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

Hybrid Necessary Principle

Mauro Garavello and Benedetto Piccoli

Abstract— **We consider a hybrid control system and general optimal control problems for this system. We suppose that the switching strategy imposes restrictions on control sets and we provide necessary conditions for an optimal hybrid trajectory, stating a Hybrid Necessary Principle (HNP). Our result generalizes various necessary principles available in the literature.**

Index Terms— **Optimal control, necessary conditions, switching strategy**

I. INTRODUCTION

This paper deals with optimal control problems for hybrid systems. Our definition of hybrid system is the one used in [8]. Roughly speaking a hybrid system is a collection of control systems called locations, possibly defined on different manifolds, and an automaton that rules the switchings between locations. The term hybrid is commonly used to indicate the presence of both continuous and discrete dynamics and in our case the continuous part is given by location controlled dynamics and the discrete part by the automaton. An optimal control problem is obtained assigning Lagrangian running costs on each location and final and switching costs. Recently optimization problems for hybrid systems have attracted a lot of attention, thus both theoretical results and applications were developed, see [2], [3], [7], [14]. For general theory of hybrid systems we refer to [1], [5]. For an optimal (classical) control problem, the main tool toward the construction of optimal trajectories, and then optimal synthesis, is the celebrated Pontryagin Maximum Principle (PMP). The strength of PMP is evident when it permits to describe completely the structure of optimal trajectories and obtain a finite dimensional reduction of the problem. A Hybrid Maximum Principle (HMP) was developed in [11], see also [8], [9]. A key role is played by the switching mechanism that permits to pass from one location to another with possible restrictions on state, time to spend in next location and feasible controls for next location. The first two kinds of restrictions do not affect the general strategy of PMP and a HMP can be proved in a similar way. However the restriction on usable controls, after location switchings, dramatically changes the possibility of constructing "needle variations" that are the basic ingredient to prove PMP, and HMP is no more applicable. More precisely, a classical needle variation is no longer prolongable after a location switching time, therefore a new class of "admissible needle variations" must be introduced. As a first example (in Section III), to construct such a kind of variation, one can prolong the family of trajectories, originating from a needle variation, via the choice of a suitable family of controls having continuity and weak differentiability (in L^1) properties.

Then in Section IV we introduce a general concept of "map of variations". The basic request is again to have weak differentiability properties, but now in the space of bounded Radon measures, seen as the dual of the space of continuous functions. We are able to prove the Hybrid Necessary Principle (HNP), using results from [6], [10], [12]. The word maximum in this context disappears, since necessary conditions are no more written in a supremum form.

Section II gives basic definitions of *Hybrid Systems* and states HMP, Section III deals with *Admissible Needle Variations* and with simple necessary conditions for hybrid systems where the switching strategy affects the choice of controls. Finally Section IV introduce the concept of *Map of Variations* and states HNP.

II. BASIC DEFINITIONS AND HMP

We start introducing the definition of hybrid system.

Definition 1: A *hybrid control system* is a 7-tuple $\Sigma = (\mathcal{Q}, M, U, f, \mathcal{U}, J, \mathcal{S})$ such that

1) $\mathcal{Q}$ is a finite set;
2) $M = \{M_q\}_{q\in\mathcal{Q}}$ is a family of smooth manifolds, indexed by $\mathcal{Q}$;
3) $U = \{U_q\}_{q\in\mathcal{Q}}$ is a family of sets;
4) $f = \{f_q\}_{q\in\mathcal{Q}}$ is a family of maps
$$f_q : M_q \times U_q \mapsto TM_q$$
(TM_q is the tangent bundle of M_q), such that $f_q(x,u) \in T_xM_q$ for every $(x,u) \in M_q \times U_q$;
5) $\mathcal{U} = \{\mathcal{U}_q\}_{q\in\mathcal{Q}}$ is a family of sets $\mathcal{U}_q$ whose members are maps $u : Dom(u) \to U_q$, defined on some interval $Dom(u) \subset \mathbb{R}$;
6) $J = \{J_q\}_{q\in\mathcal{Q}}$ is a family of subintervals of $\mathbb{R}^+$;
7) $\mathcal{S}$ is a subset of
$$\{(q,x,q',x',u(\cdot),\tau) : q,q' \in \mathcal{Q}, x \in M_q, x' \in M_{q'}, u(\cdot) \in \mathcal{U}_{q'}, \tau \in J_{q'}\}.$$

The members of $\mathcal{Q}$ are called *locations* and represent the states of the automaton. The families M, U, are, respectively, the *family of state spaces* and the *family of control spaces* of Σ. For each q, the manifold M_q, the set U_q, the map f_q and the set $\mathcal{U}_q$ are, respectively, the *state space*, the *control space*,

*Work (partially) done in the framework of the HYCON Network of Excellence, contract number FP6-IST-511368.

B. Piccoli is with I.A.C., Viale del Policlinico 137, 00161 Roma, Italy piccoli@iac.rm.cnr.it

M. Garavello is with Dipartimento di Matematica e Applicazioni, Università di Milano Bicocca, via Cozzi 53, 20125 Milano, Italy mauro.garavello@unimib.it

0-7803-9568-9/05/$20.00 ©2005 IEEE

the *controlled dynamical law* and the *class of admissible controls* at location q.

The system evolves in a location q according to the corresponding controlled dynamic and then switches as prescribed by $\mathcal{S}$. The intervals J_q indicate the lengths of time intervals on which the system can stay in location q. So, for example, if $J_q = [0, +\infty[$ then the system can evolve in location q on every interval of time.

