Stability and sensitivity analysis for optimal control problems with a first-order state constraint and application to continuation methods

Audrey Hermant

CMAP École Polytechnique and INRIA Futurs, France

23rd IFIP Conference on System Modeling and Optimization
Cracow, Poland, July 25, 2007

Joined work with J. Frédéric Bonnans
Introduction

- Stability/sensitivity analysis of optimal control problems with first-order state constraints: two approaches in the literature
  2. Reduction to finite-dimensional problem (shooting approach) Malanowski-Maurer 98.
- Second approach: stronger assumptions (stability of the structure of solutions). Can we weaken this assumption?
- Application: combined shooting/homotopy method when the structure of the trajectory is unknown.
Outline of the talk

1. Optimal control problem
   ▶ First-order optimality conditions
   ▶ Basic definitions and assumptions
2. Shooting formulation
3. Second-order sufficient condition
4. Stability and sensitivity analysis result
   ▶ Assumptions
   ▶ Shooting re-formulation (with touch points)
   ▶ Result
5. Homotopy method
Optimal control problem

\((\mathcal{P})\) \quad \min_{(u, y) \in \mathcal{U} \times \mathcal{Y}} \int_0^T \ell(u(t), y(t)) \, dt + \phi(y(T))

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

- Control & state spaces: \(\mathcal{U} \times \mathcal{Y} = L^\infty(0, T) \times W^{1, \infty}(0, T; \mathbb{R}^n)\) (scalar control, scalar state constraint).
- The time is a state variable (\(\dot{y}_n(t) = 1, \quad y_n(0) = 0 \Rightarrow y_n(t) = t\)).
- First-order state constraint:

\[
\frac{d}{dt} g(y(t)) = g'(y)f(u, y) =: g^{(1)}(u, y)
\]

depends explicitly on the control variable, i.e. \(g^{(1)}_{u} \neq 0\).
First-order optimality conditions (FOC)

- Hamiltonian $H(u, y, p) = \ell(u, y) + pf(u, y)$.

- Contact set

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

- (Pontryagin’s Minimum Principle) If a feasible trajectory $(u, y)$ is solution of $(P) +$ constraint qualification condition, there exist $(p, d\eta) \in BV([0, T]; \mathbb{R}^{n^*}) \times M[0, T]$ such that

$$-dp = H_y(u, y, p)dt + g'(y)d\eta, \quad p(T) = \phi'(y(T))$$

$$u(t) \in \arg\min_{w \in \mathbb{R}} H(w, y(t), p(t)) \quad \text{a.e.} \ [0, T]$$

$$d\eta \geq 0, \quad \text{supp}(d\eta) \subset I(g(y)).$$
Assumptions

(A1) Data $C^2$, Lipschitz continuous second-order derivatives, dynamics $f$ Lipschitz continuous, $g(y_0) < 0$.

(A2) $u$ is continuous and strengthened Legendre-Clebsch condition

$$
\exists \alpha > 0, \quad H_{uu}(u(t), y(t), p(t)) \geq \alpha, \quad \forall t \in [0, T].
$$

(A3) Regular first-order state constraint

$$
\exists \beta > 0, \quad |g_u^{(1)}(u(t), y(t))| \geq \beta \quad \text{for } t \text{ in the neighborhood of the contact set } I(g(y)).
$$

- (A1) and (A3) $\Rightarrow$ constraint qualification holds and uniqueness of multipliers $(p, d\eta)$.
- (A1)–(A3) $\Rightarrow$ $u$ and $\eta$ Lipschitz continuous on $[0, T]$.
**Structure of the trajectory**

- Number and order of boundary arcs and touch points:

(a) Boundary arc $[\tau_{en}, \tau_{ex}]$

(b) Touch point $\tau_{to}$

- $\tau_{en}$ entry point, $\tau_{ex}$ exit point.

- (A1)–(A3) $\Rightarrow$ touch points $\tau_{to}$ are nonessential: $\eta(\tau_{to}) = 0$.

