No-gap Second-order Optimality Conditions for State Constrained Optimal Control Problems

J. Frédéric Bonnans   Audrey Hermant

INRIA Rocquencourt, France

Workshop on Advances in Continuous Optimization
Reykjavik, Iceland, 30 June and 1 July 2006
Outline

1. Presentation of the problem, motivations
2. Definitions and assumptions
3. Main result
4. Application to the shooting algorithm
The Optimal Control Problem

\[(P) \quad \min_{(u,y) \in U \times Y} \int_0^T \ell(u(t), y(t)) dt + \phi(y(T)) \quad \text{subject to:} \]

\[
\dot{y}(t) = f(u(t), y(t)) \quad \text{a.e. on } [0, T] \ ; \quad y(0) = y_0 \\
g(y(t)) \leq 0 \quad \text{on } [0, T].
\]

- Control and state spaces: \( U := L^\infty(0, T; \mathbb{R}) \), \( Y := W^{1, \infty}(0, T; \mathbb{R}^n) \).

- Assumptions:

(A0) The mappings \( \ell : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R} \), \( \phi : \mathbb{R}^n \to \mathbb{R} \), \( f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n \) and \( g : \mathbb{R}^n \to \mathbb{R} \) are \( C^\infty \); \( f \) is Lipschitz continuous.

(A1) The initial condition satisfies \( g(y_0) < 0 \).
Why study Second-Order Optimality Conditions?

- **Second-order Sufficient Conditions**: analysis of convergence of numerical algorithms, stability and sensitivity analysis.

- Strong second-order sufficient conditions known, in e.g. [Malanowski-Maurer et al. 1997,1998,2001,2004] ...

- To weaken the sufficient condition, find a Second-order Sufficient Condition as close as possible to the Second-order Necessary Condition (no gap).

- No-gap Second-order conditions known for *mixed* control-state constraints [Milyutin-Osmolovskii 1998], [Zeidan 1994] ...
Abstract formulation of Optimal Control Problem

- State mapping \( \mathcal{U} \to \mathcal{Y}, \ u \mapsto y_u \), where \( y_u \) is the solution of:
  \[
  \dot{y}_u(t) = f(u(t), y_u(t)) \quad \text{for a.a. } t \in [0, T]; \quad y_u(0) = y_0
  \]

- Cost and constraint mappings \( J : \mathcal{U} \to \mathbb{R}, \ G : \mathcal{U} \to C[0, T] : \)
  \[
  J(u) = \int_0^T \ell(u(t), y_u(t)) \, dt + \phi(y_u(T)) \quad ; \quad G(u) = g(y_u).
  \]

- Abstract formulation of \((\mathcal{P})\) is:
  \[
  \min_{u \in \mathcal{U}} J(u); \quad G(u) \in K,
  \]
  where \( K \) is the cone of nonpositive continuous functions \( C_-[0, T] \).
Definitions (1/3) Structure of a trajectory

- State constraint: $g(y_u(t)) \leq 0$, $\forall t \in [0, T]$.
- Contact set:

$$I(g(y_u)) := \{ t \in [0, T] ; g(y_u(t)) = 0 \}.$$ 

- Boundary arc $[\tau_{en}, \tau_{ex}]$ → entry and exit points
- Isolated contact point $\{\tau_{to}\}$ → touch points

- Junction points:

$$\mathcal{T} := \partial I(g(y_u)).$$
Order of the state constraint \( q \): smallest number of time-derivation of the function
\[ t \mapsto g(y_u(t)), \]
so that an explicit dependence in the control variable \( u \) appears.
\[ g^{(j)}(u, y) := g^{(j-1)}_y(y)f(u, y) \quad 1 \leq j \leq q, \quad (u, y) \in \mathbb{R} \times \mathbb{R}^n \]
\[ g_u^{(j)} \equiv 0, \quad 0 \leq j \leq q - 1 \quad \text{and} \quad g_u^{(q)} \not\equiv 0. \]

Example of a state constraint of order \( q \):
\[ y^{(q)}(t) = u(t) \quad ; \quad y(t) \leq 0. \]
\[(\mathcal{P}) \quad \min J(u) \; ; \; G(u) \in K\]

