Journal of Convex Analysis
Volume 8 (2001), No. 2, 447–469

Relaxation of Some Nonlocal Integral Functionals in Weak Topology of Lebesgue Spaces

Eugene Stepanov
Computer Technology Department, St. Petersburg Institute of Fine Mechanics and Optics, 14 Sablinskaya ul., 197101 St. Petersburg, Russia
stepanov@spb.runnet.ru

Andrei Zdorovtsev
Computer Technology Department, St. Petersburg Institute of Fine Mechanics and Optics, 14 Sablinskaya ul., 197101 St. Petersburg, Russia
zdrvtsff@spb.runnet.ru

Received February 16, 2000
Revised manuscript received December 12, 2000

We are studying the relaxation of the integral functional involving argument deviations

$$I(u) := \int_\Omega f(x, u(g_1(x)), \dots, u(g_k(x)))\, dx,$$

in weak topology of a Lebesgue space $L^p(\Omega)$, $1 < p < +\infty$, with open bounded $\Omega \subset \mathbb{R}^n$. It is proven that, unlike the classical case without deviations, the relaxed functional in general cannot be obtained as convexification of the original one. However, we show that if the set functions $g_i \colon \Omega \to \Omega$ satisfies certain condition (called *unifiability*), which is just a natural extension of nonergodicity property of a single function to sets of functions, and which is automatically satisfied when $k = 1$, then the relaxed functional is equal to the convexification of the original one. We show that the unifiability requirement is essential for such a convexification result for a generic integrand. Further slightly restricting this condition, we also obtain the nice representation of the relaxed functional in terms of convexification of some new integrand, but involving in general countably many new argument deviations.

1991 Mathematics Subject Classification: 49J45, 47B38, 47H30

1. Introduction

This paper is concerned with the relaxation in the weak topology of Lebesgue spaces $L^p \equiv L^p(\Omega)$ of integral functionals $I \colon L^p \to \mathbb{R} \cup \{+\infty\} \equiv \mathbb{R}^{\cdot}$ of the type

$$I(u) := \int_\Omega f(x, u(g_1(x)), \dots, u(g_k(x)))\, dx. \tag{1}$$

Here $\Omega \subset \mathbb{R}^n$ is an open bounded set, $f \colon \Omega \times \mathbb{R}^k \to \mathbb{R}^{\cdot}$ is an *integrand* (i.e. $\Sigma \otimes \mathcal{B}(\mathbb{R}^k)$ measurable, Σ standing for the σ-algebra of Lebesgue measurable subsets of Ω, $\mathcal{B}(\mathbb{R}^k)$ standing for the Borel σ-algebra of $\mathbb{R}^k$), and all $g_i \colon \Omega \to \Omega$, $i = 1, \dots, k$ are measurable functions satisfying

$$e \subset \Omega, |e| = 0 \Rightarrow \mu_{g_i}(e) := |g_i^{-1}(e)| = 0,$$

$|\cdot|$ standing for the Lebesgue measure. The latter condition is obviously necessary to define the compositions $u \circ g_i$, for each $u \in L^p$ stands for a *class* of equivalent functions.

ISSN 0944-6532 / \$ 2.50

The functionals of the above type seem to be a rather unusual object in the classical calculus of variations. They arize rather naturally in the study of various functional differential equations (FDE's) having variational structure (an overview of the literature and of some results on the subject may be found in [11]), as well as in optimal control problems for for FDE's with deviating argument [4]. It is the latter source that is motivating for our study. In fact, it is well-known that many classical optimal control problems for ordinary or partial differential equations are easily reduced to a purely variational formulation like (1) but with $k = 1$ and $g_1 = \mathrm{id}_\Omega$ the identity in Ω (see, for instance, [3]). This appears to be possible because such problems involve only *local* operators like Nemytskiĭ or differentiation operators. However, if the underlying state equations of the optimal control setting are not local, e.g. are equations with deviating argument, one can only reduce the problem to the general variational formulation (1).

It is clear, that the problem of writing out an explicit relaxation of (1) in weak topology of L^p presents essential difficulties. In fact, in the well-studied classical case $k = 1$ with $g_1 = \mathrm{id}_\Omega$, sometimes (not quite correctly) referred to as *local*, it is known that the relaxation of the functional in weak topology of L^p is equal to its convexification, and, moreover, can be obtained by means of convexification of the integrand in the second variable. In this paper we show that for general functionals with deviating argument this is not the case. Namely, in general one even cannot expect to obtain the relaxation by convexifying the whole functional, to say nothing of convexifying the integrand. Nevertheless, restricting ourselves to a rather large class of such functionals involving "unifiable" set of deviations g_i (we give an exact definition below), we are able to prove that still the relaxed functional coincides with the convexification of the original one. Note that we assume by definition that a set consisting of a single function g_1 is automatically unifiable, and hence in the case $k = 1$ one always has the representation of a relaxed functional by means of convexification of the original one. Moreover, in this case the relaxation admits a representation in terms of some new convexified integrand. Much more difficult is the general case $k \geq 2$. Here the requirement of unifiablity is rather strong and in practice is nothing else than a generalization of the nonergodicity property of a single function to sets of functions. It is important to emphasize that it characterizes the collective behavior of functions. We show that it is quite unavoidable, if one wants to have a general representation result for a relaxed functional in terms of convexification of the original one. Also, slightly restricting this condition, we are able to prove the representation of the relaxed functional by an integral functional with deviating argument, involving a convexification of some new "integrand" and, generally speaking, a countable sequence of new argument deviations.

At last let us remark that all the results we provide hold also in more general measure situations. Namely, an open bounded subset in $\mathbb{R}^n$ with Lebesgue measure may be easily replaced by a standard measure space with finite nonatomic measure. We try however here to avoid such additional complications.

2. Notation and preliminaries

Let (X, τ) be a topological space. The relaxation of the functional $I\colon X \to \mathbb{R}^{\cdot}$ in the topology τ defined as the maximum τ-lower semicontinuous functional over X not greater than I will be denoted by $sc^-(\tau)I$. If τ is a weak topology of a normed space X then this we will denote the respective relaxation by $sc^-(w - X)I$. It is well-known (see [3]) that if X is a reflexive Banach space, then the latter relaxation coincides with the sequential one and, in particular, is uniquely characterized by the following two properties:

(i) if $u_\nu \rightharpoonup u$ weakly in X, then $sc^-(w-X)I(u) \leq \liminf_\nu I(u_\nu)$;

(ii) for every $u \in X$ there is a sequence $\{u_\nu\} \subset X$, such that $u_\nu \rightharpoonup u$ weakly in X, and $sc^-(w-X)I(u) = \lim_\nu I(u_\nu)$.

The same is true, if X is dual to a separable Banach space, while the weak topology/convergence are substituted by $*$-weak ones.

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: $X \to \mathbb{R}^{\cdot}$ is a new functional I^*: $X^* \to \mathbb{R}^{\cdot}$ defined by the formula

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

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

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

Further on we will always assume that there is at least one continuous affine functional less than I (in particular, this is true when I is nonnegative). It is well-know (see theorem I.5 of [5]) that in this case I^{**} is the convexification of I in the sense that it is the maximum lower semicontinuous in the strong topology and convex functional not greater than I.

The linear *shift* (composition, inner superposition) operators, is defined formally by the relationship

$$(T_g u)(x) := u(g(x)),$$

where g: $\Omega \to \Omega$ is a given function. The operator T_g is well-defined by the above formula over Lebesgue spaces, if and only if the generating function g is measurable and satisfies the additional requirement

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

which we will always consider fulfilled in the sequel. Although unlike the Nemytskiĭ operator, the shift T_g is not *local* (see [16] for the precise definition of a local operator), it has many properties similar to those of local operators. For instance, the shift T_g in Lebesgue spaces is never compact and posesses some deterioration properties, i.e. maps smaller spaces into larger ones [1, 10]. However, under only the condition (2) the operator T_g maps L^∞ in itself and is automatically bounded over this space. We find it important to remark that in spite of its intrinsically variational nature, the question of relationship between the relaxation of the functional (1) in weak topology of L^p and its convexification is intimately related to the operator theory. In fact, it depends on an interplay between the properties of *Nemytskiĭ* operator, and linear shifts.

3. Relaxation versus convexification

The study of the relaxation of the functional (1) presents an essential difficulty in comparison with the classical "local" case without argument deviations. In fact, in the latter case (i.e. when $k = 1$, $g_1(x) = x$) it is well-known that under some not very restrictive conditions on the functional and/or on the integrand the relaxation of the functional in the weak topology of L^p coincides with its convexification and, moreover, can be obtained by convexifying the integrand in the last variable (see Chapter IX of [12]). However, we will show here that there is no hope to obtain the relaxation of a general functional (1) by means of any convexification technique. First of all, consider the example below which involves very pathological argument deviations.

