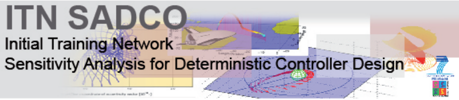
Second-order sensitivity relations for the Mayer problem with differential inclusions

Teresa Scarinci

ITN-Sadco
University of Rome Tor Vergata, Department of Mathematics
Institut de Mathématiques de Jussieu, UPMC

Workshop "New Perspectives in Optimal Control and Games", Rome, Nov. 10-12, 2014

Joint work with: P. Cannarsa & H. Frankowska



Outline

Section 1

Mayer problems for differential inclusions

Mayer problems with differential inclusions

Given

- $\phi : \mathbb{R}^n \to \mathbb{R}$, $\phi \in Lip(\mathbb{R}^n)$
- $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ set-valued map such that

$$\textbf{(SH)} \begin{cases} \bullet F(x) \text{ is nonempty, convex, compact for each } x \in \mathbb{R}^n, \\ \bullet F \text{ is loc. Lipschitz w.r.t. the Hausdorff metric,} \\ \bullet \exists r > 0 \text{ s. t. } \max\{|v| : v \in F(x)\} \leq r(1 + |x|)\, \forall x. \end{cases}$$

- $T \in \mathbb{R}$, $t \leq T$, $x \in \mathbb{R}^n$.

Denote by $y(\cdot; t, x)$ any absolutely continuous arc *(admissible trajectory)* such that

$$\begin{cases} \dot{x}(s) \in F(x(s)), \text{ a.e. } s \in [t, T], \\ x(t) = x. \end{cases} \qquad \text{(DI)}$$

Definition

Mayer problem $\mathcal{P}(t, x)$:

Minimize $\phi(y(T; t, x))$ over all $y(\cdot) \in W^{1,1}(\mathbb{R}^n; \mathbb{R})$ that solves (DI).

Maximum principle (in Hamiltonian form)

Necessary condition for an admissible trajectory to be optimal

$$\min_{y(\cdot;t_0,x_0):\ \dot y\in F(y)} \phi(y(T;t_0,x_0)) = \phi(x(T)).$$

Theorem

Assume (SH), ϕ loc. Lipschitz.
If $x(\cdot)$ is optimal for $\mathcal{P}(t_0,x_0)$, then there exists p (dual arc) a.c. so that

$$\begin{cases} (-\dot p(s), \dot x(s)) \in \partial H(x(s), p(s)),\ a.e.\ \ s \in [t_0, T], \\ -p(T) \in \partial\phi(x(T)). \end{cases}$$

- **Hamiltonian**: $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined as $H(x,p) = \sup_{v\in F(x)}\langle v, p\rangle$.
- $\partial\phi(x) = \overline{co}\,\{\zeta \in \mathbb{R}^n :\ p = \lim_{n\to\infty} \nabla\phi(x_n),\ x_n \to x\}$.
- the adjoint system above encodes the maximum principle

$$H(x(t), p(t)) = \langle p(t), \dot x(t)\rangle \text{ for a.e. } t \in [t_0, T].$$

Semiconcavity results

Further assumptions in Hamiltonian form:

$$(H)\begin{cases} \text{for each non empty, convex and compact subset } K \subseteq \mathbb{R}^n, \\ (i)\ \exists\, c \geq 0 \text{ so that }, \forall p \in S^{n-1}, x \mapsto H(x,p) \text{ is semiconvex on } K \text{ with constant } c, \\ (ii)\ \nabla_p H(x,p) \text{ exists and is Lipschitz continuous in } x \text{ on } K, \text{ uniformly for } p \in S^{n-1}. \end{cases}$$

Value function: $V : (-\infty, T] \times \mathbb{R}^n \to \mathbb{R}$ defined by: for all $(t_0, x_0) \in (-\infty, T] \times \mathbb{R}^n$

$$V(t_0, x_0) = \inf \left\{ \phi(x(T)) : x \in W^{1,1}\left([t_0, T]; \mathbb{R}^n\right),\ \dot{x} \in F(x),\ x(t_0) = x_0 \right\}.$$

Semiconcavity

Let $\Omega \subset \mathbb{R}^n$ be open set. $u : \Omega \to \mathbb{R}$ is *semiconcave* if it is continuous in Ω and $\exists\, c$ s. t.

$$u(x+h) + u(x-h) - 2u(x) \leq c|h|^2,$$

for all $x, h \in \mathbb{R}^n$ such that $[x-h, x+h] \subset \Omega$.

