CONTINUOUS REPRESENTATION FOR A CLASS OF OPTIMAL HYBRID CONTROL PROBLEMS

Muriel Gapaillard

Laboratoire d'Ingénierie des Systèmes Automatisés, LISA
CNRS FRE 2656
62, avenue Notre-Dame du Lac, 49000, Angers, France
muriel.gapaillard@iut-nantes.univ-nantes.fr

Abstract: A class of optimal hybrid control problems involving model and state jump systems is considered. The calculation of optimal switching times amounts to minimizing a cost function. A continuous representation of these problems – based on convergence results – is introduced. This representation simplifies the study of the initial problem as it is shown, while applying variational formalism to determine the expression of the gradients of the cost function with respect to the switching times. An application example is presented and the implementation of a descent method confirms the validity of the continuous representation.

Keywords: continuous systems, optimal control, switching times, minimization, variational analysis, gradients.

1. INTRODUCTION

The study of optimal hybrid control problems must take into account the discontinuities occurring in the hybrid systems (see for instance Branicky, 1998 and Sussmann, 1999). On a practical point of view, these discontinuities raise difficulties in solving optimal control problems efficiently (Shaikh and Caines, 2003; Egerstedt *et al.*, 2003; Xu and Antsaklis, 2004). For example, while using the formalism of the calculus of variations (Bryson and Ho, 1968), discontinuities in the adjoint state of the hybrid system may occur (Cébron *et al.*, 1999).

The hybrid systems involved here are controlled – model and state – jump systems. Given an initial state, the problem consists in finding switching times that enable to best approximate a desired state at the final time. The calculation of the optimal switching times is set as a minimization problem of a cost function.

This paper introduces a representation of this class of optimal hybrid control problems by continuous optimal control problems. The interest of this representation – based on convergence theorems – lies in the possibility to solve hybrid problems avoiding difficulties generated by the discontinuities.

The class of studied optimal hybrid control problems is introduced in Section 2. The associated continuous problems are presented and convergence results are given (Section 3). Variational formalism is then applied to both problems, and expressions of the gradients with respect to the switching times are established (Section 4). Finally, an application example for an optimal hybrid control problem with two switching times and nonlinear subsystems is given. The optimal switching times are obtained by use of the gradients and of a descent method. This numerical application shows the practical validity of the continuous representation (Section 5).

2. HYBRID PROBLEM PRESENTATION

2.1 Studied Systems

The type of systems considered is the following

$$\begin{cases} \dot{\mathbf{x}}(t) &= \mathbf{f}_1\big(t,\mathbf{x}(t)\big) \quad \text{if} \quad t \in [t_0,\tau_1[\\ \mathbf{x}(t_0) &= \mathbf{x}_0 \\ \dot{\mathbf{x}}(t) &= \mathbf{f}_2\big(t,\mathbf{x}(t)\big) \quad \text{if} \quad t \in [\tau_1,\tau_2[\\ \mathbf{x}(\tau_1) &= \mathbf{x}(\tau_1^-) + \boldsymbol{\delta}_1 \\ \dot{\mathbf{x}}(t) &= \mathbf{f}_1\big(t,\mathbf{x}(t)\big) \quad \text{if} \quad t \in [\tau_2,t_\mathrm{f}] \\ \mathbf{x}(\tau_2) &= \mathbf{x}(\tau_2^-) + \boldsymbol{\delta}_2 \end{cases} \tag{1}$$

where $\mathbf{f}_1$ and $\mathbf{f}_2$ are functions defined over $[t_0,t_\mathrm{f}] \times \mathbb{R}^d$, $d \geq 1$, with values in $\mathbb{R}^d$, that are supposed to be Lipschitzian with respect to their second variable and continuous. Values t_0 and t_f are given in $\mathbb{R}^+$ and jumps $\boldsymbol{\delta}_1$, $\boldsymbol{\delta}_2$ and initial state $\mathbf{x}_0$ are given in $\mathbb{R}^d$.

System (1) admits a global unique solution defined over $[t_0,t_\mathrm{f}]$. This solution will be denoted $\mathbf{x}$.

Remark 1. It is possible to consider a finite number of switching times τ_i and of functions $\mathbf{f}_i$, except for the complexity of writing. Furthermore, the results presented here would be valid even if a control $\mathbf{u}$ is introduced in functions $\mathbf{f}_i$.

2.2 Cost Function

The problem amounts to finding switching times τ_1 and τ_2 that allow the final state $\mathbf{x}(t_\mathrm{f})$ to be the closest possible to a desired state $\mathbf{x}_\mathrm{d}$ fixed in $\mathbb{R}^d$. It is set as a minimization of the cost function J defined as follows

$$\forall(\tau_1,\tau_2) \in \mathcal{T} \quad J(\tau_1,\tau_2) = \frac{1}{2}\left\|\mathbf{x}(t_\mathrm{f}) - \mathbf{x}_\mathrm{d}\right\|^2 , \tag{2}$$

where $\|.\|$ is the Euclidean norm of $\mathbb{R}^d$ and

$$\mathcal{T} = \Big\{(\tau_1,\tau_2) \in [t_0,t_\mathrm{f}] \times [t_0,t_\mathrm{f}], \quad \tau_1 \leq \tau_2\Big\} .$$

The problem then amounts to finding $(\tau_1^\star,\tau_2^\star)$ belonging to $\mathcal{T}$ such that

$$\forall(\tau_1,\tau_2) \in \mathcal{T} \quad J(\tau_1^\star,\tau_2^\star) \leq J(\tau_1,\tau_2) .$$

