Geometric Approach to Second-Order Sufficient Optimality Conditions in Optimal Control

Daniel Hoehener

Geometric Approach to Second-Order Sufficient Optimality Conditions in Optimal Control*

Daniel Hoehener[†]

November 16, 2012

Abstract

We consider the Bolza optimal control problem with control constraints and propose new second-order sufficient optimality conditions. These conditions are applicable under much weaker regularity assumptions than those known in the literature. In particular, the reference control may be merely measurable and the dynamics may be discontinuous in the time variable. Our approach is based on a new decomposition result for control constraints, involving second-order tangents.

Résumé

Approche géométrique des conditions d'optimalité suffisantes du second ordre pour des problèmes de contrôle optimal. Pour un problème de contrôle optimal sous forme de Bolza avec des contraintes sur les contrôles, nous proposons des nouveaux conditions d'optimalités suffisantes du second ordre. Ces conditions demandent peu de régularité des données. En particulier le contrôle de référence peut être seulement mesurable et les dynamiques peuvent être discontinues par rapport au temps. Notre approche est basée sur un nouveau résultat de décomposition pour les contraintes sur les contrôles. Cette décomposition utilise des tangents du second ordre.

1 Introduction

This paper is devoted to second-order sufficient optimality conditions for the following optimal control problem:

$$\text{Minimize } \varphi\left(x(1)\right) + \int_0^1 \ell\left(t, x(t), u(t)\right) dt, \tag{P}$$

over absolutely continuous $x\colon [0,1] \to \mathbb{R}^n$ and measurable $u\colon [0,1] \to \mathbb{R}^m$ satisfying,

$$\begin{cases} \dot{x}(t) = f(t, x(t), u(t)), \quad u(t) \in U(t), \qquad a.e. \\ x(0) = x_0, \end{cases} \tag{1}$$

where the maps $f\colon [0,1] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $\ell\colon [0,1] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $\varphi\colon \mathbb{R}^n \to \mathbb{R}$, the set-valued map $U\colon [0,1] \rightsquigarrow \mathbb{R}^m$ and the initial state $x_0 \in \mathbb{R}^n$ are given.

Denote by $W^{1,1}([0,1];\mathbb{R}^n)$ the space of absolutely continuous mappings. The set of all measurable mappings $u\colon [0,1] \to \mathbb{R}^m$ such that $u(t) \in U(t)$ a.e. will be denoted by $\mathcal{U}$. A *process* (x, u) comprises

*Financial support by the European Commission (FP7-PEOPLE-2010-ITN, Grant Agreement no. 264735-SADCO) is gratefully acknowledged

[†]Combinatoire & Optimisation, Institut de Mathématiques de Jussieu (UMR 7586), Université Pierre et Marie Curie, 4 place Jussieu, 75252 Paris cedex 05, France, email: hoehener@math.jussieu.fr

a *control function* $u \in \mathcal{U}$ and a *state trajectory* $x \in W^{1,1}([0,1];\mathbb{R}^n)$ satisfying the differential equation in (1). Under standard assumptions the state trajectory x is uniquely defined by a control function u and the control system in (1). Consequently, the cost functional $J\colon \mathcal{U} \to \mathbb{R}$,

$$J(u) := \varphi\left(x_u(1)\right) + \int_0^1 \ell\left(t, x_u(t), u(t)\right) dt, \tag{2}$$

where $x_u \in W^{1,1}([0,1];\mathbb{R}^n)$ is the solution of (1) corresponding to the control u, is well defined.

Definition 1.1 *A process* $(\bar{x}, \bar{u})$ *is a* (strict) weak local minimizer *if there exists* $\bar{\rho} > 0$ *such that* $\bar{u}$ *minimizes* J *(strictly) over all* $\tilde{u} \in \mathcal{U}$ *satisfying* $\|\tilde{u} - \bar{u}\|_\infty \leq \bar{\rho}$.

