French-German-Spanish Conference on Optimization
Gijón, June 18-21, 2024
(pp. 37-45)

On the equivalence of some relaxations of optimal control problems on unbounded time domains

Ilya Dikariev[1], **Sabine Pickenhain**[2]

1. dikarill@b-tu.de Brandenburg Technical University Cottbus-Senftenberg, Germany
2. sabine.pickenhain@b-tu.de Brandenburg Technical University Cottbus-Senftenberg, Germany

1. Introduction

Relaxation methods are a general concept for solving problems that lack convexity. There are several such methods, and we consider three of them: Γ-regularization by [3], Young measures by [4], and convex combinations by [2]. For bounded time domains, the comparisons are mostly done by Roubíček in [8], considering different generalizations of Young measures.

We consider the relaxations for unbounded time domains and/or unbounded control sets. We establish sufficient conditions under which all these three types of relaxations are equivalent to each other. Furthermore, we give an example showing that in some cases the relaxations differ.

The equivalence to the relaxation of the problem via convex combinations is convenient for computations. This type of formulation does not introduce any new mathematical objects such as Radon measures or bipolars, but rather involves no more that functions, derivatives, and so on.

In the scenario where two problems $(\mathrm{PR}_1), (\mathrm{PR}_2)$ are equivalent, one can establish the existence of an optimal solution for the first by proving the existence for the other, and vice versa. In the subject "Existence Theorem for Relaxed Control Problems on Infinite Time Horizon Utilizing Weight Functions" on the conference (FGS2024, Gijón), we present existence results for relaxed optimal control problems utilizing Young measures technique. In this manner, one can automatically derive existence results for other equivalent relaxations.

In the following, we present only the proofs that are not contained in the cited works, or that need modification.

Definition 1.1 Let $(\mathrm{P}_1), (\mathrm{P}_2)$ be two abstract optimization problems with admissible sets A_1, A_2 and real valued objectives J_1, J_2:

$$\begin{aligned} & J_1(x) \longrightarrow Min \\ s.t. \quad & x \in A_1 \end{aligned} \quad (\mathrm{P}_1), \qquad \begin{aligned} & J_2(y) \longrightarrow Min \\ s.t. \quad & y \in A_2 \end{aligned} \quad (\mathrm{P}_2).$$

We call the problems $(\mathrm{P}_1), (\mathrm{P}_2)$ **equivalent** if there are two mappings $\iota_1 : A_1 \to A_2, \iota_2 : A_2 \to A_1$ with the property $J_2(\iota_1(x)) \le J_1(x)$ (resp. $J_1(\iota_2(y)) \le J_2(y)$) for all $x \in A_1$ (resp. $y \in A_2$).

It follows from this definition that the mappings $\iota_{1,2}$ map minimizing sequences (optimal solution) of J_1 to minimizing sequences (optimal solution) of J_2 and vice versa.

Lemma 1.2 *Let the problems* $(\mathrm{P}_1), (\mathrm{P}_2)$ *be equivalent with corresponding mappings* ι_1, ι_2. *Furthermore, let* $\{x_i\}_{i\in\mathbb{N}}$ *be a minimizing sequence of* $J_1(x)$. *Then* $\iota_1(x_i)$ *represents a minimizing sequence of* $J_2(x)$. *Moreover, if* x^* *is an optimal solution of* (P_1), *then* $\iota_1(x^*)$ *forms an optimal solution of* (P_2).

Proof *We denote as* y_i *the images* $\iota_1(x_i)$ *and assume that there exists* $\bar{y} \in A_2$ *with* $J_2(\bar{y}) < \inf_{i\in\mathbb{N}} J_2(y_i)$. *We then obtain a contradiction to* $\{x_i\}$ *being a minimizing sequence because the image* $\iota_2(\bar{y})$ *is admissible for* (P_1), *i.e. lies in* A_1, *and*

$$\forall i \in \mathbb{N}: \quad J_1(\iota_2(\bar{y})) \le J_2(\bar{y}) < J_2(y_i) \le J_1(x_i).$$

The second statement is rather trivial. One consider the existence of an admissible solution $\bar{y} \in A_2$ *with* $J_2(\bar{y}) < J_2(\iota_1(x^*))$, *and we obtain a contradiction to* $J_1(x^*) = \inf_{x\in A_1} J_1(x)$:

$$J_1(\iota_2(\bar{y})) \le J_2(\bar{y}) < J_2(\iota_1(x^*)) \le J_1(x^*).$$

□

We relax an optimal control problem of following type:

$$\begin{aligned} J(x,u) &= \int_\Omega r(t,x(t),u(t))\mathrm{d}t \longrightarrow Min, \\ \dot{x}(t) &= f(t,x(t),u(t)) \text{ a.e. on } \Omega,\ x(t_0) = x_0, \\ x &\in W_p^{1,n}(\Omega,\nu), \\ u(t) &\in U \subseteq \mathbb{R}^m \text{ a.e. on } \Omega, \end{aligned} \tag{P}$$

where $f(t,\xi,\mathrm{v})$ is a Carathéodory function $\Omega \times \mathbb{R}^{n+m} \to \mathbb{R}^n$, $r(t,\xi,\mathrm{v})$ is a real valued normal integrand $\Omega \times \mathbb{R}^{n+m} \to \mathbb{R}$, and U is some closed set of $\mathbb{R}^m$. We call a variable v the **control variable** and ξ the **state variable.** We use weighted Sobolev spaces as a state space, and the weight ν and the exponent p are supposed to be chosen in a way that $W_p^{1,n}(\Omega,\nu)$ forms a Banach space and such that for every element $x \in W_p^{1,n}(\Omega,\nu)$ there exists an absolutely continuous representative[1]. In following, we do not distinguish between elements from $x \in W_p^{1,n}(\Omega,\nu)$ and their absolute continuous representatives.

