International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

Convergence of Minima of Integral Functionals, with Applications to Optimal Control and Stochastic Optimization

Lucchetti, R. & Wets, R.J.-B.

IIASA Working Paper

WP-90-081

December 1990

Working Paper

Convergence of Minima of Integral Functionals, with Applications to Optimal Control and Stochastic Optimization

Roberto Lucchetti
Roger J-B Wets

WP-90-81
December 1990

IIASA

International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria

Telephone: (0 22 36) 715 21 *0 □ Telex: 079 137 iiasa a □ Telefax: (0 22 36) 71313

Convergence of Minima of Integral Functionals, with Applications to Optimal Control and Stochastic Optimization

Roberto Lucchetti†
Roger J-B Wets‡

WP-90-81
December 1990

†visiting University of California
‡Research supported by grants of the National Science Foundation and the Air Force Office of Scientific Research

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: (0 22 36) 715 21*0 □ Telex: 079 137 iiasa a □ Telefax: (0 22 36) 71313

Foreword

Epi-convergence of integral functionals is derived under new conditions that can be used in the infinite dimensional case. Applications include: the convergence of the solutions of approximating optimal control problems and of stochastic optimization problems.

Alexander B. Kurzhanski
Chairman
System and Decision Sciences Program

CONVERGENCE OF MINIMA OF INTEGRAL FUNCTIONALS, WITH APPLICATIONS TO OPTIMAL CONTROL AND STOCHASTIC OPTIMIZATION

Roberto Lucchetti†
Dipartimento di Matematica, Università di Milano

Roger J-B Wets‡
Mathematics, University of California-Davis

Keywords: epi-convergence, statistical consistency, uniform convergence of measures, expectation functionals, integral functionals, optimal control, mid-course maneuver problems

Date: December 10, 1990

† visiting University of California

‡ Research supported by grants of the National Science Foundation and the Air Force Office of Scientific Research

To justify the use of probability distributions based on a (necessarily) limited number of samples to calculate estimates for the optimal solutions of stochastic optimization problems, or to obtain consistency results for statistical estimators that may have to be chosen under constraints, Dupačová and Wets [**3**] showed that when the probability measures derived from the samples converge narrowly (weakly), the problems epi-converge (in a probabilistic sense); in turn, this implies, in a sens that can be made precise, the convergence of optimal solutions or statistical estimators. The technique requires proving the ***epi-convergence*** of integral functionals. This is also the concern of this paper, but in a more general setting.

The results of Dupačová and Wets [**3**, theorems **3**.7 and **3**.9] are for expectation functionals defined on finite dimensional spaces, and the conditions they use to obtain epi-consistency (almost sure epi-convergence), in particular a Lipschitz-like continuity condition on the integrand (criterion function), does not suit well the infinite dimensional setting. This excludes application of their results to certain dynamic optimization problems, in particular continuous-time stochastic control problems and to nonparametric estimation in statistics. In [**5**], King and Wets obtain an epi-consistency result that is valid for reflexive Banach spaces but only for convergent sequences of empirical probability measures; their proof relies on the law of large numbers for random sets.

In this paper, we introduce a totally different technique that relies on substantially different assumptions than those in [**3**, **4**, **9**]. Even in the finite dimensional case, our results are not included in those of Dupačová and Wets [**3**].

1. Framework and general results.

We work with the following framework: let $(S, \mathcal{A}, \mu)$ be a measure space with (S, d) a Polish (complete separable metric) space, d the metric, $\mathcal{A}$ the Borel field on S and μ a bounded (nonnegative) measure; without loss of generality, we may as well assume that all measures have been appropriately scaled so that we can restrict our attention henceforth to the case when all measures are *probability measures*. We are also given a sequence of probability measures $\{\mu^\nu, \nu \in \mathbb{N}\}$, also defined on $(S, \mathcal{A})$, that converge narrowly (weakly) to μ.

The *integral functionals* Ef and $E^\nu f$ defined on the separable reflexive Banach space X, are constructed as follows:

$$Ef(x) := \int_S f(x,s)\, \mu(ds),$$

$$E^\nu f(x) := \int_S f(x,s)\, \mu^\nu(ds),$$

where the integrand f is a proper extended real-valued function defined on $X \times S$; proper means that $f > -\infty$ and the effective domain of f is nonempty:

$$\operatorname{dom} f := \{(x,s) \in X \times S \,|\, f(x,s) < \infty\} \neq \emptyset.$$

The integrand f corresponds to the essential objective function of an optimal control problem, of a stochastic optimization problem, or to the criterion function (maximum likelihood function, for example) in statistical estimation problems; in both cases f takes the value $+\infty$ when (x,s) is a pair that fails to satisfy the (implicit) constraints.

We are looking for a theorem that would tell us that

$$\operatorname*{epi-lim}_{\nu \to \infty} E^\nu f = Ef$$

from which would follow that the solutions and optimal values converge, see [**1, 3**] for further motivation. We begin with a list of our conditions.

Condition 1. *f is a normal integrand such that* $\operatorname{dom} f = D \times S$ *with D a nonempty weakly closed subset of X.*

Recall that a closed-valued set-valued mapping Γ with domain S and with values in the subsets of X, $s \mapsto \Gamma(s) : S \rightrightarrows X$ is *measurable*, if for any open set $G \subset X$, $\Gamma^{-1}(G) = \{s \in S \,|\, \Gamma(s) \cap G \neq \emptyset\} \in \mathcal{A}$. A *normal integrand* f defined on $X \times S$ is an extended real-valued function whose epigraphical set-valued mapping,

$$s \mapsto \operatorname{epi} f(\cdot, s) : S \rightrightarrows X \times \mathbb{R}$$

is a closed-valued measurable set-valued mapping; recall $\operatorname{epi} g = \{(x,\alpha) \,|\, g(x) \leq \alpha\}$; this concept is due to Rockafellar [**10**]. Because an extended real-valued function is lsc if and only if its epigraph is closed, it follows that

(i) for all $s \in S$, the function $x \mapsto f(x,s)$ is lower semicontinuous.