Example 3.1. *Let $\Omega = (0,1)$. Each $x \in \Omega$ can be written in a binary system*

$$x = (0, x_1 x_2 x_3 \dots x_j \dots)_2, \text{ where } x_j \in \{0,1\},\ j \in \mathbb{N}.$$

Define the maps

$$\begin{aligned} g_1(x) &:= (0, x_1 x_3 \dots x_{2j-1} \dots)_2, \\ g_2(x) &:= (0, x_2 x_4 \dots x_{2j} \dots)_2, \end{aligned}$$

It is easy to prove following the lines of a similar example in [6] that these maps are measure preserving (in particular, they satisfy the condition (2)) and, what is more important, the respective families of σ-subalgebrae $g_i^{-1}(\Sigma)$ generated by g_i, $i = 1, 2$, are independent. Define the functional $I\colon L^2 \to \mathbb{R}$ by the formula

$$I(u) := -\int_\Omega u(g_1(x))u(g_2(x))\,dx. \tag{3}$$

In probabilistic terminology, $-I$ is the expectation of the product of two independent random variables $u \circ g_i$. Hence,

$$I(u) = -\int_\Omega u(g_1(x))\,dx \int_\Omega u(g_2(x))\,dx = -\left(\int_\Omega u(x)\,dx\right)^2,$$

where the measure preserving property of g_i has been used. Thus I is weakly lower semicontinuous (and even weakly continuous) but not convex.

Clearly, in the above example there is nothing specific to an interval $(0,1)$. In fact, it has been shown in [6] using the theorem on isomorphism for measure spaces [14], that in a standard probability space (Ω, Σ, μ) for each fixed $k \in \mathbb{N}$ there exists a finite collection of measure preserving maps

$$g_i : \ \Omega \to \Omega, \qquad i = 1, \dots, k$$

such that a family of σ-subalgebrae $\{(g_i)^{-1}(\Sigma)\}_{i=1}^k$ is independent in totality. This gives a possibility to construct, following the lines of the above example, more general weakly continuous functionals with deviating argument which are neither convex nor concave. However, such examples might be rather confusing. In fact, it is worth emphasizing that the maps g_i satisfying the above condition have some rather pathological properties, e.g. they can be injective only on a set of zero measure, while the preimage of almost any point is nondenumerable. Such properties possess, for instance, Wiener shifts along the trajectories of a standard Brownian motion. One might have think then that for nicer argument deviations (e.g. piecewise injective ones) everything works like in the classical "local" case. Unfortunately this is not true, even for very nice argument deviations, as we will show below. Therefore, what we will need is in fact not a property of a single function, but a collective property of the set of functions $g_1, \dots, g_k$. We will call such nice property "unifiability" and will prove that for functionals of the type (1) involving unifiable set of argument deviations the relaxation in weak topology of L^p still coincides with convexification as in the "local" case.

In the following definition we introduce preliminary notions necessary to further specify what exactly is meant by unifiable function set.

Definition 3.2. We say that a measurable function $g\colon \Omega \to \Omega$

(i) is piecewise injective, 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;
(ii) satisfies ω-condition, if it is piecewise injective, satisfies (2), and the respective inverses $\gamma_j\colon g(\Omega_j) \to \Omega_j$, $\gamma_j \circ g|_{\Omega_j} = \mathrm{id}$, satisfy (2).

The notion of ω-condition has been introduced by M.E. Drakhlin in [9] to study a class of shifts in Lebesgue spaces which have "nice" representation of adjoints.

Remark. Note that a piecewise injective function $g\colon \Omega \to \Omega$ satisfying (2) can always be changed on a set of zero measure in order to satisfy the ω-condition. In fact, consider the measures μ_j on Ω defined by

$$\mu_j(e) := |g^{-1}(e) \cap \Omega_j|$$

for every measurable $e \subset \Omega$. Since μ_j are absolutely continuous with respect to the Lebesgue measure, we may consider the respective Radon-Nikodym derivatives $\frac{d\mu_j}{d\mathcal{L}^n}$ and set

$$E_j := g^{-1}\left(\left\{y \,:\, \frac{d\mu_j}{d\mathcal{L}^n}(y) = 0\right\}\right) \bigcap \Omega_j.$$

Setting now

$$\tilde{g}(x) := \begin{cases} g(x), & x \in \Omega \setminus \bigcup_j E_j, \\ x, & x \in \bigcup_j E_j, \end{cases}$$

and observing that $|E_j| = 0$, one arrives at a conclusion that $g(x) = \tilde{g}(x)$ a.e. in Ω, while $\tilde{g}$ satisfies the ω-condition. Therefore, everywhere in the sequel we identify the piecewise injective functions satisfying (2) with the functions satisfying ω-condition.

We would like to mention that a function which does not satisfy ω-condition is a real pathology, and in fact given such a function one can easily construct an example of a functional of the type (1) with the same properties as in example 3.1. To show this, let us enlist some properties of such functions.

Proposition 3.3. *The following statements about the function $g\colon \Omega \to \Omega$ satisfying (2) are equivalent:*

(i) *g is piecewise injective;*
(ii) *for almost all $x \in \Omega$ the inverse image $g^{-1}(x)$ is at most countable.*

Also, if g is not piecewise injective, then there is a measurable $E \subset \Omega$, $|E| > 0$, and a function $h\colon E \to \Omega$ such that the σ-algebrae $g^{-1}(\Sigma) \cap E$ and $h^{-1}(\Sigma)$ are independent.

Remark. It is clear that in particular, every locally Lipschitz continuous function is piecewise injective.

Proof. $(i) \Rightarrow (ii)$ is trivial. We prove thus $\neg(i) \Rightarrow \neg(ii)$. Suppose that (i) does not hold. Restricting, if necessary, to a set of full measure in Ω, we may suppose, that g is Borel measurable. Let $E' \subset \Omega$, $|E'| > 0$ be such that the restriction of g to E' is antiinjective, that is, for every measurable $e \subset E'$ one has $|e| = 0$ whenever $g\,|_e$ is injective. Suppose also without loss of generality that $|E'| = 1$ (otherwise just renorm the measure). By the proposition 2.2 from [15], which is a consequence of Maharam theorem on homogeneous measure algebrae, there exist then a Borel set $E \subset E'$, $|E| = |E'| = 1$, a compact metric space M, a nonatomic Borel probability measure $\mathbb{P}$ on M, and a Borel measurable map $\tau\colon E \to M$ such that the following conditions hold

(a) the map $g \times \tau\colon E \to \Omega \times M$ defined by

$$(g \times \tau)(x) := (g(x), \tau(x))$$

is invertible;

(b) the image measure μ' over $\Omega \times M$ defined by

$$\mu'(B) := |(g \times \tau)^{-1}(B)|$$

for all Borel sets $B \subset \Omega \times M$ satisfies

$$\mu' = \mu_g \otimes \mathbb{P}.$$

Note that here and below g is identified with its restriction to E, so that $g\colon E \to \Omega$. Consider now the projection maps $p_M\colon \Omega \times M \to M$ and $p_\Omega\colon \Omega \times M \to \Omega$. Clearly,

$$g(x) = p_\Omega((g \times \tau)(x)), \qquad \tau(x) = p_M((g \times \tau)(x)). \tag{4}$$

From this representation of g and from (a) follows that for all $x \in E$ the inverse image $g^{-1}(x)$ is uncountable, since so is M. In other words, this means that *(ii)* does not hold.

A similar argument proves the last claim of the proposition. In fact, by the isomorphism theorem [14] there is a bijection $j\colon M \to \Omega$ providing the isomorphism of the respective measure spaces (with Borel σ-algebrae). Define now $h\colon E \to \Omega$ by $h := j \circ \tau$. We will prove that the σ-algebrae $h^{-1}(\mathcal{B})$ and $g^{-1}(\mathcal{B})$, where $\mathcal{B}$ is the Borel σ-algebra of Ω, are independent.

To verify the latter claim, choose arbitrarily $B_1, B_2 \in \mathcal{B}$. We will abbreviate $j^{-1}(B_2)$ by $\hat{B}_2$. One has then

$$\begin{aligned}|g^{-1}(B_1) \cap h^{-1}(B_2)| &= \mu'((g \times \tau)(g^{-1}(B_1) \cap \tau^{-1}(\hat{B}_2))) = \\ &\mu'(p_\Omega^{-1}(B_1) \cap p_M^{-1}(\hat{B}_2)) = (\mu_g \otimes \mathbb{P})(p_\Omega^{-1}(B_1) \cap p_M^{-1}(\hat{B}_2)) = \\ &(\mu_g \otimes \mathbb{P})(B_1 \times \hat{B}_2) = \mu_g(B_1)\mathbb{P}(\hat{B}_2),\end{aligned}$$

where (a), (b) and (4) were used. Using the same formula with $B_1 := \Omega$, we have

$$|h^{-1}(B_2)| = \mathbb{P}(\hat{B}_2), \tag{5}$$

and recalling that $\mu_g(B_1) := |g^{-1}(B_1)|$, we get finally

$$|g^{-1}(B_1) \cap h^{-1}(B_2)| = |g^{-1}(B_1)| \cdot |h^{-1}(B_2)|,$$

thus showing the desired independence. $\square$

Now the proposition 3.3 provides the way to generalize the example 3.1. In fact, it immediately implies that if $g\colon \Omega \to \Omega$ satisfies (2) but fails to satisfy the ω-condition, then the functional $I\colon L^\infty \to \mathbb{R}$ defined by

$$I(u) := -\int_\Omega 1_E(x)u(g(x))u(h(x))\,dx,$$

where 1_E stands for the characteristic function of E, is $*$-weakly continuous in L^∞ but not convex.

