Pointwise Second Order Optimality Conditions in Optimal Control

Hélène Frankowska

CNRS and UNIVERSITÉ PIERRE et MARIE CURIE

In collaboration with D. Hoehener

International Conference on Variational Analysis, Optimisation and Quantitative Finance in Honor of Terry Rockafellar's 80th Birthday

Université de Limoges, May 18 - 22, 2015

Diary of Nice Days

- 1980-1981 PhD student at SISSA
- May 1981 - met Terry Rockafellar at IIASA for one day. (Cellina's idea) First discussion. **Thank you, Terry.**
- May 1984 - A recommendation letter by Terry for a CR position at CNRS. Got this first position. **Thank you, Terry.**
- December 1984 - Terry reported on my Doctorat d'Etat (habilitation) and came for habilitation. **Thank you, Terry.**
- May 1986 - First stay in Terry's home. **Thank you, Terry.**

Outline of the talk

1. **Mayer Optimal Control Problem**
   - Maximum Principle
   - Second Order Tangents and Normals
   - Second Order Maximum Principle
2. **Jacobson's Necessary Condition**
   - Partially Singular Controls
   - Control Constraints Described by Inequalities
3. **Sensitivity Relations**
   - Second Order Jets
   - Propagation of Second Order Jets

Mayer's Optimal Control Problem

Consider the minimization problem

$$\text{Minimize } \varphi(x(1)) \tag{P}$$

over absolutely continuous $x \in W^{1,1}([0,1];\mathbb{R}^n)$ satisfying

$$\begin{cases} \dot{x}(t) = f(t, x(t), u(t)) & \text{a.e. in } [0,1] \\ u(t) \in U(t) & \text{a.e. in } [0,1] \\ x(0) \in K_0 \end{cases}$$

where $\varphi\colon \mathbb{R}^n \to \mathbb{R}$, $f\colon [0,1] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $U\colon [0,1] \rightsquigarrow \mathbb{R}^m$ is a measurable set-valued map with closed nonempty images and $K_0 \subset \mathbb{R}^n$ is closed.
Let $\bar{x}$ be a **minimizer** and $\bar{u}$ be a corresponding **control**.

Standard Assumptions (SA)

- For all $(t, x, u) \in [0,1] \times \mathbb{R}^n \times \mathbb{R}^m$, $f(\cdot, x, u)$ is measurable, $f(t, x, \cdot)$ is continuous and there exists $a_1 > 0$ such that $\sup_{u \in U(t)} |f(t, x, u)| \leq a_1(|x| + 1)$;
- $\forall\ R > 0$, there exists an integrable map $k_R : [0,1] \to \mathbb{R}_+$ such that for a.e. $t \in [0,1]$ and all $u \in U(t)$

$$|f(t, x, u) - f(t, y, u)| \leq k_R(t)\,|x - y|\,,\ \forall x,\ y \in RB;$$

- $\exists\ a_2 > 0,\ \rho > 0,\ f(t, \cdot, u) \in C^2$ on $\bar{x}(t) + \rho B$, $\sup_{u \in U(t)} \|f_x(t, \bar{x}(t), u)\| \leq a_2$ for a.e. $t \in [0,1]$;
- $\exists\ \ell \in L^1([0,1]; \mathbb{R}_+)$, s.t. for a.e. $t \in [0,1]$, $\forall\ u \in U(t)$ $\|f_x(t, x, u) - f_x(t, y, u)\| \leq \ell(t)\,|x - y|$, $\forall\ x, y \in \bar{x}(t) + \rho B$;
- φ is differentiable

First Order Necessary Conditions

The **Hamiltonian** $\mathcal{H}\colon [0,1]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ is defined by

$$\mathcal{H}(t,x,p,u)=\langle p, f(t,x,u)\rangle$$

The maximum principle states that the solution $\bar{p}\in W^{1,1}([0,1];\mathbb{R}^n)$ of the **adjoint system**

$$\begin{cases} -\dot{p}(t)=\mathcal{H}_x(t,\bar{x}(t),p(t),\bar{u}(t)) & \text{a.e.}\\ -p(1)=\nabla\varphi(\bar{x}(1)) \end{cases}$$

satisfies the **transversality** condition

$$\bar{p}(0)\in N^{\flat}_{K_0}(\bar{x}(0))$$

and the **maximality** condition for a.e. $t\in[0,1]$:

$$\max_{u\in U(t)}\langle\bar{p}(t), f(t,\bar{x}(t),u)-f(t,\bar{x}(t),\bar{u}(t))\rangle=0$$

The maximum is attained by $\bar{u}(t)$. (Pontryagin and al.)

