RELAXATION OF OPTIMAL CONTROL PROBLEMS WITH SHIFTED CONTROLS

EUGENE STEPANOV

Abstract. The paper studies the representation of relaxed settings of optimal control problems for nonlocal differential equations with deviating argument involving possibly many shifts in control variables. As opposed to the more general approach of J. Rosenblueth, J. Warga, R. Vinter and Q. Zhu, the relaxation is accomplished in original function spaces rather than in much larger spaces of special Young measures. To find relaxed setting, one uses the straightforward convexification method similar to that used in optimal control problems for ordinary (local) differential equations. It is shown that such an approach is essentially limited, since it can work only in the cases when the shifts in control variables satisfy the special "unifiability" condition, which can be regarded as some commensurability or "collective nonergodicity" property. Slightly restricting further the unifiability requirement one is able to obtain the exact representation of relaxed control problem setting. The latter will be again an optimal control problem but involving, generally speaking, infinitely many new shifts both in the state equation and in the cost functional.

1. Introduction

Consider the following sample optimal control problem for a differential equation with deviating argument

$$\inf_{y\in Y,\, u\in U}\left\{\int_\Omega \left(\Psi(t,y(t))+\frac{|u(t)|^2}{2}\right)dt\ :\ \begin{array}{l}\dot y(t)=a(t,y(h_1(t)))+b(t,y(h_2(t)))\psi(t,u(g(t))),\\ y(t_0)=y_0\end{array}\right\},\tag{1}$$

where t is time, y is the state variable and u is control, $\Omega\subset\mathbb{R}$ is assumed to be open and bounded (in most applications it is just an interval (t_0,t_1)), h_1, h_2 and g: $\Omega\to\Omega$ are given measurable functions called *shifts*, while g is also supposed to satisfy

$$|e|=0\Rightarrow \mu_g(e):=|g^{-1}(e)|=0,\tag{2}$$

$|\cdot|$ standing for the Lebesgue measure, and at last Ψ, a, b, ψ: $\Omega\times\mathbb{R}\to\bar{\mathbb{R}}\equiv\mathbb{R}\cup\{+\infty\}$ are Carathèodory functions. The space of states $Y=W^{1,1}(\Omega)$, and the space of controls $U=L^p(\Omega)$ are Sobolev and Lebesgue spaces respectively equipped with their weak topologies. The scope of this paper is to give the exact representation of the relaxed setting for the above problem as well as for more general ones involving differential equation with deviating argument in the control variable. It is in fact well known that the main difficulty arises due to the presence of argument deviations (shifts) in control variables. This paper studies such problems which can involve

Date: July 6, 2000.
1991 *Mathematics Subject Classification.* 49J45; 47B38, 47H30.

many different argument shifts in controls, and hence are even more general than the above simple example.

The general approach to such problems is developed in the recent series of papers of J. Rosenblueth, J. Warga, R. Vinter and Q. Zhu (see [15] and references therein), and is based on looking for relaxation in the space of special Young measures. This approach allows to obtain rather easily the relaxed setting, but in quite an abstract form which is impossible to apply unless one is able to characterize the set of relaxed Young measures [15]. Thus under this approach the difficulty of solving the original problem (i.e. finding the relaxation) is practically reduced to finding a manageable characterization of the latter set, which is quite a complicated problem on its own. Another, though minor, disadvantage of this approach is that it deals normally with relaxation not in the whole space of controls but only in its subset consisting of functions bounded in L^∞ norm (unless one develops for this purpose some modification of classical Young measures, say, Di Perna–Majda Young measures [11, Chapter 3]).

In this paper a much less general though straightforward approach to relaxation is developed, which deals with relaxation directly in the original spaces without passing through the larger spaces of special Young measures. It stems from the classical convexification method like that used in [3] to deal with optimal control problems for ordinary differential equations. Namely, by introducing an auxiliary variable, the original optimal control problem is reduced to a purely variational setting. Due to the presence of argument shifts in controls, this setting cannot however be reduced to a so-called "local" integral functional like it is done in the case of optimal control problems for ODE's. Nevertheless, under some special condition on the shifts involved, called further *unifiability*, which can be regarded as some commensurability or "collective nonergodicity" property, it is still possible to find the relaxation of the latter variational setting by means of convexification technique similar to that one routinely uses in the case of local functionals (and hence also for optimal control problems for differential equations involving only local operators, like ODE's). Note that a vaguely similar approach based on convexification has already beeen employed by J. Warga in section VI.5.5 of [14] but for very particular (so called "composition power") shifts. The approach used in this paper is fairly general and, moreover, we will be able to explore its natural limits. In fact, it will be shown that the unifiability condition on shifts in control variables in general is both necessary and sufficient for the convexification method to be applicable.

Under some further restriction of the unifiability condition on shifts we will be able to find the exact representation of the relaxed setting for the control problem studied, which in general is, roughly speaking, again an optimal control problem but with both the cost functional and state equation involving possibly a countable number of different shifts in control variables. The "state equation" constraint will be a differential inclusion, but we find it more convenient to write it as a couple of equations. The proof of the respective result, though being very technical, provides the construction of the respective relaxed setting which involves a series of operations of convexification (some done in infinite dimensional spaces) and of taking conditional expectation. Thus the approach developed in this paper is constructive as opposed to that of [15] based on relaxed Young measures, though clearly less general than the latter.

2. Notation and preliminaries

Let (X, τ) be a topological space. The relaxation of the functional $I\colon X \to \bar{\mathbb{R}}$ in the topology τ defined as the maximum τ-lower semicontinuous functional over X not greater than I will be denoted by $sc^{-}(\tau)I$.

Let (X, X^*) be dual pair of locally convex topological spaces with the pairing denoted by $< \cdot, \cdot >$. The Fenchel conjugate of a functional $I\colon X \to \bar{\mathbb{R}}$ is a new functional $I^*\colon X^* \to \bar{\mathbb{R}}$ defined by the formula

$$I^*(u') = \sup_{u \in X} < u, u' > -I(u).$$

The second Fenchel conjugate $I^{**}\colon X \to \bar{\mathbb{R}}$ is given by

$$I^{**}(u) = \sup_{u' \in X^*} < u, u' > -I^*(u').$$

It is well-know that I^{**} is the convexification of I in the sense that it is the maximum (strongly) lower semicontinuous and convex functional not greater than I.

Let $(\Omega, \Sigma(\Omega), \mu_\Omega)$ be a standard measure space (i.e. Ω Polish space, μ_Ω be a Borel finite measure), while $\Sigma(\Omega)$ stand for the completion of Borel σ-algebra $\mathcal{B}(\Omega)$ with respect to μ_Ω. Here and in the sequel $\mathcal{B}(X)$ will stand for the Borel σ-algebra of the metric space X. If Ω is a Borel subset of $\mathbb{R}^n$, then we will always silently assume μ_Ω to be the respective Lebesgue measure and $\Sigma(\Omega)$ to be the Lebesgue σ-algebra. If $(\Omega, \Sigma(\Omega), \mu_\Omega)$ and $(\Theta, \mathcal{B}(\Theta), \mu_\Theta)$ are measure spaces and $g\colon \Omega \to \Theta$ is a measurable map, then we will refer to functions measurable with respect to $g^{-1}(\mathcal{B}(\Theta))$ as to g-measurable functions.

An *integrand* is a $(\Sigma(\Omega) \otimes \mathcal{B}(X))$-measurable function $f\colon \Omega \times X \to \bar{\mathbb{R}}$, where X is a normed space. An integrand $f\colon \Omega \times X \to \bar{\mathbb{R}}$ is called *normal*, if $f(x, \cdot)$ is lower semicontinuous for a.e. $x \in \Omega$. The Fenchel conjugate of an integrand $f\colon \Omega \times X \to \bar{\mathbb{R}}$ is the integrand $f^*\colon \Omega \times X^* \to \bar{\mathbb{R}}$ defined by

$$f^*(x, y') = \sup_{y \in X} < y, y' > -f(x, y).$$

The second conjugate is given by $f^{**}(x, y) = (f^*)^*(x, y)$.

Also, we will often use the following construction. Let $D = \{D(x)\}_{x \in \Omega}$ be a family of normed spaces. We will say that a function $u\colon \Omega \to \bigcup_{x \in \Omega} D(x)$ acts into the family D, if for all $x \in \Omega$ one has $u(x) \in D(x)$. We also define a function of two variables, defined over the family D. For this purpose we will identify the family $D = \{D(x)\}_{x \in \Omega}$ with the set

$$D = \{(x, y) : x \in \Omega, y \in D(x)\}.$$

and will say that a function $f : D \to \bar{\mathbb{R}}$ is defined over the family D. The conjugate of a function $f\colon D \to \bar{\mathbb{R}}$ acting over the family D can be defined in an analogous way:

$$f^*(x, y') = \sup_{y \in D(x)} < y, y' > -f(x, y).$$

The second conjugate is again given by $f^{**}(x, y) = (f^*)^*(x, y)$.

Let us now introduce some function spaces which will appear in the sequel. Given a vector $\bar{p} = \{p_i\}_{i=1}^k$, where $k \in \mathbb{N}$ and $1 \leq p_i \leq +\infty$, we denote

$$\begin{array}{rcl} l_k^{\bar{p}} & := & \prod_{i=1}^k l^{p_i}, \\ L_k^{\bar{p}}(\Omega) & := & \prod_{i=1}^k L^{p_i}(\Omega), \\ L_k^{\bar{p}}(\Omega; \prod_{i=1}^k X_i) & := & \prod_{i=1}^k L^{p_i}(\Omega; X_i), \end{array}$$

where all X_i are normed spaces and $L^{p_i}(\Omega, X_i)$ are the respective Lebesgue spaces of X_i-valued functions. Further, given a sequence of $\mathcal{N}_i \subset \mathbb{N}$, $i = 1, \dots, k$ and of weight coefficients $\{w_i^\nu\}_{i,\nu=1}^{k,\infty} \subset [0, +\infty]$ we denote:

$$\begin{aligned} l_k^{\bar p}[\{\mathcal{N}_i\}_{i=1}^k] &:= \left\{ y \in \prod_{i=1}^k \mathbb{R}^{\mathcal{N}_i} \, : \, \sum_{\nu\in\mathcal{N}_i} |y_i^\nu|^{p_i} < +\infty, \ i = 1, \dots, k \right\}, \\ l_k^{\bar p}[\{w_i^\nu\}_{i=1,\nu\in\mathcal{N}_i}^k] &:= \left\{ y \in \prod_{i=1}^k \mathbb{R}^{\mathcal{N}_i} \, : \, \sum_{\nu\in\mathcal{N}_i} |w_i^\nu y_i^\nu|^{p_i} < +\infty, \ i = 1, \dots, k \right\}. \end{aligned}$$