3. CONTINUOUS REPRESENTATION AND CONVERGENCE RESULTS

3.1 Continuous Systems

Let us consider functions $\mathbf{f}_1$ and $\mathbf{f}_2$ from system (1). Since it is always possible to define the continuous function $\tilde{\mathbf{f}}$ over $[t_0,t_\mathrm{f}] \times \mathbb{R}^d \times [1,2]$ by

$$\begin{aligned} &\forall(t,\mathbf{x},\nu) \in [t_0,t_\mathrm{f}] \times \mathbb{R}^d \times [1,2] \\ &\quad \tilde{\mathbf{f}}(t,\mathbf{x},\nu) = (2-\nu)\mathbf{f}_1(t,\mathbf{x}) + (\nu-1)\mathbf{f}_2(t,\mathbf{x}) , \end{aligned}$$

assume that there exists an interval $[\nu_1,\nu_2]$ and a continuous function $\tilde{\mathbf{f}}$ defined over $[t_0,t_\mathrm{f}] \times \mathbb{R}^d \times [\nu_1,\nu_2]$ verifying

$$\tilde{\mathbf{f}}(t,\mathbf{x},\nu_1) = \mathbf{f}_1(t,\mathbf{x})$$

and

$$\tilde{\mathbf{f}}(t,\mathbf{x},\nu_2) = \mathbf{f}_2(t,\mathbf{x}) .$$

System (1) is therefore equivalent to

$$\begin{cases} \nu(t) &= \nu_1 \quad \text{if} \quad t \in [t_0,\tau_1[\\ \nu(t) &= \nu(\tau_1^-) + \nu_2 - \nu_1 \quad \text{if} \quad t \in [\tau_1,\tau_2[\\ \nu(t) &= \nu(\tau_2^-) + \nu_1 - \nu_2 \quad \text{if} \quad t \in [\tau_2,t_\mathrm{f}] \\ \dot{\mathbf{x}}(t) &= \tilde{\mathbf{f}}\big(t,\mathbf{x}(t),\nu(t)\big) \quad \text{if} \quad t \in [t_0,\tau_1[\\ \mathbf{x}(t_0) &= \mathbf{x}_0 \\ \dot{\mathbf{x}}(t) &= \tilde{\mathbf{f}}\big(t,\mathbf{x}(t),\nu(t)\big) \quad \text{if} \quad t \in [\tau_1,\tau_2[\\ \mathbf{x}(\tau_1) &= \mathbf{x}(\tau_1^-) + \boldsymbol{\delta}_1 \\ \dot{\mathbf{x}}(t) &= \tilde{\mathbf{f}}\big(t,\mathbf{x}(t),\nu(t)\big) \quad \text{if} \quad t \in [\tau_2,t_\mathrm{f}] \\ \mathbf{x}(\tau_2) &= \mathbf{x}(\tau_2^-) + \boldsymbol{\delta}_2 . \end{cases} \tag{3}$$

By setting $\mathbf{X} = (\mathbf{x},\nu)$ and $\mathbf{F} = (\tilde{\mathbf{f}},0)$, system (3) yields

$$\begin{cases} \dot{\mathbf{X}}(t) &= \mathbf{F}\big(t,\mathbf{X}(t)\big) \quad \text{if} \quad t \in [t_0,\tau_1[\\ \mathbf{X}(t_0) &= \mathbf{X}_0 \\ \dot{\mathbf{X}}(t) &= \mathbf{F}\big(t,\mathbf{X}(t)\big) \quad \text{if} \quad t \in [\tau_1,\tau_2[\\ \mathbf{X}(\tau_1) &= \mathbf{X}(\tau_1^-) + \boldsymbol{\mu}_1 \\ \dot{\mathbf{X}}(t) &= \mathbf{F}\big(t,\mathbf{X}(t)\big) \quad \text{if} \quad t \in [\tau_2,t_\mathrm{f}] \\ \mathbf{X}(\tau_2) &= \mathbf{X}(\tau_2^-) + \boldsymbol{\mu}_2 \end{cases}$$

where $\boldsymbol{\mu}_1 = (\boldsymbol{\delta}_1,\nu_2-\nu_1)$, $\boldsymbol{\mu}_2 = (\boldsymbol{\delta}_2,\nu_1-\nu_2)$ and $\mathbf{X}_0 = (\mathbf{x}_0,\nu_1)$. Associated continuous system can be defined by (see (Gapaillard, 2004) and Section 3.3)

$$\begin{cases} \dot{\mathbf{X}}(t) &= \mathbf{F}\big(t,\mathbf{X}(t)\big) + \dot{H}_n(t-\tau_1)\boldsymbol{\mu}_1 \\ &\qquad + \dot{H}_n(t-\tau_2)\boldsymbol{\mu}_2, \quad t \in [t_0,t_\mathrm{f}] \\ \mathbf{X}(t_0) &= \mathbf{X}_0 + H_n(t_0-\tau_1)\boldsymbol{\mu}_1 \\ &\qquad + H_n(t_0-\tau_2)\boldsymbol{\mu}_2 \end{cases} \tag{4}$$

where n is a positive integer and H_n is a regular approximation of the Heaviside function, defined by

$$\forall t \in \mathbb{R} \quad H_n(t) = \frac{1}{1+\exp(-nt)} .$$

The global unique solution of system (4), defined over $[t_0,t_\mathrm{f}]$, will be denoted $\mathbf{X}_n$.

3.2 Cost Function

The continuous problem – associated to the hybrid problem introduced in Section 2 – is set as a minimization of the cost function J_n defined by

$$\forall(\tau_1,\tau_2)\in\mathcal{T}$$
$$J_n(\tau_1,\tau_2)=\frac{1}{2}\,\|\mathbf{X}_n(t_\mathrm{f})-\mathbf{X}_\mathrm{d}\|^2 \quad (5)$$

where $\mathbf{X}_\mathrm{d}=(\mathbf{x}_\mathrm{d},\nu_1)$.
The problem amounts here to finding $(\tau_{1,n}^\star;\tau_{2,n}^\star)$ belonging to $\mathcal{T}$ such that

$$\forall(\tau_1,\tau_2)\in\mathcal{T}\quad J_n(\tau_{1,n}^\star;\tau_{2,n}^\star)\;\le\;J_n(\tau_1,\tau_2)\;.$$

3.3 Convergence Results

System (4) is considered as the continuous representation of system (1). This representation is based on the following convergence theorem.