We stop with the examples of "pathological" functions and introduce the following definition which makes precise the idea of how nice the set of functions g_i should be in order that the relaxation of the functional (1) in weak topology of L^p be obtained by means of convexification. Note that in view of the above examples this should be restrictive, but, as we will show below, only for $k \geq 2$, while if $k = 1$, i.e. when the functional involves only one deviation, everything is similar to the "local" case without argument deviations. This means that when $k = 1$, then the relaxation of the functional (1) in weak topology of L^p coincides with the convexification of this functional whatever the deviation g_1 is.

Definition 3.4. The set of functions $g_1, \ldots, g_k\colon \Omega \to \Omega$ is called unifiable if there is a measurable function $\gamma\colon \Omega \to \Omega$ (called unifier) satisfying (2) such that

$$\gamma(g_1(x)) = \ldots = \gamma(g_k(x))$$

for a.e. $x \in \Omega$. If a unifier can be chosen to satisfy also the ω-condition, then the respective set of functions is called ω-unifiable. The set consisting of a single function ($k = 1$) is always said to be ω-unifiable with unifier id_Ω.

It is important to emphasize that unlike ω-condition, unifiability is not an individual property of a function but a property of sets of functions. In fact, not every set of functions even satisfying the ω-condition is unifiable. The simplest example of two nonunifiable functions is given by $g_1 := \mathrm{id}_\Omega$ and $g_2\colon \Omega \to \Omega$ any *ergodic* map. In fact, the ergodicity of g_2 means that the only function γ satisfying $\gamma(x) = \gamma(g_2(x))$ for a.e. $x \in \Omega$ is a constant [13]. A classical particular example of the ergodic map is the rotation of a circle $\Omega = S^1$ (equipped with the one-dimensional Hausdorff measure $|\cdot|$) by the angle $2\pi\alpha$, where α is irrational. Therefore, the unifiability can be understood as the natural extension of *nonergodicity* property to sets of functions.

It is rather easy to show now that the property of unifiability is unavoidable in order to be able to represent the relaxation of the functional (1) in weak topology of L^p as the convexification of the latter. Consider for this purpose the following construction.

Example 3.5. *Consider the simplest nonunifiable function pair $g_1 := id_\Omega$ and $g_2\colon \Omega \to \Omega$ any* ergodic *map. Define $f\colon \mathbb{R}^2 \to \mathbb{R}^\cdot$ by*

$$f(y_1, y_2) := \begin{cases} 0, & y_1 = y_2 \geq 1 \text{ or } y_1 = y_2 \leq 0, \\ +\infty, & \text{elsewhere.} \end{cases}$$

Consider the functional

$$I(u) := \int_\Omega f(u(g_1(x)), u(g_2(x)))\,dx$$

and observe that $I(u) = 0$, only if $u \circ g_1 = u \circ g_2$ a.e. in Ω, but this can happen only if $u = const$ in view of ergodicity of g_2. Hence

$$I(u) = \begin{cases} 0, & u = const \geq 1 \text{ or } u = const \leq 0, \\ +\infty, & \text{elsewhere.} \end{cases}$$

Now observe that I is weakly lower semicontinuous, i.e. $I = sc^-(w - L^p)I$, but not convex, namely $I \neq I^{**}$. As a matter of a fact,

$$I^{**}(u) = \int_\Omega f^{**}(u(g_1(x)), u(g_2(x)))\, dx.$$

It is easy to note that both functions g_1 and g_2 involved can themselfes be very nice, even injective.

The following theorem is the first principal result of the paper.

Theorem 3.6. *Let $1 < p \leq +\infty$. If the set of functions $g_1, \dots, g_k \colon \Omega \to \Omega$ is unifiable, then*

$$sc^-(w - L^p)I(u) = I^{**}(u).$$

In the case $p = +\infty$ the relaxation is meant in the $*$-weak topology.

Let us emphasize again that according to the definition of unifiability 3.4 in the case $k = 1$ this theorem implies no extra condition on g_1.

Proof. The announced claim will be shown, if we prove that $sc^-(w - L^p)I$ is convex. To show the latter, consider arbitrary $\{u_1, u_2\} \subset L^p$ and let the sequences $\{u^1_\nu\}$, $\{u^2_\nu\}$ be such that

$$u^i_\nu \rightharpoonup u_i \text{ weakly in } L^p, \text{ while } sc^-(w - L^p)I(u_i) = \lim_\nu I(u^i_\nu), \qquad i = 1, 2.$$

Pick up a countable dense set $\{p_j\} \subset L^{p'}$, and fix an arbitrary $\lambda \in (0, 1)$. Let $\gamma \colon \Omega \to \Omega$ be a unifier of the set $g_1, \dots, g_k$ (in particular, identity, if $k = 1$). Furthermore, set $g := \gamma \circ g_1$. Consider for each fixed $\nu \in \mathbb{N}$ the following systems of equations with respect to $c_\nu \in L^\infty$:

$$\begin{cases} \displaystyle\int_\Omega p_j(x)(T_\gamma c_\nu)(x) u^1_\nu(x)\, dx = \lambda \int_\Omega p_j(x) u^1_\nu(x)\, dx, \\ \displaystyle\int_\Omega p_j(x)(T_\gamma(1 - c_\nu))(x) u^2_\nu(x)\, dx = (1 - \lambda) \int_\Omega p_j(x) u^2_\nu(x)\, dx, \\ \displaystyle\int_\Omega (T_g c_\nu)(x) f(x, u^1_\nu(g_1(x)), \dots, u^1_\nu(g_k(x)))\, dx = \\ \qquad \displaystyle\lambda \int_\Omega f(x, u^1_\nu(g_1(x)), \dots, u^1_\nu(g_k(x)))\, dx, \\ \displaystyle\int_\Omega (T_g(1 - c_\nu))(x) f(x, u^2_\nu(g_1(x)), \dots, u^2_\nu(g_k(x)))\, dx = \\ \qquad \displaystyle(1 - \lambda) \int_\Omega f(x, u^2_\nu(g_1(x)), \dots, u^2_\nu(g_k(x)))\, dx, \end{cases} \tag{6}$$

where $j = 1, \dots, \nu$. This system represents, in fact, a momentum problem with respect to the unknown c_ν, since it is linear and finite-dimensional. Since it admits at least one solution $c_\nu \equiv \lambda$, satisfying $0 < c_\nu < 1$ a.e. in Ω, then by the Lyapunov convexity theorem (theorem in the Appendix to Chapter IV, Section 4 of [5]) it has also another solution $c_\nu = 1_{e_\nu}$, where 1_A stands for a characteristic function of a measurable $A \subset \Omega$. Now, by the first and second equations in (6) one has

$$1_{\gamma^{-1}(e_\nu)} u_\nu^1 + 1_{\Omega\setminus\gamma^{-1}(e_\nu)} u_\nu^2 \rightharpoonup \lambda u_1 + (1-\lambda) u_2$$

weakly in L^p, and hence, by definition of $sc^-(w - L^p)I$ with the help of the third and the fourth equations of (6) we obtain

$$\begin{aligned} sc^-(w-L^p)I(\lambda u_1 + (1-\lambda)u_2) \leq \liminf_\nu I(1_{\gamma^{-1}(e_\nu)} u_\nu^1 + 1_{\Omega\setminus\gamma^{-1}(e_\nu)} u_\nu^2) = \\ \lambda \lim_\nu I(u_\nu^1) + (1-\lambda)\lim_\nu I(u_\nu^2) = \\ \lambda\, sc^-(w-L^p)I(u_1) + (1-\lambda)\, sc^-(w-L^p)I(u_2), \end{aligned}$$

concluding the proof for the case $p \neq +\infty$. If $p = +\infty$, then the word-to-word restating of the same proof substituting the weak topology by $*$-weak one shows the desired result. □

Let us remark now that the theorem 3.6 provides conditions under which the relaxation of the functional (1) in weak topology of L^p coincides with the convexification of the latter for a generic integrand f. In this case, as indicated in the example 3.5 the unifiability condition is essential. It is however not necessarily so for particular integrands. For instance, it is only a matter of a slight modification of the proof of the theorem 3.6 (namely, of the system of equations (6)) to show that if

$$f(x, y_1, \ldots, y_k) = f_1(x, y_1, \ldots, y_l) + f_2(x, y_{l+1}, \ldots, y_k)$$

for some $l \in \{1, \ldots, k\}$, then the statement of the theorem 3.6 remains valid under only the condition that each of the two sets of functions $g_1, \ldots, g_l$ and $g_{l+1}, \ldots, g_k$ be unifiable itself.

