Control and Cybernetics

vol. **33** (2004) No. 4

Variational formalism applied to control of autonomous switching systems

by

M. El Bagdouri[1], B. Cébron[2], M. Sechilariu[2] and J. Burger[2]

[1] Laboratoire SeT, UTBM, 90010 Belfort, France
[2] LISA, Université d'Angers, 62 avenue Notre Dame du Lac
49000 Angers, France
e-mail: mohammed.elbagdouri@utbm.fr

Abstract: The formalism of the calculus of variations is applied to determine an optimal control of a class of Hybrid Dynamical Systems. This class consists of autonomous switching systems where jumps of the state are taken into account. It is shown that model switching involves discontinuities in the adjoint state of the system. The expression of the gradient of the cost function, with respect to the control, allows for the calculation of an optimal control by implementing a descent method. An illustrative linear quadratic example is given, which allows to conclude that the method can be easily implemented.

Keywords: control theory, switching systems, variational calculation.

1. Introduction

Hybrid Dynamical Systems (HDS) can be roughly defined as continuous systems with several modes of operation in which an event causes the mode to change. The way the events occur is described by means of a more or less complex Discrete Event System (DES) (Antsaklis and Nerode, 1998). The events that cause switching can be of two types (Branicky, 1998): first, events that are triggered by the continuous part of the system, thus inducing *autonomous* switching (Cébron et al., 1999a); second, those which are triggered by the discrete part of the system and thus induce *controlled* switching (Cébron et al., 1999b). Therefore, HDS can be considered as continuous systems interacting with a DES (Van Der Schaft and Schumacher, 2000). The latter can be modelled by means of either an automaton (Nerode, 1993; Brockett, 1993) or a Petri network (Andreu et al., 1996; Daubas et al., 1994).

In this article, we study the optimal control of a specific class of HDS called autonomous switching systems. One of the main problems raised by the calculation of an optimal control of an autonomous switching system is to know how discontinuities occurring in the continuous system should be taken into account. Generally, these discontinuities come in the form of model changes associated with jumps of the state vector (Branicky, 1995). The calculus of the optimal control of switching systems has already been studied, by Bryson and Ho (1975), for instance, who used variational formalism, and more recently, within a more general theoretical framework, by Vinter (1993), by Sussmann (1999), and by Riedinger et al. (1999), who state the conditions of application of the Maximum Principle (MP) to wide classes of systems. Application of MP to optimal control calculations can be performed by means of Hamilton-Jacobi-Bellman equations and linear programming techniques (see, e.g., Hedlund and Rantzer, 2002).

Our purpose in this article is to calculate the optimal control of an autonomous switching system using the classical tools of the calculus of variations, and to show how it is practically possible to compute this control by means of descent methods which involve the expression of the gradient of the cost function. It is to be noted that although the calculation can be considered as a classical one, Bryson and Ho (1975) explained (p. 101) that "Finding solutions to such problems is, in general, quite involved. The method of steepest descent may be used to solve such problem numerically", (see, e.g., Xu and Antsaklis, 2003 and 2004). Since derivability is needed, we present the calculation with the machinery of the variational formalism.

The article is organised as follows: in the section entitled Background Preliminaries, we refer to notations and definitions; then we study the minimisation of a cost function in which both model switching and jumps of the state are taken into account. The calculations are first performed in the general case where the form of the state equation is not particularised. They lead to the expression of both the adjoint system and the gradient of the cost function, which is to be minimised. The characteristic of HDS is noticeable in the expression of the adjoint state that displays jumps when switching occurs. The results are an extension of Bryson and Ho's (1975) which are applied to HDS. Next, we consider an application to a linear quadratic control where the cost function corresponding to a pursuit problem and the expression of its gradient allow a descent method to be implemented. The algorithm of the optimal control calculation is described in detail, which leads, at each step of the descent algorithm, to the resolution of a linear system. In the last section, a numerical example is given, which allows to conclude that the method can be easily implemented.

2. Background preliminaries

In this section, we refer to variational formalism (Ciarlet, 1990) and express the general variation of a function which will be used later to calculate an optimal control of HDS. At the initial instant t_0 and final instant t_{f}, we consider the set of functions E defined by:

$$E = \big\{ \mathbf{q} : [t_0, t_\mathrm{f}] \to \mathbb{R}^n \text{ of class } \mathcal{C}^1 \mid \mathbf{q}(t_0) = \mathbf{q}_0 \text{ and } \mathbf{q}(t_\mathrm{f}) = \mathbf{q}_\mathrm{f} \big\}.$$

Let $L(\mathbf{q}(t), \dot{\mathbf{q}}(t), t)$ and $K(\mathbf{q}_0, \mathbf{q}_\mathrm{f}, t_0, t_\mathrm{f})$ be two functions of class $\mathcal{C}^1$ defined into $\mathbb{R}$, and J the functional defined from E into $\mathbb{R}$ by:

$$J(\mathbf{q}) = \int_{t_0}^{t_\mathrm{f}} L\big(\mathbf{q}(t), \dot{\mathbf{q}}(t), t\big)\ dt + K\left(\mathbf{q}_0, \mathbf{q}_\mathrm{f}, t_0, t_\mathrm{f}\right). \tag{1}$$

The principle of the calculus of variations consists in writing the variations of functional J with respect to the perturbations of a nominal trajectory and in characterising the optimal trajectory as being that whose perturbations induce no variation of J. We denote $L(t) = L\big(\mathbf{q}(t), \dot{\mathbf{q}}(t), t\big)$ and $K = K(\mathbf{q}_0, \mathbf{q}_\mathrm{f}, t_0, t_\mathrm{f})$ and consider the variation of the functional J in the general case where the boundaries $t_0, t_\mathrm{f}, \mathbf{q}_0$ and $\mathbf{q}_\mathrm{f}$ are free.