For $q, q' \in \mathcal{Q}$, we write

$$\mathcal{S}_{q,q'} \stackrel{\text{def}}{=} \{(x, x') \in M_q \times M_{q'} : (q, x, q', x', u(\cdot), \tau) \in \mathcal{S} \text{ for some } u(\cdot) \in \mathcal{U}_{q'} \text{ and } \tau \in J_{q'}\}.$$

The sets $\mathcal{S}_{q,q'}$ are called the *switching sets* of Σ from location q to location q'. Moreover, for $q, q' \in \mathcal{Q}$ and $x \in M_q$, $x' \in M_{q'}$, we write

$$\mathcal{U}_{q,x,q',x'} \stackrel{\text{def}}{=} \{u(\cdot) \in \mathcal{U}_{q'} : (q, x, q', x', u(\cdot), \tau) \in \mathcal{S} \text{ for some } \tau \in J_{q'}\}.$$

The set $\mathcal{U}_{q,x,q',x'}$ contains the controls we can use at location q' if there is a switching from the point x of M_q to the point x' of $M_{q'}$.

Definition 2: A hybrid state is a triplet (q, x, τ), where $q \in \mathcal{Q}$ is the location, $x \in M_q$ is the state of the control system and $\tau \in [0, \sup J_q)$ is the time since last switching. We denote by $\mathcal{HS}$ the set of all hybrid states.

The evolution of the hybrid system is as follows. Given a hybrid initial state $(q_1, x_0, 0)$, at time t_0, on some time interval $[t_0, t_1[$, with $t_1 - t_0 \in J_{q_1}$, the system evolves according to

$$\begin{cases} q(t) \equiv q_1 \\ \dot{x}(t) = f_{q_1}(x(t), u_1(t)), & x(t_0) = x_0 \\ \dot{\tau}(t) = 1, & \tau(t_0) = 0 \end{cases}$$

for some $u_1(\cdot) \in \mathcal{U}_{q_1}$ such that $Dom(u_1) \supset [t_0, t_1]$. This means that the system remains in location q_1 until $\tau = t_1 - t_0$ and it evolves on M_{q_1} according to the dynamic $f_{q_1}(x(t), u_1(t))$ for the control $u_1(\cdot) \in \mathcal{U}_{q_1}$. If the solution to the previous system can be prolonged on the whole interval $[t_0, t_1]$, then we can choose another hybrid state $(q_2, x_1, 0)$, a control $u_2(\cdot) \in \mathcal{U}_{q_2}$ and t_2 such that $(q_1, x(t_1), q_2, x_1, u_2(\cdot), t_2 - t_1) \in \mathcal{S}$ and let the system evolve in location q_2 following the corresponding controlled dynamics on the interval $[t_1, t_2]$:

$$\begin{cases} q(t) \equiv q_2 \\ \dot{x}(t) = f_{q_2}(x(t), u_2(t)), & x(t_1) = x_1 \\ \dot{\tau}(t) = 1, & \tau(t_1) = 0. \end{cases}$$

Then we can proceed in the same way with a location switching and so on. Notice that the time t_1 (t_2 and so on) can be chosen freely in J_{q_1} (respectively J_{q_2} and so on), hence it represents a control for the hybrid system.

We assume that if $u \in \mathcal{U}_q$ then every time translation of u is in $\mathcal{U}_q$, more precisely we assume

(A1) If $u \in \mathcal{U}_q$ for some $q \in \mathcal{Q}$, then for every $\sigma \in \mathbb{R}$ the control $\tilde{u}(t) = u(t + \sigma)$ satisfies $\tilde{u} \in \mathcal{U}_q$.

Hence we can always assume that $t_0 = 0$.

Let us now give a precise definition of trajectories, cost functionals and optimal control problems.

Definition 3: A trajectory is a map $\mathbf{X} : [0, T] \to \mathcal{HS}$, $\mathbf{X}(t) = (q(t), x(t), \tau(t))$, such that the following holds. There exist $0 = t_0 < t_1 < \ldots < t_\nu = T$ such that, if $i \in \{1, \ldots, \nu\}$, then $q(\cdot)$ is constant in $[t_{i-1}, t_i[$ and equal to $q_i \in \mathcal{Q}$, $\tau(t) = t - t_{i-1}$ on $[t_{i-1}, t_i[$, $t_i - t_{i-1} \in J_{q_i}$. Moreover, for every $i \in \{1, \ldots, \nu\}$, there exists $u_i \in \mathcal{U}_{q_i}$ such that:

- $x_i(\cdot) := x|_{]t_{i-1}, t_i[}(\cdot)$ is an absolutely continuous function in $]t_{i-1}, t_i[$, continuously prolongable to $[t_{i-1}, t_i]$;
- $\frac{d}{dt} x_i(t) = f_{q_i}(x_i(t), u_i(t))$ for a.e. $t \in]t_{i-1}, t_i[$;
- $(x_i(t_i), x_{i+1}(t_i)) \in \mathcal{S}_{q_i, q_{i+1}}$ if $i = 1, \ldots, \nu - 1$;
- $u_{i+1} \in \mathcal{U}_{q_i, x_i(t_i), q_{i+1}, x_{i+1}(t_i)}$ if $i = 1, \ldots, \nu - 1$.

Remark 1: In this setting, for a Cauchy type problem, it is not appropriate to choose first a sequence of controls and then determine the trajectory associated to it, because a priori the sequence could not be admissible, in the sense that there could exist no trajectory corresponding to it. This is due to the fact that in every location q, it is possible to use, as controls, only a subset of $\mathcal{U}_q$, depending on the switching strategy.

Definition 4: If Σ is a hybrid system, then a *Lagrangian* for Σ is a family $L = \{L_q\}_{q \in Q}$, $L_q : M_q \times U_q \to \mathbb{R}$ such that, for every trajectory $\mathbf{X}$, for every $i \in \{1, \ldots, \nu\}$ and for every control u_i associated to x_i, the function $t \mapsto L_{q_i}(x_i(t), u_i(t))$ is integrable in $]t_{i-1}, t_i[$.

Definition 5: If Σ is a hybrid system, then a switching cost function is a family $\Phi = \{\Phi_{q,q'}\}_{(q,q') \in \mathcal{Q} \times \mathcal{Q}}$ such that each $\Phi_{q,q'}$ is a real valued function defined on $\mathcal{S}_{q,q'}$.