Theorem 1. Let $\mathbf{x}_n$ be the first component, with values in $\mathbb{R}^d$, of $\mathbf{X}_n$.
1. When n tends to infinity, $\mathbf{x}_n(t)$ converges towards $\mathbf{x}(t)$ for all $t\in[t_0,t_\mathrm{f}]\setminus\{\tau_1,\tau_2\}$;
2. When n tends to infinity, the sequence $(\mathbf{x}_n)_{n>0}$ converges towards $\mathbf{x}$ in $L^p\big(]t_0,t_\mathrm{f}[\big)$, $p\ge 1$.

Proof. This theorem is a generalization of results given in (Gapaillard, 2004) and could be proved the same way.

The cost functions J and J_n defined by (2) and (5) are continuous over $\mathcal{T}$ (Gapaillard, 2003). Therefore, they reach their minimum values at $(\tau_1^\star,\tau_2^\star)$ and $(\tau_{1,n}^\star,\tau_{2,n}^\star)$ respectively.
Now, since the sequence $(\tau_{1,n}^\star,\tau_{2,n}^\star)_{n>1}$ is bounded, there exists a subsequence – which we do not relabel – that converges towards $(\bar\tau_1,\bar\tau_2)$ in $\mathcal{T}$.
In the following, we suppose that $\tau_2^\star<t_\mathrm{f}$ and $\bar\tau_2<t_\mathrm{f}$.

Proposition 1. The sequence $\big(J_n(\tau_{1,n}^\star,\tau_{2,n}^\star)\big)_{n>0}$ tends to $J(\bar\tau_1,\bar\tau_2)$.

Proof. The same mathematical tools as in (Gapaillard, 2004) could be used here to prove this result (see this reference for technical details).

Theorem 2. If $(\tau_1^\star,\tau_2^\star)$ is supposed to be unique, then the pairs $(\tau_1^\star,\tau_2^\star)$ and $(\bar\tau_1,\bar\tau_2)$ coincide.

Proof. Let $n>0$. Definition of $(\tau_{1,n}^\star,\tau_{2,n}^\star)$ allows to write

$$\forall(\tau_1,\tau_2)\in\mathcal{T}\quad J_n(\tau_{1,n}^\star,\tau_{2,n}^\star)\;\le\;J_n(\tau_1,\tau_2)\;. \quad (6)$$

Point 1 of theorem 1 leads to

$$\forall(\tau_1,\tau_2)\in\mathcal{T}\quad \tau_2\neq t_\mathrm{f}\Rightarrow\{J_n(\tau_1,\tau_2)\xrightarrow[n\to+\infty]{}J(\tau_1,\tau_2)\}\;.$$

According to proposition 1, equation (6) implies therefore

$$\forall(\tau_1,\tau_2)\in\mathcal{T},\ \tau_2\neq t_\mathrm{f},\quad J(\bar\tau_1,\bar\tau_2)\;\le\;J(\tau_1,\tau_2)\;,$$

and assumptions $\tau_2^\star<t_\mathrm{f}$ and $\bar\tau_2<t_\mathrm{f}$ allow to write

$$(\tau_1^\star,\tau_2^\star)=(\bar\tau_1,\bar\tau_2)\;,$$

since $(\tau_1^\star,\tau_2^\star)$ is supposed to be unique.
The claim is confirmed.

4. CALCULUS OF VARIATIONS

In this section, the formalism of the calculus of variations is applied to both hybrid and continuous problems. This enables to obtain the expressions of the gradients of J and J_n with respect to τ_1 and τ_2. These expressions allow to use a descent method to minimize the cost functions (see Section 5).

4.1 Applied to the Hybrid Problem

Consider the solution $\mathbf{x}$ of system (1). Terms $\mathbf{x}(\tau_i^-)$ and $\mathbf{x}(\tau_i)$, $i=1,2$, verify the following equations

$$\boldsymbol{\psi}_i\big(\mathbf{x}(\tau_i^-),\mathbf{x}(\tau_i)\big)=0, \quad (7)$$

where functions $\boldsymbol{\psi}_i$, $i=1,2$, are defined by

$$\forall(\mathbf{x}^-,\mathbf{x})\in\mathbb{R}^d\times\mathbb{R}^d\quad \boldsymbol{\psi}_i(\mathbf{x}^-,\mathbf{x})=\mathbf{x}-\mathbf{x}^--\boldsymbol{\delta}_i.$$

Consider the cost function J introduced in Section 2 and adjoin to J the state equations given in (1) with multiplier function $\boldsymbol{\lambda}$, and the constraints (7) with multipliers $\boldsymbol{\eta}_i$, $i=1,2$.
Then define a function $\tilde J$ by

$$\begin{aligned}&\forall(\tau_1,\tau_2)\in\mathcal{T}\quad \tilde J(\tau_1,\tau_2)=\tfrac{1}{2}\|\mathbf{x}(t_\mathrm{f})-\mathbf{x}_\mathrm{d}\|^2\\
&+\int_{t_0}^{\tau_1}\boldsymbol{\lambda}^\top(t)\left[\dot{\mathbf{x}}(t)-\mathbf{f}_1\big(t,\mathbf{x}(t)\big)\right]dt+\boldsymbol{\eta}_1^\top\boldsymbol{\psi}_1\big(\mathbf{x}(\tau_1^-),\mathbf{x}(\tau_1)\big)\\
&+\int_{\tau_1}^{\tau_2}\boldsymbol{\lambda}^\top(t)\left[\dot{\mathbf{x}}(t)-\mathbf{f}_2\big(t,\mathbf{x}(t)\big)\right]dt+\boldsymbol{\eta}_2^\top\boldsymbol{\psi}_2\big(\mathbf{x}(\tau_2^-),\mathbf{x}(\tau_2)\big)\\
&+\int_{\tau_2}^{t_\mathrm{f}}\boldsymbol{\lambda}^\top(t)\left[\dot{\mathbf{x}}(t)-\mathbf{f}_1\big(t,\mathbf{x}(t)\big)\right]dt\;.\end{aligned}$$

