Control and Cybernetics

vol. **37** (2008) No. 2

Sufficient optimality conditions for a bang–singular extremal in the minimum time problem *

by

Laura Poggiolini and Gianna Stefani

Dipartimento di Matematica Applicata "G. Sansone",
Università degli Studi di Firenze, Italy

Abstract: The paper gives second order sufficient conditions for the strong local optimality of a bang–singular extremal in a minimum time problem. The conditions are given in terms of regularity assumptions on the extremal and of the coercivity of the extended second variation associated to the minimum time problem with fixed end-points on the singular arc. The conditions are *close* to the necessary ones in the usual sense, namely we require strict inequalities where necessary conditions have mild inequalities.

Keywords: minimum time, second order sufficient conditions, bang–singular arc, Hamiltonian formalism.

1. Introduction

This paper is part of a research project aiming to use a Hamiltonian approach to study second order conditions in optimal control. Indeed, we prove sufficient second order conditions for the strong local optimality of a bang–singular extremal for the minimum time problem between fixed end–points with dynamics given by

$$\dot{\xi}(t) = f_0(\xi(t)) + u f_1(\xi(t)) \tag{1}$$

$$u \in [-1, 1] \tag{2}$$

and constrained to

$$\xi(0) = \widehat{x}_0 \,, \quad \xi(T) = \widehat{x}_f . \tag{3}$$

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

We study the strong local optimality of a reference triplet $(\widehat{T}, \widehat{\xi}, \widehat{u})$, according to the following

*Submitted: December 2007; Accepted: April 2008.

Definition 1 *The trajectory* $\widehat{\xi}$ *is* a strong local minimizer *if there is a neighborhood* $\mathcal{V}$ *of its graph in* $\mathbb{R} \times M$ *such that* $\widehat{\xi}$ *is a minimizer with respect to those admissible trajectories whose graph belongs to* $\mathcal{V}$*, independently of the values of the associated controls.*

Note that this kind of optimality is local with respect to both time and space.

In this paper we consider the strong local optimality of a reference trajectory associated to a bang–singular control, namely we study a reference triplet $(\widehat{T}, \widehat{\xi}, \widehat{u})$ satisfying (1), (2), (3) and such that $\widehat{u}$ has the following structure:

$$\widehat{u}(t) \equiv u_1 \in \{-1, 1\} \qquad \forall t \in [0, \widehat{\tau}) \tag{4}$$

$$\widehat{u}(t) \in (-1, 1) \qquad \forall t \in (\widehat{\tau}, \widehat{T})\,, \tag{5}$$

$\widehat{\tau}$ is called *switching time* of the reference control $\widehat{u}$.

Thanks to the structure of $\widehat{u}$ we are able to prove that the sufficient conditions for local optimality of $(\widehat{T}, \widehat{\xi}, \widehat{u})$ with end–point constraint (3) are sufficient also for the minimum time problem with the same dynamics and relaxing the final constraint to the integral curve Γ_f of the controlled vector field f_1 through the final point of the reference trajectory. Namely, we give second order sufficient optimality conditions for the problem

$$\text{minimize} \quad T \tag{6}$$

subject to (1), (2) and

$$\xi(0) = \widehat{x}_0, \qquad \xi(T) \in \Gamma_f := \{\exp s f_1(\widehat{x}_f) \colon s \in \mathbb{R}\}. \tag{7}$$

In this paper we consider the case when there is a *feedback singular control*, i.e. there is a smooth function $v_S \colon M \to \mathbb{R}$ such that any singular state–extremal ξ associated to the dynamics (1) is an integral line of the vector field $f_S = f_0 + v_S f_1$ and any singular control can be written as $v_S \circ \xi$. Note that this case is generic in dimension 2 and 3, see the subsequent Section 2.5. Under this assumption we have that both the reference trajectory $\widehat{\xi}$ and the reference control $\widehat{u}$ are smooth on the singular arc and that the reference vector field can be written as

$$\widehat{f}_t = \begin{cases} h_1 := f_0 + u_1 f_1 & t \in [0, \widehat{\tau}) \\ f_S & t \in (\widehat{\tau}, \widehat{T}]. \end{cases}$$

The Hamiltonian approach to sufficient conditions in Optimal Control corresponds to the classical construction in Calculus of Variations of a field of non–intersecting state extremals covering a neighborhood of the reference trajectory, see for example Giaquinta and Hildebrandt (1996). The idea is to compare the costs of admissible trajectories, independently of the associated controls, by lifting them to the cotangent bundle.

When the maximized Hamiltonian $H^{\max}$ is sufficiently smooth, the field of extremals can be obtained by projecting on the state manifold the flow of $\overrightarrow{H}^{\max}$ emanating from a suitable Lagrangian sub–manifold. Then the coercivity of the second variation permits to invert the projection and lift the admissible trajectories, see Agrachev et al. (1998B, 2002) and the references therein.

An important feature is that the comparison of costs can also be obtained with a Hamiltonian, which is greater than or equal to $H^{\max}$. This possibility has been employed in Stefani (2004 and 2007) to prove second order conditions for a totally singular trajectory.

In addition to the first order necessary optimality conditions, in the form of Pontryagin Maximum Principle (Assumption 1), we make regularity assumptions on the bang arc (Assumption 2), and the *strengthened generalized Legendre condition (SGLC)* (Assumption 3). Moreover, to overcome the difficulties introduced by the junction point $\widehat{\xi}(\widehat{\tau})$ between the bang and the singular arc, we introduce a second–order regularity condition at the junction point (Assumption 4). This condition, under SGLC, is proved to be equivalent to the discontinuity of $\widehat{u}$ at time $\widehat{\tau}$. These regularity conditions allow us to define a new Hamiltonian, greater than or equal to $H^{\max}$, smooth enough to apply the Hamitonian approach.

As far as the second variation is concerned, we require the coercivity of the *extended second variation* for the minimum time problem with fixed end–points along the singular arc, as defined in Section 4.

Some partial result was already obtained in Poggiolini and Stefani (2005, 2006), but there the conditions were not completely satisfactory, since stronger second order conditions were required.

For references on singular trajectories see Goh (1966), Gabasov and Kirillova (1972), Gardner Moyer (1973), Dmitruk (1977, 1983), Agrachev and Sachkov (2004), and the references therein. To our knowledge sufficient conditions for strong local optimality do not appear in the literature even for totally singular trajectories.

The plan of the paper is as follows: in Section 2 we give the notation used, the assumptions and the main result of the paper, in Section 3 we describe the Hamiltonian approach to strong optimality, we define a suitable time–dependent Hamiltonian and we prove its properties which allow us to use the method in the present case, in Section 4 we define the coordinate–free extended second variation J_E'' and describe its properties. Finally, in Section 5 we prove the main result.

2. Statement of the results

2.1. Notation

In this paper we use some basic element of the theory of symplectic manifolds related to the cotangent bundle T^*M. For a general introduction see

Arnold (1980), for specific application to Control Theory we refer to Agrachev and Gamkrelidze (1997) and Agrachev and Sachkov (2004). Let us recall some basic facts and introduce some specific notations.

Denote by $\pi\colon T^*M \to M$ the canonical projection, the space $T^*_{\pi\ell}M$ is canonically embedded in $T_\ell T^*M$ as the space of tangent vectors to the fibers.

The canonical Liouville one–form $\boldsymbol{s}$ on T^*M and the associated canonical symplectic two–form $\boldsymbol{\sigma} = \mathrm{d}\boldsymbol{s}$ make it possible to associate to any, possibly time-dependent, smooth Hamiltonian $H_t\colon T^*M \to \mathbb{R}$, a Hamiltonian vector field $\overrightarrow{H}_t$, by

$$\boldsymbol{\sigma}(v, \overrightarrow{H}_t(\ell)) = \langle \mathrm{d}H_t(\ell)\, , \, v\rangle, \quad \forall v \in T_\ell T^*M.$$

In this paper, time $\widehat{\tau}$ plays a special role, hence we consider all the flows as starting at time $\widehat{\tau}$. In particular, we denote the flow of $\overrightarrow{H}_t$ from time $\widehat{\tau}$ to time t by

$$\mathcal{H} : (t, \ell) \mapsto \mathcal{H}(t, \ell) = \mathcal{H}_t(\ell).$$

Since no misunderstanding can occur, we shall also say that $\mathcal{H}$ is the flow of the Hamiltonian H_t. We keep this notation throughout the paper, namely the overhead arrow denotes the vector field associated to a Hamiltonian and the script letter denotes its flow from time $\widehat{\tau}$.

