Second-order optimality conditions for state-constrained optimal control problems

J. Frédéric Bonnans

INRIA-Saclay & CMAP, Ecole Polytechnique, France

Joint work with Audrey Hermant, INRIA-Saclay & CMAP

NPA 2008, Beijing, April 7-9

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

INRIA FUTURS

CMAP

Data of the academic example

$$(\mathcal{P}) \qquad \mathrm{Min} \int_0^1 \left(\tfrac{1}{2}u^2(t) + g(t)y(t)\right) \mathrm{d}t$$

$$\text{s.t.} \qquad \dot{y}(t) = u(t), \qquad y(0) = y(1) = 0, \qquad y(t) \geq h$$

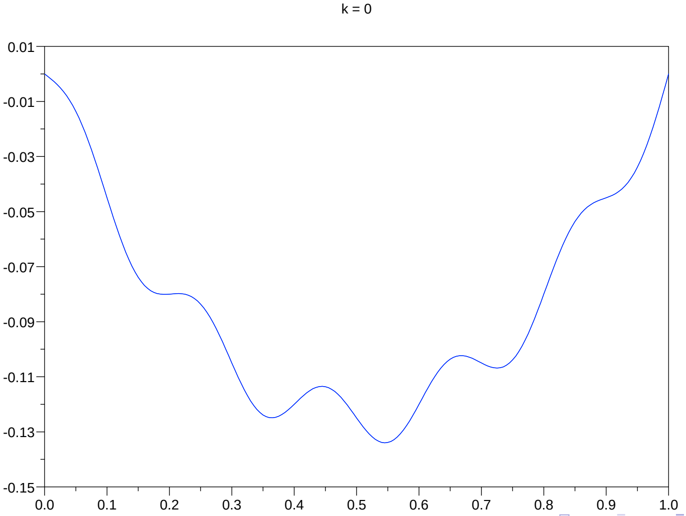
with

$$g(t) := (c - \sin(\alpha t))g_0, \qquad c > 0, \ \ \alpha > 0.$$

Time viewed as second state variable ($\dot{\tau} = 1$)
$\mu = (h - h_0)/(h_1 - h_0)$ homotopy parameter;
$h_0 = \min \bar{y}(t)$, where $\bar{y}$ is the solution of unconstrained problem
$h_1 = h$ target value; numerical values are

$$g_0 := 10, \qquad \alpha = 10\pi, \qquad c = 0.1, \qquad h_1 = -0.001.$$

Unconstrained problem: optimal state



Neigborhood of limiting problem: when $\mu > 0$ is small

For $\mu > 0$ the state constraint is active (convex problem)

The contact set could be then for small $\mu > 0$:

1. One point
2. A small interval
3. A non connected set

Your guess ?

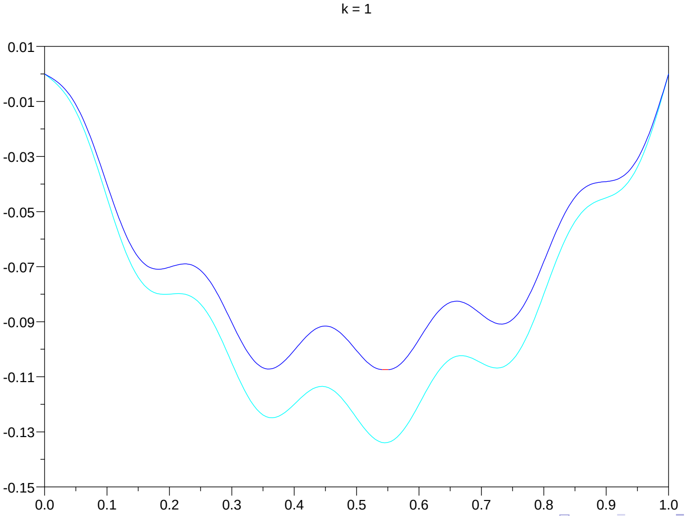
Neigborhood of limiting problem: when $\mu > 0$ is small