Definition 6: If Σ is a hybrid system, then an endpoint cost function is a family $\varphi = \{\varphi_{q,q'}\}_{(q,q') \in \mathcal{Q} \times \mathcal{Q}}$ such that each $\varphi_{q,q'}$ is a real valued function defined on $M_q \times M_{q'}$.

If $L = \{L_q\}_{q \in \mathcal{Q}}$ is a Lagrangian, $\Phi = \{\Phi_{q,q'}\}_{(q,q') \in \mathcal{Q} \times \mathcal{Q}}$ is a switching cost function, $\varphi = \{\varphi_{q,q'}\}_{(q,q') \in \mathcal{Q} \times \mathcal{Q}}$ is an endpoint cost function for the hybrid control system Σ, then we can define the corresponding *cost functional* C, by letting

$$C(\mathbf{X}) = \sum_{j=1}^{\nu} \int_{t_{j-1}}^{t_j} L_{q_j}(x_j(t), u_j(t))\, dt$$
$$+ \sum_{j=1}^{\nu-1} \Phi_{q_j, q_{j+1}}(x_j(t_j), x_{j+1}(t_j)) + \varphi_{q_1, q_\nu}(x_1(t_0), x_\nu(t_\nu)),$$

where $\mathbf{X}$ is a trajectory for Σ.

Definition 7: Given a hybrid control system Σ, a cost functional C and two non empty subsets $\mathcal{N}_{in}, \mathcal{N}_{fin}$ of $\mathcal{HS}$, we call with $\mathcal{P}$ the problem of minimizing $C(\mathbf{X})$ over all trajectories $\mathbf{X}$ for Σ such that:

i) $(q_1, x_1(t_0), 0) \in \mathcal{N}_{in}$;
ii) $(q_\nu, x_\nu(t_\nu), t_\nu - t_{\nu-1}) \in \mathcal{N}_{fin}$.

Remark 2: Note that there could be no trajectory satisfying boundary data. However, we expect that in many applications the set $\mathcal{N}_{fin}$ should be chosen so to impose restriction only on the final location q and point x. So if $(q, x, t) \in \mathcal{N}_{fin}$ then $\mathcal{N}_{fin}$ should contain also all the points

(q, x, s) with $s \le \sup J_{q_\nu}$ (with possible equality only if $\sup J_{q_\nu} \in J_{q_\nu}$).

The *Maximum Principle* gives a necessary condition for a trajectory $\mathbf{X}$ to be a solution of $\mathcal{P}$. The set of variations involves trajectories having the same *history* (see [8]) of the candidate optimal one, that is having the same switching strategy.

Definition 8: If Σ is a hybrid system and L is a Lagrangian for Σ, then we say that (ψ, ψ_0) is an adjoint pair along a trajectory $\mathbf{X}$ if:

1) $\psi = (\psi_1, \dots, \psi_\nu)$ is such that, for every $i \in \{1, \dots, \nu\}$, $\psi_i : [t_{i-1}, t_i] \to T^* M_{q_i}$ is an absolutely continuous function, $\psi_i(t) \in T^*_{x_i(t)} M_{q_i}$ and
$$\dot{\psi}_i(t) = - < \psi_i(t), \frac{\partial}{\partial x} f_{q_i}(x_i(t), u_i(t)) > + \psi_0 \frac{\partial}{\partial x} L_{q_i}(x_i(t), u_i(t))$$
for a.e. $t \in [t_{i-1}, t_i]$;
2) $\psi_0 \in \mathbb{R}^+$.

In order to state the switching condition, we need a concept of a tangent cone. In this paper, as in [12], we use the notion of a Boltyanskii approximating cone.

Definition 9: Let S be a subset of a smooth manifold $\mathcal{X}$ and let $\bar{s} \in S$. A Boltyanskii approximating cone to S at $\bar{s}$ is a closed convex cone K in the tangent space $T_{\bar{s}}\mathcal{X}$ such that there exists a neighborhood W of 0 in $T_{\bar{s}}\mathcal{X}$ and a continuous map $\omega : W \cap K \to S$ with the property that $\omega(0) = \bar{s}$ and $\omega(w) = \bar{s} + w + o(\|w\|)$ as $w \to 0$ via values in $W \cap K$.

Definition 10: If Σ is a hybrid system, L is a Lagrangian and Φ is a switching cost function, then we say that an adjoint pair (ψ, ψ_0) along a trajectory $\mathbf{X}$ satisfies the switching condition if
$$(-\psi_i(t_i), \psi_{i+1}(t_i)) - \psi_0 \nabla \Phi_{q_i, q_{i+1}}(x_i(t_i), x_{i+1}(t_i)) \in K_i^\perp$$
for every $i \in \{1, \dots, \nu - 1\}$, where K_i is a Boltyanskii approximating cone to the set $\mathcal{S}_{q_i, q_{i+1}}$ at the point $(x_i(t_i), x_{i+1}(t_i))$ and $K_i^\perp$ is its polar cone.

Definition 11: If (ψ, ψ_0) is an adjoint pair along $\mathbf{X}$, and H_i is given by
$$\sup\{< \psi_i(t), f_{q_i}(x_i(t), u) > - \psi_0 L_{q_i}(x_i(t), u) : u \in U_{q_i}\},$$
then we say that (ψ, ψ_0) satisfies the Hamiltonian value condition if, for every $i \in \{1, \dots, \nu - 1\}$,

- if $t_i - t_{i-1} \in \mathrm{Int}(J_{q_i})$, then $H_i = H_\nu = 0$;
- if $t_i - t_{i-1}$ is the left endpoint of J_{q_i}, but J_{q_i} is nontrivial, then $H_i \le 0$;
- if $t_i - t_{i-1}$ is the right endpoint of J_{q_i}, but J_{q_i} is nontrivial, then $H_i \ge 0$.

As explained in the introduction for "simple" switching constraints a Hybrid Maximum Principle is valid. The condition ensuring this is precisely the following:

Assumption (H). For every $q, q' \in \mathcal{Q}$, $x \in M_q$, $x' \in M_{q'}$, we have $\mathcal{U}_{q,x,q',x'} = \mathcal{U}_{q'}$.