Theorem (Cannarsa-Wolenski, 2011)

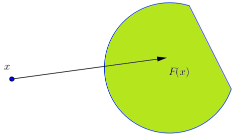
Assume (SH), (H), ϕ semiconcave. Then V is semiconcave on $(-\infty, T] \times \mathbb{R}^n$.

On the assumption (H)

- The semiconvexity of the map $x \mapsto H(x,p)$ is equivalent to the *mid-point property* of the multifunction F on K, that is, ($dist^+_{\mathcal{H}}$: Hausdorff semidistance)

$$dist^+_{\mathcal{H}}\,(2F(x), F(x+z)+F(x-z)) \le c \mid z \mid^2, \quad \forall x, z: \; x, x \pm z \in K.$$

- The existence of $\nabla_p H(x,p)$, $p \neq 0$ is equivalent to the fact that $F_p(x) := argmax\{\langle v, p\rangle, v \in F(x)\} = \{\nabla_p H(x,p)\}$, that is, $H(x,p) = \langle \nabla_p H(x,p), p\rangle$, $p \neq 0$.



We don't have flat parts on the values of F.

- The mid-point property together with the following geometric assumption on F

$$\begin{cases} \text{For every compact } K \subset \mathbb{R}^n, \text{ there exists a constant } c' = c'(K) > 0 \text{ such that for all} \\ x \in K,\; p \in \mathbb{R}^n, \text{ we have: } v_p \in F_p(x) \Rightarrow \langle v - v_p, p\rangle \le -c'|p||v - v_p|,\ \forall v \in F(x), \end{cases}$$

are sufficient to have that $\nabla_p H(x,p)$ exists and is Lipschitz in x on any compact K.

- Splitting Lemma: under assumptions (SH) and (H), we have that

$$\partial H(x,p) = \partial_x^- H(x,p) \times \partial_p^- H(x,p), \quad \forall (x,p) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Weak solutions

HJB equation

$$\begin{cases} -\partial_t V(t,x) + H(x, -V_x(t,x)) = 0 & (-\infty, T) \times \mathbb{R}^n \\ V(T,x) = \phi(x) & x \in \mathbb{R}^n \end{cases} \tag{1}$$

- has no global smooth solution in general
- the value function of the Mayer problem is the unique viscosity solution (semiconcave function that satisfies the equation a.e.)

Heuristically, system of characteristics:

$$\begin{cases} \dot{x} = H_p, & x(T) = z \\ \dot{p} = -H_x, & p(T) = \nabla\phi(z) \end{cases}$$

As long as V is smooth

$$\Rightarrow \quad (H(x,p), -p) = \nabla V$$

Goal: To obtain nonsmooth sensitivity relations

Some known results for the Mayer problem with differential inclusion.

- Clarke-Vinter (1987) $-p(t) \in \partial_x V(t, x(t))$,
- Vinter (1988) $(H(x(t), p(t)), -p(t)) \in \partial V(t, x(t))$.

First-order sensitivity relations

Theorem

Assume

- *(SH), (H),*
- *ϕ be locally Lipschitz.*

Let $x : [t_0, T] \to \mathbb{R}^n$ be optimal for $\mathcal{P}(t_0, x_0)$ and p be such that

$$-p(T) \in \partial^+ \phi(x(T)), \quad \begin{cases} -\dot{p}(t) & \in & \partial_x^- H(x(t), p(t)) \\ \dot{x}(t) & \in & \partial_p^- H(x(t), p(t)) \end{cases} \quad \text{a.e. in } [t_0, T] .$$

Then $(H(x(t), p(t)), -p(t)) \in \partial^+ V(t, x(t))$ for all $t \in [t_0, T]$.

