Regularity Properties of the Hamiltonian in Optimal Control

Richard B. Vinter

Imperial College London

(Joint work with Michele Palladino)

'Variational Analysis, Optimization and Quantitative Finance',

In Celebration of Terry Rockafellar's 80th Birthday

18-22 May 2015, Limoges, France

Happy Birthday, Terry!

Outline of the talk

- Regularity properties of the Hamiltonian: why are these useful?
- The Hamiltonian is constant (when data does not depend on time)
- The Hamiltonian is Lipschitz when the data is Lipschitz in time
- These are two examples of a principle: the Hamitonial inherits the regularity of the data w.r.t. time
- A new example: The Hamiltonian has bounded variation if the data has bounded variation w.r.t. time
- Applications and open questions.

The Optimal Control Problem

$$
(P)\begin{cases}
\text{Minimize } g(x(S),x(T)) \\
\text{over } x(.)\in W^{1,1}([S,T],\mathbb{R}^n) \text{ s.t.} \\
\dot{x}(t) \in F(t,x(t)) \text{ a.e.}, \\
h(t,x(t)) \le 0\,, \quad \text{for all } t\in[S,T] \\
(x(S),x(T)) \in C\,,
\end{cases}
$$

($g:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$, $C\subset\mathbb{R}^n\times\mathbb{R}^n$ (closed) and $F(.,.):[S,T]\times\mathbb{R}^n\rightsquigarrow\mathbb{R}^n$.)

Note:

- ‘Differential inclusion’ formulation
- State constraint ‘$h(x(t))\le 0$’

Take a minimizer $\bar{x}(.)$.

Hypotheses

(H1): $F(.,.)$ is closed valued, $F(.,x)$ is $\mathcal{L}$ - measurable for each $x \in \mathbb{R}^n$

(H2): There exist $c > 0$ and $k > 0$ and $\bar{\delta} > 0$ such that

$$F(t,x) \subset F(t,x') + k(|x - x'|)B \quad \text{and} \quad F(t,x) \in c\mathbb{B}\ .$$

for all $x, x' \in \bar{x}(t) + \bar{\delta}B,\ v \in F(t,x)$, a.e. $t \in [S,T]$.

(H3): h(.) is continuously differentiable.

Define Hamiltonian $H(.,.,.) : [S,T] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$

$$H(t,x,p) := \sup_{v \in F(t,x)} p \cdot v$$

The Hamiltonian Inclusion

Theorem (Measurable Time Dependence). Take a minimizer $\bar{x}(.)$. Assume $(H1)$-$(H3)$.

Then there exist $p(.) \in W^{1,1}$, $\mu(.) \in NBV^+(S,T)$ and $\lambda \geq 0$ such that

(i): $\text{supp}\{\mu\} \subset \{t \,|\, h(\bar{x}(t)) = 0\}$

(ii) $(p(.), \lambda, \mu(.)) \neq (0,0,0)$,

(iii) $(-\dot{p}(t), \dot{\bar{x}}(t)) \in \text{co}\, \partial_{x,p} H(t, \bar{x}(t), q(t))$ a.e. ,

(iv) $(q(S), -q(T)) \in \lambda \partial g(\bar{x}(S), \bar{x}(T)) + N_C(\bar{x}(S), \bar{x}(T))$,

where $q(t) = \begin{cases} p(S) & \text{if } t = S \\ p(t) + \int_{[S,t]} \nabla h(\bar{x}(s))\mu(ds) & \text{if } t \in (S,T] \end{cases}$

(Gives Pontryagin Max. Principle when $F(t,x) = f(t,x,U)$.)

Regular Time Dependence: Nec. Conditions

- $F(t,x)$ is indep. of $t \implies H(t,\bar{x}(t),q(t)) =$ const. on open interval (S,T)
- $F(.,x)$ is Lipschitz $\implies t \to H(t,\bar{x}(t),q(t))$ is Lipschitz on open interval (S,T)

Not obvious because

$$H(t,\bar{x}(t),q(t)) = \sup_{v \in F(t,x)} \left(p(t) + \int_{[S,t]} \nabla h(\bar{x}(t))\mu(ds)\right) \cdot v$$

and $\mu(.)$ may have jumps!