Assumption (H) says that in every location $q \in \mathcal{Q}$ we can use always all the controls in $\mathcal{U}_q$. Thus the admissible controls do not depend on the location switchings and the classical "needle variations" are still admissible variations.

Hybrid Maximum Principle. *Consider the problem $\mathcal{P}$ and assume (H). Let $\mathbf{X}$ be a solution for $\mathcal{P}$. Then, under suitable assumptions, there exists an adjoint pair (ψ, ψ_0) along $\mathbf{X}$ that satisfies the switching condition, the Hamiltonian maximization, nontriviality, transversality, and Hamiltonian value conditions for $\mathcal{P}$.*

There are some technical assumptions to require for the Hybrid Maximum Principle to hold true. These are specified in [9], [11], [12].

III. Simple necessary conditions

We present some introductory results about necessary conditions for optimality for hybrid systems that do not satisfy assumption (H). We postpone to the next section the statement and the proof of the *Hybrid Necessary Principle*. So this section is intended as a clarifying introduction to the subject of the next section.

Remark 3: Assumption (H) is not verified by many mechanical control systems. For example, to describe a car with gears, one can use a hybrid system, where each location corresponds to a gear of the car and the control is the acceleration. In this case it is clear that, when a switching from a low gear to a higher one happens, not all the controls can be used; see [4].

Suppose that every M_q is equal to $\mathbb{R}^{d_q}$ for some $d_q \in \mathbb{N}$, $d_q \ge 1$ and that every U_q is a compact subset of $\mathbb{R}^l$ for some $l \in \mathbb{N}$, $l \ge 1$. So $f_q : \mathbb{R}^{d_q} \times U_q \to \mathbb{R}^{d_q}$ and assume that
$$f_q \in C^2(\mathbb{R}^{d_q} \times U_q; \mathbb{R}^{d_q}). \tag{1}$$
Moreover, consider the case
$$\mathcal{U}_q = L^{p_q}_{loc}(\mathbb{R}; U_q) \tag{2}$$
for some $1 \le p_q \le +\infty$ and
$$L_q \in C^2(\mathbb{R}^{d_q} \times U_q; \mathbb{R}). \tag{3}$$
The symbol $L^{p_q}_{loc}(\mathbb{R}; U_q)$ denotes the set of functions from $\mathbb{R}$ to U_q belonging to $L^{p_q}(K; U_q)$ for every compact subset K of $\mathbb{R}$.

Remark 4: In order to avoid too many technicalities, we prefer to consider simplified hypotheses about the manifolds, the vector fields and the lagrangians. However it is possible to prove the results in a similar way using weaker assumptions.

Needle variations are the basic tool to prove the *Pontryagin Maximum Principle* in non-hybrid setting and the *Hybrid Maximum Principle* in hybrid setting. Needle variations consist in modifying the supposed optimal control in a small interval of times and to understand how the trajectory and the cost vary. In our case, since the choice of admissible controls depends by the switching strategy, needle variations do not produce admissible trajectories.

For simplicity, we consider only admissible needle variations of the following type: the control is the same of the candidate optimal trajectory until a certain time $\bar{\tau}$, then we produce a constant variation for a small interval of times and finally, in the following locations, we consider controls satisfying the switching conditions and some continuity and differentiability properties.

Definition 12: Fix a trajectory $\mathbf{X}$ and $i \in \{1, \dots, \nu\}$. We say that the family of trajectories $\mathbf{X}^\varepsilon = (q, x^\varepsilon, \tau)$, $\mathbf{X}^\varepsilon : [0, T] \to \mathcal{HS}$ $(\varepsilon > 0)$ is an **admissible needle variation** at location i if

1) $\mathbf{X}^0 \equiv \mathbf{X}$;
2) $\mathbf{X}^\varepsilon(t) = \mathbf{X}(t)$ for every $t \in [0, t_{i-1}]$;
3) the curves $\varepsilon \mapsto x_j^\varepsilon(t_{j-1})$ are differentiable at $\varepsilon = 0^+$ for every $j \in \{1, \dots, \nu\}$;
4) there exists a time $\bar\tau \in [t_{i-1} + \varepsilon, t_i]$ such that

$$u_i^\varepsilon(t) = \begin{cases} u_i(t), & t \in [t_{i-1}, \bar\tau - \varepsilon[, \\ \omega, & t \in [\bar\tau - \varepsilon, \bar\tau[, \\ u_i(t), & t \in [\bar\tau, t_i], \end{cases} \tag{4}$$

for some $\omega \in U_{q_i}$, where the symbol u_j^ε $(j \in \{1, \dots, \nu\})$ denotes the control at location j of x_j^ε;
5) for every $j \in \{i+1, \dots, \nu\}$, $u_j^\varepsilon \to u_j$ strongly in $L^1([t_{j-1}, t_j])$ as $\varepsilon \to 0^+$ and $\frac{u_j^\varepsilon - u_j}{\varepsilon} \rightharpoonup \theta_j$ weakly in $L^1([t_{j-1}, t_j])$ as $\varepsilon \to 0^+$ for some $\theta_j \in L^1([t_{j-1}, t_j])$.

Remark 5: Notice that, in Definition 12, we require that $\mathbf{X}^\varepsilon$, when $\varepsilon > 0$, is a family of trajectories. This means that, for a fixed $\varepsilon > 0$, $\mathbf{X}^\varepsilon$ is a trajectory and hence, by Definition 3,

$$(x_j^\varepsilon(t_j), x_{j+1}^\varepsilon(t_j)) \in \mathcal{S}_{q_j, q_{j+1}}$$

for every $j \in \{1, \dots, \nu - 1\}$ and

$$u_{j+1}^\varepsilon \in \mathcal{U}_{q_j, x_j(t_j), q_{j+1}, x_{j+1}(t_j)}$$

for every $j \in \{1, \dots, \nu - 1\}$.

