Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Optimal Control for Switched Distributed Delay Systems with Refractory Period

Erik I. Verriest
School of ECE
Georgia Institute of Technology
Atlanta, GA 30332-250
erik.verriest@ee.gatech.edu

Abstract— **The optimal switching of a multi-mode distributed time delay system is considered. The class of multi-mode systems consists of systems where the control variables are the switching times in a sequence of fixed vector fields. We assume that the systems considered all have a refractory period. Necessary conditions for a stationary solution are derived and shown to extend those reported for systems with a single or commensurate delays [5]. A method based on excision and regularization to determine the *optimal mode sequence* is presented.**

I. INTRODUCTION

The switching control problem for a finite dimensional multi-mode system involves control actions at discrete instants. The control variables are the switching times between the fixed systems $\dot{x} = f_i(x,t)$, where $i \in \{1,\ldots,m\} = \Xi$. We assume that at the switching times, the state is carried over from one mode to the next in a continuous fashion.

In general, also the particular *sequence* of vector fields needs to be optimized. Such a global control is then parameterizable by the number of switches, $N-1$, a "word" of length N with alphabet, Ξ, and the sequence of switching times $\{T_1,\ldots,T_{N-1}\}$. This global problem involves comparison of $N(N-1)^{m-1}$ individual time optimized strategies, (itself manifestly leading to an obvious combinatorial explosion). Optimal sequencing can also be obtained by the excision method [5].

The objective of this work is to derive the necessary conditions for optimality for modes modelled by functional differential equations. It is assumed that the systems considered have a refractory period, in the sense that once an action is taken, it takes a non-infinitesimal amount of time before a subsequent action can be taken. Refractory periods are ubiquitous, not only in physiological systems but also in many technological systems, (e.g., time required to recharge a capacitor). A refractory time provides a safeguard towards unwanted high frequency switching. The paper extends the results of [5] to systems with *distributed* delays. We believe the derivation of the optimality conditions via a classical variational approach [1] to be somewhat more straightforward than the one carried out in [2] for the finite dimensional case. However, the presence of delays adds a nontrivial twist to the original problem posed in [9]. It should also be noted that the optimal switching problem bears some relation to the optimal impulsive control problem in [6].

Necessary conditions for the optimal switching policy with fixed mode sequence are determined in Section 2. In Section 3, a regularization method is presented to determine the optimal mode sequence as well.

II. VARIATIONAL APPROACH TO OPTIMAL SWITCHING

Consider a distributed delay system with a maximal delay, τ. As usual, x_t denotes the data $\{\, x(t+\theta) \mid -\tau \le \theta \le 0 \,\}$ [3]. First we investigate a *fixed* sequence of vector fields: $f_i(x_t); i = 1,\ldots,N$, satisfying the usual Lipschitz conditions to guarantee well-posedness of the problem. The state space for this multi-mode delay system is $C([-\tau,0],\mathbb{R}^n)$, and the *instants of switching* are the sole control variables. As in the delay free problem, we assume that the entire state x_t is carried over from one mode to the next at the switch, thus preserving the continuity.

The vector $x(t) \in \mathbb{R}^n$ is called the *partial state* at t. Obviously, continuity of x implies continuity of the state x_t. The problem is to minimize the performance index

$$J = \int_0^T L(x,\xi)\, dt + \Phi(x(T)) \qquad (1)$$

for a fixed terminal time T by an optimal choice of the switching times. Here, $\xi(t)$ is a discrete state, taking values in the finite set, Ξ, and denotes the operating mode at time t. If $\xi(t) = a$, then the dynamical system at t is the autonomous system $\dot{x}(t) = f^{(a)}(x_t)$. Denote the nominal switching times by $T_i; i = 1,\ldots,N-1$, and define $T_0 = 0$ and $T_N = T$ (assumed fixed). For simplicity, of notation, set $L(x,\xi(t)) = L_i(x)$ and $f^{(\xi(t))} = f_i$ in the interval (T_{i-1},T_i). The performance index (1) expands to

$$\sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} L_i(x)\, dt + \Phi(x(T)), \text{ with } \dot{x} = f_i(x_t) \qquad (2)$$

for $T_{i-1} \le t \le T_i$. Consider now arbitrary, independent perturbations of the nominal T_i with scale parameter, ϵ, which we will let eventually tend to zero, i.e., $T_i \to T_i + \epsilon\theta_i$. Adjoining the dynamical constraints with *different* Lagrange multipliers, defined in each appropriate subinterval, will not alter the value of J. Assume further that optimal values T_i exist, giving the nominal performance index, $\overline{J}_0$, equal to

$$\sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} [L_i(x) + \lambda_i'(f_i(x_t) - \dot{x})]\, dt + \Phi(x(T)) \qquad (3)$$

