Strong local optimality for generalised L^1 optimal control problems

F. C. Chittaro, L. Poggiolini

LIS (Toulon)

23rd International Symposium on Mathematical Programming
July 3rd, 2018

The problem

M n-dimensional smooth manifold, f_0, f_1 smooth vector fields on M
$\psi : M \to \mathbb{R}$ smooth function (cost)
$T \geqslant 0$, $q_0, q_f \in M$

$$\text{minimise} \int_0^T |u(t)\psi(\xi(t))| \, dt \tag{1}$$

under the constraints

$$\begin{cases} \dot{\xi} = (f_0 + u(t)f_1) \circ \xi(t) \\ \xi(0) = q_0, \quad \xi(T) = q_f \\ \mathbf{u}(\cdot) \in L^\infty([0,T],[-1,1]). \end{cases} \tag{2}$$

Definition

An admissible pair $(\widehat{\xi}, \widehat{u})$ for (2) is a *strong*-local minimiser if there exists a neighbourhood $\mathcal{U}$ in $[0,T] \times M$ of the graph of $\widehat{\xi}$ such that
$\int_0^T |\widehat{u}(t)\psi(\widehat{\xi}(t))|dt \leqslant \int_0^T |u(t)\psi(\xi(t))|dt$
for every admissible pair (ξ, u) of (2) with $\text{graph}\{\xi\} \subset \mathcal{U}$.

Necessary Optimality Conditions: PMP

$h(p,q,u) : T^*M \times U \to \mathbb{R}$ control-dependent Hamiltonian

$$h(p,q,u) = \langle p, f_0(q)\rangle + u\langle p, f_1(q)\rangle + p_0|u\psi(q)|$$

Pontryagin Maximum Principle: if $(\widehat{\xi}, \widehat{u})$ is optimal, then $\widehat{\xi}$ is the projection on M of a solution $(\widehat{p}(t), \widehat{\xi}(t))$ (*extremal*) of the Hamiltonian system associated with the Hamiltonian h.
Moreover $(\widehat{p}(t), \widehat{\xi}(t))$ satisfies the following conditions

$$(\widehat{p}(t), p_0) \neq (0,0) \quad \forall t, \quad p_0 \leqslant 0 \tag{3}$$

$$\frac{d\widehat{p}(t)}{dt} = -\frac{\partial h}{\partial q}\Big|_{\widehat{p}(t),\widehat{\xi}(t),\widehat{u}(t)} \quad \frac{d\widehat{\xi}(t)}{dt} = \frac{\partial h}{\partial p}\Big|_{\widehat{p}(t),\widehat{\xi}(t),\widehat{u}(t)} \quad \text{a.e.t} \tag{4}$$

$$h(\widehat{p}(t), \widehat{\xi}(t), \widehat{u}(t)) = \max_{u\in U} h(\widehat{p}(t), \widehat{\xi}(t), u) \quad \text{a.e.t} \tag{5}$$

In the following, $H_{\max}(p,q) = \max_{u\in U} h(p,q,u)$

Necessary Optimality Conditions: PMP (cont'd)

The *normal*[1] maximised Hamiltonian associated with (1)-(2) is

$$H_{\max}(p,q) = \max_{u\in[-1,1]} F_0(p,q) + uF_1(p,q) - |u||\psi(q)|, \quad F_i(p,q) = \langle p, f_i(q)\rangle \quad (6)$$

$\mathbf{F_1(p,q) - \psi(q)}$ and $\mathbf{F_1(p,q) + \psi(q)}$ are called *switching functions* and determine the value of the control realising the maximum in (6).

We can distinguish three possible cases:

- if $|F_1(p,q)| > |\psi(q)|$, then the maximising control is either 1 or -1 (*bang*).
- if $|F_1(p,q)| = |\psi(q)|$, then the maximising control is not uniquely determined by the PMP (*singular*).