Note that in the latter spaces, if a weight coefficient $w_i^\nu = \infty$, then necessarily $y_i^\nu = 0$, while if $w_i^\nu(s) = 0$, then the coordinate y_i^ν may be arbitrary independently of the other coordinates. The duals of the introduced spaces are well-known, namely,

$$\begin{aligned} &(l_k^{\bar p})^* = l_k^{\bar p'}, \ (l_k^{\bar p}[\{\mathcal{N}_i\}_{i=1}^k])^* = l_k^{\bar p'}[\{\mathcal{N}_i\}_{i=1}^k] \text{ and} \\ &(l_k^{\bar p}[\{w_i^\nu\}_{i=1,\dots,k,\,\nu\in\mathcal{N}_i}])^* = l_k^{\bar p'}[\{1/w_i^\nu\}_{i=1,\dots,k,\,\nu\in\mathcal{N}_i}] \end{aligned}$$

where $\bar p' := \{p_i'\}_{i=1}^k$, and $1/p_i + 1/p_i' = 1$ for all $i = 1, \dots, k$. The pairings applied here are, respectively,

$$< y, y' > = \sum_{i=1}^k \sum_{\nu=1}^\infty y_i^\nu y_i'^\nu \text{ and } < y, y' >_* = \sum_{i=1}^k \sum_{\nu\in\mathcal{N}_i} y_i^\nu y_i'^\nu,$$

the first one for $l_k^{\bar p}$ and the latter one for $l_k^{\bar p}[\{\mathcal{N}_i\}_{i=1}^k]$ and $l_k^{\bar p}[\{w_i^\nu\}_{i=1,\dots,k,\,\nu\in\mathcal{N}_i}]$. At last, we recall that one always has $(L_k^{\bar p}(\Omega))^* = L_k^{p'}(\Omega)$, while

$$(L_k^{\bar p}(\Omega; \prod_{i=1}^k X_i))^* = L_k^{p'}(\Omega; \prod_{i=1}^k X_i^*)$$

only when all X_i^* possess the Radon-Nikodym property (see Theorem IV.1.1 in [6, p. 98]). The pairing applied for these function spaces is $[u', u] := \int_\Omega < u'(x), u(x) > d\mu_\Omega(x)$.

3. Unifiability of function matrices

Here we introduce and discuss the notion of unifiablity for function matrices, which is of utmost importance for our further results.

Let $(\Omega, \Sigma(\Omega), \mu_\Omega)$ and $(\Theta, \Sigma(\Theta), \mu_\Theta)$ be standard measure spaces, the latter with nonatomic measure. In applications, it is usually enough to consider Ω and Θ to be open bounded subsets of $\mathbb{R}^n$ equipped with Lebesgue σ-algebrae, the respective measures being just the Lebesgue measure. The functions $g\colon \Omega \to \Theta$ generating "argument shifts" will always be assumed to satisfy

$$\mu_\Theta(e) = 0 \Rightarrow \mu_g(e) := \mu_\Omega(g^{-1}(e)) = 0. \tag{3}$$

In fact, this condition is necessary and sufficient to define correctly the compositions $u \circ g$ for u being a class of measurable functions.

First of all consider the auxiliary notion to be used in the sequel, which was first introduced in [7].

Definition 3.1. *We say that a measurable function $g\colon \Omega \to \Theta$ satisfying (3), satisfies ω-condition, if there exists a disjoint at most countable covering of Ω by measurable sets $\Omega = \sqcup_j \Omega_j$ such that over each Ω_j the function is injective, and the respective inverses $\gamma_j\colon g(\Omega_j) \to \Omega_j$, $\gamma_j \circ g|_{\Omega_j} = id$, satisfy (3).*

Now we pass to the principal new definition of this section. Consider the the $k \times l$ matrix of measurable functions $\{g_{ij}\}_{i,j=1}^{k,l}$: $\Omega \to \Theta$, all of which are assumed to satisfy (3).

Definition 3.2. *A vector row of measurable (with respect to $\Sigma(\Theta)$) functions*

$$\gamma = (\gamma_1, \ldots, \gamma_k), \qquad \gamma_i : \Theta \to (0,1),\ i = 1, \ldots, k$$

is said to be unifier *of the function matrix $\{g_{ij}\}_{i,j=1}^{k,l}$, if all the functions $\gamma_i \circ g_{ij}$: $\Omega \to \mathbb{R}$ coincide μ_Ω-a.e. over Ω.*

Definition 3.3. *The function matrix $\{g_{ij}\}_{i,j=1}^{k,l}$ is called unifiable, if it admits such a unifier $(\gamma_1, \ldots, \gamma_l)$ that*

$$\mu_\Theta(\gamma_i^{-1}(y)) = 0, \qquad i = 1, \ldots, k \tag{4}$$

for all $y \in (0,1)$. Moreover, if the unifier $(\gamma_1, \ldots, \gamma_l)$ can be chosen so that each γ_i satisfy the ω-condition, then the function matrix is called ω-unifiable.

Note that it is not restrictive to suppose that the components of the unifier act in Ω rather than in $\mathbb{R}$, since by the theorem on isomorphism of measure spaces there always exists a measure preserving bijection j: $(0,1) \to \Omega$, with the help of which any unifier h: $\Omega \to (0,1)$ can be turned into the new unifier $j \circ h$: $\Omega \to \Omega$. The ω-condition is clearly preserved by this operation. In the sequel we will mainly deal with even a slightly simpler situation when $\Omega = \Theta$.

To understand better the introduced notion of unifiablity, let us study it in particular cases of matrices consisting of either a single row or a single column of functions. Namely,

- the vector row of functions $(g_1, \ldots, g_l)$: $\Omega \to \Theta$ is unifiable, if there is a measurable function γ: $\Theta \to (0,1)$ satisfying
$$\mu_\Theta(\gamma^{-1}(y)) = 0, \qquad i = 1, \ldots, k \tag{5}$$
for all $y \in (0,1)$, and such that
$$\gamma(g_1(x)) = \ldots = \gamma(g_l(x))$$
holds for μ_Ω-a.e. $x \in \Omega$.
- the vector column of functions $(g_1, \ldots, g_k)^T$: $\Omega \to \Theta$ (T stands for transposition, in order to distinguish between rows and columns) is unifiable, if there is a vector of measurable functions $(\gamma_1, \ldots, \gamma_k)$: $\Theta \to (0,1)$, all satisfying (5) such that
$$\gamma_1(g_1(x)) = \ldots = \gamma_k(g_k(x))$$
for μ_Ω-a.e. $x \in \Omega$.

We first note that the unifiablity condition does not put a very strong restriction on vector columns of functions. In fact, for a vector column to be unifiable it is sufficient, for instance, that its components be injective. To be more precise, the following statement holds.

Proposition 3.1. *A vector column of injective functions $(g_1, \ldots, g_k)^T$: $\Omega \to \Theta$ satisfying (3) is ω-unifiable.*

PROOF: Up to restricting the functions g_i to the Borel subset of full measure in Ω, we may assume that the respective inverses g_i^{-1} satisfy (3). Hence it is enough to take $\gamma_i := g_i^{-1}$, $i = 1, \ldots, k$. $\square$

However, the unifiablity condition appears to be much stronger for vector rows. Obviously, if a vector row is unifiable, then its transpose is unifiable too (but now as a column). In fact, although every vector column of injective functions is unifiable, not every vector row of injective functions is so, as the following example shows.

Example 3.1. Let $\Omega = \Theta$ and consider $g_1 = \mathrm{id}$, g_2 any *ergodic* map. Then vector row of functions (g_1, g_2) is not unifiable. In fact, the ergodicity of g_2 means that the only function γ satisfying $\gamma(t) = \gamma(g_2(t))$ for a.e. $t \in \Omega$ is a constant [9]. A classical particular example of the ergodic map is the rotation of the circle $\Omega = S^1$ (equipped with the one-dimensional Hausdorff measure $|\cdot|$) by the angle $2\pi\alpha$ where α is irrational. It can also be stated in terms of transformation of the interval $\Omega = (0,1)$ as $g(t) = \{t + \alpha\}$ where $\{\cdot\}$ stands for the fractional part function. Note that in both cases the functions involved are very "nice", in particular, invertible.

It follows from the above example that for $\Omega = \Theta = (0,1)$ the vector row of constant shifts $g_i(t) = \{t + \alpha_i\}$, where $\alpha_i \in (0,1)$ is unifiable, if and only if all α_i are commensurable, i.e. if $\alpha_i/\alpha_j \in \mathbb{Q}$ for all i, j. On the other hand, a vector column of such functions is always unifiable in view of proposition 3.1.

It is important to mention, that though the unifiability condition for vector columns is less restrictive than that for vector rows, still not all vector columns of functions are unifiable, as the example below shows.

Example 3.2. Let $\Theta = \Omega = (0,1)$. Set

$$\begin{aligned} g_1(x) &:= 1/2 + 1/4 \cos 2\pi x, \\ g_2(x) &:= 1/2 + 1/4 \cos 2\pi (x - \alpha/2), \end{aligned}$$

where α is some irrational number. Suppose that there exists a unifier (γ_1, γ_2) of the vector column $(g_1, g_2)^T$ and consider the function $g = \gamma_1 \circ g_1 = \gamma_2 \circ g_2$. Let us take an arbitrary point $x \in \Omega$. First, since $g_1(x) = g_1(1-x)$, we have $g(x) = g(1-x)$. Second, since $g_2(1-x) = g_2(\{1 - (1-x) + \alpha\}) = g_2(\{x + \alpha\})$, we have $g(1-x) = g(\{x+\alpha\})$. Summing up, we obtain $g(x) = g(\{x + \alpha\})$, which means that g is a constant function, and hence γ_1 and γ_2 are constant over $g_1(\Omega)$ and $g_1(\Omega)$ respectively.

The unifiability condition for generic matrices just combines that for vector columns and for vector rows, and in fact provides very strong relationship between functions in each row and rather week relationship between functions in each column.

The importance of the notion of unifiability for function matrices for our purposes can be seen from the theorem below proven in a formally slightly weaker form in [12].

Theorem 3.1. *Let $1 \le p_i \le +\infty$, $i = 1, \ldots, k$. If the function matrix $\{g_{ij}\}_{i,j=1}^{k,l}$ is unifiable, then for the functional $I\colon L_k^{\bar{p}} \to \bar{\mathbb{R}}$ defined by*

$$I(u) := \int_\Omega f(x, \{u_i(g_{ij}(x))\}_{i,j=1}^{k,l})\, d\mu_\Omega(x), \tag{6}$$

where $u := (u_1, \ldots, u_k)$ is a vector with components $u_i \in L^{p_i}(\Theta)$, $i = 1, \ldots, k$, $f\colon \Omega \times \mathbb{R}^{kl} \to \bar{\mathbb{R}}$ is an integrand one has

$$sc^-(w - L_k^{\bar{p}}) I(u) = I^{**}(u),$$

where in the case some $p_i = +\infty$ the relaxation is meant in the $*$-weak topology.

If the function matrix $\{g_{ij}\}_{i,j=1}^{k,l}$ is not unifiable, then there is an integrand $f\colon \Omega\times\mathbb{R}^{kl}\to\bar{\mathbb{R}}$ such that the respective functional I defined by (6) is weakly lower semicontinuous over $L_k^{\bar p}$ (resp. $*$-weakly lower semicontinuous, if some $p_i=+\infty$) and not convex.