- Assumption:

(A4) The trajectory has finitely many boundary arcs and touch points, and $g(y(T)) < 0$. 
Shooting formulation

- Alternative multipliers \((p_1, \eta_1)\) uniquely associated with \((p, \eta)\)

\[
\begin{align*}
\eta_1(t) &= 0 \quad \text{on interior arcs} \\
\eta_1(t) &= \int_t^{\tau_{\text{ex}}} \, d\eta(s) \quad \text{on boundary arcs} \\
p_1(t) &= p(t) - g'(y(t))\eta_1(t).
\end{align*}
\]

- Alternative Hamiltonian

\[
\mathcal{H}(u, y, p_1, \eta_1) = H(u, y, p_1) + \eta_1 g^{(1)}(u, y).
\]

- Shooting parameters: \(p_0\) (initial costate), jump parameters of the costate at entry points \(\nu_1\), entry and exit points \(\tau_{\text{en}}, \tau_{\text{ex}}\).

Vector of shooting parameters

\[
\theta = (p_0, \nu_1, \tau_{\text{en}}, \tau_{\text{ex}})^T.
\]
Shooting formulation (continued)

Shooting mapping

\[ F : \begin{pmatrix} p_0 \\ \nu_1 \\ \tau_{en} \\ \tau_{ex} \end{pmatrix} \rightarrow \begin{pmatrix} p_1(T) - \phi'(y(T)) \\ g(y(\tau_{en})) \\ g^{(1)}(u(\tau_{en}^-), y(\tau_{en})) \\ g^{(1)}(u(\tau_{ex}^+), y(\tau_{ex})) \end{pmatrix} \]

where \((u, y, p_1, \eta_1)\) is solution of

\[-\dot{p}_1 = \mathcal{H}_y(u, y, p_1, \eta_1), \quad p_1(0) = p_0 \]
\[ 0 = \mathcal{H}_u(u, y, p_1, \eta_1) \]
\[ 0 = g^{(1)}(u, y) \] on boundary arcs
\[ 0 = \eta_1 \] on interior arcs
\[ [p_1(\tau_{en})] = -\nu_1 g'(y(\tau_{en})) \] at entry points.
Shooting formulation (continued)

- Under assumptions (A2)-(A4), \((u, y)\) solution of the first-order optimality condition of \((\mathcal{P})\), iff

(i) There exists a vector of shooting parameters \(\theta\) such that \(\mathcal{F}(\theta) = 0\) and \((u, y)\) is the trajectory associated with \(\theta\);

(ii) The additional conditions below are satisfied

\[
g(y(t)) \leq 0 \quad \text{on interior arcs} \quad \hat{\eta}_1(t) \leq 0 \quad \text{on boundary arcs.}
\]

- One can check that

\[
\nu_1 = \eta_1(\tau_{en}^+) \geq 0.
\]

- Shooting algorithm: find (using Newton's method) \(\theta\) such that \(\mathcal{F}(\theta) = 0\) and check afterward the additional conditions.
Strong second-order sufficient condition

- Extended critical cone \( \hat{C} \): set of
  \[(v, z) \in L^2(0, T) \times H^1(0, T; \mathbb{R}^n) \] satisfying
  \[
  \dot{z} = f_u(u, y)v + f_y(u, y)z \quad \text{a.e. } [0, T], \quad z(0) = 0 \quad (1)
  
  \[
  g'(y(t))z(t) = 0 \quad \text{on } \text{supp}(d\eta).
  
- Remark: omitted constraint in the critical cone
  \[
  g'(y(t))z(t) \leq 0 \quad \text{on } l(g(y)) \setminus \text{supp}(d\eta).
  
