SELF-DUAL APPROXIMATIONS TO FULLY CONVEX IMPULSIVE SYSTEMS

Cristopher Hermosilla
Department of Mathematics, Louisiana State University, Baton Rouge, LA70803-4918, USA

Peter Wolesnki
Department of Mathematics, Louisiana State University, Baton Rouge, LA70803-4918, USA

Abstract. Fully convex optimal control problems contain a Lagrangian that is jointly convex in the state and velocity variables. Problems of this kind have been widely investigated by Rockafellar and collaborators if the Lagrangian is coercive and without state constraints. A lack of coercivity implies the dual has nontrivial state constraints, and vice versa (that is, they are dual concepts in convex analysis). We consider a framework using Goebel's self-dualizing technique that approximates both the primal and dual problem simultaneously and maintains the duality relationship. Previous results are applicable to the approximations, and we investigate the limiting behavior as the approximations approach the original problem. A specific example is worked out in detail.

1. INTRODUCTION

We are concerned with fully convex optimal control problems, that is, variational problems whose costs are represented by convex functions. Problems of this kind have been widely investigated by Rockafellar's school in several contexts; we mention for instance [1]-[8]. Consider the fully convex control problem with the optimization taken over arcs of Bounded Variation (BV). The primal/dual problems have a symmetric character containing both state constraints and permitting the arcs to have impulses. We plan to establish a link between impulsive problems ([3]) and Absolutely Continuous (AC) ones ([5]) by employing an approximation scheme based on the Goebel's self-dualizing technique [9]. Our ongoing research plan is summarized in the following conjecture:

Conjecture 1. *The optimal solutions to an impulsive Bolza problem over BV can be obtained as the limit of a sequence of primal/dual optimal solutions to a family of approximating AC Bolza problems.*

We describe in this paper how the approximation scheme should work and provide evidence with a detailed example. The main goal motivating this conjecture is to develop a Hamilton-Jacobi theory for

BV problems based on the results in [5] which were obtained under a classical set of assumptions.

The paper is organized as follows: we first present the problem where the minimization is over all absolutely continuous arcs, and then we exhibit an extension to impulsive systems. Secondly, we introduce an approximate scheme for impulsive systems. Finally, we work out an example in detail to show the capability of the approximate scheme.

1.1. **Notation, basic definitions, and preliminaries.** Suppose $f : E \to \mathbb{R} \cup \{\pm\infty\}$ is a function with E being a topological vector space. The effective domain of f is the set $\text{dom}(f) := \{x \in E \mid\ f(x) < +\infty\}$. Then f is called (i) *proper* if $\text{dom}(f) \neq \emptyset$ and $f(x) > -\infty$ for all $x \in E$; (ii) *convex* if $\text{epi}(f) := \{(x, r) \in E \times \mathbb{R} : f(x) \leq r\}$ is a convex set, and (iii) *lower semicontinuous* if $\text{epi}(f)$ is a closed set. The set of functions satisfying (i)-(iii) is denoted by $\Gamma_0(E)$.

When $E = \mathbb{R}^n$, $|\cdot|$ denotes the Euclidean norm and $\langle \cdot, \cdot \rangle$ is the Euclidean inner product on $\mathbb{R}^n$. Let $f \in \Gamma_0(\mathbb{R}^n)$. The Legendre-Fenchel conjugate is

$$f^*(y) := \sup\{\langle x, y\rangle - f(x) \mid\ x \in \text{dom}(f)\},$$

belongs to $\Gamma_0(\mathbb{R}^n)$, and satisfies $(f^*)^* = f$. The recession function f_∞ is given by

$$f_\infty(d) := \sup\{f(x + d) - f(x) \mid\ x \in \text{dom}(f)\}$$

for a direction $d \in \mathbb{R}^n$, and also belongs to $\Gamma_0(\mathbb{R}^n)$ with the additional property of being positively homogeneous. It is the support function of $\text{dom}(f^*) : f_\infty(d) = \sup_{y \in \text{dom}(f^*)}\langle d, y\rangle$; see [10, 11] for details. The subdifferential is the set

$$\partial f(x) := \{y \in \mathbb{R}^n \mid f(z) \geq f(x) + \langle y, z - x\rangle,\ \forall z \in \mathbb{R}^n\}.$$

A function $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is called concave-convex if $\forall x,\ y \in \mathbb{R}^n$, $h_y(\cdot) = -h(\cdot, y)$ and $h_x(\cdot) = h(x, \cdot)$ are both convex functions. The (concave-convex) subdifferential of h is

$$\partial h(x, y) := \left[-\partial(-h_y)(x)\right] \times \partial h_x(y).$$

The distance function to $S \subseteq \mathbb{R}^n$ is denoted by $\text{dist}(x, S)$ with the convention that $\text{dist}(x, \emptyset) = +\infty$. Furthermore, the indicator function of S is denoted by I_S and the The normal cone to S at $x \in S$ is $\mathcal{N}_S(x) := \partial I_S(x)$.

We suppose $T > 0$ is fixed. An arc is just a function $x : [0, T] \to \mathbb{R}^n$ and, when appropriate, is extended to a function defined on $\mathbb{R}$ by setting $x(t) = x(0)$ when $t < 0$ and $x(t) = x(T)$ when $t > T$. The space of continuous, absolutely continuous, bounded variation arcs are denoted by **C**, **AC**, and **BV**, respectively. For $p \in [1, +\infty]$, $\mathbf{L^p}$ denotes the Lebesgue p-integrable arcs, and $\mathbf{W^{1,p}}$ its subspace whose arcs have their derivatives in $\mathbf{L^p}$.