For $\mu > 0$ the state constraint is active (convex problem)

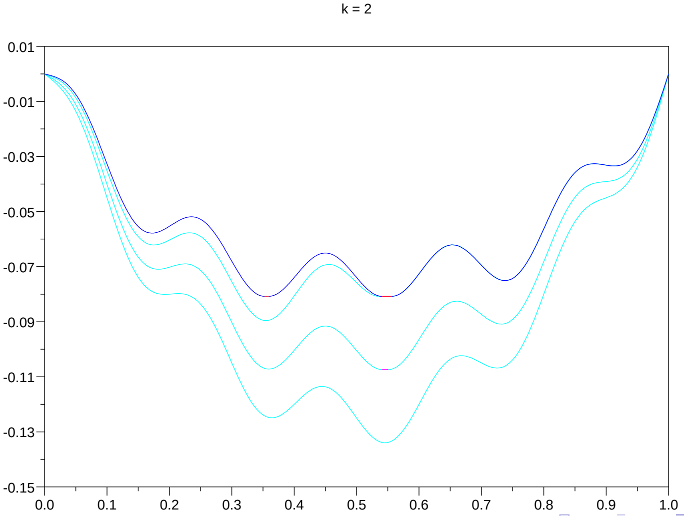
- Structural result: the contact set is an interval
- Quantitative result: first-order expansion of value of extreme points of that interval !

Next: numerical results using a shooting algorithm that will be presented later.

