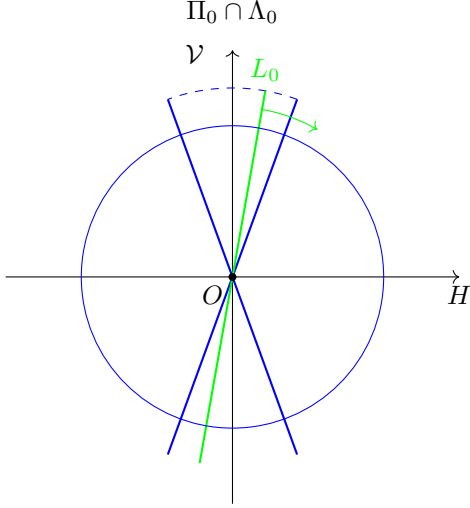
(a) Good choice of L_0, corresponding to $S_0 > 0$
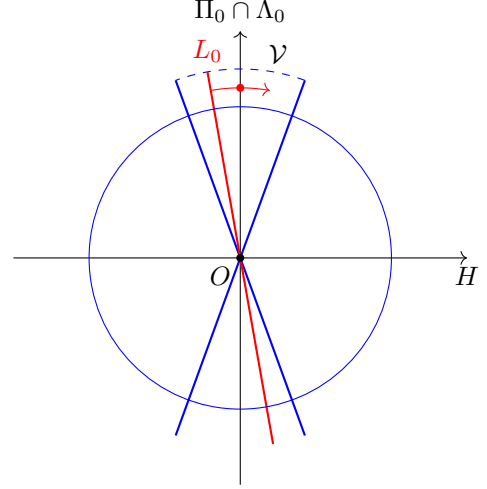
(b) Bad choice of L_0, corresponding to $S_0 < 0$

Figure 3: Graphical explanation on how the initial subspace L_0 should be chosen, for $n = 1$ (so, $L(T_{\lambda_0}(T^*M)) \simeq S^1$). Supposing that a curve with positive velocity moves clockwise, you see that if we choose L_0 in the right-half of $\mathcal{V}$ (case (a)), then $B_t L_0$ does not cross the vertical space Π_0 for small times. If instead we choose L_0 in the left-half of $\mathcal{V}$, then there can be an intersection with Π_0 in arbitrarily small time.

and in particular $S_t > 0$. Therefore, L_t stays in the chart $H(0)^{\pitchfork}$ and we have $S_t > 0$ for $t \in [0, \varepsilon]$ for ε sufficiently small, and this implies that S_t do not cross the vertical subspace $\bar{\Pi}_0$, since it corresponds to the zero matrix.

Now, if in addition we choose $S_0 > 0$ close enough to zero, then by continuity of the flow B_t there will be no conjugate points also for $t \in [\varepsilon, T]$. Indeed, suppose, by contradiction that there is a sequence of matrices $S^{(k)} > 0$, $\lim_{k \to \infty} S^{(k)} = 0$, for which the corresponding curve of Lagrange subspaces $(\Xi_t^{(k)})_{t \in [0,T]}$ has a non-trivial intersection with the vertical space at time $t_k \in [0, T]$. Then, since $[0, T]$ is compact, there is a limit point t_∞ of the sequence $(t_k)_{k \in \mathbb{N}}$. Moreover, by continuity the flow B_t is continuous, also the curve $B_t \Lambda_0$ has a non-trivial intersection with the vertical space at time t_∞.

So, we have proven that

Proposition 4.7. *If $C_t = \{0\}$, then, for $\varepsilon > 0$ small enough, there is a sufficiently small neighbourhood $\bar{\mathcal{V}}$ of $\bar{\Pi}_0(0)$ and a Lagrangian subspace $\bar{L}_0 \in \bar{\mathcal{V}}$ such that $B_t \bar{L}_0 \cap \Pi_0 = \{0\}$ for $t \in [0, T]$.*

We have to consider now the case $C_t \neq \{0\}$. It is possible to reduce it to the previous case $C_t = \{0\}$ passing to the quotient $(C_t)^{\angle}/C_t$. Indeed, the family of subspaces C_t is monotone non-increasing:

$$C_{t_2} \subset C_{t_1} \text{ if } t_1 \leq t_2.$$

We set $C_0 = \bar{\Pi}_0$. The family C_t is continuous from the left (see, for instance, [3]). Denote its points of discontinuity as

$$0 \leq s_1 < s_2 < \ldots s_k \leq T.$$

Hence, the family C_t is constant on the segments $(s_i, s_{i+1}]$. We can then decompose each C_t as

$$C_t = E_{i+1} \oplus E_{i+1} \oplus \cdots \oplus E_k, \quad t \in (s_i, s_{i+1}].$$

16

so that $C_t = E_1 \oplus \cdots \oplus E_k$ for $t \in (0, s_1]$, $C_t = E_2 \oplus \cdots \oplus E_k$ for $t \in (s_1, s_2]$, and so on. We can construct the corresponding dual splitting in Γ

$$H = F_1 \oplus \cdots \oplus F_k, \quad \sigma_{\lambda_0}(E_i, F_j) = 0, \text{ if } i \neq j.$$

We have the following Lemma [7, Lemma 21.3.].

Lemma 4.8. *Fix H_0 any Lagrangian subspace in Γ such that $\bar{\Pi}_0 \cap H_0 = \{0\}$. Then, there are $H_1, \ldots, H_k$ Lagrangian subspaces, $H_i \cap \bar{\Pi}_0 = 0$, and $\varepsilon_1 \geq \cdots \geq \varepsilon_k > 0$ such that, for any Lagrangian subspace $\bar{L}_0$, $\bar{L}_0 \cap H_0 = \{0\}$, corresponding to the matrix $\varepsilon \mathrm{Id}$ in the $(\bar{\Pi}_0, H_0)$ local chart, we have*

1. $B_t \bar{L}_0 \cap \bar{\Pi}_0 = \{0\}$, for $t \in [0, s_i]$;

2. $B_t \bar{L}_0 \cap H_i = \{0\}$, for $t \in [0, s_i]$ and $S_0 > 0$ in the $(\bar{\Pi}_0, H_i)$ local chart.

From this Lemma, for $i = k$, we obtain that $B_t \bar{L}_0 \cap \bar{\Pi}_0 = \{0\}$, for $t \in [0, T]$, provided that $\bar{L}_0$ corresponds to the matrix $S_0 = \varepsilon \mathrm{Id}$, for $\varepsilon < \varepsilon_k$, in the $(\bar{\Pi}_0, H_0)$ local chart. Thus, if we take any of these $\bar{L}_0$, then $L_0 := \mathbb{R} Z_I(t) \oplus \bar{L}_0$ is a Lagrangian subspace in $T_{\lambda_0}(T^*M)$ satisfying (4.9), which was exactly what we wanted to prove to complete the proof of the Theorem. $\qquad \square$

Notice that, in the proof of Lemma 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 $H^\pitchfork$, for some Lagrangian subspace H, and this is always possible by means of small perturbation of the H_i.

5 Examples

5.1 An example: the Heisenberg group with a drift

We are going now to show an example of Problem 1 modelled on the Heisenberg group. We will show that there are some nontrivial singular extremals, but, thanks to the necessary condition of Theorem 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

$$X = \partial_x - \frac{y}{2} \partial_z,$$
$$Y = \partial_y + \frac{x}{2} \partial_z,$$
$$Z = \partial_z.$$

The three vector fields satisfy
$$[X, Y] = Z, \qquad [X, Z] = 0, \qquad [Y, Z] = 0.$$
We consider the following control system

$$\dot{q} = \alpha X(q) + \beta Y(q) + \gamma Z(q) + u_1(t) X(q) + u_2(t) Y(q),$$
$$q(0) = q_0,$$

where $\alpha, \beta, \gamma \in \mathbb{R}$, $u_1, u_2 : [0, T] \to \overline{B_1(0)} \subset \mathbb{R}^2$. As before, we consider the cost functional $J(u) = \|u\|_{L^1}$. The Hamiltonian function for this problem is

$$H(\lambda, u) = h_0(\lambda) + \big\langle u, h_I(\lambda) \big\rangle - |u|$$

17

where

$$h_0(\lambda) = \langle \lambda, \alpha X(q) + \beta Y(q) + \gamma Z(q) \rangle = \alpha h_X(\lambda) + \beta h_Y(\lambda) + \gamma h_Z(\lambda),$$
$$h_I(\lambda) = \big(h_X(\lambda), h_Y(\lambda)\big),$$

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

$$(u_1, u_2) = r(t) \frac{(h_X, h_Y)}{\sqrt{h_X^2 + h_Y^2}}.$$

So, the singular control u is such that $h_X^2 + h_Y^2 = 1$ along the whole trajectory. If we differentiate in time this equation and simplify, we obtain

$$\frac{d}{dt}[h_c(\lambda_t)] = h_{0c}(\lambda_t) = \big(\alpha h_Y(\lambda_t) - \beta h_X(\lambda_t)\big)h_Z(\lambda_t) = 0,$$

$$\frac{d^2}{dt^2}[h_c(\lambda_t)] = \{h_0 + rh_c, h_{0c}\} = (\alpha^2 + \beta^2 + r(\alpha h_X + \beta h_Y))h_Z^2 = 0.$$

If $h_Z = 0$ identically, then the two previous equations are trivial. 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 equation 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 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.

