Fully Convex Control problems with state constraints and impulses

Peter R. Wolenski

Louisiana State University

International Workshop on Nonlinear Analysis and Optimization in Celebration of Piermarco's 60^{th}
Rome, July 3–7, 2017

Outline

1. I. Fully Convex Control **(FCC)** problems
2. II. State constraints and impulses.
3. III. FCC problems of Bounded Variation
4. IV. Goebel's self-dual envelope
5. V. Our approach to impulse Hamilton-Jacobi **(HJ)** theory
6. VI. A simple example

I. Fully Convex Control (FCC) problems

Consider the "calculus of variations" problem

$$\boldsymbol{\mathcal{P}(\tau,\xi)}\textbf{:} \qquad \inf_{x(\cdot)} \left\{ \ell\big(x(0)\big) \;+\; \int_0^\tau L\big(x(t),\dot{x}(t)\big)\,dt \right\} \text{ with } x(\tau)=\xi.$$

where the infimum is taken over absolutely continuous arcs $x(\cdot)$.
Classically, the data is assumed to be smooth. A Fully Convex Control **(FCC)** problem has data $\ell(\cdot)$ and $L(\cdot,\cdot)$ convex, closed, and proper $(\equiv \mathcal{F})$
.

GOAL:

Develop FCC with state constraints/impulses.

I. Fully Convex Control (FCC) problems

Consider the "calculus of variations" problem

$$\boldsymbol{\mathcal{P}(\tau,\xi)}\textbf{:} \qquad \inf_{x(\cdot)} \left\{ \ell\big(x(0)\big) \; + \int_0^\tau L\big(x(t), \dot{x}(t)\big)\, dt \right\} \text{ with } x(\tau) = \xi.$$

where the infimum is taken over absolutely continuous arcs $x(\cdot)$.
Classically, the data is assumed to be smooth. A Fully Convex Control **(FCC)** problem has data $\ell(\cdot)$ and $L(\cdot,\cdot)$ convex, closed, and proper $(\equiv \mathcal{F})$
.

GOAL:

Develop FCC with state constraints/impulses.

Review of existing FCC theory

(A) Existence theory: Our hypotheses imply the existence of an optimal solution $\bar{x}(\cdot)$.

(B) Optimality conditions: The **E**uler-**L**agrange inclusion has the (nonsmooth) convex statement:

$$\exists\, \bar{y}(\cdot) \text{ s.t. } \big(\dot{\bar{y}}(t), \bar{y}(t)\big) \in \partial_{x,v} L\big(\bar{x}(t), \dot{\bar{x}}(t)\big), \qquad \textbf{(E-L)}$$

which is equivalent to the canonical or **H**amiltonian **I**nclusion:

$$\big(-\dot{\bar{y}}(t), \dot{\bar{x}}(t)\big) \in \partial H\big(x(t), y(t)\big). \qquad \textbf{(HI)}$$

The **T**ransversality **C**ondition is

$$\bar{y}(0)) \in \partial \ell\big(\bar{x}(0)\big). \qquad \textbf{(TC)}$$

$$H(x,y) := \sup_{v \in \mathbb{R}^n} \Big\{ \langle y, v \rangle - L(x,v) \Big\}.$$

Review of existing FCC theory

(A) Existence theory: Our hypotheses imply the existence of an optimal solution $\bar{x}(\cdot)$.

(B) Optimality conditions: The **E**uler-**L**agrange inclusion has the (nonsmooth) convex statement:

$$\exists\, \bar{y}(\cdot) \text{ s.t. } \big(\dot{\bar{y}}(t), \bar{y}(t)\big) \in \partial_{x,v} L\big(\bar{x}(t), \dot{\bar{x}}(t)\big), \qquad \textbf{(E-L)}$$

which is equivalent to the canonical or **H**amiltonian **I**nclusion:

$$\big(-\dot{\bar{y}}(t), \dot{\bar{x}}(t)\big) \in \partial H\big(x(t), y(t)\big). \qquad \textbf{(HI)}$$

The **T**ransversality **C**ondition is

$$\bar{y}(0)\big) \in \partial \ell\big(\bar{x}(0)\big). \qquad \textbf{(TC)}$$

$$H(x,y) := \sup_{v \in \mathbb{R}^n} \Big\{ \langle y, v \rangle - L(x,v) \Big\}.$$

(C) Duality: The dual data consists of:

$$\begin{aligned} M(y,w) &:= L^*(w,y) = \sup_{(x,v)\in\mathbb{R}^{2n}} \left\{ \langle w,x\rangle + \langle y,v\rangle - L(x,v) \right\} \\ m(\eta') &:= \ell^*(\eta') = \sup_{\xi'\in\mathbb{R}^n} \left\{ \langle \xi',\eta'\rangle - \ell(\xi') \right\} \end{aligned}$$

and an associated dual problem is

$$\boldsymbol{\mathcal{Q}(\tau,\eta)}: \qquad \inf_{y(\cdot)} \left\{ m\big(y(0)\big) \;+\; \int_0^\tau M\big(y(t),\dot y(t)\big)\,dt \right\} \text{ with } y(\tau)=\eta.$$

The dual Hamilton $\widetilde{H}(\cdot,\cdot)$ does not introduce new data:

$$\widetilde{H}(y,x) := \sup_{w\in\mathbb{R}^n} \left\{ \langle x,w\rangle - M(y,w) \right\} = -H(x,y).$$

Theorem (Rockafellar 1970's)

Suppose $\bar{x}(\cdot)$ *and* $\bar{y}(\cdot)$ *are primal/dual feasible and satisfy the (equivalent)* **(TC)**:

$$\bar{y}(0) \in \partial\ell\big(\bar{x}(0)\big) \quad \textit{and} \quad \bar{x}(0) \in \partial m\big(\bar{y}(0)\big).$$