It is very important to remark that due to the requisite continuity of the state, a change in T_j say will have an effect on all the modes $i > j$. This happens because the change from T_j to $T_j + \epsilon\theta_j$ (keeping everything else the same) now changes the final (partial) state in the j-th mode from $x(T_j)$ to ($\theta_j > 0$ is assumed, the other case being similar)

$$x(T_j + \epsilon\theta_j) = x(T_j) + \dot{x}(T_j^-)\,\epsilon\theta_j.$$

Note that the *left* derivative is taken with mode j, and this is

$$\dot{x}(T_j^-) = f_j(x_{T_j}).$$

If however, no perturbation were made to T_j, then the value of the partial state $x(T_j + \epsilon\theta_j)$ would have been

$$x(T_j + \epsilon\theta_j) = x(T_j) + \dot{x}(T_j^+)\,\epsilon\theta_j.$$

This gives a difference in the state at the beginning of the $j+1$-st mode of

$$\Delta_{T_j} x = [f_j(x_{T_j}) - f_{j+1}(x_{T_j})]\,\epsilon\theta_j. \tag{4}$$

As each subsequent switch will add such a term, it is clear that the effects of all such perturbations will accumulate in subsequent modes. Keeping track of all these effects will complicate the derivation requiring the explicit computation of perturbations as done in [2]. In keeping with the philosophy of calculus of variations, we shall avoid having to keep track of these by introducing a *sequence* of induced variations, $\{\eta_j\}$, in the same way as we introduced independent Lagrange multipliers λ_j for each mode, i.e., in each subinterval. Equivalently, we may model the induced partial state variation $\eta(t)$ as a possibly *discontinuous* function with discontinuities at the switching times. The same also holds for the costates $\lambda(t)$. These costates can then be *chosen* in a very convenient way in order to avoid computation of the induced variations.

Defining the Hamiltonian functionals,

$$H_i(x_t, \lambda) = L_i(x) + \lambda_i' f_i(x_t), \tag{5}$$

we find for a neighboring solution (with η possibly discontinuous)

$$\begin{aligned}
\overline{J}_\epsilon &= \Phi(x(T) + \epsilon\eta(T)) + \\
&\quad + \sum_{i=1}^N \int_{T_{i-1}+\epsilon\theta_{i-1}}^{T_i+\epsilon\theta_i} [H_i(x_t + \epsilon\eta_t) - \lambda_i'(\dot{x} + \epsilon\dot{\eta})]\,dt \\
&= \Phi(x(T)) + \sum_{i=1}^N \int_{T_{i-1}}^{T_i} [H_i(x) - \lambda_i'\dot{x}]\,dt + \\
&\quad + \epsilon \sum_{i=1}^N \int_{T_{i-1}}^{T_i} [\mathbf{D}_x H_i(x_t, \lambda_i; \eta_t) - \lambda_i'\dot{\eta}]\,dt + \\
&\quad + \epsilon \left.\frac{\partial\Phi}{\partial x}\right|_T \eta(T) + \sum_{i=1}^N \int_{T_i}^{T_i+\epsilon\theta_i} [H_i - \lambda_i'\dot{x}]\,dt + \\
&\quad - \sum_{i=1}^N \int_{T_{i-1}}^{T_{i-1}+\epsilon\theta_{i-1}} [H_i - \lambda_i'\dot{x}]\,dt.
\end{aligned} \tag{6}$$

In the above, $\mathbf{D}_x$ is the Fréchet derivative defined by

$$\mathbf{D}_x H_i(x_t, \lambda, ; \eta_t) = \lim_{\epsilon\to 0} \frac{H_i(x_t + \epsilon\eta_t, \lambda) - H_i(x_t, \lambda)}{\epsilon}. \tag{7}$$

Note that actually $\theta_0 = \theta_N = 0$, since initial and final time were considered fixed.

A. Progressively Distributed Delay Mode Systems

We shall leave the full generality of the problem behind and consider from now on only systems having progressive distributed delay dynamics, by which we mean system modes representable by the functional differential equations

$$\dot{x}(t) = \int_{0_-}^{\tau_+} G_i(x(t-\sigma), \sigma)\,dB_i(\sigma). \tag{8}$$