In this note we present a new approach to second-order sufficient conditions. We consider set-valued maps $U(\cdot)$ where the boundary of $U(t)$ is smooth. Then we use the geometry of the sets $U(t)$ in order to show that there exists a neighborhood of the reference control $\bar{u}$, such that for every $\tilde{u} \in \mathcal{U}$ in this neighborhood, one can approximate $\tilde{u} - \bar{u}$ sufficiently well by first- and second-order tangents. Unlike in [2], [4], [6], where the sufficient conditions were obtained by using abstract results, our approach permits a direct and geometric proof. So far direct methods required the reference control $\bar{u}$ to be continuous or piecewise continuous, see for instance [5]. With our approach we are able to state sufficient conditions when $\bar{u}$ is merely measurable. Moreover the dynamics can be discontinuous in the time variable and no special structure of the Hamiltonian is required. In particular we do not impose a Legendre-Clebsch condition. Our sufficient optimality conditions form no-gap second-order optimality conditions with the necessary optimality conditions of [3, Thm. 3.2]. Finally, this approach can be extended to sets $U(t)$ having nonsmooth boundary.

2 Notations and assumptions

As usual $B(x, \rho)$ stands for the open ball with center $x \in \mathbb{R}^n$ and radius $\rho > 0$. ∂K denotes the boundary of a set $K \subset \mathbb{R}^n$ and $T_K(x)$ the Bouligand tangent cone to K at x, see [1]. Throughout the paper $(\bar{x}, \bar{u}) \in W^{1,1}([0,1];\mathbb{R}^n) \times \mathcal{U}$ denotes a fixed reference process. In order to simplify the notations, we will abbreviate $(t, \bar{x}(t), \bar{u}(t))$ by $[t]$, thus for instance $f[t] := f(t, \bar{x}(t), \bar{u}(t))$. Furthermore we define,

$$\mathcal{T}_B := \{t \in [0,1] \mid \bar{u}(t) \in \partial U(t)\}.$$

For a map $\phi\colon [0,1] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$ we denote by $\phi_x(t_0, x_0, u_0)$, $\phi_u(t_0, x_0, u_0)$, $\phi_{xx}(t_0, x_0, u_0)$, $\phi_{xu}(t_0, x_0, u_0)$, $\phi_{uu}(t_0, x_0, u_0)$ its partial- and second-order partial derivatives with respect to x and/or u at (t_0, x_0, u_0), assuming these derivatives exist. The derivative, resp. the Hessian, of the map $(x, u) \mapsto \phi(t, x, u)$ evaluated at (x_0, u_0) is denoted by $\phi'(t, x_0, u_0)$, resp. $\phi''(t, x_0, u_0)$.
We require the following regularity of the data: There exists $\rho_0 > 0$ such that,

(A1) (a) $\forall x \in \mathbb{R}^n, \forall u \in \mathbb{R}^m$, $f(\cdot, x, u)$ and $\ell(\cdot, x, u)$ are measurable and for almost all $t \in [0,1]$, $f(t, \cdot, \cdot)$, $\ell(t, \cdot, \cdot)$ are continuously differentiable on $B(\bar{x}(t), \rho_0) \times B(\bar{u}(t), \rho_0)$. Moreover, $\varphi \in \mathcal{C}^2(\mathbb{R}^n;\mathbb{R})$;

(b) $\forall R > 0$, there exists $k_R \in L^1([0,1];\mathbb{R}_+)$ such that for a.e. $t \in [0,1]$, $f(t, \cdot, u)$ and $\ell(t, \cdot, u)$ are $k_R(t)$-Lipschitz on RB for all $u \in B(\bar{u}(t), \rho_0)$;

(c) There exists $\gamma \in L^1([0,1];\mathbb{R}_+)$ such that for a.e. $t \in [0,1]$ and all $x \in \mathbb{R}^n$,

$$|f(t, x, u)| \leq \gamma(t)(1 + |x|), \quad \forall u \in B(\bar{u}(t), \rho_0);$$

(d) $\forall R > 0$, the mapping $t \mapsto \sup_{(x,u) \in RB} |\ell(t, x, u)|$ is integrable on $[0,1]$;

(A2) $f_x[\cdot]$ is integrable and $\exists a_1 > 0$ s.t. for a.e. $t \in [0,1]$, $\forall x, y \in B(\bar{x}(t), \rho_0)$, $\forall u, v \in B(\bar{u}(t), \rho_0)$,

$$\|f_u[t]\| \leq a_1 \quad \text{and} \quad \|f'(t,x,u) - f'(t,y,v)\| \leq a_1(|x-y| + |u-v|);$$

