ON LOCAL STATE OPTIMALITY OF PONTRYAGIN EXTREMALS IN THE MINIMUM TIME PROBLEM

Laura Poggiolini [*,1] **Marco Spadini** [*] **Gianna Stefani** [*,1]

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

Abstract: We give regularity and second order sufficient condition for (time, state)–local optimality and state–local optimality of a normal Pontryagin extremal in the minimum time problem, and describe the Hamiltonian approach to prove these different kinds of strong local optimality.

Keywords: Optimal control, second order sufficient conditions, Pontryagin extremal

1. INTRODUCTION

We consider the problem of strong local optimality for a Pontryagin extremal in the minimum time problem between two fixed end–points.

We assume that the state space is a n–dimensional smooth manifold M while the dynamics is assumed to be affine with respect to the control. The control space is some convex compact polyhedron in $\mathbb{R}^m$ and the vector fields $f_i\colon M \to TM$, $i = 0, 1, \dots, m$ are C^∞. Namely, we consider the following problem

$$
\begin{aligned}
&\text{minimize} \quad T \quad \text{subject to}\\
&\dot\xi(t) = f_0(\xi(t)) + \sum_{i=1}^{m} u_i(t) f_i(\xi(t)) \quad t \in [0,T]\\
&\xi(0) = x_0 \qquad \xi(T) = x_f\\
&u(t) \in \Omega \qquad t \in [0,T].
\end{aligned}
$$

We assume we are given a reference triplet $(\widehat{T}, \widehat{\xi}, \widehat{u})$ satisfying Pontryagin Maximum Principle (PMP), in normal form, with adjoint covector

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

and show that under some regularity conditions and the coercivity of a suitable second variation, then such triplet is a *strong local optimizer* for the problem.

By strong local optimizer we mean that we compare the time $\widehat{T}$ with the time T of all those admissible triplets (T, ξ, u) such that ξ is, in some sense, *near* $\widehat{\xi}$, regardless of any distance between $\widehat{u}$ and u. To be more precise, we consider two different kinds of strong local optimality:

- $(\widehat{T}, \widehat{\xi}, \widehat{u})$ is optimal among all those triplets (T, ξ, u) such that the graph of ξ belongs to a neighbourhood of the graph of $\widehat{\xi}$ in $\mathbb{R} \times M$ i.e. the optimality is local with respect to both state and final time. We call this type of local optimality *(time, state)–local*.
- $\widehat{\xi}$ is optimal among all those triplets (T, ξ, u) such that the range $\{\xi(t) \colon t \in [0,T]\}$ of ξ belongs to a neighbourhood of the range of $\widehat{\xi}$ in M i.e. the optimality is local only with respect to the state. We call this type of local optimality *state–local*.

Note that state–local optimality implies (time, state)–local optimality but the two notions are not equivalent, see (Stefani and Zezza, 2003).

The regularity conditions we require take different forms according to the possible different structure of the reference control function $\widehat{u}$, while the second variation changes both with the structure of $\widehat{u}$ and with the kind of strong local optimality one wishes to prove. The proof of both kinds of optimality is carried

[1] Partially supported by PRIN 2006019927_002, Controllabilità stabilizzabilità e ottimizzazione di sistemi dinamici non lineari

out by means of Hamiltonian methods which we are now going to describe.

2. HAMILTONIAN METHODS

Let $\pi\colon T^*M \to M$ be the canonical projection of the cotangent bundle onto the state space. We denote by $\widehat{H}$ the Hamiltonian associated to the reference control:

$$\widehat{H}\colon (t,\ell) \in [0,\widehat{T}] \times T^*M \mapsto \widehat{H}_t(\ell) = \\ = \langle \ell\,,\, f_0(\pi\ell) + \sum_{i=1}^{m} \widehat{u}_i(t) f_i(\pi\ell)\rangle \in \mathbb{R}$$

and by $H^{\max}$ the maximized Hamiltonian of the control system:

$$H^{\max}\colon \ell \in T^*M \mapsto H^{\max}(\ell) = \\ = \max_{u\in\Omega} \langle \ell\,,\, f_0(\pi\ell) + \sum_{i=1}^{m} u_i f_i(\pi\ell)\rangle \in \mathbb{R}$$

For any possibly time–dependent Hamiltonian G_t, we denote by the symbol $\overrightarrow{G}_t$ the associated Hamiltonian vector field.