This means that unifiability of the function matrix $\{g_{ij}\}_{i,j=1}^{k,l}$ is necessary and sufficient for the relaxation in the weak topology of $L_k^{\bar p}$ of the functional I defined by (6) to be equal to its convexification. The method exploited in this paper consists in reducing the original optimal control problem to a purely variational setting similar to (6), and then relaxing the latter in weak topology of respective spaces by means of convexification. That is why unifiability condition will appear quite naturally in our further results. It is hence not just the case that similar, though less general "commensurability" or "compatibility" conditions appear under the approach of [15] (there such conditions are stated in a quite different language and for the particular case of the shifts being constant delays).

However, to find the explicit representation of the convexified functional one needs further information about the unifier. We will use for this purpose the ω-unifiability condition which seems quite a natural restriction of the notion of unifiability in this context. In fact, it requires that the all the components of the unifier be "nice" (at least piecewise injective) functions. One has to note however, that unifiability in general does not imply ω-unifiability.

Example 3.3. Let $\Omega=\Theta=(0,1)$ and consider the Borel isomorphism $j\colon\Omega\to\Omega^2$ (the existence of the latter is guaranteed by the isomorphism theorem of Borel measure spaces [10]). Now define the measurable functions g_1, $g_2\colon\Omega\to\Omega$ by the relationships

$$g_1(x):=x \text{ and } g_2(x):=j^{-1}(q(j(x))),$$

where $q\colon\Omega^2\to\Omega^2$ is given by

$$q(x_1,x_2):=(x_1,\{x_2+\alpha\})$$

for some irrational $\alpha\in(0,1)$.

Clearly, the function row (g_1,g_2) is unifiable, say, by the function $p_1\circ j\colon\Omega\to\Omega$, where $p_1\colon\Omega^2\to\Omega$ is a projection $p_1(x_1,x_2):=x_1$. Note that $p_1\circ j$ is measure preserving and hence satisfies (3). On the other hand, g_1 and g_2 are not ω-unifiable.

We prove the latter assertion by contradiction. In fact, let $\gamma\colon\Omega\to\Omega$ be a unifier of g_1 and g_2 satisfying the ω-condition. Clearly then

$$\gamma(y)=\gamma(g_2^\nu(y))$$

for a.e. $y\in\Omega$ and for all $\nu\in\mathbb{N}$. Let $\Theta\subset\Omega$ be any injectivity set of γ of positive measure. We first prove then that γ must be also injective on each $\Theta_\nu:=g_2^\nu(\Theta)$. In fact, if the latter is not true, i.e. when there exist $y_1\neq y_2$ in Θ_ν such that $\gamma(y_1)=\gamma(y_2)$, then for x_1 and x_2 from Θ such that $y_i=g_2^\nu(x_i)$, $i=1,2$ we have from the above formula $\gamma(x_1)=\gamma(x_2)$, while $x_1\neq x_2$ contradicting the injectivity of γ over Θ. The next step is to show that all Θ_ν are pairwise disjoint. In fact, if for some $\mu>\nu$ there is an $x\in\Theta_\nu\cap\Theta_\mu$, then, since $\Theta_\mu=g_2^{\mu-\nu}(\Theta_\nu)$, there is a $z\in\Theta_\nu$ satisfying $x=g_2^{\mu-\nu}(z)$. Hence, $\gamma(x)=\gamma(z)$, while clearly $x\neq z$, because g_2 has no periodic points, contradicting the injectivity of γ over Θ_ν. At last it remains to observe that all Θ_ν are disjoint and have equal measures. This implies

$$|\Theta|=|\Theta_\nu|=0,$$

which contradicts the definition of Θ.

At last is seems worth mentioning that according to our definition a matrix consisting of a single function (i.e. $k = l = 1$) is always ω-unifiable.

4. Conditional expectations

In what follows we will extensively use conditional expectations of integrands with respect to some σ-subalgebrae of the original σ-algebrae. The purpose of this short section is to draw attention to such operations, which should be properly understood. In fact, preserving the notation of the previous section, let $f\colon \Omega\times Y \to \bar{\mathbb{R}}$ be an integrand, Y being a Banach space, and $g\colon \Omega \to \Theta$ be a measurable function satisfying (3). Then formally for each $y \in Y$ the conditional expectation $\tilde{f}(\cdot, y) := E(f(\cdot, y); g^{-1}(\Sigma(\Omega))$ with respect to the σ-subalgebra generated by g is defined. One then expects for every measurable $u\colon \Theta \to Y$ (see p. 49 in vol. 1 of [9])

$$\tilde{f}(x, u(g(x))) = E(f(\cdot, u(g(\cdot))); g^{-1}(\Sigma(\Omega)))(x)$$

for a.e. $x \in \Omega$, and hence

$$\int_\Omega \tilde{f}(x, u(g(x)))\, d\mu_\Omega(x) = \int_\Omega f(x, u(g(x)))\, d\mu_\Omega(x). \tag{7}$$

If interpreted directly, this is however not true. In fact, the operator of conditional expectation is defined between classes of a.e. equal measurable functions, while different functions $\tilde{f}(\cdot, y)$ from the same class generate different integrals $\int_\Omega \tilde{f}(x, u(g(x)))\, d\mu_\Omega(x)$ for the same u.

Nevertheless, the above construction can be understood properly in the sense that in the class of functions given by the mentioned conditional expectation there is a representative ψ satisfying the desired relationship (7) (cfr. theorem III.2 of [2] where the existence of the respective version of the conditional expectation has has been proven for generic σ-algebrae though only for convex integrands). We will denote such a "nice" version of conditional expectation $E_g f$.

Proposition 4.1. *Let $f\colon \Omega \times Y \to \bar{\mathbb{R}}$ be an integrand, $g\colon \Omega \to \Theta$ be a measurable function satisfying (3). Then there is an integrand $E_g\, f\colon \Omega \times Y \to \bar{\mathbb{R}}$ such that for all $y \in Y$ one has*

$$(E_g\, f)(\cdot, y) \in E(f(\cdot, y); g^{-1}(\Sigma(\Omega))(\cdot)$$

(in the sense of belonging to a class of μ_Ω-a.e. equal functions), and (7) holds.

PROOF: The arguments are quite standard, so we just briefly indicate the main lines. Assuming without loss of generality that $\mu_\Omega(\Omega) = 1$, one has according to the disintegration theorem (theorem 1.0.8 of [1]) the existence of a measurable function $x \in \Omega_0 \mapsto \pi_x$ where π_x is a probability measure over Ω (called also *fiber measure*) and $\Omega_0 \subset \Omega$ is a set of full measure, such that $\pi_x(g^{-1}(x)) = 1$ for all $x \in \Omega_0$ and

$$\mu_\Omega(A) = \int_\Omega \pi_x(A)\, d\mu_g(x)$$

for all measurable A. It is a matter of simple calculations to verify that the function

$$(E_g f)(x, y) := \int_\Omega f(\xi, y)\, d\pi_{g(x)}(\xi)$$

satisfies the conditions of the statement being proven. In fact,

$$\begin{aligned}\int_\Omega f(x,u(g(x)))\,d\mu_\Omega(x) &= \int_\Omega d\mu_g(x)\int_\Omega f(\xi,u(g(\xi)))\,d\pi_x(\xi) = \\ &\int_\Omega d\mu_g(x)\int_\Omega f(\xi,u(x))\,d\pi_x(\xi),\end{aligned}$$

since π_x is concentrated on $g^{-1}(x)$. But then

$$\begin{aligned}\int_\Omega d\mu_g(x)\int_\Omega f(\xi,u(x))\,d\pi_x(\xi) &= \int_\Omega d\mu_\Omega(x)\int_\Omega f(\xi,u(g(x)))\,d\pi_{g(x)}(\xi) = \\ &\int_\Omega (E_g f)(x,u(g(x)))\,d\mu_\Omega(x),\end{aligned}$$

which shows the claim. □

The importance of the introduced notion of the special version of conditional expectation can be seen from the following simple representation result for the relaxation of a functional containing just one argument deviation.

Theorem 4.1. *Let*

$$I(u) := \int_\Omega f(x,u(g(x)))\,d\mu_\Omega(x),$$

where $f\colon \Omega\times\mathbb{R}\to\bar{\mathbb{R}}$ be a nonnegative normal integrand such that $I(u)<+\infty$ for every $u\equiv const$, and $g\colon \Omega\to\Theta$ be a measurable function satisfying (3). Then

$$sc^-(w-L^p(\Theta))\,I(u) = I^{**}(u) = \int_\Omega (E_g\,f)^{**}(x,u(g(x)))\,d\mu_\Omega(x).$$

PROOF: Let us remark that since the matrix consisting of a single function g is always ω-unifiable (one can take just the identity function as a unifier), then the first part of the statement, namely, the equality between relaxation in weak topology and convexification of the whole functional, follows immediately from theorem 3.1.

To prove the remaining assertion, note that (7) means

$$I(u) = \int_\Omega (E_g f)(x,u(g(x)))\,d\mu_\Omega(x), \tag{8}$$

and assume without loss of generality $f(x,0)=0$. Since $(E_g f)(\cdot,y)$ is measurable with respect to $g^{-1}(\Sigma(\Omega))$, there is a $\varphi\colon\ \Omega\times\mathbb{R}\to\bar{\mathbb{R}}$ such that $(E_g f)(x,y) = \varphi(g(x),y)$, and hence changing the variables in (8) we obtain that

$$I(u) = \int_\Theta \frac{d\mu_g}{d\mu_\Theta}(t)\varphi(t,u(t))\,d\mu_\Theta(t).$$

Applying now the proposition 2.3 of [8, chapter IX], we have

$$sc^-(w-L^p)I(u) = \int_\Theta \frac{d\mu_g}{d\mu_\Theta}(t)\varphi^{**}(t,u(t))\,d\mu_\Theta(t),$$

which further implies

$$\begin{aligned}sc^-(w-L^p)I(u) &= \int_\Omega \varphi^{**}(g(x),u(g(x)))\,d\mu_\Omega(x) = \\ &\int_\Omega (E_g f)^{**}(x,u(g(x)))\,d\mu_\Omega(x),\end{aligned}$$

concluding the proof. □

Let us observe that in general in the condition of the above theorem

$$sc^-(w-L^p)I(u) \neq \int_\Omega f^{**}(x,u(g(x)))\,d\mu_\Omega(x).$$

In fact, consider the following

Example 4.1. Let $\Omega=\Theta=(0,1)$ and

$$f(x,y):=(1-\exp(-y^2))\chi_{(0,1/2]}(x)+(\exp(y^2)-1)\chi_{(1/2,1)}(x),$$

while

$$g(x):=\begin{cases} 2x, & 0<x\le 1/2,\\ 2-2x, & 1/2<x<1.\end{cases}$$

It is a matter of simple computation to show

$$I(u):=\int_\Omega f(x,u(g(x)))\,d\mu_\Omega(x)=\int_0^1 \sinh u(x)^2\,d\mu_\Omega(x),$$

and hence $sc^-(w-L^p)I(u)=I(u)$. Using the fact that g_1 is measure preserving, we find out

$$sc^-(w-L^p)I(u)=I(u)=\frac{1}{2}\int_\Omega \sinh u(g(x))^2\,d\mu_\Omega(x),$$

while

$$\int_\Omega f^{**}(x,u(g(x)))\,d\mu_\Omega(x)=\int_\Omega(\exp u(g(x))^2-1)\,d\mu_\Omega(x).$$

Let us show that the discussed "nice" version of conditional expectation in some cases can be easily constructed. This is the case, for instance, when the function $g\colon \Omega\to\Theta$ satisfies the ω-condition.

