Pure and Applied Functional Analysis
Volume 1, Number 2, 2016, 217–230

Yokohama Publishers
ISSN 2189-3764 ONLINE JOURNAL

ON THE SUFFICIENCY OF PONTRYAGIN'S MAXIMUM PRINCIPLE

M. MARGARIDA A. FERREIRA AND GEORGI V. SMIRNOV

Abstract. This work focus on sufficient conditions of optimality for an optimal control problem. A *refined maximum principle condition* which guarantees weak local optimality of control processes for affine control systems with a polyhedral set of controls is introduced. This refined maximum principle condition expresses that the control is uniquely defined for almost all instants of time and the behavior of adjoint variables is rather regular. Several examples are presented to illustrate such sufficient condition.

1. Introduction

In this work we focus on sufficient conditions of optimality for an optimal control problem. Traditionally sufficient conditions involve second order derivatives (see, for example, [3] and the references therein). However, for certain classes of problems the Pontryagin maximum principle [4] is by itself a sufficient condition. This is well known for linear control problems with convex cost and convex constraints. This sufficiency of the Pontryagin maximum principle is also observed in certain classes of non convex problems of hydro-electric power stations management [1].

In the optimal control context, a problem satisfying conditions of existence theorems can be solved determining all trajectories satisfying Pontryagin's maximum principle and choosing the optimal one. In practice this can be very hard and it would be important to understand if a given trajectory is optimal or not. Here, we introduce a *refined maximum principle condition* that for affine control systems with a polyhedral set of controls guarantees weak local optimality of control processes. This refined maximum principle condition means that the control is uniquely defined for almost all instants of time and the behavior of adjoint variables is rather regular. We illustrate this sufficient condition with different examples.

2. Preliminary considerations

We start by introducing some notation that shall be used throughout the paper. The Euclidean norm of a point x and the inner product between x, y are denoted respectively by $|x|$ and $\langle x, y\rangle$. The norm $|\cdot|_p$ means the L_p-norm, with $1 \leq p \leq \infty$. The set of non negative real numbers is represented by R_+ and $C([a,b];D)$ denotes the set of continuous functions $f : [a,b] \to D$. Given a matrix A, the transpose of

2010 *Mathematics Subject Classification.* 49K15.
Key words and phrases. Pontryagin's maximum principle, sufficient condition.

A is represented by A^* and the identity matrix is represented by I. The Lebesgue measure of a given set C is represented by meas(C).

A set $C \subset R^m$ is a *polyhedral set* if it defined as the set of solutions of a linear inequalities system, i.e., $C = \{x \in R^m : \ \langle x, c_i \rangle \leq \alpha_i, i = \overline{1,k}\}$, where c_i is a fixed vector in R^m and α_i is a fixed real constant, for every $i = \overline{1,k}$. Bounded polyhedral sets are referred to as *polyhedrons*.

Now we proceed with an analysis of Pontryagin's maximum principle as a sufficient condition and recall that such maximum principle guarantees directional optimality of control processes. Consider the following optimal control problem in Mayer form

$$\phi(x(T)) \to \inf, \tag{2.1}$$

$$\dot{x} = f(t,x) + g(t,x)u, \ \ u \in U, \tag{2.2}$$

$$x(0) = x_0, \tag{2.3}$$

where T is fixed and x_0 is a given point in R^n. Here, and throughout the paper, $\phi : R^n \to R$, $f : R \times R^n \to R^n$ and $g : R \times R^n \to R^{n \times m}$ are twice continuously differentiable functions and $U \subset R^m$ is a polyhedral set.

As usual in the optimal control framework, we refer to a measurable function $u : [0,T] \to R^m$ that satisfies $u(t) \in U$, *a.e.*, as a *control function* $u(\cdot)$. A *control process* $(u(\cdot), x(\cdot))$ (some times refereed simply by *process*) comprises a control function $u(\cdot)$ and a state *trajectory* $x(\cdot)$ that is a solution to the differential equation $\dot{x} = f(t,x) + g(t,x)u(t)$.

Denote by $x(\cdot, \bar{u}(\cdot))$ the solution to the Cauchy problem

$$\dot{x} = f(t,x) + g(t,x)(\hat{u} + \bar{u}), \ \ x(0) = x_0. \tag{2.4}$$

Set $\hat{x}(\cdot) = x(\cdot, 0)$. Then we have, for some *difference* function $r : [0,T] \times R^m \to R^n$,

$$x(t, \bar{u}(\cdot)) = \hat{x}(t) + \bar{x}(t) + r(t, \bar{u}(\cdot)), \tag{2.5}$$

where $\bar{x}(\cdot)$ is the solution to the Cauchy problem

$$\dot{\bar{x}} = (\nabla_x f(t, \hat{x}) + \nabla_x (g(t, \hat{x})\hat{u}))\bar{x} + g(t, \hat{x})\bar{u}, \ \ \bar{x}(0) = 0. \tag{2.6}$$

From the Filippov theorem [2] we obtain

$$|r(t, \bar{u}(\cdot))| \leq (\text{const}) \int_0^T \rho(t, \bar{u}(\cdot))dt, \tag{2.7}$$

where $\rho(t, \bar{u}(\cdot))$ represents the following distance

$$\begin{aligned} \rho(t, \bar{u}(\cdot)) &= |\dot{\hat{x}}(t) + \dot{\bar{x}}(t) - f(t, \hat{x}(t) + \bar{x}(t)) - g(t, \hat{x}(t) + \bar{x}(t))(\hat{u}(t) + \bar{u}(t)))| \\ &\leq \frac{1}{2} \max |\nabla^2 f(t,x)||\bar{x}(t)|^2 + \frac{1}{2} \max |\nabla^2 g(t,x)||\bar{x}(t)|^2 (|\hat{u}(t)| + |\bar{u}(t)|) \\ &\qquad + \max |\nabla g(t,x)||\bar{x}(t)||\bar{u}(t)|. \end{aligned}$$