The following are equivalent:

(a) $\bar{x}(\cdot)$ *is optimal in* $\boldsymbol{\mathcal{P}(\tau,\xi)}$ *and* $\bar{y}(\cdot)$ *satisfies the primal* **(E-L)** *inclusion*

$$\big(\dot{\bar{y}}(t), \bar{y}(t)\big) \in \partial_{x,v} L\big(\bar{x}(t), \dot{\bar{x}}(t)\big).$$

(b) $\bar{y}(\cdot)$ *is optimal in* $\boldsymbol{\mathcal{Q}(\tau,\eta)}$ *and* $\bar{x}(\cdot)$ *satisfies the dual* **(E-L)** *inclusion*

$$\big(\dot{\bar{x}}(t), \bar{x}(t)\big) \in \partial_{y,w} M\big(\bar{y}(t), \dot{\bar{y}}(t)\big)$$

(c) $\big(\bar{x}(\cdot), \bar{y}(\cdot)\big)$ *satisfy the primal canonical* **(HI)** *inclusion*

$$\big(-\dot{\bar{y}}(t), \dot{\bar{x}}(t)\big) \in \partial H\big(\bar{x}(t), \bar{y}(t)\big).$$

(d) $\big(\bar{x}(\cdot), \bar{y}(\cdot)\big)$ *satisfy the dual canonical* **(HI)** *inclusion*

$$\big(-\dot{\bar{x}}(t), \dot{\bar{y}}(t)\big) \in \partial \widetilde{H}\big(\bar{y}(t), \bar{x}(t)\big).$$

(D) Hamilton-Jacobi (HJ) theory **(RTR & PW, 2001)**] Recall the primal problem

$$\boldsymbol{\mathcal{P}(\tau,\xi)}\text{:} \qquad \min\left\{\ell\big(x(0)\big) \;+\; \int_0^\tau L\big(x(t),\dot{x}(t)\big)\,dt\right\} \text{ with } x(\tau)=\xi.$$

Let $V(\tau,\xi)$ be the value of the problem $\mathcal{P}(\tau,\xi)$, and if this function is differentiable, then it satisfies the Hamilton-Jacobi equation:

$$\begin{aligned} \frac{\partial}{\partial\tau}V(\tau,\xi) &= H\Big(\xi,\nabla_\xi V(\tau,\xi)\Big) \\ V(0,\xi) &= \ell(\xi). \end{aligned}$$

Under general (FCC) assumptions, the function

$$\xi \mapsto V(\tau,\xi) =: V_\tau(\xi)$$

is convex, and one can expect a more general formulation.

Value function duality

Everything thus far applies equally to the dual data, where $\big(\ell(\cdot), L(\cdot,\cdot)\big)$ are replaced by $\big(m(\cdot), M(\cdot,\cdot)\big)$. We have a value function $W(\tau,\eta)$ that satisfies the same properties as $V(\tau,\xi)$. As before, the convex function $\eta \mapsto W(\tau,\eta)$ is written as $W_\tau(\cdot)$.

Theorem

For $\tau \geq 0$, the value functions $V_\tau(\cdot)$ and $W_\tau(\cdot)$ are dual to each other:

$$W_\tau(\eta) = \sup_{\xi\in\mathbb{R}^n}\Big\{\langle\xi,\eta\rangle - V_\tau(\xi)\Big\} \quad \textit{and} \quad V_\tau(\xi) = \sup_{\eta\in\mathbb{R}^n}\Big\{\langle\xi,\eta\rangle - W_\tau(\eta)\Big\}$$

This implies the subgradients of these maps are related by

$$\eta \in \partial V_\tau(\xi) \iff \xi \in \partial W_\tau(\eta) \iff V_\tau(\xi) + W_\tau(\eta) = \langle\xi,\eta\rangle.$$

Method of characteristics (co-state inclusion is an equality)

Recall that primal/dual pair of feasible arcs $\big(x(\cdot),\, y(\cdot)\big)$ satisfy the Hamilton Inclusion **(HI)** and Transversality Condition **(TC)**

$$
\begin{aligned}
-\dot{y}(t) &\in \partial_x H\big(x(t), y(t)\big) && (1)\\
\dot{x}(t) &\in \partial_y H\big(x(t), y(t)\big) && (2)\\
y(0) &\in \partial \ell\big(x(0)\big)
\end{aligned}
$$

if and only if $x(\cdot)$ solves $\mathcal{P}(\tau, \xi)$ and $y(\cdot)$ solves $\mathcal{Q}(\tau, \eta)$.
Moreover, $t \mapsto H\big(x(t), y(t)\big)$ is constant.

The Hamiltonian flow (or Reachable set-valued map) consists of a 1-parameter family of mappings $\mathcal{R}_\tau(\cdot,\cdot) : \mathbb{R}^{2n} \rightrightarrows \mathbb{R}^{2n}$, $\tau > 0$, so that

$$
\mathcal{R}_\tau(\xi_0, \eta_0) := \Big\{(\xi, \eta) : \exists\ \big(x(\cdot), y(\cdot)\big) \text{satisfying } (1), (2) \text{ with } \big(x(0), y(0)\big) = (\xi_0, \eta_0) \text{ and } \big(x(\tau), y(\tau)\big) = (\xi, \eta)\Big\}
$$

Let the graph of ∂V_τ be defined by

$$\mathrm{gr}\big(\partial V_\tau\big) = \big\{(\xi,\eta)\;\big|\;\eta \in \partial V_\tau(\xi)\big\} \;\subset\; \mathbb{R}^n \times \mathbb{R}^n.$$

The case $\tau = 0$ is when

$$\mathrm{gr}\big(\partial V_0\big) = \mathrm{gr}\big(\partial \ell\big) = \big\{(\xi,\eta)\;\big|\;\eta \in \partial \ell(\xi)\big\}.$$