- **Lagrangian** \( L : \mathcal{U} \times \mathcal{M}[0, T] \to \mathbb{R} \)

\[
L(u, \eta) := J(u) + \langle \eta, G(u) \rangle \\
= \int_0^T \ell(u(t), y_u(t))dt + \phi(y_u(T)) + \int_0^T g(y_u(t))d\eta(t)
\]

- **Hamiltonian** \( H : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^{n*} \to \mathbb{R}, \)

\[
H(u, y, p) := \ell(u, y) + pf(u, y).
\]

- **Costate** \( p_{u, \eta} \) : the solution in \( BV([0, T]; \mathbb{R}^{n*}) \) of:

\[
-dp_{u, \eta} = H_y(u, y_u, p_{u, \eta})dt + g_y(y_u)d\eta \; ; \; p_{u, \eta}(T) = \phi_y(y_u(T)).
\]
(A2) Strong convexity of the Hamiltonian w.r.t. the control variable: \( \exists \alpha > 0 \) such that

\[
\alpha \leq H_{uu}(w, y_u(t), p_u, \eta(t^-)) \quad \text{for all } w \in \mathbb{R} \text{ and } t \in [0, T].
\]

(A3) Constraint Regularity: \( \exists \gamma, \varepsilon > 0 \) such that

\[
\gamma \leq |g_u^{(q)}(u(t), y_u(t))| \quad \text{for a.a. } t, \text{ dist}\{t ; l(g(y_u))\} \leq \varepsilon.
\]

(A4) Finite set of junctions points \( T \), and \( g(y_u(T)) < 0 \).
\( u \in \mathcal{U} \) is a \textbf{stationary point} of \((\mathcal{P})\), if there exists a \textbf{Lagrange multiplier} \( \eta \in \mathcal{M}_+ [0, T] \) such that

\[
\begin{aligned}
D_u L(u, \eta) &= H_u(u(\cdot), y_u(\cdot), p_{u,\eta}(\cdot)) = 0, \text{ a.e. on } [0, T] \\
\eta &\in N_K(G(u)).
\end{aligned}
\]
Junctions conditions results

- \( u \in \mathcal{U} \) is a stationary point of \((\mathcal{P})\), if there exists a Lagrange multiplier \( \eta \in \mathcal{M}_+[0, T] \) such that

\[
\begin{align*}
D_u L(u, \eta) &= H_u(u(\cdot), y_u(\cdot), p_{u, \eta}(\cdot)) = 0, \text{ a.e. on } [0, T] \\
\eta &\in N_K(G(u)).
\end{align*}
\]

Proposition (Jacobson, Lele and Speyer, 1971)

Let \((u, \eta) \in \mathcal{U} \times \mathcal{M}_+[0, T] \) a stationary point and its (unique) Lagrange multiplier, satisfying (A2)-(A4). Then:

- \( u \) and \( \eta \) are \( C^\infty \) on \([0, T] \setminus \mathcal{T} \) ⇒ \( d\eta = \eta_0 dt + \sum_{\tau \in \mathcal{T}} \nu_{\tau} \delta_{\tau} \)
- \( u, \ldots, u^{(q-2)} \) are continuous at junctions times;
- If \( q \) is odd, \( u^{(q-1)} \) and \( \eta \) are continuous at entry/exit times;
- If \( q = 1 \), \( \eta \) is continuous at touch points.
### Proposition (Jacobson, Lele and Speyer, 1971)

Let \((u, \eta) \in \mathcal{U} \times \mathcal{M}_+ [0, T]\) a stationary point and its (unique) Lagrange multiplier, satisfying (A2)-(A4). Then:

- \(u\) and \(\eta\) are \(C^\infty\) on \([0, T] \setminus \mathcal{T}\) \(\Rightarrow\) \(d\eta = \eta_0 dt + \sum_{\tau \in \mathcal{T}} \nu_\tau \delta_\tau\)

- \(u, \ldots, u^{(q-2)}\) are continuous at junctions times;

- If \(q\) is odd, \(u^{(q-1)}\) and \(\eta\) are continuous at entry/exit times;

- If \(q = 1\), \(\eta\) is continuous at touch points.