Also:

'$F(t,x)$ is convex for each (t,x)'

$\implies$ above relations are true on closed interval $[S,T]$

(Refinement due to Aseev and Arutyunov, '94)

Idea of Proof

By considering transformation of the independent variable

$$\sigma(s) = \int_{[S,t]} (1 + w(s))ds, \; w(s) \in [1 - \epsilon, 1 + \epsilon]$$

show that $(\bar{x}(s), \bar{z}(s) = s)$ is minimizer for problem with dynamics:

$$(\dot{x}(s), \dot{z}(s)) \in \{((1 + w)v, (1 + w)) \,|\, v \in F(z(s), x(s)), -\epsilon \leq w \leq \epsilon\}$$

- Richer class of variations (perturb state trajectories also by 'scaling' time variable) yield extra information:

  $$H(t, \bar{x}(t, q(t)) = r(s) \quad \text{for } t \in (S, T) \tag{1}$$

  for some Lipschitz continuous function $r(.)$ satisfying

  $$\dot{r} \in \partial_t H(t, \bar{x}(t, q(t))$$

- additional analysis to extend to (1) to all $[S, T]$.
- $F(t, x)$ must be Lipschitz continuous in both variables, because time is now a state variable.

Regularity of the Hamiltonian: Open Questions

Recall

- $F(t,x)$ is independent of $t \quad \Longrightarrow \quad H(t,\bar{x}(t),q(t)) = c$
- $F(.,x)$ is Lipschitz $\quad \Longrightarrow \quad t \to H(t,\bar{x}(t),q(t))$ is Lipschitz

Interpretation:

'The Hamiltonian inherits the time-regularity properties of the dynamics'

Does the Hamiltonian inherit other forms of continuity?

'$t \to F(.,x)$ is continuous' $\stackrel{?}{\Longrightarrow}$ 'Hamiltonian is continous?'

We answer related questions . .

Why is Regularity Useful?

- Lagrangian mechanics constancy of Hamiltonian gives invariants of motion.

  $x(.)$ moves in a conservative force field ($F(x) = \nabla\phi(x)$). Motion $\bar{x}(.)$ renders stationary 'the action':

  $$\int \left(\phi(x(t)) - \frac{1}{2}\dot{x}^2(t) \right) dt$$

  Hamiltonian is $\phi(\bar{x}(t)) + \frac{1}{2}\dot{\bar{x}}^2(t)$ (conservation of energy)
- Optimality conditions on singular arcs
- Conditions for regularity of optimal controls
- Existence of non-degenerate multipliers

Functions of Bounded Variation

Classical concept:

$r(.) : [S, T] \to \mathbb{R}$ has bounded variation means

$$\eta(T) \; < \; +\infty$$

in which

$$\eta(t) := \sup_{\mathcal{T}(t)}\{ \sum_{i=0}^{N-1} |r(t_{i+1} - r(t_i)\;|\}$$

(Sup taken over all partitions $\mathcal{T}(t)$ ($\{t_0 = S, \ldots, t_N = t\}$) of $[S, t]$.)

$\eta(t)$ is called the cummulative variation function

- $\eta(.)$ is monotone increasing
- $\eta(.)$ has a countable number of continuity points
- $\eta(.)$ has everywhere left and right limits

Generalization to Multifunctions

Take a multifunction $F : [S, T] \times R^n \rightsquigarrow \mathbb{R}^n$ and an F- trajectory $\bar{x}(.)$.