Introduce now Hamiltonians $\mathcal{H}_i$, $i=1,2$, defined as follows

$$\forall(t,\mathbf{x},\boldsymbol{\lambda})\in[t_0,t_\mathrm{f}]\times\mathbb{R}^d\times\mathbb{R}^d$$
$$\mathcal{H}_i(t,\mathbf{x},\boldsymbol{\lambda})=\boldsymbol{\lambda}^\top\mathbf{f}_i(t,\mathbf{x}).$$

In order to simplify the following equations, $\mathcal{H}_i(t)$ will be used instead of $\mathcal{H}_i\big(t,\mathbf{x}(t),\boldsymbol{\lambda}(t)\big)$, $i=1,2$.
It results

$$\begin{aligned}&\forall(\tau_1,\tau_2)\in\mathcal{T}\quad \tilde J(\tau_1,\tau_2)=\tfrac{1}{2}\|\mathbf{x}(t_\mathrm{f})-\mathbf{x}_\mathrm{d}\|^2\\
&+\int_{t_0}^{\tau_1}\left[\boldsymbol{\lambda}^\top(t)\dot{\mathbf{x}}(t)-\mathcal{H}_1(t)\right]dt+\boldsymbol{\eta}_1^\top\boldsymbol{\psi}_1\big(\mathbf{x}(\tau_1^-),\mathbf{x}(\tau_1)\big)\\
&+\int_{\tau_1}^{\tau_2}\left[\boldsymbol{\lambda}^\top(t)\dot{\mathbf{x}}(t)-\mathcal{H}_2(t)\right]dt+\boldsymbol{\eta}_2^\top\boldsymbol{\psi}_2\big(\mathbf{x}(\tau_2^-),\mathbf{x}(\tau_2)\big)\\
&+\int_{\tau_2}^{t_\mathrm{f}}\left[\boldsymbol{\lambda}^\top(t)\dot{\mathbf{x}}(t)-\mathcal{H}_1(t)\right]dt\;.\end{aligned}$$

The variation of $\tilde{J}$ is given by (see Bryson and Ho, 1968; Cébron, 2000)

$$\begin{aligned}
\delta\tilde{J} &= \big(\mathbf{x}(t_\mathrm{f})-\mathbf{x}_\mathrm{d}\big)^\top\delta\mathbf{x}(t_\mathrm{f})\\
&+\textstyle\int_{t_0}^{\tau_1}\left[\boldsymbol{\lambda}^\top(t)\delta\dot{\mathbf{x}}(t)-\frac{\partial\mathcal{H}_1}{\partial\mathbf{x}}(t)\delta\mathbf{x}(t)\right]\,dt-\mathcal{H}_1(\tau_1^-)\delta\tau_1\\
&\qquad+\boldsymbol{\eta}_1^\top\frac{\partial\boldsymbol{\psi}_1}{\partial\mathbf{x}^-}\big(\mathbf{x}(\tau_1^-),\mathbf{x}(\tau_1)\big)\delta\mathbf{x}(\tau_1^-)\\
&\qquad+\boldsymbol{\eta}_1^\top\frac{\partial\boldsymbol{\psi}_1}{\partial\mathbf{x}}\big(\mathbf{x}(\tau_1^-),\mathbf{x}(\tau_1)\big)\delta\mathbf{x}(\tau_1)\\
&+\textstyle\int_{\tau_1}^{\tau_2}\left[\boldsymbol{\lambda}^\top(t)\delta\dot{\mathbf{x}}(t)-\frac{\partial\mathcal{H}_2}{\partial\mathbf{x}}(t)\delta\mathbf{x}(t)\right]\,dt+\mathcal{H}_2(\tau_1)\delta\tau_1\\
&\quad-\mathcal{H}_2(\tau_2^-)\delta\tau_2+\boldsymbol{\eta}_2^\top\frac{\partial\boldsymbol{\psi}_2}{\partial\mathbf{x}^-}\big(\mathbf{x}(\tau_2^-),\mathbf{x}(\tau_2)\big)\delta\mathbf{x}(\tau_2^-)\\
&\qquad+\boldsymbol{\eta}_2^\top\frac{\partial\boldsymbol{\psi}_2}{\partial\mathbf{x}}\big(\mathbf{x}(\tau_2^-),\mathbf{x}(\tau_2)\big)\delta\mathbf{x}(\tau_2)\\
&+\textstyle\int_{\tau_2}^{t_\mathrm{f}}\left[\boldsymbol{\lambda}^\top(t)\delta\dot{\mathbf{x}}(t)-\frac{\partial\mathcal{H}_1}{\partial\mathbf{x}}(t)\delta\mathbf{x}(t)\right]\,dt+\mathcal{H}_1(\tau_2)\delta\tau_2\ .
\end{aligned}$$

By integrating terms $\boldsymbol{\lambda}^\top\delta\dot{\mathbf{x}}$ by parts, it follows

