Lectures on Young Measures *

Erik J. Balder †
CEREMADE, Université de Paris-Dauphine

Contents

*Based on lectures given at Université de Paris-Dauphine (June 1994); these notes will appear in the series *Cahiers de Mathématiques de la Décision* of CEREMADE.

†On leave from Mathematical Institute, University of Utrecht, Netherlands

Preface

These notes were motivated by a series of lectures which I gave in June 1994 as an invited professor at the Centre de Recherche de Mathématiques de la Décision (CEREMADE) of Université de Paris-Dauphine. I am very grateful to CEREMADE for having been given this opportunity.

Much of the actual writing took place in Paris, but a trip to Vietnam, undertaken in September 1994, provided me with an occasion to improve many details in a a second round of work. For this reason I extend my thanks to the International Cooperation Bureau of the University of Utrecht, the Institute of Mathematics in Hanoi and the University of Hue for their support and hospitality. Invited lectures given in December 1994 at the Department of Applied Mathematics II of the University of Sevilla, inspired a final round of revisions to provide more transparency and to remove some unnecessary interdependence. I also wish to thank Michel Valadier for kindly providing me with a number of very insightful comments and suggestions.

It turned out to be possible to base a substantial part of the lectures on my own work on Young measures. Nevertheless, the task of presenting the material in a coherent and transparent fashion challenged me to come up with some fundamental refinements of the theory. This included finding answers to questions that had been open to at least myself. The comments in section 10 give an account of this.

A distinguishing feature of these notes is that they extend the classical narrow convergence theory for probability measures to the domain of Young measures, in a very direct and general fashion, via the auxiliary notion of K-convergence. This approach, of which a more rudimentary version can be found in [Ba17], has Komlós' theorem (Theorem 1.2) as a point of departure. The major results on classical narrow convergence, recapitulated in section 2, are then systematically expanded into results on narrow convergence for Young measures (sections 3 and 4). This natural route leads to the first applications of the theory to economical and game-theoretical models in section 4. For readers with a primary interest in the application of Young measures to functional analysis and optimal control it should be possible to skip to section 5 after the introductory section 1; certainly sections 6–8 can then be followed with only an occasional glance at the earlier material. Some familiarity with convex analysis and measurable multifunctions is expected from the reader, but quite a few advanced instruments have been recalled or derived in Appendix A. The subject of outer integration is treated separately in Appendix B.

A minor drawback of the K-convergence approach is that it fails to do full justice to the purely nonsequential aspects of narrow convergence for Young measures, but this is a small sacrifice in favor of directness, generality and transparency. At any rate, these lecture notes demonstrate that most of the usual applications suffer no loss of coverage. My choice for K-convergence also dispenses with the need for a detailed study of the approximation of normal integrands. Consequently, I have opted to ignore the subject of measurable regularization altogether. Extreme point considerations of Young measures, less natural for a sequentially oriented presentation, have also been avoided. Partly because of this and partly in view of the intended readership, no applications of Young measures to statistical decision theory have been included either.

Heelsum, January 1995

E.J.B.

1 Introduction

Contents: Historical remarks, introductory examples, Komlós' theorem, Young relaxation, functional relaxation

Young measures were originally conceived as 'generalized curves' by L.C. Young [Y1] in order to complete sets of ordinary curves in the calculus of variations. Young thus generalized the solution concept, and this proved to be the proper response to Hilbert's twentieth problem, because a broad class of problems in the calculus of variations could be shown to have solutions in the form of generalized curves. L.C. Young and E.J. McShane [Mc1] also showed that, under suitable convexity, the existence of optimal generalized curves entailed the existence of optimal ordinary curves. In the same decade, the development of a general decision theory for mathematical statistics by A. Wald [Wal] led to rather related existence results. (For instance, under certain convexity conditions essential completeness statements were obtained that had much to do with the discovery of Young and McShane observed above.) Nevertheless, this development of decision theory, which was substantially influenced by the advent of game theory [NM] and the general theory of weak convergence of probability measures [Pr, Var, L2], did not interact in any serious way with the calculus of variations. Thus, the different kinds of questions that were being asked continued to mask the fact that, in essence, these areas could be approached from one and the same theoretical perspective. This situation did not change until the seventies. By then, the interest in classical statistical decision theory had dwindled, and in optimal control theory the theory of relaxed controls (the continuation of Young's theory of generalized curves) [War, Gh, Gam] had a very limited influence, even upon the existence theory itself. However, a number of questions concerning existence and functional relaxation, with origins in applied analysis and mathematical economics, inspired developments on the interface of convex analysis, variational analysis and measure theory [AP, Ro, ET]. A paper of H. Berliocchi and J.-M. Lasry [BL] made the significance of Young measures for such questions quite patent. This inspired the present author to a series of papers which form the foundations of these lectures. These not only established that Young measure theory (in particular, relaxed control theory) could be merged harmoniously with the theory of statistical decision functions and topological measure theory [Ba1, Ba2, Ba3], but led also to the discovery of several new features, such as support properties of limit Young measures [Ba3], tensor product continuity [Ba11], K-convergence aspects [Ba15, Ba16, Ba17], extreme point properties [BL, Ba20, Ba21] and measurable regularization of integrands [Ba3]. In addition, new ways were found to apply Young measures and their by-products, and some of those methods are also discussed here.

The fundamental idea behind Young measures is that, in some sense, they form a *completion* of the ordinary functions. In the completion, the ordinary functions correspond to the *Dirac* Young measures (Definition 3.2). When the completion is equipped with the narrow convergence topology (Definition 3.3), the following two major advantages over classical modes of convergence, by their very nature restricted to ordinary functions, emerge:

i. *Narrow convergence is more informative.*

ii. *Narrow convergence is more readily available.*

Of course, such desirable properties exact a sacrifice: Narrow convergence, as a rule, is only available for *subsequences*. Moreover, a practical complication may be that the limiting

Young measure, which forms a generalized narrow limit for the (sub)sequence, is non-Dirac, and can therefore only be connected indirectly with an ordinary function. Nevertheless, restriction to subsequences does not hamper existence or closure-type arguments, and the typical non-Dirac nature of narrow limits is an immediate consequence of the greater precision achieved by narrow convergence. To illustrate these ideas, let us give an unexpected twist to the following classical situation:

Example 1.1 (Rademacher functions) Let $\Omega := [0,1]$ be equipped with the Lebesgue σ-algebra $\mathcal{A} := \mathcal{L}([0,1])$ and the Lebesgue measure λ_1. Let (f_n) be the sequence of *Rademacher functions* $f_n : [0,1] \to \mathbf{R}$, given by

$$f_n(\omega) := \text{sign } \sin 2^n \pi\omega.$$

Let (ϵ_{f_n}) be the corresponding sequence of Dirac Young measures $\epsilon_{f_n} : [0,1] \to \mathcal{P}(\mathbf{R})$, defined by [1]

$$\epsilon_{f_n}(\omega) := \epsilon_{f_n(\omega)} := \text{Dirac measure at } f_n(\omega).$$

Observe that the f_n, viewed as random variables on the Lebesgue unit interval, are independent and identically distributed. The strong law of large numbers applies [Bi2], but this merely gives

$$\frac{1}{N}\sum_{n'=1}^{N} f_{n'}(\omega) \to 0 \text{ in } \mathbf{R} \text{ as } N \to \infty \text{ for a.e. } \omega \text{ in } \Omega \tag{1.1}$$

for every subsequence $(f_{n'})$ of (f_n) (as usual, 'a.e.' stands for 'almost every' or 'almost everywhere', depending upon the context). While this implies weak convergence of (f_n) to $f_0 \equiv 0$ (the null function) in $\mathcal{L}^1_{\mathbf{R}}$, it is more subtle to realize that for every bounded continuous $c : \mathbf{R} \to \mathbf{R}$ the compositions $(c \circ f_n)$ also form an independent and identically distributed sequence of random variables. The same (scalar) strong law of large numbers, applied countably many times, [2] gives

$$\frac{1}{N}\sum_{n'=1}^{N} \epsilon_{f_{n'}}(\omega) \Rightarrow \frac{1}{2}[\epsilon_1 + \epsilon_{-1}] \text{ in } \mathcal{P}(\mathbf{R}) \text{ as } N \to \infty \text{ for a.e. } \omega \text{ in } \Omega \tag{1.2}$$

for every subsequence $(\epsilon_{f_{n'}})$ of (ϵ_{f_n}); here $\Rightarrow$ denotes *classical* narrow (alias weak) convergence of probability measures (see Definition 2.1). Now (1.2) is a special manifestation of what we shall come to know as *K-convergence* of (ϵ_{f_n}) to the non-Dirac Young measure δ_0, given by $\delta_0(\omega) \equiv \frac{1}{2}[\epsilon_1 + \epsilon_{-1}]$; *a fortiori,* this implies narrow convergence of (ϵ_{f_n}) to δ_0 in the sense of Young measures (see Corollary 3.14). $\square$

Comparison of these successive applications of the strong law of large numbers shows the significance of Young measures for limit phenomena: in (1.1) the scores of 1's and -1's registered at an arbitrary ω cancel each other out on average, but in (1.2) the Young measure device marks those scores – as ϵ_1 and ϵ_{-1} respectively – and averaging leaves much clearer evidence of the values taken by the f_n's. To see why narrow convergence and K-convergence of Young measures are typically *subsequence-oriented*, it is already enough to consider the sequence (f'_n) defined by taking $f'_n := f_n$ if n is odd and $f'_n := 2f_n$ if n is even. While (1.1) continues to hold for (f'_n), which therefore also converges weakly to 0 in

[1] $\mathcal{P}(\mathbf{R})$ stands for the set of all probability measures on $\mathbf{R}$.

[2] Here the countable determination property (Proposition 2.2) is used.

$\mathcal{L}^1_{\mathbf{R}}$, the more general narrow limit statement (1.2) no longer holds: the subsequence with odd indices gives pointwise the limit $\frac{1}{2}[\epsilon_1 + \epsilon_{-1}]$, but the even indices now give the different limit $\frac{1}{2}[\epsilon_2 + \epsilon_{-2}]$ in (1.2). Hence, $(\epsilon_{f'_n})$ fails to K-converge and, consequently, it also fails to converge narrowly (see Corollary 3.16).

The second advantage of Young measure convergence, easy availability, will be noticeable when we present the rather broad notion of *parametrized tightness* for Young measures and the associated version of Prohorov's theorem (Theorem 4.7). The following rather amazing fact will then be demonstrated: statements like (1.2) still hold in the presence of parametrized tightness, provided that – once again – we no longer expect them to hold for an entire sequence of Young measures, but only for a *subsequence.* This possibility, which leads immediately to the notion of K-convergence for Young measures, stems basically from replacing the scalar strong law of large numbers, as used above, by the following deep and beautiful result of J. Komlós [K, Cha1], which can be seen as its analogue by the renowned *subsequence principle* [Cha2]. Let $(\Omega, \mathcal{A}, \mu)$ be an abstract measure space (possibly not even σ-finite), and denote the space of all integrable functions from Ω into $\mathbf{R}$ by $\mathcal{L}^1_{\mathbf{R}} := \mathcal{L}^1_{\mathbf{R}}(\Omega, \mathcal{A}, \mu)$.

Theorem 1.2 (Komlós) *Let (ϕ_n) be a sequence in $\mathcal{L}^1_{\mathbf{R}}$ such that*

$$\sup_{n \in \mathbf{N}} \int_\Omega |\phi_n| d\mu < +\infty.$$

Then there exist a subsequence $(\phi_{n'})$ of (ϕ_n) and a function $\phi_* \in \mathcal{L}^1_{\mathbf{R}}$ such that for every further subsequence $(\phi_{n''})$ of $(\phi_{n'})$

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n''=1}^{N} \phi_{n''}(\omega) = \phi_*(\omega) \text{ for a.e. } \omega \text{ in } \Omega.$$

Observe here that ϕ_* is universal with respect to the possible choices of a subsequence $(\phi_{n''})$ from $(\phi_{n'})$, but that the associated exceptional μ-null set in the limit statement is allowed to vary with the subsequence.

It is tempting to conclude this introduction by indicating, within the context of Example 1.1, the significance which Young measures have for existence and relaxation questions, both in problems arising in the calculus of variations and optimal control and in problems in mechanics.

Example 1.3 (nonexistence of optimal controls) Let $\Omega := [0, 1]$ be as in Example 1.1. The following example of an optimal control problem goes back to Bolza:

$$(\mathcal{P}): \ \inf J(f) := \int_{[0,1]} [y_f(\omega)^2 + (f(\omega)^2 - 1)^2] d\omega$$

over all $f \in \mathcal{L}^1_{\mathbf{R}}$, with y_f given by $y_f(\omega) := \int_0^\omega f(\omega')d\omega'$. Then $(\mathcal{P})$ has no optimal solution: since $J \geq 0$ and $J(f_n) \to 0$ for the Rademacher functions (f_n) defined in Example 1.1, it follows that $\inf(\mathcal{P}) = 0$ (observe that $0 \leq y_{f_n}(\omega) \leq 2^{-n}$). But there cannot be $f \in \mathcal{L}^1_{\mathbf{R}}$, with $J(f) = 0$, since it would have to satisfy $y_f \equiv 0$ and $|f(\omega)| = 1$ a.e. at the same time. In an effort to establish existence at a generalized level, one defines the *Young relaxation* of $(\mathcal{P})$ as the following optimization problem:

$$(\mathcal{P}_{rel}): \ \inf I(\delta) := \int_{[0,1]} [y_\delta(\omega)^2 + \int_{\mathbf{R}} (x^2 - 1)^2 \delta(\omega)(dx)] d\omega$$

over all Young measures δ from $[0,1]$ into $\mathbf{R}$, that is, over all measurable $\delta : \Omega \to \mathcal{P}(\mathbf{R})$. Here $y_\delta(\omega) := \int_0^\omega [\int_{\mathbf{R}} x\, \delta(\omega')(dx)]d\omega'$ (observe the systematic convexification in the control argument and note that y_δ is well-defined whenever $I(\delta) < +\infty$). Then it is easy to inspect that the Young measure limit δ_0 of Example 1.1 is the essentially unique optimal solution of $(\mathcal{P}_{rel})$, since $y_{\delta_0} \equiv 0$ and $I(\delta_0) = 0$. Observe also that (ϵ_{f_n}), which corresponds to the minimizing sequence (f_n) for $(\mathcal{P})$, is not only a minimizing sequence for $(\mathcal{P}_{rel})$, but also converges to the true and essentially unique minimizer δ_0. In contrast, even though $\inf(\mathcal{P}) = \inf(\mathcal{P}_{rel}) = 0$ and (f_n) was seen to have $f_0 \equiv 0$ as its weak limit in $\mathcal{L}^1_{\mathbf{R}}$, the infimum of $(\mathcal{P})$ is not attained. The difference can be explained as follows. The relaxed objective I is lower semicontinuous for the narrow Young measure topology (Example 5.11), but the original objective J fails to be lower semicontinuous for the weak topology on $\mathcal{L}^1_{\mathbf{R}}$. This can be ascribed entirely to the nonconvexity of the control part $x \mapsto (x^2-1)^2$ of the integrand. Indeed, in what may be seen as an alternative effort to remedy the lower semicontinuity situation, the weakly lower semicontinuous hull $\bar{J}$ of J turns out to be as follows (Example 9.5):

$$\bar{J}(f) = \int_{[0,1]} [y_f(\omega)^2 + \max(f(\omega)^2 - 1, 0)^2] d\omega,$$

so $(x^2-1)^2$ is replaced by its convexification $\max(x^2-1,0)^2$. Thus, the *functional relaxation* $(\mathcal{P}_{frel})$ of $(\mathcal{P})$, defined by

$$(\mathcal{P}_{frel}) : \inf_{f \in \mathcal{L}^1_{\mathbf{R}}} \bar{J}(f),$$

satisfies $\inf(\mathcal{P}_{frel}) = \inf(\mathcal{P}) = 0$ and it does have the the zero function f_0 as its optimal solution. □

Thus, Young relaxation convexifies in the variables of control-type by enlarging the class of participating functions, whereas functional relaxation keeps itself to the original functions, but replaces the original objective integral functional by its lower semicontinuous hull. Of course, this causes the two types of relaxation to be rather different in general. However, in certain frequently studied situations, such as when the intended lower semicontinuity is weak lower semicontinuity on an $\mathcal{L}^1$-space, as in the above example, the latter amounts to a convexification of the objective integrand in a control-type variable. In this case close connections exist between the two types of relaxation, as we shall see in section 9. Most of these lecture notes will be concerned with Young relaxation; for a systematic treatment of functional relaxation the reader is referred to [Bu, ET, Val2].

2 Narrow convergence of probability measures

Contents: narrow convergence, portmanteau theorem, product narrow convergence, supports, tightness, Prohorov's theorem, Suslin spaces

Let (S,ρ) be a separable metric space [3] and let $\mathcal{C}_b(S)$ $[\mathcal{C}_u(S)]$ denote the set of all continuous [uniformly continuous] bounded functions from S into $\mathbf{R}$. Further, let $\mathcal{B}(S)$ be the Borel σ-algebra on S and let $\mathcal{P}(S)$ denote the set of all probability measures on $(S,\mathcal{B}(S))$. In this section first some well-known results on narrow convergence in $\mathcal{P}(S)$ are recapitulated [Bi1, DM]. This is followed by a synthesizing result for narrow convergence in $\mathcal{P}(S)$. Via K-convergence methods, all concepts and results of this section are to be generalized systematically into results on the narrow convergence of Young measures. [4]

Definition 2.1 (narrow convergence and topology) A sequence (ν_n) in $\mathcal{P}(S)$ is said to converge *narrowly* to ν_0 in $\mathcal{P}(S)$ (notation: $\nu_n \Rightarrow \nu_0$) if

$$\lim_n \int_S c\, d\nu_n = \int_S c\, d\nu_0 \text{ for every } c \in \mathcal{C}_b(S).$$

Correspondingly, the *narrow topology* on $\mathcal{P}(S)$ is defined as the weakest topology on $\mathcal{P}(S)$ for which all functionals $\nu \mapsto \int_S c\, d\nu$, $c \in \mathcal{C}_b(S)$, are continuous. □

Our next result is better known in the traditional expositions on narrow convergence [Bi1] in the form of convergence-determining classes of *sets*; it depends upon the separability assumption for S and it implies that $\mathcal{P}(S)$ is metrizable for the narrow topology.

Proposition 2.2 (narrow convergence determination) *There exists a countable subset $\mathcal{C}_0$ of $\mathcal{C}_u(S)$ such that for every sequence (ν_n) in $\mathcal{P}(S)$*

$$\lim_n \int_S c\, d\nu_n = \int_S c\, d\nu_0 \textit{ for all } c \in \mathcal{C}_0$$

implies that $\nu_n \Rightarrow \nu_0$. In particular, $\mathcal{C}_0$ separates the points of $\mathcal{P}(S)$.

PROOF. Let (x_i) be countable and dense in S. Consider the countable collection $\mathcal{B}$ consisting of all finite unions of open balls $B(x_i;r)$ with center x_i and rational radius $r > 0$. Then every open set is a nondecreasing union of sets from $\mathcal{B}$. Associate to each B in $\mathcal{B}$ the sequence $(c_{B,j})$ in $\mathcal{C}_u(S)$, where $c_{B,j}(x) := \phi(j\rho(x,B))$ with $\phi : \mathbf{R}_+ \to \mathbf{R}_+$ given by $\phi(\xi) := 1$ if $\xi = 0$, $\phi(\xi) := 0$ if $\xi \geq 1$ and $\phi(\xi) := 1-\xi$ if $0 \leq \xi \leq 1$. The resulting collection $\mathcal{C}_0$ of all such functions $c_{B,j}$ is countable. Let G be an arbitrary open subset of S. Then G is a countable union of balls $B(x_i;r_i)$, whence of the corresponding balls $B(x_i;r_i-1/j)$. By the hypothesis and monotone convergence it follows easily that $\liminf_n \nu_n(G) \geq \nu_0(G)$. By Theorem 2.3 below, this gives $\nu_n \Rightarrow \nu_0$. □

The following result on narrow convergence ([As, 4.5.1],[Bi1, Theorem 2.1]) is quite useful:

[3] Several definitions in this section continue to make sense when S is equipped with a more general topology; later this observation will be used without further comments or explanation.

[4] In order of appearance: concepts and results from Definition 2.1 up to Theorem 2.8 are generalized in section 3; those from Definition 2.10 up to Theorem 2.15 in section 4; finally, Theorem 2.20 is generalized in section 5.

Theorem 2.3 (portmanteau theorem) *Let (ν_n) and ν_0 be in $\mathcal{P}(S)$. The following are equivalent:*

(a) $\nu_n \Rightarrow \nu_0$.

(b) $\lim_n \int_S c\, d\nu_n = \int_S c\, d\nu_0$ *for every* $c \in \mathcal{C}_u(S)$.

(c) $\liminf_n \nu_n(G) \geq \nu_0(G)$ *for every open* $G \subset S$.

(d) $\liminf_n \int_S q\, d\nu_n \geq \int_S q\, d\nu_0$ *for every lower semicontinuous function* $q : S \to (-\infty, +\infty]$ *which is bounded from below.*

Let $\hat{\mathbf{N}} := \mathbf{N} \cup \{\infty\}$ be the usual Alexandrov-compactification of the natural numbers; observe that this is a compact metrizable space. Recall also that $\epsilon_n \in \mathcal{P}(\hat{\mathbf{N}})$ denotes the Dirac measure concentrated at n, $n \in \hat{\mathbf{N}}$.

Corollary 2.4 (product narrow convergence) *Let (ν_n) and ν_0 be in $\mathcal{P}(S)$.*

(i) If $\nu_n \Rightarrow \nu_0$ in $\mathcal{P}(S)$, then $\nu_n \times \epsilon_n \Rightarrow \nu_0 \times \epsilon_\infty$ in $\mathcal{P}(S \times \hat{\mathbf{N}})$.

(ii) If $\frac{1}{N}\sum_{n=1}^N \nu_n \Rightarrow \nu_0$ in $\mathcal{P}(S)$, then $\frac{1}{N}\sum_{n=1}^N \nu_n \times \epsilon_n \Rightarrow \nu_0 \times \epsilon_\infty$ in $\mathcal{P}(S \times \hat{\mathbf{N}})$.

PROOF. Let $c \in \mathcal{C}_u(S \times \hat{\mathbf{N}})$. Then for any $\eta > 0$ uniform continuity implies that there exists $p \in \mathbf{N}$ such that $\sup_{x \in S} |c(x,n) - c(x,\infty)| < \eta$ for $n \geq p$. This gives $|\int_S c(x,n)\nu_n(dx) - \int_S c(x,\infty)\nu_n(dx)| \leq \eta$ for $n \geq p$, so (i) follows by Theorem 2.3(b). Likewise, (ii) follows from $|N^{-1}\sum_{n=1}^N \int_S [c(x,n) - c(x,\infty)]\nu_n(dx)| \leq 2(p-1)N^{-1}\sup_{S\times\hat{\mathbf{N}}}|c| + (N-p+1)N^{-1}\eta$ for $N \geq p$ and letting N go to infinity. □

Neither Theorem 2.3 nor its Corollary 2.4 depends upon separability of S; in contrast, the following partial extension of Corollary 2.4 does [Bi1, Theorem 3.2]:

Theorem 2.5 (product narrow convergence) *Let S' be another separable metric space. Let $\nu_n \Rightarrow \nu_0$ in $\mathcal{P}(S)$ and let $\nu'_n \Rightarrow \nu'_0$ in $\mathcal{P}(S')$. Then*

$$\nu_n \times \nu'_n \Rightarrow \nu_0 \times \nu'_0 \text{ in } \mathcal{P}(S \times S').$$

Definition 2.6 (support of probabilities) The *support* of $\nu \in \mathcal{P}(S)$ is defined by

$$\text{supp } \nu := \cap\{F : F \subset S,\ F \text{ closed},\ \nu(F) = 1\}.$$

□

To study the support properties of narrow limit probability measures, the following notion is indispensible.

Definition 2.7 (Kuratowski limes superior) The *limes superior* of a sequence of subsets (A_n) of S is defined as the set, denoted by $\text{Ls}_n A_n$, of all $x \in S$ for which there exists a subsequence $(A_{n'})$ of (A_n), and elements $a_{n'} \in A_{n'}$, such that $x = \lim_{n'} a_{n'}$. □

Observe that when all A_n are singletons (say $A_n = \{a_n\}$), then $\text{Ls}_n A_n$ is just the set of all limit points of (a_n). Observe also that the Kuratowski limes superior has the following simple form:

$$\text{Ls}_n A_n = \cap_{p=1}^\infty \text{cl } \cup_{n \geq p} A_n. \tag{2.1}$$

by elementary considerations based on using the metric ρ (here 'cl' stands for closure). For nonmetrizable S such a simplification is *not* possible (the above identity being a strict inclusion in general).

Theorem 2.8 (support of narrow limits) *Let either* $\nu_n \Rightarrow \nu_0$ *or* $\frac{1}{N}\sum_{n=1}^{N} \nu_n \Rightarrow \nu_0$ *in* $\mathcal{P}(S)$*. Then*

$$\text{supp } \nu_0 \subset \text{Ls}_n \text{supp } \nu_n.$$

Proof. By Corollary 2.4 we have either $\nu_n \times \epsilon_n \Rightarrow \nu_0 \times \epsilon_\infty$ or $\frac{1}{N}\sum_{n=1}^{N} \nu_n \times \epsilon_n \Rightarrow \nu_0 \times \epsilon_\infty$ in $\mathcal{P}(S \times \hat{\mathbf{N}})$, as the case may be. Let us apply Theorem 2.3(d) to $q_0 : S \times \hat{\mathbf{N}} \to \{0, +\infty\}$ given by

$$q_0(x,k) := \begin{cases} 0 & \text{if } x \in \text{supp } \nu_k \text{ and } k < \infty \\ 0 & \text{if } x \in \text{Ls}_n\text{supp } \nu_n \text{ and } k = \infty \\ +\infty & \text{otherwise} \end{cases}$$

Then q_0 is lower semicontinuous at every point (x,k) of $S \times \hat{\mathbf{N}}$; for $k < \infty$ this follows by closedness of the supports, and for $k = \infty$ by Definition 2.7 and closedness of $\text{Ls}_n\text{supp } \nu_n$ (by (2.1)); observe that sequential arguments suffice by the metrizability of $S \times \hat{\mathbf{N}}$. Therefore, Theorem 2.3(c) gives $\int_S q_0(x,\infty)\nu_0(dx) = 0$ in each case, since $\int_S q_0(x,n)\nu_n(dx) = 0$ for all n (see Proposition 2.18 below). So the result follows by the definition of q_0. □

Standard proofs of Theorem 2.8 can also be given, but these rely on the formula (2.1) and are less useful for what is to follow in Theorem 2.20 below. Next, we enhance the portmanteau theorem by slightly extending the above proof of Theorem 2.8:

Theorem 2.9 (enhanced portmanteau theorem) *Let* (ν_n) *and* ν_0 *be in* $\mathcal{P}(S)$*. The following are equivalent:*

(a) $\nu_n \Rightarrow \nu_0$.

(e) $\liminf_n \int_S q\, d\nu_n \geq \int_S q\, d\nu_0$ *for every function* $q : S \to (-\infty, +\infty]$ *which is lower semicontinuous on* supp ν_0*, relative to the set* $\cup_{n=0}^{\infty}\text{supp } \nu_n$*, and which is bounded from below.*

Proof. (a) ⇒ (e): By Corollary 2.4, (a) implies $\nu_n \times \epsilon_n \Rightarrow \nu_0 \times \epsilon_\infty$ in $\mathcal{P}(S \times \hat{\mathbf{N}})$. Let q be as stated; similar to the proof Theorem 2.8, we define a function $q_1 : S \times \hat{\mathbf{N}} \to \{0, +\infty\}$ by

$$q_1(x,k) := \begin{cases} q(x) & \text{if } x \in \text{supp } \nu_k \text{ and } k < \infty \\ q(x) & \text{if } x \in \text{Ls}_n\text{supp } \nu_n \text{ and } k = \infty \\ +\infty & \text{otherwise} \end{cases}$$

Then it is not hard to see that the given lower semicontinuity property of q is *equivalent* to q_1 being lower semicontinuous on all of $S \times \hat{\mathbf{N}}$. An application of Theorem 2.3(d) gives $\liminf_n \int_S q_1(\cdot,n)d\nu_n \geq \int_S g_1(\cdot,\infty)d\nu_0$; by definition of q_1 the left side equals $\liminf_n \int_S q d\nu_n$ (use Definition 2.6), and the right side equals $\int_S q\, d\nu_0$ (use Theorem 2.8). Finally, the implication (e) ⇒ (a) follows immediately (for instance, observe that (e) is stronger than (d) of Theorem 2.3). □

Definition 2.10 (inf-compactness) A function $h : S \to (-\infty, +\infty]$ is *inf-compact* if

$$\{x \in S : h(x) \leq \beta\} \text{ is compact for every } \beta \in \mathbf{R}.$$

□

The following definitions of the classical notion of tightness are equivalent; of these, the first is well-known (e.g., cf. [Bo, Excercise 10, p. 109]) and the second is standard [Bi1, As]. Equivalence of these two definitions is also a consequence of the more general equivalence result in Proposition 4.3.