Remark 1 *If $M = \mathbb{R}^n$, then*

$$T^*M = (\mathbb{R}^n)^* \times \mathbb{R}^n = \left\{(p_1, \dots, p_n, q^1, \dots, q^n),\ p_i,\ q^i \in \mathbb{R}\right\},$$

and $\boldsymbol{s} = \sum_{i=1}^n p_i \,\mathrm{d}q^i, \quad \boldsymbol{\sigma} = \sum_{i=1}^n \mathrm{d}p_i \wedge \mathrm{d}q^i.$

With this notation, $\mathcal{H}_t(p_0, q_0) = (\mu(t), \xi(t))$ is the solution at time t of the Hamiltonian system

$$\begin{cases} \dot{\mu}_i(t) = -\dfrac{\partial}{\partial q^i} H_t(\mu(t), \xi(t)) \\ \dot{\xi}^i(t) = \dfrac{\partial}{\partial p_i} H_t(\mu(t), \xi(t)) \\ \mu(\widehat{\tau}) = p_0, \qquad \xi(\widehat{\tau}) = q_0. \end{cases}$$

For a general manifold M, this is the notation in local coordinates.

Finally, recall that any vector field f on the manifold M defines, by lifting to the cotangent bundle, a Hamiltonian

$$F\colon \ell \in T^*M \mapsto \langle \ell\, , \, f(\pi\ell)\rangle \in \mathbb{R},$$

in coordinates $\ell = (p, q)$ and $F(p, q) = \langle p\, , \, f(q)\rangle$.

We denote by $F_0,\ F_1,\ F_S,\ H_1$ the Hamiltonians associated to $f_0,\ f_1,\ f_S,\ h_1$, respectively and by

$$F_{i_1 i_2 \ldots i_k} := \{F_{i_1}, \{\ldots\{F_{i_{k-1}}, F_{i_k}\}\ldots\}, \quad i_1, \ldots, i_k \in \{0, 1, S\}$$

the Hamiltonian associated to $f_{i_1 i_2 \ldots i_k} := [f_{i_1}, [\ldots[f_{i_{k-1}}, f_{i_k}]\ldots]$, where $\{\cdot,\cdot\}$ denotes the Poisson parentheses between Hamiltonians and $[\cdot,\cdot]$ denotes the Lie brackets between vector fields.

The flow from time $\widehat{\tau}$ of the reference vector field $\widehat{f}_t$ is a map locally defined in a neighborhood of $\widehat{x} := \widehat{\xi}(\widehat{\tau})$. We denote it as

$$\widehat{S}_t \colon x \mapsto \begin{cases} \exp(t-\widehat{\tau})h_1(x) & \text{if } \ t \in [0, \widehat{\tau}] \\ \exp(t-\widehat{\tau})f_S(x) & \text{if } \ t \in [\widehat{\tau}, \widehat{T}] \end{cases}$$

while

$$\widehat{F}_t = \begin{cases} H_1 & \text{if } \ t \in [0, \widehat{\tau}] \\ F_S & \text{if } \ t \in [\widehat{\tau}, \widehat{T}] \end{cases}$$

denotes the time–dependent reference Hamiltonian defined by $\widehat{f}_t$.

We also use the following notation from differential geometry: $L_f\alpha$ is the Lie derivative of a function α with respect to the vector field f. Moreover, if G is a C^1 map from a manifold X in a manifold Y, we denote its tangent map at a point $x \in X$ as T_xG. If the point x is clear from the context, we also write $T_xG = G_*$.

2.2. Pontryagin Maximum Principle

In this section we recall the first order optimality condition which the reference couple must satisfy.

We call *extremal* any curve in the cotangent bundle which satisfies PMP and *state–extremal* its projection on the state space, namely we give the following

DEFINITION 2 *Let $u \in L^\infty([0,T], \mathbb{R})$ be an admissible control, we call* extremal *of the control system* (1)-(2) *any non trivial trajectory $\lambda \colon [0,T] \to T^*M$ of the Hamiltonian vector field $\overrightarrow{F}_0 + u(t)\overrightarrow{F}_1$, which satisfies*

$$F_0 \circ \lambda(t) + u(t) F_1 \circ \lambda(t) = \max_{u \in [-1,1]} \{F_0 \circ \lambda(t) + u F_1 \circ \lambda(t)\} \quad a.e.\ t \in [0,T]$$

and we call its projection on the state space $\xi := \pi\lambda \colon [0,T] \to M$ a state extremal*.*

We require the reference trajectory to be a state extremal, namely $(\widehat{\xi}, \widehat{u})$ satisfies the following

Assumption 1 (PMP) *There exist* $p_0 \in \{0,1\}$ *and a solution* $\widehat{\lambda}\colon t \in [0,\widehat{T}] \mapsto \widehat{\lambda}(t) \in T^*M$ *of the Hamiltonian system*

$$\dot{\lambda}(t) = \overrightarrow{F}_t \circ \lambda(t)$$

such that $\pi\widehat{\lambda} = \widehat{\xi}$, $\widehat{\lambda}(0) \neq 0$ *and*

$$\widehat{F}_t(\widehat{\lambda}(t)) = \max\{\langle\widehat{\lambda}(t)\,,\, f_0(\widehat{\xi}(t)) + uf_1(\widehat{\xi}(t))\rangle\colon\ |u| \leq 1\} = p_0 \quad a.e.\ t \in [0,\widehat{T}].$$

We assume that PMP holds in the normal form, i.e. $p_0 = 1$.

$\widehat{\lambda}\colon [0,\widehat{T}] \to T^*M$ is called *adjoint covector* and we denote $\widehat{\ell} := \widehat{\lambda}(\widehat{\tau})$.

Given the structure of the reference control $\widehat{u}$, as defined by equations (4)-(5), PMP implies

$$u_1 F_1 \circ \widehat{\lambda}(t) \geq 0 \quad \forall t \in [0,\widehat{\tau}], \tag{8}$$

$$F_1 \circ \widehat{\lambda}(t) = \langle\widehat{\lambda}(t)\,,\, f_1(\widehat{\xi}(t)\rangle = 0 \quad \forall t \in [\widehat{\tau},\widehat{T}]. \tag{9}$$

By differentiation of equation (9) we also have

$$F_{01} \circ \widehat{\lambda}(t) = \langle\widehat{\lambda}(t)\,,\, f_{S1}(\widehat{\xi}(t))\rangle \equiv 0 \quad t \in [\widehat{\tau},\widehat{T}] \tag{10}$$

$$(F_{001} + \widehat{u}(t)F_{101}) \circ \widehat{\lambda}(t) = \langle\widehat{\lambda}(t)\,, f_{SS1}(\widehat{\xi}(t))\rangle = 0 \quad t \in (\widehat{\tau},\widehat{T}) \tag{11}$$

while equations (8) and (10) give

$$\lim_{t \to \widehat{\tau}^-} \frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(u_1 F_1 \circ \widehat{\lambda}\right)(t) = u_1\left(F_{001} + u_1 F_{101}\right)(\widehat{\ell}) \geq 0. \tag{12}$$

It is also known that a necessary condition for the local optimality of a Pontryagin extremal is the *generalized Legendre condition* (GLC) along the singular arc:

$$F_{101} \circ \widehat{\lambda}(t) = \langle\widehat{\lambda}(t)\,,\, [f_1,[f_S,f_1]](\widehat{\xi}(t))\rangle \geq 0 \quad t \in [\widehat{\tau},\widehat{T}]\,,$$

see, for example, Agrachev and Sachkov (2004), Corollary 20.18 p. 318; for a classical result see Gabasov and Kirillova (1972).

2.3. Regularity conditions

The regularity conditions we impose consist in requiring strict inequalities where necessary conditions yield mild inequalities.

Assumption 2 (Regularity along the bang arc)

$$u_1 F_1 \circ \widehat{\lambda}(t) > 0 \quad \forall t \in [0,\widehat{\tau}).$$

ASSUMPTION 3 (STRONG GENERALIZED LEGENDRE CONDITION (SGLC))

$$F_{101} \circ \widehat{\lambda}(t) > 0, \quad t \in [\widehat{\tau}, \widehat{T}]. \tag{SGLC}$$

ASSUMPTION 4 (REGULARITY AT THE JUNCTION POINT)

$$u_1 \left(F_{001} + u_1 F_{101}\right)(\widehat{\ell}) > 0.$$

REMARK 2 *We point out that, if SGLC holds, then Assumption 4 is equivalent to the discontinuity of $\widehat{u}$ at time $\widehat{\tau}$: in fact, from equation* (11) *and the continuity of $\widehat{\lambda}$, we have*

$$u_1 \left(F_{001} + \widehat{u}(\widehat{\tau}+)F_{101}\right)(\widehat{\ell}) = u_1 \lim_{t \to \widehat{\tau}+} \Big(F_{001} \circ \widehat{\lambda}(t) + \widehat{u}(t) F_{101} \circ \widehat{\lambda}(t)\Big) = 0.$$

Hence, thanks to Assumption 3, we have $u_1 F_{001}(\widehat{\ell}) + F_{101}(\widehat{\ell}) > 0$ if and only if $\lim_{t \to \widehat{\tau}+} \widehat{u}(t) \neq u_1$.