Theorem

The flow mapping transforms $gr\big(\partial\ell\big)$ *onto* $gr\big(\partial V_\tau\big)$. *That is,*

$$gr\big(\partial V_\tau\big) \;=\; \mathcal{R}_\tau\Big(gr\big(\partial\ell\big)\Big) \quad \textit{for all } \tau \ge 0.$$

Remark

In more general non-FCC problems, only the existence of a co-state inclusion occurs, or that

$$\mathrm{gr}\big(\partial V_\tau\big) \;\bigcap\; \mathcal{R}_\tau\Big(\mathrm{gr}\big(\partial\ell\big)\Big) \ne \emptyset \quad \text{for all } \tau \ge 0.$$

The Hamilton-Jacobi equation

The function $(\tau, \xi) \mapsto V(\tau, \xi)$ is not jointly convex, and so a more general subgradient is required. Suppose $f : \mathbb{R}^m \to \bar{\mathbb{R}}$ and $z \in \mathbb{R}^m$. A vector $\zeta \in \mathbb{R}^m$ belongs to the subgradient set $\partial f(z)$ provided

$$f(z') \geq f(z) + \langle \zeta, z' - z \rangle + \mathrm{o}(|z' - z|).$$

This does not conflict with our notation for subgradients of convex functions, for in that case they are the same.

Theorem (HJ equation)

The subgradients of $V(\cdot, \cdot)$ *have the property*

$$(\sigma, \eta) \in \partial V(\tau, \xi) \iff \sigma = -H(\xi, \eta),$$

and in particular, $V(\cdot, \cdot)$ *satisfies the Hamilton-Jacobi equation*

$$\sigma + H(\xi, \eta) = 0 \quad \textit{for all} \quad (\sigma, \eta) \in \partial V(\tau, \xi) \quad \textit{when } \tau \geq 0.$$

The Hamilton-Jacobi equation

The function $(\tau, \xi) \mapsto V(\tau, \xi)$ is not jointly convex, and so a more general subgradient is required. Suppose $f : \mathbb{R}^m \to \bar{\mathbb{R}}$ and $z \in \mathbb{R}^m$. A vector $\zeta \in \mathbb{R}^m$ belongs to the subgradient set $\partial f(z)$ provided

$$f(z') \geq f(z) + \langle \zeta, z' - z \rangle + \mathrm{o}(|z' - z|).$$

This does not conflict with our notation for subgradients of convex functions, for in that case they are the same.

Theorem (HJ equation)

The subgradients of $V(\cdot, \cdot)$ have the property

$$(\sigma, \eta) \in \partial V(\tau, \xi) \quad \Longleftrightarrow \quad \sigma = -H(\xi, \eta),$$

and in particular, $V(\cdot, \cdot)$ satisfies the Hamilton-Jacobi equation

$$\sigma + H(\xi, \eta) = 0 \quad \textit{for all} \quad (\sigma, \eta) \in \partial V(\tau, \xi) \quad \textit{when } \tau \geq 0.$$

Assumptions

(primal) Lagrangian formulation:

(A1) The running function $L(\cdot,\cdot)$ belongs to $\mathcal{F}(\mathbb{R}^n \times \mathbb{R}^n)$.

(A2) The set $F(x) := \operatorname{dom} L(x,\cdot)$ is not empty and $\exists\, \rho > 0$ satisfying

$$\operatorname{dist}\big(0, F(x)\big) \le \rho\,\big(1 + |x|\big)\ \forall x.$$

(A3) $\exists\, \alpha,\, \beta > 0$ and a coercive function $\theta : [0,\infty) \to \mathbb{R}$ so that

$$L(x,v) \ge \theta\Big(\max\{0, |v| - \alpha|x|\}\Big) - \beta|x|.$$

(primal) Hamiltonian formulation:

(A1) $(x,y) \mapsto H(x,y)$ is concave/convex.

(A2) $\exists\, \delta > 0,\ \gamma > 0$ and a finite concave function $\psi(\cdot)$ with

$$H(x,y) \ge \psi(x) - \big(\gamma|x| + \delta\big)|y| \quad \forall\, x,\, y \in \mathbb{R}^n)$$

(A3) $\exists\, \delta' > 0,\ \gamma' > 0$ and a finite convex function $\phi(\cdot)$ with

$$H(x,y) \le \phi(y) + \big(\gamma'|y| + \delta'\big)|x| \quad \forall\, x,\, y \in \mathbb{R}^n)$$

II. State constraints and impulses.

Suppose $X \subseteq \mathbb{R}^n$ is closed convex, and the state constraint is added to problem $\mathcal{P}$:

$$x(t) \in X$$

General nonlinear theory suggests the adjoint arc may have a "jump" when the optimal arc activates the constraint.

Philosophy of Convex Analysis:

Primal/dual problems should be symmetric and treated equally

Thus the primal problem should admit impulses as well.

II. State constraints and impulses.

Suppose $X \subseteq \mathbb{R}^n$ is closed convex, and the state constraint is added to problem $\mathcal{P}$:

$$x(t) \in X$$

General nonlinear theory suggests the adjoint arc may have a "jump" when the optimal arc activates the constraint.

Philosophy of Convex Analysis:

Primal/dual problems should be symmetric and treated equally

Thus the primal problem should admit impulses as well.

Duality and Recession functions

For $f(\cdot) \in \mathcal{F}(\mathbb{R}^k)$, recall the recession function $f_\infty(\cdot) \in \mathcal{F}(\mathbb{R}^k)$ is given by

$$f_\infty(d) = \sup_{x \in \mathrm{dom}(f)} \Big\{ f(x+d) - f(x) \Big\} = \sup_{e \in \mathrm{cl}\ \mathrm{dom}(f^*)} \langle e, d \rangle$$

and is the support function of $\mathrm{dom}(f^*)$.