Since

$$|\bar{x}(t)| \leq (\text{const}) \int_0^T |\bar{u}(t)|dt, \tag{2.8}$$

we have

$$|r(t,\bar{u}(\cdot))| \le (\text{const})\left(\int_0^T |\bar{u}(t)|dt\right)^2. \tag{2.9}$$

Let $\hat{u}(t) \in U$ and $\hat{u}(t) + \alpha\bar{u}(t) \in U$, $t \in [0,T]$, $\alpha \in [0,\alpha_0]$. Then, since

$$\lim_{\alpha\to 0} \alpha^{-1} r(t,\alpha\bar{u}(\cdot)) = 0,$$

we obtain

$$\phi(x(T,\alpha\bar{u}(\cdot))) = \phi(\hat{x}(T)) + \langle \nabla\phi(\hat{x}(T)), \alpha\bar{x}(T) + r(T,\alpha\bar{u}(\cdot))\rangle$$

$$+\frac{1}{2}\langle \alpha\bar{x}(T) + r(T,\alpha\bar{u}(\cdot)), \nabla^2\phi(x_\alpha)(\alpha\bar{x}(T) + r(T,\alpha\bar{u}(\cdot)))\rangle, \tag{2.10}$$

where $x_\alpha = (1-\theta)\hat{x}(T) + \theta x(T,\alpha\bar{u}(\cdot))$ for some $\theta \in [0,1]$.

Let $\Phi(t,s)$ be the fundamental matrix of the system

$$\dot{\bar{x}} = (\nabla_x f(t,\hat{x}) + \nabla_x(g(t,\hat{x})\hat{u}))\bar{x}. \tag{2.11}$$

Set $p(t) = -\Phi^*(T,t)\nabla\phi(\hat{x}(T))$. Then we have

$$\langle \nabla\phi(\hat{x}(T)), \bar{x}(T)\rangle = -\int_0^T \langle p(t), g(t,\hat{x}(t))\bar{u}(t)\rangle dt. \tag{2.12}$$

Assume that

$$\int_0^T \langle p(t), g(t,\hat{x}(t))\bar{u}(t)\rangle dt < 0 \tag{2.13}$$

whenever $\bar{u}(t) \in (U - \hat{u}(t))$ and $\bar{u}(\cdot) \neq 0$. Then we get

$$\phi(x(T,\alpha\bar{u}(\cdot))) > \phi(\hat{x}(T))$$

for all $\alpha > 0$ sufficiently small, i.e. the condition

$$\langle p(t), g(t,\hat{x}(t))u\rangle < \langle p(t), g(t,\hat{x}(t))\hat{u}(t)\rangle, \quad u \in U,\ u \neq \hat{u}(t), \tag{2.14}$$

implies that $\hat{u}(\cdot)$ is a directional minimizer. Condition (2.14) can be interpreted as the maximum principle uniquely defining the control.

Two main questions arise. What shall we impose more to guarantee local optimality of $\hat{u}$, not merely directional optimilaty, and what is the class of control systems for which such condition applies. We will work in the class of affine control systems and polyhedral set of controls. The following two examples show that this class is a natural one and one hardly expect to enlarge it.

Examples. Consider the optimal control problem for a non affine system:

$$\begin{aligned} &x_2(1) - 2x_1^2(1) \to \inf,\\ &\dot{x}_1 = u,\\ &\dot{x}_2 = u^2,\\ &u \in [-1,1]\\ &x_1(0) = x_2(0) = 0. \end{aligned}$$

The control $\hat{u} \equiv 0$ satisfies the maximum principle. However it is not a local minimizer. Indeed, consider the control $u \equiv \epsilon$. For the corresponding trajectory, we have

$$x_2(1) - 2x_1^2(1) = -\epsilon^2 < 0.$$

Hence for the systems of general form $\dot{x} = f(t, x, u)$, the maximum principle cannot be a sufficient condition.

The following example shows that even for affine control systems the maximum principle is not a sufficient optimality condition. Consider the optimal control problem

$$\begin{aligned} &x_1(1) \to \min, \\ &\dot{x}_1 = x_3 - x_2^2, \\ &\dot{x}_2 = u_2, \\ &\dot{x}_3 = u_3, \\ &u_2^2 + (u_3 - 1)^2 \leq 1, \\ &x_i(0) = 0, i = 1, 2, 3. \end{aligned}$$

The control $(\hat{u}_2(t), \hat{u}_3(t)) \equiv (0, 0)$ is a directional minimizer. Indeed, the solution to the adjoin system is $p_1 = -1$, $p_2 = 0$, $p_3 = t - 1$, and the maximum condition takes the form $(t - 1)u_3 < 0$, whenever $u_3 > 0$. However, the control $(\hat{u}_2(t), \hat{u}_3(t))$ is not a minimizer. Indeed, consider the control $u_2 = \sqrt{1 - (\epsilon - 1)^2}$, $u_3 = \epsilon$. Obviously we have $x_2 = t\sqrt{1 - (\epsilon - 1)^2}$, $x_3 = t\epsilon$, $x_1(1) = -\epsilon/6 + \epsilon^2/3 < 0$, $\epsilon \ll 1$.