[1] i.e. $p_0 < 0$

Necessary Optimality Conditions: PMP (cont'd)

The *normal*[1] maximised Hamiltonian associated with (1)-(2) is

$$H_{\max}(p,q) = \max_{u\in[-1,1]} F_0(p,q) + uF_1(p,q) - |u||\psi(q)|, \quad F_i(p,q) = \langle p, f_i(q)\rangle \quad (6)$$

$F_1(p,q) - \psi(q)$ and $F_1(p,q) + \psi(q)$ are called *switching functions* and determine the value of the control realising the maximum in (6).

We can distinguish three possible cases:

- if $|F_1(p,q)| > |\psi(q)|$, then the maximising control is either 1 or -1 (*bang*).
- if $|F_1(p,q)| = |\psi(q)|$, then the maximising control is not uniquely determined by the PMP (*singular*).
- if $|F_1(p,q)| < |\psi(q)|$, the maximising control is 0 (*inactivated*).

[1] i.e. $p_0 < 0$

Aim of the talk

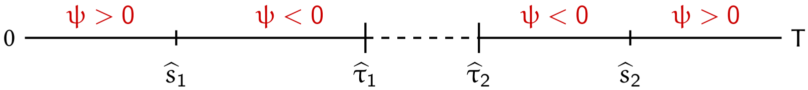
Establishing **sufficient optimality condition** via Hamiltonian methods for a *bang-inactivted-bang* extremal with control of the form

$$\widehat{u}(t) = \begin{cases} u_1 & t \in [0, \widehat{\tau}_1), \\ 0 & t \in (\widehat{\tau}_1, \widehat{\tau}_2), \\ u_3 & t \in (\widehat{\tau}_2, T], \end{cases} \tag{7}$$

$0 < \widehat{\tau}_1 < \widehat{\tau}_2 < T, u_1, u_3 \in \{1, -1\}$.

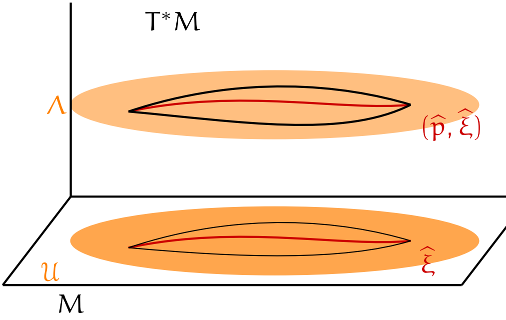
Working hypothesis (just for the talk):

- $u_1 = 1, u_3 = -1$
- $\psi(\widehat{\xi}(t))$ *vanishes only once in* $[0, \widehat{\tau}_1]$ *and only once in* $[\widehat{\tau}_2, T]$. *In particular,*



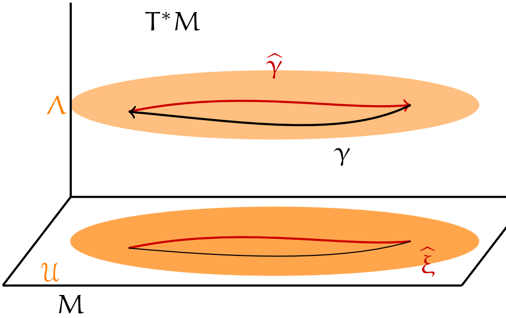
Hamiltonian methods: main ideas

- General setting: φ is smooth, $H_{max} = \max_u \langle p, \dot{q} \rangle - \varphi(q, u)$ is C^2
- Define a Lagrangian submanifold $\Lambda \subset T^*M$ such that
  - $(\widehat{p}(t), \widehat{\xi}(t)) \in \Lambda \ \forall t \in [0, T]$
  - the projection on M is a **diffeomorphism** from Λ to $\mathcal{U}$
  - $\mathbf{p}\,\mathbf{dq} - \mathbf{H}_{max}(\mathbf{p}, \mathbf{q})\,\mathbf{dt}$ is *exact* on $[0, T] \times \Lambda$
- for every admissible pair (ξ, u) such that the range of ξ is in $\mathcal{U}$, lift $\xi(t)$ to some curve $(p(t), \xi(t)) \in \Lambda$.



