

OPTIMAL IMPULSIVE CONTROL OF DELAY SYSTEMS

FLORENT DELMOTTE¹, ERIK I. VERRIEST¹ AND MAGNUS EGERSTEDT¹

Abstract. In this paper, we solve an optimal control problem using the calculus of variation. The system under consideration is a switched autonomous delay system that undergoes jumps at the switching times. The control variables are the instants when the switches occur, and a set of scalars which determine the jump amplitudes. Optimality conditions involving analytic expressions for the partial derivatives of a given cost function with respect to the control variables are derived using the calculus of variation. A locally optimal impulsive control strategy can then be found using a numerical gradient descent algorithm.

Mathematics Subject Classification. 49K15, 49K25

Received July 27, 2006. Revised May 9, 2007.

Published online January 30, 2008.

1. INTRODUCTION

Systems with impulsive inputs have been studied in [2–4], and more recently in [22]. The optimal control problem has also recently received some attention in [7], where a timing problem for the delay-free case is discussed. In the present paper, optimization with respect to the strengths of the impulses as well as their timing, is considered, and necessary conditions are established for switched systems in which the number of switches as well as the sequence of the different dynamical modes are given. Here, the presence of delays also adds a nontrivial twist to the original problem posed in [21].

This optimal impulsive control problem is related to the optimal switching problem [5,9,16]. The paper extends the results of [16] to systems with delays, but derives the optimality conditions *via* a classical variational approach.

Finally, the problem formulation and results of this paper resemble previous work by the authors in [17] and draws its motivation from problems in epidemiology in which this type of switched, delayed systems appear [18,19]. However, an improved methodology and use of notations allows us to overcome two assumptions, and hence bring more generality to the problem: here, the system does not require a refractory period (some non-infinitesimal amount of time between two subsequent actions), and the system dynamics can involve any finite number of delays.

The problem formulation and notations are presented in Section 2. Necessary conditions for the optimal impulsive control of a single delay system are determined in Section 3. An illustrative example is given in Section 4. The results are finally extended to multi-delay systems in Section 5.

Keywords and phrases. Optimal control, impulse control, switched systems, delay systems, calculus of variation.

¹ School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30302, USA;
florent@ece.gatech.edu; erik.verriest@ece.gatech.edu; magnus@ece.gatech.edu

2. PROBLEM FORMULATION

The dynamical system discussed in this paper is modelled by an autonomous point-delay system, *i.e.* a system with discrete delays. Here, and for simplicity, we consider only one delay, *i.e.* a system of the form

$$\dot{x}(t) = f(x(t)) + g(x_\tau(t)),$$

where we use the short notation $x_\tau(t)$ for $x(t-\tau)$. This one-delay assumption, which keeps us from an excessively technical derivation, does not prevent the generalization to any finite number of delays, which will be discussed later. We assume that the control consists of a sequence of amplitude parameters u_i and discrete instants T_i . The effect of this impulsive control is modelled by

$$x(T_i^+) = x(T_i^-) + G_i(x(T_i^-), u_i, T_i), \quad i = 1, \dots, N-1, \quad (2.1)$$

where $N-1$ is the number of jumps and $\{G_i\}_{i=1}^{N-1}$ is a set of given amplitude functions.

As the system dynamics may change because of the impulsive inputs (*e.g.*, due to loss of mass in spacecraft trajectory applications), we let x be given by the differential equation

$$\dot{x}(t) = f_i(x(t)) + g_i(x_\tau(t)), \quad t \in (T_{i-1}, T_i), \quad i = 1, \dots, N, \quad (2.2)$$

where $\{f_i\}_{i=1}^N$ and $\{g_i\}_{i=1}^N$ are given vector fields, T_0 and T_N are given initial and final times, and $x(\theta)$ is known for $T_0 - \tau < \theta < T_0$. (Note here that we use t_f as the given final time but at the same time we will, for indexing reasons, replace this by T_N whenever this can be done without obscuring the presentation.) Now, equation (2.2) can be written in a more general form

$$\dot{x}(t) = f_{\xi(t)}(x(t)) + g_{\xi(t)}(x_\tau(t)), \quad \forall t \in (T_0, T_N), \quad (2.3)$$

where ξ is a discrete state counting the number of impulses, *i.e.* $\xi(t) = i$, if $T_{i-1} < t < T_i$.

The parameters u_i and instants T_i are the control variables to be chosen such that a performance index

$$J = \sum_{i=1}^N \int_{T_{i-1}}^{T_i} L_i(x(t)) dt + \sum_{i=1}^{N-1} K_i(x(T_i^-), u_i, T_i) + \Phi(x(t_f)) \quad (2.4)$$

is optimized. In some applications, the added generality of a running cost, L , depending on the number of past impulses, may be of interest. The K_i are discrete costs associated with the control, and Φ is the terminal cost at the fixed terminal time. Note that, in view of the above, we may set

$$\Phi(x(T_N)) = K_N(x(T_N), 0, T_N), \quad (2.5)$$

thus including the terminal cost in the sum of the control costs. This is useful in the more general problem of a free endpoint.

3. VARIATIONAL APPROACH

As stated, the problem is a *parameter optimization* problem. However, solving it as such requires the explicit solution of the state equations and their dependencies on the u_i and T_i . We therefore solve the problem using

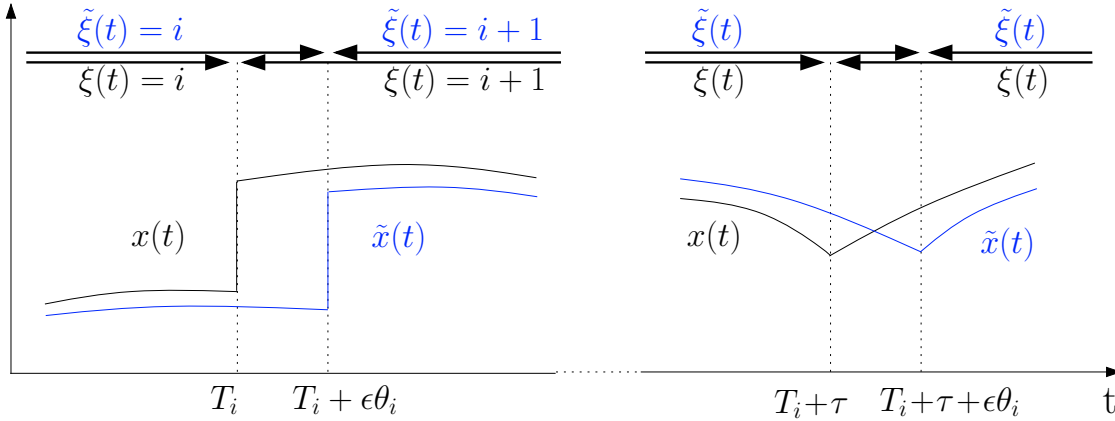


FIGURE 1. Compared trajectories.

