

## On second order sufficient optimality conditions for a bang-singular arc

---

**Laura Poggiolini**<sup>\*†</sup>

*Università di Firenze*

*E-mail:* laura.poggiolini@unifi.it

**Gianna Stefani**<sup>‡</sup>

*Università di Firenze*

*E-mail:* gianna.stefani@unifi.it

We consider a bang-singular normal extremal for a Mayer problem associated to a single-input control system affine in the control. We give sufficient conditions for the extremal to be a strong local minimizer. The conditions include the coercivity of the *extended second variation* associated to the problem restricted to the singular arc and *regularity conditions* on the bang arc. Remark that the non-negativity of the *extended second variation* is a necessary condition, when the normalized adjoint covector is unique, and its coercivity is sufficient to ensure that the singular arc is a local minimizer for the restricted problem. The paper is an attempt to give second order sufficient conditions for an extremal including a finite number of both bang and singular arcs.

*Control Systems: Theory, Numerics and Applications*

*30 March - 1 April 2005*

*Rome*

---

\*Speaker.

†A footnote may follow.

‡A footnote may follow.

## 1. Introduction

The paper describes a work in progress aiming to give sufficient conditions for a bang-singular Pontryagin extremal to be a strong local optimizer for the Mayer Problem

$$\text{minimize } \alpha(\xi(0)) + \beta(\xi(T_f))$$

subject to

$$\begin{aligned} \dot{\xi}(t) &= f_0(\xi(t)) + u f_1(\xi(t)) \quad t \in [0, T_f] \\ \xi(0) &\in N_0, \quad \xi(T_f) \in N_f, \\ u &\in [-1, 1]. \end{aligned}$$

The state space is a smooth  $n$ -dimensional manifold  $M$ , and  $f_0, f_1 : M \rightarrow TM$  are smooth vector fields, by smooth we mean  $C^\infty$ .

Suppose to have a *Pontryagin normal extremal*, i.e. a couple  $(\hat{\xi}, \hat{u})$ , with associated adjoint covector

$$\hat{\lambda} : t \in [0, T_f] \rightarrow \hat{\lambda}(t) \in T^*M,$$

satisfying Pontryagin Maximum Principle (PMP). Recall that

$$\pi \hat{\lambda}(t) = \hat{\xi}(t), \quad \forall t \in [0, T_f],$$

where  $\pi : T^*M \rightarrow M$  is the canonical projection, and  $\hat{\lambda}$  is a solution of the Hamiltonian system associated to the reference Hamiltonian - with appropriate transversality conditions - and it satisfies the maximization property. To be more precise, if we denote by  $H_i$ ,  $i = 0, 1$ , the Hamiltonian function associated to the vector field  $f_i$ , that is

$$H_i : \ell \in T^*M \rightarrow \langle \ell, f_i(\pi \ell) \rangle, \quad i = 0, 1,$$

then  $\hat{\lambda}$  satisfies the Hamiltonian system associated to

$$\hat{H}_t = H_0 + \hat{u}(t) H_1,$$

and

$$\begin{aligned} \hat{\lambda}(0) &= d\alpha(\hat{\xi}(0)) \text{ on } T_{\hat{\xi}(0)} N_0, \quad \hat{\lambda}(T_f) = -d\beta(\hat{\xi}(T_f)) \text{ on } T_{\hat{\xi}(T_f)} N_f, \\ \hat{u}(t) H_1(\hat{\lambda}(t)) &= \max_{u \in [-1, 1]} u H_1(\hat{\lambda}(t)) \quad \text{a.e. } t \in [0, T_f]. \end{aligned}$$

Therefore if  $H_1(\hat{\lambda}(t)) \neq 0$ , then

$$\hat{u}(t) = \text{sgn} \left( H_1(\hat{\lambda}(t)) \right).$$

An arc of an extremal is called "bang", if the associated control is either 1 or  $-1$  and it is called singular if  $H_1$  is zero on the associated adjoint covector.

This paper is part of a general research program aiming to extend the use of Hamiltonian methods in the study of second order sufficient conditions for  $\hat{\xi}$  to be a strong local minimizer for

the problem. Here *strong local optimizer* means *optimal with respect to a neighborhood of  $\widehat{\xi}$  in  $C([0, T_f], M)$* , independently on the value of the control.

If  $\widehat{\xi}$  is bang-bang with a finite number of switching times, say  $t_1, \dots, t_r$ , then it is possible to define a second variation  $J'$  associated to the finite-dimensional problem obtained by moving the switching times. In this case PMP implies

$$H_1(\widehat{\lambda}(t_i)) = 0 \quad i = 1, \dots, r,$$

and, under the *strict bang-bang Legendre condition*

$$\langle \widehat{\lambda}(t_i), [f_0, f_1](\widehat{\xi}(t_i)) \rangle \neq 0 \quad i = 1, \dots, r,$$

$J'' > 0$  is a sufficient condition for the strong local optimality of the trajectory, see [2] and the reference therein.

If  $\widehat{\xi}$  is totally singular, then it is possible to define an *extended second variation*, whose coercivity is again a sufficient condition for the strong local optimality of the trajectory, see [6].

Here we study the case when  $\widehat{\xi}$  is the concatenation of a bang and a singular arc, namely we suppose that there is  $T_b \in (0, T_f)$  such that

$$\begin{aligned} \widehat{u}(t) &\equiv 1 & \forall t \in (0, T_b), \\ \widehat{u}(t) &\in (-1, 1) & \forall t \in (T_b, T_f). \end{aligned}$$

Nothing changes if the bang control is identically equal to  $-1$ .

Generally speaking, the Hamiltonian approach to sufficient conditions to strong optimality consists in using the coercivity of a suitable *second variation* to construct a field of non-intersecting state extremals (super-extremals) covering a neighborhood of the given trajectory. Such a field is obtained by projecting on the state manifold the flow associated to the maximized Hamiltonian  $H^{max}$  (of a Hamiltonian  $H \geq H^{max}$ ) emanating from a suitable horizontal Lagrangean sub-manifold. The sub-manifold can be obtained, if the second variation is coercive, adding a suitable penalty which allows to reduce the problem to another one without constraints on the initial point. Moreover the field of extremals can be constructed via the coercivity of the second variation of the free-fixed problem. In the case of variable final point, the existence of a field of non-intersecting extremals reduces the problem to a finite dimensional one and further conditions, coming from the coercivity of the second variation, give the result.