Second Order Tangents

Adjacent tangent to $K \subset \mathbb{R}^n$ at $x \in K$ is

$$T_K^\flat(x) := \{u \in \mathbb{R}^n \mid \lim_{h \to 0+} \frac{\mathrm{dist}_K(x+hu)}{h} = 0\}$$

Second order adjacent set to K at x in the direction $u \in \mathbb{R}^n$

$$T_K^{\flat(2)}(x;u) := \{v \in \mathbb{R}^n \mid \lim_{h \to 0+} \frac{\mathrm{dist}_K(x+hu+h^2v)}{h^2} = 0\}$$

$v \in T_K^{\flat(2)}(x;u)$ if and only if $\exists\, \psi : [0,\tau) \to K$ such that

$$\psi(0) = x, \quad \dot\psi(0) = u, \quad \ddot\psi(0) = 2v$$

$\psi(\cdot)$ is Rockafellar's second order inner arc
(goes back to Ben Tal and Zowe 1980.
See also R. Cominetti 1990 and Aubin-HF 1990.)

Second Order Normals

Weak point of this notion -

$$\{(u,v) \mid v \in T_K^{\flat(2)}(x;u)\} \text{ is not closed.}$$

First order normal cone $N_K^\flat(x) := \left[T_K^\flat(x)\right]^-$

We associate with every $q \in N_K^\flat(x)$ the **second order "normals"**

$$N_K^{\flat(2)}(x;q) := \{Q \in \mathbf{S}(\mathrm{n}) \mid \langle q,w\rangle + \frac{1}{2}Qyy \le 0,$$
$$\forall y \in T_K^\flat(x) \cap \{q\}^\perp,\ \forall w \in T_K^{\flat(2)}(x;y)\}$$

where $\mathbf{S}(\mathrm{n})$ denotes the set of symmetric $n \times n$ matrices.

Second Order Normals

Weak point of this notion -

$$\{(u,v) \mid v \in T_K^{\flat(2)}(x;u)\} \text{ is not closed.}$$

First order normal cone $N_K^\flat(x) := \left[T_K^\flat(x)\right]^-$

We associate with every $q \in N_K^\flat(x)$ the **second order "normals"**

$$N_K^{\flat(2)}(x;q) := \{Q \in \mathbf{S}(\mathrm{n}) \mid \langle q, w\rangle + \frac{1}{2}Qyy \le 0,$$
$$\forall y \in T_K^\flat(x) \cap \{q\}^\perp,\ \forall w \in T_K^{\flat(2)}(x;y)\}$$

where $\mathbf{S}(\mathrm{n})$ denotes the set of symmetric $n \times n$ matrices.

Maximizing Controls

The set of **maximizing controls** at $t \in [0,1]$ is defined by

$$\overline{U}(t) := \arg\max_{u \in U(t)} \mathcal{H}(t, \bar{x}(t), \bar{p}(t), u)$$

Then $\bar{u}(t) \in \overline{U}(t)$ a.e.