2.4. Main result

The remaining sufficient conditions are derived from the sub–problem obtained by keeping the reference final point and the reference bang control fixed and allowing the singular control to vary, namely we study the minimum time problem subject to (1) for $t \in [\widehat{\tau}, \widehat{T}]$ and

$$\xi(\widehat{\tau}) = \widehat{x}\,, \quad \xi(T) = \widehat{x}_f. \tag{13}$$

For this problem we require the second order conditions for singular trajectories stated in Stefani (2004), which will be described in Section 4.

ASSUMPTION 5 (2ND ORDER CONDITION ON THE SINGULAR ARC) *We assume the coercivity of the extended second variation J''_E associated to the singular trajectory $\widehat{\xi}_{|[\widehat{\tau}, \widehat{T}]}$ of the minimum time problem with fixed end-points $\widehat{x}$ and $\widehat{x}_f$, as defined in Section 4, see also Stefani (2004 and 2007).*

THEOREM 1 (MAIN THEOREM) *Suppose that $\widehat{\xi}$ is a normal bang–singular state-extremal and that the regularity conditions (Assumptions 2–4) are satisfied. If the extended second variation J''_E on the singular arc is coercive, then $\widehat{\xi}$ is a strict strong local minimizer for the minimum time problem between $\widehat{x}_0$ and the integral curve Γ_f of f_1 through $\widehat{x}_f$.*

2.5. Geometry near the singular arc

In this section we assume SGLC and we describe some properties of the Hamiltonians linked to our system near the singular arc of the adjoint covector.

For $t \in [\widehat{\tau}, \widehat{T}]$, $\widehat{u}(t) = \dfrac{-F_{001}}{F_{101}}(\widehat{\lambda}(t))$ and $\widehat{\lambda}_{|[\widehat{\tau},\widehat{T}]}$ is called a singular extremal of the first kind, see for example Zelikin and Borizov (1994).

By (9) and (10), any singular extremal of the first kind belongs to the $(2n-2)$–dimensional symplectic manifold

$$\mathcal{S} = \{\ell \colon F_1(\ell) = F_{01}(\ell) = 0, F_{101}(\ell) > 0\}$$

which is contained in the hyper-surface

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

where the maximized Hamiltonian

$$H^{\max} \colon \ell \mapsto \max\{F_0(\ell) + uF_1(\ell) \colon u \in [-1, 1]\}$$

coincides with F_0.

Note that $\mathcal{S}$ and Σ are independent of the control constraints but, by (2), any singular extremal of our problem belongs to

$$\mathcal{S} \cap \left\{\ell \colon \frac{F_{001}}{F_{101}}(\ell) \in (-1, 1)\right\}.$$

By SGLC it is easy to prove the following result:

Lemma 2.1 *There is a neighborhood $\mathcal{U}$ of $\mathcal{S}$ in T^*M where the following claims hold true.*

1. *Σ is an hyper-surface containing the symplectic manifold $\mathcal{S}$ and it separates in $\mathcal{U}$ the regions defined by: $H^{\max} = F_0 + F_1$, $\quad H^{\max} = F_0 - F_1$.*
2. *$\overrightarrow{F}_1$ is tangent to Σ and transversal to $\mathcal{S}$, while $\overrightarrow{F}_{01}$ is transversal to Σ.*
3. *The maps $(s, \ell) \mapsto \exp s\overrightarrow{F}_1(\ell)$ and $(\tau, s, \ell) \mapsto \exp \tau\overrightarrow{F}_{01} \circ \exp s\overrightarrow{F}_1(\ell)$ are local diffeomorphisms from $\mathbb{R} \times \mathcal{S}$ to Σ and from $\mathbb{R} \times \mathbb{R} \times \mathcal{S}$ to T^*M respectively.*

Property *(3)* in Lemma 2.1 yields the possibility of defining the smooth function $\upsilon \colon \mathcal{U} \to \mathbb{R}$ as

$$\upsilon := \frac{-F_{001}}{F_{101}} \quad \text{on} \quad \mathcal{S} \tag{14}$$

and then extending it constant first on the integral lines of $\overrightarrow{F}_1$ and then on those of $\overrightarrow{F}_{01}$. In this way we get *the Hamiltonian of singular extremal of the first kind* by defining

$$K := F_0 + \upsilon\, F_1.$$

In this paper we consider the case when the ratio F_{001}/F_{101} restricted to $\mathcal{S}$ is a function υ_S depending only on $\pi\ell \in M$. Indeed, in this case the Hamiltonian

$K_{|\mathcal{S}}$ is the lift of the vector field f_S and we say that *any singular control is feedback.*

Note that this case is generic in dimensions 2 and 3, as proved in the following lemma.

LEMMA 2.2 *Let $\ell \in \mathcal{S}$, $x = \pi\ell$ and $f_1(x) \neq 0$.*

- *If $\dim M = 2$, then $v(\ell)$ depends only on x.*
- *If $\dim M = 3$, then $v(\ell)$ depends only on x either if $f_1(x)$, $f_{01}(x)$ are linearly independent or if $f_1(x)$, $f_{101}(x)$, $f_{001}(x)$ are linearly dependent.*

Therefore, in dimension two, f_1 non-null on M implies that any singular control of the first kind is feedback, while, in dimension three, the existence of a feedback singular control of the first kind is implied either by f_1, f_{01} linearly independent of M or by f_1, f_{101}, f_{001} linearly dependent on M.

Proof. $f_1(x) \neq 0$ implies f_1, f_{101} linearly independent at x. If $\dim M = 2$, we choose $\{\omega_1, \omega_2\}$ as the dual base of $\{f_1(x), f_{101}(x)\}$ and it is easy to see that $v(\ell) = -\langle \omega_2\,,\, f_{001}(x)\rangle$.

Analogously, if $\dim M = 3$ and f_1, f_{01} are linearly independent at x, then we choose $\{\omega_1, \omega_2, \omega_3\}$ as the dual base of $\{f_1(x), f_{101}(x), f_{01}(x)\}$ and we get again $v(\ell) = -\langle \omega_2\,,\, f_{001}(x)\rangle$.

If f_1, f_{01} are linearly dependent at x, then we choose $\{\omega_1, \omega_2, \omega_3\}$ as the dual base of a base of type $\{f_1(x), f_{101}(x), g\}$ and we get $\ell \in \mathcal{S}$ if $\ell = a_2\omega_2 + a_3\omega_3$ with $a_2 > 0$. Therefore $v(\ell) = -\langle \omega_2\,,\, f_{001}(x)\rangle - \frac{a_3}{a_2}\langle \omega_3\,,\, f_{001}(x)\rangle$ and the claim follows. ∎

3. Hamiltonian approach to strong local optimality

The classical Hamiltonian approach to prove sufficient conditions for strong optimality is based on the construction of a field of non–intersecting state extremals covering a neighborhood of the given trajectory: i.e. by projecting on the state space the flow of $\overrightarrow{H}^{\max}$ starting from a suitable horizontal Lagrangian sub–manifold. Indeed, if $H = H^{\max}$ is C^2 and we can find a Horizontal Lagrangian sub–manifold Λ such that

$$\mathrm{id} \times \pi\mathcal{H}\colon [0,T] \times \Lambda \to [0,T] \times M\,, \quad (t,\ell) \mapsto (t, \pi\mathcal{H}_t(\ell)) \tag{15}$$

is a diffeomorphism, then we can use symplectic arguments to compare the costs by lifting admissible trajectories to the cotangent bundle, independently of the associated controls, and finally to prove sufficient conditions. In this case $\mathrm{id} \times \pi\mathcal{H}$ is a C^1 map and the coercivity of the second variation allows us to define a suitable manifold Λ for which $\mathrm{id} \times \pi\mathcal{H}$ is locally one–to–one, see Agrachev et al. (1998B).

The same ideas have been used in Agrachev et al. (2002) for bang–bang trajectories. In this case $\mathcal{H}$ is not C^1 but suitable regularity conditions at the switching points of the reference extremal imply that $\mathcal{H}$ is piecewise C^∞. Again the coercivity of a suitable second variation permits us to find Λ, for which $\mathrm{id} \times \pi\mathcal{H}$ is a piecewise diffeomorphism.

