Control and Cybernetics

vol. **39** (2010) No. 4

Singular extremals in multi-input time-optimal problems: a sufficient condition*

by

Francesca C. Chittaro[1] **and Gianna Stefani**[2]

[1] LSS-Supélec, Gif-sur-Yvette, France
[2] Dipartimento di Matematica Applicata, Firenze, Italy

Abstract: In this paper we study second order sufficient conditions for the strong-local optimality of singular Pontryagin extremals. In particular, we focus on the minimum-time problem for a control-affine system with vector inputs. We use Hamiltonian methods to prove that the coercivity of a suitably-defined second variation - plus an involutivity assumption on the distribution of the controlled fields - is a sufficient condition for the strong optimality of a candidate extremal.

Keywords: second variation, singular extremal, sufficient condition, Hamiltonian methods.

1. Introduction

We consider the minimum-time problem for a multi-input control affine system on a smooth n-dimensional manifold M, namely we study the problem

$$\min T \tag{1}$$

subject to

$$\begin{cases} \dot{\xi} = (f_0 + \sum_{i=1}^m u_i f_i) \circ \xi(t) \\ \xi(0) = \widehat{x}_0, \quad \xi(T) = \widehat{x}_f \\ \mathbf{u} = (u_1, \ldots, u_m) \in U \subset \mathbb{R}^m \end{cases} . \tag{2}$$

where $f_i,\ i = 0, \ldots, m$ are smooth vector fields on M; the points $\widehat{x}_0,\ \widehat{x}_f \in M$ are fixed. We remark that for smooth we mean C^∞, although the result holds true for C^2 data.

We study the strong-local optimality of a reference triple $\left(\widehat{\xi}, \widehat{T}, \widehat{\mathbf{u}}\right)$ that satisfies the control system (2), and such that $\widehat{\mathbf{u}} \in int\, U$. Strong optimality means

*Submitted: October 2009; Accepted: June 2010.

that the reference triple is optimal with respect to "neighbouring trajectories", independently on the values of control. We consider a type of optimality local with respect to both state and final time, i.e. according to the following definition:

DEFINITION *An admissible trajectory* $\widehat{\xi} : [0, \widehat{T}] \to M$ *is strongly locally optimal if it is optimal with respect to a neighbourhood of the graph of* $\widehat{\xi}$ *in* $\mathbb{R} \times M$.

Pontryagin Maximum Principle establishes a first-order necessary condition to be fulfilled by the reference triple. The aim of this paper is to give second-order *sufficient* conditions for a totally-singular state-extremal $\widehat{\xi}$ to be a strong-local minimiser: that is, the reference triple satisfies PMP and the reference control takes values in the interior of U.

Sufficient second order conditions for weak and Pontryagin minima in the singular case can be found in Dmitruk (1977, 1983, 2008) and references therein. For a single-input control-affine system, the strong-local optimality of a Mayer problem is studied in Stefani (2008), while the strong optimality for the minimum-time problem is studied in Stefani (2004), see also Poggiolini and Stefani (2008 and 2009).

A classical approach to second order conditions is to consider the so-called *second variation*, i.e. an accessory linear-quadratic control problem. See for example Hestenes (1966), Páles and Zeidan (1994) for a classical formulation, and Agrachev et al. (1998a), Agrachev and Sachkov (2004) for an intrinsic version which can be also used when the systems evolve on a manifold.

For the case under study, both the classical second variation and the intrinsic version are totally degenerate; we then require the coercivity of a suitable extended second variation, obtained starting from the coordinate-free second variation defined in Agrachev et al. (1998a) and applying an intrinsic version of the so-called Goh transformation (Goh, 1966; Dmitruk, 2008), in the spirit of Stefani (2004, 2008).

We prove the result under the further assumption that the controlled vector fields $f_1, \ldots, f_m$ generate an involutive distribution (see Subsection 2.2 for the precise definition).

We consider this result as a first step to understand strong-local optimality of singular trajectory in the multi-input case. It is the opinion of the authors that the Hamiltonian approach is particularly effective in studying strong optimality; in fact, it consists in lifting singular trajectories to the cotangent bundle (independently of the values of the associate control) and to use the lifted trajectories to compare the costs. In the standard theory, the trajectory to be lifted belongs to a neighbourhood of the reference trajectory constituted by a field of non-intersecting state-extremals, obtained by projecting suitable solutions of the Hamiltonian system associated to the maximised Hamiltonian F_{max}, see for example Agrachev and Sachkov (2004). When the extremal is singular, F_{max} cannot be used, then we define a Hamiltonian greater than or equal to F_{max}, as suggested by the approach used in Stefani (2004, 2008).

Remark 1 *We do not make any assumption on the control set U. Indeed, U compact convex assures that an optimal solution exists whenever $\widehat{x}_f$ is reachable from $\widehat{x}_0$ by means of the solutions of the control system. However, we do not need the compactness assumption in proving the theorem, and we think that the result may be useful also in case of unbounded controls, in order to find minimising trajectory sequences in the spirit of Jurdjevic (see Jurdjevic, 1997, Chapter 7, and also Remark 15 in Section 7).*

The plan of the paper is the following:

2. Notations and preliminary results: here we recall some basic facts on differential geometry and Hamiltonian formalism, and we state the Pontryagin Maximum Principle.

3. Statement of the results: in this section we state and discuss the main result of the paper; we, moreover, recall the necessary conditions for optimality of a singular extremal and we illustrate the definition and the properties of the main tool we use in this paper, the *second variation*.

4. The Hamiltonian approach: here we illustrate the Hamiltonian approach; we state the sufficient condition for optimality and we define the super-Hamiltonian. We perform all the proofs in the case of $m = 2$ controlled fields.

5. Proof of the Main Theorem: here we complete the proof of Theorem 2.

6. The case with several controls: this section is devoted to the generalisation of the proof of the main result to the case of m controls, $m \leq n-1$.

7. An example: here we illustrate the result with an example.

8. Final remarks: in this last section we give some remarks on the result and on possible developments of research in this field.

Some details of the proofs are the subject of the Appendices.

2. Notations and preliminary results

2.1. Notations

For any vector field f on a manifold M, we indicate by $L_f\varphi(x)$ its action on the smooth function φ, that is $L_f\varphi : x \in M \mapsto \langle d\varphi(x), f(x)\rangle$, where the symbol $\langle\cdot,\cdot\rangle$ denotes the dual action of T_q^*M on T_qM. We recall, moreover, that the Lie bracket of two vector fields f, g on M is the vector field $[f,g]$ that acts on smooth functions on M in the following way:

$$[f,g](\varphi) := L_f\left(L_g(\varphi)\right) - L_g\left(L_f(\varphi)\right), \qquad \varphi \in C^\infty(M).$$

If in coordinates (or in the case $M = \mathbb{R}^n$) we have that if $f(q) = \sum_{i=1}^n f_i(q)\frac{\partial}{\partial x_i}$ and $g(q) = \sum_{i=1}^n g_i(q)\frac{\partial}{\partial x_i}$, then $[f,g](q) = \sum_{i,j=1}^n \left(f_i(q)\frac{\partial g_j}{\partial x_i}(q) - g_i(q)\frac{\partial f_j}{\partial x_i}(q)\right)\frac{\partial}{\partial x_j}$.

For any smooth manifold M, T^*M denotes its cotangent bundle, and $\pi : T^*M \to M$ the canonical projection onto the base manifold. It is well known that the cotangent bundle is a smooth manifold of dimension $2n$; if we put on the manifold local coordinates $\{q_1, \dots, q_n\}$ in a neighbourhood of a point $\boldsymbol{q} \in M$, the 1-forms $\{dq_1(\boldsymbol{q}), \dots, dq_n(\boldsymbol{q})\}$ constitute a basis for the cotangent space $T^*_{\boldsymbol{q}}M$, therefore there are induced local coordinates $\{p_1, \dots, p_n; q_1, \dots, q_n\}$ on T^*M, in the following way: for any 1-form $\omega \in T^*M$,

$$\omega \simeq (p_1, \dots, p_n; x_1, \dots, x_n) \Leftrightarrow \omega = \sum_{i=1}^{n} p_i dq_i(x).$$

Obviously, the action of the projection is $\pi : (p_1, \dots, p_n; q_1, \dots, q_n) \mapsto (q_1, \dots, q_n)$.

In the paper we will largely use the symplectic structure of the cotangent bundle T^*M; it is well known, in fact, that to any smooth manifold a skew-symmetric non-degenerate two form is canonically associated, called the *standard symplectic form*, which is constructed in this way: for any $\ell \in T^*_{\pi\ell}M$, we define the *canonical Liouville form*

$$s_\ell \in T^*_\ell\left(T^*M\right), \quad s_\ell := \ell \circ \pi_*, \qquad \ell \in T^*_{\pi\ell}M,$$

where we recall that $\pi_* : T_\ell(T^*M) \to T_{\pi\ell}M$. In coordinates, we can write $\ell = (p_1, \dots, p_n; q_1, \dots, q_n)$ and then we get that $s_\ell = \sum_{i=1}^n p_i dq_i$.

σ denotes the standard symplectic form $\sigma_\ell = ds_\ell$, and possesses the coordinates expression

$$\sigma_\ell = \sum_{i=1}^{n} dp_i \wedge dq_i;$$

we can prove that it is a skew-symmetric non-degenerate 2-form, thus it endows the cotangent bundle with a symplectic structure. Given the two vectors in $T_\ell(T^*M)$ written in coordinates as $X = \sum_{i=1}^n X_q^i \frac{\partial}{\partial q_i} + X_p^i \frac{\partial}{\partial p_i}$ and $Y = \sum_{i=1}^n Y_q^i \frac{\partial}{\partial q_i} + Y_p^i \frac{\partial}{\partial p_i}$, the symplectic form acts as

$$\sigma_\ell\left(X, Y\right) = \sum_{i=1}^{n} X_p^i Y_q^i - X_q^i Y_p^i.$$

We recall, moreover, that an n-dimensional subspace V is said to be *Lagrangian* if the symplectic form vanishes on it, $\sigma\big|_V = 0$. Moreover, an n-dimensional submanifold Λ is called Lagrangian if $T_\ell\Lambda$ is Lagrangian for any $\ell \in \Lambda$.

A smooth function $F : T^*M \to \mathbb{R}$ is called a *Hamiltonian* on T^*M; we recall that its associated Hamiltonian vector field $\vec{F}$ is defined by

$$\sigma_\ell(\cdot, \vec{F}) = dF(\ell), \qquad \ell \in T^*M;$$

in coordinates, $\vec{F}(\ell) = \sum_{i=1}^n \frac{\partial F}{\partial p_i}\frac{\partial}{\partial q_i} - \frac{\partial F}{\partial q_i}\frac{\partial}{\partial p_i}$.

Given two smooth functions $F, G : T^*M \to \mathbb{R}$, their Poisson bracket is defined as

$$\{F, G\}(\ell) = \sigma_\ell(\vec{F}, \vec{G});$$

a straight computation shows that $\{F,G\}(\ell) = \sum_{i=1}^{n} \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i} - \frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i}$.

For any vector field f on M, the lifted Hamiltonian is defined as $F(\ell) = \langle \ell, f(\pi\ell)\rangle,\ \ell \in T^*_{\pi\ell}M$; if f has the coordinate expression $f = \sum_{i=1}^{n} f^i \frac{\partial}{\partial q_i}$ and $\ell = (p_1,\dots,p_n; q_1,\,\dots,q_n)$, then $F(\ell) = \sum_{i=1}^{n} p_i f^i$.

In the following, we will use the notation: $F_{ij} = \{F_i, F_j\}$ and $F_{ijk} = \{F_i,\{F_j,F_k\}\},\ i,j,k \in \{0,\dots,m\}$, and it turns out that

$$F_{ij}(\ell) = \langle \ell, [f_i,f_j](\pi\ell)\rangle, \qquad F_{ijk} = \langle \ell, [f_i,[f_j,f_k]](\pi\ell)\rangle.$$

Finally, we recall that a Hamiltonian vector field $\vec{F}$ defines a flow via the differential equation

$$\dot{\ell}(t) = \vec{F} \circ \ell(t);$$

which is the Hamiltonian system associated to F. We will use the *script* typesetting to denote the flow generated by the Hamiltonian vector fields, e.g. $\mathcal{F}_t$ (or $\mathcal{H}_t$) denotes the flow generated by $\vec{F}$ (or $\vec{H}$) from the time 0 to t. We recall that, for any function $F : T^*M \to \mathbb{R}$ and any Hamiltonian H, we have

$$\frac{d}{dt} F \circ \exp(t\vec{H}) = \frac{d}{dt} F \circ \mathcal{H}_t = \{H,F\} \circ \mathcal{H}_t.$$

2.2. Involutive distributions

In this subsection we recall some facts on the properties of the vector distribution we are going to use. As a reference, we cite the textbooks of Abraham and Marsden (1978) and Lee (2006).

First of all, let us recall that a smooth vector distribution D of dimension k is a k-dimensional sub-bundle of the tangent bundle, $D \subset TM$. If we put $D_q = D \cap T_qM,\ q \in M$, by definition D_q is a k-dimensional subspace of the tangent space T_qM, and the subspaces D_q vary smoothly with respect to $q \in M$.

If D is a smooth distribution of dimension k, then we can locally find a *local frame* for the distribution; this means that for any $q \in M$ there are a neighbourhood U of q and k smooth vector fields $X_1,\dots,X_k : U \to TM$ such that $\{X_1(q),\dots,X_k(q)\}$ is a basis of D_q for any $q \in U$. In this situation, we say that the distribution D is *(locally) spanned* by the vector fields $X_1,\dots,X_k$.

A distribution is said to be *involutive* if for any two smooth local sections of D (i.e. two locally defined vector fields $X, Y : U \to TM$ such that $X(q), Y(q) \in D_q$ for any $q \in U$), their Lie bracket is also a local section contained in D; in other words, if $\{X_1,\dots,X_k\}$ is a local frame for D, then for any pair Y, Z of vector fields in D we can locally find k smooth functions $\alpha_i,\ i = 1,\dots,k$, such that

$$[Y,Z](q) = \sum_k \alpha_k(q) X_k(q).$$

Remark 2 *We remark that, given an involutive distribution D locally spanned by the vectors $\{X_1, \dots, X_k\}$, it is always possible to find locally some functions $\alpha_i^j : M \to \mathbb{R}$, $i, j = 1, \dots, k$, such that the vector fields*

$$\phi_i(q) := \sum_{j=1}^{k} \alpha_i^j(q) X_j(q), \qquad i = 1, \dots, k,$$

form a commuting *local frame.*

It is a well-known consequence of the Frobenius Theorem (see Abraham and Marsden, 1978; Lee, 2006) that if D is an involutive distribution on M, then M is foliated by *integral manifolds* of D: that is, if the distribution is involutive, then for any $q \in M$ there is an immersed k-dimensional submanifold N such that $q \in N$ and $T_{q'}N = D_{q'}$ for any $q' \in N$.

From now on, we assume that the controlled vector fields $f_1, \dots, f_m$ span an m-dimensional distribution, which will be denoted by $\mathcal{D}$.

Proposition 1 *These three conditions are equivalent:*

1. *$\mathcal{D}$ is involutive on M;*
2. *the set $\left\{\vec{F}_1, \dots, \vec{F}_m\right\}$ is involutive on the set $\{F_1 = \dots = F_m = 0\} \subset T^*M$;*
3. *for any $i, j = 1, \dots, m$, $F_{ij} = 0$ on the set $\{F_1 = \dots = F_m = 0\}$.*

Proof. Let us prove that (1) $\Rightarrow$ (2). Fix $i, j = 1, \dots, m$ and write, locally, $[f_i, f_j](q) = \sum_{k=1}^m \alpha_k(q) f_k(q)$; then $F_{ij}(\ell) = \langle \ell, [f_i, f_j](q)\rangle = \sum_{k=1}^m \alpha_k(\pi\ell) F_k(\ell)$, and

$$dF_{ij}(\ell) = \sum_{k=1}^{m} \alpha_k(\pi\ell) dF_k(\ell) + \sum_{k=1}^{m} F_k(\ell) d\alpha_k(\pi\ell);$$

on the set $\{F_1 = \dots = F_m = 0\}$, this expression reduces to

$$dF_{ij}(\ell) = \sum_{k=1}^{m} \alpha_k(\pi\ell) dF_k(\ell),$$

and therefore to

$$[\vec{F}_i, \vec{F}_j](\ell) = \vec{F}_{ij}(\ell) = \sum_{k=1}^{m} \alpha_k(\pi\ell) \vec{F}_k(\ell).$$

Now we prove that (2) $\Rightarrow$ (3). For $\ell \in \{F_1 = \dots = F_m = 0\}$, we have that $[\vec{F}_i, \vec{F}_j](\ell) = \sum_{k=1}^m \alpha_k(\ell) \vec{F}_k(\ell)$ for some functions α_k. Since $[f_i, f_j](\pi\ell) = \pi_*[\vec{F}_i, \vec{F}_j] = \sum_{k=1}^m \alpha_k(\ell) f_k(\pi\ell)$, we are done.