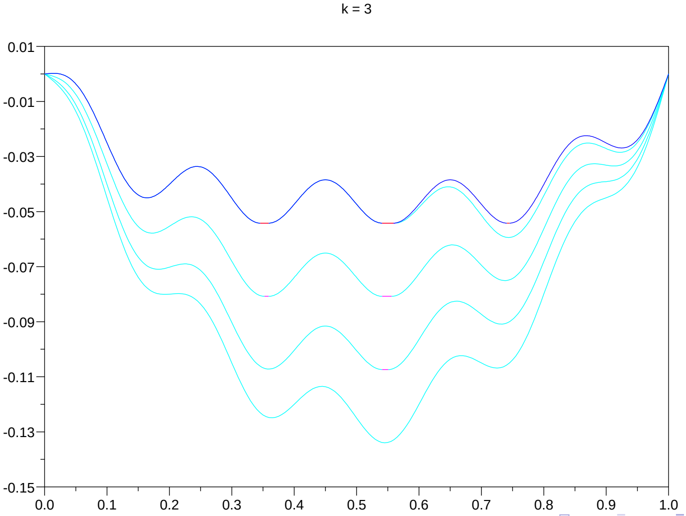
Numerical results II



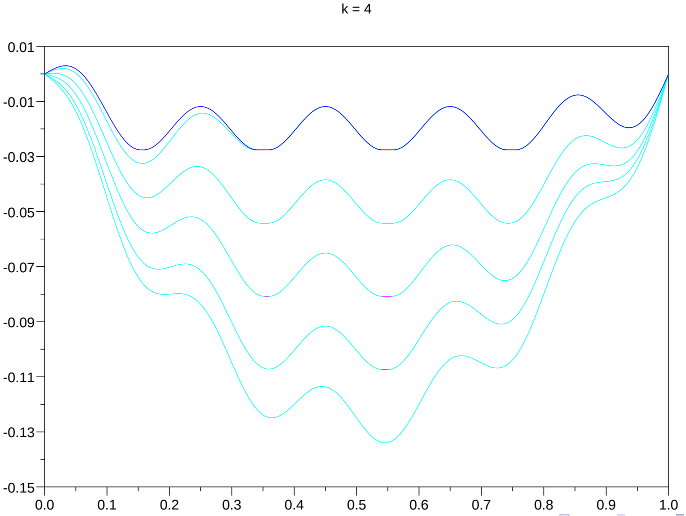
Numerical results III



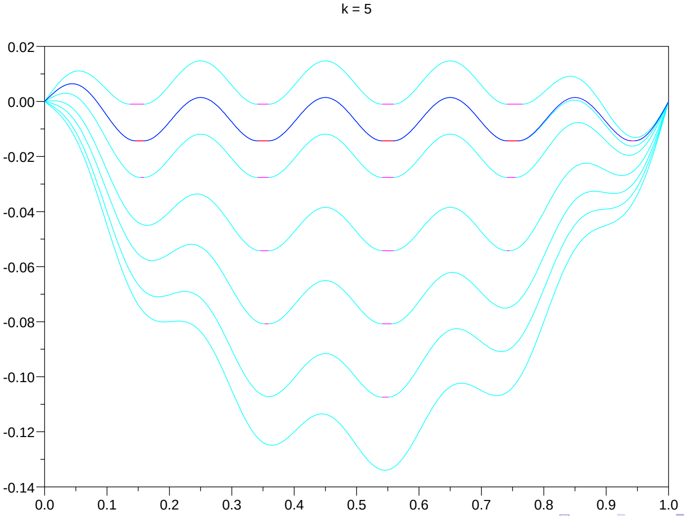
Numerical results IV



Numerical results V



Numerical results VI



Setting: general unconstrained problem

- State equation: $y(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$

$$\dot{y}(t) = f(u(t), y(t)) \qquad \text{p.p. } t \in [0, T], \qquad y(0) = y_0 \quad (1)$$

- Cost function: integral + final term

$$J(u, y) = \int_0^T \ell(u(t), y(t))\mathrm{d}t + \phi(y(T)). \qquad (2)$$

- Optimal control problem

$$\operatorname*{Min}_{(u,y)} J(u, y) \quad \text{s.t. } (1). \qquad (P)$$

- C^∞, Lipschitz data f, ℓ, ϕ.

Functional spaces and costate equation

- Control space: $\mathcal{U} := L^\infty(0, T; \mathbb{R}^m)$
- State and costate space
  $\mathcal{Y} := W^{1,\infty}(0, T; \mathbb{R}^n), \quad \mathcal{P} := W^{1,\infty}(0, T; \mathbb{R}^{n*})$
  where
  $W^{1,\infty}(0, T; \mathbb{R}^n) = \{y \in L^\infty(0, T; \mathbb{R}^n);\ \dot{y} \in L^\infty(0, T; \mathbb{R}^n)\}.$
- Hamiltonian $H(u, y, p) := \ell(u, y) + pf(u, y)$
- Costate equation

$$\begin{aligned} -\dot{p}(t) &= H_y(u(t), y(t), p(t)) \quad \text{p.p. } t \in [0, T], \\ p(T) &= \phi'(y(T)). \end{aligned}$$

Pontryaguin's principle

- $S(P)$ Solution set of (P)
- Pontryaguin's Minimum principle (PMP):
  $H(u(t), y(t), p(t)) = \min_v H(v, y(t), p(t))$ a.a. t
- Weak PMP:
  $H_u(u(t), y(t), p(t)) = 0$ a.a. t

Theorem: If $u \in S(P)$, and y and p are the associated state and costate, then it satisfies the PMP.

Consequence: weak PMP and also:

$$H_{uu} := H_{uu}(u(t), y(t), p(t)) \text{ is semidefinite positive.}$$

Elimination of control

- Assume: **(A1)** Strong Legendre condition (W/SLC)

$$H_{uu}(u(t), y(t), p(t)) \succeq \alpha I_d, \qquad \text{for some } \alpha > 0$$

- By IFT: weak PMP locally equivalent to:

$$u(t) = \Upsilon(y(t), p(t))$$

with Υ of class C^∞

Shooting mapping

- TPBVP Two Point Boundary Value Problem

$$\begin{array}{rcll} \dot{y} & = & f(\Upsilon(y,p),y) & \text{p.p. } [0,T], \quad y(0) = y_0 \\ -\dot{p} & = & H_y(\Upsilon(y,p),y,p) & \text{p.p. } [0,T], \quad p(T) = \phi_y(y(T)). \end{array}$$