2. FULLY CONVEX BOLZA PROBLEMS

Consider a Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and an end-point cost $\ell : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. The cost functional defined for $x(\cdot) \in \mathbf{AC}$ is given by

$$J_P(x) := \int_0^T L(x(t), \dot{x}(t))dt + \ell(x(0), x(T)).$$

The fully convex Bolza problem has the form

$$(P_0) \qquad \text{Minimize } J_P(x) \quad \text{over all } x \in \mathbf{AC}$$

By allowing L to take infinite values, we are handling implicitly constraints over the state of the system

$$X := \{x \in \mathbb{R}^n \mid \ \exists v \in \mathbb{R}^n,\ L(x, v) \in \mathbb{R}\}.$$

Note $J_P(x) \in \mathbb{R}$ implies $x(t) \in X$ for almost all $t \in [0, T]$. The formulation also includes dynamical and end-point constraints, however, this plays no major role in here.

Our discussion is focused on the fully convex case, where it is assumed that

$$(H_0) \qquad L \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^n) \quad \text{and} \quad \ell \in \Gamma_0(\mathbb{R}^n \times \mathbb{R}^n)$$

The convex framework yields a duality theory utilizing convex conjugate data. The dual problem to (P_0) is

$$(D_0) \qquad \text{Minimize } J_D(y) \quad \text{over all } y \in \mathbf{AC},$$

where the functional J_D is defined via

$$J_D(y) := \int_0^T L^*(\dot{y}(t), y(t))dt + \ell^*(y(0), -y(T)).$$

Analogously, this formulation hides the state constraints

$$Y := \{y \in \mathbb{R}^n \mid \ \exists w \in \mathbb{R}^n,\ L^*(w, y) \in \mathbb{R}\}.$$

The Hamiltonian $H_L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ is given by

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

and is a concave-convex function. The Hamiltonian H_{L^*} associated to L^* satisfies $H_{L^*}(y, x) = -H_L(x, y)$ for any $x, y \in \mathbb{R}^n$, and so is not a new object per se. Thus we are justified to write H for H_L.

2.1. **Optimal control problem.** The abstract framework of the fully convex Bolza problems allows to treat important problems in optimal control as for instance Linear-Quadratic regulators. Indeed, let $X \subseteq \mathbb{R}^n$ and $U \subseteq \mathbb{R}^m$ be two convex closed nonempty set, A be a $n \times n$ matrix, B be a $n \times m$ matrix and, Q and R be two symmetric positive semi-definite matrices of dimension n and m, respectively. We aim at minimizing

$$\frac{1}{2}\int_0^T [x(t)^\tau Qx(t) + u(t)^\tau Ru(t)]dt$$

subject to the dynamical constraint

$$\dot{x}(t) = A(t)x(t) + B(t)u(t),$$

and the input constraint

$$u(t) \in U, \qquad \text{for a.e. } t \in [0,T].$$

Under these circumstances, the accumulative cost for the Bolza problem is given by $L(x,v) = x^TQx + \mathcal{L}(x,v)$, where

$$\mathcal{L}(x,v) = \begin{cases} \inf\limits_{u\in U}\{u^\tau Ru \mid \ v = Ax + Bu\} & \text{if } x \in X \\ +\infty & \text{otherwise} \end{cases}$$

with the convention that $\inf(\emptyset) = +\infty$. It is clear (H_0) holds.

2.2. **Duality relations.** The relation between the primal and dual Bolza problems is well-understood under the following assumption

$$(H_1) \qquad \begin{cases} (1)\ \exists\rho > 0 \text{ so that for any } x \in \mathbb{R}^n : \\ \qquad \operatorname{dist}(0, \operatorname{dom}(L(x,\cdot))) \le \rho(1 + |x|). \\ (2)\ \exists\alpha,\beta \in \mathbb{R} \text{ and } \theta : \mathbb{R} \to \mathbb{R} \text{ nondecreasing,} \\ \text{proper and coercive on } [0,+\infty) \text{ so that :} \\ \qquad L(x,v) \ge \theta(\max\{0, |v| - \alpha|x|\}) - \beta|x| \end{cases}$$

Note that $(H_1)(1)$ implies that for any $x \in \mathbb{R}^n$ there is some $v \in \mathbb{R}^n$ so that $L(x,v) \in \mathbb{R}$, and in particular means that no primal state constraint is involved (X equals $\mathbb{R}^n$). Additionally, $(H_1)(2)$ implies the recession function L_∞ is $I_{\{0\}}$, and so there is no dual state constraints (Y equals $\mathbb{R}^n$).

Our approximating primal and dual Lagrangians will actually be finite, and the next proposition summarizes the relation between the value of (P_0) and (D_0) in this case. Further results are available (see [5, Theorem 4.5] for details) but are not relevant here.

Proposition 2.1. *Suppose that (H_0) and (H_1) hold.*

(1) *If L is finite, then* $\operatorname{val}(P_0) \in \mathbb{R}$ *and there is an optimal solution* $y \in \mathbf{AC}$ *for* (D_0).
(2) *If L^* is finite, then* $\operatorname{val}(D_0) \in \mathbb{R}$ *and there is an optimal solution* $x \in \mathbf{AC}$ *for* (P_0).