The described above Hamiltonian methods suggest to analyze, as a first step, the case when the initial point is free, therefore in this paper we study the problem

$$\text{minimize } \alpha(\xi(0)) + \beta(\xi(T_f)) \tag{1.1}$$

subject to

$$\begin{aligned} \dot{\xi}(t) &= f_0(\xi(t)) + u(t)f_1(\xi(t)) \quad t \in [0, T], \\ \xi(0) &\in M, \quad \xi(T_f) \in N_f, \\ u(t) &\in [-1, 1]. \end{aligned} \tag{1.2}$$

A consequence of PMP is that

$$H_1(\widehat{\lambda}(t)) = \langle \widehat{\lambda}(t), f_1(\xi(t)) \rangle \equiv 0 \quad \forall t \in [T_b, T_f], \quad (1.3)$$

$$H_{01}(\widehat{\lambda}(t)) = \langle \widehat{\lambda}(t), [f_0, f_1](\widehat{\xi}(t)) \rangle \equiv 0 \quad \forall t \in [T_b, T_f], \quad (1.4)$$

where equation (1.4) is obtained from (1.3) differentiating with respect to  $t$ .

We state the following regularity assumptions

**Assumption 1.** *The bang arc is regular in the interval  $[0, T_b)$ , i.e.*

$$H_1(\widehat{\lambda}(t)) = \langle \widehat{\lambda}(t), f_1(\widehat{\xi}(t)) \rangle > 0 \quad \forall t \in [0, T_b). \quad (1.5)$$

Differentiating with respect to time at  $t = T_b$  and using (1.3) and (1.4), we obtain

$$\langle \widehat{\lambda}(T_b), ([f_0, [f_0, f_1]] + [f_1, [f_0, f_1]]) (\widehat{\xi}(T_b)) \rangle \geq 0,$$

and we strengthen the condition requiring

**Assumption 2.**

$$\langle \widehat{\lambda}(T_b), ([f_0, [f_0, f_1]] + [f_0, [f_0, f_1]]) (\widehat{\xi}(T_b)) \rangle > 0.$$

The optimality of the reference trajectory implies that the singular arc is optimal for the following problem

$$\text{minimize } \alpha \circ \exp(-T_b f^+) (\xi(T_b)) + \beta(\xi(T_f)) \quad (1.6)$$

subject to

$$\begin{aligned} \dot{\xi}(t) &= f_0(\xi(t)) + u(t)f_1(\xi(t)) \quad t \in [T_b, T_f] \\ \xi(T_b) &\in M, \quad \xi(T_f) \in N_f, \\ u(t) &\in [-1, 1]. \end{aligned} \quad (1.7)$$

Remark that the *extended second variation* of this problem, defined in [6], has to be non-negative if the adjoint covector is unique up to a positive constant. Therefore a natural assumption is to require that it is coercive. This last condition ensures that the singular arc is optimal for the problem (1.6) - (1.7).

Allowing the trajectory to vary only on  $[0, T_b]$ , we obtain that the bang arc is optimal for the free-fixed sub-problem

$$\text{minimize } \alpha(\xi(0)) \quad (1.8)$$

subject to

$$\begin{aligned} \dot{\xi}(t) &= f_0(\xi(t)) + u(t)f_1(\xi(t)) \quad t \in [0, T_b] \\ \xi(0) &\in M, \quad \xi(T_b) = \widehat{x}_b, \\ u(t) &\in [-1, 1]. \end{aligned} \quad (1.9)$$

It is easily seen that this implies

$$L_{f_1} L_{f_1} \alpha_s(\widehat{x}_b) \geq 0.$$

Remark that the strict version of the above inequality is a consequence of the coercivity of the extended second variation of problem (1.6) - (1.7), see Condition 1 in Section 3. Moreover Assumptions 1 - 2 and  $L_{f_1} L_{f_1} \alpha_s(\widehat{x}_b) > 0$  are sufficient conditions for the local optimality of the bang arc with respect to problem (1.8) - (1.9), see Proposition 2.

Our sufficient conditions for problem (1.1) - (1.2) include an extra regularity assumption:

**Assumption 3.** *There is a neighborhood  $V$  of  $\widehat{\xi}(T_b)$  in  $M$  such that, if  $y \in V$  and  $L_{f_1}(\alpha \circ \exp - T_b f^+)(y) = 0$ , then*

$$L_{[f_0, f_1]}(\alpha \circ \exp - T_b f^+)(y) \leq 0.$$

The main result of the paper is the following:

**Theorem 1.** *Assume  $(\widehat{\xi}(\cdot), \widehat{u}(\cdot))$  is a bang-singular Pontryagin extremal of the optimal control problem (1.1) - (1.2). Let Assumptions 1, 2 and 3 be satisfied. If the extended second variation of the restricted problem (1.6) - (1.7), defined in (3.3) - (3.4), is coercive along the singular arc, then  $(\widehat{\xi}(\cdot), \widehat{u}(\cdot))$  is a strict strong local minimizer.*

## 2. Notation and preliminary remarks

We denote the reference vector field by  $\widehat{f}$ , i.e.

$$\widehat{f}_t = \begin{cases} f^+ \equiv f_0 + f_1 & \text{if } t \in [0, T_b] \\ f_0 + \widehat{u}(t)f_1 & \text{if } t \in [T_b, T_f] \end{cases} \quad (2.1)$$

and its flow from time 0 by

$$\widehat{S}: (x, t) \mapsto \widehat{S}_t(x).$$

We denote the iterated Lie brackets of the vector fields  $f_0$  and  $f_1$  by

$$f_{i_1 i_2 \dots i_k} \equiv [f_{i_1}, [\dots [f_{i_{k-1}}, f_{i_k}] \dots]] \quad i_j \in \{0, 1\}, \quad j = 1, \dots, k,$$

and the associated Hamiltonian functions by

$$H_{i_1 i_2 \dots i_k}: \ell \mapsto \langle \ell, f_{i_1 i_2 \dots i_k}(\pi \ell) \rangle, \quad j = 1, \dots, k.$$