Definition 2.11 (tightness) A sequence (ν_n) in $\mathcal{P}(S)$ is said to be *tight* if there exists an inf-compact function $h : S \to [0, +\infty]$ such that

$$\sup_n \int_S h \, d\nu_n < +\infty.$$

□

Definition 2.12 (tightness) A sequence (ν_n) in $\mathcal{P}(S)$ is said to be *tight* if for every $\epsilon > 0$ there exists a compact set $K_\epsilon \subset S$ such that

$$\sup_n \nu_n(S \backslash K_\epsilon) \leq \epsilon.$$

□

The next result establishes the importance of tightness as a criterion for relative narrow compactness [Bi1, Theorem 6.1].

Theorem 2.13 (Prohorov's theorem) *Let (ν_n) in $\mathcal{P}(S)$ be tight. Then there exist a subsequence $(\nu_{n'})$ of (ν_n) and $\nu_* \in \mathcal{P}(S)$ such that $\nu_{n'} \Rightarrow \nu_*$.*

Frequently, the space S must have an additional property, which we shall introduce now:

Definition 2.14 (Suslin space) A Hausdorff topological space is said to be a *Suslin space* if it is the continuous image of a separable metric and complete space. □

Note that a Suslin space is evidently separable. Many useful spaces (e.g., Euclidean spaces, compact metric spaces, separable Banach spaces with strong or weak topology) are Suslin. For our study of narrow convergence of Young measures the following result will be of crucial importance.

Theorem 2.15 (necessity of tightness) *Suppose that S is Suslin. Let $\nu_n \Rightarrow \nu_0$ in $\mathcal{P}(S)$. Then (ν_n) is tight.*

Proof. Since S is Suslin, it follows by [DM, III.69] that each ν_n (viewed as a singleton) is tight. Hence, the result holds by [Bi1, Theorem 8, Appendix III]. □

Let τ be another – possibly nonmetrizable – topology on S. We shall write $S_\tau := (S, \tau)$ to indicate when S is equipped with τ; for extra safety, we shall often write S_ρ, etc., when the original ρ-topology is used on S (for instance, $\mathcal{B}(S_\tau)$ denotes the σ-algebra of τ-Borel sets, and $\mathcal{B}(S_\rho)$ denotes what used to be simply $\mathcal{B}(S)$). Correspondingly, we speak of τ-closedness, τ-compactness, τ-lower semicontinuity, etc., and these also have *sequential* counterparts: sequential τ-closedness, sequential τ-compactness, sequential τ-lower continuity, etc. The same distinction can be made for the Kuratowski limes superior: the definition of the *sequential* τ-limes superior is based on formula Definition 2.7 and runs as follows:

Definition 2.16 (Kuratowski sequential limes superior) The *sequential τ-limes superior* of a sequence of subsets (A_n) of S is defined as the set, denoted by $\tau\text{-Ls}_n A_n$, of all $x \in S$ for which there exists a subsequence $(A_{n'})$ of (A_n), and corresponding elements $a_{n'} \in A_{n'}$, such that $x = \tau\text{-lim}_{n'} a_{n'}$. □

In contrast, the definition of the *topological* τ-limes superior is based on (2.1):

Definition 2.17 (Kuratowski topological limes superior) The *τ-limes superior* of a sequence (A_n) of subsets of S is defined by

$$\tau\text{-LS}_n A_n := \cap_{p=1}^{\infty} \tau\text{-cl } \cup_{n \geq p} A_n.$$

□

Definition 2.6 of the support extends immediately into the definition of the *τ-support* of a probability measure. The following proposition will be of use; [5] in fact, we have already used it in the proof of Theorem 2.8 above, where S_ρ is Lindelöf.

Proposition 2.18 *Let $\nu \in \mathcal{P}(S)$ be carried by a set in $\mathcal{B}(S_\tau)$ with the Lindelöf property. Then ν is carried by the support set τ-supp ν.*

Proof. There exists $B \in \mathcal{B}(S_\tau)$, with the Lindelöf property, such that $\nu(B) = 1$. By Definition 2.6 (for S_τ), the complement of τ-supp ν is the union of all τ-open sets G with $\nu(G) = 0$. Therefore, $C := B \backslash \tau$-supp ν is the union of all sets $B \cap G$, where $G \subset S$ is τ-open and ν-null, so by the given Lindelöf property C is the union of a countable subcollection of such sets. Therefore, C is also ν-null, which implies $\nu(\tau\text{-supp } \nu) = \nu(B) = 1$. □

Corollary 2.19 *Let $\nu \in \mathcal{P}(S)$ be τ-tight. Then ν is carried by the support set τ-supp ν.*

Proof. Tightness implies that ν is carried by a countable union of τ-compact sets, i.e., by a set with the Lindelöf property. □

We set the stage for future developments by synthesizing Theorems 2.3, 2.8 and 2.13 into the following result:

Theorem 2.20 (synthesis) *Suppose that τ is at least as strong as the ordinary ρ-topology on S and that $\mathcal{B}(S_\tau) = \mathcal{B}(S_\rho)$. Let (ν_n) be a sequence in $\mathcal{P}(S)$ such that there exists a sequentially τ-inf-compact function $h : S \to [0, +\infty]$ for which*

$$\sup_n \int_S h \, d\nu_n < +\infty.$$

Then there exist a subsequence $(\nu_{n'})$ of (ν_n) and $\nu_* \in \mathcal{P}(S)$ such that

$$\liminf_{n'} \int_S q \, d\nu_{n'} \geq \int_S q \, d\nu_*$$

for every sequentially τ-lower semicontinuous function $q : S \to (-\infty, +\infty]$ which is bounded from below. Moreover,

$$\tau\text{-supp } \nu_* \subset \tau\text{-seq-cl } \tau\text{-Ls}_n \tau\text{-supp } \nu_n.$$

[5] I thank M. Valadier for pointing out the necessity of including such a result.

Proof. *A fortiori*, h is inf-compact on the metric space S_ρ. Hence, (ν_n) is tight in $\mathcal{P}(S_\rho)$. So by Theorem 2.13 there exist a subsequence $(\nu_{n'})$ of (ν_n) and $\nu_* \in \mathcal{P}(S)$ such that $\nu_{n'} \Rightarrow \nu_*$ in $\mathcal{P}(S_\rho)$. Now let q be as stated. Define $q_\epsilon : S \to (-\infty, +\infty]$ by $q_\epsilon(x,k) := q(x,k) + \epsilon h(x)$, $\epsilon > 0$. Evidently, q_ϵ is sequentially τ-inf-compact on S. This implies *a fortiori* that q_ϵ is inf-compact on the metrizable space S_ρ. Hence, q_ϵ is certainly lower semicontinuous on S_ρ. Theorem 2.3(d) can now be applied to q_ϵ, giving $\liminf_{n'} \int_S q \, d\nu_{n'} + \epsilon\sigma \geq \int_S q_\epsilon \, d\nu_*$, where $\sigma := \sup_n \int_S h \, d\nu_n$. By letting ϵ go to zero and using $q_\epsilon \geq q$ (by nonnegativity of h), we find the proclaimed inequality. Finally, for $S \times \hat{\mathbf{N}}$ instead of S, we can apply the above lower semicontinuity result to $q_0 : S \times \hat{\mathbf{N}} \to \{0, +\infty\}$, given by

$$q_0(x,k) := \begin{cases} 0 & \text{if } x \in \tau\text{-supp } \nu_k \text{ and } k < \infty \\ 0 & \text{if } x \in \tau\text{-seq-cl } \tau\text{-Ls}_n \tau\text{-supp } \nu_n \text{ and } k = \infty \\ +\infty & \text{otherwise} \end{cases}$$

which is sequentially lower semicontinuous on $S_\tau \times \hat{\mathbf{N}}$. For $(\nu_{n'})$ as above, Corollary 2.4 implies that $\nu_{n'} \times \epsilon_{n'} \Rightarrow \nu_* \times \epsilon_\infty$ in $\mathcal{P}(S_\rho \times \hat{\mathbf{N}})$. Therefore, the desired support property follows similar to the way in which Theorem 2.8 was derived from Theorem 2.3(d). □

Remark 2.21 If S_τ is a Suslin space, then the condition $\mathcal{B}(S_\tau) = \mathcal{B}(S_\rho)$ in Theorem 2.20 follows automatically from the fact that τ is at least as strong as the ρ-topology [Schw, Corollary 2, p. 101]. □

Remark 2.22 It is easy to see that when τ is a topology on S which is at least as strong as the ordinary ρ-topology, then τ-compactness implies sequential τ-compactness, using the fact that ρ-compactness and sequential ρ-compactness coincide, plus the fact that on any τ-compact the τ-topology coincides with the ρ-topology. Thus, if h in Theorem 2.20 is τ-inf-compact, it is already sequentially τ-inf-compact. □

Remark 2.23 (i) In Theorem 2.20 the following inclusion is obvious (cf. Theorem 2.8):

$$\tau\text{-seq-cl } \tau\text{-Ls}_n \tau\text{-supp } \nu_n \subset \tau\text{-LS}_n \tau\text{-supp } \nu_n \subset \rho\text{-Ls}_n \rho\text{-supp } \nu_n.$$

(ii) If in Theorem 2.20 there exists a τ-compact set containing all supports supp ν_n, $n \in \mathbf{N}$, then

$$\tau\text{-seq-cl } \tau\text{-Ls}_n \tau\text{-supp } \nu_n = \tau\text{-Ls}_n \tau\text{-supp } \nu_n$$

by Remark 2.22. □

3 Narrow and K-convergence of Young measures

Contents: Young measures, narrow convergence, semimetrizability, integrands, outer integral functionals, K-convergence, support theorem, Fatou-Vitali theorem, portmanteau theorem, tensor product continuity

Let $(\Omega, \mathcal{A}, \mu)$ be a finite measure space,[6] and let (S, ρ) be a separable metric space.

Definition 3.1 (Young measure) A *Young measure* from Ω into S is a function $\delta : \Omega \to \mathcal{P}(S)$ such that $\delta(\cdot)(B) : \omega \mapsto \delta(\omega)(B)$ is $\mathcal{A}$-measurable for every $B \in \mathcal{B}(S)$. The set of all Young measures from Ω into S is denoted by $\mathcal{R}(\Omega, \mathcal{A}, \mu; S)$ (or $\mathcal{R}$ for short). □

Thus, a Young measure can be regarded as a *measurably parametrized* probability measure. Other names for Young measures are, depending on the context: *transition probabilities, Markov kernels, randomized decision functions, relaxed control functions*, etc. For some elementary measure-theoretical properties of Young measures the reader is referred to [N, III.2] or [As, 2.6] (see also Appendix A). To see how the present section extends results from the theory treated in section 2, observe that when $\mathcal{A}$ is trivial ($\mathcal{A} = \{\emptyset, \Omega\}$), then $\mathcal{R}$ consists only of constant functions and can be identified with the set $\mathcal{P}(S)$ of probability measures on S.

Definition 3.2 (Dirac Young measure) A Young measure $\delta \in \mathcal{R}$ is *Dirac* if there exists a function $f : \Omega \to S$ such that

$$\delta(\omega) = \epsilon_{f(\omega)} := \text{Dirac measure at the point } f(\omega), \omega \in \Omega.$$

The function f is then a measurable function from Ω into S. In this special case δ is denoted by $\delta = \epsilon_f$. The set of all Dirac Young measures is denoted by $\mathcal{R}_D(\Omega, \mathcal{A}, \mu; S)$ (or $\mathcal{R}_D$) and the set of all measurable functions from $(\Omega, \mathcal{A})$ into $(S, \mathcal{B}(S))$ by $\mathcal{L}^0(\Omega, \mathcal{A}, \mu; S)$ (or $\mathcal{L}^0_S$). □

The fundamental idea behind Young measure theory is that, in some sense, $\mathcal{R}$ forms a *completion* of $\mathcal{L}^0_S$. Indeed, by Definition 3.2 the latter set can be identified with $\mathcal{R}_D$; it can thus be viewed as a subset of $\mathcal{R}$. To make this completion useful, the notion of narrow convergence of Young measures is very important.

Definition 3.3 (narrow convergence and topology) A sequence (δ_n) in $\mathcal{R}$ is said to converge *narrowly* to δ_0 in $\mathcal{R}$ (notation: $\delta_n \Longrightarrow \delta_0$) if for every $A \in \mathcal{A}$ and $c \in \mathcal{C}_b(S)$

$$\lim_n \int_A [\int_S c(x)\delta_n(\omega)(dx)]\mu(d\omega) = \int_A [\int_S c(x)\delta_0(\omega)(dx)]\mu(d\omega).$$

Correspondingly, the *narrow topology* on $\mathcal{R}$ is defined as the weakest topology on $\mathcal{R}$ for which all functionals $\delta \mapsto \int_A [\int_S c(x)\delta(\omega)(dx)]\mu(d\omega)$, $A \in \mathcal{A}$, $c \in \mathcal{C}_b(S)$, are continuous. □

Note the difference in notation between the narrow convergences of Definitions 2.1 and 3.3 (short and long arrow respectively). In the above form, the definition of narrow convergence is classical in statistical decision theory [Wal, L1]. It merges two completely different classical modes of convergence:

[6] Much of what is done here immediately extends to a σ-finite measure space $(\Omega, \mathcal{A}, \mu)$.

Remark 3.4 (equivalence) The following are obviously equivalent:

(a) $\delta_n \Longrightarrow \delta_0$ in $\mathcal{R}$.

(b) For every $A \in \mathcal{A}$ with $\mu(A) > 0$

$$[\mu \otimes \delta_n](A \times \cdot)/\mu(A) \Rightarrow [\mu \otimes \delta_0](A \times \cdot)/\mu(A) \text{ in } \mathcal{P}(S)$$

(see Theorem A.1 for the product measure notation).

(c) For every $c \in \mathcal{C}_b(S)$

$$\int_S c(x)\delta_n(\cdot)(dx) \overset{*}{\rightharpoonup} \int_S c(x)\delta_0(\cdot)(dx) \text{ in } \mathcal{L}^\infty_{\mathbf{R}},$$

where '$\overset{*}{\rightharpoonup}$' denotes convergence in the topology $\sigma(\mathcal{L}^\infty_{\mathbf{R}}, \mathcal{L}^1_{\mathbf{R}})$. □

The following example implies the narrow convergence result found in Example 1.1.

Example 3.5 (periodic functions) Let $(\Omega, \mathcal{A}, \mu)$ be $([0,1], \mathcal{L}([0,1]), \lambda_1)$ (cf. Example 1.1). Let $f_1 \in \mathcal{L}^1_{\mathbf{R}}$ be arbitrary; it can be extended periodically from $\Omega = [0,1]$ to all of $\mathbf{R}$. We define $f_{n+1}(\omega) := f_1(2^n\omega)$. Then $\epsilon_{f_n} \Longrightarrow \delta_0$, where $\delta_0 \in \mathcal{R}([0,1], \mathcal{L}([0,1]), \lambda_1)$ is given by

$$\delta_0(\omega) \equiv \lambda_1^{f_1}.$$

Here $\lambda^{f_1} \in \mathcal{P}(\mathbf{R})$ is the image of λ_1 under the mapping f_1, given by $\lambda^{f_1}(B) := \lambda(f_1^{-1}(B))$. To prove this, let $c \in \mathcal{C}_b(\mathbf{R})$ be arbitrary, and let A be first of the form $A = [0, \beta]$, $\beta > 0$. By $f_{n+1}(\omega) = f_1(2^n\omega)$ a simple change of variable gives

$$\int_A c(f_{n+1}(\omega))d\omega = \int_0^\beta c(f_1(2^n\omega))d\omega = 2^{-n}\int_0^{2^n\beta} c(f_1(\omega'))d\omega',$$

and by periodicity of f_1 the latter expression equals $\beta \int_0^1 c(f_1(\omega'))d\omega' = \lambda_1(A) \int_{\mathbf{R}} c(x)\lambda_1^{f_1}(dx)$ in the limit. So we have shown that

$$\lim_{n\to\infty} \int_A c(f_n(\omega))d\omega = \int_A [\int_{\mathbf{R}} c(x)\delta_0(\omega)(dx)]d\omega \tag{3.1}$$

for $A = [0, \beta]$. By subtraction, (3.1) continues to be valid for A's of the type $A = (\alpha, \beta]$, and, by addition, also for A's that are a finite disjoint union of such intervals. Finally, note that for any $A \in \mathcal{A}$ and any $\epsilon > 0$ there exists A', a finite union of intervals $(\alpha, \beta]$, such that the symmetric difference of A and A' has Lebesgue measure at most ϵ. Since then $|\int_A c(f_n) - \int_{A'} c(f_n)| \le \epsilon \sup_{x\in S} |c(x)|$, we conclude that (3.1) continues to hold for the general case. □

The narrow limit of a sequence of Young measures is essentially unique; this is a simple consequence of the next result.

Proposition 3.6 (essential separation) *For every δ, δ' in $\mathcal{R}$ the following are equivalent:*

(a) For every $A \in \mathcal{A}$ and $c \in \mathcal{C}_u(S)$

$$\int_A [\int_S c(x)\delta(\omega)(dx)]\mu(d\omega) = \int_A [\int_S c(x)\delta'(\omega)(dx)]\mu(d\omega).$$

(b) $\delta(\omega) = \delta'(\omega)$ for a.e. ω in Ω.

PROOF. $(a) \Rightarrow (b)$: By Proposition 2.2, $\mathcal{C}_u(S)$ has a countable subset $\mathcal{C}_0$ which separates the points of $\mathcal{P}(S)$. Now (a) implies that $\int_S c(x)\delta(\cdot)(dx) = \int_S c(x)\delta'(\omega)(\cdot)$ a.e. for every $c \in \mathcal{C}_0$, so (b) follows easily. The implication $(b) \Rightarrow (a)$ is trivial. □

The essentially sequential setup chosen for these lecture notes leads to the occasional use of a semimetric on $\mathcal{R}$, which is defined in the next result; when it is applicable, this allows us to identify corresponding sequential and purely topological notions for the narrow topology.

Theorem 3.7 (semimetrizability of narrow convergence) *If the σ-algebra $\mathcal{A}$ on Ω is countably generated, then the narrow topology on $\mathcal{R}$ is semimetrizable.*

PROOF. There exists an at most countable algebra which generates $\mathcal{A}$. Let (A_j) be an enumeration of this algebra. Let $\mathcal{C}_0 =: (c_i)$ be narrow convergence-determining on $\mathcal{P}(S)$, as in Proposition 2.2. Now define

$$d_{\mathcal{R}}(\delta, \delta') := \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} 2^{-i-j}(\mu(A_j) \sup_{x \in S} |c_i(x)|)^{-1} |I_{g_{i,j}}(\delta) - I_{g_{i,j}}(\delta')|,$$

where the integrand $g_{i,j}$ is given by $g_{i,j}(\omega, x) := 1_{A_j}(\omega)c_i(x)$. It is easy to see that $d_{\mathcal{R}}$ is a semimetric on $\mathcal{R}$. By [As, 1.3.11] it follows that for any sequence (δ_n) in $\mathcal{R}$, the convergence $d_{\mathcal{R}}(\delta_n, \delta_0) \to 0$ implies that

$$\lim_n \int_A [\int_S c_i(x)\delta_n(\omega)(dx)]\mu(d\omega) = \int_A [\int_S c_i(x)\delta_0(\omega)(dx)]\mu(d\omega).$$

for every $A \in \mathcal{A}$ and every $c_i \in \mathcal{C}_0$. By the given property of $\mathcal{C}_0$, this implies $\delta_n \Longrightarrow \delta_0$ (use Remark 3.4). The converse implication is immediate. □

The integrands $(\omega, x) \mapsto 1_A(\omega)c(x)$ in Definition 3.3 are too simple to be of use in optimal control theory, mathematical economics or applied analysis.

Definition 3.8 (integrands) (i) An *integrand* on $\Omega \times S$ is a function $g : \Omega \times S \to (-\infty, +\infty]$ such that for every $\omega \in \Omega$ the function $g(\omega, \cdot)$ on S is $\mathcal{B}(S)$-measurable.

(ii) A *lower semicontinuous [continuous] [[inf-compact]] integrand* on $\Omega \times S$ is an integrand g on $\Omega \times S$ such that $g(\omega, \cdot)$ is lower semicontinuous [continuous] [[inf-compact]] on S for every ω in Ω.

(iii) An integrand g on $\Omega \times S$ is said to be *integrably bounded below* if there exists $\phi \in \mathcal{L}^1_{\mathbf{R}}$ such that $g(\omega, x) \geq \phi(\omega)$ for all $\omega \in \Omega$ and $x \in S$. □

For *outer integrals* over $(\Omega, \mathcal{A}, \mu)$, i.e., integrals involving functions which may or may not be measurable, we refer to Appendix B. From now on also the integrals of measurable functions over Ω will follow the formula given in Lemma B.2; this coincides with a well-known convention in measure theory, which extends quasi-integration [CV, VII.7].

Definition 3.9 (outer integral functional) Let $g : \Omega \times S \to (-\infty, +\infty]$ be an integrand on $\Omega \times S$. For any $\delta \in \mathcal{R}$ the double integral

$$I_g(\delta) := \int_\Omega^* [\int_S g(\omega, x)\delta(\omega)(dx)]\mu(d\omega)$$

is well defined, provided that for each ω the inner integral of the function $g(\omega, \cdot)$ (which is $\mathcal{B}(S)$-measurable by Definition 3.8) over S is determined by the formula in Lemma B.2, of course now with $\delta(\omega)$ as the measure over which integration takes place. For $\delta = \epsilon_f \in \mathcal{R}_D$ this simplifies into the following formula:

$$J_g(f) := I_g(\epsilon_f) = \int_\Omega^* g(\omega, f(\omega))\mu(d\omega).$$

□

Usually an approximation procedure is followed to study narrow convergence for the various integrands mentioned in Definition 3.8 [BL, Ba3, Ba11]. In these lecture notes the scope of narrow convergence is extended via the following notion of convergence, introduced and studied in a much wider context in [Ba15, Ba16].

Definition 3.10 (K-convergence) A sequence (δ_n) in $\mathcal{R}$ is said to *K-converge* to δ_0 in $\mathcal{R}$ (notation: $\delta_n \xrightarrow{K} \delta_0$) if for every subsequence $(\delta_{n'})$ of (δ_n)

$$\frac{1}{N}\sum_{n'=1}^{N} \delta_{n'}(\omega) \Rightarrow \delta_0(\omega) \text{ as } N \to \infty \text{ for a.e. } \omega \text{ in } \Omega.$$

□

This convergence notion is *nontopological*, which explains why narrow convergence is of such importance, even though for *sequences* of Young measures K-convergence turns out to be a superior tool. Observe once more that the short arrow '$\Rightarrow$' in the defining pointwise limit statement above denotes narrow convergence in $\mathcal{P}(S)$ (cf. Definition 2.1). Note also that only the exceptional null set is allowed to vary with the subsequence $(\delta_{n'})$, in agreement with the observations following Theorem 1.2.

As an immediate consequence of Theorem 2.8 and the above definition we have the following result:

Theorem 3.11 (support of K-limits) *Let $\delta_n \xrightarrow{K} \delta_0$ in $\mathcal{R}$. Then*

$$\text{supp } \delta_0(\omega) \subset \text{Ls}_n \text{supp } \delta_n(\omega) \textit{ for a.e. } \omega \textit{ in } \Omega.$$

PROOF. Apply Theorem 2.8 pointwise. □

A quite useful property of K-convergence, which is very indicative of the significance of K-convergence for narrow convergence of Young measures, is as follows:

Lemma 3.12 (Fatou for K-convergence) *Let (δ_n) and δ_0 be in $\mathcal{R}$. Then $\delta_n \xrightarrow{K} \delta_0$ implies*

$$\liminf_n I_g(\delta_n) \geq I_g(\delta_0)$$

for every lower semicontinuous integrand g on $\Omega \times S$ which is integrably bounded from below.

PROOF. Set $\alpha := \liminf_n I_g(\delta_n)$; then there exists a subsequence $(\delta_{n'})$ such that $\alpha = \lim_{n'} I_g(\delta_{n'})$. By Definition 3.10, $\frac{1}{N}\sum_{n'=1}^{N} \delta_{n'}(\omega) \Rightarrow \delta_0(\omega)$ in $\mathcal{P}(S)$ for a.e. ω. Define $\psi_N(\omega) := \frac{1}{N}\sum_{n'=1}^{N} \int_S g(\omega, x)\delta_{n'}(\omega)(dx)$ and $\psi_0(\omega) := \int_S g(\omega, x)\delta_0(\omega)(dx)$; then $\liminf_N \psi_N(\omega) \geq \psi_0(\omega)$ for a.e. ω, since $g(\omega, \cdot)$ fulfills the conditions of Theorem 2.9(d). An application of the Fatou-result in Proposition B.4 gives $\liminf_N \int_\Omega^* \psi_N \, d\mu \geq \int_\Omega^* \psi_0 \, d\mu$. Here the right-hand side is equal to $I_g(\delta_0)$, and the left-hand side is at most α, by Lemma B.5 and our choice of $(\delta_{n'})$. □

By using Corollary 2.4 in a rather unexpected way, we obtain the following Fatou-Vitali strengthening of Lemma 3.12:

Theorem 3.13 (Fatou-Vitali for K-convergence) *Let (δ_n) and δ_0 be in $\mathcal{R}$. Then $\delta_n \stackrel{K}{\rightarrow} \delta_0$ implies*

$$\liminf_n I_g(\delta_n) \geq I_g(\delta_0)$$

for every integrand g on $\Omega \times S$ which is such that for every $\omega \in \Omega$

$$g(\omega, \cdot) \text{ is lower semicontinuous on } \mathrm{supp}\ \delta_0(\omega), \text{ relative to } \cup_{n=0}^{\infty} \mathrm{supp}\ \delta_n(\omega)$$

and such that there exists a uniformly integrable sequence (ϕ_n) in $\mathcal{L}^1_{\mathbf{R}}$ with

$$g(\omega, x) \geq \phi_n(\omega) \text{ for every } \omega \in \Omega \text{ and } x \in \mathrm{supp}\ \delta_n(\omega)$$

for every $n \in \mathbf{N}$.