In both cases, we have $\operatorname{val}(P_0) + \operatorname{val}(D_0) = 0$.

Proof. Note that the assumptions imply that any $a, b \in \mathbb{R}^n$ with a, b belonging to $\operatorname{dom}(\ell)$ or $\operatorname{dom}(\ell^*)$ can be joint by an arc, whose integral cost is finite, respectively. The result follows then as a direct consequence of [1, Theorem 1]. □

We recall the optimality conditions for a primal arc $x \in \mathbf{AC}$ to be a solution of (P_0). This involves the existence of a dual arc $y \in \mathbf{AC}$ in which the Hamiltonian system

$$(-\dot{y}(t), \dot{x}(t)) \in \partial H(x(t), y(t)), \quad \text{a.e. } t \in [0, T] \tag{1}$$

plus appropriate transversality conditions is satisfied. Explicitly, we have

Proposition 2.2 ([5, Theorem 4.1])**.** *Suppose that* (H_0) *holds. Let* $x, y \in \mathbf{AC}$ *be two given arcs. Then,* (x, y) *is a trajectory of* (1) *satisfying the transversality condition*

$$(y(0), -y(T)) \in \partial \ell(x(0), x(T)) \tag{2}$$

if and only if x and y are optimal solutions of (P_0) *and* (D_0)*, respectively, and* $\operatorname{val}(P_0) + \operatorname{val}(D_0) = 0$.

2.3. **State constraints and extended problems.** When state constraints are involved, it is expected the adjoint arc will have jumps when the constraint is active, and this naturally leads to the dual problem minimizing over $\mathbf{BV}$ rather than $\mathbf{AC}$. The philosophy of convex analysis is that symmetry between primal and dual problems should be adhered to, and hence the primal problem should be extended to minimizing over arcs of bounded variation as well. For these reasons the variational problems were extended by Rockafellar [3] to $\mathbf{BV}$ in the following manner: For $x \in \mathbf{BV}$, we write $dx(t) = \dot{x}(t)dt + \xi_x(t)d\mu(t)$ for some singular measure (w.r.t. the Lebesgue measure) μ. The extended primal problem is

$$\min_{x \in \mathbf{BV}} J_P(x) + \int_0^T L_\infty(0, \xi_x(t)) d\mu(t) \tag{P}$$

Similarly the extended dual problem is

$$\min_{y \in \mathbf{BV}} J_D(y) + \int_0^T L^*_\infty(\xi_y(t), 0) d\mu(t) \tag{D}$$

The formulations are invariant w.r.t. the singular measure μ taken. Moreover, it is clear that if $L_\infty(0,\cdot) = I_{\{0\}}$, then (P) recovers the formulation (P_0); similar comments apply to the dual problem, which is the case when the dual has no state constraints.

Optimality conditions are available here as well. According to [3, Theorem 2], if $\text{val}\,(P_0), \text{val}\,(D_0) \in \mathbb{R}$, then the arcs $x, y \in \mathbf{BV}$ are optimal solutions of (P) and (D), respectively, if and only if (1) and (2) hold together with the following:

$$\text{val}\,(P) + \text{val}\,(D) = 0 \tag{3}$$

$$x(t-), x(t+) \in \overline{X}, \quad \forall t \in [0,T] \tag{4}$$

$$y(t-), y(t+) \in \overline{Y}, \quad \forall t \in [0,T] \tag{5}$$

$$\xi_x(t) \in \mathcal{N}_{\overline{Y}}(y(t-)) \cap \mathcal{N}_{\overline{Y}}(y(t+)), \quad \mu\text{-a.e. } t \in [0,T] \tag{6}$$

$$\xi_y(t) \in \mathcal{N}_{\overline{X}}(x(t-)) \cap \mathcal{N}_{\overline{X}}(x(t+)), \quad \mu\text{-a.e. } t \in [0,T] \tag{7}$$

3. APPROXIMATE PROBLEMS

Let us from now on fix a Lagrangian L and an end-points cost ℓ that satisfy (H_0) but not (H_1). This means that either the (P_0) or (D_0) dual has state constraints. We would like now to provide an scheme to approximate our constrained problem by a sequence of problems without state constraints.

This type of approach can be found in the literature in different contexts. For example, it is used in [12, Chapter 4] to obtain optimality conditions, that generalize to the infinite dimensional case the ones already provided in [2, 3]. We also mention [13], where the author deals with the dynamics of elastic shocks, a problem in which the acceleration of the state is expected to be a measure.

The key tool used in the quoted works is the so-called Moreau-Yosida approximate on a Hilbert space $(\mathbb{H}, \|\cdot\|)$ of a function $f \in \Gamma_0(\mathbb{H})$, which is defined for any $\lambda > 0$ via

$$[f]_\lambda(x) := \inf_{z \in \mathbb{H}} \left\{ f(z) + \frac{1}{2\lambda}\|z - x\|^2 \right\}.$$

Thus, at first sight it may seem a good idea to take the Moreau-Yosida approximate of L and see how far we can reach. Nevertheless, by doing so, we are only precluding state constraints on the primal problem and not on the dual. The reason is that, as pointed out in [14, Theorem 2.1], the recession function of L and its Moreau-Yosida approximate are the same. Indeed, if $(\mathbb{H}, \|\cdot\|)$ is a finite dimensional Hilbert space and $f \in \Gamma_0(\mathbb{H})$, then for any $\lambda > 0$

$$f_\infty(d) = ([f]_\lambda)_\infty(d), \quad \forall d \in \mathbb{H}.$$

This fact shows that another kind of approximation needs to be used, one that essentially erases state constraints on the primal and dual problems at the same time.