Several ways to define '$t \to F(t, .)$' has bounded variation

Definition. $t \to F(t, .)$ has bounded variation along $\bar{x}(.)$ if $\eta(T) < +\infty$, where

$$\eta(t) := \sup_{\mathcal{T}} \left\{ \sum_{i=0}^{N-1} \sup \{ d_H(F(t_{i+1}, x), F(t_i, x)) \mid x \in G \} \right\} .$$

Take supremum over partitions $\mathcal{T} = \{t_0 = S, \ldots, t_N = t\}$ of $[S, t]$.

$$G := \{\bar{x}(t) \mid t \in [S, T]\}$$

$\eta(.)$ is called the cummulative variation of $t \to F(t, .)$

Precedents: Moreau's Sweeping Processes

Take a closed, convex multifunction $C(.) : [S, T] \to \mathbb{R}^n$.
Sweeping processes are state trajectories for

$$\begin{cases} -\dot{x}(t) \in N_{C(t)}(x(t)) \\ x_0 = x_0 \end{cases}$$

(Moreau, 1973)

Hypotheses: $\sup_{\mathcal{T}} \left\{ \sum_{i=0}^{N-1} \sup_{v \in C(t_{i+1})} d_{C(t_i)}(v)) \right\} < \infty$.

(Supremum over partitions $\mathcal{T} = \{t_0 = S, \ldots, t_N = t\}$ of $[S, T]$.)

(Early example of use of BV multifunctions)

Properties

Take a multifunction $F(.,.)$; of bounded variation along $\bar{x}(.)$.

Write $\eta(.) =$ cummulative variation function. Then

$$d_H(F(t,x'),F(s,x')) \le \eta(t) - \eta(s)$$

for all $[s,t] \subset [S,T]$ and $x' \in x([S,T])$.

Multifunctions of bounded variation have many 'classical' properties:

(a): Take any $s \in [S,T)$ and $t \in (S,T]$. The one-sided limits

$$F(s^+,x) := \lim_{s' \downarrow s} F(s',x) \quad \text{and} \quad F(t^-,x) := \lim_{t' \uparrow t} F(t',x)$$

exist for every $x \in G$.

(b): There exists a countable set $\mathcal{A}$ such that,

$$\lim_{t' \to t} d_H(F(t',x),F(t,x)) = 0 .$$

for every $t \in (S,T) \backslash \mathcal{A}$ and $x \in G$

Examples of Multifunctions having Bounded Variation

Class of multifunctions $t \to F(.,.)$ with bdd. var. (along some $\bar{x}(.)$) is much larger than the class of Lip. multifunctions $t \to F(t,.)$.

Examples of Mutifunctions having bounded variation include:

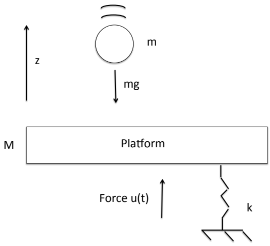
- $F(.,.)$'s with a finite number of fractional singularities, e.g.

$$F(t,x) = \sum_{i=1}^{N} |t - t_i|^{\frac{1}{2}} \tilde{F}_i(x) \quad (\tilde{F}(.) \text{ 'smooth' })$$

- $F(.,.)$'s with a finite number of interior discontinuities
- $F(.,.)$'s with end-time discontinuities

BUT some Hölder $t \to F(t,.)$'s do not have bounded variation.

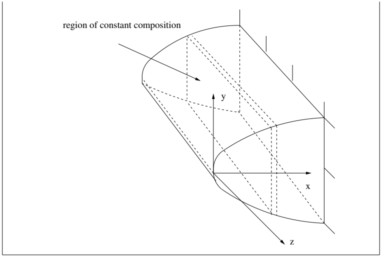
Example (BV Data)



$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -k & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) - \begin{bmatrix} 1 \\ 0 \end{bmatrix} d(t) \,.$$

(Controlled differential equation, in which the control satisfies $|u(t)| \leq K$, and which can be reformulated as a differential inclusion with BV time dependence.)

Example: Optimize bending rigidity in cantelever



- For uniform downward force on leading edge, choose distribution of two materials to maximize bending rigidity R:

$$R = force/displacement$$

- Solve variational problem, in which horizontal displacement x is time-like variable.
- For rounded leading edge, Lagrangian is non-autonomous with a fractional singularity of at 'time' x.