- Shooting function $\mathbb{R}^{n*} \mapsto \mathbb{R}^{n*}$: $p_0 \mapsto p(T) - \phi_y(y(T))$, where (y,p) solution of the Cauchy problem

$$\begin{array}{rcll} \dot{y} & = & f(\Upsilon(y,p),y) & \text{p.p. } t \in [0,T], \quad y(0) = y_0 \\ -\dot{p} & = & H_y(\Upsilon(y,p),y,p) & \text{p.p. } [0,T], \quad p(0) = p_0. \end{array}$$

Definition

- Shooting mapping well posed at solution point p_0 if it has an invertible Jacobian at this point. Then:
- By IFT: well-posedness under small perturbation
- Newton's method converges locally quadratically

- Question: Is it satisfied under weak hypotheses ?

Tangent quadratic problem

$$\operatorname*{Min}_{v} J'(u)v + \tfrac{1}{2}J''(u)(v,v) \qquad (TQP)$$

Quadratic Growth condition QGC: for some $\varepsilon > 0$ and $\alpha > 0$

$$J(u+v) \geq J(u) + \alpha\|v-u\|_2^2 \quad \text{if} \ \ \|v-u\|_\infty < \varepsilon.$$

- Second Order Necessary Condition (SONC): $v = 0$ solution of (TQP) (interpretation)
- Second Order Sufficient Condition (SOSC): $v = 0$ unique solution of (TQP) and Strong Legendre Condition (definition)

Theorem: SOSC equivalent to QGC. These conditions imply that the shooting algorithm is well-posed.

State constrained problem: data

- State constraint:

$$g_i(y(t)) \leq 0, \quad t \in [0, T],\ i = 1, \dots, r. \tag{3}$$

- Same cost function: integral + final term

$$J(u, y) = \int_0^T \ell(u(t), y(t))\mathrm{d}t + \phi(y(T)). \tag{4}$$

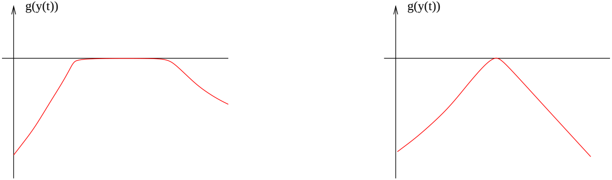
- Optimal control problem

$$\operatorname*{Min}_{(u,y)} J(u, y) \quad \text{s.t. (1) and (3).} \tag{P}$$

- C^∞, Lipschitz data f, ℓ, ϕ, g.

Constraint structure

- Contact set: $\{t \in [0,T]\ ;\ g(y(t)) = 0\}$.



boundary arc $[\tau_{en}, \tau_{ex}]$ (isolated) touch point $\{\tau_{to}\}$

- Question: If known structure: number, ordering of boundary arcs and touch points;
  Then can we design a shooting algorithm ?
  Will it be well-posed ?

Junction points

- Set of junction points: closure of end-points of interior arcs
- Regular junction point: end-point of two arcs, of three types:
- Entry, exit points: end-points of a boundary arc
- Touch point: isolated contact points

Homotopy

- : Constraint structure generally unknown
- Possible approach: start from an unconstrained perturbed problem and compute a path with endpoints the solutions of the perturbed and original problem.
- Example: $g^{\mu}(y) = g(y) - (1-\mu)K$, K "large".
- Motivates the local study of structural changes.
- Work by Oberle, Gergaud, Caillau, Martinon ...