If in addition to (i),

(ii) $(x,s) \mapsto f(x,s)$ is $\mathcal{B} \otimes \mathcal{A}$-measurable,

then, it is easy to verify that f is a normal integrand. In fact, if $(S, \mathcal{A}, \mu)$ is complete, conditions (i) and (ii) are not just sufficient, but also necessary, for f to be a normal integrand. In a probabilistic context, normal integrands are called *random lsc (lower semicontinuous) functions.*

The condition on the domain of f, means that $s \mapsto \operatorname{dom} f(\cdot, s)$ is constant. This condition does not restrict in any way the practical applicability of the results. If $\operatorname{dom} f(\cdot, s)$

depends on s, as would be the case in general, we would have to define the integral to accommodate an integrand that takes on the value ∞. The natural convention used in such a situation (minimization) is to define $E\{f(x,\cdot)\} = \infty$ whenever $\mu[f(x,s) = \infty] > 0$. Thus, if S is the support of the measure μ and as long as $s \mapsto \operatorname{dom} f(\cdot, s)$ satisfies some "natural" assumptions [**16**]; if, for example, for all $s \in S$, $\operatorname{dom} f(\cdot, s)$ is closed, then $D = \cap_{s\in S} \operatorname{dom} f(\cdot, s)$, provided that $f(x,\cdot)$ is summable for all $x \in D$. Having $\operatorname{dom} f$ closed is consistent with our next condition, although not implied by it.

Condition 2. *For all $\alpha \in \mathbb{R}$, the (inf-)level sets of f:*

$$\operatorname{lev}_\alpha f = \{(x,s) \,|\, f(x,s) \leq \alpha\} \subset (D \times S),$$

are sequentially closed with respect to the product of the weak topology on X and the metric topology on S.

This is a *lower semicontinuity* condition on f.

Condition 3. *f is bounded on bounded subsets of $D \times S$, more precisely, $|f|$ is bounded on every product set $H \times K$, where $H \subset D$ and $K \subset S$ are bounded sets.*

Let $\mathbb{B}_X$ be the unit ball in X. Note that condition 3 implies that for all $r \in \mathbb{R}_+$ and all $s \in S$,

$$w_r(s) := \inf\{f(x,s) \,|\, x \in r\,\mathbb{B}_X\}$$

is finite.

Like conditions 1 and 3, the next condition is mostly technical in nature.

Condition 4. *There exists a family*

$$\{u_r : S \to \overline{\mathbb{R}} \,|\, u_r \text{ usc }, u_r \leq w_r,\ r \in \mathbb{R}_+\};$$

usc = upper semicontinuous. There exists a measurable function $h : S \to \overline{\mathbb{R}}$ such that for all s, $h \leq u_r$ and $\int_S h\, d\mu > -\infty$.

Note that condition 4 implies that $f(x,s) \geq h(s)$ for all $x \in X$. Also, observe that in most situations the functions u_r can be chosen to be w_r. If it is known in advance that f admits a lower bound, such as $f \geq \gamma, \gamma \in \mathbb{R}$, we simply set $u_r \equiv \gamma$ for all r in $\mathbb{R}_+$.

Note that it does not follow from this condition that h is bounded below on bounded subsets of S.

The remaining conditions are concerned with the interplay between f and the probability measures μ and μ^ν.

Condition 5. *The strict (inf-)level sets of the functions* $\{f(x,\cdot), x \in D\}$

$$\operatorname{lev}^{<}_{\alpha} f(x,\cdot) := \{s \in S \,|\, f(x,s) < \alpha\}$$

are μ-continuity sets, i.e., their boundaries are μ-null sets.

In the derivation of the main results, for a number of technical reasons, it is possible to substitute for this latter condition, the following one that could be easier to verify, but does not necessarily hold in many interesting applications.

Condition 5(alt). *The functions* $\{f(x,\cdot) : S \to \mathbb{R}, x \in X\}$ *are continuous.*

The last condition is a combination of a tightness-like condition, involving bounded rather than compact sets, and a uniform integrability condition.

Condition 6. *Let*

$$\mathcal{W} := \{f(x,\cdot), x \in D\}.$$

We assume that the probability measures

$$\mathcal{M} := \{\mu; \mu^{\nu}, \nu \in \mathbb{N}\},$$

are $\mathcal{W}$-tight, by which one means that given any function $w \in \mathcal{W}$, to every $\varepsilon > 0$ there corresponds a bounded set B_{ε} such that for all Q in $\mathcal{M}$,

$$\int_{S \setminus B_{\varepsilon}} |w(s)|\, Q(ds) < \varepsilon.$$

The desired epi-convergence result will follow from these conditions. We begin by a number of preliminary lemmas.