We abbreviate $[t] := (t, \bar{x}(t), \bar{p}(t), \bar{u}(t))$,

$$\Delta f(t, \bar{x}(t), u) := f(t, \bar{x}(t), u) - f(t, \bar{x}(t), \bar{u}(t))$$

$$\Delta f_x(t, \bar{x}(t), u) := f_x(t, \bar{x}(t), u) - f_x(t, \bar{x}(t), \bar{u}(t))$$

Define

$$\overline{F}(t) := \mathrm{co}\,\{(\Delta f(t, \bar{x}(t), u), \Delta f_x(t, \bar{x}(t), u)) \mid u \in \overline{U}(t)\}$$

Maximizing Controls

The set of **maximizing controls** at $t \in [0,1]$ is defined by

$$\overline{U}(t) := \arg\max_{u \in U(t)} \mathcal{H}(t, \bar{x}(t), \bar{p}(t), u)$$

Then $\bar{u}(t) \in \overline{U}(t)$ a.e.
We abbreviate $[t] := (t, \bar{x}(t), \bar{p}(t), \bar{u}(t))$,

$$\Delta f(t, \bar{x}(t), u) := f(t, \bar{x}(t), u) - f(t, \bar{x}(t), \bar{u}(t))$$

$$\Delta f_x(t, \bar{x}(t), u) := f_x(t, \bar{x}(t), u) - f_x(t, \bar{x}(t), \bar{u}(t))$$

Define

$$\overline{F}(t) := \mathrm{co}\,\{(\Delta f(t, \bar{x}(t), u), \Delta f_x(t, \bar{x}(t), u)) \mid u \in \overline{U}(t)\}$$

Second Order Maximum Principle

Theorem

Assume φ is twice differentiable. Then the matrix solution W of the second order adjoint system

$$\begin{cases} \dot{W}(t) = -\mathcal{H}_{px}[t]W(t) - W(t)\mathcal{H}_{xp}[t] - \mathcal{H}_{xx}[t] \\ W(1) = -\varphi''(\bar{x}(1)) \end{cases}$$

satisfies the second order transversality condition

$$W(0) \in N^{\flat(2)}_{K_0}(\bar{x}(0); \bar{p}(0))$$

and the second order maximality condition

$$\max_{(v,M)\in\overline{F}(t)} \left\langle M^T \bar{p}(t) + W(t)v, v \right\rangle = 0, \quad a.e.\ in\ [0,1]$$

Additional Regularity Assumptions (ARA)

This is similar to the first order maximum principle :

$$\max_{u \in U(t)} \langle \bar{p}(t), \Delta f(t, \bar{x}(t), u) \rangle = 0 \quad \Leftrightarrow \quad \max_{v \in co\, \Delta f(t, \bar{x}(t), U(t))} \langle \bar{p}(t), v \rangle = 0$$

- For a.e. $t \in [0,1]$, $f(t, \cdot, \cdot) \in C^2$ on $(\bar{x}(t) + \rho B) \times (\bar{u}(t) + \rho B)$;
- $\exists\, a_3 > 0$ and integrable $\ell_2 \colon [0,1] \to \mathbb{R}_+$ s.t. for a.e. $t \in [0,1]$, $\|f_u(t, \bar{x}(t), \bar{u}(t))\| \le a_3$, $\forall\, x, y \in \bar{x}(t) + \rho B$, $\forall\, u, v \in \bar{u}(t) + \rho B$

$$\|f'(t,x,u) - f'(t,y,v)\| \le a_3(|x-y| + |u-v|)$$
$$\|f''(t,x,u) - f''(t,y,v)\| \le \ell_2(t)(|x-y| + |u-v|)$$

$f'(t,x,u)$ and $f''(t,x,u)$ denote the Jacobian, resp. Hessian, of $f(t,\cdot,\cdot)$

Define the set of non-singular times $A := \{t \in [0,1] \mid \mathcal{H}_u[t] \neq 0\}$

Additional Regularity Assumptions (ARA)

This is similar to the first order maximum principle :

$$\max_{u \in U(t)} \langle \bar{p}(t), \Delta f(t, \bar{x}(t), u) \rangle = 0 \quad \Leftrightarrow \quad \max_{v \in co\, \Delta f(t, \bar{x}(t), U(t))} \langle \bar{p}(t), v \rangle = 0$$