Multipliers are measures

- Lagrange multiplier $\eta \in M(0, T)$
- Lagrangian function

$$L(u, \eta) := J(u) + \int_0^T g(y_u(t))\mathrm{d}\eta(t)$$

Slater qualification condition: $G(u) = g(y_u)$

$$G(u) + G'(u)v < 0 \text{ on } [0, T], \text{ for some } v \in \mathcal{U}.$$

- Complementarity conditions

$$N(u) := \left\{ \eta \in M(0, T)_+; \quad \int_0^T g(y_u(t)) = 0 \right\}.$$

Costate equation

- Costate equation

$$\begin{aligned} -\mathrm{d}p(t) &= H_y(u(t), y(t), p(t))\mathrm{d}t + \mathrm{d}\eta(t) g'(y(t)), \quad \text{p.p. } t \\ p(T) &= \phi'(y(T)). \end{aligned}$$

- Weak Pontryaguin principle (WPMP)
  $H_u(u(t), y(t), p(t)) = 0$ for a.a. t; $\quad \eta \in N(u)$.
- Then: call η a Lagrange multiplier; denote $\eta \in \Lambda(u)$.

Pontryaguin's principle

- $S(P)$ Solution set of (P)
- Minimum principle (PMP):
  $H(u(t), y(t), p(t)) = \underset{v}{\mathrm{Min}}\, H(v, y(t), p(t))$ a.a. t
  for some $\eta \in \Lambda(u)$.

Theorem: Let $u \in S(P)$ be qualified, y associated state. Then
(i) The set $\Lambda(u)$ is non empty and bounded.
(ii) There exists $\eta \in N(u)$ for which the PMP holds.

Consequence: For the (p, η) satisfying the PMP, we have
$H_{uu} := H_{uu}(u(t), y(t), p(t))$ semidefinite positive.

Order of the state constraint

- Total derivative of a scalar state constraint:

$$g^{(1)}(u,y) := g'(y)f(u,y).$$

While result does not depend on u, we can continue:

$$g^{(i+1)}(u,y) := g^{(i)}(y)f(u,y).$$

Constraint order: q smallest number such that

$$g_u^{(q)}(u,y) \neq 0$$

Well-posed constraint order: when

$$g_u^{(q)}(u,y) \neq 0, \quad \text{for all } (u,y)$$

Algebraic variables

- Two algebraic variables
  u, $\dot\eta$ (density, if it exists)
- Algebraic relations: interior arc

$$H_u(u(t), y(t), p(t)) = 0; \quad \dot\eta = 0.$$

- Algebraic relations: boundary arc

$$H_u(u(t), y(t), p(t)) = 0; \quad g^{(q)}(u, y) = 0.$$

Not well-posed in the latter case: $\dot\eta$ does not appear.

First step of the alternative formulation I

- Costate equation

$$\begin{aligned} -\mathrm{d}p(t) &= H_y(u(t), y(t), p(t))\mathrm{d}t + \mathrm{d}\eta(t) g'(y(t)), \quad \text{p.p. } t \\ p(T) &= \phi'(y(T)). \end{aligned}$$

- Write costate dynamics as:

$$-\mathrm{d}(p + \eta g'(y)) = [H_y(u, y, p) - \eta g''(y) f(u, y)]\mathrm{d}t$$

- First alternative costate and multiplier:
$p^1 = p + \eta g'(y); \quad \eta^1 = -\eta$

First step of the alternative formulation II

- The alternative costate p^1 has bounded derivatives. It is solution of the differential equation

$$\begin{aligned} -\dot{p}^1 &= \ell_y(u,y) + pf_y(u,y) + \eta^1 g''(y)f(u,y) \\ &= \ell_y(u,y) + p^1 f_y(u,y) + \eta^1 [g'(y)f_y(u,y) + g''(y)f(u,y)] \end{aligned}$$

- The bracket on r.h.s. is a partial derivative w.r.t. y:

$$\begin{aligned} g^{(1)}(u,y) &= g'(y)f(u,y) \\ g_y^{(1)}(u,y) &= g'(y)f_y(u,y) + g''(y)f(u,y). \end{aligned}$$

- We recognize a Hamiltonian system!