Lemma 7. *Let $g : S \longrightarrow (0,1)$ be an $\mathcal{A}$-measurable function. Let $k \in \mathbb{N}$ and for $i = 0,\dots,k$, let*

$$A_i = \{s \in S \,|\, g(s) > \frac{i}{k}\}$$

$$B_i = \{s \in S \,|\, g(s) \geq \frac{i}{k}\}.$$

Then

$$\frac{1}{k}\sum_{i=1}^{k} \mu(B_i) \leq \int_S g(s)\,\mu(ds) \leq \frac{1}{k} + \frac{1}{k}\sum_{i=1}^{k} \mu(A_i).$$

Proof. The proof is reminiscent of the one used to obtain the Portemanteau theorem.

$$\begin{aligned}
\frac{1}{k}\sum_{i=1}^{k}\mu(B_i) &= \sum_{i=1}^{k}\frac{i-1}{k}[\mu(B_{i-1})-\mu(B_i)]\\
&= \sum_{i=1}^{k}\frac{i-1}{k}\mu(B_{i-1}\setminus B_i)\\
&= \sum_{i=1}^{k}\frac{i-1}{k}\mu\{s \mid \frac{i-1}{k}\leq g(s)<\frac{i}{k}\}\\
&\leq \int_S g(s)\,\mu(ds)\\
&\leq \sum_{i=1}^{k}\frac{i}{k}\mu\{s \mid \frac{i-1}{k}< g(s)\leq\frac{i}{k}\}\\
&= \sum_{i=1}^{k}\frac{i}{k}\mu(A_{i-1}\setminus A_i)\\
&\leq \frac{1}{k}+\frac{1}{k}\sum_{i=1}^{k}\mu(A_i).
\end{aligned}$$

And this completes the proof. (Note that because $A_i \subset B_i$, the inequalities remain valid with A_i or B_i on both sides.) □

Theorem 8. *Let $g : S \longrightarrow \mathbb{R}$ be measurable and bounded on a bounded measurable subset E of S. Suppose that the upper level sets of the restriction of g to E are μ-continuity sets, i.e., for all $\alpha \in \mathbb{R}$, the boundaries of the sets*

$$\text{lev}_{\alpha}^{\geq} f \cap E = \{s \in E \mid g(s) \geq \alpha\}$$

carry no μ-measure. Then, provided the probability measures $\{\mu^\nu, \nu \in \mathbb{N}\}$ narrowly converge to μ,

$$\lim_{\nu\to\infty}\int_E g(s)\,\mu^\nu(ds) = \int_E g(s)\,\mu(ds).$$

Proof. Without loss of generality we can assume that $g(E) \subset (0,1)$. Pick a (large) integer k, and observe that from lemma 7 it follows that

$$\int_E g(s)\,\mu^\nu(ds) \leq k^{-1}\left(1+\sum_{i=1}^{k}\mu^\nu(B_i)\right),$$

where, as in lemma 7,

$$B_i = \{s \in E \mid g(s) \geq \frac{i}{k}\}.$$

Moreover, $\lim_{\nu\to\infty} \mu^\nu(B_i) = \mu(B_i)$ since by assumption, the sets $\{B_i, i = 0, \ldots, k\}$ are μ-continuity sets. Hence for ν sufficiently large, $\mu^\nu(B_i) < \mu(B_i) + k^{-1}$, and thus

$$\int_E g(s)\,\mu^\nu(ds) < \frac{1}{k}\sum_{i=1}^{k} \mu(B_i) + \frac{2}{k} < \int_E g(s)\,\mu(ds) + \frac{2}{k}.$$

From which it follows that

$$\limsup_{\nu\to\infty} \int_E g(s)\,\mu^\nu(ds) \le \int_E g(s)\,\mu(ds).$$

Interchanging the role of μ^ν and μ in the preceding relations, yields the desired result. □

Theorem 8'. *Let $g : S \longrightarrow \mathbb{R}$ be continuous and bounded on a bounded μ-continuity subset E of S and suppose the probability measures $\{\mu^\nu, \nu \in \mathbb{N}\}$ narrowly converge to μ. Then,*

$$\lim_{\nu\to\infty} \int_E g(s)\,\mu^\nu(ds) = \int_E g(s)\,\mu(ds).$$

Proof. Because $\mathrm{bdy}(\mathrm{cl}\,E) \subset \mathrm{bdy}E$, and thus $\mathrm{cl}\,E$ is also a μ-continuity set, we may as well assume that E is closed. Define B_i as in the proof of theorem 8 and note that these sets are thus also closed. Now appeal to the classical Portemanteau theorem giving $\limsup_\nu \mu^\nu(B_i) \le \mu(B_i)$ to conclude

$$\limsup_{\nu\to\infty} \int_E g(s)\,\mu^\nu(ds) \le \int_E g(s)\mu(ds).$$

Now apply lemma 7, with $\mathrm{int}\,E$ the underlying space (instead of S) and the collection of open sets A_i. Again appeal to the Portemanteau theorem, using the fact that the sets A_i are open, obtains

$$\liminf_{\nu\to\infty} \int_{\mathrm{int}\,E} g(s)\,\mu^\nu(ds) \ge \int_{\mathrm{int}\,E} g(s)\,\mu(ds)$$

which completes the proof (noticing that $\mu(\mathrm{bdy}E) = 0$). □

The next result is well known for the case when the integrand g is a continuous bounded function; again refer to the Portemanteau theorem (for a version that fits our needs, cf. [**6**]).