- Quadratic cost
  \[
  Q(v, z) = \int_0^T D^2_{(u,y)(u,y)}H(u, y, p)((v, z), (v, z))\,dt
  
  + \int_0^T g''(y)(z, z)d\eta + \phi''(y(T))(z(T), z(T))
  \]
Strong second-order sufficient condition (continued)

Remark: Equivalent expression of the quadratic cost over the set of \((v, z)\) satisfying the linearized state equation (1)

\[
Q(v, z) = \int_0^T D^2_{(u,y)(u,y)} \mathcal{H}(u, y, p_1, \eta_1)((v, z), (v, z)) dt \\
+ \phi''(y(T))(z(T), z(T)) \\
+ \sum_{\tau_{en}} \nu_1 g''(y(\tau_{en}))(z(\tau_{en}), z(\tau_{en})).
\]

Strong second-order sufficient condition

\[
Q(v, z) > 0, \quad \forall (v, z) \in \hat{C} \setminus \{0\}.
\]
Assumptions

(A5) Uniform strict complementarity on boundary arcs

\[ \exists \gamma > 0, \quad \frac{d\eta}{dt}(t) \geq \gamma \quad \text{on boundary arcs.} \]

- Remark: (A5) implies the following tangentiality conditions at entry/exit points:

\[ \lim_{t \to \tau_{en}, \tau_{ex}^+} \frac{d^2}{dt^2} g(y(t)) < 0. \]

- Under assumptions (A1)–(A5) and if there is no touch points, solutions and multipliers are differentiable by application of the implicit function theorem to the shooting mapping (Malanowski-Maurer 98).
Assumptions

- **With touch points**, the structure of the trajectory is not stable (strict complementarity does not hold at touch points).

- Key assumption for touch points:  
  \[(A6) \text{ for all touch points } \tau_{to},\]
  \[
  \frac{d^2}{dt^2} g(y(t))\big|_{t=\tau_{to}} < 0.
  \]
Perturbed optimal control problem

\[(\mathcal{P}^\mu) \quad \min_{(u,y) \in U \times Y} J^\mu(u, y) := \int_0^T \ell(u(t), y(t), \mu)dt + \phi(y(T), \mu)\]
\[\text{s.t.} \quad \dot{y}(t) = f(u(t), y(t), \mu) \quad \text{a.e.} \ [0, T], \quad y(0) = y_0(\mu),\]
\[g(y(t), \mu) \leq 0, \text{ on } [0, T].\]

▶ Data depend on a parameter $\mu \in M$ ($M$ open subset of a Banach space)

▶ $(\mathcal{P}^\mu)$ is a stable extension of $(\mathcal{P})$, if
  > there exists $\bar{\mu}$ such that $(\mathcal{P}) \equiv (\mathcal{P}\bar{\mu})$
  > data $C^2$ w.r.t. $(u, y, \mu)$, with Lipschitz continuous second-order derivatives, $f(\cdot, \cdot, \mu)$ Lipschitz continuous, uniformly w.r.t. $\mu$.

▶ $(\bar{u}, \bar{y})$ local solution of $(\mathcal{P}\bar{\mu}) \equiv (\mathcal{P})$, satisfying (A1)–(A6), with shooting parameters $\bar{\theta}$. 
Structural stability analysis

Theorem

For all stable extension \((\mathcal{P}^\mu)\) of \((\mathcal{P})\), there exists \(\delta > 0\) such that for all \((u, y)\) solution of the FOC of \((\mathcal{P}^\mu)\) with \(\|\mu - \bar{\mu}\|, \|u - \bar{u}\|_\infty < \delta\), the structure of \((u, y)\) is as follows:

- In the neighborhood of each **boundary arc** of \((\bar{u}, \bar{y})\), \((u, y)\) has a **single boundary arc**;

- In the neighborhood of each **touch point** of \((\bar{u}, \bar{y})\), \((u, y)\) has either a **single boundary arc**, or a **single touch point**, or \(g(y(t), \mu) < 0\).
Touch points are boundary arcs of zero length

- Consider touch points $\tau_{to}$ as boundary arcs of zero length, i.e.

$$\nu_1 = 0, \quad \tau_{en} = \tau_{to} = \tau_{ex}.$$ 

Introduce them in the shooting mapping, and apply the implicit function theorem to $F(\theta, \mu) = 0$.

- Problem: the solution of the perturbed problem may be such that exit points are smaller than entry points.