Proof. Define $\delta_n \otimes \epsilon_n$ in $\mathcal{R}(\Omega, \mathcal{A}, \mu; S \times \hat{\mathbf{N}})$ by

$$(\delta_n \otimes \epsilon_n)(\omega) := \delta_n(\omega) \times \epsilon_n,$$

etc. (see also Definition 3.22 below). Then $\delta_n \otimes \epsilon_n \stackrel{K}{\rightarrow} \delta_0 \otimes \epsilon_\infty$ in $\mathcal{R}(\Omega, \mathcal{A}, \mu; S \times \hat{\mathbf{N}})$ by Definition 3.10 and Corollary 2.4. The integrand analogue of q_1, as used in the proof of Theorem 2.9, is $g_1 : \Omega \times S \times \hat{\mathbf{N}} \rightarrow (-\infty, +\infty]$, given by

$$g_1(\omega, x, k) := \begin{cases} g(\omega, x) & \text{if } x \in \mathrm{supp}\ \delta_k(\omega) \text{ and } k < \infty \\ g(\omega, x) & \text{if } x \in \mathrm{Ls}_n \mathrm{supp}\ \delta_n(\omega) \text{ and } k = \infty \\ +\infty & \text{otherwise} \end{cases}$$

Just like q_1 was lower semicontinuous, we have here that $g_1(\omega, \cdot)$ is lower semicontinuous on $S \times \hat{\mathbf{N}}$ for every ω; that is, g_1 is a lower semicontinuous integrand on $\Omega \times (S \times \hat{\mathbf{N}})$. By de la Vallée-Poussin's Theorem A.3, the uniform integrability of (ϕ_n) implies the existence of a function $h' : \mathbf{R}_+ \rightarrow \mathbf{R}_+$, convex, continuous and nondecreasing, with $\lim_{\xi \rightarrow \infty} h'(\xi)/\xi = +\infty$ and

$$\sigma' := \sup_n \int_\Omega h'(|\phi_n(\omega)|) \mu(d\omega) < +\infty.$$

For $\epsilon > 0$ define $g_{1,\epsilon} : \Omega \times S \times \hat{\mathbf{N}} \rightarrow (-\infty, +\infty]$ by

$$g_{1,\epsilon}(\omega, x, k) := \begin{cases} g_1(\omega, x, k) + \epsilon h'(|\phi_k(\omega)|) & \text{if } k < \infty \\ g_1(\omega, x, \infty) + \epsilon \liminf_n h'(|\phi_n(\omega)|) & \text{if } k = \infty \end{cases}$$

It is not hard to see that also $g_{1,\epsilon}$ is a lower semicontinuous integrand on $\Omega \times (S \times \hat{\mathbf{N}})$. Moreover, the given bound $g(\omega, \cdot) \geq \phi_n(\omega)$ on $\mathrm{supp}\ \delta_n(\omega)$ for each ω implies $g_{1,\epsilon}(\omega, \cdot, n) \geq \gamma_\epsilon$ on S for every $n < \infty$. Here $\gamma_\epsilon := \inf_{\xi \in \mathbf{R}} [\xi + \epsilon h'(|\xi|)]$ is a *finite* real number (by superlinear growth of h'). Thus, $g_{2,\epsilon} := \max(g_{1,\epsilon}, \gamma_\epsilon)$ is a lower semicontinuous integrand on $\Omega \times (S \times \hat{\mathbf{N}})$ which is bounded below and has $I_{g_{1,\epsilon}}(\delta_n \otimes \epsilon_n) = I_{g_{2,\epsilon}}(\delta_n \otimes \epsilon_n)$ for all $n < \infty$. Therefore, an application of Lemma 3.12 (with $S \times \hat{\mathbf{N}}$ instead of S) gives

$$\liminf_n I_{g_{1,\epsilon}}(\delta_n \otimes \epsilon_n) \geq I_{g_{2,\epsilon}}(\delta_0 \otimes \epsilon_\infty) \geq I_{g_1}(\delta_0 \otimes \epsilon_\infty),$$

where the last inequality follows by $g_{2,\epsilon} \geq g_{1,\epsilon} \geq g_1$. Also, $I_{g_{1,\epsilon}}(\delta_n \otimes \epsilon_n) \leq I_{g_1}(\delta_n \otimes \epsilon_n) + \epsilon\sigma'$ by the subadditivity of outer integration (Lemma B.5). By letting ϵ go to zero, it follows that $\liminf_n I_{g_1}(\delta_n \otimes \epsilon_n) \geq I_{g_1}(\delta_0 \otimes \epsilon_\infty)$. Quite similar to the final step in the proof of Theorem 2.9, this gives $\liminf_n I_g(\delta_n) \geq I_g(\delta_0)$, this time by invoking Theorem 3.11 instead of Theorem 2.8. □

Corollary 3.14 (K-convergence implies narrow convergence) *Let (δ_n) and δ_0 be in $\mathcal{R}$. Then $\delta_n \stackrel{K}{\to} \delta_0$ implies $\delta_n \Longrightarrow \delta_0$.*

PROOF. Choose any $A \in \mathcal{A}$ and $c \in \mathcal{C}_b(S)$; define $g(\omega, x) := 1_A(\omega)c(x)$. Then both g and $-g$ are lower semicontinuous integrands on $\Omega \times S$, bounded below by a constant. So a twofold application of Lemma 3.12 gives the the limit statement in Definition 3.3. □

Crucial for our further analysis of narrow convergence of Young measures is the following result (see Theorem A.1 for the notation). It is a direct consequence of the Prohorov-type Theorem 4.7, and will be proven in section 4. The reader need not fear circular arguments: the proofs of both Theorem 4.7 and Theorem 3.15 *only* involve Theorems 1.2, 2.13, 2.15 and some simple arguments.

Theorem 3.15 *Let (δ_n) be a sequence in $\mathcal{R}$ such that*

$$([\mu \otimes \delta_n](\Omega \times \cdot)/\mu(\Omega)) \text{ is tight in } \mathcal{P}(S).$$

Then there exist a subsequence $(\delta_{n'})$ of (δ_n) and $\delta_* \in \mathcal{R}$ such that $\delta_{n'} \stackrel{K}{\to} \delta_*$.

The next result explains the precise relationship between narrow and K-convergence in $\mathcal{R}$ when the metric space S is Suslin.

Corollary 3.16 (narrow convergence in terms of K-convergence) *Suppose that S is Suslin. Let (δ_n) and δ_0 be in $\mathcal{R}$. The following are equivalent:*

(a) $\delta_n \Longrightarrow \delta_0$.

(b) Every subsequence $(\delta_{n'})$ of (δ_n) has a further subsequence $(\delta_{n''})$ such that $\delta_{n''} \stackrel{K}{\to} \delta_0$.

PROOF. *(a)* ⇒ *(b)*: We have $[\mu \otimes \delta_n](\Omega \times \cdot)/\mu(\Omega) \Rightarrow [\mu \otimes \delta_0](\Omega \times \cdot)/\mu(\Omega)$ by Remark 3.4. By Theorem 2.15 this implies that $([\mu \otimes \delta_n](\Omega \times \cdot)/\mu(\Omega))$ is tight. Thus, by Theorem 3.15 for every subsequence $(\delta_{n'})$ of (δ_n) there exist a further subsequence $(\delta_{n''})$ and $\delta_* \in \mathcal{R}$ such that $\delta_{n''} \stackrel{K}{\to} \delta_*$. By Corollary 3.14, this implies $\delta_{n''} \Longrightarrow \delta_*$. By Proposition 3.6 , it follows that $\delta_*(\omega) = \delta_0(\omega)$ for a.e. ω. So we conclude that $\delta_{n''} \stackrel{K}{\to} \delta_0$.

(b) ⇒ *(a)*: Suppose that *(a)* were not true. This implies that there exist $A \in \mathcal{A}$ and $c \in \mathcal{C}_b(S)$ such that for $g(\omega, x) := 1_A(\omega)c(x)$ one would have $\alpha := \liminf_n I_g(\delta_n) < I_g(\delta_0)$. Then there would be a subsequence $(\delta_{n'})$ of (δ_n) such that $I_g(\delta_{n'}) \to \alpha$. By hypothesis, there would be a further subsequence $(\delta_{n''})$ with $\delta_{n''} \stackrel{K}{\to} \delta_0$. Applied to this subsequence, Lemma 3.12 would give $\alpha \geq I_g(\delta_0)$, a contradiction. □

Even though narrow convergence implies K-convergence for *subsequences*, as made precise in the above corollary, a narrowly convergent sequence does not have to K-converge as a whole:

Example 3.17 (narrow convergence does not imply K-convergence) Consider the sequence (f_n) of Rademacher functions from Example 1.1. Define the following sequence (f'_n) in $\mathcal{L}^1_{\mathbf{R}}$: for each $m \in \mathbf{N}$ define $f'_n := f_m$ for $2^{m-1} \leq n \leq 2^m - 1$. From Examples 1.1 and 3.5 it is clear that $\epsilon_{f'_n} \Longrightarrow \delta_0$. It is easy to check the following: for $N = 2^m - 1$

$$\frac{1}{N}\sum_{n=1}^{N} \epsilon_{f'_n}(\omega) = \frac{1}{2^m - 1}\sum_{n=1}^{2^{m-2}-1} \epsilon_{f'_n}(\omega) + \frac{2^{m-2}}{2^m - 1}\epsilon_{f_{m-1}}(\omega) + \frac{2^{m-1}}{2^m - 1}\epsilon_{f_m}(\omega).$$

By Corollary 3.14 we know that if $(\epsilon_{f'_n})$ were to K-converge to some Young measure, it would have to be $\delta_0 \equiv \frac{1}{2}[\epsilon_1 + \epsilon_{-1}]$ a.e. But by the above expression this is not possible, since $2^{m-i}/(2^m - 1) \to 2^{-i}$ for $i = 1, 2$, and $\lambda_1(\{\omega \in \Omega : f_m(\omega) = f_{m-1}(\omega)\}) = 4^{-1}$ for all $m \in \mathbf{N}$. □

Theorem 3.11 and Corollary 3.16 immediately give the following result:

Corollary 3.18 (support of narrow limits) *Suppose that S is Suslin. Let $\delta_n \Longrightarrow \delta_0$ in $\mathcal{R}$. Then*

$$\text{supp } \delta_0(\omega) \subset \text{Ls}_n \text{supp } \delta_n(\omega) \textit{ for a.e. } \omega \textit{ in } \Omega.$$

Directly from Corollary 3.16 and the Fatou-Vitali-type Theorem 3.13 comes the following extension of the portmanteau Theorems 2.3 [Ba11, Ba17].

Theorem 3.19 (portmanteau theorem) *Suppose that S is Suslin. Let (δ_n) and δ_0 be in $\mathcal{R}$. The following are equivalent:*

(a) $\delta_n \Longrightarrow \delta_0$.

(b) $\lim_n \int_A[\int_S c(x)\delta_n(\omega)(dx)]\mu(d\omega) = \int_A[\int_S c(x)\delta_0(\omega)(dx)]\mu(d\omega)$ *for every* $A \in \mathcal{A}$, $c \in \mathcal{C}_u(S)$.

(c) $\liminf_n I_g(\delta_n) \geq I_g(\delta_0)$ *for every lower semicontinuous integrand g on $\Omega \times S$ which is integrably bounded from below.*

(d) $\liminf_n I_g(\delta_n) \geq I_g(\delta_0)$ *for every integrand g on $\Omega \times S$ such that for every $\omega \in \Omega$*

$$g(\omega, \cdot) \textit{ is lower semicontinuous on } \text{supp } \delta_0(\omega), \textit{ relative to } \cup_{n=0}^{\infty} \text{supp } \delta_n(\omega),$$

and such that there exists a uniformly integrable sequence (ϕ_n) in $\mathcal{L}^1_{\mathbf{R}}$ with

$$g(\omega, x) \geq \phi_n(\omega) \textit{ for every } \omega \in \Omega \textit{ and } x \in \text{supp } \delta_n(\omega)$$

for every $n \in \mathbf{N}$.

Proof. $(a) \Leftrightarrow (b)$: The equivalence follows immediately from the equivalence of (a) and (b) in Theorem 2.3 and Remark 3.4.

$(d) \Rightarrow (c) \Rightarrow (b)$: *A fortiori.* Note that (b) follows by applying (c) to both $g(\omega, x) := 1_A(\omega)c(x)$ and $g'(\omega, x) := -1_A(\omega)c(x)$, with A and c as in Definition 3.3.

$(a) \Rightarrow (d)$: Immediate by Corollary 3.16 and Theorem 3.13. □

Observe that $(a) \Rightarrow (d)$, which is by far the most useful implication of the above theorem, constitutes a theorem of Fatou-Vitali type for narrow convergence. We can now easily examine narrow convergence when it is restricted to the set $\mathcal{R}_D$ of Dirac Young measures:

Definition 3.20 (convergence in measure) A sequence (f_n) in $\mathcal{L}^0_S$ is said to *converge in measure* to $f_0 \in \mathcal{L}^0_S$ (notation: $f_n \overset{\mu}{\to} f_0$) if for every $\epsilon > 0$

$$\lim_n \mu(\{\omega \in \Omega : \rho(f_n(\omega), f_0(\omega)) > \epsilon\}) = 0.$$

□

Proposition 3.21 (narrow convergence in $\mathcal{R}_D$) *Suppose that S is Suslin. Let (f_n) and f_0 be in $\mathcal{L}^0_S$. Then the following are equivalent:*

(a) $\epsilon_{f_n} \Longrightarrow \epsilon_{f_0}$ *in* $\mathcal{R}$.

(b) $f_n \overset{\mu}{\to} f_0$ *in* $\mathcal{L}^0_S$.

PROOF. $(a) \Rightarrow (b)$: Let $\epsilon > 0$ be arbitrary. Define a lower semicontinuous integrand on $\Omega \times S$ by

$$g(\omega, x) := \begin{cases} -1 & \text{if } \rho(x, f_0(\omega)) \geq \epsilon \\ 0 & \text{otherwise} \end{cases}$$

By Theorem 3.19(c), $\liminf_n J_g(f_n) \geq J_g(f_0) = 0$; i.e., $\limsup_n \mu(\{\rho(f_n, f_0) \geq \epsilon\}) = 0$.

$(b) \Rightarrow (a)$: We invoke Theorem 3.19(b). Let $A \in \mathcal{A}$ and $c \in \mathcal{C}_u(S)$ be arbitrary. For each $\gamma > 0$ there exists $\epsilon > 0$ such that $|c(x) - c(x')| \leq \gamma$ whenever $\rho(x, x') < \epsilon$. Then

$$\int_A |c(f_n) - c(f_0)| \leq \gamma\mu(\Omega) + 2\mu(\{\omega \in \Omega : \rho(f_n(\omega), f_0(\omega)) \geq \epsilon\}) \sup_{x \in S} |c(x)|,$$

and $\epsilon_{f_n} \Longrightarrow \epsilon_{f_0}$ follows. □

Definition 3.22 (tensor product) Let $(\Omega', \mathcal{A}', \mu')$ be another finite measure space and S' another separable metric space. Then $\mathcal{A} \times \mathcal{A}'$ is the product σ-algebra on $\Omega \times \Omega'$ and $\mu \times \mu'$ the product measure. Also, $S \times S'$ is a separable metric space. The *tensor product* $\delta \otimes \delta'$ of $\delta \in \mathcal{R}(\Omega, \mathcal{A}, \mu; S)$ and $\delta' \in \mathcal{R}(\Omega', \mathcal{A}', \mu'; S')$ is defined by

$$(\delta \otimes \delta')(\omega, \omega') := \delta(\omega) \times \delta'(\omega');$$

that is, $(\delta \otimes \delta')(\omega, \omega')$ is the product of the probability measures $\delta(\omega)$ and $\delta'(\omega')$. Clearly, $\delta \otimes \delta'$, thus defined, belongs to $\mathcal{R}(\Omega \times \Omega', \mathcal{A} \times \mathcal{A}', \mu \times \mu'; S \times S')$. □

In a more rudimentary form, namely, for a trivial measure space $(\Omega', \mathcal{A}', \mu')$ (for instance, with Ω' a singleton), the narrow convergence behavior of the tensor product used above was already introduced in the proof of Theorem 3.13.

Theorem 3.23 (tensor product narrow convergence) *Let $(\Omega', \mathcal{A}', \mu')$ be another finite measure space and S' another separable metric space. Let $\delta_n \Longrightarrow \delta_0$ in $\mathcal{R}(\Omega, \mathcal{A}, \mu; S)$ and $\delta'_n \Longrightarrow \delta'_0$ in $\mathcal{R}(\Omega', \mathcal{A}', \mu'; S')$. Then $\delta_n \otimes \delta'_n \Longrightarrow \delta_0 \otimes \delta'_0$ in $\mathcal{R}(\Omega \times \Omega', \mathcal{A} \times \mathcal{A}', \mu \times \mu'; S \times S')$.*

Lemma 3.24 *For every $\tilde{A} \in \mathcal{A} \times \mathcal{A}'$ and every ϵ there exist finitely many disjoint measurable rectangles $A_i \times A'_i$ in $\mathcal{A} \times \mathcal{A}'$, $i = 1, \cdots, m$, such that the symmetric difference of $\tilde{A}$ and $\cup_{i=1}^m A_i \times A'_i$ has $\mu \times \mu'$-measure at most ϵ.*

PROOF. The algebra consisting of finite disjoint unions of measurable rectangles generates $\mathcal{A} \times \mathcal{A}'$; hence, the result follows by [As, 1.3.11]. □

PROOF of Theorem 3.23. Let $\tilde{A} \in \mathcal{A} \times \mathcal{A}'$ and $c \in \mathcal{C}(S \times S')$, and set $g(\omega, \omega', x, x') := 1_{\tilde{A}}(\omega, \omega')c(x, x')$. Since uniform limits of continuous functions are continuous, the result obtained in Lemma 3.24 enables to just consider the case $\tilde{A} = A \times A'$, with $A \in \mathcal{A}$ and $A' \in \mathcal{A}'$. We may suppose $\mu(A) > 0$, $\mu'(A') > 0$. Then $I_g(\delta_n \otimes \delta'_n) = \mu(A)\mu'(A') \int_{S \times S'} c d(\nu_n \times \nu'_n)$, where $\nu_n := [\mu \otimes \delta_n](A \times \cdot)/\mu(A)$ and $\nu'_n := [\mu \otimes \delta'_n](A' \times \cdot)/\mu'(A')$ satisfy $\nu_n \Rightarrow \nu_0$ and $\nu'_n \Rightarrow \nu'_0$, in view of Remark 3.4. By Theorem 2.5 this gives $I_g(\delta_n \otimes \delta'_n) \to I_g(\delta_0 \otimes \delta'_0)$. □

As shown by the following counterexample, Theorem 3.23 need not hold when the measure on $(\Omega \times \Omega', \mathcal{A} \times \mathcal{A}')$ is no longer a product measure, even when μ and μ' are its marginals.

Example 3.25 Take for $(\Omega, \mathcal{A})$ and $(\Omega', \mathcal{A}')$ the space $([0,1], \mathcal{B}([0,1]))$. Let (f_n) be the sequence of Rademacher functions on Ω and let (f'_n) be the sequence of Rademacher functions on Ω' (see Example 1.1). Equip $\tilde{\Omega} := [0,1]^2$ with $\tilde{\mathcal{A}} := \mathcal{B}([0,1]^2)$ and with $\tilde{\mu}$, the uniform measure concentrated on the diagonal of $[0,1]^2$. By Example 1.1, we have $\epsilon_{f_n} \Longrightarrow \delta_0$ in $\mathcal{R}(\Omega, \mathcal{A}, \mu; \mathbf{R})$ and $\epsilon_{f'_n} \Longrightarrow \delta_0$ in $\mathcal{R}(\Omega', \mathcal{A}', \mu'; \mathbf{R})$, but we do not have $\epsilon_{f_n} \otimes \epsilon_{f'_n} \Longrightarrow \delta_0 \otimes \delta_0$ in $\mathcal{R}(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{\mu}; \mathbf{R}^2)$. To see the latter, apply Definition 3.3 with $A := \tilde{\Omega}$ and $c(x, x') := xx'$; then in Definition 3.3 the limit on the left equals 1, but the expression on the right is equal to 0. $\square$

4 Relative compactness of Young measures

Contents: parametrized tightness, Prohorov's theorem for Young measures, existence of Bayesian Nash, Cournot-Nash and Stackelberg equilibria

This section introduces two equivalent notions of parametrized tightness for Young measures. Tightness is shown to be a criterion for the relative compactness of sequences of Young measures vis à K-convergence (and *a fortiori* sequential Young measure convergence). As in section 3, $(\Omega, \mathcal{A}, \mu)$ is a finite measure space and (S, ρ) a separable metric space.

Definition 4.1 (parametrized tightness) A sequence (δ_n) of Young measures in $\mathcal{R}$ is said to be *tight* if there exists a nonnegative inf-compact integrand h on $\Omega \times S$ (cf. Definition 3.8) such that

$$\sup_n I_h(\delta_n) < +\infty.$$

□

This definition, given in [Ba3], extends earlier notions of tightness [L1, BL, JM]. In [Jaw] it was shown to be equivalent to the following definition:

Definition 4.2 (parametrized tightness) A sequence (δ_n) of Young measures in $\mathcal{R}$ is said to be *tight* if for every $\epsilon > 0$ there exists a multifunction $\Gamma_\epsilon : \Omega \to 2^S$, with $\Gamma_\epsilon(\omega)$ compact for every $\omega \in \Omega$, such that

$$\sup_n \int_\Omega^* \delta_n(\omega)(S \backslash \Gamma_\epsilon(\omega))\mu(d\omega) \leq \epsilon.$$

□

Proposition 4.3 (equivalence) *Definitions 4.1 and 4.2 are equivalent.*

Proof. First, let (δ_n) be tight according to Definition 4.1. I.e., there exists an inf-compact integrand $h \geq 0$ such that $0 \leq \sigma := \sup_n I_h(\delta_n) < +\infty$. For each $\epsilon > 0$, let $\Gamma_\epsilon(\omega)$ be the set of all $x \in S$ for which $h(\omega, x) \leq \sigma/\epsilon$; then $\Gamma_\epsilon(\omega)$ is compact for every ω. Also, for every n

$$\frac{\sigma}{\epsilon} \int_\Omega^* \delta_n(\omega)(S \backslash \Gamma_\epsilon(\omega))\mu(d\omega) \leq I_h(\delta_n) \leq \sigma,$$

and this proves that Definition 4.2 holds.

Conversely, let Γ_m be the multifunction corresponding to $\epsilon = 3^{-m}$, $m \in \mathbf{N}$. With no loss of generality we may suppose that $\Gamma_m(\omega) \subset \Gamma_{m+1}(\omega)$ for every ω and m (otherwise, we could always take finite unions of the Γ_m). Now set $\Gamma_0 \equiv \emptyset$ and define

$$h(\omega, x) := \begin{cases} 2^m & \text{if } x \in \Gamma_m(\omega) \backslash \Gamma_{m-1}(\omega),\ m \in \mathbf{N} \\ +\infty & \text{if } x \notin \cup_m \Gamma_m(\omega) \end{cases}$$

Then $h(\omega, \cdot)$ is inf-compact on S for every ω and $\sup_n I_h(\delta_n) \leq 6$. □

Definitions 4.1 and 4.2 clearly extend Definitions 2.11 and 2.12 respectively. We illustrate the usefulness of each of these definitions:

Example 4.4 (parametrized tightness) Suppose that $(f_n) \subset \mathcal{L}^1_{\mathbf{R}^d}$ is uniformly bounded in L^1-norm:

$$\sup_{n \in \mathbf{N}} \int_\Omega |f_n(\omega)| \mu(d\omega) < +\infty.$$

Then the corresponding sequence (ϵ_{f_n}) in $\mathcal{R}_D$ is tight: simply use $h(\omega, x) := |x|$ in Definition 4.1. □

Example 4.5 (parametrized tightness) Let (f_n) and f_0 in $\mathcal{L}^0_{\mathbf{R}^d}$ be such that $f_n \stackrel{\mu}{\to} f_0$. Then the corresponding sequence (ϵ_{f_n}) in $\mathcal{R}_D$ is tight. Indeed, for any $\epsilon > 0$ there exists n_0 such that $\mu(\{|f_n - f_0| \geq 1\}) \leq \epsilon$ for all $n \geq n_0$. Now choose $\alpha \geq 1$ sufficiently large to have $\max_{1 \leq n \leq n_0} \mu(\{|f_n - f_0| \geq \alpha\}) \leq \epsilon$. Then it is clear that $\sup_n \mu(\{|f_n - f_0| \geq \alpha\}) \leq \epsilon$. By Definition 4.2 the desired tightness holds for $\Gamma_\epsilon(\omega) := \{x \in \mathbf{R}^d : |x - f_0(\omega)| \leq \alpha\}$. □

Remark 4.6 Let (δ_n) be in $\mathcal{R}$. If the separable metric space S is in addition either complete or locally compact, then the following are equivalent:

(*a*) $([\mu \otimes \delta_n](\Omega \times \cdot)/\mu(\Omega))$ is tight in the sense of Definitions 2.11, 2.12.

(*b*) (δ_n) is tight in the sense of Definitions 4.1, 4.2.

Of course, (*a*) ⇒ (*b*) is *a fortiori*. Also, (*b*) ⇒ (*a*) holds by observing that Theorem 4.7 below implies that $([\mu \otimes \delta_n](\Omega \times \cdot)/\mu(\Omega))$ is sequentially relatively compact. Therefore, under the additional hypothesis for S, it follows that $([\mu \otimes \delta_n](\Omega \times \cdot)/\mu(\Omega))$ is tight in $\mathcal{P}(S)$ by [Bi1, Theorem 6.2] or [Bo, 5.5] ('converse Prohorov theorem'). Even when S has one of these additional properties it is still useful to have the broad parametrized tightness criterion at our disposal, because direct verification of tightness of the form (a) is not always easy in actual examples. □

Theorem 4.7 (Prohorov's theorem for Young measures) *Let (δ_n) be a tight sequence in $\mathcal{R}$. Then there exist a subsequence $(\delta_{n'})$ of (δ_n) and $\delta_* \in \mathcal{R}$ such that $\delta_{n'} \stackrel{K}{\to} \delta_*$. In particular, this implies $\delta_{n'} \Longrightarrow \delta_*$.*

PROOF. Let $\mathcal{C}_0 := (c_i)$ be a countable subset of $\mathcal{C}_u(S)$ which determines narrow convergence on $\mathcal{P}(S)$, as guaranteed by Proposition 2.2. For every $i \in \mathbf{N}$ we have

$$\sup_n \int_\Omega |\phi_{i,n}| d\mu \leq \mu(\Omega) \sup_{x \in S} |c_i(x)| < +\infty,$$

where $\phi_{i,n}(\omega) := \int_S c_i(x) \delta_n(\omega)(dx)$. Let h be as in Definition 4.1. By Lemma B.3 there exists for each $n \in \mathbf{N}$ a function $\phi_{0,n} \in \mathcal{L}^1_{\mathbf{R}}$ such that $\phi_{0,n}(\omega) \geq \int_S h(\omega, x) \delta_n(\omega)(dx)$ for all $\omega \in \Omega$ and with $\int_\Omega \phi_{0,n} \, d\mu = I_h(\delta_n)$. Applying the Komlós Theorem 1.2 in a diagonal procedure, we obtain a subsequence $(\delta_{n'})$ of (δ_n) and functions $\phi_{i,*} \in \mathcal{L}^1_{\mathbf{R}}$, $i \in \mathbf{N} \cup \{0\}$, such that $\lim_N \frac{1}{N} \sum_{n''=1}^N \phi_{i,n''} = \phi_{i,*}$ a.e. for every further subsequence $(\delta_{n''})$ and for all $i \in \mathbf{N} \cup \{0\}$. It follows therefore that for every such subsequence $(\delta_{n''})$ for a.e. ω in Ω

$$\lim_N \int_S h(\omega, x) \frac{1}{N} \sum_{n''=1}^N \delta_{n''}(\omega)(dx) = \phi_{0,*}(\omega) < +\infty, \tag{4.1}$$

$$\lim_N \int_S c_i(x) \frac{1}{N} \sum_{n''=1}^N \delta_{n''}(\omega)(dx) = \phi_{i,*}(\omega) \text{ for all } i \in \mathbf{N}. \tag{4.2}$$

Let us first consider $(\delta_{n'})$ itself as the subsequence in question. Fix ω outside the exceptional null set M involved in (4.1)–(4.2). Then (4.1) implies that the sequence (ν_N) in $\mathcal{P}(S)$, defined by $\nu_N := \frac{1}{N}\sum_{n'=1}^{N}\delta_{n'}(\omega)$, is tight in the classical sense. By Theorem 2.13, there exists at least one convergent subsequence of (ν_N) and a corresponding narrow limit ν_* in $\mathcal{P}(S)$ (these may still depend upon ω at this stage). Then (4.2) implies

$$\int_S c_i(x)\nu_*(dx) = \phi_{i,*}(\omega) \text{ for all } i \in \mathbf{N}.$$

But (4.2) causes the same identity to hold for any other narrow limit point of (ν_N). Since this identity uniquely determines ν_* (for (c_i) separates the measures in $\mathcal{P}(S)$), $\nu_N \Rightarrow \nu_*$ follows for the *entire* sequence. Define $\delta_*(\omega) := \nu_*$. We conclude that $\sum_{n'=1}^{N}\delta_{n'}(\omega) \Rightarrow \delta_*(\omega)$ on $\Omega\backslash M$. Therefore, if we define δ_* to be equal to an arbitrary, but fixed element from $\mathcal{P}(S)$ on M, it follows from Proposition A.2 that δ_* belongs to $\mathcal{R}$. The whole argument following (4.2) can also be repeated if one starts out with an arbitrary subsequence $(\delta_{n''})$ of $(\delta_{n'})$, instead of $(\delta_{n'})$ itself. Except for the change in the exceptional null set M, for which the definition of K-convergence allows, nothing changes. □

The deferred proof of Theorem 3.15, which is an immediate consequence of Theorem 4.7 and played a special role in section 3, can now be given.