Alternative costate equation

- First alternative Hamiltonian

$$H^1(u,y,p^1,\eta^1) := \ell(u,y) + p^1 f(u,y) + \eta^1 g^{(1)}$$

- Alternative costate equation

$$-\dot{p}^1 = H^1_y(u,y,p^1,\eta^1); \quad p^1(T) = cst + \phi'(y(T)).$$

- Alternative Pontryaguin's principle: since

$$H^1(u,y,p^1,\eta^1) := \ell(u,y) + (p^1 + \eta^1 g'(y)) f(u,y) = H(u,y,p),$$

Weak/strong Pontryaguin's principle is invariant, e. g.:

$$H^1_u(u,y,p^1,\eta^1) = 0$$

First alternative algebraic relations

- Boundary arcs (e.g. when all constraints active): obtain

$$g^{(q)}(u,y)=0; \quad H_u(u,y,p^1)+\eta^1 g_u^{(1)}=0.$$

- Case of scalar control, scalar first-order state constraint: Elimination of algebraic variables holds !

$$u=\Psi^q(y); \quad \eta^1=-H_y(u,y,p^1)/g_u^{(1)}$$

- Unconstrained arcs

$$u=\Psi(y,p^1,\eta^1); \quad \dot{\eta}^1=0.$$

General first-order constraints

- Boundary arcs, all constraints active: obtain

$$H_u(u,y,p^1) + \eta^1 g_u^{(1)}(u,y) = 0; \quad g^{(1)}(u,y) = 0.$$

- Jacobian: $\begin{pmatrix} H^1_{uu} & (g_u^{(1)})^\top \\ g_u^{(1)} & 0 \end{pmatrix}$
- Invertible iff

$$g_u^{(1)}(u,y) \ \text{onto.}; \quad H^1_{uu} = H_{uu} \text{ invertible on } \operatorname{Ker} g_u^{(1)}.$$

- So under weak hypotheses we can eliminate algebraic variables, even in the "vector case"

References for alternative formulation

- Bryson Denham, Dreyfus (1963): provided the idea
- Maurer (1979), unpublished: rigorous derivation
- Several related works by Maurer and Malanowski
- Ref. FB and A. Hermant, INRIA Rep. 6199, 2007 Equivalence with PMP, general vector case

Continuity of control I

- Hyp $H_{uu}(\cdot, y(t), p(t))$ unif. invertible
- u^-, u^+ values just before, after time τ.
- Jump of multiplier at time τ:

$$[p(\tau)] = -\nu g'(y(\tau)); \quad \nu := -[\eta(\tau)]$$

- Is u continuous ? assume H strongly convex w.r.t. u

$$\Delta := H_u(u^-, y, p^+) - H_u(u^-, y, p^-) = -\nu g'(y(\tau)) f_u(u^-, y).$$

- u continuous iff $\nu g'(y(\tau)) f_u(u^-, y) = 0$.
- Holds if all constraints of order > 1. What about order 1 ?

Continuity of control II

- Contribution of first-order terms: take $\ell = 0$

$$\begin{aligned} 0 &= H_u(u^+, y, p^+) - H_u(u^-, y, p^-) \\ &= \int_0^1 [H_{uu}()[u] + [p] f_u()] \mathrm{d}t \end{aligned}$$

- Since $H_{uu}()$ is uniformly positive:

$$\begin{aligned} \alpha |[u]|^2 &\leq \int_0^1 H_{uu}()([u],[u]) \mathrm{d}t \\ &= \nu g'(y) \int_0^1 f_u()[u] \mathrm{d}t = \nu g'(y)[f] \\ &= \nu [g^{(1)}] \leq 0 \end{aligned}$$

therefore u is continuous.

Contribution of mixed state-control constraint

- Mixed state-control constraint

$$c(u, y) \leq 0.$$

- Similar computations give:

$$\begin{aligned} \alpha|[u]|^2 &\leq \int_0^1 H_{uu}()([u],[u])\mathrm{d}t \\ &= \nu g'(y) \int_0^1 f_u()[u]\mathrm{d}t - [\lambda] \int_0^1 c_u(u,y)[u]\mathrm{d}t \\ &= \nu g'(y)[f] - [\lambda][c(u,y)] \\ &= \nu[g^{(1)}] - [\lambda][c(u,y)] \leq 0 \end{aligned}$$

therefore again u is continuous.