NS conditions for optimality

Assume, moreover, that ϕ is locally semiconcave. Let $x : [t_0, T] \to \mathbb{R}^n$ be an admissible trajectory for $\mathcal{P}(t_0, x_0)$. There exists $p(t) \in \mathbb{R}^n$ such that: a.e. in $[t_0, T]$

$$\begin{aligned} \langle p(t), \dot{x}(t) \rangle &= H(x(t), p(t)) && \textbf{maximum principle} \\ (H(x(t), p(t)), -p(t)) &\in \partial^+ V(t, x(t)), && \textbf{full sensitivity relation} \end{aligned}$$

if and only if x is optimal.

First-order sensitivity relations

Let $f : \Omega \to \mathbb{R}$, $x \in \Omega$ and $p \in \mathbb{R}^n$. We say that p is a proximal supergradient of f at x, $p \in \partial^{+,P} f(x)$, if $\exists\ c,\ \rho \geq 0$ such that

$$f(y) - f(x) - \langle p, y - x \rangle \leq -c|y - x|^2,\ \forall y \in B(x, \rho).$$

Theorem (A partial sensitivity relation)

Assume

- *(SH), (H),*
- *ϕ be locally Lipschitz.*

Let $x : [t_0, T] \to \mathbb{R}^n$ be optimal for $\mathcal{P}(t_0, x_0)$ and p be such that

$$-\overline{p}(T) \in \partial^{+,pr} \phi(\overline{x}(T)), \quad \begin{cases} -\dot{p}(t) & \in & \partial_x^- H(\overline{x}(t), p(t)) \\ \dot{\overline{x}}(t) & \in & \partial_p^- H(\overline{x}(t), p(t)) \end{cases} \quad a.e.\ in\ [t_0, T].$$

Then $-p(t) \in \partial_x^{+,pr} V(t, x(t))$ for all $t \in [t_0, T]$.

Applications

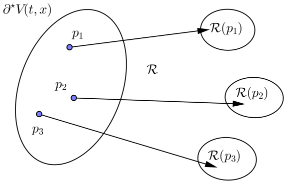
Assume (SH), (H), $\phi \in C^1(\mathbb{R}^n)$. For any $\overline{p} = (\overline{p}_t, \overline{p}_x) \in \partial^* V(t,x)$, let $\mathcal{R}(\overline{p})$ be the set of all trajectories $y(\cdot)$ that are solution of

$$\begin{cases} \dot{y}(s) \in \partial_p H(y(s), p(s)) & \text{a.e. in } [t,T], \quad y(t) = x, \\ -\dot{p}(s) \in \partial_x^- H(y(s), p(s)) & \text{a.e. in } [t,T], \quad p(t) = -\overline{p}_x, \end{cases}$$

and are optimal for $\mathcal{P}(t,x)$. Then the set-valued map $\mathcal{R} : p \in \partial^* V(t,x) \to \mathcal{R}(p)$ has nonempty values. If in addition

- $F(x)$ are not singletons and, if $n > 1$, they have a C^1 boundary,
- $\mathbb{R}^+ p \cap \left(\partial_x^- H(x,p) - \partial_x^- H(x,p) \right) = \emptyset \quad \forall p \neq 0$, (ex: $H(\cdot, p) \in C^1$),

then the set-valued map $\mathcal{R}$ is strongly injective:
$p_1 \neq p_2 \Rightarrow \mathcal{R}(p_1) \cap \mathcal{R}(p_2) = \emptyset$.



Corollary

For every $(t,x) \in (-\infty, T] \times \mathbb{R}^n$,

- there exist at least as many optimal solutions of $\mathcal{P}(t,x)$ as elements of $\partial^* V(t,x)$.
- take ϕ semiconcave. If V fails to be differentiable at a point (t,x), then there exist two or more optimal trajectories starting from (t,x).

$\zeta \in \partial^* f(x)$ if $\nabla f(x_i) \to \zeta, x_i \to x$.