It can be shown (see, e.g., Bérest, 1997) that the variation of the functional defined by (1), in any direction $\mathbf{h}$ of class $\mathcal{C}^1$, can be written as follows:

$$\begin{aligned}
\delta J = &\int_{t_0}^{t_f} \left(\frac{\partial L}{\partial \mathbf{q}}(t) - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}(t)\right)\right)^\top \cdot\ \mathbf{h}(t)\ dt \\
&+ \left(-L(t_0) + \dot{\mathbf{q}}^\top(t_0)\,\frac{\partial L}{\partial \dot{\mathbf{q}}}(t_0) + \frac{\partial K}{\partial t_0}\right)\delta t_0 + \left(-\frac{\partial L}{\partial \dot{\mathbf{q}}}(t_0) + \frac{\partial K}{\partial \mathbf{q}_0}\right)^\top \cdot\ \delta \mathbf{q}_0 \\
&+ \left(L(t_\mathrm{f}) - \dot{\mathbf{q}}^\top(t_\mathrm{f})\,\frac{\partial L}{\partial \dot{\mathbf{q}}}(t_\mathrm{f}) + \frac{\partial K}{\partial t_\mathrm{f}}\right)\delta t_\mathrm{f} + \left(\frac{\partial L}{\partial \dot{\mathbf{q}}}(t_\mathrm{f}) + \frac{\partial K}{\partial \mathbf{q}_\mathrm{f}}\right)^\top \cdot\ \delta \mathbf{q}_\mathrm{f}.
\end{aligned} \tag{2}$$

When performing the calculus of δJ, expressed by equation (2), we are led to carry out an integration by parts where L is required to be of class $\mathcal{C}^2$. We can remove this assumption with the aid of the Du Bois-Raymond Lemma (Bérest, 1997).

If we set:

$$\begin{aligned}
H(t) &= -L(t) + \dot{\mathbf{q}}^\top(t)\ \cdot\ \frac{\partial L}{\partial \dot{\mathbf{q}}}(t) \\
\mathbf{p}(t) &= \frac{\partial L}{\partial \dot{\mathbf{q}}}(t),
\end{aligned}$$

which are called Hamilton's function (or Hamiltonian) and conjugate moment respectively, the variation of J can be written as follows:

$$\begin{aligned}
\delta J = &\int_{t_0}^{t_f} \left(\frac{\partial L}{\partial \mathbf{q}}(t) - \dot{\mathbf{p}}(t)\right)^\top \cdot\ \mathbf{h}(t)\ dt + \left(H(t_0) + \frac{\partial K}{\partial t_0}\right)\ \delta t_0 \\
&+ \left(-\mathbf{p}(t_0) + \frac{\partial K}{\partial \mathbf{q}_0}\right)^\top \cdot \delta \mathbf{q}_0 + \left(-H(t_\mathrm{f}) + \frac{\partial K}{\partial t_\mathrm{f}}\right)\delta t_\mathrm{f} + \left(\mathbf{p}(t_\mathrm{f}) + \frac{\partial K}{\partial \mathbf{q}_\mathrm{f}}\right)^\top \cdot \delta \mathbf{q}_\mathrm{f}.
\end{aligned} \tag{3}$$

3. Optimal control of autonomous switching systems

We study a class of HDS which can be considered as continuous processes occurring within a time interval $[t_0, t_\mathrm{f}]$, in which discrete events induce some changes in the state and structure of the model. We begin with the formulation of the control problem in the general case of an autonomous switching system, and then study the case of a system with a single switching of a model on a curve with a jump of the state.

3.1. Formulation of the control problem

Let the state $\mathbf{x}(t)$ belong to $\mathbb{R}^n$. We assume that the state and structure of HDS change within the time interval $[t_0, t_\mathrm{f}]$, at switching instants $\tau_1, \tau_2, \ldots, \tau_m$, which are supposed to be in finite numbers and verify:

$$\tau_0 < \tau_1 < \tau_2 < \cdots < \tau_m < \tau_{m+1},$$

where we set $\tau_0 = t_0$ and $\tau_{m+1} = t_\mathrm{f}$.

We write $\xi_i^- = \mathbf{x}(\tau_i^-)$, $\xi_i^+ = \mathbf{x}(\tau_i^+)$, $i \in \{1, \ldots, m\}$ and assume that this system is governed, within each time interval (τ_{i-1}, τ_i), by a state equation of the form:

$$\begin{aligned} &\dot{\mathbf{x}}(t) = \mathbf{f}_i\big(\mathbf{x}(t), \mathbf{u}(t), t\big) \quad \text{for} \quad t \in (\tau_{i-1}, \tau_i), \quad i \in \{1, \ldots, m+1\} \\ &\mathbf{x}(t_0) = \mathbf{x}_0, \end{aligned} \tag{4}$$

where $\mathbf{f}_i$ is a function of class $\mathcal{C}^1$, defined into $\mathbb{R}^n$, over the interval $(\tau_{i-1}, \tau_i]$.

The optimal control problem consists in finding a continuous control function $\mathbf{u}^\star$, defined over the whole time interval $[t_0, t_\mathrm{f}]$, which minimises the following cost function:

$$J(\mathbf{u}) = \sum_{i=1}^{m+1} \int_{\tau_{i-1}}^{\tau_i} F_i\big(\mathbf{x}(t), \mathbf{u}(t), t\big)\ dt + \sum_{i=1}^{m} G_i\left(\boldsymbol{\xi}_i^-, \boldsymbol{\xi}_i^+, \tau_i\right) + K(\mathbf{x}_\mathrm{f}, t_\mathrm{f}), \tag{5}$$

where F_i is a function of class $\mathcal{C}^1$, defined over the interval (τ_{i-1}, τ_i). Each function G_i, of class $\mathcal{C}^1$, is called the switching cost; the function K, of class $\mathcal{C}^1$, is called the final cost.