**Consequence:** the time-derivatives of \(t \mapsto g(y_u(t))\) are continuous at entry/exit points until order \(\hat{q}\), with \(\hat{q} := 2q - 2\) if \(q\) is even, and \(\hat{q} = 2q - 1\) if \(q\) is odd.

- A touch point \(\tau\) is said to be essential, if \(\tau \in \text{supp}(\eta)\) (equivalently, if \(\nu_\tau \neq 0\) or if \(\eta\) is discontinuous at \(\tau\)).
(A5)(i) Non-Tangentiality condition at entry/exit points:

\[ (-1)^{q+1} \frac{d^{q+1}}{dt^{q+1}} g(y_u(t)) \big|_{t=\tau_{en}} < 0 \quad ; \quad \frac{d^{q+1}}{dt^{q+1}} g(y_u(t)) \big|_{t=\tau_{ex}} < 0 \]

(A5)(ii) Reducibility Condition at essential touch points \((q \geq 2)\):

\[ \frac{d^2}{dt^2} g(y_u(t)) \big|_{t=\tau_{to}^{ess}} = g^{(2)}(u(\tau_{to}^{ess}), y_u(\tau_{to}^{ess})) < 0 \]

(A6) **Strict Complementarity** on boundary arcs:

\[ \text{int} \{ g(y_u) \} = \bigcup [\tau_{en}, \tau_{ex}] \subset \text{supp}(\eta) \]
Main Result

- For $u \in \mathcal{U}$ and $v \in L^2(0, T)$, the linearized state $z_{u,v}$ is the solution in $H^1(0, T; \mathbb{R}^n)$ of

$$
\dot{z}_{u,v} = f_u(u, y_u)v + f_y(u, y_u)z_{u,v} \text{ on } [0, T] ; \quad z_{u,v}(0) = 0.
$$

Note that $(DG(u)v)(t) = g_y(y_u(t))z_{u,v}(t)$.

- For $u$ a stationary solution with multiplier $\eta$, the critical cone is

$$
C_2(u) := \{ v \in L^2 ; \ DG(u)v \in T_K(G(u)) ; \ DJ(u)v \leq 0 \}
= \{ v \in L^2 ; \ DG(u)v \in T_K(G(u)) ; \ supp(\eta) \subset I^2_{u,v} \}
$$

with the second-order contact set:

$$
I^2_{u,v} := \{ t \in I(g(y_u)) ; \ g_y(y_u(t))z_{u,v}(t) = 0 \}.
$$
Main Result

Theorem (No-gap Second-order Necessary Condition)

Let $u \in \mathcal{U}$ a local optimal solution of $(\mathcal{P})$, with (unique) Lagrange multiplier $\eta$, satisfying (A1)-(A6). Denote by $T_{to}^{\text{ess}}$ the set of essential touch points of the trajectory $(u, y_u)$ and $\nu_\tau = [\eta(\tau)]$. Then, for all $v \in C_2(u)$:

$$D^2_{u u} L(u, \eta)(v, v) - \sum_{\tau \in T_{to}^{\text{ess}}} \nu_\tau \frac{(g^{(1)}_y(y_u(\tau))z_{u,v}(\tau))^2}{g^{(2)}(u(\tau), y_u(\tau))} \geq 0.$$
Main Result

Theorem (No-gap Second-order Necessary Condition)

Let \( u \in \mathcal{U} \) a local optimal solution of \((P)\), with (unique) Lagrange multiplier \( \eta \), satisfying (A1)-(A6). Denote by \( T_{\text{ess}} \) the set of essential touch points of the trajectory \((u, y_u)\) and \( \nu_\tau = [\eta(\tau)] \).