(A3) (a) For almost all $t \in [0,1]$, $f(t,\cdot,\cdot)$ is twice differentiable on $B(\bar{x}(t), \rho_0) \times B(\bar{u}(t), \rho_0)$;

(b) $\exists a_2 > 0$ such that for a.e. $t \in [0,1]$, $\forall x, y \in B(\bar{x}(t), \rho_0)$, $\forall u, v \in B(\bar{u}(t), \rho_0)$,

$$\|f''[t]\| \leq a_2 \quad \text{and} \quad \|f''(t,x,u) - f''(t,y,v)\| \leq a_2(|x-y| + |u-v|);$$

(A4) For all $t \in [0,1]$, $U(t) := \{u \in \mathbb{R}^m \mid g(t,u) \leq 0\}$, where the function $g\colon [0,1] \times \mathbb{R}^m \to \mathbb{R}$ satisfies,

(a) $g(\cdot, u)$ is measurable for all $u \in \mathbb{R}^m$ and there exists $K_g > 0$ such that for a.e. $t \in [0,1]$, $g(t,\cdot)$ is twice continuously differentiable on $B(\bar{u}(t), \rho_0)$ and for all $u \in B(\bar{u}(t), \rho_0)$,

$$\|g_{uu}(t,u) - g_{uu}(t,\bar{u}(t))\| \leq K_g\,|u - \bar{u}(t)|\,, \qquad \|g_{uu}(t,\bar{u}(t))\| \leq K_g;$$

(b) There exists $\mu > 0$ such that $|\nabla_u g(t,\bar{u}(t))| \geq \mu$ for a.e. $t \in \mathcal{T}_B$. For convenience we assume that $|\nabla_u g(t,\bar{u}(t))| = 1$ for a.e. $t \in \mathcal{T}_B$.

When (A2) or (A3) are imposed, then we always assume that ℓ satisfies the same assumptions as f. Note that if (A4)(a) is satisfied, then one may also parameterize $U(t)$ with the oriented distance $b_{U(t)}(\cdot) = \mathrm{d}(\cdot; U(t)) - \mathrm{d}(\cdot; \mathbb{R}^m \setminus U(t))$, where $\mathrm{d}(x;K) := \inf_{k \in K} |x-k|$. In this case (A4)(b) is automatically verified.

3 Main results

Define the Hamiltonian $\mathcal{H}\colon [0,1] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ by $\mathcal{H}(t,x,p,u) := \langle p, f(t,x,u)\rangle - \ell(t,x,u)$.

Theorem 3.1 *Let $(\bar{x}, \bar{u})$ be a weak local minimizer of problem (P) and (A1), (A2), (A4) hold true. Then there exists $p \in W^{1,1}([0,1];\mathbb{R}^n)$ such that for almost every $t \in [0,1]$,*

1. $-\dot{p}(t) = \mathcal{H}_x(t,\bar{x}(t),p(t),\bar{u}(t)), \qquad p(1) = -\nabla\varphi(\bar{x}(1))$*;*
2. $\mathcal{H}_u(t,\bar{x}(t),p(t),\bar{u}(t))u \leq 0$*, for all* $u \in T_{U(t)}(\bar{u}(t))$*.*

A map $p \in W^{1,1}([0,1];\mathbb{R}^n)$ satisfying (1)-(2) is called *adjoint state.* A triple $(\bar{x},\bar{u},p) \in W^{1,1}([0,1];\mathbb{R}^n) \times \mathcal{U} \times W^{1,1}([0,1];\mathbb{R}^n)$ is called *extremal* if $(\bar{x},\bar{u})$ is a process and p is the corresponding adjoint state. For a given extremal $(\bar{x},\bar{u},p)$, we will use the abbreviation $[t] := (t,\bar{x}(t),p(t),\bar{u}(t))$ when evaluating the Hamiltonian or its (partial) derivatives at $(t,\bar{x}(t),p(t),\bar{u}(t))$.
For any $w \in L^\infty([0,1];\mathbb{R}^m)$ the solution $y \in W^{1,1}([0,1];\mathbb{R}^n)$ of the linear system,