However, as we shall see in the next section, for affine systems with polyhedral sets of controls it is possible in many situations to introduce a refined version of the maximum principle as a sufficient optimality condition.

3. Refined maximum principle

We say that the control $\hat{u}(\cdot)$ satisfies a *refined maximum principle*, if there exist a non-negative measurable function $\sigma : [0, T] \to R_+$ and constants $\gamma > 0$ and $a_0 > 0$ such that

(1) $\max_{u \in U}(\langle g(t, \hat{x}(t))(u - \hat{u}(t)), p(t)\rangle + \sigma(t)|u - \hat{u}(t)|) \leq 0$;
(2) $\text{meas}\{t \in [0, T] \mid \sigma(t) < a\} \leq \gamma a$, whenever $a \in [0, a_0]$,
(3) $\text{meas}\{t \in [0, T] \mid \sigma(t) = a\} = 0$, whenever $a > 0$.

Observe that the first inequality implies the maximum principle condition

$$\max_{u \in U}(\langle g(t, \hat{x}(t))(u - \hat{u}(t)), p(t)\rangle \leq 0.$$

Let us see some examples. Assume that $g = I$.

(1) Let $U = \{(u_1, u_2) \in R^2 \mid u_1^2 + u_2^2 \leq 1\}$, $\hat{u}(t) = (\cos t, \sin t)$, and $p(t) = (\cos t, \sin t)$, $t \in [0, 2\pi]$. In this case $\sigma(t) \equiv 0$ and the refined maximum condition is not satisfied.
(2) Let $U = \{(u, 0) \in R^2 \mid u \in [-1, 1]\}$, $\hat{u}(t) = \text{sign} \cos t$, and $p(t) = (\cos t, \sin t)$, $t \in [0, 2\pi]$. In this case $\sigma(t) = |\cos t|$ and the refined maximum condition is satisfied with $\gamma = 2\pi$ and $a_0 = 1$, for example.

The following result establishes conditions useful to verify the refined maximum principle property. Let $U = \mathrm{co}\{u_1, \dots, u_M\}$, $0 = t_0 < t_1 \dots < t_L = T$, $\hat{u}(t) = u_{m_l}$, $t \in]t_l, t_{l+1}[$, $l = \overline{0, L-1}$, $q(t) = (g(t, \hat{x}(t)))^* p(t)$, $\mathcal{M}_l = \{m \mid \langle q(t_l), u_m \rangle = \max_{u \in U} \langle q(t_l), u \rangle\}$, $l = \overline{0, L}$.

Lemma 3.1. *Assume that $q(\cdot)$ is a continuous and piece-wise continuously differentiable function, that the maximum principle uniquely defines the control $\hat{u}(\cdot)$ (in the sense that $\langle q(t), u - u_{m_l} \rangle < 0$, $\forall t \in]t_l, t_{l+1}[, \forall u \in U$, $u \neq u_{m_l}$), and*

$$\max_{\substack{m \in \mathcal{M}_l \\ m \neq m_l}} \langle \dot{q}(t_l + 0), u_m - u_{m_l} \rangle < -2\sigma_0, \; l = \overline{0, L-1}, \tag{3.1}$$

$$\min_{\substack{m \in \mathcal{M}_l \\ m \neq m_{l-1}}} \langle \dot{q}(t_l - 0), u_m - u_{m_{l-1}} \rangle > 2\sigma_0, \; l = \overline{1, L}. \tag{3.2}$$

Then the refined maximum principle condition is satisfied.

Proof. Let $\Delta t > 0$, $u = \sum_{m=1}^{M} \lambda_m u_m$, $u \neq u_{m_l}$, $\lambda_m \geq 0$, $\sum_{m=1}^{M} \lambda_m = 1$. Using (3.1) we have

$$\langle q(t_l + \Delta t), u - u_{m_l} \rangle = \langle q(t_l + \Delta t), \sum_{m \neq m_l} \lambda_m (u_m - u_{m_l}) \rangle \tag{3.3}$$

$$= \sum_{\substack{m \in \mathcal{M}_l \\ m \neq m_l}} \lambda_m (\langle q(t_l), u_m - u_{m_l} \rangle + \Delta t \langle \dot{q}(t_l + 0), u_m - u_{m_l} \rangle + o(\Delta t)) \tag{3.4}$$

$$+ \sum_{m \notin \mathcal{M}_l} \lambda_m \langle q(t_l + \Delta t), u_m - u_{m_l} \rangle \tag{3.5}$$

$$\leq - \sum_{\substack{m \in \mathcal{M}_l \\ m \neq m_l}} \lambda_m \sigma_0 \Delta t + \sum_{m \notin \mathcal{M}_l} \lambda_m \langle q(t_l + \Delta t), u_m - u_{m_l} \rangle, \tag{3.6}$$

whenever Δt is small enough.