3.2. Autonomous switching with a jump of the state

We consider the case where there is only one switching instant τ, and suppose that the state $\mathbf{x}$ of the system jumps from $\boldsymbol{\xi}^- = \mathbf{x}(\tau^-)$ to $\boldsymbol{\xi}^+ = \mathbf{x}(\tau^+)$ and that the derivative $\dot{\mathbf{x}}$ is discontinuous over the time interval $[t_0, t_\mathrm{f}]$, at the instant τ. State equations (4) come down to:

$$\begin{aligned} &\dot{\mathbf{x}}(t) = \mathbf{f}_1\big(\mathbf{x}(t), \mathbf{u}(t), t\big) \quad \text{for} \quad t \in [t_0, \tau) \\ &\dot{\mathbf{x}}(t) = \mathbf{f}_2\big(\mathbf{x}(t), \mathbf{u}(t), t\big) \quad \text{for} \quad t \in (\tau, t_\mathrm{f}] \\ &\mathbf{x}(t_0) = \mathbf{x}_0. \end{aligned} \tag{6}$$

Hence, the cost function (5) is written in the form:

$$\begin{aligned} J(\mathbf{u}) &= \int_{t_0}^{\tau} F_1\big(\mathbf{x}(t),\mathbf{u}(t),t\big)\ dt + G(\boldsymbol{\xi}^-,\boldsymbol{\xi}^+,\tau) \\ &+ \int_{\tau}^{t_f} F_2\big(\mathbf{x}(t),\mathbf{u}(t),t\big)\ dt + K\big(\mathbf{x}_\mathrm{f},t_\mathrm{f}\big). \end{aligned} \tag{7}$$

We make the following assumptions about the boundaries of the problem:

i) t_0, t_f and $\mathbf{x}_0$ are fixed;

ii) $\mathbf{x}_\mathrm{f}$ is free.

We suppose that the model switching points $\boldsymbol{\xi}^-$ and $\boldsymbol{\xi}^+$ of the system are located on curves whose equations are:

$$\boldsymbol{\phi}(\boldsymbol{\xi}^-,\tau) = 0 \tag{8}$$

$$\boldsymbol{\psi}(\boldsymbol{\xi}^+,\tau) = 0. \tag{9}$$

We apply the results of the calculus of variations to determine the control $\mathbf{u}^\star$ that minimises the cost function (7) with constraints (6), (8) and (9). To do so, we set $\mathbf{q} = [\mathbf{x}\,,\,\mathbf{u}]^\top$ and notice that the functions F_1 and F_2 do not depend on $\dot{\mathbf{q}}$. We then transform the problem by introducing multipliers which consist of an adjoint state $\boldsymbol{\lambda}$, associated with the state equation (6), and multipliers $\boldsymbol{\mu}$ and $\boldsymbol{\eta}$, associated with the switching curves (8) and (9), respectively. We then define a functional $\tilde{J}$ by:

$$\begin{aligned} \tilde{J}(\mathbf{u}) = &\int_{t_0}^{\tau} \Big[F_1\big(\mathbf{x}(t),\mathbf{u}(t),t\big) + \boldsymbol{\lambda}^\top(t)\cdot\Big(\dot{\mathbf{x}}(t) - \mathbf{f}_1(\mathbf{x}(t),\mathbf{u}(t),t)\Big)\Big]\ dt \\ &+ G(\boldsymbol{\xi}^-,\boldsymbol{\xi}^+,\tau) + \boldsymbol{\mu}^\top\cdot\boldsymbol{\phi}(\boldsymbol{\xi}^-,\tau) + \boldsymbol{\eta}^\top\cdot\boldsymbol{\psi}(\boldsymbol{\xi}^+,\tau) \\ &+ \int_{\tau}^{t_f} \Big[F_2\big(\mathbf{x}(t),\mathbf{u}(t),t\big) + \boldsymbol{\lambda}^\top(t)\cdot\Big(\dot{\mathbf{x}}(t) - \mathbf{f}_2(\mathbf{x}(t),\mathbf{u}(t),t)\Big)\Big]\, dt + K(\mathbf{x}_\mathrm{f},t_\mathrm{f}). \end{aligned}$$

If we set:

$$L_i\big(\mathbf{x}(t),\mathbf{u}(t),\dot{\mathbf{x}}(t),t\big) = F_i\big(\mathbf{x}(t),\mathbf{u}(t),t\big) + \boldsymbol{\lambda}^\top(t)\cdot\big(\dot{\mathbf{x}}(t) - \mathbf{f}_i(\mathbf{x}(t),\mathbf{u}(t),t)\big) \qquad i\in\{1\,,2\},$$

we get:

$$\begin{aligned} \tilde{J}(\mathbf{u}) = &\int_{t_0}^{\tau} L_1\big(\mathbf{x}(t),\mathbf{u}(t),\dot{\mathbf{x}}(t),t\big)\ dt\ +\ G(\boldsymbol{\xi}^-,\boldsymbol{\xi}^+,\tau) \\ &+\ \boldsymbol{\mu}^\top\cdot\boldsymbol{\phi}(\boldsymbol{\xi}^-,\tau)\ +\ \boldsymbol{\eta}^\top\cdot\boldsymbol{\psi}(\boldsymbol{\xi}^+,\tau) \\ &+\ \int_{\tau}^{t_\mathrm{f}} L_2\big(\mathbf{x}(t),\mathbf{u}(t),\dot{\mathbf{x}}(t),t\big)\ dt\ +\ K(\mathbf{x}_\mathrm{f},t_\mathrm{f}). \end{aligned}$$

It is to be pointed out that $\tilde{J}(\mathbf{u})$ is written as a sum of two functionals, the first one being associated with a final cost and the second one with both an initial cost and a final cost. Therefore, we can calculate the general variation of each of these functionals by using expression (3). In this case, the Hamiltonian and the conjugate moment are as follows:

$$\begin{aligned} H_i &= -L_i + \dot{\mathbf{x}}^\top \cdot \frac{\partial L_i}{\partial \dot{\mathbf{x}}} + \dot{\mathbf{u}}^\top \cdot \frac{\partial L_i}{\partial \dot{\mathbf{u}}} = -F_i + \boldsymbol{\lambda}^\top \cdot \mathbf{f}_i \qquad i \in \{1\,,2\} \\ \mathbf{p}(t) &= \left(\frac{\partial L_i}{\partial \dot{\mathbf{x}}}\,,\ \frac{\partial L_i}{\partial \dot{\mathbf{u}}}\right)^\top = \big(\boldsymbol{\lambda}(t)\,,\,0\big)^\top . \end{aligned}$$