Condition (1) obviously implies (3). By contradiction, assume that condition (3) is satisfied, but there is a pair of indices i, j such that $[f_i, f_j](q) \notin \mathcal{D}_q$; in other words, there is an ℓ belonging to the orthogonal complement of $\mathcal{D}_q$ in T_q^*M such that $\langle \ell, f_{ij}(q)\rangle \neq 0$. The statement is proved by noting that such an orthogonal complement is given by $\{F_1 = \dots = F_m = 0\} \cap T_q^*M$. ■

2.3. Pontryagin Maximum Principle

Let M be a smooth manifold, and let us consider the time-optimal problem subject to the control system (2). Let us recall that all triples $(\xi, T, \mathbf{u})$, with $\xi : [0,T] \to M$ and $\mathbf{u} : [0,T] \to U$, that satisfy the control system are called *admissible triples*, and the trajectories ξ *admissible trajectories*.

We assume that there is an admissible triple $(\widehat{\xi}, \widehat{T}, \widehat{\mathbf{u}})$, which will be referred to as the *reference triple*, that satisfies the Pontryagin Maximum Principle.

We recall Pontryagin Maximum Principle stated in its Hamiltonian form (see for example Agrachev and Sachkov, 2004). We consider the following Hamiltonian functions: the control-dependent Hamiltonian

$$\begin{aligned} h(\ell, u) &= \langle \ell, f_0(\pi\ell)\rangle + \sum_{i=1}^m u_i \langle \ell, f_i(\pi\ell)\rangle = \\ &= F_0(\ell) + \sum_{i=1}^m u_i F_i(\ell), \qquad \ell \in T^*M, \quad \mathbf{u} \in U, \end{aligned}$$

the (time-dependent) *reference Hamiltonian* $\widehat{F}_t : T^*M \to \mathbb{R},\ t \in [0, \widehat{T}]$,

$$\widehat{F}_t(\ell) := F_0(\ell) + \sum_{i=1}^m \widehat{u}_i(t) F_i(\ell),$$

and the *maximised Hamiltonian*

$$F_{max}(\ell) := \sup_{\mathbf{u}\in U} h(\ell, u).$$

$\overrightarrow{\widehat{F}}_t$ denotes the (non-autonomous) Hamiltonian vector field associated to $\widehat{F}_t$, and $\ell \mapsto \widehat{\mathcal{F}}_t(\ell)$ the solution at the time t of the Hamiltonian system generated by $\overrightarrow{\widehat{F}}_t(\ell)$, with initial condition $\widehat{\mathcal{F}}_0(\ell) = \ell$.

The statement of the Pontryagin Maximum Principle is the following

Theorem 1 (PMP) *If the triple $(\widehat{\xi}, \widehat{T}, \widehat{\mathbf{u}})$ is optimal, then there exist a constant $p_0 \geq 0$ and a Lipschitzian curve in the cotangent bundle*

$$t \mapsto \widehat{\lambda}(t) \in T^*M, \qquad t \in [0, \widehat{T}]$$

such that

$$\pi \circ \widehat{\lambda}(t) = \widehat{\xi}(t) \qquad t \in [0, \widehat{T}] \tag{3}$$

$$\widehat{\lambda}(t) \neq 0 \qquad t \in [0, \widehat{T}] \tag{4}$$

$$\frac{d}{dt}\widehat{\lambda}(t) = \overrightarrow{\widehat{F}}_t \circ \widehat{\lambda}(t) \tag{5}$$

$$\widehat{F}_t(\widehat{\lambda}(t)) = \max_{u\in U} h(\widehat{\lambda}(t), u) \tag{6}$$

$$\widehat{F}_t(\widehat{\lambda}(t)) \equiv p_0. \tag{7}$$

The curve $\widehat{\lambda}$ is called the *Pontryagin extremal* associated to the admissible triple $(\widehat{\xi}, \widehat{T}, \widehat{\mathbf{u}})$, and its projection onto the base manifold a *state extremal*. When $M = \mathbb{R}^n$, $\widehat{\lambda} = (\widehat{\mu}, \widehat{\xi})$, where $\widehat{\mu} : [0, \widehat{T}] \to \mathbb{R}^{n*}$ is called *adjoint covector* and satisfies the adjoint equation $\dot{\widehat{\mu}}(t) = -\langle \widehat{\mu}(t), D(f_0 + \sum_{i=1}^m \widehat{u}_i(t) f_i)(\widehat{\xi}(t)) \rangle$.

An extremal satisfying PMP is called *normal* if $p_0 > 0$, while *abnormal* if $p_0 = 0$.

We assume that the reference extremal $\widehat{\lambda}$ is a *normal* Pontryagin extremal, and we normalise the term p_0 putting $p_0 = 1$. Since the reference control $\widehat{\mathbf{u}}$ belongs to the *interior* of the control set U, then by (6) we get

$$F_i(\widehat{\lambda}(t)) = 0 \qquad \text{for } t \in [0, \widehat{T}], \quad i = 1, \dots, m. \tag{8}$$

In the following, we will call $\widehat{\ell}_0 = \widehat{\lambda}(0)$ the initial point of the extremal.

3. Statement of the results

3.1. Necessary conditions

The study of second order condition for optimality provides additional necessary optimality conditions for singular extremals, which are well known in literature (see e.g. Goh, 1966; Gabasov and Kirillova, 1972, and the textbook by Agrachev and Sachkov, 2004). In this Subsection we are focusing on them.

First of all, as a direct consequence of PMP and the singularity condition (equation (8)), we see that:

$$0 = \frac{d}{dt} F_i \circ \widehat{\lambda}(t) = \{\widehat{F}_t, F_i\}(\widehat{\lambda}(t)) = F_{0i}(\widehat{\lambda}(t)) + \sum_{j=1}^m \widehat{u}_j(t) F_{ji}(\widehat{\lambda}(t)) \quad i = 1, \dots, m. \tag{9}$$

Moreover, standard theory of singular extremals states the following *necessary* conditions for the triple $(\widehat{\lambda}, \widehat{T}, \widehat{\mathbf{u}})$ to be optimal (see Goh, 1966; Gabasov and Kirillova, 1972; Agrachev and Sachkov, 2004):

Goh condition

$$\left\{ \frac{\partial h}{\partial u_i}, \frac{\partial h}{\partial u_j} \right\} (\widehat{\lambda}(t), u) \Big|_{u=\widehat{u}(t)} = F_{ij} \circ \widehat{\lambda}(t) = 0, \quad i, j = 1, \dots, m, \quad t \in [0, \widehat{T}]; \tag{10}$$

Generalised Legendre Condition (GLC), the quadratic form

$$\begin{aligned} &(v_1, \dots, v_m) \mapsto \left\{ \left\{ h, \sum_{i=1}^m v_i \frac{\partial h}{\partial u_i} \right\}, \sum_{j=1}^m v_j \frac{\partial h}{\partial u_j} \right\} (\widehat{\lambda}(t), u) \Big|_{u=\widehat{u}(t)} = \\ &= \left(\sum_{i,j=1}^m v_i v_j F_{ij0}(\widehat{\lambda}(t)) + \sum_{i,j,k=1}^m v_i v_j \widehat{u}_k F_{ijk}(\widehat{\lambda}(t)) \right) \leq 0, \\ &v \in \mathbb{R}^m, \quad t \in [0, \widehat{T}]. \end{aligned} \tag{11}$$

Lemma 1 *If the distribution $\mathcal{D}$ is involutive, then Goh condition is automatically satisfied and the generalised Legendre quadratic form appearing in equation* (11) *reduces to*

$$\mathbb{L}_\ell : v = (v_1, \ldots, v_m) \mapsto \mathbb{L}_\ell[v]^2 = \sum_{i,j=1}^{m} v_i v_j F_{ij0}(\ell) \tag{12}$$

for any $\ell \in \{F_i = 0, F_{ij} = 0,\ i, j = 1, \ldots, m\}$.

Proof. By the involutivity of the distribution, the Lie brackets $[f_i, f_j]$ and $[f_i, [f_j, f_k]]$ belong to the distribution $\mathcal{D}$ for any i, j, k, therefore, by equation (8), we get that the F_{ij} vanish along the reference extremal. ■

In the following we will use $\mathbb{L}_\ell$ to indicate both the quadratic form and its associated matrix.

Summarising, in our case the necessary conditions for the singular extremals are

- $F_i \circ \widehat{\lambda} \equiv 0$, $i = 1, \ldots, m$;
- $\mathbb{L}_{\widehat{\lambda}(t)}[v]^2 \leq 0$ for $t \in [0, \widehat{T}]$.

3.2. The extended second variation

In this section, we define a suitable second variation for the problem under study, following the ideas forwarded in Agrachev et al. (1998a), and further developed in subsequent works (see Agrachev et al., 2002; Poggiolini and Stefani, 2004; Stefani, 2004, 2008).

For $t \in [0, \widehat{T}]$, we define the evolution function $\widehat{S}_t : M \to M$ by the action $\widehat{S}_t : x_0 \mapsto \xi(t)$, where ξ satisfies the equation $\dot{\xi} = f_0(\xi) + \sum_{i=1}^{m} \widehat{u}_i f_i(\xi)$ with $\xi(0) = x_0$ (in particular, $\widehat{S}_t(\widehat{x}_0) = \widehat{\xi}(t)$), and the *pull-back fields*

$$g_t^i := \widehat{S}_{t*}^{-1} f_i \circ \widehat{S}_t : V(\widehat{x}_0) \to TM, \quad i = 1, \ldots, m, \quad t \in [0, \widehat{T}],$$

where $V(\widehat{x}_0)$ is a neighbourhood of $\widehat{x}_0$. In coordinates,

$$g_t^i(x) = \left[D\widehat{S}_t(x)\right]^{-1} f_i(\widehat{S}_t(x)).$$

We choose, moreover, a function $\widehat{\beta} : M \to \mathbb{R}$ that satisfies the following equality:

$$d\widehat{\beta}(\widehat{x}_0) = -\widehat{\ell}_0. \tag{13}$$

As in Stefani (2004), to compute the second variation we reduce the minimum-time problem to a Mayer problem, in which we take the final time as a new variable and also the cost.

Consider the Mayer problem on the fixed time interval $[0, \widehat{T}]$:

$$\min T(\widehat{T})$$

subject to

$$\begin{cases} \dot{T} = 0 \\ \dot{\xi} = \frac{T}{\widehat{T}}\left(f_0 + \sum_{i=1}^m v_i f_i\right) \circ \xi(t) \\ \xi(0) = \widehat{x}_0, \quad \xi(\widehat{T}) = \widehat{x}_f, \qquad T\ free \\ \mathbf{u} = (u_1, \ldots, u_m) \in U \subset \mathbb{R}^m \end{cases}.$$

If $(\widehat{\xi}, \widehat{T}, \widehat{\mathbf{u}})$ is the candidate optimal triple of the original minimum-time problem, then for the Mayer problem the reference control is still $\widehat{\mathbf{u}}$ and the state-extremal is $(\widehat{T}, \widehat{\xi})$, with the associate Pontryagin extremal $\left(\left(-\frac{t}{\widehat{T}}, \widehat{T}\right), \widehat{\lambda}\right)$, where $\widehat{\lambda}$ is the normal Pontryagin extremal associated to $\widehat{\xi}$.

Evaluating the second variation of the Mayer problem, as defined in Agrachev et al. (1998a) we obtain, thanks to its special form, the second variation, defined by

$$J''[\delta u]^2 = \int_0^{\widehat{T}} \sum_{i=1}^m \delta u_i L_{\delta\eta} L_{g_t^i}(\widehat{\beta}(\widehat{x}_0))\, dt, \tag{14}$$

where $\delta u \in L^2([0, \widehat{T}], \mathbb{R}^m)$ and $\delta\eta(t) \in T_{\widehat{x}_0}M$ satisfy the following system:

$$\begin{cases} \dot{\delta\eta} = \sum_{i=1}^m \delta u_i(t) g_t^i(\widehat{x}_0) \\ \delta\eta(0) = \delta\eta(\widehat{T}) = 0. \end{cases} \tag{15}$$

We remark that (15) is the linearisation of the system satisfied by $\eta(t) = \widehat{S}_t^{-1}(\xi(t))$; in the linearised system, $\delta T = 0$, since $\widehat{\lambda}$ is normal.

We underline that, if the problem is stated on $\mathbb{R}^n$, then the second variation defined in (14)-(15) and expressed without the pull-back system reduces to the classical one, as noted in Agrachev et al. (1998a), Corollary 2.

Remark 3 *If $\delta\eta$ satisfies the system* (15)*, then the expression for the second variation does not depend on the particular choice of $\widehat{\beta}$ with the property* (13) *(see Agrachev et al., 1998a). Then J'' is well-defined and coordinate free.*

We now perform an integration by parts, that can be regarded as an intrinsic version of a Goh transformation (see Goh, 1966; Dmitruk, 2008), to transform the singular second variation into a non-singular one, which is coordinate-free, too.

Define for $i = 1, \ldots, m$

$$w_i(t) := \int_t^{\widehat{T}} \delta u_i(s)\, ds, \qquad w_0^i = w_i(0), \tag{16}$$

and the new variable $\varphi : [0, \widehat{T}] \times M \mapsto \varphi_x(t) \in T_x M$ as the solution of

$$\dot{\varphi}_x(t) = \sum_{i=1}^{m} \delta u_i(t) g_t^i(x) \qquad \varphi_x(\widehat{T}) = 0.$$

In this way, the control variation δu is represented by the pair $(w_0, w(\cdot))$. Integrating by parts equation (14), we get

$$\begin{aligned} J''[\delta u]^2 &= -\int_0^{\widehat{T}} L_{\dot{\varphi}_x(t)} L_{\varphi_x(t)} \widehat{\beta}(\widehat{x}_0)\, dt \\ &= -\frac{1}{2} \int_0^{\widehat{T}} L_{[\dot{\varphi}_x(t), \varphi_x(t)]} \widehat{\beta}(\widehat{x}_0)\, dt. \end{aligned} \tag{17}$$

We now put $Z_x(t) \in T_x M,\ t \in [0, \widehat{T}]$, as $Z_x(t) := \varphi_x(t) + \sum_{i=1}^m w_i(t)$ $g_t^i(x)$. Substituting into (17) and integrating by parts, we get

$$\begin{aligned} J''[\delta u]^2 = \sum_{i,j=1}^{m} &\left[\frac{1}{2} w_0^i w_0^j L_{f_i} L_{f_j} \widehat{\beta}(\widehat{x}_0) + \frac{1}{2} \int_0^{\widehat{T}} w_i(t) w_j(t) L_{[\dot{g}_t^i, g_t^j]} \widehat{\beta}(\widehat{x}_0)\, dt \right] + \\ &+ \sum_{i=1}^{m} \int_0^{\widehat{T}} w_i(t) L_{\zeta(t)} L_{g_t^i} \widehat{\beta}(\widehat{x}_0)\, dt. \end{aligned}$$

where $\zeta(t) \in T_{\widehat{x}_0} M$ is defined as $\zeta(t) := Z_{\widehat{x}_0}(t)$. It is easy to see that $\zeta(t)$ satisfies the equation

$$\dot{\zeta}(t) = \sum_{i=1}^{m} w_i(t) \dot{g}_t^i(\widehat{x}_0), \tag{18}$$

and the boundary conditions

$$\zeta(0) = \sum_{i=1}^{m} w_0^i f_i(\widehat{x}_0), \qquad \zeta(\widehat{T}) = 0. \tag{19}$$

This last expression is the second variation of the linear-quadratic problem for the state variable ζ; we can write it as:

$$\begin{aligned} J''[\delta u]^2 = \sum_{i,j=1}^{m} \frac{1}{2} w_0^i w_0^j L_{f_i} L_{f_j} \widehat{\beta}(\widehat{x}_0) + \frac{1}{2} \int_0^{\widehat{T}} &\sum_{i,j=1}^{m} w_i(t) w_j(t) R_{ij}(t) + \\ + 2 \sum_{i=1}^{m} w_i(t) Q_i(t) \zeta(t)\, dt& \end{aligned} \tag{20}$$

subject to (19), where

$$R_{ij}(t) = L_{[\dot{g}_t^i, g_t^j]} \widehat{\beta}(\widehat{x}_0) = -(\mathbb{L}_{\widehat{\lambda}(t)})_{ij} \qquad Q_i(t) = L_{(\cdot)} L_{\dot{g}_t^i} \widehat{\beta}(\widehat{x}_0).$$

Actually, the functional (20) is defined only on those pairs (w_0, w) related with a control variation δu via equation (16). We extend the second variation to the whole $\mathbb{R}^m \times L^2([0,\widehat{T}],\mathbb{R}^m)$, in order to require coercivity. This is possible because the functional is continuous and the map $\delta u \mapsto (w_0, w)$ stated in (16) from $L^2([0,\widehat{T}],\mathbb{R}^m)$ to $\mathbb{R}^m \times L^2([0,\widehat{T}],\mathbb{R}^m)$ is continuous and has dense image.