PMP says that $\widehat{\lambda}$ is an absolutely continuous lift of the reference trajectory: $\pi\widehat{\lambda}(t) = \widehat{\xi}(t)$, $t \in [0,\widehat{T}]$, which solves the Hamiltonian system defined by the reference Hamiltonian

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\lambda}(t) = \overrightarrow{\widehat{H}}_t(\widehat{\lambda}(t)), \quad \text{a.e. } t \in [0,\widehat{T}]$$

and such that the maximality condition

$$\widehat{H}_t(\widehat{\lambda}(t)) = H^{\max}(\widehat{\lambda}(t))$$

holds for any $t \in [0,\widehat{T}]$, i.e. for any $t \in [0,\widehat{T}]$ we have

$$\widehat{u}(t) \in \operatorname{argmax}\Big\{ \langle \widehat{\lambda}(t)\,,\, (f_0 + \sum_{i=1}^{m} u_i f_i)(\widehat{\xi}(t))\rangle, \\ u = (u_1,\dots,u_m) \in \Omega \Big\}.$$

2.1 *(Time, state)–local optimality*

The Hamiltonian approach to prove (time, state)–local optimality is based on the construction of a new, possibly time dependent, Hamiltonian

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

having the following properties

- $H_t \geq H^{\max}$,
- $H_t(\widehat{\lambda}(t)) = \widehat{H}_t(\widehat{\lambda}(t))$,
- $\dfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\lambda}(t) = \overrightarrow{H}_t(\widehat{\lambda}(t))$,
- the Hamiltonian vector field $\overrightarrow{H}_t$ exists and its flow
$$\mathcal{H}\colon (t,\ell) \in [0,\widehat{T}] \times T^*M \mapsto \mathcal{H}_t(\ell) \in T^*M$$
is well defined locally around $\widehat{\lambda}(0)$

and on the definition of a horizontal Lagrangian manifold defined in a neighborhood $\mathcal{O}(x_0)$ of x_0 in M

$$\Lambda = \{\mathrm{d}\alpha(x)\colon x \in \mathcal{O}(x_0)\}$$

such that

- $\mathrm{d}\alpha(x_0) = \widehat{\lambda}(0)$
- the one–form $\omega := \mathcal{H}^*(p\,\mathrm{d}q - H_t\,\mathrm{d}t)$ is exact on $[0,\widehat{T}] \times \Lambda$
- $\mathrm{id}\times\pi\mathcal{H}\colon (t,\ell) \in [0,\widehat{T}] \times \Lambda \mapsto (t,\pi\mathcal{H}_t(\ell)) \in [0,\widehat{T}] \times M$ is one–to–one onto a neighborhood of the graph of $\widehat{\xi}$.

Remark that $(t,\widehat{\xi}(t)) = (\mathrm{id}\times\pi\mathcal{H})(t,\widehat{\lambda}(0))$.

Let us show how this construction leads to the result. Define

$$\mathcal{V} := (\mathrm{id}\times\pi\mathcal{H})([0,\widehat{T}] \times \Lambda) \\ \psi := (\mathrm{id}\times\pi\mathcal{H})^{-1}\colon \mathcal{V} \to [0,\widehat{T}] \times \Lambda$$

and let $(T,\,u,\,\xi)$ be an admissible triplet such that the graph of ξ is in $\mathcal{V}$. Assume, by contradiction, that $T < \widehat{T}$. We can obtain a closed path in $\mathcal{V}$ with a concatenation of the following paths:

- $\Xi\colon t \in [0,T] \mapsto (t,\xi(t)) \in \mathcal{V}$,
- $\Phi\colon t \in [T,\widehat{T}] \mapsto (t,x_f) \in \mathcal{V}$,
- $\widehat{\Xi}\colon t \in [0,\widehat{T}] \mapsto (t,\widehat{\xi}(t)) \in \mathcal{V}$, ran backward in time.

Since, the one–form $\omega := \mathcal{H}^*(p\,\mathrm{d}q - H_t\,\mathrm{d}t)$ is exact on $\mathcal{V}$, then we have

$$0 = \oint \omega = \int_{\psi(\Xi)} \omega + \int_{\psi(\Phi)} \omega - \int_{\psi(\widehat{\Xi})} \omega.$$

From the over–maximality properties of H_t we get

$$\int_{\psi(\widehat{\Xi})} \omega = 0 \qquad \int_{\psi(\Xi)} \omega \leq 0$$

so that

$$\int_{\psi(\Phi)} \omega \geq 0. \tag{1}$$

Since

$$\int_{\psi(\varphi)} \omega = \int_T^{\widehat{T}} \mathcal{H}^*(-H_t\,\mathrm{d}t)$$

and

$$H_t(\mathcal{H}_t(\psi(t,x_f))) = \\ H_{\widehat{T}}(\mathcal{H}_{\widehat{T}}(\psi(\widehat{T},x_f))) + O(1) = 1 + O(t-\widehat{T}), \tag{2}$$

inequality (1) and Taylor expansion (2) yield

$$0 \leq \int_T^{\widehat{T}} -\Big(1 + O(t-\widehat{T})\Big)\,\mathrm{d}t \\ = T - \widehat{T} + O\left((T-\widehat{T})^2\right)$$

which implies $T = \widehat{T}$ or T much smaller than $\widehat{T}$.