Smoothness of control at junction points

- Scalar state constraint of order 1 or 2: u continuous.
- Scalar state constraint of order $q \geq 3$:
  $q - 2$ continuous derivatives ($q - 1$ if q is odd).
- Ref: Jakobson et al., 1971; Maurer, 1979.
- vector case much more involved, see FB and A. Hermant, 2007.
- No example of "generic" regular junction known when $q \geq 3$.

Sensitivity: Framework

$$
\begin{array}{ll}
(P^\mu) & \displaystyle\min_{(u,y)\in\mathcal{U}\times\mathcal{Y}} \int_0^T \ell^\mu(u(t),y(t))\mathrm{d}t + \phi^\mu(y(T)) \\
\text{s.c.} & \dot{y}(t) = f^\mu(u(t),y(t)) \text{ p.p. } [0,T]\ ;\ y(0) = y_0^\mu \\
 & g^\mu(y(t)) \leq 0 \quad \text{on } [0,T].
\end{array}
$$

- Rem. : scalar control $u(t) \in \mathbb{R}$.
- μ : perturbation parameter
- Hyp **(A0)** smooth data: C^∞, Lipschitz **(A1)** $g^{\mu_0}(y_0^{\mu_0}) < 0$.

Hypotheses I

$(\bar{u}, \bar{y})$ solution for $\mu = \mu_0$, with multipliers $(\bar{p}, \bar{\eta})$.

(A2) $H^{\mu_0}(\cdot, \bar{y}(t), \bar{p}(t))$ uniformly strongly convex

(A3) (Order 1 constraint) for all t:

$$|g_u^{(1)}(u(t), y(t))| \geq \gamma > 0.$$

Hypotheses II

(A4) $(\bar{u}, \bar{y})$ has a finite number of regular junctions.

(A5) Strict complementarity on boundary arcs:

$$\frac{\mathrm{d}\bar{\eta}(t)}{\mathrm{d}t} \geq \beta > 0, \qquad \text{on interior boundary arcs.}$$

(A6) For all touch point (isolated contact point) τ,

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} g(\bar{y}(t))|_{t=\tau} < 0.$$

Notion of quadratic growth condition

We say that the Quadratic Growth Condition (QGC) holds) holds if, for all C^2-perturbation (P^μ) of (P^{μ_0}), there exists a neighborhood (V_u, V_μ) of $(\bar{u}, \mu_0)$, such that for $\mu \in V_\mu$, there exists a unique local solution (u^μ, y^μ) of $(\mathcal{P}^\mu)$ with $u^\mu \in V_u$ satisfying the QGC $\exists c, r > 0$ such that

$$J^\mu(u, y) \geq J^\mu(u^\mu, y^\mu) + c(\|u - u^\mu\|_2^2 + \|y - y^\mu\|_{1,2}^2),$$

$$\forall\, (u, y) \text{ feasible for } (P^\mu),\ \|u - \bar{u}\|_\infty + \|y - \bar{y}\|_{1,\infty} < r.$$

Main result: statement

Theorem

Let $(\bar{u}, \bar{y}) = (u^{\mu_0}, y^{\mu_0})$ local solution of (P^{μ_0}) satisfying (A1)-(A6). The the following statements are equivalent:

(i) *The QGC holds*

(ii) *The following second-order sufficient condition is satisfied: The tangent linear-quadratic problem (defined later) has $v = 0$ as unique solution.*

Under these conditions: local uniqueness of local solutions in $\mathcal{U}$.

Also: Boundary arcs are stable,

Touch point remain so, vanish or become boundary arcs.