4. Representation of the relaxed functional

As we already mentioned in the section 3, in the classical "local" case without argument deviations (i.e. when $k = 1$, $g_1 = \mathrm{id}_\Omega$) one easily obtains under some rather weak requirements on the integrand also a representation of the relaxed functional by means of certain convexification of the integrand. Intuitively, in view of the theorem 3.6 we may expect a similar result in our general case only if the argument deviations g_i involved are unifiable, (which in particular is automatic, if $k = 1$). This happens to be true, as we show below, with the two corrections. First, the notion of ω-unifiability should be involved here instead of simple unifiability. Second, even if $k = 1$ but g_1 is not an identity function, then as we will show, the weak lower semicontinuity of the functional (1) does not imply any kind of convexity of the integrand in the last variable. Hence even in such a simple case, as opposed to the local setting, one cannot hope to obtain the relaxation by convexifying the original integrand. nevertheless, we will show that everything works well if one uses instead the convexification of some new integrand obtained in quite a constructive manner. Therefore, even the simple setting with $k = 1$ in comparison with the "local" case presents essential difficulties. One encounters even more difficulties in the general case $k \geq 2$. We show then that in a relaxed setting not only convexification of a new "integrand" should be involved, but also the argument deviations will change, and in general one will have a countable number of the latter.

Further on we make the following natural additional assumptions:

(a) the integrand $f\colon \Omega \times \mathbb{R}^k \to \mathbb{R}^{\cdot}$ is normal, that is, $f(x, \cdot)$ is lower semicontinuous for a.e. $x \in \Omega$, and nonnegative;

(b) the functional I is finite over every constant function, that is

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

Both requirements are not very restrictive, and can further be weakened.

First of all we prove the following simple representation result for functionals involving only one deviation.

Theorem 4.1. *Let $1 \leq p < +\infty$. If $k = 1$, then there is a new integrand $\psi\colon \Omega \times \mathbb{R} \to \mathbb{R}^{\cdot}$ such that*

$$sc^{-}(w - L^p)I(u) = \int_\Omega \psi^{**}(x, u(g_1(x)))\, dx,$$

where ψ^{**} stands for the convexification of ψ in the last variable, that is,

$$\begin{array}{rcl} \psi^*(x, y_1') & := & \sup_{y_1 \in \mathbb{R}} \left(y_1 y_1' - \psi(x, y_1) \right), \\ \psi^{**}(x, y_1) & := & (\psi^*)^*(x, y_1). \end{array}$$

Remark. It will be clear from the proof that the new integrand ψ can be obtained in a quite constructive way from the original one by taking the conditional expectation of the latter with respect to the σ-algebra generated by the function g_1.

Proof. Without loss of generality assume $f(x, 0) = 0$ and consider

$$\psi(x, y) := E(f(\cdot, y); g_1^{-1}(\Sigma))(x)$$

to be the conditional expectation of $f(\cdot, y)$ with respect to the σ-algebra $g_1^{-1}(\Sigma)$. Since $f(\cdot, y) \in L^1$ by assumption, then by a known property of a conditional expectation (see p. 49 in vol. 1 of [13])

$$E(f(\cdot, u(g_1(\cdot))); g_1^{-1}(\Sigma))(x) = \psi(x, u(g_1(x))),$$

and therefore

$$I(u) = \int_\Omega \psi(x, u(g_1(x)))\, dx. \tag{7}$$

Let us remark that the above construction should be appropriately understood. In fact, the operator of conditional expectation is defined between *classes* of a.e. equal measurable functions, while different functions $\psi(\cdot, z)$ belonging to the same class can generate different integral functionals of the type (7). The above construction should hence be understood 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).

Since by definition of ψ there is a $\varphi\colon \Omega \times \mathbb{R} \to \mathbb{R}^{\cdot}$ such that $\psi(x, y) = \varphi(g_1(x), y)$, then changing the variables in (7) we obtain that

$$I(u) = \int_\Omega \frac{d\mu_{g_1}}{d\mathcal{L}^n}(x) \varphi(x, u(x))\, dx.$$

Applying now the proposition 2.3 of [12, chapter IX], we have

$$sc^-(w-L^p)I(u)=\int_\Omega \frac{d\mu_{g_1}}{d\mathcal{L}^n}(x)\varphi^{**}(x,u(x))\,dx,$$

which further implies

$$sc^-(w-L^p)I(u)=\int_\Omega \varphi^{**}(g_1(x),u(g_1(x)))\,dx=\int_\Omega \psi^{**}(x,u(g_1(x)))\,dx,$$

concluding the proof. □

Let us observe that in general $sc^-(w-L^p)I(u)\neq\int_\Omega f^{**}(x,u(g_1(x)))\,dx$. In fact, consider the following

Example 4.2. *Let* $\Omega=(0,1)$ *and*

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

while

$$g_1(x):=\begin{cases} 2x, & 0<x\le 1/2,\\ 2-2x, & 1/2<x<1.\end{cases}$$

Then $g_1^{-1}(\Sigma)$ *consists of all Lebesgue measurable subsets of* $(0,1)$ *symmetric with respect to* $1/2$*, and so the conditional expectation of* f *in the first variable is given by*

$$\psi(x,y):=\frac{1}{2}\big((1-\exp(-y^2)-(\exp(y^2)-1)\big)=\sinh y^2.$$

Note that it is already convex. Hence by the theorem 4.1 one has for the relaxation of the functional

$$I(u):=\int_\Omega f(x,u(g_1(x)))\,dx$$

the following representation:

$$sc^-(w-L^p)I(u)=\int_\Omega \sinh u(g_1(x))^2\,dx.$$

But since g_1 *is measure preserving, we find out*

$$sc^-(w-L^p)I(u)=\int_\Omega \sinh u(x)^2\,dx.$$

On the other hand, it is a matter of simple computation to show

$$I(u)=\int_0^1 \sinh u(x)^2\,dx,$$

and hence

$$sc^-(w-L^p)I(u)=I(u)=\int_\Omega \sinh u(g_1(x))^2\,dx,$$

while

$$\int_\Omega f^{**}(x,u(g_1(x)))\,dx=\int_\Omega(\exp u(g_1(x))^2-1)\,dx.$$

Now we pass to a difficult general case of functionals involving many deviations.

Theorem 4.3. *If the set of functions $g_1, \dots, g_k\colon \Omega \to \Omega$ is ω-unifiable, then there is a sequence of measurable subsets $\{\Omega_\nu\}_{\nu=1}^\infty$ of Ω of positive measure, a new set of measurable functions $\{\hat{g}_\nu\}_{\nu=1}^\infty\colon \Omega_\nu \to \Omega$ satisfying (2) and a set of a.e. positive and finite measurable weight functions $\{\omega_\nu\}_{\nu=1}^\infty\colon \Omega_\nu \to (0, +\infty)$ defining the spaces*

$$\Delta(x) := \left\{ \hat{y} \in \mathbb{R}^{\{\nu \,:\, x\in\Omega_\nu\}} : |\hat{y}|_{x,p}^p := \sum_{\nu \,:\, x\in\Omega_\nu} |\omega_\nu(x)\hat{y}_\nu|^p < +\infty \right\}$$

such that

$$sc^-(w - L^p)I(u) = \int_\Omega \psi^{**}(x, u(\hat{g}_1(x)), \dots, u(\hat{g}_\nu(x)), \dots)\, dx,$$

for some function $\psi\colon \{(x,\hat{y}) : x \in \Omega, \hat{y} \in \Delta(x)\} \to \bar{\mathbb{R}}$. Here

$$\begin{array}{rcl} \psi^*(x, \hat{y}') & := & \sup_{\hat{y}\in\Delta(x)} (< \hat{y}, \hat{y}' >_x - \phi(x,\hat{y})) \ \textit{ for } \hat{y}' \in \Delta'(x), \\ \psi^{**}(x,\hat{y}) & := & (\psi^*)^*(x,\hat{y}), \end{array}$$

while

$$\Delta'(x) := \left\{ \hat{y} \in \mathbb{R}^{\{\nu \,:\, x\in\Omega_\nu\}} : \sum_{\nu \,:\, x\in\Omega_\nu} \left| \frac{1}{\omega_\nu(x)} \hat{y}_\nu \right|^p < +\infty \right\}$$

and $< \cdot, \cdot >_x$ stands for the duality between $\Delta(x)$ and $\Delta'(x)$ defined by

$$< \hat{y}, \hat{y}' >_x := \sum_{\nu \,:\, x\in\Omega_\nu} \hat{y}_\nu \hat{y}'_\nu.$$

Again, it is important to note that both the new "integrand" ψ and the new argument deviations $\{\hat{g}_\nu\}$ can be obtained in a quite constructive way (although not that simple as in the case of just one deviation). For the reader who is not interested in following horribly technical constructions of the proof we enlist several useful additional features of the construction in the remark below.