Here the G_i are smooth differentiable maps $\mathbf{R}^n \times [0, \tau] \mapsto \mathbf{R}^n$ and $B_i(\cdot)$ has bounded variation. The integration boundaries 0_- and τ_+ indicate that possible crisp (or point-) delays for $\sigma = 0$ (the no-delay term) and $\sigma = \tau$ are included. However, the form (8) excludes *correlations* such as the product $x(t)x(t-\tau)$. This class includes the separable mode systems (with crisp point delays) [5]. At this point we remark that the above restriction is only made for notational purposes in what is to follow. Full generality can be achieved since as we shall see it suffices to apply the perturbation to the *linear* variational problem. Thus, the Hamiltonians, $H_i(x_t, \lambda_i)$ and final cost adjoined with the state continuity constraints are respectively defined by

$$L_i(x(t)) + \lambda_i'(t) \left[\int_0^\tau G_i(x(t-\sigma), \sigma)\,dB_i(\sigma)\right] \tag{9}$$

$$\Psi(x, \mu) = \Phi(x) + \sum_{i=1}^N \mu_i'[x(T_i^+) - x(T_i^-)]. \tag{10}$$

B. Delay effect of a single switch

We shall say that a function y is C_k at t_0, if the k-th derivative of y is continuous at t_0, but the $(k+1)$-st is not. Obviously, this implies that the derivatives of order i are all continuous at t_0 for $i \leq k$.

Assume that a single controlled switch occurs at time T, switching from mode i to $i+1$. This makes $\dot{x}$ discontinuous at T. Consequently, x has a 'kink' (i.e., is non-differentiable) at T. From the continuity assumption, x is C_0 at T. But then $\int_0^\tau G(x(t-\sigma), \sigma)dB(\sigma)$ is C_0 at $t = T + \tau$ if $B(\theta) = b\delta(\theta - \tau)$ (see [5]), and C_1 if B is smooth. In turn this implies that $\dot{x}$ is *at least* C_0 at $T + \tau$, inducing again at least C_1 behavior in x at $T + \tau$ and at least C_2 behavior in $\int GdB$ at time $T + \tau$, and so on. We summarize the chain of events (where $\subseteq C_k$ means "at least C_k"):

Lemma 1: *A controlled switch at time T induces the following behavior in the delay system* (8)

	T	$T+\tau$	$T+2\tau$	$\dots$	$T+k\tau$
$\dot{x}(t)$	jump	$\supseteq C_0$	$\supseteq C_1$		$\supseteq C_{k-1}$
$x(t)$	C_0	$\supseteq C_1$	$\supseteq C_2$		$\supseteq C_k$
$\int_0^\tau G(x_t)dB$		$\supseteq C_1$	$\supseteq C_2$		$\supseteq C_k$

C. Effect of variation in a single switching time

Let us first recall a simple result:

Lemma 2: *If y is C_k at t_0, then the variation of y in an interval of length θ about t_0 is of order $k+1$ in θ.*

Consider now for smooth functions G a functional of the form

$$\mathcal{I}(T) = \int_0^{T+k\tau} \int_0^{\tau} q(x(t-\sigma), \sigma)\, dB(\sigma)\, dt \qquad (11)$$

where x satisfies the dynamics (8) with a single switch at time T. Observe (by Lemma 1) that this induces C_0 behavior at T, C_1 behavior at $T+\tau$,...,C_m behavior at $T+m\tau$ in $x(t)$.

Partition the integral as

$$\mathcal{I}(T) = \int_0^{T} y(t)\, dt + \sum_{i=1}^{k} \int_{T+(i-1)\tau}^{T+i\tau} y(t)\, dt.$$

where

$$y(t) = \int_0^{\tau} q(x(t-\sigma), \sigma)\, dB(\sigma).$$

If the switch occurred ϵ time units later than T, then $\mathcal{I}(T)$ becomes $\mathcal{I}(T+\epsilon)$, given by

$$\int_0^{T+\epsilon} y(t)\, dt + \sum_{i=1}^{k} \int_{T+(i-1)\tau+\epsilon}^{T+i\tau+\epsilon} [y(t) + \epsilon\eta_i(t)]\, dt.$$

The induced variation in the integral is $\mathcal{I}(T+\epsilon) - \mathcal{I}(T)$ and (assuming first $\epsilon > 0$) after some regrouping of terrms, reduces to

$$\begin{aligned}
&\int_T^{T+\epsilon} [y_-(t) - y_+(t)]\, dt + \\
&+ \sum_{i=1}^{k-1} \int_{T+i\tau}^{T+i\tau+\epsilon} [y_-(t) + \epsilon\eta_i - y_+(t) - \epsilon\eta_{i+1}]\, dt + \\
&+ \int_{T+k\tau}^{T+k\tau+\epsilon} [y_-(t) + \epsilon\eta_k(t)]\, dt + \\
&+ \sum_{i=1}^{k} \int_{T+(i-1)\tau}^{T+i\tau} \epsilon \mathbf{D}_{\eta_i}[y(t)]\, dt. \qquad (12)
\end{aligned}$$

The subscript $\pm$ reminds us that the integrand needs to be computed on the left $(-)$ and right(+) hand side of the point. The analysis for $\epsilon < 0$ is analogous.