Then, for all \( v \in C_2(u) \):

\[
D_{uu}^2 L(u, \eta)(v, v) - \sum_{\tau \in T_{\text{ess}}} \nu_\tau \frac{(g^{(1)}_y(y_u(\tau))z_{u,v}(\tau))^2}{g^{(2)}(u(\tau), y_u(\tau))} \geq 0.
\]

- Additional term (in blue), called the curvature term [Kawasaki, 1988].
- Only essential touch points have a contribution to the curvature term (the contribution of boundary arcs is null).
Theorem (No-gap Second-order Sufficient Condition)

Let \((u, \eta) \in \mathcal{U} \times \mathcal{M}_+(0, T)\) a stationary point and its multiplier, satisfying (A1)-(A6). The following assertions are equivalent:

(i) For all \(v \in C^2(u) \setminus \{0\}\),

\[
D^2_{uu} L(u, \eta)(v, v) - \sum_{\tau \in T_{to}^{ess}} \nu_{\tau} \frac{(g_y^{(1)}(y_u(\tau))z_{uu, v}(\tau))^2}{g^{(2)}(u(\tau), y_u(\tau))} > 0.
\]

(ii) \(u\) is a local optimal solution of \((P)\) satisfying the quadratic growth condition: there exists \(\beta, r > 0\) such that

\[
J(\tilde{u}) \geq J(u) + \beta \|\tilde{u} - u\|_2^2 \quad \text{for all} \quad G(\tilde{u}) \in K, \quad \|\tilde{u} - u\|_\infty < r.
\]
Reduction Approach (cf. semi-infinite programming)

- Let $x_0 \in W^{2,\infty}(0, T)$ having a local maximum at $t_0 \in (0, T)$, the latter being reducible:
  $$\ddot{x}_0 \text{ is continuous at } t_0 \quad \text{and} \quad \ddot{x}_0(t_0) < 0.$$

- Then there exists $\varepsilon, \delta > 0$ such that for all $x \in W^{2,\infty}(0, T)$,
  $$\|x - x_0\|_{2,\infty} \leq \delta,$$
  $x$ attains its unique maximum on $[t_0 - \varepsilon, t_0 + \varepsilon]$ at time $t_x$, and
  $$x(t) \leq 0 \text{ on } [0, T] \iff \begin{cases} 
  x(t) \leq 0 \text{ on } [0, T] \setminus (t_0 - \varepsilon, t_0 + \varepsilon) \\
  x(t_x) \leq 0.
  \end{cases}$$

- The additional term comes from the second-order derivative of the mapping $x \mapsto x(t_x)$. 

J. Frédéric Bonnans, Audrey Hermant

No-gap Second-order Optimality Conditions for State Constrained Optimal Control Problems
Application to the shooting algorithm

- **Unconstrained case:**

$$\begin{cases} 
\dot{y} &= f(u, y) \quad ; \quad y(0) = y_0 \\
-\dot{p} &= H_y(u, y, p) \quad ; \quad p(T) = \phi_y(y(T)) \\
0 &= H_u(u, y, p).
\end{cases}$$

- By (A2),

$$0 = H_u(u(t), y(t), p(t)) \iff u(t) = \Upsilon(y(t), p(t)).$$

- **Shooting mapping** $\mathcal{F}: \mathbb{R}^n \mapsto \mathbb{R}^n,$

$$p_0 \mapsto p(T) - \phi_y(y(T)),$$

with $(y, p)$ solution of:

$$\begin{cases} 
\dot{y} &= f(\Upsilon(y, p), y) \quad ; \quad y(0) = y_0 \\
-\dot{p} &= H_y(\Upsilon(y, p), y, p) \quad ; \quad p(0) = p_0.
\end{cases}$$
Assume that $q \geq 2$ and there is one isolated contact point $\tau$. Then the shooting mapping is defined by $\mathcal{F} : \mathbb{R}^{n+2} \rightarrow \mathbb{R}^{n+2}$,

$$
\begin{pmatrix}
    p_0 \\
    \nu \\
    \tau
\end{pmatrix}
\mapsto
\begin{pmatrix}
    p(T) - \phi_y(y(T)) \\
    g(y(\tau)) \\
    g^{(1)}(y(\tau))
\end{pmatrix},
$$

where $(y, p)$ is solution of:

$$
\begin{cases}
    \dot{y} = f(\Upsilon(y, p), y) & \text{on } [0, T]; \\
    -\dot{p} = H_y(\Upsilon(y, p), y, p) & \text{on } [0, \tau) \cup (\tau, T]; \\
    [p(\tau)] = -\nu g_y(y(\tau)).
\end{cases}
$$