The reference time-dependent Hamiltonian function is

$$\widehat{H}_t = \begin{cases} H_0 + H_1 & t \in [0, T_b] \\ H_0 + \widehat{u}(t)H_1 & t \in [T_b, T_f] \end{cases},$$

while the maximized Hamiltonian

$$H: \ell \mapsto \max\{H_0(\ell) + uH_1(\ell), u \in [-1, 1]\}$$

is given by

$$H: \ell \mapsto \begin{cases} (H_0 + H_1)(\ell) & H_1(\ell) > 0 \\ H_0(\ell) = \widehat{H}_t(\ell) & H_1(\ell) = 0 \\ (H_0 - H_1)(\ell) & H_1(\ell) < 0. \end{cases}$$

For each, possibly time-dependent, Hamiltonian function  $H_t$ , we denote the associated Hamiltonian vector field by

$$\vec{H}_t : T^*M \rightarrow TT^*M.$$

With this notation  $\widehat{\lambda} : [0, T_f] \rightarrow T^*M$  is solution of

$$\dot{\widehat{\lambda}}(t) = \vec{H}_t(\widehat{\lambda}(t)) \text{ on } [0, T_f]$$

and the transversality conditions are

$$\widehat{\lambda}(0) = d\alpha(\widehat{x}_0), \quad \widehat{\lambda}(T_f) = -d\beta(\widehat{x}_f) \text{ on } T_{\widehat{x}_f}N_f,$$

where  $\widehat{x}_0 \equiv \widehat{\xi}(0)$  and  $\widehat{x}_f \equiv \widehat{\xi}(T_f)$ . Without loss of generality we may assume

$$\widehat{\lambda}(T_f) = -d\beta(\widehat{x}_f) \text{ on } T_{\widehat{x}_f}M. \quad (2.2)$$

Let us also define

$$\widehat{\ell}_0 \equiv \widehat{\lambda}(0), \quad \widehat{\ell}_b \equiv \widehat{\lambda}(T_b), \quad \widehat{\ell}_f \equiv \widehat{\lambda}(T_f), \quad \widehat{x}_b \equiv \widehat{\xi}(T_b) = \pi\widehat{\ell}_b.$$

Since the assumptions of Theorem 1 imply the *strict generalized Legendre condition* (SGLC):

$$H_{101}(\widehat{\lambda}(t)) > 0, \quad \forall t \in [T_b, T_f], \quad (2.3)$$

see Remark 1 in Section 3, then we can describe the following geometric picture near the adjoint covector. Define

$$\begin{aligned} \Sigma &= \{\ell \in T^*M : H_1(\ell) = 0\}, \\ S &= \{\ell \in T^*M : H_1(\ell) = H_{01}(\ell) = 0, H_{101}(\ell) > 0\}, \end{aligned}$$

then  $\widehat{\lambda}([T_b, T_f]) \subset S \subset \Sigma$ ,  $\vec{H}_1$  is tangent to  $\Sigma$  and transversal to  $S$  and  $\vec{H}_{01}$  is transversal to  $\Sigma$ . Moreover, differentiating (1.4), we get

$$H_{001}(\widehat{\lambda}(t)) + \widehat{u}(t)H_{101}(\widehat{\lambda}(t)) = 0, \quad \forall t \in [T_b, T_f],$$

hence

$$\widehat{u}(t) = \frac{-H_{001}}{H_{101}}(\widehat{\lambda}(t)), \quad \forall t \in [T_b, T_f].$$

### 3. The singular arc

We now consider sub-problem (1.6) - (1.7), namely we consider only the admissible couples  $(\xi, u)$  such that  $u(t) = \widehat{u}(t) = 1$  for any  $t \in [0, T_b]$ . For this problem we use the results of [6].

Shifting time, denote

$$\begin{aligned} T_s &= T_f - T_b, \\ \widehat{u}_s : t \in [0, T_s] &\rightarrow \widehat{u}(t + T_b), \\ \widehat{\xi}_s : t \in [0, T_s] &\rightarrow \widehat{\xi}(t + T_b), \\ \alpha_s &= \exp -T_b f^+, \\ \widehat{\lambda}_s : t \in [0, T_s] &\rightarrow \widehat{\lambda}(t + T_b), \\ \widehat{\beta}_s &= \beta \circ \widehat{S}_{T_f} \circ \widehat{S}_{T_b}^{-1}, \\ \gamma_s &= \alpha_s + \widehat{\beta}_s. \end{aligned}$$

By the properties of the adjoint covector, using (2.2), we get  $d\gamma_s(\widehat{x}_b) = 0$ , so that

$$\gamma_s'' \equiv d^2\gamma_s(\widehat{x}_b)$$

is a well defined quadratic form on  $T_{\widehat{x}_b}M$ .

With the above stated notations, the sub-problem (1.6) - (1.7) is equivalent to

$$\text{minimize } \alpha_s(\xi(0)) + \beta(\xi(T_s)) \quad (3.1)$$

subject to

$$\begin{aligned} \dot{\xi}(t) &= f_0(\xi(t)) + u(t)f_1(\xi(t)) \quad t \in [0, T_s] \\ \xi(0) &\text{ free, } \xi(T_s) \in N_f, \\ u(t) &\in [-1, 1] \end{aligned} \quad (3.2)$$

with reference couple  $(\widehat{\xi}_s, \widehat{u}_s)$  and adjoint covector  $\widehat{\lambda}_s$ . Moreover letting

$$\begin{aligned} \widehat{N}_f &= \widehat{S}_{T_b} \circ \widehat{S}_{T_f}^{-1}(N_f), \\ h_t(x) &= (\exp T_b f^+)_*^{-1} \left( \widehat{S}_{t+T_b} \right)_*^{-1} f_1(\widehat{S}_{t+T_b} \circ \exp(-T_b f^+)(x)) \quad t \in [0, T_s], \end{aligned}$$

the second variation of problem (3.1), (3.2), as defined in [1], is given by the quadratic form on  $T_{\widehat{x}_b}M \times L^2([0, T_s])$  defined below.