First-order sensitivity relations

Theorem (A partial sensitivity relation)

Assume

- *(SH), (H),*
- *ϕ be locally Lipschitz.*

Let $x : [t_0, T] \to \mathbb{R}^n$ be optimal for $\mathcal{P}(t_0, x_0)$ and p be such that

$$-p(t_0) \in \partial^{-,pr} V(t_0, x_0), \quad \begin{cases} -\dot{p}(t) & \in & \partial_x^- H(x(t), p(t)) \\ \dot{x}(t) & \in & \partial_p^- H(x(t), p(t)) \end{cases} \quad \text{a.e. in } [t_0, T]\,.$$

Then, $-p(t) \in \partial_x^{-,pr} V(t, x(t))$ for all $t \in [t_0, T]$.

Corollary

Same assumptions as before. Let $x(\cdot)$ be optimal for $\mathcal{P}(t_0, x_0)$, and suppose that $\partial^+ \phi(x(T)) \neq \emptyset$. Let $p(\cdot)$ be such that

$$\begin{cases} \dot{x}(s) & \in & \partial_p^- H(x(s), p(s)), & x(T) & = & x(T) \\ -\dot{p}(s) & \in & \partial_x^- H(x(s), p(s)), & -p(T) & \in & \partial^+ \phi(x(T)) \end{cases} \quad \text{for a.e. } s \in [t_0, T]\,. \tag{2}$$

If $\partial_x^- V(t_0, x_0) \neq \emptyset$, then $V(t, \cdot)$ is differentiable at $x(t)$ with $\nabla_x V(t, x(t)) = -p(t)$ for all $t \in [t_0, T]$.

Definition (Crandall, Ishii, Lions)

Let $f : \mathbb{R}^n \to [-\infty, +\infty]$, $x \in dom(f)$ and $(q, Q) \in \mathbb{R}^n \times S(n)$.

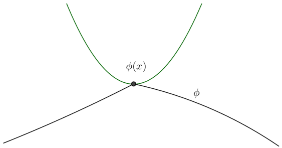
- superjet of f at x, $(q, Q) \in J^{2,+}f(x)$: there exists $\delta > 0$ such that $\forall y \in B(x, \delta)$,

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

- subjet of f at x, $(q, Q) \in J^{2,-}f(x)$: there exists $\delta > 0$ such that $\forall y \in B(x, \delta)$,

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

Equivalently, $J^{2,-}f(x) := -J^{2,+}(-f(x))$.



$(q, Q) \in J^{2,+}f(x)$
if and only if
$\exists \phi \in C^2(\mathbb{R}^n; \mathbb{R})$ such that $f \leq \phi$, $f(x) = \phi(x)$, and $(\nabla\phi(x), \nabla^2\phi(x)) = (q, Q)$

Remarks

Some properties:

- $J^{2,+}f(x)$ is a convex subset of $\mathbb{R}^n \times S(n)$.
- if $f \leq g$ and $f(\widehat{x}) = g(\widehat{x})$ for some $\widehat{x} \in \mathbb{R}^n$, then $J^{2,+}g(\widehat{x}) \subset J^{2,+}f(\widehat{x})$.
- if $(q, Q) \in J^{2,+}f(x)$, then $(q, Q') \in J^{2,+}f(x)$ for all $Q' \in S(n)$ such that $Q' \geq Q$. Thus, the set $J^{2,+}f(x)$ is either empty or unbounded.
- f is semiconcave on a neighborhood of x with semiconcavity constant c, then for any $p \in \partial^+ f(x)$ we have $(p, cI_n) \in J^{2,+}f(x)$.
- if $J^{2,+}f(x)$ and $J^{2,-}f(x)$ are both nonempty, then f is differentiable at x and, for any $(q_1, Q_1) \in J^{2,+}f(x)$ and $(q_2, Q_2) \in J^{2,-}f(x)$, we have that $q_1 = q_2 = \nabla f(x)$ and $Q_1 \geq Q_2$. Moreover, $J^{2,+}f(x) \cap J^{2,-}f(x)$ is nonempty if and only if it is a singleton (q, Q), and in that case

$$f(y) = f(x) + \langle q, y - x \rangle + \frac{1}{2}\langle Q(y - x), y - x \rangle + o(|\ y - x\ |) = 0.$$

First-order sensitivity relations from the point of view of jets

Theorem

Assume

- *(SH), (H),*
- ϕ *be locally Lipschitz.*

Let $x : [t_0, T] \to \mathbb{R}^n$ *be optimal for* $\mathcal{P}(t_0, x_0)$ *and let* $p : [t_0, T] \to \mathbb{R}^n$ *be s. t.:*