Additional conditions: $g(y(t)) \leq 0$ and $\nu \geq 0$. 
Application to the shooting algorithm

- Shooting algorithm well-posed ⇔ the Jacobian of the shooting mapping $D\mathcal{F}(p_0, \nu, \tau)$ is invertible.
- Solution of $D\mathcal{F}(p_0, \nu, \tau)(\pi_0, \gamma, \sigma) = 0$?

$$(PQ) \min_{v, z \in L^2 \times H^1} \frac{1}{2} \left\{ \int_0^T D_{(u,y)(u,y)}^2 H(u, y, p)((v, z), (v, z)) dt \right.$$ \\
$$+ \phi_{yy}(y(T))(z(T), z(T)) + \nu g_{yy}(y(\tau))(z(\tau), z(\tau))$$ \\
$$- \nu \frac{(g_y^{(1)}(y(\tau))z(\tau))^2}{g^{(2)}(u(\tau), y(\tau))} \right\}$$

subject to

$$\begin{cases}
\dot{z} = f_y(u, y)z + f_u(u, y)\nu ; & z(0) = 0 \\
g_y(y(\tau))z(\tau) = 0.
\end{cases}$$
The solution of $DF(p_0, \nu, \tau)(\pi_0, \gamma, \sigma) = 0$ is as follows:

- $\pi_0$ initial costate associated with a stationary solution $(\nu, z)$ of $(PQ)$
- $\gamma$ multiplier associated with the punctual constraint $g_y(y(\tau))z(\tau) = 0$, and

$$\sigma = -\frac{g_y^{(1)}(y(\tau))z(\tau)}{g^{(2)}(u(\tau), y(\tau))}.$$

The no-gap second-order sufficient condition implies that $(\nu, z) = 0$ is the only stationary solution of $(PQ)$, and hence, $(\pi_0, \gamma, \sigma) = 0$ implies the shooting algorithm is well-posed.

Similar results when boundary arcs are present and $q \leq 2$ can be derived.
Proof

\[ \mathcal{F} : \begin{pmatrix} p_0 \\ \nu \\ \tau \end{pmatrix} \mapsto \begin{pmatrix} p(T) - \phi_y(y(T)) \\ g(y(\tau)) \\ g^{(1)}(y(\tau)) \end{pmatrix}, \]

\[ \begin{cases} \dot{y} = f(\Upsilon(y, p), y) & \text{on } [0, T]; \quad y(0) = y_0 \\ -\dot{p} = H_y(\Upsilon(y, p), y, p) & \text{on } [0, \tau) \cup (\tau, T]; \quad p(0) = p_0 \\ [p(\tau)] = -\nu g_y(y(\tau)). \end{cases} \]

- Differentiate, and obtain:

\[ 0 = g_y^{(1)}(y(\tau))z(\tau) + \sigma g^{(2)}(u(\tau), y(\tau)) \quad \Rightarrow \quad \sigma \]

\[ [\pi(\tau)] = -\nu g_{yy}(y(\tau))z(\tau) - \gamma g_y(y(\tau)) - \nu \sigma g_y^{(1)}(y(\tau)) \]

\[ = -\nu g_{yy}(y(\tau))z(\tau) - \gamma g_y(y(\tau)) \]

\[ + \nu \frac{g_y^{(1)}(y(\tau))z(\tau) g^{(2)}(u(\tau), y(\tau))}{g^{(2)}(u(\tau), y(\tau))} g_y^{(1)}(y(\tau)). \]
We give necessary and sufficient second-order optimality conditions for optimal control problems with a state constraint of arbitrary order $q$.

We compute the curvature term: only essential touch points have a contribution to the curvature term; the contribution of boundary arcs is zero.

Application of this no-gap second-order optimality conditions: well-posedness of the shooting algorithm.