2. Preliminaries

Let us start with some definitions from [7] and [3]. Let X be a set from a Euclidean space of finite dimension, and let $\Omega \subseteq \mathbb{R}$ be an open set. Furthermore we utilize following conventions

$$\sup\emptyset = -\infty,\ \inf\emptyset = +\infty.$$

Moreover, we denote the convex hull and the closed convex hull of some set A by $\operatorname{co} A$ and $\overline{\operatorname{co}}\, A$ resp.

Definition 2.1 The function $g : \Omega \times X \to \overline{\mathbb{R}}$ is a **normal integrand** if

1. $g(t,\cdot) : X \to \overline{\mathbb{R}}$ is a l.s.c. function for a.a. $t \in \Omega$,
2. there exists a measurable function $\tilde{g} : \Omega \times X \to \overline{\mathbb{R}}$ such that $\tilde{g}(t,\cdot) = g(t,\cdot)$ for a.a. $t \in \Omega$.

Definition 2.2 The function $g : \Omega \times X \to \mathbb{R}$ is a **Carathéodory function** if

1. $g(t,\cdot) : X \to \mathbb{R}$ is a continuous function for a.a. $t \in \Omega$,
2. there exists a measurable function $\tilde{g} : \Omega \times X \to \mathbb{R}$ such that $\tilde{g}(t,\cdot) = g(t,\cdot)$ for a.a. $t \in \Omega$.

Lemma 2.3 *Let $g : \Omega\times(\mathbb{R}^n\times\mathbb{R}^l) \to \overline{\mathbb{R}}$, $(t,\xi,\mathrm{v}) \mapsto g(t,\xi,\mathrm{v})$ be some normal integrand and x some measurable mapping $\Omega \to \mathbb{R}^n$. Then the function $g \circ x$ defined as $g \circ x : (t,\mathrm{v}) \mapsto g(t,x(t),\mathrm{v})$ is a normal integrand on $\Omega \times \mathbb{R}^l$. In this sense we can identify*

Proof *Follows immediately from [7, Cor.2B].* □

Definition 2.4 Let $\Gamma : \Omega \to \mathcal{P}(X)$ be some set valued mapping. We call Γ **measurable** if for every closed set $A \subset X$ the set

$$\Gamma^{-1}(A) := \{t \in \Omega \,|\, \Gamma(t) \cap A \neq \emptyset\}$$

is measurable. We call Γ **closed-valued** if for every $t \in \Omega$ the set $\Gamma(t)$ is closed. Further we define $\operatorname{dom}\Gamma := \{t \in \Omega \,|\, \Gamma(t) \neq \emptyset\}$.

Lemma 2.5 *For a measurable closed valued multifunction $\Gamma : \Omega \to \mathcal{P}(\mathbb{R}^n)$ there exists at least one* ***measurable selection****, i.e. a function $u : \operatorname{dom}\Gamma \to \mathbb{R}^n$ with $u(t) \in \Gamma(t)$ for all $t \in \operatorname{dom}\Gamma$.*

Now we introduce the Γ-regularization (see [3, p.14]).

Definition 2.6 Let Y be a real convex space, and $g : Y \to \overline{\mathbb{R}}$. We call a pointwise supremum of continuous affine functions $Y \to \mathbb{R}$, that are everywhere less than g, a **Γ-regularization** g^{**} of g[2].

The Γ-regularization is always l.s.c. and convex, [3, Prop.3.1.].

Now we cite a sufficient condition for the invariance of a normal integrand property under Γ-regularization ([3, p.246, Prop.2.1]).

[1]See [5,6].
[2]As g^{**} we denote a bipolar of g, which for local convex spaces coincides with Γ-regularization, [3].

Lemma 2.7 *Let $g(t,\xi,\mathrm{v})$ be a normal integrand on $\Omega \times \mathbb{R}^{n+l}$ and satisfies $\Phi(\|\mathrm{v}\|) \leq g(t,\xi,\mathrm{v})$, where the function $\Phi : [0,\infty) \to \overline{\mathbb{R}}$ is convex, increasing, l.s.c. and fulfills $\lim_{z\to\infty} \frac{\Phi(z)}{z} = +\infty$. Then the Γ-regularization $g^{**}(t,\xi,\mathrm{v})$ is a normal integrand on $\Omega \times \mathbb{R}^{n+l}$ and satisfies $\Phi(\|\mathrm{v}\|) \leq g^{**}(t,\xi,\mathrm{v})$.*

Lemma 2.8 *The integrand $g(t,\xi,\mathrm{v})$ is normal iff $\chi_K(t)g(t,\xi,\mathrm{v})$ is normal for every $K \in \mathrm{comp}(\Omega)$.*

Proof *One direction of this statement is obvious. For the other one we remark that the supremum $g(t,\xi,\mathrm{v}) = \sup_{i\in J} g_i(t,\xi,\mathrm{v})$ over some countable family J of normal integrands is normal, [7, Prop.2L]. Since $\mathbb{R}$ is the union of countably many compact subsets the statement of the lemma follows immediately.* ☐

3. Equivalence of Γ-regularization and Convex combinations

Let us define a relaxation of a problem (P) in the sense of Γ-regularization (PRG) and in the sense of convex combinations (PRC).