The first integral, $\int_T^{T+\epsilon}[y_-(t) - y_+(t)]\, dt$, is of order 2 in ϵ. Indeed, $y_-(T) = \int_0^\tau q(x(T-\sigma), \sigma) dB(\sigma)$, and involves the dynamics present *before* the switch. Likewise $y_+(T)$ must be computed with the dynamics reigning *after* the switch. The integral near $T+\tau$ contributes a perturbation proportional to at least second degree in ϵ. Likewise, the induced perturbations at $T+i\tau$ are of order at least $i+1$ in ϵ. All these terms will be neglected.

This leaves the $\mathbf{D}_x[y(t)]$ - integrals, where $\mathbf{D}_x$ denotes the Fréchet derivative. Hence

$$\mathbf{D}\mathcal{I}(T) = \sum_{i=1}^{k} \int_{T+(i-1)\tau}^{T+i\tau} \mathbf{D}_x[y(t)]\, dt. \qquad (13)$$

This was the situation if there is only a single switch. If another switch occurs, say at T_1, then for some k, $T+k\tau < T_1 < T+(k+1)\tau$, and the G-term would induce another ϵ perturbation. Therefore the bookkeeping of all perturbation terms will be quite complicated, especially in view of the fact that that all possibilities (of relative positions of switching instants) need to be taken into account.

D. Refractory period

Let us now assume the existence of a refractory period. In particular, we consider the case where this refractory time exceeds the delay time, τ. In this case, the problem greatly simplifies, and the aforementioned complexity disappears, as only two adjacent intervals need to be considered [5]. Indeed, the delayed effect of the k-th switch hits before the $k+1$-st switch. This yields a situation which is akin to the case with a single switch in the previous section.

As in (11), with the Hamiltonian functionals defined in (9), we express the first variation in the performance index as the limit for $\epsilon \to 0$ of

$$\delta J = \lim_{\epsilon \to 0} \frac{\overline{J}_\epsilon - \overline{J}}{\epsilon}.$$

We start with (3) for the progressively distributed mode systems mode form

$$\begin{aligned}
&\sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \Big[L_i(x) + \lambda_i' \Big(\int_{0-}^{\tau+} G_i(x(t{-}\sigma), \sigma)\, dB_i(\sigma) - \dot{x} \Big) \Big] dt \\
&\qquad + \Phi(x(T)) + \sum_{i=1}^{N} \mu_i' [x(T_i^+) - x(T_i^-)], \qquad (14)
\end{aligned}$$

in which we separate the delayed terms

$$\begin{aligned}
\overline{J} \;\; = \;\; & \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} [L_i(x) - \lambda_i' \dot{x}]\, dt \\
& + \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \lambda_i' \Big(\int_{0-}^{\tau+} G_i(x)\, dB_i \Big)\, dt \\
& + \Phi(x(T)) + \sum_{i=1}^{N} \mu_i' [x(T_i^+) - x(T_i^-)] \qquad (15)
\end{aligned}$$

The integrals of the delayed terms in (15) may be rearranged by permuting the order of integration as follows:

$$\begin{aligned}
&\int_{T_{i-1}}^{T_i} \lambda_i'(t) \left(\int_{0-}^{\tau+} G_i(x(t-\sigma),\sigma)\, dB_i(\sigma) \right) dt \\
&= \int_{T_{i-2}}^{T_{i-1}} \chi_{[T_{i-1}-\tau, T_{i-1}]} \left(\int_{T_{i-1}-t}^{\tau} (\lambda_i^\tau)'\, dB_i \right) G_i(x(t))\, dt + \\
&\quad + \int_{T_{i-1}}^{T_i} \chi_{(T_{i-1}, T_i-\tau)} \left(\int_0^{\tau} (\lambda_i^\tau)'\, dB_i \right) G_i(x(t))\, dt + \\
&\quad + \int_{T_{i-1}}^{T_i} \chi_{[T_i-\tau, T_i]} \left(\int_0^{T_i-t} (\lambda_i^\tau)'\, dB_i \right) G_i(x(t))\, dt.
\end{aligned}$$

In these expressions, $\chi_{\mathcal{I}}$ is the indicator function of the interval $\mathcal{I}$, and the advanced term $\lambda(t+\sigma)$ is denoted by λ^τ.