Then, from now on, we will consider the *extendend second variation*, which is the second variation J'' extended by continuity on $\mathbb{R}^m \times L^2([0,\widehat{T}],\mathbb{R}^m)$. We will refer to it as $J''_{\rm ext}$, i.e. $J''_{\rm ext}[(w_0,w)]^2$ is defined as the right-hand side of (20). We remark that $J''_{\rm ext}$ is defined on the whole $\mathbb{R}^m \times L^2([0,\widehat{T}],\mathbb{R}^m)$, it is invariant only on the subspace:

$$\widetilde{\mathcal{W}} := \{(w_0,w) \in \mathbb{R}^m \times L^2([0,\widehat{T}],\mathbb{R}^m) \text{ that satisfy equations } (18)-(19)\}, \quad (21)$$

while its extension on the whole space depends on the choice of $\widehat{\beta}$.

In the following, when we will speak about coercivity of $J''_{\rm ext}$ we will mean on $\widetilde{\mathcal{W}}$, with respect to the norm induced by $\mathbb{R}^m \times L^2([0,\widehat{T}],\mathbb{R}^m)$.

Remark 4 *From the above formula, we see that the coercivity of the second variation implies that the Legendre quadratic form* (12) *is negative definite, i.e. that there is an* $\alpha > 0$ *such that:*

$$\mathbb{L}_{\widehat{\lambda}(t)}[v]^2 \leq -\alpha|v|^2, \quad v \in \mathbb{R}^m. \quad (22)$$

Equation (22) *is known as Strengthened Generalised Legendre Condition, or SGLC (see Agrachev and Sachkov, 2004).*

Remark 5 *It is not difficult to see that the coercivity of* J''_{ext} *on* $\widetilde{\mathcal{W}}$ *implies that the controlled vector fields are linearly independent at* $\widehat{x}_0$.

Remark 6 *We remark that, thanks to the linear independence of the controlled vector fields at* $\widehat{x}_0$, $\mathbb{R}^m$ *is isomorphic to* $V = \operatorname{span}\{f_1(\widehat{x}_0), \ldots, f_m(\widehat{x}_0)\} \subset T_{\widehat{x}_0}M$, *therefore the accessory problem associated to* J''_{ext} *is a standard one defined on the finite-dimensional space* $T_{\widehat{x}_0}M$. *Sufficient conditions for the coercivity of such quadratic forms in Hamiltonian setting can be found in Stefani and Zezza (1997).*

3.3. The result

In this section we state our main result and we give the main ideas for the proof in the Hamiltonian setting.

Theorem 2 *Let* $\widehat{\xi}$ *be a totally-singular state-extremal with associate normal Pontryagin extremal* $\widehat{\lambda}$ *for the minimum time problem* (1)-(2). *Assume that the distribution* $\mathcal{D}$ *is involutive. Let the extendend second variation* J''_{ext}, *as defined in Subsection 3.2, be coercive on* $\widetilde{\mathcal{W}}$ *(see* (21)*).*

Then $\widehat{\xi}$ *is a strict strong-local minimiser, according to Definition 1.*

Remark 7 *We remark here that the hypotheses of Theorem 2 imply that the necessary conditions are automatically satisfied by the reference extremal: in particular, the involutivity of the distribution $\mathcal{D}$ implies Goh condition (see Lemma 1), and the coercivity of the second variation implies SGLC.*

A standard technique to prove the optimality of a candidate extremal is a generalisation of the method of *fields of extremals* of the Calculus of Variations (see Giaquinta and Hildebrandt, 1996); for its use in optimal control see, for instance, Agrachev and Sachkov (2004). Briefly, this technique usually consists in covering a neighbourhood of the candidate trajectory with non-intersecting state-extremals, i.e. with trajectories on M that are the projection of the solutions of the Hamiltonian system associated to the maximised Hamiltonian F_{max} emanating from a suitable Lagrangian submanifold. If it is possible to invert the projection and lift to the cotangent bundle the admissible trajectories of the control problem, we can compare the costs evaluated on them (see for instance Agrachev et al., 1998b, 2002, and references therein).

In the case of singular extremals, the Hamiltonian vector field is multi-valued: indeed, all the Hamiltonians of the form $F_0 + \sum_{i=1}^m u_i F_i$, $\mathbf{u} \in U$, coincide and realise the maximum along the singular extremal. Moreover, no selection of such multi-valued Hamiltonian vector fields is suitable to construct the field of non-intersecting state-extremals.

To overcome this problem, we define a Hamiltonian H_0 such that $F_0 \leq H_0$ on the set

$$\Sigma = \{\ell \in T^*M : F_1(\ell) = \cdots = F_m(\ell) = 0\},$$

in such a way that we can find a suitable selection K^S of the new multi-valued super-Hamiltonian $H_0 + \sum_{i=1}^m u_i F_i$, such that the reference extremal is a solution of the Hamiltonian system associated to K^S, and such that the corresponding vector field is smooth and tangent to Σ. In this way, we can apply the method used in Agrachev and Sachkov (2004), Agrachev et al. (1998a,b). This super-Hamiltonian is obtained in the spirit of Stefani (2004, 2008).

4. The Hamiltonian approach

In this section we apply the Hamiltonian approach to sufficient condition, and use it to the problem under consideration. We perform our construction in the case of $m = 2$; the case with general m is a straight generalisation, and will be illustrated in Section 6.

We remark that in this section we assume the following regularity assumptions:

- the distribution $\mathcal{D}$ spanned by the controlled vector fields is involutive;
- Strengthened Generalised Legendre condition holds along the reference extremal, therefore it holds by continuity in a (full-measure) neighbourhood $\mathcal{U}$ of the extremal;

we recall that these conditions are already included in the hypotheses of Theorem 2.

4.1. Preliminary discussion

In this subsection, we study the geometry of the problem in the neighbourhood $\mathcal{U}$ of the reference extremal.

We recall that

$$\Sigma = \{\ell \in T^*M : F_1(\ell) = F_2(\ell) = 0\},$$

and we define the subset of Σ

$$\mathcal{S} = \{\ell \in T^*M : F_1(\ell) = F_2(\ell) = F_{01}(\ell) = F_{02}(\ell) = 0\} \cap \mathcal{U} \tag{23}$$

(i.e., the set where the Hamiltonians $F_1,\ F_2,\ F_{01},\ F_{02}$ vanish and SGLC holds) and we further notice that the reference extremal $\widehat{\lambda}(t) \in \mathcal{S}$ for $t \in [0, \widehat{T}]$. We remark that $\mathcal{S}$ contains all the singular extremals of the control system under consideration, which satisfy SGLC.

LEMMA 2 *In the neighbourhood $\mathcal{U}$, where SGLC is satisfied, the following statements hold:*

1. *$\vec{F}_1$ and $\vec{F}_2$ are tangent to Σ and linearly independent;*
2. *$\vec{F}_1$ and $\vec{F}_2$ are transversal to $\mathcal{S}$;*
3. *$\vec{F}_{01}$ and $\vec{F}_{02}$ are transversal to Σ (and therefore to $\mathcal{S}$), and the vectors $\{\vec{F}_1, \vec{F}_2, \vec{F}_{01}, \vec{F}_{02}\}$ are linearly independent;*
4. *$\mathcal{S}$ is a symplectic submanifold of dimension $2n-4$ contained in Σ.*

Proof. First of all, notice that:

$$T_\ell\Sigma = \ker(dF_1(\ell)) \cap \ker(dF_2(\ell)) \qquad \ell \in \Sigma$$
$$T_\ell\mathcal{S} = \ker(dF_1(\ell)) \cap \ker(dF_2(\ell)) \cap \ker(dF_{01}(\ell)) \cap \ker(dF_{02}(\ell)) \qquad \ell \in \mathcal{S}.$$

Since $\langle dF_i, \vec{F}_i \rangle \equiv 0,\ i = 1, 2$, and $\langle dF_1, \vec{F}_2 \rangle = -\langle dF_2, \vec{F}_1 \rangle = F_{12} = 0$ on Σ, we can deduce that $\vec{F}_1$ and $\vec{F}_2$ are tangent to Σ. They are linearly independent by SGLC; in fact, assume without loss of generality that for some $\ell \in \mathcal{U}$, $\vec{F}_1(\ell) = -\mu\vec{F}_2(\ell)$. Then we have that

$$F_{110}(\ell) = \sigma_\ell\left(\vec{F}_1, \vec{F}_{10}\right) = -\mu\ \sigma_\ell\left(\vec{F}_2, \vec{F}_{10}\right) = -\mu F_{210}(\ell)$$
$$F_{120}(\ell) = \sigma_\ell\left(\vec{F}_1, \vec{F}_{20}\right) = -\mu\ \sigma_\ell\left(\vec{F}_2, \vec{F}_{20}\right) = -\mu F_{220}(\ell),$$

and then the matrix $\mathbb{L}_\ell$ is degenerate, which contradicts SGLC.

(2) follows by the fact that $\langle dF_{01}, \vec{F}_1 \rangle = -F_{110}$ and $\langle dF_{02}, \vec{F}_2 \rangle = -F_{220}$ are nonvanishing on $\mathcal{S}$ (by SGLC).

The same fact implies (3); linear independence is again a consequence of SGLC (the proof is analogous to the one above).

To complete the proof, we notice that the restriction $\sigma\Big|_{T_\ell \mathcal{S}}$, $\ell \in \mathcal{S}$, is nondegenerate; then $\mathcal{S}$ is a symplectic submanifold. ■

REMARK 8 *Since we proved that the vector fields $\vec{F}_1$, $\vec{F}_2$ are linearly independent in $\mathcal{U}$, we call $\widetilde{\mathcal{D}}$ the distribution (defined on $\mathcal{U}$) spanned by them. We notice that $\widetilde{\mathcal{D}}$ is involutive on Σ (see Proposition 1).*

LEMMA 3 *There is a neighbourhood $\mathcal{V}$ of the range of $\widehat{\lambda}$ in Σ such that for any $\ell \in \mathcal{V}$ there is a unique triple $(\ell_{\mathcal{S}}, t_1, t_2) \in (\mathcal{S} \cap \mathcal{V}) \times \mathbb{R}^2$ such that*

$$\ell = \psi(\ell_{\mathcal{S}}, t_1, t_2) := \exp\left(-t_2\vec{F}_2\right) \circ \exp\left(-t_1\vec{F}_1\right)(\ell_{\mathcal{S}}); \tag{24}$$

moreover, there exists an $\epsilon > 0$ such that the map $\psi : (\mathcal{S} \cap \mathcal{V}) \times [-\epsilon, \epsilon]^2 \to \mathcal{V}$ is a global diffeomorphism.

Proof. The whole proof is an easy consequence of Lemma 2, point (3), and the compactness of the interval $[0, \widehat{T}]$. ■

REMARK 9 *We can use the same argument to prove that the map $\widetilde{\psi} : \Sigma \times \mathbb{R}^2 \to T^*M$ defined as $\widetilde{\psi}(\ell_\Sigma, \tau_1, \tau_2) = \exp(-\tau_2\vec{F}_{02}) \circ \exp(-\tau_1\vec{F}_{01})(\ell_\Sigma)$ is a local diffeomorphism; therefore, there are an $\epsilon' > 0$ and a neighbourhood $\widetilde{\mathcal{V}}$ in T^*M of the reference extremal such that $\widetilde{\psi}$ is a* global *diffeomorphism from $\mathcal{V} \times [-\epsilon', \epsilon']^2$ to $\widetilde{\mathcal{V}}$.*

Without loss of generality, we can assume that $\mathcal{U} = \widetilde{\mathcal{V}}$.

LEMMA 4 *Under SGLC, every singular Pontryagin extremal belongs to $\mathcal{S}$ and is a Hamiltonian trajectory of the* feed-back Hamiltonian

$$F^S(\ell) = F_0(\ell) + \nu_1(\ell)F_1(\ell) + \nu_2(\ell)F_2(\ell), \quad \ell \in \mathcal{U}, \tag{25}$$

where the feed-back controls *are defined by*

$$\begin{pmatrix} \nu_1(\ell) \\ \nu_2(\ell) \end{pmatrix} = \mathbb{L}_\ell^{-1} \begin{pmatrix} F_{001}(\ell) \\ F_{002}(\ell) \end{pmatrix}, \quad \ell \in \mathcal{S}, \tag{26}$$

and then extended constant to the whole $\mathcal{U}$. With this definition, we have that $L_{\vec{F}_i}\nu_j = 0,\ i, j = 1, 2$.

In particular, $\widehat{u}_i(t) = \nu_i(\widehat{\lambda}(t)),\ i = 1, 2,\ t \in [0, \widehat{T}]$, which proves that $\widehat{\mathbf{u}}$ is smooth.

Proof. We define the feed-back controls on $\mathcal{S}$ as in (26); then, we choose two linearly independent commuting vector fields X_1, X_2 on $\mathcal{U}$ that span the distribution $\widetilde{\mathcal{D}}$, in such a way that

$$T_\ell\Sigma = T_\ell\mathcal{S} \oplus \mathbb{R}X_1(\ell) \oplus \mathbb{R}X_2(\ell), \qquad \ell \in \mathcal{S};$$

we can extend $\nu_i,\ i = 1, 2$, to $\mathcal{V}$ by putting it constant along the integral lines of X_1 and X_2.

Analogously, we choose two linearly independent commuting vector fields $Y_1,\ Y_2$ on $\mathcal{U}$ such that

$$T_\ell(T^*M) = T_\ell\Sigma \oplus \mathbb{R}Y_1(\ell) \oplus \mathbb{R}Y_2(\ell), \qquad \ell \in \mathcal{V},$$

and we extend $\nu_i,\ i = 1, 2$, to $\mathcal{U}$ by putting it constant along the integral lines of Y_1 and Y_2. The rest of the proof is an easy consequence of the singularity conditions (9). ∎

Remark 10 *The matrix $\mathbb{L}$ is symmetric on $\mathcal{S}$, in other words, $F_{120} = F_{210}$. In fact, if $\ell \in \mathcal{S}$, then, for small t, $\exp(t\vec{F}^S)(\ell) \in \mathcal{S}$, where $\vec{F}^S$ is the vector field associated to the feedback Hamiltonian defined in equation* (25)*; this happens since $\vec{F}^S$ is tangent to $\mathcal{S}$, by definition of ν.*

Then we have that

$$0 = \frac{d}{dt}F_{12}\left(\exp(t\vec{F}^S)(\ell)\right) = \{F^S, F_{12}\} = F_{012},$$

since F_{112} and F_{212} vanish on $\mathcal{S}$. By the Jacobi identity, $F_{120} = F_{210}$ on $\mathcal{S}$.

4.2. Construction of the modified Hamiltonian

Lemma 5 *Possibly restricting $\mathcal{V}$, we can define functions ϑ_1 and ϑ_2 in such a way that*

$$\exp(\vartheta_1(\ell)\vec{F}_1) \circ \exp(\vartheta_2(\ell)\vec{F}_2)(\ell) \in \mathcal{S} \qquad \forall\ \ell \in \mathcal{V} \subset \Sigma. \tag{27}$$

For any $\ell \in \mathcal{S} \cap \mathcal{V}$ we have that

$$\begin{pmatrix} D\vartheta_1(\ell) \\ D\vartheta_2(\ell) \end{pmatrix}[\delta\ell] = \mathbb{L}_\ell^{-1}\begin{pmatrix} DF_{01}(\ell) \\ DF_{02}(\ell) \end{pmatrix}(\ell)\,[\delta\ell] \qquad \forall\ \delta\ell \in T_\ell\Sigma. \tag{28}$$

Extending $\vartheta_1,\ \vartheta_2$ constant to the whole $\mathcal{U}$, we obtain:

$$D\vartheta_1(\ell)\,[\delta\ell] = D\vartheta_2(\ell)\,[\delta\ell] = 0 \qquad \forall\ \delta\ell \in \mathbb{R}\vec{F}_{01}(\ell) \oplus \mathbb{R}\vec{F}_{02}(\ell), \quad \ell \in \mathcal{S}. \tag{29}$$

Proof. Since $\vec{F}_1$ and $\vec{F}_2$ are tangent to Σ, then equation (27) is satisfied if and only if

$$F_{01}(\exp(\vartheta_1(\ell)\vec{F}_1) \circ \exp(\vartheta_2(\ell)\vec{F}_2)(\ell)) = 0 \tag{30}$$

$$F_{02}(\exp(\vartheta_1(\ell)\vec{F}_1) \circ \exp(\vartheta_2(\ell)\vec{F}_2)(\ell)) = 0 \tag{31}$$

and

$$\vartheta_1(\ell) = \vartheta_2(\ell) = 0 \qquad \forall\ \ell \in \mathcal{S}.$$

We then consider the function $\Phi : \mathcal{V} \times \mathbb{R}^2 \to \mathbb{R}^2$, defined as

$$\Phi(\ell, t_1, t_2) = \begin{pmatrix} F_{01}(\exp(t_1\vec{F}_1) \circ \exp(t_2\vec{F}_2)(\ell)) \\ F_{02}(\exp(t_1\vec{F}_1) \circ \exp(t_2\vec{F}_2)(\ell)) \end{pmatrix}.$$

We notice that $D_{(t_1,t_2)}\Phi\Big|_{(\ell,0,0)} = -\mathbb{L}_\ell$, which is a non-degenerate matrix in a neighbourhood of the reference trajectory (by SGLC). Therefore, we can apply the Implicit Function Theorem to define (locally in a neighbourhood of the reference trajectory in Σ) the functions $\vartheta_1(\ell)$, $\vartheta_2(\ell)$ that satisfy (30) and (31), plus the boundary conditions. We can extend the two functions ϑ_1 and ϑ_2 to $\mathcal{U}$ with the same technique that we used in Lemma 4.