Proof of Theorem 3.15. As observed before, tightness of $([\mu \otimes \delta_n](\Omega \times \cdot)/\mu(\Omega))$, in the sense of Definitions 2.11–2.12 implies tightness of (δ_n) in the sense of Definitions 4.1–4.2. So Theorem 4.7 gives the desired result immediately. □

Example 4.8 Let $(f_n) \subset \mathcal{L}^1_{\mathbf{R}^d}$ be such that $\sup_n \int_\Omega |f_n| d\mu < +\infty$. Then there exist a subsequence $(f_{n'})$ of (f_n) and δ_* in $\mathcal{R}(\Omega, \mathcal{A}, \mu; \mathbf{R}^d)$ such that $\epsilon_{f_{n'}} \xrightarrow{K} \delta_*$ and *a fortiori* $\epsilon_{f_{n'}} \Longrightarrow \delta_*$. Moreover, for a.e. ω the barycenter $f_*(\omega) := \text{bar } \delta_*(\omega)$ exists; this is defined by

$$\text{bar } \delta_*(\omega) := \int_{\mathbf{R}^d} x\, \delta_*(\omega)(dx)$$

and satisfies

$$f_*(\omega) \in \text{co } \text{Ls}_n f_n(\omega).$$

The *parametrized barycenter* f_* of δ_*, thus defined, has a modification belonging to $\mathcal{L}^1_{\mathbf{R}^d}$. To derive these facts, observe that (ϵ_{f_n}) is tight in $\mathcal{R}(\Omega, \mathcal{A}, \mu; \mathbf{R}^d)$ by uniform L^1-boundedness of (f_n) (set $h(\omega, x) := |x|$ in Definition 4.1). Hence, by Theorem 4.7 there exist a subsequence $(f_{n'})$ of (f_n) and δ_* in $\mathcal{R}$ such that $\epsilon_{f_{n'}} \xrightarrow{K} \delta_*$. By Theorem 3.11, this gives $\delta_*(\omega)(\text{Ls}_n f_n(\omega)) = 1$ for a.e. ω. Also, applying Lemma 3.12 to $g(\omega, x) := |x|$ gives

$$\int_\Omega [\int_{\mathbf{R}^d} |x| \delta_*(\omega)(dx)]\mu(d\omega) = I_g(\delta_*) \leq \liminf_{n'} \int_\Omega |f_{n'}| d\mu < +\infty.$$

From this it follows that the inner integral on the left is finite for a.e. ω, and this implies existence of the barycenter $f_*(\omega)$ of $\delta_*(\omega)$; by Theorem A.13 and the support property of $\delta_*(\omega)$, it follows that $f_*(\omega) \in \text{co } \text{Ls}_n f_n(\omega)$. The above inequality also gives $\int_\Omega |f_*| d\mu \leq I_g(\delta_*) < +\infty$ by Jensen's inequality, so f_* belongs to $\mathcal{L}^1_{\mathbf{R}^d}$. □

Example 4.9 Let $(f_n) \subset \mathcal{L}^1_{\mathbf{R}^d}$ be such that (f_n) is uniformly integrable. Then the subsequence $(f_{n'})$ of (f_n) and $\delta_* \in \mathcal{R}(\Omega, \mathcal{A}, \mu; \mathbf{R}^d)$ of Example 4.8 are such that $f_{n'} \rightharpoonup f_*$, where, as in the previous example, $f_* \in \mathcal{L}^1_{\mathbf{R}^d}$ is the parametrized barycenter of δ_*. Here ' $\rightharpoonup$ ' denotes convergence in the weak topology $\sigma(\mathcal{L}^1_{\mathbf{R}^d}, \mathcal{L}^\infty_{\mathbf{R}^d})$, i.e., it is claimed that $\lim_{n'} \int_\Omega f_{n'} \cdot b\, d\mu = \int_\Omega f_* \cdot b\, d\mu$ for every $b \in \mathcal{L}^\infty_{\mathbf{R}^d}$. To see this, define a continuous integrand

g_b on $\Omega \times \mathbf{R}^d$ by $g_b(\omega, x) :=< x, b(\omega) >$. Then $g_b(\omega, f_{n'}(\omega)) \geq \phi_{n'}(\omega) := -\|b\|_\infty |f_{n'}(\omega)|$, with $(\phi_{n'})$ obviously forming a uniformly integrable sequence. Hence, an application of Theorem 3.13 gives

$$\liminf_{n'} \int_\Omega f_{n'} \cdot b \, d\mu = \liminf_{n'} J_{g_b}(f_{n'}) \geq I_{g_b}(\delta_*) = \int_\Omega f_* \cdot b \, d\mu,$$

and the same inequality also holds for $-b$ instead of b. □

Combining Examples 4.8 and 4.9 gives immediately the following result, due to Z. Artstein [Ar2, Proposition C] (the proof given in [Ar2] is totally different).

Proposition 4.10 (Artstein) *Let (f_n) converge weakly (i.e., in $\sigma(\mathcal{L}^1_{\mathbf{R}^d}, \mathcal{L}^\infty_{\mathbf{R}^d})$) to f_0 in $\mathcal{L}^1_{\mathbf{R}^d}$. Then*

$$f_0(\omega) \in \text{co Ls}_n f_n(\omega) \text{ for a.e. } \omega \text{ in } \Omega.$$

Example 4.11 Let the σ-algebra $\mathcal{A}$ on Ω be countably generated and let $\Gamma : \Omega \to 2^S$ be such that $\Gamma(\omega)$ is compact for every $\omega \in \Omega$. Let $\mathcal{R}_\Gamma$ be the set consisting of all $\delta \in \mathcal{R}(\Omega, \mathcal{A}, \mu; S)$ such that $\delta(\omega)(\Gamma(\omega)) = 1$ for a.e. ω. Then $\mathcal{R}_\Gamma$ is compact for the narrow topology. Indeed, $\mathcal{R}_\Gamma$ meets Definition 4.2 immediately, so Theorem 4.7 gives that $\mathcal{R}_\Gamma$ is relatively sequentially compact, both for K- and for narrow convergence. Moreover, $\mathcal{R}_\Gamma$ is obviously closed for K-convergence; hence, it is also closed for narrow convergence. By Theorem 3.7, this implies that $\mathcal{R}_\Gamma$ is compact for the narrow topology. □

The following continuation of Example 4.11 additionally considers tensor product continuity [Ba11] (see [BR] for an extension):

Example 4.12 (existence of Bayesian Nash equilibria) Let $\Omega := \Pi_{i=1}^m \Omega_i$ and $\mathcal{A} := \times_{i=1}^m \mathcal{A}_i$, where the measurable space $(\Omega_i, \mathcal{A}_i)$ is the *type space* [Ha] (a kind of private observation space) of player i, $i = 1, \cdots, m$. Let ν on $(\Omega, \mathcal{A})$ be a probability measure which is absolutely continuous with respect to the product measure $\mu := \nu_1 \times \cdots \times \nu_m$, where ν_i is the marginal probability induced by ν on Ω_i. Also, let $S := \Pi_{i=1}^m S_i$ be the product of m compact metric *action spaces*. A *strategy* for player i is a Young measure $\delta_i \in \mathcal{R}(\Omega_i, \mathcal{A}_i, \nu_i; S_i)$, $1 \leq i \leq m$. An m-tuple $\delta. := (\delta_1, \cdots, \delta_m)$, $\delta_i \in \mathcal{R}(\Omega_i, \mathcal{A}_i, \nu_i; S_i)$, $1 \leq i \leq m$, is called a *strategy profile*; this gives the following *expected utility* to player j:

$$U_j(\delta.) := \int_\Omega [\int_S u_j(\omega, x) \otimes_{i=1}^m \delta_i(\omega)(dx)] \nu(d\omega),$$

where u_j is a ν-integrably bounded continuous integrand on $\Omega \times S$, which is $\mathcal{A} \times \mathcal{B}(S)$-measurable. Here $\otimes_{i=1}^m \delta_i \in \mathcal{R}(\Omega, \mathcal{A}, \nu; S)$ is a m-fold tensor product, obtained by applying Definition 3.22 m times. Then there exists a strategy profile $(\delta^*_\cdot)$ such that for $j = 1, \cdots, m$

$$U_j(\delta^*_\cdot) \geq U_j(\delta^*_1, \cdots, \delta^*_{j-1}, \delta_j, \delta^*_{j+1}, \cdots, \delta^*_m) \text{ for all } \delta_j \in \mathcal{R}(\Omega_j, \mathcal{A}_j, \nu_j; S_j). \quad (4.3)$$

Thus, $\delta^*_\cdot$ constitutes a *Nash equilibrium* profile. Its existence is seen as follows. First, by Proposition A.12 there exist countably generated sub-σ-algebras $\mathcal{A}_{0,i} \subset \mathcal{A}_i$, $1 \leq i \leq m$, for which $u_1, \cdots, u_m$ are $\mathcal{A}_0 \times \mathcal{B}(S)$-measurable; here $\mathcal{A}_0 := \times_{i=1}^m \mathcal{A}_{0,i}$ (observe that the proof of Proposition A.12 immediately extends to the present product situation). Correspondingly, let $\mathcal{R}_i$ be the set of all Young measures from $(\Omega_i, \mathcal{A}_{0,i}, \nu_i)$ into S_i; then $\mathcal{R}_i$ is narrowly semimetrizable and compact by Example 4.11. Secondly, we observe that

$U_j(\delta.) = I_{v_j}(\otimes_{i=1}^m \delta_i)$, where the integrand is $v_j(\omega, x) := u_j(\omega, x)\frac{d\nu}{d\mu}(\omega)$. Hence, by Theorems 3.23 and 3.19(c) (the latter being applied twice), it follows that U_j is continuous on $\Pi_{i=1}^m \mathcal{R}_i$. By Ky Fan's inequality (Theorem A.5) this gives that there exists a profile $\delta_\cdot^* \in \Pi_{i=1}^m \mathcal{R}_i$ such that

$$\chi(\delta_\cdot^*, \delta.) - \chi(\delta_\cdot^*, \delta_\cdot^*) \leq 0 \text{ for all profiles } \delta.,$$

which implies that (4.3) holds for all $\delta_j \in \mathcal{R}_j$. Here

$$\chi(\delta'_\cdot, \delta.) := \sum_{i=1}^m [U_i(\delta'_\cdot) - U_i(\delta'_1, \cdots, \delta'_{i-1}, \delta_i, \delta'_{i+1}, \cdots, \delta'_m)].$$

Finally, a standard conditioning argument for the Young measures ('existence of regular conditional distributions' [Bi2]), plus a standard approximation argument for $\mathcal{L}^1_{C(S)}(\Omega, \mathcal{A}, \mu)$, gives that (4.3) also holds for all δ_j in $\mathcal{R}(\Omega_j, \mathcal{A}_j, \nu_j; S_j)$. □

The next application is from [Ba18] (see [Ba22] for further extensions):

Example 4.13 (existence of Cournot-Nash equilibria) Let $\mu(\Omega) = 1$, let S be a nonempty compact metric space and let $u : \Omega \times S \times \mathcal{P}(S) \to [-\infty, +\infty]$ be a $\mathcal{A} \times \mathcal{B}(S \times \mathcal{P}(S))$-measurable utility function, such that $u(\omega, \cdot, \cdot)$ is upper semicontinuous for each $\omega \in \Omega$ and $u(\omega, x, \cdot)$ is narrowly continuous on $\mathcal{P}(S)$ for each $(\omega, x) \in \Omega \times S$. Here $\mathcal{P}(S)$ is equipped with the Borel σ-algebra for the narrow topology. Then there exists $\delta_* \in \mathcal{R}(\Omega, \mathcal{A}, \mu; S)$ such that for a.e. ω in Ω

$$\delta_*(\omega)(\{x \in S : u(\omega, x, (\mu \otimes \delta_*)(\Omega \times \cdot)) = \max_{x' \in S} u(\omega, x', (\mu \otimes \delta_*)(\Omega \times \cdot)\}) = 1.$$

The product probability $\mu \otimes \delta_*$ is known in economics as a *Cournot-Nash equilibrium distribution.* The model above extends the original one of A. Mas-Colell [Ma]; see [Ba18, Ba22] for more details and several generalizations. The proof of the above existence result is as follows. First, by Proposition A.12 there exists a countably generated sub-σ-algebra $\mathcal{A}_0$ of $\mathcal{A}$ such that u is also $\mathcal{A}_0 \otimes \mathcal{B}(S)$-measurable. By Example 4.11 the set $\mathcal{R}_0 := \mathcal{R}(\Omega, \mathcal{A}_0, \mu; S)$ is narrowly semimetrizable and compact. Define

$$\chi'(\delta', \delta) := \int_\Omega [\int_S \arctan(u(\omega, x, (\mu \otimes \delta')(\Omega \times \cdot))\delta(\omega)(dx)]\mu(d\omega)$$

By virtue of Theorem 3.7 the functional χ' is easily seen to be upper semicontinuous by Theorems 3.19(c) and 3.23 (this argument also uses the metrizability of $\mathcal{P}(S)$), while for any fixed $\delta \in \mathcal{R}_0$, $\chi'(\cdot, \delta)$ is continuous by a simple application of the dominated convergence theorem. We can apply Theorem A.5 to the functional χ on $\mathcal{R}_0 \times \mathcal{R}_0$, defined by $\chi(\delta', \delta) := \chi'(\delta', \delta) - \chi'(\delta', \delta')$, and this gives the existence of δ_* in $\mathcal{R}_0$ such that

$$\chi'(\delta_*, \delta_*) = \sup_{\delta \in \mathcal{R}_0} \chi'(\delta_*, \delta).$$

Finally, a well-known measurable selection argument gives

$$\sup_{\delta \in \mathcal{R}_0} \chi'(\delta_*, \delta) = \int_\Omega \arctan(\sup_{x \in S} u(\omega, x, (\mu \otimes \delta_*)(\Omega \times \cdot)))\, \mu(d\omega),$$

from which the result follows (see [Ba18, Ba22] for more details). □

Next, we discuss an application of K-convergence to Stackelberg-type equilibria. This kind of result has a continuum of constraints, and it would seem quite hard to find proofs which avoid K-convergence.

Example 4.14 (existence of Stackelberg equilibria) Let $\mu(\Omega) = 1$ and let $\Gamma : \Omega \to 2^S$ be compact-valued. Let u_l, u_f be respectively an upper semicontinuous and a continuous integrand on $\Omega \times S$ (these represent utility functions for the two players). Consider the following optimization problem

$$(\mathcal{P}): \ \max \int_\Omega^* [\int_S u_l(\omega, x)\delta(\omega)(dx)]\mu(\omega)$$

over all $\delta \in \mathcal{R}_\Gamma$ (see Example 4.11 for the notation) that meet the following *incentive compatibility* condition: For a.e. $\omega \in \Omega$

$$\int_S u_f(\omega, x)\delta(\omega)(dx) \geq \int_S u_f(\omega, x)\delta(\omega')(dx)$$

for a.e. $\omega' \in \Omega$. We suppose that at least one $\delta \in \mathcal{R}_\Gamma$ meets these conditions (for instance, in case $\Gamma(\omega) \equiv S$ one can take any *constant* δ for this). In this model of a Stackelberg-type game with incomplete information, Ω stands for the space of possible *types* of the second player ('follower'), and μ represents how the first player ('leader') believes they are distributed. The incentive compatibility condition can be interpreted as a condition which forces the follower to report his/her type faithfully. Because of obvious tightness (cf. Example 4.11), the application of the relative K-compactness Theorem 4.7 to a minimizing sequence sequence for $(\mathcal{P})$, followed by the Fatou Lemma 3.13, gives the existence of an optimal solution (observe that the feasible set, which consists of the incentive compatible δ's in $\mathcal{R}_\Gamma$, is closed for K-convergence). The above model is a substantial simplification of [Pa, Ba23], where the reader can find a much more complete and subtle account, including quite special measurable selection arguments. □

5 Synthesis and applications

Contents: growth dominance properties, synthesis, fine twinning, existence of optimal relaxed and ordinary controls

In this section we present a synthesis of previous material which should be adequate to understand most of the applications which are to follow, *without* having read all of the preceding sections. Instead, all that is required is an understanding of the statement of Theorem 5.5 (its proof heavily depends upon the aforegoing, of course), complemented by an occasional glance at an earlier definition or result. In particular, Definitions 3.8 and 3.9 should be consulted for a proper understanding of the following notation:

$$I_g(\delta) := \int_\Omega^* [\int_S g(\omega, x)\delta(\omega)(dx)]\mu(d\omega).$$

Also, the notions of the Kuratowski limes superior, presented in Definitions 2.7, 2.16 and 2.17, should be kept in mind.

Let $(\Omega, \mathcal{A}, \mu)$ be a finite measure space and (S, ρ) a separable metric space. Let τ be another topology on S, possibly nonmetrizable. We begin with a growth dominance condition from [BL], which is extended below:

Definition 5.1 (property (γ)) For functions $h : \Omega \times S \to [0, +\infty]$ and $g : \Omega \times S \to (-\infty, +\infty]$ we say that g has *property (γ) with respect to h* if for every $\epsilon > 0$ there exists $\phi_\epsilon \in \mathcal{L}^1_{\mathbf{R}}$ such that

$$g(\omega, x) + \epsilon h(\omega, x) \geq \phi_\epsilon(\omega) \text{ for all } \omega \in \Omega \text{ and } x \in S.$$

□

Intuitively speaking, this expresses that the negative part of g does not nearly grow as fast as h. We cite a familiar example, associated with de la Vallée-Poussin's Theorem A.3, which will serve us frequently.

Example 5.2 (superlinear growth) Let S be $\mathbf{R}^d$ and let $h' : \mathbf{R}_+ \to \mathbf{R}_+$ be such that $\lim_{\xi\to\infty} h'(\xi)/\xi = +\infty$ (superlinear growth). Then g, defined by $g(\omega, x) := -|x|$, has property (γ) with respect to h, given by $h(\omega, x) := h'(|x|)$. □

A growth dominance condition which is more general than property (γ) is given next:

Definition 5.3 (property (γ')) Let (δ_n) be a sequence in $\mathcal{R}$. For functions $h : \Omega \times S \to [0, +\infty]$ and $g : \Omega \times S \to (-\infty, +\infty]$ we say that g has *property (γ') with respect to h and (δ_n)* if for every $\epsilon > 0$ there exists a uniformly integrable sequence $(\phi_{n,\epsilon})$ in $\mathcal{L}^1_{\mathbf{R}}$ such that

$$g(\omega, x) + \epsilon h(\omega, x) \geq \phi_{n,\epsilon}(\omega) \text{ for all } \omega \in \Omega \text{ and } x \in \tau\text{-supp } \delta_n(\omega).$$

for every $n \in \mathbf{N}$ □

Property (γ') combines two different and well-known lower boundedness conditions for integrands:

Remark 5.4 Let $h : \Omega \times S \to [0, +\infty]$ and $g : \Omega \times S \to (-\infty, +\infty]$ and let $(\delta_n) \subset \mathcal{R}$.

(*i*) If there exists a uniformly integrable sequence (ϕ_n) in $\mathcal{L}^1_{\mathbf{R}}$ for which $g(\omega, x) \geq \phi_n(\omega)$ for all $x \in \text{supp } \delta_n(\omega)$, $\omega \in \Omega$, $n \in \mathbf{N}$, then g has property (γ') with respect to h and (δ_n).

(*ii*) If g has property (γ) with respect to h, then g also has property (γ') with respect to h and (δ_n). □

One of the most useful and versatile results of these lecture notes is the following generalization of Theorem 2.20. Just as that theorem formed a synthesis of Theorems 2.3, 2.8 and 2.13, the present result synthesizes Theorems 3.11, 3.13 and 4.7. The fact that it is not restricted to the ρ-topology on S, will be particularly relevant in the next section.

Theorem 5.5 (synthesis) *Suppose that τ is at least as strong as the ρ-topology on S and that $\mathcal{B}(S_\tau) = \mathcal{B}(S_\rho)$. Let (δ_n) be a sequence in $\mathcal{R}$ for which there exists a nonnegative sequentially τ-inf-compact integrand h on $\Omega \times S$ with*

$$\sup_n I_h(\delta_n) < +\infty.$$

Then there exist a subsequence $(\delta_{n'})$ of (δ_n) and a Young measure $\delta_* \in \mathcal{R}$ such that

$$\liminf_{n'} I_g(\delta_{n'}) \geq I_g(\delta_*) \tag{5.1}$$

for every sequentially τ-lower semicontinuous integrand g on $\Omega \times S$ which has property (γ') with respect to h and (δ_n). Moreover,

$$\tau\text{-supp } \delta_*(\omega) \subset \tau\text{-seq-cl } \tau\text{-Ls}_n \tau\text{-supp } \delta_n(\omega) \textit{ for a.e. } \omega \textit{ in } \Omega.$$

Proof. *A fortiori, $h(\omega, \cdot)$* is inf-compact on S_ρ for every $\omega \in \Omega$; this means that (δ_n) is tight in $\mathcal{R}$ in the sense of Definitions 4.1–4.2. By Theorem 4.7, there exist a subsequence $(\delta_{n'})$ of (δ_n) and δ_* in $\mathcal{R}$ such that $\delta_{n'} \xrightarrow{K} \delta_*$ in $\mathcal{R}$. By Definition 3.10 and Corollary 2.4 this implies $\delta_{n'} \otimes \epsilon_{n'} \xrightarrow{K} \delta_* \otimes \epsilon_\infty$. Let $g : \Omega \times S \to (-\infty, +\infty]$ and $\epsilon > 0$ be arbitrary, with the properties mentioned in the statement. In particular, let $(\phi_{n,\epsilon})$ be as guaranteed by g's property (γ') with respect to h. We now combine essential elements from the proofs of Theorems 2.20 and 3.13. By de la Vallée-Poussin's Theorem A.3, there exists a function $h'_\epsilon : \mathbf{R}_+ \to \mathbf{R}_+$, convex, continuous and nondecreasing, with $\lim_{\xi \to \infty} h'_\epsilon(\xi)/\xi = +\infty$ and

$$\sup_n \int_\Omega h'_\epsilon(|\phi_{n,\epsilon}(\omega)|)\mu(d\omega) \leq \epsilon$$

(the bound ϵ follows by elementary normalization). Set $g_\epsilon := g + \epsilon h$ and define $g_{1,\epsilon} : \Omega \times S \times \hat{\mathbf{N}} \to (-\infty, +\infty]$ as follows:

$$g_{1,\epsilon}(\omega, x, k) := \begin{cases} g_\epsilon(\omega, x) + h'_\epsilon(|\phi_{k,\epsilon}(\omega)|) & \text{if } x \in \tau\text{-supp } \delta_k(\omega) \text{ and } k < \infty \\ g_\epsilon(\omega, x) + \liminf_n h'_\epsilon(|\phi_{n,\epsilon}(\omega)|) & \text{if } x \in \tau\text{-seq-cl } \tau\text{-Ls}_n \tau\text{-supp } \delta_n(\omega), k = \infty \\ +\infty & \text{otherwise} \end{cases}$$

Similar to the proof of Theorem 3.13, property (γ') implies $g_{1,\epsilon}(\omega, x, k) \geq \gamma_\epsilon$ for all $\omega \in \Omega$, $x \in S$ and $k < \infty$. Here $\gamma_\epsilon := \inf_{\xi \in \mathbf{R}}[\xi + h'_\epsilon(|\xi|)]$ is finite by the superlinear growth of h'_ϵ. We claim that for every $\omega \in \Omega$, $g_{1,\epsilon}(\omega, \cdot, \cdot)$ is lower semicontinuous on $S_\rho \times \hat{\mathbf{N}}$. Indeed, let $k_j \to k_0$ in $\hat{\mathbf{N}}$ and $x_j \to \bar{x}$ in S_ρ. Set $\alpha := \liminf_j g_{1,\epsilon}(\omega, x_j, k_j)$ and suppose $\alpha < +\infty$ to avoid trivialities. There exist subsequences $(k_{j'})$ and $(x_{j'})$ such that $g_{1,\epsilon}(\omega, x_{j'}, k_{j'}) \to \alpha$ in

$\mathbf{R}$. First, consider the case where $k_0 = \infty$ and all other k_j's are finite. Then, by definition of $g_{1,\epsilon}(\omega,\cdot,\cdot)$, $x_{j'} \in \text{supp}\ \delta_{k_{j'}}(\omega)$ for all j' and also $h(\omega, x_{j'}) \leq (\alpha + 1 - \gamma_\epsilon)/\epsilon$, for j' large enough, by the lower bound established above. Hence, a subsequence of $(x_{j'})$ τ-converges to $\bar{x}$; therefore, $\bar{x}$ belongs to $\tau\text{-Ls}_n\text{supp}\ \delta_n(\omega)$. By the given sequential τ-lower semicontinuity of $g(\omega,\cdot)$ and the definition of $g_{1,\epsilon}$, it is now easy to conclude that $\alpha \geq g_{1,\epsilon}(\omega, \bar{x}, \infty)$. For the remaining possible cases (i.e., one with $k_0 < \infty$ and one with $k_j = \infty$ eventually) similar arguments can be given, which we leave to the reader. Next, we define $g_{2,\epsilon} := \max(g_{1,\epsilon}, \gamma_\epsilon)$; by the above, this is a lower semicontinuous integrand on $\Omega \times (S_\rho \times \hat{\mathbf{N}})$, bounded from below, with $I_{g_{1,\epsilon}}(\delta_n \times \epsilon_n) = I_{g_{2,\epsilon}}(\delta_n \otimes \epsilon_n)$ for all $n < \infty$. In a move which parallels the proof of Theorem 2.8 and which will give the desired support properties for δ_*, we consider first the special case $g = \bar{g}$, with $\bar{g} :\equiv 0$. By Corollary 2.19 we have $I_{\bar{g}_{1,\epsilon}}(\delta_n \times \epsilon_n) = I_{\bar{g}_{2,\epsilon}}(\delta_n \otimes \epsilon_n) = 0$ for every $n \in \mathbf{N}$ [because $I_h(\delta_n) < +\infty$ implies for a.e. ω that $\int_S h(\omega, x)\, \delta_n(\omega)(dx) < +\infty$, i.e., that $\delta_n(\omega)$ is τ-tight (Definition 2.11)]. Hence, application of the Fatou Lemma 3.12 gives

$$I_{\bar{g}_{1,\epsilon}}(\delta_* \otimes \epsilon_\infty) \leq I_{\bar{g}_{2,\epsilon}}(\delta_* \otimes \epsilon_\infty) \leq \epsilon\sigma + \epsilon < +\infty,$$

with $\sigma := \sup_n I_h(\delta_n)$. This immediately gives that $\bar{g}_{1,\epsilon}(\omega, \cdot, \infty)$ is finite on $\text{supp}\ \delta_*(\omega)$ for a.e. ω, so the proclaimed support property follows by the definition of $\bar{g}_{1,\epsilon}$. Similarly, and now running in line with the proofs of Theorem 2.20 and 3.13, we obtain for general g

$$I_g(\delta_*) \leq I_{g_{1,\epsilon}}(\delta_* \otimes \epsilon_\infty) \leq I_{g_{2,\epsilon}}(\delta_* \otimes \epsilon_\infty) \leq \liminf_{n'} I_g(\delta_{n'}) + \epsilon\sigma + \epsilon,$$

by also using the support property of δ_*, just derived. Letting ϵ converge to zero gives $I_g(\delta_*) \leq \liminf_{n'} I_g(\delta_{n'})$. □

Remark 5.6 (addenda) The inequality (5.1) can be extended into

$$\liminf_{n'} I_g(\delta_{n'} \otimes \epsilon_{n'}) \geq I_g(\delta_* \otimes \epsilon_\infty) \tag{5.2}$$

for every sequentially lower semicontinuous integrand g on $\Omega \times (S_\tau \times \hat{\mathbf{N}})$ which has property (γ') with respect to h and (δ_n) (by the latter we formally mean that g has property (γ') with respect to $\tilde{h}$ and (δ_n), where $\tilde{h}$ is the inf-compact integrand on $\Omega \times (S_\tau \times \hat{\mathbf{N}})$ defined by $\tilde{h}(\omega, x, k) := h(\omega, x)$) To see this, one can either expand the proof of Theorem 5.5 or simply apply Theorem 5.5 as we have it, with S replaced by $S \times \hat{\mathbf{N}}$, to the sequence $(\tilde{\delta}_n)$ given by $\tilde{\delta}_n := \delta_n \otimes \epsilon_n$, and to deduce that the corresponding $\tilde{\delta}_*$ must be of the form $\tilde{\delta}_* = \delta_* \otimes \epsilon_\infty$ for some $\delta_* \in \mathcal{R}$ [one does so by taking $g(\omega, x, k) := c(k)$, $c \in \mathcal{C}_b(\hat{\mathbf{N}})$, in (5.1)]. □

In complete analogy to Remark 2.23, the following can be observed:

Remark 5.7 *(i)* In Theorem 5.5 for every $\omega \in \Omega$

$$\tau\text{-seq-cl}\ \tau\text{-Ls}_n\tau\text{-supp}\ \delta_n(\omega) \subset \tau\text{-LS}_n\tau\text{-supp}\ \delta_n(\omega) \subset \rho\text{-Ls}_n\rho\text{-supp}\ \delta_n(\omega).$$

(ii) In Theorem 5.5

$$\tau\text{-seq-cl}\ \tau\text{-Ls}_n\tau\text{-supp}\ \delta_n(\omega) = \tau\text{-Ls}_n\tau\text{-supp}\ \delta_n(\omega),$$

for every $\omega \in \Omega$ for which there exists a τ-compact set containing all supports $\text{supp}\ \delta_n(\omega)$, $n \in \mathbf{N}$. □

Proposition 5.8 (existence of optimal randomized decision rules) *Let g_{m+1} be a nonnegative inf-compact integrand on $\Omega \times S$, and let $g_0, g_1, \cdots, g_m$ be lower semicontinuous integrands on $\Omega \times S$, having property (γ) with respect to g_{m+1}. Let $\alpha := (\alpha_1, \cdots, \alpha_{m+1})$ be in $\mathbf{R}^{m+1}$. Then the optimization problem*

$$(\mathcal{P}): \inf_{\delta \in \mathcal{R}} \{ I_{g_0}(\delta) : I_{g_1}(\delta) \leq \alpha_1, \cdots, I_{g_{m+1}}(\delta) \leq \alpha_{m+1} \}$$

has an optimal solution, provided that the feasible set is nonempty.