classical variational methods instead [6]. For this, we will analyze the cost variation between two systems: an “unperturbed” system x with control instants T_i and strength u_i , and a “perturbed” system \tilde{x} for which arbitrary, independent perturbations are added to the control variables *i.e.* $T_i \rightarrow T_i + \epsilon\theta_i$, and $u_i \rightarrow u_i + \epsilon\nu_i$, with $\epsilon \rightarrow 0$. Figure 1 shows how x and \tilde{x} differ, with a focus on the intervals $(T_i, T_i + \epsilon\theta_i)$ and $(T_i + \tau, T_i + \tau + \epsilon\theta_i)$. Outside these intervals, where the variation is continuous and ϵ -small, we will write $\tilde{x}(t) = x(t) + \epsilon\eta(t)$. The following equations describe the differences between the two systems:

- on $(T_i, T_i + \epsilon\theta_i)$,

$$\xi(t) = i + 1 \tag{3.1}$$

$$\tilde{\xi}(t) = i \tag{3.2}$$

$$x(t) = x(T_i^+) + \mathcal{O}(\epsilon) \tag{3.3}$$

$$\tilde{x}(t) = x(T_i^-) + \mathcal{O}(\epsilon) \tag{3.4}$$

$$\tilde{x}_\tau(t) = x_\tau(t) + \mathcal{O}(\epsilon) = x_\tau(T_i) + \mathcal{O}(\epsilon) \tag{3.5}$$

$$\dot{\tilde{x}}(t) = \dot{x}(T_i^-) + \mathcal{O}(\epsilon) \tag{3.6}$$

- on $(T_i + \tau, T_i + \tau + \epsilon\theta_i)$,

$$\tilde{\xi}(t) = \xi(t) = \xi(T_i + \tau) \tag{3.7}$$

$$\tilde{x}(t) = x(t) + \mathcal{O}(\epsilon) = x(T_i + \tau) + \mathcal{O}(\epsilon) \tag{3.8}$$

$$x_\tau(t) = x(T_i^+) + \mathcal{O}(\epsilon) \tag{3.9}$$

$$\tilde{x}_\tau(t) = x(T_i^-) + \mathcal{O}(\epsilon) \tag{3.10}$$

$$\tilde{x}_\tau(t) = x_\tau(t) + \mathcal{O}(\epsilon) = x_\tau(T_i) + \mathcal{O}(\epsilon) \tag{3.11}$$

$$\dot{\tilde{x}}(t) = \dot{x}(T_i + \tau^-) + \mathcal{O}(\epsilon) \tag{3.12}$$

- everywhere else,

$$\tilde{\xi}(t) = \xi(t) \tag{3.13}$$

$$\tilde{x}(t) = x(t) + \epsilon\eta(t) \tag{3.14}$$

$$\tilde{x}_\tau(t) = x_\tau(t) + \epsilon\eta_\tau(t) \tag{3.15}$$

$$\dot{\tilde{x}}(t) = \dot{x}(t) + \epsilon\dot{\eta}(t). \tag{3.16}$$

Note that we have assumed $T_0 < T_1 < \dots < T_N - 1 < T_N = t_f$, so that, as $\epsilon \rightarrow 0$, there is no possible overlap between any two sets $(T_i, T_i + \epsilon\theta_i)$ and $(T_j, T_j + \epsilon\theta_j)$. Similarly, we have assumed no overlap between any two sets $(T_i + \tau, T_i + \tau + \epsilon\theta_i)$ and $(T_j, T_j + \epsilon\theta_j)$. Finally, equation (3.7) does not require that on $(T_i + \tau, T_i + \tau + \epsilon\theta_i)$, $\tilde{\xi}(t) = \xi(t) = i + 1$. This means that we allow for subsequent impulses happening before $T_i + \tau$ (i.e. $T_{i+1} < T_i + \tau$). In other words, we do not require a refractory period of τ seconds after each jump, as was the case in [17].

We now analyze the induced variation in the performance index. Since at the jump times the functions $f_i(x(t))$ jump, equation (2.3) cannot be satisfied at these times. Hence, one must take care in adjoining the dynamical constraints only in the *open* subintervals defined by the jump times. This will be done with a Lagrange multiplier, $\lambda(t)$, defined in the subintervals between the state jumps. In addition, the jump constraints (2.1) are adjoined with Lagrange multipliers, μ_i , to the discrete summation in (2.4).

$$\begin{aligned} \bar{J}_0 &= \sum_{i=1}^N \int_{T_{i-1}}^{T_i} [L_{\xi(t)}(x(t)) + \lambda'(t)(f_{\xi(t)}(x(t)) + g_{\xi(t)}(x_\tau(t)) - \dot{x}(t))] dt \\ &\quad + \sum_{i=1}^N [K_i(x(T_i^-), u_i, T_i) + \mu'_i(G_i(x(T_i^-), u_i, T_i) - x(T_i^+) + x(T_i^-))]. \end{aligned} \tag{3.17}$$

For simplicity, we set $K_i = K_i(x(T_i^-), u_i, T_i)$, $G_i = G_i(x(T_i^-), u_i, T_i)$, and $\Delta x|_{T_i} = x(T_i^+) - x(T_i^-)$. We also define the *Hamiltonian functionals*,

$$H_\xi(x, x_\tau, \lambda) \stackrel{\text{def}}{=} L_\xi(x) + \lambda'[f_\xi(x) + g_\xi(x_\tau)], \tag{3.18}$$

and the *Lagrangian constraints*,

$$M_i \stackrel{\text{def}}{=} K_i + \mu'_i G_i. \tag{3.19}$$

Finally, in order to lighten the equations, we drop the time-dependency notation '(t)' wherever possible. Equation (3.17) now writes

$$\begin{aligned} \bar{J}_0 &= \sum_{i=1}^N \int_{T_{i-1}}^{T_i} [H_\xi(x, x_\tau, \lambda) - \lambda' \dot{x}] dt + \sum_{i=1}^N [M_i(x(T_i^-), u_i, T_i) - \mu'_i \Delta x|_{T_i}] \\ &= J_0^{(1)} + J_0^{(2)}. \end{aligned} \tag{3.20}$$