- $$-p(T) \in \partial^{+,pr}\phi(x(T)), \quad \begin{cases} -\dot{p}(t) & \in & \partial_x^- H(x(t), p(t)) \\ \dot{x}(t) & \in & \partial_p^- H(x(t), p(t)) \end{cases} \quad \text{a.e. in } [t_0, T].$$

  Then, $\exists c_1 > 0$ *such that* $(-p(t), c_1 I_n) \in J_x^{2,+} V(t, x(t))$ *for all* $t \in [t_0, T)$*.*

- $$-p(t_0) \in \partial^{-,pr} V(t_0, x_0), \quad \begin{cases} -\dot{p}(t) & \in & \partial_x^- H(x(t), p(t)) \\ \dot{x}(t) & \in & \partial_p^- H(x(t), p(t)) \end{cases} \quad \text{a.e. in } [t_0, T].$$

  Then, $\exists c_2 > 0$ *such that* $(-p(t), c_2 I_n) \in J_x^{2,-} V(t, x(t))$ *for all* $t \in [t_0, T)$*.*

Application: propagation of the local regularity of V

Theorem

Assume

- *(SH),*
- $H \in C^2(\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\}))$, $\phi \in C^2(\mathbb{R}^n)$.

Let $(t_0, x_0) \in (-\infty, T] \times \mathbb{R}^n$ *and let* $\overline{x}$ *be optimal for* $\mathcal{P}(t_0, x_0)$ *such that* $\nabla\phi(x(T)) \neq 0$. *If* $\partial_x^{-,pr} V(t_0, x_0) \neq \emptyset$, *then* $V(t, \cdot)$ *is of class* C^2 *in a neighborhood of* $x(t), \forall t \in [t_0, T]$. *Moreover,* $(\nabla_x V(t, x(t)), \nabla^2_{xx} V(t, x(t)) = (-p(t), -R(t)))$, *where* $H_{px}[t] := H_{px}(x(t), p(t))$ *and*

$$\dot{R}(t) + H_{px}[t]R(t) + R(t)H_{xp}[t] + R(t)H_{pp}[t]R(t) + H_{xx}[t] = 0, \quad R(T) = \nabla^2\phi(\overline{x}(T)).$$

Example (if $\nabla\phi(\overline{x}(T)) = 0$?)

- Final cost $\phi : \mathbb{R} \to \mathbb{R}, \ \phi(z) = z^2$.
- Dynamic: $\dot{x} \in [-1, 1] \Longrightarrow H(p) =| p |$.

$$(t, x) \in [0, T] \times \mathbb{R}, \quad V(t, x) = \begin{cases} (x + t - T)^2, & \text{if } x \geq T - t, \\ (x - t + T)^2, & \text{if } x \leq t - T, \\ 0, & \text{otherwise.} \end{cases}$$

If $x = \pm(T - t)$:

- V is not twice differentiable at (t, x),
- $\phi'(y(T; t, x)) = 0$,
- $\partial_x^{-,pr} V(t, x) = \{0\}$

New goal: to obtain new sensitivity relations in form of sub/super jets

Theorem

Assume

- *(SH),*
- $H \in C^2(\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\}))$,
- $\phi : \mathbb{R}^n \to \mathbb{R}$ *locally Lipschitz.*

Let $x : [t_0, T] \to \mathbb{R}^n$ *be optimal for* $\mathcal{P}(t_0, x_0)$ *and let* $(q, Q) \in J^{2,+}\phi(x(T))$ *with* $q \neq 0$.

- *Let* $\overline{p}(\cdot)$ *be such that*

$$-\overline{p}(T) = q, \quad \begin{cases} -\dot{p}(t) & \in & \partial_x^- H(x(t), p(t)) \\ \dot{x}(t) & \in & \partial_p^- H(x(t), p(t)) \end{cases} \quad a.e. \text{ in } [t_0, T],$$

- *let* R *be the solution of the Riccati equation*

$$\begin{cases} \dot{R}(t) + H_{px}[t]R(t) + R(t)H_{xp}[t] + R(t)H_{pp}[t]R(t) + H_{xx}[t] = 0, \\ R(T) = -Q, \end{cases} \tag{3}$$

defined on $[a, T]$, *for some* $a \in [t_0, T)$.