**Proposition 1.** For any  $\delta e = (\delta x, v) \in T_{\widehat{x}_b}M \times L^2([0, T_s])$  the second order approximation of the cost is

$$J_s''[\delta e]^2 = \frac{1}{2} \gamma_s''[\delta x]^2 + \int_0^{T_s} v(t) L_{\eta(t)} L_{h_t} \widehat{\beta}_s(\widehat{x}_b) dt$$

where  $\eta$  satisfies

$$\begin{aligned} \dot{\eta}(t) &= v(t) h_t(\widehat{x}_b), \\ \eta(0) &= \delta x, \quad \eta(T) \in T_{\widehat{x}_b} \widehat{N}_f. \end{aligned}$$

Remark that  $J_s''$  does not depend on the choice of  $\beta$  with the property (2.2).

After the transformation

$$\Phi: v \in L^2([0, T_s]) \rightarrow (w_0, w) = \left( \int_0^{T_s} v(s) ds, t \rightarrow \int_t^{T_s} v(s) ds \right) \in \mathbb{R} \times L^2([0, T_s])$$

we obtain an equivalent formulation for the second variation. Since  $\Phi$  is dense and continuous, we consider its extension to all the space  $T_{\widehat{x}_b}M \times \mathbb{R} \times L^2([0, T_s])$ , which we call *extended second variation*  $J_e''$ , and which is given by

$$\begin{aligned} J_e''[\delta e_{se}]^2 &= \frac{1}{2} \gamma_s''(\widehat{x}_b)[\delta x]^2 + \frac{w_0^2}{2} L_{f_1} L_{f_1} \widehat{\beta}_s(\widehat{x}_b) + w_0 L_{\delta x} L_{f_1} \widehat{\beta}_s(\widehat{x}_b) \\ &\quad + \frac{1}{2} \int_0^{T_s} w^2(t) H_{101}(\widehat{\lambda}_s(t)) dt + \int_0^{T_s} L_{\zeta(t)} L_{w(t)h_t} \widehat{\beta}_s(\widehat{x}_s) dt \end{aligned} \quad (3.3)$$

where  $\zeta$  satisfies

$$\begin{aligned} \zeta(0) &= \delta x + w_0 f_1(\hat{x}_b) \\ \dot{\zeta}(t) &= w(t) \dot{h}_t(\hat{x}_b) \quad t \in [0, T_s] \\ \zeta(T_s) &\in T_{\hat{x}_b} \hat{N}_f. \end{aligned} \tag{3.4}$$

Remark that

$$\dot{h}_t = (\exp T_b f^+)_*^{-1} \left( \hat{S}_{t+T_b} \right)_*^{-1} [f_0, f_1] \circ \hat{S}_{t+T_b} \circ \exp(-T_b f^+) \quad t \in [0, T_s].$$

It is not difficult to prove, see also [6], that the Assumption “ $J'_\ell$  is coercive” is equivalent to the following two conditions:

**Condition 1.**  $L_{f_1} L_{f_1} \alpha_s(\hat{x}_b) > 0$ .

**Condition 2.** The coercivity of the quadratic form

$$\begin{aligned} J''[\delta y, w_0, w] &= \frac{1}{2} \gamma_s''(\hat{x}_b) [\delta y]^2 + \frac{w_0^2}{2} L_{f_1} L_{f_1} \hat{\beta}_s(\hat{x}_b) + w_0 L_{\delta y} L_{f_1} \hat{\beta}_s(\hat{x}_b) \\ &+ \frac{1}{2} \int_0^{T_s} w^2(t) H_{101}(\hat{\lambda}_s(t)) dt + \int_0^{T_s} L_{\zeta(t)} L_{w(t) \dot{h}_t} \hat{\beta}_s(\hat{x}_b) dt \end{aligned} \tag{3.5}$$

where  $\delta y$  and  $\zeta$  satisfy

$$\begin{aligned} L_{\delta y} L_{f_1} \alpha_s(\hat{x}_b) &= 0 \\ \zeta(0) &= \delta y + w_0 f_1(\hat{x}_b), \\ \dot{\zeta}(t) &= w(t) \dot{h}_t(\hat{x}_b) \quad t \in [0, T_s] \\ \zeta(T_s) &\in T_{\hat{x}_b} \hat{N}_f. \end{aligned} \tag{3.6}$$

**Remark 1.** Condition 1 implies that  $J''$  is an accessory linear-quadratic problem on  $T_{\hat{x}_b} M$ . Moreover the coercivity of  $J''$  implies SGLG

$$H_{101}(\hat{\lambda}_s(t)) > 0 \quad \forall t \in [0, T_s].$$

**Remark 2.** Since the initial point of our problem is free, then the normalized adjoint covector is unique, hence the conditions

$$L_{f_1} L_{f_1} \alpha_s(\hat{x}_b) \geq 0 \quad \text{and} \quad J'' \geq 0$$

are necessary for the optimality of  $\hat{\xi}_s$  in problem (3.1) - (3.2), and hence in the original problem.

If we define

$$M_s \equiv \{x \in M : L_{f_1} \alpha_s(x) = 0\}$$

it is clear that  $\hat{x}_b \in M_s$  and, by Condition 1, that  $f_1$  is transversal to  $M_s$  in  $\hat{x}_b$ . Therefore there exist a neighborhood  $O(\hat{x}_b)$  of  $\hat{x}_b$  in  $M$ , a positive number  $\bar{b}$  and a neighborhood  $O_s(\hat{x}_b)$  of  $\hat{x}_b$  in  $M_s$  such that the functions

$$b : O(\hat{x}_b) \rightarrow (-\bar{b}, \bar{b}) \quad \text{and} \quad y : O(\hat{x}_b) \rightarrow O_s(\hat{x}_b)$$

are well defined by

$$x = \exp b(x) f_1(y(x)) \quad \forall x \in O(\hat{x}_b).$$

In [6] it is proved that defining  $\tilde{\alpha}_s: O(\hat{x}_b) \rightarrow \mathbb{R}$  by

$$\tilde{\alpha}_s(x) = \alpha_s(y(x)),$$

then  $J''$  turns out to be the standard second variation of the following problem

$$\text{minimize } \tilde{\alpha}_s(\xi(0)) + \beta(\xi(T_s))$$

subject to

$$\begin{aligned} \dot{\xi}(t) &= \widehat{f}_{T_b+t}(\xi(t)) + w(t)[f_0, f_1](\xi(t)) + \frac{w^2(t)}{2}[f_1, [f_0, f_1]](\xi(t)) \quad t \in [0, T_s] \\ \xi(0) &\text{ free, } \xi(T_s) \in N_f, \\ u(t) &\in [-1, 1], \end{aligned}$$

with reference couple  $(w, \xi) = (0, \widehat{\xi}_s)$ .