If $m \notin \mathcal{M}_l$, then we have $\langle q(t_l), u_m - u_{m_l} \rangle < 0$. Since $a(t) = \langle q(t), u_m - u_{m_l} \rangle$ is continuous and $a(t_l) < 0$, it comes $a(t_l + \Delta t) = \langle q(t_l + \Delta t), u_m - u_{m_l} \rangle < 0$, whenever Δt is small enough. As a consequence we can write $\langle q(t_l + \Delta t), u_m - u_{m_l} \rangle < -\sigma_2$, for some $\sigma_1 > 0$. So, we obtain

$$\langle q(t_l + \Delta t), u - u_{m_l} \rangle \leq - \sum_{\substack{m \in \mathcal{M}_l \\ m \neq m_l}} \lambda_m \sigma_0 \Delta t - \sum_{m \notin \mathcal{M}_l} \lambda_m \sigma_1$$

$$\leq - \sum_{\substack{m \in \mathcal{M}_l \\ m \neq m_l}} \lambda_m \sigma_0 \Delta t - \sum_{m \notin \mathcal{M}_l} \lambda_m \sigma_1 \Delta t \; \leq \; - \sum_{m \neq m_l} \lambda_m \sigma_2 \Delta t$$

for some $\sigma_2 > 0$, whenever Δt is small enough. Since

$$|u - u_{m_l}| = \left| \sum_{m \neq m_l} \lambda_m (u_m - u_{m_l}) \right| \leq \sum_{m \neq m_l} \lambda_m \max_{m \neq m_l} |u_m - u_{m_l}|,$$

we get

$$\begin{aligned}\langle q(t_l+\Delta t), u-u_{m_l}\rangle &\leq -\sigma_2 \frac{\left|\sum\limits_{m\neq m_l}\lambda_m(u_m-u_{m_l})\right|}{\max\limits_{m\neq m_l}|u_m-u_{m_l}|}\Delta t\\ &= -\sigma_2\frac{|u-u_{m_l}|}{\max\limits_{m\neq m_l}|u_m-u_{m_l}|}\Delta t.\end{aligned}$$

Analogously from (3.2) we obtain

$$\langle q(t_l-\Delta t), u-u_{m_{l-1}}\rangle \leq -\sigma_3\frac{|u-u_{m_l}|}{\max\limits_{m\neq m_{l-1}}|u_m-u_{m_{l-1}}|}\Delta t.$$

Hence the function $\sigma(t)$ from the refined maximum principle can be defined as

$$\sigma(t)=\bar\sigma\left(\frac{t_{l+1}-t_l}{2}-\left|t-\frac{t_{l+1}+t_l}{2}\right|\right),\quad t\in[t_l,t_{l+1}],\ l=\overline{0,L-1},$$

where $\bar\sigma>0$ is sufficiently small. □

Main inequality. The following lemma provides an inequality which will be of particular relevance to prove sufficient conditions of optimality.

Lemma 3.2. *Let σ be as in the refined maximum principle condition. Then*

$$\int_0^T \sigma(t)w(t)dt - c\left(\int_0^T w(t)dt\right)^2 \geq 0 \tag{3.7}$$

whenever $|w(\cdot)|_\infty < 1/(2\gamma c)$.

Proof. Consider the following optimal control problem

$$\begin{aligned}&\int_0^T \sigma(t)w(t)dt - cy^2(T)\to\inf,\\ &\dot y = w,\ \ w\in[0,\epsilon],\ \ y(0)=0.\end{aligned}$$

Assume that the optimal control $\hat w(\cdot)$ is different from zero. There exist $\lambda\geq 0$ and an absolutely continuous function $\psi(\cdot)$ such that

$$\begin{aligned}&\dot\psi = 0,\ \ \psi(T)=2\lambda c\hat y(T);\\ &\max_{w\in[0,\epsilon]}(\psi(t)-\lambda\sigma(t))w = (\psi(t)-\lambda\sigma(t))\hat w(t),\\ &\lambda+|\psi(\cdot)|>0.\end{aligned}$$

Obviously $\lambda\neq 0$. Set $\lambda=1$. Hence

$$\hat w(t)=\begin{cases}\epsilon & \psi(t)>\sigma(t),\\ 0 & \psi(t)<\sigma(t),\end{cases}$$

and

$$\psi(t)\equiv 2c\hat y(T)=2c\epsilon\mu,$$

where

$$M=\operatorname{meas}\{t\mid \sigma(t)<\psi(t)\}.$$

Thus, we have

$$M = \text{meas}\{t \mid \sigma(t) < 2c\epsilon M\} \le 2\gamma c\epsilon M.$$

Taking $\epsilon < 1/(2\gamma c)$ we obtain a contradiction. Thus the optimal control is zero. This implies (3.7) □

Sufficiency of the refined maximum principle. Consider the optimal control problem (2.1)-(2.3).

Theorem 3.3. *Let $(\hat{u}(\cdot), \hat{x}(\cdot))$ be an admissible control process satisfying the refined maximum principle, were $p(\cdot)$ is solution to the Cauchy problem*

$$\dot{p}(t) = -\left(\nabla(f(t,\hat{x}(t)) + g(t,\hat{x}(t))\hat{u}(t)\right)^* p(t), \quad p(T) = -\nabla\phi(\hat{x}(T)).$$

Then $(\hat{u}(\cdot), \hat{x}(\cdot))$ is a weak local minimizer in the following sense: there exists $\epsilon > 0$ such that, for any admissible control process $(u(\cdot), x(\cdot))$ satisfying $|u(\cdot) - \hat{u}(\cdot)|_\infty < \epsilon$, the inequality $\phi(x(T)) \ge \phi(\hat{x}(T))$ holds.