3.1. **Self-dual approximate scheme.** For any $\lambda \in (0,1)$, we denote by L_λ and L^*_λ the self-dual approximate of L and L^* introduced in [9], that is,

$$L_\lambda(x,v) := (1-\lambda^2)[L]_\lambda(x,v) + \frac{\lambda}{2}(|x|^2 + |v|^2),$$
$$L^*_\lambda(w,y) := (1-\lambda^2)[L^*]_\lambda(w,y) + \frac{\lambda}{2}(|y|^2 + |w|^2).$$

This approximate enjoys several favorable properties inherited from the Moreau-Yosida approximate: it is everywhere continuously differentiable with its gradient being $\frac{1}{\lambda}$-Lipschitz continuous on the entire space. Furthermore, it turns out that, L_λ and L^*_λ are conjugate to each other, from where comes the name; see [9] for details. Also, because of the quadratic terms in their definitions, we have that

$$(L_\lambda)_\infty = +\infty \quad \text{and} \quad (L^*_\lambda)_\infty = +\infty, \quad \text{on } \mathbb{R}^n \times \mathbb{R}^n \setminus \{0\}.$$

This fact yields to a primal and a dual approximate problems without state constraints. Moreover, the Hamiltonian associated with L_λ

$$H_\lambda(x,y) := \sup_{v\in\mathbb{R}^n} \{\langle v,y\rangle - L_\lambda(x,v)\}$$

agrees with the self-dual approximate of H is the concave-convex sense, that is,

$$H_\lambda(x,y) = (1-\lambda^2)[H]_\lambda(x,y) + \frac{\lambda}{2}(|y|^2 - |x|^2).$$

Here, $[H]_\lambda$ stands for the Moreau-Yosida approximate of H in the concave-convex sense, which is given by

$$[H]_\lambda(x,y) := \sup_{p\in\mathbb{R}^n} \inf_{q\in\mathbb{R}^n} \left[H(p,q) + \frac{|q-y|^2 - |p-x|^2}{2\lambda} \right].$$

The self-dual approximate of H is also everywhere continuously differentiable and its gradient is $\frac{1}{\lambda}$-Lipschitz continuous on the entire space; this is because the Moreau-Yosida approximate satisfies these properties (cf. [15]). It is a relatively simple matter to show L_λ satisfies (H_1).

We now consider the approximated Bolza problem:

$$(P_\lambda) \qquad \min_{x\in\mathbf{AC}} \int_0^T L_\lambda(x(t),\dot{x}(t))dt + \ell(x(0),x(T))$$

and its dual counterpart

$$(D_\lambda) \qquad \min_{y\in\mathbf{AC}} \int_0^T L_\lambda^*(\dot y(t), y(t))dt + \ell^*(y(0), -y(T)).$$

Note that the end-point costs are the same as in the original primal and dual problems. Furthermore, by Proposition 2.1 we have that both problems (P_λ) and (D_λ) have solutions and that strong duality holds, that is,

$$\mathrm{val}\,(P_\lambda) + \mathrm{val}\,(D_\lambda) = 0 \tag{8}$$

Letting $x_\lambda(\cdot)$ and $y_\lambda(\cdot)$ be respectively optimal solutions of (P_λ) and (D_λ), by Proposition 2.2, we have that they satisfy the Hamiltonian systems

$$(-\dot y_\lambda(t), \dot x_\lambda(t)) = \nabla H_\lambda(x_\lambda(t), y_\lambda(t)), \quad \forall t \in (0,T) \tag{9}$$

as well as the transversality condition

$$(y_\lambda(0), -y_\lambda(T)) \in \partial\ell(x_\lambda(0), x_\lambda(T)) \tag{10}$$

The Lipschitz continuity of ∇H_λ implies that $x_\lambda(\cdot)$ and $y_\lambda(\cdot)$ are uniquely determined provided that the initial or final conditions are fixed, that is, $x_\lambda(0)$ and $y_\lambda(0)$ are given. We also have that the primal/dual pair satisfies the Euler-Lagrange equations

$$(\dot y_\lambda(t), y_\lambda(t)) = \nabla L_\lambda(x_\lambda(t), \dot x_\lambda(t)), \quad \forall t \in (0,T) \tag{11}$$

$$(x_\lambda(t), \dot x_\lambda(t)) = \nabla L_\lambda^*(\dot y_\lambda(t), y_\lambda(t)), \quad \forall t \in (0,T) \tag{12}$$

4. An example

We consider a one-dimensional example to illustrate the approach we propose to answer Conjecture 1. Let

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

This problems includes $X = [-1,1]$ as implicit state constraints. Also, for $x_0, x_T \in X$ fixed, we consider

$$\ell(a,b) = \begin{cases} 0 & \text{if } a = x_0,\ b = x_T, \\ +\infty & \text{otherwise.} \end{cases}$$