Coercivity and no state constraints are dual concepts:

$f(\cdot)$ is coercive (superlinear growth)	$\Leftrightarrow$	$\mathrm{dom}(f_\infty) = \{0\}$	$\Leftrightarrow$	$\mathrm{dom}(f^*) = \mathbb{R}^k$ (no dual state constraints)
$\mathrm{dom}(f) = \mathbb{R}^k$ (no primal state constraints)	$\Leftrightarrow$	$\mathrm{dom}(f^*_\infty) = \{0\}$	$\Leftrightarrow$	$f^*(\cdot)$ is coercive (superlinear growth)

III. FCC problems of Bounded Variation

Rockafellar (1974) proposed the following Bolza-type problem:

$$\mathcal{P}\textbf{:}\quad \inf\Big\{\ell\big(x(0),x(T)\big) \;+\; \int_0^T L\big(x(t),\dot{x}(t)\big)\,dt \;+\; \int_{[0,T]} L_\infty\big(\xi_x(t)\big)\,d\mu(dt)\Big\}.$$

The optimization is over $x(\cdot)\in$ **BV** ($=$ arcs of bounded variation), where

$$dx = \dot{x}(t)\,dt + \xi_x(t)\,d\mu(t)$$

and the *recession function* (independent of $(x,v)\in \operatorname{dom} L(\cdot,\cdot)$) is given by

$$L_\infty(d) := \sup_{\lambda>0}\left\{\frac{L(x,v+\lambda d)-L(x,v)}{\lambda}\right\}.$$

Implicit state constraint: $x(t)\in X := cl\big\{x : \operatorname{dom} L(x,\cdot)\neq\emptyset\big\}$.

The "jump" directions: $\xi_x \in \operatorname{dom} L_\infty(\cdot)$.

III. FCC problems of Bounded Variation

Rockafellar (1974) proposed the following Bolza-type problem:

$$\mathcal{P}\textbf{:}\quad \inf\Big\{\ell\big(x(0),x(T)\big) + \int_0^T L\big(x(t),\dot{x}(t)\big)\,dt + \int_{[0,T]} L_\infty\big(\xi_x(t)\big)\,d\mu(dt)\Big\}.$$

The optimization is over $x(\cdot) \in$ **BV** ($=$ arcs of bounded variation), where

$$dx = \dot{x}(t)\,dt + \xi_x(t)\,d\mu(t)$$

and the *recession function* (independent of $(x,v) \in \mathrm{dom}\, L(\cdot,\cdot)$) is given by

$$L_\infty(d) := \sup_{\lambda>0}\left\{\frac{L(x,v+\lambda d) - L(x,v)}{\lambda}\right\}.$$

Implicit state constraint: $x(t) \in X := cl\big\{x : \mathrm{dom}\, L(x,\cdot) \neq \emptyset\big\}$.
The "jump" directions: $\xi_x \in \mathrm{dom}\, L_\infty(\cdot)$.

Dual problem

The *real* dual problem is given by

$$\mathcal{Q}\textbf{:}\quad \inf\Big\{ m\big(y(0), y(T)\big) + \int_0^T M\big(y(t), \dot{y}(t)\big)\,dt + \int_{[0,T]} M_\infty\big(\xi_y(t)\big)\,d\mu(dt)\Big\}.$$

The optimization is over $y(\cdot) \in$ **BV** with $dy = \dot{y}(t)\,dt + \xi_y(t)\,d\mu(t)$, and where $M(y,w) := L^*(w,y)$, $m(y_0, y_T) := \ell^*(y_0, -y_T)$, and

$$M_\infty(e) := \sup_{\lambda>0}\left\{\frac{M(y, w+\lambda e) - M(y,w)}{\lambda}\right\}.$$

Dual implicit state constraint:

$$y(t) \in Y := cl\big\{y : \mathrm{dom}\, M(y,\cdot) \neq \emptyset\big\}.$$

The dual "jump" directions: $\xi_y \in \mathrm{dom}\, M_\infty(\cdot)$.

Dual problem

The *real* dual problem is given by

$$\mathcal{Q}: \quad \inf\Big\{ m\big(y(0), y(T)\big) + \int_0^T M\big(y(t), \dot{y}(t)\big)\, dt + \int_{[0,T]} M_\infty\big(\xi_y(t)\big)\, d\mu(dt) \Big\}.$$

The optimization is over $y(\cdot) \in$ **BV** with $dy = \dot{y}(t)\, dt + \xi_y(t)\, d\mu(t)$, and where $M(y, w) := L^*(w, y)$, $m(y_0, y_T) := \ell^*(y_0, -y_T)$, and

$$M_\infty(e) := \sup_{\lambda > 0} \left\{ \frac{M(y, w + \lambda e) - M(y, w)}{\lambda} \right\}.$$

Dual implicit state constraint:

$$y(t) \in Y := cl\big\{ y : \operatorname{dom} M(y, \cdot) \neq \emptyset \big\}.$$

The dual "jump" directions: $\xi_y \in \operatorname{dom} M_\infty(\cdot)$.