- For a.e. $t \in [0,1]$, $f(t,\cdot,\cdot) \in C^2$ on $(\bar{x}(t) + \rho B) \times (\bar{u}(t) + \rho B)$;
- $\exists\ a_3 > 0$ and integrable $\ell_2 \colon [0,1] \to \mathbb{R}_+$ s.t. for a.e. $t \in [0,1]$, $\|f_u(t, \bar{x}(t), \bar{u}(t))\| \le a_3$, $\forall\, x, y \in \bar{x}(t) + \rho B$, $\forall\, u, v \in \bar{u}(t) + \rho B$

$$\|f'(t,x,u) - f'(t,y,v)\| \le a_3(|x-y| + |u-v|)$$
$$\|f''(t,x,u) - f''(t,y,v)\| \le \ell_2(t)(|x-y| + |u-v|)$$

$f'(t,x,u)$ and $f''(t,x,u)$ denote the Jacobian, resp. Hessian, of $f(t,\cdot,\cdot)$

Define the set of non-singular times $A := \{t \in [0,1] \mid \mathcal{H}_u[t] \neq 0\}$

Necessary Condition for Partially Singular Controls

Theorem

Assume (SA), (ARA) and let W be the solution of the second order adjoint system. Then for a.e. $t \in [0,1]$ and for every $u \in T^{\flat}_{U(t)}(\bar{u}(t))$ satisfying one of the following conditions

(i) $t \in A$, $\mathcal{H}_u[t]u = 0$ *and* $\mathcal{H}_u[t]v + \frac{1}{2}\mathcal{H}_{uu}[t]uu = 0$ *for some* $v \in T^{\flat(2)}_{U(t)}(\bar{u}(t); u)$

(ii) $t \in [0,1]\backslash A$ *and* $\mathcal{H}_{uu}[t]uu = 0$

the Jacobson inequality holds true

$$f_u[t]^T \left(\mathcal{H}_{ux}[t] + W(t)f_u[t]\right) uu \leq 0.$$

Example: Controls Given by Inequality Constraints

Assume (SA), (ARA) and that

$$U(t) := \bigcap_{j=1}^{s} \left\{ u \in \mathbb{R}^m \;\middle|\; c^j(t,u) \leq 0 \right\}$$

where $c^j : [0,1] \times \mathbb{R}^m \to \mathbb{R}$ are measurable in t, $c^j(t,\cdot) \in C^2$.

$$I(t) := \{ j \mid c^j(t, \bar{u}(t)) = 0 \}$$

Assume $\{\nabla_u c^j(t, \bar{u}(t))\}_{j \in I(t)}$ are linearly independent for a.e. t. Then

$$T^{\flat}_{U(t)}(\bar{u}(t)) = \{ u \in \mathbb{R}^m \mid \langle \nabla_u c_j(t, \bar{u}(t)), u \rangle \leq 0 \;\; \forall\, j \in I(t) \}$$

Second Order Condition via Lagrange Multipliers

Corollary

Let W be as in the second order maximum principle.
Then there exist measurable, uniquely defined (up to a set of measure zero) $\alpha_j : [0,1] \to \mathbb{R}_+$, $j = 1, ..., r$ *such that for a.e.* $t \in [0,1]$

(i) $\alpha_j(t)c^j(t, \bar{u}(t)) = 0$ *for all* $j \in \{1, \ldots, s\}$;

(ii) $\mathcal{H}_u[t] = \sum_{j=1}^{s} \alpha_j(t)\nabla_u c^j(t, \bar{u}(t))$;

(iii) $\max_{u \in U_0(t)} f_u[t]^T \left(\mathcal{H}_{ux}[t] + W(t)f_u[t]\right) uu = 0$,

$$U_0(t) := \{u \in T^\flat_{U(t)}(\bar{u}(t)) \mid \mathcal{H}_u[t]u = 0 \text{ and}$$

$$(\mathcal{H}_{uu}[t] - \sum_{j=1}^{s} \alpha_j(t)c^j_{uu}(t, \bar{u}(t)))uu = 0\}.$$

Second Order Jets

Let $f \colon \mathbb{R}^n \to [-\infty, \infty]$ and $x \in dom(f)$. A pair $(q, Q) \in \mathbb{R}^n \times \mathbf{S}(\mathrm{n})$ is a **superjet** of f at x if for some $\delta > 0$ and for all $y \in x + \delta B$,

$$f(y) \leq f(x) + \langle q, y - x \rangle + \frac{1}{2} Q(y - x)(y - x) + o(|y - x|^2)$$

The set of all superjets of f at x is denoted by $J^{2,+}f(x)$.
Similarly, $(q, Q) \in \mathbb{R}^n \times \mathbf{S}(\mathrm{n})$ is a **subjet** of f at x if the above holds with $\leq$ replaced by $\geq$
The set of all subjets of f at x is denoted by $J^{2,-}f(x)$.