Hamiltonian methods: main ideas

- General setting: φ is smooth, $H_{max} = \max_u \langle p, \dot{q} \rangle - \varphi(q, u)$ is C^2
- Define a Lagrangian submanifold $\Lambda \subset T^*M$ such that
  - $(\widehat{p}(t), \widehat{\xi}(t)) \in \Lambda \ \forall t \in [0, T]$
  - the projection on M is a **diffeomorphism** from Λ to $\mathcal{U}$
  - $\mathbf{pdq - H_{max}(p, q)dt}$ is *exact* on $[0, T] \times \Lambda$
- for every admissible pair (ξ, u) such that the range of ξ is in $\mathcal{U}$, lift $\xi(t)$ to some curve $(p(t), \xi(t)) \in \Lambda$.



Hamiltonian methods: main ideas (cont'd)

- let

$$\widehat{\gamma} = \{(t, \widehat{p}(t), \widehat{\xi}(t)) : t \in [0, T]\} \qquad \gamma = \{(t, p(t), \xi(t)) : t \in [T, 0]\}$$

and $\Gamma = \gamma \cup \widehat{\gamma}$
Since Γ is a closed circuit in $[0, T] \times \Lambda$, then

$$\oint_\Gamma p\,dq - H_{\max}(p, q)\,dt = 0$$

- since $\displaystyle\int_0^T \varphi(\widehat{\xi}(t), \widehat{u}(t))dt = \int_{\widehat{\gamma}} p\,dq - H_{\max}(p, q)\,dt$

$$\int_0^T \varphi(\xi(t), u(t))dt = \int_\gamma p\,dq - h(p, q, u(t))\,dt \geqslant \int_\gamma p\,dq - H_{\max}(p, q)\,dt$$

then $\int_0^T \varphi(\widehat{\xi}(t), \widehat{u}(t))dt \leqslant \int_0^T \varphi(\xi(t), u(t))dt$.

Hamiltonian methods for bang-inactivated-bang arcs

Main points for the application of Hamiltonian methods:

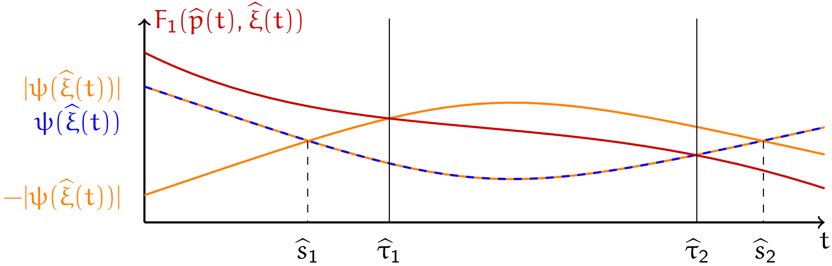
- to ensure that maximised Hamiltonian of PMP is well defined, continuous and piecewise smooth (in $\mathcal{U}$), we need some *regularity conditions*
- given any function $\alpha : M \to \mathbb{R}$ with $d\alpha(\widehat{\xi}(0)) = \widehat{p}(0)$, its graph $\Lambda_0 = \{d\alpha(x) : x \in M\} \subset T^*M$ is a Lagrangian submanifold with injective projection onto M.
- let Λ_t be the image at the time t of Λ_0 via the Hamiltonian flow; following [ASZ98] [2] Λ_t projects **injectively** onto $\mathcal{U}$ as long as the *second variation* associated with (1)-(2) is **coercive**
- as for bang-bang extremals (see for instance [ASZ02,PS04] [3]), the second variation for (1)-(2) is **finite-dimensional** (only variations of the switching times are considered)

[2] ASZ98 A. A. Agrachev, G. Stefani, P. Zezza Int. J Control 71 (5), 1998

[3] ASZ02 A.A. Agrachev, G. Stefani, P. Zezza, SIAM J CONTROL OPT, 41(4), 2002.
PS04 L. Poggiolini, G. Stefani, Syst Control Lett, 53(2), 2004.