Let $\ell \in \mathcal{V}$ and $\delta\ell \in T_\ell\Sigma$; from equations (30) and (31) we obtain

$$0 \equiv D\left[F_{01} \circ \exp(\vartheta_1(\ell)\vec{F}_1) \circ \exp(\vartheta_2(\ell)\vec{F}_2)(\ell)\right][\delta\ell]$$
$$0 \equiv D\left[F_{02} \circ \exp(\vartheta_1(\ell)\vec{F}_1) \circ \exp(\vartheta_2(\ell)\vec{F}_2)(\ell)\right][\delta\ell]\,;$$

after long but straight computations, for $\ell_{\mathcal{S}} \in \mathcal{S}$ we have that

$$0 \equiv [DF_{01}(\ell_{\mathcal{S}}) + F_{101}(\ell_{\mathcal{S}})D\vartheta_1(\ell_{\mathcal{S}}) + F_{201}(\ell_{\mathcal{S}})D\vartheta_2(\ell_{\mathcal{S}})]\,[\delta\ell]$$
$$0 \equiv [DF_{02}(\ell_{\mathcal{S}}) + F_{102}(\ell_{\mathcal{S}})D\vartheta_1(\ell_{\mathcal{S}}) + F_{202}(\ell_{\mathcal{S}})D\vartheta_2(\ell_{\mathcal{S}})]\,[\delta\ell]$$

and hence equation (28). Equation (29) comes from the fact that we extended ϑ_1, ϑ_2 constant. ■

Remark 11 *By the Implicit Function Theorem it follows that $\vartheta_i \equiv 0$, $i = 1, 2$, on $\mathcal{S}$, and hence $D\vartheta_i\Big|_{T_\ell\mathcal{S}} = 0$, $i = 1, 2$.*

We now define a new smooth Hamiltonian H_0 on $\mathcal{U}$ that will allow us to use the Hamiltonian approach. We remark that we are interested only in the values of H_0 on Σ, and that they are obtained by transporting $F_0\Big|_{\mathcal{S}}$ along suitable trajectories tangent to Σ.

Definition *We define the Hamiltonian $H_0 : \mathcal{U} \to \mathbb{R}$ as*

$$H_0(\ell) = F_0 \circ \exp(\vartheta_1(\ell)\vec{F}_1) \circ \exp(\vartheta_2(\ell)\vec{F}_2)(\ell), \qquad \ell \in \mathcal{U}. \tag{32}$$

We define, moreover, the function $\chi : \mathcal{U} \to \mathbb{R}$ as

$$\chi(\ell) = H_0(\ell) - F_0(\ell) = F_0 \circ \exp(\vartheta_1(\ell)\vec{F}_1) \circ \exp(\vartheta_2(\ell)\vec{F}_2)(\ell) - F_0(\ell). \tag{33}$$

Theorem 3 *Let H_0 and χ be the functions defined above. Then:*

1. *For $\ell \in \Sigma$, we have*

$$\vec{H}_0(\ell) = \Big(\exp(-\vartheta_2(\ell)\vec{F}_2) \circ \exp(-\vartheta_1(\ell)\vec{F}_1)\Big)_* \vec{F}_0$$
$$\circ \Big(\exp(\vartheta_1(\ell)\vec{F}_1) \circ \exp(\vartheta_2(\ell)\vec{F}_2)\Big)(\ell).$$

2. The vector field $\vec{H}_0$ is tangent to Σ.
3. For $\ell \in \mathcal{S}$, $D\chi(\ell) = 0$, hence $D^2\chi(\ell)$ is well defined and a non-negative quadratic form on $T_\ell\Sigma$ whose kernel is $T_\ell\mathcal{S}$.
4. $F_0 \le H_0$ on Σ, and $F_0 = H_0$ on $\mathcal{S}$.

Proof. We use the following notation:

$$\Theta(\ell) = \exp(\vartheta_1(\ell)\vec{F}_1) \circ \exp(\vartheta_2(\ell)\vec{F}_2)(\ell), \quad \ell \in \Sigma,$$

in order to make the computations more comprehensible.

In order to prove (1), it is sufficient to prove that

$$dH_0 = dF_0(\Theta(\ell)) \circ \exp(\vartheta_1(\ell)\vec{F}_1)_* \circ \exp(\vartheta_2(\ell)\vec{F}_2)_* \qquad \forall\, \ell \in \Sigma. \tag{34}$$

In fact, calling $X\big|_\ell$ a generic vector in $T_\ell(T^*M)$, we have that

$$\begin{aligned}\sigma\left(X\big|_\ell, \vec{H}_0(\ell)\right) &= \langle dH_0(\ell), X\big|_\ell\rangle = \\ &= \langle dF_0(\Theta(\ell)), \exp(\vartheta_1(\ell)\vec{F}_1)_* \circ \exp(\vartheta_2(\ell)\vec{F}_2)_*(X\big|_\ell)\rangle = \\ &= \sigma\left(\exp(\vartheta_1(\ell)\vec{F}_1)_* \circ \exp(\vartheta_2(\ell)\vec{F}_2)_*(X\big|_\ell), \vec{F}_0\big|_{\Theta(\ell)}\right) = \\ &= \sigma\left(X\big|_\ell, \exp(-\vartheta_2(\ell)\vec{F}_2)_* \circ \exp(-\vartheta_1(\ell)\vec{F}_1)_*(\vec{F}_0\big|_{\Theta(\ell)})\right).\end{aligned}$$

By nondegeneracy of the symplectic form and genericity of X, we got (1).

To prove (34), consider

$$\begin{aligned}dH_0(\ell) = dF_0(\Theta(\ell)) &\circ \exp(\vartheta_1(\ell)\vec{F}_1)_*\big|_{\exp(\vartheta_2(\ell)\vec{F}_2)(\ell)} \circ \exp(\vartheta_2(\ell)\vec{F}_2)_*\big|_\ell + \\ &+ \langle dF_0(\Theta(\ell)), \vec{F}_1(\Theta(\ell))\rangle d\vartheta_1(\ell) + \\ &+ \langle dF_0(\Theta(\ell)), \exp(\vartheta_1(\ell)\vec{F}_1)_*\left[\vec{F}_2(\exp(\vartheta_2(\ell)\vec{F}_2)(\ell))\right]\rangle d\vartheta_2(\ell).\end{aligned}$$

The distribution $\widetilde{\mathcal{D}}$ being involutive, $\exp(\vartheta_1(\ell)\vec{F}_1)_*\vec{F}_2(\exp(\vartheta_2(\ell)\vec{F}_2)(\ell))$ is tangent to the distribution, that is, there are $\alpha,\ \beta \in \mathbb{R}$ such that

$$\exp(\vartheta_1(\ell)\vec{F}_1)_*\left[\vec{F}_2(\exp(\vartheta_2(\ell)\vec{F}_2)(\ell))\right] = \alpha\vec{F}_1(\Theta(\ell)) + \beta\vec{F}_2(\Theta(\ell)).$$

Therefore we can conclude that

$$\begin{aligned}dH_0(\ell) = dF_0(\Theta(\ell)) &\circ \exp(\vartheta_1(\ell)\vec{F}_1)_*\big|_{\exp(\vartheta_2(\ell)\vec{F}_2)(\ell)} \circ \exp(\vartheta_2(\ell)\vec{F}_2)_*\big|_\ell + \\ &- F_{01}(\Theta(\ell))d\vartheta_1(\ell) - (\alpha F_{01}(\Theta(\ell)) + \beta F_{02}(\Theta(\ell)))\, d\vartheta_2(\ell) = \\ &= dF_0(\Theta(\ell)) \circ \exp(\vartheta_1(\ell)\vec{F}_1)_*\big|_{\exp(\vartheta_2(\ell)\vec{F}_2)(\ell)} \circ \exp(\vartheta_2(\ell)\vec{F}_2)_*\big|_\ell.\end{aligned}$$

Thesis (2) comes straightforwardly from (1), since we have that $\vec{F}_0$ is tangent to Σ on $\mathcal{S}$ (by straight computations), therefore $\vec{F}_0\big|_{\Theta(\ell)} \in T_{\Theta(\ell)}\Sigma$, and

$$\exp(-\vartheta_2(\ell)\vec{F}_2)_* \circ \exp(-\vartheta_1(\ell)\vec{F}_1)_* : T_{\Theta(\ell)}\Sigma \to T_\ell\Sigma.$$

Point (3) can be proved by means of straight but long computations, that can be found in Appendix A, Lemma A.2.

A straight consequence of (3) is that χ has a minimum on $\mathcal{S}$; since $\chi\big|_{\mathcal{S}} = 0$, we get that $\chi \geq 0$ and hence the thesis of point (4). ∎

REMARK 12 *We remark here that* $D^2\chi(\ell)$ *coincides with* $-\mathbb{L}_\ell$ *on* $\widetilde{\mathcal{D}}$, *in the following sense: for any* $X \in T_\ell\Sigma$, $\ell \in \mathcal{S}$, *we can write* $X = X_{\mathcal{S}} + \alpha\vec{F}_1(\ell) + \beta\vec{F}_2(\ell)$, $X_{\mathcal{S}} \in T_\ell\mathcal{S}$, *and then we have that*

$$D^2\chi(\ell)[X]^2 = -\mathbb{L}_\ell[(\alpha,\beta)]^2$$

(see Appendix A for the proof).

COROLLARY 1 (SEE STEFANI, 2008) *Let* $\boldsymbol{v}_t = (v_t^1, v_t^2) : [0,\widehat{T}] \times \mathcal{U} \to \mathbb{R}^2$ *be a function such that* $v_t^i(\widehat{\lambda}(t)) = \widehat{u}_i(t)$, $t \in [0,\widehat{T}]$, $i = 1,2$. *Consider the Hamiltonian* $H^{\boldsymbol{v}_t} = F_0 + v_t^1 F_1 + v_t^2 F_2 + \chi$. *Then*

1. $H^{\boldsymbol{v}_t} \geq F_0$ *on* Σ.
2. $\overrightarrow{H^{\boldsymbol{v}_t}}$ *is tangent to* Σ.
3. $\widehat{\lambda}$ *is a trajectory of the Hamiltonian system associated to* $\overrightarrow{H^{\boldsymbol{v}_t}}$.

In particular, these facts hold for the Hamiltonians $H_t = \widehat{F}_t + \chi = H_0 + \widehat{u}_1(t)F_1 + \widehat{u}_2(t)F_2$ *and* $K^S = F^S + \chi = H_0 + \nu_1 F_1 + \nu_2 F_2$.

REMARK 13 *The problem is symmetric, in the sense that Proposition 3 - and therefore the whole result - holds also for the super-Hamiltonian*

$$\widetilde{H}_0(\ell) = F_0 \circ \exp(\widetilde{\vartheta}_2(\ell)\vec{F}_2) \circ \exp(\widetilde{\vartheta}_1(\ell)\vec{F}_1)(\ell),$$

where $\widetilde{\vartheta}_1$, $\widetilde{\vartheta}_2$ *are suitably defined by*

$$F_{01}(\exp(\widetilde{\vartheta}_2(\ell)\vec{F}_2) \circ \exp(\widetilde{\vartheta}_1(\ell)\vec{F}_1)(\ell)) = 0$$
$$F_{02}(\exp(\widetilde{\vartheta}_2(\ell)\vec{F}_2) \circ \exp(\widetilde{\vartheta}_1(\ell)\vec{F}_1)(\ell)) = 0 \qquad \forall\, \ell \in \mathcal{V}$$

and

$$\widetilde{\vartheta}_1(\ell) = \widetilde{\vartheta}_2(\ell) = 0 \qquad \forall\, \ell \in \mathcal{S}.$$

4.3. The sufficient condition

As already said, we intend to prove the optimality of the extremal using a method that generalises the method of Fields of Extremals of Calculus of Variations. We state and prove this method in this section.

We recall that we denote with K^S the Hamiltonian

$$K^S(\ell) = H_0(\ell) + \nu_1(\ell)F_1(\ell) + \nu_2(\ell)F_2(\ell), \qquad \ell \in \mathcal{U}. \tag{35}$$

Theorem 4 *Let $\widehat{\lambda}$ be a normal singular Pontryagin extremal for the minimum-time problem. Let K^S be the Hamiltonian defined in equation* (35)*, and let $\mathcal{K}_t^S$ denote its flow.*

If there exists a Lagrangian submanifold $\Lambda \subset \Sigma$ such that $\widehat{\ell}_0 \in \Lambda$ and

$$\ker \pi_* \mathcal{K}_{t*}^S \cap T_{\widehat{\ell}_0}\Lambda = \{0\}, \qquad t \in [0, \widehat{T}], \tag{36}$$

then $(\widehat{\xi}, \widehat{T})$ is a strict strong-local minimiser.

Proof. We give the proof by steps:

(i) (36) implies the local invertibility of the map $\pi \circ \mathcal{K}_t^S : \Lambda \to M$, for any t. Thanks to the compactness of the interval $[0, \widehat{T}]$, the map $\mathrm{id} \times \pi \circ \mathcal{K}_t^S : [0, \widehat{T}] \times \Lambda \to [0, \widehat{T}] \times M$ is also a diffeomorphism covering the graph of $\widehat{\xi}$. That is, we can find a neighbourhood $\mathcal{O}$ of $\widehat{\ell}_0$ in Λ and a neighbourhood $\mathfrak{U}$ of the range of $\widehat{\xi}$ in M such that the map

$$(t, \ell) \in [0, \widehat{T}] \times \mathcal{O} \mapsto (t, \pi \circ \mathcal{K}_t^S(\ell)) \in [0, \widehat{T}] \times \mathfrak{U}$$

is invertible, with smooth inverse.

(ii) Recall that s is the canonical Liouville form. Then the 1-form

$$\omega(t, \ell) = \mathcal{K}_t^{S*} s - K^S \circ \mathcal{K}_t^S(\ell)\, dt$$

is *exact* on $[0, \widehat{T}] \times \Lambda$ (see for instance Agrachev and Sachkov, 2004, Section 17.1.1).

(iii) Assume that there exists a solution $(\xi, T, \mathbf{u})$ of system (2) with $T \le \widehat{T}$, whose graph is contained in $\mathcal{U}$.

If we define the following paths in $[0, \widehat{T}] \times \Lambda$:

$$\begin{aligned}
\widehat{\mu}(t) &= (t, \widehat{\ell}_0) \qquad t \in [0, \widehat{T}] \\
\mu(t) &= (t, (\pi \circ \mathcal{K}_t^S)^{-1} \xi(t)) \qquad t \in [0, T] \\
\mu_0(t) &= (t, (\pi \circ \mathcal{K}_t^S)^{-1}(\widehat{x}_f)) \qquad t \in [T, \widehat{T}],
\end{aligned}$$

we get

$$0 = \int_{\mu} \omega + \int_{\mu_0} \omega - \int_{\widehat{\mu}} \omega. \tag{37}$$

We call $\ell(t) = (\pi \circ \mathcal{K}_t^S)^{-1} \xi(t)$ and $\lambda(t) = \mathcal{K}_t^S \circ \ell(t)$; we notice that $\ell(t) \in \Lambda$, $\forall\, t \in [0, T]$, and that $\ell(0) = \widehat{\ell}_0$.

Equation (37) writes as:

$$
\begin{aligned}
0 = &\int_0^T \langle \mathcal{K}_t^S \circ \ell(t), \dot{\xi}(t)\rangle - K^S \circ \mathcal{K}_t^S(\ell(t))\, dt + \int_{\mu_0} \omega + \\
&- \int_0^{\widehat{T}} \langle \widehat{\lambda}(t), f_0(\widehat{\xi}(t)) + \widehat{u}_1(t) f_1(\widehat{\xi}(t)) + \widehat{u}_2(t) f_2(\widehat{\xi}(t))\rangle - K^S(\widehat{\lambda}(t))\, dt \leq \\
&\leq \int_{\mu_0} \omega
\end{aligned}
$$

because

$$
\begin{aligned}
&\langle \mathcal{K}_t^S \circ \ell(t), \dot{\xi}(t)\rangle - K^S \circ \mathcal{K}_t^S(\ell(t)) = h(\mathcal{K}_t^S \circ \ell(t), u_\xi(t)) - K^S \circ \mathcal{K}_t^S(\ell(t)) \leq \\
&\leq h(\mathcal{K}_t^S \circ \ell(t), u_\xi(t)) - F_{max} \circ \mathcal{K}_t^S(\ell(t)) \leq 0,
\end{aligned}
$$

while $\langle \widehat{\lambda}(t), f_0(\widehat{\xi}(t)) + \widehat{u}_1(t) f_1(\widehat{\xi}(t)) + \widehat{u}_2(t) f_2(\widehat{\xi}(t))\rangle = K^S(\widehat{\lambda}(t))$.