$$\begin{aligned} &J_{(\mathrm{PRG})}(x) = \int_\Omega g^{**}(t,x(t),\dot{x}(t))\mathrm{d}t \to Min,\\ &g(t,\xi,\eta) = \inf\{r(t,\xi,\mathrm{v}) \mid \mathrm{v} \in U \subseteq \mathbb{R}^m, f(t,\xi,\mathrm{v}) = \eta\},\\ &\qquad x \in W_p^{1,n}(\Omega,\nu),\ x(t_0) = x_0. \end{aligned} \tag{PRG}$$

The Γ-regularization g^{**} is obtained from g resp. to variable η. It follows from the definition of g in (PRG) that the function $g(t,x(t),\dot{x}(t))$ takes the value $+\infty$ for every t with $\forall \mathrm{v} \in U : \dot{x}(t) \neq f(t,x(t),\mathrm{v})$. Thus, we know that for any admissible solution x, the set

$$\{t \in \Omega \mid \forall \mathrm{v} \in U : \dot{x}(t) \neq f(t,x(t),\mathrm{v})\}$$

forms a negligible set (set of measure zero).

$$\begin{aligned} &J_{(\mathrm{PRC})}(x,\lambda,u) = \int_\Omega \sum_{i=1}^{n+1} \lambda_i(t) r(t,x(t),u_i(t))\mathrm{d}t \to Min,\\ &\dot{x}(t) = \sum_{i=1}^{n+1} \lambda_i(t) f(t,x(t),u_i(t)) \text{ a.e. on } \Omega,\ x(t_0) = x_0,\\ &x \in W_p^{1,n}(\Omega,\nu),\\ &\lambda(t) \in E^n := \mathrm{co}\{e_1,\dots,e_{n+1}\} \text{ a.e. on } \Omega,\\ &u_i(t) \in U \subseteq \mathbb{R}^m \text{ a.e. on } \Omega,\\ &u_i, \lambda_i - \text{measurable for } i = 1 \dots n+1. \end{aligned} \tag{PRC}$$

Notice that the set E^n is n-dimensional, being the convex hull of $n+1$ points of dimension n.

From now on we define the function $\Psi(t,z) : \Omega \times [0,+\infty) \to \overline{\mathbb{R}}$ as a non-decreasing, convex l.s.c. in z function with the property

$$\lim_{z\to\infty} \frac{\Psi(t,z)}{z} = +\infty \text{ uniformly on every } K \in \mathrm{comp}(\Omega). \tag{C}$$

The integrand r satisfies a growth condition (G) if holds

$$\Psi(t,\|\mathrm{v}\|) \leq r(t,\xi,\mathrm{v}) \tag{G}$$

with Ψ satisfying (C).

Lemma 3.1 *We consider the problem* (PRG)*. Let the integrand $r(t,\xi,\mathrm{v})$ satisfy growth condition* (G)*. Let the function f be a Carathéodory-function, and U be a closed set. Then the functions $g(t,\xi,\eta)$ and its Γ-regularization $g^{**}(t,\xi,\eta)$ are normal integrands on $\Omega \times \mathbb{R}^{2n}$.*

Proof *Let K be some compact subset of Ω. We use a variant of Scorzà-Dragoni theorem for normal integrands, [3, Thm.1.1]. We show that*

$$\forall \varepsilon > 0\ \exists K_\varepsilon \subset K :\ |K \setminus K_\varepsilon| \leq \varepsilon \text{ and } g\big|_{K_\varepsilon \times \mathbb{R}^{2n}} \text{ l.s.c.} \tag{3.3}$$

Since r is a normal integrand we can establish condition (3.3) *for* $r(t,\xi,\mathrm{v})$ *restricted to* $K_\varepsilon \times \mathbb{R}^{n+m}$*, instead of* g*.*

We consider some sequence $\{(t_i,\xi_i,\eta_i)\} \subset K_\varepsilon \times \mathbb{R}^{2n}$ *converging to* $(\bar{t},\bar{\xi},\bar{\eta})$ *and show* $g(\bar{t},\bar{\xi},\bar{\eta}) \leq \varliminf_{i\to\infty} g(t_i,\xi_i,\eta_i)$*. We only need to show the inequality for the case that the limes inferior is a real number from* $[0,\infty)$*. We take a subsequence, that represents the limes inferior. For simplicity let the sequence be again* $\{(t_i,\xi_i,\eta_i)\}$ *and we have*

$$\lim_{i\to\infty} g(t_i,\xi_i,\eta_i) = \alpha < +\infty. \tag{3.4}$$

For sufficiently large indexes i we have $g(t_i,\xi_i,\eta_i) < +\infty$*, which means*

$$\left\{ r(t_i,\xi_i,\mathrm{v}) \,\middle|\, \mathrm{v} \in U, f(t_i,\xi_i,\mathrm{v}) = \eta_i \right\} \neq \emptyset.$$

For every (t_i,ξ_i) *the level sets of* $r(t_i,\xi_i,\cdot) : U \to \mathbb{R}$ *are compact since we have* $\Psi(t_i,\|\mathrm{v}\|) \leq r(t_i,\xi_i,\mathrm{v})$ *and the function* $\Psi(t,\|\mathrm{v}\|)$ *fulfills* (C)*. Since the function r is l.s.c. in* v *and the preimage* $f^{-1}(t,\xi,\cdot)(\eta_i)$ *is closed we obtain for every* (t_i,ξ_i,η_i) *a* $\mathrm{v}_i \in U$ *with*

$$g(t_i,\xi_i,\eta_i) = r(t_i,\xi_i,\mathrm{v}_i) \text{ and } f(t_i,\xi_i,\mathrm{v}_i) = \eta_i.$$