Lemma 9. *Suppose $g : S \longrightarrow \mathbb{R}$ is measurable and bounded on bounded subsets of S such that the strict level sets $\mathrm{lev}_\alpha^<$ of g are μ-continuity sets. Suppose, that the probability measures μ^ν converge narrowly to the measure μ, and that $\mathcal{M} = \{\mu; \mu^\nu, \nu \in \mathbb{N}\}$ are $\mathcal{W}$-tight with $\mathcal{W} = \{g\}$. Then*

$$\lim_{\nu\to\infty} \int_S g(s)\,\mu^\nu(ds) = \int_S g(s)\,\mu(ds).$$

Proof. For fixed $\varepsilon > 0$, we choose a ball B_ε with the following properties: for all $Q \in \mathcal{M}$,

$$\int_{S\setminus B_\varepsilon} |g(s)|\, Q(ds) < \frac{\varepsilon}{3} \quad \text{and} \quad B_\varepsilon \text{ is a } \mu\text{-continuity } set.$$

This is possible because of the $\mathcal{W}$-tightness condition and the fact that the function $r \mapsto \mathbb{B}_r$ can only have a countable number of discontinuities. From theorem 8, using the formula $\text{bdry}(A \cap B) \subset \text{bdry}\, A \cup \text{bdry}\, B$, for ν sufficiently large, we obtain:

$$\begin{aligned} \Big|\int_S g(s)\, \mu^\nu(ds) - &\int_S g(s)\, \mu(ds)\Big| \\ &\leq \int_{S\setminus B_\varepsilon} |g(s)|\, \mu^\nu(ds) + \int_{S\setminus B_\varepsilon} |g(s)|\, \mu(ds) \\ &\quad + \Big|\int_{B_\varepsilon} g(s)\, \mu^\nu(ds) - \int_{B_\varepsilon} g(s)\, \mu(ds)\Big| < \varepsilon. \end{aligned}$$

The proof is completed by letting ε tend to 0. □

The (topological) *limit superior* of a sequence $\{C^\nu \subset S, \nu \in \mathbb{N}\}$ is as usual defined by:

$$\begin{aligned} \text{Lim sup}_{\nu\to\infty} C^\nu &:= \{x = \lim_{k\to\infty} x^k \,\big|\, \forall k,\, x^k \in C^{\nu_k}; \{\nu_k\} \subset \mathbb{N}\} \\ &= \cap\{\text{cl}\, C \,\big|\, \forall x \in E,\, d(x, C) \leq \liminf_{\nu\to\infty} d(x, C^\nu)\}. \end{aligned}$$

Note that the limit superior of a sequence is always closed. For more about set limits, consult [**1**, **11**].

Lemma 10. *Suppose $\{C; C^\nu \subset S, \nu \in \mathbb{N}\}$ are $\mathcal{A}$-measurable and $C \supset \text{Lim sup}_\nu C^\nu$. Then $\liminf_\nu \mu^\nu(S \setminus C^\nu) \geq \mu(S \setminus C)$.*

Proof. Let $\varepsilon A = \{s \,|\, d(s, A) < \varepsilon\}$. It is easy to see that if $A \supset \text{Lim sup}_{\nu\to\infty} A^\nu$, then $\text{Lim sup}_{\nu\to\infty}(A^\nu \setminus \varepsilon A) = \emptyset$ for all $\varepsilon > 0$, cf. [**11**, theorem 2.2] for example. Hence $\text{Lim sup}_{\nu\to\infty}(C^\nu \setminus \varepsilon C) = \emptyset$ for all $\varepsilon > 0$. This implies [**6**],

$$\limsup_{\nu\to\infty} \mu^\nu(C^\nu) \leq \mu(C).$$

Taking complements yields the desired result. □

Lemma 11. *Suppose $\{g; g^\nu : S \to (0, \infty]\}$ such that for all s and any sequence $\{s^\nu, \nu \in \mathbb{N}\}$ converging to s,*

$$\liminf_{\nu\to\infty} g^\nu(s^\nu) \geq g(s),$$

equivalently $g \leq \text{epi-lim}_{\nu\to\infty} g^\nu$. *Suppose moreover that these functions are equi-bounded on bounded subsets of* S. *Then*

$$\liminf_{\nu\to\infty} \int_S g^\nu(s)\,\mu^\nu(ds) \geq \int_S g(s)\,\mu(ds).$$

Proof. Let E be an arbitrary bounded subset of S. Without loss of generality, we can assume that on E the range of the functions $\{g; g^\nu, \nu \in \mathbb{N}\}$ is included in $(0,1)$. Applying lemma 7, we obtain the following inequalities:

$$\int_E g(s)\,\mu(ds) \leq \frac{1}{k}(1 + \sum_{i=1}^{k} \mu(A_i)),$$

and

$$\int_E g^\nu(s)\,\mu^\nu(ds) \geq \frac{1}{k}\sum_{i=1}^{k} \mu^\nu(A_i^\nu),$$

with the sets A_i and A_i^ν as in lemma 7.

Observe that by assumption,

$$\text{Lim}\sup_{\nu\to\infty}(S \setminus A_i^\nu) \subset (S \setminus A_i).$$

We can appeal to lemma 10 to claim that $\liminf_{\nu\to\infty} \mu^\nu(A_i^\nu) \geq \mu(A_i)$, which with the preceding inequalities proves the assertion on the bounded set E. Thus, for all bounded sets E,

$$\int_E g^\nu(s)\,\mu^\nu(ds) \leq \int_S g^\nu(s)\,\mu^\nu(ds),$$

and hence

$$\liminf_{\nu\to\infty} \int_S g^\nu(s)\,\mu^\nu(ds) \geq \int_E g(s)\,\mu(ds).$$

Taking the supremum over all sets E completes the proof. □

Lemma 11 provides immediately the following result when the functions $\{g^\nu, \nu \in \mathbb{N}\}$ *continuously* converge to g; a special case of this theorem was proved in [8].