2.2 *State–local optimality*

In order to prove state–local optimality, one still has to construct a Hamiltonian with the properties described above, but the Lagrangian manifold Λ has to be replaced with a $(n-1)$–dimensional manifold, say $\widetilde{\Lambda}$, contained in $\{\ell \in T^*M \colon H(\ell) = 1\}$ and which is still horizontal:

$$\begin{aligned}\widetilde{\Lambda} := \{\ell \in T^*M \colon H(\ell) = 1,\ \ell = \mathrm{d}\widetilde{\alpha}(x), \\ x \in \mathcal{O}(x_0)\}, \quad \dim \widetilde{\Lambda} = n-1.\end{aligned}$$

$\widetilde{\Lambda}$ must also satisfy the following properties:

- $\mathrm{d}\widetilde{\alpha}(x_0) = \widehat{\lambda}(0)$
- the one–form $\widetilde{\omega} := \mathcal{H}^*(p\,\mathrm{d}q)$ is exact on $[0, \widehat{T}] \times \widetilde{\Lambda}$
- $\pi\mathcal{H} \colon (t, \ell) \in [0, \widehat{T}] \times \widetilde{\Lambda} \mapsto \pi\mathcal{H}_t(\ell) \in M$ is one–to–one onto a neighborhood of the range of $\widehat{\xi}$.

Remark that $\widehat{\xi}(t) = \pi\mathcal{H}(t, \widehat{\lambda}(0))$ and that a necessary condition for the invertibility of $\pi\mathcal{H} \colon [0, T] \times \widetilde{\Lambda} \to M$ is the injectivity of $\widehat{\xi}$, a condition which is not needed to prove (time, state)–local optimality.

Let us show how this construction leads to the result. Define

$$\begin{aligned}\widetilde{\mathcal{V}} &:= \pi\mathcal{H}([0, \widehat{T}] \times \widetilde{\Lambda}) \\ \psi &:= (\pi\mathcal{H})^{-1} \colon \widetilde{\mathcal{V}} \to [0, \widehat{T}] \times \widetilde{\Lambda}\end{aligned}$$

and let $(T,\, u,\, \xi)$ be an admissible triplet such that the range of ξ is in $\widetilde{\mathcal{V}}$. Assume, by contradiction, that $T < \widehat{T}$. We can obtain a closed path in $\widetilde{\mathcal{V}}$ concatenating ξ and $\widehat{\xi}$ ran backward in time. Since the one–form $\widetilde{\omega} := \mathcal{H}^*(p\,\mathrm{d}q)$ is exact on $\widetilde{\mathcal{V}}$, then we have

$$\begin{aligned}0 = \oint \omega &= \int_{\widetilde{\psi}(\xi)} \widetilde{\omega} - \int_{\widetilde{\psi}(\widehat{\xi})} \widetilde{\omega} \\ &= \int_0^T \langle \mathcal{H}(\widetilde{\psi}(\xi(t)))\,,\, \dot{\xi}(t) \rangle \,\mathrm{d}t - (\widehat{T} - 0) \\ &\le \int_0^T H(\mathcal{H}(\widetilde{\psi}(\xi(t))))\,\mathrm{d}t - \widehat{T} = T - \widehat{T}.\end{aligned}$$

which implies $T = \widehat{T}$.

3. THE PURE BANG–BANG CASE

The Hamiltonian approach was first used in (Agrachev *et al.*, 2002) where the authors prove strong optimality for a Mayer problem on a fixed time interval, but the proof can be adapted to prove (time, state)–local optimality for the minimum time problem. Then in (Poggiolini and Stefani, 2004) the authors prove state–local optimality for the minimum time problem (the procedure was then extended in (Poggiolini, 2006) to cover a general Bolza problem). In all these papers the reference control $\widehat{u}$ has a *pure bang–bang structure*. By pure bang–bang control we mean that the reference control $\widehat{u}$ is piecewise constant, takes values on the vertexes of Ω and all its switches are simple, i.e. if $\widehat{t}$ is a switching time for $\widehat{u}$, then $\operatorname{convhull}\{\widehat{u}(\widehat{t}-0),\ \widehat{u}(\widehat{t}+0)\}$ is an edge of Ω.

In both papers the regularity assumptions the authors ask for are the following

- $\widehat{u}$ is a *regular control* i.e.
  (1) if t is not a switching time for $\widehat{u}$, then
  $$\begin{aligned}\operatorname{argmax}\Big\{\langle \widehat{\lambda}(t)\,,\, (f_0 + \sum_{i=1}^m u_i f_i)(\widehat{\xi}(t))\rangle, \\ u = (u_1, \dots, u_m) \in \Omega\Big\}\end{aligned}$$
  is the singleton $\{\widehat{u}(t)\}$.
  (2) if $\widehat{t}$ is a switching time for $\widehat{u}$, then
  $$\begin{aligned}\operatorname{argmax}\Big\{\langle \widehat{\lambda}(t)\,,\, (f_0 + \sum_{i=1}^m u_i f_i)(\widehat{\xi}(t))\rangle, \\ u = (u_1, \dots, u_m) \in \Omega\Big\}\end{aligned}$$
  is the edge of Ω $\operatorname{convhull}\{\widehat{u}(\widehat{t}-0),\ \widehat{u}(\widehat{t}+0)\}$.
- the *strong bang–bang Legendre condition* holds:
  $$\frac{\mathrm{d}}{\mathrm{d}t}\left(\widehat{H}_{\widehat{t}+0} - \widehat{H}_{\widehat{t}-0}\right) \circ \widehat{\lambda}(t)\Big|_{t=\widehat{t}} > 0.$$
  Remark that the maximality condition in PMP implies that this derivative cannot be negative.
- In order to prove state–local optimality we also make the obvious additional request that $\widehat{\xi}$ is a simple curve