Again in view of (C) *we obtain that all of* v_i *lie in some compact subset of U, and finally we obtain a subsequence* $(t_{i_j},\xi_{i_j},\mathrm{v}_{i_j})$ *converging to* $(\bar{t},\bar{\xi},\bar{\mathrm{v}})$ *and in view of continuity of f and l.s.c. of r on* $K_\varepsilon \times \mathbb{R}^{n+m}$ *we have*

$$f(\bar{t},\bar{\xi},\bar{\mathrm{v}}) = \bar{\eta},$$
$$r(\bar{t},\bar{\xi},\bar{\mathrm{v}}) \leq \varliminf_{j\to\infty} r(t_{i_j},\xi_{i_j},\mathrm{v}_{i_j}).$$

From latter inequality and definition of g we obtain

$$g(\bar{t},\bar{\xi},\bar{\eta}) \leq r(\bar{t},\bar{\xi},\bar{\mathrm{v}}) \leq \varliminf_{j\to\infty} r(t_{i_j},\xi_{i_j},\mathrm{v}_{i_j}) = \varliminf_{j\to\infty} g(t_{i_j},\xi_{i_j},\eta_{i_j}) = \alpha.$$

The last limes inferior is equal to α *because of* (3.4)*. Thus, we obtain that* $g(t,\xi,\mathrm{v})$ *is a normal integrand on* $\Omega\times\mathbb{R}^{n+m}$*. Finally, lemma 2.8 together with lemma 2.7 deliver that* $g^{**}(t,\xi,\mathrm{v})$ *is a normal integrand on* $\Omega\times\mathbb{R}^{n+m}$ *as well.* ☐

Lemma 3.2 *Let the integrand r satisfy growth condition* (G)*. Moreover, let* $x : \Omega \to \mathbb{R}^n, y : \Omega \to \mathbb{R}^n$ *be measurable. Then there exist* $n+1$ *measurable functions* $y_i : \Omega \to \mathbb{R}^n, i = 1 \dots n+1$ *and* $\lambda : \Omega \to E^n$*, such that we have for almost every* $t \in \Omega$*:*

$$g^{**}(t,x(t),y(t)) = \sum_{i=1}^{n+1} \lambda_i(t) g(t,x(t),y_i(t)), \tag{3.5}$$
$$y(t) = \sum_{i=1}^{n+1} \lambda_i(t) y_i(t).$$

Proof *From lemma 3.1 follows that* $g^{**}(t,\xi,\eta)$ *is a normal integrand on* $\Omega \times \mathbb{R}^{2n}$ *and corollary [7, Cor.2B] delivers that* $g^{**}(t,x(t),\eta)$ *and* $g(t,x(t),\eta)$ *are both normal integrands on* $\Omega\times\mathbb{R}^n$*, and due to [3, Prop.3.1.] we obtain representation* (3.5)*.* ☐

Lemma 3.3 *Let x be an admissible solution of* (PRG)*, and the integrand r satisfy growth condition* (G)*. Then there exist functions* $u : \Omega \to U^{n+1}$ *and* $\lambda : \Omega \to E^n$ *such that the triple* (x,λ,u) *is admissible for* (PRC) *and* $J_{(\mathrm{PRG})}(x) = J_{(\mathrm{PRC})}(x,\lambda,u)$*.*

Proof *From lemma 3.2 we obtain measurable functions* $\lambda_i(t), y_i(t), i = 1 \dots n+1$*, which fulfill*

$$g^{**}(t,x(t),\dot{x}(t)) = \sum_{i=1}^{n+1} \lambda_i(t) g(t,x(t),y_i(t)),$$
$$\dot{x}(t) = \sum_{i=1}^{n+1} \lambda_i(t) y_i(t).$$

Now we need to define a proper selection u_i, for every function y_i, to fulfill the state equation

$$\dot{x}(t) = \sum_{i=1}^{n+1} \lambda_i(t) f(t, x(t), u_i(t)).$$

For every $y_i(t)$ we define a set valued mapping

$$\Gamma_i(t) := \{ \mathrm{v} \in U \,|\, r(t, x(t), \mathrm{v}) = g(t, x(t), y_i(t)) \}. \tag{3.6}$$

The function $g(t, x(t), y_i(t)) : \Omega \to \overline{\mathbb{R}}$ is measurable (lemma 2.3) and by [7, Thm.2]] we obtain that $\Gamma_i(t)$ are measurable set valued mappings with closed values, and for every t with $\Gamma_i(t) \neq \emptyset$ (follows from growth condition (G) *as in proof of lemma 3.1). That is the case for every $t \in \operatorname{dom} \Gamma_i$ because of coercivity of Ψ in z (see the proof of lemma 3.1). The set $\Omega \setminus \bigcap_{i=1}^{n+1} \operatorname{dom} \Gamma_i$ is negligible, because $\bar{x}$ is an admissible solution with $J(\bar{x}) < +\infty$. The same theorem [7, Thm.2]] delivers that there exists a measurable selection $u_i(t)$ for every i such that $u_i(t) \in \Gamma_i(t)$ and $y_i(t) = f(t, x(t), u_i(t))$ for all $t \in \operatorname{dom} \Gamma_i(t)$. And finally using* (3.6) *we get:*

$$g^{**}(t, x(t), \dot{x}(t)) = \sum_{i=1}^{n+1} \lambda_i(t) r(t, x(t), u_i(t)), \tag{3.7}$$

$$\dot{x}(t) = \sum_{i=1}^{n+1} \lambda_i(t) f(t, x(t), u_i(t))$$

for almost all $t \in \Omega$. The solution (x, λ, u) with $\lambda = (\lambda_1, \ldots, \lambda_{n+1}), u = (u_1, \ldots, u_{n+1})$ is then an admissible solution of (PRC) *and, because of* (3.7)*, we have $J_{(PRC)}(x, \lambda, u) = J_{(PRG)}$.* □