Therefore

$$
\begin{aligned}
0 \leq \int_{\mu_0} \omega = - \int_T^{\widehat{T}} K^S \circ \mathcal{K}_t^S \circ (\pi \circ \mathcal{K}_t^S)^{-1}(\widehat{x}_f)\, dt = - \int_T^{\widehat{T}} (1 + O(t))\, dt = \\
= T - \widehat{T} + o(\widehat{T} - T),
\end{aligned}
$$

since $K^S(\widehat{\lambda}(t)) = F_{max}(\widehat{\lambda}(t)) = 1,\ t \in [0, \widehat{T}]$. A contradiction, therefore $T = \widehat{T}$.

Let us now prove that the minimum is strict: assume that there is an admissible curve $\xi(t)$ that satisfies the system (2), with $\xi(0) = \widehat{x}_0$ and $\xi(\widehat{T}) = \widehat{x}_f$.

Define μ, μ_0 and $\widehat{\mu}$ as above; since $T = \widehat{T}$, $\int_{\mu_0} \omega \equiv 0$ and then equality (37) reduces to

$$
\begin{aligned}
0 &= \int_0^{\widehat{T}} \langle \lambda(t), \dot{\xi}(t)\rangle - K^S(\lambda(t))\, dt - \int_0^{\widehat{T}} \langle \widehat{\lambda}(t), \dot{\widehat{\xi}}(t)\rangle - K^S(\widehat{\lambda}(t))\, dt = \\
&= \int_0^{\widehat{T}} \langle \lambda(t), \dot{\xi}(t)\rangle - K^S(\lambda(t))\, dt,
\end{aligned}
$$

which implies that $\langle \lambda(t), \dot{\xi}(t)\rangle - K^S(\lambda(t)) \equiv 0$, that is $\lambda(t) \in \mathcal{S}$, since $\lambda(t) \in \Sigma$.

Let us now compute the derivative of λ:

$$
\begin{aligned}
\dot{\lambda}(t) = &\vec{K}^S \circ \lambda(t) + \mathcal{K}_{t*}^S \left(-\pi \circ \mathcal{K}_t^S\right)_*^{-1} \left(\pi_* \vec{K}_t^S\right) \circ \lambda(t) + \mathcal{K}_{t*}^S \left(\pi \circ \mathcal{K}_t^S\right)_*^{-1} \dot{\xi}(t) = \\
&= \left(\vec{F}_0 + \nu_1 \vec{F}_1 + \nu_2 \vec{F}_2\right) \circ \lambda(t) + \\
&+ \mathcal{K}_{t*}^S \left(\pi \circ \mathcal{K}_t^S\right)_*^{-1} (f_0 + \nu_1 f_1 + \nu_2 f_2 - f_0 - u_1 f_1 - u_2 f_2) \circ \xi(t) = \\
&= \left(\vec{F}_0 + \nu_1 \vec{F}_1 + \nu_2 \vec{F}_2\right) \circ \lambda(t) + \mathcal{K}_{t*}^S \left(\pi \circ \mathcal{K}_t^S\right)_*^{-1} ((\nu_1 - u_1) f_1(\xi(t)) + \\
&\qquad + (\nu_2 - u_2) f_2(\xi(t))) .
\end{aligned}
$$

By Lemma B.2 in Appendix B we have that $\mathcal{K}^S_{t*}\left(\pi\circ\mathcal{K}^S_t\right)^{-1}_* f_i(\xi(t)) = \vec{F}_i(\lambda(t)),\ i=1,2.$

Therefore

$$\dot\lambda(t) = \vec{F}_0(\lambda(t)) + u_1(t)\vec{F}_1(\lambda(t)) + u_2(t)\vec{F}_2(\lambda(t));$$

we notice, moreover, that, since $\lambda\in\mathcal{S}$, the control $(u_1(t),u_2(t))$ satisfies the equation of the feedback control (26), which implies that $\dot\lambda(t) = \vec{K}^S_t(\lambda(t))$, that is

$$\lambda(t) = \mathcal{K}^S_t\circ\widehat{\ell}_0.$$

Then it coincides with $\widehat{\lambda}(t)$. ■

5. Proof of the main theorem

Let us recall that $H_t : \mathcal{U}\to\mathbb{R}$ is defined as

$$H_t = H_0 + \widehat{u}_1(t)F_1 + \widehat{u}_2(t)F_2, \tag{38}$$

and denote with $\vec{H}_t$ and $\mathcal{H}_t$, respectively, its associated Hamiltonian vector field and Hamiltonian flow. In this section we will prove that if the extended second variation is coercive, then it is possible to define a Lagrangian submanifold $\Lambda\subset\Sigma$ such that $\widehat{\ell}_0\in\Lambda$ and

$$\ker\pi_*\mathcal{H}_{t*}\cap T_{\widehat{\ell}_0}\Lambda = 0, \qquad t\in[0,\widehat{T}]. \tag{39}$$

This completes the proof of Theorem 2, thanks to the following Lemma:

Lemma 6 *If* $\ker(\pi_*\mathcal{H}_{t*}\big|_{T_{\widehat{\ell}_0}\Lambda}) = 0$, *then* $\ker(\pi_*\mathcal{K}^S_{t*}\big|_{T_{\widehat{\ell}_0}\Lambda}) = 0$.

The proof can be found in Appendix B.

5.1. Coercivity of the second variation

Let $H''_t : T^*_{\widehat{x}_0}M\times T_{\widehat{x}_0}M\to\mathbb{R}$ be the quadratic Hamiltonian function associated to the second variation, defined in Subsection 5.2, and let $\mathcal{H}''_t$ be its associated linear flow.

In this subsection we prove that the coercivity of J''_{ext} on $\widetilde{\mathcal{W}}$ allows us to add a penalty, so that J''_{ext} is coercive on a larger subspace $\mathcal{W}$, which correspond to a free-fixed problem. We also prove that the coercivity of the second variation on $\mathcal{W}$ is equivalent to the condition

$$\ker\pi_*\mathcal{H}''_{t*}\cap L = 0 \qquad \forall\, t\in[0,\widehat{T}],$$

for a suitably defined subspace $L\subset T^*_{\widehat{x}_0}M\times T_{\widehat{x}_0}M$.

Thanks to the linear independence of the controlled fields at the point $\widehat{x}_0$, we can define suitable local coordinates, in order to get rid of the *finite dimensional* term in equation (20).

Let us in fact define, locally in a neighbourhood of $\widehat{x}_0$ in M, a vector field ϕ_2 such that

$$[f_1, \phi_2] = 0 \quad \text{and} \quad \operatorname{span}\{f_1, \phi_2\} = \mathcal{D};$$

we can then choose coordinates $(q_1, \ldots, q_n)$ on a neighbourhood of $\widehat{x}_0$ in M in such a way that

$$\begin{aligned}
&\widehat{x}_0 = (0, \ldots, 0) \\
&f_1 = \frac{\partial}{\partial q_1} \qquad \phi_2 = \frac{\partial}{\partial q_2};
\end{aligned}$$

therefore, we can locally write

$$f_2 = \mu_1 \frac{\partial}{\partial q_1} + \mu_2 \frac{\partial}{\partial q_2},$$

for two locally defined functions μ_1, μ_2, with $\mu_2(x) \neq 0$ for any x in a neighbourhood of $\widehat{x}_0$.

Then, we have that the covector $\widehat{\ell}_0$ can be written as $\widehat{\ell}_0 = \sum_{i=3}^n \widehat{p}_i dq^i$.

Let us now recall that the choice of $\widehat{\beta}$ is free, provided that $d\widehat{\beta}(\widehat{x}_0) = -\widehat{\ell}_0$; we then choose

$$\widehat{\beta}(x) = \sum_{i=3}^{n} -\widehat{p}_i q_i, \tag{40}$$

for x in a neighbourhood of $\widehat{x}_0$. This choice certainly satisfies the required condition, and, moreover, we have that the term $\sum_{i,j=1}^2 \frac{1}{2} w_0^i w_0^j L_{f_i} L_{f_j} \widehat{\beta}(\widehat{x}_0)$ in the second variation vanishes.

By means of the local coordinates chosen above, we define a local function $\alpha : M \to \mathbb{R}$ as

$$\alpha = \sum_{i=3}^{n} \left(\widehat{p}_i q_i + \rho q_i^2 \right), \quad \rho > 0; \tag{41}$$

with such definition, we have that

$$\begin{aligned}
&d\alpha(\widehat{x}_0) = \sum_{i=3}^{n} \widehat{p}_i dq^i = \widehat{\ell}_0, \\
&L_{f_1}\alpha(y) \equiv L_{f_2}\alpha(y) \equiv 0;
\end{aligned}$$

which implies that $d(\alpha + \widehat{\beta})(\widehat{x}_0) = 0$ and therefore

$$\gamma'' := D^2(\alpha + \widehat{\beta})(\widehat{x}_0)$$

is a well defined quadratic form on $T_{\widehat{x}_0}M$ such that

$$\gamma''[f_1(\widehat{x}_0), T_{\widehat{x}_0}M] = \gamma''[f_2(\widehat{x}_0), T_{\widehat{x}_0}M] = 0.$$

Moreover, by setting $V = \operatorname{span}\left\{\frac{\partial}{\partial q^i},\ i \geq 3\right\}$, we have

$$\gamma''\Big|_V > 0.$$

Remark that $\ker\gamma'' = \mathcal{D}_{\widehat{x}_0}$, and that $T_{\widehat{x}_0}M = V \oplus \mathcal{D}_{\widehat{x}_0}$.

LEMMA 7 *If the extended second variation is coercive, then*

$$\ker \pi_* \mathcal{H}''_{t*} \cap L = 0 \quad \forall\, t \in [0, \widehat{T}],$$

where

$$L = \operatorname{span}\left\{\left(-d\alpha(\widehat{x}_0)\left(\frac{\partial}{\partial q_i}, \cdot\right), \frac{\partial}{\partial q_i}\right),\ i = 1, \ldots, n\right\}.$$

Proof. By standard results in the theory of quadratic forms (Theorems 13.2 and 13.3 in Hestenes, 1951), there is a $\mu > 0$ such that $\frac{1}{2}\mu\gamma'' + J''_{\text{ext}}$ is coercive on the set $\mathcal{W}$ of the pairs

$$\mathcal{W} = \{\delta e = (\delta x, w) \in T_{\widehat{x}_0}M \times L^2([0, \widehat{T}], \mathbb{R}^2)\}$$

such that

$$\begin{aligned} &\dot{\zeta}(t) = w_1(t)\dot{g}^1_t(\widehat{x}_0) + w_2(t)\dot{g}^2_t(\widehat{x}_0) \\ &\zeta(0) = \delta x \quad \zeta(\widehat{T}) = 0. \end{aligned} \tag{42}$$

In other words, $\zeta(0)$ is *free* and $\zeta(\widehat{T})$ is *fixed.* With no loss to generality, we can put $\mu = 1$, so that the modified second variation,

$$J''_\alpha[\delta e]^2 = \frac{1}{2}\gamma''[(\delta x, w)]^2 + \sum_{i,j=1}^{2} \frac{1}{2}\int_0^{\widehat{T}} w_i(t)w_j(t)R_{ij}(t) + 2w_i(t)Q_i(t)\zeta(t)\, dt, \tag{43}$$

is coercive on $\mathcal{W}$.

The coercivity of J''_α on $\mathcal{W}$ is equivalent to

$$\ker \pi_* \mathcal{H}''_{t*}\Big|_L = \{0\},$$

see, for example, Stefani and Zezza (1997). ∎

5.2. Hamiltonian linear flows

In this subsection we study the relation between the quadratic Hamiltonian function $H''_t : T^*_{\widehat{x}_0}M \times T_{\widehat{x}_0}M \to \mathbb{R}$ associated to the second variation, defined in Subsection 3.2, and the linear Hamiltonian flow defined below. For more details, see Agrachev et al. (1998a).

Define the *pull-back Hamiltonian* $G_t : T^*M \to \mathbb{R}$ as

$$G_t = \left(H_t - \widehat{F}_t\right) \circ \widehat{\mathcal{F}}_t = \chi \circ \widehat{\mathcal{F}}_t,$$

whose associated Hamiltonian flow is

$$\mathcal{G}_t = \widehat{\mathcal{F}}_t^{-1} \circ \mathcal{H}_t$$

(the relevant proof can be found in Agrachev and Gamkrelidze, 1997).

Since $DG_t(\widehat{\ell}_0) = 0$, then $D^2G_t(\widehat{\ell}_0)$ is well-defined and the flow

$$\mathcal{G}_{t*} : T_{\widehat{\ell}_0}(T^*M) \to T_{\widehat{\ell}_0}(T^*M)$$

is the Hamiltonian flow associated to $\frac{1}{2}D^2G_t(\widehat{\ell}_0)$ (see for instance Agrachev et al., 1998a, and references therein for details).

We notice (see Agrachev et al., 1998a) that the tangent space $T_{\widehat{\ell}_0}(T^*M)$ is isomorphic to the product $T^*_{\widehat{x}_0}M \times T_{\widehat{x}_0}M$, via the anti-symplectic isomorphism

$$\begin{aligned} \iota : T^*_{\widehat{x}_0}M \times T_{\widehat{x}_0}M &\to T_{\widehat{\ell}_0}(T^*M) \\ (\omega, \delta x) &\mapsto -\omega + d(-\widehat{\beta})_*\delta x. \end{aligned}$$

To define H''_t, consider the Hamiltonian $h''_t : T^*_{\widehat{x}_0}M \times T_{\widehat{x}_0}M \times \mathbb{R}^2 \to \mathbb{R}$, defined as

$$h''_t(\omega, \delta x, \boldsymbol{v}) = \sum_{i=1}^{2} v_i \langle \omega, \dot{g}^i_t(\widehat{x}_0)\rangle + v_i L_{\delta x} L_{\dot{g}^i_t}\widehat{\beta}(\widehat{x}_0) + \frac{1}{2}\sum_{i,j=1}^{2} v_i v_j L_{[\dot{g}^i_t, g^j_t]}\widehat{\beta}(\widehat{x}_0),$$

where $\boldsymbol{v} = (v_1, v_2)$. The minimising Hamiltonian is defined by

$$\frac{\partial}{\partial \boldsymbol{v}} h''_t = 0,$$

i.e.

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = -R(t)^{-1} \begin{pmatrix} \langle \omega, \dot{g}^1_t(\widehat{x}_0)\rangle + L_{\delta x} L_{\dot{g}^1_t}\widehat{\beta}(\widehat{x}_0) \\ \langle \omega, \dot{g}^2_t(\widehat{x}_0)\rangle + L_{\delta x} L_{\dot{g}^2_t}\widehat{\beta}(\widehat{x}_0) \end{pmatrix}$$

(here we remark that the matrix $R(t) = -\mathbb{L}_{\widehat{\lambda}(t)}$ is symmetric).

Then,

$$
\begin{aligned}
H_t''(\omega,\delta x) &= h_t''(\omega,\delta x,\boldsymbol{v}(\omega,\delta x))\\
&= -\frac{1}{2}\left(\langle\omega,\dot g_t^1\rangle + L_{\delta x}L_{\dot g_t^1}\widehat\beta(\widehat x_0), \langle\omega,\dot g_t^2\rangle + L_{\delta x}L_{\dot g_t^2}\widehat\beta(\widehat x_0)\right) R(t)^{-1}\times\\
&\quad\times\begin{pmatrix}\langle\omega,\dot g_t^1\rangle + L_{\delta x}L_{\dot g_t^1}\widehat\beta(\widehat x_0)\\ \langle\omega,\dot g_t^2\rangle + L_{\delta x}L_{\dot g_t^2}\widehat\beta(\widehat x_0)\end{pmatrix}.
\end{aligned}
$$

We denote with $\vec H_t''$ and $\mathcal{H}_t''$, respectively, the Hamiltonian vector field and the Hamiltonian flow associated to H_t''.