The fact that the Lagrangian has nontrivial recession direction $(0,d)$ implies that the dual problem has state constraints as well, namely

$Y = [-1,1]$. This affirmation can also be verified when looking at the dual Lagrangian

$$L^*(w,y) = I_{[-1,1]}(y) + \begin{cases} \frac{1}{2}w^2 & |w| \leq 1, \\ |w| - \frac{1}{2} & |w| > 1. \end{cases}$$

In this setting, the dual end-points cost is the linear map

$$\ell^*(c,d) = x_0 c + x_T d.$$

Note that in this case the transversality condition (2) doesn't provide any information. The Hamiltonian associated with the primal problem is then

$$H(x,y) = I_{[-1,1]}(y) - \frac{1}{2}x^2 - I_{[-1,1]}(x).$$

Using, the optimality conditions provided in [3, Theorem 2], that is, (1)-(7), we get that for the case $x_0 \cdot x_T < 0$ the arc

$$x^*(t) = \begin{cases} x_0 & \text{if } t = 0, \\ 0 & \text{if } t \in (0,T), \\ x_T & \text{if } t = T \end{cases} \tag{13}$$

is an optimal solution together with the dual arc

$$y^*(t) = \operatorname{sign}(x_T) = -\operatorname{sign}(x_0), \quad t \in [0,T] \tag{14}$$

By definition we have that (4)-(5) are satisfied. Moreover, (3) holds with

$$\operatorname{val}(P) = |x_0| + |x_T| = -\operatorname{val}(D)$$

Letting μ be the atomic measure supported at $\{0,T\}$, we have that $\xi_{x^*}(t) = -x_0\delta_0(t) + x_T\delta_T(t)$ and $\xi_{y^*}(t) = 0$, and of course $\dot{x}^*(t) = \dot{y}^*(t) = 0$ for any $t \in [0,T]$. From these remarks it is easy to see that (1), (6) and (7) are satisfied.

4.0.1. *Approximated solutions.* The self-dual regularization of the Lagrangians provides the following formulas

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

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

Note that the primal problem (P_λ) is strictly convex, and so, the data of the problem determine in a unique way the arc $t \mapsto x_\lambda(t)$. This is not the case for the dual problem (D_λ), because the end-points cost is linear. Therefore, we might expect that the (D_λ) has several solutions $t \mapsto y_\lambda(t)$ associated with the unique primal solution of (P_λ) (although uniquely determined by $y_\lambda(0)$).

From the structure of the problem, it is easy to see that

$$H_\lambda(x,y) = L_\lambda^*(0,y) - L_\lambda(x,0).$$

The level sets of this function has been sketched in Fig. 1. Besides, we get that if $|x| \le 1+\lambda$ and $|y| > 1$

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

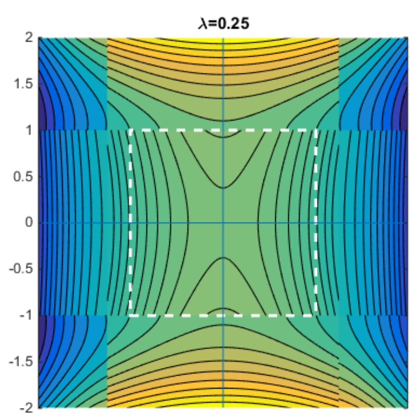

FIGURE 1. Level sets of the Self-dual approximate Hamiltonian

Hence, if $y_\lambda(0)$ is taken so that $H_\lambda(x_0, y_\lambda(0)) > \frac{\lambda}{2}$ then $|y_\lambda(t)| > 1$ for any $t \in [0,T]$ because the Hamiltonian is constant along the optimal arcs. Setting $k_\lambda^T := \frac{(1-\lambda^2)\lambda T}{2}$, the values of (P_λ) and (D_λ) are

$$\begin{aligned}\mathrm{val}\,(P_\lambda) &= \frac{1}{2}\|x_\lambda\|_{L^2}^2 + \frac{\lambda}{2}\|\dot{x}_\lambda\|_{L^2}^2 + (1-\lambda^2)\|\dot{x}_\lambda\|_{L^1} - k_\lambda^T,\\ \mathrm{val}\,(D_\lambda) &= \frac{1}{2\lambda}\||y_\lambda| - 1 + \lambda^2\|_{L^2}^2 + \frac{1}{2}\|\dot{y}_\lambda\|_{L^2}^2 + k_\lambda^T\\ &\quad + x_0 y_\lambda(0) - x_T y_\lambda(T).\end{aligned}$$

4.1. **Some estimates and convergence of optimal arcs.** Note that $x(t) = x_0 + \frac{t}{T}(x_T - x_0)$ is a feasible trajectory for (P_λ) and we can assume that $|x_T - x_0| > \lambda T$ (taking λ small enough), so by definition of (P_λ) we get

$$\mathrm{val}\,(P_\lambda) \le \frac{1}{2}\|x\|_{L^2}^2 + \frac{1}{2}\|\dot{x}\|_{L^2}^2 + \|\dot{x}\|_{L^1}.$$

Since $\mathrm{val}\,(P_\lambda) \ge 0$, we get that x_λ and $\sqrt{\lambda}\dot{x}_\lambda$ are uniformly bounded in $\mathbf{L^2}$ and $\dot{x}_\lambda$ is uniformly bounded in $\mathbf{L^1}$. Also, under these circumstances the Hamiltonian system is

$$\begin{cases}\dot{x}_\lambda(t) &= \frac{1}{\lambda}(|y_\lambda(t)| - 1 + \lambda^2)\ \frac{y_\lambda(t)}{|y_\lambda(t)|},\\ \dot{y}_\lambda(t) &= x_\lambda(t),\end{cases} \qquad \forall t \in (0,T)$$

From where we can infer that $\dot{y}_\lambda$ is uniformly bounded in $\mathbf{L^2}$ and $|y_\lambda| \to 1$ in $\mathbf{L^2}$ and in $\mathbf{L^1}$.