Regularity Assumptions

- the reference trajectory satisfies PMP with $p_0 = -1$ (normal extremal)
- the function ψ does not vanish at the switching points: $\psi(\widehat{\xi}(\widehat{\tau}_i)) \neq 0, \ i = 1, 2.$
- $L_{f_0+\widehat{u}(t)f_1}\psi(\widehat{\xi}(t)) \neq 0 \ \forall t \in [0, T] \Rightarrow$ along the reference trajectories, ψ annihilates at a *finite* number of points.
- the switching functions have the following behaviour:



Regularity Assumptions (cont'd)

- $$\begin{aligned} u_1 F_1((\widehat{p}(t),\widehat{\xi}(t))) &> \left|\psi(\widehat{\xi}(t))\right|, \ t \in (0,\widehat{\tau}_1), \\ \left|F_1((\widehat{p}(t),\widehat{\xi}(t)))\right| &< \left|\psi(\widehat{\xi}(t))\right|, \ t \in (\widehat{\tau}_1,\widehat{\tau}_2), \\ u_3 F_1((\widehat{p}(t),\widehat{\xi}(t))) &> \left|\psi(\widehat{\xi}(t))\right|, \ t \in (\widehat{\tau}_2,T). \end{aligned}$$

  $\Rightarrow$ Along each arc the reference control is the only one that maximises over $u \in [-1,1]$ the control-dependent Hamiltonian evaluated along $\widehat{\lambda}$.
- $\dfrac{d}{dt}\big(F_1 - |\psi|\big)(\widehat{p}(t),\widehat{\xi}(t))|_{t=\widehat{\tau}_1} < 0$
- $\dfrac{d}{dt}\big(F_1 + |\psi|\big)(\widehat{p}(t),\widehat{\xi}(t))|_{t=\widehat{\tau}_2} > 0$

Remark: in the equations here above, weak inequalities are a necessary condition for the optimality of the reference extremal (PMP).

Construction of the field of extremals

At the initial point of the reference extremal $\bullet = (\widehat{p}(0), \widehat{\xi}(0))$, the maximised Hamiltonian flow is given by

$$\mathcal{H}_t(\bullet) = \begin{cases} \exp(t(\vec{F}_0 + \vec{F}_1 - D\psi))(\bullet) & t \in [0, \widehat{s}_1) \\ \exp((t - \widehat{s}_1)(\vec{F}_0 + \vec{F}_1 + D\psi)) \circ \mathcal{H}_{\widehat{s}_1}(\bullet) & t \in (\widehat{s}_1, \widehat{\tau}_1) \\ \exp((t - \widehat{\tau}_1)(\vec{F}_0)) \circ \mathcal{H}_{\widehat{\tau}_1}(\bullet) & t \in (\widehat{\tau}_1, \widehat{\tau}_2) \\ \exp((t - \widehat{\tau}_2)(\vec{F}_0 - \vec{F}_1 - D\psi)) \circ \mathcal{H}_{\widehat{\tau}_2}(\bullet) & t \in (\widehat{\tau}_2, \widehat{s}_2) \\ \exp((t - \widehat{s}_2)(\vec{F}_0 - \vec{F}_1 + D\psi))) \circ \mathcal{H}_{\widehat{s}_2}(\bullet) & t \in (\widehat{s}_2, T] \end{cases}$$

How to define $\mathcal{H}_t$ in a neighbourhood of $\bullet$ in T^*M?