The presence of a singular arc prevents the existence of a sufficiently smooth flow of $\overrightarrow{H}^{\max}$; on the other hand one can observe that a comparison of the costs can be obtained also with a Hamiltonian H which is greater than or equal to $H^{\max}$. For this reason we are led to introduce a time–dependent Hamiltonian H_t and the notion of *almost–extremal*. We call almost–extremal a solution of the Hamiltonian system associated to a Hamiltonian H_t with the following properties

$$H_t \geq H^{\max}, \quad H_t \circ \widehat{\lambda} = \widehat{H}_t \circ \widehat{\lambda}, \quad \dot{\widehat{\lambda}} = \overrightarrow{H}_t \circ \widehat{\lambda}. \tag{16}$$

See Stefani (2004) for the application to a totally singular arc and Subsection 3.2 for further details on the bang–singular case.

The paradigm is as follows: the regularity assumptions allow us to choose an *over–maximized* Hamiltonian satisfying (16) and whose flow is C^1. The coercivity of the second variation gives the possibility of defining a manifold Λ such that map (15) is invertible.

This general paradigm can be followed to prove sufficient conditions also in other situations.

3.1. The Hamiltonian χ

We use the strategy adopted in Stefani (2007) to overcome the problems arising from the existence of a singular arc, that is we add, *near* the singular arc of the adjoint covector, a positive Hamiltonian χ to the reference one, see Section 3.2. This possibility is given by Regularity Assumption 3 and is described in the following lemma.

LEMMA 3.1 *By possibly restricting $\mathcal{U}$, it is possible to define a smooth function $\rho\colon \mathcal{U} \subset T^*M \to \mathbb{R}$, with the following properties*

i. The Hamiltonian $\chi = \frac{\rho}{2} F_{01}^2$ is such that the Hamiltonian vector field $\overrightarrow{F}_0 + \overrightarrow{\chi}$ is tangent to Σ.

ii. For any $\ell \in \mathcal{S}$, $\rho(\ell) = \dfrac{1}{F_{101}(\ell)}$, hence, without loss of generality, we can suppose $\rho > 0$

iii. ρ can be chosen so that

$$\chi(\ell) = F_0 \circ \exp \vartheta(\ell) \overrightarrow{F}_1(\ell) - F_0(\ell)$$

where $\vartheta(\ell)$ is defined by

$$F_{01}(\exp \vartheta(\ell) \overrightarrow{F}_1(\ell)) = 0 \quad \text{and} \quad \vartheta(\ell) = 0\,, \ \forall \ell \in \mathcal{S},$$

hence

$$(\overrightarrow{F}_0 + \overrightarrow{\chi})(\ell) = (\exp(-\vartheta(\ell))\overrightarrow{F}_1)_* \overrightarrow{F}_0 \circ (\exp \vartheta(\ell) \overrightarrow{F}_1)(\ell)\,, \quad \forall \ell \in \Sigma.$$

Proof. The proof is given in Stefani (2007) with a reversed inequality, indeed in that paper the first order conditions are considered in the form of Lagrange multipliers rule, which leads to an adjoint covector which is opposite to the one we consider. ∎

In the following lemma we collect the main properties of the Hamiltonian χ in connection with the Hamiltonians H_1 and

$$H_S := K + \chi = F_0 + \upsilon F_1 + \chi, \tag{17}$$

which allow us to pursue our paradigm.

LEMMA 3.2 *The following statements hold true:*

1. $H_S \geq H^{\max}$ *on* Σ *and* $\overrightarrow{H}_S$ *is tangent to* Σ
2. $H_1 + \chi \geq H^{\max}$ *on* Σ *and* $\overrightarrow{H}_1 + \overrightarrow{\chi}$ *is tangent to* Σ
3. $\vec{\chi}$ *is null and* $\overrightarrow{K} = \overrightarrow{F}_S$ *on* $\mathcal{S}$ *hence,* $\widehat{\lambda}_{|[\widehat{\tau},\widehat{T}]}$ *is a trajectory of* $\overrightarrow{H}_S$.
4. $\overrightarrow{F}_1$ *is invariant both with respect to the flow of* $\overrightarrow{H}_S$ *and with respect to the flow of* $\overrightarrow{H}_1 + \overrightarrow{\chi}$*, namely let* H *be either equal to* H_S *or equal to* $H_1 + \chi$*, then*
$$\overrightarrow{F}_1 \circ \mathcal{H}_t(\ell) = \mathcal{H}_{t*}\overrightarrow{F}_1(\ell)\,, \quad \ell \in \Sigma$$
5. $\langle \ell\,,\, \pi_* \overrightarrow{\chi}(\ell)\rangle = \chi(\ell)$ *for all* $\ell \in \Sigma$.

Proof. Claims 1, 2, 3 follow easily from the definitions and Lemma 3.1.

To prove claim 4, take $\ell \in \Sigma$, notice that $L_{\overrightarrow{F}_1}\upsilon \equiv 0$ by definition, and compute

$$\partial_t \left(\mathcal{H}_{t*}^{-1}\overrightarrow{F}_1 \circ \mathcal{H}_t\right)(\ell) = \mathcal{H}_{t*}^{-1}\left[\overrightarrow{F}_0 + \overrightarrow{\chi}, \overrightarrow{F}_1\right] \circ \mathcal{H}_t(\ell).$$

From claim iii. of Lemma 3.1 we obtain on Σ

$$\begin{aligned}
[\overrightarrow{F}_0 + \overrightarrow{\chi}, \overrightarrow{F}_1] &= \left[\exp(-\vartheta\overrightarrow{F}_1)_*\overrightarrow{F}_0 \circ \exp(\vartheta\overrightarrow{F}_1)\,,\, \overrightarrow{F}_1\right] = \\
&= \exp(-\vartheta\overrightarrow{F}_1)_*\{\overrightarrow{F}_{01} - L_{\overrightarrow{F}_1}\vartheta\overrightarrow{F}_{10}\} \circ \exp(\vartheta\overrightarrow{F}_1) = 0
\end{aligned}$$

since $L_{\overrightarrow{F}_1}\vartheta \equiv -1$ on Σ.

The proof of claim 5 follows easily from iii. of Lemma 3.1. ∎

3.2. Almost extremals

In this section we prove that regularity Assumptions 2, 3 and 4 imply the existence of a Hamiltonian H_t with the desired properties (16) on its flow starting from Σ.

The flow of H_1 backward in time emanating from $\ell \in \Sigma$ at time $\widehat{\tau}$ behaves differently according to the sign of $u_1F_{01}(\ell)$. Namely, if $u_1F_{01}(\ell) \leq 0$ then

$u_1 F_1(\exp(t-\widehat{\tau})\overrightarrow{H}_1)(\ell) > 0$ for $t < \widehat{\tau}$, hence H_1 is the maximized Hamiltonian along its own flow. If $u_1 F_{01}(\ell) > 0$ then $u_1 F_1(\exp(t-\widehat{\tau})\overrightarrow{H}_1)(\ell) < 0$ for $t < \widehat{\tau}$, hence H_1 is no longer the maximized Hamiltonian. We overcome this problem by substituting H_1 with $H_1 + \chi$ for these *bad* points so that the flow is kept on Σ until it reaches a *good* point, as precised in the following lemma.

LEMMA 3.3 *Let Assumptions 3 and 4 be satisfied. Then*

1. *There exists a neighborhood $\mathcal{O}$ of $\widehat{\ell}$ in Σ such that the implicit equation*

$$\begin{cases} \tau(\ell) = \widehat{\tau}\,, \\ F_{01} \circ \exp(\tau(\ell) - \widehat{\tau})(\overrightarrow{H}_1 + \overrightarrow{\chi})(\ell) = 0 \end{cases}$$

defines a smooth function

$$\tau \colon \ell \in \mathcal{O} \mapsto \tau(\ell) \in \mathbb{R}.$$

2. $\langle \mathrm{d}\tau(\ell)\,,\,\delta\ell\rangle = \dfrac{-\boldsymbol{\sigma}\left(\delta\ell, \overrightarrow{F}_{01}(\ell)\right)}{(F_{001} + u_1 F_{101})\,(\ell)} \qquad \forall \ell \in \mathcal{O} \cap \mathcal{S}.$

3. $F_{01}(\ell) = 0 \Longleftrightarrow \tau(\ell) = \widehat{\tau}$ *and* $\operatorname{sgn}(u_1 F_{01}(\ell)) = \operatorname{sgn}(\widehat{\tau} - \tau(\ell))$.