Moreover we require the existence of a location q_i, $i \in \{1, \dots, \nu\}$, in which a variation originates. In particular we demand that, in the fixed location q_i, the variation is a classical needle variation and so the expression of the control u_i^ε is given in (4). In another location q_j, $j \in \{1, \dots, \nu\}$, $j \neq i$, we have the following possibilities.

1) If $j < i$, then $u_j^\varepsilon = u_j$ and $x_j^\varepsilon = x_j$ since the variation originates in location q_i.
2) If $j > i$, then we need some regularity properties of the control with respect to the parameter ε. These properties are described in 5 of Definition 12. Recall that for HMP we may choose $u_j^\varepsilon = u_j$ so that 5 is trivially satisfied.

For an admissible needle variation $\mathbf{X}^\varepsilon$ we define:

$$v_j(t) = \frac{d}{d\varepsilon} x_j^\varepsilon(t)\big|_{\varepsilon=0}. \tag{5}$$

The following lemmas holds. For a proof see [6].

Lemma 1: Let $\mathbf{X}^\varepsilon$ be an admissible needle variation. Then x^ε converges to x uniformly as ε goes to 0.

Lemma 2: Let $\mathbf{X}^\varepsilon$ be an admissible needle variation. Then $v_j \equiv 0$ if $j < i$, $v_i(t) = 0$ if $t_{i-1} \le t < \bar\tau$,

$$\begin{cases} \dot v_i(t) = D_x f_{q_i}(x_i(t), u_i(t)) v_i(t), \\ v_i(\bar\tau) = f_{q_i}(x_i(\bar\tau), \omega) - f_{q_i}(x_i(\bar\tau, u_i(\bar\tau))), \end{cases} \tag{6}$$

in the i-location if $\bar\tau \le t \le t_i$, while

$$\begin{cases} \dot v_j(t) = D_u f_{q_j}(x_j(t), u_j(t)) \theta_j(t) \\ \qquad + D_x f_{q_j}(x_j(t), u_j(t)) v_j(t) \\ v_j(t_{j-1}) = \frac{d}{d\varepsilon} x_j^\varepsilon(t_{j-1})\big|_{\varepsilon=0} \end{cases} \tag{7}$$

if $j > i$.

Remark 6: The evolution equation for v_j in general is an affine equation, since a term depending on θ_j appears. For hybrid systems with assumption (H), we may consider usual needle variations and so the resulting equation for v_j is linear.

Remark 7: It is useful to recall that equation (7) is valid only if $j > i$, i.e. only if the variation is originated in a previous location. Therefore, to prove (7) only properties 3 and 5 of Definition 12 are needed.

Let us evaluate the variation of the Lagrangian cost. Define

$$G_\varepsilon(t) := \sum_{h=1}^{j-1} \int_{t_{h-1}}^{t_h} L_{q_h}(x_h^\varepsilon(s), u_h^\varepsilon(s)) ds + \int_{t_{j-1}}^{t} L_{q_j}(x_j^\varepsilon(s), u_j^\varepsilon(s)) ds$$

when $t_{j-1} \le t < t_j$, and $w(t) := \frac{d}{d\varepsilon} G_\varepsilon(t)\big|_{\varepsilon=0^+}$.

Lemma 3: Let $\bar\tau \in]t_{i-1}, t_i[$ be the time at which an admissible needle variation originates. If $t \in]\bar\tau, t_i[$, then w satisfies the following differential equation:

$$\begin{cases} \dot w(t) = \frac{\partial}{\partial x} L_{q_i}(x_i(t), u_i(t)) v_i(t) \\ w(\bar\tau) = L_{q_i}(x_i(\bar\tau), \omega) - L_{q_i}(x_i(\bar\tau), u_i(\bar\tau)). \end{cases}$$

Moreover if $i < j \le \nu$, then we have:

$$\begin{cases} \dot w(t) = \frac{\partial}{\partial x} L_{q_j}(x_j(t), u_j(t)) v_j(t) \\ \qquad + \frac{\partial}{\partial u} L_{q_j}(x_j(t), u_j(t)) \theta_j(t), \ t_{j-1} < t < t_j \\ w(t_{j-1}) = \lim_{t \to t_{j-1}^-} w(t). \end{cases}$$

The following proposition follows from the previous lemmas.

Proposition 1: Let $\mathbf{X}$ be a trajectory and let $\mathbf{X}^\varepsilon$ be an admissible needle variation. Then, for every adjoint pair (ψ, ψ_0) along $\mathbf{X}$ and for every $j \in \{1, \dots, \nu\}$ the function

$$\psi_j(t) \cdot v_j(t) - \psi_0 w(t) + q_j(t) \tag{8}$$

is constant in $[t_{j-1}, t_j]$, where v_j is defined by (5) and q_j is any function defined by

$$\dot q_j(t) = -\psi_j(t) \frac{\partial}{\partial u} f_{q_j}(x_j(t), u_j(t)) \theta_j(t) + \psi_0 \frac{\partial}{\partial u} L_{q_j}(x_j(t), u_j(t)) \theta_j(t)$$

if $j > i$, while $q_j \equiv 0$ otherwise.

Proof: If $j < i$, then $v_j \equiv 0$, $q_j \equiv 0$ and $w(t) = 0$ for every $t \in [0, t_j]$.

If $j = i$, where q_i is the location at which the admissible needle variation originates, then

$$\frac{d}{dt} [\psi_i(t) \cdot v_i(t) - \psi_0 w(t)] = \dot\psi_i(t) \cdot v_i(t) + \psi_i(t) \cdot \dot v_i(t) - \psi_0 \dot w(t) = 0$$

and so we have the thesis when $j = i$.