Proof. Indeed, from (2.8), (2.9), (2.12), and the refined maximum principle we have

$$\begin{aligned}
\phi(x(T,\bar{u}(\cdot))) &\ge \phi(\hat{x}(T)) + \langle \nabla\phi(\hat{x}(T)), \bar{x}(T)\rangle - (\text{const})\left(\int_0^T |\bar{u}(t)|dt\right)^2 \\
&= \phi(\hat{x}(T)) - \int_0^T \langle p(t), g(t,\hat{x}(t))\bar{u}(t)\rangle dt - (\text{const})\left(\int_0^T |\bar{u}(t)|dt\right)^2 \\
&\ge \phi(\hat{x}(T)) + \int_0^T \sigma(t)|\bar{u}(t)|dt - (\text{const})\left(\int_0^T |\bar{u}(t)|dt\right)^2.
\end{aligned}$$

Applying Lemma 3.2, we obtain the result. □

Let us see some examples showing the relevance of different aspects of the refined maximum principle.

Examples. Consider the optimal control problem

$$\begin{aligned}
&x_1(1) \to \min, \\
&\dot{x}_1 = x_3 - x_2^2, \\
&\dot{x}_2 = u_2, \\
&\dot{x}_3 = u_3, \\
&|u_2| + |u_3 - 1| \le 1, \\
&x_i(0) = 0, i = 1,2,3.
\end{aligned}$$

The control $(\hat{u}_2(t), \hat{u}_3(t)) \equiv (0,0)$ is a weak minimizer. Indeed, the solution to the adjoin system is $p_1 = -1$, $p_2 = 0$, $p_3 = t - 1$, and the refined maximum condition is satisfied:

$$p_3 u_3 = (t-1)u_3 \le -(1-t)\frac{|u_2| + |u_3|}{2}.$$

Here $\sigma(t) = (1-t)/2$.

The second condition in the refined maximum principle is essential. To illustrate that, consider the optimal control problem

$$\begin{aligned} & x_1(\pi/2) \to \min, \\ & \dot{x}_1 = x_3 - x_2^2, \\ & \dot{x}_2 = x_3 + u, \\ & \dot{x}_3 = -x_2, \\ & u \in [-1, 0], \\ & x_i(0) = 0, i = 1, 2, 3. \end{aligned}$$

The control $u \equiv 0$ satisfies the maximum principle with $p_1 = -1$, $p_2 = 1 - \sin t$, $p_3 = -\cos t$. The function σ in this case is $(1 - \sin t)$. Such function does not satisfy the second condition. Take the control functions sequence

$$u_n(t) = \begin{cases} 0, & t \in [0, \pi/2 - 1/n[, \\ -1, & t \in [\pi/2 - 1/n, \pi/2]. \end{cases}$$

Then, for the corresponding trajectory, we have

$$x_1(\pi/2) = -\frac{1}{6n^3} + o\left(\frac{1}{n^3}\right),$$

i.e., $u \equiv 0$ is not a local minimizer.

The following example shows that the L_∞-norm in Theorem 3.3 can not be replaced by L_p-norm, with $1 \leq p < \infty$. Consider the optimal control problem

$$\begin{aligned} & x_1(1) - x_2^2(1) \to \min, \\ & \dot{x}_1 = x_2, \\ & \dot{x}_2 = u, \\ & u \in [0, 1], \\ & x_i(0) = 0, i = 1, 2. \end{aligned}$$

The zero control satisfies the refined maximum principle with $p_1 = -1$ and $p_2 = t - 1$. Take the control functions sequence

$$u_n(t) = \begin{cases} 0, & t \in [0, 1 - 1/n[, \\ 1, & t \in [1 - 1/n, 1]. \end{cases}$$

Then, for the corresponding trajectories, we have

$$x_1(1) - x_2^2(1) = \int_{1-1/n}^{1} \int_{1-1/n}^{t} ds dt - \left(\int_{1-1/n}^{1} dt \right)^2 = -\frac{1}{2n^2}.$$

4. Problems with constraints

The generalization of the refined maximum principle to optimal control problems which accommodate not only final and initial but also path-wise state constraints is now established.

Consider the problem

$$\phi(x(T)) \to \inf \tag{4.1}$$

$$\dot{x} = f(t,x) + g(t,x)u, \;\; u \in U \tag{4.2}$$

$$x(t) \in C, \tag{4.3}$$

$$x(0) \in C_0, \;\; x(T) \in C_1. \tag{4.4}$$

Theorem 4.1. *Let $(\hat{u}(\cdot), \hat{x}(\cdot))$ be an admissible control process. Assume that there exist a function of bounded variation $p(\cdot)$ and a vector valued Borel measure μ defined in $[0,T]$, satisfying the following conditions:*

$$dp(t) = -\nabla_x (f(t,\hat{x}(t)) + g(t,\hat{x}(t))\hat{u}(t))^* p(t)dt + d\mu(t), \tag{4.5}$$

$$\int_0^T \langle x(t) - \hat{x}(t), d\mu(t)\rangle \le 0, \text{ for all admissible trajectories } x(\cdot) \tag{4.6}$$

$$\langle p(0), c_0 - \hat{x}(0)\rangle \le -(\text{const})|c_0 - \hat{x}(0)|^{2-\epsilon}, \;\; c_0 \in C_0, \; \epsilon \in]0,1], \tag{4.7}$$

$$\langle -p(T) - \nabla\phi(\hat{x}(T)), c_1 - \hat{x}(T)\rangle \le 0, \;\; c_1 \in C_1. \tag{4.8}$$

Moreover the refined maximum principle is satisfied. Then $(\hat{u}(\cdot), \hat{x}(\cdot))$ is weakly locally optimal.