Proposition 4.2. *Let $f\colon \Omega\times Y\to\bar{\mathbb{R}}$ be an integrand, and a function $g\colon\Omega\to\Theta$ satisfy the ω-condition. Then one can take*

$$(E_g\,f)(x,y):=\sum_{\nu=1}^{\infty}\frac{1}{\frac{d\mu_{g_\nu}}{d\mu_\Theta}(g_\nu(g(x)))\frac{d\mu_g}{d\mu_\Theta}(g(x))}f(g_\nu(g(x)),y)).$$

PROOF: We choose an arbitrary measurable function $u\colon\Theta\to X$ and consider $I(u)=\int_\Omega f(x,u(g(x)))\,d\mu_\Omega(x)$. By substituting $x:=g_\nu(t)$ over each $\Omega_\nu\subset\Omega$ one has

$$I(u)=\sum_{\nu=1}^{\infty}\int_{\Omega_\nu} f(x,u(g(x)))\,d\mu_\Omega(x)=\sum_{\nu=1}^{\infty}\int_\Theta\frac{1}{\frac{d\mu_{g_\nu}}{d\mu_\Theta}(g_\nu(t))}f(g_\nu(t),u(t))\,d\mu_\Theta(t).$$

Note that we assumed $\frac{d\mu_{g_\nu}}{d\mu_\Theta}(g_\nu(t)):=+\infty$ for $t\notin g(\Omega_\nu))$. Aplying now the substitution $t:=g(x)$ one obtains further

$$I(u)=\int_\Omega\bar f(x,u(x))\,d\mu_\Omega(x),$$

where

$$\bar f(x,y):=\sum_{\nu=1}^{\infty}\frac{1}{\frac{d\mu_{g_\nu}}{d\mu_\Theta}(g_\nu(g(x)))\frac{d\mu_g}{d\mu_\Theta}(g(x))}f(g_\nu(g(x)),u(g(x)))$$

Clearly, since $\bar f(\cdot,y)$ is g-measurable, then $\bar f=\mathrm{E}_g\,f$. $\square$

Of special interest for further applications is the case when $f = \chi_A$ is the indicator function of some $A \subset \Omega \times Y$, where

$$\chi_A(z) := \begin{cases} 0, & z \in A, \\ +\infty, & z \notin A. \end{cases}$$

Corollary 4.1. *Let $A \in \Sigma(\Omega) \times \mathcal{B}(Y)$. Then $E_g\, \chi_A = \chi_{e_g\, A}$, where the set $e_g\, A$ is defined by*

$$e_g\, A := \{(x, y) \in A \,:\, (x', y) \in A \text{ whenever } g(x') = g(x)\}.$$

PROOF: One has

$$\mathrm{e}_g\, A = \bigcap_{\nu=1}^{\infty} ((g^\nu, \mathrm{id}) \circ (g, \mathrm{id}))(A),$$

which concludes the proof. □

5. General setting and solution

Consider a rather general setting of an optimal control problem for differential equation with deviating argument

$$\inf_{y \in Y,\, u \in U} \left\{ \begin{array}{l} \displaystyle\int_\Omega (\Psi(t, y(t)) + \varphi(t, u(t)))\; dt \;: \\ \dot{y}(t) = a(t, y(h_1(t))) + b(t, y(h_2(t)))\psi(t, \{u_i(g_{ij}(t))\}_{i,j=1}^{k,l}), \\ y(t_0) = y_0, \end{array} \right\} \tag{9}$$

where y is the state variable, $u = (u_1, \dots, u_k)$ is the control vector, $\Omega = (t_0, t_1) \subset \mathbb{R}$, h_1, h_2, g_{ij}: $\Omega \to \Omega$ are measurable and all g_{ij} satisfy (3) while φ: $\Omega \times \mathbb{R}^{kl} \to \bar{\mathbb{R}}$ is the nonnegative normal integrand, and Ψ, a, b: $\Omega \times \mathbb{R} \to \bar{\mathbb{R}}$ and ψ: $\Omega \times \mathbb{R}^{kl} \to \bar{\mathbb{R}}$ are Carathéodory functions. The space of states $Y = W^{1,1}(\Omega)$ and the space of controls $U = L_k^{\bar{p}}(\Omega)$, $\bar{p} = \{p_i\}_{i=1}^k$, are Sobolev and Lebesgue spaces respectively. In order to avoid further complications, we sty within the reflexive space framework hence assuming $1 < p_i < +\infty$ for all $i = 1, \dots, k$.

Let us consider the following condition

(Q) There are such numbers $\lambda > 0$, $\beta \geq 0$ and $q \in [1, +\infty)$ that

$$\lambda \int_\Omega \left(\sum_{i=1}^k |u_i(t)|^{p_i} + \left| \psi(t, \{u_i(g_{ij}(t))\}_{i,j=1}^{k,l}) \right|^q, \right) dt - \beta \leq \max\left(\int_\Omega \sum_{i=1}^k |u_i(t)|^{p_i}\, dt, \int_\Omega \varphi(t, u(t))\, dt \right)$$

for every $u \in L_k^{\bar{p}}(\Omega)$.

Note that when $\varphi \equiv 0$, i.e. the cost functional does not depend on the control variable, then the condition (Q) is then reduced to the following simpler one

(Q') There are such numbers $\beta \geq 0$, $q \in [1, +\infty)$ and a function $\alpha \in L^1(\Omega)$ that

$$|\psi(t, u_1, \dots, u_k)| \leq \alpha(t) + \beta \sum_{i=1}^k |u_i|^{p_i/q}.$$

for a.e. $t \in \Omega$ and for all $u \in \mathbb{R}$ (that is, the Nemytskiĭ operator generated by ψ sends bounded sets of $L_k^{\bar{p}}(\Omega)$ into bounded sets of $L^q(\Omega)$).

We denote $V := L^q$ with the exponent $q \in [1, +\infty)$ determined by the above condition.

We proclaim the following theorem, which constitutes the main result of the paper. It says that under the above natural condition (Q) plus ω-unifiability of the function matrix obtained by extending $\{g_{ij}\}_{i,j=1}^{k,l}$ by a column of identity functions, the relaxed setting for the problem (9) can be obtained by the convexification method (it is here that we use the unifiability). Moreover, the relaxed formulation is, roughly speaking, again an optimal control problem but with cost functional and state equations involving up to countably many argument deviations. Note that in principle this result is constructive, namely, the proof indicates the precise construction of the relaxed setting.

Theorem 5.1. *Suppose that the function matrix $\{g_{ij}\}_{i,j=1}^{k,l+1}$, where $g_{i,l+1} = id$ for all $j = 1, \ldots, k$, is ω-unifiable. Then under condition (Q) there exist*

(i) *two sequences of measurable sets $\Omega_i^\nu \subset \Omega$ and $\Omega^\nu \subset \Omega$, $\nu \in \mathbb{N}$, some of them possibly empty,*

(ii) *two sequences of measurable functions $g^\nu\colon \Omega^\nu \to \Omega$ and $g_i^\nu\colon \Omega_i^\nu \to \Omega$ all satisfying ω-condition,*

(iii) *the functions $\hat{\varphi}$, $\hat{\psi}\colon \Delta \to \bar{\mathbb{R}}$, the former being also nonnegative and lower semicontinuous in the second variable (i.e. $\varphi(t, \cdot)$ is l.s.c. for a.e. $t \in \Omega$), where the set Δ corresponds to the family*

$$\Delta(t) = l_k^{\bar{p}}[\{\omega_i^\nu(t)\}_{t\in\Omega_i^\nu}] \times l^q[\{\omega^\nu(t)\}_{t\in\Omega^\nu}],$$

with some measurable functions $\omega_i^\nu\colon \Omega_i^\nu \to [0, +\infty]$, $\omega^\nu\colon \Omega^\nu \to [0, +\infty]$,

such that the problem (9) has a relaxed setting represented by

$$\inf_{y\in Y,\, u\in U, v\in V} \left\{ \int_\Omega \left(\Psi(t, y(t)) + \hat{\varphi}(t, \{u_i(g_i^\nu(t))\}_{t\in\Omega_i^\nu}, \{v(g^\nu(t))\}_{t\in\Omega^\nu}) \right) dt \ : \begin{array}{l} \hat{\psi}(t, \{u_i(g_i^\nu(t))\}_{t\in\Omega_i^\nu}, \{v(g^\nu(t))\}_{t\in\Omega^\nu}) = 0, \\ \dot{y}(t) = a(t, y(h_1(t))) + b(t, y(h_2(t)))v(t),\ y(t_0) = y_0. \end{array} \right\} \tag{10}$$

Note that the relaxed setting in the form given by the above theorem involves an extra state equation

$$\hat{\psi}(t, \{u_i(g_i^\nu(t))\}_{t\in\Omega_i^\nu}, \{v(g^\nu(t))\}_{t\in\Omega^\nu}) = 0.$$

It will be however clear from the proof that $\hat{\psi}$ is just an indicator function of some subset of Δ, and hence the use of such extra equation is just a convenient form to write a differential inclusion.

The proof of the following corollary concerning a simpler setting with the cost functional independent of the control vector can be easily obtained following the lines of that of the above theorem. Note that it involves condition on the original function matrix extended by the row (rather than by the column, as above) of identity functions.

Corollary 5.1. *If in (9) the cost functional does not depend on the control vector, i.e. $\varphi \equiv 0$, the function matrix $\{g_{ij}\}_{i,j=1}^{k+1,l}$ obtained from the original one by adding a row of identity functions, i.e. $g_{k+1,l} = id$, is ω-unifiable and condition (Q') holds, then the problem (9) has the relaxed setting given by (10) with $\hat{\varphi} \equiv 0$.*

Proof of the theorem 5.1.