Now if $j > i$, then

$$\frac{d}{dt}\left[\psi_j(t)\cdot v_j(t) - \psi_0 w(t) + q_j(t)\right]$$
$$= \dot{\psi}_j(t)\cdot v_j(t) + \psi_j(t)\cdot \dot{v}_j(t) - \psi_0 \dot{w}(t) + \dot{q}_j(t) = 0.$$

So, the proof is finished. ■

Let us deduce necessary conditions from the previous analysis. For simplicity, consider optimal control problems where the cost is formed only by the lagrangian part, that is the switching cost and the endpoint cost vanish. We suppose that $\mathbf{X}$ is an optimal trajectory and we consider an admissible needle variation $\mathbf{X}^\varepsilon$. Clearly, by optimality, $C(\mathbf{X}) \leq C(\mathbf{X}^\varepsilon)$. This implies that

$$w(T) \geq 0.$$

Let us consider an adjoint pair (ψ, ψ_0) along $\mathbf{X}$ with the properties that, for every $j \in \{1, \ldots, \nu\}$, $\psi_j(t_j)\cdot v_j(t_j) \leq 0$. Thus

$$\psi_\nu(t_\nu)\cdot v_\nu(t_\nu) - \psi_0 w(t_\nu) \leq 0.$$

For every $q_\nu(\cdot)$ defined as in Proposition 1 with $q_\nu(t_\nu) \leq 0$, it holds:

$$\psi_\nu(t)\cdot v_\nu(t) - \psi_0 w(t) + q_\nu(t) \leq 0$$

for every $t \in [t_{\nu-1}, t_\nu]$. Therefore in the $\nu - 1$ location we have

$$\psi_{\nu-1}(t_{\nu-1})\cdot v_{\nu-1}(t_{\nu-1}) - \psi_0 w(t_{\nu-1}) + q_\nu(t_{\nu-1}) \leq 0$$

and so

$$\psi_{\nu-1}(t)\cdot v_{\nu-1}(t) - \psi_0 w(t) + q_\nu(t_{\nu-1}) + q_{\nu-1}(t) \leq 0$$

for every $q_{\nu-1}$ with $q_{\nu-1}(t_{\nu-1}) \leq 0$ and for every $t \in [t_{\nu-2}, t_{\nu-1}]$. Iterating this argument we conclude that

$$\psi_j(t)\cdot v_j(t) - \psi_0 w(t) + \sum_{l=j+1}^{\nu} q_l(t_{l-1}) + q_j(t) \leq 0 \quad (9)$$

for every $j \in \{1, \ldots, \nu\}$, $t \in [t_{j-1}, t_j]$ and for every function q_l with $q_l(t_l) \leq 0$.

Equation (9) gives a necessary condition for optimality when the hybrid system does not satisfy assumption (H).

IV. Hybrid Necessary Principle

This section deals with a general *Hybrid Necessary Principle* when assumption (H) does not hold. Again for sake of simplicity we assume (1).

Let $\mathbf{X}$ be an optimal trajectory for the problem $\mathcal{P}$ and let $\bar{\varepsilon} > 0$. We denote with K a cone in $\mathbb{R}^{d_1} \times \ldots \times \mathbb{R}^{d_\nu}$, with $v = (v_1, \ldots, v_\nu)$ an element of K and with $(u_1, \ldots, u_\nu)$ the controls of the candidate optimal hybrid trajectory $\mathbf{X}$. The aim of next definition is to give a rigorous description of all variations we are able to consider. In analogy with [12], we treat variations depending by two parameters: ε and v. ε is a real positive number, while v belongs to a cone in a finite dimensional manifold. The reader can think v as the parameter responsible for the variation of the initial points of each trajectory x_j, $j \in \{1, \ldots, \nu\}$, and ε as parameterizing the control variation.

Definition 13: **(Map of variations)**. A map V defined on $[0, \bar{\varepsilon}] \times K$, $V(\varepsilon, v) = (x_1^{(\varepsilon,v)}, u_1^{(\varepsilon,v)}, \ldots, x_\nu^{(\varepsilon,v)}, u_\nu^{(\varepsilon,v)})$, is called a map of variations if, for every $(\varepsilon, v) \in [0, \bar{\varepsilon}] \times K$, the following hold:

1) for every $i \in \{1, \ldots, \nu\}$, $u_i^{(\varepsilon,v)} \in \mathcal{U}_{q_i}$ and $u_i^{(\delta\varepsilon,\delta v)} \to u_i$ in $L^1(t_{i-1}, t_i)$ as $\delta \to 0^+$;
2) for every $i \in \{1, \ldots, \nu\}$, $x_i^{(\varepsilon,v)} :]t_{i-1}, t_i[\to \mathbb{R}^{d_i}$ is an absolutely continuous function continuously prolongable to $[t_{i-1}, t_i]$ such that $\frac{d}{d\delta} x_i^{(\delta\varepsilon,\delta v)}(t_{i-1})\big|_{\delta=0} = v_i$ and
$$\frac{d}{dt} x_i^{(\varepsilon,v)}(t) = f_{q_i}(x_i^{(\varepsilon,v)}(t), u_i^{(\varepsilon,v)}(t)) \text{ for a.e. } t;$$
3) for every $i \in \{1, \ldots, \nu - 1\}$, the control $u_{i+1}^{(\varepsilon,v)}$ belongs to $\mathcal{U}_{q_i, x_i^{(\varepsilon,v)}(t_i), q_{i+1}, x_{i+1}^{(\varepsilon,v)}(t_i)}$;
4) the maps
$$(\varepsilon, v) \mapsto x_i^{(\varepsilon,v)}(t_i)$$
and
$$(\varepsilon, v) \mapsto \gamma(\varepsilon, v) := \sum_{i=1}^{\nu} \int_{t_{i-1}}^{t_i} L_{q_i}(x_i^{(\varepsilon,v)}(t), u_i^{(\varepsilon,v)}(t))dt$$
are differentiable at $(0, 0)$ in the direction $\mathbb{R}^+ \times K$;
5) for every $i \in \{1, \ldots, \nu\}$, there exist a vector-valued Radon measure $\alpha_{i,f,V}^{(\varepsilon,v)}$ and a scalar Radon measure $\alpha_{i,L,V}^{(\varepsilon,v)}$ such that
$$\frac{f_{q_i}(x_i, u_i^{(\delta\varepsilon,\delta v)}) - f_{q_i}(x_i, u_i)}{\delta} \rightharpoonup^* \alpha_{i,f,V}^{(\varepsilon,v)}$$
and
$$\frac{L_{q_i}(x_i, u_i^{(\delta\varepsilon,\delta v)}) - L_{q_i}(x_i, u_i)}{\delta} \rightharpoonup^* \alpha_{i,L,V}^{(\varepsilon,v)}$$
as $\delta \downarrow 0$, where the convergence is intended in the weak* topology of the space of Radon measure, seen as the dual of continuous functions.