$$\begin{aligned}
\delta\tilde{J} &= \left[\big(\mathbf{x}(t_\mathrm{f})-\mathbf{x}_\mathrm{d}\big)^\top+\boldsymbol{\lambda}^\top(t_\mathrm{f})\right]\delta\mathbf{x}(t_\mathrm{f})\\
&+\left[\boldsymbol{\lambda}^\top(\tau_1^-)+\boldsymbol{\eta}_1^\top\frac{\partial\boldsymbol{\psi}_1}{\partial\mathbf{x}^-}\big(\mathbf{x}(\tau_1^-),\mathbf{x}(\tau_1)\big)\right]\delta\mathbf{x}(\tau_1^-)\\
&+\left[-\boldsymbol{\lambda}^\top(\tau_1)+\boldsymbol{\eta}_1^\top\frac{\partial\boldsymbol{\psi}_1}{\partial\mathbf{x}}\big(\mathbf{x}(\tau_1^-),\mathbf{x}(\tau_1)\big)\right]\delta\mathbf{x}(\tau_1)\\
&+\left[\boldsymbol{\lambda}^\top(\tau_2^-)+\boldsymbol{\eta}_2^\top\frac{\partial\boldsymbol{\psi}_2}{\partial\mathbf{x}^-}\big(\mathbf{x}(\tau_2^-),\mathbf{x}(\tau_2)\big)\right]\delta\mathbf{x}(\tau_2^-)\\
&+\left[-\boldsymbol{\lambda}^\top(\tau_2)+\boldsymbol{\eta}_2^\top\frac{\partial\boldsymbol{\psi}_2}{\partial\mathbf{x}}\big(\mathbf{x}(\tau_2^-),\mathbf{x}(\tau_2)\big)\right]\delta\mathbf{x}(\tau_2)\\
&\qquad-\textstyle\int_{t_0}^{\tau_1}\left[\dot{\boldsymbol{\lambda}}^\top(t)+\frac{\partial\mathcal{H}_1}{\partial\mathbf{x}}(t)\right]\delta\mathbf{x}(t)\,dt\\
&\qquad-\textstyle\int_{\tau_1}^{\tau_2}\left[\dot{\boldsymbol{\lambda}}^\top(t)+\frac{\partial\mathcal{H}_2}{\partial\mathbf{x}}(t)\right]\delta\mathbf{x}(t)\,dt\\
&\qquad-\textstyle\int_{\tau_2}^{t_\mathrm{f}}\left[\dot{\boldsymbol{\lambda}}^T(t)+\frac{\partial\mathcal{H}_1}{\partial\mathbf{x}}(t)\right]\delta\mathbf{x}(t)\,dt\\
&+\left[\mathcal{H}_2(\tau_1)-\mathcal{H}_1(\tau_1^-)\right]\delta\tau_1+\left[\mathcal{H}_1(\tau_2)-\mathcal{H}_2(\tau_2^-)\right]\delta\tau_2\ .
\end{aligned}$$

The multipliers $\boldsymbol{\lambda}$ and $\boldsymbol{\eta}_i$ are chosen to cause the coefficients of $\delta\mathbf{x}$ to vanish.
From

$$\frac{\partial\psi_i}{\partial\mathbf{x}^-}\big(\mathbf{x}(\tau_i^-),\mathbf{x}(\tau_i)\big)=-1,\ \ i=1,2,$$

and

$$\frac{\partial\psi_i}{\partial\mathbf{x}}\big(\mathbf{x}(\tau_i^-),\mathbf{x}(\tau_i)\big)=1,\ \ i=1,2,$$

it results

$$\boldsymbol{\eta}_i=\boldsymbol{\lambda}(\tau_i^-)=\boldsymbol{\lambda}(\tau_i),\ \ i=1,2.$$

The following adjoint system is therefore given by

$$\begin{cases}
\dot{\boldsymbol{\lambda}}(t) &= -\dfrac{\partial\mathcal{H}_1}{\partial\mathbf{x}}^\top\big(t,\mathbf{x}(t),\boldsymbol{\lambda}(t)\big)\quad\text{if } t\in[\tau_2,t_\mathrm{f}]\\
\boldsymbol{\lambda}(t_\mathrm{f}) &= \mathbf{x}_\mathrm{d}-\mathbf{x}(t_\mathrm{f})\\
\dot{\boldsymbol{\lambda}}(t) &= -\dfrac{\partial\mathcal{H}_2}{\partial\mathbf{x}}^\top\big(t,\mathbf{x}(t),\boldsymbol{\lambda}(t)\big)\quad\text{if } t\in[\tau_1,\tau_2[\\
\boldsymbol{\lambda}(\tau_2^-) &= \boldsymbol{\lambda}(\tau_2)\\
\dot{\boldsymbol{\lambda}}(t) &= -\dfrac{\partial\mathcal{H}_1}{\partial\mathbf{x}}^\top\big(t,\mathbf{x}(t),\boldsymbol{\lambda}(t)\big)\quad\text{if } t\in[t_0,\tau_1[\\
\boldsymbol{\lambda}(\tau_1^-) &= \boldsymbol{\lambda}(\tau_1),
\end{cases}\tag{8}$$

and the gradients of the cost function J with respect to the switching times τ_i, $i=1,2$, are

$$\begin{aligned}
\nabla J_{\tau_1} &= \mathcal{H}_2\big(\tau_1,\mathbf{x}(\tau_1),\boldsymbol{\lambda}(\tau_1)\big)-\mathcal{H}_1\big(\tau_1^-,\mathbf{x}(\tau_1^-),\boldsymbol{\lambda}(\tau_1^-)\big)\ ;\\
\nabla J_{\tau_2} &= \mathcal{H}_1\big(\tau_2,\mathbf{x}(\tau_2),\boldsymbol{\lambda}(\tau_2)\big)-\mathcal{H}_2\big(\tau_2^-,\mathbf{x}(\tau_2^-),\boldsymbol{\lambda}(\tau_2^-)\big)\ .
\end{aligned}$$

4.2 Applied to the Continuous Problem

Consider the cost function J_n defined by (5). Adjoin to J_n the state equation (4) with multiplier function $\boldsymbol{\lambda}$ and define a function $\tilde{J}_n$ by

$$\begin{aligned}
\forall(\tau_1,\tau_2)\in\mathcal{T}\quad \tilde{J}_n(\tau_1,\tau_2) &= \frac{1}{2}\left\|\mathbf{X}_n(t_\mathrm{f})-\mathbf{X}_\mathrm{d}\right\|^2\\
&+\int_{t_0}^{t_\mathrm{f}}\boldsymbol{\lambda}^\top(t)\dot{\mathbf{X}}_n(t)\,dt\\
-\int_{t_0}^{t_\mathrm{f}}\boldsymbol{\lambda}^\top(t)\Big[\mathbf{F}\big(t,\mathbf{X}_n(t)\big)&+\dot{H}_n(t-\tau_1)\boldsymbol{\mu}_1\\
&+\dot{H}_n(t-\tau_2)\boldsymbol{\mu}_2\Big]\,dt\ .
\end{aligned}$$