Lemma 3.4 *Let (x, λ, u) be an admissible solution of* (PRC) *and the integrand r satisfy the growth condition* (G)*. Then x is an admissible solution of* (PRG) *and $J_{(PRG)}(x) \leq J_{(PRC)}(x, \lambda, u)$.*

Proof *From the definition of function g in* (PRG) *we obtain*

$$g(t, x(t), f(t, x(t), u_i(t))) \leq r(t, x(t), u_i(t)) \text{ a.e.}$$

We make use of [3, Lemma 3.3.] and get

$$g^{**}(t, x(t), \dot{x}(t)) \leq \sum_{i=1}^{n+1} \lambda_i(t) g\big(t, x(t), f(t, x(t), u_i(t))\big) \leq \sum_{i=1}^{n+1} \lambda_i(t) r(t, x(t), u_i(t))$$

with $\lambda(t) \in E^n$ a.e. on Ω. □

Lemmas 3.4 and 3.3 imply immediately the equivalence of problems (PRC) and (PRG) in the sense of definition 1.1.

4. Equivalence of Young measures and Convex combinations

We first extend the notion of Young measure, as stated in [4], to unbounded domains Ω and sets U, which are closed, but not necessarily bounded.

Definition 4.1 We call a family of Radon measures[3] $\mu = \{\mu_t\}_{t \in \Omega}$ on U a **generalized control** and write $\mu \in \mathcal{M}_U$ if it fulfills:

i) $\operatorname{supp} \mu_t \subseteq U$ for almost all $t \in \Omega$,

ii) μ_t is a probability measure for almost all $t \in \Omega$,

iii) for every $g \in C_c(\Omega \times U)$ the function

$$h(t) = \langle \mu_t, g(t, \mathrm{v}) \rangle := \int_U g(t, \mathrm{v}) d\mu_t(\mathrm{v})$$

is measurable.

[3]For the theory of Radon measures we refer to [1].

Now we are ready to define a relaxation in the sense of Young measures (also known as Gamkrelidze controls):

$$\begin{aligned} J_{(\mathrm{PRY})}(x,\mu) &= \int_\Omega \langle \mu_t, r(t,x(t),\mathrm{v})\rangle \,\mathrm{d}t \;\longrightarrow\; Min, \\ \dot{x}(t) &= \langle \mu_t, f(t,x(t),\mathrm{v})\rangle \text{ a.e. on } \Omega,\; x(t_0)=x_0, \\ x &\in W_p^{1,n}(\Omega,\nu), \\ \mu &\in \mathcal{M}_U. \end{aligned} \tag{PRY}$$

For further explanations we need following two definitions of **orientor fields**

$$P(t,\xi) = \left\{ \begin{pmatrix} r(t,\xi,\mathrm{v}) \\ f(t,\xi,\mathrm{v}) \end{pmatrix} \;\Big|\; \mathrm{v}\in U \right\},$$

$$P_{\mathcal{M}}(t,\xi) = \left\{ \left\langle \hat{\mu}, \begin{pmatrix} r(t,\xi,\mathrm{v}) \\ f(t,\xi,\mathrm{v}) \end{pmatrix} \right\rangle \;\Big|\; \operatorname{supp}\hat{\mu}\subseteq U,\; \hat{\mu} \text{ – probability measure} \right\}.$$

The following lemma is a modification of [4, Assertion 2.1.].

Lemma 4.2 *Let U be some closed subset of $\mathbb{R}^m$, the function $g : U \to \mathbb{R}^n$ be continuous, and let $H^k \subset \mathbb{R}^n$ be some hyperplane of dimension k, where $1 \le k \le n$. Let the probability measure $\hat{\mu}$ on U be such that $\operatorname{supp}\hat{\mu} \subseteq g^{-1}(H^k)$. Further, let the point $p := \langle \hat{\mu}, g\rangle$ lie in H^k and not in $\operatorname{co} P$, where P represents the orientor field*

$$P := \{g(\mathrm{v}) \mid \mathrm{v}\in U\}.$$

Then there exists a hyperplane H^{k-1} of dimension $k-1$, such that $p \in H^{k-1}$ and $\operatorname{supp}\hat{\mu} \subseteq g^{-1}(H^{k-1})$.

Proof *Since $g^{-1}(H^k)$ contains a support of the probability measure it is not empty. We conclude that $\operatorname{co} P \cap H^k$ is convex and not empty as well. We define a $k-1$-dimensional hyperplane, denoted by $H^{k-1} \subset H^k$, that separates the point p and the set $\operatorname{co} P \cap H^k$. Furthermore, p lies in H^{k-1}.*

Let $\mathcal{X}(\mathrm{v})$ be the characteristic function of the preimage $g^{-1}(H^k)$:

$$\mathcal{X}(\mathrm{v}) := \begin{cases} 1, & g(\mathrm{v}) \in H^{k-1} \\ 0, & g(\mathrm{v}) \notin H^{k-1} \end{cases}.$$

The preimage $g^{-1}(H^k)$ is closed, as it is the preimage of a closed set under continuous mapping. Consequently, the function $\mathcal{X} : U \to \mathbb{R}$ is u.s.c.