Weak local optimality of $(\hat{u}(\cdot), \hat{x}(\cdot))$ must be interpreted in the sense defined before: there exists $\epsilon > 0$ such that for any admissible control process $(u(\cdot), x(\cdot))$ satisfying $|u(\cdot) - \hat{u}(\cdot)|_\infty < \epsilon$ the inequality $\phi(x(T)) \ge \phi(\hat{x}(T))$ holds.

The theorem do not impose any restriction on the support of the measure μ, usually present in necessary conditions. This restriction is already contained in condition (4.6). Observe also that this condition can usually be easily checked in practical examples.

Proof. Let $(\hat{u}(\cdot) + \bar{u}(\cdot), x(\cdot))$ be an admissible control process. We have

$$\phi(x(T)) \ge \phi(\hat{x}(T)) + \langle \nabla\phi(\hat{x}(T)), x(T) - \hat{x}(T)\rangle - (\text{const})|x(T) - \hat{x}(T)|^2.$$

Using (4.8) we obtain

$$\phi(x(T)) \ge \phi(\hat{x}(T)) - \langle p(T), x(T) - \hat{x}(T)\rangle - (\text{const})|x(T) - \hat{x}(T)|^2.$$

Let $\bar{x}(\cdot)$ be a solution to the Cauchy problem

$$\dot{\bar{x}} = \nabla_x (f(t,\hat{x}(t)) + g(t,\hat{x}(t))\hat{u}(t))\bar{x} + g(t,\hat{x}(t))\bar{u}, \;\; \bar{x}(0) = x(0) - \hat{x}(0).$$

Observe that

$$\rho(t) = |\dot{\hat{x}}(t) + \dot{\bar{x}}(t) - (f(t,\hat{x}+\bar{x}) + g(t,\hat{x}+\bar{x})(\hat{u}+\bar{u}))| \le (\text{const})(|\bar{x}|^2 + |\bar{x}||\bar{u}|).$$

Since

$$|\bar{x}| \le (\text{const}) \left(|x(0) - \hat{x}(0)| + \int_0^T |\bar{u}(t)|dt \right),$$

applying the Filippov Theorem, we get

$$|x(T) - (\hat{x}(T) + \bar{x}(T))| \;\; \le \;\; (\text{const}) \int_0^T \rho(t)dt$$

$$\leq \quad (\text{const}) \left(|x(0) - \hat{x}(0)|^2 + \left(\int_0^T |\bar{u}(t)| dt \right)^2 \right)$$

From this and an obvious inequality

$$|x(T) - \hat{x}(T)| \leq (\text{const}) \left(|x(0) - \hat{x}(0)| + \int_0^T |\bar{u}(t)| dt \right)$$

we obtain

$$\phi(x(T)) \geq$$
$$\phi(\hat{x}(T)) - \langle p(T), \bar{x}(T) \rangle - (\text{const}) \left(|x(0) - \hat{x}(0)|^2 + \left(\int_0^T |\bar{u}(t)| dt \right)^2 \right).$$

Let Φ be the fundamental matrix of the system

$$\dot{\bar{x}} = \nabla_x (f(t, \hat{x}(t)) + g(t, \hat{x}(t)) \hat{u}(t)) \bar{x}.$$

Since

$$p(t) = \Phi^*(T, t) p(T) - \int_t^T \Phi^*(s, t) d\mu(s),$$

we have

$$-\langle p(T), \bar{x}(T) \rangle = -\langle p(T), \Phi(T, 0)(x(0) - \hat{x}(0)) + \int_0^T \Phi(T, t) g(t, \hat{x}(t)) \bar{u}(t) dt \rangle$$
$$= -\langle p(0) + \int_0^T \Phi^*(s, 0) d\mu(s), x(0) - \hat{x}(0) \rangle - \int_0^T \langle p(t), g(t, \hat{x}(t)) \bar{u}(t) \rangle dt$$
$$- \int_0^T \langle \int_t^T \Phi^*(s, t) d\mu(s), g(t, \hat{x}(t)) \bar{u}(t) \rangle dt$$
$$= -\langle p(0) + \int_0^T \Phi^*(s, 0) d\mu(s), x(0) - \hat{x}(0) \rangle - \int_0^T \langle p(t), g(t, \hat{x}(t)) \bar{u}(t) \rangle dt$$
$$- \int_0^T \langle \int_0^s \Phi^*(s, t) g(t, \hat{x}(t)) \bar{u}(t) dt, d\mu(s) \rangle$$
$$= -\langle p(0), x(0) - \hat{x}(0) \rangle - \int_0^T \langle p(t), g(t, \hat{x}(t)) \bar{u}(t) \rangle dt - \int_0^T \langle \bar{x}(s), d\mu(s) \rangle$$
$$= -\langle p(0), x(0) - \hat{x}(0) \rangle - \int_0^T \langle p(t), g(t, \hat{x}(t)) \bar{u}(t) \rangle dt - \int_0^T \langle x(s) - \hat{x}(s), d\mu(s) \rangle$$
$$+ \int_0^T \langle x(s) - (\hat{x}(s) + \bar{x}(s)), d\mu(s) \rangle$$

Using (4.7), (4.6) and the refined maximum principle we have

$$\phi(x(T)) \geq \phi(\hat{x}(T)) + (\text{const}) |x(0) - \hat{x}(0)|^{2-\epsilon} + \int_0^T \sigma(t) |\bar{u}(t)| dt$$
$$-(\text{const}) \left(|x(0) - \hat{x}(0)|^2 + \left(\int_0^T |\bar{u}(t)| dt \right)^2 \right).$$

Applying the Lemma 3.2, we obtain the result. □

Note. Condition (4.7) is satisfied if C_0 is a point or a polyhedron and $\hat{x}(0)$ is a vertex. As the following example shows, parameter ϵ in (4.7) cannot be zero.