5.2 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 1.3 and 1.6 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

$$\dot{q} = f_0(q) + u_1(t)f_1(q) + u_2(t)f_2(q). \tag{5.1}$$

The brackets of these vector fields are (we define $f_{ij} = [f_i, f_j]$ and $f_{ijk} = [f_i, [f_j, f_k]]$)

$$f_{01} = 0, \qquad\qquad f_{02} = \alpha x \partial_z, \qquad\qquad f_{12} = x\partial_z,$$
$$f_{001} = 0, \qquad\qquad f_{002} = \alpha^2 \partial_z, \qquad\qquad f_{101} = 0,$$
$$f_{102} = \alpha \partial_z, \qquad\qquad f_{201} = 0, \qquad\qquad f_{202} = 0.$$

18

Let us also denote
$$p_x(\lambda) = \langle \lambda, \partial_x \rangle, \qquad p_y(\lambda) = \langle \lambda, \partial_y \rangle, \qquad p_z(\lambda) = \langle \lambda, \partial_z \rangle.$$

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 either 1) $h_2 = 0$ or 2) $h_{02} = 0$. In case 1) we must have

$$h_{c0c} = h_1 h_{12} h_{02} = \alpha p_x p_z^2,$$

$$r = -\frac{h_{00c}}{h_{c0c}} = -\frac{h_{02}}{h_1 h_{12}} = -\frac{\alpha}{p_x} = -\alpha p_x \frac{1}{p_x^2}$$

As a result,

$$h_{c0c} > 0 \iff r < 0,$$

so 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 have that the

$$h_{c0c} = h_1 h_2 h_{102} = \alpha p_x p_y p_z > 0.$$

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

$$\dot{x} = u_1 + \alpha, \qquad\qquad\qquad \dot{p}_x = -u_2 x p_z,$$

$$\dot{y} = u_2 + \beta, \qquad\qquad\qquad \dot{p}_y = 0,$$

$$\dot{z} = u_2 \frac{x^2}{2} + \gamma, \qquad\qquad\qquad \dot{p}_z = 0.$$

19

Therefore, if $x = 0$, the only possibility for p is to be constant.

The trajectory of (5.1) corresponding to the control u starting from $(0, y_0, z_0)$ is

$$x(t) = 0,$$

$$y(t) = \left(\beta - \frac{\alpha p_y}{p_x}\right) t + y_0,$$

$$z(t) = \gamma t + z_0.$$

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 $h_{c0c} = \frac{\alpha}{2} |\sin(2\theta)| > 0$.

So, to resume, for any triple $(\alpha, \beta, \gamma) \in [-1, 1] \times \mathbb{R}^2$, the control system (5.1) has an infinite family of singular extremals satisfying strong generalized Legendre condition. Their corresponding controls are constant and they can be parametrized as

$$u = (-\alpha, -\alpha \tan\theta), \quad \begin{cases} \text{for } \pi - \arccos\alpha < \theta < \pi + \arccos\alpha & \text{if } \alpha > 0, \\ \text{for } -\arccos|\alpha| < \theta < \arccos|\alpha| & \text{if } \alpha < 0. \end{cases}$$

5.3 Left-invariant problem in $\mathrm{SU}(2)$

In this section we are going to illustrate an example where we use all the optimality conditions developed in this paper. 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:

$$\dot{U} = U\big(\alpha A + \beta B + \gamma C + u_1(t) A + u_2(t) B\big), \quad U(0) = \mathrm{Id},$$

where $U \in \mathrm{SU}(2)$, $\alpha, \beta, \gamma \in \mathbb{R}$, $(u_1, u_2) : [0, T] \to \overline{B_1(0)} \subset \mathbb{R}^2$ 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:

$$[A, B] = C, \qquad\qquad [B, C] = A, \qquad\qquad [C, A] = B. \qquad (5.2)$$

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 leaving on the dual of the Lie algebra $\mathfrak{su}(2)$

$$h_A(\lambda) = \langle \lambda, A \rangle, \quad h_B(\lambda) = \langle \lambda, B \rangle, \quad h_C(\lambda) = \langle \lambda, C \rangle,$$
$$h_0(\lambda) = \langle \lambda, F_0 \rangle, \quad h_I(\lambda) = \big(h_A(\lambda), h_B(\lambda)\big), \quad \lambda \in (\mathfrak{su}(2))^*.$$

20

The Pontryagin (left-invariant) Hamiltonian function for our optimal control problem is

$$H(\lambda, u) = h_0(\lambda) + \langle u, h_I(\lambda) \rangle - |u|$$
$$= \alpha h_A(\lambda) + \beta h_B(\lambda) + \gamma h_C(\lambda) + r\left(\Big\langle v, \big(h_A(\lambda), h_B(\lambda)\big)\Big\rangle - 1\right).$$

From the commutation rules (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 corresponds to

$$|h_I| = 1 \iff 2h_c = h_A^2 + h_B^2 = 1.$$

The first derivative of h_c is

$$\dot{h}_c = \{h_0, h_c\} = \left\{\alpha h_A + \beta h_B + \gamma h_C, \frac{1}{2}(h_A^2 + h_B^2)\right\} = (\alpha h_B - \beta h_A)\{h_A, h_B\} + \gamma\big(h_A\{h_C, h_A\} + h_B\{h_C, h_B\}\big) =$$

$$= (\alpha h_B - \beta h_A)h_C.$$

We can further compute

$$\dot{h}_A = \{h_0 + rh_c, h_A\} = -\beta h_C + \gamma h_B - r(t)h_B h_C,$$
$$\dot{h}_B = \{h_0 + rh_c, h_B\} = \alpha h_C - \gamma h_A + r(t)h_A h_C,$$
$$\dot{h}_C = \{h_0 + rh_c, h_C\} = -\alpha h_B + \beta h_A - r(t)h_A h_B + r(t)h_B h_A = \beta h_A - \alpha h_B.$$

And the second derivative of h_c is

$$\ddot{h}_c = (\alpha \dot{h}_B - \beta \dot{h}_A)h_C + (\alpha h_B - \beta h_A)\dot{h}_C$$
$$= \begin{cases} \big((\alpha^2 + \beta^2)h_C + (rh_C - \gamma)(\alpha h_A + \beta h_B)\big)h_C, & \text{if } \alpha h_B \equiv \beta h_A \\ 0 & \text{if } h_C \equiv 0. \end{cases}$$

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

$$u_* = r(h_A, h_B) = \left(\frac{\gamma}{h_C} - (\pm 1) \right) (\alpha, \beta). \tag{5.3}$$

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 (5.3).

Denote $u_* = \left(u_*^{(1)}, u_*^{(2)} \right)$ 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

$$U(t) = e^{tX}, p = const.$$

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 is again a basis for SU(2) and, since we have assumed $\alpha^2 + \beta^2 = 1$,

$$[P_1, P_2] = C, \quad [P_2, C] = P_1, \quad [K, P_1] = P_2.$$

So, since the commutator relations are the same as (5.2), we can assume $\alpha = 1, \beta = 0$. Thus, we have

$$h_I = (1, 0), \quad h_I^\perp = (0, 1),$$
$$\langle h_I, \vec{h}_I \rangle = \vec{h}_A, \quad \langle h_I^\perp, \vec{h}_I \rangle = \vec{h}_B.$$

Given an admissible variation v, we have that

$$v(t) = \rho(t)h_I + w(t) = \rho(t)(1, 0) + \omega(t)(0, 1).$$

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

$$(\Phi_t)_*^{-1} = e^{t \operatorname{ad} \vec{H}_{u_*}}.$$

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

$$\vec{H}_1 = \vec{h}_B, \quad \vec{H}_2 = \frac{1}{\sqrt{1 + \kappa^2}}(-\kappa \vec{h}_A + \vec{h}_C), \quad \vec{H}_3 = \vec{H}_{u_*} = \frac{\gamma}{\kappa}(\vec{h}_A + \kappa \vec{h}_C), \quad \partial_{H_1}, \quad \partial_{H_2}, \quad \partial_{H_3}.$$

22

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 $\mathrm{ad}\vec{H}_3 = \frac{\gamma}{\kappa}\mathrm{ad}(\vec{h}_A + \kappa\vec{h}_C)$ acts as follows:

$$[\vec{H}_3, \vec{H}_1] = k\vec{H}_2, \qquad\qquad\qquad [\vec{H}_3, \partial_{H_1}] = 0,$$
$$[\vec{H}_3, \vec{H}_2] = -k\vec{H}_1, \qquad\qquad\qquad [\vec{H}_3, \partial_{H_2}] = 0,$$
$$[\vec{H}_3, \vec{H}_3] = 0, \qquad\qquad\qquad\quad [\vec{H}_3, \partial_{H_3}] = 0,$$

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 $e^{t\mathrm{ad}\vec{H}_3}$ 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

$$e^{t\mathrm{ad}\vec{H}_3} = \begin{pmatrix} \cos(kt) & -\sin(kt) & 0 & 0 & 0 & 0 \\ \sin(kt) & \cos(kt) & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Now, since $Z_t h_I = e^{t\mathrm{ad}\vec{H}_3}(\vec{h}_A)$ and $[\vec{H}_{u_*}, \vec{h}_A] = \gamma\vec{h}_B$, we have

$$\frac{d}{dt}(Z_t h_I) = e^{t\mathrm{ad}\vec{H}_3}(\mathrm{ad}\vec{H}_3(\vec{h}_A)) = \gamma e^{t\mathrm{ad}\vec{H}_3}\vec{h}_B.$$