Similarly, for the perturbed system, we get

$$\begin{aligned} \bar{J}_\epsilon &= \sum_{i=1}^N \int_{T_{i-1} + \epsilon\theta_{i-1}}^{T_i + \epsilon\theta_i} [H_{\tilde{\xi}}(\tilde{x}, \tilde{x}_\tau, \lambda) - \lambda' \dot{\tilde{x}}] dt + \sum_{i=1}^N [M_i(\tilde{x}(T_i + \epsilon\theta_i^-), u_i + \epsilon\nu_i, T_i + \epsilon\theta_i) - \mu'_i \Delta \tilde{x}|_{T_i + \epsilon\theta_i}] \\ &= J_\epsilon^{(1)} + J_\epsilon^{(2)}. \end{aligned} \tag{3.21}$$

Note that on the ϵ -small intervals $(T_i, T_i + \epsilon\theta_i)$ and $(T_i + \tau, T_i + \tau + \epsilon\theta_i)$, the discrepancies between ξ and $\tilde{\xi}$, x and \tilde{x} , or x_τ and \tilde{x}_τ also yield a discrepancy in the Hamiltonian. Therefore, we suggest to carefully split the integral terms $J_0^{(1)}$ and $J_\epsilon^{(1)}$ of equations (3.20) and (3.21), then plug with equations (3.1)–(3.16). We get

$$\begin{aligned}
J_0^{(1)} &= \sum_{i=1}^N \int_{T_{i-1}}^{T_{i-1} + \epsilon\theta_{i-1}} [H_\xi(x, x_\tau, \lambda) - \lambda' \dot{x}] dt + \sum_{i=1}^N \int_{T_{i-1} + \epsilon\theta_{i-1}}^{T_{i-1} + \tau} [H_\xi(x, x_\tau, \lambda) - \lambda' \dot{x}] dt \\
&\quad + \sum_{i=1}^N \int_{T_{i-1} + \tau}^{T_{i-1} + \tau + \epsilon\theta_{i-1}} [H_\xi(x, x_\tau, \lambda) - \lambda' \dot{x}] dt + \sum_{i=1}^N \int_{T_{i-1} + \tau + \epsilon\theta_{i-1}}^{T_i} [H_\xi(x, x_\tau, \lambda) - \lambda' \dot{x}] dt \\
&= \sum_{i=1}^N \epsilon\theta_{i-1} [H_i(x(T_{i-1}^+), x_\tau(T_{i-1}), \lambda(T_{i-1}^+)) - \lambda(T_{i-1}^+) \dot{x}(T_{i-1}^+)] + \sum_{i=1}^N \int_{T_{i-1} + \epsilon\theta_{i-1}}^{T_{i-1} + \tau} [H_\xi(x, x_\tau, \lambda) - \lambda' \dot{x}] dt \\
&\quad + \sum_{i=1}^N \epsilon\theta_{i-1} [H_{\xi(T_{i-1} + \tau)}(x(T_{i-1} + \tau), x(T_{i-1}^+), \lambda(T_{i-1} + \tau^+)) - \lambda(T_{i-1} + \tau^+) \dot{x}(T_{i-1} + \tau^+)] \\
&\quad + \sum_{i=1}^N \int_{T_{i-1} + \tau + \epsilon\theta_{i-1}}^{T_i} [H_\xi(x, x_\tau, \lambda) - \lambda' \dot{x}] dt + o(\epsilon), \tag{3.22}
\end{aligned}$$

and

$$\begin{aligned}
J_\epsilon^{(1)} &= \sum_{i=1}^N \int_{T_{i-1} + \epsilon\theta_{i-1}}^{T_{i-1} + \tau} [H_{\tilde{\xi}}(\tilde{x}, \tilde{x}_\tau, \lambda) - \lambda' \dot{\tilde{x}}] dt + \sum_{i=1}^N \int_{T_{i-1} + \tau}^{T_{i-1} + \tau + \epsilon\theta_{i-1}} [H_{\tilde{\xi}}(\tilde{x}, \tilde{x}_\tau, \lambda) - \lambda' \dot{\tilde{x}}] dt \\
&\quad + \sum_{i=1}^N \int_{T_{i-1} + \tau + \epsilon\theta_{i-1}}^{T_i} [H_{\tilde{\xi}}(\tilde{x}, \tilde{x}_\tau, \lambda) - \lambda' \dot{\tilde{x}}] dt + \sum_{i=1}^N \int_{T_i}^{T_i + \epsilon\theta_i} [H_{\tilde{\xi}}(\tilde{x}, \tilde{x}_\tau, \lambda) - \lambda' \dot{\tilde{x}}] dt \\
&= \sum_{i=1}^N \int_{T_{i-1} + \epsilon\theta_{i-1}}^{T_{i-1} + \tau} [H_\xi(x + \epsilon\eta, x_\tau + \epsilon\eta_\tau, \lambda) - \lambda'(\dot{x} + \epsilon\dot{\eta})] dt \\
&\quad + \sum_{i=1}^N \epsilon\theta_{i-1} [H_{\xi(T_{i-1} + \tau)}(x(T_{i-1} + \tau), x(T_{i-1}^-), \lambda(T_{i-1} + \tau^+)) - \lambda(T_{i-1} + \tau^+) \dot{x}(T_{i-1} + \tau^-)] \\
&\quad + \sum_{i=1}^N \int_{T_{i-1} + \tau + \epsilon\theta_{i-1}}^{T_i} [H_\xi(x + \epsilon\eta, x_\tau + \epsilon\eta_\tau, \lambda) - \lambda'(\dot{x} + \epsilon\dot{\eta})] dt \\
&\quad + \sum_{i=1}^N \epsilon\theta_i [H_i(x(T_i^-), x_\tau(T_i), \lambda(T_i^+)) - \lambda(T_i^+) \dot{x}(T_i^-)] + o(\epsilon). \tag{3.23}
\end{aligned}$$

Now, a Taylor expansion to first order in ϵ , a change in variable, and the fact that $\theta_0 = \theta_N = 0$ (as T_0 and T_N are fixed initial and final times) give us the first part of the directional derivative:

$$\begin{aligned} \frac{J_\epsilon^{(1)} - J_0^{(1)}}{\epsilon} &= \sum_{i=1}^N \int_{T_{i-1} + \epsilon\theta_{i-1}}^{T_{i-1} + \tau} [D_x H_\xi - \lambda' \dot{\eta}] dt + \sum_{i=1}^N \int_{T_{i-1} + \tau + \epsilon\theta_{i-1}}^{T_i} [D_x H_\xi - \lambda' \dot{\eta}] dt \\ &+ \sum_{i=1}^{N-1} \theta_i [H_{\xi(T_i + \tau)}(x(T_i + \tau), x(T_i^-), \lambda(T_i + \tau^+)) - \lambda(T_i + \tau^+) \dot{x}(T_i + \tau^-)] \\ &+ \sum_{i=1}^{N-1} \theta_i [H_i(x(T_i^-), x_\tau(T_i), \lambda(T_i^+)) - \lambda(T_i^+) \dot{x}(T_i^-)] \\ &- \sum_{i=1}^{N-1} \theta_i [H_{i+1}(x(T_i^+), x_\tau(T_i), \lambda(T_i^+)) - \lambda(T_i^+) \dot{x}(T_i^+)] \\ &- \sum_{i=1}^{N-1} \theta_i [H_{\xi(T_i + \tau)}(x(T_i + \tau), x(T_i^+), \lambda(T_i + \tau^+)) - \lambda(T_i + \tau^+) \dot{x}(T_i + \tau^+)] + o(1). \end{aligned} \tag{3.24}$$