The variation of $\tilde{J}_n$ is given by

$$\begin{aligned}
\delta\tilde{J}_n &= \left[\mathbf{X}_n(t_\mathrm{f})-\mathbf{X}_\mathrm{d}\right]^\top\delta\mathbf{X}_n(t_\mathrm{f})\\
&+\int_{t_0}^{t_\mathrm{f}}\left[\boldsymbol{\lambda}^\top(t)\delta\dot{\mathbf{X}}_n(t)-\frac{\partial\mathcal{H}_n}{\partial\mathbf{X}}(t)\delta\mathbf{X}_n(t)\right]\,dt\\
&\quad-\int_{t_0}^{t_\mathrm{f}}\left[\frac{d\mathcal{H}_n}{d\tau_1}(t)\delta\tau_1+\frac{d\mathcal{H}_n}{d\tau_2}(t)\delta\tau_2\right]\,dt
\end{aligned}$$

where the Hamiltonian $\mathcal{H}_n$ is defined by

$$\begin{aligned}
&\forall(t,\mathbf{X},\boldsymbol{\lambda})\in[t_0,t_\mathrm{f}]\times\mathbb{R}^{d+1}\times\mathbb{R}^{d+1}\\
&\quad\mathcal{H}_n(t,\mathbf{X},\boldsymbol{\lambda})=\boldsymbol{\lambda}^\top\Big[\mathbf{F}(t,\mathbf{X})+\dot{H}_n(t-\tau_1)\boldsymbol{\mu}_1\\
&\qquad\qquad+\dot{H}_n(t-\tau_2)\boldsymbol{\mu}_2\Big]
\end{aligned}$$

and where the notation $\mathcal{H}_n(t)$ was used instead of $\mathcal{H}_n\big(t,\mathbf{X}_n(t),\boldsymbol{\lambda}(t)\big)$.
Integrating $\boldsymbol{\lambda}^\top\delta\dot{\mathbf{X}}_n$ by parts implies

$$\begin{aligned}
\delta\tilde{J}_n &= \left[\mathbf{X}_n(t_\mathrm{f})-\mathbf{X}_\mathrm{d}+\boldsymbol{\lambda}(t_\mathrm{f})\right]^\top\delta\mathbf{X}_n(t_\mathrm{f})\\
&\quad-\int_{t_0}^{t_\mathrm{f}}\left[\dot{\boldsymbol{\lambda}}^\top(t)+\frac{\partial\mathcal{H}_n}{\partial\mathbf{X}}(t)\right]\delta\mathbf{X}_n(t)\,dt\\
&\quad-\int_{t_0}^{t_\mathrm{f}}\left[\frac{d\mathcal{H}_n}{d\tau_1}(t)\delta\tau_1+\frac{d\mathcal{H}_n}{d\tau_2}(t)\delta\tau_2\right]\,dt\ .
\end{aligned}$$

The multiplier $\boldsymbol{\lambda}$ is chosen to cause the coefficients of $\delta\mathbf{X}_n$ to vanish. It leads to the following adjoint system

$$\begin{cases}
\dot{\boldsymbol{\lambda}}(t) &= -\dfrac{\partial\mathcal{H}_n}{\partial\mathbf{X}}^\top\big(t,\mathbf{X}_n(t),\boldsymbol{\lambda}(t)\big),\quad t\in[t_0,t_\mathrm{f}]\\
\boldsymbol{\lambda}(t_\mathrm{f}) &= \mathbf{X}_\mathrm{d}-\mathbf{X}_n(t_\mathrm{f})\ .
\end{cases}\tag{9}$$

The gradients of the cost function J_n with respect to the switching times τ_i, $i=1,2$, are defined by

$$\begin{aligned}
\nabla(J_n)_{\tau_1} &= -\int_{t_0}^{t_\mathrm{f}}\frac{d\mathcal{H}_n}{d\tau_1}\big(t,\mathbf{X}_n(t),\boldsymbol{\lambda}(t)\big)\,dt\ ;\\
\nabla(J_n)_{\tau_2} &= -\int_{t_0}^{t_\mathrm{f}}\frac{d\mathcal{H}_n}{d\tau_2}\big(t,\mathbf{X}_n(t),\boldsymbol{\lambda}(t)\big)\,dt\ .
\end{aligned}$$

This section showed, on a theoretical perspective, the simplification introduced by the continuous representation exposed in Section 2. The variational formalism applied to the hybrid problem indeed involves heavy calculation and leads to a switched adjoint system. On the other hand, the variation of $\tilde{J}_n$ is easier to establish and the adjoint system does not include model switchings.

5. APPLICATION EXAMPLE

It remains to be verified that continuous representation does not lead to numerical difficulties. The following example answers this question.

5.1 Optimal Hybrid Control Problem

In order to define a system of type (1), consider here the following differential equation which describes the movement of a pendulum with friction

$$\ddot{x} = -\sin x - \varepsilon \dot{x}, \quad \varepsilon > 0.$$

Replace this equation by a first order differential system and define functions $\mathbf{f}_i$ of system (1), $i = 1, 2$, by

$$\forall (t, \mathbf{x}) \in [t_0, t_\mathrm{f}] \times \mathbb{R}^2$$
$$\mathbf{f}_i(t, \mathbf{x}) = \begin{pmatrix} -\sin x_2 - \varepsilon_i\, x_1 \\ x_1 \end{pmatrix}, \quad i = 1, 2\ ,$$

where $\mathbf{x} = (x_1, x_2)$ and $\varepsilon_i > 0$, $i = 1, 2$.