LEMMA 8 *The two linear Hamiltonian flows $\mathcal{G}_{t*}$ and $\mathcal{H}_t''$ are equivalent, i.e. there hold:*

$$
H_t'' = -\frac{1}{2}D^2G_t(\widehat\ell_0)\circ\iota \tag{44}
$$

$$
\mathcal{H}_t'' = \iota^{-1}\circ\mathcal{G}_{t*}\circ\iota. \tag{45}
$$

Proof. We split the proof of (44) in two steps: in the first one, we prove that

$$
D^2G_t(\widehat\ell_0)[\delta\ell]^2 = \sum_{i,j=1}^{2}(R^{-1}(t))_{ij}[DF_{0i}(\widehat\lambda(t))\circ\mathcal{H}_{t*}\delta\ell][DF_{0j}(\widehat\lambda(t))\circ\mathcal{H}_{t*}\delta\ell]; \tag{46}
$$

in the second one, that actually

$$
H_t'' = -\frac{1}{2}\sum_{i,j=1}^{2}(R(t)^{-1})_{ij}(DF_{0i}(\widehat\lambda(t))\circ\mathcal{H}_{t*}\circ\iota)(DF_{0j}(\widehat\lambda(t))\circ\mathcal{H}_{t*}\circ\iota). \tag{47}
$$

Once (44) proved, (45) follows straightforwardly.

Let us then prove (44). From $G_t = \chi\circ\widehat{\mathcal{F}}_t$ and the fact that $DG_t(\widehat\ell_0) = 0$, we get that $D^2G_t(\widehat\ell_0)$ is well defined and $D^2G_t(\widehat\ell_0) = D^2\chi(\widehat{\mathcal{F}}_t\circ\widehat\ell_0)\circ\widehat{\mathcal{F}}_{t*}\otimes\widehat{\mathcal{F}}_{t*}$.

Applying Lemma A.2 from Appendix A, we obtain for $\delta\ell\in T_{\widehat\ell_0}(T^*M)$ that

$$
D^2G_t(\widehat\ell_0)[\delta\ell]^2 = \sum_{i,j=1}^{2}(R^{-1}(t))_{ij}[DF_{0i}(\widehat\lambda(t))\circ\mathcal{F}_{t*}\delta\ell][DF_{0j}(\widehat\lambda(t))\circ\mathcal{F}_{t*}\delta\ell].
$$

Now let $\delta\ell = \iota(\delta\omega,\delta x) = -\delta\omega + d(-\widehat\beta)_*\delta x$, where $(\delta\omega,\delta x)\in T^*_{\widehat x_0}M\times T_{\widehat x_0}M$. Since

$$
\begin{aligned}
DF_{0i}(\widehat\ell_0)\,[\iota(\delta\omega,\delta x)] &= L_{\delta-\omega+d(-\widehat\beta)_*\delta x}\langle\ell,[f_0,f_i]\rangle\Big|_{\ell=\widehat\ell_0} =\\
&= \langle-\delta\omega,[f_0,f_i]\rangle + L_{d(-\widehat\beta)_*\delta x}\langle\ell,[f_0,f_i]\rangle\Big|_{\ell=\widehat\ell_0} =\\
&= \langle-\delta\omega,[f_0,f_i]\rangle + L_{\delta x}\langle-d\widehat\beta(x),[f_0,f_i]\rangle\Big|_{x=\widehat x_0},
\end{aligned}
$$

then

$$\begin{aligned}
&DF_{0i}(\widehat{\lambda}(t))\left[\widehat{\mathcal{H}}_{t*}\iota(\delta\omega,\delta x)\right]= \\
&= L_{-\widehat{\mathcal{H}}_{t*}\delta\omega+\widehat{\mathcal{H}}_{t*}d(-\widehat{\beta})_*\delta x}\left(\langle \ell,[f_0,f_i]\circ\widehat{S}_t(\widehat{x}_0)\rangle\right)\Big|_{\ell=\widehat{\lambda}(t)}= \\
&= \left(\langle -\delta\omega,\widehat{S}_{t*}^{-1}[f_0,f_i]\circ\widehat{S}_t(x)\rangle + L_{\delta x}\langle\widehat{\mathcal{F}}_t d(-\widehat{\beta})(x),[f_0,f_i]\circ\widehat{S}_t(x)\rangle\right)\Big|_{x=\widehat{x}_0}= \\
&= -\langle\delta\omega,\dot{g}_t^i(\widehat{x}_0)\rangle - L_{\delta x}L_{\dot{g}_t^i}\widehat{\beta}(\widehat{x}_0).
\end{aligned}$$

Hence, equation (47) is satisfied; we get the thesis. ■

The proof of the main Theorem is completed after the following

LEMMA 9 *Let α be the function defined in* (41)*, and set*

$$\Lambda = \{d\alpha(x) : x\in M\}\,.$$

Λ is a Lagrangian submanifold of Σ containing $\widehat{\ell}_0$ and such that (39) *is fulfilled.*

Proof. Since $L_{f_i}\alpha\equiv 0,\ i=1,2,\ \Lambda\subset\Sigma$. Moreover, it is not difficult to see that the statement is proved after noting that

$$\iota L = T_{\widehat{x}_0}\Lambda.$$

Therefore,

$$\ker\pi_*\mathcal{H}_t''\Big|_L = \ker\pi_*\iota^{-1}\circ\mathcal{G}_{t*}\Big|_{T_{\widehat{x}_0}\Lambda} = \{0\},$$

which implies (39), since $\pi_*\widehat{\mathcal{F}}_{t*}^{-1}=\widehat{S}_{t*}^{-1}\pi_*$. ■

6. The case with several controls

In this section we show how to adapt the construction of Sections 4 and 5 to the case of generic $m\le n-1$.

The proofs are completely analogous to the ones in the case of $m=2$; we just write here the definition of the objects we construct, and claim their properties, without repeating the proofs.

Let us notice that the number m of controlled fields shall be less or equal than $n-1$; in fact, by hypothesis we assume that they are linearly independent, therefore $m\le n$. Moreover, if $m=n$, then we can write $f_0(\widehat{x}_0)=\sum_{i=1}^m\alpha_i f_i(\widehat{x}_0)$, and therefore $F_{max}(\widehat{\lambda}(t))=F_0(\widehat{\lambda}(t))=\sum_{i=1}^m\alpha_i F_i(\widehat{\lambda}(t))=0$, which contradicts the fact that the reference extremal is normal.

6.1. Definition of the super-Hamiltonian

To construct the super-Hamiltonian, we just repeat the same arguments as above.

First of all, we recall that the coercivity of the second variation implies that SGLC holds along the reference extremal, and therefore in a full-measure neighbourhood $\mathcal{U}$ of the reference extremal; we define the subset of Σ

$$\mathcal{S} = \{\ell \in T^*M : F_i(\ell) = 0,\ F_{0i}(\ell) = 0, \quad i = 1, \dots, m\} \cap \mathcal{U}.$$

We claim that the statements of Lemma 2 hold also in this case, that means that

1. $\vec{F}_i$ is tangent to Σ for any $i = 1, \dots, m$;
2. $\vec{F}_i$ is transversal to $\mathcal{S}$ for any $i = 1, \dots, m$;
3. $\vec{F}_{0i}$ is transversal to Σ (and therefore to $\mathcal{S}$) for any $i = 1, \dots, m$;
4. the vectors $\{\vec{F}_1, \dots, \vec{F}_m, \vec{F}_{01}, \dots, \vec{F}_{0m}\}$ are linearly independent;
5. $\mathcal{S}$ is a symplectic submanifold of dimension $2(n-m)$ contained in Σ.

Lemmas 3 and 5 generalise in the following way:

Lemma 10 *There is a neighbourhood $\mathcal{V}$ of the range of $\widehat{\lambda}$ in Σ such that for any $\ell \in \mathcal{V}$ there is a unique (m+1)-tuple $(\ell_{\mathcal{S}}, t_1, \dots, t_m) \in \mathcal{S} \times \mathbb{R}^m$ such that*

$$\ell = \psi(\ell_{\mathcal{S}}, t_1, \dots, t_m) = \exp\left(-t_m \vec{F}_m\right) \circ \cdots \circ \exp\left(-t_1 \vec{F}_1\right)(\ell_{\mathcal{S}});$$

moreover, there exists an $\epsilon > 0$ such that the map $\psi : (\mathcal{S} \cap \mathcal{V}) \times [-\epsilon, \epsilon]^m \to \Sigma$ is a global diffeomorphism (over its image).

Lemma 11 *In the neighbourhood $\mathcal{V}$ we can define the functions $\vartheta_i,\ i = 1, \dots, m$, in such a way that*

$$\begin{aligned}
&\exp(\vartheta_1(\ell)\vec{F}_1) \circ \cdots \circ \exp(\vartheta_m(\ell)\vec{F}_m)(\ell) \in \mathcal{S} \qquad \forall\, \ell \in \mathcal{V}, \\
&\vartheta_1(\ell) = \ldots = \vartheta_m(\ell) = 0 \qquad \ell \in \mathcal{S}.
\end{aligned}$$

We then define the Hamiltonians $H_0 : \mathcal{U} \to \mathbb{R}$ and $\chi : \mathcal{U} \to T^*M$ as

$$\begin{aligned}
&H_0(\ell) = F_0 \circ \exp(\vartheta_1(\ell)\vec{F}_1) \circ \cdots \circ \exp(\vartheta_m(\ell)\vec{F}_m)(\ell), \\
&\chi(\ell) = H_0(\ell) - F_0(\ell) = F_0 \circ \exp(\vartheta_1(\ell)\vec{F}_1) \circ \cdots \circ \exp(\vartheta_m(\ell)\vec{F}_m)(\ell) - F_0(\ell).
\end{aligned}$$

We can show that:

1. For $\ell \in \Sigma$, we have

$$\begin{aligned}
\vec{H}_0(\ell) = &\left(\exp(-\vartheta_m(\ell)\vec{F}_m) \circ \cdots \circ \exp(-\vartheta_1(\ell)\vec{F}_1)\right)_* \vec{F}_0 \\
&\circ \left(\exp(\vartheta_1(\ell)\vec{F}_1) \circ \ \cdots \circ \exp(\vartheta_m(\ell)\vec{F}_m)\right)(\ell).
\end{aligned}$$

2. The vector field $\vec{H}_0$ is tangent to Σ.
3. For $\ell \in \mathcal{S}$, $D\chi(\ell) = 0$, hence $D^2\chi(\ell)$ is well defined and a quadratic form on $T_\ell\Sigma$, whose kernel is $T_\ell\mathcal{S}$ and such that $D^2\chi(\ell)$, coincides with $-\mathbb{L}_\ell$ on $\widetilde{\mathcal{D}} = \mathbb{R}\vec{F}_1 \oplus \cdots \oplus \mathbb{R}\vec{F}_m$.
4. $F_0 \leq H_0$ on Σ, and $F_0 = H_0$ on $\mathcal{S}$.

As for the sufficient condition (Theorem 4), its statement holds in this case, too.

6.2. Proof of the result

As done in the previous case, we can define local coordinates in a neighbourhood of $\widehat{x}_0$ in such a way that

- $\widehat{x}_0 = (0, \dots, 0)$;
- $f_1 = \frac{\partial}{\partial q_1}$ and $f_i = \sum_{j=1}^m \mu_{ij} \frac{\partial}{\partial q_j}$, $i = 2, \dots, m$, where μ_{ij} are locally defined smooth functions;
- the covector $\widehat{\ell}_0$ can be written as $\widehat{\ell}_0 = \sum_{i=m+1}^n \widehat{p}_i dq_i$.

In these coordinates, we can choose the function $\widehat{\beta}$ as

$$\widehat{\beta}(x) = \sum_{i=m+1}^{n} -\widehat{p}_i q_i,$$

and this guarantees that the finite-dimensional term in the second variation vanishes.

We define the local function $\alpha : M \to \mathbb{R}$ as

$$\alpha = \sum_{i=m+1}^{n} \widehat{p}_i q_i + \rho q_i^2, \quad \rho > 0.$$

We put $\gamma'' = D^2(\alpha + \widehat{\beta})(\widehat{x}_0)$. Repeating the same argument as above, we obtain that the modified second variation

$$J''_\alpha[\delta e]^2 = \frac{1}{2}\gamma''[(\delta x, w)]^2 + \sum_{i,j=1}^{m} \frac{1}{2} \int_0^{\widehat{T}} w_i(t) w_j(t) R_{ij}(t) + 2 w_i(t) Q_i(t) \zeta(t)\, dt$$

is coercive on the set $\mathcal{W}$ of the admissible pairs

$$\mathcal{W} = \{\delta e = (\delta x, w) \in T_{\widehat{x}_0}M \times L^2([0, \widehat{T}], \mathbb{R}^m)\}$$

such that

$$\dot{\zeta}(t) = \sum_{i=1}^{m} w_i(t) \dot{g}_t^i(\widehat{x}_0)$$

$$\zeta(0) = \delta x \qquad \zeta(\widehat{T}) = 0.$$

This last condition is equivalent to

$$\ker \pi_* \mathcal{H}''_{t*}\Big|_L = \{0\},$$

where

$$L = \operatorname{span}\left\{\left(-d\alpha(\widehat{x}_0)\left(\frac{\partial}{\partial q_i}, \cdot\right), \frac{\partial}{\partial q_i}\right),\ i = 1, \dots, n\right\}.$$

Putting then

$$\Lambda = \{d\alpha(x) : x \in M\},$$

we get the thesis, since $T_{\widehat{x}_0}\Lambda = \iota L$.

7. An example

In this section we provide an example, illustrating the abstract result. This example is academic, but useful for making the theory more concrete.

Let us consider the following control problem on $\mathbb{R}^3$:

$$\min T$$

subject to

$$\begin{cases} \dot{x}_1 &= u_1 + u_2 x_1 \\ \dot{x}_2 &= u_2 \\ \dot{x}_3 &= 1 - x_1^2 - x_2^2 \end{cases}, \tag{48}$$

with the initial condition

$$\boldsymbol{x}(0) = (0, 0, x_3^0), \qquad \boldsymbol{x}(T) = (0, 0, x_3^1), \quad x_3^1 > x_3^0.$$

The controls may assume values on the whole $\mathbb{R}^2$.

Explicitly, the drift and the controlled vector fields are:

$$f_0(\boldsymbol{x}) = \left(1 - x_1^2 - x_2^2\right)\frac{\partial}{\partial x_3} \qquad f_1(\boldsymbol{x}) = \frac{\partial}{\partial x_1} \qquad f_2(\boldsymbol{x}) = x_1\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}.$$

Since $[f_1, f_2] = f_1$, the distribution $\mathcal{D}$ is involutive.

7.1. Preliminary analysis of the system

The lifted Hamiltonians are

$$F_0(\boldsymbol{p}, \boldsymbol{x}) = p_3\left(1 - x_1^2 - x_2^2\right) \qquad F_1(\boldsymbol{p}, \boldsymbol{x}) = p_1 \qquad F_2(\boldsymbol{p}, \boldsymbol{x}) = p_1 x_1 + p_2,$$

and their Poisson brackets

$$F_{01}(\boldsymbol{p},\boldsymbol{x}) = 2x_1p_3 \qquad F_{02}(\boldsymbol{p},\boldsymbol{x}) = 2(x_1^2 + x_2)p_3.$$

Therefore, the submanifold Σ is

$$\begin{aligned}\Sigma &= \{(\boldsymbol{p},\boldsymbol{x}) : p_1 = 0,\ x_1p_1 + p_2 = 0\}\\ &= \{(\boldsymbol{p},\boldsymbol{x}) : p_1 = p_2 = 0\}\,.\end{aligned}$$

By computations, we can prove that the matrix $\mathbb{L}_{(\boldsymbol{p},\boldsymbol{x})}$ (equation (12)) is negative definite in the half-space $\{p_3 > 0\}$. Then

$$\begin{aligned}\mathcal{S} &= \{(\boldsymbol{p},\boldsymbol{x}) : F_1 = F_2 = F_{01} = F_{02} = 0\} \cap \{p_3 > 0\}\\ &= \{(0,0,p_3;0,0,x_3) : p_3 > 0\}.\end{aligned}$$

Remark 14 *Since any singular optimal trajectory is the projection of an extremal with values in $\mathcal{S}$, then the minimum-time problem for this dynamics may have a singular solution only if the end-points lie on x_3-axis, with $x_3^1 > x_3^0$.*

Remark 15 *If the control set is the whole $\mathbb{R}^2$, the minimum-time problem between fixed points has a standard solution if and only if the points belong to the x_3-axis and $x_3^1 > x_3^0$. On the other hand, we can easily check that the infimum of the time needed to join two points belonging to the same plane parallel to the (x_1,x_2)-plane, with $x_3^1 > x_3^0$, is zero, since the controls may be unbounded (see also Jurdjevic, 1997, Sec. 2.1, Theorem 6). One could solve the problem between (x_1^0,x_2^0,x_3^0) and (x_1^1,x_2^1,x_3^1) allowing jumps, i.e. by jumping from (x_1^0,x_2^0,x_3^0) to $(0,0,x_3^0)$ in "zero time", then following the system from $(0,0,x_3^0)$ to $(0,0,x_3^1)$, and finally jumping again from $(0,0,x_3^1)$ to (x_1^1,x_2^1,x_3^1) in "zero time".*

7.2. Studying the extremal

It is not difficult to see that the curve

$$\widehat{\lambda}(t) := (\widehat{\boldsymbol{p}}(t),\widehat{\boldsymbol{x}}(t)) = (0,0,1;0,0,x_3^0 + t), \quad t \in [0, x_3^1 - x_3^0] \tag{49}$$

is a singular extremal for the minimum-time problem subject to (48), associated to the control $\widehat{\mathbf{u}} = (0,0)$. In fact, $\boldsymbol{p}$ satisfies the adjoint equation and $\widehat{F}_t(\widehat{\boldsymbol{p}}(t),\widehat{\boldsymbol{x}}(t)) = p_3 \equiv 1$.