Remark. In the conditions of the above theorem

1. If there is a unifier of the functions $g_1, \dots, g_k$ with a *finite* number (say, $l \in \mathbb{N}$) of injectivity sets, then in fact in the sequence of new deviations only l are different.
2. The images of the functions $\hat{g}_\nu$ are pairwise disjoint (up to the set of zero measure).
3. Let $T(x) := \{\nu \in \mathbb{N} : x \in \Omega_\nu\}$ and $\omega(x) := \{\omega_\nu(x)\}_{\nu\in T(x)}$ is the measure over the at most countable set $T(x)$. In this notation $\Delta(x)$ is the Banach space $l^p(T(x); \omega^{1/p}(x))$. Note that $T(x) = \emptyset$ implies $\Delta(x) = \{0\}$. The assertion of the above theorem can be written out as

$$sc^-(w - L^p)I(u) = \int_\Omega \psi^{**}(x, \{u(\hat{g}_\nu(x))\}_{\nu \,:\, x\in T(x)})\, dx,$$

4. Specific measurable properties of the new "integrand" ψ can be stated once one introduces some kind of measure structure in the "fibration" Δ (in fact, we put the word "integrand" in quotes here because strictly speaking the notion of an integrand

on Δ is not defined). We avoid here introducing any such precise notion, since for our purpose it is enough to mention that the expression under the integration sign in the representation formula for the relaxed functional is a measurable function for any measurable u: $\Omega \to \mathbb{R}$ (which can also be taken as the definition of a notion of "integrand" in this case).

The rest of the section will be dedicated to the proof of the theorem 4.3.

Proof of the Theorem 4.3:

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

$$T(u)(x) := f(x, u(g_1(x)), \ldots, u(g_k(x))).$$

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

Lemma 4.4. *There is a measurable set $\Theta \subset \Omega$, $|\Theta| > 0$, a measurable function g: $\Omega \to \Theta$, a sequence of measurable subsets $\{\Omega_\nu\}_{\nu=1}^\infty$ of Ω of positive measure, $\Omega_\nu \in g^{-1}(\Sigma)$, and a set of measurable functions $\{\gamma_\nu\}$: $g(\Omega_\nu) \to \Omega$, satisfying (2), a set of a.e. positive and finite measurable weight functions $\{w_\nu\}_{\nu=1}^\infty$: $g(\Omega_\nu) \to (0, +\infty)$ defining the spaces*

$$D(x) := \left\{ \hat{y} \in \mathbb{R}^{\{\nu : x \in g(\Omega_\nu)\}} : \sum_{x \in g(\Omega_\nu)} |w_\nu(x) \hat{y}_\nu|^p < +\infty \right\},$$

a function $\hat{f}$: $\Omega \times l^p \to \mathbb{R}^{\cdot}$, two linear continuous operators

$$\begin{array}{rrcl} P: & L^p(\Theta; l^p) & \to & L^p(\Omega), \\ \hat{P}: & L^p(\Omega) & \to & L^p(\Theta; l^p), \end{array}$$

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

$$\begin{array}{rrcl} H(x): & l^p & \to & K(x), \\ \hat{H}(x): & K(x) & \to & l^p, \end{array}$$

where $K(x) = \{y \in \mathbb{R}^k : g_i(x) = g_{i'}(x) \Rightarrow y_i = y_{i'}\}$, and at last two operator functions of $x \in \Theta$, the values of which are linear continuous operators

$$\begin{array}{rrcl} A(x): & l^p & \to & D(x), \\ \hat{A}(x): & D(x) & \to & l^p, \end{array}$$

satisfying 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* $u \in L^p$, $\hat{u} \in L^p(\Theta; l^p)$ *and for a.e.* $x \in \Omega$;

(ii) *one has* $\hat{f}(x, \hat{y}) = f(x, H(x)\hat{y})$, *while* $\hat{f}$ *is a normal integrand whenever so is* f;

(iii) $P\hat{P} = Id_{L^p}$ *and* $A(x)\hat{A}(x) = Id_{D(x)}$ *for a.e.* $x \in \Omega$*, where* Id_X *stands for identity in* X*, while*

$$(\hat{P}u)(x) = \hat{A}(x)\{u(\gamma_\nu(x))\}_{x\in g(\Omega_\nu)} \tag{8}$$

for all $u \in L^p$ *and a.e.* $x \in \Omega$*, while* $\hat{A}(x)$ *is an isometry;*

(iv) $H(x)\hat{A}(g(x))A(g(x)) = H(x)$ *for 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

$$H(x)\hat{u}(g(x)) = ((P\hat{u})(g_1(x)), \ldots, (P\hat{u})(g_k(x))) \text{ a.e. in } \Omega \tag{9}$$

for all $\hat{u} \in L^p(\Theta; l^p)$.

By *(i)* of the lemma 4.4 one has

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

Step 2. Consider a new functional $\hat{I}$: $L^p(\Theta; l^p) \to \bar{\mathbb{R}}$ defined by the formula

$$\hat{I}(\hat{u}) = \int_\Omega \hat{f}(x, \hat{u}(g(x)))\,dx. \tag{10}$$

We will prove that there exists a new integrand $\hat{\psi}$: $\Omega \times l^p \to \bar{\mathbb{R}}$ such that

$$sc^-(w - L^p(\Theta; l^p))\hat{I}(\hat{u}) = \int_\Omega \hat{\psi}^{**}(x, \hat{u}(g(x)))\,dx, \tag{11}$$

where

$$\begin{aligned} \hat{\psi}^*(x, \hat{y}') &:= \sup_{\hat{y}\in l^p}\Big(< \hat{y}, \hat{y}' > - \hat{\psi}(x, \hat{y})\Big) \text{ for } \hat{y}' \in l^{p'}, \\ \hat{\psi}^{**}(x, \hat{y}) &:= (\hat{\psi}^*)^*(x, \hat{y}), \end{aligned}$$

$1/p + 1/p' = 1$ and $< \cdot, \cdot >$ stands for the duality between l^p and $l^{p'}$. 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

$$\hat{I}(u) = \int_\Omega \hat{\psi}(x, \hat{u}(g(x)))\,dx = \int_\Theta \frac{d\mu_g}{d\mathcal{L}^n}(s)\hat{\varphi}(s, \hat{u}(s))\,ds,$$

where $\hat{\psi}$: $\Omega \times l^p \to \bar{\mathbb{R}}$ is defined by

$$\hat{\psi}(x, y) := E(\hat{f}(\cdot, y); g^{-1}(\Sigma \cap \Theta))(x)$$

and $\hat{\varphi}$: $\Theta \times l^p \to \bar{\mathbb{R}}$ satisfies $\hat{\varphi}(g(x), y) = \hat{\psi}(x, y)$. We substitute now the proposition 2.3 of [12, chapter IX] with the theorem VII.7 of [5] in order to assert that

$$sc^-(w - L^p(\Theta; l^p))\hat{I}(\hat{u}) = \int_\Theta \frac{d\mu_g}{d\mathcal{L}^n}(s)\hat{\varphi}^{**}(s, \hat{u})(s)\,ds,$$

and hence

$$sc^-(w - L^p(\Theta; l^p))\hat{I}(\hat{u}) = \int_\Omega \hat{\psi}^{**}(x, \hat{u}(g(x)))\, dx,$$

which shows the claim.

Step 3. We use the lemma below with $X = L^p(\Omega)$, $\hat{X} = L^p(\Theta; l^p)$, the functional $\hat{I}$ given by (10), and I the original functional to prove that

$$sc^-(w - L^p)I(u) = sc^-(w - L^p(\Theta; l^p))\hat{I}(\hat{P}u) = \int_\Omega \hat{\psi}^{**}(x, (\hat{P}u)(g(x)))\, dx, \tag{12}$$

the latter equality being valid in view of (11).

Lemma 4.5. *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,

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

Proof. To deduce $(**)$ from $(*)$ it is enough to substitute $\hat{u} = \hat{P}u$ into $(*)$. Hence, we are to show $(*)$, i.e.

(i) If $\hat{u}_\nu \rightharpoonup \hat{u}$ weakly in $\hat{X}$, then $sc^-(w - X)I(P\hat{u}) \le \liminf_\nu \hat{I}(\hat{u}_\nu)$.

(ii) For every $\hat{u} \in \hat{X}$ there is a sequence $\{\hat{u}_\nu\} \subset \hat{X}$, such that $\hat{u}_\nu \rightharpoonup \hat{u}$ weakly in $\hat{X}$, and $sc^-(w - X)I(P\hat{u}) = \lim_\nu \hat{I}(\hat{u}_\nu)$.