Step 1. Denoting

$$\begin{aligned}
\Lambda \;&:=\; \Big\{(u,y)\in U\times Y\ :\ \dot y(t)=a(t,y(h_1(t)))+\\
&\qquad b(t,y(h_2(t)))\psi(t,\{u_i(g_{ij}(t))\}_{i,j=1}^{k,l}),\ y(t_0)=y_0\Big\},\\
F(u,y) \;&:=\; \begin{cases} \displaystyle\int_\Omega (\Psi(t,y(t))+\varphi(t,u(t)))\,dt, & (u,v)\in\Lambda,\\ +\infty, & \text{otherwise,}\end{cases}
\end{aligned}$$

we see that the problem consists in finding the relaxation $\mathrm{sc}^-(U\times Y)\,F$. Introduce an auxiliary variable $v\in V=L^q(\Omega)$ and define a new functional

$$\Phi(u,v,y):=\begin{cases} F(u,y), & v(t)=\psi(t,\{u_i(g_{ij}(t))\}_{i,j=1}^{k,l})\text{ a.e. in }\Omega,\\ +\infty, & \text{otherwise.}\end{cases}$$

Let us denote now

$$\begin{aligned}
\bar\Lambda \;&:=\; \Big\{(u,v,y)\in U\times V\times Y\ :\\
&\qquad \dot y(t)=a(t,y(h_1(t)))+b(t,y(h_2(t)))v(t),\ y(t_0)=y_0,\Big\}\\
J(u,v) \;&:=\; \int_\Omega \Big(\varphi(t,u(t))+\chi[\psi(t,z)=v](t,\{u_i(g_{ij}(t))\}_{i,j=1}^{k,l},v(t))\Big)\,dt,\\
J_0(y) \;&:=\; \int_\Omega \Psi(t,y(t))\,dt,
\end{aligned}$$

where for the indicator function χ the following notation is chosen:

$$\chi[P(r)](r'):=\begin{cases} 0, & P(r')\text{ holds ,}\\ +\infty, & \text{otherwise .}\end{cases}$$

Then the problem (9) can be rewritten as

$$\begin{aligned}
\inf_{y\in Y,\,u\in U}\Big\{\int_\Omega \Big(&\Psi(t,y(t))+\varphi(t,u(t))+\\
&\chi[\psi(t,z)=v](t,\{u_i(g_{ij}(t))\}_{i,j=1}^{k,l},v(t))\Big)\,dt\ :\\
&\dot y(t)=a(t,y(h_1(t)))+b(t,y(h_2(t)))v(t),\ y(t_0)=y_0.\Big\}
\end{aligned}\tag{11}$$

The condition (Q) implies that whenever $(u_\nu,y_\nu)\in U\times Y$ is a converging sequence with $F(u_\nu,y_\nu)$ bounded, then the sequence $v_\nu\in V$ defined by

$$v(t):=\psi(t,\{u_i(g_{ij}(t))\}_{i,j=1}^{k,l})$$

is precompact in the weak topology. Hence, if we consider V to be endowed with its weak topology, then by the general theorem on behaviour of Γ-limits under introduction of an auxiliary variable from [3] the problem of finding $\mathrm{sc}^-(U\times Y)\,F$ is reduced to finding $\mathrm{sc}^-(U\times V\times Y)\,\Phi$, namely,

$$\mathrm{sc}^-(U\times Y)\,F(u,y)=\inf_{v\in V}\mathrm{sc}^-(U\times V\times Y)\,\Phi(u,v,y).$$

We note that $\Phi(u,v,y)=J_0(y)+J(u,v)+\chi_{\bar\Lambda}$ and by the general theorems on Γ-limits of sums of functionals from [5] one has that

$$\mathrm{sc}^-(U\times V\times Y)\,\Phi(u,v,y)=J_0(y)+\mathrm{sc}^-(U\times V)\,J(u,v)+\chi_{\bar\Lambda},$$

and it remains hence to characterize $\mathrm{sc}^-(U\times V)\,J$.

Step 2. Define the functions $\tilde\varphi$, $\tilde\psi$: $\Omega\times\mathbb{R}^{k(l+1)}\times\mathbb{R}^{l+1}\to\bar{\mathbb{R}}$ by setting

$$\begin{aligned}
\tilde\varphi(t,z,w) \;&:=\; \varphi(t,\{z_{i,l+1}\}_{i=1}^{k}),\\
\tilde\psi(t,z,w) \;&:=\; w_1-\psi(t,\{z_j\}_{i,j=1}^{k,l}).
\end{aligned}$$

Then

$$J(u,v)=\int_\Omega \tilde\phi(t,\{u_i(g_{ij}(t))\}_{i,j=1}^{k,l+1},\{v(g_{k+1,j}(t)))\}_{j=1}^{l+1})\,dt,$$

where all $g_{i,l+1}:=\mathrm{id}$ and $g_{k+1,j}:=\mathrm{id}$, while

$$\tilde\phi(t,z,w):=\tilde\varphi(t,z,w)+\chi[\tilde\psi(t,z,w)=0](t,z,w).$$

Observe now that the $(k+1)\times(l+1)$ function matrix $\{g_{ij}\}_{i,j=1}^{k+1,l+1}$ is still ω-unifiable, since adding a row of identity functions to the $k\times(l+1)$ matrix $\{g_{ij}\}_{i,j=1}^{k,l+1}$ obviously does not change the ω-unifiability. Therefore to calculate $\mathrm{sc}^-(U\times V)\,J$, we may now employ the technique used in the proof of the theorem 6.1.

Step 3. Using the notation of the lemma 6.1, define a function $\phi_\infty\colon \Omega\times l_k^{\bar p}\times l^q\to\bar{\mathbb{R}}$ by setting

$$\phi_\infty(t,\hat z,\hat w):=\tilde\phi(t,H(t)(\hat z,\hat w))$$

and notice that $\phi_\infty=\varphi_\infty+\chi[\psi_\infty=0]$, where

$$\begin{array}{rcl}\varphi_\infty(t,\hat z,\hat w) & := & \tilde\varphi(t,H(t)(\hat z,\hat w)),\\ \psi_\infty(t,\hat z,\hat w) & := & \tilde\psi(t,H(t)(\hat z,\hat w)).\end{array}$$

One should find the conditional expectation $\mathrm{E}_g\,\phi_\infty$ of the function ϕ_∞. For this purpose one can use proposition 4.2 and corollary 4.1. Let $\varphi_{\infty,g}:=E_g\varphi_\infty$ and $\psi_{\infty,g}\colon\Omega\times l_k^{\bar p}\times l^q\to\bar{\mathbb{R}}$ stand for any function such that

$$\{\psi_{\infty,g}=0\}=e_g\{\psi_\infty=0\}.$$

Now define the function $\bar\phi\colon\Delta\to\bar{\mathbb{R}}$ by setting

$$\bar\phi(t,\bar z,\bar w):=(E_g\phi_\infty)(t,\hat A(g(t))(\bar z,\bar w)),$$

noting that $\bar\phi=\bar\varphi+\chi[\bar\psi=0]$, where

$$\begin{array}{rcl}\bar\varphi(t,\bar z,\bar w) & := & \varphi_{\infty,g}(t,\hat A(g(t))(\bar z,\bar w)),\\ \bar\psi(t,\bar z,\bar w) & := & \psi_{\infty,g}(t,\hat A(g(t))(\bar z,\bar w)).\end{array}$$

It remains to find the convexification of the function $\bar\phi$ in the pair of variables $(\bar z,\bar w)$. Clearly, $\bar\phi^{**}\equiv+\infty$ over the set $\overline{\mathrm{co}}\{\bar\psi=0\}$, where $\overline{\mathrm{co}}$ stands for the closed convex hull. Therefore, we can write

$$\hat\phi:=\bar\phi^{**}=\hat\varphi+\chi[\hat\psi=0],$$

where

$$\hat\psi=\chi_{\overline{\mathrm{co}}\{\bar\psi=0\}},\qquad \hat\varphi=\hat\phi-\hat\psi,$$

while $\infty-\infty=0$ is assumed. Now we can write out

$$\begin{array}{rcl}\mathrm{sc}^-(U\times V)\,J & = & \int_\Omega \hat\varphi(t,\{u_i(g_i^\nu(t))\}_{t\in\Omega_i^\nu},\{v(g^\nu(t))\}_{t\in\Omega^\nu})\,dt+\\ & & \chi[\hat\psi(t,\{u_i(g_i^\nu(t))\}_{t\in\Omega_i^\nu},\{v(g^\nu(t))\}_{t\in\Omega^\nu})=0](u,v).\end{array}$$

Recalling the result of Step 1, we come to the problem (10). □

Let us now give some examples.

Example 5.1. Consider the sample optimal control problem (1) involving only one argument deviation in control variable. Here we have $k=1$, $l=1$ and $g_{11}=g$. In general, the "extended" 1×2 matrix of functions $g_{11}=g$, $g_{12}=\mathrm{id}$ may be not ω-unifiable. For instance, it is not even unifiable, if g is ergodic. But if this matrix is ω-unifiable (which is a bit stronger than just nonergodicity of g), we can apply theorem 5.1 and assert that under the condition (Q) the relaxed problem looks like

$$\inf_{y\in Y,\, u\in U, v\in V} \left\{ \int_\Omega \left(\Psi(t,y(t)) + \hat{\varphi}(t, \{u(g_1^\nu(t))\}_{t\in\Omega_1^\nu}, \{v(g^\nu(t))\}_{t\in\Omega^\nu})\right)\, dt \ : \begin{array}{l} \hat{\psi}(t, \{u(g_1^\nu(t))\}_{t\in\Omega_1^\nu}, \{v(g^\nu(t))\}_{t\in\Omega^\nu}) = 0, \\ \dot{y}(t) = a(t, y(h_1(t))) + b(t, y(h_2(t)))v(t),\ y(t_0) = y_0. \end{array} \right\}$$

Example 5.2. Consider a particular case of (1) when the cost functional of the original problem is independent of the control variable. Note that in this case the 2×1 matrix of functions $g_{11} = g$, $g_{21} = \mathrm{id}$ is always unifiable (just take (id, g) as a unifier), but it is ω-unifiable whenever g satisfies ω-condition.