Hence, the pair $(\widehat{\boldsymbol{p}}(t),\widehat{\boldsymbol{x}}(t);\widehat{\mathbf{u}})$ satisfies PMP, with appropriate boundary conditions.

Lemma 12 *The second variation is coercive on the space of admissible variations $\widetilde{\mathcal{W}}$.*

Proof. Since the reference control is identically null, the reference flow acts as

$$\widehat{S}_t : (x_1^0,x_2^0,x_3^0) \mapsto (x_1^0,x_2^0,x_3^0 + (1 - (x_1^0)^2 - (x_2^0)^2)t).$$

We compute the feedback fields g_t^i, $i = 1, 2$; explicitly, we get

$$g_t^1(\boldsymbol{x}) = x_1 \frac{\partial}{\partial x_1} + 2x_1 t \frac{\partial}{\partial x_3} \qquad g_t^2(\boldsymbol{x}) = x_1 \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2} + 2(x_1 + x_2) t \frac{\partial}{\partial x_3},$$

and

$$\dot{g}_t^1(\boldsymbol{x}) = 2x_1 \frac{\partial}{\partial x_3} \qquad \dot{g}_t^2(\boldsymbol{x}) = 2(x_1 + x_2) \frac{\partial}{\partial x_3},$$

hence $\dot{g}_t^1(\widehat{\boldsymbol{x}}(0)) = \dot{g}_t^2(\widehat{\boldsymbol{x}}(0)) = 0$.

This implies that the admissible $\zeta(t)$ satisfies the system

$$\begin{cases} \dot{\zeta}(t) = 0 \\ \zeta(0) = w_1^0 f_1(0) + w_2^0 f_2(0), \quad \zeta(T) = 0, \qquad (w_1^0, w_2^0) \in \mathbb{R}^2, \end{cases}$$

meaning that $\zeta(t) \equiv 0$. Therefore, the set of admissible variations is $\widetilde{\mathcal{W}} = \{0\} \times L^2([0, \widehat{T}], \mathbb{R}^2)$.

The extended second variation is

$$J''_{\text{ext}}[(w_0, w)]^2 = \int_0^T w_1^2(t) + w_2^2(t)\, dt = \|\boldsymbol{w}\|_{L^2}^2,$$

which is coercive on $\widetilde{\mathcal{W}}$. ■

8. Final remarks

This paper is a part of a research project, in which we intend to use the Hamiltonian approach to establish second-order optimality conditions for optimal control problems. There are immediate generalisation of this result, that we intend to study in the future, and also many interesting issues in this research field.

One direct generalisation of this result will concern a further relaxation of the hypotheses on the controlled fields, that is, the case in which we do not ask anything on their Lie brackets. This case is quite natural: the minimum-length problem in subriemannian geometry belongs to this class.

A natural step would be to study the optimality conditions for singular extremals of the Mayer problem. In this case, sufficient conditions for weak and Pontryagin optimality of singular extremals have already been obtained in Dmitruk (1977, 1983, 2008).

Further investigation will concern the statement of second-order optimality condition for concatenations of bang-singular arcs. For the single-input case, see Poggiolini and Stefani (2008, 2009).

Another development will be to consider a stronger notion of optimality, that is, *strong state-local optimality* of the minimum-time problem, where *state-local* means “in a neighbourhood of the *range* of the reference trajectory”. For precise definition, see Poggiolini and Stefani (2004), where this type of local strong optimality was first considered. In order to obtain this type of optimality we shall follow the ideas from Poggiolini and Stefani (2009), where the single-input case is considered and sufficient conditions of strong state-local optimality are proved.

To obtain the suitable second variation in this case, we have to reduce the time-optimal problem to a Mayer problem on $[0, \widehat{T}]$, but in this case we extend both the state space and the control space: indeed, we add a new variable T, which is also the cost and a new control $u_0 : [0, \widehat{T}] \to (0, \infty)$. With the same techniques, used in Subsection 3.2, we obtain an extended second variation defined on $\mathbb{R}^{m+1} \times L^2([0, \widehat{T}], \mathbb{R}^m)$ and given by

$$J''_{\text{ext}}[(\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_m, w)]^2 = \frac{1}{2} L_{(\varepsilon_0 f_0 + \sum_{i=1}^m \varepsilon_i f_i)} L_{(\varepsilon_0 f_0 + \sum_{i=1}^m \varepsilon_i f_i)} \widehat{\beta}(\widehat{x}_0) + \\ + \sum_{i=1}^m \int_0^1 w_i(t) L_{\zeta(t)} L_{\dot{g}_t^i} \widehat{\beta}(\widehat{x}_0) + \frac{1}{2} \sum_{i,j=1}^m \int_0^1 w_i(t) w_j(t) L_{[g_t^i, \dot{g}_t^j]} \widehat{\beta}(\widehat{x}_0)\, dt,$$

where the variable $\zeta(t) \in T_{\widehat{x}_0} M$ satisfies the following problem:

$$\dot{\zeta}(t) = \sum_{i=1}^m w_i(t) \dot{g}_t^i(\widehat{x}_0), \tag{50}$$

$$\zeta(0) = \varepsilon_0 f_0(\widehat{x}_0) + \sum_{i=1}^m \varepsilon_i f_i(\widehat{x}_0), \qquad \zeta(\widehat{T}) = 0. \tag{51}$$

The space of admissible variations is then a subspace $\widetilde{\mathcal{W}}^{(m+1)}$ of $\mathbb{R}^{m+1}$ $\times L^2([0, 1], \mathbb{R}^m)$ defined as

$$\widetilde{\mathcal{W}}^{(m+1)} := \{(\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_m, w) \in \mathbb{R}^{m+1} \times L^2([0, 1], \mathbb{R}^m) \\ \text{that satisfy equations } (50) - (51)\}.$$

It is immediate to see that the space $\widetilde{\mathcal{W}}$ of admissible variations for the time-optimal problem (given by equation (21)) coincides with the subspace $\widetilde{\mathcal{W}}^{(m+1)} \cap \{\varepsilon_0 = 0\}$.

Appendix A

Lemma A.1 *$D\chi(\ell) = 0$ for any $\ell \in \mathcal{S}$.*

Proof. Recall that

$$\chi(\ell) = H_0(\ell) - F_0(\ell) = F_0 \circ \exp(\vartheta_1(\ell) \vec{F}_1) \circ \exp(\vartheta_2(\ell) \vec{F}_2)(\ell) - F_0(\ell).$$

Recall the definition of the map $\psi : \mathcal{S} \times [-\epsilon, \epsilon]^2 \to \Sigma$ (equation (24)), define the map $\phi : \Sigma \cap \mathcal{U} \to \mathcal{S} \times \mathbb{R}^2$ (locally in a neighbourhood of the reference extremal) by

$$\phi(\ell) = \Big(\exp(\vartheta_1(\ell)\vec{F}_1) \circ \exp(\vartheta_2(\ell)\vec{F}_2)(\ell), \vartheta_1(\ell), \vartheta_2(\ell)\Big)\,; \tag{52}$$

notice that ϕ is the local inverse of ψ. We, moreover, define the projection:

$$\pi_0 : \mathcal{S} \times [-\epsilon,\epsilon]^2 \to \mathcal{S} \times [-\epsilon,\epsilon]^2, \qquad \pi_0(\ell_\mathcal{S}, t_1, t_2) = (\ell_\mathcal{S}, 0, 0).$$

We can write χ as

$$\chi = [F_0 \circ \psi \circ \pi_0 - F_0 \circ \psi] \circ \phi,$$

therefore

$$D\chi = DF_0 \circ D\psi \circ D\pi_0 \circ D\phi - DF_0 \circ D\psi \circ D\phi.$$

Notice that $D\pi_0 = \begin{pmatrix} \mathrm{id} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, and that

$$DF_0 \circ D\psi = \left(\frac{\partial(F_0\circ\psi)}{\partial \ell_\mathcal{S}}, \frac{\partial(F_0\circ\psi)}{\partial t_1}, \frac{\partial(F_0\circ\psi)}{\partial t_2} \right),$$

therefore

$$D\chi = \left(\frac{\partial(F_0\circ\psi)}{\partial \ell_\mathcal{S}}\Big|_{\psi(\ell_\mathcal{S},0,0)} - \frac{\partial(F_0\circ\psi)}{\partial \ell_\mathcal{S}}\Big|_{\psi(\ell_\mathcal{S},-t_1,-t_2)}, \right.$$
$$\left. -\frac{\partial(F_0\circ\psi)}{\partial t_1}\Big|_{\psi(\ell_\mathcal{S},-t_1,-t_2)}, -\frac{\partial(F_0\circ\psi)}{\partial t_2}\Big|_{\psi(\ell_\mathcal{S},-t_1,-t_2)} \right) \circ D\phi.$$

On $\mathcal{S}$ we then have

$$D\chi(\ell_\mathcal{S}) = \underbrace{\left(\frac{\partial(F_0\circ\psi)}{\partial \ell_\mathcal{S}}\Big|_{\ell_\mathcal{S}} - \frac{\partial(F_0\circ\psi)}{\partial \ell_\mathcal{S}}\Big|_{\ell_\mathcal{S}}, -F_{01}\Big|_{\ell_\mathcal{S}}, -F_{02}\Big|_{\ell_\mathcal{S}} \right)}_{=0} \circ D\phi = 0.$$

LEMMA A.2 *$D^2\chi(\ell_\mathcal{S}) = -\mathbb{L}_{\ell_\mathcal{S}}^{-1}[(\langle DF_{01}(\ell_\mathcal{S}), \delta\ell\rangle, \langle DF_{02}(\ell_\mathcal{S}), \delta\ell\rangle)]^2$ for any $\ell_\mathcal{S} \in \mathcal{S}$ and $\delta\ell \in T_{\ell_\mathcal{S}}\Sigma$.*

Proof. Let us first prove that $D^2\chi(\ell_\mathcal{S}) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -F_{110}(\ell_\mathcal{S}) & -F_{120}(\ell_\mathcal{S}) \\ 0 & -F_{120}(\ell_\mathcal{S}) & -F_{220}(\ell_\mathcal{S}) \end{pmatrix} [D\phi \otimes D\phi]\Big|_{\ell_\mathcal{S}}$ for any $\ell_\mathcal{S} \in \mathcal{S}$, where ϕ is defined in (52).

Formally

$$D^2\chi(\ell) = D\left[DF_0 \circ D\phi \circ D\pi_0 - DF_0 \circ D\psi\right]\Big|_{\phi(\ell)} \circ [D\phi \otimes D\phi]\Big|_{\ell} +$$
$$+ \left[DF_0 \circ D\psi \circ D\pi_0 - DF_0 \circ D\psi\right]\Big|_{\phi(\ell)} \circ D^2\phi\Big|_{\ell},$$

and then on $\mathcal{S}$ we have

$$\begin{aligned}D^2\chi\Big|_{\ell_{\mathcal{S}}} &= D\left[DF_0\circ D\psi\circ D\pi_0 - DF_0\circ D\psi\right]\Big|_{\ell_{\mathcal{S}}}\circ\left[D\phi\otimes D\phi\right]+\\ &+\underbrace{\left[DF_0\circ D\psi\circ D\pi_0 - DF_0\circ D\psi\circ D\pi_-\right]\Big|_{\ell_{\mathcal{S}}}}_{=0}\circ D^2\phi =\\ &= D\left[DF_0\circ D\psi\circ D\pi_0 - DF_0\circ D\psi\circ D\pi_-\right]\Big|_{\ell=\ell_{\mathcal{S}}}\circ\left[D\phi\otimes D\phi\right].\end{aligned}$$

In other words

$$D^2\chi\Big|_{\ell_{\mathcal{S}}}=\left[\begin{pmatrix}\frac{\partial^2(F_0\circ\psi)}{\partial\ell_{\mathcal{S}}^2} & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}-\begin{pmatrix}\frac{\partial^2(F_0\circ\psi)}{\partial\ell_{\mathcal{S}}^2} & \frac{\partial^2(F_0\circ\psi)}{\partial\ell_{\mathcal{S}}\partial t_1} & \frac{\partial^2(F_0\circ\psi)}{\partial\ell_{\mathcal{S}}\partial t_2}\\ \frac{\partial^2(F_0\circ\psi)}{\partial\ell_{\mathcal{S}}\partial t_1} & \frac{\partial^2(F_0\circ\psi)}{\partial t_1^2} & \frac{\partial^2(F_0\circ\psi)}{\partial t_2\partial t_1}\\ \frac{\partial^2(F_0\circ\psi)}{\partial\ell_{\mathcal{S}}\partial t_2} & \frac{\partial^2(F_0\circ\psi)}{\partial t_2\partial t_1} & \frac{\partial^2(F_0\circ\psi)}{\partial t_2^2}\end{pmatrix}\right]\Big|_{\phi(\ell_{\mathcal{S}})}\circ\left[D\phi\otimes D\phi\right]\Big|_{\ell_{\mathcal{S}}}.$$

Let us compute the derivatives:

$$\begin{aligned}&\frac{\partial}{\partial t_1}F_0\circ\exp(-t_2\vec{F}_2)\circ\exp(-t_1\vec{F}_1)(\ell_{\mathcal{S}}) =\\ &=\langle dF_0,\exp(-t_2\vec{F}_2)_*(-\vec{F}_1)\rangle\Big|_{\exp(-t_2\vec{F}_2)\circ\exp(-t_1\vec{F}_1)(\ell_{\mathcal{S}})},\end{aligned}$$

then, on $\mathcal{S}$,

$$\frac{\partial}{\partial t_1}F_0\circ\exp(-t_2\vec{F}_2)\circ\exp(-t_1\vec{F}_1)(\ell_{\mathcal{S}})\Big|_{t_1=t_2=0}=\sigma_{\ell_{\mathcal{S}}}\left(\vec{F}_0,\vec{F}_1\right)=F_{01}(\ell_{\mathcal{S}}).$$

As for t_2, we have

$$\frac{\partial}{\partial t_2}F_0\circ\exp(-t_2\vec{F}_2)\circ\exp(-t_1\vec{F}_1)(\ell_{\mathcal{S}})\Big|_{t_1=t_2=0}=F_{02}(\ell_{\mathcal{S}}).$$

Moreover

$$\frac{\partial^2}{\partial t_1^2}F_0\circ\exp(-t_2\vec{F}_2)\circ\exp(-t_1\vec{F}_1)(\ell_{\mathcal{S}})\Big|_{t_1=t_2=0}=F_{110}(\ell_{\mathcal{S}})$$

and

$$\frac{\partial^2}{\partial t_2^2}F_0\circ\exp(-t_2\vec{F}_2)\circ\exp(-t_1\vec{F}_1)(\ell_{\mathcal{S}})\Big|_{t_1=t_2=0}=F_{220}(\ell_{\mathcal{S}}).$$

The mixed derivative is

$$\frac{\partial^2}{\partial t_2\partial t_1}F_0\circ\exp(-t_2\vec{F}_2)\circ\exp(-t_1\vec{F}_1)(\ell)\Big|_{t_1=t_2=0}=\sigma\left(\vec{F}_0,[\vec{F}_2,\vec{F}_1]\right)(\ell)=F_{120}(\ell);$$

since F_{01} and F_{02} vanish identically on $\mathcal{S}$, we have that

$$\frac{\partial^2}{\partial \ell_{\mathcal{S}} \partial t_1} F_0 \circ \exp(-t_2 \vec{F}_2) \circ \exp(-t_1 \vec{F}_1)(\ell_{\mathcal{S}}) \Big|_{t_1=t_2=0} = \frac{\partial}{\partial \ell_{\mathcal{S}}} F_{10}(\ell_{\mathcal{S}}) \equiv 0$$
$$\frac{\partial^2}{\partial \ell_{\mathcal{S}} \partial t_2} F_0 \circ \exp(-t_2 \vec{F}_2) \circ \exp(-t_1 \vec{F}_1)(\ell_{\mathcal{S}}) \Big|_{t_1=t_2=0} = \frac{\partial}{\partial \ell_{\mathcal{S}}} F_{20}(\ell_{\mathcal{S}}) \equiv 0.$$

Let us now compute $D\phi$:

$$\begin{aligned} D\phi(\ell)[\delta\ell] = \Big(& \exp(\vartheta_1(\ell)\vec{F}_1)_* \circ \exp(\vartheta_2(\ell)\vec{F}_2)_*[\delta\ell] + D\vartheta_1[\delta\ell]\vec{F}_1 + \\ & + D\vartheta_2[\delta\ell]\vec{F}_2, D\vartheta_1[\delta\ell], D\vartheta_2[\delta\ell]\Big) = \\ & = \Big(\delta\ell_{\mathcal{S}} + \alpha\vec{F}_1 + \beta\vec{F}_2 - \alpha\vec{F}_1 - \beta\vec{F}_2, D\vartheta_1[\delta\ell], D\vartheta_2[\delta\ell]\Big) = \\ & = (\delta\ell_{\mathcal{S}}, D\vartheta_1[\delta\ell], D\vartheta_2[\delta\ell]), \end{aligned}$$

where $\delta\ell = \delta\ell_{\mathcal{S}} + \alpha\vec{F}_1 + \beta\vec{F}_2$ is the decomposition of $\delta\ell$ in $T_\ell\mathcal{S} \oplus \widetilde{\mathcal{D}}$.