By taking a direction of variation $\mathbf{h}(t) = \big(\boldsymbol{\varphi}(t)\,,\,\boldsymbol{\nu}(t)\big)^\top$ where $\boldsymbol{\varphi}$ belongs to the space of states and $\boldsymbol{\nu}$ to the space of controls, and by applying equation (3), we get the expression of the general variation of $\tilde{J}(\mathbf{u})$ which is given by:

$$\begin{aligned} \delta\tilde{J} = &\int_{t_0}^{\tau} \left[\left(\frac{\partial L_1}{\partial \mathbf{x}}(t) - \dot{\boldsymbol{\lambda}}(t)\right)^\top \cdot \boldsymbol{\varphi}(t) + \left(\frac{\partial L_1}{\partial \mathbf{u}}(t)\right)^\top \cdot \boldsymbol{\nu}(t)\right] dt \\ &+ \left(-H_1(\tau^-) + \frac{\partial G}{\partial \tau} + \boldsymbol{\mu}^\top \cdot \frac{\partial \boldsymbol{\phi}}{\partial \tau}(\boldsymbol{\xi}^-,\tau) + \boldsymbol{\eta}^\top \cdot \frac{\partial \boldsymbol{\psi}}{\partial \tau}(\boldsymbol{\xi}^+,\tau)\right) \delta\tau \\ &+ \left(\boldsymbol{\lambda}(\tau^-) + \frac{\partial G}{\partial \boldsymbol{\xi}^-} + \left[\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{\xi}^-}(\boldsymbol{\xi}^-,\tau)\right]^\top \cdot \boldsymbol{\mu}\right)^\top \cdot \delta\boldsymbol{\xi}^- \\ &+ \int_{\tau}^{t_\mathrm{f}} \left[\left(\frac{\partial L_2}{\partial \mathbf{x}}(t) - \dot{\boldsymbol{\lambda}}(t)\right)^\top \cdot \boldsymbol{\varphi}(t) + \left(\frac{\partial L_2}{\partial \mathbf{u}}(t)\right)^\top \cdot \boldsymbol{\nu}(t)\right] dt \\ &+ H_2(\tau^+)\,\delta\tau + \left(-\boldsymbol{\lambda}(\tau^+) + \frac{\partial G}{\partial \boldsymbol{\xi}^+} + \left[\frac{\partial \boldsymbol{\psi}}{\partial \boldsymbol{\xi}^+}(\boldsymbol{\xi}^+,\tau)\right]^\top \cdot \boldsymbol{\eta}\right)^\top \cdot \delta\boldsymbol{\xi}^+ \\ &+ \left(\boldsymbol{\lambda}(t_\mathrm{f}) + \frac{\partial K}{\partial \mathbf{x}_\mathrm{f}}(\mathbf{x}_\mathrm{f}, t_\mathrm{f})\right)^\top \cdot \delta\mathbf{x}_\mathrm{f}. \end{aligned}$$

Since we have

$$\begin{aligned} \frac{\partial L_i}{\partial \mathbf{x}} &= \frac{\partial F_i}{\partial \mathbf{x}} - \left[\frac{\partial \mathbf{f}_i}{\partial \mathbf{x}}\right]^\top \cdot \boldsymbol{\lambda} = -\frac{\partial H_i}{\partial \mathbf{x}} \qquad i \in \{1,2\} \\ \frac{\partial L_i}{\partial \mathbf{u}} &= \frac{\partial F_i}{\partial \mathbf{u}} - \left[\frac{\partial \mathbf{f}_i}{\partial \mathbf{u}}\right]^\top \cdot \boldsymbol{\lambda} = -\frac{\partial H_i}{\partial \mathbf{u}} \qquad i \in \{1,2\}, \end{aligned}$$

the general variation $\delta\tilde{J}$ is written in the form:

$$
\begin{aligned}
\delta\tilde{J} =& \int_{t_0}^{\tau} \left[\left(-\frac{\partial H_1}{\partial \mathbf{x}}(t) - \dot{\boldsymbol{\lambda}}(t) \right)^\top \cdot \boldsymbol{\varphi}(t) + \left(-\frac{\partial H_1}{\partial \mathbf{u}}(t) \right)^\top \cdot \boldsymbol{\nu}(t) \right] dt \\
&+ \int_{\tau}^{t_f} \left[\left(-\frac{\partial H_2}{\partial \mathbf{x}}(t) - \dot{\boldsymbol{\lambda}}(t) \right)^\top \cdot \boldsymbol{\varphi}(t) + \left(-\frac{\partial H_2}{\partial \mathbf{u}}(t) \right)^\top \cdot \boldsymbol{\nu}(t) \right] dt \\
&+ \left(-H_1(\tau^-) + H_2(\tau^+) + \frac{\partial G}{\partial \tau} + \boldsymbol{\mu}^\top \cdot \frac{\partial \boldsymbol{\phi}}{\partial \tau}(\boldsymbol{\xi}^-, \tau) + \boldsymbol{\eta}^\top \cdot \frac{\partial \boldsymbol{\psi}}{\partial \tau}(\boldsymbol{\xi}^+, \tau) \right) \delta\tau \\
&+ \left(\boldsymbol{\lambda}(\tau^-) + \frac{\partial G}{\partial \boldsymbol{\xi}^-} + \left[\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{\xi}^-}(\boldsymbol{\xi}^-, \tau) \right]^\top \cdot \boldsymbol{\mu} \right)^\top \cdot \delta\boldsymbol{\xi}^- \\
&+ \left(-\boldsymbol{\lambda}(\tau^+) + \frac{\partial G}{\partial \boldsymbol{\xi}^+} + \left[\frac{\partial \boldsymbol{\psi}}{\partial \boldsymbol{\xi}^+}(\boldsymbol{\xi}^+, \tau) \right]^\top \cdot \boldsymbol{\eta} \right)^\top \cdot \delta\boldsymbol{\xi}^+ \\
&+ \left(\boldsymbol{\lambda}(t_\mathrm{f}) + \frac{\partial K}{\partial \mathbf{x}_\mathrm{f}} \right)^\top \cdot \delta\mathbf{x}_\mathrm{f}.
\end{aligned}
$$