Main result (continued)

Theorem (End of statement)

... If (i) or (ii) is satisfied, then $\mu \mapsto (u^\mu, y^\mu, p^\mu, \eta^\mu)$ *is locally Lipschitz in*

$$\mathcal{U} \times \mathcal{Y} \times L^\infty(0,T;\mathbb{R}^{n*}) \times L^\infty(0,T;\mathbb{R})$$

and directionally differentiable in

$$L^r(0,T) \times W^{1,r}(0,T;\mathbb{R}^n) \times L^r(0,T;\mathbb{R}^{n*}) \times L^r(0,T)$$

for all $1 \leq r < \infty$. *The directional derivative in direction* d *is the unique solution of a certain linear quadratic problem* (P_d).

The linear quadratic problem

Space of linearized control and states

$$\mathcal{V} := L^2(0,T) \supset \mathcal{U}; \qquad \mathcal{Z} := H^1(0,T;\mathbb{R}^n) \supset \mathcal{Y}.$$

$d = \mu - \mu_0$: "given" direction of perturbation.

$$(\mathcal{P}_d) \qquad \min_{(v,z)\in\mathcal{V}\times\mathcal{Z}} \quad \tfrac{1}{2}\Big\{\int_0^T D^2_{(u,y,\mu)^2} H^{\mu_0}(\bar{u},\bar{y},\bar{p})(v,z,d)^2 \mathrm{d}t$$

$$+ D^2\phi^{\mu_0}(\bar{y}(T))(z(T),d)^2 + \int_0^T D^2 g^{\mu_0}(\bar{y},\mu_0)(z,d)^2 \mathrm{d}\bar{\eta}(t)\Big\}$$

s.c. $\quad \dot{z}(t) = Df^{\mu_0}(\bar{u},\bar{y})(v,z,d)$ sur $[0,T]$, $z(0) = Dy_0^{\mu_0} d$

$Dg^{\mu_0}(\bar{y})(z,d) = 0$ on boundary arcs of $(\bar{u},\bar{y})$

$Dg^{\mu_0}(\bar{y}(\tau))(z(\tau),d) \leq 0$, $\forall\, \tau$ isolated contact point of $(\bar{u},\bar{y})$.

Algorithmic consequences

- If no isolated touch point: Newton's method well-defined (with the "shooting parameters, see paper)
- Convergent homotopy algorithm taking into account transitions
- Touch point viewd as zero lenght boundary arc
- Backtracking over μ if Newton's method non convergent.

Expression of linearization of entry times

Linearize

$$\hat{g}^{(1)}(\bar{u}(\bar{t}^{en}), \bar{y}(\bar{t}^{en}), \mu_0) = 0$$

Denote by v, z, σ^{en} the directional derivative of control, state, entry point w.r.t. a variation of μ in direction d, then

$$\sigma^{en} = -\frac{D\hat{g}^{(1)}(\bar{u}(\bar{t}^{en}), \bar{y}(\bar{t}^{en}), \mu_0)(v(\bar{t}^{en-}), z(\bar{t}^{en}), d)}{\frac{\mathrm{d}}{\mathrm{d}t} g^{(1)}(\bar{u}, \bar{y})|_{t=\bar{t}^{en-}}}$$

Challenges

What happens when:

- A boundary arc splits into two ?
- Two boundary arcs split into one ?
- second-order derivative at a touch point is zero ?

References

- J.F. B. and A.Hermant, *Stability and Sensitivity Analysis for Optimal Control Problems with a First-order State Constraint*. INRIA report RR-5889. ESAIM:COCV, to appear.
- J.F. B. and A.Hermant, *Second-order Analysis for Optimal Control Problems with Pure State Constraints and Mixed Control-State Constraints*. Annales de l'I.H.P. - Nonlinear Analysis, to appear.

Available on

http://www.cmap.polytechnique.fr/∼bonnans/papers.html

References

J.F.B. and A. Shapiro: *Perturbation analysis of optimization problems*. Springer, 2000.

Chinese edition: Science Press, to appear