The Hamiltonian's "saddle" equivalent class

The Hamiltonian $H(\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is given by

$$H(x,y) = \sup_{v\in\mathbb{R}^n}\Big\{\langle y, v\rangle - L(x,v)\Big\} = \inf_{w\in\mathbb{R}^n}\Big\{\langle x, w\rangle + M(y,w)\Big\}.$$

Under assumption **(A1)**, $H(\cdot,\cdot)$ is a concave/convex saddle function. Since it is not necessarily finite-valued, one has to deal with its equivalence class (and all the headaches this brings). Each element in an equivalence class agrees on the effective domain $X \times Y \subseteq \mathbb{R}^n \times \mathbb{R}^n$, where

$$X := \big\{x : \exists y \in \mathbb{R}^n \text{ with } H(x,y) > -\infty\big\}$$

(primal state constraint)

$$Y := \big\{y : \exists x \in \mathbb{R}^n \text{ with } H(x,y) < \infty\big\}$$

(dual state constraint)

$$L_\infty(\xi_x) = \sup_{y\in Y}\langle \xi_x, y\rangle \quad \text{and} \quad M_\infty(\xi_y) = \sup_{x\in X}\langle \xi_y, x\rangle$$

Optimality conditions (Extended **(HI)**)

A pair $(x(\cdot), y(\cdot))$ of **BV** arcs is feasible for $\mathcal{P}/\mathcal{Q}$ imply

$$x(t\pm) \in \mathrm{cl}(X) \quad \text{and} \quad y(t\pm) \in \mathrm{cl}(Y) \quad \text{a.e. } t \in [0, T].$$

Recall $x(\cdot)$ and $y(\cdot)$ have decompositions:

$$\begin{aligned} dx &= \dot{x}(t)\, dt + \xi_x(t) d\mu(t) \\ dy &= \dot{y}(t)\, dt + \xi_y(t) d\mu(t) \end{aligned}$$

They satisfy the **extended** Hamiltonian inclusion **(HI)** provided

$$\begin{aligned} \begin{pmatrix} -\dot{y}(t) \\ \dot{x}(t) \end{pmatrix} &\in \begin{pmatrix} \partial_x H\big(x(t), y(t)\big) \\ \partial_y H\big(x(t), y(t)\big) \end{pmatrix} \quad \text{a.e. } t \in [0, T] \\ \xi_x(t) &\in N_{\mathrm{cl}(Y)}\big(y(t+)\big) \cap N_{\mathrm{cl}(Y)}\big(y(t-)\big) \quad \mu - \text{a.e. } t \in [0, T] \\ \xi_y(t) &\in N_{\mathrm{cl}(X)}\big(x(t+)\big) \cap N_{\mathrm{cl}(X)}\big(x(t-)\big) \quad \mu - \text{a.e. } t \in [0, T] \end{aligned}$$

Theorem (Rockafellar 1976)

A pair $(x(\cdot), y(\cdot))$ of **BV** *arcs satisfy the* **extended** *Hamiltonian equations (plus appropriate transversality conditions) if and only if $x(\cdot)$ solves $\mathcal{P}$ and $y(\cdot)$ solves $\mathcal{Q}$.*

Rockafellar left the subject of FCC with state constraints/impulses at this point, but we did develop an **HJ** theory for FCC under **(A1)-(A3)**.

More specific goal:

Apply existing FCC results in HJ theory to an approximation of state constraint/impulse problems.

Theorem (Rockafellar 1976)

A pair $(x(\cdot), y(\cdot))$ of **BV** *arcs satisfy the* **extended** *Hamiltonian equations (plus appropriate transversality conditions) if and only if $x(\cdot)$ solves $\mathcal{P}$ and $y(\cdot)$ solves $\mathcal{Q}$.*

Rockafellar left the subject of FCC with state constraints/impulses at this point, but we did develop an **HJ** theory for FCC under **(A1)-(A3)**.

More specific goal:

Apply existing FCC results in HJ theory to an approximation of state constraint/impulse problems.

Elimination of state constraints

The Moreau-Yosida envelope (for $\lambda > 0$):

$$\big[e_\lambda L\big](x, v) := \inf_{(x',v')\in\mathbb{R}^{n\times n}} \left\{ L(x', v') + \frac{1}{\lambda}\big\|(x', v') - (x, v)\big\|^2 \right\},$$

The idea is to replace $L(\cdot,\cdot)$ by $\big[e_\lambda L\big](\cdot,\cdot)$ and let $\lambda \downarrow 0$.

Positives:

- Eliminates state constraints.
- Convexity is preserved with C^{1+} data satisfying **(A1)-(A2)**.
- Good approximation: $\big[e_\lambda L\big](\cdot,\cdot) \to L(\cdot,\cdot)$ epigraphically as $\lambda \downarrow 0$.

Big Negatives:

- Recession is the same, $\big[e_\lambda L\big]_\infty(\cdot) = L_\infty(\cdot)$: **(A3)** may not hold, and dual constraints persist.
- Duality is lost, $\big[e_\lambda L\big]^*(\cdot,\cdot) \neq \big[e_\lambda L^*\big](\cdot,\cdot)$: Dual problem is very complicated and existing FCC theory is not readily applicable.

Elimination of state constraints

The Moreau-Yosida envelope (for $\lambda > 0$):

$$\big[e_\lambda L\big](x,v) := \inf_{(x',v')\in\mathbb{R}^{n\times n}} \left\{ L(x',v') + \frac{1}{\lambda}\big\|(x',v')-(x,v)\big\|^2 \right\},$$

The idea is to replace $L(\cdot,\cdot)$ by $\big[e_\lambda L\big](\cdot,\cdot)$ and let $\lambda \downarrow 0$.

Positives:

- Eliminates state constraints.
- Convexity is preserved with C^{1+} data satisfying **(A1)-(A2)**.
- Good approximation: $\big[e_\lambda L\big](\cdot,\cdot) \to L(\cdot,\cdot)$ epigraphically as $\lambda \downarrow 0$.

Big Negatives:

- Recession is the same, $\big[e_\lambda L\big]_\infty(\cdot) = L_\infty(\cdot)$: **(A3)** may not hold, and dual constraints persist.
- Duality is lost, $\big[e_\lambda L\big]^*(\cdot,\cdot) \neq \big[e_\lambda L^*\big](\cdot,\cdot)$: Dual problem is very complicated and existing FCC theory is not readily applicable.