To prove (i), assume $\hat{u}_\nu \rightharpoonup \hat{u}$ weakly in $\hat{X}$ and note that since $P\hat{u}_\nu \rightharpoonup P\hat{u}$ weakly in X, then

$$sc^-(w - X)I(P\hat{u}) \le \liminf_\nu I(P\hat{u}_\nu) = \liminf_\nu \hat{I}(\hat{u}_\nu).$$

To prove (ii) note that for each $\hat{u} \in \hat{X}$ there is a sequence $\{u_\nu\} \subset X$ satisfying $u_\nu \rightharpoonup P\hat{u}$ weakly in X and

$$sc^-(w - X)I(P\hat{u}) = \lim_\nu I(u_\nu).$$

Setting $\hat{u}_\nu := \hat{P}u_\nu$, we obtain $u_\nu = P\hat{u}_\nu$ and, therefore, (ii) holds. □

Step 4. Set $\omega_i^\nu(x) = w_i^\nu(g(x))$ and observe that $\Delta(x) = D(g(x))$, since all $\Omega_\nu \in g^{-1}(\Sigma)$ according to the lemma 4.4. Define the function $\bar{f}\colon \{(x, \hat{y}) : \hat{y} \in \Delta(x)\} \to \bar{\mathbb{R}}$ by the relationship

$$\bar{f}(x, \hat{y}) := \hat{f}(x, \hat{A}(g(x))\hat{y}).$$

Note that

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

In fact, this claim follows easily from

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

The latter is valid due to *(iv)* of the lemma 4.4 and the definition of $\bar{f}$.

Now define the function ψ by

$$\psi(x,\hat{y}) = E(\bar{f}(\cdot,\hat{y}); g^{-1}(\Sigma))(x).$$

Clearly, since $A(g(\cdot))\hat{y}$ is g-measurable, one has

$$\hat{\psi}(x,\hat{y}) = \psi(x, A(g(x))\hat{y}).$$

Applying the lemma below with $\hat{X} = l^p$, $X = \Delta(x)$, $\hat{F} = \hat{\psi}(x,\cdot)$, $F = \psi(x,\cdot)$, $\hat{H} = \hat{A}(g(x))$, $H = A(g(x))$, we obtain

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

and in particular

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

which concludes the proof.

Lemma 4.6. *Let X and $\hat{X}$ be normed spaces, $H\colon \hat{X} \to X$ and $\hat{H}\colon X \to \hat{X}$ be linear bounded operators satisfying $H\hat{H} = Id_X$. If $F\colon X \to \mathbb{R}^{\cdot}$ and $\hat{F}\colon \hat{X} \to \mathbb{R}^{\cdot}$ are such nonnegative functionals that $\hat{F}(\hat{u}) = F(H\hat{u})$ for every $\hat{u} \in \hat{X}$, then*

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

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

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

Remark. Clearly, for the above lemma to be valid, it is enough to require instead of nonnegativity of F that there exist at least one continuous affine functional less than F.

Proof. The condition $\hat{F}(\hat{u}) = F(H\hat{u})$ is equivalent to the fact that

$$\operatorname{epi}\hat{F} = (H \times Id_{\mathbb{R}})^{-1}(\operatorname{epi} F),$$

where epi stands for the epigraph of the functional, the operator $P \times Id_{\mathbb{R}}\colon \hat{X} \times \mathbb{R} \to X \times \mathbb{R}$ is given by $(H \times Id_{\mathbb{R}})(\hat{u},\lambda) := (P\hat{u},\lambda)$. Since the functionals $\hat{F}^{**}$ and F^{**} are determined by the conditions

$$\operatorname{epi}\hat{F}^{**} = \overline{\operatorname{co}}\operatorname{epi}\hat{F}, \qquad \operatorname{epi} F^{**} = \overline{\operatorname{co}}\operatorname{epi} F,$$

where $\overline{\operatorname{co}}$ stands for the closed convex hull, then to show the statement it is enough to prove

$$\overline{\operatorname{co}}\operatorname{epi}\hat{F} = (H \times Id_{\mathbb{R}})^{-1}(\overline{\operatorname{co}}\operatorname{epi} F).$$

The latter is true in view of $(H \times Id_{\mathbb{R}})^{-1}\overline{\operatorname{co}}E = \overline{\operatorname{co}}(H \times Id_{\mathbb{R}})^{-1}E$ for every $E \subset \hat{X} \times \mathbb{R}$, which holds since $H \times Id_{\mathbb{R}}$ is surjective and posseses the bounded right inverse $\hat{H} \times Id_{\mathbb{R}}$. $\square$

We now concentrate on the proof of the lemma 4.4 which constitutes the heart of the proved results.

Proof of the Lemma 4.4:

Step 1. Preparatory constructions. According to the lemma A.2 we may consider that the unifier $\gamma\colon \Omega \to \Omega$ satisfies the condition $\gamma(x) \in O(x)$ for a.e. $x \in \Omega$ (the orbits $O(x)$ are defined in Appendix A). Let $g(x) := \gamma(g_1(x))$ and observe that this function satisfies (2).

We set now $\Theta := \gamma(\Omega)$ and remark that Θ is measurable with $|\Theta| > 0$. We have $\operatorname{card} O(x) \cap \Theta = 1$ for every $x \in \Omega$. In fact, $\gamma(x) \in O(x) \cap \Theta$ by construction. What is more, if for some $y \in O(x) \cap \Theta$ holds $y \neq \gamma(x)$, then there is an $x' \in \Omega$ such that $y = \gamma(x')$, due to the definition of Θ. Now $x' \not\sim x$ (otherwise by construction of γ one would have $y = \gamma(x)$), which leads to a contradiction $y = \gamma(x') \notin O(x)$ since $\gamma(x') \sim x'$ by construction of γ.

Since γ satisfies the ω-condition, we may consider its measurable injectivity sets $\Theta_\nu \subset \Omega$, $\nu \in \mathbb{N}$ with $\Omega = \sqcup_\nu \Theta_\nu$. Without loss of generality we will assume that the sequence $\{\Theta_\nu\}$ is countable (if γ is injective only on a finite number of pieces, just fill the rest of the sequence with empty sets).

Note that the orbit $O(x)$ of every $x \in \Omega$ intersects with Θ_ν by at most one point, i.e. $\operatorname{card} O(x) \cap \Theta_\nu \leq 1$, where card stands for cardinality of a set. In fact, supposing the existence of $y_1 \neq y_2$, in $O(x) \cap \Theta_\nu$, we would have $\gamma(y_1) \neq \gamma(y_2)$ since γ is injective over each Θ_ν, which contradicts the fact that γ is constant over each equivalence class by construction. Moreover, if $x \in O(\Theta_\nu)$, then there is an $x' \sim x$, $x' \in \Theta_\nu$, and hence $O(x) \cap \Theta_\nu \neq \emptyset$ whenever $\Theta_\nu \neq \emptyset$. Let $\Omega_\nu := g^{-1}(O(\Theta_\nu) \cap \Theta)$. Now it is possible to define the functions $\gamma_\nu\colon g(\Omega_\nu) \to \Theta_\nu$ by the relationship

$$\gamma_\nu(x) := O(x) \cap \Theta_\nu.$$

Denote by μ_{γ_ν} the measure over Θ_ν defined by $\mu_{\gamma_\nu}(e) := |\gamma_\nu^{-1}(e)|$ and let

$$a_\nu(x) := \left(\frac{d\mu_{\gamma_\nu}}{d\mathcal{L}^n}(x)\right)^{1/p}, \qquad x \in \Theta_\nu.$$

This definition is correct since the Radon-Nikodym derivative $\frac{d\mu_{\gamma_\nu}}{d\mathcal{L}^n}(x) > 0$ for a.e. $x \in \Theta_\nu$. In fact, otherwise there would exist a measurable $e \subset \Theta_\nu$ with $|e| > 0$ such that $|\gamma_\nu^{-1}(e)| = 0$, contradicting with

$$e \subset \gamma_\nu(\gamma_\nu^{-1}(e)) \subset O(\gamma_\nu^{-1}(e))$$

which implies $|e| = 0$ due to the lemma A.2.

At last define the function $\bar{\nu}\colon \Omega \to \mathbb{N}$ by setting $\bar{\nu}(x) := \nu$ whenever $x \in \Theta_\nu$. Note that $x \in \Theta_{\bar{\nu}(x)}$ for every $x \in \Omega$.

In the sequel we will need the relationship

$$\gamma_{\bar{\nu}(x)}(\gamma(x)) = x \text{ a.e. in } \Omega, \tag{13}$$

as well as

$$\gamma(\gamma_\nu(x)) = x \text{ a.e. in } \Theta \cap O(\Theta_\nu), \tag{14}$$

In fact, (13) is proved by

$$\gamma_{\bar{\nu}(x)}(\gamma(x)) = O(\gamma(x)) \cap \Theta_{\bar{\nu}(x)} = O(x) \cap \Theta_{\bar{\nu}(x)} = x,$$

because $x \in \Theta_{\bar{\nu}(x)}$ and $x \in O(x)$ simultaneously. To prove the relationship (14) note that $\gamma(\gamma_\nu(x)) \in O(x)$ by construction of γ and γ_ν, whereas $\gamma(\gamma_{\bar{\nu}(x)}(x)) \in \Theta$. Observing $\Theta \cap O(x) = x$ one concludes the proof of the claim.