$(L(x, y, u) \sim |x|^{\alpha}, 0 < \alpha < 1.)$

Hamiltonians of Bounded Variation

Theorem (Palladino + V, 2014). Take a minimizer $\bar{x}(.)$. Assume

- $F(.,.)$ is convex valued
- $t \to F(t,.)$ has bounded variation along $\bar{x}(.)$ with cummulative variation $\eta(.)$.

Then the multipliers $(p(.), \mu(.), \lambda)$ can be chosen

to satisfy the following additional condition:

- $|H(t, \bar{x}(t), q(t)) - H(s, \bar{x}(s), q(s))| \leq K \times (\eta(t) - \eta(s))$ for all intervals $[s, t] \subset [S, T]$.

i.e. 'Hamiltonian has inherits BV property from data, and has some cummulative variation (modulo scaling)'.

M Palladino and R B Vinter, 'Regularity of the Hamiltonian along Optimal Trajectories', SIAM J. Control and Opt., to appear.

R. B. Vinter, 'Multifunctions of Bounded Variation', submitted

Idea of Proof

Approximate (P) by Autonomous Multistage Problem on the partition: $\{t_0 = S, \ldots, t_N = T\}$:

$$(P') \begin{cases} \text{Minimize } g(x(T)) \\ \text{over } x(.) : [S,T] \to \mathbb{R}^n \text{ s.t.} \\ \dot{x}(t) \in \sum_{i=0}^{N-1} F(t_i, x(t)) \chi_{[t_i, t_{i+1})}(t) \text{ a.e.} \\ \text{and} \\ h(x(t)) \le 0 \,, \quad \text{for all } t \in [S,T] \\ x(S) = x_0, \; x(T) \in C \,. \end{cases}$$

Strengthened nec. conditions for multiprocess problem give:

$$|H(t_i, \bar{x}(t_{i+1}), q(t_{i+1})) - H(t_i, \bar{x}(t_i), q(t_i))| \approx 0$$

$$\Longrightarrow$$

$$|H(t_{i+1}, \bar{x}(t_{i+1}), q(t_{i+1})) - H(t_i, \bar{x}(t_i), q(t_i)) \le K(\eta(t_i) - \eta(t_i)) \ldots$$

(desired link between Hamiltonian and cummulative var. fn.!)

1st Application

Calculus of Variations:

$$(Q)\begin{cases}\text{Minimize } \int_S^T L(t,x(t),\dot{x}(t))dt \\ \text{over } x(.)\in W^{1,1}([0,1];\mathbb{R}^n) \text{ s.t.} \\ x(S)=x_0 \text{ and, } x(T)=x_1\ .\end{cases}$$

(Q) has a minimizer $\bar{x}(.)$ when:

(HE): **(i):** $L(.,x,v)$ is $\mathcal{L}\times\mathcal{B}^{n\times n}$ measurable and $L(t.,.,.)$ is lower semicontinuous for each $t\in[S,T]$.

(ii): $L(t,x,.)$ is convex for each $(t,x)\in\mathbb{R}^n\times\mathbb{R}^n$.

(iii): There exists a convex function $\theta(.):\mathbb{R}^+\to\mathbb{R}^+$ and a number α such that $\lim_{r\uparrow\infty}\theta(r)/r=+\infty$, and

$$L(t,x,v)\geq\theta(|v|)-\alpha|x| \text{ for all } (t,x,v)\in[S,T]\times\mathbb{R}^n\times\mathbb{R}^n\,.$$

Ball Mizel Example - Non Lipschitz Minimizer

$$\begin{cases} \text{Minimize } \int_0^1 \left\{ r\dot{x}^2(t) + (x^3(t) - t^2)^2 \dot{x}^{14}(t) \right\} dt \\ \text{over } x \in W^{1,1}([0,1]; R) \text{ satisfying} \\ x(0) = 0, \quad x(1) = k. \end{cases}$$

Here, $r > 0$ and $k > 0$ are constants, linked by the relationship

$$r \; = \; (2k/3)^{12}(1 - k^3)(13k^3 - 7).$$

$\exists\ \epsilon > 0$ s.t., $\forall\ k \in (1-\epsilon, 1)$, the arc $\bar{x}(t) \; := \; kt^{2/3}$ is unique minimizer.