Here $D_x H_\xi$ is the functional derivative of H_ξ

$$D_x H_\xi = \lim_{\epsilon \rightarrow 0} \frac{H_\xi(x + \epsilon\eta, x_\tau + \epsilon\eta_\tau, \lambda) - H_\xi(x, x_\tau, \lambda)}{\epsilon}. \tag{3.25}$$

Everywhere in (3.24), we replace the Hamiltonian with its expression from (3.18). We also take a first order Taylor approximation to compute $D_x H_\xi$, and we integrate by parts the $\lambda' \dot{\eta}$ terms:

$$\begin{aligned} \frac{J_\epsilon^{(1)} - J_0^{(1)}}{\epsilon} &= \sum_{i=1}^N \int_{T_{i-1} + \epsilon\theta_{i-1}}^{T_{i-1} + \tau} \left[\frac{\partial L_\xi}{\partial x} \eta + \lambda' \left(\frac{\partial f_\xi}{\partial x} \eta + \frac{\partial g_\xi}{\partial x_\tau} \eta_\tau \right) + \dot{\lambda}' \eta \right] dt - [\lambda' \eta]_{T_{i-1} + \epsilon\theta_{i-1}}^{T_{i-1} + \tau} \\ &+ \sum_{i=1}^N \int_{T_{i-1} + \tau + \epsilon\theta_{i-1}}^{T_i} \left[\frac{\partial L_\xi}{\partial x} \eta + \lambda' \left(\frac{\partial f_\xi}{\partial x} \eta + \frac{\partial g_\xi}{\partial x_\tau} \eta_\tau \right) + \dot{\lambda}' \eta \right] dt - [\lambda' \eta]_{T_{i-1} + \tau + \epsilon\theta_{i-1}}^{T_i} \\ &+ \sum_{i=1}^{N-1} \theta_i [L_i(x(T_i^-)) - L_{i+1}(x(T_i^+))] \\ &+ \lambda(T_i^+) (f_i(x(T_i^-)) + g_i(x(T_i - \tau))) \\ &- \lambda(T_i^+) (f_{i+1}(x(T_i^+)) + g_{i+1}(x(T_i - \tau))) \\ &+ \lambda(T_i + \tau) (g_{\xi(T_i + \tau)}(x(T_i^-)) - g_{\xi(T_i + \tau)}(x(T_i^+))) \\ &+ \lambda(T_i^+) \Delta \dot{x}|_{T_i} + \lambda(T_i + \tau^+) \Delta \dot{x}|_{T_i + \tau} + o(1). \end{aligned} \tag{3.26}$$

Before looking at the second part of the directional derivative, we derive a few useful equations:

$$\begin{aligned} \tilde{x}(T_i + \epsilon\theta_i^-) &= \tilde{x}(T_i^-) + \epsilon\theta_i \dot{\tilde{x}}(T_i^-) + o(\epsilon) \\ &= x(T_i^-) + \epsilon\eta(T_i^-) + \epsilon\theta_i \dot{x}(T_i^-) + o(\epsilon), \end{aligned} \tag{3.27}$$

$$\begin{aligned} \Delta \tilde{x}|_{T_i + \epsilon\theta_i} &= \tilde{x}(T_i + \epsilon\theta_i^+) - \tilde{x}(T_i + \epsilon\theta_i^-) \\ &= x(T_i + \epsilon\theta_i^+) + \epsilon\eta(T_i + \epsilon\theta_i^+) - \tilde{x}(T_i + \epsilon\theta_i^-) \\ &= x(T_i^+) + \epsilon\theta_i \dot{x}(T_i^+) + \epsilon\eta(T_i + \epsilon\theta_i^+) - x(T_i^-) - \epsilon\eta(T_i^-) - \epsilon\theta_i \dot{x}(T_i^-) + o(\epsilon) \\ &= \Delta x|_{T_i} + \epsilon\Delta\eta|_{T_i} + \epsilon\theta_i \Delta \dot{x}|_{T_i} + o(\epsilon). \end{aligned} \tag{3.28}$$

Similarly, at $t = T_i + \tau + \epsilon\theta_i$, we get

$$\Delta\tilde{x}|_{T_i+\tau+\epsilon\theta_i} = \Delta x|_{T_i+\tau} + \epsilon\Delta\eta|_{T_i+\tau} + \epsilon\theta_i\Delta\dot{x}|_{T_i+\tau} + o(\epsilon), \quad (3.29)$$

and because both $x(t)$ and $\tilde{x}(t)$ are continuous at $T_i + \tau$,

$$\Delta\eta|_{T_i+\tau} + \theta_i\Delta\dot{x}|_{T_i+\tau} + o(1) = 0. \quad (3.30)$$

With these, the nonintegral terms in (3.21) expand to

$$\begin{aligned} J_\epsilon^{(2)} &= \sum_{i=1}^N [M_i(\tilde{x}(T_i + \epsilon\theta_i^-), u_i + \epsilon\nu_i, T_i + \epsilon\theta_i) - \mu'_i\Delta\tilde{x}|_{T_i+\epsilon\theta_i}] \\ &= \sum_{i=1}^N [M_i(x(T_i^-) + \epsilon\eta(T_i^-) + \epsilon\theta_i\dot{x}(T_i^-), u_i + \epsilon\nu_i, T_i + \epsilon\theta_i) - \mu'_i(\Delta x|_{T_i} + \epsilon\Delta\eta|_{T_i} + \epsilon\theta_i\Delta\dot{x}|_{T_i}) + o(\epsilon)] \\ &= \sum_{i=1}^N [M_i(x(T_i^-), u_i, T_i) - \mu'_i\Delta x|_{T_i}] \\ &\quad + \epsilon \sum_{i=1}^N \left[\frac{\partial M_i}{\partial x}(\eta(T_i^-) + \theta_i\dot{x}(T_i^-)) + \frac{\partial M_i}{\partial u}\nu_i + \frac{\partial M_i}{\partial T}\theta_i - \mu'_i\Delta\eta|_{T_i} - \mu'_i\theta_i\Delta\dot{x}|_{T_i} \right] + o(\epsilon), \end{aligned} \quad (3.31)$$

where $\frac{\partial M_i}{\partial \alpha}$ is the partial derivative of $M_i(x, u, T, \mu)$ taken at $(x(T_i^-), u_i, T_i)$. We can extract the second part of the directional derivative

$$\frac{J_\epsilon^{(2)} - J_0^{(2)}}{\epsilon} = \sum_{i=1}^N \left[\frac{\partial M_i}{\partial x}(\eta(T_i^-) + \theta_i\dot{x}(T_i^-)) + \frac{\partial M_i}{\partial u}\nu_i + \frac{\partial M_i}{\partial T}\theta_i - \mu'_i\Delta\eta|_{T_i} - \mu'_i\theta_i\Delta\dot{x}|_{T_i} \right]. \quad (3.32)$$