We denote by $\mathcal{V}$ the set of all maps of variations.

Remark 8: The main reasoning in the proof of the HNP follows the classical approach. We consider *feasible* cones generated by final points of admissible variations and *profitable* cones formed by points that realize a cost lower than that of the candidate trajectory. To have optimality these two cones must be separated, i.e. there exists a hyperplane that separates the cones. From these considerations we deduce necessary conditions.

In Definition 12 (map of variations), we require various assumptions. In particular, the assumptions 1, 2 and 3 guarantee that maps of variations produce admissible trajectories for our hybrid system. Moreover assumption 4 implies the existence of the cone generated by the variations and, finally, assumption 5 is necessary in order to have differentiability properties of trajectories. Notice that we consider only weak differentiability properties of trajectories. We remand to the paper by Piccoli and Sussmann [10] for the general theory about differentiability of trajectories with respect a family of parameters.

For $V \in \mathcal{V}$, consider the map $\widetilde{C}_V$ defined on $[0,\bar{\varepsilon}] \times K$ into $(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}) \times \cdots (\mathbb{R}^{d_\nu} \times \mathbb{R}^{d_1}) \times \mathbb{R}$ given by

$$\begin{aligned}(\varepsilon, v) \quad \mapsto \quad & ((x_1^{(\varepsilon,v)}(t_1), x_2^{(\varepsilon,v)}(t_1)), \ldots, \\ & (x_\nu^{(\varepsilon,v)}(t_\nu), x_1^{(\varepsilon,v)}(t_0)), \gamma(\varepsilon, v)).\end{aligned}$$

By Definition 13, $\widetilde{C}_V$ is differentiable at $(0,0)$. Therefore we may define the cone

$$K_V := D\widetilde{C}_V(0,0)([0,\bar{\varepsilon} \times K]). \qquad (10)$$

Definition 14: If $V \in \mathcal{V}$, (ψ, ψ_0) is an adjoint pair along $\mathbf{X}$ and $(\psi_1^-, \ldots, \psi_\nu^-) \in \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_\nu}$, then we say that the *covector inequality* holds if

$$\begin{aligned}-\psi_0 \sum_{i=1}^{\nu} \int_{t_{i-1}}^{t_i} \frac{\partial}{\partial x} L_{q_i}(x_i(t), u_i(t)) \int_{t_{i-1}}^{s} M_i(s,r) d\alpha_{i,f,V}^{(\varepsilon,v)}(r) ds \\ + \sum_{i=1}^{\nu} \left(\psi_i^- \cdot v_i + \psi_i(t_i) \int_{t_{i-1}}^{t_i} M_i(t_i, s) d\alpha_{i,f,V}^{(\varepsilon,v)}(s) \right) \\ + \sum_{i=1}^{\nu} \psi_i(t_{i-1}) v_i - \psi_0 \sum_{i=1}^{\nu} \int_{t_{i-1}}^{t_i} d\alpha_{i,L,V}^{(\varepsilon,v)}(s) \le 0, \qquad (11)\end{aligned}$$

where $M_i(t,s)$ is the fundamental matrix solution to the linear system $\dot{y}(t) = \frac{\partial}{\partial x} f_{q_i}(x_i(t), u_i(t)) y(t)$.

Definition 15: If $V \in \mathcal{V}$, (ψ, ψ_0) is an adjoint pair along $\mathbf{X}$, $(\psi_1^-, \ldots, \psi_\nu^-) \in \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_\nu}$ and K_i is a Boltyanskii approximating cone to $\mathcal{S}_{q_i, q_{i+1}}$ at $(x_i(t_i), x_{i+1}(t_i))$ for every $i \in \{1, \ldots, \nu - 1\}$, then we say that the *general switching condition* holds if

$$((-\psi_i(t_i), -\psi_{i+1}^-) - \psi_0 \nabla \Phi_{q_i, q_{i+1}}(x_i(t_i), x_{i+1}(t_i))) \in K_i^\perp, \qquad (12)$$

for every $i \in \{1, \ldots, \nu - 1\}$, where $K_i^\perp$ denotes the polar of the cone K_i.

Definition 16: If $V \in \mathcal{V}$, (ψ, ψ_0) is an adjoint pair along $\mathbf{X}$, $(\psi_1^-, \ldots, \psi_\nu^-) \in \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_\nu}$ and K_ν is a Boltyanskii approximating cone to $\mathcal{S}_{q_\nu, q_1}$ at $(x_\nu(t_\nu), x_1(t_0))$, then we say that the *general endpoint condition* holds if

$$((-\psi_\nu(t_\nu), -\psi_1^-) - \psi_0 \nabla \varphi_{q_\nu, q_1}(x_\nu(t_\nu), x_1(t_0))) \in K_\nu^\perp, \qquad (13)$$

where $K_\nu^\perp$ is the polar of the cone K_ν.

The following theorem holds.

Theorem 1: **(Hybrid Necessary Principle)**. Let $\mathbf{X}$ be an optimal trajectory for problem $\mathcal{P}$. For every convex cone $\widehat{K}$ contained in $\cup_{V \in \mathcal{V}} K_V$, where K_V is the cone of feasible directions given by the map of variation V, there exist an adjoint pair (ψ, ψ_0) along $\mathbf{X}$ and $(\psi_1^-, \ldots, \psi_\nu^-) \in \mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_\nu}$ such that the covector inequality, the general switching condition and the general endpoint condition hold for every $V \in \mathcal{V}$, $(\varepsilon, v) \in [0, \bar{\varepsilon}] \times K$ such that $D\widetilde{C}_V(0)(\varepsilon, v) \in \widehat{K}$.

