CHALMERS | UNIVERSITY OF GOTHENBURG

PREPRINT 2009:10

Continuous-discrete optimal control problems

KJELL HOLMÅKER

Department of Mathematical Sciences
Division of Mathematics
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Göteborg Sweden 2009

Preprint 2009:10

Continuous-discrete optimal control problems

Kjell Holmåker

Department of Mathematical Sciences
Division of Mathematics
Chalmers University of Technology and University of Gothenburg
SE-412 96 Göteborg, Sweden
Göteborg, February 2009

Preprint 2009:10
ISSN 1652-9715

Matematiska vetenskaper
Göteborg 2009

Continuous-discrete optimal control problems

Kjell Holmåker

Abstract

A continuous-discrete control system is studied where the state variable x may have a jump discontinuity at certain times τ_k. The size of the jump is determined by the actual value of x and the choice of a control parameter z_k. Between the jump times, x satisfies an equation $\dot{x} = f(t, x, u(t))$ for some choice of the control function u. There may also be constraints on the values $x(\tau_k)$ and the jumps. A cost functional depending on these quantities is to be minimized. Necessary conditions for optimality are derived. This is done by formulating the problem as a special case of a general mathematical programming problem in an infinite-dimensional space and applying a general multiplier rule. The times τ_k may be fixed or allowed to vary. As an application necessary conditions are obtained for a multiprocess problem, i.e., a problem where at τ_k there is a transition to a different state space.

1 Introduction

Differential equations with impulse effects can be used to model situations where some quantities can change instantly. In a typical case a state variable x evolves in time according to an equation $\dot{x} = f(t, x)$ except at times τ_k, where it makes a jump of size $J_k(x(\tau_k))$. A solution that exists on an interval $[T_0, T_1]$ with jumps at interior points $\tau_k,\ k = 1, \ldots, r$, with

$$T_0 = \tau_0 < \tau_1 < \cdots < \tau_r < \tau_{r+1} = T_1,$$

is assumed to be continuous from the left at τ_k for $k = 1, \ldots, r+1$ and continuous from the right at T_0. On $(\tau_k, \tau_{k+1}]$ $(0 \le k \le r)$ it is absolutely continuous and satisfies $\dot{x}(t) = f(t, x(t))$ a.e. with $x(\tau_k^+) = x(\tau_k + 0) = x(\tau_k) + J_k(x(\tau_k))$ for $1 \le k \le r$. We write $\Delta x|_{\tau_k} = x(\tau_k^+) - x(\tau_k) = J_k(x(\tau_k))$ and say that x satisfies

$$\begin{aligned} \dot{x} &= f(t, x), \quad t \ne \tau_k, \\ \Delta x|_{\tau_k} &= J_k(x(\tau_k)). \end{aligned}$$

Equations of this type are treated for example in [1].

In this paper we will consider continuous-discrete control systems

$$\begin{aligned} \dot{x} &= f(t, x, u(t)), \quad t \ne \tau_k, \\ \Delta x|_{\tau_k} &= J_k(x(\tau_k), z_k), \end{aligned}$$

with control functions $u(t) \in \Omega(t)$ and control parameters $z_k \in Z_k$. The set $J_k(x(\tau_k), Z_k)$ of allowed jumps is assumed to be convex. We have a number of constraints for the values $x(\tau_k)$ and the jumps $J_k(x(\tau_k), z_k)$, and we want to minimize a functional depending on these quantities. In Section 5 we derive necessary conditions for optimality, first when the switching times τ_k are fixed, and then when they are allowed to vary. The first result we obtain by formulating the problem as a special case of a general optimization problem and applying the necessary conditions for this problem. This general problem is discussed in Section 3. To be able to apply this we need to know how the solution $x(t)$ is affected by certain variations of $u(t)$ and z_k. In Section 4 a known result is extended to systems with impulse effects. We also need a solution formula for the linear case which is given in Section 2. In the case where τ_k may vary, a certain transformation of the time variable will reduce the problem to one with fixed switching times, and the previously obtained results can be applied.

Our results contain as special cases the Pontryagin maximum principle for problems with interior constraints ($J_k = 0$) and the discrete maximum principle as it appears in [13] ($f = 0$). In a related type of problems there are different descriptions (including different state spaces) on different time intervals. Then we cannot talk about jumps, but there can be equations relating the values at the endpoints of the time intervals. Such problems – sometimes called multiprocess problems – are discussed in Section 6, where it is shown how the results from Section 5 can be used to derive necessary conditions.

It seems that this particular combination of an ordinary control and discrete controls has not been studied much. Some authors (e.g. [2], [12] and [14]) have considered problems with only discrete controls. In [14] there is the same convexity assumption as in this paper, and their result is included in our Theorem 4.

2 Linear systems with impulse effects

Consider the following linear system in $\mathbb{R}^n$ with impulse effects:

$$\dot{x} = A(t)x + b(t), \quad t \neq \tau_k, \tag{2.1}$$

$$\Delta x|_{\tau_k} = x(\tau_k^+) - x(\tau_k) = C_k x(\tau_k) + d_k, \tag{2.2}$$

$$x(T_0) = x_0. \tag{2.3}$$

Here the elements of the $n \times n$-matrix $A(t)$ and the n-vector $b(t)$ belong to $L^1(T_0, T_1)$, and C_k is an $n \times n$-matrix.

We will derive an expression for the solution (equation (2.7) below). Let $\Phi(t)$ be the fundamental matrix at T_0 of the system $\dot{x} = A(t)x$ (i.e., $\dot{\Phi}(t) = A(t)\Phi(t)$ a.e., $\Phi(T_0) = E$, E is the $n \times n$ identity matrix). Define

$$\Psi_k = \Phi^{-1}(\tau_k)(E + C_k)\Phi(\tau_k), \quad k = 1, \dots, r, \tag{2.4}$$

$$\Psi_{kj} = \begin{cases} \Psi_k \cdots \Psi_{j+1} & \text{if } 0 \le j < k \le r, \\ E & \text{if } 0 \le j = k \le r. \end{cases}$$

For $T_0 \le s \le t \le T_1$ we define a piecewise constant function $\Psi(t, s)$:

$$\Psi(t, s) = \Psi_{00} = E, \text{ if } T_0 \le s \le t \le \tau_1;$$

if $\tau_k < t \le \tau_{k+1}$ ($k = 1, \dots, r$), then

$$\Psi(t, s) = \begin{cases} \Psi_{k0} & \text{if } T_0 \le s \le \tau_1, \\ \Psi_{k1} & \text{if } \tau_1 < s \le \tau_2, \\ \quad \vdots & \\ \Psi_{k,k-1} = \Psi_k & \text{if } \tau_{k-1} < s \le \tau_k, \\ \Psi_{kk} = E & \text{if } \tau_k < s \le t. \end{cases}$$

Note that

$$\Psi(t, s) = \begin{cases} \Psi_{kj} = \Psi_k \Psi(\tau_k, s) & \text{if } \tau_j < s \le \tau_{j+1}, \ \tau_k < t \le \tau_{k+1}, \ 0 \le j < k \le r, \\ \Psi_{kk} = E & \text{if } \tau_k < s \le t \le \tau_{k+1}, \ 0 \le k \le r, \end{cases} \tag{2.5}$$

$$\Psi(t, T_0) = \begin{cases} \Psi_{k0} = \Psi_k \Psi(\tau_k, T_0) & \text{if } \tau_k < t \le \tau_{k+1}, \ 1 \le k \le r, \\ \Psi_{00} = E & \text{if } T_0 \le t \le \tau_1. \end{cases} \tag{2.6}$$

Then the solution of (2.1)–(2.3) can be written

$$x(t) = \Phi(t)\Psi(t, T_0)x_0 + \int_{T_0}^{t} \Phi(t)\Psi(t, s)\Phi^{-1}(s)b(s)\, ds + \sum_{T_0 < \tau_j < t} \Phi(t)\Psi(t, \tau_j^+)\Phi^{-1}(\tau_j)d_j \tag{2.7}$$

for all $t \in [T_0, T_1]$. To see this, let $y(t) = \Phi^{-1}(t)x(t)$. On (τ_k, τ_{k+1}) $(0 \le k \le r)$ we have $\dot{y}(t) = \Phi^{-1}(t)b(t)$ a.e. Further, $y(T_0) = x_0$ and (using (2.4))

$$\begin{aligned} y(\tau_k^+) &= \Phi^{-1}(\tau_k)x(\tau_k^+) = \Phi^{-1}(\tau_k)[(E + C_k)x_k(\tau_k) + d_k] \\ &= \Phi^{-1}(\tau_k)(E + C_k)\Phi(\tau_k)y(\tau_k) + \Phi^{-1}(\tau_k)d_k \\ &= \Psi_k y(\tau_k) + \Phi^{-1}(\tau_k)d_k \quad \text{for } k \ge 1. \end{aligned} \tag{2.8}$$

For $T_0 \le t \le \tau_1$ we have (see (2.6))

$$y(t) = x_0 + \int_{T_0}^{t} \Phi^{-1}(s)b(s)\,ds = \Psi(t, T_0)x_0 + \int_{T_0}^{t} \Psi(t, s)\Phi^{-1}(s)b(s)\,ds.$$

For $\tau_1 < t \le \tau_2$ we then obtain from (2.8) using (2.5)–(2.6)

$$\begin{aligned} y(t) &= y(\tau_1^+) + \int_{\tau_1}^{t} \Phi^{-1}(s)b(s)\,ds = \Psi_1 y(\tau_1) + \Phi^{-1}(\tau_1)d_1 + \int_{\tau_1}^{t} \Phi^{-1}(s)b(s)\,ds \\ &= \Psi(t, T_0)x_0 + \Psi_1 \int_{T_0}^{\tau_1} \Psi(\tau_1, s)\Phi^{-1}(s)b(s)\,ds + \Phi^{-1}(\tau_1)d_1 + \int_{\tau_1}^{t} \Phi^{-1}(s)b(s)\,ds \\ &= \Psi(t, T_0)x_0 + \int_{T_0}^{t} \Psi(t, s)\Phi^{-1}(s)b(s)\,ds + \Psi(t, \tau_1^+)\Phi^{-1}(\tau_1)d_1. \end{aligned}$$

By induction we get in a similar way

$$y(t) = \Psi(t, T_0)x_0 + \int_{T_0}^{t} \Psi(t, s)\Phi^{-1}(s)b(s)\,ds + \sum_{j=1}^{k} \Psi(t, \tau_j^+)\Phi^{-1}(\tau_j)d_j$$

for $\tau_k < t \le \tau_{k+1}$, $1 \le k \le r$. This proves (2.7).

See [1] for further properties of systems of the form (2.1)–(2.3).

3 A general extremal problem

Necessary conditions for many optimization problems, including optimal control problems, can be derived by applying the necessary conditions for a generally formulated extremal problem. A simple version is the following mathematical programming problem in an infinite-dimensional space.

Problem ($\mathcal{P}$). Let X be a real linear space, and let there be given real-valued functions $\varphi_{-q}, \ldots, \varphi_0, \ldots, \varphi_p$ (p and q are non-negative integers) defined on X, and a subset A of X. Minimize $\varphi_0(x)$ subject to the constraints

$$\varphi_i(x) = 0, \quad i = 1, \ldots, p, \tag{3.1a}$$

$$\varphi_i(x) \le 0, \quad i = -q, \ldots, -1, \tag{3.1b}$$

$$x \in A. \tag{3.1c}$$

Assume that x_0 is a solution of Problem ($\mathcal{P}$), that is, x_0 satisfies (3.1) and $\varphi_0(x_0) \le \varphi_0(x)$ for all x such that (3.1) is satisfied. Under suitable conditions it is possible to derive a necessary condition for optimality in the form of a generalized multiplier rule. What we need is some differentiability properties of φ_i at x_0 and some way of approximating A near x_0. When the multiplier rule is applied to an optimal control problem this approximation of A comes from a certain perturbation result in the theory of differential equations. We make the following assumption:

Assumption 1. There exist a set $M \subseteq X$ and functions $h_i \colon X \to \mathbb{R}$, $-q \le i \le p$, with the following properties: For every finite collection $y_1, \ldots, y_N$ of points in M there is an $\varepsilon_0 > 0$ such that for each $\varepsilon \in (0, \varepsilon_0]$ and each $\beta \in S_N$, where

$$S_N = \{\beta = (\beta_1, \ldots, \beta_N) \in \mathbb{R}^N : \beta_j \ge 0 \text{ for } j = 1, \ldots, N,\ \sum_{j=1}^{N} \beta_j = 1\}, \tag{3.2}$$

there exists a point $\rho_{\beta,\varepsilon} \in A$ such that, with $y_\beta = \sum_{j=1}^{N} \beta_j y_j \in \operatorname{co} M$,

(i) $\varphi_i(\rho_{\beta,\varepsilon})$ is continuous with respect to $\beta \in S_N$ for $1 \le i \le p$;

(ii) for $1 \le i \le p$, h_i is linear, and

$$\lim_{\varepsilon \to 0^+} \frac{\varphi_i(\rho_{\beta,\varepsilon}) - \varphi_i(x_0)}{\varepsilon} = h_i(y_\beta);$$

(iii) for $-q \le i \le 0$, h_i is convex, and

$$\limsup_{\varepsilon \to 0^+} \frac{\varphi_i(\rho_{\beta,\varepsilon}) - \varphi_i(x_0)}{\varepsilon} \le h_i(y_\beta);$$

and the convergence in (ii) and (iii) is uniform with respect to $\beta \in S_N$.