So that

$$\mathcal{Z}_t v = -\frac{d}{dt}(Z_t h_I)\phi(t) + (Z_t h_I^\perp)\omega(t) = \big(\omega(t) - \gamma\phi(t)\big)e^{t\mathrm{ad}\vec{H}_3}\vec{h}_B = \big(\omega(t) - \gamma\phi(t)\big)e^{t\mathrm{ad}\vec{H}_3}\vec{H}_1 =$$

$$= \big(\omega(t) - \gamma\phi(t)\big)\big(\cos(kt)\vec{H}_1 + \sin(kt)\vec{H}_2\big).$$

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

$$\eta = \eta_1\vec{H}_1 + \eta_2\vec{H}_2 + \eta_3\vec{H}_3 + \eta_1^*\partial_{H_1} + \eta_2^*\partial_{H_2} + \eta_3^*\partial_{H_3}. \tag{5.4}$$

The symplectic form in this basis is

$$\sigma(\partial_{H_\alpha}, \vec{H}_\beta) = \langle dH_\beta, \partial_{H_\alpha}\rangle = \delta_{\alpha\beta},$$
$$\sigma(\partial_{H_\alpha}, \partial_{H_\beta}) = 0,$$
$$\sigma(\vec{H}_\alpha, \vec{H}_\beta) = H_{\alpha\beta}.$$

23

We can calculate the Poisson brackets in a straightforward manner

$$\{H_1, H_2\}(\lambda_0) = \left\{h_B, \frac{1}{\sqrt{1+\kappa^2}}(-\kappa h_A + h_C)\right\}(\lambda_0) = \frac{1}{\sqrt{1+\kappa^2}}(\kappa h_C(\lambda_0) + h_A(\lambda_0)) = \sqrt{1+\kappa^2},$$

$$\{H_1, H_3\}(\lambda_0) = \left\{h_B, \frac{\gamma}{\kappa}(h_A + \kappa h_C)\right\}(\lambda_0) = \gamma(-h_C(\lambda_0)/\kappa + h_A(\lambda_0)) = 0,$$

$$\{H_2, H_3\}(\lambda_0) = \left\{\frac{1}{\sqrt{1+\kappa^2}}(-\kappa h_A + h_C), \frac{\gamma}{\kappa}(h_A + \kappa h_C)\right\}(\lambda_0) = \frac{\gamma(1+\kappa^2)}{\kappa\sqrt{1+\kappa^2}}h_B(\lambda_0) = 0.$$

Hence, we have

$$\sigma(\cos(kt)\vec{H}_1 + \sin(kt)\vec{H}_2, \eta) =$$
$$= \sigma\left(\cos(kt)\vec{H}_1 + \sin(kt)\vec{H}_2, \eta_1\vec{H}_1 + \eta_2\vec{H}_2 + \eta_3\vec{H}_3 + \eta_1^*\partial_{H_1} + \eta_2^*\partial_{H_2} + \eta_3^*\partial_{H_3}\right)$$
$$= \cos(kt)\eta_2 H_{12} - \cos(kt)\eta_1^* - \sin(kt)\eta_1 H_{12} - \sin(kt)\eta_2^*$$
$$= \sqrt{1+\kappa^2}(\cos(kt)\eta_2 - \sin(kt)\eta_1) - (\cos(kt)\eta_1^* + \sin(kt)\eta_2^*).$$

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

$$l_t^{-1}\sigma(\mathcal{Z}_t\cdot, \eta) = \left(\frac{\gamma\kappa}{r} - \kappa^2\right)^{-1} \begin{pmatrix} \gamma\kappa & \kappa \\ \kappa & \frac{1}{r} \end{pmatrix} \begin{pmatrix} 1 \\ -\gamma \end{pmatrix} \psi(t, \eta) = -\frac{\psi(t, \eta)}{\kappa}\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Hence

$$\dot{\eta} = -\mathcal{Z}_t l_t^{-1}\sigma(\mathcal{Z}_t\cdot, \eta) = \tag{5.5}$$
$$= -\frac{\gamma}{\kappa}\left(\sqrt{1+\kappa^2}(\eta_2\cos(kt) - \eta_1\sin(kt)) - \eta_1^*\cos(kt) - \eta_2^*\sin(kt)\right)\left(\cos(kt)\vec{H}_1 + \sin(kt)\vec{H}_2\right).$$

First, we are going find six linearly independent solutions to equation (5.5) and then we are going to determine solutions corresponding to conjugate times. If we compare equation (5.5) with (5.4), we immediately see that $\dot{\eta}_1^* = \dot{\eta}_2^* = \dot{\eta}_3^* = \dot{\eta}_3 = 0$. To determine a solution of equation (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

$$\text{span}\left\{(0,0,1,0,0,0), (0,0,0,0,0,1), (0,1,0,\sqrt{1+\kappa^2},0,0), (1,0,0,0,-\sqrt{1+\kappa^2},0)\right\}$$

are zeros of the r.h.s. of (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:

$$\begin{pmatrix} \dot{\eta}_1 \\ \dot{\eta}_2 \end{pmatrix} = k \begin{pmatrix} ab & -a^2 \\ b^2 & -ab \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}. \tag{5.6}$$

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 variable

$$\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}, \tag{5.7}$$

and rewrite system (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 (5.7) is the identity for $t = 0$, we have $\xi(0) = (\eta_1(0), \eta_2(0))$ and

$$\begin{pmatrix} \eta_1(t) \\ \eta_2(t) \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} \eta_1(0) \\ -kt\eta_1(0) + \eta_2(0) \end{pmatrix} = \begin{pmatrix} \big(\cos(kt) + kt\sin(kt)\big)\eta_1(0) - \sin(kt)\eta_2(0) \\ \big(\sin(kt) - kt\cos(kt)\big)\eta_1(0) + \cos(kt)\eta_2(0) \end{pmatrix}.$$

So, we have obtained a family of six linearly independent solution for the equation (5.5). We still have to determine a solution with the appropriate boundary conditions, which are

$$\eta(0) \in \big(T_{\lambda_0}(T_{q_0}^* M) + \mathbb{R}Z_I(0)\big) \cap \mathbb{R}Z_I(0)^{\measuredangle},$$
$$\eta(T) \in T_{\lambda_0}(T_{q_0} M). \tag{5.8}$$

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

$$\big(T_{\lambda_0}(T_{q_0}^* M) + \mathbb{R}Z_I(0)\big) = \text{span}\{\partial_{H_1}, \partial_{H_2}, \partial_{H_3}, \vec{h}_A\},$$
$$\mathbb{R}Z_I(0)^{\measuredangle} = \text{span}\{\vec{h}_A, \vec{H}_3, \partial_{H_1}, c_3\partial_{H_2} + c_2\partial_{H_3}, \vec{H}_1 - \sqrt{1+\kappa^2}\partial_{H_2}\}.$$

Therefore,

$$\big(T_{\lambda_0}(T_{q_0}^* M) + \mathbb{R}Z_I(0)\big) \cap \mathbb{R}Z_I(0)^{\measuredangle} = \text{span}\{\vec{h}_A, \partial_{H_1}, c_3\partial_{H_2} + c_2\partial_{H_3}\}.$$

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 (5.8). 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) and a time-varying solution.

- If $\eta(0) = \vec{h}_A = c_2\vec{H}_2 - c_3\vec{H}_3$, then

$$\eta(t) = -\sin(kt)c_2\vec{H}_1 + \cos(kt)c_2\vec{H}_2 - c_3\vec{H}_3.$$

25

- If $\eta(0) = \sqrt{1 + \kappa^2}\partial_{H_1} = (\sqrt{1 + \kappa^2}\partial_{H_1} + \vec{H}_2) - \vec{H}_2$, then

$$\eta(t) = (\sqrt{1 + \kappa^2}\partial_{H_1} + \vec{H}_2) + \sin(kt)\vec{H}_1 - \cos(kt)\vec{H}_2 =$$
$$= \sqrt{1 + \kappa^2}\partial_{H_1} + \sin(kt)\vec{H}_1 + (1 - \cos(kt))\vec{H}_2.$$

- 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

$$\eta(t) = -\sqrt{1 + \kappa^2}c_2\partial_{H_3} + c_3(\vec{H}_1 - \sqrt{1 + \kappa^2})\partial_{H_2}) - c_3\big((\cos(kt) + kt\sin(kt))\vec{H}_1 + (\sin(kt) - kt\cos(kt))\vec{H}_2\big) =$$

$$= -\sqrt{1 + \kappa^2}c_2\partial_{H_3} - \sqrt{1 + \kappa^2}c_3\partial_{H_2} + c_3(1 - \cos(kt) - kt\sin(kt))\vec{H}_1 - c_3(\sin(kt) - kt\cos(kt))\vec{H}_2.$$

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 (5.8), we can restrict to this case. Hence, we have to find $a_2, a_3 \in \mathbb{R}$ such that

$$a_3c_3\big(1 - \cos(kt) - kt\sin(kt)\big) + a_2\sin(kt) = 0,$$
$$-a_3c_3\big(\sin(kt) - kt\cos(kt)\big) + a_2(1 - \cos(kt)) = 0.$$

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:

$$\big(1 - \cos(kt) - kt\sin(kt)\big)(1 - \cos(kt)) + \big(\sin(kt) - kt\cos(kt)\big)\sin(kt) = 0,$$

which simplified is

$$2 - 2\cos(kt) - kt\sin(kt) = 0,$$

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 for $t = 2\pi/k$ and for $\bar{t}$ that satisfies

$$\tan\left(\frac{k\bar{t}}{2}\right) = \frac{k\bar{t}}{2}.$$

6 Second variation and the generalized Legendre condition

6.1 Definition of the end-point map and the Hessian

When dealing with control problems, a central object is the Endpoint-map.