To write out the relaxed problem for this particular case, we first recall what the new functions g_1^ν, g^ν will look like. It is clear that the vector column of functions (g, id) is unifiable by, say, the unifier (id, g). In general, g has at most countable number of injectivity sets Θ^ν while id has only one injectivity set $\Theta_1 = \Omega$. Thus, we introduce functions g^ν: $\Omega^\nu \to \Omega$, $\Omega^\nu = g^{-1}(g(\Theta^\nu))$ by $g^\nu = \hat{g}^\nu \circ g$ where $\hat{g}^\nu$ are the inverses of the injective restrictions of g, and a function $g_1^1 : \Omega \to \Omega$ by $g_1^1 = g$. Now, whatever the measurable function g satisfying ω-condition is, under only the condition (Q') the relaxed problem is represented by

$$\inf_{y\in Y,\, u\in U, v\in V} \left\{ \int_\Omega \Psi(t,y(t))\, dt \ : \begin{array}{l} \hat{\psi}(t, u(g(t)), \{v(g^\nu(t))\}_{t\in\Omega^\nu}) = 0, \\ \dot{y}(t) = a(t, y(h_1(t))) + b(t, y(h_2(t)))v(t),\ y(t_0) = y_0. \end{array} \right\}$$

In particular, when g is injective, then the relaxed problem becomes

$$\inf_{y\in Y,\, u\in U, v\in V} \left\{ \int_\Omega \Psi(t,y(t))\, dt \ : \begin{array}{l} \hat{\psi}(t, u(g(t)), v(t)) = 0, \\ \dot{y}(t) = a(t, y(h_1(t))) + b(t, y(h_2(t)))v(t),\ y(t_0) = y_0. \end{array} \right\}$$

The latter setting can be further essentially simplified, if one manages to reduce the equations by getting rid of the extra v variable. For instance, in the simplest case $b \equiv 1$ the relaxed setting becomes

$$\inf_{y\in Y,\, u\in U} \left\{ \int_\Omega \Psi(t,y(t))\, dt \ : \ \hat{\psi}\left(t, u(g(t)), \dot{y}(t) - a(t, y(h_1(t)))\right),\ y(t_0) = y_0. \right\}$$

Example 5.3. We consider a setting with a control represented by a vector, with different deviations in each of the components involved:

$$\inf_{y\in Y,\, u\in U} \left\{ \int_\Omega \left(\Psi(t,y(t)) + \frac{|u_1(t)|^2}{2} + \frac{|u_2(t)|^2}{2}\right)\, dt \ : \begin{array}{l} \dot{y}(t) = a(t, y(h_1(t))) + b(t, y(h_2(t)))\psi(t, u_1(g_1(t)), u_2(g_2(t))), \\ y(t_0) = y_0 \end{array} \right\}, \tag{12}$$

The state of controls is $U = L_2^{\bar{p}}(\Omega)$.

Here we have $k = 2$, $l = 1$, $g_{11} = g_1$, $g_{21} = g_2$. In general, the matrix of functions $g_{11} = g_1$, $g_{12} = \mathrm{id}$, $g_{21} = g_2$, $g_{22} = \mathrm{id}$ may be not ω-unifiable. This is the case, for instance, when one of the functions is ergodic (then this matrix is not even unifiable). But if it is ω-unifiable (this is a bit stronger than nonergodicity of each of the functions g_1 and g_2), then under the condition (Q) the relaxed setting can be written out as

$$\inf_{y\in Y,\, u\in U, v\in V} \left\{ \int_\Omega \Big(\Psi(t,y(t)) + \hat\varphi(t, \{u_1(g_1^\nu(t))\}_{t\in\Omega_1^\nu}, \{u_2(g_2^\nu(t))\}_{t\in\Omega_2^\nu}, \{v(g^\nu(t))\}_{t\in\Omega^\nu}) \Big)\, dt \;: \begin{array}{l} \hat\psi(t, \{u_1(g_1^\nu(t))\}_{t\in\Omega_1^\nu}, \{u_2(g_2^\nu(t))\}_{t\in\Omega_2^\nu}, \{v(g^\nu(t))\}_{t\in\Omega^\nu}) = 0, \\ \dot y(t) = a(t, y(h_1(t))) + b(t, y(h_2(t)))v(t),\ y(t_0) = y_0. \end{array} \right\}$$

We see that the difference with the example 5.1 is rather formal. In particular, if the cost functional does not depend on control variables, to write out the relaxed setting (which will be exactly as above, but with $\hat\varphi \equiv 0$) it will be enough to require only the fulfillment of the condition (Q') and the ω-unifiability of the 3×1 matrix of functions $g_{11} = g_1$, $g_{21} = g_2$, $g_{31} = \mathrm{id}$. We will only write out this setting in the case when both g_1 and g_2 are injective, since in this case it is particularly simple:

$$\inf_{y\in Y,\, u\in U, v\in V} \left\{ \int_\Omega \Psi(t,y(t))\, dt \;: \begin{array}{l} \hat\psi(t, u_1(g_1(t)), u_2(g_2(t)), v(t)) = 0, \\ \dot y(t) = a(t, y(h_1(t))) + b(t, y(h_2(t)))v(t),\ y(t_0) = y_0. \end{array} \right\}$$

6. Relaxation versus convexification in variational setting

The purpose of this section is to prove the general theorem on relaxation of integral functionals $I\colon L_k^{\bar p}(\Theta) \to \bar{\mathbb{R}}$ defined by

$$I(u) := \int_\Omega f(x, \{u_i(g_{ij}(x))\}_{i,j=1}^{k,l})\, d\mu_\Omega(x), \tag{13}$$

in weak topology of the respective Lebesgue space. Here we consider $(\Omega, \Sigma(\Omega), \mu_\Omega)$ and $(\Theta, \mathcal{B}(\Theta), \mu_\Theta)$ to be standard measure spaces. The result proved in this section serves as a basis for the proof of the principal theorem 5.1 on relaxation of optimal control problems.

Further on we make the following additional assumptions:

(a) the integrand $f\colon \Omega\times\mathbb{R}^{kl} \to \bar{\mathbb{R}}$ is normal and nonnegative;
(b) the functional I is finite over every constant function, that is

$$I(u) < +\infty \text{ for every } u \equiv y \in \mathbb{R}^k.$$

Both requirements are not very restrictive, and can further be weakened. We also assume further on $1 < p_i < +\infty$ for all $i = 1, \dots, k$.

Theorem 6.1. *If a function matrix $\{g_{ij}\}_{i,j=1}^{k,l}\colon \Omega \to \Theta$ is ω-unifiable, then there exist*

(1) *a sequence of measurable sets $\{\Omega_i^\nu\}_{i,\nu=1}^{k,\infty}$, $\Omega_i^\nu \subset \Omega$,*
(2) *a sequence of measurable functions $\{g_i^\nu\}_{i,\nu=1}^{k,\infty}$, $g_i^\nu\colon \Omega_i^\nu \to \Theta$,*
(3) *a sequence of nonnegative measurable functions $\{\omega_i^\nu\}_{i,\nu=1}^{k,\infty}\colon \Omega_i^\nu \to [0,+\infty]$, defining a family of spaces of weighted l^p type $\{\Delta(x)\}_{x\in\Omega}$,*

$$\Delta(x) = l_k^{\bar p}[\{\omega_i^\nu(x)\}_{x\in\Omega_i^\nu}],$$

(4) *a function $\bar\phi : \Delta \to \bar{\mathbb{R}}$ defined over the family Δ,*

such that

$$sc^-(w - L_k^{\bar p}(\Theta))\, I(u) = \int_\Omega \bar\phi^{**}(x, \{u_i(g_i^\nu(x))\}_{x\in\Omega_i^\nu})\, d\mu_\Omega(x).$$

PROOF:

Step 1. Without loss of generality we suppose that $f(x,0,\dots,0) \equiv 0$. Denote

$$T(u)(x) := f(x, \{u_i(g_{ij}(x))\}_{i,j=1}^{k,l}).$$

Our intention is to represent the formal operator T acting over the Lebesgue space $L_k^{\bar p}(\Theta)$ as a composition of a Nemytskiĭ operator with *one* linear shift. To fulfill this, we need to pass from the Lebesgue space of scalar functions $L_k^{\bar p}(\Theta)$ to the Lebesgue space of functions with values in an infinite-dimensional Banach space. The respective construction is provided by the lemma below.

Lemma 6.1. *If the function matrix $\{g_{ij}\}_{i,j=1}^{k,l}\colon \Omega \to \Theta$, is ω-unifiable, with a unifier $(\gamma_1,\dots,\gamma_l)\colon \Theta \to (0,1)$, then there exist*

1. *a sequence of measurable sets $\{\Lambda_i^\nu\}_{i,\nu=1}^{k,\infty}$, $\Lambda_i^\nu \subset (0,1)$,*
2. *a sequence of measurable functions $\{\gamma_i^\nu\}_{i,\nu=1}^{k,\infty}\colon \Lambda_i^\nu \to \Theta$,*
3. *a sequence of measurable functions $\{w_i^\nu\}_{i,\nu=1}^{k,\infty}\colon \Lambda_i^\nu \to [0,+\infty]$, defining a family of spaces of weighted l^p type $\{D(s)\}_{s\in(0,1)}$, $D(s) = l_k^{\bar p}[\{w_i^\nu(s)\}_{s\in\Lambda_i^\nu}]$,*
4. *a normal integrand $\hat f\colon \Omega\times l_k^{\bar p} \to \bar{\mathbb{R}}$,*
5. *two linear continuous operators*

$$\begin{array}{rrcl} P: & L_k^{\bar p}((0,1);l_k^{\bar p}) & \to & L_k^{\bar p}(\Theta),\\ \hat P: & L_k^{\bar p}(\Theta) & \to & L_k^{\bar p}((0,1);l_k^{\bar p}), \end{array}$$

6. *two operator-functions of $x\in\Omega$, the values of which are linear continuous operators*

$$\begin{array}{rrcl} H(x): & l_k^{\bar p} & \to & K(x)\subset \mathbb{R}^{kl},\\ \hat H(x): & K(x) & \to & l_k^{\bar p}, \end{array}$$

   where $K(x) = \{y\in\mathbb{R}^{kl} : g_{ij'}(x) = g_{ij'}(x) \Rightarrow y_{ij'} = y_{ij'}\}$,

7. *two operator-functions of $s\in(0,1)$, the values of which are linear continuous operators*

$$\begin{array}{rrcl} \bar A(s): & l_k^{\bar p} & \to & D(s),\\ \hat A(s): & D(s) & \to & l_k^{\bar p}, \end{array}$$

   $\hat A(s)$ being also an isometry,

that satisfy the following conditions:

- (i) *$T(P\hat u)(x) = \hat f(x, \hat u(g(x)))$ and, in particular, $T(u)(x) = \hat f(x, (\hat P u)(g(x)))$ for all $u\in L_k^{\bar p}(\Theta)$, $\hat u\in L_k^{\bar p}((0,1);l_k^{\bar p})$ and for μ_Ω-a.e. $x\in\Omega$, where*

$$g := \gamma_i \circ g_{ij};$$

- (ii) *$\hat f(x,\hat y) = f(x, H(x)\hat y)$, while $\hat f$ is a normal integrand whenever so is f;*
- (iii) *$(\hat P u)(s) = \hat A(s)\{u_i(\gamma_i^\nu(s))\}_{s\in\Lambda_i^\nu}$ for all $u\in L_k^{\bar p}(\Theta)$ and a.e. $s\in(0,1)$;*
- (iv) *$P\hat P = Id_{L_k^{\bar p}(\Theta)}$, $H(x)\hat H(x) = Id_{K(x)}$ for μ_Ω-a.e. $x\in\Omega$, $\bar A(s)\hat A(s) = Id_{D(s)}$ for a.e. $s\in(0,1)$ where Id_X stands for the identity of X;*
- (v) *$H(x)\hat A(g(x))\bar A(g(x)) = H(x)$ for μ_Ω-a.e. $x\in\Omega$.*

REMARK. As it will be clear from the proof of the above lemma, *(i)* will follow directly from a stronger assertion

$$H(x)\hat{u}(g(x)) = \{(P\hat{u})_i(g_{ij}(x))\}_{i,j=1}^{k,l} \qquad \mu_\Omega - \text{a.e. in } \Omega \tag{14}$$

for all $\hat{u} \in L_k^{\bar{p}}((0,1); l_k^{\bar{p}})$.