PROOF. Let (δ_n) be any minimizing sequence for $(\mathcal{P})$; then $I_{g_0}(\delta_n) \to \inf(\mathcal{P})$. By the $m+1$-st constraint, (δ_n) meets the tightness condition of Theorem 5.5 (take for τ the ρ-topology). Hence, there exist a subsequence $(\delta_{n'})$ and $\delta_* \in \mathcal{R}$ for which the statement of Theorem 5.5 holds. It is then immediate that δ_* is feasible and that $I_{g_0}(\delta_*) = \inf(\mathcal{P})$. □

Proposition 5.9 (existence of optimal nonrandomized decision rules) *Let g_0 be a nonnegative inf-compact integrand on $\Omega \times S$, and let $g_1, \cdots, g_m$ be continuous integrands on $\Omega \times S$, $\mathcal{A} \times \mathcal{B}(S)$-measurable, with $|g_1|, \cdots, |g_m|$ having property (γ) with respect to g_0. Suppose that for all $\beta := (\beta_1, \cdots, \beta_m)$ in a neighborhood of $\alpha := (\alpha_1, \cdots, \alpha_m) \in \mathbf{R}^m$ there exists $\delta \in \mathcal{R}$ such that $I_{g_1}(\delta) = \beta_1, \cdots, I_{g_m}(\delta) = \beta_m$ and $I_{g_0}(\delta) < +\infty$. Then for the optimization problem*

$$(\mathcal{P}): \inf_{\delta \in \mathcal{R}} \{ I_{g_0}(\delta) : I_{g_1}(\delta) = \alpha_1, \cdots, I_{g_m}(\delta) = \alpha_m \}$$

there exists $\lambda \in \mathbf{R}^m$ such for that every $\delta \in \mathcal{R}$ the following are equivalent:
(a) δ is optimal for $(\mathcal{P})$.
(b) For a.e. ω in Ω

$$\text{supp } \delta(\omega) \subset \{ x \in S : g_\lambda(\omega, x) = \min_S g_\lambda(\omega, \cdot) \},$$

where $g_\lambda := g_0 + \sum_{i=1}^m \lambda_i g_i$.

PROOF. First, by Proposition 5.8 applied to $g'_{2m+1} := g_0$ and $g'_i := g_i$, $g'_{m+i} := -g_i$, there exists an optimal solution for $(\mathcal{P})$ (observe that the constraint $I_{g_0}(\delta) \leq \inf(\mathcal{P}) + 1$ can trivially be added). The nonemptiness condition clearly forms a stability condition for the canonical perturbations of the convex optimization problem $(\mathcal{P})$ in the sense of [ET, Definition III.2.2]. Hence, there exists a Lagrange multiplier $\lambda \in \mathbf{R}^m$ such that

$$\inf(\mathcal{P}) = \inf_{\delta \in \mathcal{R}} I_{g_\lambda}(\delta).$$

By a standard measurable selection argument, based on Theorem A.4, this gives the desired result. □

For applications of Proposition 5.9 to *fine phase mixing* in multicomponent fluid mechanics we refer to [BK]. There nonclassical solutions can arise, which correspond to optimal non-Dirac solutions in Proposition 5.9. A rather similar phenomenon, where Young measures also play an important descriptive role, is found in nonlinear elasticity, in connection with *fine twinning* in crystals. Instead of presenting these ideas in a detailed mechanical model, we only discuss the following artificial example from [BK], which is rather related to the initial Example 1.3.

Example 5.10 (fine twinning) Let $\Omega := (0,1)^2$ be equipped with $\mathcal{A} := \mathcal{L}((0,1)^2)$ and $\mu := \lambda_2$. Consider the following optimization problem

$$(\mathcal{P}): \ \inf J(f) := \int_\Omega [(f_{\omega_1}^2 - 1)^2 + f_{\omega_2}^2] d\lambda_2$$

over all $f \in W^{1,2}(\Omega)$ with $f(\omega_1, 0) = 0$ (here J represents some energy functional). Then $\inf(\mathcal{P}) = 0$, but $(\mathcal{P})$ does not have an optimal solution. To see the former, define $f_1 : (0,1) \times (0,+\infty) \to \mathbf{R}$ by

$$f_1(\omega_1,\omega_2) := \begin{cases} \omega_1 \min(\omega_2, 1) & \text{if } 0 \le \omega_1 \le \frac{1}{2} \\ (1-\omega_1)\min(\omega_2,1) & \text{if } \frac{1}{2} \le \omega_1 \le 1 \end{cases}$$

and extend $f_1(\omega_1,\omega_2)$ periodically with period 1 in ω_1 to $\mathbf{R} \times (0,+\infty)$. Define (f_n) by $f_n(\omega_1,\omega_2) := n^{-1} f_1(n\omega_1, n\omega_2)$. Then $J(f_n) \to 0$, by uniform boundedness of the gradients of the functions f_n and by $(f_n)_{\omega_2}(\omega_1,\omega_2) = 0$ for all $\omega_2 \ge n^{-1}$. This immediately gives $\inf(\mathcal{P}) = 0$. However, $J(f) = 0$ for $f \in W^{1,2}(\Omega)$ would imply $f_{\omega_2} = 0$ a.e., whence $f = 0$ in view of the boundary condition. In turn this would give $J(f) = J(0) = 1$, which gives a contradiction. It is the highly oscillatory behavior of (f_n) which in mechanical reality can be observed to lead to twinning (with a period of at most a few atomic spacings). Quite similar to Example 1.3, one can now formulate the corresponding *Young-relaxation* of $(\mathcal{P})$ and prove that $\delta_*(\omega_1,\omega_2) \equiv \frac{1}{2}\epsilon_{(1,0)} + \frac{1}{2}\epsilon_{(-1,0)}$ is its unique optimal solution. In some sense this can be seen as a condensed description of the oscillatory behavior of the minimizing sequence (f_n). $\square$

Example 5.11 (existence of optimal relaxed controls) Let $(\Omega, \mathcal{A}, \mu)$ be the Lebesgue unit interval $([0,1], \mathcal{L}([0,1]), \lambda_1)$. Let h be a nonnegative inf-compact integrand on $[0,1] \times S$. Consider the optimal control problem

$$(\mathcal{P}): \ \inf_{f \in \mathcal{L}_S^0} J(f) := \int_{[0,1]}^* g_0(\omega, f(\omega), y_f(\omega)) d\omega + e(y_f(1)),$$

under the constraint

$$J_h(f) := \int_{[0,1]}^* h(\omega, f(\omega)) d\omega \le 1. \tag{5.3}$$

Here $y_f : [0,1] \to \mathbf{R}^m$ in $\mathcal{AC}([0,1]; \mathbf{R}^m)$ is the absolutely continuous solution of the ordinary differential equation

$$\dot{y}(\omega) = c(\omega, f(\omega), y(\omega)) \text{ for a.e. } \omega \text{ in } \Omega,$$

and we use initial condition $y(0) = y_0$. Further, $c : [0,1] \times S \times \mathbf{R}^m \to \mathbf{R}^m$ is measurable, continuous in its second and third arguments and for every $\epsilon > 0$ there exists $\psi_\epsilon \in \mathcal{L}_{\mathbf{R}}^1$ with

$$|c(\omega, x, y)| \le \epsilon h(\omega, x) + \psi_\epsilon(\omega) \tag{5.4}$$

for all $\omega \in [0,1]$, $x \in S$ and $y \in \mathbf{R}^m$ (hence each component function c^i of $c := (c^1, \ldots, c^m)$ is a continuous integrand on $[0,1] \times (S \times \mathbf{R}^m)$ such that c^i and $-c^i$ have property (γ) with respect to h uniformly in the variable y – rather less demanding boundedness conditions are indicated at the end of this example). Also, g_0 is a lower semicontinuous integrand on $[0,1] \times (S \times \mathbf{R}^m)$, and we suppose that g_0 has property (γ) with respect to h, uniformly in the variable y. Finally, $e : \mathbf{R}^m \to (-\infty, +\infty]$ is supposed to be lower semicontinuous.

Even though $(\mathcal{P})$ may not have any optimal solution (see Example 1.3), the following *Young relaxation* $(\mathcal{P}_{rel})$ of $(\mathcal{P})$ has an optimal solution:

$$(\mathcal{P}_{rel}): \inf_{\delta \in \mathcal{R}} J_{rel}(\delta) := \int^*_{[0,1]} [\int_S g_0(\omega, x, y_\delta(\omega))\delta(\omega)(dx)]d\omega + e(y_\delta(1)),$$

under the constraint

$$I_h(\delta) := \int^*_\Omega [\int_S h(\omega, x)\delta(\omega)(dx)]\mu(d\omega) \leq 1, \tag{5.5}$$

which is relaxed correspondingly. Here $y_\delta : [0,1] \to \mathbf{R}^m$ in $\mathcal{AC}([0,1];\mathbf{R}^m)$ is the absolutely continuous solution of the ordinary differential equation

$$\dot{y}(\omega) = \int_S c(\omega, x, y(\omega))\delta(\omega)(dx) \text{ for a.e. } \omega \text{ in } \Omega,$$

with initial condition $y(0) = y_0$.

To see the existence of an optimal solution of $(\mathcal{P}_{rel})$, observe first that by (5.5) any minimizing sequence (δ_n) of $(\mathcal{P}_{rel})$ is automatically tight. So Theorem 5.5 applies: let $(\delta_{n'})$ and $\delta_* \in \mathcal{R}$ be as stated in that theorem. We claim that $\sup_{[0,1]} |y_{\delta_{n'}} - y_{\delta_*}| \to 0$. To prove this claim, observe that by the property (γ) condition for the c^i's the collection $(y_{\delta_{n'}})$ is equicontinuous and equibounded. Hence there exist a subsequence $(\delta_{n''})$ of $(\delta_{n'})$ and $y_* \in \mathcal{C}[0,1]$ such that $\sup_{[0,1]} |y_{\delta_{n''}} - y_*| \to 0$ (Arzela-Ascoli theorem [War]). Fix ω' and $1 \leq i \leq m$ arbitrarily; then

$$y^i_{\delta_{n''}}(\omega') = y^i_0 + \int_0^{\omega'} [\int_S c^i(\omega, x, y_{\delta_{n''}}(\omega))\delta(\omega)(dx)]d\omega.$$

Now define

$$g(\omega, x, k) := \begin{cases} 1_{[0,\omega']}(\omega)c^i(\omega, x, y_{\delta_k}(\omega)) & \text{if } k < \infty \\ 1_{[0,\omega']}(\omega)c^i(\omega, x, y_{\delta_*}(\omega)) & \text{if } k = \infty \end{cases}$$

By the above, g is a continuous integrand on $[0,1] \times (S \times \hat{\mathbf{N}})$. Hence, a twofold application of (5.2) implies

$$\lim_{n''}(y^i_{\delta_{n''}}(\omega') - y^i_0) = \lim_{n''} I_g(\delta_{n''} \otimes \epsilon_{n''}) = I_g(\delta_* \otimes \epsilon_\infty) = y^i_{\delta_*}(\omega') - y^i_0,$$

by virtue of Remark 5.6. So we also have $y_{\delta_{n''}} \to y_{\delta_*}$ pointwise. This shows that $y_* = y_{\delta_*}$, so the claim has been proven. We now finish by applying (5.2) to the lower semicontinuous integrand g' on $[0,1] \times (S \times \hat{\mathbf{N}})$, defined by

$$g'(\omega, x, k) := \begin{cases} g_0(\omega, x, y_{\delta_k}(\omega)) & \text{if } k \in \mathbf{N} \\ g_0(\omega, x, y_{\delta_*}(\omega)) & \text{if } k = \infty \end{cases}$$

This proves that $\liminf_{n''} J_{rel}(\delta_{n''}) \geq J_{rel}(\delta_*)$. Since $(\delta_{n''})$ is a minimizing sequence for $(\mathcal{P}_{rel})$, we conclude that $\inf(\mathcal{P}_{rel})$ is attained by δ_*.

In some control problems the constraints (5.3) and (5.5) do not appear explicitly. One obvious case is the one where an inequality such as $g_0(\omega, x, y) \geq h(\omega, x) + \psi(\omega)$ holds for some $\psi \in \mathcal{L}^1_{\mathbf{R}}$. Provided that the problem is nontrivial (i.e., $\inf(\mathcal{P}) < +\infty$), this yields (5.3) for some suitable multiple of h (consider a minimizing sequence). In this way we could use $h(\omega, x) := (x^2 - 1)^2$ and $\psi = 0$ in Example 1.3, which is a special case of the present example. Let us also observe that the boundedness condition (5.4) for $|c|$, used until now, could be replaced by the following condition: There exist $\phi \in \mathcal{L}^1_{\mathbf{R}}$, $\alpha > 0$ and a nondecreasing function $\chi : \mathbf{R}_+ \to \mathbf{R}_+$ such that

$$|y \cdot c(\omega, x, y)| \leq \phi(\omega)(|y|^2 + 1) + \alpha h(\omega, x),$$

and for every $\epsilon > 0$ there exists $\psi_\epsilon \in \mathcal{L}^1_{\mathbf{R}}$ with

$$|c(\omega, x, y)| \leq \psi_\epsilon(\omega)\chi(|y|) + \epsilon h(\omega, x)$$

for all $\omega \in [0, 1]$, $x \in S$ and $y \in \mathbf{R}^m$. To see that this suffices, imitate the corresponding proof for a simpler situation in [War, p. 341], using Gronwall's inequality.

Finally, we note that the classical situation in relaxed control theory [War] is captured by taking a compact-valued multifunction $\Gamma : \Omega \to 2^S$ and by defining h by

$$h(\omega, x) := \begin{cases} 0 & \text{if } x \in \Gamma(\omega) \\ +\infty & \text{if } x \notin \Gamma(\omega) \end{cases}$$

In that case the constraint (5.5) for δ amounts precisely to the requirement $\delta(\omega)(\Gamma(\omega)) = 1$ a.e. (see also Example 4.11). □

Example 5.12 (existence of optimal controls) We continue Example 5.11 by an enquiry into the conditions needed to ensure the existence of an optimal solution for the original problem $(\mathcal{P})$. Let $\delta_* \in \mathcal{R}$ be the optimal solution for $(\mathcal{P}_{rel})$, found in Example 5.11. Denote y_{δ_*} as y_*. Suppose that in addition to Example 5.11 the following convexity condition holds:

$$Q(\omega, y_*(\omega)) \text{ is convex for a.e. } \omega \in \Omega,$$

where the *orientor field* multifunction $Q : [0, 1] \times \mathbf{R}^m \to 2^{\mathbf{R}^m \times \mathbf{R} \times \mathbf{R}_+}$ is given by letting $Q(\omega, y)$ be the set of all $(\xi, \eta, \zeta) \in \mathbf{R}^m \times \mathbf{R} \times \mathbf{R}_+$ for which

$$\xi = c(\omega, x, y),\ \eta \geq g_0(\omega, x, y),\ \zeta \geq h(\omega, x)$$

for some $x \in S$. We also suppose in addition to Example 5.11 that

$$g_{00} \text{ and } h \text{ are } \mathcal{A} \times \mathcal{B}(S)\text{-measurable},$$

where $g_{00}(\omega, x) := g_0(\omega, x, y_*(\omega))$. From Theorem A.13($ii$) and the convexity condition above it follows directly that

$$\int_S (c(\omega, x, y_*(\omega)), g_{00}(\omega, x), h(\omega, x))\delta_*(\omega)(dx) \in Q(\omega, y_*(\omega)) \text{ for a.e. } \omega \text{ in } \Omega.$$

Therefore, for a.e. ω in Ω there exists $x_\omega \in S$ with

$$\dot{y}_*(\omega) = c(\omega, x_\omega, y_*(\omega)),\ \textstyle\int_S g_{00}(\omega, x)\delta_*(\omega)(dx) \geq g_{00}(\omega, x_\omega),\ \int_S h(\omega, x)\delta_*(\omega)(dx) \geq h(\omega, x_\omega).$$

By the additional measurability hypothesis, an application of the measurable implicit function Theorem A.4 gives the existence of $f_* \in \mathcal{L}^0_S$ such that the above (in)equalities also hold a.e. for $f_*(\omega)$ instead of x_ω. Of course, this implies that $\dot{y}_{f_*} = \dot{y}_*$ a.e. (whence $y_{f_*} = y_*$) and that $J(f_*) \leq J_{rel}(\delta_*)$ and $J_h(f_*) \leq I_h(\delta_*) \leq 1$. Since $\inf(\mathcal{P}_{rel}) \leq \inf(\mathcal{P})$ is obvious (consider δ's in $\mathcal{R}_D$), it follows that f_* is optimal for $(\mathcal{P})$.

Compared to $(\mathcal{P}_{rel})$, the extra measurability hypotheses, introduced above to arrive at existence for $(\mathcal{P})$, would seem puzzling and unaesthetical. Fortunately, there is a natural resolution of the apparent dilemma: Following [Ba3], one can introduce so-called *measurable regularizations* of the semicontinuous integrands g_{00} and h (in [Bu] a systematic use is made of this device). Although this subject is not treated in these lecture notes, we mention that one can demonstrate the existence of $\mathcal{A} \times \mathcal{B}(S)$-measurable lower semicontinuous integrands $\tilde{g}_{00}$ and $\tilde{h}$ on $\Omega \times S$, with $\tilde{g}_{00} \geq g_{00}$ and $\tilde{h} \geq h$, such that $J_{\tilde{g}_{00}}(f) = J_{g_{00}}(f)$ and $J_{\tilde{h}}(f) = J_h(f)$ for every $f \in \mathcal{L}^0_S$. Therefore, the additional measurability hypothesis adopted here is not really needed. □

6 Applications to weak L^1-convergence

Contents: Dunford-Pettis-type criteria, L^1-versions of tightness, oscillation restriction, existence and well-posedness

In this section several consequences of the theory developed in previous sections will be studied in connection with weak and strong convergence of Bochner integrable functions. Throughout this section $(E, \|\cdot\|)$ will be a separable Banach space. At times E is equipped with its weak topology (w-topology), and at other times with its strong (i.e., norm) topology (s-topology). To treat these topologies on E simultaneously, we let τ be a topology on E that is not weaker than the w-topology and not stronger than the s-topology. Observe already that $\mathcal{B}(E_s) = \mathcal{B}(E_w) = \mathcal{B}(E_\tau)$, so that for $S := E$ the definition of the set $\mathcal{P}(S)$ and hence the Definition 3.1 of a Young measure are topology-independent.

Recall the definition of Bochner integrability [DU, War]:

Definition 6.1 (Bochner integrability) A function $f : \Omega \to E$ is said to be *Bochner integrable* (with respect to μ) if f is measurable with respect to $\mathcal{A}$ and $\mathcal{B}(E)$ and

$$\|f\|_1 := \int_\Omega \|f(\omega)\| \mu(d\omega) < +\infty.$$

The set of all such functions is denoted by $\mathcal{L}^1(\Omega, \mathcal{A}, \mu; E)$ (or $\mathcal{L}^1_E$). □

The central result of this section is as follows; it is a direct consequence of Theorem 5.5 (see also Examples 4.8–4.9).

Theorem 6.2 *Let (f_n) be in $\mathcal{L}^1_E$ uniformly L^1-bounded and such that there exists a non-negative sequentially τ-inf-compact integrand h on $\Omega \times E$ with*

$$\sup_n \int_\Omega^* h(\omega, f_n(\omega))\mu(d\omega) < +\infty. \tag{6.1}$$

Then there exist a subsequence $(f_{n'})$ and a Young measure $\delta_* \in \mathcal{R}$ such that

$$\liminf_{n'} \int_\Omega^* g(\omega, f_{n'}(\omega))\mu(d\omega) \geq \int_\Omega^* [\int_E g(\omega, x)\delta_*(\omega)(dx)]\mu(d\omega). \tag{6.2}$$

for every sequentially τ-lower semicontinuous integrand g on $\Omega \times E$ which has property (γ') with respect to h and (ϵ_{f_n}). Moreover,

$$\tau\text{-supp } \delta_*(\omega) \subset \tau\text{-seq-cl } \tau\text{-Ls}_n f_n(\omega) \textit{ for a.e. } \omega \in \Omega$$

and the parametrized barycenter f_* *of δ_*, given by*

$$f_*(\omega) := \text{bar } \delta_*(\omega) := \int_E x \, \delta_*(\omega)(dx),$$

exists almost everywhere and has a modification belonging to $\mathcal{L}^1_E$.

PROOF. *Step 1.* With an eye towards Remark 6.5 below, we distinguish two cases. In the first case τ is the s-topology on $S := E$. Then we apply Theorem 5.5 with ρ the norm-metric (i.e., the τ and ρ-topology coincide). In the second case τ is any other topology (but at least as strong as the w-topology). We then apply Theorem 5.5 with $\rho := d_w$, an auxiliary metric which is defined as follows: By separability of E and the Alaoglu-Bourbaki theorem the closed unit ball U^* of E's dual E^*, is metrizable and compact for the topology $\sigma(E^*, E)$. Therefore, U^* contains a countable, $\sigma(E^*, E)$-dense subset (x_i^*); we fix it and define a metric d_w on E by

$$d_w(x, x') := \sum_{i=1}^{\infty} 2^{-i} | < x - x', x_i^* > |.$$

Observe that even the w-topology on E is at least as strong as the d_w-topology. Observe also that then $\mathcal{B}(E_\tau) = \mathcal{B}(E_\rho)$, as required in Theorem 5.5. This can either be seen *ad hoc* or as a consequence of Remark 2.21.

Step 2. We now come to the actual application of Theorem 5.5 in the two cases. By Step 1 and (6.1), the sequence (ϵ_{f_n}) in $\mathcal{R}$ satisfies precisely the conditions of Theorem 5.5. Thus, there exist a subsequence $(f_{n'})$ of (f_n) and δ_* in $\mathcal{R}$ such that (5.1) holds. This gives immediately (6.2) and the proclaimed support property. Specialized to $g(\omega, x) := \|x\|$, the former also yields $I_g(\delta_*) \leq \liminf_{n'} \int_\Omega \|f_{n'}\| d\mu < +\infty$. Therefore, the existence a.e. of the barycenter bar $\delta_*(\omega)$ follows by Theorem A.13, just as in Example 4.8. On the exceptional null set we set f_* equal to 0. Then $f_* : \Omega \to E$ is clearly measurable and belongs to $\mathcal{L}_E^1$. □

Remark 6.3 (addenda) The proof above shows Theorem 6.2 to be a specialization of Theorem 5.5. By virtue of Remark 5.6, this implies that

$$\liminf_{n'} \int_\Omega^* g(\omega, f_{n'}(\omega), n')\mu(d\omega) \geq \int_\Omega^* [\int_E g(\omega, x, \infty)\delta_*(\omega)(dx)]\mu(d\omega)^7$$

for every sequentially lower semicontinuous integrand g on $\Omega \times (E_\tau \times \hat{\mathbf{N}})$ which has property (γ') with respect to h and (ϵ_{f_n}). Of course, this inequality remains trivially valid if $(f_{n'})$ is replaced by any subsequence $(f_{n''})$ of $(f_{n'})$. □

The following notion will frequently be used below:

Definition 6.4 (ball-compactness) A subset of E is said to be τ-*ball-compact* if its intersection with every closed ball of E is τ-compact. □

Observe that for the w-topology the entire space E, and all its w-closed subsets, are w-ball-compact if E is *reflexive.*

Remark 6.5 If in Theorem 6.2 τ is the s-topology, then the support property for δ_* reads

$$s\text{-supp } \delta_*(\omega) \subset s\text{-Ls}_n f_n(\omega) \text{ for a.e. } \omega \in \Omega,$$

since τ is then metrizable (cf. Definitions 2.16 and 2.17). On the other hand, if τ is not the s-topology (but at least as strong as the w-topology), then in general one can only say in the proclaimed support property that

$$\tau\text{-seq-cl } w\text{-Ls}_n f_n(\omega) \subset \tau\text{-LS}_n f_n(\omega) \subset d_w\text{-Ls}_n f_n(\omega) \text{ for a.e. } \omega \in \Omega,$$

using Remark 5.7(i). Moroever, since a w-convergent sequence is automatically bounded (by the Banach-Steinhaus theorem), Remark 5.7(ii) can be strengthened as follows [He]: If for a.e. ω there exists a τ-ball-compact subset of E which contains $(f_n(\omega))$, then

$$\tau\text{-seq-cl } w\text{-Ls}_n f_n(\omega) = \tau\text{-Ls}_n f_n(\omega) \text{ for a.e. } \omega \in \Omega.$$

□

[7] I.e., $\liminf_{n'} I_g(\epsilon_{f_{n'}} \otimes \epsilon_{n'}) \geq I_g(\delta_* \otimes \epsilon_\infty)$ in our usual notation.

As is sometimes illustrated by results in the literature that are related to Theorem 6.2, it is possible to use *ad hoc* variants of the tightness condition (6.1). Sometimes these are more general in appearance, but still turn out to be a version of Definition 4.1. The remark below directs attention to some instances where this occurs.

Remark 6.6 (versions of tightness) (*a*) For uniformly L^1-bounded (f_n) in $\mathcal{L}^1_E$ such that

$$(f_n(\omega)) \text{ is contained in a } \tau\text{-ball-compact set for a.e. } \omega,$$

we have that (6.1) is fulfilled for a sequentially τ-inf-compact integrand h, as required. Indeed, we can define

$$h(\omega, x) := \begin{cases} \|x\| & \text{if } x \in R(\omega), \\ +\infty & \text{otherwise,} \end{cases}$$

where $R(\omega)$ stands for the τ-ball-compact set containing $(f_n(\omega))$, used in the above hypothesis. Obviously, τ-ball-compactness of the set $R(\omega)$ is exactly equivalent to τ-inf-compactness of $h(\omega, \cdot)$, and this implies sequential τ-inf-compactness (see Remark 2.22 and Step 1 of the proof of Theorem 6.2). Also, $\sup_n I_h(\epsilon_{f_n}) = \sup_n \int_\Omega \|f_n(\omega)\| \mu(d\omega)$, so by uniform L^1-boundedness this implies the validity of (6.1).

(*b*) For uniformly L^1-bounded (f_n) in $\mathcal{L}^1_E$, the following condition is introduced in [Ca3, pp. 52-53]: For each $\epsilon > 0$ let there exist a multifunction $L_\epsilon : \Omega \to 2^E$, containing 0 and having τ-ball-compact values, [8] and a sequence (Ω^n_ϵ) of measurable subsets of Ω, such that for every n

$$f_n(\omega) \in L_\epsilon(\omega) \text{ for all } \omega \in \Omega^n_\epsilon \text{ and } \int_{\Omega \setminus \Omega^n_\epsilon} \|f_n\| d\mu \leq \epsilon.$$

To see that the above condition certainly implies (6.1), we construct h as follows (similar to the second part of the proof of Proposition 4.3). Denote L_ϵ corresponding to $\epsilon = 3^{-m}$ by L_m and assume without loss of generality that the sequence (L_m) is pointwise nondecreasing. Now we define

$$h(\omega, x) := \begin{cases} 2^m \|x\| & \text{if } x \in L_m(\omega) \setminus L_{m-1}(\omega),\ m \in \mathbf{N} \\ +\infty & \text{if } x \notin \cup_m L_m(\omega) \end{cases}$$

Then $h(\omega, \cdot)$ is τ-inf-compact on E for each ω and $\sup_n J_h(f_n) < +\infty$. The latter fact is immediate from the construction (observe that, by the second hypothesis, $\int_{A_n} \|f_n\| = 0$ for each n, where $A_n := \Omega \setminus \cup_m \Omega^n_{3^{-m}}$). To see the τ-inf-compactness, fix $\omega \in \Omega$ and $\beta \in \mathbf{R}$ and observe that when a sequence [9] (x_k) in E satisfies $h(\omega, x_k) \leq \beta$ for all k, then for each k there is m_k such that $\|x_k\| \leq 2^{-m_k}\beta$ and $x_k \in L_{m_k}(\omega) \setminus L_{m_k - 1}(\omega)$. Now (m_k) can be either bounded or unbounded. In the unbounded case there is a subsequence along which the x_k converge in norm to zero. In the bounded case, infinitely many x_k's belong to $L_{\tilde{m}}(\omega)$ for a certain index $\tilde{m}$. Since $\|x_k\| \leq \beta 2^{-\tilde{m}}$ for those x_k's, it is evident that (x_k) has a subsequence which τ-converges to some x_* in $L_{\tilde{m}}(\omega)$. Hence, both $\|x_*\| \leq \beta 2^{-\tilde{m}}$ and $h(\omega, x_*) \leq 2^{\tilde{m}} \|x_*\|$, which shows that $h(\omega, x_*) \leq \beta$. This proves that (6.1) holds.