Elimination of state constraints

The Moreau-Yosida envelope (for $\lambda > 0$):

$$\big[e_\lambda L\big](x,v) := \inf_{(x',v')\in\mathbb{R}^{n\times n}} \left\{ L(x',v') + \frac{1}{\lambda}\big\|(x',v')-(x,v)\big\|^2 \right\},$$

The idea is to replace $L(\cdot,\cdot)$ by $\big[e_\lambda L\big](\cdot,\cdot)$ and let $\lambda \downarrow 0$.

Positives:

- Eliminates state constraints.
- Convexity is preserved with C^{1+} data satisfying **(A1)-(A2)**.
- Good approximation: $\big[e_\lambda L\big](\cdot,\cdot) \to L(\cdot,\cdot)$ epigraphically as $\lambda \downarrow 0$.

Big Negatives:

- Recession is the same, $\big[e_\lambda L\big]_\infty(\cdot) = L_\infty(\cdot)$: **(A3)** may not hold, and dual constraints persist.
- Duality is lost, $\big[e_\lambda L\big]^*(\cdot,\cdot) \neq \big[e_\lambda L^*\big](\cdot,\cdot)$: Dual problem is very complicated and existing FCC theory is not readily applicable.

IV. Goebel's self-dual envelope

Raf Goebel modified the Moreau-Yosida convolution by considering

$$\big[s_\lambda L\big](x,v) := (1-\lambda^2)\big[e_\lambda L\big](x,v) + \frac{\lambda}{2}\big\|(x,v)\big\|^2$$

Major Advantages:

- All the **positive qualities** of $\big[e_\lambda L\big](\cdot,\cdot)$ are maintained
- $\big[s_\lambda L\big](\cdot,\cdot)$ satisfies **(A1)-(A3)** (so existing FCC theory applies)
  In particular, applying s_λ eliminates both the state constraint and recession at the same time!
- Conjugation and applying s_λ commute:

$$\big[s_\lambda L\big]^*(\cdot,\cdot) = \big[s_\lambda L^*\big](\cdot,\cdot)$$

This implies duality relationships are maintained in the approximation.

V. Our approach to impulse **(HJ)** theory:

Replace $L(\cdot,\cdot)$ by $L_\lambda(\cdot,\cdot) := \big[s_\lambda L\big](\cdot,\cdot)$ and let $\lambda \downarrow 0$; i.e. Let $\mathcal{P}_\lambda$ be the primal problem with data $L_\lambda(\cdot,\cdot)$. The dual problem is denoted by $\mathcal{Q}_\lambda$, and is the same as problem $\mathcal{Q}$ with data $M_\lambda(\cdot,\cdot) := \big[s_\lambda L\big]^* = s_\lambda\big(L^*\big)$.

The Moreau-Yosida envelope in the concave/convex sense is

$$\begin{aligned}\big[e_\lambda H\big](x,y) &= \inf_{y'}\sup_{x'}\left\{H(x',y') + \frac{1}{2\lambda}\left[\|y'-y\|^2 - \|x'-x\|^2\right]\right\}\\ &= \sup_{x'}\inf_{y'}\left\{H(x',y') + \frac{1}{2\lambda}\left[\|y'-y\|^2 - \|x'-x\|^2\right]\right\}.\end{aligned}$$

The Hamiltonian $H_\lambda(\cdot,\cdot)$ associated with $L_\lambda(\cdot,\cdot)$ turns out to be

$$H_\lambda(x,y) = (1-\lambda^2)\big[e_\lambda H\big](x,y) + \frac{\lambda}{2}\left[\|y\|^2 - \|x\|^2\right],$$

the Goebel envelope applied to $H(\cdot,\cdot)$ in the concave/convex sense.

A research plan:

Develop **HJ** *theory for state constrained/impulse* **FCC** *problems by approximating* $\mathcal{P}/\mathcal{Q}$ *by* $\mathcal{P}_\lambda/\mathcal{Q}_\lambda$ *, apply the known results to the approximate problems, and finally pass to the limit as* $\lambda \downarrow 0$ *to capture the behavior in the original problem.*

We (i.e. Cristopher Hermosilla and I) are currently working on this in its full generality. We published detailed results of the specific example below as a 2016 IEEE conference paper. Another paper on Linear-Quadratic models will appear in a 2017 IFAC proceeding. More substantial journal articles are in preparation.

A research plan:

Develop **HJ** *theory for state constrained/impulse* **FCC** *problems by approximating* $\mathcal{P}/\mathcal{Q}$ *by* $\mathcal{P}_\lambda/\mathcal{Q}_\lambda$ *, apply the known results to the approximate problems, and finally pass to the limit as* $\lambda \downarrow 0$ *to capture the behavior in the original problem.*

We (i.e. Cristopher Hermosilla and I) are currently working on this in its full generality. We published detailed results of the specific example below as a 2016 IEEE conference paper. Another paper on Linear-Quadratic models will appear in a 2017 IFAC proceeding. More substantial journal articles are in preparation.

Convergence theorem

Theorem

Suppose both primal and dual state constraints have nonempty interior. Then the primal/dual solutions $(x_\lambda(\cdot), y_\lambda(\cdot))$ *of* $\mathcal{P}_\lambda/\mathcal{Q}_\lambda$ *converge (in the appropriate sense) to a primal/dual solution* $(x(\cdot), y(\cdot))$ *of* $\mathcal{P}/\mathcal{Q}$ *as* $\lambda \searrow 0$. *Also the optimal values* $V_\lambda(\tau, \xi)$ *converge to the optimal value* $V(\tau, \xi)$.