By *(i)* of the lemma 6.1 one has

$$I(u) = \int_\Omega \hat{f}(x, (\hat{P}u)(g(x)))\, d\mu_\Omega(x).$$

Step 2. Consider a new functional $\hat{I}\colon L_k^{\bar{p}}((0,1); l_k^{\bar{p}}) \to \bar{\mathbb{R}}$ defined by the formula

$$\hat{I}(\hat{u}) = \int_\Omega \hat{f}(x, \hat{u}(g(x)))\, d\mu_\Omega(x). \tag{15}$$

We will prove that

$$\mathrm{sc}^-(w - L_k^{\bar{p}}((0,1); l_k^{\bar{p}}))\, \hat{I}(\hat{u}) = \int_\Omega \hat{\phi}^{**}(x, \hat{u}(g(x)))\, d\mu_\Omega(x), \tag{16}$$

where $\hat{\phi} = \mathrm{E}_g\, \hat{f}$. The proof is only a slight modification of that of the theorem 4.1 in order to deal with integrands defined over infinite-dimensional spaces. In fact, first in a complete analogy with the mentioned proof we show that

$$\hat{I}(\hat{u}) = \int_\Omega \hat{\phi}(x, \hat{u}(g(x)))\, d\mu_\Omega(x) = \int_0^1 \frac{d\mu_g}{d\mathcal{L}^1}(s) \hat{\varphi}(s, \hat{u}(s))\, ds,$$

where $\mathcal{L}^1$ stands for the one-dimensional Lebesgue mesaure and $\hat{\varphi}\colon (0,1) \times l_k^{\bar{p}} \to \bar{\mathbb{R}}$ satisfies $\hat{\varphi}(g(x), \hat{y}) = \hat{\phi}(x, \hat{y})$. We substitute then the proposition 2.3 of [8, chapter IX] with the theorem VII.7 of [4] in order to assert that

$$\mathrm{sc}^-(w - L_k^{\bar{p}}((0,1); l_k^{\bar{p}}))\, \hat{I}(\hat{u}) = \int_0^1 \frac{d\mu_g}{d\mu_\Theta}(s) \hat{\varphi}^{**}(s, \hat{u})(s)\, ds,$$

and hence

$$\mathrm{sc}^-(w - L_k^{\bar{p}}((0,1); l_k^{\bar{p}}))\, \hat{I}(\hat{u}) = \int_\Omega \hat{\phi}^{**}(x, \hat{u}(g(x)))\, d\mu_\Omega(x).$$

Step 3. We use the lemma below which is a reformulation of lemma 4.2 from [13] with $X = L_k^{\bar{p}}(\Theta)$, $\hat{X} = L_k^{\bar{p}}((0,1); l_k^{\bar{p}})$, the functional $\hat{I}$ given by (15), and I the original functional to prove that

$$\begin{aligned} \mathrm{sc}^-(w - L_k^{\bar{p}}(\Theta))\, I(u) = \mathrm{sc}^-(w - L_k^{\bar{p}}((0,1); l_k^{\bar{p}}))\, \hat{I}(\hat{P}u) = \\ \int_\Omega \hat{\phi}^{**}(x, (\hat{P}u)(g(x)))\, d\mu_\Omega(x), \end{aligned} \tag{17}$$

the latter equality being valid in view of (16).

Lemma 6.2. *Let X and $\hat{X}$ be normed spaces, $P\colon \hat{X} \to X$ and $\hat{P}\colon X \to \hat{X}$ be linear bounded operators satisfying $P\hat{P} = Id_X$. If $I\colon X \to \bar{\mathbb{R}}$ and $\hat{I}\colon \hat{X} \to \bar{\mathbb{R}}$ are such functionals that $I(P\hat{u}) = \hat{I}(\hat{u})$ for every $\hat{u} \in \hat{X}$, then*

$$sc^-(w - X)\, I(P\hat{u}) = sc^-(w - \hat{X})\, \hat{I}(\hat{u}). \tag{$*$}$$

In particular,

$$sc^-(w - X)\, I(u) = sc^-(w - \hat{X})\, \hat{I}(\hat{P}u). \tag{$**$}$$

Step 4. Set $\Omega_i^\nu = g^{-1}(\Lambda_i^\nu)$, $\Delta(x) = D(g(x))$, and $\omega_i^\nu(x) = w_i^\nu(g(x))$. Define a function $\bar\phi\colon \Delta \to \bar{\mathbb{R}}$ by the relationship

$$\bar\phi(x,\bar y) = \hat\phi(x, \hat A(g(x))\bar y).$$

Note that then always $\hat\phi(x,\hat y) = \bar\phi(x, \bar A(g(x))\hat y)$. In fact, to verify this claim, it is enough to show that

$$\hat\phi(x,\hat y) = \hat\phi(x, \hat A(g(x))A(g(x))\hat y).$$

The latter, by virtue of the g-measurability of the operator function $\hat A(g(x))A(g(x))$, follows from

$$\hat f(x,\hat y) = \hat f(x, \hat A(g(x))A(g(x))\hat y),$$

which is equivalent to

$$f(x, H(x)\hat y) = f(x, H(x)\hat A(g(x))\bar A(g(x))\hat y).$$

and holds by *(v)* of Lemma 6.1.

Further on, applying the lemma below which reformulates lemma 4.3 from [13] with $\hat X = l_k^{\bar p}$, $\bar X = D(g(x))$, $\hat F = \hat\phi(x,\cdot)$, $\bar F = \bar\phi(x,\cdot)$, $\hat A = \hat A(g(x))$, $\bar A = \bar A(g(x))$, we can state that

$$\hat\phi^{**}(x,\hat y) = \bar\phi^{**}(x, \bar A(g(x))\hat y)$$

and, in particular, that

$$\bar\phi^{**}(x,\bar y) = \hat\phi^{**}(x, \hat A(g(x))\bar y).$$

Lemma 6.3. *Let $\bar X$ and $\hat X$ be normed spaces, $\bar A\colon \hat X \to \bar X$ and $\hat A\colon \bar X \to \hat X$ be linear bounded operators satisfying $\bar A\hat A = Id_X$. If $\bar F\colon \bar X \to \mathbb{R}$ and $\hat F\colon \hat X \to \mathbb{R}$ are such nonnegative functionals that $\hat F(\hat u) = \bar F(\bar A\hat u)$ for every $\hat u \in \hat X$, then*

$$\hat F^{**}(\hat u) = \bar F^{**}(\bar A\hat u).$$

In particular, for every $\bar u \in \bar X$ one has

$$\bar F^{**}(\bar u) = \hat F^{**}(\hat A\bar u).$$

Step 5. Setting now $g_i^\nu(x) = \gamma_i^\nu(g(x))$, we obtain

$$\mathrm{sc}^-(w - L_k^{\bar p}(\Theta))\, I(u) = \int_\Omega \hat\phi^{**}(x, (\hat P u)(g(x)))\, d\mu_\Omega(x) =$$

$$\int_\Omega \hat\phi^{**}(x, \hat A(g(x))\{u_i(g_i^\nu(x))\}_{i,\nu=1}^{k,\infty})\, d\mu_\Omega(x) = \int_\Omega \bar\phi^{**}(x, \{u_i(g_i^\nu(x))\}_{i,\nu=1}^{k,\infty})\, d\mu_\Omega(x),$$

which ends the proof. □

We now concentrate on the proof of the lemma 6.1 which constitutes the heart of the proved results.

PROOF OF THE LEMMA 6.1:

Step 1. Since the functions γ_i satisfy the ω-condition, we can denote their injectivity sets by Θ_i^ν, $\nu = 1, 2, \ldots$, and set $\Lambda_i^\nu = \gamma_i(\Theta_i^\nu)$. The restrictions $\gamma_i|_{\Theta_i^\nu}$ are invertible and possess inverses $\gamma_i^\nu : \Lambda_i^\nu \to \Theta_i^\nu$.

Consider an arbitrary μ_Θ-nullset $e \subset \Theta_i^\nu$. The preimage $(\gamma_i^\nu)^{-1}(e) = \gamma_i(e)$, the latter being of zero Lebesgue measure since γ_i satisfies the ω-condition. Denote by $\mu_{\gamma_i^\nu}$ the induced measure over Θ_i^ν defined by $\mu_{\gamma_i^\nu}(e) := |(\gamma_i^\nu)^{-1}(e)|$. Since $\mu_{\gamma_i^\nu}$ is absolutely continuous with respect to μ_Θ, we can define functions $a^\nu\colon \Theta_i^\nu \to \mathbb{R}$ by

$$a_i^\nu(t) := \left(\frac{d\mu_{\gamma_i^\nu}}{d\mu_\Theta}(t)\right)^{1/p_i}.$$

Note that $a_i^\nu(t) > 0$ for μ_Θ-a.e. $x \in \Theta_i^\nu$ as well as the Radon-Nikodym derivative $\frac{d\mu_{\gamma^\nu}}{d\mu_\Theta}(t)$. In fact, otherwise there would exist a measurable $e \subset \Theta_i^\nu$ with $\mu_\Theta(e) > 0$ such that $|(\gamma_i^\nu)^{-1}(e)| = 0$, contradicting with $e \subset \gamma_i^{-1}((\gamma_i^\nu)^{-1}(e))$ and hence

$$\mu_\Theta(e) \leq \mu_\Theta(\gamma_i^{-1}((\gamma_i^\nu)^{-1}(e))) = 0.$$

At last define the functions $\bar{\nu}_i\colon \Theta \to \mathbb{N}$, $i = 1, \dots, k$, by setting $\bar{\nu}_i(t) := \nu$ whenever $t \in \Theta_i^\nu$. Note that $t \in \Theta_i^{\bar{\nu}_i(t)}$ for every $t \in \Theta$.

Step 2. We introduce now formally the operators $H(x)$, $\hat{H}(x)$ as indicated in the statement of the lemma being proved. The operators P, $\hat{P}$ will be defined on the next step of the proof. Define $H(x)\colon l_k^{\bar{p}} \to K(x)$ by the relationship

$$H(x)\hat{y} := \left\{ a_i^{\bar{\nu}_i(g_{ij}(x))}(g_{ij}(x))\hat{y}_i^{\bar{\nu}_i(g_{ij}(x))} \right\}_{i,j=1}^{k,l}.$$

It is easy to check that $H(x)$ maps $l_k^{\bar{p}}$ into $K(x)$ and is continuous.

The operator $\hat{H}(x)\colon K(x) \to K(x)$ is introduced by the relationship

$$\hat{H}_i^\nu(x)y := \begin{cases} y_{ij}/a_i^\nu(g_i(x)), & \text{if there exists } 1 \leq j \leq l : \nu = \bar{\nu}_i(g_{ij}(x)), \\ 0, & \text{otherwise.} \end{cases}$$

This definition is correct, since if for some $\nu \in \mathbb{N}$ and $x \in \Omega$ one has $g_{ij}(x) \in \Theta_i^\nu$ and $g_{ij'}(x) \in \Theta_i^\nu$, then $y_{ij} = y_{ij'}$. In fact, $\gamma_i(g_{ij}(x)) = \gamma_i(g_{ij'}(x))$, and, since Θ_i^ν is an injectivity set for γ_i, we have $g_{ij}(x) = g_{ij'}(x)$. Thus, $y_{ij} = y_{ij'}$ by definition of $K(x)$.