Hence, by rearranging terms the expression (15) reduces to (for simplicity, we also denoted $\chi_{i+1}^+ = \chi_{(T_i, T_{i+1}-\tau)}$ and $\chi_{i+1}^- = \chi_{[T_i-\tau, T_i]}$)

$$\begin{aligned}
\overline{J} = &\sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \left[L_i(x) - \lambda_i'\dot{x} \right] dt + \\
&+ \int_{-\tau}^{0} \left(\int_{-t}^{\tau} (\lambda_1^\tau)' dB_1 \right) G_1(x)\, dt + \\
&+ \sum_{i=1}^{N-1} \int_{T_{i-1}}^{T_i} \chi_{i+1}^- \left(\int_{T_i-t}^{\tau} (\lambda_{i+1}^\tau)' dB_{i+1} \right) G_{i+1}(x)\, dt + \\
&+ \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \chi_i^+ \left(\int_0^{\tau} (\lambda_i^\tau)' dB_i \right) G_i(x)\, dt + \\
&+ \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \chi_{i+1}^- \left(\int_0^{T_i-\tau} (\lambda_i^\tau)' dB_i \right) G_i(x)\, dt + \\
&+ \Phi(x(T)) + \sum_{i=1}^{N} \mu_i' [x(T_i^+) - x(T_i^-)]. \qquad (16)
\end{aligned}$$

By adding $\lambda_{N+1}(t) = 0$ for $t > T_N$, the sum from $i = 0$ to $N-1$, may be changed to the sum from $i = 0$ to N, thus,

$$\begin{aligned}
\overline{J} \;=\; &\int_{-\tau}^{0} \left(\int_{-t}^{\tau} (\lambda_1^\tau)' dB_1 \right) G_1\, dt + \\
&+ \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \left[L_i(x) + M_i(x) - \lambda_i'\dot{x} \right] dt + \\
&+ \Phi(x(T)) + \sum_{i=1}^{N} \mu_i' [x(T_i^+) - x(T_i^-)], \qquad (17)
\end{aligned}$$

where

$$M_i(x) = \left[\chi^+ \mathcal{I}_i^{(1)} + \chi_{i+1}^- (\mathcal{I}_i^{(2)} + \mathcal{I}_i^{(3)}) \right],$$

and

$$\mathcal{I}_i^{(1)}(x) = \left(\int_{0_-}^{\tau_+} (\lambda_i^\tau)' dB_i \right) G_i(x) \qquad (18)$$

$$\mathcal{I}_i^{(2)}(x) = \left(\int_{0}^{T_i-\tau} (\lambda_i^\tau)' dB_i \right) G_i(x) \qquad (19)$$

$$\mathcal{I}_i^{(3)}(x) = \left(\int_{T_i-t}^{\tau} (\lambda_{i+1}^\tau)' dB_{i+1} \right) G_{i+1}(x) \qquad (20)$$

The integrals in (16) involve expressions of the form

$$\overline{\mathcal{K}} = \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \left[K_i(x) - \lambda_i'\dot{x} \right] dt, \qquad (21)$$

for which we now consider *independent* perturbations of the switching times. To that effect introduce the variables $\{\theta_i\}$ and a scale parameter ϵ and the induced perturbations η_i:

$$\begin{aligned}
\overline{\mathcal{K}}_\epsilon &= \sum_{i=1}^{N} \int_{T_{i-1}+\epsilon\theta_{i-1}}^{T_i+\epsilon\theta_i} \left[K_i(x+\epsilon\eta_i) - \lambda_i'(\dot{x}+\epsilon\dot{\eta}_i) \right] dt \\
&= \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \left[K_i(x) - \lambda_i'\dot{x} \right] dt + \\
&\quad + \sum_{i=1}^{N} \left(\int_{T_i}^{T_i+\epsilon\theta_i} \left[K_i(x) - \lambda_i'\dot{x} \right]_- dt + \right. \\
&\qquad\qquad \left. - \int_{T_{i-1}}^{T_{i-1}+\epsilon\theta_{i-1}} \left[K_i(x) - \lambda_i'\dot{x} \right]_+ dt \right) \\
&\quad + \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \left(\frac{\partial K_i}{\partial x} + \dot{\lambda}_i' \right) \epsilon\eta_i\, dt. \\
&\quad + \epsilon \sum_{i=1}^{N} \left[-\lambda_i'(T_i^-)\eta_i(T_i^-) + \lambda_i'(T_{i-1}^+)\eta_i(T_{i-1}^+) \right],
\end{aligned}$$

where as usual, we integrated by parts. Subtracting the nominal value, dividing by ϵ and taking the limit gives the induced variation $\delta\mathcal{K}$

$$\begin{aligned}
\delta\mathcal{K} = &\sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \left(\frac{\partial K_i}{\partial x} + \dot{\lambda}_i'(t) \right) \eta(t)\, dt + \\
&+ \sum_{i=1}^{N} \theta_i [K_i - \lambda_i'\dot{x}]_{T_i^-} - \sum_{i=0}^{N-1} \theta_i [K_{i+1} - \lambda_{i+1}'\dot{x}]_{T_i^+} \\
&- \sum_{i=1}^{N} \lambda_i'(T_i^-)\eta_i(T_i^-) + \sum_{i=0}^{N-1} \lambda_{i+1}'(T_i^+)\eta_{i+1}(T_i^+) \qquad (22)
\end{aligned}$$