Now, by adding equations (3.26) and (3.32), and rearranging terms, we get (in the limit) an expression for the total directional derivative of J :

$$\begin{aligned} \delta J &= \lim_{\epsilon \rightarrow 0} \frac{J_\epsilon - J_0}{\epsilon} \\ &= \sum_{i=1}^{N-1} \frac{\partial M_i}{\partial u}\nu_i + \sum_{i=1}^{N-1} \left[\frac{\partial M_i}{\partial x}\dot{x}(T_i^-) + \frac{\partial M_i}{\partial T} + (\lambda(T_i^+) - \mu'_i)\Delta\dot{x}|_{T_i} + L_i(x(T_i^-)) - L_{i+1}(x(T_i^+)) \right. \\ &\quad \left. + \lambda(T_i^+)'(f_i(x(T_i^-)) + g_i(x(T_i - \tau))) - \lambda(T_i^+)'(f_{i+1}(x(T_i^+)) + g_{i+1}(x(T_i - \tau))) \right. \\ &\quad \left. + \lambda(T_i + \tau)'(g_{\xi(T_i+\tau)}(x(T_i^-)) - g_{\xi(T_i+\tau)}(x(T_i^+))) \right] \theta_i + \sum_{i=1}^N \int_{T_{i-1}}^{T_i+\tau} \left[\frac{\partial L_\xi}{\partial x}\eta + \lambda' \left(\frac{\partial f_\xi}{\partial x}\eta + \frac{\partial g_\xi}{\partial x_\tau}\eta_\tau \right) + \dot{\lambda}'\eta \right] dt \\ &\quad + \sum_{i=1}^N \int_{T_{i-1}+\tau}^{T_i} \left[\frac{\partial L_\xi}{\partial x}\eta + \lambda' \left(\frac{\partial f_\xi}{\partial x}\eta + \frac{\partial g_\xi}{\partial x_\tau}\eta_\tau \right) + \dot{\lambda}'\eta \right] dt \\ &\quad + \sum_{i=1}^N \left[\frac{\partial M_i}{\partial x} + \mu'_i - \lambda'(T_i^-) \right] \eta(T_i^-) + \sum_{i=1}^N [-\mu'_i + \lambda(T_i^+)'] \eta(T_i^+) + \sum_{i=1}^N [\lambda(T_i + \tau^+) - \lambda'(T_i + \tau^-)] \eta(T_i + \tau^-). \end{aligned} \quad (3.33)$$

The next step consists of choosing λ and μ_i so that all η terms disappear.

First, consider the expression under the integral terms. After a change in variable $t - \tau \rightarrow t'$ for the η_τ terms, this expression becomes $[\frac{\partial L_{\xi(t)}(x(t))}{\partial x} + \lambda(t)' \frac{\partial f_{\xi(t)}(x(t))}{\partial x} + \lambda(t + \tau)' \frac{\partial g_{\xi(t+\tau)}(x(t))}{\partial x_\tau} + \dot{\lambda}(t)']\eta(t)$. Also note that on the last interval, this expression is still true if we set

$$\lambda(t) = 0 \text{ on } (t_f, t_f + \tau). \tag{3.34}$$

To avoid computation of η on these intervals, we need to choose λ so that on the intervals $(T_{i-1}, T_{i-1} + \tau)$ and $(T_{i-1} + \tau, T_i)$, ($i = 1, \dots, N$),

$$\dot{\lambda}(t)' = -\frac{\partial L_{\xi(t)}(x(t))}{\partial x} - \lambda(t)' \frac{\partial f_{\xi(t)}(x(t))}{\partial x} - \lambda(t + \tau)' \frac{\partial g_{\xi(t+\tau)}(x(t))}{\partial x_\tau}. \tag{3.35}$$

To avoid computation of η at T_i^+ , we choose the discrete Lagrange multipliers μ_i so that

$$\mu_i' = \lambda(T_i^+)', \quad i = 1, \dots, N. \tag{3.36}$$

To avoid computation of η at T_i^- , we choose λ to be discontinuous at the instants T_i , with

$$\lambda(T_i^-)' = \lambda(T_i^+)' + \frac{\partial M_i}{\partial x}, \quad i = 1, \dots, N. \tag{3.37}$$

Finally, to avoid computation of η at $T_i + \tau^-$, we choose λ to be continuous at the instants $T_i + \tau$

$$\lambda'(T_i + \tau^-) = \lambda'(T_i + \tau^+), \quad i = 1, \dots, N. \tag{3.38}$$

Now that all η terms have disappeared, we are left with an expression of the form

$$\delta J = \sum_{i=1}^{N-1} \frac{\partial J}{\partial u_i} \nu_i + \sum_{i=1}^{N-1} \frac{\partial J}{\partial T_i} \theta_i. \tag{3.39}$$

By identification, we get

$$\frac{\partial J}{\partial u_i} = \frac{\partial M_i}{\partial u}, \tag{3.40}$$

and

$$\begin{aligned} \frac{\partial J}{\partial T_i} &= \frac{\partial M_i}{\partial x} \dot{x}(T_i^-) + \frac{\partial M_i}{\partial T} + (\lambda(T_i^+)' - \mu_i') \Delta \dot{x}|_{T_i} \\ &\quad + L_i(x(T_i^-)) - L_{i+1}(x(T_i^+)) \\ &\quad + \lambda(T_i^+)'(f_i(x(T_i^-)) + g_i(x(T_i - \tau))) \\ &\quad - \lambda(T_i^+)'(f_{i+1}(x(T_i^+)) + g_{i+1}(x(T_i - \tau))) \\ &\quad + \lambda(T_i + \tau)'(g_{\xi(T_i+\tau)}(x(T_i^-)) - g_{\xi(T_i+\tau)}(x(T_i^+))) \\ &= (\lambda(T_i^-)' - \lambda(T_i^+)'(f_i(x(T_i^-)) + g_i(x(T_i - \tau)))) + \frac{\partial M_i}{\partial T} \\ &\quad + L_i(x(T_i^-)) - L_{i+1}(x(T_i^+)) \\ &\quad + \lambda(T_i^+)'(f_i(x(T_i^-)) + g_i(x(T_i - \tau))) \\ &\quad - \lambda(T_i^+)'(f_{i+1}(x(T_i^+)) + g_{i+1}(x(T_i - \tau))) \\ &\quad + \lambda(T_i + \tau)'(g_{\xi(T_i+\tau)}(x(T_i^-)) - g_{\xi(T_i+\tau)}(x(T_i^+))). \end{aligned} \tag{3.41}$$