Example. Consider the problem

$$\begin{aligned} &x_2(1) - x_1^2(1) \to \inf, \\ &\dot{x}_1 = u_1, \\ &\dot{x}_2 = u_2, \\ &|u_1| + |u_2| \le 1, \\ &x_1^2(0) + (1 - x_2(0))^2 \le 1. \end{aligned}$$

The process $\hat{u}_1 \equiv 0$, $\hat{u}_2 \equiv -1$, $\hat{x}_1 \equiv 0$, $\hat{x}_2 = -t$, satisfies the maximum principle and the transversality condition

$$\langle p(0), c_0 - \hat{x}(0)\rangle \le -(\text{const})|c_0 - \hat{x}(0)|^2, \quad c_0 \in C_0.$$

(Indeed, $-x_2(0) \le -(x_1^2(0) + x_2^2(0))/2$.) However the process is not locally optimal.

5. Illustrative examples

Here we show how the above theorem can be used to analyse optimality of control processes.

Example 1. Consider a rocket car equipped with two rocket jets and moving along a straight line. Its motion is modelled by the following equations:

$$\begin{aligned} &\dot{x} = v, \\ &\dot{v} = \frac{u_1 - u_2}{m}, \\ &\dot{m} = -k(u_1 + u_2), \\ &u_1,\ u_2 \in [0, 1], \\ &x(0) = x_0,\ v(0) = v_0,\ m(0) = m_0, \\ &x(T) = v(T) = 0. \end{aligned}$$

Here $x(t)$ is the position at time t, $v(t)$ the velocity, $m(t)$ the mass of the car (changing as fuel is burned), $u_1(t)$ and $u_2(t)$ are the thrusts, and k is a constant. It is necessary to maximize the amount of fuel at the end of the motion, i.e., it is necessary to maximize $m(T)$. We introduce two new controls $w_1 = u_1 - u_2$ and $w_2 = u_1 + u_2$. The system takes the form

$$\begin{aligned} &\dot{x} = v, \\ &\dot{v} = \frac{w_1}{m}, \\ &\dot{m} = -kw_2, \\ &|w_1| + |w_2 - 1| \le 1. \end{aligned}$$

The adjoint system has the form

$$\begin{aligned} \dot{p}_1 &= 0, & p_1(T) &= \lambda_1, \\ \dot{p}_2 &= -p_1, & p_2(T) &= \lambda_2, \\ \dot{p}_3 &= \tfrac{p_2 w_1}{m^2}, & p_3(T) &= 1. \end{aligned}$$

The maximum principle reads

$$\max_{|w_1|+|w_2-1|\leq 1}\left(\frac{p_2 w_1}{m}-p_3 k w_2\right)=\left(\frac{p_2\hat{w}_1}{m}-p_3 k\hat{w}_2\right).$$

Let $x_0>0$ and $v_0>0$. The admissible processes with $m(T)>0$ and the control

$$(w_1(t),w_2(t))=\begin{cases}(-1,1), & t\in[0,\tau_1],\\ (0,0), & t\in]\tau_1,\tau_2],\\ (1,1), & t\in]\tau_2,T],\end{cases}$$

with $0<\tau_1<\tau_2<T$, is optimal. Indeed, from the form of control we see that p_2 is an increasing function, so λ_1 must be negative. Also, we have

$$q(t)=g^*(t,(x(t),v(t),m(t)))p(t)=\begin{pmatrix}\dfrac{p_2(t)}{m(t)}\\ -kp_3(t)\end{pmatrix}$$

and

$$\dot{q}(t)=\begin{pmatrix}\dfrac{-\lambda_1 m(t)+kp_2(t)w_2(t)}{m^2(t)}\\ -k\dfrac{p_2(t)w_1(t)}{m^2(t)}\end{pmatrix}.$$

Therefore conditions (3.1) and (3.2) of Lemma 3.1, at point τ_1, can be written as

$$-\frac{\lambda_1}{m(\tau_1)}(-1-0)=\frac{\lambda_1}{m(\tau_1)}<0$$

$$\frac{-\lambda_1 m(\tau_1)+kp_2\tau_1}{m(\tau_1)}(0+1)+\frac{kp_2\tau_1}{m^2(\tau_1)}(0-1)=-\frac{\lambda_1}{m(\tau_1)}>0$$

and, at point τ_2, as

$$\frac{-\lambda_1+kp_2(\tau_2)}{m^2(\tau_2)}(0-1)+\frac{-kp_2\tau_2}{m^2(\tau_2)}(0-1)=\frac{\lambda_1}{m(\tau_2)}<0$$

$$\frac{-\lambda_1}{m(\tau_2)}(1-0)=-\frac{\lambda_1}{m(\tau_2)}>0$$

Thus, from Theorem 4.1 we see that the process is optimal.

Example 2. Consider the optimal control problem

$$\begin{aligned}&\min\ x_3(T)\\ &\dot{x}_1=u_1,\\ &\dot{x}_2=u_2,\\ &\dot{x}_3=x_2-\rho x_1^2,\\ &|u_1|+|u_2|\leq 1,\\ &x_2(t)\geq c,\\ &x_1(0)=-a,\ x_1(T)=a,\\ &x_2(0)=x_2(T)=b,\end{aligned}$$

$$x_3(0) = 0,$$

where $T = 2a + 2(b - c)$. The constants ρ, a, b and c are positive, with $b > c$.