Given a desired state $\mathbf{x}_\mathrm{d}$ in $\mathbb{R}^2$, the problem to be solved is the one described in Section 2.2. In order to minimize the cost function J, it will be useful to have expressions of the gradients of J with respect to the switching times.
Hamiltonians $\mathcal{H}_i$, $i = 1, 2$, are here defined as follows

$$\forall (t, \mathbf{x}, \boldsymbol{\lambda}) \in [t_0, t_\mathrm{f}] \times \mathbb{R}^2 \times \mathbb{R}^2$$
$$\mathcal{H}_i(t, \mathbf{x}, \boldsymbol{\lambda}) = -\lambda_1 \sin x_2 - \lambda_1 \varepsilon_i x_1 + \lambda_2 x_1, \quad i = 1, 2,$$

where $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ and $\mathbf{x} = (x_1, x_2)$.
Adjoint system (8) is therefore defined by

$$\begin{cases} \dot{\boldsymbol{\lambda}}(t) &= \begin{pmatrix} \lambda_1(t)\varepsilon_1 - \lambda_2(t) \\ \lambda_1(t)\cos x_2(t) \end{pmatrix} \quad \text{if} \quad t \in [\tau_2, t_\mathrm{f}] \\ \boldsymbol{\lambda}(t_\mathrm{f}) &= \mathbf{x}_\mathrm{d} - \mathbf{x}(t_\mathrm{f}) \\ \dot{\boldsymbol{\lambda}}(t) &= \begin{pmatrix} \lambda_1(t)\varepsilon_2 - \lambda_2(t) \\ \lambda_1(t)\cos x_2(t) \end{pmatrix} \quad \text{if} \quad t \in [\tau_1, \tau_2[\\ \boldsymbol{\lambda}(\tau_2^-) &= \boldsymbol{\lambda}(\tau_2) \\ \dot{\boldsymbol{\lambda}}(t) &= \begin{pmatrix} \lambda_1(t)\varepsilon_1 - \lambda_2(t) \\ \lambda_1(t)\cos x_2(t) \end{pmatrix} \quad \text{if} \quad t \in [t_0, \tau_1[\\ \boldsymbol{\lambda}(\tau_1^-) &= \boldsymbol{\lambda}(\tau_1) \end{cases}$$

where $\mathbf{x}(t) = (x_1(t), x_2(t))$ and $\boldsymbol{\lambda}(t) = (\lambda_1(t), \lambda_2(t))$.
The gradients of the cost function J – defined by (2) – with respect to the switching times τ_i, $i = 1, 2$, are

$$\nabla J_{\tau_1} = \boldsymbol{\lambda}^\top(\tau_1)\left[\mathbf{f}_2\big(\tau_1, \mathbf{x}(\tau_1^-) + \boldsymbol{\delta}_1\big) - \mathbf{f}_1\big(\tau_1, \mathbf{x}(\tau_1^-)\big)\right]\ ;$$
$$\nabla J_{\tau_2} = \boldsymbol{\lambda}^\top(\tau_2)\left[\mathbf{f}_1\big(\tau_2, \mathbf{x}(\tau_2^-) + \boldsymbol{\delta}_2\big) - \mathbf{f}_2\big(\tau_2, \mathbf{x}(\tau_2^-)\big)\right]\ .$$

5.2 Associated Continuous Problem

Suppose $\varepsilon_2 > \varepsilon_1$ and define function $\tilde{\mathbf{f}}$ from system (3) by

$$\forall (t, \mathbf{x}, \nu) \in [t_0, t_\mathrm{f}] \times \mathbb{R}^2 \times [\varepsilon_1, \varepsilon_2]$$
$$\tilde{\mathbf{f}}(t, \mathbf{x}, \nu) = \begin{pmatrix} -\sin x_2 - \nu\, x_1 \\ x_1 \end{pmatrix}\ ,$$

where $\mathbf{x} = (x_1, x_2)$.
The continuous system of type (4) considered here is defined by

$$\begin{cases} \dot{\mathbf{X}}(t) &= \begin{pmatrix} \tilde{\mathbf{f}}\big(t, \mathbf{X}(t)\big) \\ 0 \end{pmatrix} + \dot{H}_n(t - \tau_1)\begin{pmatrix} \boldsymbol{\delta}_1 \\ \varepsilon_2 - \varepsilon_1 \end{pmatrix} \\ & \quad + \dot{H}_n(t - \tau_2)\begin{pmatrix} \boldsymbol{\delta}_2 \\ \varepsilon_1 - \varepsilon_2 \end{pmatrix}, \quad t \in [t_0, t_\mathrm{f}] \\ \mathbf{X}(t_0) &= \begin{pmatrix} \mathbf{x}_0 \\ \varepsilon_1 \end{pmatrix} + H_n(t_0 - \tau_1)\begin{pmatrix} \boldsymbol{\delta}_1 \\ \varepsilon_2 - \varepsilon_1 \end{pmatrix} \\ & \quad + H_n(t_0 - \tau_2)\begin{pmatrix} \boldsymbol{\delta}_2 \\ \varepsilon_1 - \varepsilon_2 \end{pmatrix} \end{cases}$$

where $\mathbf{X} = (\mathbf{x}, \nu)$.
The problem consists in minimizing the cost function J_n, defined by (5), and – as done for the hybrid problem – the expressions of the gradients of J_n with respect to τ_1 and τ_2 will be given.
Definition of Hamiltonian $\mathcal{H}_n$ leads to

$$\forall (t, \mathbf{X}, \boldsymbol{\lambda}) \in [t_0, t_\mathrm{f}] \times \mathbb{R}^3 \times \mathbb{R}^3$$
$$\mathcal{H}_n(t, \mathbf{X}, \boldsymbol{\lambda}) = -\lambda_1 \sin x_2 - \lambda_1 x_3 x_1 + \lambda_2 x_1$$
$$+\dot{H}_n(t - \tau_1)\boldsymbol{\lambda}^\top \begin{pmatrix} \boldsymbol{\delta}_1 \\ \varepsilon_2 - \varepsilon_1 \end{pmatrix} + \dot{H}_n(t - \tau_2)\boldsymbol{\lambda}^\top \begin{pmatrix} \boldsymbol{\delta}_2 \\ \varepsilon_1 - \varepsilon_2 \end{pmatrix}$$

where $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$ and $\mathbf{X} = (x_1, x_2, x_3)$.
It allows to write the adjoint system given by (9) as follows