Construction of the field of extremals

At the initial point of the reference extremal $\bullet = (\widehat{p}(0), \widehat{\xi}(0))$, the maximised Hamiltonian flow is given by

$$\mathcal{H}_t(\bullet) = \begin{cases} \exp(t(\vec{F}_0 + \vec{F}_1 - D\psi))(\bullet) & t \in [0, \widehat{s}_1) \\ \exp((t - \widehat{s}_1)(\vec{F}_0 + \vec{F}_1 + D\psi)) \circ \mathcal{H}_{\widehat{s}_1}(\bullet) & t \in (\widehat{s}_1, \widehat{\tau}_1) \\ \exp((t - \widehat{\tau}_1)(\vec{F}_0)) \circ \mathcal{H}_{\widehat{\tau}_1}(\bullet) & t \in (\widehat{\tau}_1, \widehat{\tau}_2) \\ \exp((t - \widehat{\tau}_2)(\vec{F}_0 - \vec{F}_1 - D\psi)) \circ \mathcal{H}_{\widehat{\tau}_2}(\bullet) & t \in (\widehat{\tau}_2, \widehat{s}_2) \\ \exp((t - \widehat{s}_2)(\vec{F}_0 - \vec{F}_1 + D\psi))) \circ \mathcal{H}_{\widehat{s}_2}(\bullet) & t \in (\widehat{s}_2, T] \end{cases}$$

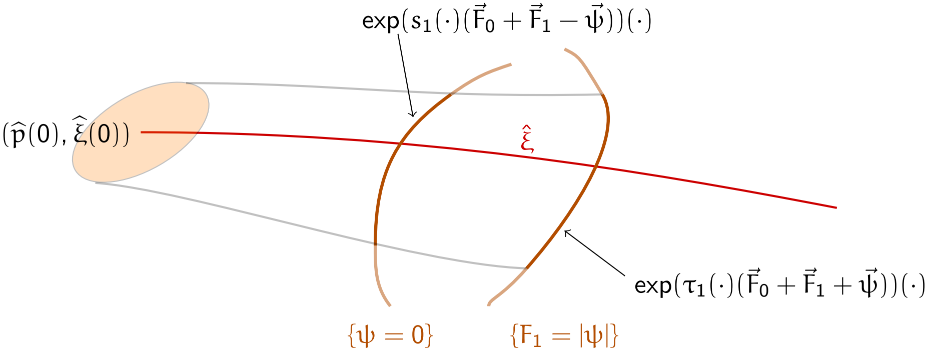
How to define $\mathcal{H}_t$ in a neighbourhood of $\bullet$ in T^*M?

Lemma

For (p, q) in a sufficiently small neighbourhood of $\bullet$, the "switching times" s_1, s_2, τ_1, τ_2 are smooth functions of (p, q).

Proof based on the Implicit Function Theorem and regularity assumptions

Construction of the field of extremals (cont'd)



Construction of the field of extremals (cont'd)

For every (p, q) in a sufficiently small neighbourhood of $\bullet$, the maximised Hamiltonian flow is given by

$$\mathcal{H}_t(p,q) = \begin{cases} \exp(t(\vec{F}_0 + \vec{F}_1 - D\psi))(p,q) & t \in [0, s_1(p,q)) \\ \exp((t - s_1)(\vec{F}_0 + \vec{F}_1 + D\psi)) \circ \mathcal{H}_{s_1(p,q)}(p,q) & t \in (s_1(p,q), \tau_1(p,q) \\ \exp((t - \tau_1)(\vec{F}_0)) \circ \mathcal{H}_{\tau_1(p,q)}(p,q) & t \in (\tau_1(p,q), \tau_2(p,q) \\ \exp((t - \tau_2)(\vec{F}_0 - \vec{F}_1 - D\psi)) \circ \mathcal{H}_{\tau_2(p,q)}(p,q) & t \in (\tau_2(p,q), s_2(p,q) \\ \exp((t - s_2)(\vec{F}_0 - \vec{F}_1 + D\psi))) \circ \mathcal{H}_{s_2(p,q)}(p,q) & t \in (s_2(p,q), T] \end{cases}$$