As previously said, the second variation one has to consider is different according to which kind of optimality one wishes to prove i.e. even if one studies the same quadratic form, nevertheless one asks for its positivity on different linear spaces.

We consider a finite dimensional sub–problem of the given one by allowing only for those controls that have the same sequence of values of the reference one. Namely let

$$0 := \widehat{t}_0 < \widehat{t}_1 < \ldots < \widehat{t}_r < \widehat{t}_{t+1} := \widehat{T}$$

be the switching times of $\widehat{u}$ and let

$$v^j := \widehat{u}|_{(\widehat{t}_{j-1}, \widehat{t}_j)} \quad j = 1, \dots, r+1.$$

For any small perturbation $(t_1, \dots, t_{r+1}) = (\widehat{t}_1 + \varepsilon_1, \dots, \widehat{t}_{r+1} + \varepsilon_{r+1})$ of the vector of switching times let u_ε be the control function such that

$$u_\varepsilon|_{(t_{j-1}, t_j)} := v^j \quad j = 1, \dots, r+1$$

where $t_0 := 0$. We consider the following sub–problem

$$\text{minimize} \quad t_{r+1} \quad \text{subject to} \tag{3}$$

$$\begin{aligned}\dot{\xi}(t) = f_0(\xi(t)) + \sum_{i=1}^m v_i^j(t) f_i(\xi(t)) \\ t \in (t_{j-1}, t_j), \quad j = 1, \dots, r+1\end{aligned} \tag{4}$$

$$\xi(0) = x_0, \tag{5}$$

$$\xi(t_{r+1}) = x_f. \tag{6}$$

Introducing the reference time increments

$$\widehat{\theta}_j := \widehat{t}_j - \widehat{t}_{j-1} \quad j = 1, \ldots, r+1,$$

the perturbed time increments

$$\theta_j := t_j - t_{j-1} = \widehat{\theta}_j + \varepsilon_j - \varepsilon_{j-1} \quad j = 1, \ldots, r+1,$$

$\varepsilon_0 := 0$ and denoting by $S(\theta)$ the solution of (4)–(5), then sub–problem (3)–(4)–(5)–(6) can be rewritten as

$$\text{minimize} \quad T(\theta) := \sum_{j=1}^{r+1} \theta_j \quad \text{subject to} \quad S(\theta) = x_f.$$

It is easily seen that PMP implies

$$\frac{\partial}{\partial \theta_j} T(\theta) - \langle \widehat{\lambda}(\widehat{T}) \,,\, \frac{\partial}{\partial \theta_j} S(\theta) \rangle \Big|_{\theta = \widehat{\theta}} = 0$$

for any $j = 1, \ldots, r+1$ i.e. the Lagrange multiplier rule is satisfied, $-\widehat{\lambda}(\widehat{T})$ being the multiplier.

We define the *second variation at the switching points* as the quadratic form

$$J'' \colon \theta \in \operatorname{Ker} \mathrm{D}S(\widehat{\theta}) \mapsto -\widehat{\lambda}(\widehat{\theta})\, \mathrm{D}^2 S(\widehat{\theta})[\theta]^2 \in \mathbb{R}$$

In order to prove (time, state)–local optimality we ask for the positivity of J'' on

$$W := \{\theta \in \mathbb{R}^{r+1} \colon \theta \in \operatorname{Ker} \mathrm{D}S(\widehat{\theta}),\ \theta_{r+1} = 0\}$$

while, to prove state–local optimality we ask for the positivity of J'' on the whole space $\operatorname{Ker} \mathrm{D}S(\widehat{\theta})$.

4. THE BANG–BANG CASE WITH MULTIPLE SWITCHES

A second step in considering this problem is the presence of multiple switches i.e. $\widehat{u}$ is still bang–bang but it is no longer pure, that is, at some switching time $\widehat{t}$ $\operatorname{convhull}\{\widehat{u}(\widehat{t}+0),\ \widehat{u}(\widehat{t}-0)\}$ is not an edge of Ω. To our knowledge only the problem of L^1–weak local optimality has been addressed in the general case, see (Sarychev, 1997), but in proving strong local optimality (and in the very close field of sensitivity analysis, see (Felgenhauer, 2007)) studies are still at an early stage.