Remark. If X is normed, then the function h_i in (ii) might be the Hadamard derivative, i.e.,

$$\lim_{\substack{\varepsilon \to 0^+ \\ z \to y}} \frac{\varphi_i(x_0 + \varepsilon z) - \varphi_i(x_0)}{\varepsilon} = h_i(y),$$

and similarly in (iii). In this case $\rho_{\beta,\varepsilon}$ should be a point in A such that $\rho_{\beta,\varepsilon} = x_0 + \varepsilon y_\beta + o(\varepsilon)$. The set M may be considered as a set of directions in which it is possible to approximate A near x_0. However, we only need the properties of the composite functions $\varphi_i(\rho_{\beta,\varepsilon})$. It may also be easier to verify (i), (ii) and (iii) directly than treating $\varphi_i(x)$ and $\rho_{\beta,\varepsilon}$ separately.

Theorem 1. *Let x_0 be a solution of Problem $(\mathcal{P})$ and let Assumption 1 be satisfied. Then there exist real numbers $\lambda_{-q}, \dots, \lambda_p$, not all zero, such that*

$$\sum_{i=-q}^{p} \lambda_i h_i(y) \le 0 \quad \text{for all } y \in \operatorname{co} M, \tag{3.3}$$

$$\lambda_i \le 0, \quad -q \le i \le 0, \tag{3.4}$$

$$\lambda_i \varphi_i(x_0) = 0, \quad -q \le i \le -1. \tag{3.5}$$

Proof of this theorem can be found in [10]; see also [7] and [9].

4 A perturbation result for differential equations

We are going to study control systems $\dot{x} = f(t, x, u(t))$ on a fixed time interval $I = [T_0, T_1]$. For the admissible controls $u(\cdot)$ there is a constraint of the form $u(t) \in \Omega(t)$ for all $t \in I$. We say that a function $(t, x) \mapsto F(t, x)$ with values in $\mathbb{R}^r$ for some r, defined for $t \in I$ and $x \in \mathbb{R}^n$ is of *Carathéodory-type* (on I) if it is measurable in t and continuous in x, and if for each compact set $K \subset \mathbb{R}^n$ there exists a function $\rho \in L^1(I)$ such that $|F(t, x)| \le \rho(t)$ for all $t \in I$, $x \in K$. We say that F belongs to the class $\mathcal{F}$ if F is differentiable with respect to x, and F and $\frac{\partial F}{\partial x}$ are of Carathéodory-type. [$\frac{\partial F}{\partial x}$ is the $r \times n$-matrix with elements $\frac{\partial F_i}{\partial x_j}$.] For $f(t, x, u)$ and $\Omega(t)$ we make the following assumption.

Assumption 2.

(i) f is a mapping from $I \times \mathbb{R}^n \times \mathbb{R}^m$ to $\mathbb{R}^n$. For each $t \in I$, $\Omega(t)$ is a non-empty subset of $\mathbb{R}^m$.

(ii) f is differentiable with respect to x, and f and $\frac{\partial f}{\partial x}$ are measurable in t (for fixed x and u) and continuous in x and u separately (for fixed t).

(iii) The set of admissible controls $u(\cdot)$ is

$$\mathcal{U} = \{u(\cdot) : u(\cdot) \text{ is measurable on } I,\ u(t) \in \Omega(t) \text{ for all } t \in I, \text{the function } (t, x) \mapsto f(t, x, u(t)) \text{ belongs to the class } \mathcal{F}\}.$$

(iv) There exists a countable family $\{u_j\}_{j=1}^{\infty}$ of functions $u_j \in \mathcal{U}$ such that the set $\{u_j(t)\}_{j=1}^{\infty}$ is dense in $\Omega(t)$ for every $t \in I$. [It can be shown that this is the case, for example, if the set-valued mapping $t \mapsto \Omega(t)$ is measurable in the sense that the set $\{(t,u) \in \mathbb{R}^{m+1} : t \in I, u \in \Omega(t)\}$ belongs to the σ-algebra in $\mathbb{R}^{m+1}$ that is generated by the Lebesgue sets in I and the Borel sets in $\mathbb{R}^m$ (assuming that $\mathcal{U} \neq \emptyset$ and that $(t,x) \mapsto f(t,x,v(t))$ belongs to $\mathcal{F}$ for bounded measurable functions v); see [11].]

For applications to optimal control the following result about certain perturbations of a given control u_0 is of central importance.

Theorem 2. *Assume that $u_0 \in \mathcal{U}$ and that x_0 is a piecewise continuous function on $[T_0, T_1]$ that is a solution of $\dot{x} = f(t,x,u_0(t))$ on a subinterval $I' = [\tau, \tau'] \subseteq I$. Let for $\xi \in \mathbb{R}^n$ and $u \in \mathcal{U}$, $v(t;\xi,u)$ be the solution of*

$$\begin{cases} \dot{v} = \dfrac{\partial f}{\partial x}(t, x_0(t), u_0(t))v + f(t, x_0(t), u(t)) - f(t, x_0(t), u_0(t)), \\ v(\tau) = \xi. \end{cases} \tag{4.1}$$

Let $\xi_j \in \mathbb{R}^n$ and $u_j \in \mathcal{U}$, $j = 1, \ldots, N$, be given. For each $\beta \in S_N$ (see (3.2)) and each $\varepsilon \in (0,1)$ there exist pairwise disjoint sets $A_j = A_j(\beta,\varepsilon) \subseteq I$, $j = 0, 1, \ldots, N$, each a finite union of intervals, such that $\cup_{j=0}^{N} A_j = I$ and such that the following is true. Define the control $u_{\beta,\varepsilon} \in \mathcal{U}$ by

$$u_{\beta,\varepsilon}(t) = u_j(t) \quad \text{for } t \in A_j(\beta,\varepsilon),\ j = 0, 1, \ldots, N,$$

and let $x_{\beta,\varepsilon}$ be the solution of

$$\dot{x} = f(t, x, u_{\beta,\varepsilon}(t)), \qquad x(\tau) = x_0(\tau) + \varepsilon \sum_{j=1}^{N} \beta_j \xi_j.$$

Then there exists an $\varepsilon_0 \in (0,1)$ such that for all $\beta \in S_N$ and all $\varepsilon \in (0, \varepsilon_0]$, $x_{\beta,\varepsilon}(\cdot)$ exists on all I' and satisfies

$$x_{\beta,\varepsilon}(t) = x_0(t) + \varepsilon \sum_{j=1}^{N} \beta_j v(t;\xi_j,u_j) + r(t;\beta,\varepsilon), \tag{4.2}$$

where $r(t;\beta,\varepsilon)/\varepsilon \to 0$ as $\varepsilon \to 0^+$, uniformly with respect to t and β. Furthermore, for fixed ε, $x_{\beta,\varepsilon}(t)$ is continuous in β, uniformly with respect to t.

Proof of this theorem can be found in [7] and [8] (in slightly different notation) or [10]. An important part of the proof is a lemma by Halkin (see [7]; a proof is also given in [10]) which states that the sets $A_j(\beta,\varepsilon)$ can be chosen so that

$$\begin{aligned} &\mathrm{m}(A_0) = (1-\varepsilon)(T_1 - T_0), \quad \mathrm{m}(A_j) = \varepsilon\beta_j(T_1 - T_0),\ j = 1, \ldots, N, \\ &\mathrm{m}(A_j(\beta,\varepsilon) \,\triangle\, A_j(\beta',\varepsilon)) \to 0 \text{ as } \beta \to \beta', \quad \beta,\, \beta' \in S_N,\ j = 0, \ldots, N, \end{aligned}$$

and

$$\begin{aligned} \Big|(1-\varepsilon) \int_{T_0}^{t} f(\tau, x_0(\tau), u_0(\tau))\, d\tau + \varepsilon \sum_{j=1}^{N} \beta_j \int_{T_0}^{t} f(\tau, x_0(\tau), u_j(\tau))\, d\tau \\ - \int_{T_0}^{t} f(\tau, x_0(\tau), u_{\beta,\varepsilon}(\tau))\, d\tau\Big| < \varepsilon^2 \quad \text{for all } t \in I. \end{aligned} \tag{4.3}$$

From (4.3) and a general result about the effect of perturbations of the initial value and the right-hand side of a differential equation Theorem 2 follows (see [10]).

Now we want to extend Theorem 2 to problems with impulse effects. We have $T_0 = \tau_0 < \tau_1 < \cdots < \tau_r < \tau_{r+1} = T_1$, τ_k fixed, $u \in \mathcal{U}$, and consider the equation

$$\begin{aligned} \dot{x} &= f(t, x, u(t)), \quad t \neq \tau_k, \\ \Delta x|_{\tau_k} &= J_k(x(\tau_k), z_k), \quad k = 1, \dots, r. \end{aligned}$$

The parameters z_k are chosen from some sets Z_k. We assume that $J_k(\cdot, z_k) \in C^1(\mathbb{R}^n)$ for each k and $z_k \in Z_k$, and that the set $J_k(x, Z_k)$ is convex for each k and x. Denote

$$\bar{z} = (z_1, \dots, z_r) \in Z_1 \times \cdots \times Z_r = \bar{Z}.$$

Let $u_0 \in \mathcal{U}$, $\bar{z}_0 = (z_{0,1}, \dots, z_{0,r}) \in \bar{Z}$ and a corresponding solution x_0 be given, i.e.,

$$\begin{aligned} \dot{x}_0 &= f(t, x_0, u_0(t)), \quad t \neq \tau_k, \\ \Delta x_0|_{\tau_k} &= J_k(x_0(\tau_k), z_{0,k}). \end{aligned}$$

Let $\xi_j \in \mathbb{R}^n$, $u_j \in \mathcal{U}$, and $\bar{z}_j \in \bar{Z}$, $j = 1, \dots, N$, be given. For $\varepsilon \in (0, 1)$, $\beta \in S_N$, find sets $A_j(\beta, \varepsilon)$ and define $u_{\beta,\varepsilon} \in \mathcal{U}$ as described above. Let $x_{\beta,\varepsilon}$ be the solution of

$$\begin{aligned} \dot{x} &= f(t, x, u_{\beta,\varepsilon}(t)), \\ x(T_0) &= x_0(T_0) + \varepsilon \sum_{j=1}^{N} \beta_j \xi_j. \end{aligned}$$

If ε is sufficiently small, $x_{\beta,\varepsilon}(t)$ exists on $[T_0, \tau_1]$ and satisfies (according to Theorem 2)

$$x_{\beta,\varepsilon}(t) = x_0(t) + \varepsilon \sum_{j=1}^{N} \beta_j v(t; \xi_j, u_j) + r_1(t; \beta, \varepsilon), \quad t \in [T_0, \tau_1],$$

where $v(\cdot; \xi, u)$ satisfies (4.1) with $\tau = T_0$, and $r_1(t; \beta, \varepsilon)/\varepsilon \to 0$ as $\varepsilon \to 0^+$ uniformly with respect to $t \in [T_0, \tau_1]$ and $\beta \in S_N$; we say that $r_1(t; \beta, \varepsilon) = o(\varepsilon)$ uniformly w.r.t. t and β. Furthermore, $x_{\beta,\varepsilon}(\tau_1)$ is continuous in β. Since $J_1(x_{\beta,\varepsilon}(\tau_1), Z_1)$ is convex, there exists $z_{\beta,\varepsilon,1} \in Z_1$ such that

$$J_1(x_{\beta,\varepsilon}(\tau_1), z_{0,1}) + \varepsilon \sum_{j=1}^{N} \beta_j [J_1(x_{\beta,\varepsilon}(\tau_1), z_{j,1}) - J_1(x_{\beta,\varepsilon}(\tau_1), z_{0,1})] = J_1(x_{\beta,\varepsilon}(\tau_1), z_{\beta,\varepsilon,1}).$$

We continue the definition of $x_{\beta,\varepsilon}$ by letting the jump at τ_1 be $J_1(x_{\beta,\varepsilon}(\tau_1), z_{\beta,\varepsilon,1})$. The new initial value at τ_1 is $x_{\beta,\varepsilon}(\tau_1^+) = x_{\beta,\varepsilon}(\tau_1) + J_1(x_{\beta,\varepsilon}(\tau_1), z_{\beta,\varepsilon,1})$. We have

$$J_1(x_{\beta,\varepsilon}(\tau_1), z_{0,1}) = J_1(x_0(\tau_1), z_{0,1}) + \varepsilon \frac{\partial J_1}{\partial x}(x_0(\tau_1), z_{0,1}) \sum_{j=1}^{N} \beta_j v(\tau_1; \xi_j, u_j) + o(\varepsilon)$$

uniformly w.r.t. β. For $j = 1, \dots, N$ it is enough to note that $J_1(x_{\beta,\varepsilon}(\tau_1), z_{j,1}) = J_1(x_0(\tau_1), z_{j,1}) + O(\varepsilon)$ uniformly w.r.t. β. Thus

$$\begin{aligned} x_{\beta,\varepsilon}(\tau_1^+) &= x_0(\tau_1) + \varepsilon \sum_{j=1}^{N} \beta_j v(\tau_1; \xi_j, u_j) + J_1(x_0(\tau_1), z_{0,1}) \\ &\quad + \varepsilon \frac{\partial J_1}{\partial x}(x_0(\tau_1), z_{0,1}) \sum_{j=1}^{N} \beta_j v(\tau_1; \xi_j, u_j) \\ &\quad + \varepsilon \sum_{j=1}^{N} \beta_j [J_1(x_0(\tau_1), z_{j,1}) - J_1(x_0(\tau_1), z_{0,1})] + o(\varepsilon) \qquad (4.4) \\ &= x_0(\tau_1^+) + \varepsilon \big(E + \frac{\partial J_1}{\partial x}(x_0(\tau_1), z_{0,1})\big) \sum_{j=1}^{N} \beta_j v(\tau_1; \xi_j, u_j) \\ &\quad + \varepsilon \sum_{j=1}^{N} \beta_j [J_1(x_0(\tau_1), z_{j,1}) - J_1(x_0(\tau_1), z_{0,1})] + o(\varepsilon) \end{aligned}$$

uniformly w.r.t. β. From the definition of $J_1(x_{\beta,\varepsilon}(\tau_1), z_{\beta,\varepsilon,1})$ we see that it is continuous in β; thus $x_{\beta,\varepsilon}(\tau_1^+)$ is continuous in β.