(*c*) In [ACV, p. 174] a sequence (f_n) in $\mathcal{L}^1_E$ is defined to be $\mathcal{R}_\tau$-*tight* if for each $\epsilon > 0$ there exist a multifunction $L_\epsilon : \Omega \to 2^S$, having τ-ball-compact values, such that for every $\epsilon > 0$

$$\sup_n \mu(\{\omega \in \Omega : f_n(\omega) \notin L_\epsilon(\omega)\}) \leq \epsilon.$$

[8] In addition, these values are supposed to be convex in [Ca3].

[9] Sequential arguments are allowed to prove inf-compactness; for $\tau = w$ this is by virtue of the Eberlein-Šmulian theorem.

Now in all applications in [ACV] this notion goes hand in hand with the uniform L^1-boundedness hypothesis. This joint appearance makes $\mathcal{R}_\tau$-tightness into a special version of (6.1). To demonstrate this, we argue as follows: As before, denote by L_m the L_ϵ that corresponds to $\epsilon = 3^{-m}$, and assume without loss of generality that the sequence (L_m) is pointwise nondecreasing. This time, we define h by

$$h(\omega, x) := \begin{cases} 2^m + \|x\| & \text{if } x \in L_m(\omega)\backslash L_{m-1}(\omega),\ m \in \mathbf{N} \\ +\infty & \text{if } x \notin \cup_m L_m(\omega) \end{cases}$$

Then $h(\omega, \cdot)$ is τ-inf-compact on E for each ω and $\sup_n I_h(\epsilon_{f_n}) < +\infty$. Again, the latter inequality follows quite simply from the construction (see the proof of Proposition 4.3), in view of uniform L^1-boundedness. To check inf-compactness, fix $\omega \in \Omega$ and $\beta \in \mathbf{R}$ arbitrarily, and let (x_k) in E satisfy $h(\omega, x_k) \leq \beta$ for all k. Then for each k there exists m_k for which $2^{m_k} + \|x_k\| \leq \beta$ and $x_k \in L_{m_k}(\omega)\backslash L_{m_k-1}(\omega)$. Clearly, (m_k) is bounded, so infinitely many x_k's must belong to $L_{\tilde{m}}(\omega)$ for a certain index $\tilde{m}$. For those same x_k's we have $2^{\tilde{m}} + \|x_k\| \leq \beta$, so it is evident that (x_k) has a subsequence which τ-converges to some x_* in $L_{\tilde{m}}(\omega)$ (by τ-ball-compactness of the latter set). Therefore, $h(\omega, x_*) \leq 2^{\tilde{m}} + \|x_*\|$. Since $2^{\tilde{m}} + \|x_*\| \leq \beta$, we conclude that $h(\omega, x_*) \leq \beta$. So also in this case (6.1) holds. □

Next, we discuss some Dunford Pettis-like results which follow directly from Theorem 6.2. Recall [Ion, IV] that the dual of $\mathcal{L}^1_E$ is the space $\mathcal{L}^\infty_{E^*}[E]$ of all E-scalarly measurable bounded – i.e., for the dual norm – functions from Ω into E^*.

Definition 6.7 (weak convergence) A sequence (f_n) is said to converge *weakly* to a function f_0 in $\mathcal{L}^1_E$ (notation: $f_n \rightharpoonup f_0$) if

$$\lim_n \int_\Omega < f_n, b > d\mu = \int_\Omega < f_0, b > d\mu$$

for every $b \in \mathcal{L}^\infty_{E^*}[E]$. The corresponding topology $\sigma(\mathcal{L}^1_E, \mathcal{L}^\infty_{E^*}[E])$ on $\mathcal{L}^1_E$ is called the *weak* topology. □

Theorem 6.8 (weak compactness criterion) *Let (f_n) be in $\mathcal{L}^1_E$, such that $(\|f_n\|)$ is uniformly integrable in $\mathcal{L}^1_{\mathbf{R}}$ and such that there exists a nonnegative sequentially w-inf-compact integrand h on $\Omega \times E$ with*

$$\sup_n \int_\Omega^* h(\omega, f_n(\omega))\mu(d\omega) < +\infty.$$

Then Theorem 6.2 applies with $\tau = w$, and we have in addition in Theorem 6.2 that $f_{n'} \rightharpoonup f_*$.

Proof. By Theorem A.3, the uniform integrability hypothesis for $(\|f_n\|)$ implies the existence of a function $h' : \mathbf{R}_+ \to \mathbf{R}_+$, convex, continuous, nondecreasing, with $\lim_{\xi\to\infty} h'(\xi)/\xi = +\infty$ and $\sup_n \int_\Omega h'(\|f_n\|)d\mu < +\infty$. Define $\bar{h}(\omega, x) := h(\omega, x) + h'(\|x\|)$. It is easy to verify that $\bar{h}$ is a sequentially w-inf-compact integrand on $\Omega \times E$ and that (6.1) holds for $h := \bar{h}$. Of course, uniform L^1-boundedness holds by the uniform integrability. By Theorem 6.2, with $\tau := w$, there exist a subsequence $(f_{n'})$, a Young measure $\delta_* \in \mathcal{R}$ and its parametrized barycenter $f_* \in \mathcal{L}^1_E$ with the properties as stated in Theorem 6.2. For any $b \in \mathcal{L}^\infty_{E^*}[E]$ a w-continuous integrand g_b on $\Omega \times E$ is defined by $g_b(\omega, x) :=< x, b(\omega) >$. Clearly, $-|g_b|$ has property (γ) with respect to $\bar{h}$ (cf. Example 5.2), so a twofold application of (6.2), combined with the barycentric property of f_* pointwise, implies $\lim_{n'} J_{g_b}(f_{n'}) = I_{g_b}(\delta_*) = J_{g_b}(f_*)$. By Definition 6.7, this finishes the proof. □

An analogue of Theorem 6.8 turns out to hold if the w-topology is replaced by the s-topology; in a more rudimentary form, this result appears in [Jal] (see also [BGJ]).

Theorem 6.9 *Let (f_n) be in $\mathcal{L}^1_E$, such that $(\|f_n(\omega)\|)$ is uniformly integrable and such that there exists a nonnegative s-inf-compact integrand h on $\Omega \times E$ with*

$$\sup_n \int^*_\Omega h(\omega, f_n(\omega))\mu(d\omega) < +\infty.^{10}$$

Then Theorem 6.2 applies with $\tau = s$, and we have in addition in Theorem 6.2 that for every $A \in \mathcal{A}$

$$\| \int_A (f_{n'} - f_*)d\mu\| \to 0.$$

Proof. Let $\bar{h}$ be defined as in the proof of Theorem 6.8; this time, $\bar{h}$ is an s-inf-compact integrand on $\Omega \times E$. By Theorem 6.2, applied with $\tau := s$, there exist a subsequence $(f_{n'})$, a Young measure $\delta_* \in \mathcal{R}$ and its parametrized barycenter $f_* \in \mathcal{L}^1_E$ with the properties as stated in Theorem 6.2 and Remark 6.3. Fix $A \in \mathcal{A}$ arbitrarily. Let $\alpha := \limsup_{n'} \| \int_A (f_{n'} - f_*)d\mu\|$. There exists a subsequence $(f_{n''})$ of $(f_{n'})$ with $\alpha = \lim_{n''} \| \int_A (f_{n''} - f_*)d\mu\|$. By the Hahn-Banach theorem, for each n'' there exists $x^*_{n''}$ in U^* (the closed unit ball of E^*), such that

$$\| \int_A (f_{n''} - f_*)d\mu\| =< \int_A (f_{n''} - f_*)d\mu, x^*_{n''} >= \int_A < (f_{n''} - f_*), x^*_{n''} > d\mu.$$

As observed at the beginning of this section, U^* is compact and metrizable for $\sigma(E^*, E)$. So $(x^*_{n''})$ has a subsequence $(x^*_{n'''})$ which $\sigma(E^*, E)$-converges to some $x^*_\infty \in U^*$. We now apply Remark 6.3 to the lower semicontinuous integrand g on $\Omega \times (E_s \times \hat{\mathbf{N}})$ given by $g(\omega, x, k) := 1_A(\omega) < f_*(\omega) - x, x^*_k >$ (observe that g has property (γ) with respect to h). This gives

$$-\alpha = \lim_{n'''} \int_\Omega g(\cdot, f_{n'''}, n''')d\mu \geq \int_\Omega [\int_E g(\omega, x, \infty)\delta_*(\omega)(dx)]\mu(d\omega) \geq 0,$$

where the last inequality comes from $\int_E g(\omega, x, \infty)\delta_*(\omega)(dx) \geq g(\omega, f_*(\omega), \infty) = 0$, which is a consequence of Jensen's inequality (for $g(\omega, \cdot, \infty)$ is convex). Since also $\alpha \geq 0$, we conclude $\alpha = 0$. □

Theorems 6.8–6.9 can be extended by replacing the uniform integrability hypothesis by a uniform $\mathcal{L}^1_E$-boundedness condition.

Definition 6.10 **(w^2-convergence)** A sequence (f_n) is said to *w^2-converge* to a function f_0 in $\mathcal{L}^1_E$ (notation: $f_n \overset{w^2}{\to} f_0$) if there exists a sequence (A_p) in $\mathcal{A}$, $\mu(\Omega \backslash A_p) \to 0$, such that for every $p \in \mathbf{N}$ and every $b \in \mathcal{L}^\infty_{E^*}[E]$

$$\lim_n \int_{A_p} < f_n, b > d\mu = \int_{A_p} < f_0, b > d\mu$$

(i.e., for every p, $f_n |_{A_p} \rightharpoonup f_0 |_{A_p}$). □

[10] In the terminology of section 4: (ϵ_{f_n}) is tight in $\mathcal{R}(\Omega, \mathcal{A}, \mu; E_s)$.

The following sequential relative compactness criterion for w^2-convergence in $\mathcal{L}^1_{\mathbf{R}}$ is due to V.F. Gaposhkin [Gap]. It has been rediscovered by several authors (e.g., see [BC, Sl]), and its proof is rather elementary.

Theorem 6.11 (Gaposhkin's biting theorem) *Let (ϕ_n) be in $\mathcal{L}^1_{\mathbf{R}}$ such that*

$$\sup_n \int_\Omega |\phi_n| d\mu < +\infty.$$

Then there exists a sequence (A_p) in $\mathcal{A}$, $\mu(\Omega \backslash A_p) \to 0$, such that

$$(|\phi_n|) \text{ is uniformly integrable over } A_p \text{ for each } p.$$

In particular, there exist a subsequence $(\phi_{n'})$ and a function ϕ_* such that $\phi_{n'} \overset{w^2}{\to} \phi_*$ in $\mathcal{L}^1_{\mathbf{R}}$.

It is easy to see from Examples 4.8–4.9 (applied to $f_n := \phi_n$) that the function ϕ_* in Theorem 6.11 coincides a.e. with the parametrized barycenter of the Young measure δ_* in those examples. It is also evident that weak convergence in $\mathcal{L}^1_E$ implies w^2-convergence (take $A_p \equiv \Omega$) and that the two notions coincide in the presence of uniform integrability. By Theorem 6.11, the following result is obtained as a direct modification of Theorem 6.8

Theorem 6.12 *Let (f_n) be uniformly L^1-bounded in $\mathcal{L}^1_E$ and such that there exists a non-negative sequentially w-inf-compact integrand h on $\Omega \times E$ with*

$$\sup_n \int_\Omega^* h(\omega, f_n(\omega))\mu(d\omega) < +\infty.$$

Then Theorem 6.2 applies with $\tau = w$, and we have in addition in Theorem 6.2 that $f_{n'} \overset{w^2}{\to} f_*$.

Next, we study the role which Young measures can play to differentiate between strong and weak convergence. Recall first the formal definition of strong convergence in $\mathcal{L}^1_E$:

Definition 6.13 (strong convergence) A sequence (f_n) in $\mathcal{L}^1_E$ is said to converge *strongly* to a function f_0 in $\mathcal{L}^1_E$ (notation: $f_n \overset{L^1}{\to} f_0$) if

$$\lim_n \int_\Omega \|f_n(\omega) - f_0(\omega)\| \mu(d\omega) = 0.$$

□

It is well-known that the difference between strong and weak convergence in $\mathcal{L}^1_{\mathbf{R}}$ lies in the nonoccurrence (respectively occurrence) of persistent – i.e. continuing 'in the limit' – oscillations. For instance, see [DM, II.26] for an elementary result of this nature. Likewise, we can expect that in $\mathcal{L}^1_E$ weak convergence will imply strong convergence if the (pointwise) oscillations around the limit function can be restricted. One such restriction, which is also necessary for strong convergence, is due to M. Girardi [Gi]:

Definition 6.14 (Bocce criterion) A sequence (f_n) in $\mathcal{L}^1_E$ is said to fulfill the *Bocce criterion* if for every $\epsilon > 0$, every $A \in \mathcal{A}$, $\mu(A) > 0$, and every subsequence $(f_{n'})$ of (f_n) there exists $B \subset A$, $\mu(B) > 0$, such that

$$\liminf_{n'} \int_B \|f_{n'}(\omega) - m_B(f_{n'})\| \mu(d\omega) \leq \epsilon \mu(B).$$

Here $m_B(f_n) := \frac{1}{\mu(B)} \int_B f_n \, d\mu$. □

Theorem 6.15 (Girardi) *Let (f_n) be in $\mathcal{L}^1_E$. The following are equivalent:*

(*a*) $f_n \rightharpoonup f_0$, *(ϵ_{f_n}) is tight in $\mathcal{R}(\Omega, \mathcal{A}, \mu; E_s)$ and (f_n) fulfills the Bocce criterion.*

(*b*) $f_n \xrightarrow{L^1} f_0$.

The following are also equivalent, provided that $\sup_n \int_\Omega \|f_n\|\, d\mu < +\infty$*:*

(*a′*) $f_n \xrightarrow{w^2} f_0$, *$(\epsilon_{f_n})$ is tight in $\mathcal{R}(\Omega, \mathcal{A}, \mu; E_s)$ and (f_n) fulfills the Bocce criterion.*

(*b′*) $f_n \xrightarrow{\mu} f_0$.

Lemma 6.16 *Let $\phi_* \in \mathcal{L}^1_{\mathbf{R}}$, $\phi_* \geq 0$, be such that for every $A \in \mathcal{A}$, $\mu(A) > 0$, and every $\epsilon > 0$ there exists $B \in \mathcal{A}$, $B \subset A$ and $\mu(B) > 0$, such that*

$$\int_B \phi_*\, d\mu \leq \epsilon \mu(B).$$

Then $\phi_* = 0$ a.e.

Proof. If for any $\alpha > 0$ the set $A := \{\omega \in \Omega : \phi_*(\omega) \geq \alpha\}$ would have positive measure, then the subset B of A, corresponding to $\epsilon := \frac{1}{2}\alpha$ would give $\alpha\mu(B) \leq \int_B \phi_* \leq \epsilon\mu(B)$. This produces a contradiction. □

Proof of Theorem 6.15. $(a) \Rightarrow (b)$: It is enough to prove that every subsequence of (f_n) – and without loss of generality we take (f_n) itself for this – has a further subsequence which converges strongly to f_0. The hypothesized weak convergence $f_n \rightharpoonup f_0$ implies the uniform integrability of $(\|f_n\|)$ [BD, Theorem 4.1]. Hence, we may invoke Theorem 6.2 (with $\tau = s$). Let $(f_{n'})$ and $\delta_* \in \mathcal{R}$ be as assured by Theorem 6.2. Since s-tightness implies *a fortiori* w-tightness, Theorems 6.8 and 6.9 both apply to $(f_{n'})$. Theorem 6.8 gives that $f_{n'} \rightharpoonup f_*$, with f_* the parametrized barycenter of δ_*. Since already $f_{n'} \rightharpoonup f_0$ by hypothesis, this implies that $f_* = f_0$ a.e. Hence, Theorem 6.9 gives $\|m_B(f_{n'}) - m_B(f_0)\| \to 0$ for every $B \in \mathcal{A}$ with $\mu(B) > 0$. Next, observe that the above also yields the uniform integrability of $(\|f_{n'} - f_0\|)$. Hence, there exists a subsequence $(f_{n''})$ of $(f_{n'})$ and $\phi_* \in \mathcal{L}^1_{\mathbf{R}}$ such that $\|f_{n''} - f_0\| \rightharpoonup \phi_*$ in $\mathcal{L}^1_{\mathbf{R}}$. Let $A \in \mathcal{A}$ and $\epsilon > 0$ be arbitrary, with $\mu(A) > 0$. Let $B \subset A$ be as in Definition 6.14, corresponding to A, $\epsilon/2$ and $(f_{n''})$. Then

$$\epsilon\mu(B)/2 \geq \liminf_{n''} \int_B \|f_{n''} - f_0\|\, d\mu - \int_B \|f_0 - m_B(f_0)\|\, d\mu = \int_B \phi_*\, d\mu - \int_B \|f_0 - m_B(f_0)\|\, d\mu,$$

by the triangle inequality and $\|m_B(f_{n''}) - m_B(f_0)\| \to 0$. Approximating f_0 by simple functions and applying Egorov's theorem, we see easily that ϕ_* meets the conditions of Lemma 6.16. So $\phi_* = 0$ a.e., and this gives $\int_\Omega \|f_{n''} - f_0\|\, d\mu \to 0$, i.e. $f_{n''} \xrightarrow{L^1} f_0$.

$(b) \Rightarrow (a)$: Both weak convergence and the Bocce criterion hold elementarily. As for the s-tightness of (ϵ_{f_n}), observe that the hypothesis implies $\epsilon_{f_n} \Longrightarrow \epsilon_{f_0}$ by Proposition 3.21. Therefore, $\mu \otimes \epsilon_{f_n}/\mu(\Omega) \Rightarrow \mu \otimes \epsilon_{f_0}/\mu(\Omega)$ in $\mathcal{P}(E_s)$. Since E_s is complete, it follows that $(\mu \otimes \epsilon_{f_n}/\mu(\Omega))$ is tight by the converse Prohorov theorem [Bi1, Theorem 6.2]; so by Remark 4.6, (ϵ_{f_n}) is s-tight.

$(a') \Rightarrow (b')$: By Definition 6.10, the proof is a variant of the proof of $(a) \Rightarrow (b)$ (replace Ω by any set A_p).

$(b') \Rightarrow (a')$: By Theorems 6.11 and 6.12, the proof is a variant of the above proof of the implication $(b) \Rightarrow (a)$. □

Observe that the Bocce criterion restricts oscillations by locally forcing certain first moments to go to zero. A quite different way to restrict oscillations is to require that the values of the limit function are in some sense extreme with respect to the values taken by the functions of the sequence [Vi]. To describe such results, the following convergence mode, introduced in [Ba9], is useful. Observe that it is not stronger than strong convergence and not weaker than weak convergence in $\mathcal{L}^1_E$.

Definition 6.17 (limited convergence) A sequence (f_n) in $\mathcal{L}^1_E$ is said to converge *limitedly* to a function f_0 in $\mathcal{L}^1_E$ (notation: $f_n \overset{lim}{\to} f_0$) if

$$\lim_n \int_\Omega g(\omega, f_n(\omega) - f_0(\omega))\mu(d\omega) = 0$$

for every integrand g on $\Omega \times E$ such that

(i) $g(\omega, 0) = 0$ for every $\omega \in \Omega$.

(ii) $g(\omega, \cdot)$ is w-continuous on E for every $\omega \in \Omega$.

(iii) $|g(\omega, x)| \leq C\|x\| + \phi(\omega)$ on $\Omega \times E$ for some $C \geq 0$ and $\phi \in \mathcal{L}^1_{\mathbf{R}}$.

(iv) $g(\cdot, f(\cdot))$ is $\mathcal{A}_\mu$-measurable for every $f \in \mathcal{L}^1_E$.

□

Here $\mathcal{A}_\mu$ stands for the μ-completion of the σ-algebra $\mathcal{A}$. Clearly, limited convergence implies weak convergence in $\mathcal{L}^1_E$ (use $g_b(\omega, x) := < x, b(\omega) >$, $b \in \mathcal{L}^\infty_{E^*}[E]$). When $E = \mathbf{R}^d$, one can take $g(\omega, x) := |x|$ (Euclidean norm) to see that limited and strong convergence in $\mathcal{L}^1_{\mathbf{R}^d}$ coincide (observe that in infinite dimensions this is no longer possible, since $x \mapsto \|x\|$ is just w-lower semicontinuous). To see precisely how far limited convergence is removed from strong convergence, the following equivalence, due to V. Jalby [Jal, BGJ, PV], is relevant (since the proof is rather technical, [11] we advise the reader that this equivalence result plays no further role below).

Proposition 6.18 (strong is limited convergence plus tightness) *Let* (f_n) *and* f_0 *be in* $\mathcal{L}^1_E$. *The following are equivalent:*

(a) $f_n \overset{lim}{\to} f_0$ *and* (ϵ_{f_n}) *is tight in* $\mathcal{R}(\Omega, \mathcal{A}, \mu; E_s)$.

(b) $f_n \overset{L^1}{\to} f_0$.

Proof. $(a) \Rightarrow (b)$: As noticed above, $f_n \overset{lim}{\to} f_0$ implies $f_n \rightharpoonup f_0$. So $(\|f_n - f_0\|)$ is uniformly integrable [BD, Theorem 4.1]. Hence, by de la Vallée Poussin's Theorem A.3 there exists a function $h' : \mathbf{R}_+ \to \mathbf{R}_+$, convex, continuous and nondecreasing, such that $\lim_{\xi\to\infty} h'(\xi)/\xi = +\infty$ and

$$\sigma' := \sup_n \int_\Omega h'(\|f_n - f_0\|)\mu(d\omega) < +\infty.$$

By completeness of $(E, \|\cdot\|)$, the given parametrized tightness of (ϵ_{f_n}) is equivalent to tightness of the sequence $([\mu \otimes \epsilon_{f_n}](\Omega \times \cdot)/\mu(\Omega))$ in $\mathcal{P}(E_s)$ in the sense of Definition 2.11 (cf. Remark 4.6). Therefore, an s-inf-compact function $h : E \to [0, +\infty]$, exists such that

$$\sigma := \sup_n \int_\Omega h(f_n)\, d\mu < +\infty.$$

We define for $\epsilon > 0$

$$g_\epsilon(\omega, x) := -\|x\| + \epsilon h(x + f_0(\omega)) + \epsilon h'(\|x\|).$$

[11] Actually, it contains the conerstones of an approximation of measurable lower semicontinuous integrands that, until now, we have managed to avoid in these lectures; cf. [Ba3, Ba11, Ba14].

Set $\gamma_\epsilon := \inf_{\xi\in\mathbf{R}}[\xi + \epsilon h'(|\xi|)]$; this is a finite real number by the superlinear growth of h'. The function $g_\epsilon(\omega,\cdot)$ is s-inf-compact for every $\omega \in \Omega$ (it is clearly s-lower semicontinuous, and for each $\beta \in \mathbf{R}$ the set of all $x \in E$ with $g_\epsilon(\omega, x) \leq \beta$ is contained the set of all $x \in E$ for which $h(x + f_0(\omega)) \leq (\beta - \gamma_\epsilon)/\epsilon$). This implies that $g_\epsilon(\omega,\cdot)$ is also d_w-inf-compact, and *a fortiori* d_w-lower semicontinuous. We now define for $m \in \mathbf{N}$

$$g_\epsilon^m(\omega, x) := \inf_{x'\in E}[md_w(x, x') + g_\epsilon(\omega, x')].$$

Then for each $\omega \in \Omega$ the function $g_\epsilon^m(\omega,\cdot)$ is d_w-continuous – whence w-continuous – by the triangle inequality for d_w. By [As, A6.6] we have $\lim_m g_\epsilon^m = g_\epsilon$ pointwise on $\Omega \times E$ for each $\epsilon > 0$, the limit being monotone. Truncation from above does not essentially alter this: we also have $\tilde{g}_\epsilon^m \uparrow g_\epsilon$ on E, where $\tilde{g}_\epsilon^m := \min(g_\epsilon^m, m)$. It is now easy to see that $\bar{g}_\epsilon^m$, defined by $\bar{g}_\epsilon^m(\omega, x) := \tilde{g}_\epsilon^m(\omega, x) - \tilde{g}_\epsilon^m(\omega, 0)$ meets (*i*)–(*iii*) of Definition 6.17. Also, (*iv*) of that definition holds by combining III.39 (a result which comes essentially from applying a celebrated projection theorem) and III.14 of [CV]. By applying Definition 6.17 to the $\bar{g}_\epsilon^m$ and applying the monotone convergence theorem to the $\tilde{g}_\epsilon^m$, it follows that

$$\liminf_n \int_\Omega g_\epsilon(\omega, f_n(\omega) - f_0(\omega))\mu(d\omega) \geq \int_\Omega g_\epsilon(\omega, 0)\mu(d\omega) \geq 0.$$

For $\alpha := \limsup_n \int_\Omega \|f_n - f_0\|\, d\mu$ this immediately implies $-\alpha + \epsilon\sigma + \epsilon\sigma' \geq 0$, and letting ϵ go to zero gives $\alpha \leq 0$.

(*b*) $\Rightarrow$ (*a*): *A fortiori*, strong convergence gives $f_n \overset{\mu}{\to} f_0$. So by Proposition 3.21 we get $\epsilon_{f_n} \Longrightarrow \epsilon_{f_0}$ in $\mathcal{R}(\Omega, \mathcal{A}, \mu; E_s)$. So $[\mu \otimes \epsilon_{f_n}](\Omega \times \cdot)/\mu(\Omega) \Rightarrow [\mu \otimes \epsilon_{f_0}](\Omega \times \cdot)/\mu(\Omega)$ in $\mathcal{P}(E_s)$. So Theorem 2.15 and Definition 4.1 imply that (ϵ_{f_n}) is tight in $\mathcal{R}(\Omega, \mathcal{A}, \mu; E_s)$. Secondly, the stated limited convergence follows immediately by the Lebesgue-Vitali theorem, since the sequence $(\|f_n\|)$ is uniformly integrable. □

The following result was originally given by A. Visintin [Vi] and subsequently refined and extended in [Ba9, Ba17]; see also [Val4].

Theorem 6.19 (Visintin) *Let (f_n) be in $\mathcal{L}_E^1$, such that*

$$f_n \rightharpoonup f_0 \text{ in } \mathcal{L}_E^1$$

and such that there exists a nonnegative w-inf-compact integrand h on $\Omega \times E$ with

$$\sup_n \int_\Omega^* h(\omega, f_n(\omega))\mu(d\omega) < +\infty.$$

Suppose that

$$f_0(\omega) \text{ is extreme point of } \mathrm{cl}\ \mathrm{co}\ w\text{-seq-cl}\ w\text{-}\mathrm{Ls}_n f_n(\omega) \text{ for a.e. } \omega \text{ in } \Omega.$$

Then $f_n \overset{lim}{\to} f_0$. Moreover, if $E = \mathbf{R}^d$ then already the weaker condition

$$f_0(\omega) \text{ is extreme point of } \mathrm{co}\ w\text{-}\mathrm{Ls}_n f_n(\omega) \text{ for a.e. } \omega \text{ in } \Omega.$$

implies $f_n \overset{L^1}{\to} f_0$.