Note that for $i = 0$, $\theta_0 = 0$ and for $i = N$, $\theta_N = 0$, since initial and final time were fixed.

The $\{\theta_i\}$-induced perturbation of the non-integral term in (16), $\Psi(x,\mu) = \Phi(x) + \sum \mu_i' \Delta_{T_i} x$, follows from (4),

$$\begin{aligned}
\delta\Psi = \frac{\partial\Phi}{\partial x}\eta_N(T) + \sum_{i=1}^{N-1} \mu_i' \big[&(\dot{x}(T_i^-) - \dot{x}(T_i^+))\theta_i + \\
&+ \eta_i(T_i^-) - \eta_{i+1}(T_i^+) \big]. \qquad (23)
\end{aligned}$$

Combine the two perturbations $\delta\mathcal{K}$ and $\delta\Psi$ and choose λ_i in the intervals $[T_{i-1}, T_i]$ to solve

$$\dot{\lambda}_i = -\left(\frac{\partial K_i}{\partial x}\right)'. \tag{24}$$

This yields $\delta\mathcal{K} + \delta\Psi$ in the form

$$\sum_{i=1}^{N-1} \left[A_i\theta_i + B_i'\eta(T_i^+) + C_i'\eta(T_i^-)\right] + \tag{25}$$

$$+\lambda_1(0^+)\eta_1(0^+) + \left(\frac{\partial \Phi}{\partial x} - \lambda_N'(T_N^-)\right)\eta(T_N^-),$$

where

$$\begin{aligned} A_i &= L_i(x(T_i^-)) - L_{i+1}(x(T_i^+)) + \\ &\quad +\mu_i'\left[\dot{x}(T_i^-) - \dot{x}(T_i^+)\right] && (26)\\ B_i &= -\mu_i + \lambda_{i+1}(T_i^+) && (27)\\ C_i &= \mu_i - \lambda_i(T_i^-) && (28) \end{aligned}$$

With the initial data given $\eta(0^+)$ must be zero. Computation of the requisite perturbations $\{\eta_i\}$ is avoided if one *chooses*

$$\lambda_i(T_i^-) = \mu_i = \lambda_{i+1}(T_i^+), \tag{29}$$

with final condition

$$\lambda_N(T_N^-) = \left(\frac{\partial \Phi}{\partial x}\right)^T, \tag{30}$$

thus specifying the boundary conditions for the differential equations (24). This implies that we can choose the *costates* λ_i in $[T_{i-1}, T_i]$ to concatenate to a *continuous* functions in $[0, T]$.

It follows that the first order variation of J reduces to

$$\delta J = \sum_{i=1}^{N-1} A_i\theta_i. \tag{31}$$

E. Main Result

The above derivations are now put together to obtain necessary conditions for the general progressively distributed delay system Since the θ_i are independent, necessary conditions for optimality are the vanishing of the A_i in (31). In view of the choice (29) of the. multipliers μ_i and boundary conditions, it gives for $i = 1$ to $N-1$

$$H_i(x_{T_i}, \lambda(T_i)) = H_{i+1}(x_{T_i}, \lambda(T_i)). \tag{32}$$

Simply stated, it means the continuity of the Hamiltonian functional H at the switching times.

In formulating our main theorem below, we will assume that the vector fields $G_i(x(t-\sigma), \sigma)$ as well as the functions $L_i(x)$ are smooth, and we let $N-1$ be the total number of switches, with $T_0 = 0$ and $T_N = t_f$ being fixed.