Now let $v(t;\xi,u,\bar{z})$ for $\xi \in \mathbb{R}^n$, $u \in \mathcal{U}$, and $\bar{z} \in \bar{Z}$ be the solution of the linear system

$$\begin{cases} \dot{v} = \dfrac{\partial f}{\partial x}(t, x_0(t), u_0(t))v + f(t, x_0(t), u(t)) - f(t, x_0(t), u_0(t)), \quad t \neq \tau_k, \\ \Delta v|_{\tau_k} = \dfrac{\partial J_k}{\partial x}(x_0(\tau_k), z_{0,k})v(\tau_k) + J_k(x_0(\tau_k), z_k) - J_k(x_0(\tau_k), z_{0,k}), \quad k = 1, \dots, r, \\ v(T_0) = \xi. \end{cases}$$

Let $v_j(t) = v(t;\xi_j, u_j, \bar{z}_j)$. Note that $v_j(t)$ coincides with $v(t;\xi_j, u_j)$ above for $T_0 \leq t \leq \tau_1$. Since

$$v_j(\tau_1^+) = v_j(\tau_1) + \frac{\partial J_1}{\partial x}(x_0(\tau_1), z_{0,1})v_j(\tau_1) + J_1(x_0(\tau_1), z_{j,1}) - J_1(x_0(\tau_1), z_{0,1}),$$

we have from (4.4)

$$x_{\beta,\varepsilon}(\tau_1^+) = x_0(\tau_1^+) + \varepsilon \sum_{j=1}^{N} \beta_j v_j(\tau_1^+) + o(\varepsilon).$$

Next, define $x_{\beta,\varepsilon}(t)$ for $\tau_1 < t \leq \tau_2$ as the solution of

$$\begin{cases} \dot{x} = f(t, x, u_{\beta,\varepsilon}(t)), \\ x(\tau_1^+) = x_{\beta,\varepsilon}(\tau_1) + J_1(x_{\beta,\varepsilon}(\tau_1), z_{\beta,\varepsilon,1}) = x_0(\tau_1^+) + \varepsilon \displaystyle\sum_{j=1}^{N} \beta_j v_j(\tau_1^+) + o(\varepsilon). \end{cases}$$

It exists on $(\tau_1, \tau_2]$ if ε is sufficiently small and satisfies (according to Theorem 2)

$$x_{\beta,\varepsilon}(t) = x_0(t) + \varepsilon \sum_{j=1}^{N} \beta_j v_j(t) + o(\varepsilon)$$

as $\varepsilon \to 0^+$ uniformly w.r.t. t and β. This follows since v_j satisfies

$$\dot{v}_j(t) = \frac{\partial f}{\partial x}(t, x_0(t), u_0(t))v_j(t) + f(t, x_0(t), u_j(t)) - f(t, x_0(t), u_0(t))$$

on $(\tau_1, \tau_2]$ (a.e.). The extra term $o(\varepsilon)$ in the initial value only contributes an extra $o(\varepsilon)$ in the solution (an application of the Gronwall inequality). Also, $x_{\beta,\varepsilon}(t)$ is continuous in β. We now define $z_{\beta,\varepsilon,2} \in Z_2$ such that

$$J_2(x_{\beta,\varepsilon}(\tau_2), z_{0,2}) + \varepsilon \sum_{j=1}^{N} \beta_j [J_2(x_{\beta,\varepsilon}(\tau_2), z_{j,2}) - J_2(x_{\beta,\varepsilon}(\tau_2), z_{0,2})] = J_2(x_{\beta,\varepsilon}(\tau_2), z_{\beta,\varepsilon,2}).$$

The jump at τ_2 is $J_2(x_{\beta,\varepsilon}(\tau_2), z_{\beta,\varepsilon,2})$. Then we can proceed as above and define $x_{\beta,\varepsilon}(t)$ on $(\tau_2, \tau_3]$, $(\tau_3, \tau_4]$,$\dots$, $(\tau_r, T_1]$. We have then defined $\bar{z}_{\beta,\varepsilon} = (z_{\beta,\varepsilon,1}, \dots, z_{\beta,\varepsilon,r}) \in \bar{Z}$, and $x_{\beta,\varepsilon}(\cdot)$ is the solution of

$$\begin{cases} \dot{x} = f(t, x, u_{\beta,\varepsilon}(t)), \quad t \neq \tau_k, \\ \Delta x|_{\tau_k} = J_k(x(\tau_k), z_{\beta,\varepsilon,k}), \quad k = 1, \dots, r, \\ x(T_0) = x_0(T_0) + \varepsilon \displaystyle\sum_{j=1}^{N} \beta_j \xi_j. \end{cases}$$

As end result we obtain

$$x_{\beta,\varepsilon}(t) = x_0(t) + \varepsilon \sum_{j=1}^{N} \beta_j v_j(t) + r(t;\beta,\varepsilon), \tag{4.5}$$

where $r(t;\beta,\varepsilon)/\varepsilon \to 0$ as $\varepsilon \to 0^+$, uniformly w.r.t. t and β, and $x_{\beta,\varepsilon}(t)$ is continuous in β uniformly w.r.t. t (for fixed ε).

5 A continuous-discrete optimal control problem

5.1 Fixed switching times

Consider the control system

$$\dot{x} = f(t, x, u(t)), \quad t \neq \tau_k, \quad u(\cdot) \in \mathcal{U}, \tag{5.1}$$

$$\Delta x|_{\tau_k} = J_k(x(\tau_k), z_k), \quad k = 1, \ldots, r, \quad \bar{z} \in \bar{Z}. \tag{5.2}$$

The notation and general assumptions are as in Section 4. We have a number of constraints for $x(\tau_k)$ $(0 \le k \le r+1)$ and $x(\tau_k^+)$ or equivalently $\Delta x|_{\tau_k} = J_k(x(\tau_k), z_k)$ $(1 \le k \le r)$:

$$g_i(x(\tau_0), x(\tau_1), \Delta x|_{\tau_1}, \ldots, x(\tau_r), \Delta x|_{\tau_r}, x(\tau_{r+1})) = 0, \quad i = 1, \ldots, p, \tag{5.3}$$

$$g_i(x(\tau_0), x(\tau_1), \Delta x|_{\tau_1}, \ldots, x(\tau_r), \Delta x|_{\tau_r}, x(\tau_{r+1})) \le 0, \quad i = -q, \ldots, -1. \tag{5.4}$$

We want to minimize a functional

$$g_0(x(\tau_0), x(\tau_1), \Delta x|_{\tau_1}, \ldots, x(\tau_r), \Delta x|_{\tau_r}, x(\tau_{r+1})). \tag{5.5}$$

We assume that the functions g_i, defined for $(x_0, x_1, y_1, \ldots, x_r, y_r, x_{r+1}) \in (\mathbb{R}^n)^{2r+2}$, are continuously differentiable.

Assume that $(u_0(\cdot), \bar{z}_0, x_0(\cdot))$ is an optimal triple for the problem of minimizing (5.5) subject to (5.1)–(5.4). Let us work in the linear space $X = PC_n(I) \times (\mathbb{R}^n)^r$, where $PC_n(I)$ is the linear space of piecewise continuous n-vector functions on $I = [T_0, T_1]$ that are continuous except perhaps at τ_k, $k = 1, \ldots, r$, where they are continuous from the left. Elements of $(\mathbb{R}^n)^r$ are denoted by $\bar{y} = (y_1, \ldots, y_r)$. Let

$$\begin{aligned} A = \{(x, \bar{y}) \in X : {}& x(\cdot) \text{ is a solution of (5.1)–(5.2)},\ y_k = J_k(x(\tau_k), z_k),\ k = 1, \ldots, r, \\ & \text{for some } u \in \mathcal{U},\ \bar{z} \in \bar{Z}\}, \\ \varphi_i(x, \bar{y}) &= g_i(x(\tau_0), x(\tau_1), y_1, \ldots, x(\tau_r), y_r, x(\tau_{r+1})), \quad -q \le i \le p, \\ \bar{y}_0 &= (y_{0,1}, \ldots, y_{0,r}), \qquad \text{where } y_{0,k} = J_k(x_0(\tau_k), z_{0,k}). \end{aligned}$$

Then $(x_0, \bar{y}_0) \in A$ is a solution of the problem of minimizing $\varphi_0(x, \bar{y})$ subject to $(x, \bar{y}) \in A$, $\varphi_i(x, \bar{y}) = 0$ for $i = 1, \ldots, p$, $\varphi_i(x, \bar{y}) \le 0$ for $i = -q, \ldots, -1$. Thus our control problem is formulated as a special case of Problem $(\mathcal{P})$ in Section 3. We must verify Assumption 1 in this case.

Let $v(t; \xi, u, \bar{z})$ be as in Section 4, and define

$$w_k(\xi, u, \bar{z}) = \frac{\partial J_k}{\partial x}(x_0(\tau_k), z_{0,k}) v(\tau_k; \xi, u, \bar{z}) + J_k(x_0(\tau_k), z_k) - J_k(x_0(\tau_k), z_{0,k}),\ k = 1, \ldots, r.$$

Let

$$M = \{(x, \bar{y}) \in X : x(\cdot) = v(\cdot; \xi, u, \bar{z}),\ y_k = w_k(\xi, u, \bar{z}) \text{ for some } \xi \in \mathbb{R}^n,\ u \in \mathcal{U},\ \bar{z} \in \bar{Z}\},$$

and let elements $(v_j, \bar{y}_j)$, $j = 1, \ldots, N$, in M be given. Then

$$v_j(\cdot) = v(\cdot; \xi_j, u_j, \bar{z}_j) \text{ and } y_{j,k} = w_k(\xi_j, u_j, \bar{z}_j) \text{ for some } \xi_j \in \mathbb{R}^n,\ u_j \in \mathcal{U},\ \bar{z}_j \in \bar{Z}.$$

For $\beta \in S_N$ and $\varepsilon > 0$ sufficiently small we define $u_{\beta,\varepsilon}$, $x_{\beta,\varepsilon}$, and $\bar{z}_{\beta,\varepsilon}$ as is described in Section 4. We showed there (see (4.5)) that

$$x_{\beta,\varepsilon}(t) = x_0(t) + \varepsilon \sum_{j=1}^N \beta_j v_j(t) + o(\varepsilon), \tag{5.6}$$

as $\varepsilon \to 0^+$, uniformly w.r.t. t and β. Further, $x_{\beta,\varepsilon}(t)$ is continuous in β. Also, $\bar{z}_{\beta,\varepsilon}$ satisfies

$$J_k(x_{\beta,\varepsilon}(\tau_k), z_{0,k}) + \varepsilon \sum_{j=1}^N \beta_j [J_k(x_{\beta,\varepsilon}(\tau_k), z_{j,k}) - J_k(x_{\beta,\varepsilon}(\tau_k), z_{0,k})] = J_k(x_{\beta,\varepsilon}(\tau_k), z_{\beta,\varepsilon,k}). \tag{5.7}$$

Define an element $\bar{y}_{\beta,\varepsilon} \in (\mathbb{R}^n)^r$ with components

$$y_{\beta,\varepsilon,k} = J_k(x_{\beta,\varepsilon}(\tau_k), z_{\beta,\varepsilon,k}).$$

It follows from (5.7) that $y_{\beta,\varepsilon,k}$ is continuous in β. We also have from (5.6) and (5.7)

$$\begin{aligned} y_{\beta,\varepsilon,k} &= J_k(x_0(\tau_k), z_{0,k}) + \varepsilon \frac{\partial J_k}{\partial x}(x_0(\tau_k), z_{0,k}) \sum_{j=1}^{N} \beta_j v_j(\tau_k) \\ &\quad + \varepsilon \sum_{j=1}^{N} \beta_j [J_k(x_0(\tau_k), z_{j,k}) - J_k(x_0(\tau_k), z_{0,k})] + o(\varepsilon) \\ &= y_{0,k} + \varepsilon \sum_{j=1}^{N} \beta_j w_k(\xi_j, u_j, \bar{z}_j) + o(\varepsilon) = y_{0,k} + \varepsilon \sum_{j=1}^{N} \beta_j y_{j,k} + o(\varepsilon), \end{aligned}$$

so that

$$\bar{y}_{\beta,\varepsilon} = \bar{y}_0 + \varepsilon \sum_{j=1}^{N} \beta_j \bar{y}_j + o(\varepsilon) \tag{5.8}$$

uniformly w.r.t. β.