From the results in [6], the coercivity of  $J''$  implies, for any solution of system (3.2) belonging to a suitable neighborhood of the reference trajectory,

$$\tilde{\alpha}_s(\xi(0)) + \beta(\xi(T_s)) \geq \tilde{\alpha}_s(\widehat{x}_b) + \beta(\widehat{x}_f) = \alpha_s(\widehat{x}_b) + \beta(\widehat{x}_f),$$

moreover equality holds if and only if  $\xi = \widehat{\xi}_s$ .

**Remark 3.** Since  $J''$  is associated to a linear quadratic non singular control problem, then its coercivity can be checked via the non existence of conjugate times or through an appropriate Riccati equation, see for example [4], [7], [5] and the references therein.

Rewriting the result with respect to Problem (1.1) - (1.2), possibly restricting the neighborhood  $O(\widehat{x}_b)$  we can state the following

**Lemma 1.** Let the extended second variation defined in (3.5) - (3.6) be coercive. There is a neighborhood  $\mathcal{V}$  of the graph of  $\widehat{\xi}_{[T_b, T_f]}$  such that  $\mathcal{V} \cap \{T_b\} \times M = O(\widehat{x}_0)$  and for any solution  $\xi$  of (1.2) such that the graph of  $\xi_{[T_b, T_f]}$  is contained in  $\mathcal{V}$  the following inequality holds

$$\tilde{\alpha}_s(\xi(T_b)) + \beta(\xi(T_f)) \geq \alpha_s(\widehat{x}_b) + \beta(\widehat{x}_f).$$

Moreover the equality holds if and only if  $\xi = \widehat{\xi}$  on  $[T_b, T_f]$ .

#### 4. The bang arc

In this section we consider the geometric picture near the bang arc and its lift  $\widehat{\lambda}_{|[0, T_b]}$  to the cotangent bundle.

If for any  $t \in [0, T_b]$  we define the vector field

$$g_t \equiv \widehat{S}_{t*}^{-1} f_1 \circ \widehat{S}_t,$$

then a straightforward calculation gives

$$\dot{g}_t = \widehat{S}_{t*}^{-1}[f_0, f_1] \circ \widehat{S}_t, \quad \ddot{g}_t = \widehat{S}_{t*}^{-1}[f_0 + f_1, [f_0, f_1]] \circ \widehat{S}_t \quad \forall t \in [0, T_b].$$

**Remark 4.** Since

$$L_{g_{T_b}} L_{g_{T_b}} \alpha(\hat{x}_0) = L_{f_1} L_{f_1} \alpha_s(\hat{x}_b) > 0, \quad (4.1)$$

$g_{T_b}$  is transversal in  $\hat{x}_0$  to

$$M_0 \equiv \exp(-T_b f^+) (M_s) = \left\{ z \in M : L_{g_{T_b}} \alpha(z) = 0 \right\}.$$

Moreover Assumption 2 is equivalent to

$$L_{\ddot{g}_{T_b}} \alpha(\hat{x}_0) > 0. \quad (4.2)$$

As we pointed out in the introduction,  $L_{g_{T_b}} L_{g_{T_b}} \alpha(\hat{x}_0) \geq 0$  is a necessary condition for the optimality of the bang arc for Problem (1.8) - (1.9). This can be easily seen calculating the second variation of the free-fixed problem on the bang arc constrained to controls  $u$  such that

$$u(t) = \begin{cases} 1 & t \in [0, T_b - \varepsilon] \\ -1 & t \in [T_b - \varepsilon, T_b]. \end{cases}$$

In the following proposition we show that Assumptions 1 and 2, together with (4.1) are sufficient conditions for  $\hat{\xi}_{|[0, T_b]}$  to be a locally optimal trajectory for the free-fixed problem restricted to the bang arc.

**Proposition 2.** *If Assumptions 1, 2 and inequality (4.1) are verified, then  $\hat{\xi}_{|[0, T_b]}$  is a strict strong local optimizer for Problem (1.8) - (1.9).*

*Proof:* From (4.1) and (4.2), by continuity, there exist  $\varepsilon > 0$  and a neighborhood  $\mathcal{U}$  of  $\hat{x}_0$  in  $M$  such that the flow  $(t, x) \rightarrow \exp t f^+(x)$  is defined on  $[0, T_b] \times \mathcal{U}$ , for all  $(t, x) \in [T_b - \varepsilon, T_b] \times \mathcal{U}$  the following inequalities hold

$$L_{g_t} L_{g_t} \alpha(x) \geq \frac{1}{2} L_{g_{T_b}} L_{g_{T_b}} \alpha(\hat{x}_0) \quad (4.3)$$

$$L_{\ddot{g}_t} \alpha(x) \geq \frac{1}{2} L_{\ddot{g}_{T_b}} \alpha(\hat{x}_0) \quad (4.4)$$

$$4\varepsilon L_{g_t} L_{\dot{g}} \alpha(x) < L_{g_{T_b}} L_{g_{T_b}} \alpha(\hat{x}_0), \quad (4.5)$$

and

$$L_{g_t} \alpha(x) = H_1(\exp t \vec{H}^+(\mathrm{d}\alpha(x))) > 0 \quad \forall (t, x) \in [0, T_b - \varepsilon] \times \mathcal{U}. \quad (4.6)$$

We denote

$$K \equiv \max \{ |L_{g_t} L_{\dot{g}} \alpha(x)| : (t, x) \in [T_b - \varepsilon, T_b] \times \mathcal{U} \}.$$