- Solution: formulate the following inequality-constrained problem: Find $\theta = (p_0, \nu_1, \tau_{en}, \tau_{ex})^T$ such that

$$\begin{align*}
(GE^\mu)
\begin{cases}
p_1(T) - \phi'(y(T), \mu) &= 0 \\
g(y(\tau_{en}), \mu) &\leq 0, \quad \nu_1 \geq 0, \quad \nu_1 g(y(\tau_{en}), \mu) = 0 \\
g^{(1)}(u(\tau_{en}^-), y(\tau_{en}), \mu) &= 0 \\
g^{(1)}(u(\tau_{ex}^+), y(\tau_{ex}), \mu) &= 0.
\end{cases}
\end{align*}$$
Application of Robinson’s strong regularity theory

- If the strong second-order sufficient holds: there exist neighborhoods $V \times W$ of $(\bar{\theta}, \bar{\mu})$ such that for all $\mu \in W$, $(GE^\mu)$ has in $V$ a unique solution $\theta^\mu$, which is Lipschitz continuous w.r.t. $\mu$.
- $\theta^\mu$ is such that $\tau^\mu_{en} \leq \tau^\mu_{ex}$, for all entry and exit points and whenever $\tau^\mu_{en} < \tau^\mu_{ex}$,

$$\dot{\eta}^\mu_1 < 0 \text{ on } (\tau^\mu_{en}, \tau^\mu_{ex}).$$

- The trajectory and multipliers $(u^\mu, y^\mu, p_1^\mu, \eta_1^\mu)$ associated with $\theta^\mu$ satisfy the additional conditions, and hence are solution of the FOC of $(P^\mu)$.
- $(u^\mu, y^\mu)$ satisfies also the strong second-order sufficient condition, and hence is a local solution of $(P^\mu)$. 
Theorem

The following assertions are equivalent:

(i) The strong second-order sufficient condition holds.

(ii) For all stable extension \((\mathcal{P}^\mu)\) of \((\mathcal{P})\), there exists \(\delta > 0\) such that for all \(\|\mu - \bar{\mu}\| < \delta\), there exists a unique point \((u^\mu, y^\mu)\) solution of the FOC of \((\mathcal{P}^\mu)\) with \(\|u^\mu - \bar{u}\|_\infty < \delta\), and this point is a local solution of \((\mathcal{P}^\mu)\) satisfying the uniform quadratic growth condition

\[
\exists c, \rho > 0, \quad J^\mu(u, y) \geq J^\mu(u^\mu, y^\mu) + c\|u - u^\mu\|_2^2,
\]

for all feasible trajectory \((u, y)\) of \((\mathcal{P}^\mu)\) with \(\|u - \bar{u}\|_\infty < \rho\).
Theorem

In addition, if either (i) or (ii) holds, then

- The mapping $\mu \mapsto (u^\mu, y^\mu, p^\mu, \eta^\mu)$ is Lipschitz continuous in $L^\infty \times W^{1,\infty} \times L^\infty \times L^\infty$ and directionally differentiable in $L^r \times W^{1,r} \times L^r \times L^r$, for all $1 \leq r < +\infty$.

- The shooting parameters associated with $(u^\mu, y^\mu)$ are directionally differentiable.

The directional derivatives are obtained as the solutions and multipliers of an inequality-constrained linear quadratic problem.
Directional derivatives

\[
\begin{align*}
\min_{(v,z) \in L^2 \times H^1} & \quad \int_0^T D^2_{(u,y, \mu)}^2 H(u, y, p, \bar{\mu})((v, z, d)^2) dt \\
& + \int_0^T D^2 g(\bar{y}, \bar{\mu})((z, d)^2) d\bar{\eta} + D^2 \phi(\bar{y}(T), \bar{\mu})((z(T), d)^2) \\
\text{subject to} & \quad \dot{z} = Df(\bar{u}, \bar{y}, \bar{\mu})(v, z, d) \quad \text{a.e. } [0, T], \quad z(0) = Dy_0(\bar{\mu})d \\
& \quad Dg(\bar{y}, \bar{\mu})(z, d) = 0 \quad \text{on } \cup [\bar{\tau}_{en}, \bar{\tau}_{ex}] \\
& \quad Dg(\bar{y}(\bar{\tau}_{to}), \bar{\mu})(z(\bar{\tau}_{to}), d) \leq 0 \quad \text{for all touch point } \bar{\tau}_{to}. 
\end{align*}
\]