By writing the stationarity of $\tilde{J}$ with respect to the state $\mathbf{x}$, and by taking into account the conditions of transversality given by the integrated part of $\delta\tilde{J}$, we get the necessary conditions for optimality:

i) Equations which allow to calculate the adjoint state $\boldsymbol{\lambda}$ and the multipliers $\boldsymbol{\mu}$ and $\boldsymbol{\eta}$:

$$
\begin{aligned}
&-\frac{\partial H_1}{\partial \mathbf{x}}(t) - \dot{\boldsymbol{\lambda}}(t) = 0 \ \text{ for } \ t \in [t_0, \tau) \\
&-\frac{\partial H_2}{\partial \mathbf{x}}(t) - \dot{\boldsymbol{\lambda}}(t) = 0 \ \text{ for } \ t \in (\tau, t_\mathrm{f}] \\
&\boldsymbol{\lambda}(\tau^-) + \frac{\partial G}{\partial \boldsymbol{\xi}^-} + \left[\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{\xi}^-}(\boldsymbol{\xi}^-, \tau) \right]^\top \cdot \boldsymbol{\mu} = 0 \\
&-\boldsymbol{\lambda}(\tau^+) + \frac{\partial G}{\partial \boldsymbol{\xi}^+} + \left[\frac{\partial \boldsymbol{\psi}}{\partial \boldsymbol{\xi}^+}(\boldsymbol{\xi}^+, \tau) \right]^\top \cdot \boldsymbol{\eta} = 0 \\
&-H_1(\tau^-) + H_2(\tau^+) + \frac{\partial G}{\partial \tau} + \boldsymbol{\mu}^\top \cdot \frac{\partial \boldsymbol{\phi}}{\partial \tau}(\boldsymbol{\xi}^-, \tau) + \boldsymbol{\eta}^\top \cdot \frac{\partial \boldsymbol{\psi}}{\partial \tau}(\boldsymbol{\xi}^+, \tau) = 0 \\
&\boldsymbol{\lambda}(t_\mathrm{f}) + \frac{\partial K}{\partial \mathbf{x}_\mathrm{f}} = 0
\end{aligned}
$$

ii) Gradient of the cost function with respect to the control $\mathbf{u}$:

$$
\begin{aligned}
\nabla J_\mathbf{u}(t) &= -\frac{\partial H_1}{\partial \mathbf{u}}(t) \qquad \text{for } t \in [t_0, \tau) \\
\nabla J_\mathbf{u}(t) &= -\frac{\partial H_2}{\partial \mathbf{u}}(t) \qquad \text{for } t \in (\tau, t_\mathrm{f}].
\end{aligned}
$$

Comments: For convenience, we considered only one switching instant in this section. The extension to several switching instants does not raise any new theoretical problem. Since the switching instants over the time interval $[t_0, t_\mathrm{f}]$ are finite in number, the cost function (7) would involve terms which could be written as

$$\sum_{i=0}^{m} \int_{\tau_i}^{\tau_{i+1}} \ldots$$

and, writing the stationarity of $\tilde{J}$ would give the differential equations of the adjoint state on each time interval as well as the jumps of the adjoint state at each switching instant τ_i. As can be seen, the problem is only a technical one.

4. The trajectory pursuit problem

We apply the general results presented above to a trajectory pursuit problem with autonomous switching of the model and discontinuities in the state when switching occurs. We consider the following piecewise linear system in which the state $\mathbf{x}(t)$ belongs to $\mathbb{R}^n$, and the control input $\mathbf{u}(t)$ belongs to $\mathbb{R}^m$:

$$\begin{aligned} &\dot{\mathbf{x}}(t) = A_1\, \mathbf{x}(t) + B\, \mathbf{u}(t) \quad \text{for } t \in [t_0, \tau) \\ &\dot{\mathbf{x}}(t) = A_2\, \mathbf{x}(t) + B\, \mathbf{u}(t) \quad \text{for } t \in (\tau, t_\mathrm{f}] \\ &\mathbf{x}(t_0) = \mathbf{x}_0. \end{aligned} \tag{10}$$

Matrices A_1, A_2, and B have adequate dimensions and τ is the model switching instant. We suppose that switching points ξ^- and ξ^+ are located on straight lines whose equations are:

$$\phi(\boldsymbol{\xi}^-, \tau) = \prod_{j=1}^{n} \big(\xi_j^- - (a_j\, \tau + b_j)\big) = 0, \quad \psi(\boldsymbol{\xi}^+, \tau) = \prod_{j=1}^{n} \big(\xi_j^+ - (c_j\, \tau + d_j)\big) = 0.$$

Also, a control $\mathbf{u}^\star$ must be found which enables the state of the system to deviate the least from a desired trajectory $\mathbf{x}_\mathrm{d}$. To do so, the following cost function is introduced:

$$\begin{aligned} J(\mathbf{u}) = \frac{1}{2} \int_{t_0}^{t_f} & \Big[\big(\mathbf{x}(t) - \mathbf{x}_\mathrm{d}(t)\big)^\top Q\, \big(\mathbf{x}(t) - \mathbf{x}_\mathrm{d}(t)\big) + \mathbf{u}^\top(t)\, R\, \mathbf{u}(t) \Big]\ dt \\ & + \frac{1}{2}\, \boldsymbol{\xi}^{-\top}\, S_1\, \boldsymbol{\xi}^- + \frac{1}{2}\, \boldsymbol{\xi}^{+\top}\, S_2\, \boldsymbol{\xi}^+ + \mu\, \phi(\boldsymbol{\xi}^-, \tau) + \ \eta\, \psi(\boldsymbol{\xi}^+, \tau) + \frac{1}{2}\, \mathbf{x}_f^\top T\, \mathbf{x}_\mathrm{f}, \end{aligned}$$

where Q, S_1, S_2 and T are positive semidefinite symmetric matrices, and R a positive definite symmetric matrix.