We want to prove that  $\alpha(\xi(0)) - \alpha(\hat{x}_0) \geq 0$  for any couple  $(\xi, u)$  such that  $\hat{\xi}(T_b) = \hat{x}_b$  and  $\xi(s) \in \exp s f^+(\mathcal{U})$  for any  $s \in [0, T_b]$ . For such a couple let us define

$$v_\varepsilon(t) \equiv \int_{T_b - \varepsilon}^t (1 - u(s)) \, ds, \quad q(t) \equiv \exp(-t f^+)(\xi(t)),$$

hence

$$\dot{q}(t) = -(1 - u(t)) g_t(q(t)), \quad q(0) = \xi(0), \quad q(T_b) = \hat{x}_0,$$

and we can evaluate

$$\begin{aligned}\alpha(\xi(0)) - \alpha(\hat{x}_0) &= \int_{T_b}^0 L_{\cdot q(s)} \alpha(q(s)) \, ds = \int_0^{T_b} (1 - u(s)) L_{g_s} \alpha(q(s)) \, ds \\ &= \int_0^{T_b} (1 - u(s)) H_1(\exp s \vec{H}^+(\mathrm{d}\alpha(q(s)))) \, ds \geq \int_{T_b - \varepsilon}^{T_b} (1 - u(s)) L_{g_s} \alpha(q(s)) \, ds.\end{aligned}\quad (4.7)$$

Integrating by parts and recalling that  $q(T_b) = \hat{x}_0$  we get

$$\begin{aligned}\alpha(\xi(0)) - \alpha(\hat{x}_0) &\geq - \int_{T_b - \varepsilon}^{T_b} v_\varepsilon(s) [L_{\cdot g} \alpha(q(s)) - \dot{v}_\varepsilon(s) L_{g_s} L_{g_s} \alpha(q(s))] \, ds \\ &\geq \int_{T_b - \varepsilon}^{T_b} v_\varepsilon(s) \dot{v}_\varepsilon(s) \frac{1}{2} L_{g_{T_b}} L_{g_{T_b}} \alpha(\hat{x}_0) \, ds - \int_{T_b - \varepsilon}^{T_b} v_\varepsilon(s) L_{\cdot g} \alpha(q(s)) \, ds \\ &\geq \frac{v_\varepsilon^2(T_b)}{4} L_{g_{T_b}} L_{g_{T_b}} \alpha(\hat{x}_0) + \frac{L_{\ddot{g}_{T_b}} \alpha(\hat{x}_0)}{2} \int_{T_b - \varepsilon}^{T_b} \left( \int_{T_b - \varepsilon}^s v_\varepsilon(t) \, dt \right) \, ds \\ &\quad - K \int_{T_b - \varepsilon}^{T_b} \left( \int_{T_b - \varepsilon}^s v_\varepsilon(t) \, dt \right) \dot{v}_\varepsilon(s) \, ds\end{aligned}$$

For the last addendum, we have

$$\begin{aligned}K \int_{T_b - \varepsilon}^{T_b} \left( \int_{T_b - \varepsilon}^s v_\varepsilon(t) \, dt \right) \dot{v}_\varepsilon(s) \, ds &= K \int_{T_b - \varepsilon}^{T_b} v_\varepsilon(t) \left( \int_t^{T_b} \dot{v}_\varepsilon(s) \, ds \right) \, dt \\ &\leq K \int_{T_b - \varepsilon}^{T_b} v_\varepsilon(t) v_\varepsilon(T_b) \, dt \leq K \varepsilon v_\varepsilon^2(T_b),\end{aligned}$$

so that

$$\alpha(\xi(0)) - \alpha(\hat{x}_0) \geq \frac{v_\varepsilon^2(T_b)}{4} \left( L_{g_{T_b}} L_{g_{T_b}} \alpha(\hat{x}_0) - 4K\varepsilon \right) \geq 0 \quad (4.8)$$

because of our choice of  $\varepsilon$ . To prove that the minimum is strict let us assume  $\alpha(\xi(0)) = \alpha(\hat{x}_0)$ . If this equality holds, then from (4.8) we get  $v_\varepsilon(T_b) = 0$ , i.e.  $u(s) = 1$  for any  $s \in [T_b - \varepsilon, T_b]$ , so that  $\xi(s) \equiv \hat{\xi}(s)$  for any  $s \in [T_b - \varepsilon, T_b]$ . Since  $H_1(\exp t \vec{H}^+(\mathrm{d}\alpha(q(s)))) > 0$  for any  $s \in [0, T_b - \varepsilon]$ , equation (4.7) implies  $u(s) = 1$  also for any  $s \in [0, T_b - \varepsilon]$ .  $\square$

Let us now go back to problem (1.1) - (1.2). Without loss of generality we can assume that  $\exp(-T_b f^+)$  is well defined on  $O(\hat{x}_b)$ , therefore, defining  $O(\hat{x}_0) = \exp(-T_b f^+)(O(\hat{x}_b))$  and  $O_s(\hat{x}_0) = \exp(-T_b f^+)(O_s(\hat{x}_b))$  we have that for any  $x \in O(\hat{x}_0)$  there exists one and only one couple  $(a, z) \in (-\bar{b}, \bar{b}) \times O_s(\hat{x}_0)$  such that

$$x = \exp a g_{T_b}(z).$$

**Remark 5.** Possibly restricting  $O_s(\hat{x}_0)$ , Assumption 3 is equivalent to

$$L_{\cdot g_b} \alpha(z) \leq 0 \quad \forall z \in O_s(\hat{x}_0).$$

The map

$$\tilde{\alpha}: O(\hat{x}_0) \rightarrow \tilde{\alpha}_s \circ \exp T_b f^+.$$

is constant along the integral lines of  $g_{T_b}$ , so that

$$L_{g_{T_b}} \tilde{\alpha}(x) = 0 \text{ for any } x \in O(\hat{x}_0).$$

Moreover it is not difficult to see that  $L_{\mathbf{v}}\alpha(z) = L_{\mathbf{v}}\tilde{\alpha}(z)$  for any  $z \in O_s(\hat{x}_0)$  and for any  $\mathbf{v} \in T_zM$ . Without loss of generality we can assume that  $O(\hat{x}_0)$  is contained in the neighborhood  $\mathcal{U}$ , defined in the proof of Proposition 2; therefore from now on we assume that equations (4.3) - (4.4) - (4.5) and (4.6) hold. Moreover, possibly restricting  $\varepsilon > 0$  and  $O(\hat{x}_0)$  we assume

1.  $\alpha(x) \geq \tilde{\alpha}(x)$  for any  $x$  in  $O(\hat{x}_0)$ ;
2.  $L_{g_{T_b}}L_{g_{T_b}}\alpha(z) - \varepsilon L_{\exp(-ag_{T_b}) * g_t \circ \exp ag_{T_b}}L_{g_{T_b}}\alpha(z) > 0 \quad \forall (t, x) \in [T_b - \varepsilon, T_b] \times O(\hat{x}_0)$ ;
3.  $L_{\dot{g}_t}\tilde{\alpha}(x) > 0 \quad \forall (t, x) \in [T_b - \varepsilon, T_b] \times O(\hat{x}_0)$ ;
4.  $L_{[g_{T_b}, \dot{g}_b]}\tilde{\alpha}(x) > 0 \quad \forall x \in O(\hat{x}_0)$ .