- Unique optimal solution and multipliers \((v_d, z_d, \pi_d, \zeta_d)\).
- Directional derivatives of \(p_0\) is \(\pi_d(0)\), of \(\nu_1\) is the multipliers associated with (2)-(3), and those of entry/exit points \(\sigma_{\tau,d}\) are given by (with \(\tau^\pm = \bar{\tau}_{en}^- \text{ or } \bar{\tau}_{ex}^+\))

\[
\sigma_{\tau,d} = - \frac{Dg^{(1)}(\bar{u}, \bar{y}, \bar{\mu})(v, z, d)(\tau^\pm)}{\frac{d^2}{dt^2} g(\bar{y}, \bar{\mu})|_{t=\tau^\pm}}.
\]
Remark: control constraints

\[
\min_{u \in \mathcal{U}} \int_0^2 (u(t) - (1 - t)^2)^2 dt, \quad u(t) \geq \varepsilon.
\]

- \(\varepsilon = 0\): unconstrained trajectory \(u(t) = (1 - t)^2\), touch point at \(\tau_{to} = 1\).
- \(\varepsilon > 0\): apparition of a boundary arc
  \[\left[1 - \sqrt{\varepsilon}, 1 + \sqrt{\varepsilon}\right].\]
- Entry and exit points are not differentiable at \(\varepsilon = 0\), shooting mapping not differentiable at touch points.
Homotopy method

- Homotopy on the state constraint, $\mu \in [0, 1]$

$$g(y, \mu) := g(y) - (1 - \mu)K \quad (K > 0 \text{ large enough}).$$

$$(\mathcal{P}) \equiv (\mathcal{P}^1), (\mathcal{P}^0) = \text{unconstrained problem}.$$

- Initialize by $p_0$ (initial costate for the unconstrained problem).

- Increase $\mu$ and solve the problem by the shooting algorithm.

- If the state constraint is violated, add a boundary arc: the new associated shooting parameters are initialized by

$$\nu_1 = 0, \quad \tau_{en} = t_m = \tau_{ex}$$

with $t_m \in \arg\max g(y(\cdot), \mu)$.

- If it happens that for a boundary arc, $\tau_{ex} < \tau_{en}$, delete the corresponding boundary arc.
Remarks on the algorithm

- Homotopy method with a variable dimension of the vector of shooting parameters, adapts automatically to the change of structure of the trajectory when a boundary arc appears or disappear.

- For correctness: assumptions (A2)-(A6) have to remain satisfied from $\mu = 0$ to $\mu = 1$ (and in particular, uniform strict complementarity on boundary arcs), as well as the strong second-order sufficient condition.

- Predictor-corrector method along subintervals of $[0, 1]$ where the structure does not vary.

- Difficulty remains to correctly increase $\mu$ in practice so as to make the Newton algorithm converges.
Application on academic example

\begin{equation*}
\min_{(u,y)} \int_0^1 \left( \frac{u(t)^2}{2} + \gamma(t)y(t) \right) \, dt
\end{equation*}

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

with the coefficient in the cost function

\[ \gamma(t) = \gamma_0(c - \sin(\omega t)), \quad \gamma_0 = 10, \quad c = 0.1, \quad \omega = 10\pi. \]

▶ Remark: the previous results can be extended in presence of final state constraints if a controllability condition is satisfied in addition.
Unconstrained trajectory
Iteration

![Graph showing iteration process with values from -0.15 to 0.01 on the y-axis and values from 0.0 to 1.0 on the x-axis. The graph includes a series of oscillations and transitions through different iterations.]
Solution
Perspectives

- Test the algorithm on realistic problems.

- Extension of the results and homotopy method to higher-order state constraints?