Invertibility of the flow

Consider a Lagrangian submanifold $\Lambda_0 \subset T * M$ containing $(\widehat{p}(0), \widehat{\xi}(0))$ and with injective projection onto M. Let $\Lambda_t = \mathcal{H}_t(\Lambda_0),\ t \geqslant 0$.
The following facts hold true:

- the projection of Λ_t onto M is injective **for every $\mathbf{t} \in [0, \widehat{\boldsymbol{\tau}}_1)$**; indeed, $(\pi \circ \mathcal{H}_t)_* = \exp(t(f_0 + f_1))$
- at $t = \widehat{\tau}_1$, $\pi \circ \mathcal{H}_t$ has different left and right linearisation;
- if $(\pi \circ \mathcal{H}_t)_*$ is non-singular at $t = \widehat{\tau}_1$, then the projection of Λ_t onto M is injective **for every $\mathbf{t} \in [0, \widehat{\boldsymbol{\tau}}_2)$**;
- at $t = \widehat{\tau}_2$, $\pi \circ \mathcal{H}_t$ has different left and right linearisation;
- if $(\pi \circ \mathcal{H}_t)_*$ is non-singular at $t = \widehat{\tau}_2$, then the projection of Λ_t onto M is injective **for every $\mathbf{t} \in [0, \mathbf{T}]$**.

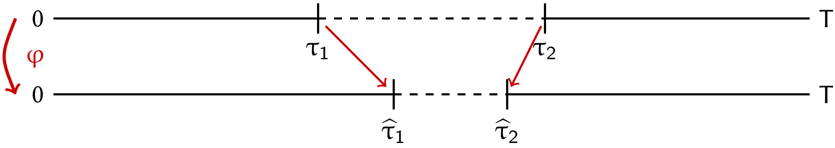
$\Rightarrow$ To ensure the bijectivity of the projection of Λ_t onto M, it is sufficient to check the non-singularity of $(\pi \circ \mathcal{H}_t)_*$ **at the switching times**.

The second variation - construction

- Consider a *finite-dimensional sub-problem* of the original one (1)-(2):

$$\begin{cases} \dot{\xi} = \begin{cases} (f_0+f_1)\circ\xi(t) & t\in[0,\tau_1) \\ (f_0)\circ\xi(t) & t\in[\tau_1,\tau_2) \\ (f_0-f_1)\circ\xi(t) & t\in[\tau_2,T] \end{cases} \\ \xi(0)=q_0, \quad \xi(T)=q_f \\ 0<\tau_1<\tau_2<T \end{cases} \tag{8}$$

- for every candidate trajectory satisfying (8), perform a piecewise-affine time-reparametrization φ



$\dot{\varphi} = 1+\frac{\theta_1}{\hat{\tau}_1}$ $\quad\dot{\varphi} = 1+\frac{\theta_2}{\hat{\tau}_2-\hat{\tau}_1}$ $\quad\dot{\varphi} = 1+\frac{\theta_3}{T-\hat{\tau}_2}$

$$\theta_i = \hat{\tau}_i - \hat{\tau}_{i-1} - \tau_i - \tau_{i-1}$$

The second variation - construction

- the OCP can be then written as

$$\min_{x,\theta} J(x,\theta)$$

$$J(x,\theta) = \alpha(x) + \widehat{\beta}(\zeta_T(x,\theta)) + \int_{[0,\widehat{\tau}_1]\cup[\widehat{\tau}_2,T]} \dot{\varphi}_\theta(t)\left|\widehat{\psi}_t(\zeta_t(x,\theta))\right| dt.$$

where
  - $\zeta_t(x,\theta)$ is the pullback of $\xi(\cdot)$ with respect to the reference flow
  - α and β are some smooth functions satisfying $d\alpha(q_0) = p_0$, $d\widehat{\beta}(q_0) = -p_T$
  - $\widehat{\psi}_t$ is ψ composed with the reference flow at the time t
- the second variation is just the Hessian of $J(x,\theta)$ finite-dimensional problem!