Some natural open questions:

- ⋆ **Do optimal/co-optimal arcs jump only at the endpoints?**
- ⋆ **Can continuous singularities appear?**
- ⋆ **What is needed to modify the HJ equation?**
- ⋆ **Does the method of characteristics carry over?** **YES!**

Convergence theorem

Theorem

Suppose both primal and dual state constraints have nonempty interior. Then the primal/dual solutions $(x_\lambda(\cdot), y_\lambda(\cdot))$ of $\mathcal{P}_\lambda/\mathcal{Q}_\lambda$ converge (in the appropriate sense) to a primal/dual solution $(x(\cdot), y(\cdot))$ of $\mathcal{P}/\mathcal{Q}$ as $\lambda \searrow 0$. Also the optimal values $V_\lambda(\tau, \xi)$ converge to the optimal value $V(\tau, \xi)$.

Some natural open questions:

- $\star$ **Do optimal/co-optimal arcs jump only at the endpoints?**
- $\star$ **Can continuous singularities appear?**
- $\star$ **What is needed to modify the HJ equation?**
- $\star$ **Does the method of characteristics carry over?** YES!

Convergence theorem

Theorem

Suppose both primal and dual state constraints have nonempty interior. Then the primal/dual solutions $(x_\lambda(\cdot), y_\lambda(\cdot))$ of $\mathcal{P}_\lambda/\mathcal{Q}_\lambda$ converge (in the appropriate sense) to a primal/dual solution $(x(\cdot), y(\cdot))$ of $\mathcal{P}/\mathcal{Q}$ as $\lambda \searrow 0$. Also the optimal values $V_\lambda(\tau, \xi)$ converge to the optimal value $V(\tau, \xi)$.

Some natural open questions:

- **Do optimal/co-optimal arcs jump only at the endpoints?**
- **Can continuous singularities appear?**
- **What is needed to modify the HJ equation?**
- **Does the method of characteristics carry over? YES!**

A simple example

Let

$$L(x, v) = \frac{1}{2}x^2 + \mathcal{I}_{[-1,1]}(x) + |v| = L_0(x) + L_1(v).$$

Then $X = [-1, 1]$ is the implicit state constraint. Nontrivial recession of $L_1(\cdot)$ implies nontrivial dual state constraints: $Y = [-1, 1]$, and in fact

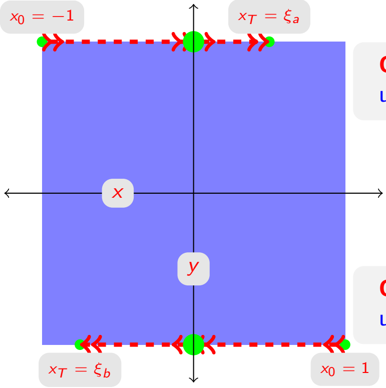
$$M(y, w) = L_1^*(y) + L_0^*(w) = \mathcal{I}_{[-1,1]}(y) + \begin{cases} \frac{1}{2}w^2 & |w| \leq 1, \\ |w| - \frac{1}{2} & |w| > 1. \end{cases}$$

$$H(x, y) = L_1^*(y) - L_0(x) = \mathcal{I}_{[-1,1]}(y) - \frac{1}{2}x^2 - \mathcal{I}_{[-1,1]}(x).$$

$L(x, v) = \frac{1}{2}x^2 + \mathcal{I}_{[-1,1]}(x) + |v|$

Case 1(a): $\ell(x_0, x_T) = \mathcal{I}_{\{-1\}}(x_0) + \mathcal{I}_{\{\xi_a\}}(x_T)$, where $\xi_a \geq 0$:

Case 1(b): $\ell(x_0, x_T) = \mathcal{I}_{\{1\}}(x_0) + \mathcal{I}_{\{\xi_b\}}(x_T)$, where $\xi_b \leq 0$:



Case 1(a): $y \equiv 1$; Jump to 0 and stay there until a jump at time T; **Cost** $= 1 + \xi_a$

Case 1(b): $y \equiv -1$; Jump to 0 and stay there until a jump at time T; **Cost** $= 1 - \xi_b$

The trajectories in Case 1 have the same cost and look the same for any $T > 0$.

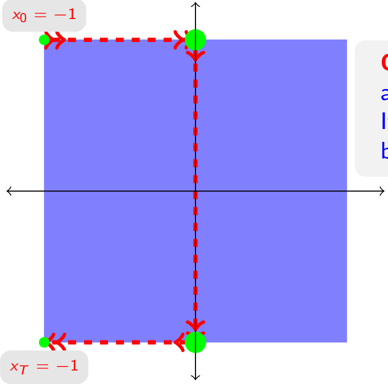
$L(x, v) = \frac{1}{2}x^2 + \mathcal{I}_{[-1,1]}(x) + |v|$

Case 2: $\ell(x_0, x_T) = \mathcal{I}_{\{-1\}}(x_0) + \mathcal{I}_{\{\xi\}}(x_T)$, where $\xi < 0$.

For definiteness, let $\xi = -1$:

The nature of the trajectories depend on $T > 0$.

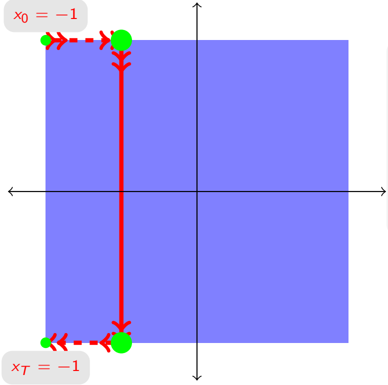
Case 2(a): $T > 4$; **Cost** $= 2$



Case 2(a): $y(\cdot)$initially jumps from 1 to -1, and simultaneously, $x(\cdot)$ jumps from -1 to 0. It stays there until time T when $x(\cdot)$ jumps back to -1. Each jump in $x(\cdot)$ costs 1.

$$L(x,v) = \frac{1}{2}x^2 + \mathcal{I}_{[-1,1]}(x) + |v|$$

Case 2(b): $T = 4$; **Cost** $= 2$.