By using the notations of the preceding section, we get:

$$\begin{aligned} F_1(t) = F_2(t) &= \frac{1}{2}\big(\mathbf{x}(t) - \mathbf{x}_\mathrm{d}(t)\big)^\top Q\,\big(\mathbf{x}(t) - \mathbf{x}_\mathrm{d}(t)\big) + \frac{1}{2}\,\mathbf{u}^\top(t)\,R\,\mathbf{u}(t) \\ H_1(t) &= -F_1(t) + \boldsymbol{\lambda}^\top(t)\cdot\big(A_1\,\mathbf{x}(t) + B\,\mathbf{u}(t)\big) \\ H_2(t) &= -F_2(t) + \boldsymbol{\lambda}^\top(t)\cdot\big(A_2\,\mathbf{x}(t) + B\,\mathbf{u}(t)\big). \end{aligned} \tag{11}$$

The equations which allow to calculate both the adjoint state $\boldsymbol{\lambda}$ and the multipliers μ and η are then:

$$\dot{\boldsymbol{\lambda}}(t) = -\frac{\partial H_1}{\partial \mathbf{x}}(t) = -A_1^\top\,\boldsymbol{\lambda}(t) + Q\,\big(\mathbf{x}(t) - \mathbf{x}_\mathrm{d}(t)\big) \quad \text{for } t \in [t_0, \tau) \tag{12}$$

$$\dot{\boldsymbol{\lambda}}(t) = -\frac{\partial H_2}{\partial \mathbf{x}}(t) = -A_2^\top\,\boldsymbol{\lambda}(t) + Q\,\big(\mathbf{x}(t) - \mathbf{x}_\mathrm{d}(t)\big) \quad \text{for } t \in (\tau, t_\mathrm{f}] \tag{13}$$

$$0 = \lambda_i(\tau^-) + \mu \prod_{\substack{j=1 \\ j \neq i}}^{n} (\,\xi_j^- - a_j\tau - b_j) + \zeta_i^- \quad i \in \{1,\ldots,n\} \quad \text{with} \quad \boldsymbol{\zeta}^- = \mathrm{S}_1\,\boldsymbol{\xi}^- \tag{14}$$

$$0 = -\,\lambda_i(\tau^+) + \eta \prod_{\substack{j=1 \\ j \neq i}}^{n} (\,\xi_j^+ - c_j\tau - d_j) + \zeta_i^+ \quad i \in \{1,\ldots,n\} \quad \text{with} \quad \boldsymbol{\zeta}^+ = \mathrm{S}_2\,\boldsymbol{\xi}^+ \tag{15}$$

$$\begin{aligned} 0 = H_2(\tau^+) - H_1(\tau^-) &- \sum_{i=1}^{n}\Bigg(a_i\,\mu \prod_{\substack{j=1 \\ j \neq i}}^{n}(\xi_j^- - a_j\,\tau - b_j)\Bigg) \\ &- \sum_{i=1}^{n}\Bigg(c_i\,\eta \prod_{\substack{j=1 \\ j \neq i}}^{n}(\xi_j^+ - c_j\,\tau - d_j)\Bigg) \end{aligned} \tag{16}$$

$$\boldsymbol{\lambda}(t_\mathrm{f}) = -T\,\mathbf{x}(t_\mathrm{f}). \tag{17}$$

The gradient of the cost function, with respect to $\mathbf{u}$, becomes:

$$\nabla J_{\mathbf{u}}(t) = -\frac{\partial H_1}{\partial \mathbf{u}}(t) = R\,\mathbf{u}(t) - B^\top\,\boldsymbol{\lambda}(t) \quad \text{for} \ \ t \in [t_0, \tau) \tag{18}$$

$$\nabla J_{\mathbf{u}}(t) = -\frac{\partial H_2}{\partial \mathbf{u}}(t) = R\,\mathbf{u}(t) - B^\top\,\boldsymbol{\lambda}(t) \quad \text{for} \ \ t \in (\tau, t_\mathrm{f}]. \tag{19}$$

A descent method which uses the expression of the gradient can now be implemented. To this end we need to calculate $\boldsymbol{\lambda}(\tau^-)$.

Calculation of $\boldsymbol{\lambda}(\tau^-)$

From equations (14) and (15) we get the following terms:

$$\mu \prod_{\substack{j=1 \\ j \neq i}}^{n} (\,\xi_j^- - a_j\tau - b_j) = -\zeta_i^- - \lambda_i(\tau^-) \quad i \in \{1,\ldots,n\} \tag{20}$$

$$\eta \prod_{\substack{j=1 \\ j \neq i}}^{n} (\,\xi_j^+ - c_j\tau - d_j) = \lambda_i(\tau^+) - \zeta_i^+ \quad i \in \{1,\ldots,n\}. \tag{21}$$

We take these terms into account in equation (16) and we replace $H_1(\tau^-)$ and $H_2(\tau^+)$ by their proper expressions (see equation (11)). We then get:

$$\begin{aligned}\boldsymbol{\lambda}^\top(\tau^-)\cdot\big(A_1\,\boldsymbol{\xi}^- + B\,\mathbf{u}(\tau) - \mathbf{a}\big) = \boldsymbol{\lambda}^\top(\tau^+)\cdot\big(A_2\,\boldsymbol{\xi}^+ + B\,\mathbf{u}(\tau) - \mathbf{c}\big)\\ +\mathbf{a}^\top\cdot\ \boldsymbol{\zeta}^- + \mathbf{c}^\top\cdot\ \boldsymbol{\zeta}^+ + \mathbf{F_1}(\tau^-) - \mathbf{F_2}(\tau^+).\end{aligned}\tag{22}$$

Model switching with jump in the state only occurs when a single component of the state vector reaches the corresponding curve of jump. Thus, from equation (21) we can have directly the multiplier η; and we have a linear system, issued from equations (20) and (22), to be solved with $n+1$ unknown variables which are the n components of $\boldsymbol{\lambda}(\tau^-)$ and the multiplier μ.

4.1. General algorithm

A general algorithm that solves the trajectory pursuit problem can be described as follows:

1. Define $\varepsilon > 0$;
2. Initialise $\mathbf{u}^k$ for $k = 0$;
3. Solve system (10) and obtain τ, $\boldsymbol{\xi}^-$, $\boldsymbol{\xi}^+$;
4. Solve (13),(17) between t_f and τ, and calculate $\boldsymbol{\lambda}(\tau^+)$;
5. Calculate $\boldsymbol{\lambda}(\tau^-)$;
6. Solve (12) between τ and t_0;
7. Calculate $\nabla J_{\mathbf{u}^k}$ according to (18) and (19);
8. Calculate $\mathbf{u}^{k+1}$ from $\mathbf{u}^k$ by means of a descent method;
9. If $|J(\mathbf{u}^{k+1})| \le \varepsilon$ then stop; otherwise go to 2.