The value function $V \colon [0, 1] \times \mathbb{R}^n \to \mathbb{R}$ is defined by

$$V(t, y) := \inf \left\{ \varphi(z(1)) \;\middle|\; z \in \mathcal{S}_{[t,1]}(y) \right\},$$

where $y \in \mathbb{R}^n$ and $\mathcal{S}_{[t,1]}(y)$ denotes the set of solutions of the control system satisfying $x(t) = y$ and defined on $[t, 1]$.

Backward Propagation of Superjets

Theorem

Assume (SA) and that a process $(\bar{x}, \bar{u})$ *satisfies*

$$V(t_0, \bar{x}(t_0)) = \varphi(\bar{x}(1))$$

Let $\Psi \in \boldsymbol{S}(\mathrm{n})$ *be so that* $(\nabla\varphi(\bar{x}(1)), \Psi) \in J^{2,+}\varphi(\bar{x}(1))$.
Consider solutions $\bar{p}$ *and* W *of first and second order adjoint systems with* $\bar{p}(1) = -\nabla\varphi(\bar{x}(1))$ *and* $W(1) = -\Psi$.
Then W *satisfies the maximality condition a.e. in* $[t_0, 1]$ *and*

$$(-\bar{p}(t), -W(t)) \in J_x^{2,+}V(t, \bar{x}(t)), \ \forall\, t \in [t_0, 1].$$

Forward Propagation of Subjets

Theorem

Assume (SA) and that a process $(\bar{x}, \bar{u})$ *satisfies*

$$V(t_0, \bar{x}(t_0)) = \varphi(\bar{x}(1))$$

Consider the adjoint state $\bar{p}$ *defined on* $[t_0, 1]$ *and assume that for some* $W_0 \in \boldsymbol{S}(\mathrm{n})$ *we have* $(-\bar{p}(t_0), -W_0) \in J_x^{2,-} V(t_0, x_0)$.
Then for the solution W *of the linear matrix equation*

$$\dot{W}(t) = -\mathcal{H}_{px}[t]W(t) - W(t)\mathcal{H}_{xp}[t] - \mathcal{H}_{xx}[t], \quad W(t_0) = W_0$$

the following sensitivity relation holds true:

$$(-\bar{p}(t), -W(t)) \in J_x^{2,-} V(t, \bar{x}(t)), \ \forall\, t \in [t_0, 1].$$

Bibliography

D. H. Jacobson, *A new necessary condition of optimality for singular control problems*, SIAM J. Control, 1969.

P. Cannarsa, H. Frankowska, and T. Scarinci, *Second-order sensitivity relations and regularity of the value function ...*, arXiv.org.

H. Frankowska, D. Hoehener and D. Tonon, *A second-order maximum principle in optimal control under state constraints*, Serdica Mathematical Journal, 2013.

D. Hoehener, *Variational approach to second-order optimality conditions for control problems with pure state constraints*, SIAM J. Control, 2012.

H. Frankowska and N. Osmolovskii, *Second-order necessary optimality conditions for the Mayer problem subject to a general control constraint,* Springer INDAM series, 2015.

H. Lou, *Second-order necessary/sufficient optimality conditions for optimal control problems ...*, Discrete Contin. Dyn. Syst. Ser. B, 2010.