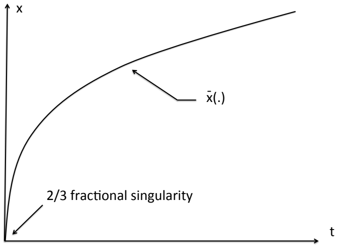


Figure : Non-Lipschitz Minimizer.

Application 1, Continued

(HE) does not guarantee that $\bar{x}(.)$ is Lipschitz. But:

Corollary. Let $\bar{x}(.)$ be a minimizer. Assume that

- (HE)
- $t \to \text{epi}\, L(t,.,.)$ has bounded variation along $(\bar{x}(.), \dot{\bar{x}}(.))$

Then $\bar{x}(.)$ is Lipschitz continuous.

Extends earlier theory:

Replace '$F(., x)$ is Lipschitz' by '$F(., x)$ has bounded variation'

Proof Technique: Use Tonelli Regularity Theory + strengthened conditions . .

Application 2. Non-degeneracy of Necessary Conditions

If the data is BV then we know

- $|H(t, \bar{x}(t), q(t)) - H(t, \bar{x}(t), q(t))| \leq K \times (\eta(t) - \eta(s))$

for all intervals $[s, t] \subset [S, T]$ (not just (S, T)).

This is an extension to BV time dependence of Arutyunov's strengthened necessary conditions.

- The strengthened condition can be used to guarantee existence of non-degenerate Lagrange multiplier in some new situations.

Extends earlier theory:

Replace '$F(., x)$ is Lipschitz' by '$F(., x)$ has bounded variation'

Linear L^∞ Distance Estimates

$$\begin{cases} \dot{x}(t) \in F(t, x(t)) \\ x(t) \in A \end{cases}$$

($F(.,.) : [S,T] \times \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ and closed set $A \subset \mathbb{R}^n$)

A Linear Distance Estimate is valid if there exists $K > 0$ with the property:

For any F-trajectory $\hat{x}(.)$ with $x(S) \in A$, then there is a feasible $x(.)$ with $x(0) = \hat{x}(.)$ and

$$||x(.) - \hat{x}(.)||_{L^\infty} \leq K \max_{t \in [S,T]} d_A(\hat{x}(t)).$$

Distance estimates have many uses.

Linear L^∞ Estimates for BV Data

Theorem (Bettiol, Frankowska, Vinter, JDE 2012).

Assume

(i): Standard Lipschitz/boundedness conditions

(ii): 'inward pointing condition':

$$\Big(\liminf_{(t',x') \stackrel{D}{\to} (t,x)} \; \mathrm{co}\, F(t',x') \Big) \cap \mathrm{int}\, T_A(x) \neq \emptyset \, .$$

(iii): $t \to F(t,x)$ is absolutely continuous from the left

Then an L^∞ Linear Estimate is valid.

(iii) can be replaced by 'F(.,x) has bounded variation'.

Concluding Remarks

This talk illustrates:

'The Hamiltonian inherits the regularity of $t \to F(,.)$'

We have seen useful instances of this principle.

$t \to F(t,.)$ is constant	$\Longrightarrow$	Hamiltonian is constant
$t \to F(t,.)$ is Lipschitz	$\Longrightarrow$	Hamiltonian is Lipschitz
$t \to F(t,.)$ has bdd. var.	$\Longrightarrow$	Hamiltonian has bbd. var. (new)

Open Question

$t \to F(t,.)$ is continuous $\overset{?}{\Longrightarrow}$ Hamiltonian is continuous