Definition 6.1 (Endpoint-map and extended Endpoint-map)**.** We will call Endpoint-map based at $q_0 \in M$ the map $F_{q_0, T}$ that to any control $u \in \mathcal{U}$ associate the Endpoint of the solution of (1.1) with starting point q_0:

$$F_{q_0, T} : \mathcal{U} \to M, \quad F_{q_0, T}(u) = q_u(T; q_0).$$

We will call extended Endpoint-map based at $q_0 \in M$ the map $E_{q_0, T}$ that to any control $u \in \mathcal{U}$ associate the couple made by $q_u(T; q_0)$ and the value $J(u)$:

$$E_{q_0, T} : \mathcal{U} \to M \times \mathbb{R}, \quad E_{q_0, T}(u) = \big(q_u(T; q_0), J(u)\big).$$

In the following, we will drop the subscript q_0 whenever there is no ambiguity.

Remark 6.2. If u is an interior point of $\mathcal{U}$, then it is possible to reformulate PMP in terms of the extended Endpoint-map. More precisely, if u is a solution to the Optimal Control Problem 1, then it is a critical point of the extended Endpoint-map, that is there is some non-trivial $(\lambda, \nu) \in T^*_{q(T)} M \times \mathbb{R}$ such that

$$\big\langle (\lambda, \nu), D_u E_T[v] \big\rangle = \nu D_u J + \lambda D_u F_T = 0, \quad \text{for every } v \in T_u \mathcal{U},$$

where D_u denotes the differential at a point $u \in \mathcal{U}$.

As for classical optimization problem on finite dimensional spaces, once we have found the critical points of the functional that we want to minimize or maximize, we can look at the second derivative in order to distinguish minimum points from maximum and saddle points. In our case, we have to look to the second variation of the extended Endpoint-map.

Now, we recall the main facts about the second variation. For a general introduction to the second variation of Endpoint-maps coming from optimal control problems, see [7], Chapter 20. See also [1], Section 2, for a concise summary on this topic, and [10, 2] for more properties of the second variation.

First, we have to introduce the notion of Hessian. Let $\mathcal{U}$ be a Banach space and $F : \mathcal{U} \to M$ be a smooth map. Let $u \in \text{int} \mathcal{U}$ be a critical point of F, that is $D_u F$ is not surjective. The *Hessian* of a map F at the point u is the bilinear map

$$\text{Hess}_u F : \ker D_u F \times \ker D_u F \to T_{F(u)} M / \text{im} D_u F$$

defined by

$$\lambda \text{Hess}(v, w) = W(V(a \circ F)),$$

where $\lambda \in (\text{im} D_u F)^\perp$, $a \in C^\infty(M)$ with $d_{F(u)} a = \lambda$, and $V, W \in \text{Vec}(\mathcal{U})$ with $V(u) = v, W(u) = w$. It is possible to show that indeed $W(V(a \circ F)) = V(W(a \circ F))$ and that this definition does not depend on the extensions V, W and a but just on the values v, w, λ (see [7]).

More explicitly, in the particular case of E_T, one has that its Hessian is

$$(\lambda, \nu) \text{Hess} E_T = \big(\nu D_u^2 J - \langle \lambda, D_u^2 F_T \rangle \big)_{|\ker D_u E_T}$$

which, if $\nu \neq 0$, also coincides with the second derivative of E_T restricted to the level set $F_T^{-1}(u)$.

Definition 6.3 (Locally open map)**.** A map $F : \mathcal{U} \to M$ is called *locally open* at $u \in \mathcal{U}$ if for every open neighbourhood O_u of u it holds that

$$F(u) \in \text{int} F(O_u).$$

Of course, if the Extended End-Point map E_T is locally open at a point u, then the control u cannot be optimal. Indeed, in this case, there are controls arbitrarily close to u realizing the same endpoint with a smaller cost.

Definition 6.4. The negative index of a quadratic form $Q : V \to \mathbb{R}$ is

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

We have the following criterion for local openness of a map.

Theorem 6.5. *Let $F : \mathcal{U} \to M$ be a smooth mapping and suppose that $u \in \mathcal{U}$ is a critical point of F such that $\text{im } D_u F$ has corank r. If*

$$\text{ind}_- \lambda \text{ Hess}_u F \geq r \quad \forall \lambda \in \text{im } D_u F^\perp,$$

then F is locally open at u.

Now, we go back to the optimal control problem. We see that, if $\text{ind}^-(\lambda, \nu) \text{ Hess}_u E_T = +\infty$ for a critical point u, then we immediately know that E_T is locally open at u, hence u is not optimal. So, our first goal is to compute the Hessian of E_T and then to find a condition that ensures $\text{ind}^-(\lambda, \nu) \text{ Hess}_u E_T < +\infty$. Moreover, in Subsection 6.4, we show how that it is possible to compute $\text{ind}^-(\lambda, \nu) \text{ Hess}_u E_T$ finding the solutions of the Jacobi equation, which together with Theorem 6.5 will prove Theorem 1.5.

From now on we fix a critical point $u \in \mathcal{U}$ of E_T and we assume that it is a singular control of the form (2.3). Moreover, we denote by λ its corresponding singular extremal, which is solution of the Equation (2.4). For sake of brevity, we will also use the notation

$$Q_T := (\lambda, \nu) \text{ Hess}_u E_T.$$

6.2 Explicit expression for the second variation form

The formula of the Hessian of the endpoint map E_T for a general control system is computed in [7]. It reads as:

$$Q_T(v) = \int_0^T \left(\frac{d^2 H(\lambda_0, u(t))}{du^2}(v(t), v(t)) + \sigma_{\lambda_0}\left(Z_t v(t), \int_0^t Z_s v(s) \, ds \right) \right) dt, \quad v \in \ker D_u E_T,$$

where the first summand is the Hessian of the PMP Hamiltonian with respect to the control variables. The map $Z_t : T_{u(t)} U \to T_{\lambda_0}(T^* M)$ was defined in (4.2), but we recall the definition here:

$$Z_t v(t) := \left\langle v(t), \left[(\Phi_t^{-1})_* \vec{h}_I\right](\lambda_0) \right\rangle,$$

where Φ_t is the flow on $T^* M$ of the Hamiltonian system (2.4) 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. In symplectic language a variation $v \in T_u \mathcal{U} = L^\infty([0, T], \mathbb{R}^m)$ is admissible if and only if (see [7])

$$v \in \ker D_u E_T \quad \Longleftrightarrow \quad \int_0^T Z_t v(t) \, dt \in \Pi_0, \tag{6.1}$$

where $\Pi_0 = T_{\lambda_0}(T_{q_0}^* M)$ as before is the vertical fibre.

It is important to note that both the form Q_T and the first differential in (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 $\langle h_I, w \rangle = 0$. A direct computation, using the short notation $Z_I(t) = Z_t h_I(t)$, shows that

$$\frac{d^2 H(\lambda_0, u(t))}{du^2}(v, v) = \frac{|v|^2}{r} - \frac{\langle h_I, v \rangle^2}{r} = \frac{|w|^2}{r}.$$

28

Then we obtain

$$Q_T(v) = \int_0^T \left(\frac{|w(t)|^2}{r} + \sigma_{\lambda_0}\left(Z_t\big(\rho(t)h_I(t) + w(t)\big), \int_0^t Z_s\big(\rho(s)h_I(t) + w(s)\big)\, ds \right) \right) dt = \tag{6.2}$$

$$= \int_0^T \frac{|w(t)|^2}{r} dt +$$

$$+ \int_0^T \sigma_{\lambda_0}\left(\rho(t)Z_I(t), \int_0^t \rho(s)Z_I(s)ds \right) dt + \tag{6.3}$$

$$+ \int_0^T \sigma_{\lambda_0}\left(\rho(t)Z_I(t), \int_0^t Z_s w(s)ds \right) dt + \tag{6.4}$$

$$+ \int_0^T \sigma_{\lambda_0}\left(Z_t w(t), \int_0^t \rho(s)Z_I(s)ds \right) dt + \tag{6.5}$$

$$+ \int_0^T \sigma_{\lambda_0}\left(Z_t w(t), \int_0^t Z_s w(s)ds \right) dt.$$

The form Q_T is not coercive in the L^2-topology. To see this, we can consider the family variations of the form $w \equiv 0$ and $\rho_\varepsilon(t) = \rho_1(t/\varepsilon)$, $0 < \varepsilon \le 1$ and for any admissible ρ_1. A direct computation shows that $Q_T = O(\varepsilon^2)$ for $\varepsilon \to 0$, 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 (6.1), we get