We can study the case when Ω is a box: $\Omega = [-1, 1]^m$, so that the multiplicity of each switching time $\widehat{t}$ can be determined just by counting the number of components of $\widehat{u}$ that switch at time $\widehat{t}$) and only one double switch occurs, all the other switches being simple. The problem of (time, state)–local optimality for a minimum time trajectory can be recovered from (Poggiolini and Spadini, to appear) where the authors study a Bolza problem on a fixed time interval. In this talk we are going to face the problem of state–local optimality. As in the pure bang–bang case the regularity hypotheses one needs are the same for proving both kinds of strong local optimality. Let

$$0 := \widehat{\theta}_{00} < \widehat{\theta}_{01} < \ldots < \widehat{\theta}_{0J_0} < \widehat{\tau} < \\ < \widehat{\theta}_{11} < \ldots < \widehat{\theta}_{1,J_1} < \widehat{\theta}_{1,J_1+1} := \widehat{T}$$

be the sequence of the switching times, $\widehat{\tau}$ being the only double switching time. To simplify the notation define $\widehat{\theta}_{0,J_0+1} = \widehat{\theta}_{10} := \widehat{\tau}$. We assume the following regularity conditions

- $\widehat{u}$ is a *regular control* i.e.
  (1) if t is not a switching time for $\widehat{u}$, then
  $$\operatorname{argmax} \Big\{ \langle \widehat{\lambda}(t) \,,\, (f_0 + \sum_{i=1}^m u_i f_i)(\widehat{\xi}(t)) \rangle, \\ u = (u_1, \ldots, u_m) \in \Omega \Big\}$$
  is the singleton $\{\widehat{u}(t)\}$.
  (2) if $t = \widehat{\theta}_{ij}$ is a simple switching time for $\widehat{u}$, then
  $$\operatorname{argmax} \Big\{ \langle \widehat{\lambda}(t) \,,\, (f_0 + \sum_{i=1}^m u_i f_i)(\widehat{\xi}(t)) \rangle, \\ u = (u_1, \ldots, u_m) \in \Omega \Big\}$$
  is the edge of Ω given by $\operatorname{convhull}\{\widehat{u}(\widehat{\theta}_{ij} - 0),\ \widehat{u}(\widehat{\theta}_{ij} + 0)\}$.
  (3) if $t = \widehat{\tau}$ is the double switching time for $\widehat{u}$, then
  $$\operatorname{argmax} \Big\{ \langle \widehat{\lambda}(t) \,,\, (f_0 + \sum_{i=1}^m u_i f_i)(\widehat{\xi}(t)) \rangle, \\ u = (u_1, \ldots, u_m) \in \Omega \Big\}$$
  is the 2–dimensional face of Ω that contains $\operatorname{convhull}\{\widehat{u}(\widehat{\tau} - 0),\ \widehat{u}(\widehat{\tau} + 0)\}$.
- the *strong bang–bang Legendre condition* holds at the simple switching times:
  $$\frac{\mathrm{d}}{\mathrm{d}t} \left(\widehat{H}_{\widehat{\theta}_{ij}+0} - \widehat{H}_{\widehat{\theta}_{ij}-0} \right) \circ \widehat{\lambda}(t) \Big|_{t = \widehat{\theta}_{ij}} > 0$$
  Remark that the maximality condition in PMP implies that this derivative cannot be negative.
- the *strong bang–bang Legendre condition for double switching times* holds at the double switching time. Let us introduce this new notion: without loss of generality we may assume that the switching components of $\widehat{u}$ at time $\widehat{\tau}$ are the first and the second one. Denote by Δ_ν, $\nu = 1, 2$ their jumps:
  $$\Delta_\nu := \widehat{u}_\nu(\widehat{\tau} + 0) - \widehat{u}_\nu(\widehat{\tau} - 0) \quad \nu = 1, 2.$$
  Then
  $$\widehat{H}_{\widehat{\tau}+0}(\ell) = \widehat{H}_{\widehat{\tau}-0}(\ell) \\ + \Delta_1 \langle \ell \,,\, f_1(\pi\ell) \rangle + \Delta_2 \langle \ell \,,\, f_2(\pi\ell) \rangle.$$
  Define
  $$K_\nu(\ell) = \widehat{H}_{\widehat{\tau}-0}(\ell) + \Delta_\nu \langle \ell \,,\, f_\nu(\pi\ell) \rangle, \quad \nu = 1, 2.$$
  We assume
  $$\left. \begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t} \left(\widehat{K}_\nu - \widehat{H}_{\widehat{\tau}-0} \right) \circ \widehat{\lambda}(t) \Big|_{t = \widehat{\tau}-0} > 0 \\ &\frac{\mathrm{d}}{\mathrm{d}t} \left(\widehat{H}_{\widehat{\tau}+0} - \widehat{K}_\nu \right) \circ \widehat{\lambda}(t) \Big|_{t = \widehat{\tau}+0} > 0 \end{aligned} \right. \quad \nu = 1, 2.$$
  Remark that, as for simple switching times, the maximality condition in PMP implies that these one-side derivatives cannot be negative.