4.1.1. *Convergence of the dual optimal arc.* By (8) we get that $|y_\lambda(0)||x_T - x_0|$ is bounded by

$$\|x_\lambda\|_{L^2}^2 + \lambda\|\dot{x}_\lambda\|_{L^2}^2 + (1-\lambda^2)\|\dot{x}_\lambda\|_{L^1} + x_T\|\dot{y}_\lambda\|_{L^1}.$$

Which means that $|y_\lambda(0)|$ is uniformly bounded and since $\dot{y}_\lambda$ is also uniformly bounded in $\mathbf{L^1}$, we get that $|y_\lambda(T)|$ is uniformly bounded. In particular, since $\{|y_\lambda|\}_{\lambda\in(0,1)}$ is bounded in $\mathbf{W^{1,2}}$, by the Sobolev injections [16, Theorem 8.8] we get that, passing into a subsequence if necessary, $|y_\lambda| \to 1$ uniformly in $\mathbf{C}$; the fact that the limit is 1 is because we already know that $|y_\lambda| \to 1$ in $\mathbf{L^1}$. Fig. 2 sketches the uniform convergence of the dual arc, with a time horizon $T = 2$ and end points $x_0 = -1$ and $x_T = 0.5$.

Moreover, if $x_0 < 0 < x_T$ then $\dot{x}_\lambda(t) > 0$ and so $y_\lambda(t) > 1$. Otherwise, we have $y_\lambda(t) < 1$. Hence, in any case we get $y_\lambda \to -\mathrm{sign}(x_0) = \mathrm{sign}(x_T)$ uniformly in $\mathbf{C}$, from where we recover (14). This also leads to

$$\ell^*(y_\lambda(0), -y_\lambda(T)) \to -|x_0| - |x_T|.$$

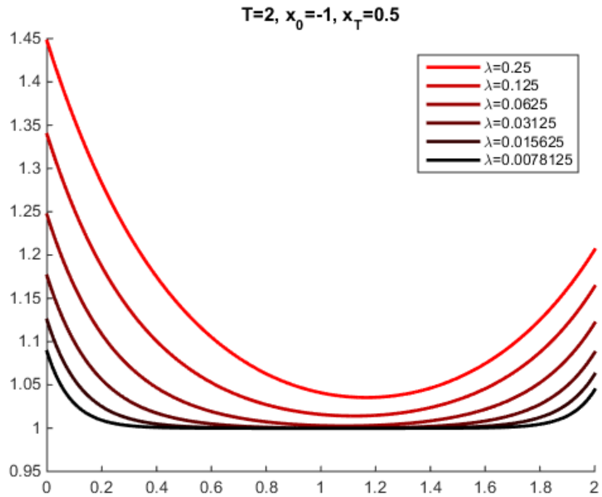

FIGURE 2. Convergence of the dual optimal trajectories

Since $x_0 \cdot x_T < 0$ and $|\dot{x}_\lambda| > 0$, there is $t_\lambda \in (0, T)$ so that

$$\operatorname{sign}(\dot{y}_\lambda(t)) = \operatorname{sign}(x_\lambda(t)) = \operatorname{sign}(x_0), \quad \forall t \in [0, t_\lambda),$$
$$\operatorname{sign}(\dot{y}_\lambda(t)) = \operatorname{sign}(x_\lambda(t)) = \operatorname{sign}(x_T), \quad \forall t \in (t_\lambda, T].$$

We assume with loss of generality that $t_\lambda \to \tau \in [0, T]$ and so, for any $\varepsilon > 0$ we have that $\|\dot{y}_\lambda\|_{L^1}$ is bounded by

$$\operatorname{sign}(x_0)(y_\lambda(\tau - \varepsilon) - y_\lambda(0) + y_\lambda(\tau + \varepsilon) - y_\lambda(T)) + \sqrt{2\varepsilon}\|\dot{y}_\lambda\|_{L^2}$$

Hence, taking limsup as $\lambda \to 0$ we get that

$$\limsup_{\lambda \to 0} \|\dot{y}_\lambda\|_{L^1} \leq \sqrt{2\varepsilon} \limsup_{\lambda \to 0} \|\dot{y}_\lambda\|_{L^2}.$$

Thus, given that $\dot{y}_\lambda$ is uniformly bounded in $\mathbf{L^2}$ and $\varepsilon > 0$ is arbitrary, we have then that $\dot{y}_\lambda \to 0$ in $\mathbf{L^1}$ and we can also assume that $\dot{y}_\lambda \to 0$ pointwise for a.e. $t \in [0, T]$. Recall that $\dot{y}_\lambda$ is also uniformly bounded in $\mathbf{L^2}$, so by the Dominated Convergence Theorem, $\dot{y}_\lambda \to 0$ in $\mathbf{L^2}$ as well.