Lemma 6.20 *Let $\nu \in \mathcal{P}(E)$ be such that* bar ν *exists and is an extreme point of* cl co supp ν*. Then ν is Dirac. Moreover, if $E = \mathbf{R}^d$ then the condition that* bar ν *exists and is an extreme point of* co supp ν *is weaker and implies that ν is Dirac.*

PROOF. Denote $x_0 := \text{bar } \nu$. We claim that ν is the Dirac measure at the point x_0. Let us consider the following implication for any $B \in \mathcal{B}(E)$:

$$x_0 \notin B \text{ implies } \nu(B) = 0. \tag{6.3}$$

Then (6.3) certainly holds when B is closed and convex. For suppose not. First of all, then $\nu(B)$ cannot be 1 (or else $x_0 \in B$ would give a contradiction). So we may define $\nu_1 := \nu/\nu(B)$ and $\nu_2 := \nu/(1-\nu(B))$. Then the decomposition $\nu = \nu(B)\nu_1 + (1-\nu(B))\nu_2$ would give $x_0 = \nu(B)\text{bar } \nu_1 + (1-\nu(B))\text{bar } \nu_2$, whence $x_0 = \text{bar } \nu_1 = \text{bar } \nu_2$. But $\text{bar } \nu_1 \in B$, so this gives contradiction. So all closed balls B satisfy (6.3), and hence also every s-open ball B. Since every s-open set of E is the countable union of s-open balls, it follows that (6.3) holds for every s-open set, and, in particular, for the complement of x_0. Finally, the statement for the case $E = \mathbf{R}^d$ follows immediately from the above by Theorem A.13(*ii*). □

PROOF of Theorem 6.19. It is enough to prove that every subsequence of (f_n) – and without loss of generality we take (f_n) itself for this – has a further subsequence which converges limitedly to f_0. Let $\bar{h}$ be defined as in the proof of Theorem 6.8. Let $(f_{n'})$ and $\delta_* \in \mathcal{R}$ be as in Theorem 6.2 (for $\tau := w$), with the parametrized barycenter $f_* := \text{bar } \delta_*$, such that $\text{supp } \delta_*(\omega) \subset w\text{-seq-cl } w\text{-Ls}_n f_n(\omega)$ for a.e. ω, and also with and $f_{n'} \rightharpoonup f_*$, by Theorem 6.8. So here $f_* := \text{bar } \delta_* = f_0$ a.e. By the given extremality condition, it is clear that pointwise the condition of Lemma 6.20 holds. We conclude that $\delta_*(\omega) = \epsilon_{f_0}(\omega)$ for a.e. ω. It remains to apply (6.2) to the w-continuous integrand $(\omega, x) \mapsto g(w, x - f_0(\omega))$, where g is as in Definition 6.17 (this integrand has property (γ) with respect to $\bar{h}$ by (*iii*) of Definition 6.17). □

A counterexample given in [Vi], combined with Proposition 6.18, suggests that Theorem 6.19 can only yield strong convergence by requiring the integrand h to be s-inf-compact; see [Ca2] for such a result. Interesting applications of Theorem 6.19 are so-called well-posedness results for integral functionals on an L^1-space; see [Vi, Ba17]:

Theorem 6.21 (existence and well-posedness) *Let g_0 be a sequentially w-inf-compact integrand on $\Omega \times E$ which is $\mathcal{A} \otimes \mathcal{B}(S)$-measurable and such that*

$$g_0(\omega, \cdot) \textit{ is strictly convex on } E \textit{ for every } \omega \in \Omega,$$

$$g_0(\omega, x) \geq h'(\|x\|) + \phi(\omega),$$

where $h' : \mathbf{R}_+ \to \mathbf{R}_+$ is convex, lower semicontinuous and nondecreasing, $\lim_{\xi\to\infty} h'(\xi)/\xi = +\infty$*, and where $\phi \in \mathcal{L}^1_{\mathbf{R}}$. Then the optimization problem*

$$(\mathcal{P}) : \inf_{f \in \mathcal{L}^1_E} J_{g_0}(f)$$

has an essentially unique optimal solution $f_* \in \mathcal{L}^1_E$. Moreover, if $\inf(\mathcal{P}) < +\infty$*, then every minimizing sequence (f_n) of $\mathcal{P})$ satisfies*

$$f_n \overset{lim}{\to} f_* \textit{ and } \lim_n \int_\Omega |g_0(\omega, f_n(\omega)) - g_0(\omega, f_*(\omega))|\mu(d\omega) = 0.$$

PROOF. We may suppose without loss of generality that $\inf(\mathcal{P}) < +\infty$. Let (f_n) be a minimizing sequence. It is simple to see that the conditions of Theorem 6.2 hold for $\tau := w$, given the properties of the integrand g_0. Hence, applying Jensen's inequality pointwise, there exist a subsequence $(f_{n'})$, a Young measure δ_* and its parametrized barycenter $f_* \in \mathcal{L}^1_E$ for which $J_{g_0}(f_*) \leq \liminf_{n'} J_{g_0}(f_{n'}) = \inf(\mathcal{P})$. This implies $J_{g_0}(f_*) = \inf(\mathcal{P})$. The essential uniqueness of the optimal solution f_* follows from the given strict convexity by a standard argument. Next, suppose that (f_m) is an arbitrary minimizing sequence, and let (f_n) be any subsequence of (f_m). Then (f_n) is also minimizing, and the above implies the existence of a subsequence $(f_{n'})$ and of $f_{**} \in \mathcal{L}^1_E$ such that $f_{n'} \rightharpoonup f_{**}$ (by Theorem 6.8). As above, it follows that f_{**} is an optimal solution of $(\mathcal{P})$, and by the essential uniqueness of the optimal solution this implies $f_{**} = f_*$ a.e. Now (6.2) applies to the integrands $g_A(\omega, x) := 1_A(\omega) g_0(\omega, x)$, $A \in \mathcal{A}$. Since f_* is the parametrized barycenter of δ_*, this gives $\liminf_{n'} J_{g_A}(f_{n'}) \geq J_{g_A}(f_*)$ for every $A \in \mathcal{A}$. Together with $J_{g_0}(f_{n'}) \to \inf(\mathcal{P})$, this implies that $(g_0(\cdot, f_{n'}(\cdot)))$ converges weakly in $\mathcal{L}^1_{\mathbf{R}}$ to $g_0(\cdot, f_*(\cdot))$. So we get $(f_{n'}, g_0(\cdot, f_{n'}(\cdot))) \rightharpoonup (f_*, g_0(\cdot, f_*(\cdot)))$. Also, the strict convexity of $g_0(\omega, \cdot)$ implies that $(f_*(\omega), g_0(\omega, f_*(\omega)))$ is an extreme point of the epigraph of the function $g_0(\omega, \cdot)$ for every $\omega \in \Omega$. Hence, we may apply Theorem 6.19, which gives $\int_\Omega |g_0(\omega, f_{n'}(\omega)) - g_0(\omega, f_*(\omega))| \mu(d\omega) \to 0$ for the first coordinate (which is one-dimensional – see the comments following Definition 6.17) and $f_{n'} \stackrel{lim}{\to} f_*$ for the second one. Since the subsequence (f_n) was taken arbitrarily from (f_m), we conclude that $(g_0(\cdot, f_m(\cdot)))$ converges strongly in $\mathcal{L}^1_{\mathbf{R}}$ to the function $g_0(\cdot, f_*(\cdot))$ and that $f_m \stackrel{lim}{\to} f_*$. $\square$

7 Applications to semicontinuity and lower closure in L^1

Contents: strong-weak lower semicontinuity, strong-weak lower closure, existence of optimal distributed and ordinary controls

In this section we shall derive generalizations of classical strong-weak lower semicontinuity annex lower closure results for integral functionals. Actually, all these results are direct consequences of Theorem 6.2, which acts as the L^1-manifestation of the synthesis Theorem 5.5. As before, let $(\Omega, \mathcal{A}, \mu)$ be a finite measure space and $(E, \|\cdot\|)$ a separable Banach space.

Theorem 7.1 (strong-weak lower semicontinuity) *Let (f_n) and f_0 be in $\mathcal{L}^1_E$, such that (f_n) is uniformly L^1-bounded, $f_n \stackrel{w^2}{\to} f_0$, and such that there exists a nonnegative sequentially w-inf-compact integrand h on $\Omega \times E$ with*

$$\sup_n \int_\Omega^* h(\omega, f_n(\omega))\mu(d\omega) < +\infty.$$

Then

$$\liminf_n \int_\Omega^* g(\omega, f_n(\omega), n)\mu(d\omega) \geq \int_\Omega^* g(\omega, f_0(\omega), \infty)\mu(d\omega)$$

for every sequentially lower semicontinuous integrand g on $\Omega \times (E_w \times \hat{\mathbf{N}})$ such that

$$g(\omega, \cdot, \infty) \text{ is convex on } E \text{ for every } \omega \in \Omega$$

and such that g has property (γ') with respect to h and (ϵ_{f_n}).

PROOF. It is enough to prove that every subsequence of (f_n) – which we take to be (f_n) itself without loss of generality – contains a further subsequence for which the stated lower semicontinuity inequality holds. The conditions for (f_n) entail that Theorem 6.2 applies; let $(f_{n'})$, $\delta_* \in \mathcal{R}$ and f_* be as asserted in that theorem. By Theorem 6.12, we have in addition that $f_{n'} \stackrel{w^2}{\to} f_*$ for the parametrized barycenter f_* of δ_*. Hence, it follows here that $f_* = f_0$ a.e. by essential uniqueness of w^2-limits. Therefore, the desired inequality follows directly from Remark 6.3 and a simple pointwise application of Jensen's inequality. □

The above formulation of the lower semicontinuity result is general enough to deal with strong modes of convergence in an additional functional variable; this results in a quite general lower semicontinuity result for weak-strong convergence:

Corollary 7.2 (strong-weak lower semicontinuity) *Let (f_n) and f_0 be in $\mathcal{L}^1_E$, such that (f_n) is uniformly L^1-bounded, $f_n \stackrel{w^2}{\to} f_0$, and*

$$(f_n(\omega)) \text{ is contained in a } w\text{-ball-compact set for a.e. } \omega \text{ in } \Omega.$$

Also, let $\bar{S}$ be a metric space and let $(\bar{f}_n)$, $\bar{f}_0$ in $\mathcal{L}^0_{\bar{S}}$ be such that $\bar{f}_n \stackrel{\mu}{\to} \bar{f}_0$. Then

$$\liminf_n \int_\Omega^* \ell(\omega, f_n(\omega), \bar{f}_n(\omega))\mu(d\omega) \geq \int_\Omega^* \ell(\omega, f_0(\omega), \bar{f}_0(\omega))\mu(d\omega)$$

for every lower semicontinuous integrand ℓ on $\Omega \times (E_w \times \bar{S})$ such that

$$\ell(\omega, \cdot, \bar{f}_0(\omega)) \text{ is convex on } E \text{ for every } \omega \in \Omega,$$

provided that there exists a uniformly integrable sequence (ϕ_n) *in* $\mathcal{L}^1_{\mathbf{R}}$ *for which for every* $n \in \mathbf{N}$

$$\ell(\omega, f_n(\omega), \bar{f}_n(\omega)) \geq \phi_n(\omega) \text{ for all } \omega \in \Omega.$$

PROOF. It is enough to prove that every subsequence of (f_n) – which we take to be (f_n) itself without loss of generality – contains a further subsequence for which the stated lower semicontinuity inequality holds. Define $\alpha := \liminf_n \int_\Omega^* \ell(\cdot, f_n, \bar{f}_n)d\mu$; then there exists a subsequence $(f_{n'}, \bar{f}_{n'})$ such that $\alpha = \lim_{n'} \int_\Omega^* \ell(\cdot, f_{n'}, \bar{f}_n)d\mu$. Since convergence in measure implies a.e. convergence of subsequences [N, Proposition II.4.3], there exist a further subsequence $(\bar{f}_{n''})$ of $(\bar{f}_{n'})$ and a null set M such that $\bar{f}_{n''}(\omega) \to \bar{f}_0(\omega)$ for all $\omega \in \Omega \backslash M$. Define g by

$$g(\omega, x, k) := \begin{cases} 1_{\Omega\backslash M}(\omega)\ell(\omega, x, \bar{f}_k(\omega)) & \text{if } k < \infty \\ 1_{\Omega\backslash M}(\omega)\ell(\omega, x, \bar{f}_0(\omega)) & \text{if } k = \infty \end{cases}$$

Then g is a sequentially lower semicontinuous integrand on $\Omega \times (E_w \times \hat{\mathbf{N}})$ having property (γ') with respect to h and (ϵ_{f_n}). Now apply Theorem 7.1, observing Remarks 6.6(a), 5.4 and 6.3; this gives the result. □

As observed earlier, if E is reflexive (in particular, when E is finite dimensional) then the above condition involving w-ball-compactness holds *automatically*, for then E itself is w-ball-compact. We illustrate the usefulness of Corollary 7.2 by an application to an existence problem for a variational problem. By use of the compactness of various Green operators listed in [ET, VIII.3], more existence results of this kind can easily be constructed.

Example 7.3 (existence of optimal distributed controls) Let $\Omega \subset \mathbf{R}^m$ be very regular open and bounded. Consider the optimization problem

$$(\mathcal{P}): \quad \inf_{f \in L^2_{\mathbf{R}}(\Omega)} J(f) := \int_\Omega^* g_0(\omega, f(\omega), y_f(\omega))d\omega,$$

Here $y_f \in H^1_0(\Omega)$ is the solution of the partial differential equation

$$\Delta y(\omega) = f(\omega) \text{ a.e.}$$

Also, $g_0 : \Omega \times \mathbf{R}^2 \to [0, +\infty]$ is a lower semicontinuous integrand on $\Omega \times \mathbf{R}^2$ such that $g_0(\omega, \cdot, y)$ is convex for every ω and y. We suppose also that there exists $\gamma > 0$ such that $g_0(\omega, x, y) \geq \gamma|x|^2$ for all y and x. Then $(\mathcal{P})$ has an optimal solution. Indeed, without loss of generality we may suppose that $\iota := \inf(\mathcal{P}) < +\infty$. Let (f_n) be a minimizing sequence in $L^2_{\mathbf{R}}$. Then $(\int_\Omega |f_n|^2)$ is evidently bounded. Since the injection of the Sobolev space $H^1_0(\Omega)$ into $L^2_{\mathbf{R}}$ is compact [ET, Example 1, p. 252], the mapping $f \mapsto y_f$ from $L^2_{\mathbf{R}}$ into $L^2_{\mathbf{R}}$ is compact. By two successive extractions there exist a subsequence $(f_{n'})$ of (f_n) and f_* in $L^2_{\mathbf{R}}$ with the following properties: $f_{n'} \rightharpoonup f_*$ in $\sigma(L^2_{\mathbf{R}}, L^2_{\mathbf{R}})$, and $y_{f_{n'}} \to y_{f_*}$ in the L^2-norm. By boundedness of Ω, we have *a fortiori* $f_{n'} \rightharpoonup f_*$ (i.e., in $\sigma(L^1_{\mathbf{R}}, L^\infty_{\mathbf{R}})$), whence $f_{n'} \overset{w^2}{\to} f_*$. Also, $(y_{f_{n'}})$ converges *a fortiori* in measure to y_{f_*}. So Corollary 7.2 gives $\iota = \lim_n J(f_n) = \lim_{n'} J(f_{n'}) \geq J(f_*)$. Hence, $\inf(\mathcal{P})$ is attained by f_*. □

Because of our use of w^2-convergence, lower closure results for strong-weak convergence of the general kind considered in Corollary 7.2 also follow from Theorem 7.1:

Corollary 7.4 (strong-weak lower closure) *Let $(E', \|\cdot\|')$ be another separable Banach space. Let (f_n) and f_0 be in $\mathcal{L}^1_E$ and let (f'_n) be in $\mathcal{L}^1_{E'}$, $f_n \overset{w^2}{\to} f_0$, such that*

$$\sup_n \int_\Omega (\|f_n(\omega)\| + \|f'_n(\omega)\|')\mu(d\omega) < +\infty,$$

$(f_n(\omega), f'_n(\omega))$ *is contained in a w-ball-compact subset of $E \times E'$ for a.e. ω in Ω.*

Then there exist a subsequence $(f_{n'}, f'_{n'})$ of (f_n, f'_n) and $f'_* \in \mathcal{L}^1_{E'}$ such that

$$\liminf_{n'} \int_\Omega^* g(\omega, f_{n'}(\omega), f'_{n'}(\omega), n')\mu(d\omega) \geq \int_\Omega^* g(\omega, f_0(\omega), f'_*(\omega), \infty)\mu(d\omega)$$

for every lower semicontinuous integrand g on $\Omega \times (E_w \times E'_w \times \hat{\mathbf{N}})$ such that

$$g(\omega, \cdot, \cdot, \infty) \text{ is convex on } E \times E' \text{ for every } \omega \in \Omega,$$

and such that g has property (γ') with respect to h and $(\epsilon_{(f_n, f'_n)})$.

Proof. By Remarks 6.6(a) and 5.4, Theorem 6.12 applies to (f'_n). Therefore, there exist a subseqence $(f'_{n'})$ of (f'_n) and $f'_* \in \mathcal{L}^1_{E'}$ such that $f'_{n'} \overset{w^2}{\to} f'_*$. Now $(f_{n'}, f'_{n'}) \overset{w^2}{\to} (f_0, f'_*)$ in $\mathcal{L}^1_{E \times E'}$, so Theorem 7.1 applies and the result follows. □

Example 7.5 (existence of optimal controls) In Example 5.12 we already showed that under a convexity condition for the orientor field Q, introduced in that same example, the original problem $(\mathcal{P})$ had an optimal solution. We did so via an existence result for the Young relaxation $(\mathcal{P}_{rel})$ of the problem $(\mathcal{P})$, obtained in Example 5.11. This time, we use Corollary 7.4 to demonstrate directly that $\iota := \inf(\mathcal{P})$ is attained. Let (f_n) be a minimizing sequence for $(\mathcal{P})$; this means that $J(f_n) = \int_{[0,1]} \eta_n + e(y_n(1)) \to \iota$ and that $\int_{[0,1]} \zeta_n \leq 1$ for all n. Here we abbreviate $\zeta_n := h(\cdot, f_n(\cdot))$, $\eta_n := g_0(\cdot, f_n(\cdot), y_n(\cdot))$ and $y_n := y_{f_n}$. If $\iota = +\infty$, the result in question is trivial; so let us suppose $\iota < +\infty$. This implies

$$\sup_n \int_{[0,1]} |(\eta_n, \zeta_n)| < +\infty.$$

The given property (γ) for the components of c is easily seen to be equivalent to $-|c|$ having property (γ) with respect to h. Therefore, $(\dot{y}_n) \subset \mathcal{L}^1_{\mathbf{R}^m}$ is uniformly integrable by the constraint (5.3) of $(\mathcal{P})$. So there exist a subsequence $(\dot{y}_{n'})$ and $\xi_\infty \in \mathcal{L}^1_{\mathbf{R}^m}$ such that

$$\dot{y}_{n'} \rightharpoonup \xi_\infty.$$

In particular, for $y_\infty \in \mathcal{AC}([0,1]; \mathbf{R}^m)$, defined by $y_\infty(\omega) := y_0 + \int_{[0,\omega]} \xi_\infty(\omega')d\omega'$, this yields

$$y_{n'}(\omega) \to y_\infty(\omega) \text{ for every } \omega.$$

Hence, $\liminf_{n'} \int_{[0,1]} \eta_{n'} \leq \iota' := \iota - e(y_\infty(1))$. Let $\ell', \ell'' : \Omega \times \mathbf{R}^m \times \mathbf{R}^2 \times \hat{\mathbf{N}} \to (-\infty, +\infty]$ be given by

$$\ell'(\omega, \xi, \eta, \zeta, k) := \begin{cases} \eta & \text{if } (\xi, \eta, \zeta) \in Q(\omega, y_k(\omega)), \\ +\infty & \text{otherwise} \end{cases}$$

$$\ell''(\omega, \xi, \eta, \zeta, k) := \begin{cases} \zeta & \text{if } (\xi, \eta, \zeta) \in Q(\omega, y_k(\omega)), \\ +\infty & \text{otherwise} \end{cases}$$

It is not hard to see that by the properties of the functions h, c and g_0 the multifunction Q has the following property, which is also referred to as *property* (K) [Ce]:

$$Q(\omega, y) = \cap_{\epsilon>0}\text{cl } \cup_{z\in\mathbf{R}^m,|y-z|<\epsilon} Q(\omega, z) \text{ for every } \omega \in \Omega \text{ and } y \in \mathbf{R}^m.$$

Indeed, for fixed $\omega \in \Omega$ we can observe that when $(\xi^k, \eta^k, \zeta^k) \in Q(\omega, y^k)$, with $(\xi^k, \eta^k, \zeta^k) \to (\xi^0, \eta^0, \zeta^0)$ and $y^k \to y^0$, then the corresponding sequence $(x^k) \subset S$ (x^k as in the definition of $Q(\omega, y^k)$) contains a convergent subsequence (by $h(\omega, x^k) \leq \zeta^k \to \zeta^0 < +\infty$ and inf-compactness of $h(\omega, \cdot)$). The rest of the proof of the above closed graph property of $Q(\omega, \cdot)$ is then quite easy. In turn, this property implies that both ℓ' and ℓ'' are lower semicontinuous integrands on $\Omega \times (\mathbf{R}^m \times \mathbf{R}^2 \times \hat{\mathbf{N}})$. Suppose that the convexity condition

$$Q(\omega, y_\infty(\omega)) \text{ is convex for every } \omega \in \Omega$$

for Q holds. Then the convexity condition of Corollary 7.4 is satisfied. By the given property (γ) of g_0 with respect to h, the remaining condition of Corollary 7.4 is also valid. Application of that corollary gives the existence of $(\eta_*, \zeta_*) \in \mathcal{L}^1_{\mathbf{R}^2}$ such that

$$\iota' \geq \int_{[0,1]} \ell'(\omega, \xi_\infty(\omega), \eta_*(\omega), \zeta_*(\omega), \infty) d\omega, \; 1 \geq \int_{[0,1]} \ell''(\omega, \xi_\infty(\omega), \eta_*(\omega), \zeta_*(\omega), \infty) d\omega.$$

Here we have used the obvious identities

$$\ell'(\omega, \dot{y}_{n'}(\omega), \eta_{n'}(\omega), \zeta_{n'}(\omega), n') = \eta_{n'}(\omega), \; \ell''(\omega, \dot{y}_{n'}(\omega), \eta_{n'}(\omega), \zeta_{n'}(\omega), n') = \zeta_{n'}(\omega).$$

Finiteness of the two integrals above implies that

$$\ell'(\omega, \xi_\infty(\omega), \eta_*(\omega), \zeta_*(\omega), \infty) < +\infty \text{ and } \ell''(\omega, \xi_\infty(\omega), \eta_*(\omega), \zeta_*(\omega), \infty) < +\infty \text{ a.e.}$$

By definition of ℓ' and ℓ'' this means that

$$(\xi_\infty(\omega), \eta_*(\omega), \zeta_*(\omega)) \in Q(\omega, y_\infty(\omega)) \text{ a.e.} \tag{7.1}$$

and that

$$\ell'(\omega, \xi_\infty(\omega), \eta_*(\omega), \zeta_*(\omega), \infty) = \eta_*(\omega), \; \ell''(\omega, \xi_\infty(\omega), \eta_*(\omega), \zeta_*(\omega), \infty) = \zeta_*(\omega) \text{ a.e.}$$

This results in $\int_{[0,1]} \eta_* \leq \iota' := \iota - e(y_\infty(1))$ and $\int_{[0,1]} \zeta_* \leq 1$. Finally, it follows from (7.1), by means of Theorem A.4, that there exists $f_* \in \mathcal{L}^0_S$ such that

$$\xi_\infty(\omega) = c(\omega, f_*(\omega), y_\infty(\omega)), \; \eta_*(\omega) \geq g_0(\omega, f_*(\omega), y_\infty(\omega)), \; \zeta_*(\omega) \geq h(\omega, f_*(\omega)) \text{ a.e.}$$

Since $\xi_\infty = \dot{y}_\infty$ a.e., we conclude that f_* meets (5.3) and satisfies $J(f_*) \leq \iota$. Hence, f_* is optimal for $(\mathcal{P})$. □

See [Ce] for a variety of existence results of the above type. See also [Ba19] for somewhat more involved lower closure and existence results for the optimal control of ordinary differential equations with so-called *recursive* objective functions.

8 Applications to lower closure without convexity in L^1

Contents: finite support equivalence, chattering control, integrals of multifunctions, Fatou-Vitali in several dimensions, closedness of attainable sets, existence of optimal controls without convexity

In previous sections we saw that, given certain convexity conditions, barycentric functions associated to Young measures can form solutions to lower closure and existence problems in L^1-spaces. In the absence of such convexity, alternatives for barycentric functions can sometimes be found. In such cases the desired solutions can actually be found by the use of extremal representing measures from Choquet theory (this can be seen as diametrically opposed to forming barycenters of Young measures). 'Hidden' convexity properties, obtained by supposing the underlying measure space $(\Omega, \mathcal{A}, \mu)$ to be nonatomic and restricting ourselves to what are essentially settings capable of producing finite-dimensional vector measures, also play a significant role in this endeavor (e.g., see [BL, Ba3, Ba21]). A more efficient approach to such lower closure results is obtained here by using what we shall call *finite support equivalence* (Theorem 8.2), followed by an application of a generalization of Lyapunov's theorem (Corollary A.11). This avoids the rather ponderous application of Choquet theory and/or extremal properties of Young measures altogether. As in sections 3-5, let S be a separable metric space.

Definition 8.1 (finitely supported Young measure) A Young measure $\delta \in \mathcal{R}$ is *finitely supported* if there exist for some $r \in \mathbf{N}$ functions $s_1, \cdots, s_r$ in $\mathcal{L}_S^0$ and *parametrized convex coefficients* $\alpha_1, \cdots, \alpha_r$ (i.e., in $\mathcal{L}_{\mathbf{R}}^\infty$, nonnegative and with $\sum_{i=1}^r \alpha_i = 1$) such that

$$\delta(\omega) = \sum_{i=1}^r \alpha_i(\omega)\epsilon_{s_i}(\omega) \text{ for all } \omega \in \Omega.$$

The set of all finitely supported Young measures is denoted by $\mathcal{R}_f$. □

Theorem 8.2 (finite support equivalence) *Suppose that S is Suslin. Let $g_1, \cdots, g_m$ be $\mathcal{A} \times \mathcal{B}(S)$-measurable functions on $\Omega \times S$. Let $\delta \in \mathcal{R}$ be such that*

$$I_{|g_j|}(\delta) < +\infty, \; j = 1, \cdots, m.$$

Then there exists $\delta' \in \mathcal{R}_f$ such that

$$\int_S g_j(\omega, x)\, \delta'(\omega)(dx) = \int_S g_j(\omega, x)\, \delta(\omega)(dx) \text{ for a.e. } \omega \in \Omega, \; j = 1, \cdots, m.$$

PROOF. Fix $\omega \in \Omega$ arbitrarily. Let $q : S \to \mathbf{R}^m$ be defined by $q(x) := (g_j(\omega, x))_{j=1}^m$. Then an application of Theorem A.13 gives that $\int_S q(x)\, \delta(\omega)(dx)$ lies in the convex hull of the set $q(S)$ (of course, here the actual application of Theorem A.13(ii) is to the image of the probability measure $\delta(\omega)$ under the mapping q [DM, II.12]). Since this is true for a.e. ω, Carathéodory's theorem plus a standard application of Theorem A.4 (cf. [CV, p. 101]) yield the existence of $m+1$ functions s_i in $\mathcal{L}_S^0$ and corresponding measurably parametrized convex coefficients α_i such that for a.e. ω

$$\int_S g_j(\omega, x)\, \delta(\omega)(dx) = \sum_{i=1}^{m+1} \alpha_i(\omega) g_j(\omega, s_i(\omega)), \; j = 1, \cdots, m.$$

Setting $\delta'(\omega) := \sum_i \alpha_i(\omega)\epsilon_{s_i}(\omega)$ completes the argument. □

Observe that the above proof also gives that the maximal cardinality of the support of $\delta'(\omega)$, equals $m+1$. As a first application, we use Example 5.11 to show how *chattering control* [Gam, War, Ce], i.e., relaxed control by means of finitely supported Young measures, comes about very naturally in optimal control.