- *non–degeneracy condition at the double switching point*
  We assume
  $$\frac{\Delta_1 f_1(\widehat{\xi}(\widehat{\tau}))}{\boldsymbol{\sigma}\big(\overrightarrow{\widehat{H}}_{\widehat{\tau}-0}, \overrightarrow{K}_1\big)(\widehat{\lambda}(\widehat{\tau}))} \neq \frac{\Delta_2 f_2(\widehat{\xi}(\widehat{\tau}))}{\boldsymbol{\sigma}\big(\overrightarrow{\widehat{H}}_{\widehat{\tau}-0}, \overrightarrow{K}_2\big)(\widehat{\lambda}(\widehat{\tau}))}$$
  Differently from the other regularity hypotheses, the non–degeneracy condition cannot be considered a strengthening of some necessary condition, but it is needed to construct the manifolds Λ and $\widetilde{\Lambda}$ with the requested properties. To be more precise, the presence of a double switching time compels the authors to prove the invertibility of the projected flow with some differential topology technique and some transversality assumption is needed. Such transversality can be granted only if the non–degeneracy condition holds.
- As in the case of simple switches, in order to prove state–local optimality we also make the additional obvious request that $\widehat{\xi}$ is a simple curve

Also, the presence of a double switch gives rise to two different second variation, and each one of them is assumed to be positive (each on different spaces and, as in the case of a pure bang–bang reference control, also according to which kind of strong optimality one wishes to prove).

As in the pure bang–bang case we allow for the switching times of the reference control function to move but now, in doing so we must distinguish between the simple switching times and the double switching time. Moving a simple switching time $\widehat{\theta}_{i\,j}$ to time $\theta_{i\,j} := \widehat{\theta}_{i\,j} + \delta_{i\,j}$ amounts to using the value $v^{ij} := \widehat{u}\big|_{(\widehat{\theta}_{i,j-1}, \widehat{\theta}_{i,j})}$ and the value $v^{i,j+1} := \widehat{u}\big|_{(\widehat{\theta}_{i,j}, \widehat{\theta}_{i,j+1})}$ of the control function in the time intervals $\big(\widehat{\theta}_{i,j-1}, \theta_{i\,j}\big)$ and $\big(\theta_{i\,j}, \widehat{\theta}_{i,j+1}\big)$, respectively. On the other hand, when we move the double switching time $\widehat{\tau}$ we must keep in mind that we are moving the switching time of two different components of the reference control function and we therefore allow for two different perturbations of $\widehat{\tau}$. We call $\tau_1 := \widehat{\tau} + \varepsilon_1$ the perturbed switching time of $\widehat{u}_1$ and $\tau_2 := \widehat{\tau} + \varepsilon_2$ the perturbed switching time of $\widehat{u}_2$. Also, let us define

$$\theta_{0,J_0+1} := \min\{\tau_1, \tau_2\} \quad \theta_{10} := \max\{\tau_1, \tau_2\}$$

and

$$\begin{aligned} w^1 &:= v^{0,J_0+1} + (\Delta_1, 0, 0, \ldots, 0) \\ w^2 &:= v^{0,J_0+1} + (0, \Delta_2, 0, \ldots, 0) \end{aligned}$$

In each *simply perturbed* interval of time we do as in the pure bang–bang case, setting

$$u_{(\varepsilon,\delta)}\big|_{(\theta_{i,j-1}, \theta_{ij})} := v^{ij} \quad i = 1, \ldots, J_i + 1,\ i = 0, 1$$

while we set

$$u_{(\varepsilon,\delta)}\big|_{(\theta_{0,J_0+1}, \theta_{10})} := \begin{cases} w^1 & \text{if } \varepsilon_1 < \varepsilon_2 \\ w^2 & \text{if } \varepsilon_2 < \varepsilon_1 \end{cases}$$

This procedure gives rise to two different finite–dimensional sub–problems P_ν, $\nu = 1, 2$ given by

$$\text{minimize } \ \theta_{1,J_1+1} \ \text{ subject to} \qquad (P_\nu\text{a})$$

$$\dot{\xi}(t) = \begin{cases} (f_0 + \sum_{k=1}^m v_k^{0j} f_k)(\xi(t)) \\ \qquad t \in (\theta_{0,j-1}, \theta_{0,j}) \\ \qquad j = 1, \ldots, J_0 + 1 \\ (f_0 + \sum_{k=1}^m w_k^{\nu} f_k)(\xi(t)) \\ \qquad t \in (\theta_{0,J_0+1}, \theta_{10}) \\ (f_0 + \sum_{k=1}^m v_k^{0j} f_k)(\xi(t)) \\ \qquad t \in (\theta_{0,j-1}, \theta_{0j}) \\ \qquad j = 1, \ldots, J_1 + 1 \end{cases} \qquad (P_\nu\text{b})$$

$$\text{and } \ \xi(0) = x_0 \qquad \xi(\theta_{1,J_1+1}) = x_f. \qquad (P_\nu\text{c})$$

We shall denote the solution, evaluated at time t, of $(P_\nu\text{b})$ emanating from x_0 at time 0 as $S_t^\nu(\delta, \varepsilon)$.