Theorem 12. *Let* $\{k; k^\nu : S \to \mathbb{R}\}$ *be a family of measurable functions with the property that given any* $s \in S$ *and any sequence* $\{s^\nu, \nu \in \mathbb{N}\}$ *converging to* s,

$$\lim_{\nu\to\infty} k^\nu(s^\nu) = k(s).$$

Suppose also that there exists a measurable function g *such that* $|k| \leq g, |k^\nu| \leq g$ *for all* ν, *and*

$$\lim_{\nu\to\infty} \int_S g(s)\,\mu^\nu(ds) = \int_S g(s)\,\mu(ds).$$

Then,

$$\lim_{\nu\to\infty}\int_S k^\nu(s)\,\mu^\nu(ds) = \int_S k(s)\,\mu(ds).$$

Proof. Apply lemma 11 to $g - k^\nu$ and $g + k^\nu$, in order to obtain

$$\liminf_{\nu\to\infty}\int_S (g+k^\nu)(s)\,\mu^\nu(ds) \ge \int_S (g+k)(s)\,\mu(ds),$$

and

$$\liminf_{\nu\to\infty}\int_S (g-k^\nu)(s)\,\mu^\nu(ds) \ge \int_S (g-k)(s)\,\mu(ds).$$

The assertion now follows from the fact that $\lim_{\nu\to\infty}\int_S g(s)\,\mu^\nu(ds) = \int_S g(s)\,\mu(ds)$. □

Note that a sufficient condition to guarantee that $\lim_{\nu\to\infty}\int_S g\,\mu^\nu(ds) = \int_S g\,\mu(ds)$ is to have g continuous plus either g bounded or the measures $\{\mu;\mu^\nu,\nu\in\mathbb{N}\}$ $\{g\}$-tight.

We are now ready to prove the main results.

Theorem 13. Pointwise convergence of integral functionals. *Suppose that f is an integrand that satisfies the* conditions 1-6, *then for all $x \in X$:*

$$\lim_{\nu\to\infty} E^\nu f(x) = Ef(x).$$

Proof. If $x \notin D$ then both sides are $+\infty$. If $x \in D$, apply lemma 9 with $g = f(x,\cdot)$. □

Theorem 14. Mosco-epi-convergence of integral functionals. *Suppose that f is an integrand that satisfies the* conditions 1-6, *then*

$$Mosco\text{-}\operatorname*{epi\text{-}lim}_{\nu\to\infty} E^\nu f = Ef.$$

Proof. There are two conditions that need to be checked for Mosco-epi-convergence (for functions defined on a reflexive Banach space)[**1**]: for all x in X:

for any sequence $\{x^\nu,\ \nu = 1,\ldots\}$ converging weakly to x, $\liminf_{\nu\to\infty} E^\nu f(x^\nu) \ge Ef(x)$,

and

there exists $\{\hat{x}^\nu,\ \nu = 1,\ldots\}$ converging strongly to x such that $\limsup_{\nu\to\infty} E^\nu f(\hat{x}^\nu) \le Ef(x)$.

In view of theorem 13, it suffices to verify the first one of these conditions, and the only interesting case is when both x and the sequence $\{x^\nu,\nu\in\mathbb{N}\}$ lie in D. The set

$$\cup\{x^\nu\}\cup\{x\} \subset r\,\mathbb{B}_X \quad \text{for some } r,$$

hence we can apply lemma 11 to the functions

$$g^\nu := f(x^\nu, \cdot) - u_r,$$
$$g := f(x, \cdot) - u_r,$$

where the functions u_r are those introduced in condition 4. The upper semicontinuity of u_r guarantees that for any sequence $\{s^\nu, \nu \in \mathbb{N}\}$ converging to s, $\liminf_{\nu\to\infty} g^\nu(s^\nu) \geq g(s)$. From lemma 11, it follows that

$$\liminf_{\nu\to\infty} \int_S [f(x^\nu, s) - u_r(s)]\, \mu^\nu(ds) \geq \int_S [f(x, s) - u_r(s)]\, \mu(ds).$$

To conclude, it will suffice to verify that

$$\limsup_{\nu\to\infty} \int_S u_r(s)\, \mu^\nu(ds) \leq \int_S u_r(s)\, \mu(ds).$$

We simply apply lemma 11 to the functions $g^\nu = g := -u_r$. □

2. An optimal control problem

We consider the following problem:

$$\text{minimize} \int_0^1 g(x(t), t)\, dt$$

where $g : \mathbb{R}^n \times [0,1] \to \mathbb{R}$ is continuous, dt is the Lebesgue measure on $[0,1]$ and $t \mapsto x(t)$ is solution of the system:

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \;\; a.e., \;\; x(0) = x_0 \tag{1}$$

where $A \in L^\infty([0,1]; \mathbb{R}^n \times \mathbb{R}^n)$, $B \in L^\infty([0,1]; \mathbb{R}^n \times \mathbb{R}^k)$, and $u : [0,1] \to \mathbb{R}^k$ is a control function in $L^2([0,1]; \mathbb{R}^k)$ such that for $i = 1, \ldots, k$, $|u_i(t)| \leq 1$ *a.e.*. It is well known that the solution, for a fixed control function u, of the preceding dynamical system is an AC-function (absolutely continuous function) x with derivatives in $L^2([0,1]; \mathbb{R}^n)$.