$$\int_0^T Z_t v(t)\, dt = \int_0^T Z_t\left(\rho(t)h_I(\lambda_t) + w(t) \right) dt =$$

$$= \phi(T)Z_I(T) - \int_0^T \phi(t)\dot{Z}_I(t)dt + \int_0^T Z_t w(t)\, dt = \tag{6.6}$$

$$= \phi(T)Z_I(T) + \int_0^T \mathcal{Z}_t v(t)dt \in \Pi_0,$$

where we have shortened

$$\mathcal{Z}_t v(t) := -\phi(t)\dot{Z}_I(t) + Z_t w(t).$$

Similarly, integrating parts in the formula for Q_T (see Appendix A) we obtain an expression:

$$Q_T(v_1, v_2) = \phi_1(T)\sigma\left(Z_I(T), \int_0^T \mathcal{Z}_t v_2(t)dt \right) + \int_0^T \langle v_1(t), l_t v_2(t)\rangle dt + \int_0^T \sigma\left(\mathcal{Z}_t v_1(t), \int_0^t \mathcal{Z}_s v_2(s)\, ds \right) dt \tag{6.7}$$

where

$$l_t := \begin{pmatrix} \frac{\mathrm{Id}}{r} & \sigma\big(Z_t \cdot, Z_I(t)\big) \\ \sigma\big(Z_t \cdot, Z_I(t)\big)^T & \sigma\big(Z_I(t), \dot{Z}_I(t)\big) \end{pmatrix}. \tag{6.8}$$

and $l_t \in \mathbb{R}^{(m+1)\times(m+1)}$ with $\mathrm{Id} \in \mathbb{R}^{m \times m}$.

From here we can notice that the operator in (6.6) and the quadratic form (6.7) will be 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.

6.3   Proof of Theorem 1.1

To prove Theorem 1.1 we restrict the form Q_T in (6.7) to the finite-dimensional subspace with $\phi(T) = 0$. Then, essentially we have

$$Q_T(v) = \int_0^T \langle v(t), l_t v(t) \rangle dt + \text{ other terms.} \tag{6.9}$$

Following the approach in [7, Chapter 20], we have to check that the form Q_T has a finite negative index.

Proposition 6.6. *If the quadratic form Q_T 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_T restricted to variations supported in $[\tau, \tau + \varepsilon]$ is strictly positive definite. In particular, Q_T has finite negative index.*

The proof is analogous to the proof of Proposition 20.1 of [7] 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 (6.9)). 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

$$\begin{pmatrix} \frac{\mathrm{Id}}{r} & a \\ a^T & \alpha \end{pmatrix} = \begin{pmatrix} \mathrm{Id} & 0 \\ ra^T & 1 \end{pmatrix} \begin{pmatrix} \frac{\mathrm{Id}}{r} & 0 \\ 0 & \alpha - r|a|^2 \end{pmatrix} \begin{pmatrix} \mathrm{Id} & ra \\ 0 & 1 \end{pmatrix}$$

where $a = \sigma\big(Z_t \cdot, Z_I(t)\big)$ and $\alpha = \sigma\big(Z_I(t), \dot{Z}_I(t)\big)$. Hence, the previous matrix is non-negative if and only if

$$\sigma\left(Z_I(t), \dot{Z}_I(t)\right) - r\|\sigma\big(Z_t \cdot, Z_I(t)\big)\|^2 \geq 0, \tag{6.10}$$

30

We point out that the form l_t is non-negative definite for $(w, \phi) \in (T_{u(t)}U \cap h_I^\perp) \times \mathbb{R}$ if and only if it is non-negative definite for $(w, \phi) \in T_{u(t)}U \times \mathbb{R}$. One implication is obvious. Concerning the other one, if you take $(w, \phi) \in T_{u(t)}U \times \mathbb{R}$, 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 (6.10) explicitly in terms of the Hamiltonians h_i. Recall, that in Proposition 4.4 we claimed that $l_t \geq 0$ is equivalent to $h_{c0c}(t) \geq 0$. So, to prove the Proposition, we only need to shows that (6.10) is equivalent to $h_{c0c}(t) \geq 0$.

Proof of Proposition 4.4. From the definition of Z_t, we have

$$Z_I(t) = (\Phi_t^{-1})_* \big(\langle h_I(\lambda_t), \vec{h}_I \rangle \big).$$

If we differentiate in t, we obtain

$$\dot{Z}_I(t) = (\Phi_t^{-1})_* \left\langle \frac{d}{dt}[h_I(\lambda_t)], \vec{h}_I \right\rangle + (\Phi_t^{-1})_* \left[\vec{h}_0 + \langle u(t), \vec{h}_I \rangle, \left\langle h_I(\lambda_t), \vec{h}_I \right\rangle \right]$$

$$= (\Phi_t^{-1})_* \left\langle \{h_0 + \langle u(t), h_I \rangle, h_I\}(\lambda_t), \vec{h}_I \right\rangle + (\Phi_t^{-1})_* \left[\vec{h}_0, \left\langle h_I(\lambda_t), \vec{h}_I \right\rangle \right]$$

$$= (\Phi_t^{-1})_* \left(\sum_{i=1}^m \left(h_{0i} + r(t) \sum_{j=1}^m h_j(\lambda_t) h_{ji} \right) \vec{h}_i + \sum_{i=1}^m h_i(\lambda_t) \vec{h}_{0i} \right) =$$

$$= (\Phi_t^{-1})_* \left(\vec{h}_{0c} + r(t) \sum_{i,j=1}^m h_j(\lambda_t) h_{ji} \vec{h}_i \right).$$

So, we can compute $\sigma_{\lambda_0} \left(Z_I(t), \dot{Z}_I(t) \right)$:

$$\sigma_{\lambda_0} \left(Z_I(t), \dot{Z}_I(t) \right) = \sigma_{\lambda(t)} \left(\sum_{i=l}^m h_l(\lambda_t) \vec{h}_l, \vec{h}_{0c} + r(t) \sum_{i,j=1}^m h_j(\lambda_t) h_{ji} \vec{h}_i \right) =$$

$$= h_{c0c} + r \sum_{i,j,l=1}^m h_j h_l h_{ji} h_{li}.$$

The second sum is (every function now is evaluated in $\lambda(t)$)

$$\sum_{i,l,j=1}^m h_i h_{jl} h_j h_{il} = \sum_{i,l=1}^m h_i h_{il} \sum_j h_{jl} h_j = \sum_{l=1}^m \left[\left(\sum_j h_{jl} h_j \right) \sum_i h_i h_{il} \right] = \sum_{l=1}^m \left(\sum_j h_{jl} h_j \right)^2$$

$$= \sum_l \{h_l, h_c\}^2 = |\{h_c, h_I\}|^2 = |h_{cI}|^2,$$

where h_{cI} is the vector whose i^{th} component is h_{ci}. Thus, putting everything together, we obtain

$$\sigma_{\lambda_0} \left(Z_I(t), \dot{Z}_I(t) \right) = h_{c0c}(\lambda_t) + r(\lambda_t) |h_{cI}(\lambda_t)|^2. \tag{6.11}$$

On the other hand, for any admissible variation v the second addendum in (6.10) is

$$\sigma_{\lambda_0}(Z_t v, Z_I(t)) = \sigma_{\lambda_0}(\langle v(t), (\Phi_t^{-1})_* \vec{h}_I \rangle, \langle h_I(\lambda_t), (\Phi_t^{-1})_* \vec{h}_I \rangle) = \sum_{i,j} v_i(t) h_j(\lambda_t) \sigma_{\lambda(t)}(\vec{h}_i, \vec{h}_j) =$$

$$= \sum_{i,j} v_i(t) h_j(\lambda_t) h_{ij}.$$

Thus

$$|\sigma_{\lambda_0}(Z_t \cdot, Z_I(t))|^2 = \sum_i \left(\sum_j h_j h_{ij} \right)^2 = |h_{cI}|^2. \tag{6.12}$$

Finally, using (6.11) and (6.12), the left hand side of Equation (6.10) reads

$$\sigma_{\lambda_0}\left(Z_I(t), \dot{Z}_I(t) \right) - r\|\sigma_{\lambda_0}\left(Z_t \cdot, Z_I(t) \right)\|^2 = h_{c0c} + r|h_{cI}|^2 - r|h_{cI}|^2 = h_{c0c},$$

which proves the result. □

6.4   Derivation of the Jacobi equation and proof of Theorem 1.5

We consider a singular extremal for which the strong Legendre condition is satisfied. In this subsection we shorten $\sigma = \sigma_{\lambda_0}$.

In Proposition 6.6, 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

$$\mathcal{V}_t := \{(\phi(t), \phi, w) \in \mathbb{R} \times L^2([0,T]) \times L^2([0,T], \mathbb{R}^{m-1}) \mid \operatorname{supp}\phi, \operatorname{supp}w \subseteq [0,t]\} \cap \ker D_u E_T$$

$$t_* := \sup\{t \in [0,T] \mid Q_{T|\mathcal{V}(t)} \text{ is positive definite}\}.$$

If $t_* < T$, we say t_* is conjugate to 0 along the trajectory λ. The remaining part of this Subsection is devoted to the proof of the equivalence between this notion of conjugate time and the one already given in 4.1 are equivalent.

We have the following result.