Now we can verify Assumption 1 (i)–(iii). We have

$$\rho(\beta, \varepsilon) = (x_{\beta,\varepsilon}(\cdot), \bar{y}_{\beta,\varepsilon}) \in A,$$

and

$$\varphi_i(\rho(\beta,\varepsilon)) = g_i(x_{\beta,\varepsilon}(\tau_0), x_{\beta,\varepsilon}(\tau_1), y_{\beta,\varepsilon,1}, \ldots, x_{\beta,\varepsilon}(\tau_r), y_{\beta,\varepsilon,r}, x_{\beta,\varepsilon}(\tau_{r+1})) \tag{5.9}$$

is continuous in β. For (ii) and (iii) we have from (5.6), (5.8)–(5.9) with

$$e_0 = (x_0(\tau_0), x_0(\tau_1), y_{0,1}, \ldots, x_0(\tau_r), y_{0,r}, x_0(\tau_{r+1})),$$

$$\begin{aligned} \lim_{\varepsilon \to 0^+} \frac{\varphi_i(\rho(\beta,\varepsilon)) - \varphi_i(x_0, \bar{y}_0)}{\varepsilon} &= \sum_{k=0}^{r+1} \frac{\partial g_i}{\partial x_k}(e_0) \sum_{j=1}^{N} \beta_j v_j(\tau_k) + \sum_{k=1}^{r} \frac{\partial g_i}{\partial y_k}(e_0) \sum_{j=1}^{N} \beta_j y_{j,k} \\ &= h_i\Big(\sum_{j=1}^{N} \beta_j (v_j, \bar{y}_j)\Big), \end{aligned}$$

where the convergence is uniform w.r.t. β, and where h_i is the linear functional on X defined by

$$h_i(x, \bar{y}) = \sum_{k=0}^{r+1} \frac{\partial g_i}{\partial x_k}(e_0) x(\tau_k) + \sum_{k=1}^{r} \frac{\partial g_i}{\partial y_k}(e_0) y_k.$$

According to Theorem 1 there exist numbers λ_i, not all zero, such that

$$\begin{aligned} \sum_{i=-q}^{p} \lambda_i h_i(x, \bar{y}) &\le 0 \quad \text{for all } (x, \bar{y}) \in \operatorname{co} M, \\ \lambda_i &\le 0 \quad \text{for } i \le 0, \\ \lambda_i g_(e_0) &= 0 \quad \text{for } i < 0. \end{aligned}$$

Take a typical element $(x, \bar{y})$ of M. Then

$$\sum_{i=-q}^{p} \lambda_i \Big\{ \sum_{k=0}^{r+1} \frac{\partial g_i}{\partial x_k}(e_0) v(\tau_k; \xi, u, \bar{z}) + \sum_{k=1}^{r} \frac{\partial g_i}{\partial y_k}(e_0) w_k(\xi, u, \bar{z}) \Big\} \le 0$$

for all $\xi \in \mathbb{R}^n$, $u \in \mathcal{U}$, $\bar{z} \in \bar{Z}$.

Write $G = \sum_{i=-q}^{p} \lambda_i g_i$, $G_k = \frac{\partial G}{\partial x_k}(e_0)$ $(0 \le k \le r+1)$, and $G_k^+ = \frac{\partial G}{\partial y_k}(e_0)$ $(1 \le k \le r)$, so that

$$\sum_{k=0}^{r+1} G_k v(\tau_k; \xi, u, \bar{z}) + \sum_{k=1}^{r} G_k^+ w_k(\xi, u, \bar{z}) \le 0. \tag{5.10}$$

With

$$\begin{aligned} A(t) &= \frac{\partial f}{\partial x}(t, x_0(t), u_0(t)), \quad b(t) = f(t, x_0(t), u(t)) - f(t, x_0(t), u_0(t)), \\ C_k &= \frac{\partial J_k}{\partial x}(x_0(\tau_k), z_{0,k}), \quad d_k = J_k(x_0(\tau_k), z_k) - J_k(x_0(\tau_k), z_{0,k}),\ k = 1, \dots, r, \end{aligned}$$

we have that $v(\cdot; \xi, u, \bar{z})$ is the solution of

$$\begin{cases} \dot{v} = A(t)v + b(t), \quad t \ne \tau_k, \\ \Delta v|_{\tau_k} = C_k v(\tau_k) + d_k = w_k(\xi, u, \bar{z}), \\ v(T_0) = \xi. \end{cases}$$

With the notation from Section 2 we have from (2.7)

$$v(\tau_k; \xi, u, \bar{z}) = \Phi(\tau_k)\Psi_{k-1,0}\xi + \int_{T_0}^{\tau_k} \Phi(\tau_k)\Psi(\tau_k, t)\Phi^{-1}(t)b(t)\,dt + \sum_{j=1}^{k-1} \Phi(\tau_k)\Psi_{k-1,j}\Phi^{-1}(\tau_j)d_j$$

for $k = 1, \dots, r+1$.

Take $u = u_0$, $\bar{z} = \bar{z}_0$ in (5.10). Then $b(t) = 0$, $d_k = 0$, and

$$\Big(G_0 + \sum_{k=1}^{r+1} G_k \Phi(\tau_k)\Psi_{k-1,0} + \sum_{k=1}^{r} G_k^+ C_k \Phi(\tau_k)\Psi_{k-1,0}\Big)\xi \le 0 \quad \text{for all } \xi \in \mathbb{R}^n.$$

Thus

$$G_0 + \sum_{k=1}^{r} (G_k + G_k^+ C_k)\Phi(\tau_k)\Psi_{k-1,0} + G_{r+1}\Phi(T_1)\Psi_{r,0} = 0. \tag{5.11}$$

Next, take $\xi = 0$, $\bar{z} = \bar{z}_0$ in (5.10). Then

$$\sum_{k=1}^{r} (G_k + G_k^+ C_k) \int_{T_0}^{\tau_k} \Phi(\tau_k)\Psi(\tau_k, t)\Phi^{-1}(t)b(t)\,dt + G_{r+1} \int_{T_0}^{T_1} \Phi(T_1)\Psi(T_1, t)\Phi^{-1}(t)b(t)\,dt \le 0$$

for all $u \in \mathcal{U}$. Define ($\chi_A(\cdot)$ is the characteristic function of the set A)

$$\eta(t) = \sum_{k=1}^{r} (G_k + G_k^+ C_k)\chi_{[T_0, \tau_k]}(t)\Phi(\tau_k)\Psi(\tau_k, t)\Phi^{-1}(t) + G_{r+1}\Phi(T_1)\Psi(T_1, t)\Phi^{-1}(t).$$

Then

$$\int_{T_0}^{T_1} \eta(t)[f(t, x_0(t), u(t)) - f(t, x_0(t), u_0(t))]\,dt \le 0 \quad \text{for all } u \in \mathcal{U}. \tag{5.12}$$

For $\tau_j < t \le \tau_{j+1}$ $(j = 0, \dots, r)$ we have

$$\eta(t) = \sum_{k=j+1}^{r} (G_k + G_k^+ C_k)\Phi(\tau_k)\Psi_{k-1,j}\Phi^{-1}(t) + G_{r+1}\Phi(T_1)\Psi_{rj}\Phi^{-1}(t).$$

In the case $j = r$ the sum is empty, i.e. zero. Thus $\eta(\cdot)$ satisfies $\dot{\eta}(t) = -\eta(t)A(t)$ a.e. on $(\tau_j, \tau_{j+1}]$. The jump at τ_j $(j = 1, \dots, r)$ is

$$\begin{aligned} \eta(\tau_j^+) - \eta(\tau_j) = \sum_{k=j+1}^{r} &(G_k + G_k^+ C_k)\Phi(\tau_k)(\Psi_{k-1,j} - \Psi_{k-1,j-1})\Phi^{-1}(\tau_j) \\ &- (G_j + G_j^+ C_j) + G_{r+1}\Phi(T_1)(\Psi_{rj} - \Psi_{r,j-1})\Phi^{-1}(\tau_j). \end{aligned}$$

If $k > j+1$, then (by (2.4))

$$\begin{aligned}\Psi_{k-1,j} - \Psi_{k-1,j-1} &= \Psi_{k-1}\cdots\Psi_{j+1} - \Psi_{k-1}\cdots\Psi_j = \Psi_{k-1,j}(E-\Psi_j)\\ &= -\Psi_{k-1,j}\Phi^{-1}(\tau_j)C_j\Phi(\tau_j).\end{aligned}$$

This is true also if $k = j+1$, since $\Psi_{jj} - \Psi_{j,j-1} = E - \Psi_j$. Thus

$$\begin{aligned}\eta(\tau_j^+) - \eta(\tau_j) = &-\sum_{k=j+1}^{r}(G_k + G_k^+C_k)\Phi(\tau_k)\Psi_{k-1,j}\Phi^{-1}(\tau_j)C_j\\ &-(G_j + G_j^+C_j) - G_{r+1}\Phi(T_1)\Psi_{rj}\Phi^{-1}(\tau_j)C_j\\ = &-\eta(\tau_j^+)C_j - G_j - G_j^+C_j.\end{aligned}$$

At T_0 and T_1 we have according to (5.11)

$$\begin{aligned}&\eta(T_0) = \eta(T_0^+) = \sum_{k=1}^{r}(G_k + G_k^+C_k)\Phi(\tau_k)\Psi_{k-1,0} + G_{r+1}\Phi(T_1)\Psi_{r0} = -G_0,\\ &\eta(T_1) = G_{r+1}.\end{aligned}$$

It is convenient to introduce $H(t,x,u) = \eta(t)f(t,x,u)$. Then $\eta(\cdot)$ satisfies

$$\dot\eta(t) = -\frac{\partial H}{\partial x}(t, x_0(t), u_0(t)) \quad \text{a.e.}$$

With $H_0(t,u) = H(t,x_0(t),u)$ (5.12) can be written

$$\int_{T_0}^{T_1} H_0(t,u(t))\,dt \le \int_{T_0}^{T_1} H_0(t,u_0(t))\,dt \qquad \text{for all } u \in \mathcal{U}. \tag{5.13}$$

From part (iv) of Assumption 2 we can obtain a pointwise inequality. Let $u_j \in \mathcal{U}$, $j = 1,2,\ldots$, be such that $\{u_j(t)\}_{j=1}^{\infty}$ is dense in $\Omega(t)$ for every $t \in I$. Let $t' \in (T_0, T_1]$ be a Lebesgue point of $t \mapsto f(t, x_0(t), u_j(t))$ for every $j = 0,1,2,\ldots$. Define for $0 < \varepsilon < t' - T_0$, $j = 1,2,\ldots$,

$$u_{j,\varepsilon}(t) = \begin{cases} u_j(t), & \text{if } t' - \varepsilon < t \le t',\\ u_0(t), & \text{otherwise on } I.\end{cases}$$

Then $u_{j,\varepsilon} \in \mathcal{U}$, and if we apply (5.13) to $u_{j,\varepsilon}$, divide by ε and let $\varepsilon \to 0^+$, we obtain

$$H_0(t', u_j(t')) \le H_0(t', u_0(t')) \quad \text{for all } j.$$

Since $\{u_j(t')\}_{j=1}^{\infty}$ is dense in $\Omega(t')$, and H_0 is continuous in u, it follows that

$$H_0(t', u) \le H_0(t', u_0(t')) \quad \text{for all } u \in \Omega(t').$$

The point t' can be almost any point in I.

Finally, take $\xi = 0$, $u = u_0$ in (5.10). Then

$$\sum_{k=2}^{r+1} G_k \sum_{j=1}^{k-1}\Phi(\tau_k)\Psi_{k-1,j}\Phi^{-1}(\tau_j)d_j + G_1^+d_1 + \sum_{k=2}^{r} G_k^+\Big(C_k\sum_{j=1}^{k-1}\Phi(\tau_k)\Psi_{k-1,j}\Phi^{-1}(\tau_j)d_j + d_k\Big) \le 0$$

After changing the order of summation and exchanging j and k we obtain

$$\sum_{k=1}^{r}\sum_{j=k+1}^{r+1} G_j\Phi(\tau_j)\Psi_{j-1,k}\Phi^{-1}(\tau_k)d_k + \sum_{k=1}^{r-1}\sum_{j=k+1}^{r} G_j^+C_j\Phi(\tau_j)\Psi_{j-1,k}\Phi^{-1}(\tau_k)d_k + \sum_{k=1}^{r} G_k^+d_k \le 0.$$

From the definition of η we see that the coefficient of d_k is $\eta(\tau_k^+) + G_k^+$. Therefore

$$\sum_{k=1}^{r} (\eta(\tau_k^+) + G_k^+)[J_k(x_0(\tau_k), z_k) - J_k(x_0(\tau_k), z_{0,k})] \leq 0 \quad \text{for all } \bar{z} \in \bar{Z}.$$

Since the z_k:s can be varied independently of each other, we have

$$(\eta(\tau_k^+) + G_k^+)[J_k(x_0(\tau_k), z_k) - J_k(x_0(\tau_k), z_{0,k})] \leq 0 \quad \text{for all } z_k \in Z_k,\ k = 1, \ldots, r.$$

We have obtained the following theorem.

Theorem 3. *Assume that* $(u_0(\cdot), \bar{z}_0, x_0(\cdot))$ *is a solution of the problem of minimizing* (5.5) *subject to* (5.1)–(5.4)*. Then there exist numbers* $\lambda_{-q}, \ldots, \lambda_0, \ldots, \lambda_p$*, not all zero, and a piecewise absolutely continuous row vector function* $\eta(\cdot)$ *such that, with* $G = \sum_{i=-q}^{p} \lambda_i g_i$ *and* $H(t, x, u) = \eta(t) f(t, x, u)$*,*

$$\begin{aligned}
\dot{\eta}(t) &= -\frac{\partial H}{\partial x}(t, x_0(t), u_0(t)) \quad a.e., \\
\Delta\eta|_{\tau_k} &= -\big(\eta(\tau_k^+) + \frac{\partial G}{\partial y_k}(e_0)\big)\frac{\partial J_k}{\partial x}(x_0(\tau_k), z_{0,k}) - \frac{\partial G}{\partial x_k}(e_0), \quad k = 1, \ldots, r, \\
\eta(T_0) &= -\frac{\partial G}{\partial x_0}(e_0), \\
\eta(T_1) &= \frac{\partial G}{\partial x_{r+1}}(e_0), \\
\lambda_i &\leq 0 \quad for\ -q \leq i \leq 0, \\
\lambda_i g_i(e_0) &= 0 \quad for\ -q \leq i \leq -1, \\
H(t, x_0(t), u_0(t)) &= \max_{u \in \Omega(t)} H(t, x_0(t), u) \quad a.e.\ on\ I, \\
\big(\eta(\tau_k^+) + \frac{\partial G}{\partial y_k}(e_0)\big)J_k(x_0(\tau_k), z_{0,k}) &= \max_{z_k \in Z_k} \big(\eta(\tau_k^+) + \frac{\partial G}{\partial y_k}(e_0)\big)J_k(x_0(\tau_k), z_k),\ k = 1, \ldots, r.
\end{aligned}$$

Remark 1. In (5.1) it is possible to have different right-hand sides on different intervals $I_k = (\tau_{k-1}, \tau_k]$, so that x satisfies $\dot{x} = f_k(t, x, u(t))$ on I_k. Just write $f(t, x, u) = \sum_{k=1}^{r+1} \chi_{I_k}(t) f_k(t, x, u)$ and apply Theorem 3. It is also possible to have control vectors u_k of different dimensions on different I_k, so that the admisible controls on I_k satisfy $u_k(t) \in \Omega_k(t) \subseteq \mathbb{R}^{m_k}$. If $m = \max_{1 \leq k \leq r+1} m_k$, we may extend u_k and $\Omega_k(t)$ with $m - m_k$ components that are set to 0. Then Theorem 3 can be applied.