Notice that P_1 is defined only for $\varepsilon_1 \leq \varepsilon_2$, while P_2 is defined only for $\varepsilon_2 \leq \varepsilon_1$, and the reference control is the one we obtain when every δ_{ij} and ε_k is zero, i.e. in a point on the boundary of the domain of P_ν. But from PMP we still have that the Lagrange multiplier rule is satisfied with $-\widehat{\lambda}(\widehat{T})$ as multiplier, hence we can consider the second variation for the constrained problems P_1 and P_2. We shall ask for their second order variations to be positive (on proper spaces) and prove the following theorems:

Theorem 1. ((Time, state)–local optimality). Assume $(\widehat{T}, \widehat{\xi}, \widehat{u})$ is a bang–bang normal Pontryagin extremal for the minimum time problem with associated covector $\widehat{\lambda}$. Assume all the switching times of $\widehat{u}$ but one are simple, while the only non–simple switching time is double.
Assume the Legendre conditions for simple and for double switching times hold. Also, assume the non degeneracy condition holds at the double switching time. Assume also that the second variation J''_ν of each problem P_ν is positive definite on

$$\{(\delta, \varepsilon) \in \operatorname{Ker} DS^\nu(0,0), \quad \delta_{1,J_1+1} = 0\}.$$

Then $(\widehat{T}, \widehat{\xi}, \widehat{u})$ is a strict (times, state)–local optimizer for the minimum time problem.

Theorem 2. (State–local optimality). Let $(\widehat{T}, \widehat{\xi}, \widehat{u})$ be a bang–bang normal Pontryagin extremal for the minimum time problem with associated covector $\widehat{\lambda}$. Assume all the switching times of $\widehat{u}$ but one are simple, while the only non–simple switching time is double.
Assume the Legendre conditions for simple and for double switching times hold. Also, assume the non degeneracy condition holds at the double switching time. Assume also that the second variation J''_ν of each problem P_ν is positive definite on $\operatorname{Ker} DS^\nu(0,0)$. Then $(\widehat{T}, \widehat{\xi}, \widehat{u})$ is a strict state–local optimizer for the minim time problem.

4.1 *A model case*

In both the bang–bang cases presented here one can actually construct a maximized Hamiltonian, that is a Hamiltonian function which actually agrees with $H^{\max}$. Such construction is based on the implicit function theorem which can be applied thanks to the regularity conditions. In this section we show how this procedure is performed when only one switch occurs and that switch is double.

Thanks to the regularity condition, in a neighborhood of $\widehat{\lambda}(t)$, $t \in [0, \widehat{\tau})$, the maximized Hamiltonian is given by

$$\widehat{H}_{\widehat{\tau}-0}(\ell) := \langle \ell\,,\, f_0(\pi\ell) + \sum_{i=1}^{m} v_i^{01} f_i(\pi\ell)\rangle.$$

Let $\exp t\overrightarrow{\widehat{H}}_{\widehat{\tau}-0}(\ell)$ be the flow of $\overrightarrow{\widehat{H}}_{\widehat{\tau}-0}(\ell)$ emanating from ℓ at time 0.

We then consider the two implicit equations, $\nu = 1, 2$

$$\begin{cases} \tau_\nu(\ell) = \widehat{\tau} \\ (K_\nu - \widehat{H}_{\widehat{\tau}-0}) \circ \exp \tau_\nu(\ell) \overrightarrow{\widehat{H}}_{\widehat{\tau}-0}(\ell) = 0 \end{cases}$$

The strong bang–bang Legendre condition for double switching times permits to define, in a neighborhood $\mathcal{O} := \mathcal{O}(\widehat{\lambda}(0))$ of $\widehat{\lambda}(0)$ in T^*M two C^1 functions:

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

Define $\vartheta_{01}(\ell) := \min\{\tau_\nu(\ell) \colon \nu = 1, 2\}$ and consider the two implicit equations, $\nu = 1, 2$

$$\begin{cases} \vartheta_{10}^\nu(\ell) = \widehat{\tau} \\ (\widehat{H}_{\widehat{\tau}+0} - K_\nu) \circ \exp \vartheta_{10}^\nu(\ell) \overrightarrow{K}_\nu(\ell) = 0. \end{cases}$$