Proof: A complete proof of the theorem is contained in [6]. In this paper we give just an idea of it.

Consider an optimal trajectory $\mathbf{X}$. If $V \in \mathcal{V}$, $\varepsilon \in [0, \bar{\varepsilon}]$ and $v \in K$, then we may consider $\mathbf{X}^{(\varepsilon,v)}(\cdot)$ a candidate hybrid trajectory obtained piecing together $x_i^{(\varepsilon,v)}$, $i = 1, \ldots, \nu$. Then $\mathbf{X}^{(\varepsilon,v)}(\cdot)$ is a trajectory if and only if

- $(x_i^{(\varepsilon,v)}(t_i), x_{i+1}^{(\varepsilon,v)}(t_i)) \in \mathcal{S}_{q_i, q_{i+1}}$ for $i = 1, \ldots, \nu - 1$;
- $(q_1, x_1^{(\varepsilon,v)}(t_0), 0) \in \mathcal{N}_{in}$;
- $(q_\nu, x_\nu^{(\varepsilon,v)}(t_\nu), t_\nu - t_{\nu-1}) \in \mathcal{N}_{fin}$.

Since $\mathbf{X}$ is optimal we have that $C(\mathbf{X}^{(\varepsilon,v)}) \ge C(\mathbf{X})$ whenever the previous conditions hold.

For every $V \in \mathcal{V}$, the cone K_V, defined in (10), is the cone of feasible directions. Let P be the set of points $((z_1, z_1'), \ldots, (z_\nu, z_\nu'), r)$ of $(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}) \times \cdots \times (\mathbb{R}^{d_\nu} \times \mathbb{R}^{d_1}) \times \mathbb{R}$ such that $(z_i, z_i') \in \mathcal{S}_{q_i, q_{i+1}}$ for every $i \in \{1, \ldots, \nu\}$ and

$$r \le C(\mathbf{X}) - \sum_{i=1}^{\nu-1} \Phi_{q_i, q_{i+1}}(z_i, z_i') - \varphi_{q_1, q_\nu}(z_\nu', z_\nu) - \sum_{i=1}^{\nu} \sigma_i(z_i, z_i')$$

where σ_i are smooth functions, $\sigma_i(x_i(t_i), x_{i+i}(t_i)) = 0$ for every $i \in \{1, \ldots, \nu\}$ and $\sigma_i(z_i, z_i')$ is strictly positive if $(z_i, z_i') \neq (x_i(t_i), x_{i+i}(t_i))$. If K_P is a Boltyanskii approximating cone to P at the point $((x_1(t_1), x_2(t_1)), \ldots, (x_\nu(t_\nu), x_1(t_0)), C_L(\mathbf{X}))$, where C_L is the Lagrangian cost, then the optimality of $\mathbf{X}$ implies that the cones K_V and K_P are weakly separated; see [13]. From this we deduce the covector inequality, the general switching condition and the general endpoint condition. ∎

References

[1] P. J. Antsaklis and A. Nerode eds., *Special issue on hybrid systems*, IEEE Trans. Automat. Control, 43 (1998), no. 4.

[2] M. S. Branicky, V. S. Borkar and S. K. Mitter, *A unified framework for hybrid control: model and optimal control theory*, IEEE Trans. Automatic Control, 43 (1998), pp. 31–45.

[3] C. Cassandras, D. L. Pepyne and Y. Wardi, *Optimal control of a class of hybrid systems*, IEEE Trans. Automat. Control, 46 (2001), pp. 398–415.

[4] C. D'Apice, R. Manzo, M. Garavello and B. Piccoli, *Hybrid optimal control: case study of a car with gears*, Internat. J. Control, 76 (2003), pp. 1272–1284.

[5] M. D. Di Benedetto and A. Sangiovanni-Vincentelli eds., *Hybrid System Computation and Control*, Lecture Notes in Computer Science 2034, Springer-Verlag, Berlin Heidelberg, 2001.

[6] M. Garavello and B. Piccoli, *Hybrid necessary principles*, SIAM J. Control Optim., to appear.

[7] S. Hedlund and A. Rantzer, *Optimal control of hybrid systems*, Proceedings of the 38th IEEE Conference on Decision and Control, Phoenix, 1999, pp. 3972–3977.

[8] B. Piccoli, *Hybrid systems and Optimal control*, Proceedings of the 37th IEEE Conference on Decision and Control, Tampa, 1998, pp. 13–18.

[9] B. Piccoli, *Necessary conditions for hybrid optimization*, Proceedings of the 38th IEEE Conference on Decision and Control, Phoenix, 1999, pp. 410–415.

[10] B. Piccoli and H. J. Sussmann, *Regular synthesis and sufficient conditions for optimality*, SIAM J. Control Optim., 39 (2000), pp. 359-410.

[11] H. J. Sussmann, *A Maximum Principle for hybrid optimal control problems*, Proceedings of the 38th IEEE Conference on Decision and Control, Phoenix, 1999.

[12] H. J. Sussmann, *A nonsmooth hybrid maximum principle*. Stability and stabilization of nonlinear systems (Ghent, 1999), 325–354, Lecture Notes in Control and Inform. Sci., 246, Springer, London, 1999.

[13] H. J. Sussmann, *Multidifferential calculus: chain rule, open mapping and transversal intersection theorems*. In *Optimal Control: Theory, Algorithms and Applications*, W. W. Hager and P. M. Pardalos Editors, Kluwer Academic Publishers, 1998, pp. 436-487.

[14] X. Xu and P. J. Antsaklis, *A dynamic programming approach for optimal control of switched systems*, Proceedings of the 39th IEEE Conference on Decision and Control, Sidney 2000, pp. 1822–1827.