Remark 2. Consider the same problem as above but with an integral term $\int_{T_0}^{T_1} f^0(t, x(t), u(t))\, dt$ added to the cost fuctional (5.5), where the real-valued function f^0 has the same properties as f. Then the necessary conditions for optimality are exactly as in Theorem 3, except that H now is defined as

$$H(t, x, u) = \lambda_0 f^0(t, x, u) + \eta(t) f(t, x, u).$$

This is proved by introducing a new state variable x^0 satisfying

$$\dot{x}^0(t) = f^0(t, x(t), u(t)) \quad \text{a.e. in } I$$

with no jumps. The integral term is then written as $x^0(\tau_{r+1}) - x^0(\tau_0)$, and Theorem 3 can be applied to the extended system.

5.2 Variable switching times

Let us now allow the times τ_k (including the initial time T_0 and the final time T_1) to vary, i.e., they become control parameters. We will transform the problem to a problem with fixed switching times, so that the previous result (Theorem 3) can be applied. In this transformation t will be treated as a state variable, so we need the same regularity in t as in x. Also $\Omega(t)$ must be constant. The functions g_i can now depend explicitly on $\bar{\tau} = (\tau_0, \ldots, \tau_{r+1})$, and the jump J_k may depend on τ_k. The assumptions from 5.1 have to be modified in the following way:

We may have different differential equations on different intervals $I_k = (\tau_{k-1}, \tau_k]$, so we assume that we are given $r+1$ functions $(t, x, u_k) \mapsto f_k(t, x, u_k)$ from $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^{m_k}$ to $\mathbb{R}^n$. We assume that f_k is continuously differentiable w.r.t. (t, x), and that f_k, $\frac{\partial f_k}{\partial t}$ and $\frac{\partial f_k}{\partial x}$ are continuous in u_k. We will also denote the first argument of f_k by τ, write $\hat{x} = (\tau, x)$, and consider f_k as a function of $\hat{x}$ and u_k. By $\mathcal{U}_{I_k}$ we denote the set of all controls $u_k(\cdot)$ that are defined and measurable on I_k and such that $u_k(t) \in \Omega_k$ for all $t \in I_k$ (Ω_k is a given set in $\mathbb{R}^{m_k}$), and the function $(t, \hat{x}) \mapsto f_k(\hat{x}, u_k(t))$ belongs to the class $\mathcal{F}$ on I_k. The jump functions $J_k(\hat{x}, z_k) = J_k(\tau, x, z_k)$ are continuously differentiable w.r.t. $\hat{x}$ (for fixed $z_k \in Z_k$), and such that the sets $J_k(\hat{x}, Z_k)$ are convex for each $\hat{x}$. The functions g_i have arguments

$$(\bar{\tau}, x_0, x_1, y_1, \ldots, x_r, y_r, y_{r+1}) \in \mathbb{R}^{r+2} \times (\mathbb{R}^n)^{2r+2},$$

and are supposed to be continuously differentiable in all arguments.

Consider the control system

$$\dot{x} = f_k(t, x, u_k(t)) \quad \text{on } I_k, \quad u_k(\cdot) \in \mathcal{U}_{I_k}, \quad k = 1, \ldots, r+1, \tag{5.14}$$

$$\Delta x|_{\tau_k} = x(\tau_k^+) - x(\tau_k) = J_k(\tau_k, x(\tau_k), z_k), \quad k = 1, \ldots, r, \quad \bar{z} \in \bar{Z}, \tag{5.15}$$

with

$$T_0 = \tau_0 \le \tau_1 \le \ldots \le \tau_r \le \tau_{r+1} = T_1.$$

We want to minimize the functional

$$g_0(\bar{\tau}, x(\tau_0), x(\tau_1), \Delta x|_{\tau_1}, \ldots, x(\tau_r), \Delta x|_{\tau_r}, x(\tau_{r+1}))$$

subject to the constraints $g_i = 0$ for $i = 1, \ldots, p$, and $g_i \le 0$ for $i = -q, \ldots, -1$, where the g_i:s have the same arguments as g_0. Note that we allow some of the times τ_k to coincide. If $\tau_k = \tau_{k+1}$, there is of course no differential equation on I_{k+1}, but we treat $x(\tau_k)$ and $x(\tau_{k+1})$ as separate quantities and have the condition $x(\tau_{k+1}) = x(\tau_k^+) = x(\tau_k) + J_k(\tau_k, x(\tau_k), z_k)$.

Assume that $\bar{\tau}_0 = (\tau_{0,0}, \tau_{0,1}, \ldots, \tau_{0,r+1})$, $u_{0,k}(\cdot) \in \mathcal{U}_{I_{0,k}}$ (where $I_{0,k} = (\tau_{0,k-1}, \tau_{0,k}]$), $k = 1, \ldots, r+1$, $\bar{z}_0 = (z_{0,0}, \ldots, z_{0,r}) \in \bar{Z}$, and $x_0(\cdot)$ are optimal, i.e., x_0 satisfies

$$\dot{x}_0 = f_k(t, x_0, u_{0,k}(t)), \quad t \in I_{0,k},$$
$$\Delta x_0|_{\tau_{0,k}} = J_k(\tau_{0,k}, x_0(\tau_{0,k}), z_{0,k}), \quad k = 1, \ldots, r;$$

the constraints $g_i = 0$ $(i > 0)$ and $g_i \le 0$ $(i < 0)$ are satisfied at

$$e_0 = (\bar{\tau}_0, x_0(\tau_{0,0}), x_0(\tau_{0,1}), y_{0,1}, \ldots, x_0(\tau_{0,r}), y_{0,r}, x_0(\tau_{0,r+1})),$$

where

$$y_{0,k} = J_k(\tau_{0,k}, x_0(\tau_{0,k}), z_{0,k});$$

and g_0 is minimized. Let

$$\mathcal{K} = \{k \in \{1, \ldots, r+1\} : m(I_{0,k}) > 0\}, \qquad \mathcal{K}_0 = \{k \in \{1, \ldots, r+1\} : m(I_{0,k}) = 0\}.$$

Define

$$H_k(\tau, x, u_k, \eta) = \eta f_k(\tau, x, u_k),$$

where η is a row vector in $\mathbb{R}^n$.

Theorem 4. *Under the assumptions above there exist numbers $\lambda_{-q}, \ldots, \lambda_p$, not all zero, and functions $\eta_0(\cdot)$ and $\eta(\cdot)$ such that, with $G = \sum_{i=-q}^{p} \lambda_i g_i$ and*

$$M_k(t) = \sup_{u_k \in \Omega_k} H_k(t, x_0(t), u_k, \eta(t)),$$

the following holds:

$$\lambda_i \le 0 \quad \text{for } i = -q, \dots, 0,$$
$$\lambda_i g_i(e_0) = 0 \quad \text{for } i = -q, \dots, -1.$$

If $k \in \mathcal{K}$, then η_0 and η are absolutely continuous on $I_{0,k}$ and satisfy

$$\dot{\eta}_0(t) = -\frac{\partial H_k}{\partial \tau}\Big|_0 \quad a.e.,$$
$$\dot{\eta}(t) = -\frac{\partial H_k}{\partial x}\Big|_0 \quad a.e.,$$

where $\cdot|_0$ means evaluation at $(t, x_0(t), u_{0,k}(t), \eta(t))$,

$$H_k(t, x_0(t), u_{0,k}(t), \eta(t)) = M_k(t) \ a.e.,$$
$$\eta_0(t) + M_k(t) = 0 \ a.e., \text{ and } \eta_0(t) + M_k(t) \le 0 \text{ for all } t \in I_{0,k}.$$

If $k \in \mathcal{K}_0$, then $\eta_0(\tau_{0,k-1}^+) = \eta_0(\tau_{0,k}) = \eta_{0,k}$, $\eta(\tau_{0,k-1}^+) = \eta(\tau_{0,k}) = \eta_k$, and

$$\eta_{0,k} + M_k \le 0, \quad \text{where } M_k = \sup_{u_k \in \Omega_k} H_k(\tau_{0,k}, x_0(\tau_{0,k}), u_k, \eta_k) = M_k(\tau_{0,k}).$$

For $k = 1, \dots, r$,

$$\eta_0(T_0) = -\frac{\partial G}{\partial \tau_0}(e_0), \qquad \eta(T_0) = -\frac{\partial G}{\partial x_0}(e_0),$$
$$\eta_0(T_1) = \frac{\partial G}{\partial \tau_{r+1}}(e_0), \qquad \eta(T_1) = \frac{\partial G}{\partial x_{r+1}}(e_0),$$
$$\Delta\eta_0|_{\tau_{0,k}} = -\big(\eta(\tau_{0,k}^+) + \frac{\partial G}{\partial y_k}(e_0)\big)\frac{\partial J_k}{\partial \tau}\Big|_{0,k} - \frac{\partial G}{\partial \tau_k}(e_0),$$
$$\Delta\eta|_{\tau_{0,k}} = -\big(\eta(\tau_{0,k}^+) + \frac{\partial G}{\partial y_k}(e_0)\big)\frac{\partial J_k}{\partial x}\Big|_{0,k} - \frac{\partial G}{\partial x_k}(e_0),$$
$$\big(\eta(\tau_{0,k}^+) + \frac{\partial G}{\partial y_k}(e_0)\big)J_k|_{0,k} = \max_{z_k \in Z_k}\big(\eta(\tau_{0,k}^+) + \frac{\partial G}{\partial y_k}(e_0)\big)J_k(\tau_{0,k}, x_0(\tau_{0,k}), z_k),$$

where $\cdot|_{0,k}$ means evaluation at $(\tau_{0,k}, x_0(\tau_{0,k}), z_{0,k})$. If, furthermore, for $k \in \mathcal{K}$, there are functions ρ_1 and ρ_2 in $L^1(I_{0,k})$ such that

$$\Big|\frac{\partial f_k}{\partial x}(s, x_0(s), u_{0,k}(t))\, f_k(s, x_0(s), u_{0,k}(\tau))\Big| \le \rho_1(t) + \rho_2(\tau) \text{ for all } s, t, \tau \in I_{0,k}, \tag{5.16}$$

then $\eta_0(t) + M_k(t) = 0$ for all $t \in I_{0,k}$.

Proof. Let v_0 be a function in $L^1(0, r+1)$ such that $v_0(s) > 0$ in $(k-1, k]$ and

$$\int_{k-1}^{k} v_0(s)\, ds = \tau_{0,k} - \tau_{0,k-1}, \quad \text{if } k \in \mathcal{K},$$

and $v_0(s) = 0$ in $(k-1, k]$ if $k \in \mathcal{K}_0$. Define

$$\tau_0(s) = \tau_{0,0} + \int_0^s v_0(\sigma)\, d\sigma, \quad s \in [0, r+1].$$

Then $\tau_0(\cdot)$ is absolutely continuous with $\dot{\tau}_0(s) = v_0(s) > 0$ a.e. in $(k-1, k]$ if $k \in \mathcal{K}$, and

$$\tau_0(k) = \tau_{0,k}, \quad k = 0, \dots, r+1.$$

Choose an element $\bar{u}_k \in \Omega_k$ and define for $s \in (k-1, k]$

$$\tilde{x}_0(s) = \begin{cases} x_0(\tau_0(s)) & \text{if } k \in \mathcal{K}, \\ x_0(\tau_{0,k}) & \text{if } k \in \mathcal{K}_0. \end{cases} \qquad \tilde{u}_{0,k}(s) = \begin{cases} u_{0,k}(\tau_0(s)) & \text{if } k \in \mathcal{K}, \\ \bar{u}_k & \text{if } k \in \mathcal{K}_0. \end{cases}$$

From the properties of absolutely continuous functions it follows that $\tilde{u}_{0,k}$ is measurable, and $\tilde{x}_0$ is absolutely continuous on the intervals $(k-1, k]$ and satisfies, if $k \in \mathcal{K}$,

$$\begin{aligned} \dot{\tilde{x}}_0(s) &= \dot{x}_0(\tau_0(s))\,\dot{\tau}_0(s) = v_0(s)\, f_k(\tau_0(s), x_0(\tau_0(s)), u_{0,k}(\tau_0(s))) \\ &= v_0(s)\, f_k(\tau_0(s), \tilde{x}_0(s), \tilde{u}_{0,k}(s)) \quad \text{a.e.} \end{aligned}$$