Case 2(b): Three types of optimal trajectories:

(i) Same as in Case 2(a).

(ii) $x(\cdot)$ jumps to $-\frac{1}{2}$ and $y(\cdot)$ moves with velocity $-\frac{1}{2}$ from 1 to -1. Then $x(\cdot)$ jumps back to -1.

(iii) $x(\cdot)$ stays at -1 for the duration.

$$L_\lambda(x,v) = \begin{cases} \frac{1}{2}|x|^2 & |x| \le 1+\lambda \\ \frac{1-\lambda^2}{2\lambda}(\lambda + (1-|x|)^2) + \frac{\lambda}{2}|x|^2 & |x| > 1+\lambda \end{cases}$$

$$+ \begin{cases} \frac{1}{2\lambda}|v|^2 & |v| \le \lambda \\ \frac{1-\lambda^2}{2}(2|v| - \lambda) + \frac{\lambda}{2}|v|^2 & |v| > \lambda \end{cases}$$

$$M_\lambda(y,w) = \begin{cases} \frac{\lambda}{2}|y|^2 & |y| \le 1 \\ \frac{1-\lambda^2}{2\lambda}(1-|y|)^2 + \frac{\lambda}{2}|y|^2 & |y| > 1 \end{cases}$$

$$+ \begin{cases} \frac{1}{2}|w|^2 & |w| \le 1+\lambda \\ \frac{1-\lambda^2}{2}(2|w| - (1+\lambda)) + \frac{\lambda}{2}|w|^2 & |w| > 1+\lambda \end{cases}$$

$$H_\lambda(x,y) = M_\lambda(y,0) - L_\lambda(x,0) = \frac{1}{2\lambda}((|y| - 1 + \lambda^2)^2 - (\lambda^4 - \lambda^2 + \lambda|x|^2)).$$

(when $|x| \leq 1 + \lambda$ and $|y| > 1$)

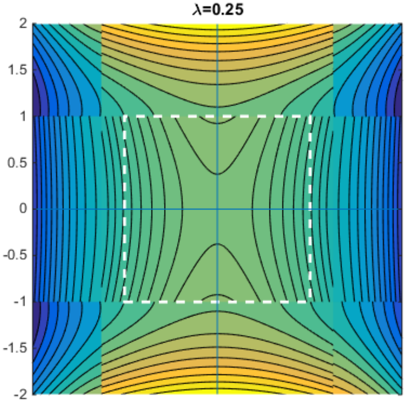


Figure: Level sets of the Self-dual approximate Hamiltonian

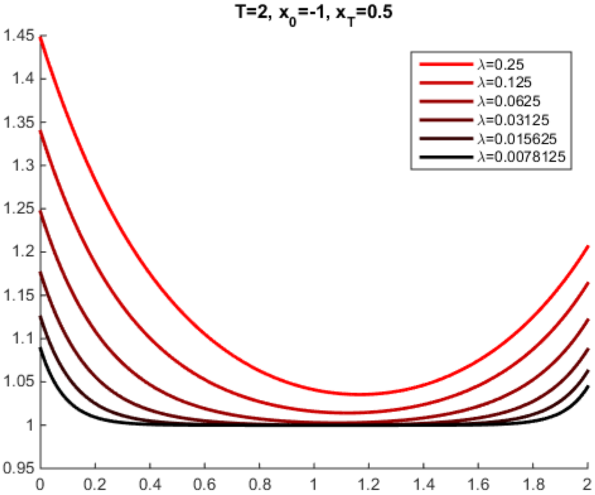


Figure: Convergence of the dual optimal trajectories

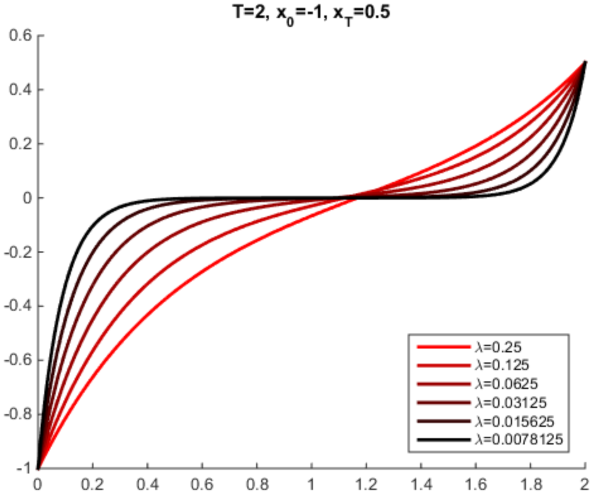


Figure: Convergence of the primal optimal trajectories

Finally ...

This is July 4, which is

America's National Holiday

But in deference to our hosts, my color scheme evolved to:

Tanti auguri, Piermarco!!

Finally ...

This is July 4, which is

America's National Holiday

But in deference to our hosts, my color scheme evolved to:

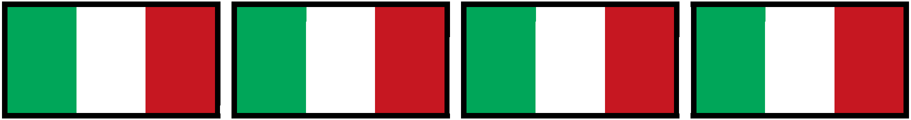
Tanti auguri, Piermarco!!

Finally ...

This is July 4, which is

America's National Holiday

But in deference to our hosts, my color scheme evolved to:



Tanti auguri, Piermarco!!