We consider the equation

$$\langle \hat{\mu}, g(\mathrm{v}) - p\rangle = 0$$

from which we deduce

$$\langle \hat{\mu}, g(\mathrm{v}) - p\rangle = \langle \hat{\mu}, \mathcal{X}(\mathrm{v})(g(\mathrm{v}) - p)\rangle + \langle \hat{\mu}, (1-\mathcal{X}(\mathrm{v}))(g(\mathrm{v}) - p)\rangle = 0.$$

Let $w \in H^k$ be a vector orthogonal to H^{k-1}, and directed towards $\operatorname{co} P(t,x) \cap H^k$. By taking a scalar product with the above equation we obtain

$$\left\langle \hat{\mu}, \mathcal{X}(\mathrm{v}) w^T (g(\mathrm{v}) - p)\right\rangle + \left\langle \hat{\mu}, (1-\mathcal{X}(\mathrm{v})) w^T (g(\mathrm{v}) - p)\right\rangle = 0. \tag{4.2}$$

For all v with $\mathcal{X}(\mathrm{v}) = 1$, the scalar product $w^T(g(\mathrm{v}) - p)$ vanishes because the points $g(\mathrm{v})$ and p lie in the hyperplane H^{k-1}, and the vector w is then orthogonal to $g(\mathrm{v}) - p$. It follows

$$\forall \mathrm{v} \in U \;\; \mathcal{X}(\mathrm{v}) w^T (g(\mathrm{v}) - p) = 0,$$

and together with (4.2) we conclude

$$\left\langle \hat{\mu}, (1-\mathcal{X}(\mathrm{v})) w^T (g(\mathrm{v}) - p)\right\rangle = 0. \tag{4.3}$$

Since $g(\mathrm{v}) \in \operatorname{co} P$, and w is directed toward $\operatorname{co} P \cap H^k$, for any $\mathrm{v} \in g^{-1}(H^k \setminus H^{k-1})$ we obtain

$$w^T (g(\mathrm{v}) - p) > 0.$$

As for v *from* $g^{-1}(H^k \setminus H^{k-1})$ *the indicator function* $\mathcal{X}$ *is equal zero we conclude*

$$\forall \mathrm{v} \in U: \ g(\mathrm{v}) \in H^k \setminus H^{k-1} \Rightarrow (1-\mathcal{X}(\mathrm{v}))w^T(g(\mathrm{v})-p) > 0. \tag{4.4}$$

Now from equation (4.3) *we become*

$$\int_U (1-\mathcal{X}(\mathrm{v}))w^T(g(\mathrm{v})-p)\mathrm{d}\hat{\mu} = \int_{g^{-1}(H^k)} (1-\mathcal{X}(\mathrm{v}))w^T(g(\mathrm{v})-p)\mathrm{d}\hat{\mu}.$$

Since $1-\mathcal{X}(\mathrm{v})$ *is l.s.c., and* $w^T(g(\mathrm{v})-p)$ *is non-negative and continuous on* $g^{-1}(H^k)$*, we deduce that the function* $\mathrm{v} \mapsto (1-\mathcal{X}(\mathrm{v}))w^T(g(\mathrm{v})-p)$ *is l.s.c. on* $g^{-1}(H^k)$*. Now we use a proposition [1, Ch.IV, §2(1), Prop.3] and conclude that the integrand* $(1-\mathcal{X}(\mathrm{v}))w^T(g(\mathrm{v})-p)$ *vanishes on* $\operatorname{supp}\hat{\mu}$*. Now, from inequality* (4.4) *it follows that* $\operatorname{supp}\hat{\mu} \cap g^{-1}(H^k \setminus H^{k-1}) = \emptyset$*. As we assumed that* $\operatorname{supp}\hat{\mu}$ *lies in* $g^{-1}(H^k)$*, we get*

$$\operatorname{supp}\hat{\mu} \subseteq g^{-1}(H^{k-1}).$$

□

Lemma 4.3 *Let* r, f *be Carathéodory functions on* $\Omega \times \mathbb{R}^{n+m}$*, then* $\operatorname{co} P(t,\xi) = P_{\mathcal{M}}(t,\xi)$ *for almost all* $t \in \Omega$*.*

Proof *Let* (t,ξ) *be arbitrary pair from* $\Omega \times \mathbb{R}^n$ *such that* $\big(r(t,\xi,\cdot), f(t,\xi,\cdot)\big)^T : U \to \mathbb{R}^{n+1}$ *is continuous. The inclusion* $P(t,\xi) \subseteq P_{\mathcal{M}}(t,\xi)$ *is obvious. Let's show* $P_{\mathcal{M}}(t,\xi) \subseteq P(t,\xi)$*.*

Let g *be a continuous function* $g : \mathrm{v} \mapsto \big(r(t,\xi,\mathrm{v}), f(t,\xi,\mathrm{v})\big)^T$*. We assume that there exists some probability measure* $\hat{\mu}$ *with* $\langle \hat{\mu}, g\rangle = p \notin \operatorname{co} P(t,\xi)$*. Using lemma 4.2 with the settings* $k := n+1, H^k := \mathbb{R}^{n+1}$ *we obtain* $\operatorname{supp}\hat{\mu} \subseteq g^{-1}(H^n)$*, where* H^n *is a hyperplane of dimension n, contains the point p, and lies in* H^{n+1}*.*