Again, the strong bang–bang Legendre condition for double switching times allows us to define, possibly restricting $\mathcal{O}$, two C^1 functions:

$$\vartheta_{10}^\nu \colon \ell \in \mathcal{O} \mapsto \tau_\nu(\ell) \in \mathbb{R}$$

and we set

$$\vartheta_{10}(\ell) := \begin{cases} \vartheta_{10}^1 & \text{if } \vartheta_{01}(\ell) = \tau_1(\ell) \\ \vartheta_{10}^2 & \text{if } \vartheta_{01}(\ell) = \tau_2(\ell) \end{cases}$$

The maximized Hamiltonian is thus given by

$$H_t(\ell) = \begin{cases} \widehat{H}_{\widehat{\tau}-0}(\ell) & \text{if } 0 \le t < \vartheta_{01}(\ell) \\ K_\nu(\ell) & \begin{array}{l}\text{if } \vartheta_{01}(\ell) < t < \vartheta_{10}(\ell) \\ \quad \text{and } \vartheta_{01}(\ell) = \tau_\nu(\ell)\end{array} \\ \widehat{H}_{\widehat{\tau}+0}(\ell) & \text{if } \vartheta_{10}(\ell) < t \le \widehat{T} \end{cases}$$

Let us show how to define $\widetilde{\Lambda}$. The positivity of both the second variations J''_ν of problems P_ν allows us to remove the constraint on the initial point of the trajectories and introduce a penalty $\alpha(\xi(0))$ on this point in such a way that the second variations $J''_{\alpha,\nu}$ of the problems $P_{\alpha,\nu}$ one thus obtain are positive definite on a linear space containing $\operatorname{Ker} \mathrm{D}S(0,0)$. Namely the problems $P_{\alpha,\nu}$ are given by

$$\text{minimize } \alpha(\xi(0)) + \theta_{11} \qquad (P_{\alpha,\nu}\text{a})$$

subject to

$$\dot{\xi}(t) = \begin{cases} (f_0 + \sum_{k=1}^m v_k^{01} f_k)(\xi(t)) \\ \qquad t \in (0, \theta_{01}) \\ (f_0 + \sum_{k=1}^m w_k^{\nu} f_k)(\xi(t)) \\ \qquad t \in (\theta_{01}, \theta_{10}) \\ (f_0 + \sum_{k=1}^m v_k^{11} f_k)(\xi(t)) \\ \qquad t \in (\theta_{10}, \theta_{11}) \end{cases} \qquad (P_{\alpha,\nu}\text{b})$$

$$\text{and } \ \xi(\theta_{11}) = x_f. \qquad (P_{\alpha,\nu}\text{c})$$

We shall denote the solution, evaluated at time t, of $(P_{\alpha,\nu}\text{b})$ emanating from x at time 0 as $\Sigma_t^\nu(x, \delta, \varepsilon)$. α can be chosen so that the second variation $J''_{\alpha,\nu}$ of $P_{\alpha,\nu}$ is positive definite on $V_\alpha^\nu := \operatorname{Ker} \mathrm{D}\Sigma_{\widehat{T}}^\nu(\widehat{x}_0, 0, 0)$. The initial manifold which grants the invertibility of $\pi\mathcal{H} \colon \widetilde{\Lambda} \to M$ is thus given by

$$\widetilde{\Lambda} := \{\ell = \mathrm{d}\alpha(x),\ x \in M,\ \widehat{H}_{\widehat{\tau}-0}(\ell) = 1\}.$$

The positivity of the second variations $J''_{\alpha,\nu}$ grant both the transversality of the manifold $\{\ell = \mathrm{d}\alpha(x),\ x \in M\}$ and of the level set $\{\ell \in T*M,\ \widehat{H}_{\widehat{\tau}-0}(\ell) = 1\}$ and the invertibility of $\pi\mathcal{H} \colon [0, \widehat{T}] \times \widetilde{\Lambda} \to M$.

5. REFERENCES

Agrachev, Andrei A., Gianna Stefani and PierLuigi Zezza (2002). Strong optimality for a bang-bang trajectory. *SIAM J. Control Optimization* **41**(4), 991–1014.

Felgenhauer, Ursula (2007). On lipschitz stable switching behavior of bang-bang extremals in parametric control. In: *23rd IFIP Conference on System Modelling and Optimization, Krakow (Poland)*.

Poggiolini, Laura (2006). On local state optimality of bang-bang extremals in a free horizon bolza problem. *Rendiconti del Seminario Matematico dell'Universitá e del Politecnico di Torino*.

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

Poggiolini, Laura and Marco Spadini (to appear). *Proceedings of WMCTF07*. Chap. Sufficient optimality conditions for a bang–bang trajectory in a Bolza problem. Business/Economics and Statistics. Springer–Verlag.

Sarychev, Andrei V. (1997). First and second order sufficient optimality conditions for bang-bang controls. *SIAM J. Control Optimization* **1**(35), 315–340.

Stefani, Gianna and Pier Luigi Zezza (2003). Time–optimality of a bang–bang trajectory with maple. In: *2nd IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control*.