Step 2. We introduce now formally the operators $H(x)$, $\hat{H}(x)$ and $A(x)$, $\hat{A}(x)$ 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 $H(x)$: $l^p \to K(x)$ by the relationship

$$H(x)\hat{y} := \big(a_{\bar{\nu}(g_1(x))}(g_1(x))\hat{y}_{\bar{\nu}(g_1(x))}, \dots, a_{\bar{\nu}(g_k(x))}(g_k(x))\hat{y}_{\bar{\nu}(g_k(x))}\big) .$$

It is easy to check that $H(x)$ maps l^p into $K(x)$ and is continuous.

The operator $\hat{H}(x)$: $K(x) \to l^p$ is introduced by

$$(\hat{H}(x)y)_\nu := \begin{cases} y_i/a_\nu(g_i(x)), & \text{if there is an } i \in \{1, \dots, k\} \,:\, \bar{\nu}(g_i(x)) = \nu, \\ 0, & \text{otherwise.} \end{cases}$$

This definition is correct since if for some $\nu \in \mathbb{N}$ and $x \in \Omega$ one has $g_i(x) \in \Theta_\nu$ and $g_j(x) \in \Theta_\nu$, then $y_i = y_j$. In fact, $g_i(x) \sim g_j(x)$ and since Θ_ν contains at most one point of each orbit, $g_i(x) = g_j(x)$, and hence $y_i = y_j$ by definition of $K(x)$. Furthermore, it is clear that $\hat{H}(x)$ maps $K(x)$ into l^p and is continuous between these spaces, because the image of every $y \in K(x)$ has only finite number of nonzero components by definition of $K(x)$.

Introduce now the operators $A(x)$ and $\hat{A}(x)$ for a.e. $x \in \Theta$. Set

$$w_\nu(x) := a_\nu(\gamma_\nu(x))$$

and define $A(x)$ by

$$A(x)_\nu \hat{y} := \hat{y}_\nu w_\nu(x),$$

and the operator $\hat{A}$ by the formula

$$\hat{A}_\nu(x)\hat{y} = \begin{cases} \hat{y}_\nu / w_\nu(x), & x \in g(\Omega_\nu), \\ 0, & \text{elsewhere.} \end{cases}$$

Here it was assumed $0 \cdot \infty = 0$ and $0/0 = 0$. It is matter of a simple excercise to verify the acting and continuity of the operators $A(x)$ and $\hat{A}(x)$, as well as the fact that $\hat{A}(x)$ is an isometry. Also the identity $A(x)\hat{A}(x) = \mathrm{Id}_{D(x)}$ is immediate.

Step 3. We dedicate a separate step to a more delicate construction of the operators P and $\hat{P}$. The operator P: $L^p(\Theta; l^p) \to L^p(\Omega)$ is defined by

$$(P\hat{u})(x) := A_{\bar{\nu}(x)}(\gamma(x))\{\hat{u}_\nu(\gamma(x))\}_{\nu=1}^{\infty} = a_{\bar{\nu}(x)}\hat{u}_{\bar{\nu}(x)}(\gamma(x)).$$

This definition is correct since for every $x \in \Omega$ one has $\gamma(x) \in \Theta$ and $x \in \Theta_{\bar{\nu}(x)}$ by construction of Θ and $\bar{\nu}$. The operator $\hat{P}$: $L^p(\Omega) \to L^p(\Theta; l^p)$ is defined by

$$(\hat{P}u)(x) = \hat{A}(x)\{u(\gamma_\nu(x))\}_{x \in g(\Omega_\nu)}.$$

We are to verify that P and $\hat{P}$ act continuously between the indicated spaces. First we concentrate on $\hat{P}$. We calculate for each $m \in \mathbb{N}$ the following integrals of finite sums

$$\begin{aligned} \int_\Theta \sum_{\nu=1}^m \left|(\hat{P}u)_\nu(x)\right|^p dx &= \sum_{\nu=1}^m \int_{\Theta\cap O(\Theta_\nu)} \left|\frac{u(\gamma_\nu(x))}{a_\nu(\gamma_\nu(x))}\right|^p dx = \\ &\sum_{\nu=1}^m \int_{\Theta_\nu} |u(x)|^p\, dx \leq ||u||_p^p, \end{aligned} \tag{15}$$

where in the last equality the change of variables has been used together with the observation

$$\gamma_\nu(\Theta \cap O(\Theta_\nu)) = \Theta_\nu. \tag{16}$$

To show (16) one has to note that γ_ν acts into Θ_ν by definition, while for every $x \in \Theta_\nu$ there is an $x' := O(x) \cap \Theta$ (this definition is correct since by construction Θ contains one and only one element of each orbit) such that $\gamma_\nu(x') = x$, because $O(\Theta) = \Omega$.

The estimate (15) implies by Beppo Levi theorem that $\hat{P}$ acts between $L^p(\Omega)$ and $L^p(\Theta; l^p)$ (the measurability of $\hat{P}u$ is a simple excercise), while

$$||\hat{P}u||_{L^p(\Theta;l^p)} = ||u||_p,$$

showing the desired continuity.

Note that by definition, the operator $\hat{P}$ satisfies (8) with

$$(A(x)\hat{y})_\nu := \begin{cases} \hat{y}_\nu / a_\nu(\gamma_\nu(x)), & x \in O(\Theta_\nu), \\ 0, & x \notin O(\Theta_\nu). \end{cases}$$

To verify that for a.e. $x \in \Omega$ one has $A(x)$: $l^p \to l^p$, we reiterate (15) with $u(x) := \sum_\nu \hat{y}_\nu 1_{\Theta_\nu}(x)$ for arbitrary $\hat{y} \in l^p$, hence showing $A(x)\hat{y} \in l^p$ for a.e. $x \in \Omega$. The same argument shows that $A(x)$ is bounded.

Now we verify acting and continuity conditions of P. For this purpose compute

$$||P\hat{u}||_p^p = \int_\Omega a_{\bar{\nu}(x)}(x)^p \left|\hat{u}_{\bar{\nu}(x)}(\gamma(x))\right|^p dx = \int_\Omega \frac{d\mu_{\gamma_{\bar{\nu}(x)}}}{d\mathcal{L}^n}(x) \left|\hat{u}_{\bar{\nu}(x)}(\gamma(x))\right|^p dx$$

Representing then the integral over Ω as a sum of integrals over Θ_ν and changing variables, we have

$$||P\hat{u}||_p^p = \sum_{\nu=1}^\infty \int_{O(\Theta_\nu)\cap\Theta} |\hat{u}_\nu(\gamma(\gamma_\nu(x)))|^p\, dx = \sum_{\nu=1}^\infty \int_{O(\Theta_\nu)\cap\Theta} |\hat{u}_\nu(x)|^p\, dx$$

due to (14). At last, the upper estimate of the latter integral shows

$$||P\hat{u}||_p^p \leq \sum_{\nu=1}^\infty \int_\Theta |\hat{u}_\nu(x)|^p\, dx = ||\hat{u}||_{L^p(\Theta;l^p)}^p,$$

concluding the proof of the claim.

Step 4. We conclude the proof of the claim *(iii)* of the lemma. by showing $P\hat{P} = Id_{L^p}$. Given a $u \in L^p$, we have

$$(P\hat{P}u)(x) = a_{\bar{\nu}(x)}(x)(\hat{P}u)_{\bar{\nu}(x)}(\gamma(x)). \tag{17}$$

Since $\gamma(x) \in O(x) \subset O(\Theta_{\bar{\nu}(x)})$, then using the definition of $\hat{P}$ one deduces from (17) and (13) that

$$(P\hat{P}u)(x) = \frac{a_{\bar{\nu}(x)}(x)}{a_{\bar{\nu}(x)}(\gamma_{\bar{\nu}(x)}(\gamma(x)))} u(\gamma_{\bar{\nu}(x)}(\gamma(x))) = u(x),$$

which shows the claim.