The second variation

Following ASZ02, it is possible to define a *two dimensional subspace* $V_0 \subset T_{q_0}M \times \mathbb{R}^3$ and quadratic form $J'' : T_{q_0}M \times \mathbb{R}^3 \to \mathbb{R}$ such that $J''|_{V_0}$ is the second variation of the problem (2)-(1). *(details in the paper)*

Theorem

If $J''|_{V_0}$ *is coercive, then the flow* $\pi \circ \mathcal{H}_t$ *is invertible for every* $t \in [0, T]$

The second variation

Following ASZ02, it is possible to define a *two dimensional subspace* $V_0 \subset T_{q_0}M \times \mathbb{R}^3$ and quadratic form $J'' : T_{q_0}M \times \mathbb{R}^3 \to \mathbb{R}$ such that $J''|_{V_0}$ is the second variation of the problem (2)-(1). *(details in the paper)*

Theorem

If $J''|_{V_0}$ *is coercive, then the flow* $\pi \circ \mathcal{H}_t$ *is invertible for every* $t \in [0, T]$

Remarks:

- the coercivity of the second variation implies the invertibility at the switching points (need of some non-smooth inversion function Theorem)
- checking the coercivity is rather simple, since the variation space is two-dimensional
- even if the second variation is constructed taking into account only the variations in the switching time, the optimality condition holds for all the admissible trajectories in a C^0-neighbourhood of the reference one.

Main Result

Theorem (F.C.C. - L. Poggiolini, 2018)

Let $\widehat{\xi}\colon [0, T] \to M$ *be an admissible trajectory for the control system* (2)-(1) *that satisfies the regularity assumptions and such that its second variation* $J''|_{V_0}$ *is coercive. Then the trajectory* $\widehat{\xi}$ *is a strict strong-local minimiser for the OCP.*

Example

Consider the optimal control problem

$$\min_{|u(\cdot)|\leqslant 1} \int_0^T |u(t)x_2(t)|\, dt$$

$$\begin{cases} \dot{x}_1 = x_2 & \\ \dot{x}_2 = u - \alpha x_2 & \alpha > 0 \\ x_1(0) = 0,\ x_2(0) = 0 & \\ x_1(T) = X > 0,\ x_2(T) = 0. & \end{cases}$$

Theorem (N. Boizot-O. Oukacha, 2016)

For every X, α *there exist times time* $0 < \hat{\tau}_1 \leqslant \hat{\tau}_2 < T_{\lim}$ *such that, if* $T_{\min} < T \leqslant T_{\lim}$[a], *then the bang-inactivated-bang trajectory with switching times* $\hat{\tau}_1, \hat{\tau}_2$ *is optimal.*

[a] $T_{\min}$ is the minimum time needed to reach X from 0

Example

Consider the optimal control problem

$$\min_{|u(\cdot)|\leqslant 1} \int_0^T |u(t)x_2(t)|\, dt$$

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = u - \alpha x_2 \qquad\qquad \alpha > 0 \\ x_1(0) = 0,\ x_2(0) = 0 \\ x_1(T) = X > 0,\ x_2(T) = 0. \end{cases}$$

Theorem (N. Boizot-O. Oukacha, 2016)

For every X, α *there exist times time* $0 < \hat{\tau}_1 \leqslant \hat{\tau}_2 < T_{\lim}$ *such that, if* $T_{\min} < T \leqslant T_{\lim}$ [a], *then the bang-inactivated-bang trajectory with switching times* $\hat{\tau}_1, \hat{\tau}_2$ *is optimal.*

[a] $T_{\min}$ is the minimum time needed to reach X from 0

- $\hat{\tau}_1 \neq \hat{\tau}_2$, then V_0 is the trivial linear space, and the second variation is coercive by definition.
- if $\hat{\tau}_1 = \hat{\tau}_2$, the reference trajectory satisfies the PMP in the abnormal form.