4. *There exists $\varepsilon > 0$ such that*

$$|\tau(\ell) - \widehat{\tau}| < \varepsilon \qquad \forall \ell \in \mathcal{O} \tag{18}$$

and such that, for any $\ell \in \mathcal{O}$ and for any t, s with $\widehat{\tau} - \varepsilon \le t \le s \le \widehat{\tau}$, the following inequality holds:

$$u_1\,(F_{001} + u_1 F_{101}) \circ \exp(t-s)\overrightarrow{H}_1 \circ \exp(s-\widehat{\tau})\left(\overrightarrow{H}_1 + \overrightarrow{\chi}\right)(\ell) > 0. \tag{19}$$

Proof. *1.* and *2.* Since the partial derivative

$$\partial_t F_{01} \circ \exp(t-\widehat{\tau})(\overrightarrow{H}_1 + \overrightarrow{\chi})(\ell)\Big|_{(t,\ell)=(\widehat{\tau},\widehat{\ell})} = (F_{001} + u_1 F_{101})\,(\widehat{\ell}) \tag{20}$$

is not zero by Assumption 4, then the implicit function theorem applies.
3. From equation (20) and Assumption 4 the function

$$t \mapsto u_1 \partial_t F_{01} \circ \exp(t-\widehat{\tau})(\overrightarrow{H}_1 + \overrightarrow{\chi})(\ell)$$

is locally strictly increasing, hence claim *3.* follows.
4. Inequalities (18)–(19) are satisfied in a neighborhood of $t = s = \widehat{\tau}$, $\ell = \widehat{\ell}$, by continuity. ∎

Now we are able to define a Hamiltonian H_t, visualized in Fig. 1, which allows us to pursue the paradigm:

$$H_t(\ell) = \begin{cases} H_1(\ell) & \text{if } \ 0 \le t \le \widehat{\tau} - \varepsilon \quad \text{or} \\ & \text{if } \ \widehat{\tau} - \varepsilon < t < \widehat{\tau},\ u_1 F_{01}(\ell) \le 0 \\ (H_1 + \chi)(\ell) & \text{if } \ \widehat{\tau} - \varepsilon < t < \widehat{\tau},\ u_1 F_{01}(\ell) > 0 \\ H_S(\ell) & \text{if } \ \widehat{\tau} \le t \le \widehat{T} \end{cases} \tag{21}$$

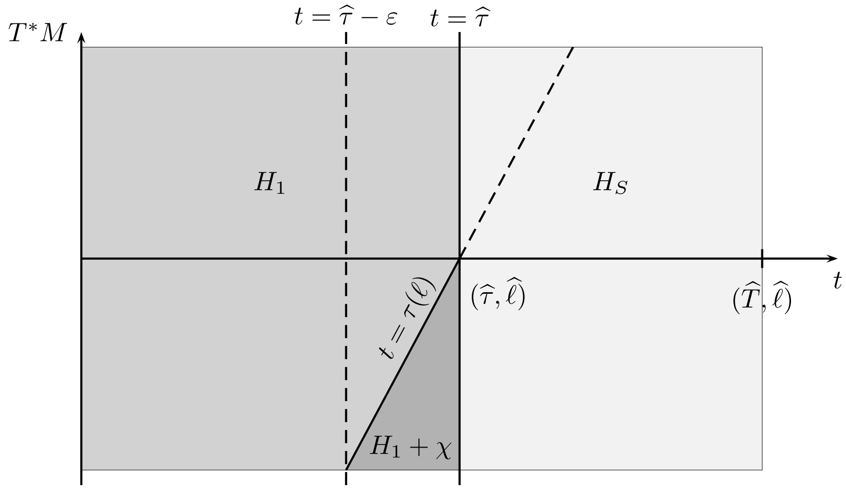

Figure 1. Construction of the over–maximized Hamiltonian

Note that the above defined Hamiltonian H_t has the required properties only on its flow starting from Σ, as shown in the following lemma.

LEMMA 3.4 *The Hamiltonian*

$$H : (t, \ell) \in [0, \widehat{T}] \times T^*M \mapsto H_t(\ell) \in \mathbb{R}$$

satisfies the following properties:

1. *H is C^1 with respect to ℓ, for any t.*
2. *$\overrightarrow{H}_t(\ell)$ is Lipschitz continuous with respect to ℓ, for any $t \in [0, \widehat{T}]$. The only discontinuities, with respect to time, of $\overrightarrow{H}_t(\ell)$ occur at $t = \widehat{\tau}$, i.e. they occur at the discontinuity of the reference control function $\widehat{u}$.*
3. *The flow $\mathcal{H}$ of $\overrightarrow{H}_t$ is $C^1([0, \widehat{T}] \times \mathcal{O})$.*
4. *The restriction of the Hamiltonian H to $(\mathrm{id} \times \mathcal{H})([0, \widehat{T}] \times \mathcal{O})$ is continuous.*
5. *$\mathcal{H}_t(\widehat{\ell}) = \widehat{\lambda}(t)$, for any $t \in [0, \widehat{T}]$.*
6. *$F_1 \circ \mathcal{H}_t(\ell) = 0$ for any $(t, \ell) \in [\min\{\tau(\ell), \widehat{\tau}\}, \widehat{T}] \times \mathcal{O}$.*
7. *Possibly restricting $\mathcal{O}$,*

$$u_1 F_1 \circ \mathcal{H}_t(\ell) > 0 \quad \forall (t, \ell) \in [0, \min\{\tau(\ell), \widehat{\tau}\}) \times \mathcal{O}. \tag{22}$$

8. *$H_t \circ \widehat{\lambda}(t) = \widehat{H}_t \circ \widehat{\lambda}(t)$ for any $t \in [0, \widehat{T}]$ and*

$$H_t \circ \mathcal{H}_t(\ell) \geq H^{\max} \circ \mathcal{H}_t(\ell) \quad \textit{for any} \quad (t, \ell) \in [0, \widehat{T}] \times \mathcal{O}.$$

Proof. *1.* and *2.* are easy corollaries of Lemma 3.3, taking into account that χ is C^∞, by Lemma 3.1.

3. For $t \in (\widehat{\tau}, \widehat{T}]$, the property is obvious, since $H_t = H_S$ is C^∞.
For $t \in [0, \widehat{\tau}]$, the only discontinuities of the map $\ell \mapsto T_\ell \mathcal{H}_t$ may occur at the (t, ℓ) such that $F_{01}(\ell) = 0$ or such that $u_1 F_{01}(\ell) \geq 0$ and $t = \tau(\ell)$. By definition we obtain

$$T_\ell \mathcal{H}_t = \begin{cases} T_\ell \exp(t - \widehat{\tau})\overrightarrow{H}_1 & \text{if } u_1 F_{01}(\ell) < 0 \\ T_\ell \exp(t - \widehat{\tau})(\overrightarrow{H}_1 + \overrightarrow{\chi}) & \text{if } u_1 F_{01}(\ell) > 0,\ t \in (\tau(\ell), \widehat{\tau}). \end{cases}$$

Moreover, if $u_1 F_{01}(\ell) > 0$ and $t \in [0, \tau(\ell))$ then, taking into account that $\overrightarrow{\chi} \circ \mathcal{H}_{\tau(\ell)}(\ell) = 0$, we get

$$T_\ell \mathcal{H}_t = T_{\mathcal{H}_{\tau(\ell)}(\ell)} \exp(t - \tau(\ell))\overrightarrow{H}_1 \circ T_\ell \exp(\tau(\ell) - \widehat{\tau})(\overrightarrow{H}_1 + \overrightarrow{\chi}).$$

An easy computation completes the proof.
4. and *5.* follow easily from *1.* and *2.*
6. is obvious, by Lemma 3.1.
7. Consider the cases

1. $\widehat{\tau} \leq \tau(\ell)$ and $t < \widehat{\tau}$. $\mathcal{H}_t(\ell) = \exp(t - \widehat{\tau})\overrightarrow{H}_1(\ell)$ and $u_1 F_{01}(\ell) \leq 0$ so that

$$u_1 F_1 \circ \mathcal{H}_t(\ell) = (t - \widehat{\tau}) u_1 F_{01}(\ell) + \frac{(t - \widehat{\tau})^2}{2} u_1 \left(F_{001} + u_1 F_{101}\right)(\ell) + o((t - \widehat{\tau})^2).$$

2. $\tau(\ell) < \widehat{\tau}$, and $t < \tau(\ell)$. $\mathcal{H}_t(\ell) = \exp(t - \tau(\ell))\overrightarrow{H}_1(\mathcal{H}_{\tau(\ell)}(\ell))$, $u_1 F_{01}(\ell) > 0$ so that

$$u_1 F_1 \circ \mathcal{H}_t(\ell) = \frac{(t - \tau(\ell))^2}{2} u_1 \left(F_{001} + u_1 F_{101}\right) \circ \mathcal{H}_{\tau(\ell)}(\ell) + o((t - \tau(\ell))^2).$$

Using (19), *7.* follows.
8. The proof is straightforward from *7.*, since χ is a non–negative function. ■