Further on, it is clear that $\hat{H}(x)$ maps $K(x)$ into $l_k^{\bar{p}}$ and is continuous between those two spaces as the image of any $y \in K(x)$ has only finite number of nonzero components by definition of $\hat{H}(x)$.

Step 3. We introduce operators $\bar{A}(s)$ and $\hat{A}(s)$ for all $s \in (0, 1)$. We set

$$w_i^\nu(s) = a_i^\nu(\gamma_i^\nu(s))$$

and define $\bar{A}(s)$ by

$$\bar{A}(s)_i^\nu \hat{y} = \hat{y}_i^\nu w_i^\nu(s),$$

and $\hat{A}$ by

$$\hat{A}_i^\nu(s)\bar{y} = \begin{cases} \bar{y}_i^\nu / w_i^\nu(s), & s \in \Lambda_i^\nu, \\ 0, & s \notin \Lambda_i^\nu. \end{cases}$$

Here we assumed $0 \cdot \infty = 0$ and $0/0 = 0$. The acting and continuity conditions of the operators $\bar{A}(s)$ and $\hat{A}(s)$ can be easily verified.

Step 4. We dedicate a separate step to a more delicate construction of the operators P and $\hat{P}$. In accordance with the statement of the lemma being proven, the operator $P\colon L_k^{\bar{p}}((0,1); l_k^{\bar{p}}) \to L_k^{\bar{p}}(\Theta)$ is defined by

$$(P\hat{u})_i(t) = \bar{A}_i^{\bar{\nu}_i(t)}(\gamma_i(t))\{\hat{u}_i^\nu(\gamma_i(t))\}_{i,\nu=1}^{k,\infty}$$

while the operator $\hat{P}\colon L^p(\Theta; \mathbb{R}^k) \to L^p((0,1); (l^{\bar{p}}k)$ is given by

$$(\hat{P}u)(s) = \hat{A}(s)\{u_i(\gamma_i^\nu(s))\}_{s \in \Lambda_i^\nu}.$$

Let us now check that P and $\hat{P}$ act continuously between the indicated spaces. Consider first the operator $\hat{P}$.

$$\begin{aligned}
||(\hat{P}u)_i||_p^p = \int_0^1 \sum_{\nu=1}^{\infty} \left|(\hat{P}u)_i^\nu(s)\right|^p ds \quad &= \quad \sum_{\nu=1}^{\infty} \int_{\Lambda_i^\nu} \left|\frac{u_i(\gamma_i^\nu(s))}{a_i^\nu(\gamma_i^\nu(s))}\right|^p ds \quad = \\
\sum_{\nu=1}^{\infty} \int_{\Theta_i^\nu} |u_i(t)|^p \, d\mu_\Theta(t) \quad &= \quad \int_{\bigcup_{\nu=1}^{\infty} \Theta_i^\nu} |u_i(t)|^p \, d\mu_\Theta(t) \quad = \\
\int_\Theta |u_i(t)|^p \, d\mu_\Theta(t) \quad &= \quad ||u_i||_p^p,
\end{aligned} \tag{18}$$

where in last but one equality the variable substitution $t = \gamma_i^\nu(s)$ was applied. In the estimate (18), we applied the Beppo Levi theorem (the measurability of $\hat{P}u$ is a simple exercise). Thus, the operator $\hat{P}$ acts between $L_k^{\bar{p}}(\Theta)$ and $L_k^{\bar{p}}((0,1); l_k^{\bar{p}})$ and even preserves the norm.

In a quite analogous way, one can verify that P maps $L_k^{\bar{p}}((0,1); l_k^{\bar{p}})$ into $L_k^{\bar{p}}(\Theta)$ and is continuous as an operator between these spaces.

Step 5. At this moment we are able to introduce the function $\hat{f}$ and verify its properties. Define $\hat{f}$: $\Omega \times l_k^{\bar{p}} \to \bar{\mathbb{R}}$ by setting

$$\hat{f}(x, \hat{y}) := f(x, H(x)\hat{y}).$$

Clearly, if $f(x, \cdot)$ is lower semicontinuous in $\mathbb{R}^{kl}$ for a.e. $x \in \Omega$, then so is $\hat{f}(x, \cdot)$ in $l_k^{\bar{p}}$. Hence, to complete the proof of *(ii)*, it remains to show that if f is an integrand, then so is $\hat{f}$. Observe that for this purpose it is enough to prove that the function

$$(x, \hat{y}) \in \Omega \times l_k^{\bar{p}} \mapsto (x, H(x)\hat{y}) \in \Omega \times \mathbb{R}^{kl}$$

is $(\Sigma(\Omega) \otimes \mathcal{B}(l_k^{\bar{p}}), \Sigma(\Omega) \otimes \mathcal{B}(\mathbb{R}^{kl}))$-measurable. To check this statement, we observe first that the function $(x, \hat{y}) \in \Omega \times l_k^{\bar{p}} \mapsto x \in \Omega$ is obviously $(\Sigma(\Omega) \otimes \mathcal{B}(l_k^{\bar{p}}), \Sigma(\Omega))$-measurable. Hence it is enough to show that for every fixed $i \in \{1, \ldots, k\}$ the function ψ_i: $\Omega \times l_k^{\bar{p}} \to \mathbb{R}$ defined by

$$\psi_{ij}(x, \hat{y}) := a_i^{\bar{\nu}_i(g_{ij}(x))}(g_{ij}(x))\hat{y}_i^{\bar{\nu}_i(g_{ij}(x))},$$

is $(\Sigma(\Omega) \otimes \mathcal{B}(l_k^{\bar{p}}), \mathcal{B}(\mathbb{R}^{kl}))$-measurable. Fixing an open set $U \subset \mathbb{R}$, we calculate

$$\begin{aligned}
\psi_{ij}^{-1}(U) \quad = \quad & \bigcup_{\nu=1}^{\infty} \left\{(x, \hat{y}) \in \Omega \times l_k^{\bar{p}} \,:\, \bar{\nu}_i(g_{ij}(x)) = \nu \text{ and } a_i^\nu(g_{ij}(x))\hat{y}_i^\nu \in U\right\} = \\
& \bigcup_{\nu=1}^{\infty} \left(g_{ij}^{-1}(\Theta_i^\nu) \times l_k^{\bar{p}}\right) \cap \left\{(x, \hat{y}) \in \Omega \times l_k^{\bar{p}} \,:\, a_i^\nu(g_{ij}(x))\hat{y}_i^\nu \in U\right\}.
\end{aligned}$$

The latter evidently belongs to $\Sigma(\Omega) \otimes \mathcal{B}(l_k^{\bar{p}})$, which proves the statement.

Step 6. Let us first prove (v). Note that

$$g(x) = \gamma_i(g_{ij}(x)) \in \gamma_i(\Theta_i^{\bar{\nu}_i(g_{ij}(x))}) = \Lambda_i^{\bar{\nu}_i(g_{ij}(x))},$$

therefore by definition of operators $\bar{A}$ and $\hat{A}$ we have

$$\left(\hat{A}(g(x))\bar{A}(g(x))\hat{y}\right)_i^{\bar{\nu}_i(g_{ij}(x))} = \hat{y}_i^{\bar{\nu}_i(g_{ij}(x))}.$$

But then

$$H(x)\hat{A}(g(x))\bar{A}(g(x))\hat{y} = H(x)\hat{y},$$

as for the operator $H(x)$ only coordinates $\hat{y}_i^{\bar{p}_i(g_{ij}(x))}$ matter.

Now we prove (14), and hence also the most crucial claim *(i)*. In fact, (14) and *(ii)* directly imply

$$T(P\hat{u})(x) = \hat{f}(x, \hat{u}(g(x))), \tag{19}$$

for all $\hat{u} \in L_k^{\bar{p}}((0,1); l_k^{\bar{p}})$. Then

$$T(u)(x) = \hat{f}(x, (\hat{P}u)(g(x)))$$

is an immediate consequence from the latter due to *(iii)*. Finally, (14) is obtained by a straightforward computation. □

References

[1] J. Aaronson. *An Introduction to Infinite Ergodic Theory*, volume 50 of *Mathematical Surveys and Monographs*. American Math. Soc., 1997.

[2] J.-M. Bismut. Intégrales convexes et probabilités. *J. Math. Anal. and Appl.*, 42:639–673, 1973. in French.

[3] G. Buttazzo. *Semicontinuity, Relaxation and Integral Representation in the Calculus of Variations*, volume 207 of *Pitman research notes in mathematics*. Longman Scientific, Harlow, 1989.

[4] C. Castaing and M. Valadier. *Convex analysis and measurable multifunctions*, volume 580 of *Lecture notes in mathematics*. Springer-Verlag, Berlin, 1977.

[5] G. Dal Maso. *An Introduction to* Γ*-convergence*. Birkhäuser, Boston, 1993.

[6] J. Diestel and L. Uhl. *Vector Measures*. American Math. Soc., Providence, R.I., 1968.

[7] M.E. Drakhlin. On one linear functional equation. In *Functional Differential Equations*, pages 91–111. Perm Polytechnical Institute, Perm, 1985. in Russian.

[8] I. Ekeland and R. Temam. *Convex Analysis and Variational Problems*, volume 1 of *Studies in Mathematics and Its Applications*. North Holland Publ. Co. Amsterdam, Oxford, 1976.

[9] I.I. Gikhman and A.V. Skorokhod. *The Theory of Stochastic Processes*. "Nauka", Moscow, 1975. in Russian. English transl.: Springer-Verlag, Berlin, 1980.

[10] R.P. Halmos. *Measure Theory*, volume 18 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, 1974.

[11] T. Roubíček. *Relaxation in optimization theory and variational calculus*. Walter de Gruyter & Co., Berlin, 1997.

[12] E. Stepanov. Relaxation and convexity of functionals with pointwise nonlocality. *Proc. AMS*, 2001. to appear.

[13] E. Stepanov and A. Zdorovtsev. Relaxation of some nonlocal integral functionals in weak topology of Lebesgue spaces. *J. Convex Anal.*, 8(2):1–23, 2001. to appear.

[14] J. Warga. *Optimal control of differential and functional equations*. Academic Press, New York, 1972.

[15] J. Warga. Problems with delayed and shifted controls. In A. Ioffe, S. Reich, and I. Shafrir, editors, *Calculus of Variations and Optimal Control*, volume 411 of *Research Notes in Mathematics*, pages 251–269, London, 1999. Chapman and Hall.

Dipartimento di Matematica, Universitá di Pisa, via Buonarroti 2, 56127 Pisa, Italy

E-mail address: `stepanov@cibs.sns.it`