4.1.2. *Convergence of the primal optimal arc.* By the Hamiltonian system and the preceding part, we have that $x_\lambda \to 0$ in $\mathbf{L^1}$, $\mathbf{L^2}$ and pointwise for a.e. $t \in [0, T]$.

Note that $\{x_\lambda\}_{\lambda \in (0,1)}$ is bounded in $\mathbf{W^{1,1}}$ (but not necessarily in $\mathbf{W^{1,2}}$). Hence, $\{x_\lambda\}_{\lambda \in (0,1)}$ is also bounded in $\mathbf{L}^\infty$ and so, by the Helly Theorem [12, Theorem 1.126], passing again into a subsequence if necessary, we can assume that there is $x^* \in \mathbf{BV}$ such that $x_\lambda \to x^*$

pointwise for a.e. $t \in [0,T]$ and $\dot{x}_\lambda$ converges in the weak-$\star$ topology of the dual space to $\mathbf{C}$, that is

$$\int_0^T \varphi(t)\dot{x}_\lambda(t)dt \to \int_0^T \varphi(t)dx^*(t) \quad \forall \varphi \in \mathbf{C}.$$

Fig. 3 sketches the pointwise convergence of the primal arc as $\lambda \to 0$ for the same date used in Fig. 2.

Note as well that $x^* = 0$ on S, a full measure subset of $[0,T]$, but since $x^* \in \mathbf{BV}$ its lateral limits are well-defined. In particular, given that S is dense in $[0,T]$ we have

$$\left.\begin{aligned} x(t+) &= \lim_{s\to t^+,\ s\in S} x^*(s), \\ x(t-) &= \lim_{s\to t^-,\ s\in S} x^*(s), \end{aligned}\right. \quad \forall t \in (0,T).$$

Therefore, it is not restrictive to assume that $x^*(t) = 0$ at any $t \in (0,T)$. Furthermore, due to the fact that $t \mapsto x_\lambda(t)$ is monotonic, we must have

$$\|\dot{x}_\lambda\|_{L^1} \to \mathrm{var}(x^*),$$

but since $x_T - x_0 = \int_0^T \dot{x}_\lambda(t)dt$ we get that

$$\|\dot{x}_\lambda\|_{L^1} = |x_T - x_0| = |x_T| + |x_0|,$$

from where we get that $\mathrm{var}(x^*) = |x_T| + |x_0|$, and so, we recover the optimal solution given by (13).

Finally, by (8) and the Hamiltonian system we have that $\sqrt{\lambda}\dot{x}_\lambda$ and $\frac{1}{\sqrt{\lambda}}(|y_\lambda| - 1)$ converges to zero in $\mathbf{L^2}$ because the square of their $\mathbf{L^2}$ norms equal

$$x_T y_\lambda(T) - x_0 y_\lambda(0) - \|x_\lambda\|_{L^2}^2 - (1-\lambda^2)\|\dot{x}_\lambda\|_{L^1} - c_\lambda^T.$$

This in turn means that

$$\mathrm{val}\,(P_\lambda) \to \mathrm{val}\,(P) \quad \text{and} \quad \mathrm{val}\,(D_\lambda) \to \mathrm{val}\,(D).$$

4.2. **Convergence of the optimality conditions.** The optimality conditions (1)-(7) can also be recovered from the abstract setting we have treated our example. Indeed, the uniform convergence of y_λ combined with the general fact that

$$[f]_\lambda(z) \geq \inf(f) + \frac{1}{2\lambda}\mathrm{dist}(z, \mathrm{dom}(f))^2$$

imply that $\mathrm{dist}(y_\lambda, Y) \to 0$ in $\mathbf{L^2}$, and so $\mathrm{dist}(y_\lambda, Y) \to 0$ uniformly on $\mathbf{C}$, from where we obtain (5). Moreover, (7) comes from the fact that, for each $t \in [0,T]$, either $x^*(t-)$ or $x^*(t+)$ belongs to the interior of X, and so the normal cone is reduced to $\{0\}$ at any $t \in [0,T]$.

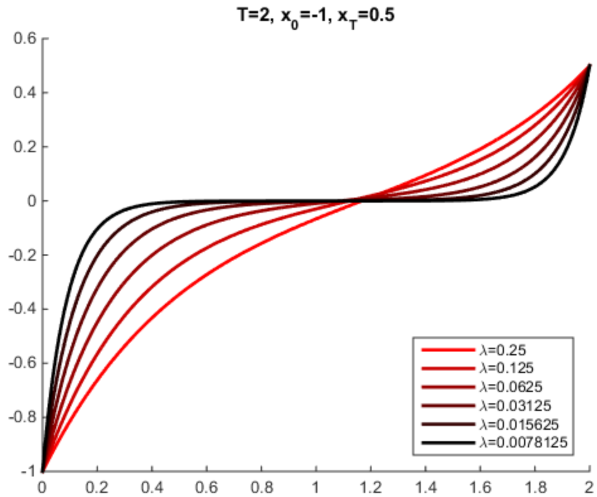

FIGURE 3. Convergence of the primal optimal trajectories

We have already mentioned that the transversality condition (2) is trivial and we have shown that the value approximate functions converge to the original ones, which gives (3). So, it remains to show that (1), (4) and (6) can also be inferred from the scheme. The techniques required are rather general, and so, applicable to more abstract situations.