4. The extended second variation

In Stefani (2004) the extended second variation J'', associated to a totally singular extremal of a minimum time problem is defined starting from the coordinate–free second variation given in Agrachev et al. (1998a). Locally around $\widehat{x}$ define the time–dependent vector field

$$g_t := \widehat{S}_{t*}^{-1} f_1 \circ \widehat{S}_t \colon \mathcal{W} \to TM, \quad t \in [\widehat{\tau}, \widehat{T}]$$

and choose a function β on $\mathcal{W}$ such that

$$\mathrm{d}\beta(\widehat{x}) = -\widehat{\lambda}(\widehat{\tau}). \tag{23}$$

Then the second variation as defined in Agrachev et al. (1998a) is a quadratic form on $L^2([\widehat{\tau}, \widehat{T}], \mathbb{R})$ realized as the following intrinsic LQ problem on the vector space $T_{\widehat{x}}M$:

$$J''[\delta u]^2 = \int_{\widehat{\tau}}^{\widehat{T}} \delta u(t)\, L_{\delta\eta(t)}\, L_{g_t}\, \beta(\widehat{x}_0)\,\mathrm{d}t$$
$$\dot{\delta\eta}(t) = \delta u(t) g_t(\widehat{x}), \quad \delta\eta(\widehat{\tau}) = \delta\eta(\widehat{T}) = 0.$$

This result is independent of the choice of the function β with property 23, see Agrachev et al. (1998A), hence we can choose β so that $L_{f_1}\beta \equiv 0$. Applying the techniques in Stefani (2004), developed in Stefani (2007), we obtain a quadratic form on $\mathbb{R}\times L^2([\widehat{\tau},\widehat{T}],\mathbb{R})$ realized as an intrinsic LQ problem on the vector space $T_{\widehat{x}}M$. Setting for $t \in [\widehat{\tau},\widehat{T}]$

$$R(t) := L_{[\dot{g}_t\,,\,g_t]}\beta(\widehat{x}) = \langle \mathrm{d}\beta(\widehat{x})\,,\,[\dot{g}_t, g_t]\rangle = \langle \widehat{\lambda}(t)\,,\,f_{101}(\widehat{\xi}(t))\rangle$$
$$Q(t) := L_{(\cdot)}\, L_{\dot{g}_t}\, \beta(\widehat{x})$$
$$\delta e := (w_0, w) \in \mathbb{R}\times L^2([\widehat{\tau},\widehat{T}],\mathbb{R})$$

we can write the *extended second variation* as

$$J''_E[\delta e]^2 = \frac{1}{2}\int_{\widehat{\tau}}^{\widehat{T}} w(t)^2 R(t) + 2w(t)Q(t)\zeta(t)\ \mathrm{d}t,$$
$$\dot{\zeta}(t) = w(t)\dot{g}_t(\widehat{x}), \quad \zeta(\widehat{\tau}) = w_0 f_1(\widehat{x}), \quad \zeta(\widehat{T}) = 0.$$

See Stefani (2007) for further details and note that the same second variation, written in symplectic form, is considered in Agrachev and Sachkov (2004). In Stefani (2007) the necessary and sufficient conditions are proved for the coercivity of J''_E from different points of view, in particular a reduction to a non singular problem is proved: denote

$$\Gamma := \{\exp s f_1(\widehat{x}), s \in \mathbb{R}\}$$

the integral curve of f_1 through $\widehat{x}$, and let α be any function such that $\mathrm{d}\alpha(\widehat{x}) = \widehat{\ell}$ and $L^2_{f_1}\alpha(\widehat{x}) = 0$. Then, J''_E is proved to be the standard second variation, relative to the extremal $\widehat{\lambda}_{|[\widehat{\tau},\widehat{T}]}$ and to the reference control $\widehat{w} \equiv 0$, of the Mayer problem

$$\text{minimize} \quad \alpha(\xi(\widehat{\tau})) \quad \text{subject to} \tag{24}$$
$$\dot{\xi}(t) = f_S(\xi(t)) + w(t) f_{S1}(\xi(t)) + \frac{1}{2}w(t)^2 f_{1S1}(\xi(t)) \tag{25}$$
$$\xi(\widehat{\tau}) \in \Gamma\,, \quad \xi(T) = \widehat{x}_f. \tag{26}$$

Note that the coercivity of J''_E implies that SGLC must be satisfied and that $f_1(\widehat{x}) \neq 0$. Moreover, easy calculations show that SGLC is the strong Legendre condition for the above non-singular problem.

Since in our case the Hamiltonian of singular extremals F_S is the lift of a vector field f_S, we can use Lemma 6.4 in Stefani (2007) to give necessary and sufficient conditions for the coercivity of J''_E in terms of the vector fields f_S and f_1.

LEMMA 4.1 *Necessary and sufficient condition for the coercivity of* J''_E.
If f_1, f_{01} *are linearly dependent at* $\widehat{x}$, *then* J''_E *is coercive if and only if*

$$f_1(\widehat{\xi}(t)) \neq 0 \text{ for all } t \in [\widehat{\tau}, \widehat{T}].$$

If f_1, f_{01} *are linearly independent at* $\widehat{x}$, *then* J''_E *is coercive if and only if*

$$f_1(\widehat{\xi}(t)) \text{ and } \widehat{S}_{t*}f_1(\widehat{x}) \text{ are linearly independent for all } t \in (\widehat{\tau}, \widehat{T}].$$

4.1. Reduction to a free initial point problem

In this section we show how the coercivity of J''_E allows us to add a penalty on the initial point of problem (24)–(25)–(26) so that the second variation remains coercive also removing the constraint on the initial point. This allows us to define the Lagrangian sub–manifold with the property required in Section 3.

LEMMA 4.2 *If* J''_E *is coercive, then there is a function* α *defined in a neighborhood* $\mathcal{W}$ *of* $\widehat{x}$, *with the following properties*

i) $L_{f_1}\alpha \equiv 0$

ii) $\mathrm{d}\alpha(\widehat{x}) = \widehat{\ell}$

iii) Let $\Lambda = \{\mathrm{d}\alpha(x) \colon x \in \mathcal{W}\}$ *be the Lagrangian sub–manifold defined by* α *and let* $\mathcal{H}$ *be the Hamiltonian flow associated to the Hamiltonian* H_t *defined in* (21), *then* $\pi_*\mathcal{H}_{t*} \colon T_{\widehat{\ell}}\Lambda \to T_{\widehat{\xi}(t)}M$ *is an isomorphism for* $t \in [\widehat{\tau}, \widehat{T}]$.

Proof. Since $f_1(\widehat{x}) \neq 0$, we can choose coordinates $(x_1, \cdots, x_n)$ at $\widehat{x}$ such that $f_1 = \frac{\partial}{\partial x_1}$, therefore $\widehat{\ell} = \sum_{i=2}^n \lambda_i \,\mathrm{d}x_i$ and we can choose β in (23) as

$$\beta = -\sum_{i=2}^n \lambda_i x_i.$$

In these coordinates choose the non-negative quadratic form on $T_{\widehat{x}}M$

$$\Omega = \frac{1}{2}\sum_{i=2}^n \mathrm{d}x_i \otimes \mathrm{d}x_i$$

and extend it to $T_{\widehat{x}}M \times L^2([\widehat{\tau}, \widehat{T}], \mathbb{R})$ by $\Omega[\delta e]^2 = \Omega[\delta x]^2$. The quadratic form defined on $T_{\widehat{x}}M \times L^2([\widehat{\tau}, \widehat{T}], \mathbb{R})$ by

$$(\delta x, w) \mapsto \Omega[\delta x]^2 + \frac{1}{2}\int_{\widehat{\tau}}^{\widehat{T}} w(t)^2 R(t) + 2w(t)Q(t)\zeta(t)\,\mathrm{d}t$$
$$\dot{\zeta}(t) = w(t)\dot{g}_t(\widehat{x}), \qquad \zeta(\widehat{\tau}) = \delta x, \qquad \zeta(\widehat{T}) = 0,$$

coincides with J''_E on the kernel of Ω, hence we can apply Theorem 13.2 of Hestenes (1951) and conclude that there is a positive constant s such that

$$J''_E + s\Omega \text{ is coercive on } W = \{\delta e \in T_{\widehat{x}}M \times L^2([\widehat{\tau}, \widehat{T}], \mathbb{R}) : \zeta(\widehat{T}) = 0\}.$$

By defining

$$\alpha(x) = \sum_{i=2}^{n} \lambda_i x_i + \frac{1}{2} s \sum_{i=2}^{n} x_i^2,$$

we get that α satisfies i), ii) and that $\gamma'' := D^2(\alpha + \beta)(\widehat{x}) = s\Omega$.