The constraint sets C, C_0, C_1 of Theorem 4.1 are represented in this problem by $C = \{(x_1, x_2, x_3) \in R^3 : x_2 \geq c\}$, $C_0 = \{(x_1, x_2, x_3) \in R^3 : x_1 = -a,\ x_2 = b \text{ and } x_3 = 0\}$ and $C_1 = \{(x_1, x_2, x_3) \in R^3 : x_1 = a \text{ and } x_2 = b\}$. The adjoint system (4.5) and (4.8) comes

$$\begin{aligned} dp_1(t) &= 2\rho\hat{x}_1(t)p_3(t) + d\mu_1(t), \\ dp_2(t) &= -p_3(t) + d\mu_2(t), \\ dp_3(t) &= d\mu_3(t), \\ p_3(T) &= -1. \end{aligned}$$

Take the admissible control process $(\hat{u}(\cdot), \hat{x}(\cdot))$ defined as:

$$\hat{x}_1(t) = \begin{cases} -a, & t \in [0, b-c[, \\ t - a - (b-c), & t \in [b-c, T-(b-c)[, \\ a, & t \in [T-(b-c), T]. \end{cases} \tag{5.1}$$

$$\hat{x}_2(t) = \begin{cases} b - t, & t \in [0, b-c[, \\ c, & t \in [b-c, T-(b-c)[, \\ t - T + b, & t \in [T-(b-c), T]. \end{cases} \tag{5.2}$$

$$\hat{u}_1(t) = \begin{cases} 0, & \text{if } t \in [0, b-c[, \\ 1, & \text{if } t \in [b-c, T-(b-c)[, \\ 0, & \text{if } t \in [T-(b-c), T]. \end{cases}$$

$$\hat{u}_2(t) = \begin{cases} -1, & \text{if } t \in [0, b-c[, \\ 0, & \text{if } t \in [b-c, T-(b-c)[, \\ 1, & \text{if } t \in [T-(b-c), T]. \end{cases}$$

Assume that $\rho < \frac{1}{2a}$ and $a > b - c$. Consider the following set of multipliers

$$\begin{aligned} p_1(t) &= (b-c)(1-2a\rho) - 2\rho \int_0^t \hat{x}_1(s)\, ds, \\ p_2(t) &= -2(b-c) + t + \int_0^t d\mu(s), \\ p_3(t) &= -1, \\ \mu(t) &= \begin{cases} 0, & t \in [0, b-c[, \\ -\dfrac{a+b-c+K}{a}(t-b+c), & t \in [b-c, T-(b-c)[, \\ -2(a-b+c), & t \in [T-(b-c), T]. \end{cases} \end{aligned}$$

The functions p_1 and p_2 can be equivalently written as

$$p_1(t) = \begin{cases} (b-c)(1-2a\rho) + 2a\rho t, & \text{if } t \in I_1, \\ (b-c)(1-2a\rho) - \rho(b-c)^2 + 2\rho(a+b-c)t - \rho t^2, & \text{if } t \in I_2, \\ (b-c)(1-2a\rho) + 4a\rho(b-c) - 2a\rho t + 4a^2\rho, & \text{if } t \in I_3. \end{cases}$$

$$p_2(t) = \begin{cases} K + t, & t \in I_1, \\ K + t - \dfrac{a+b-c+K}{a}(t-b+c), & t \in I_2, \\ K + t - 2(a+b-c-K), & t \in I_3, \end{cases}$$

with $K = -2(b-c)$ and where, for shortening, we write I_1, I_2, I_3 to denote respectively the intervals $[0, b-c[$, $[b-c, T-(b-c)[$ and $[T-(b-c), T]$.

This set of multipliers satisfies (4.5)-(4.8) of Theorem 4.1. In particular (4.6) is complied since $\mu(t)$ generates a non positive measure with support on the interval $[b-c, T-b+c]$ and on this interval $x(t) - \hat{x_2}(t) = x(t) - c \geq 0$ for any admissible trajectory x. It can also be checked that conditions under which Lemma 3.1 applies are fulfilled. In particular,

$$|\langle \dot{q}(t_l \pm 0), u_m - u_{m_l} \rangle| = |2a\rho| \neq 0.$$

We may conclude that the process $(\hat{u}(\cdot), \hat{x}(\cdot))$ is a weak local minimizer.

References

[1] M. M. A. Ferreira, A. F. Ribeiro and G.V. Smirnov, *Local minima of quadratic functionals and control of hydro-electric power stations*, JOTA-Journal of Optimization Theory and Applications, **165**, (2015), 985–1005.

[2] A. F. Filippov, *Classical solutions of differential equations with multivalued right-hand side*, SIAM J. Control **5** (1967), 609–621.

[3] N. P. Osmolovskii and H. Maurer, *Applications to Regular and Bang-Bang control: second-order necessary and sufficient conditions optimality conditions in the calculus of variations and optimal control*. SIAM, Advances in Design and Control, **24**, (2012).

[4] L. S. Pontryagin, V. G.Boltyanski, R. V. Gamkrelidze and E. F. Mischenko, *The Mathematical Theory of Optimal Processes*, Wiley-Interscience, New York, 1962.

Manuscript received 5 December 2015
revised 21 March 2016

M. M. A. Ferreira
DEEC-FEUP, University of Porto, Portugal
E-mail address: mmf@fe.up.pt

Georgi V. Smirnov
University of Minho, Portugal
E-mail address: smirnov@math.uminho.pt