$$\begin{cases} \dot{\boldsymbol{\lambda}}(t) &= \begin{pmatrix} \lambda_1(t)x_3(t) - \lambda_2(t) \\ \lambda_1(t)\cos x_2(t) \\ \lambda_1(t)x_1(t) \end{pmatrix}\ , \quad t \in [t_0, t_\mathrm{f}] \\ \boldsymbol{\lambda}(t_\mathrm{f}) &= \begin{pmatrix} \mathbf{x}_\mathrm{d} \\ \varepsilon_1 \end{pmatrix} - \mathbf{X}_n(t_\mathrm{f})\ , \end{cases}$$

where $\mathbf{X}_n(t) = (x_1(t), x_2(t), x_3(t))$ and $\boldsymbol{\lambda}(t) = (\lambda_1(t), \lambda_2(t), \lambda_3(t))$.
The gradients of J_n with respect to τ_1 and τ_2 are

$$\nabla(J_n)_{\tau_1} = \int_{t_0}^{t_\mathrm{f}} \ddot{H}_n(t - \tau_1)\boldsymbol{\lambda}^\top(t)\begin{pmatrix} \boldsymbol{\delta}_1 \\ \varepsilon_2 - \varepsilon_1 \end{pmatrix} dt\ ;$$
$$\nabla(J_n)_{\tau_2} = \int_{t_0}^{t_\mathrm{f}} \ddot{H}_n(t - \tau_2)\boldsymbol{\lambda}^\top(t)\begin{pmatrix} \boldsymbol{\delta}_2 \\ \varepsilon_1 - \varepsilon_2 \end{pmatrix} dt\ .$$

5.3 Numerical Application

Let us consider both optimal hybrid and continuous problems described above. In order to find the values of optimal switching times $\tau_1^\star$ and $\tau_2^\star$ that achieve the minimum of J, both functions J and J_n are minimized by use of the function

Table 1. Values used for the application

Parameters	Values
ε_1	0.1
ε_2	0.4
t_0	0
t_f	20
$\mathbf{x}_0$	$(0, \frac{\pi}{4})$
$\mathbf{x}_d$	$(-0.0640, 0.0216)$
δ_1	$(-0.2, 0)$
δ_2	$(0.1, 0)$

Table 2. Minimization of J

	$\tau_1^\star$	$\tau_2^\star$	Number of Iterations
Without Gradients	10.3432	15.9307	14
With Gradients	7.9998	15.0001	7

Table 3. Minimization of J_{100}

	$\tau_{1;100}^\star$	$\tau_{2;100}^\star$	Number of Iterations
Without Gradients	7.0003	14.6001	14
With Gradients	7.9918	14.9976	7

`fminunc` of MATLAB. With the Broyden-Fletcher-Goldfarb-Shanno algorithm, selected here, function `fminunc` allows us to supply the expressions of the gradients of cost functions with respect to τ_1 and τ_2. The results given by function `fminunc` with and without the expressions of the gradients have been compared.

The values used for the application are given in Table 1. In order to test the convergence of the algorithm (with or without gradients), the desired state $\mathbf{x}_d$ is chosen to be the output of the system of type (1) described in Section 5.1, where τ_1 and τ_2 are set to 8 and 15 respectively. Minimizations are then initialized by taking $\tau_1 = 7$ and $\tau_2 = 13$. The values obtained with the minimization of J are given in Table 2.

To get good approximations of $\tau_1^\star$ and $\tau_2^\star$, function J_n with $n = 100$ is considered. The values obtained with the minimization of J_{100} are given in Table 3.

The minimizations of the cost functions, which do not take into account the expressions of the gradients, give values that correspond to local minimum of J or J_n. In order to obtain approximations of the optimal switching times, it is necessary here to supply the gradients to the function `fminunc`. The results given by the minimization of J_{100}, which are close to those obtained by minimizing J, confirms the interest of reducing the study of an optimal hybrid control problem to the study of its continuous representation.

6. CONCLUDING REMARKS

This paper presented a continuous representation for a class of optimal hybrid control problems. Such a problem can be studied via an associated continuous problem. This approach is justified by convergence results. The continuous representation enables to solve with less difficulties the initial hybrid problem. An illustration of this reduction was given here by applying the formalism of the calculus of variations in order to determine optimal switching times. A numerical implementation confirmed the interest of this simplification.

A continuous control can be considered and the expression of the gradient of a cost function with respect to this control could be obtained via the calculus of variations applied to the continuous representation. It is also conceivable to use other typical results for continuous optimal control problems (see Vinter, 2000).

REFERENCES

Branicky, M.S., V.S. Borkar and S.K. Mitter (1998). A unified framework for hybrid control: Model and optimal control theory. *IEEE Transactions on Automatic Control*, **43**(1).

Bryson, E.B. and Y.C. Ho (1968). *Applied Optimal Control.* Gin and Company, Waltham.

Cébron, B., M. Sechilariu and J. Burger (1999). Optimal control of hybrid dynamical systems with hysteresis. *Proceedings of ECC.*

Cébron, B. (2000). Commande de systèmes dynamiques hybrides. *Thesis,* Université d'Angers. (`http://www.istia.univ-angers.fr/LISA/THESES/theses.html`)

Egerstedt, M., Y. Wardi and F. Delmotte (2003). Optimal Control of Switching Times in Switched Dynamical Systems. *Proceedings of IEEE-CDC.*

Gapaillard, M. (2003). Optimal control for a class of hybrid systems via a stochastic method. *Proceedings of IEEE-ICCA.*

Gapaillard, M. (2004). Représentation de systèmes hybrides par des systèmes continus. *Technical report*, LISA. (`http://shiva.istia.univ-angers.fr/~baudin/`)

Shaikh, M.S. and P.E. Caines (2003). On the Optimal Control of Hybrid Systems: Optimization of Switching Times and Combinatoric Location Schedules. *Proceedings of IEEE-ACC.*

Sussmann, H.J. (1999). A maximum principle for hybrid optimal control problems. *Proceedings of IEEE-CDC*, 425-430.

Vinter, R. (2000). *Optimal Control*. Birkhäuser, Boston.

Xu, X. and P.J. Antsaklis (2004). Optimal Control of Switched Systems Based on Parameterization of the Switching Instants. *IEEE Transactions on Automatic Control*, **49**(1).