$$\dot{y}(t) = f_x[t]y(t) + f_u[t]w(t) \ \text{ a.e.} \qquad y(0) = 0, \tag{3}$$

is called *the solution of the linearized system corresponding to w.* From now on $\mathcal{H}''(t,x,p,u)$ denotes the Hessian of the function $(x,u) \mapsto \mathcal{H}(t,x,p,u)$ and for $t \in \mathcal{T}_B$, $\pi(t,u)$ stands for the metric projection of u onto $\partial T_{U(t)}(\bar{u}(t))$. For convenience we set $\pi(t,\cdot) \equiv 0$ if $t \in [0,1] \setminus \mathcal{T}_B$.

Theorem 3.2 *Let $(\bar{x}, \bar{u}, p)$ be an extremal and (A1)-(A4) hold true. Assume there exist $\gamma_1, \gamma_2 > 0$ such that for all $u \in L^\infty([0,1];\mathbb{R}^m)$, satisfying $u(t) \in T_{U(t)}(\bar{u}(t))$ a.e. in $\mathcal{T}_B$ and $u(t) \in \partial T_{U(t)}(\bar{u}(t))$ for a.e. $t \in \{s \in [0,1] \mid |\mathcal{H}_u[s]| \geq \gamma_1\}$, we have that the following inequality holds true,*

$$\varphi''(\bar{x}(1))y(1)y(1) + \int_0^1 \big(|\mathcal{H}_u[t]|\, g_{uu}(t, \bar{u}(t))\pi(t, u(t))\pi(t, u(t)) - \mathcal{H}''[t](y(t), u(t))(y(t), u(t))\big)dt$$
$$\geq \gamma_2 \int_0^1 \left(|u(t)|^2\right) dt, \quad (4)$$

where $y \in W^{1,1}([0,1];\mathbb{R}^n)$ is the solution of the linearized control system (3) *corresponding to u. Then $(\bar{x}, \bar{u})$ is a strict weak local minimizer of problem (P).*

Remark 1 For $t \in \mathcal{T}_B$ and $u \in \partial T_{U(t)}(\bar{u}(t)) \cap S^{m-1}$, $-g_{uu}(t, \bar{u}(t))uu$ is the normal curvature of $\partial U(t)$ at $\bar{u}(t)$ with respect to u. Therefore condition (4) *is related to the curvature of $\partial U(t)$ at $\bar{u}(t)$.*

To prove Theorem 3.2 we consider an arbitrary $\tilde{u} \in \mathcal{U}$ such that $\|\tilde{u} - \bar{u}\|_\infty$ is small enough. For $t \in [0,1]\backslash\mathcal{T}_B$ set $u(t) = \tilde{u}(t) - \bar{u}(t)$ and $v(t) = 0$. On the other hand, if $t \in \mathcal{T}_B$, we set $u'(t) = \pi(t, \tilde{u}(t) - \bar{u}(t))$. Then one shows that $\tilde{u}(t) - u'(t) - \bar{u}(t) = v'(t) + r(t)$, where $v'(t)$ is such that $g_u(t, \bar{u}(t))v'(t) + \frac{1}{2}g_{uu}(t, \bar{u}(t))u'(t)u'(t) \leq 0$ holds true and $|r(t)| / |\tilde{u}(t) - \bar{u}(t)|^2 \to 0$, when $|\tilde{u}(t) - \bar{u}(t)| \to 0+$. Finally, we define $u(t) = u'(t)$, $v(t) = v'(t)$ for all $t \in \{s \in \mathcal{T}_B \mid |\mathcal{H}_u[s]| \geq \gamma_1 \text{ or } g_u(s, \bar{u}(s))v'(s) \geq 0\} =: A$ and $u(t) = u'(t) + v'(t)$, $v(t) = 0$ for all $t \in \mathcal{T}_B \setminus A$. Having this, one uses the variational analysis performed in [3] in order to conclude.