$J''_E + s\Omega$ turns out to be the standard second variation of the Mayer problem

$$\text{minimize } \alpha(\xi(\widehat{\tau})) \text{ subject to (25) and } \xi(T) = \widehat{x}_f.$$

Therefore the proof of iii) follows from Corollary 6.1 in Stefani (2007). ■

5. Proof of the Main Theorem

In this section we complete the proof of the main result. We start by proving the existence of a Lagrangian sub–manifold having the required properties with respect to the Hamiltonian defined in (21).

Lemma 5.1 *Let $\alpha\colon \mathcal{W} \to \mathbb{R}$ and Λ be the function and the associated Lagrangian sub–manifold defined in Lemma 4.2. Possibly restricting $\mathcal{W}$, the map*

$$\mathrm{id} \times \pi\mathcal{H}\colon (t, \ell) \in [0, \widehat{T}] \times \Lambda \mapsto (t, \pi\mathcal{H}_t(\ell)) \in [0, \widehat{T}] \times M$$

is one–to–one onto a neighborhood $\mathcal{V}$ of the graph of $\widehat{\xi}$ in $[0, \widehat{T}] \times M$.

Proof. Since $[0, \widehat{T}]$ is compact, it suffices to show that $\pi\mathcal{H}_t\colon \ell \mapsto \pi\mathcal{H}_t(\ell)$ is locally one–to–one around $(t, \widehat{\ell})$ for any $t \in [0, \widehat{T}]$. Since $\mathcal{H}_t$ is C^1, it suffices to show that $\pi_*\mathcal{H}_{t*}\colon T_{\widehat{\ell}}\Lambda \to T_{\widehat{\xi}(t)}M$ is an isomorphism for any $t \in [0, \widehat{T}]$. In fact, if $t \geq \widehat{\tau}$, the claim is proved in Lemma 4.2, while if $t \in [0, \widehat{\tau}]$ an easy calculation shows that $\pi_*\mathcal{H}_{t*} = \exp(t - \widehat{\tau})h_{1\,*}$. ■

We can now use symplectic arguments to complete the proof of the main theorem. First, observe that the one–form $\boldsymbol{\omega}$ defined by

$$\boldsymbol{\omega} = \mathcal{H}^*(\boldsymbol{s} - H_t \, \mathrm{d}t)$$

is exact on $[0, \widehat{T}] \times \Lambda$, see Arnold (1980).

Let (T, ξ, u) be an admissible triplet such that $\widehat{\tau} < T \leq \widehat{T}$ and such that the graph of $\xi\colon [0, T] \to M$ is in the neighborhood $\mathcal{V}$ defined in Lemma 5.1.

By assumption $\xi(T)$ belongs to the integral line of f_1 emanating from $\widehat{x}_f$ i.e. $\exists\, \overline{s} \in \mathbb{R}$ such that $\xi(T) = \exp \overline{s} f_1(\widehat{x}_f)$. Define

$$\phi\colon r \in [0, 1] \mapsto \exp \overline{s}(1 - r) f_1(\widehat{x}_f) \in M.$$

Consider the paths in $\mathcal{V}$:

1. $\Xi := (\mathrm{id}, \xi) \colon t \in [0, T] \mapsto (t, \xi(t)) \in \mathbb{R} \times M$,
2. $\Phi \colon r \in [0, 1] \mapsto \Big((\widehat{T} - T) r + T, \phi(r) \Big) \in \mathbb{R} \times M$,
3. $\widehat{\Xi} := (\mathrm{id}, \widehat{\xi}) \colon t \in [0, \widehat{T}] \mapsto (t, \widehat{\xi}(t)) \in \mathbb{R} \times M$

and denote $\psi := (\mathrm{id} \times \pi\mathcal{H})^{-1}$. Since the concatenation of $\psi \circ \Xi$, $\psi \circ \Phi$, and of $\psi \circ \widehat{\Xi}$ ran backward in time, is a closed path in $[0, \widehat{T}] \times \Lambda$, then

$$0 = \oint \omega = \int_{\psi \circ \Xi} \omega + \int_{\psi \circ \Phi} \omega - \int_{\psi \circ \widehat{\Xi}} \omega. \tag{27}$$

From the properties of H_t (point *8.* of Lemma 3.4) we get

$$\int_{\psi \circ \widehat{\Xi}} \omega = 0, \qquad \int_{\psi \circ \Xi} \omega \leq 0 \tag{28}$$

so that

$$\int_{\psi \circ \Phi} \omega \geq 0. \tag{29}$$

Since $\mathcal{H} \circ \psi \circ \Phi$ takes values in Σ, then

$$\langle \mathcal{H} \circ \psi \circ \Phi(r) \, , \, f_1(\pi\mathcal{H} \circ \psi \circ \Phi(r)) \rangle = F_1 \circ \mathcal{H} \circ \psi \circ \Phi(r) \equiv 0,$$

so that

$$0 \leq \int_{\psi \circ \Phi} \omega = -(\widehat{T} - T) \int_0^1 H_S(\mathcal{H} \circ \psi \circ \Phi(r)) \, \mathrm{d}r. \tag{30}$$

Assume that $T \leq \widehat{T}$, then (30) is equivalent to

$$\begin{aligned}
0 &\leq - \int_T^{\widehat{T}} H_S \circ \mathcal{H} \circ \psi(t, \exp \frac{\overline{s}(\widehat{T} - t)}{\widehat{T} - T} f_1(\widehat{x}_f)) \, \mathrm{d}t \\
&= \int_T^{\widehat{T}} \Big(H_S \circ \mathcal{H}_{\widehat{T}} \circ \psi(\widehat{T}, \widehat{x}_f) + O(t - \widehat{T}) \Big) \, \mathrm{d}t = \int_T^{\widehat{T}} \Big(H_S \circ \mathcal{H}_{\widehat{T}}(\widehat{\ell}) + O(t - \widehat{T}) \Big) \, \mathrm{d}t \\
&= \int_T^{\widehat{T}} - \Big(1 + O(t - \widehat{T}) \Big) \, \mathrm{d}t = T - \widehat{T} + O\left((T - \widehat{T})^2 \right)
\end{aligned}$$

which implies $T = \widehat{T}$, so that $(\widehat{\xi}, \widehat{u})$ is a (time, state)–local minimizer.

To prove that such minimum is strict, let us assume, by contradiction, that $T = \widehat{T}$. Then $\Phi(r) = \Big(\widehat{T}, \exp \overline{s}(1 - r) f_1(\widehat{x}_f) \Big)$, and $\displaystyle\int_{\psi \circ \Phi} \omega = 0$. From equation (27) we also get

$$0 = \int_{\psi \circ \Xi} \omega = \int_0^{\widehat{T}} \langle \mathcal{H} \circ \psi(t, \xi(t)) \, , \, \dot{\xi}(t) \rangle - H_t \circ \mathcal{H} \circ \psi(t, \xi(t)) \, \mathrm{d}t,$$

and, by property *8.* of Lemma 3.4,

$$\langle \mathcal{H} \circ \psi(t, \xi(t)) \, , \, (f_0 + u(t) f_1)(\xi(t)) \rangle = H_t \circ \mathcal{H} \circ \psi(t, \xi(t)) \\ = H^{\max} \circ \mathcal{H} \circ \psi(t, \xi(t)) \quad (31)$$

for any $t \in [0, \widehat{T}]$. Define $\lambda(t) := \mathcal{H} \circ \psi(t, \xi(t))$. If $\lambda(t) \in \Sigma$, then (31) implies $\lambda(t) \in \mathcal{S}$, while if $\lambda(t) \notin \Sigma$, then claim *7.* of Lemma 3.4 and equation (31) yield $u(t) = u_1$. In particular, we get $u(t) = u_1$ for any $t \in [0, \widehat{\tau} - \varepsilon]$, therefore $\lambda(t) = \widehat{\lambda}(t)$ for any $t \in [0, \widehat{\tau} - \varepsilon]$, since $\lambda(0) = \widehat{\lambda}(0)$.
For $t \in (\widehat{\tau} - \varepsilon, \widehat{T})$ we can compute

$$\dot{\lambda}(t) = \overrightarrow{H}_t \circ \lambda(t) - \mathcal{H}_{t*}(\pi \mathcal{H}_t)_*^{-1} \left\{ \pi_* \overrightarrow{H}_t \circ \lambda(t) - (f_0 + u(t) f_1) \circ \xi(t) \right\} =$$
$$= \begin{cases} \overrightarrow{H}_1 \circ \lambda(t) + (u(t) - u_1) \mathcal{H}_{t*}(\pi \mathcal{H}_t)_*^{-1} f_1 \circ \xi(t) & t \in (\widehat{\tau} - \varepsilon, \widehat{\tau}) \\ \overrightarrow{H}_S \circ \lambda(t) + (u(t) - v_S \circ \xi(t)) \mathcal{H}_{t*}(\pi \mathcal{H}_t)_*^{-1} f_1 \circ \xi(t) & t \in (\widehat{\tau}, \widehat{T}). \end{cases}$$