Proposition 6.7. *Let λ be a singular extremal of Problem 1 satisfying strong generalized Legendre condition. Suppose that $t_* \in (0,T]$. Then $Q_{t_*}|_{\mathcal{V}(t)}$ is singular.*

The proof of this fact follows very closely the proof of Proposition 21.1 in [7], 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 notations we simply rename t_* to T.

To derive the Jacobi equation, we look for admissible variations $\bar{v}$ in the kernel of Q_T:

$$Q_T(\cdot, \bar{v}) = 0.$$

On the other hand, an admissible variation must satisfy (6.6) or, equivalently,

$$\sigma\left(\cdot_{\phi(T)}\left(Z_I(T) \right) + \int_0^T \mathcal{Z}_t \cdot_v dt, \nu \right) = 0, \quad \forall \nu \in \Pi_0.$$

Therefore, $\ker D_u E_T$ can be identified with a common kernel of a finite number of linear forms. Since, by assumption $Q_T(\cdot, \bar{v}) = 0$ on $\ker D_u E_T$, it follows that there is $\nu \in \Pi_0$ such that

$$Q_T(\cdot, \bar{v}) = \sigma\left(\cdot_{\phi(T)} \left(Z_I(T) \right) + \int_0^T \mathcal{Z}_t \cdot_v \, dt, \nu \right), \quad \text{on } T_u \mathcal{U}. \tag{6.13}$$

The equation (6.13) splits into two equations:

$$\sigma\left(Z_I(T), \int_0^T \mathcal{Z}_t \bar{v} dt \right) = \sigma\left(Z_I(T), \nu \right) \tag{6.14}$$

$$\int_0^T l_t(\cdot_v, \bar{v}) dt + \int_0^T \sigma\left(\mathcal{Z}_t \cdot_v, \int_0^t \mathcal{Z}_s \bar{v} ds \right) dt = \int_0^T \sigma(\mathcal{Z}_t \cdot_v, \nu) dt \tag{6.15}$$

We denote

$$\eta(t) := \nu - \int_0^t \mathcal{Z}_t \bar{v} dt. \tag{6.16}$$

Since we have the equalities for the forms in (6.15), 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 (6.16) we find

$$\dot{\eta}(t) = -\mathcal{Z}_t \bar{v}(t) = -\mathcal{Z}_t l_t^{-1} \sigma\left(\mathcal{Z}_t \cdot, \eta(t) \right) \quad \text{for a.e. } t \in [0, T].$$

which gives us a non-autonomous linear ODE for η. Therefore, we have derived the *Jacobi equation*.

Concerning (6.14), we have

$$\sigma(\eta(T), Z_I(T)) = 0. \tag{6.17}$$

It is common to rewrite this condition using skew-orthogonal complement as in (4.1). Using this notation, we can take into account both (6.6) and (6.17) as

$$\eta(T) \in \left(\Pi_0 + \mathbb{R} Z_I(T) \right) \cap \left(\mathbb{R} Z_I(T) \right)^{\angle} = \Pi_0^{\mathbb{R} Z_I(T)}.$$

This motivates the following definition.

Definition 6.8. 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:

$$\dot{\eta}(t) = -\mathcal{Z}_t l_t^{-1} \sigma\left(\mathcal{Z}_t \cdot, \eta(t) \right) \quad \text{for a.e. } t \in [0, T], \tag{6.18}$$

$$\eta(0) \in \Pi_0, \eta(T) \in \Pi_0^{\mathbb{R} Z_I(T)}. \tag{6.19}$$

The corresponding point $\pi(\lambda(t))$ is said to be conjugate to $\pi(\lambda(0))$.

Recall that previously we have given a slightly different definition of a conjugate point in Definition 4.1. But the two definitions are actually equivalent, even though the Definition 6.8 is more convenient in practice.

Proposition 6.9. *Definitions 4.1 and 6.8 are equivalent.*

Proof. We show that solutions of Equation (6.18) satisfy boundary conditions (6.19) if and only if they satisfy also the boundary conditions (4.4). By Lemma (4.2) $\eta(t) = Z_I(t)$ 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 (6.19) or (4.4). So, if T is conjugate with 0 there must be some vector $\nu \in \Pi_0$ such that the corresponding solution to (6.18) with $\eta(0) = \nu$ is such that $\eta(T) \in \Pi_0 \cap Z_I(T)^{\angle}$. Hence, we must have $\sigma(\eta(T), Z_I(T)) = 0$. Since the Equation (6.18) is hamiltonian, we will have that if and only if $\sigma(\eta(0), Z_I(0)) = 0$. $\square$

Finally, we can prove Theorem 1.5.

Proof of Theorem 1.5. Let λ be a singular extremal for Problem 1 corresponding to the singular control u and let q be its trajectory on M. If the conjugate points along q are isolated, the previous argument shows that the negative Morse index of Q_T is equal to the number of conjugate points, counted with multiplicity. Let N be the sum of these multiplicity. By Theorem 6.5, if q is of corank r, we have that if $N \geq r$, then E_T is locally open at u. In particular, u is not a solution of the optimal control problem. $\square$

A   Proof of formula (6.9)

Here we are going to show how to obtain (6.9) from (6.2). We have just to integrate by part ρ in every integral where it appears.

For the first integral (6.3), we have

$$\int_0^T \rho_1(t)\sigma\left(Z_I(t), \int_0^t \rho_2(s)Z_I(s)ds\right)dt =$$

$$\overset{\text{int.by\_parts}}{=} \int_0^T \rho_1(t)\sigma\left(Z_I(t), \phi_2(t)Z_I(t) - \int_0^t \phi_2(s)\dot{Z}_I(s)ds\right)dt =$$

$$= -\int_0^T \rho_1(t)\sigma\left(Z_I(t), \int_0^t \phi_2(s)\dot{Z}_I(s)ds\right)dt \overset{\text{Fubini}}{=} -\int_0^T\left(\int_s^T \rho_1(t)\sigma\left(Z_I(t), \phi_2(s)\dot{Z}_I(s)\right)dt\right)ds$$

$$= -\int_0^T \phi_2(s)\sigma\left(\int_s^T \rho_1(t)Z_I(t)dt, \dot{Z}_I(s)\right)ds =$$

$$\overset{\text{int.by\_parts}}{=} -\int_0^T \phi_2(s)\sigma\left(\phi_1(T)Z_I(T) - \phi_1(s)Z_I(s) - \int_s^T \phi_1(t)\dot{Z}_I(t)dt\,,\, \dot{Z}_I(s)\right)ds =$$

$$= -\sigma\left(\phi_1(T)Z_I(T), \int_0^T \phi_2(s)\dot{Z}_I(s)ds\right) +$$

$$+ \int_0^T \phi_1(s)\phi_2(s)\sigma\left(Z_I(s), \dot{Z}_I(s)\right)ds + \int_0^T \sigma\left(\phi_1(t)\dot{Z}_I(t), \int_0^t \phi_2(s)\dot{Z}_I(s)ds\right)dt.$$

A similar computation for (6.4) yields

$$\int_0^T \sigma\left(\rho_1(t)Z_I(t), \int_0^t Z_s w_2(s)ds\right)dt \overset{\text{Fubini}}{=} \int_0^T \sigma\left(\int_s^T \rho_1(t)Z_I(t)dt, Z_s w_2(s)\right)ds$$

$$= \int_0^T \phi_1(T)\sigma\left(Z_I(T), Z_s w_2(s)\right)ds - \int_0^T \phi_1(s)\sigma\left(Z_I(s), Z_s w_2(s)\right)ds -$$

$$- \int_0^T \sigma\left(\int_s^T \phi_1(t)\dot{Z}_I(t)dt, Z_s w_2(s)\right)ds =$$

$$= \sigma\left(\phi_1(T)Z_I(T), \int_0^T Z_s w_2(s)ds\right) + \int_0^T \phi_1(s)\sigma\left(Z_s w_2(s), Z_I(s)\right)ds -$$

$$- \int_0^T \sigma\left(\phi_1(t)\dot{Z}_I(t), \int_0^t Z_s w_2(s)ds\right)dt$$

And similarly for (6.5)

35

$$\int_0^T \sigma\left(Z_t w_1(t), \int_0^t \rho_2(s) Z_I(s) ds\right) dt =$$

$$= \int_0^T \phi_2(t) \sigma\left(Z_t w_1(t), Z_I(t)\right) dt - \int_0^T \sigma\left(Z_t w_1(t), \int_0^t \phi_2(s) \dot{Z}_I(s) ds\right) dt$$

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

$$Q_1(v_1, v_2) = \int_0^T \left(\frac{\langle w_1(t), w_2(t)\rangle^2}{r(t)} + \phi_1(t)\phi_2(t)\, \sigma\left(Z_I(t), \dot{Z}_I(t)\right) + \right.$$

$$\left. + \phi_1(s)\sigma\left(Z_t w_2(t), Z_I(t)\right) + \phi_2(t)\sigma\left(Z_t w_1(t), Z_I(t)\right)\right) dt,$$

$$Q_2(v_1, v_2) = \int_0^T \left(\sigma\left(\phi_1(t)\dot{Z}_I(t), \int_0^t \phi_2(s)\dot{Z}_I(s) ds\right) - \sigma_{\lambda_0}\left(\phi_1(t)\dot{Z}_I(t), \int_0^t Z_s w_2(s) ds\right) - \right.$$

$$\left. - \sigma\left(Z_t w_1(t), \int_0^t \phi_2(s)\dot{Z}_I(s) ds\right)\right) dt,$$

$$Q_3(v_1, v_2) = \sigma\left(\phi_1(T) Z_I(T), \int_0^T \left(-\phi_2(s)\dot{Z}_I(s) + Z_s w_2(s)\right) ds\right)$$

One can recognize in Q_1 the principal part of formula (6.9).