The result is obviously true also if $k \in \mathcal{K}_0$, since then $v_0 = 0$ and $\tilde{x}_0$ is constant. For the jumps we have

$$\begin{aligned} \Delta\tilde{x}_0|_k &= \tilde{x}_0(k^+) - \tilde{x}_0(k) = x_0(\tau_0(k)^+) - x_0(\tau_0(k)) = x_0(\tau_{0,k}^+) - x_0(\tau_{0,k}) \\ &= \Delta x_0|_{\tau_{0,k}} = J_k(\tau_0(k), x_0(\tau_0(k)), z_{0,k}) = J_k(\tau_0(k), \tilde{x}_0(k), z_{0,k}) \end{aligned}$$

if $k+1 \in \mathcal{K}$, and

$$\tilde{x}_0(k^+) = \tilde{x}_0(k+1) = x_0(\tau_{0,k+1}) = x_0(\tau_{0,k}) + J_k(\tau_{0,k}, x_0(\tau_{0,k}), z_{0,k})$$

if $k+1 \in \mathcal{K}_0$. Define

$$\begin{aligned} \hat{x} &= (\tau, x) \in \mathbb{R} \times \mathbb{R}^n, \quad w_k = (v, u_k) \in \mathbb{R} \times \mathbb{R}^{m_k}, \\ \hat{x}_0(s) &= (\tau_0(s), \tilde{x}_0(s)), \quad w_{0,k}(s) = (v_0(s), \tilde{u}_{0,k}(s)), \\ \hat{f}_k(\hat{x}, w_k) &= (v, v f_k(\tau, x, u_k)), \\ \hat{J}_k(\hat{x}, z_k) &= (0, J_k(\tau, x, z_k)), \\ \hat{\Omega}_k &= (0, \infty) \times \Omega_k \cup \{(0, \bar{u}_k)\}, \\ \hat{\mathcal{U}}_k &= \{w_k(\cdot) = (v(\cdot), u_k(\cdot)) : w_k \text{ is measurable on } (k-1, k],\ w_k(s) \in \hat{\Omega}_k, \\ &\qquad \text{the function } (s, \hat{x}) \mapsto \hat{f}_k(\hat{x}, w_k(s)) \text{ belongs to the class } \mathcal{F}\}. \end{aligned}$$

Then

$$\begin{aligned} \dot{\hat{x}}_0 &= \hat{f}_k(\hat{x}_0(s), w_{0,k}(s)) \text{ a.e. in } (k-1, k], \\ \Delta\hat{x}_0|_k &= \hat{J}_k(\hat{x}_0(k), z_{0,k}). \end{aligned}$$

We also have that $w_{0,k} \in \hat{\mathcal{U}}_k$. This follows from the fact that $\rho \in L^1(I_{0,k})$ implies that the function $s \mapsto v_0(s)\rho(\tau_0(s))$ belongs to $L^1(k-1, k)$.

For $\hat{x}_k = (\tau_k, x_k) \in \mathbb{R} \times \mathbb{R}^n$, $\hat{y}_k = (\sigma_k, y_k) \in \mathbb{R} \times \mathbb{R}^n$, $-q \le i \le p$, we define

$$\hat{g}_i(\hat{x}_0, \hat{x}_1, \hat{y}_1, \ldots, \hat{x}_r, \hat{y}_r, \hat{x}_{r+1}) = g_i(\tau_0, \tau_1, \ldots, \tau_{r+1}, x_0, x_1, y_1, \ldots, x_r, y_r, x_{r+1}).$$

We have

$$\begin{aligned} &\hat{g}_i(\hat{x}_0(0), \hat{x}_0(1), \Delta\hat{x}_0|_1, \ldots, \hat{x}_0(r), \Delta\hat{x}_0|_r, \hat{x}_0(r+1)) \\ &\qquad = g_i(\bar{\tau}_0, x_0(\tau_{0,0}), x_0(\tau_{0,1}), y_{0,1}, \ldots, x_0(\tau_{0,r}), y_{0,r}, x_0(\tau_{0,r+1})) = g_i(e_0), \end{aligned}$$

or $\hat{g}_i(\hat{e}_0) = g_i(e_0)$ with the obvious definition of $\hat{e}_0$.

Let us now consider the problem of minimizing

$$\hat{g}_0(\hat{x}(0), \hat{x}(1), \Delta\hat{x}|_1, \ldots, \hat{x}(r), \Delta\hat{x}|_r, \hat{x}(r+1))$$

when $\hat{x}(\cdot)$ is a solution of

$$\dot{\hat{x}} = \hat{f}_k(\hat{x}, w_k(s)), \quad s \in (k-1, k], \quad w_k(\cdot) \in \hat{\mathcal{U}}_k, \quad k = 1, \ldots, r+1, \tag{5.17}$$

$$\Delta\hat{x}|_k = \hat{J}_k(\hat{x}(k), z_k), \quad k = 1, \ldots, r, \quad \bar{z} \in \bar{Z}, \tag{5.18}$$

subject to the constraints

$$\hat{g}_i(\hat{x}(0), \hat{x}(1), \Delta\hat{x}|_1, \ldots, \hat{x}(r), \Delta\hat{x}|_r, \hat{x}(r+1)) = 0,\ i = 1, \ldots, p, \tag{5.19}$$

$$\hat{g}_i(\hat{x}(0), \hat{x}(1), \Delta\hat{x}|_1, \ldots, \hat{x}(r), \Delta\hat{x}|_r, \hat{x}(r+1)) \le 0,\ i = -q, \ldots, -1, \tag{5.20}$$

We have seen that $\hat{x}_0(\cdot)$ with $w_{0,k}(\cdot)$ and $\bar{z}_0$ satisfies (5.17)–(5.20). We claim that this is an optimal solution. To see this, let $\hat{x}(\cdot) = (\tau(\cdot), \tilde{x}(\cdot))$ be any function sastisfying (5.17)–(5.20), so that (5.17) and (5.18) are satisfied for some $w_k(\cdot) = (v(\cdot), \tilde{u}_k(\cdot)) \in \hat{\mathcal{U}}_k$ and $\bar{z} \in \bar{Z}$. Then $\tau(\cdot)$ is absolutely continuous on $[0, r+1]$, $v \in L^1(0, r+1)$, and

$$\begin{aligned}
&\dot{\tau}(s) = v(s) \ge 0, \\
&\dot{\tilde{x}}(s) = v(s) f_k(\tau(s), \tilde{x}(s), \tilde{u}_k(s)),
\end{aligned}$$

holds a.e. on $(k-1, k]$. Let $\tau_k = \tau(k)$ and

$$\psi_k(t) = \max\{s \in (k-1, k] : \tau(s) = t\} \quad \text{for } t \in I_k = (\tau_{k-1}, \tau_k].$$

(When $k = 1$ we take the corresponding closed intervals $[0, 1]$ and $[\tau_0, \tau_1]$.) The function τ is increasing and may be constant on countably many disjoint intervals $[\alpha_j, \beta_j]$ in $[k-1, k]$. We may assume that $v(s) = 0$, $\tilde{u}_k(s) = \bar{u}_k$ on each $[\alpha_j, \beta_j]$ (change $w_k(\cdot)$ on a set of measure zero, if necessary). Define

$$x(t) = \tilde{x}(\psi_k(t)), \qquad u_k(t) = \tilde{u}_k(\psi_k(t)) \quad \text{for } t \in I_k.$$

If $G \subseteq \mathbb{R}^{m_k}$ is open, then the set $E = \{s \in (k-1, k] : \tilde{u}_k(s) \in G\}$ is measurable, and

$$\{t \in I_k : u_k(t) \in G\} = \tau(E \setminus \cup_j [\alpha_j, \beta_j))$$

is measurable, since τ (being absolutely continuous) maps measurable sets onto measurable sets. Thus $u_k(\cdot)$ is measurable with values in Ω_k. In the same way, $x(\cdot)$ is measurable (it is not difficult to see that it is in fact continuous). We have that $s \le \psi_k(\tau(s))$ for all $s \in (k-1, k]$, and if $s < \psi_k(\tau(s))$, then $v(\cdot) = 0$, $\tilde{x}(\cdot)$ is constant, and $\tilde{u}_k(\cdot) = \bar{u}_k$ on $[s, \psi_k(\tau(s))]$. Thus

$$\tilde{x}(s) = \tilde{x}(\psi_k(\tau(s))) = x(\tau(s)), \qquad \tilde{u}_k(s) = \tilde{u}_k(\psi_k(\tau(s))) = u_k(\tau(s)) \text{ for all } s.$$

Assume that $\tau_{k-1} < \tau_k$. For any function $F \in L^1(I_k)$ we have the formula

$$\int_{\tau(k-1)}^{\tau(s)} F(\tau)\, d\tau = \int_{k-1}^{s} F(\tau(\sigma)) v(\sigma)\, d\sigma, \quad s \in (k-1, k]. \tag{5.21}$$

A proof of this is found, e.g., in [16, p. 377]. An application of (5.21) gives

$$\begin{aligned}
\tilde{x}(s) &= \tilde{x}((k-1)^+) + \int_{k-1}^{s} v(\sigma) f_k(\tau(\sigma), \tilde{x}(\sigma), \tilde{u}_k(\sigma))\, d\sigma \\
&= \tilde{x}((k-1)^+) + \int_{k-1}^{s} v(\sigma) f_k(\tau(\sigma), x(\tau(\sigma)), u_k(\tau(\sigma)))\, d\sigma \\
&= \tilde{x}((k-1)^+) + \int_{\tau(k-1)}^{\tau(s)} f_k(\tau, x(\tau), u_k(\tau))\, d\tau.
\end{aligned}$$

Thus

$$x(t) = \tilde{x}(\psi_k(t)) = \tilde{x}((k-1)^+) + \int_{\tau_{k-1}}^{t} f_k(\tau, x_k(\tau), u_k(\tau))\, d\tau,$$

so that $x(\cdot)$ is absolutely continuous and satisfies

$$\dot{x}(t) = f_k(t, x(t), u_k(t)) \quad \text{a.e. in } I_k.$$

Since $w_k \in \hat{\mathcal{U}}_k$, the norms of $\hat{f}_k$ and $\frac{\partial \hat{f}_k}{\partial \hat{x}}$ are estimated on $(k-1,k]$ by a function $\hat{\rho}_k \in L^1(k-1,k)$. For almost all $t \in I_k$, $v(\psi_k(t)) > 0$, and the norms of f_k, $\frac{\partial f_k}{\partial t}$, $\frac{\partial f_k}{\partial x}$ are estimated by

$$\frac{1}{v(\psi_k(t))}\hat{\rho}_k(\psi_k(t)) = \rho_k(t).$$

From (5.21), which holds for non-negative measurable functions F, and the fact that $\psi_k(\tau(s)) = s$ in $E = \{s \in (k-1,k] : v(s) > 0\}$, we get $\int_{\tau_{k-1}}^{\tau_k} \rho_k(t)\,dt = \int_E \hat{\rho}_k(s)\,ds < \infty$. Thus $u_k(\cdot) \in \mathcal{U}_k$. If $\tau_k < \tau_{k+1}$ then

$$x(\tau_k^+) = \tilde{x}(k^+) = \tilde{x}(k) + J_k(\tau(k), \tilde{x}(k), z_k) = x(\tau_k) + J_k(\tau_k, x(\tau_k), z_k),$$

and if $\tau_k = \tau_{k+1}$ then

$$x(\tau_{k+1}) = \tilde{x}(k+1) = \tilde{x}(k^+) = x(\tau_k) + J_k(\tau_k, x(\tau_k), z_k) = x(\tau_k^+).$$

Therefore, (5.14)–(5.15) are satisfied. Finally,

$$\begin{aligned} g_i(\bar{\tau}, x(\tau_0), x(\tau_1), \Delta x|_{\tau_1}, \ldots, x(\tau_r), \Delta x|_{\tau_r}, x(\tau_{r+1})) \\ = \hat{g}_i(\hat{x}(0), \hat{x}(1), \Delta\hat{x}|_1, \ldots, \hat{x}(r), \Delta\hat{x}|_r, \hat{x}(r+1)) \quad \text{for } i = -q, \ldots, p. \end{aligned}$$

Thus x satisfies the constraints $g_i = 0$ for $i > 0$, $g_i \le 0$ for $i < 0$, and (since $(\bar{\tau}_0, x_0)$ is optimal)

$$\hat{g}_0(\ldots) = g_0(\ldots) \ge g_0(e_0) = \hat{g}_0(\hat{e}_0).$$

Now we can apply Theorem 3 (including the Remark) to the solution $(w_{0,k}(\cdot), \bar{z}_0, \hat{x}_0(\cdot))$ of the problem of minimizing $\hat{g}_0$ subject to (5.17)–(5.20). It follows that there exist numbers $\lambda_{-q}, \ldots, \lambda_p$, not all zero, and a piecewise absolutely continuous row vector function $\hat{\eta}(\cdot)$ such that, with $\hat{G} = \sum_{i=-q}^{p} \lambda_i \hat{g}_i$ and $\hat{H}_k(s, \hat{x}, w_k) = \hat{\eta}(s)\hat{f}_k(\hat{x}, w_k)$, we have

$$\begin{aligned} \dot{\hat{\eta}}(s) &= -\frac{\partial \hat{H}_k}{\partial \hat{x}}(\hat{x}_0(s), w_{0,k}(s)) \quad \text{a.e. on } (k-1,k], \quad k = 1, \ldots, r+1, \\ \lambda_i &\le 0 \quad \text{for } -q \le i \le 0, \\ \lambda_i \hat{g}_i(\hat{e}_0) &= 0 \quad \text{for } -q \le i \le -1, \\ \Delta\hat{\eta}|_k &= -\big(\hat{\eta}(k^+) + \frac{\partial \hat{G}}{\partial \hat{y}_k}(\hat{e}_0)\big)\frac{\partial \hat{J}_k}{\partial \hat{x}}(\hat{x}_0(k), z_{0,k}) - \frac{\partial \hat{G}}{\partial \hat{x}_k}(\hat{e}_0), \quad k = 1, \ldots, r, \\ \hat{\eta}(0) &= -\frac{\partial \hat{G}}{\partial \hat{x}_0}(\hat{e}_0), \\ \hat{\eta}(r+1) &= \frac{\partial \hat{G}}{\partial \hat{x}_{r+1}}(\hat{e}_0), \\ \hat{H}_k(s, \hat{x}_0(s), w_{0,k}(s)) &= \max_{w_k \in \hat{\Omega}_k} \hat{H}_k(s, \hat{x}_0(s), w_k) \quad \text{a.e. on } (k-1,k], \end{aligned}$$

$$\big(\hat{\eta}(k^+) + \frac{\partial \hat{G}}{\partial \hat{y}_k}(\hat{e}_0)\big)\hat{J}_k(\hat{x}_0(k), z_{0,k}) = \max_{z_k \in Z_k} \big(\hat{\eta}(k^+) + \frac{\partial \hat{G}}{\partial \hat{y}_k}(\hat{e}_0)\big)\hat{J}_k(\hat{x}_0(k), z_k), \quad k = 1, \ldots, r.$$