Now let us consider a sequence of probability measures μ^ν that converge narrowly to dt; for instance, finitely supported measures that would be generated by discretization of the "time" interval $[0,1]$. In terms of the general framework introduced in section 1, we have $S = [0,1]$, and X is the Sobolev space $H^1([0,1], \mathbb{R}^n)$ (with norm $\|x\|_H^2 = \int_0^1 |x(t)|^2\, dt + \int_0^1 |\dot{x}(t)|^2\, dt$). With $Q : X \times S \to \mathbb{R}^n \times S$ the projection mapping $Q(x(\cdot), t) = (x(t), t)$, the function $f : X \times S \to \overline{\mathbb{R}}$ is given by

$$f(x(t), t) = \begin{cases} (g \circ Q)(x(\cdot), t) & \text{if } x \text{ satisfies (1) for some } u, \\ \infty \text{ otherwise} . \end{cases}$$

Let us note that the operator Q has the following compactness property: for x^ν converging to x weakly in H^1 and $t^\nu \to t$, then $Q(x^\nu, t^\nu)$ converges to $Q(x,t)$. This is essentially a direct consequence of the theorem of Kondrachev [7, theorem 2.2.3] which states that if x^ν converges weakly to x in H^1 then the x^ν converge uniformly to x on $[0,1]$. From this it easily follows that x^ν converging weakly to x and $t^\nu \to t$ imply $x^\nu(t^\nu) \to x(t)$ from which the compactness property follows.

In order to apply theorem 13, all what is needed is to check the conditions 1-6. It is easy to check that f is a normal integrand, cf. [**10**]. Let

$$D := \{x \in H^1 \,|\, \exists\, u \text{ such that } x \text{ solves } (1)\ \}$$

and $\operatorname{dom} f = D \times [0,1]$. To show that D is weakly closed let us consider a sequence x^ν converging weakly to $\bar{x}$ with the x^ν in D. Let u^ν be a sequence of admissible controls that generate the trajectories x^ν. The sequence u^ν lies in the unit ball in L^2, and thus admits a subsequence converging weakly to a control $\bar{u}$. Weak convergence in L^2 implies a.e.-pointwise convergence, hence $\bar{u}_i(t) \leq 1$ a.e., i.e. $\bar{u}$ is admissible. Passing to the limit (for the Cauchy problem) shows that $\bar{x}$ is the solution of (1) for the control $\bar{u}$. This takes care of condition 1.

Condition 2 follows from the continuity of g and the fact that weak convergence of a sequence x^ν to x and $t^\nu \to t$ implies $Q(x^\nu, t^\nu) \to Q(x,t)$.

For condition 3 we need only to observe that for all $\hat{x}$ in D, $|\hat{x}_i(t) - x_{0i}| \leq \rho$ a.e. for some $\rho \in \mathbb{R}_+$. That and the continuity of g is all what is needed.

The same argument shows that both conditions 4 and 6 are satisfied. And again the continuity of g in conjunction with the continuity of the solutions x of (1) is enough to take care of condition 5(alt).

3. A mid-course maneuver problem

This example is also an optimal control problem. We consider a system to be steered from some initial state x_0 to a final state x_T but with a twist. The state x_0 is only known in probability and cannot be directly observed. After an initial phase during which the evolution of the system is tracked, we are allowed to make a mid-course correction. One refers to the class of such problems as *mid-course maneuver problems* [**14, 15**].

The problem can be formulated as follows: let U and W be reflexive Banach spaces, V a Hilbert space, $\Xi = \mathbb{R}^N$ and $\mathcal{U}$ a closed bounded convex subset of U. Let

$S : U \times \Xi \to W$ bounded linear in u and linear in ξ for $u \in \mathcal{U}$,

$p : \Xi \to W$ a linear operator ,

$R : V \to W$ a bounded linear operator ,

$l : U \to \overline{\mathbb{R}}$ a convex continuous function ,

$q : V \times \Xi \to [0, \infty]$,

μ an absolutely continuous probability measure on Ξ.

We consider the following optimization problem:

$$\min_{u \in \mathcal{U}} \left\{ l(u) + \int_{\Xi} \inf_{v} \{ q(v, \xi) \,|\, R(v) = p(\xi) - S(u, \xi) \} \, d\mu \right\}.$$

A typical example would be:

$$S(u, \xi) = \Phi(0, T)\xi + \int_0^t \Phi(0, T)\Phi(t, \tau)B(\tau)u(\tau)\, d\tau,$$

$$R(v) = \int_t^T \Phi(t, T)\Phi(T, \tau)B(\tau)v(\tau)\, d\tau,$$

$\Phi(\cdot, \cdot)$ the fundamental solution of the linear dynamical system:

$$(d/d\tau)\Phi(\tau, \tau') = A(\tau)\Phi(\tau, \tau'), \ \Phi(\tau, \tau) = I,$$

$A(\tau) : [0, t] \to \mathbb{R}^n \times \mathbb{R}^n$ continuous ,

$B(\tau) : [0, t] \to \mathbb{R}^n \times \mathbb{R}^d$ continuous ,

ξ random initial state at time $t = 0$,

$p \equiv x_T$ (the terminal state) ,

$\mathcal{U} \subset U$ the space of admissible controls , $u : [0, T] \to \mathbb{R}^d, |u(t)| \leq 1$, a.e.

$U = L^2([0, t]; \mathbb{R}^d), V = L^2([t, T]; \mathbb{R}^d)$

$$l(u) = \int_0^t |u(\tau)|^2 \, d\tau,$$

$$q(v) = \int_t^T |v(\tau)|^2 \, d\tau.$$

The preceding problem would model the situation when we seek to bring a system from an unknown state $x(0) = \xi$ to a final state $x(T) = p$ given the dynamics of the system: $\dot{x} = A(\tau)x + B(\tau)w$ using a control w, a concatenation of a control u for the time period up to t and a control v for the time period that follows t. The system is tracked up to time t and the actual state of the system is observed at time t. The choice of the control v can be viewed as a corrective action (a recourse decision). The performance is measured by the value of $l(u) + q(v)$, i.e., in terms of energy usage. The objective is to minimize this function.