After reorganization, and using (3.18)

$$\frac{\partial J}{\partial T_i} = \frac{\partial M_i}{\partial T} + H_i(T_i^-) - H_{i+1}(T_i^+) + \lambda(T_i + \tau)'(g_{\xi(T_i+\tau)}(x(T_i^-)) - g_{\xi(T_i+\tau)}(x(T_i^+))). \quad (3.42)$$

We summarize these results in the theorem below.

Theorem 3.1. *Given smooth functions $\{f_i\}_{i=1}^{N-1}$ and $\{g_i\}_{i=1}^{N-1}$ from \mathbb{R}^n to \mathbb{R}^n , $\{L_i\}_{i=1}^{N-1}$ from \mathbb{R}^n to \mathbb{R} , $\{G_i\}_{i=1}^{N-1}$ and $\{K_i\}_{i=1}^N$ from $\mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$ to \mathbb{R}^n , a necessary condition for the impulsive system with dynamic equations*

$$\dot{x}(t) = f_{\xi(t)}(x(t)) + g_{\xi(t)}(x_{\tau}(t)), \quad t \in (T_0, T_N), \quad x(t) \text{ given on for } t \in [T_0 - \tau, T_0], \quad (3.43)$$

$$\xi(t) = i, \quad t \in (T_{i-1}, T_i), \quad i = 1, \dots, N, \quad (3.44)$$

$$x(T_i^+) = x(T_i^-) + G_i(x(T_i^-), u_i, T_i), \quad i = 1, \dots, N - 1, \quad (3.45)$$

to minimize the performance index

$$J = \sum_{i=1}^N \int_{T_{i-1}}^{T_i} L_i(x(t)) dt + \sum_{i=1}^N K_i(x(T_i^-), u_i, T_i) \quad (3.46)$$

is that the control variables $\{T_i\}_{i=1}^{N-1}$ and $\{u_i\}_{i=1}^{N-1}$ satisfy:

Define:

$$H_i = L_i + \lambda_i[f_i(x) + g_i(x_{\tau})] \quad (3.47)$$

$$M_i = K_i + \mu_i G_i. \quad (3.48)$$

Euler-Lagrange equations:

$$\dot{\lambda}(t)' = -\frac{\partial L_{\xi(t)}(x(t))}{\partial x} - \lambda(t)' \frac{\partial f_{\xi(t)}(x(t))}{\partial x} - \lambda(t + \tau)' \frac{\partial g_{\xi(t+\tau)}(x(t))}{\partial x_{\tau}}. \quad (3.49)$$

Boundary conditions:

$$\begin{aligned} \lambda(t) &= 0, & \text{on } (t_f, t_f + \tau) \\ \lambda(T_i^-)' &= \lambda(T_i^+) + \frac{\partial M_i}{\partial x}, & i = 1, \dots, N. \end{aligned} \quad (3.50)$$

Multipliers:

$$\mu_i' = \lambda(T_i^+)', \quad i = 1, \dots, N. \quad (3.51)$$

Optimality conditions:

$$\frac{\partial M_i}{\partial u} = 0, \quad (3.52)$$

$$\frac{\partial M_i}{\partial T} + H_i(T_i^-) - H_{i+1}(T_i^+) + \lambda(T_i + \tau)'(g_{\xi(T_i+\tau)}(x(T_i^-)) - g_{\xi(T_i+\tau)}(x(T_i^+))) = 0. \quad (3.53)$$

Remark 3.1. The calculus of variations leads to necessary conditions for optimality of smooth unconstrained controls: the cancellation of all partial derivatives of the cost function with respect to the control variables. Analytic solutions to (3.52)–(3.53) may be quite hard to achieve. Instead, the expressions for the partial derivatives of J can be used in a numerical gradient descent algorithm. At each iteration, for the given control variables, the algorithm computes numerically the state and mode trajectories $x(t)$ and $\xi(t)$ forward in time,

the costate $\lambda(t)$ backward in time, the partial derivatives $\frac{\partial J}{\partial u_i}$ and $\frac{\partial J}{\partial T_i}$, finally updates the control variables in the direction of negative gradient. The algorithm converges to a local minimum of the performance index J . An example is given in the next section.

Remark 3.2. Note that equations (3.52) and (3.53) correspond respectively to evaluations of the sensitivities of J with respect to u_i and T_i . As such, they can be used for providing descent directions in numerical optimization algorithms.

Remark 3.3. In the case of a delay free system, we get the requisite necessary conditions by letting $g_i = 0$ in the above theorem. See also [7,10,12–15]. An approach to the minimization of a functional on a Banach space by steepest descent methods is reviewed in [11]. A related problem was treated in [9].

4. ILLUSTRATIVE EXAMPLE

Consider the scalar delay system

$$\begin{aligned} x(t) &= 1 & t \in (-1, 0) \\ \dot{x}(t) &= x(t)/2 & t \in (0, T_1) \\ \dot{x}(t) &= x(t-1)/2 & t \in (T_1, T_2) \\ \dot{x}(t) &= x(t)/2 & t \in (T_2, 3) \end{aligned}$$

with two impulses

$$\begin{aligned} x(T_1^+) &= x(T_1^-) + u_1 \\ x(T_2^+) &= x(T_2^-) + u_2 \end{aligned}$$

and performance index

$$J = \frac{1}{2} \int_0^3 (x(t) - 1)^2 dt + \frac{u_1^2}{T_1} + \frac{u_2^2}{T_2}.$$

Using the same notations we have used so far, we let:

$$\begin{aligned} T_0 = 0, \quad f_1(x) = x/2, \quad g_1(x) = 0, \quad L_1(x) &= \frac{1}{2}(x - 1)^2, \quad G_1(x(T_1^-), u_1, T_1) = u_1, \quad K_1(x(T_1^-), u_1, T_1) = \frac{u_1^2}{T_1}, \\ T_3 = 3, \quad f_2(x) = 0, \quad g_2(x) = x/2, \quad L_2(x) &= \frac{1}{2}(x - 1)^2, \quad G_2(x(T_2^-), u_2, T_2) = u_2, \quad K_2(x(T_2^-), u_2, T_2) = \frac{u_2^2}{T_2}, \\ \tau = 1, \quad f_3(x) = x/2, \quad g_3(x) = 0, \quad L_3(x) &= \frac{1}{2}(x - 1)^2, \quad K_3(x(T_3^-)) = 0. \end{aligned}$$