With such a choice of  $\varepsilon$  and  $O(\hat{x}_0)$ , let us define

$$M: (t, a, z) \in [0, T_b] \times (-\bar{b}, \bar{b}) \times O_s(\hat{x}_0) \rightarrow H_1(\exp t \vec{H}^+(\mathrm{d}\alpha(\exp ag_{T_b}(z)))) \in \mathbb{R},$$

$$\tilde{M}: (t, a, z) \in [0, T_b] \times (-\bar{b}, \bar{b}) \times O_s(\hat{x}_0) \rightarrow H_1(\exp t \vec{H}^+(\mathrm{d}\tilde{\alpha}(\exp ag_{T_b}(z)))) \in \mathbb{R}.$$

Assumption 3 is crucial in the proofs of the following two ‘‘twin’’ Lemmata.

**Lemma 2.** *If Assumption 3 holds, then*

$$M(t, a, z) \geq 0 \quad \text{for any } (t, a, z) \in (T_b - \varepsilon, T_b) \times [0, \bar{b}] \times O_s(\hat{x}_0).$$

*Proof:* Let us write the Taylor expansion of  $M$  in a neighborhood of  $(t, a, z) = (T_b, 0, \hat{x}_0)$ .

$$\begin{aligned} M(t, a, z) &= L_{g_t}\alpha(\exp ag_{T_b}(z)) = L_{g_{T_b}}\alpha(\exp ag_{T_b}(z)) \\ &+ (t - T_b)L_{\dot{g}_b}\alpha(\exp ag_{T_b}(z)) + \frac{(t - T_b)^2}{2}L_{\ddot{g}_\theta}\alpha(\exp ag_{T_b}(z)) \quad (\text{for some } \theta \in (t, T_b)) \\ &= L_{g_{T_b}}\alpha(z) + aL_{g_{T_b}}L_{g_{T_b}}\alpha(\exp \beta_1 g_{T_b}(z)) + (t - T_b)L_{\dot{g}_b}\alpha(\exp ag_{T_b}(z)) \\ &+ \frac{(t - T_b)^2}{2}L_{\ddot{g}_\theta}\alpha(\exp ag_{T_b}(z)) \quad (\text{for some } \beta_1 \text{ intermediate between } 0 \text{ and } a) \\ &= a \left[ L_{g_{T_b}}L_{g_{T_b}}\alpha(\exp \beta_1 g_{T_b}(z)) + (t - T_b)L_{g_{T_b}}L_{\dot{g}_b}\alpha(\exp \beta_2 g_{T_b}(z)) \right] \\ &+ (t - T_b) \left[ L_{\dot{g}_b}\alpha(z) + \frac{t - T_b}{2}L_{\ddot{g}_\theta}\alpha(\exp ag_{T_b}(z)) \right] \\ &(\text{for some } \beta_2 \text{ intermediate between } 0 \text{ and } a). \end{aligned}$$

Hence Assumption 3 and equation (4.5) give the result.  $\square$

**Lemma 3.** *If Assumption 3 holds, then*

$$\tilde{M}(t, a, z) \geq 0 \quad \text{for any } (t, a, z) \in (T_b - \varepsilon, T_b) \times [-\bar{b}, 0] \times O_s(\hat{x}_0).$$

*Proof:* Repeating the proof of Lemma 2 we get

$$\tilde{M}(t, a, z) = (t - T_b) \left[ 2aL_{[g_{T_b}, \dot{g}_b]}\tilde{\alpha}(\exp \beta_2 g_{T_b}(z)) + \frac{t - T_b}{2}L_{\ddot{g}_\theta}\tilde{\alpha}(\exp ag_{T_b}(z)) + 2L_{\dot{g}_b}\tilde{\alpha}(z) \right] \quad (4.9)$$

for some  $\theta \in (T_b - \varepsilon, T_b)$  and some  $\beta_2$  intermediate between 0 and  $a$ . Due to the restriction on  $\varepsilon > 0$  and  $O(\hat{x}_0)$ , we get the result.  $\square$

The following Lemma describes the behaviour of the trajectories of control system (1.2) and its proof does not require Assumption 3.

**Lemma 4.** *Let  $(\xi, u)$  be an admissible couple such that  $\xi(t) \in \text{expt}f^+(O(\hat{x}_0))$  for any  $t \in [0, T_b]$ . Define  $q(t) \equiv \exp(-tf^+)(\xi(t)) \in O(\hat{x}_0)$  and let  $(a(t), z(t)) \in (-\bar{b}, \bar{b}) \times O_s(\hat{x}_0)$  be defined by  $q(t) = \exp a(t)g_{T_b}(z(t))$ . Then  $a: t \in [T_b - \varepsilon, T_b] \rightarrow \mathbb{R}$  is a monotone non-increasing function.*

*Proof:* Differentiating the identity  $z(t) = \exp(-a(t)g_{T_b})(q(t))$  we get

$$\dot{z}(t) = -\dot{a}(t)g_{T_b}(z(t)) - (1 - u(t))\exp(-a(t)g_{T_b})_*g_t \circ \exp a(t)g_{T_b}(z(t))$$