Since, whenever $\lambda(t) \in \Sigma$, $\mathcal{H}_{t*}(\pi \mathcal{H}_t)_*^{-1} f_1 \circ \pi \lambda(t) = \overrightarrow{F}_1 \circ \lambda(t)$ (by point *4* of Lemma 3.2), in any case we get

$$\dot{\lambda}(t) = (\overrightarrow{F}_0 + u(t) \overrightarrow{F}_1) \circ \lambda(t) \quad \forall t \in (\widehat{\tau} - \varepsilon, \widehat{T}),$$

which, together with (31), means that λ satisfies PMP on $[\widehat{\tau} - \varepsilon, \widehat{T}]$. Since $\lambda(\widehat{\tau} - \varepsilon) = \widehat{\lambda}(\widehat{\tau} - \varepsilon)$, we easily get $\lambda = \widehat{\lambda}$ on $[0, \widehat{\tau}]$. Moreover, on $[\widehat{\tau}, \widehat{T}]$, λ is a Pontryagin extremal and its range is in $\mathcal{S}$, so that λ is a solution to the Cauchy problem

$$\dot{\lambda}(t) = \left(\overrightarrow{F}_0 - \frac{F_{001}}{F_{101}} \overrightarrow{F}_1 \right) \circ \lambda(t) \qquad \lambda(\widehat{\tau}) = \widehat{\ell}$$

on the interval $[\widehat{\tau}, \widehat{T}]$. Hence $\lambda(t) = \widehat{\lambda}(t)$ for any $t \in [\widehat{\tau}, \widehat{T}]$. Projecting on the state manifold M, we finally get $\xi(t) = \widehat{\xi}(t)$.

6. An example

Consider the Dodgem car problem:

minimize $\quad T$

subject to

$$\begin{array}{lll} \dot{x}_1(t) = \cos(x_3) & x_1(0) = 0 & x_1(T) = b_1 \\ \dot{x}_2(t) = \sin(x_3) & x_2(0) = 0 & x_2(T) = b_2 \\ \dot{x}_3(t) = u & x_3(0) = 0 & x_3(T) \in \mathbb{R} \\ u \in [-1, 1] & & \end{array}$$

where $x \equiv (x_1, x_2, x_3) \in \mathbb{R}^3$.
We have

$$f_0(x) = \begin{pmatrix} \cos(x_3) \\ \sin(x_3) \\ 0 \end{pmatrix}, \qquad f_1(x) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

hence, with a simple computation, we also get

$$f_{01}(x) = \begin{pmatrix} \sin(x_3) \\ -\cos(x_3) \\ 0 \end{pmatrix}, \qquad f_{101}(x) = f_0(x), \qquad f_{001}(x) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

and the associated Hamiltonians are

$$\begin{aligned} &F_0(p,x) = F_{101}(p,x) = p_1 \cos(x_3) + p_2 \sin(x_3) \qquad && F_1(p,x) = p_3 \\ &F_{01}(p,x) = p_1 \sin(x_3) - p_2 \cos(x_3) && F_{001}(p,x) = 0. \end{aligned}$$

Since, along any singular arc of a normal Pontryagin extremal, the condition $F_0(\widehat{\lambda}(t)) = 1$ must hold, then SGLC is satisfied and the only admissible singular control is $u(t) \equiv 0$, so that any singular arc is driven by $f_S = f_0$, and our theory applies.

Let us check that along any singular arc the hypotheses of Lemma 4.1 are satisfied. In fact, since $f_0(x)$ and $f_{01}(x)$ are linearly independent at any point of $\mathbb{R}^3$, we just need to show that $f_1(x)$ and $\exp(tf_0)_* f_1 \circ \exp(-tf_0)(x)$ are linearly independent for any $t \neq 0$.
A simple computation shows that

$$\begin{aligned} \exp(tf_0)(x) &= \begin{pmatrix} x_1 + t\cos(x_3) \\ x_2 + t\sin(x_3) \\ x_3 \end{pmatrix} \\ \exp(tf_0)_* f_1 \circ \exp(-tf_0)(x) &= \begin{pmatrix} -t\sin(x_3) \\ t\cos(x_3) \\ 1 \end{pmatrix}, \end{aligned}$$

which proves our claim.

Any bang–singular arc satisfying PMP in normal form is a strong local optimizer for the problem. It is known, see Craven (1995), Teo et al. (1991), Fleming and Rishel (1975), Dmitruk (2007), that if $b_1^2 + (b_2 - 1)^2 > 1$, then any minimum time extremal is bang–singular with junction point on the cylinder $x_1^2 + (x_2 - 1)^2 = 1$.

References

AGRACHEV, A.A. and GAMKRELIDZE, R.V. (1997) Symplectic methods for optimization and control. B. Jacubczyk and W. Respondek, eds., *Geometry of Feedback and Optimal Control.* Pure and Applied Mathematics, Marcel Dekker, New York, 1–58.

AGRACHEV, A.A. and SACHKOV, Y.L. (2004) *Control Theory from the Geometric Viewpoint*. Springer-Verlag.

AGRACHEV, A.A., STEFANI, G. and ZEZZA, P. (1998A) An invariant second variation in optimal control. *Internat. J. Control*, **71**, 5, 689–715.

AGRACHEV, A.A., STEFANI, G. and ZEZZA, P. (1998B) Strong minima in optimal control. *Proc. Steklov Inst. Math.*, **220**, 4–26, translation from *Tr. Mat. Inst. Steklova* **220**, 8–22.

AGRACHEV, A.A., STEFANI, G. and ZEZZA, P (2002) Strong Optimality for a Bang–Bang Trajectory. *SIAM Journal on Control and Optimization* **41–4**, 991–1014.

ARNOLD, V.I. (1980) *Mathematical Methods in Classical Mechanics*. Springer, New York.

CRAVEN, B.D. (1995) *Control and Optimization*. Chapman & Hall.

DMITRUK, A.V. (2007) Optimal synthesis in the dodgem car problem. *Personal comunication*.

DMITRUK, A.V. (1977) Quadratic conditions for a weak minimum for singular regimes in optimal control problems. *English transl., Soviet Math. Dokl.*, **18**, 2.

DMITRUK, A.V. (1983) Quadratic conditions for a Pontryagin minimum in an optimal control problem, linear in the control, with a constraint on the control. *English transl., Soviet Math. Dokl.*, **28**, 2.

FLEMING, W.H. and RISHEL, R.W. (1975) *Deterministic and Stochastic Control*. Springer–Verlag, Berlin.

GABASOV, V. and KIRILLOVA, F.M. (1972) High order necessary conditions for optimality. *SIAM J. Control Optimization*, **10**, 127–188.

GARDNER MOYER, H. (1973) Sufficient Conditions for a Strong Minimum in Singular Control Problems. *SIAM Journal on Control*, **11** (4), 620–636.

GIAQUINTA, M.and HILDEBRANDT, S. (1996) *Calculus of Variations, I and II*. Springer, Berlin.

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

HESTENES, M.R. (1951) Applications of the theory of quadratic forms in Hilbert space to the calculus of variations. *Pacific J. Math*, **1**, 525–581.

POGGIOLINI, L. and STEFANI, G. (2005) Minimum time optimality for a bang –singular arc: second order sufficient conditions. In: *44th IEEE Conference on Decision and Control and European Control Conference ECC 2005*, on CD–ROM. IEEE Control Systems Society.

POGGIOLINI, L. and STEFANI, G. (2006) Minimum time optimality of a partially singular arc: second order conditions. *Proceedings of the 3rd Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control (LHMNLC'06)*. International Federation of Automatic Control (IFAC).

STEFANI, G. (2007) Strong optimality of singular trajectories. In: *F. Ancona, A. Bressan, P. Cannarsa, F. Clarke, P.R. Wolenski, eds., Geometric control and non-smooth analysis*. Applied Sciences, **76**, World Scientific Publishing Co.

STEFANI, G. (2004) Minimum–time optimality of a singular arc: second order sufficient conditions. *43th IEEE Conference on Decision and Control.*

TEO, K.L., GOH, C.J. and WONG, K.H. (1991) *A Unified Computational Approach to Optimal Control Problems.* Longman Scientific and Technical Company, England.

ZELIKIN, M.I. and BORIZOV, V.F. (1994) *Theory of Chattering Control.* Systems & Control: Foundations & Applications. Birkhäuser, Boston, Basel, Berlin.