Returning to the general formulation of the stochastic optimization problem, let

$$f(u, \xi) = \begin{cases} \inf_v \{q(v, \xi) \mid R(v) = p(\xi) - S(u, \xi)\}\, d\mu \text{ if } u \in \mathcal{U}, \\ +\infty \text{ otherwise.} \end{cases}$$

With this definition, the problem at hand can thus be formulated as

$$\text{minimize } l(u) + \int_\Xi f(u, \xi)\, \mu(d\xi) \text{ for } u \in \mathcal{U}.$$

We are interested in replacing the probability measure μ by measures μ^ν (possibly discrete probability measures) and thus rely on the results in section 1 to claim epi-convergence, more precisely Mosco-epi-convergence. Because epi-convergence is preserved under addition of a continuous convex function, we can ignore the additive term $l(u)$. Thus it will be enough to check if

$$\int_\Xi f(u, \xi)\, \mu^\nu(\xi) \text{ Mosco-epi-converges to } \int_\Xi f(u, \xi)\, \mu(d\xi).$$

We are going to assume that the problem at hand possesses the following properties, and we shall see that in turn they imply conditions 1-6:

(a) R is onto; the problem is said to have *complete recourse*.

(b) q is proper, convex, lsc and bounded on bounded subsets of $V \times \Xi$.

(c) $\lim_{\|v\| \to \infty} \inf_{\xi \in B} q(v, \xi) = \infty$ for B any bounded subset of Ξ.

We are going to sketch the proof that these conditions are enough to guarantee that conditions 1-5 are satisfied. Thus any sequence of probability measures satisfying condition 6 will engender the epi-convergence of the functionals (theorem 14). We begin with a couple of preliminary lemmas. Henceforth, we assume that (a)-(c) hold.

Lemma 15. *Let u^ν be a sequence in $\mathcal{U}$ that converges weakly to u, and ξ^ν a sequence in Ξ converging to ξ. Suppose v^ν is a sequence in V such that*

$$R(v^\nu) = p(\xi^\nu) - S(u^\nu, \xi^\nu), \tag{2}$$

$$q(v^\nu, \xi^\nu) \leq \inf\{q(v, \xi^\nu) \,|\, R(v) = p(\xi^\nu) - S(u^\nu, \xi^\nu)\} + (1/\nu). \tag{3}$$

Then the sequence v^ν is bounded.

Proof. Consider the restriction of R to $\ker(R)^\perp$, the orthogonal complement of the kernel of R. The map $R : \ker(R)^\perp \to W$ is an isomorphism. Thus, there exists $\kappa > 0$ such that $\|R(z)\| \geq \kappa\|z\|$ for all $z \in \ker(R)^\perp$. Now consider the system

$$R(v) = p(\xi) - S(u, \xi), v \in \ker(R)^\perp$$

for $u \in \mathcal{U}$ and $\xi \in B$ a bounded subset of Ξ. For all $(u, \xi) \in \mathcal{U} \times B$, the system has a unique solution $v(u, \xi)$ and there exists $\alpha > 0$ such that $\|v(u, \xi)\| \leq \alpha$. Thus there exists at least one sequence $\bar{v}^\nu$ that is bounded and satisfies (2).

Because q is bounded on bounded sets (assumption (b)), it follows that there exist θ and η such that

$$\eta \leq \inf\{q(v, \xi^\nu) \,|\, R(v) = p(\xi^\nu) - S(u^\nu, \xi^\nu)\} \leq q(\bar{v}^\nu, \xi^\nu) \leq \theta.$$

It now suffices to appeal to assumption (c) to conclude that a sequence that satisfies both (2) and (3) must be bounded. □

Lemma 16. *The function $f : U \times \Xi \to \overline{\mathbb{R}}$ as defined above, is a convex normal integrand.*

Proof. Lower semicontinuity of f with respect to u follows directly from lemma 15. The proof will certainly be complete if we show that $(u, \xi) \mapsto f(u, \xi)$ is (Borel-)measurable as we do next. Note that for $\alpha \in \mathbb{R}$

$$\begin{aligned} \{(u, \xi) \,|\, f(u, \xi) < \alpha\} &= \mathrm{prj}_{U\times\Xi}\{(u, v, \xi) \,|\, v \in \Gamma(\xi, u), q(v, \xi) < \alpha\} \\ &= \mathrm{prj}_{U\times\Xi}\{\{(u, v, \xi) \,|\, q(v, \xi) < \alpha\} \cap \mathrm{gph}\,\Gamma\} \end{aligned}$$

where

$$\Gamma(\xi, u) = \{v \,|\, R(v) = p(\xi) - S(u, \xi)\}.$$

Now, $\mathrm{gph}\,\Gamma$ is convex and thus a measurable subset of $U \times \Xi$, and since $\{(u, v, \xi) \,|\, q(v, \xi) < \alpha\}$ is measurable (q is itself a normal integrand), the measurability of $\{(u, \xi) \,|\, f(u, \xi) < \alpha\}$ now follows from the projection theorem for measurable sets, cf. [**2,** lemma III.39, theorem III.23]. □

Condition 1 follows directly from the preceding lemma and the fact that $\mathrm{dom}\, f = \mathcal{U} \times \Xi$; certainly $\mathcal{U}$ is weakly closed.