Theorem 3:
The separable mode switched system in equation (8), with fixed mode sequence, minimizes the performance index J in (1) if the switching times T_i are chosen as follows:
Euler-Lagrange Equations:

$$\begin{aligned} \dot{\lambda}_i &= -\left(\frac{\partial L_i}{\partial x}\right)^T - \left(\frac{\partial G_i}{\partial x}\right)^T \left(\chi_i^+ \mathcal{I}_1^T + \chi_{i+1}^- \mathcal{I}_2^T\right) + \\ &\quad - \left(\frac{\partial G_{i+1}}{\partial x}\right)^T \chi_{i+1}^- \mathcal{I}_3^T && (33) \end{aligned}$$

where the $\mathcal{I}_1$ and $\mathcal{I}_2$, defined in (20) depend on λ_i^τ,and $\mathcal{I}_3$ on λ_{i+1}^τ. Also, $T_{i-1} < t < T_i$, $i = 1, \ldots, N-1$, with $\chi_i^+(t) = 1$ if $t \in [T_{i-1}, T_i - \tau]$ and 0 otherwise, $\chi_{i+1}^-(t) = 1$ if $t \in [T_i - \tau, T_i]$ and 0 otherwise, and $\lambda_i^\tau = \lambda_i(t+\tau)$. Moreover,

$$\dot{\lambda}_N = -\left(\frac{\partial L_N}{\partial x}\right)^T - \left(\frac{\partial G_N}{\partial x}\right)^T \left(\chi_i^+ \mathcal{I}_1^T + \chi_{i+1}^- \mathcal{I}_2^T\right) \tag{34}$$

Boundary Conditions:

$$\begin{aligned} \lambda_N(T_N) &= \left(\frac{\partial \Phi}{\partial x}\right)^T && (35)\\ \lambda_i(T_i^-) &= \lambda_{i+1}(T_i^+) && (36) \end{aligned}$$

Optimality Conditions:

$$H_i(x_{T_i}, \lambda_i(T_i)) = H_{i+1}(x_{T_i}, \lambda_{i+1}(T_i)). \tag{37}$$

Proof: All one has to do is to recall what the function K was and realize that both indicator functions in its definition evaluate to 1 at the switching point T_i. ■

F. Point delay systems

The separable mode case discussed in [5] is retrieved by taking the $B_i(\theta)$ to be step functions. For systems with a single delay, let

$$\int_{0-}^{\tau^+} G_i(x(t-\sigma), \sigma)\, dB_i(\sigma) = f_i(x(t) + g_i(x(t-\tau)). \tag{38}$$

Systems with multiple delays are treated by multi-level step functions B_i, whereas a system with commensurate delays can be recast as a system with a simple delay using state augmentation (see [5]).

III. Optimal Sequencing

So far, it was assumed that the sequence of modes was fixed. If m modes are available, then with $N-1$ switches, $N(N-1)^{m-1}$ possible mode sequences exist, and their optimized performance needs to be evaluated and compared in order to find the global optimal switched control. Instead we refine an alternative method, first presented in [5].

A. Excision

In the absence of a refractory period, consider the fixed mode sequence by cycling the modes: $1 \rightarrow 2 \rightarrow \ldots \rightarrow m \rightarrow 1 \rightarrow 2 \rightarrow \ldots \rightarrow m \rightarrow 1 \rightarrow 2 \rightarrow \ldots$, and optimize their switching times. If it is found that the optimal switching time sequence has $T_{i-1} = T_i$ for some i, it means that mode i only occurs for duration 0, and therefore should be excised. Performing all such excisions will leave the *optimal mode sequence* associated with *at most* $N-1$ switches. Thus by formulating first the optimal control problem for a fixed mode sequence, no generality is lost in the global optimal control problem if an upper bound on the number of switches is imposed.

B. Regularization of the refractory period

With a refractory period, the time between switches is bounded below and the excision method is not applicable. There is however a way around this. First the problem in Section 2 can be generalized by adding a switching cost term to the performance index. Thus let (1) be replaced by

$$J = \int_0^T L(x, \xi)\, dt + \sum_{i=1}^{N} \Phi_i(x(T_i), \{T_j\}_{j=1}^N) \qquad (39)$$

which leads to a more general set of boundary and optimality conditions in Theorem 3 (see [7] for the case of impulsive control.) Consider a switching cost function of the form

$$\Phi_i((x(T_i), \{T_j\}_{j=1}^N) = \phi(T_i - T_{i-1}).$$

The existence of a true refractory period, τ, is then equivalent to $\phi = \phi_r$ with $\phi_r(\theta) = 0$ for $\theta > \tau$, and $\phi_r(\theta) = \infty$ for $\theta < \tau$. It follows that one can regularize the refractory period phenomenon by considering instead of the above ϕ a more general smooth cost function: Examples are (with $\phi_0 > 0$ and $\omega > 0$) $\phi_b(t) = \phi_0[1 + \omega t^2]^{-1}$ (butterworth), $\phi_e(t) = \phi_0 \exp(-\omega t)$ (exponential), $\phi_f(t) = \phi_0[1+\exp(\omega t)]$ (fermi), $\phi_g(t) = \phi_0 \exp(-t^2/2\omega^2)$ (gaussian), $\phi_i(t) = \phi_0 t^{-\omega}$ (inverse). Note that only the latter has a n infinite cost associated with $\theta = 0$. More generally, let $\psi_n(\theta) > 0$ be such that $\lim_{n\to\infty} \psi_n = \phi_r$, the relaxed optimal switching control $\{T_i^{(n)}\}_{i=1}^N$ is expected to approach the optimal sequence $\{T_i\}_{i=1}^N$ for the control with the refractory period constraint. If for a given n, one or more of the constraints $T_i^{(n)} - T_{i-1}^{(n)} > \tau$ do not hold, increase n and start over.