Then $(-p(t), -R(t)) \in J_x^{2,+}V(t, x(t))$ *for all* $t \in [a, T]$.

Propagation of the subjet

Theorem

Assume

- *(SH),*
- $H \in C^2(\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\}))$,
- $\phi : \mathbb{R}^n \to \mathbb{R}$ *locally Lipschitz.*

Let $x : [t_0, T] \to \mathbb{R}^n$ *be optimal for* $\mathcal{P}(t_0, x_0)$ *and let* $(-p(t_0), -R_0) \in J_x^{2,-} V(t_0, x_0)$.

- *Let* $\overline{p}(\cdot)$ *be such that*

$$-\overline{p}(t_0) = p_0, \quad \begin{cases} -\dot{p}(t) & \in & \partial_x^- H(x(t), p(t)) \\ \dot{x}(t) & \in & \partial_p^- H(x(t), p(t)) \end{cases} \quad a.e. \text{ in } [t_0, T],$$

- *let* R *be the solution of the Riccati equation*

$$\begin{cases} \dot{R}(t) + H_{px}[t]R(t) + R(t)H_{xp}[t] + R(t)H_{pp}[t]R(t) + H_{xx}[t] = 0, \\ R(t_0) = R_0, \end{cases}$$

  defined on $[t_0, b]$, *for some* $b \in (t_0, T]$.

Then $(-p(t), -R(t)) \in J_x^{2,-} V(t, x(t))$ *for all* $t \in [t_0, b]$.

Sufficient condition to have no blow-up in finite time

Example: the case where ϕ is locally semiconcave

Set $\bar{t} = \sup\{t \in [t_0, T] : \ R(\cdot) \text{ is defined on } [t_0, t]\}$.

- Since $H_{pp}[t] \geq 0$ on $[t_0, T]$, we have that

$$\dot{R}(t) \leq -H_{px}[t]R(t) - R(t)H_{xp}[t] - H_{xx}[t].$$

  Then, since the solution of the linear equation

$$\dot{\tilde{R}} + H_{px}[t]\tilde{R}(t) + \tilde{R}(t)H_{xp}[t] + H_{xx}[t] = 0, \ \tilde{R}(t_0) = R_0$$

  is well-defined on $[t_0, T]$, a constant $c_1 > 0$ exists such that $R(t) \leq c_1 I$, for any $t \in [t_0, \bar{t})$.
- Since V is locally semiconcave, for some $c_2 > 0$ and all $t \in [t_0, T]$, $(-\bar{p}(t), c_2 I) \in J_x^{2,+} V(t, \cdot)$. Hence $-R(t) \leq c_2 I$ on $[t_0, \bar{t})$.

Therefore, $\|R(t)\| \leq \max\{c_1, c_2\}$ on $[t_0, \bar{t})$. R can be extended to the maximal interval of existence $[t_0, T]$.

Example (Hamiltonian of class $C^2(\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\}))$)

Consider a dynamic $\dot{x} = f(x, u)$ given by the *affine* in control system:

$$f(x) = h(x) + g(x)u, \ u \in U$$

where $U \subset \mathbb{R}^m$ is the closed unit ball, $m \geq n$. Suppose that $h : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^{n \times m}$ are of class C^2, and that the matrix $g(x)$ has full rank for all $x \in \mathbb{R}^n$.

$$H(x, p) = \langle p, h(x) \rangle + \mid g(x)^* p \mid,$$

is of class $C^2(\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\}))$.
Similarly, one can consider strictly convex sets U with sufficiently smooth boundary.

References

- P. Cannarsa, H. Frankowska and T. Scarinci, Sensitivity relations for the Mayer problem with differential inclusions. *ESAIM: COCV* (2014).
- P. Cannarsa, H. Frankowska and T. Scarinci, Second-order sensitivity relations and regularity of the value function for Mayer's problem in optimal control. On Arxiv.