Since the cost function J is not convex in the general case, the convergence of the algorithm cannot be proved. However, in all cases we considered, no difficulties due to local minima occurred. If such is the case, methods avoiding convergence to local minima could be implemented, such as random perturbation of the gradient, as studied e.g. in Pogu and Souza (1994), which yield almost sure convergence to a global minimum.

5. Numerical example

We consider system (10) with $n = 2$ and $m = 1$ and a trajectory pursuit problem with autonomous model switching and jump of the state. Matrices A_1, A_2 and B are:

$$A_1 = \begin{pmatrix} -2 & -1.5 \\ 1 & 0 \end{pmatrix},\quad A_2 = \begin{pmatrix} -0.7 & -0.5 \\ 1 & 0 \end{pmatrix},\quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

We take $t_0 = 0$ and $t_\mathrm{f} = 10$, and the initial condition is chosen such that:

$$\mathbf{x}_0 = \begin{pmatrix} 0 & 0 \end{pmatrix}^\top.$$

The numerical function which triggers the model switching is written as:

$$\phi(\boldsymbol{\xi}^-, \tau) = \left(\xi_1^- - (-0.3\,\tau + 2.3)\right)\left(\xi_2^- - (-0.3\,\tau + 2.3)\right).$$

The model switching occurs when $\boldsymbol{\phi}(\boldsymbol{\xi}, \tau) = 0$, i.e. when one component of the state reaches, at the instant τ, the straight line of equation $x(t) = -0.3\,t + 2.3$.

The numerical function which gives the value of the state after the jump is the following:

$$\psi(\boldsymbol{\xi}^+, \tau) = \left(\xi_1^+ - (-0.5\,\tau + 3.8)\right)\left(\xi_2^+ - (-0.5\,\tau + 3.8)\right).$$

The component of the state, which triggers the model switching, must be just after the jump on the straight line of equation $x(t) = -0.5\,t + 3.8$.

Matrices Q, R, S and T are the following:

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad R = 10^{-5}, \quad S = T = 0.$$

In order to test the method, we consider a control $\tilde{u}$ and apply it to system (10) whose output is considered to be the desired state $\mathbf{x}_\mathrm{d}$. After algorithm has converged on $u^\star$, we must have $u^\star \simeq \tilde{u}$. We choose the control $\tilde{u}$ as follows:

$$\tilde{u}(t) \;=\; (0.25 \times t)^{1/2}.$$

The algorithm is initialised by taking $u^0(t) = 0$ for $t \in [t_0\,, t_\mathrm{f}]$ and the stop test of iterations is performed with $\varepsilon = 10^{-4}$. The descent method involved is the BFGS method. Fig. 1 shows the values of the cost function with respect to the count of iterations.

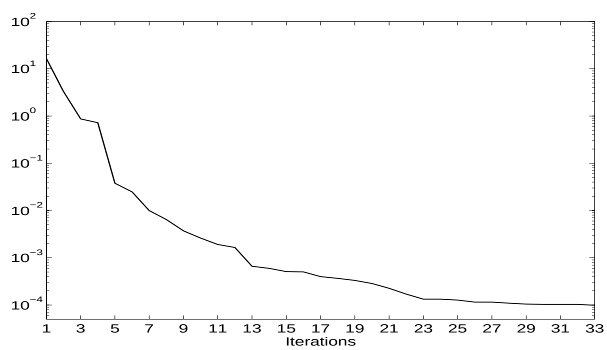

Figure 1. Values of the cost function with respect to iteration number

Fig. 2 shows the control $\tilde{u}$ as well as the optimal control $u^\star$ obtained after 33 iterations. We notice that these two curves are merely identical. The main discrepancies come about for $t = t_\mathrm{f}$, which is due to the fact that the adjoint state is initialised with $\boldsymbol{\lambda}(t_\mathrm{f}) = T\,x(t_\mathrm{f}) = 0$. Another discrepancy also takes place when switching occurs, since for $t = \tau = 3.83$, the adjoint state is discontinuous (see Fig. 5).

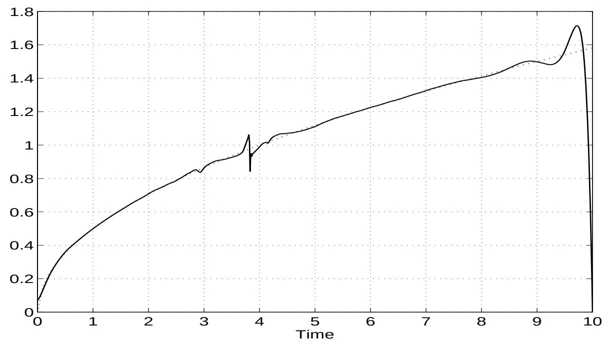

Figure 2. Control $\tilde{u}$ (dashed line) and optimal control $u^{\star}$ (solid line)

We can see in Figs. 3 and 4 that it is the second component of the state that reaches the straight line $\mathbf{x}(t) = -0.3\,t + 2.3$ and causes model switching and jump of the state.