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

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

Clearly, if $f(x, \cdot)$ is lower semicontinuous in $\mathbb{R}^k$ for a.e. $x \in \Omega$, then so is $\hat{f}(x, \cdot)$ in l^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

$$(x, \hat{y}) \in \Omega \times l^p \mapsto (x, H(x)\hat{y}) \in \Omega \times \mathbb{R}^k$$

is $(\Sigma \otimes \mathcal{B}(l^p), \Sigma \otimes \mathcal{B}(\mathbb{R}^k))$-measurable. To show this claim we first note that the function $(x, \hat{y}) \in \Omega \times l^p \mapsto x \in \Omega$ is clearly $(\Sigma \otimes \mathcal{B}(l^p), \Sigma)$-measurable. It suffices now to show that for each fixed $i \in \{1, \ldots, k\}$ the function $\psi_i\colon \Omega \times l^p \to \mathbb{R}$ defined by

$$\psi_i(x, \hat{y}) := a_{\bar{\nu}(g_i(\cdot))}(g_i(\cdot))\hat{y}_{\bar{\nu}(g_i(\cdot))},$$

is $(\Sigma \otimes \mathcal{B}(l^p), \mathcal{B})$-measurable. Fixing an open $U \in \mathbb{R}$, we compute

$$\begin{aligned} \psi_i^{-1}(U) &= \bigcup_{\nu=1}^{\infty} \{(x, \hat{y}) \in \Omega \times l^p \, : \, \bar{\nu}(g_i(x)) = \nu \text{ and } a_\nu(g_i(x))\hat{y}_\nu \in U\} = \\ & \bigcup_{\nu=1}^{\infty} \left(g_i^{-1}(\Theta_\nu) \times l^p\right) \cap \{(x, \hat{y}) \in \Omega \times l^p \, : \, a_\nu(g_i(x))\hat{y}_\nu \in U\}, \end{aligned}$$

the latter clearly belonging to $\Sigma \otimes \mathcal{B}(l^p)$, which shows the claim.

Step 6. We prove *(iv)*. Note that

$$g(x) = \gamma(g_i(x)) \in \gamma(\Theta_{\bar{\nu}(g_i(x))}) = g(\Omega_{\bar{\nu}(g_i(x))}),$$

hence by definition of $A(x)$ and $\hat{A}(x)$ one has

$$\left(\hat{A}(g(x))\bar{A}(g(x))\hat{y}\right)_{\bar{\nu}(g_i(x))} = \hat{y}_{\bar{\nu}(g_i(x))}.$$

This implies

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

since the construction of $H(x)$ involves only the coordinates $\hat{y}_{\bar{\nu}(g_i(x))}$.

Step 7. At last, we prove (9), and hence, also the most crucial claim *(i)*. In fact, from (9) and (*ii*) immediately follows

$$T(P\hat{u})(x) = \hat{f}(x, \hat{u}(g(x))), \tag{18}$$

for every $\hat{u} \in L^p(\Theta; l^p)$. Then

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

is the immediate consequence of the latter due to *(iii)*. At last, (9) is obtained by a straightforward computation.

A. Some remarks on ω-unifiability

Let us remark first that there are unifiable sets of functions which are not ω-unifiable. In fact, consider the following

Example A.1. *Let $\Omega = (0,1)$ 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 [14]). 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, g_1 and g_2 are 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 (2). 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.

Introduce now some auxiliary notions to be used in the paper. For any $x \in \Omega$ define the *orbit* $O(x) \subset \Omega$ by

$$O(x) := \bigcup_{l \in \mathbb{N}} O_l(x),$$

where

$$\begin{array}{rcl} O_0(x) & := & \{x\}, \\ O_1(x) & := & \{y \in \Omega \,:\, g_i^{-1}(y) \cap g_j^{-1}(x) \neq \emptyset \text{ for some } i \neq j\}, \\ O_l(x) & := & O_1(O_{l-1}(x)). \end{array}$$

Observe that O defines an equivalence relation $x \sim y \Leftrightarrow x \in O(y)$.

Using the above notion of orbits it is rather easy to formulate simple sufficient conditions of ω-unifiability, in terms of the orbit structure, by means of selecting a unifier from the orbits. Although we do not use such unifiablity criteria in the paper, we we will need the following simple technical lemma show that in the case of ω-unifiability one can always find a unifier as a selector from the orbits.

Lemma A.2. *Let the set of functions $g_1, \ldots, g_k$: $\Omega \to \Omega$ be ω-unifiable by a unifier δ: $\Omega \to \Omega$. Then there is an ω-unifier γ: $\Omega \to \Omega$ of this set, satisfying $\gamma(x) \in O'(x)$ for a.e. $x \in \Omega$, where $O'(x) := \delta^{-1}(\delta(x))$.*

Proof. Let $\{\Omega_j\}_{j=1}^m$, where either $m \in \mathbb{N}$ or $m = +\infty$, be the sequence of the injectivity sets of a unifier δ. Let

$$\Theta := \Omega_1 \cup \left(\bigcup_{j=2}^m \Omega_j \setminus O' \left(\bigcup_{l=1}^{j-1} \Omega_l \right) \right).$$

Since for a.e. $x \in \Omega$ the set Θ obviously contains one and only one point of $O'(x)$, we can define a function γ: $\Omega \to \Omega$ by the formula

$$\gamma(x) := \Theta \cap O'(x).$$

It is measurable since $\gamma^{-1}(B) = O'(\Theta \cap B) = \delta^{-1}(\delta(\Theta \cap B))$ is measurable for every Borel set $B \subset \Omega$. At last, γ is also an ω-unifier since $O(x) \subset O'(x)$ for a.e. $x \in \Omega$. $\square$

We remark that whenever the set of functions $g_1, \ldots, g_k$ is ω-unifiable, then the respective orbits satisfy $|O(e)| = 0$ for all sets $e \subset \Omega$ with $|e| = 0$. Indeed, this is immediately implied by $O(x) \subset O'(x) = \delta^{-1}(\delta(x))$ (see the notations of the lemma A.2). Note however, that this is not sufficient for ω-unifiability, even when the set of functions $g_1, \ldots, g_k$ is unifiable. In fact, the lemma below implies that the same holds whenever each of the functions $g_1, \ldots, g_k$ satisfies ω-condition. At the same time, the example A.1 provides a unifiable but not ω-unifiable pair of functions, each of which is even injective.

Lemma A.3. *Let $g_1, \ldots, g_k$ satisfy ω-condition. Then the respective set-valued map O: $\Omega \multimap \Omega$ maps measurable sets into measurable ones and satisfies $|O(e)| = 0$ for every $e \subset \Omega$ with $|e| = 0$.*

Proof. It is enough to show that O_1 maps measurable sets into measurable ones and if $|e| = 0$, then $|O_1(e)| = 0$. The latter holds, because

$$O_1(e) = \bigcup_{i \neq j \in \{1, \ldots, k\}} g_i(g_j^{-1}(e)), \tag{19}$$

while $|e| = 0$ implies both $|g_i^{-1}(e)| = 0$ in virtue of (2) and $|g_i(e)| = 0$ in virtue of ω-condition. □

References

[1] N. V. Azbelev, V. P. Maksimov, L. F. Rakhmatullina: Introduction to the Theory of Functional Differential Equations, Nauka, Moscow (1991) (Russian); English transl. of the 1st part: Introduction to the Theory of Linear Functional Differential Equations, World Federation Publishers (1996).

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

[3] G. Buttazzo: Semicontinuity, relaxation and integral representation in the calculus of variations, Pitman Research Notes in Mathematics 207, Longman Scientific, Harlow (1989).

[4] G. Buttazzo, M. E. Drakhlin, L. Freddi, E. Stepanov: Homogenization of optimal control problems for functional differential equations with deviating argument, J. Optimization Theory and Appl. 93 (1997) 103–119.

[5] C. Castaing, M. Valadier: Convex Analysis and Measurable Multifunctions, Lecture Notes Math. 580, Springer-Verlag, Berlin (1977).

[6] A. V. Chistyakov: A pathological counterexample for hypothesis on non-Fredholm property in algebras of weighted shift operators, Izvestiya VUZ, Matematika 39 (1995) 76–86 (Russian); English transl.: Russian Mathematics 39, no. 10 (1995) 73–83.

[7] D. L. Cohn: Measure Theory, Birkhäuser, Boston (1980).

[8] J. Diestel, L. Uhl: Vector Measures, American Math. Soc., Providence, R. I. (1968).

[9] M. E. Drakhlin: On one linear functional equation, in: Functional Differential Equations, Perm Polytechnical Institute, Perm (1985) 91–111 (Russian).

[10] M. E. Drakhlin: An inner superposition operator in spaces of summable functions, Izvestiya VUZ, Matematika 30 (1986) 18–24 (Russian).

[11] M. E. Drakhlin, E. Litsyn, E. Stepanov: Variational methods for a class of nonlocal functionals, Computers and Math. with Appl. 37 (1999) 79–100.

[12] I. Ekeland, R. Temam: Convex Analysis and Variational Problems, Studies in Mathematics and its Applications 1, North Holland Publ. Co., Amsterdam et al. (1976).

[13] I. I. Gikhman, A. V. Skorokhod: The Theory of Stochastic Processes, "Nauka", Moscow (1975) (Russian); English transl.: Springer-Verlag, Berlin (1980).

[14] R. P. Halmos: Measure Theory, Graduate Texts in Mathematics 18, Springer-Verlag, Berlin (1974).

[15] N. J. Kalton: Isomorphisms between L_p-function spaces when $p < 1$, J. Funct. Anal. 42 (1981) 299–337 .

[16] I. V. Shragin: Abstract Nemyckiĭ operators are locally defined operators, Soviet Acad. Sci. Dokl. Math., 227 (1976) 47–49 (Russian); English transl.: Soviet Math. Doklady 17 (1976) 354–357.