A numerical gradient descent algorithm using Theorem 3.1 is applied to minimize the cost J . This algorithm is initiated with control variables $T_1 = 1, T_2 = 2, u_1 = 0, u_2 = 0$. Figure 2(a) shows how J quickly converges to a local minimum as the control variables, in Figures 2(a) and (b), are updated through twenty iterations of the algorithm. Figure 2(d) shows the evolution of the state x and costate λ of the system at the last iteration, with control variables $T_1 \approx 0.64, T_2 \approx 1.27, u_1 \approx -0.33, u_2 \approx -0.58$, and minimum cost $J \approx 0.61$. Note how, in this particular example, the optimal control switching times T_1 and T_2 are within τ seconds of each other, *i.e.* $T_2 < T_1 + \tau$. The previous work by the author in [17], which requires a refractory period of τ seconds between each switching times, can only provide a suboptimal solution.

To confirm our results, we ran a brute force computation using Mathematica. Assuming $T_1 < \tau < T_2 < T_1 + \tau$, we computed J and its partial derivatives as functions of (T_1, T_2, u_1, u_2) . An analytic solution for the cancellation of the derivatives could not be found within a reasonable amount of time (one hour), but a numerical search for the minimization of J returned the same solution with $T_1 = 0.645824, T_2 = 1.27059, u_1 = -0.335215, u_2 = -0.578733$. Note that the better resolution comes from the fact that in the brute force computation, the exact value of the gradient is computed, whereas, in our algorithm, the precision on J and the gradient depends on the time increment used to numerically compute $x(t)$ and $\lambda(t)$. However, our method provides an expression

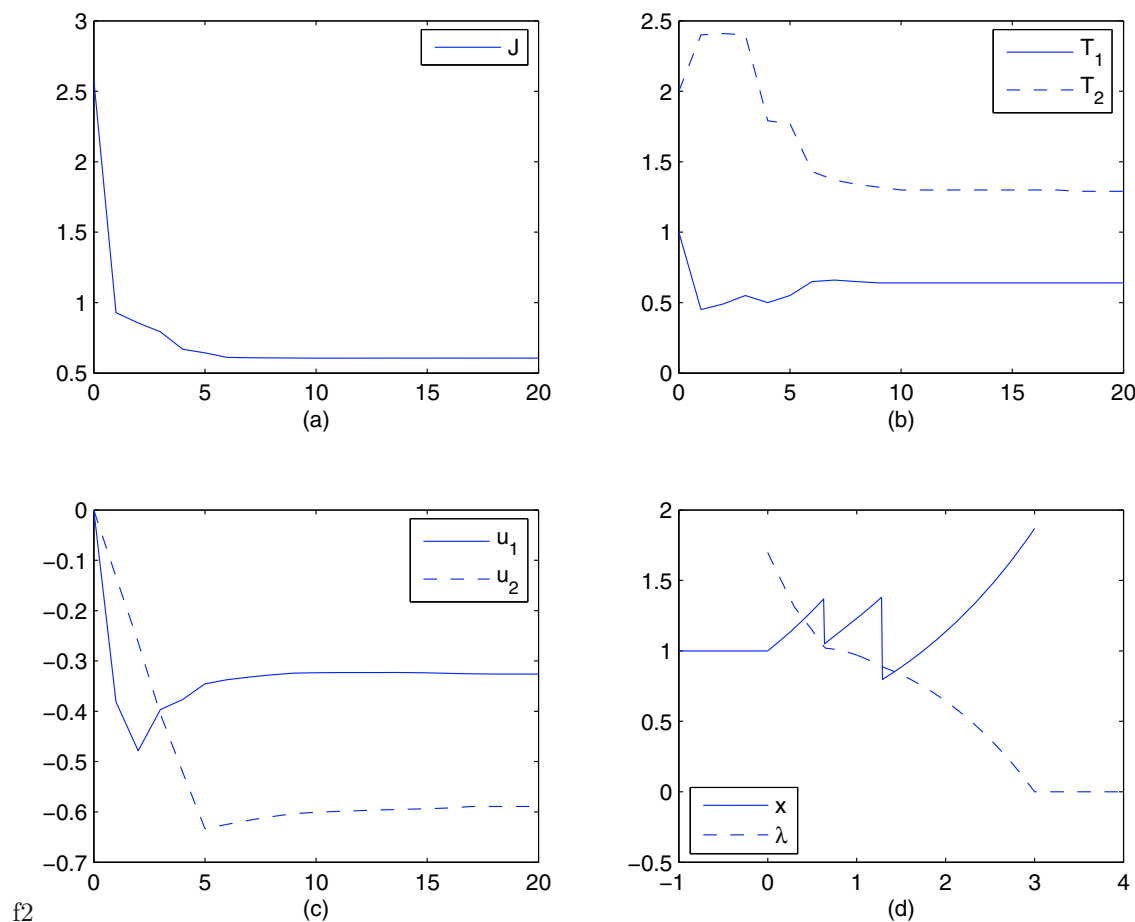


FIGURE 2. Application example: (a) convergence of the performance index J to a minimum; (b) convergence of the switching times T_1 and T_2 ; (c) convergence of the jump amplitudes u_1 and u_2 ; (d) state and costate trajectories of the system at the twentieth iteration of the algorithm.

that can be used for any control (provided that $T_0 < T_1 < T_2 < \dots < T_N$), whereas the brute force computation assumes how the T_i , $T_i + k\tau$ and $k\tau$ ($k \in \mathbb{N}$) are ordered. In this particular example, there are only seven possible orders, but as N and t_N/τ increase, the exponential growth in possible orders prevents the use of a brute force computation.

5. MULTI-DELAY SYSTEMS

In Section 3, optimality conditions were derived only in the case of a single-delay system. Indeed, the authors believe that the inclusion of many delays would have added an unnecessary burden to the already quite technical derivation of the theorem. However, this inclusion does not require more insight and tricks than in the single-delay case, so we show here how the results should be modified to apply in the case of a multi-delay system.