Remark 2 By (A4) and Theorem 3.1 we have that for a.e. $t \in [0,1]$, $|\mathcal{H}_u[t]|\, g(t, \bar{u}(t)) = 0$ and $\mathcal{H}_u[t] - |\mathcal{H}_u[t]|\, g_u(t, \bar{u}(t)) = 0$. Thus, in the case when there exists $\gamma_1 > 0$ such that,

$$\lambda\left(\{t \in [0,1] \mid |\mathcal{H}_u[t]| > 0\} \setminus \{t \in [0,1] \mid |\mathcal{H}_u[t]| \geq \gamma_1\}\right) = 0, \quad (5)$$

here $\lambda(\cdot)$ is the Lebesgue measure, (4) *becomes the classical sufficient condition of [4]. Note that in this paper the authors impose a strengthened Legendre-Clebsch condition and assume that (A4)(b) holds on a neighborhood of $\mathcal{T}_B$, in order to obtain these results for an essentially bounded $\bar{u}$. In [6] the same assumptions as in [4] are made and in addition it is required that* (5) *is satisfied. Finally, in [2] the authors consider also pure state constraints but they impose an even stronger version of the strengthened Legendre-Clebsch condition. That is they need that there exists $\gamma > 0$ such that for almost all $t \in [0,1]$, $\mathcal{H}_{uu}[t]vv - |\mathcal{H}_u[t]|\, g_{uu}(t, \bar{u}(t))vv \leq -\gamma |v|^2$ for all $v \in \mathbb{R}^m$.*

Example 1 Consider $\varphi \equiv 0$, $n = m = 2$, $x = (x_1, x_2)^T$, $u = (u_1, u_2)^T$, $x_0 = 0$ and let $f\colon \mathbb{R}^4 \to \mathbb{R}^2$, $\ell\colon \mathbb{R}^4 \to \mathbb{R}$, $g\colon \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x_1, x_2, u_1, u_2) = (u_1, u_1 + u_2)^T$, $\ell(x_1, x_2, u_1, u_2) = -\frac{1}{4}x_1^2 - u_1^4 - u_2 + 1$ and $g(u_1, u_2) = u_1^2 + u_2 - 1$. Then for $\bar{u}(t) \equiv (0,1)^T$, $\bar{x}(t) = (0, t)^T$ and $p(\cdot) \equiv 0$, $(\bar{x}, \bar{u}, p)$ is an extremal. Let $\gamma_1 = 1$ and observe that $\{t \in [0,1] \mid |\mathcal{H}_u[t]| \geq 1\} = [0,1]$. Consider $u \in L^\infty([0,1];\mathbb{R}^2)$ as in Theorem 3.2, that is for almost all $t \in [0,1]$, $u(t) \in \mathbb{R} \times \{0\}$. Using this we find that,

$$\int_0^1 (|\mathcal{H}_u[t]|\, g_{uu}(t, \bar{u}(t))u(t)u(t) - \mathcal{H}_{xx}[t]y(t)y(t))\, dt \geq \frac{3}{2}\int_0^1 |u(t)|^2\, dt,$$

where $y(\cdot)$ is the solution of (3) *corresponding to u. Since $\mathcal{H}_{xu}[t] = 0$ and $\mathcal{H}_{uu}[t] = 0$ a.e., we have that $(\bar{x}, \bar{u})$ is a strict weak local minimizer.*

Note that in this example the strengthened Legendre-Clebsch condition of [2], see Remark 2, is not satisfied.

References

[1] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis, Systems Control Found. Appl.*, Birkhäuser Boston, Inc., Boston, MA, 1990.

[2] J. F. Bonnans and A. Hermant, *Second-order analysis for optimal control problems with pure state constraints and mixed control-state constraints*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 26 (2009), pp. 561–598.

[3] D. Hoehener, *Variational approach to second-order optimality conditions for control problems with pure state constraints*, SIAM J. Control Optim., 50 (2012), pp. 1139–1173.

[4] H. Maurer and H. J. Oberle, *Second order sufficient conditions for optimal control problems with free final time: The Riccati approach*, SIAM J. Control Optim., 41 (2002), pp. 308–403.

[5] A. A. Milyutin and N. P. Osmolovskii, *Calculus of Variations and Optimal Control*, American Mathematical Society, Providence, RI, 1998.

[6] V. Zeidan, *The Riccati equation for optimal control problems with mixed state-control constraints: necessity and sufficiency*, SIAM J. Control Optim., 32 (1994), pp. 1297–1321.