First of all, (1) can be obtained from (9) using the pointwise convergence of x_λ and y_λ, and the a.e. pointwise convergence of $\dot{x}_\lambda$ and $\dot{y}_\lambda$, combined with the fact that $H_\lambda \to H$ in the hypo/epi convergence sense; see [15].

Secondly, (4) and (6) are obtained as a consequence of (12) and [17, Corallary 5A]. Indeed, integrating (12) and passing into the limit, we get that the measure $dx^* \in \partial\Phi(y^*)$ where

$$\Phi(y) := \int_0^T L^*(\dot{y}^*(t), y(t))dt, \quad \forall y \in \mathbf{C}.$$

Note that $\Phi \in \Gamma_0(\mathbf{C})$ and so the definition of $\partial\Phi(y^*)$ follows the same logic as the subdifferential of $f \in \Gamma_0(\mathbb{R}^n)$ but replacing the Euclidean inner product with the duality product between $\mathbf{C}$ and its topological dual space.

5. CONCLUSIONS AND FUTURE WORKS

We have shown evidence that the approximate scheme we have proposed for fully convex Bolza problems is a suitable tool for studying impulsive systems derived from state constrained problems. The techniques we have exhibited to treat our example are part of an abstract approach we are currently investigating to get a general result to link impulsive systems with classical ones throughout the Goebel's self-dualizing technique. As we have also pointed out, some techniques can be adapted to abstract settings provided that appropriate estimates are obtained. This we believe will lead to a positive answer to Conjecture 1.

The main purpose of establishing the link we described above is understanding the generalized characteristic methods [5] in the case of impulsive systems. It is also not clear how to interpret the Hamilton-Jacobi equation in this setting, without making use of the so-called graph completion technique. For this reason, we also expect that the self-dual approximation technique will provide some insights into the Hamilton-Jacobi theory for impulsive problem.

References

[1] R. T. Rockafellar, *Existence and duality theorems for convex problems of Bolza*, Trans. Amer. Math. Soc. (1971) Vol. 159, pp. 1–40.

[2] R. T. Rockafellar, *State Constraints in Convex Control Problems of Bolza*, SIAM J. Control Optim. (1972) Vol. 10, No. 4, pp. 691–715.

[3] R. T. Rockafellar, *Dual problems of Lagrange for arcs of Bounded Variation*, Calculus of Variations and Control Theory, DL Russell, ed., Academic Press, New York (1976), pp. 155–192.

[4] P. Loewen and R. T. Rockafellar, *Optimal Control of Unbounded Differential Inclusions*, SIAM J. Control Optim. (1994) Vol. 32, No. 2, pp. 442–470.

[5] R. T. Rockafellar and P.R. Wolenski, *Convexity in Hamilton-Jacobi Theory I: Dynamics and duality*, SIAM J. Control Optim. (2000) Vol. 39, No. 5, pp. 1323–1350.

[6] R. Goebel and R. T. Rockafellar, *Generalized conjugacy in Hamilton-Jacobi theory for fully convex Lagrangians*, J. Convex Anal. (2002) Vol. 9, No. 2, pp. 463–473.

[7] R. Goebel, *Convex optimal control problems with smooth Hamiltonians*, SIAM J. Control Optim. (2005) Vol. 43, No. 5, pp. 1787–1811.

[8] T. Pennanen and A. Perkkiö, *Duality in convex problems of Bolza over functions of bounded variation*, SIAM J. Control Optim. (2014) Vol. 52, No. 3, pp. 1481–1498.

[9] R. Goebel, *Self-dual smoothing of convex and saddle functions*, J. Convex Anal.(2008), Vol. 15 No. 1, pp. 179–190.

[10] R. T. Rockafellar, *Convex Analysis*, Princeton University Press (1970).

[11] A. Auslender and M. Teboulle *Asymptotic Cones and Functions in Optimization and Variational Inequalities*, Springer New York (2003).

[12] V. Barbu and T. Precupanu, T. *Convexity and optimization in Banach spaces*, Springer Netherlands (2012).

[13] Michelle Schatzman, *A class of nonlinear differential equations of second order in time*, Nonlinear Anal. Theor. (1978) Vol. 2, No. 3, pp 355–373.

[14] C. Lemaréchal and C. Sagastizábal, *Practical aspect of the Moreau-Yosida regularization: Theoretical preliminaries*, SIAM J. Control Optim. (1997) Vol. 7, No. 2, pp. 367–385.

[15] H. Attouch, D. Azé and R. Wets, *On continuity properties of the partial Legendre-Fenchel transform: convergence of sequences of augmented Lagrangian functions, Moreau-Yosida approximates and subdifferential operators, In:* Fermat Days 85: Mathematics for Optimization*, J.-B. Hiriart-Urruty (ed.)*, North-Holland, Amsterdam (1986), Vol. 129, pp 1–42.

[16] H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations*, Springer-Verlag New York (2010).

[17] R. T. Rockafellar, *Integrals which are convex functionals II*, Pac. J. Math. (1971) Vol. 39, No. 2, pp. 439–469.