B   Proof of Proposition 4.3

First, we derive a general formula for the left hand side of (4.7):

$$H_S(t, \ell) - H(u(t), \ell) = h_S(\ell) - h_0(\ell) + r(t)\left(|h_I(\ell)| - \langle h_I(t), h_I(\ell)\rangle\right).$$

The first derivative is

$$d_\ell\left(H_S(t, \cdot) - H(u(t), \cdot)\right) = d_\ell\left(h_S - h_0\right) + r(t)\left(\frac{1}{|h_I(l)|}\langle h_I(l), d_\ell h_I\rangle - \langle h_I(t), d_\ell h_I\rangle\right),$$

and the second derivative at λ_0 is

$$d_{\lambda_0}^2\left[\left(H_S(t, \cdot) - H(u(t), \cdot)\right) \circ \Phi_t\right][\eta_1, \eta_2] = d_{\lambda(t)}^2\left(H_S(t, \cdot) - H(u(t), \cdot)\right)[\xi_1, \xi_2] =$$

$$= d^2\left(h_S - h_0\right)[\xi_1, \xi_2] + r(t)\left(\langle dh_I[\xi_1], dh_I[\xi_2]\rangle - \langle h_I, dh_I[\xi_1]\rangle\langle h_I, dh_I[\xi_2]\rangle\right), \quad \text{(B.1)}$$

where $\xi_i = (\Phi_t)_* \eta_i$, $i = 1, 2$ and every differential and function in the last line is evaluated at $\lambda(t)$. Here we have used the linearity of $h_I(l)$ and that $\lambda_t \in \Sigma$. Now, to compute explicitly both sides of (4.7), for every fixed $t \in [0, T]$, we consider the splitting

$$T_{\lambda_0}(T^*M) = (\Phi_t^{-1})_*\left(\mathbb{R}\vec{h}_c(\lambda(t)) \oplus T_{\lambda(t)}\mathcal{S}\right) = \mathbb{R}Z_I(t) \oplus (\Phi_t^{-1})_*\left(T_{\lambda(t)}\mathcal{S}\right).$$

36

So, by bilinearity, it suffices to check (4.7) for

$$(\eta_1, \eta_2) \in \left\{ \big(Z_I(t), Z_I(t)\big), \big(Z_I(t), (\Phi_t^{-1})_* X\big), \big((\Phi_t^{-1})_* X, (\Phi_t^{-1})_* X\big) \right\},$$

where $X \in T_{\lambda(t)}\mathcal{S}$. Thanks to this choice, the l.h.s. of (4.7) will be evaluated at $\xi_i = (\Phi_t)_* \eta_i \in \{\vec{h}_c(\lambda(t)), X\}$, simplifying the two differentials. We compute the right and the left hand-side for each case separately. First, consider the case $(\eta_1, \eta_2) = \big(Z_I(t), Z_I(t)\big)$. We start by plugging it into the right hand side of (4.7). First, we have

$$\sigma_{\lambda_0}\big(\mathcal{Z}_t \cdot, Z_I(t)\big) = \sigma_{\lambda_0}\big(Z_t \cdot w, Z_I(t)\big) - \cdot_\phi \sigma_{\lambda_0}\big(\dot{Z}_I(t), Z_I(t)\big)$$

Comparing this with the matrix expression for l_t (see (6.8)) we find

$$\sigma_{\lambda_0}\big(\mathcal{Z}_t \cdot, Z_I(t)\big) = l_t \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Hence

$$\big\langle \sigma_{\lambda_0}\big(\mathcal{Z}_t \cdot, Z_I(t)\big), l_t^{-1} \sigma_{\lambda_0}\big(\mathcal{Z}_t \cdot, Z_I(t)\big) \big\rangle = -\sigma_{\lambda_0}\big(\dot{Z}_I(t), Z_I(t)\big). \tag{B.2}$$

Now, let us compute the left-hand side. Since by construction h_S is constant along integral curves of $\vec{h}_c$, we get

$$d^2_{\lambda(t)}\big(h_S - h_0\big)[(\Phi_t)_* Z_I(t), (\Phi_t)_* Z_I(t)] = -d^2_{\lambda(t)} h_0 [\vec{h}_c(\lambda_t), \vec{h}_c(\lambda_t)] = -h_{cc0}(t) = h_{c0c}(t),$$

So, substituting $(\eta_1, \eta_2) = (Z_I(t), Z_I(t))$ into (B.1), we obtain

$$d^2_{\lambda(t)}\big(H_S(t, \cdot) - H(u(t), \cdot)\big)[\vec{h}_c]^2 = h_{c0c}(t) + r(t)\left(\big|d_{\lambda(t)} h_I[\vec{h}_c]\big|^2 - \big\langle h_I(t), d_{\lambda(t)} h_I[\vec{h}_c]\big\rangle^2\right), \tag{B.3}$$

where every function and vector field is evaluated at $\lambda(t)$. Notice that

$$\big\langle h_I(t), d_{\lambda(t)} h_I[\vec{h}_c]\big\rangle = d_{\lambda(t)} h_c[\vec{h}_c] = 0.$$

Since $dh_I[\vec{h}_c] = h_{cI}$, by applying (6.11) we get

$$d^2_{\lambda(t)}\big(H_S(t, \cdot) - H(u(t), \cdot)\big)[\vec{h}_c]^2 = h_{c0c}(t) + r(t)\left(\big|d_{\lambda(t)} h_I[\vec{h}_c]\big|^2\right) = \sigma_{\lambda_0}\big(Z_I(t), \dot{Z}_I(t)\big),$$

and the equality between (B.2) and (B.3) follows.

We consider now the case $(\eta_1, \eta_2) = (Z_I(t), (\Phi_t^{-1})_* X)$. We have that

$$\big\langle \sigma_{\lambda_0}\big(\mathcal{Z}_t \cdot, (\Phi_t^{-1})_* X\big), l_t^{-1} \sigma_{\lambda_0}\big(\mathcal{Z}_t \cdot, Z_I(t)\big)\big\rangle = \Big\langle \sigma_{\lambda_0}\big(\mathcal{Z}_t \cdot, (\Phi_t^{-1})_* X\big), \begin{pmatrix} 0 \\ 1 \end{pmatrix}\Big\rangle = -\sigma_{\lambda_0}\big(\dot{Z}_I(t), (\Phi_t^{-1})_* X\big) = \tag{B.4}$$

$$= \sigma_{\lambda_0}\left((\Phi_t^{-1})_* X \, , \, (\Phi_t^{-1})_* [\vec{h}_{0c} + r(t) \sum_{i,j=1}^m h_i(t) h_{ij}(t) \vec{h}_j]\right) =$$

$$= r(t) \sum_{i,j=1}^m h_i(t) h_{ij}(t) \, d_{\lambda(t)} h_j[X] =$$

$$= r(t)\langle d_{\lambda(t)} h_I[\vec{h}_c], d_{\lambda(t)} h_I[X]\rangle,$$

37

where we used the simplification $\sigma_{\lambda(t)}(X, \vec{h}_{0c}) = \langle d_{\lambda(t)}h_{0c}, X\rangle = 0$, since $X \in T_{\lambda(t)}\mathcal{S} = \ker d_{\lambda(t)}h_{0c}$.

For the left-hand side of (B.1) we note that, since $h_S(l) = h_0(l)$ for all $l \in \mathcal{S}$, we have

$$d^2_{\lambda(t)}(h_S - h_0)[X, X] = 0.$$

Also, we have that

$$d^2_{\lambda(t)}(h_S - h_0)[\vec{h}_c(\lambda(t)), X] = 0.$$

This follows because h_S, by its definition, is constant along the flow of $\vec{h}_c$, hence

$$d^2_{\lambda(t)}h_S[\vec{h}_c(\lambda(t)), X] = d_{\lambda(t)}(d_{\lambda(t)}h_S[\vec{h}_c])[X] = 0.$$

On the other hand, we have $d^2_{\lambda(t)}h_0[\vec{h}_c(\lambda(t)), X] = d_{\lambda(t)}h_{0c}[X] = 0$. Moreover, it holds that

$$\langle h_I(t), d_{\lambda(t)}h_I[\vec{h}_c]\rangle = \{h_c, h_c\}(t) = 0.$$

Hence, we have

$$d^2_{\lambda(t)}(H_S(t, \cdot) - H(u(t), \cdot))[\vec{h}_c(\lambda(t)), X] = r(t)\langle d_{\lambda(t)}h_I[\vec{h}_c], d_{\lambda(t)}h_I[X]\rangle$$
$$= \left\langle \sigma_{\lambda_0}(\mathcal{Z}_t\cdot, (\Phi_t^{-1})_*X), l_t^{-1}\sigma_{\lambda_0}(\mathcal{Z}_t\cdot, Z_I(t))\right\rangle.$$

So, the equality between l.h.s. and r.h.s. of (4.7) in the case $(\eta_1, \eta_2) = (Z_I(t), (\Phi_t^{-1})_*X)$ follows.