Merci pour votre attention

Critical Cone of Tangent Controls

Let $(\bar{x}, \bar{u})$ be a strong local minimizer and $\bar{p}$ be the corresponding adjoint state. Define the **local critical cone** at $\bar{u} \in \mathcal{U}$ by

$$\mathcal{C}_{loc}(\bar{u}) := \{u(\cdot) \in L^1 \mid u(t) \in T^\flat_{U(t)}(\bar{u}(t)) \text{ and } \mathcal{H}_u[t]u(t) = 0 \text{ a.e. } \}$$

and let

$$M^{(2)}(\bar{u}) := \{(u, v) \in L^\infty | u \in \mathcal{C}_{loc}(\bar{u}), v(t) \in T_U^{\flat(2)}(\bar{u}(t); u(t)) \text{ a.e. in } A\}$$

Consider the **linearized system**

$$\begin{cases} \dot{y}(t) = f_x[t]y(t) + f_u[t]u(t) & \text{a.e. in } [0, 1] \\ y(0) = y_0, \end{cases}$$

where $f_x[t] := f_x(t, \bar{x}(t), \bar{u}(t))$ and $f_u[t]$ is defined in a similar way.

Quadratic functional Φ

$$\forall\ (u(\cdot), v(\cdot)) \in L^2([0,1];\mathbb{R}^m) \times L^1([0,1];\mathbb{R}^m),\ y_0,\ w_0 \in \mathbb{R}^n,$$

$$\Phi(u, v, y_0, w_0) := \langle -\bar{p}(0), w_0 \rangle + \frac{1}{2}\varphi''(\bar{x}(1))y(1)y(1) -$$

$$\int_0^1 (\mathcal{H}_u[t]v(t) + \frac{1}{2}\mathcal{H}_{xx}[t]y(t)y(t) + \mathcal{H}_{xu}[t]y(t)u(t) + \frac{1}{2}\mathcal{H}_{uu}[t]u(t)u(t))dt$$

where $y(\cdot)$ is the solution of the linearized system for $y_0,\ u$.

Theorem (D. Hoehener, 2012; HF and N. Osmolovskii 2015 under weaker assumptions)

If (SA), (ARA) hold true, then

$$\Phi(u, v, 0, 0) \geq 0 \quad \textit{for all} \quad (u, v) \in M^{(2)}(\bar{u}).$$

Integral Necessary Conditions

The above theorem does not allow to deduce pointwise conditions except under strong assumption on $U(t)$ because the sets

$$\{(u,v) \mid u \in T^{\flat}_{U(t)}(\bar{u}(t)),\ v \in T^{\flat(2)}_{U}(\bar{u}(t);u)\}$$

are not closed, in general. Define the **second order jets**

$$J^2_{K_0}(\bar{x}(0)) := \Big\{(y,w) \in \mathbb{R}^{2n} \,\Big|\, \forall\, h_i \to 0+,\ \exists\,(y_i,w_i) \to (y,w),\ \langle \bar{p}(0), y_i\rangle = 0,\ \bar{x}(0) + h_i y_i + h_i^2 w_i \in K_0, \Big\}.$$

$$J^2(t) := \{(u,v) \mid \forall\, h_i \to 0+,\ \exists\,(u_i,v_i) \to (u,v),\ \mathcal{H}_u[t]u_i = 0,\ \bar{u}(t) + h_i u_i + h_i^2 v_i \in U(t)\}.$$

Then $t \rightsquigarrow J^2(t)$ is measurable with closed images.

New Integral Necessary Conditions

Admissible variations of $\bar{u}(\cdot)$

$$\mathcal{M}^{(2)}(\bar{u}) := \{(u,v) \in L^2 \times L^1 \mid u \in \mathcal{C}_{loc}(\bar{u}),$$
$$(u(t), v(t)) \in J^2(t) \text{ a.e. in } A\}$$

Then $M^2(\bar{u}) \subset \mathcal{M}^{(2)}(\bar{u})$.

Theorem

If (SA), (ARA) hold true, then $\forall\, (u,v) \in \mathcal{M}^{(2)}(\bar{u})$, $\forall\, (y_0, w_0) \in J^2_{K_0}(\bar{x}(0))$

$$\Phi(u, v, y_0, w_0) \geq 0.$$