We now set $k := n$ *and utilize the lemma 4.2 once again. After altogether* $n+1$ *repetitions we obtain that p lies in the hyperplane* H^0 *of dimension zero, and* $\operatorname{supp}\hat{\mu} \subseteq g^{-1}(H^0)$*. Since* $p \in H^0$ *and* $\dim H^0 = 0$ *we obtain* $H^0 = \{p\}$ *and* $g^{-1}(H^0) = g^{-1}(p)$*.*

The measure $\hat{\mu}$ *is a probability measure which implies* $\operatorname{supp}\hat{\mu} \neq \emptyset$*. Together with* $\operatorname{supp}\hat{\mu} \subseteq g^{-1}(p)$ *we obtain* $g^{-1}(p) \neq \emptyset$*, that means that there exists* $\mathrm{v} \in U$ *with* $P(t,\xi) \supseteq g(\mathrm{v}) = p$*, and we get a contradiction.*

□

Lemma 4.4 *Let* r, f *be Carathéodory functions on* $\Omega \times \mathbb{R}^{n+m}$ *and* (x,μ) *be an admissible solution of* (PRY)*. Then there exists an admissible solution* (x,λ,u) *of* (PRC) *with* $J_{(\mathrm{PRC})}(x,\lambda,u) \leq J_{(\mathrm{PRY})}(x,\mu)$*.*

Proof *Let us define a vector-valued function* $g : \Omega \times \mathbb{R}^m \to \mathbb{R}^{1+n}$*,* $g : (t,\mathrm{v}) \mapsto (r(t,x(t),\mathrm{v}), f(t,x(t),\mathrm{v}))^T$*. We now use lemma 4.3, and for almost all* $t \in \Omega$ *we obtain*

$$\langle \mu, g(t,\mathrm{v})\rangle = \sum_{i=1}^{n+2} \hat{\lambda}_i g(t,u_i), \quad \hat{\lambda} \in E^{n+1}, \ u_{1,\dots,n+2} \in U. \tag{4.5}$$

Now we prove that we can diminish the dimension of E^{n+1}*. We formulate following optimization problem:*

$$\begin{aligned} & c^T\tilde{\lambda} \to Min, \\ s.t. \quad & A\tilde{\lambda} = d, \\ & \tilde{\lambda} \in E^{n+1}, \end{aligned}$$

where

$$c := \begin{pmatrix} r(t,x(t),u_1) \\ \dots \\ r(t,x(t),u_{n+2}) \end{pmatrix}, \ A := \big(f(t,x(t),u_1), \dots, f(t,x(t),u_{n+2})\big), \ d := A\hat{\lambda}, \tag{4.6}$$

with v_i *and* $\hat{\lambda}$ *from* (4.5)*. Since c and* $\tilde{\lambda}$ *are non-negative, there exists an optimal solution* $\tilde{\lambda}^*$ *of* (4.6)*. The constraints of* (4.6) *define a convex polyhedron, therefore* $\tilde{\lambda}^*$ *lies on its boundary. It means, that there exists at least one index* $1 \leq k \leq n+2$ *with* $\tilde{\lambda}^*_k = 0$*, and it follows* $(\tilde{\lambda}^*_{i=1\dots,n+2,i\neq k}) \in E^n$*. We then obtain*

$$\textit{for a.a. } t \in \Omega \, \exists \lambda \in E^n \quad \sum_{i=1}^{n+1} \lambda_i g(t,u_i) \leq \sum_{i=1}^{n+2} \hat{\lambda}_i g(t,u_i) = \langle \mu, g(t,\mathrm{v})\rangle. \tag{4.7}$$

Now we define the set-valued mapping

$$\Gamma(t) = \{(\lambda, u) \in E^n \times U^{n+1} \mid F(t,\lambda,u) = \langle \mu, f(t,x(t),\mathrm{v})\rangle, \\ F_1(t,\lambda,u) \le \langle \mu, r(t,x(t),\mathrm{v})\rangle\}$$

with $F(t,\lambda,u) = \sum_{i=1}^{n+1} \lambda_i f(t,x(t),u_i)$ *and* $F_1(t,\lambda,u) = \sum_{i=1}^{n+1} \lambda_i r(t,x(t),u_i)$. *F and* F_1 *are Carathéodory functions. The sets* $\Gamma(t)$ *are not empty for a.a.* $t \in \Omega$ *because of (4.7). Theorem [7, Thm.2J] delivers that* Γ *is measurable, and by lemma 2.5 we get functions*

$$\lambda : \Omega \to E^n, \\ u_i : \Omega \to U,\ i = 1,\dots,n+1$$

that are measurable and $(\lambda(t), u(t)) \in \Gamma(t)$ *for a.a.* $t \in \Omega$. *Finally, we obtain* $J_{(\mathrm{PRC})}(x,\lambda,u) \le J_{(\mathrm{PRY})}(x,\mu)$. ☐

Lemma 4.5 *Let r be a normal integrand and f be a Carathéodory function on* $\Omega \times \mathbb{R}^{n+m}$, *and let* (x,λ,u) *be an admissible solution of* (PRC). *Then, there exists an admissible solution* (x,μ) *of* (PRY) *such that* $J_{(\mathrm{PRY})}(x,\mu) = J_{(\mathrm{PRC})}(x,\lambda,u)$.