Also differentiating the identity  $L_{g_{T_b}}\alpha(z(t)) = 0$ , we get  $L_{z(t)}L_{g_{T_b}}\alpha(z(t)) = 0$ . Hence we have

$$-\dot{a}(t)L_{g_{T_b}}L_{g_{T_b}}\alpha(z(t)) - (1 - u(t))L_{\exp(-a(t)g_{T_b})_*g_t \circ \exp a(t)g_{T_b}}L_{g_{T_b}}\alpha(z(t)) = 0$$

which gives

$$\dot{a}(t) = -(1 - u(t)) \left\{ 1 + (t - T_b) \frac{L_{\exp(-a(t)g_{T_b})_*g_t \circ \exp a(t)g_{T_b}}L_{g_{T_b}}\alpha(z(t))}{L_{g_{T_b}}L_{g_{T_b}}\alpha(z(t))} \right\}$$

which is non-positive due to the restrictions on  $\varepsilon$  and  $O(\hat{x}_0)$ .  $\square$

## 5. Proof of Theorem 1

Let  $(\xi, u)$  be an admissible couple such that  $\xi(t) \in \text{expt}f^+(O(\hat{x}_0))$  for any  $t \in [0, T_b]$  and the graph of  $\xi|_{[T_b, T_f]}$  belongs to the neighborhood  $\mathcal{V}$  defined in Lemma 1. We want to show

$$\alpha(\xi(0)) - \alpha(\hat{x}_0) \geq \tilde{\alpha}_s(\xi(T_b)) - \alpha_s(\hat{x}_b), \quad (5.1)$$

or, equivalently with the notation of Lemma 4,

$$\alpha(q(0)) \geq \tilde{\alpha}(q(T_b)).$$

Since

$$\alpha(q(0)) - \alpha(q(T_b)) = \int_{T_b}^0 \frac{d}{ds}\alpha(q(s)) ds = \int_0^{T_b} (1 - u(s))M(s, a(s), z(s)) ds,$$

then, if  $M(s, a(s), z(s)) \geq 0, \forall s \in [0, T_b]$ , we get

$$\alpha(q(0)) \geq \alpha(q(T_b)) \geq \tilde{\alpha}(q(T_b)).$$

If there exists  $\bar{t} \in (T_b - \varepsilon, T_b)$  such that  $M(\bar{t}, a(\bar{t}), z(\bar{t})) = 0$  then, from Lemma 2,  $a(\bar{t}) < 0$ ,

$$\alpha(q(0)) \geq \alpha(q(\bar{t})) + \int_0^{\bar{t}} (1 - u(s))M(s, a(s), z(s)) ds \geq \alpha(q(\bar{t})) \geq \tilde{\alpha}(q(\bar{t}))$$

so that

$$\tilde{\alpha}(q(\bar{t})) - \tilde{\alpha}(q(T_b)) = \int_{\bar{t}}^{T_b} (1 - u(s))\tilde{M}(s, a(s), z(s)) ds. \quad (5.2)$$

Since  $a(s)$  is decreasing by Lemma 4 and  $a(\bar{t}) \leq 0$ , applying Lemma 3 equation (5.1) is proved.

Considering also the result of Lemma 1 we obtain

$$\begin{aligned} \alpha(\xi(0)) + \beta(\xi(T_f)) - (\alpha(\hat{x}_0) + \beta(\hat{x}_f)) &= \\ &= [\alpha(\xi(0)) - \tilde{\alpha}_s(\xi(T_b))] + [\tilde{\alpha}_s(\xi(T_b)) + \beta(\xi(T_f)) - \alpha_s(\hat{x}_b) - \beta(\hat{x}_f)] \geq 0. \end{aligned} \quad (5.3)$$

To complete the proof we have to show that the minimum is strict. Assume equality holds in (5.3), by equation (5.1) and Lemma 1 we obtain

$$\tilde{\alpha}_s(\xi(T_b)) + \beta(\xi(T_f)) - \alpha_s(\hat{x}_b) - \beta(\hat{x}_f) = 0 \quad (5.4)$$

$$\alpha(\xi(0)) - \tilde{\alpha}_s(\xi(T_b)) = 0 \quad (5.5)$$

From equation (5.4) and Lemma 1 we obtain

$$\xi = \hat{\xi} \text{ on } [T_b, T_f]$$

hence  $\xi(T_b) = \hat{x}_b$ . Finally, from equation (5.5) and Proposition 2 we obtain

$$\xi = \hat{\xi} \text{ on } [0, T_b].$$

## 6. Final remarks

The result of Theorem 1 is not completely satisfactory because of Assumption 3. Indeed it is not clear if it is "almost necessary" like the other ones. Assumption 3 allows us to prove that "near" the singular point  $\hat{\lambda}(T_b)$ , the flow associated to the Hamiltonian  $H^+$  starting from the union of two suitable "half Lagrangean sub-manifolds" remains in the region where  $H^+ = H^{max}$  for all  $t \in [0, T_b]$ , see Lemmata 2, 3. Therefore the flow of the vector field  $f^+$  starting from a neighborhood of  $\hat{\xi}(0)$  is a flow of non intersecting extremals covering a neighborhood of the bang arc.

So far we have not been able to find a flow with such properties without Assumption 3. More investigation is in progress in order to understand the role of Assumption 3. However, by means of this result it is possible to prove sufficient conditions for a bang-singular arc in the minimum time problem and in the Bolza problem with fixed end-points requiring the extended second variation on the singular arc to be coercive on a larger space.

## References

- [1] Andrei A. Agrachev, Gianna Stefani, and PierLuigi Zezza. An invariant second variation in optimal control. *Internat. J. Control*, 71(5):689–715, 1998.
- [2] Andrei A. Agrachev, Gianna Stefani, and PierLuigi Zezza. Strong optimality for a bang-bang trajectory. *SIAM J. Control Optimization*, 41:991–1014, 2002.
- [3] Magnus R. Hestenes. Application of the theory of quadratic forms in Hilbert spaces to calculus of variations. *Pacific J. Math.* 1:521–581, 1951.
- [4] Brian P. Molinari. Non-negativity of a quadratic functional. *SIAM J. Control*, 13:262–271, 1975.
- [5] Gianna Stefani and PierLuigi Zezza. Constrained Regular LQ-Control Problems. *SIAM J. Control Optimization*, 35:876–900, 1997.
- [6] Gianna Stefani. Optimality conditions for a singular extremal: the single-input case. *Preprint*, 2004.
- [7] Vera Zeidan. The Riccati equation for optimal control problems with mixed state-control constraints: Necessity and sufficiency. *SIAM J. Control Optimization*, 32:1297–1321, 1994.