To check condition 2, we need to show that if u^ν (in $\mathcal{U}$) converges weakly to u and ξ^ν converges to ξ, then

$$\liminf_{\nu\to\infty} f(u^\nu,\xi^\nu) \geq f(u,\xi).$$

It suffices to consider a sequence in the effective domain of f since otherwise the lim inf term is $+\infty$. Let v^ν be such a sequence that also satisfies (2) and (3). Lemma 15 tells us that such a sequence is bounded. Hence some subsequence converges weakly to $\bar{v}$ in v. Because, $S(u^\nu,\xi^\nu)$ converges weakly to $S(u,\xi)$, it follows that $R(v^\nu)$ converges weakly to $p(\xi) - S(u,\xi)$. The graph of R is a closed, convex subset of $V \times W$, hence weakly closed. Consequently, $R(\bar{v}) = p(\xi) - S(u,\xi)$ demonstrating that $\bar{v}$ is an admissible solution. Moreover,

$$f(u,\xi) \leq q(\bar{v},\xi) \leq \liminf_{\nu\to\infty} q(v^\nu,\xi^\nu) = \liminf_{\nu\to\infty} f(u^\nu,\xi^\nu)$$

where the second inequality follows from assumption (b) and the last equality from (3).

For condition 3, note that $f \geq 0$, and the argument used in the proof of lemma 15 shows that f admits an upper bound on bounded sets.

Condition 4 is trivially satisfied with $h = 0$.

It is routine to show that for $u \in \mathcal{U}$ the function $\xi \mapsto f(u,\xi)$ is convex, cf. [**15**], for example. Hence its level sets are convex. Now, in Euclidean space, every convex set is a μ-continuity set since by assumption μ is absolutely continuous with respect to the Lebesgue measure [**8, 6**]. This is all what is needed for condition 5.

Remark. Because this example is mostly here to illustrate the use one can make of the results of section 1 in this context, we have not provided the most general conditions under which one can ensure that conditions 1-5 are satisfied. For example, linearity of S with respect to u could be replaced simply by weak continuity, that is really all that gets used here. Similarly, the restrictive convexity assumptions on q (that are automatically satisfied if q does not depend on ξ and is convex in v) and the linearity of p and S with respect to ξ are only needed to guarantee condition 5. One would expect that in most concrete situations it will be possible to argue that this condition is satisfied without appeal to convexity, or alternatively use 5(alt).

References

[1] Hedy Attouch, *Variational Convergence for Functions and Operators*, Pitman, London, 1984, Applicable Mathematics Series .

[2] Charles Castaing & Michel Valadier, *Convex Analysis and Measurable Multifunctions*, Springer Verlag Lecture Notes in Mathematics 580, Berlin, 1977.

[3] Jitka Dupačová & Roger J-B Wets, "Asymptotic behavior of statistical estimators and of optimal solutions for stochastic optimization problems," *The Annals of Statistics* **16** (1988), 1517–1549.

[4] Peter Kall, "On approximations and stability in stochastic programming," in *Parametric Optimization and Related Topics*, J. Guddat, H.Th. Jonge, B. Kummer & F. Nozicka, eds., Akademie-Verlag, Berlin, 1987.

[5] Alan J. King & Roger J-B Wets, "Epi-consistency of convex stochastic programs," *Stochastics and Stochastics Reports* (1989 (to appear)), IBM Research Division, RC 14640 (#65642).

[6] Roberto Lucchetti, Gabriella Salinetti & Roger J-B Wets, "Uniform convergence of probability measures: topological criteria," manuscript, University of California-Davis, November 1988.

[7] Jurg T. Marti, *Introduction to Sobolev Spaces and Finite Element Solution of Elliptic Boundary Value Problems*, Academic Press, London, 1986.

[8] R. Ranga Rao, "Relations between weak and uniform convergence of measures with applications," *Annals of Mathematical Statistics* **33** (1962), 659–680.

[9] Stephen M. Robinson & Roger J-B Wets, "Stability in two-stage stochastic programming," *SIAM J. on Control and Optimization* **25** (1987), 1409–1416.

[10] R.T. Rockafellar, "Integral functionals, normal integrands, and measurable selections," in *Nonlinear Operators and the Calculus of Variations*, J-P. Gossez, ed., Springer-Verlag Lecture Notes in Mathematics, no. 543, Berlin, 1976, 157–207.

[11] Gabriella Salinetti & Roger J-B Wets, "On the convergence of closed-valued measurable multifunctions," *Transactions of the American Mathematical Society* **266** (1981), 275–289.

[12] Gabriella Salinetti & Roger J-B Wets, "On the convergence in distribution of measurable multifunctions (random sets), normal integrands, stochastic processes and stochastic infima," *Mathematics of Operations Research* **11** (1986), 385–419.

[13] Gabriella Salinetti & Roger J-B Wets, "Weak convergence of probability measures revisited," IIASA WP-87-30, April 1987.

[14] C.W. Sartuse, "Rocket vehicle guidance," in *Inertial Guidance*, F.R. Pitman, ed., J. Wiley, New York, 1962.

[15] Richard Van Slyke & Roger J-B Wets, "Stochastic programs in abstract spaces," in *Stochastic Optimization and Control*, H. Karreman, ed., J. Wiley, New York, 1968, 24–45.

[16] Roger J-B Wets, "Induced constraints for stochastic optimization problems," in *Techniques of Optimization*, A. Balakrishnan, ed., Academic Press, New York, 1972.