Proof *The proof is straightforward: define* $\mu_t := \sum_{i=1}^{n+1} \lambda_i(t)\delta_{u_i(t)}$, *and it can be readily shown that* $\mu := \{\mu_t\}_{t\in\Omega}$ *constitutes a generalized control according to definition 4.1.* ☐

Now, under the more restrictive conditions of lemma 4.4 we obtain the equivalence of problems (PRY) and (PRC).

5. Example

We will now provide an example to illustrate how the relaxations differ.

$$\begin{aligned} J(x,u) &= \int_0^\infty [e^{-u^2(t)} + x^2(t)]e^{-t}\mathrm{d}t \longrightarrow Min, \\ \dot{x}(t) &= \frac{1}{1+u^2(t)},\ \text{a.e. on } (0,\infty),\ x(0) = 0, \\ & x \in W_2^1((0,\infty), e^{-t}), \\ & u(t) \in \mathbb{R} \text{ a.e. on } (0,\infty), \\ & u - \text{measurable.} \end{aligned} \tag{PEX}$$

To get the Γ-regularization we first calculate the function g according to (PRG).

$$g(t,\xi,\eta) = \inf\left\{(e^{-\mathrm{v}^2} + \xi^2)e^{-t} \mid \mathrm{v} \in \mathbb{R}, \frac{1}{1+\mathrm{v}^2} = \eta\right\} = \begin{cases} +\infty, & \eta \le 0, \\ (e^{1-\frac{1}{\eta}} + \xi^2)e^{-t}, & \eta > 0. \end{cases}$$

Now we can easily calculate the Γ-regularized function according to the definition 2.6:

$$g^{**}(t,\xi,\eta) = \begin{cases} +\infty, & \eta < 0, \\ \xi^2 e^{-t}, & \eta \ge 0. \end{cases} \tag{5.1}$$

We insert this function, g^{**}, into the formulation (PRG) and conclude that the problem

$$J_{(\mathrm{PRG})}(x) = \int_0^\infty g^{**}(t,x(t),\dot{x}(t))\mathrm{d}t \to Min, \\ x \in W_2^1((0,\infty), e^{-t}),\ x(0) = 0,$$

where the function g^{**} is taken from (5.1), possesses an optimal solution $x^* \equiv 0$ with $J_{(\mathrm{PRG})}(x^*) = 0$.

On the other hand, since the integrand $r(t,\xi,\mathrm{v}) = (e^{-\mathrm{v}^2} + \xi^2)e^{-t}$ is always greater than zero, for any probability measure $\hat{\mu}$, we obtain $\langle \hat{\mu}, (e^{-\mathrm{v}^2} + \xi^2)e^{-t}\rangle > 0$. This implies that for any generalized control μ, we have

$$J_{(\mathrm{PRY})}(x,\mu) = \int_0^\infty \langle \mu_t, e^{-\mathrm{v}^2} + x^2(t)\rangle e^{-t}\mathrm{d}t > 0.$$

At the same time, the sequence of generalized controls[4] $\mu_k := \{\delta_{kt}\}_{t\in\Omega}$ and corresponding solutions $x_k(t) := \frac{1}{k}\arctan(kt)$ of the initial value problem of (PEX) form a null sequence $J_{(\mathrm{PRY})}(x_k, \mu_k)$

$$J_{(\mathrm{PRY})}(x_k, \mu_k) = \int_0^\infty \langle \delta_{kt}, e^{-\mathrm{v}^2} + x_k^2(t)\rangle e^{-t}\mathrm{d}t = \int_0^\infty \left(e^{-k^2t^2} + \frac{1}{k^2}\arctan^2(kt)\right)e^{-t}\mathrm{d}t < \frac{\sqrt{\pi}}{2k} + \frac{\pi^2}{4k^2} \overset{k\to\infty}{\longrightarrow} 0.$$

We conclude that there is no optimal solution for either the relaxations of the type of Young measures or the convex combinations, according to lemma 4.4. Furthermore, because the condition (G) cannot be satisfied, we are unable to extract any admissible solutions for other types of relaxations discussed here from Γ-regularization.

Acknowledgments

The authors were supported by DFG Grants PI209/8-3 and LY149/2-3.

References

[1] N. Bourbaki. *Integration I: Chapters 1-6*. "Springer", 2004.

[2] D.A. Carlson. Nonconvex and relaxed infinite – horizon optimal control problems. *Journal of optimization theory and applications*, 78(3):465–491, 1993.

[3] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems*. SIAM, 1999.

[4] R.V. Gamkrelidze. *Principles of Optimal Control Theory*. Plenum Press, New York and London, 1978.

[5] A. Kufner. *Weighted Sobolev Spaces*. John Wiley & Sons, 1985.

[6] A. Kufner and B. Opic. How to define reasonably weighted sobolev spaces. *Commentationes Mathematicae Universitatis Carolinae*, 025(3):537–554, 1984.

[7] R. Tyrrell Rockafellar. Integral functionals, normal integrands and measurable selections. In J.P. Gossez, E.J. Lami Dozo, J. Mawhin, and L. Waelbroeck, editors, *Lecture Notes in Mathematics: Nonlinear Operators and Calculus of Variations*, volume 543, pages 157–207. Springer, 1985.

[8] T. Roubíček. *Relaxation in Optimization Theory and Variational Calculus*. De Gruyter, Berlin, Boston, 2020.

[4]We denote by δ_x a Dirac-measure, concentrated at the point x.