The statement follows from equation (28) in Lemma 5. ■

Appendix B

LEMMA B.1 *The flow of the Hamiltonian H_t preserves the distribution $\widetilde{\mathcal{D}}$ along the reference extremal $\widehat{\lambda}$, and the flow of K^S preserves the distribution $\widetilde{\mathcal{D}}$ along all its trajectories that are contained in $\mathcal{S}$.*

Proof. Let us put $X_i(t) = \mathcal{H}_{t*}^{-1}\vec{F}_i \circ \mathcal{H}_t(\widehat{\ell}_0)$, $i = 1, 2$; we have that

$$\begin{aligned} \frac{d}{dt}X_i(t) &= \mathcal{H}_{t*}^{-1}\left[\vec{H}_t, \vec{F}_i\right] \circ \mathcal{H}_t(\widehat{\ell}_0) = \\ &= \mathcal{H}_{t*}^{-1}\left[\vec{H}_0 + \widehat{u}_1(t)\vec{F}_1 + \widehat{u}_2(t)\vec{F}_2, \vec{F}_i\right] \circ \mathcal{H}_t(\widehat{\ell}_0) = \\ &= (-1)^i \widehat{u}_j(t)\mathcal{H}_{t*}^{-1}[\vec{F}_i(t), \vec{F}_j(t)] \circ \mathcal{H}_t(\widehat{\ell}_0) \qquad i \neq j \\ &= \alpha_1^i(t)X_1(t) + \alpha_2^i(t)X_2(t), \end{aligned}$$

for some $\alpha_j^i(t)$, since $\widetilde{\mathcal{D}}$ is involutive. Indeed,

$$\begin{aligned} &[\vec{H}_0, \vec{F}_1](\ell) = \left[\exp(-\vartheta_2\vec{F}_2)_* \circ \exp(-\vartheta_1\vec{F}_1)_*\vec{F}_0 \circ \exp(\vartheta_1\vec{F}_1) \circ \exp(\vartheta_2\vec{F}_2), \vec{F}_1\right](\ell) = \\ &= \left[\exp(-t_2\vec{F}_2)_* \circ \exp(-t_1\vec{F}_1)_*\vec{F}_0 \circ \exp(t_1\vec{F}_1) \circ \exp(t_2\vec{F}_2), \vec{F}_1\right]\Big|_{t_i=\vartheta_i}(\ell) + \\ &- L_{\vec{F}_1}(\vartheta_2)(\ell)\frac{\partial}{\partial t_2}\Big(\exp(-t_2\vec{F}_2)_* \circ \exp(-t_1\vec{F}_1)_*\vec{F}_0 \circ \exp(t_1\vec{F}_1) \circ \exp(t_2\vec{F}_2)\Big)\Big|_{t_i=\vartheta_i}(\ell) + \\ &- L_{\vec{F}_1}(\vartheta_1)(\ell)\frac{\partial}{\partial t_1}\Big(\exp(-t_2\vec{F}_2)_* \circ \exp(-t_1\vec{F}_1)_*\vec{F}_0 \circ \exp(t_1\vec{F}_1) \circ \exp(t_2\vec{F}_2)\Big)\Big|_{t_i=\vartheta_i}(\ell) = \end{aligned}$$

$$= \left[\exp(-t_2\vec{F}_2)_* \circ \exp(-t_1\vec{F}_1)_*\vec{F}_0 \circ \exp(t_1\vec{F}_1) \circ \exp(t_2\vec{F}_2), \vec{F}_1\right]\Big|_{t_i=\vartheta_i}(\ell)+$$
$$-L_{\vec{F}_1}(\vartheta_2)(\ell)\exp(-t_2\vec{F}_2)_*\left[\vec{F}_2, \exp(-t_1\vec{F}_1)_*\vec{F}_0 \circ \exp(t_1\vec{F}_1)\right] \circ \exp(t_2\vec{F}_2)(\ell)\Big|_{t_i=\vartheta_i}+$$
$$-L_{\vec{F}_1}(\vartheta_1)(\ell)\exp(-t_2\vec{F}_2)_* \circ \exp(-t_1\vec{F}_1)_*\left[\vec{F}_1, \vec{F}_0\right] \circ \exp(t_1\vec{F}_1) \circ \exp(t_2\vec{F}_2)\Big|_{t_i=\vartheta_i}(\ell).$$

If $\ell \in \mathcal{S}$, we have that $\vartheta_i(\ell)=0$ and from equation (28) we see that $L_{\vec{F}_1}(\vartheta_1)(\ell)= -1$ and $L_{\vec{F}_1}(\vartheta_2)(\ell) = 0$, then

$$[\vec{H}_0, \vec{F}_1](\ell) = [\vec{F}_0, \vec{F}_1](\ell) + [\vec{F}_1, \vec{F}_0](\ell) = 0.$$

Analogously, we prove that $[\vec{H}_0, \vec{F}_2](\ell) = 0$ on $\mathcal{S}$.

Let now $\eta \in T^*_{\widehat{\ell}_0}(T^*M)$ be a 1-form such that $\langle \eta, \vec{F}_i\rangle = 0\ \forall\ i$, and call $z_i(t) = \langle \eta, X_i(t)\rangle$; we have that

$$\begin{pmatrix}\dot z_1(t)\\ \dot z_2(t)\end{pmatrix} = \begin{pmatrix}\alpha_1^1(t) & \alpha_1^2(t)\\ \alpha_2^1(t) & \alpha_2^2(t)\end{pmatrix}\begin{pmatrix}z_1(t)\\ z_2(t)\end{pmatrix},$$

i.e. the functions $z_i(t)$ satisfy a differential first-order linear system with initial conditions $z_i(0) = 0$. Since this happens for any 1-form that vanishes on $\widetilde{\mathcal{D}}$, we shall conclude that $X_i(t) \in \widetilde{\mathcal{D}}\ \forall\ t,\ i = 1, 2$. We get the thesis.

The same argument proves that

$$\mathcal{K}^{S\,-1}_{t*}\vec{F}_i \circ \mathcal{K}^S_t(\ell) \in \widetilde{\mathcal{D}}_\ell \quad \forall\ \ell \in \mathcal{S}, \quad i = 1, 2.$$

Proof (Proof of Lemma 6). Let us consider the Hamiltonian $P_t = (K^S - H_t) \circ \mathcal{H}_t$ restricted to Σ, with the associated Hamiltonian flow $\mathcal{P}_t = \mathcal{H}_t^{-1} \circ \mathcal{K}^S_t$.

Since $DP_t(\widehat{\ell}_0) = 0$, then $D^2P_t(\widehat{\ell}_0)$ is well-defined and it turns out that $\mathcal{P}_{t*} = \mathcal{H}^{-1}_{t*} \circ \mathcal{K}^S_{t*} : T_{\widehat{\ell}_0}(T^*M) \to T_{\widehat{\ell}_0}(T^*M)$ is the Hamiltonian flow associated to $\frac{1}{2}D^2P_t(\widehat{\ell}_0)$.

By computation, we get that $\mathcal{P}_{t*}$ is the Hamiltonian flow associated to the vector field

$$\sum_i \langle d\nu_i, \mathcal{H}_{t*}\cdot\rangle\ \mathcal{H}^{-1}_{t*}\vec{F}_i(\widehat{\lambda}(t))\ \in \widetilde{\mathcal{D}}_{\widehat{\ell}_0}. \tag{53}$$

Let now $\delta\ell \in T_{\widehat{\ell}_0}\Lambda \cap \ker \mathcal{K}^S_{t*}$, and write it as $\delta\ell = \delta\ell_{\mathcal{S}} + \delta\ell_{\widetilde{\mathcal{D}}}$.

By previous lemma and the fact that $\mathcal{K}^S_t$ preserves $\mathcal{S}$ we can conclude that

$$\mathcal{K}^S_{t*} : T_{\widehat{\ell}_0}\mathcal{S} \to T_{\widehat{\ell}_0}\mathcal{S} \quad \text{and} \quad \mathcal{K}^S_{t*} : \mathcal{D}_{\widehat{\ell}_0} \to \mathcal{D}_{\widehat{\ell}_0}.$$

This implies that

$$\pi_*\mathcal{K}^S_{t*}\delta\ell_{\mathcal{S}} = 0 \qquad \pi_*\mathcal{K}^S_{t*}\delta\ell_{\widetilde{\mathcal{D}}} = 0. \tag{54}$$

It is easy to show that $\mathcal{P}_{t*}\delta\ell_{\widetilde{\mathcal{D}}} = \delta\ell_{\widetilde{\mathcal{D}}}$, that is, $\mathcal{K}^S_{t*}\delta\ell_{\widetilde{\mathcal{D}}} = \mathcal{H}_{t*}\delta\ell_{\widetilde{\mathcal{D}}}$. Therefore, since by hypothesis $\ker\mathcal{H}_{t*} \cap T_{\widehat{\ell}_0}\Lambda = 0$, then $\delta\ell_{\widetilde{\mathcal{D}}} = 0$.

Since the flow $\mathcal{P}_{t*}$ is generated by the vector field (53), we get that the component belonging to $\mathcal{S}$ of $\mathcal{P}_{t*}\delta\ell_{\mathcal{S}}$ is constant and therefore $\mathcal{P}_{t*}\delta\ell_{\mathcal{S}} = \delta\ell_{\mathcal{S}} + \delta\ell'_{\widetilde{\mathcal{D}}}(t)$. Then

$$\begin{aligned}\pi_*\mathcal{K}^S_{t*}(\delta\ell_{\mathcal{S}}) &= \pi_*\mathcal{H}_{t*}(\delta\ell_{\mathcal{S}}) + \pi_*\mathcal{H}_{t*}(\delta\ell'_{\widetilde{\mathcal{D}}}(t)) = \\ &= \pi_*\mathcal{H}_{t*}\delta\ell_{\mathcal{S}} + \sum_i \alpha_i(t) f_i(\widehat{\xi}(t)) = \\ &= \pi_*\mathcal{H}_{t*}\left(\delta\ell_{\mathcal{S}} + \sum_i \beta_i(t) F_i(\widehat{\ell}_0)\right),\end{aligned}$$

for some coefficients $\alpha_i(t), \beta_i(t)$.

Therefore, if $\pi_*\mathcal{K}^S_{t*}(\delta\ell_{\mathcal{S}})=0$, then it shall be $\pi_*\mathcal{H}_{t*}\Big(\delta\ell_{\mathcal{S}}+\sum_i \beta_i(t)F_i(\widehat{\ell}_0)\Big)=0$. By hypothesis, this can happen only if $\delta\ell_{\mathcal{S}} + \sum_i \beta_i(t)F_i(\widehat{\ell}_0) = 0$. Hence it shall be $\delta\ell_{\mathcal{S}} = 0$. ∎

Lemma B.2 *Under the hypotheses of Lemma 6, the following equality holds*

$$\left(\pi\circ\mathcal{K}^S_t\right)^{-1}_* f_i(\xi(t)) = \mathcal{K}^{S-1}_{t*}\vec{F}_i(\lambda(t)),$$

where $\lambda(t) = \mathcal{K}^S_t \circ \ell(t),\ \ell(t) \in \Lambda \cap \mathcal{S}$, *and* $\xi(t) = \pi\circ\lambda(t)$.

Proof. First of all we write

$$\begin{aligned}f_i(\xi(t)) &= \pi_*\vec{F}_i(\lambda(t)) = \\ &= \pi_*\circ\mathcal{K}^S_{t*}\circ\mathcal{K}^{S-1}_{t*}\vec{F}_i(\lambda(t)).\end{aligned}$$

We know that flow $\mathcal{K}^S_{t*}$ maps $T_{\ell(t)}\mathcal{S}$ into $T_{\lambda(t)}\mathcal{S}$, and, moreover, that it preserves the distribution $\widetilde{\mathcal{D}}$. Since $F_i(\lambda(t)) \in \widetilde{\mathcal{D}}_{\lambda(t)}$, we have that $\mathcal{K}^{S-1}_{t*}\vec{F}_i(\lambda(t)) \in \widetilde{\mathcal{D}}_{\ell(t)} \subset T_{\ell(t)}\Lambda$.

Since $(\pi\circ\mathcal{K}^S_t)_*\Big|_{T_{\widehat{\ell}_0}\Lambda}$ is an isomorphism, then also $(\pi\circ\mathcal{K}^S_t)_*\Big|_{T_{\ell(t)}\Lambda}$ is an isomorphism for $\ell(t) \in \Lambda$ close to $\widehat{\ell}_0$, and then we can rewrite the equation above as

$$\left(\pi\circ\mathcal{K}^S_t\right)^{-1}_* f_i(\xi(t)) = \mathcal{K}^{S-1}_{t*}\vec{F}_i(\lambda(t)).$$

References

Abraham, R. and Marsden, J. (1978) *The Foundation of Mechanics.* Addison-Wesley, Reading.

Agrachev, A.A. and Gamkrelidze, R. (1998) Symplectic Methods for Optimization and Control. In: B. Jakubczyk and W. Respondek, eds., *Geometry of Feedback and Optimal Control*. Marcel Dekker, 1–58.

Agrachev, A.A. and Sachkov, Y. (2004) *Control Theory from the Geometric Viewpoint*. Springer-Verlag, Berlin.

Agrachev, A.A., Stefani, G. and Zezza, P. (1998a) An invariant second variation in optimal control. *Int. J. Control* **71** (5), 689–715.

Agrachev, A.A., Stefani, G. and Zezza, P. (1998b) Strong minima in optimal control. *Proc. Stekhlov Inst. Math.* **220**, 4–26.

Agrachev, A.A., Stefani, G. and Zezza, P. (2002) Strong Optimality for a Bang-Bang Trajectory. *SIAM J. Control and Optimization* **41** (4), 991–1014.

Dmitruk, A.V. (1977) Quadratic condition for a weak minimum for singular regimes in optimal control problems. *Soviet Math. Dokl.* **18**.

Dmitruk, A.V. (1984) Quadratic conditions for a Pontryagin minimum in an optimal control problem linear in the control, with a constraint on the control. *Soviet Math. Dokl.* **28**.

Dmitruk, A.V. (2008) Jacobi type conditions for singular extremals. *Control and Cybernetics* **37** (2), 285–306.

Gabasov, R. and Kirillova, F.M. (1972) High order necessary conditions for optimality. *SIAM J. Control and Optimization* **10**, 127–168.

Giaquinta, M. and Hildebrandt, S. (1996) *Calculus of Variations - I and II*. Springer-Verlag, Berlin.

Goh, B.S. (1966) The second variation for singular Bolza problems. *SIAM J. Control and Optimization* **4**, 309–325.

Hestenes, M.R. (1951) Application of the theory of quadratic forms in Hilbert spaces to the calculus of variations. *Pac. J. Math.* **1**, 525–581.

Hestenes, M.R. (1966) *Calculus of Variations and Optimal Control*. Wiley & Sons, New York.

Jurdjevic, V. (1997) *Geometric Control Theory*. Cambridge University Press, Cambridge.

Lee, J.M. (2006) *Introduction to Smooth Manifolds*. Springer-Verlag, Berlin.

Páles, Z. and Zeidan, V. (1994) Non-smooth optimum problems with constraints. *SIAM J. Control and Optimization* **32**, 1476–1502.

Poggiolini, L. and Stefani, G. (2004) State-local optimality of a bang-bang trajectory: a Hamiltonian approach. *Systems and Control Letters* **53**, 269–279.

Poggiolini, L. and Stefani, G. (2008) Sufficient optimality conditions for a bang-singular extremal in the minimum time problem. *Control and Cybernetics* **37**, 469–490.

Poggiolini, L. and Stefani, G. (2009) Sufficient optimality conditions for a bang-singular extremal in the minimum time problem. Submitted to *J. Dynamical and Control Systems*.

Stefani, G. (2004) Minimum-time optimality of a singular arc: second order sufficient conditions *Proc. of CDC04*. IEEE, Control Systems Society.

Stefani, G. (2008) Strong optimality of singular trajectories. In: F. Ancona, A. Bressan, P. Cannarsa, F. Clarke and P. R. Wolenski, eds., *Geometry Control and Nonsmooth Analysis*,

Stefani, G. and Zezza, P. (2002) Constrained regular LQ-control problems. *SIAM J. Control and Optimization* **35**,3, 876–900.