Once this is understood, it is clear that a slight variation from this scheme can actually allow the case $T_i - T_{i-1} = 0$, so that the excision method can be used to determine the *optimal mode sequence* as well. All one has to do is to adapt the sequence of regularizing cost functions to smooth functions $\overline{\psi}_n > 0$ with a notch near zero, e.g., let $\overline{\psi}_n(t) < \epsilon_1/n$ for $t < \epsilon_2$/n, $\overline{\psi}_n < \epsilon_3/n$ for $t > \tau + \epsilon_4/n$ and $\overline{\psi}_n > n/\epsilon_5$ if $\epsilon_6/n < t < \tau - \epsilon_7/n$. Now the optimal interval between switchings will either fall to the right of the hump (i.e., exceed the refractory period), or to the left of it, nudging it closer and closer to zero as n increases. In this case, it means that the corresponding mode in the interval $(T_{i-1}^{(n)}, T_i^{(n)})$ should be excised. Such a regularization lets one solve for the optimal sequence with refractory period. There remains one caveat: The refractory period may have been brought in to avoid high frequency switching. With the above regularization this safeguard would be lost. A remedy is to associate a *nonzero* but small cost, γ, with a switch close to zero. Note that one never should excise more than m adjacent intervals (since then the mode has returned to the one before the first switch). Hence, excising a (redundant) cycle $1 \rightarrow \ldots \rightarrow m$ is associated with a cost $m\gamma$, whereas sustaining the present mode incurs no additional cost in the same (infinitesimal) interval.

IV. Conclusions

We derived necessary conditions for stationarity of the performance index of a multi-mode distributed delay system controlled by switchings between a prespecified mode sequence. This is a first step in the complete optimal control of a multi-mode system, where also the *optimal sequence* of the modes needs to be found. To avoid a combinatorial search, we proposed an adaptation of the excision method, first presented in [5] for systems without refractory period constraints, via regularization of an equivalent cost function. Alternatively, a regularization method as for instance presented in [4] could be invoked to obtain a first approximation.

V. REFERENCES

[1] Bryson, A.E., and Y.C. Ho, (1975). *Applied Optimal Control*. Hemisphere.

[2] Egerstedt, M., Y. Wardi, and F. Delmotte, (2003). Optimal Control of Switching Times in Switched Dynamical Systems. *Proc. 42nd Conference on Decision and Control*, Maui, HI, 2138-2143.

[3] Hale, J.K. and S. M. Verduyn Lunel (1993). *Introduction to functional differential equations*, Applied Math. Sciences Series, **99**, Springer-Verlag, New York.

[4] Verriest, E.I., (2003). Regularization Method for Optimally Switched and Impulsive Systems with Biomedical Applications. *Proceedings of the 42th IEEE Conference on Decision and Control*, Maui, HI, 2156-2161.

[5] Verriest, E.I. and F. Delmotte, (2005). Optimal Control for Switched Point Delay Systems with Refractory Period. *Proceedings of the 16th IFAC World Congress,*, Prague, CZ, July 4-8, 2005.

[6] Verriest, E.I., F. Delmotte and M. Egerstedt, (2004). Optimal Impulsive Control for Point Delay Systems with Refractory Period. *Preprints of the 5th IFAC Workshop on Time Delay Systems*, Leuven, Belgium, September 2004.

[7] Verriest, E.I., F. Delmotte and M. Egerstedt, (2004). Optimal Impulsive Control for Point Delay Systems with Refractory Period. *Preprints of the IFAC Workshop on Time Delay Systems,* TDS-04, Leuven, Belgium, September 2004.

[8] Wardi, Y., Boccador, M. Egerstedt, M. and Verriest, E.I., (2004). Optimal Control of Switching Surfaces. *Proceedings of the 43-d IEEE Conference on Decision and Control*, pp. 1854-1859, Paradise Island, Bahamas, Dec. 14-17, 2004.

[9] Xu, X. and P. Antsaklis, (2002). Optimal Control of Switched Autonomous Systems. *Proc. 41st Conference on Decision and Control*, Las Vegas, NV, 4401-4406.