It remains the case $\eta_1 = \eta_2 = (\Phi_t^{-1})_*X \in (\Phi_t^{-1})_*(T_{\lambda(t)}\mathcal{S})$.

Claim B.1. *It holds that*

$$\sigma_{\lambda_0}(\mathcal{Z}_t\cdot, (\Phi_t^{-1})_*X) = l_t\begin{pmatrix} w(t) \\ 0 \end{pmatrix},$$

where $w(t) = r(t)d_{\lambda(t)}h_I[X]$.

Indeed, we have that

$$l_t\begin{pmatrix} w(t) \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{\langle \cdot, w(t)\rangle}{r(t)} \\ \sigma_{\lambda_0}(Z_t w(t), Z_I(t)) \end{pmatrix}.$$

So, in order to prove our claim, we have to check that the following two equalities hold true:

$$\sigma_{\lambda_0}(Z_t\cdot, (\Phi_t^{-1})_*X) = \langle \cdot, d_{\lambda(t)}h_I[X]\rangle, \tag{B.5}$$

$$-\sigma_{\lambda_0}(\dot{Z}_I(t), (\Phi_t^{-1})_*X) = \sigma_{\lambda_0}\left(Z_t[r(t)d_{\lambda(t)}h_I[X]], Z_I(t)\right). \tag{B.6}$$

Concerning the first one, we have that, for any $y \in \mathbb{R}^m$:

$$\sigma_{\lambda_0}(Z_t y, (\Phi_t^{-1})_*X) = \sum_j y_j \sigma_{\lambda_0}\left((\Phi_t^{-1})_*\vec{h}_j, (\Phi_t^{-1})_*X\right) = \sum_j y_j d_{\lambda(t)}h_j[X] = \langle y, d_{\lambda(t)}h_I[X]\rangle.$$

So, (B.5) is proved. For the second one, we have already seen (see (B.4)) that the l.h.s. is

$$\sigma_{\lambda_0}((\Phi_t^{-1})_*X, \dot{Z}_I(t)) = r(t)\langle d_{\lambda(t)}h_I[\vec{h}_c], d_{\lambda(t)}h_I[X]\rangle.$$

38

While the r.h.s. is

$$\sigma_{\lambda_0}\Big(Z_t\big[r(t)d_{\lambda(t)}h_I[X]\big], Z_I(t)\Big) =$$

$$= r(t)\sigma_{\lambda_0}\Big(\sum_{i=1}^{m} d_{\lambda(t)}h_i[X](\Phi_t^{-1})_*\vec{h}_i, \sum_{j=1}^{m} h_j(t)(\Phi_t^{-1})_*\vec{h}_j\Big) =$$

$$= r(t)\sum_{i,j=1}^{m} d_{\lambda(t)}h_i[X]h_j(t)h_{ij}(t) =$$

$$= r(t)\big\langle d_{\lambda(t)}h_I[\vec{h}_c]\,,\, d_{\lambda(t)}h_I[X]\big\rangle.$$

So, also (B.6) follows and the claim is proved.

Now, we are ready to prove (4.7) for $\eta_1 = \eta_2 = (\Phi_t^{-1})_*X \in (\Phi_t^{-1})_*(T_{\lambda(t)}\mathcal{S})$. As a consequence of the Claim that we have just proved, the r.h.s. of (4.7) reduces to

$$\Big\langle \sigma_{\lambda_0}\big(\mathcal{Z}_t\cdot, (\Phi_t^{-1})_*X\big), l_t^{-1}\sigma_{\lambda_0}\big(\mathcal{Z}_t\cdot, (\Phi_t^{-1})_*X\big)\Big\rangle = r(t)\sigma_{\lambda_0}\Big(Z_t\big[d_{\lambda(t)}h_I[X]\big], (\Phi_t^{-1})_*X\Big),$$

and a similar computation to the one above shows that (we write $w(t) = r(t)d_{\lambda(t)}h_I[X]$ for brevity)

$$\sigma_{\lambda_0}\Big(Z_t w(t), (\Phi_t^{-1})_*X\Big) = \sum_{j=1}^{m} w_j(t)\sigma_{\lambda_0}\Big((\Phi_t^{-1})_*\vec{h}_j, (\Phi_t^{-1})_*X\Big) =$$

$$= \sum_{j=1}^{m} w_j(t)\langle d_{\lambda(t)}h_j, X\rangle = \langle w(t), d_{\lambda(t)}h_I[X]\rangle =$$

$$= r(t)\big|d_{\lambda(t)}h_I[X]\big|^2.$$

Now, the l.h.s. of (4.7) is

$$d^2\big(H_S(t,\cdot) - H(u(t),\cdot)\big)[X,X] = r(t)\Big(\big|d_{\lambda(t)}h_I[X]\big|^2 - \big\langle h_I, d_{\lambda(t)}h_I[X]\big\rangle^2\Big),$$

Since $d_{\lambda(t)}h_I[X]$ must be an admissible variation $w(t)$ orthogonal to $h_I(t)$, the term $\big\langle h_I(t), d_{\lambda(t)}h_I[X]\big\rangle$ in the previous expression vanishes. Thus, the desired equality follows.

References

[1] A. A. Agrachev. "Spectrum of the second variation". In: *Tr. Mat. Inst. Steklova* 304 (2019), pp. 32–48. ISSN: 0371-9685,3034-1809. DOI: 10.4213/tm3960. URL: https://doi.org/10.4213/tm3960.

[2] Andrei Agrachev, Stefano Baranzini, and Ivan Beschastnyi. "Index theorems for graph-parametrized optimal control problems". In: *Nonlinearity* 36.5 (2023), pp. 2792–2838. ISSN: 0951-7715,1361-6544. DOI: 10.1088/1361-6544/acc5d4. URL: https://doi.org/10.1088/1361-6544/acc5d4.

[3] Andrei Agrachev and Ivan Beschastnyi. "Jacobi fields in optimal control: Morse and Maslov indices". In: *Nonlinear Analysis* 214 (2022), p. 112608. ISSN: 0362-546X. DOI: https://doi.org/10.1016/j.na.2021.112608. URL: https://www.sciencedirect.com/science/article/pii/S0362546X21002248.

[4] Andrei Agrachev and Ivan Beschastnyi. "Jacobi Fields in Optimal Control: One-dimensional Variations". In: *Journal of Dynamical and Control Systems* (2020).

[5] Andrei A. Agrachev and Carolina Biolo. "Optimality of Broken Extremals". In: *Journal of Dynamical and Control Systems* 25.2 (Aug. 2018), pp. 289–307. ISSN: 1573-8698. DOI: 10.1007/s10883-018-9416-9. URL: http://dx.doi.org/10.1007/s10883-018-9416-9.

[6] 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. DOI: 10.1137/16M110304X. eprint: https://doi.org/10.1137/16M110304X. URL: https://doi.org/10.1137/16M110304X.

[7] 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.

[8] 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. DOI: 10.1137/S036301290138866X. URL: https://doi.org/10.1137/S036301290138866X.

[9] 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. DOI: 10.3934/naco.2012.2.511. URL: https://doi.org/10.3934/naco.2012.2.511.

[10] Stefano Baranzini. "Operators arising as second variation of optimal control problems and their spectral asymptotics". In: *J. Dyn. Control Syst.* 29.3 (2023), pp. 659–689. ISSN: 1079-2724,1573-8698. DOI: 10.1007/s10883-022-09605-1. URL: https://doi.org/10.1007/s10883-022-09605-1.

[11] 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. DOI: 10.1371/journal.pcbi.1000194. URL: https://doi.org/10.1371/journal.pcbi.1000194.

[12] Bernard Bonnard, Jean-Baptiste Caillau, 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. DOI: 10.1051/cocv:2007012.

[13] Ana Cannas da Silva. *Introduction to symplectic and Hamiltonian geometry.* Publicações Matemáticas do IMPA. [IMPA Mathematical Publications]. Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2003, pp. x+130. ISBN: 85-244-0195-8. DOI: 10.1007/978-3-0348-8071-8. URL: https://doi.org/10.1007/978-3-0348-8071-8.

[14] Z. Chen, J.-B. Caillau, and Y. Chitour. "L^1-Minimization for Mechanical Systems". In: *SIAM Journal on Control and Optimization* (2016).

[15] 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. DOI: 10.1051/cocv/2015026. URL: https://doi.org/10.1051/cocv/2015026.

[16] 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. DOI: 10.1051/cocv/2020023. URL: https://doi.org/10.1051/cocv/2020023.

[17] A. A. Agrachev D. Barilari U. Boscain. *A comprehensive introduction to sub-Riemannian geometry.* Cambridge University Press, 2019.

[18] 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. DOI: 10.1134/S0081543810010098. URL: https://doi.org/10.1134/S0081543810010098.

[19] 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. DOI: 10.1137/S0363012902402578. URL: https://doi.org/10.1137/S0363012902402578.

[20] 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.

[21] 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. DOI: 10.1007/s10883-011-9127-y. URL: https://doi.org/10.1007/s10883-011-9127-y.

[22] 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. DOI: 10.1016/j.sysconle.2004.05.005. URL: https://doi.org/10.1016/j.sysconle.2004.05.005.

[23] 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).

[24] 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. DOI: 10.1142/9789812776075_0018. URL: https://doi.org/10.1142/9789812776075_0018.

[25] 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.