If $k \in \mathcal{K}$, i.e., $m(I_{0,k}) > 0$, we write

$$(\eta_0(t), \eta(t)) = \hat{\eta}(\tau_0^{-1}(t)), \quad t \in I_{0,k} = (\tau_{0,k-1}, \tau_{0,k}].$$

The functions η_0 and η are absolutely continuous on $I_{0,k}$ and satisfy

$$\begin{aligned} \dot{\eta}_0(t) &= -\eta(t)\frac{\partial f_k}{\partial t}(t, x_0(t), u_{0,k}(t)) = -\frac{\partial H_k}{\partial \tau}\Big|_0, \\ \dot{\eta}(t) &= -\eta(t)\frac{\partial f_k}{\partial x}(t, x_0(t), u_{0,k}(t)) = -\frac{\partial H_k}{\partial x}\Big|_0 \end{aligned}$$

a.e. on $I_{0,k}$. With $s = \tau_0^{-1}(t)$ the first maximum condition can be written

$$v_0(\tau_0^{-1}(t))[\eta_0(t) + \eta(t)f_k(t, x_0(t), u_{0,k}(t))] = \max_{(v,u_k)\in\hat{\Omega}_k} v[\eta_0(t) + \eta(t)f_k(t, x_0(t), u_k)] \quad \text{a.e.}$$

Since $v_0(\tau_0^{-1}(t)) > 0$, this means that almost everywhere on $I_{0,k}$

$$\begin{aligned} \eta_0(t) + \eta(t)f_k(t, x_0(t), u_{0,k}(t)) &= 0, \\ \eta_0(t) + \eta(t)f_k(t, x_0(t), u_k) &\le 0 \quad \text{for all } u_k \in \Omega_k, \end{aligned} \tag{5.22}$$

Since the left-hand side of (5.22) (as a function of t for fixed u_k) is continuous from the left in $(\tau_{0,k-1}, \tau_{0,k}]$ (5.22) is true for all $t \in I_{0,k}$. With $H_{0,k}(t, u_k) = \eta(t)f_k(t, x_0(t), u_k)$ and $M_k(t) = \sup_{u_k\in\Omega_k} H_{0,k}(t, u_k)$ we have

$$\eta_0(t) + M_k(t) \le 0 \text{ for all } t, \qquad \eta_0(t) + M_k(t) = \eta_0(t) + H_{0,k}(t, u_{0,k}(t)) = 0 \ \text{ a.e.}$$

If $k \in \mathcal{K}_0$, i.e., $m(I_{0,k}) = 0$, $\hat{\eta}$ is constant on $(k-1, k]$; we denote this constant by $(\eta_{0,k}, \eta_k)$. On the t-side we write

$$\eta_0(\tau_{0,k-1}^+) = \eta_0(\tau_{0,k}) = \eta_{0,k}, \qquad \eta(\tau_{0,k-1}^+) = \eta(\tau_{0,k}) = \eta_k.$$

The maximum condition becomes (since $v_0 = 0$ and $\hat{x}_0$ is constant on $(k-1, k]$)

$$\max_{(v,u_k)\in\hat{\Omega}_k} v[\eta_{0,k} + \eta_k f_k(\tau_{0,k}, x_0(\tau_{0,k}), u_k)] = 0,$$

so that

$$\eta_{0,k} + M_k \le 0, \quad \text{where } M_k = \sup_{u_k\in\Omega_k} \eta_k f_k(\tau_{0,k}, x_0(\tau_{0,k}), u_k) = M_k(\tau_{0,k}).$$

The conditions at the points $\tau_{0,k}$ and the second maximum condition become (with $T_0 = \tau_{0,0}$, $T_1 = \tau_{0,r+1}$)

$$\begin{aligned} \eta_0(T_0) &= -\frac{\partial G}{\partial \tau_0}(e_0), \qquad \eta(T_0) = -\frac{\partial G}{\partial x_0}(e_0), \\ \eta_0(T_1) &= \frac{\partial G}{\partial \tau_{r+1}}(e_0), \qquad \eta(T_1) = \frac{\partial G}{\partial x_{r+1}}(e_0), \\ \Delta\eta_0|_{\tau_{0,k}} &= -\big(\eta(\tau_{0,k}^+) + \frac{\partial G}{\partial y_k}(e_0)\big)\frac{\partial J_k}{\partial \tau}\Big|_{0,k} - \frac{\partial G}{\partial \tau_k}(e_0), \\ \Delta\eta|_{\tau_{0,k}} &= -\big(\eta(\tau_{0,k}^+) + \frac{\partial G}{\partial y_k}(e_0)\big)\frac{\partial J_k}{\partial x}\Big|_{0,k} - \frac{\partial G}{\partial x_k}(e_0), \\ \big(\eta(\tau_{0,k}^+) + \frac{\partial G}{\partial y_k}(e_0)\big)J_k|_{0,k} &= \max_{z_k\in Z_k}\big(\eta(\tau_{0,k}^+) + \frac{\partial G}{\partial y_k}(e_0)\big)J_k(\tau_{0,k}, x_0(\tau_{0,k}), z_k), \end{aligned}$$

for $k = 1, \dots, r$.

Now assume that (5.16) holds and suppose that $\eta_0(t') + M_k(t') < 0$ for some $t' \in I_{0,k}$, $k \in \mathcal{K}$. We have, since $u_{0,k}(t) \in \Omega_k$,

$$\eta_0(t') + H_{0,k}(t', u_{0,k}(t)) \le \eta_0(t') + M_k(t') < 0 \quad \text{for all } t \in I_{0,k}.$$

Since η_0 is continuous from the left, we have

$$\eta_0(t) < \eta_0(t') + a, \qquad \text{where } a = -\frac{1}{2}[\eta_0(t') + M_k(t')] > 0$$

for all t in a neighbourhood to the left of t'. For almost all t in this neighbourhood $H_{0,k}(t, u_{0,k}(t)) = -\eta_0(t)$, and

$$H_{0,k}(t, u_{0,k}(t)) - H_{0,k}(t', u_{0,k}(t)) > -\eta_0(t') - a + 2a + \eta_0(t') = a. \tag{5.23}$$

For all t and almost all τ in $I_{0,k}$ we have

$$\frac{\partial H_{0,k}}{\partial t}(\tau, u_{0,k}(t)) = \eta(\tau)\frac{\partial f_k}{\partial t}(\tau, x_0(\tau), u_{0,k}(t)) + \eta(\tau)\frac{\partial f_k}{\partial x}(\tau, x_0(\tau), u_{0,k}(t))f_k(\tau, x_0(\tau), u_{0,k}(\tau)) \\ - \eta(\tau)\frac{\partial f_k}{\partial x}(\tau, x_0(\tau), u_{0,k}(\tau))f(\tau, x_0(\tau), u_{0,k}(t)).$$

Since $u_{0,k} \in \mathcal{U}_k$, there is a $\rho \in L^1(I_{0,k})$ such that

$$\Big|\frac{\partial f_k}{\partial t}(\tau, x_0(\tau), u_{0,k}(t))\Big| \leq \rho(t) \quad \text{for all } t, \tau \in I_{0,k}.$$

If $|\eta(t)| \leq C$, then it it follows from (5.16) that

$$\Big|\frac{\partial H_{0,k}}{\partial t}(\tau, u_{0,k}(t))\Big| \leq C[\rho(t) + \rho_1(t) + \rho_2(\tau) + \rho_1(\tau) + \rho_2(t)] = C[\rho_3(t) + \rho_4(\tau)],$$

where $\rho_3 = \rho + \rho_1 + \rho_2 \in L^1(I_{0,k})$, $\rho_4 = \rho_1 + \rho_2 \in L^1(I_{0,k})$. From this and (5.23) we obtain

$$a \leq \left|\int_{t'}^{t} \frac{\partial H_{0,k}}{\partial t}(\tau, u_{0,k}(t))\, d\tau\right| \leq C|t - t'|\rho_3(t) + C\left|\int_{t'}^{t} \rho_4(\tau)\, d\tau\right|$$

for t in some neighbourhood to the left of t'. If $t \neq t'$ is so close to t' that $C\left|\int_{t'}^{t} \rho_4(\tau)\, d\tau\right| \leq \frac{a}{2}$, we get

$$\frac{a}{2|t - t'|} \leq C\rho_3(t),$$

which contradicts the fact that $\rho_3 \in L^1(I_{0,k})$. Thus $\eta_0(t) + M_k(t) = 0$ for all $t \in I_{0,k}$, if (5.16) holds. This also holds at T_0 if $1 \in \mathcal{K}$. The theorem is proved.

Remark. If we assume that there is a function $\tilde{\rho} \in L^2(I_{0,k})$ such that

$$\big|f_k(s, x_0(s), u_{0,k}(t))\big| + \Big|\frac{\partial f_k}{\partial x}(s, x_0(s), u_{0,k}(t))\Big| \leq \tilde{\rho}(t) \quad \text{for all } s, t \in I_{0,k},$$

then (5.16) is satisfied with $\rho_1 = \rho_2 = \frac{1}{2}\tilde{\rho}^2$.

6 A multiprocess problem

Some control systems may require different descriptions on different time intervals, such as a multistage rocket or a robot arm picking up or dropping a load. We would then have one control process $\dot{x}_k(t) = f_k(t, x_k(t), u_k(t))$ on each interval $t \in (\tau_{k-1}, \tau_k)$, and together they constitute a multiprocess (see [3] and [4]). We also have some relations between the endpoint values. The term hybrid system has been used for a more general situation where the transition from one state space to another is determined by some discrete mechanism. See [5] and [15] for a general description of hybrid systems.

In this section we consider a multiprocess with equality and inequality functional constraints on $x_k(\tau_{k-1})$ and $x_k(\tau_k)$ and the problem of minimizing a functional of the same type. We can derive necessary conditions for optimality by an application of the theorems in the previous section. A similar problem is treated in [3], where the proofs use methods from non-smooth analysis. In the case of variable τ_k we need higher regularity in the t-dependence than in [3], but as a result we get more information, in particular that the maximized Hamiltonian is (piecewise) absolutely continuous. A problem of this type but restricted to piecewise continuous controls is treated in [6].

Let us assume that on each fixed interval $I_k = [\tau_{k-1}, \tau_k]$, $k = 1, \ldots, r+1$, we have a control system

$$\dot{x}_k = f_k(t, x_k, u_k(t)), \tag{6.1}$$

$$u_k(t) \in \Omega_k(t), \tag{6.2}$$

satisfying Assumption 2 in Section 4 on $I_k \times \mathbb{R}^{n_k} \times \mathbb{R}^{m_k}$. The set of admissible controls is denoted by $\mathcal{U}_k$. We have a number of constraints of the form

$$g_i(x_1(\tau_0), x_1(\tau_1), x_2(\tau_1), \dots, x_{r+1}(\tau_r), x_{r+1}(\tau_{r+1})) = 0, \quad i = 1, \dots, p, \tag{6.3}$$
$$g_i(x_1(\tau_0), x_1(\tau_1), x_2(\tau_1), \dots, x_{r+1}(\tau_r), x_{r+1}(\tau_{r+1})) \le 0, \quad i = -q, \dots, -1. \tag{6.4}$$

We want to minimize a functional

$$g_0(x_1(\tau_0), x_1(\tau_1), x_2(\tau_1), \dots, x_{r+1}(\tau_r), x_{r+1}(\tau_{r+1})). \tag{6.5}$$

The functions g_i are defined on $\mathbb{R}^{n_1} \times \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_2} \times \cdots \times \mathbb{R}^{n_{r+1}} \times \mathbb{R}^{n_{r+1}}$ and are continuously differentiable in all variables. The argument of g_i is denoted by

$$(x_1^0, x_1^1, x_2^0, \dots, x_{r+1}^0, x_{r+1}^1).$$

A total solution is a sequence $(x_1, u_1, \dots, x_{r+1}, u_{r+1})$, where each pair (x_k, u_k) satisfies (6.1)–(6.2) on I_k, and the constraints (6.3)–(6.4) are satisfied. Assume that $(x_{0,1}, u_{0,1}, x_{0,2}, \dots, x_{0,r+1}, u_{0,r+1})$ is an optimal solution, i.e., a solution that minimizes (6.5). As an application of Theorem 3 we obtain the following result (where $e_0 = (x_{0,1}(\tau_0), x_{0,1}(\tau_1), \dots, x_{0,r+1}(\tau_r), x_{0,r+1}(\tau_{r+1})))$:

Theorem 5. *There exist numbers $\lambda_{-q}, \dots, \lambda_0, \dots, \lambda_p$, not all zero, and absolutely continuous functions η_k on I_k, $k = 1, \dots, r+1$, such that, with $G = \sum_{i=-q}^{p} \lambda_i g_i$ and $H_k(t, x_k, u_k) = \eta_k(t) f_k(t, x_k, u_k)$,*

$$\begin{aligned}
\dot{\eta}_k(t) &= -\frac{\partial H_k}{\partial x_k}(t, x_{0,k}(t), u_{0,k}(t)) \quad \textit{a.e. on } I_k, \\
\eta_k(\tau_{k-1}) &= -\frac{\partial G}{\partial x_k^0}(e_0), \\
\eta_k(\tau_k) &= \frac{\partial G}{\partial x_k^1}(e_0), \\
H_k(t, x_{0,k}(t), u_{0,k}(t)) &= \max_{u_k \in \Omega_k(t)} H_k(t, x_{0,k}(t), u_k) \quad \textit{a.e. on } I_k, \\
\lambda_i &\le 0 \quad \textit{for } -q \le i \le 0, \\
\lambda_i g_i(e_0) &= 0 \quad \textit{for } -q \le i \le -1.
\end{aligned}$$

Proof. Let $\bar{x} = (x_1, x_2, \dots, x_{r+1}) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \cdots \times \mathbb{R}^{n_{r+1}} = \mathbb{R}^N$ ($N = \sum_{k=1}^{r+1} n_k$), and let

$$\bar{f}_k(t, \bar{x}, u_k) = (0, \dots, f_k(t, x_k, u_k), \dots, 0) \in \mathbb{R}^N,$$

where only block number k is different from 0. Let us study the system

$$\dot{\bar{x}} = \bar{f}_k(t, \bar{x}, u_k(t)) \quad \text{in } I_k, \quad k = 1, \dots, r+1 \tag{6.6}$$

with no jumps at τ_k. If

$$\bar{x}_k = (x_{1,k}, x_{2,k}, \dots, x_{r+1,k}) \in \mathbb{R}^N, \quad k = 0, \dots, r+1,$$

define, for $-q \le i \le p$,

$$\bar{g}_i(\bar{x}_0, \bar{x}_1, \dots, \bar{x}_{r+1}) = g_i(x_{1,0}, x_{1,1}, \dots, x_{r+1,r}, x_{r+1,r+1}).$$

Then $\bar{x}_0 = (x_{0,1}, \dots, x_{0,k+1})$ with $(u_{0,1}, \dots, u_{0,r+1})$ is a solution of the problem of minimizing

$$\bar{g}_0(\bar{x}(\tau_0), \dots, \bar{x}(\tau_{r+1}))$$

subject to (6.1) and

$$\begin{aligned}
\bar{g}_i(\bar{x}(\tau_0), \dots, \bar{x}(\tau_{r+1})) &= 0, \quad i = 1, \dots, p, \\
\bar{g}_i(\bar{x}(\tau_0), \dots, \bar{x}(\tau_{r+1})) &\le 0, \quad i = -q, \dots, 0.
\end{aligned}$$

Let us apply Theorem 3 (with the Remark) to this problem. Thus, there exist numbers $\lambda_{-q}, \dots \lambda_p$, not all zero, and a piecewise absolutely continuous function $\bar{\eta}$ such that, with $\bar{G} = \sum_{i=-q}^{p} \lambda_i \bar{g}_i$ and $\bar{H}_k(t, \bar{x}, u_k) = \bar{\eta}(t) \bar{f}_k(t, \bar{x}, u_k)$,

$$
\begin{aligned}
\dot{\bar{\eta}} &= -\frac{\partial \bar{H}_k}{\partial \bar{x}}\Big|_0 \quad \text{on } (\tau_{k-1}, \tau_k],\ k = 1, \dots, r+1, \\
\Delta \bar{\eta}|_{\tau_k} &= -\frac{\partial \bar{G}}{\partial \bar{x}_k}\Big|_0, \quad k = 1, \dots, r, \\
\bar{\eta}(T_0) &= -\frac{\partial \bar{G}}{\partial \bar{x}_0}\Big|_0, \\
\bar{\eta}(T_1) &= \frac{\partial \bar{G}}{\partial \bar{x}_{r+1}}\Big|_0, \\
\bar{H}_k(t, \bar{x}_0(t), u_{0,k}(t)) &= \max_{u_k \in \Omega_k(t)} \bar{H}_k(t, \bar{x}_0(t), u_k) \quad \text{a.e. in } I_k \\
\lambda_i &\le 0 \quad \text{for } -q \le i \le 0, \\
\lambda_i \bar{g}_i(\bar{e}_0) &= 0 \quad \text{for } -q \le i \le -1.
\end{aligned}
$$

Write

$$\bar{\eta} = (\eta_1, \eta_2, \dots, \eta_{r+1}),$$

and redefine η_k at τ_k $(1 \le k \le r)$ so that it becomes continuous from the right there. Then $\bar{H}_k(t, \bar{x}, u_k) = \eta_k(t) f_k(t, x_k, u_k) = H_k(t, x_k, u_k)$ for $t \in I_k$, η_k satisfies $\dot{\eta}_k = -\frac{\partial H_k}{\partial x_k}$ on I_k, and η_k is constant on the other intervals. We have $\Delta \eta_j|_{\tau_k} = -\frac{\partial \bar{G}}{\partial x_{j,k}} = 0$ except when $j = k$ and $j = k+1$, $\Delta \eta_k|_{\tau_k} = -\frac{\partial \bar{G}}{\partial x_{k,k}} = -\frac{\partial G}{\partial x_k^1}$, and $\Delta \eta_{k+1}|_{\tau_k} = -\frac{\partial \bar{G}}{\partial x_{k+1,k}} = -\frac{\partial G}{\partial x_{k+1}^0}$, $k = 1, \dots, r$. From this the statements in the theorem follow.

In the case of variable τ_k we can apply Theorem 4 in the same way. Each function f_k is now continuously differentiable w.r.t. (t, x_k), and Ω_k is constant. The functions g_i may depend explicitly on $\bar{\tau} = (\tau_0, \dots, \tau_{r+1})$ also. Assume that $(\bar{\tau}_0, x_{0,1}, u_{0,1}, \dots, x_{0,r+1}, u_{0,r+1})$, where $\bar{\tau}_0 = (\tau_{0,0}, \tau_{0,1}, \dots, \tau_{0,r+1})$, is a solution. Let $I_{0,k} = [\tau_{0,k-1}, \tau_{0,k}]$, and

$$\mathcal{K} = \{k \in \{1, \dots, r+1\} : m(I_{0,k}) > 0\}, \qquad \mathcal{K}_0 = \{k \in \{1, \dots, r+1\} : m(I_{0,k}) = 0\}.$$

Assume also that there are functions $\tilde{\rho}_k \in L^2(I_{0,k})$, $k \in \mathcal{K}$, such that

$$\Big| f_k(s, x_{0,k}(s), u_{0,k}(t)) \Big| + \Big| \frac{\partial f_k}{\partial x_k}(s, x_{0,k}(s), u_{0,k}(t)) \Big| \le \tilde{\rho}_k(t) \quad \text{for all } s, t \in I_{0,k}.$$

Define

$$H_k(\tau, x_k, u_k, \eta_k) = \eta_k f_k(\tau, x_k, u_k),$$

where η_k is a row vector in $\mathbb{R}^{n_k}$. Let

$$e_0 = (\bar{\tau}_0, x_{0,1}(\tau_{0,0}), x_{0,1}(\tau_{0,1}), \dots, x_{0,r+1}(\tau_{0,r}), x_{0,r+1}(\tau_{0,r+1})).$$

Theorem 6. *There exist numbers $\lambda_{-q}, \dots, \lambda_p$, not all zero, and functions $\eta_{0,k}$ and η_k on $I_{0,k}$ such that, with $G = \sum_{i=-q}^{p} \lambda_i g_i$, the following holds.*

$$
\begin{aligned}
\lambda_i &\le 0 \quad \textit{for } i = -q, \dots, 0, \\
\lambda_i g_i(e_0) &= 0 \quad \textit{for } i = -q, \dots, -1.
\end{aligned}
$$

If $k \in \mathcal{K}$, then $\eta_{0,k}$ and η_k are absolutely continuous on $I_{0,k}$ and satisfy

$$
\begin{aligned}
\dot{\eta}_{0,k}(t) &= -\frac{\partial H_k}{\partial \tau}\Big|_{0,k} \quad \textit{a.e.}, \\
\dot{\eta}_k(t) &= -\frac{\partial H_k}{\partial x_k}\Big|_{0,k} \quad \textit{a.e.},
\end{aligned}
$$

where $\cdot|_{0,k}$ *means evaluation at* $(t, x_{0,k}(t), u_{0,k}(t), \eta_k(t))$*. If*

$$M_k(t) = \sup_{u_k \in \Omega_k} H_k(t, x_{0,k}(t), u_k, \eta_k(t)),$$

then

$$M_k(t) = -\eta_{0,k}(t) \qquad \text{for all } t \in I_{0,k},$$
$$M_k(t) = H_k(t, x_{0,k}(t), u_{0,k}(t), \eta_k(t)) \qquad \text{a.e. in } I_{0,k}.$$

If $k \in \mathcal{K}_0$*, then* $\eta_{0,k}(\tau_{0,k-1}) = \eta_{0,k}(\tau_{0,k}) = \eta_{0,k}$*,* $\eta_k(\tau_{0,k-1}) = \eta_k(\tau_{0,k}) = \eta_k$*, and*

$$M_k = \sup_{u_k \in \Omega_k} H_k(\tau_{0,k}, x_{0,k}, u_k, \eta_k) \le -\eta_{0,k}.$$

Furthermore,

$$\eta_k(\tau_{0,k-1}) = -\frac{\partial G}{\partial x_k^0}(e_0), \qquad \eta_k(\tau_{0,k}) = \frac{\partial G}{\partial x_k^1}(e_0) \qquad \text{for } 1 \le k \le r+1,$$
$$\eta_{0,1}(\tau_{0,0}) = -\frac{\partial G}{\partial \tau_0}(e_0), \qquad \eta_{0,r+1}(\tau_{0,r+1}) = \frac{\partial G}{\partial \tau_{r+1}}(e_0),$$
$$\eta_{0,k+1}(\tau_{0,k}) - \eta_{0,k}(\tau_{0,k}) = -\frac{\partial G}{\partial \tau_k}(e_0) \qquad \text{for } 1 \le k \le r.$$

Remark. Consider the same problem but with integral terms $\sum_{k=1}^{r+1} \int_{\tau_{k-1}}^{\tau_k} f_k^0(t, x_k(t), u_k(t))\, dt$ added to the cost functional, where f_k^0 has the same properties as f_k. Theorems 5 and 6 still hold, except that H_k is defined as $\lambda_0 f_k^0(t, x_k, u_k) + \eta_k f_k(t, x_k, u_k)$.

References

[1] D. D. Bainov and P. S. Simeonov, *Systems with Impulse Effects: Stability Theory and Applications*, Ellis Horwood Ltd., 1989.

[2] S. A. Belbas and W. H. Schmidt, *Optimal control of Volterra equations with impulses*, Applied Mathematics and Computation, 166 (2005), 696-723.

[3] F. H. Clarke and R. B. Vinter, *Optimal multiprocesses*, SIAM J. Control Optim., 27 (1989), 1072–1091.

[4] F. H. Clarke and R. B. Vinter, *Applications of optimal multiprocesses*, SIAM J. Control Optim., 27 (1989), 1048–1071.

[5] M. Garavello and B. Piccoli, *Hybrid necessary principle*, SIAM J. Control Optim., 43 (2005), 1867–1887.

[6] W. H. Getz and D. H. Martin, *Optimal control systems with state variable jump discontinuities*, Journal of Optimization Theory and Applications, 31 (1980), 195–205.

[7] H. Halkin, *Necessary conditions in mathematical programming and optimal control theory*, Optimal Control Theory and its Applications, I, Lecture Notes in Economics and Mathematical Systems, 105, Springer-Verlag, Berlin-Heidelberg-New York, 1974.

[8] H. Halkin, *Necessary conditions for optimal control problems with differentiable or nondifferentiable data*, Mathematical Control Theory, Lecture Notes in Mathematics, 680, Springer-Verlag, Berlin-Heidelberg-New York, 1978.

[9] H. Halkin and L. W. Neustadt, *General necessary conditions for optimization problems*, Proc. Nat. Acad. Sciences, 56 (1966), 1066–1071.

[10] K. Holmåker, *Optimal Control Theory, Part One*, Department of Mathematics, Chalmers University of Technology and the University of Göteborg, 1992.

[11] K. Holmåker, *Optimal Control Theory, Part Two, Measurable Set-valued Functions*, Department of Mathematics, Chalmers University of Technology and the University of Göteborg, 1992.

[12] B. M. Miller and E. Ya. Rubinovich, Impulsive control in continuous and discrete-continuous systems, Kluwer, New York, 2003.

[13] L. W. Neustadt, *Optimization. A theory of necessary conditions*, Princeton University Press, New York, 1976.

[14] R. Rempala and J. Zabczyk, *On the maximum principle for deterministic impulse control problems*, Journal of Optimization Theory and Applications, 59 (1988), 281–288.

[15] H. J. Sussmann, *A maximum principle for hybrid optimal control problems*, Proc. 38th IEEE Int. Conf. Decision and Control, Phoenix, AZ, 1999, 425–430.

[16] E. C. Titchmarsh, *The Theory of Functions*, Oxford University Press, 1939.