Suppose a finite set of delays $\{\tau_z\}_{z=1}^p$ and a multi-delay switched autonomous system such that

$$\dot{x}(t) = f_i(x(t)) + \sum_{z=1}^p g_{z,i}(x_{\tau_z}(t)), \quad t \in (T_{i-1}, T_i), \quad i = 1, \dots, N \quad (48b)$$

where for $i = 1, \dots, N$, $f_i, g_{1,i}, g_{2,i}, \dots, g_{p,i}$ are smooth functions from \mathbb{R}^n to \mathbb{R}^n . Also suppose that the discrete jumps and cost function are the same as in the previous section. We list here all the modifications that should be made so that the previous theorem now applies to the multi-delay system:

$$H_i = L_i + \lambda_i \left[f_i(x) + \sum_{z=1}^p g_{z,i}(x_{\tau_z}) \right], \quad (52b)$$

$$\dot{\lambda}(t)' = -\frac{\partial L_{\xi(t)}(x(t))}{\partial x} - \lambda(t)' \frac{\partial f_{\xi(t)}(x(t))}{\partial x} - \sum_{z=1}^p \lambda(t + \tau_z)' \frac{\partial g_{z,\xi(t+\tau_z)}(x(t))}{\partial x_{\tau_z}}, \quad (54b)$$

$$\frac{\partial J}{\partial T_i} = \frac{\partial M_i}{\partial T} + H_i(T_i^-) - H_{i+1}(T_i^+) + \sum_{z=1}^p \lambda(T_i + \tau_z)' (g_{z,\xi(T_i+\tau_z)}(x(T_i^-)) - g_{z,\xi(T_i+\tau_z)}(x(T_i^+))) = 0. \quad (58b)$$

6. CONCLUSIONS

We derived necessary conditions for stationarity of the performance index of an impulsively controlled system with prespecified number of pulses. The paper extends the results in [17] to multi-delay systems, without refractory period. To the best of the authors' knowledge, this problem has not been treated before in the literature on optimal control for systems with delays. In fact, a great deal of research is still going on in the delay free case [7,12,15,16]. The problem is also related to the optimal control problem for hybrid and switched systems [5,9,16]. This is a first step in the complete optimal control of such a system, where also the *optimal number* of impulses needs to be found. In principle, all possible N should be optimized for the impulsive controls, and the value of the performance index compared to find the global optimum. Whereas this quickly leads to large number of problems to be solved, regularization methods as for instance presented in [16] could be invoked to obtain a first approximation and thus narrow down the search.

As an illustration, we have set up the optimal control problem for a simple scalar system, where a numerical solution was found using a gradient descent algorithm. A more realistic application can be found in [18,19] where the authors use a similar method to determine a pulse vaccination strategy for controlling the epidemic spread of different diseases. The equations for disease dynamics are inherently nonlinear [1], and the delays enter through incubation periods [8,20]. Other possible applications include traffic control and cancer treatments, two areas where impulsive control and delay systems are commonly found.

REFERENCES

- [1] R.M. Anderson and R.M. May, Directly transmitted infectious diseases: Control by vaccination. *Science* **215** (1982) 1053–1060.
- [2] D.D. Bainov and P.S. Simeonov, *Systems with Impulse Effect: Stability, Theory and Applications*. Ellis Horwood Limited, Chichester, West Sussex (1989).
- [3] D.D. Bainov and P.S. Simeonov, *Impulsive Differential Equations: Periodic Solutions and Applications*, *Pitman Monographs and Surveys in Pure and Applied Mathematics* **66**. Longman Scientific, Harlow (1993).
- [4] D.D. Bainov and P.S. Simeonov, *Impulsive Differential Equations: Asymptotic Properties of the Solutions*, *Series on Advances in Mathematics for Applied Sciences* **28**. World Scientific (1995).
- [5] M.S. Branicky, V.S. Borkar and S.K. Mitter, A unified framework for hybrid control: Model and optimal control theory. *IEEE Trans. Automatic Control* **43** (1998) 31–45.
- [6] A.E. Bryson and Y.C. Ho, *Applied Optimal Control*. Routledge (1975).
- [7] J. Chudoung and C. Beck, The minimum principle for deterministic impulsive control systems, in *Proceedings of the 40th IEEE Conference on Decision and Control* **4**, Orlando, FL (2001) 3569–3574.
- [8] K.L. Cooke and P. van den Driessche, Analysis of an seirs epidemic model with two delays. *J. Math. Biology* **35** (1996) 240–260.

- [9] M. Egerstedt, Y. Wardi and F. Delmotte, Optimal control of switching times in switched dynamical systems, in *Proceedings of the 42nd IEEE Conference on Decision and Control*, Maui, Hawaii (2003) 2138–2143.
- [10] E.G. Gilbert and G.A. Harasty, A class of fixed-time fuel-optimal impulsive control problems and an efficient algorithm for their solution. *IEEE Trans. Automatic Control* **AC-16** (1971) 1–11.
- [11] H.E. Gollwitzer, *Applications of the method of steepest descent to optimal control problems*. Master's thesis, University of Minnesota, USA (1965).
- [12] J.C. Luo and E.B. Lee, Time-optimal control of the swing using impulse control actions, in *Proceedings of the 1998 American Control Conference* **1** (1998) V200–204.
- [13] R. Rishel, Application of an extended Pontryagin principle. *IEEE Trans. Automatic Control* **11** (1966) 167–170.
- [14] G.N. Silva and R.B. Vinter, Optimal impulsive control problems with state constraints, in *Proceedings of the 32nd IEEE Conference on Decision and Control* **4** (1993) 3811–3812.
- [15] H.J. Sussmann, A maximum principle for hybrid optimal control problems, in *Proceedings of the 38th IEEE Conference on Decision and Control* **1** (1999) 425–430.
- [16] E.I. Verriest, Regularization method for optimally switched and impulse systems with biomedical applications, in *Proceedings of the 42nd IEEE Conference on Decision and Control* (2003).
- [17] E.I. Verriest, F. Delmotte and M. Egerstedt, Optimal impulsive control for point delay systems with refractory period, in *IFAC Workshop on Time-Delay Systems*, Leuven, Belgium (2004).
- [18] E.I. Verriest, F. Delmotte and M. Egerstedt, Control of epidemics by vaccination, in *Proceedings of the 2005 American Control Conference* **2** (2005) 985–990.
- [19] E.I. Verriest, F. Delmotte and M. Egerstedt, Control strategies for epidemics by vaccination. *Automatica* (submitted).
- [20] W. Wendi and M. Zhen, Global dynamics of an epidemic model with time delay. *Nonlinear Analysis: Real World Applications archive* **3** (2002) 365–373.
- [21] X. Xu and P. Antsaklis, Optimal control of switched autonomous systems, in *Proceedings of the 41st IEEE Conference on Decision and Control*, Las Vegas, NV (2002) 4401–4406.
- [22] T. Yang, Impulsive control. *IEEE Trans. Automatic Control* **44** (1999) 1081–1083.