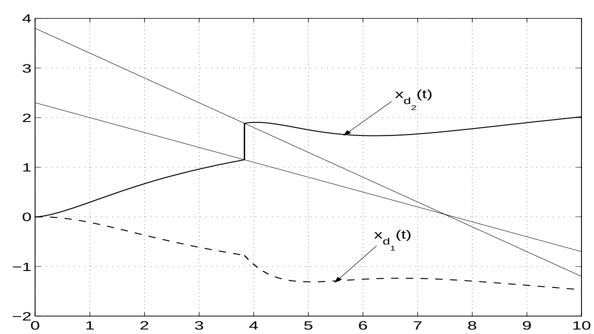

Figure 3. Desired trajectory

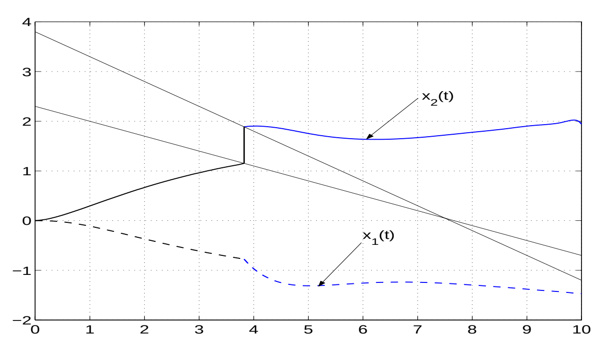

Figure 4. Output of the system

Fig. 5 shows the changes of the adjoint state. Actually, as switching takes place on the second component, according to equation (14), and since $S_1 = 0$, we have $\lambda_1(\tau^-) = 0$; also, according to equation (15) and since $S_2 = 0$, we must have $\lambda_1(\tau^+) = 0$. This result, as well as the discontinuity of the second component of the adjoint state, can be observed in Fig. 5.

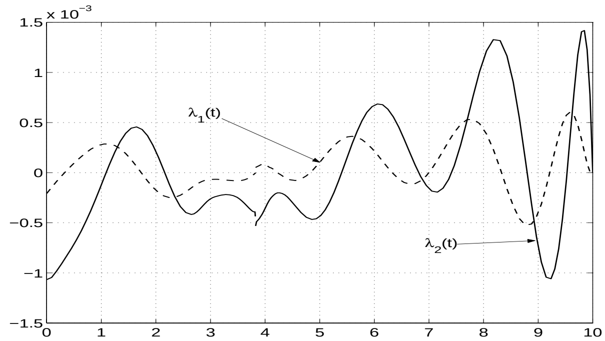

Figure 5. Adjoint state

Similar results, that are not presented here, have been obtained with a system of order $n = 10$ and $m = 5$.

6. Conclusion

We applied the formalism of the calculus of variations to determine an optimal control of a class of HDS. This class consists of autonomous switching systems that switch when the state of the system reaches a given curve.

We demonstrated and calculated the discontinuity of the adjoint state that takes place when model switching occurs. These discontinuities are due either to the switching cost, denoted by the matrix S in the calculations, or by the jump of the state when it occurs. These calculations were applied to a numerical example that validated both the theoretical results and the applicability of the method.

References

Andreu, D., Pascal, J.C. and Valette, R. (1995) Interaction of discrete and continuous parts of a batch process control system. *Proc. of ADEDOPS (Analysis and Design of Event-Driven Operation in Process Systems)*, London.

Antsaklis, P.J., Stiver, J.A. and Lemmon, M.D. (1993) Interface and Controller Design for Hybrid Control Systems. *Lecture Notes in Computer Science, Hybrid System*, Springer-Verlag.

Antsaklis, P.J. and Nerode, A. (1998) Hybrid control systems: An introductory discussion to the special issue. *IEEE Trans. Autom. Contr.* **43** (4), 457, Special issue on hybrid systems.

Bérest, P. (1997) *Calcul des variations, application à la mécanique et à la physique*. Ellipses, Paris.

Branicky, M.S. (1995) *Studies in Hybrid Systems: Modeling, Analysing, and Control*. PhD thesis, Massachussetts Institute of Technology, Cambridge.

Branicky, M.S. (1998) Multiple Lyapunov function and others analysis tools for switched and hybrid systems. *IEEE Trans. Autom. Contr.* **43** (4), 475, Special issue on hybrid systems.

Brockett, R.W. (1993) Hybrid model for motion control systems. In: H. Trentelman and J. Willems, eds., *Essays in Control: Perspectives in the Theory and Its Applications*. Birkhäuser, Boston, 29-53.

Bryson, A.E. and Ho, Y.C. (1975) *Applied Optimal Control*. Gin and Co., Waltham.

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

Cébron, B., Sechilariu, M. and Burger J. (1999b) Comparison of two calculation methods of the optimal switching instants of a hybrid dynamical system, SMC'99. *Proc. of IEEE-SMC*, 165-170.

Ciarlet, P.G. (1990) *Introduction à l'analyse numérique matricielle et à l'optimisation*. Masson, Paris.

Daubas, B., Pagès, A. and Pingaud, H. (1994) Combined simulation of hybrid processes. *Proc. of IEEE-SMC*, 320.

Hedlund, S. and Rantzer, A. (2002) Convex Dynamic Programming for Hybrid Systems, *IEEE Trans. on Autom. Contr.* **47** (9), 1536-1540.

Nerode, A. and Kohn, W. (1993) Models for hybrid systems: Automata, topologies, controllability, observability. *Lecture Notes in Computer Science, Hybrid System*, Springer-Verlag.

Pogu, M. and Souza de cursi, J.E. (1994) Global Optimization by Random Perturbation of the Gradient Method with a Fixed Parameter. *J. of Global Optimization* **5**, 159-180.

Riedinger, P., Kratz, F., Iung, C. and Zanne, C. (1999) Linear quadratic optimization for hybrid systems. *Proc. of the 38th IEEE Conference on Decision and Control*, Phoenix, Arizona.

Sussmann, H.J. (1999) A maximum principle for hybrid optimal control problems. *Proc. of the 38th IEEE Conference on Decision and Control*, Phoenix, Arizona.

Van Der Schaft, A. and Schumacher, H. (2000) An introduction to Hybrid Dynamical Systems. *Lecture Notes in Control and Information Sciences* **251**, Springer-Verlag.

VINTER, R. (1993) Convex duality and nonlinear optimal control. *SIAM J. Control and Optimization* **31** (2), 518-538.

XU, X. and ANTSAKLIS, P.J. (2003) Optimal Control of Hybrid Autonomous Systems with State Jumps. *Proc. of the American Control Conference*, Denver, Colorado,
USA, 5191-5196.

XU, X. and ANTSAKLIS, P.J. (2004) Optimal Control of Switched Systems Based on Parametrization of the Switching Instants. *IEEE Trans. on Autom. Contr.* **49** (1).