Example 8.3 (chattering control equivalence) The following relaxed optimal control problem was studied in Example 5.11

$$(\mathcal{P}_{rel}): \inf_{\delta\in\mathcal{R}} J_{rel}(\delta) := \int^*_{[0,1]}[\int_S g_0(\omega,x,y_\delta(\omega))\delta(\omega)(dx)]d\omega + e(y_\delta(1)),$$

under the constraint $I_h(\delta) \leq 1$. Here $y_\delta : [0,1] \to \mathbf{R}^m$ in $\mathcal{AC}([0,1];\mathbf{R}^m)$ is given by

$$\dot{y}(\omega) = \int_S c(\omega,x,y(\omega))\delta(\omega)(dx) \text{ for a.e. } \omega \text{ in } \Omega$$

and the initial condition $y(0) = y_0$. For any $\delta \in \mathcal{R}$ with $J_{rel}(\delta) < +\infty$ an application of Theorem 8.2 to $g_{m+1}(\omega,x) := g_0(\omega,x,y_\delta(\omega))$ and $g_j(\omega,x) := c^j(\omega,x,y_\delta(\omega))$, $1 \leq j \leq m$, gives the existence of δ' in $\mathcal{R}_f$ such that a.e.

$$\int_S g_0(\omega,x,y_\delta(\omega))\,\delta'(\omega)(dx) = \int_S g_0(\omega,x,y_\delta(\omega))\,\delta(\omega)(dx) \text{ and } \dot{y}_{\delta'}(\omega) = \dot{y}_\delta(\omega).$$

By the initial condition, the latter identity gives that the trajectories y_δ and $y_{\delta'}$ coincide completely. Combined with the first identity, this gives $J_{rel}(\delta') = J_{rel}(\delta)$. So $(\mathcal{P}_{rel})$ is equivalent to the optimization of J_{rel} over the subset $\mathcal{R}_f$ of $\mathcal{R}$. □

Definition 8.4 (selectors) An *integrable [measurable] selector* of a multifunction $F : \Omega \to 2^{\mathbf{R}^d}$ is a function $f \in \mathcal{L}^1_{\mathbf{R}^d}$ $[\mathcal{L}^0_{\mathbf{R}^d}]$ such that $f(\omega) \in F(\omega)$ for a.e. ω in Ω. The set (possibly empty) of all integrable [measurable] selectors of F is denoted by $\mathcal{L}^1_F$ $[\mathcal{L}^0_F]$. □

Definition 8.5 (integral of a multifunction) The *integral* of a (possibly nonmeasurable) multifunction $F : \Omega \to 2^{\mathbf{R}^d}$, denoted by $\int_\Omega F\,d\mu$, is the (possibly empty) set of all integrals of integrable selectors of F:

$$\int_\Omega F\,d\mu := \{\int_\Omega f\,d\mu : f \in \mathcal{L}^1_F\}.$$

□

One of the principal results of this section is as follows:

Theorem 8.6 (Fatou-Vitali in several dimensions) *Let (f_n) in $\mathcal{L}^1_{\mathbf{R}^d}$ be uniformly L^1-bounded and such that*

$$a := \lim_{n\to\infty} \int_\Omega f_n\,d\mu \textit{ exists in } \mathbf{R}^d.$$

Let C be the convex cone consisting of all $z \in \mathbf{R}^d$ for which $(\min(f_n \cdot z, 0))$ is uniformly integrable. Then for $L : \Omega \to 2^{\mathbf{R}^d}$ given by $L(\omega) := \mathrm{Ls}_n f_n(\omega)$:

$$a \in \int_\Omega L\,d\mu - C^0,$$

where C^0 is the negative polar cone associated to C. [12]

[12] I.e., C^0 consists of all $x \in \mathbf{R}^d$ such that $x \cdot z \leq 0$ for all $z \in C$.

Following [Ba21], we shall derive Theorem 8.6 from the following auxiliary result (the use of finite support equivalence simplifies the proof considerably).

Lemma 8.7 *Suppose that* $(\Omega, \mathcal{A}, \mu)$ *is nonatomic. Let* $F : \Omega \to 2^{\mathbf{R}^d}$ *be a multifunction with measurable graph, and let* $\delta \in \mathcal{R}(\Omega, \mathcal{A}, \mu; \mathbf{R}^d)$ *be such that*

$$\int_\Omega [\int_{\mathbf{R}^d} |x|\delta(\omega)(dx)]\mu(d\omega) < +\infty \textit{ and } \delta(\omega)(F(\omega)) = 1 \textit{ for a.e. } \omega \textit{ in } \Omega.$$

Then

$$\int_\Omega \text{bar } \delta(\omega)\mu(d\omega) \in \int_\Omega F\, d\mu.$$

PROOF. We apply Theorem 8.2 to $g_0(\omega, x) := 0$ if $x \in F(\omega)$ and $g_0(\omega, x) := +\infty$ if $x \notin F(\omega)$, to $g_j(\omega, x) := x^j$ (the j-th coordinate of x), $1 \leq j \leq d$, and to $g_{d+1}(\omega, x) := |x|$. It follows that there exists δ' in $\mathcal{R}_f(\Omega, \mathcal{A}, \mu; \mathbf{R}^d)$ such that a.e. $\delta'(\omega)(F(\omega)) = 1$, bar $\delta'(\omega) =$ bar $\delta(\omega)$ and $\int_{\mathbf{R}^d} |x|\delta'(\omega)(dx) = \int_{\mathbf{R}^d} |x|\delta(\omega)(dx)$. By Definition 8.1, δ' has the form $\Sigma_{i=1}^r \alpha_i \epsilon_{s_i}$ for some $r \in \mathbf{N}$ (actually, here $r = d+3$ by an earlier observation, but this is of no importance). The first of the a.e. identities just obtained implies that for a.e. ω one has $s_i(\omega) \in F(\omega)$ whenever $\alpha_i(\omega) > 0$; thus, without any loss of generality we may suppose that all s_i are in $\mathcal{L}^0_F$. Integrating the other a.e. identities over Ω gives

$$\int_\Omega \text{bar } \delta\, d\mu = \int_\Omega \sum_{i=1}^r \alpha_i s_i\, d\mu \tag{8.1}$$

and

$$\int_\Omega \sum_{i=1}^r \alpha_i |s_i| d\mu = \int_\Omega [\int_{\mathbf{R}^d} |x|\delta(\omega)(dx)]\mu(d\omega) < +\infty.$$

The latter identity allows application of the extended Lyapunov theorem, as contained in Corollary A.11. This guarantees existence of a measurable partition $B_1, \cdots, B_{d+2}$ of Ω such that each s_i is integrable over B_i and $\int_\Omega \sum_i \alpha_i s_i\, d\mu = \sum_i \int_{B_i} s_i\, d\mu$. By $\sum_i 1_{B_i} s_i \in \mathcal{L}^1_F$, the desired result follows from (8.1). □

PROOF of Theorem 8.6. Let Ω^{na} and $(\bar{A}_j)$ be as in Proposition A.7. By the hypotheses, we have $\sup_n \sum_j \mu(\bar{A}_j)|a_{n,j}| < +\infty$, where $a_{n,j}$ denotes the constant value which f_n must take a.e. on the atom $\bar{A}_j$. By a suitable diagonal procedure, this implies that there exists a subsequence $(f_{n'})$ of (f_n) for which $a_{\infty,j} := \lim_{n' \to \infty} a_{n',j}$ exists for every j. Let us define $F : \Omega \to 2^{\mathbf{R}^d}$ by $F(\omega) := \text{Ls}_{n'} f_{n'}(\omega) \subset L(\omega)$. The above implies that for every j

$$F(\omega) = \{a_{\infty,j}\} \text{ for a.e. } \omega \in \bar{A}_j. \tag{8.2}$$

By Theorem 6.2 (see Example 4.8), there exist a subsequence $(f_{n''})$ and $\delta_* \in \mathcal{R}(\Omega, \mathcal{A}, \mu; \mathbf{R}^d)$ such that $a \cdot z = \liminf_{n''} \int_\Omega f_{n''} \cdot z\, d\mu \geq \int_\Omega \text{bar } \delta_* \cdot z\, d\mu$ for every $z \in C$ entirely similar to the proof of Theorem 6.8 (or, alternatively, Example 4.9). Thus, we arrive at

$$\int_\Omega \text{bar } \delta_*\, d\mu - a \in C^0. \tag{8.3}$$

The support property in Theorem 6.2 also guarantees $\delta_*(\omega)(F(\omega)) = 1$ a.e. By Lemma 8.7, the established properties of δ_* imply that $\int_{\Omega^{na}} \text{bar } \delta_*(\omega)\mu(d\omega)$ belongs to $\int_{\Omega^{na}} F\, d\mu$. Now by (8.2), $\int_\Omega F\, d\mu = \int_{\Omega^{pa}} f_\infty\, d\mu + \int_{\Omega^{na}} F\, d\mu$, where $f_\infty(\omega) := a_{\infty,j}(\omega)$ for a.e. $\omega \in \bar{A}_j$ (the atoms $\bar{A}_j$ are pairwise disjoint modulo null sets). So by Richter's theorem (Corollary A.10) it follows that

$$\text{co} \int_\Omega F \, d\mu = \int_{\Omega^{pa}} f_\infty \, d\mu + \int_{\Omega^{na}} F \, d\mu = \int_\Omega F \, d\mu.$$

We conclude that $\int_\Omega \text{bar}\, \delta_* \, d\mu$ belongs to $\int_\Omega F \, d\mu$. Because of (8.3), this concludes the proof. □

Corollary 8.8 *Let* (f_n) *in* $\mathcal{L}^1_{\mathbf{R}^d}$ *be such that* $(\min(f_n^i, 0))$ *is uniformly integrable,* $i = 1, \cdots, d$, *and*

$$a := \lim_{n\to\infty} \int_\Omega f_n \, d\mu \textit{ exists in } \mathbf{R}^d.$$

Then for $L(\omega) := \text{Ls}_n f_n(\omega)$ *there exists* $f_* \in \mathcal{L}^1_L$ *such that*

$$a \geq \int_\Omega f_* \, d\mu.$$[13]

PROOF. The conditions imply $\sup_n \int_\Omega |f_n| d\mu < +\infty$, so the result follows directly from Theorem 8.6 and Definition 8.5, with $C = \mathbf{R}^d_+$ and $C^0 = \mathbf{R}^d_-$. □

The next result should be compared with Proposition 4.10:

Corollary 8.9 *Let* (f_n) *in* $\mathcal{L}^1_{\mathbf{R}^d}$ *be uniformly integrable and such that*

$$a := \lim_{n\to\infty} \int_\Omega f_n \, d\mu \textit{ exists in } \mathbf{R}^d.$$

Then for $L(\omega) := \text{Ls}_n f_n(\omega)$ *there exists* $f_* \in \mathcal{L}^1_L$ *such that*

$$a = \int_\Omega f_* \, d\mu.$$

PROOF. Again a direct application of Theorem 8.6 and Definition 8.5, this time with $C = \mathbf{R}^d$ and $C^0 = \{0\}$. □

Theorem 8.6 was given in [Ba4, Ba3]. See also [BH] for comprehensive extensions of these results to multifunctions with unbounded values, both in finite and infinite dimensions. Corollary 8.8 is from [Ba3, Ba4], and Corollary 8.9 from [Ar2].

Example 8.10 (closedness of attainable sets) Consider again the optimal control problem $(\mathcal{P})$ of Example 5.11. Suppose that the function $c : [0,1] \times S \times \mathbf{R}^m \to \mathbf{R}^m$ has the following *semilinear* structure:

$$c(\omega, x, y) = A(\omega)y + d(\omega, x),$$

where A belongs to $\mathcal{L}^1_{\mathbf{R}^{m\times m}}$ and $d : [0,1] \times S \to \mathbf{R}^m$ is measurable and has as its component functions d^i continuous integrands on $[0,1] \times S$ such that d^i and $-d^i$ have property (γ) with respect to the inf-compact integrand h in (5.3). As in Example 8.3, we shall also assume that g and h are measurable. By the semilinear structure of c, the trajectory y_f corresponding to the control function $f \in \mathcal{L}^0_S$, can be expressed explicitly as follows [War, II.4.8]:

$$y_f(\omega) = \Lambda(\omega)y_0 + \Lambda(\omega) \int_0^\omega \Lambda(\omega')^{-1} d(\omega', f(\omega'))d\omega'.$$

[13] '$\geq$' is the usual, coordinatewise partial ordering in $\mathbf{R}^d$.

Here $\Lambda \in \mathcal{AC}([0,1]; \mathbf{R}^{m\times m})$ is the *fundamental solution*, determined by $\dot{\Lambda} = A\Lambda$ and $\Lambda(0) =$ $m \times m$-identity matrix. Using Corollary 8.8, we prove that

$$D := \{y_f(1) : f \in \mathcal{L}^0_S, \int_{[0,1]} h(\omega, f(\omega))d\omega \leq 1\},$$

i.e., the attainable set at time 1, is closed (whence compact by the properties of d). Indeed, let $y_{f_n}(1) \to a$, for arbitrary (f_n) in $\mathcal{L}^0_S$ with $\int_{[0,1]} \zeta_n \leq 1$ for all n; here $\zeta_n(\omega) := h(\omega, f_n(\omega))$. Evidently, there exist a subsequence $(f_{n'})$ and $\alpha \in \mathbf{R}$ such that $\int_{[0,1]} \zeta_{n'} \to \alpha$. Define $\bar{f}_{n'} \in \mathcal{L}^1_{\mathbf{R}^{m+1}}$ by $\bar{f}_{n'} := (\lambda_{n'}, \zeta_{n'})$, where $\lambda_{n'}(\omega) := \Lambda(1)\Lambda(\omega)^{-1}d(\omega, f_{n'}(\omega))$. The given growth dominance property is equivalent to $-|d|$ having property (γ) with respect to h and implies that $(\lambda_{n'})$ is uniformly integrable. Now $\int_{[0,1]} \bar{f}_{n'} \to (a - \Lambda(1)y_0, \alpha)$, so Theorem 8.6 (with $C \supset \mathbf{R}^m \times \mathbf{R}_+$) gives the existence of $\bar{f}_* =: (\lambda_*, \zeta_*) \in \mathcal{L}^1_{\mathbf{R}^{m+1}}$ such that

$$\int_{[0,1]} \lambda_* = a - \Lambda(1)y_0, \int_{[0,1]} \zeta_* \leq \alpha \text{ and } \bar{f}_*(\omega) \in \mathrm{Ls}_{n'} \bar{f}_{n'}(\omega) \text{ a.e.}$$

By the latter property there exists for a.e. ω a subsequence $(f_{n''})$ of $(f_{n'})$, possibly depending upon ω, such that

$$\lim_{n''} d(\omega, f_{n''}(\omega)) = \lambda_*(\omega), \lim_{n''} h(\omega, f_{n''}(\omega)) = \zeta_*(\omega).$$

Because $\zeta_*(\omega)$ is finite, this implies for a.e. ω the existence of $x_\omega \in S$ such that $h(\omega, x_\omega) \leq \zeta_*(\omega)$, given the inf-compactness of $h(\omega, \cdot)$ [x_ω is limit point of a convergent subsequence of $(f_{n''}(\omega))$]. For the same x_ω we get also $d(\omega, x_\omega) = \lambda_*(\omega)$, by continuity of $d(\omega, \cdot)$. By Theorem A.4 there exists $f_* \in \mathcal{L}^0_S$ such that

$$d(\omega, f_*(\omega)) = \lambda_*(\omega), \ h(\omega, f_*(\omega)) \leq \zeta_*(\omega) \text{ a.e.}$$

We conclude that

$$y_{f_*}(1) = \Lambda(1)y_0 + \int_{[0,1]} \lambda_* = a, \int_{[0,1]} h(\omega, f_*(\omega))d\omega \leq \int_{[0,1]} \zeta_* \leq \alpha \leq 1,$$

which proves that D is closed. □

Example 8.11 (existence of optimal controls without convexity) The argument given in Example 8.10 can easily be expanded to prove the existence of an optimal solution for $(\mathcal{P})$, provided that the objective integrand g_0 does not explicitly depend on y (hence we simply write $g_0(\omega, x)$, etc.). In this case, let (f_n) be a minimizing sequence for $(\mathcal{P})$, i.e., such that $J(f_n) = \int_{[0,1]} \eta_n + e(y_n(1)) \to \iota := \inf(\mathcal{P})$. Here $\eta_n(\omega) := g_0(\omega, f_n(\omega))$ and $y_n := y_{f_n}$. Let $\beta := \liminf_n \int_{[0,1]} \eta_n$. There exist a subsequence $(f_{n'})$, $\alpha \in [0,1]$ and $a \in \mathbf{R}^m$ such that $\int_{[0,1]}(\eta_{n'}, \zeta_{n'}) \to (\beta, \alpha)$ and $y_{n'}(1) \to a$ [note that $(y_n(1))$ is bounded, by combining the explicit represeesentation for $y_f(1)$ in the previous example, (5.3) and the fact that $-|d|$ has property (γ) with respect to h]. Observe that $\beta \leq \iota - e(a)$, by using the lower semicontinuity of e. Now $\int_{[0,1]} \tilde{f}_{n'} \to (\beta, a - \Lambda(1)y_0, \alpha)$, where $(\tilde{f}_{n'})$ in $\mathcal{L}^1_{\mathbf{R}^{m+2}}$ is defined by $\tilde{f}_{n'} := (\eta_{n'}, \lambda_{n'}, \zeta_{n'})$, with $\lambda_{n'}$ as in Example 8.10. Quite similar to the previous example, applying Theorem 8.6 leads to the existence of $(\eta_*, \lambda_*, \zeta_*) \in \mathcal{L}^1_{\mathbf{R}^{m+2}}$ such that

$$\int_{[0,1]} \eta_* \leq \beta, \int_{[0,1]} \lambda_* = a - \Lambda(1)y_0, \int_{[0,1]} \zeta_* \leq \alpha.$$

Also, a similar application of Theorem A.4 gives the existence of $f_* \in \mathcal{L}_S^0$ with

$$g_0(\omega, f_*(\omega)) \leq \eta_*(\omega), d(\omega, f_*(\omega)) = \lambda_*(\omega),\ h(\omega, f_*(\omega)) \leq \zeta_*(\omega) \text{ a.e..}$$

Combined with the above (in)equalities, this gives $\int_{[0,1]} g_0(\omega, f_*(\omega))d\omega + e(a) \leq \iota$, $a = y_{f_*}(1)$ and $\int_{[0,1]} h(\omega, f_*(\omega))d\omega \leq 1$. Thus, f_* satisfies the constraint (5.3) and $J(f_*) \leq \iota$. □

In strong contrast to Examples 5.11, 5.12 and 7.5, the above existence result requires no convexity properties for the orientor field Q. This is a classical subject in optimal control theory; see [Ba5] for more involved applications, based on the same approach as given above. When g_0 has a semilinear form comparable to the one considered for c in Examples 8.10 and 8.11, the above approach can also be maintained by direct substitution of the formula for y_f as given in Example 8.10; see [Ce] for details. A more general approach to the subject of existence without convexity can be found in [Ba20]. There c is semilinear, as above, but g_0 is *semiconcave* in the variable y. This approach uses a Bauer-type extremum principle [Cho], and is based on the fact that the set $\mathcal{R}_D$ of Dirac Young measures forms the extreme point boundary of $\mathcal{R}$. The main results in [Ba20] include the classical existence results of this type [Ce] and generalize some recent results with semiconcave g_0 (e.g., [CC, Ra]).

9 Applications to denseness and functional relaxation

Contents: Dirac equivalence, denseness theorem, limiting bang-bang theorem, functional relaxation in optimal control

As in sections 3–5, let S be a separable metric space. We begin this section by stating an existence result which complements Proposition 5.8. It extends similar results in [AP, BL, Ar1, Ba1, Ba4] for an existence result arising in economics.

Theorem 9.1 (existence of optimal nonrandomized decision rules) *Suppose that $(\Omega, \mathcal{A}, \mu)$ is nonatomic and that S is Suslin. Let $g_{m+1} : \Omega \times S \to [0, +\infty]$ be an inf-compact integrand on $\Omega \times S$, and let $g_0, g_1, \cdots, g_m$ be lower semicontinuous integrands on $\Omega \times S$, having property (γ) with respect to g_{m+1}. Suppose that $g_0, g_1, \cdots, g_{m+1}$ are $\mathcal{A} \times \mathcal{B}(S)$-measurable. Then the optimization problem*

$$(\mathcal{Q}) : \inf_{f \in \mathcal{L}_S^0} \{J_{g_0}(f) : J_{g_1}(f) \leq 1, \cdots, J_{g_{m+1}}(f) \leq 1\}$$

has an optimal solution, provided that its feasible set is nonempty.

One of the principal results of this section, from which we shall also derive Theorem 9.1, is as follows:

Theorem 9.2 (Dirac equivalence) *Suppose that $(\Omega, \mathcal{A}, \mu)$ is nonatomic and that S is Suslin. Let $g_1, \cdots, g_m$ be $\mathcal{A} \times \mathcal{B}(S)$-measurable integrands on $\Omega \times S$. Let $\delta \in \mathcal{R}$ be such that*

$$I_{|g_j|}(\delta) < +\infty, \; j = 1, \cdots, m.$$

Then there exists $f \in \mathcal{L}_S^0$ such that

$$J_{g_j}(f) = I_{g_j}(\delta), \; j = 1, \cdots, m.$$

Proof. We can apply Theorem 8.2 to the collection of integrands consisting of all g_j and $|g_j|$, $1 \leq j \leq m$. This gives the existence of s_i in $\mathcal{L}_S^0$ and measurable convex coefficients α_i in $\mathcal{L}_{\mathbf{R}}^\infty$, $1 \leq i \leq r$, $r \in \mathbf{N}$, such that $\delta' := \sum_i \alpha_i \epsilon_{s_i}$ satisfies

$$I_{g_j}(\delta') = I_{g_j}(\delta), \; I_{|g_j|}(\delta') = I_{|g_j|}(\delta), \; j = 1, \cdots, m,$$

upon integration over Ω. An application of Lyapunov's theorem (Corollary A.11) yields the existence of a measurable partition $B_1, \cdots, B_r$ of Ω such that each f_i is integrable over B_i and $\int_\Omega \sum_{i=1}^r \alpha_i f_i \, d\mu = \sum_{i=1}^r \int_{B_i} f_i \, d\mu$, where $f_i(\omega) := (g_j(\omega, s_i(\omega)))_{j=1}^m$. Define $f \in \mathcal{L}_S^0$ by setting $f(\omega) := s_i(\omega)$ if $\omega \in B_i$. Then the desired statement holds. □

Proof of Theorem 9.1. By Proposition 5.8 there exists $\delta_* \in \mathcal{R}$, $I_{g_j}(\delta_*) \leq 1$, $1 \leq j \leq m$, such that $I_{g_0}(\delta_*) \leq J_{g_0}(f)$ for all f that are feasible for $(\mathcal{Q})$ [i.e., for which ϵ_f is feasible for $(\mathcal{P})$ – note that $(\mathcal{P})$ of Proposition 5.8 is precisely the Young relaxation of $(\mathcal{Q})$]. Let $1 \leq j \leq m$ be arbitrary. Then simple manipulations involving Definition 5.1 show that $I_{g_j^-}(\delta_*) < +\infty$ (by $I_{g_{m+1}}(\delta_*) \leq 1$). Also, $I_{g_j^+}(\delta_*) < +\infty$ is immediate by $I_{g_j}(\delta_*) \leq 1$ (see Definition B.1), and we have trivially $I_{|g_{m+1}|}(\delta_*) \leq 1$, since g_{m+1} is nonnegative. Therefore, the integrability conditions of the Dirac equivalence Theorem 9.2 hold for $\delta := \delta_*$; it follows that there exists $f_* \in \mathcal{L}_S^0$ such that $J_{g_j}(f_*) = I_{g_j}(\delta_*)$ for $0 \leq j \leq m+1$. This immediately gives optimality of f_* for $(\mathcal{Q})$. □

Alternatively, Theorem 9.1 could also be proven by an application of Corollary 8.8 [Ba4]. However, the above proof actually yields a little more, since it shows that it is already enough to require nonemptiness for the feasible set of $(\mathcal{P})$ (i.e., the Young relaxation of $(\mathcal{Q})$). The proofs in [BL, Ba3, Ba4, Ba20] all use extreme point arguments. It is an open question whether the results of [Ba20] can als be obtained using the equivalence Theorems 8.2 and 9.2.

Our next results show that in the nonatomic case the set $\mathcal{R}_D$ of Dirac Young measures is dense in $\mathcal{R}$. This is a well-known feature of Young measures (e.g., see [War]). A very general denseness result of this nature was given in [Ba6].

Theorem 9.3 (denseness theorem) *Suppose that $(\Omega, \mathcal{A}, \mu)$ is nonatomic and that S is Suslin. Let $g_1, \cdots, g_m$ be m measurable integrands on $\Omega \times S$, all integrably bounded from below. For every $\delta \in \mathcal{R}$ there exists a sequence $(f_n) \subset \mathcal{L}^0_S$ such that $\epsilon_{f_n} \Longrightarrow \delta$ and for every n*

$$J_{g_j}(f_n) = I_{g_j}(\delta) \text{ for all } 1 \leq j \leq m \text{ with } I_{g_j}(\delta) < +\infty.$$

PROOF. Choose any $\delta \in \mathcal{R}$. By the fact that $\mathcal{P}(S)$ is separable and metrizable for the narrow convergence topology [DM, III.60], Proposition A.2 implies that the σ-algebra generated by δ is countable. In conjunction with Proposition A.12 this shows that there is a countably generated sub-σ-algebra $\mathcal{A}_0$ such that δ belongs to the set $\mathcal{R}(\mathcal{A}_0)$ of Young measures from $(\Omega, \mathcal{A}_0, \mu)$ into S, and such that all g_j are $\mathcal{A}_0 \times \mathcal{B}(S)$-measurable, and $(\Omega, \mathcal{A}_0, \mu)$ is nonatomic. Let $g_{i,j}(\omega, x) := 1_{A_j}(\omega) c_i(x)$ be as in the proof of Theorem 3.7, but this time with (A_j) being an enumeration of the algebra that generates the sub-σ-algebra $\mathcal{A}_0$. Let $n \in \mathbf{N}$ be arbitrary. Without loss of generality we shall suppose that $I_{g_k}(\delta) < +\infty$ for all k (rather than omitting indices for which this does not hold). By integrable boundedness from below of the g_k's, this entails $I_{|g_k|}(\delta) < +\infty$ for all $1 \leq k \leq m$. By Theorem 9.2 there exists $f_n \in \mathcal{L}^0_S$ such that

$$J_{g_{i,j}}(f_n) = I_{g_{i,j}}(\delta),\ 1 \leq i, j \leq n,\ J_{g_k}(f_n) = I_{g_k}(\delta),\ 1 \leq k \leq m.$$

From the definition of $d_{\mathcal{R}(\mathcal{A}_0)}$ it follows easily that $d_{\mathcal{R}(\mathcal{A}_0)}(\epsilon_{f_n}, \delta) \to 0$. By Theorem 3.7 this gives $\epsilon_{f_n} \Longrightarrow \delta$ in $\mathcal{R}(\mathcal{A}_0)$ and $J_{g_k}(f_n) = I_{g_k}(\delta)$ for $1 \leq k \leq m$. It remains to show that $\epsilon_{f_n} \Longrightarrow \delta$ in $\mathcal{R}$: Let $A \in \mathcal{A}$ and $c \in \mathcal{C}_b(S)$ be arbitrary. Denote by ϕ_A a version of the conditional expectation of 1_A with respect to $\mathcal{A}_0$ and μ. Denoting $g(\omega, x) := 1_A(\omega) c(x)$ and $g'(\omega, x) := \phi_A(\omega) c(x)$, one obviously has $J_g(f_n) = J_{g'}(f_n)$ for all n and $I_g(\delta) = I_{g'}(\delta)$. Therefore, the convergence result follows immediately by Definition 3.3. □

The next result extends a limiting bang-bang result from [Val1].

Theorem 9.4 *Suppose that $(\Omega, \mathcal{A}, \mu)$ is nonatomic. Let $F : \Omega \to 2^{\mathbf{R}^d}$ be a multifunction with measurable graph and let $\delta \in \mathcal{R}(\Omega, \mathcal{A}, \mu; \mathbf{R}^d)$ be such that*

$$\int_\Omega [\int_{\mathbf{R}^d} |x| \delta(\omega)(dx)] \mu(d\omega) < +\infty \text{ and } \delta(\omega)(F(\omega)) = 1 \text{ for a.e. } \omega.$$

Then there exists a sequence $(f_n) \subset \mathcal{L}^1_F$ such that $f_n \rightharpoonup \text{bar } \delta$ in $\mathcal{L}^1_{\mathbf{R}^d}$.

PROOF. The integrability hypothesis is equivalent to having $\int_{\mathbf{R}^d} |x| \nu(dx) < +\infty$, where the finite measure ν on $(\mathbf{R}^d, \mathcal{B}(\mathbf{R}^d))$ is given by $\nu := [\mu \otimes \delta](\Omega \times \cdot)$. Hence, by de la Vallée-Poussin's criterion (Theorem A.3) there exists a function $h' : \mathbf{R}_+ \to \mathbf{R}_+$, convex, continuous and nondecreasing, such that $\lim_{\xi\to\infty} h'(\xi)/\xi = +\infty$ and $\int_{\mathbf{R}^d} h'(|x|)\nu(dx) < +\infty$. Define $h(\omega, x) := h'(|x|)$; then the previous inequality can be reformulated as $I_h(\delta) < +\infty$. Let g_F be the lower semicontinuous integrand on $\Omega \times \mathbf{R}^d$ given by

$$g_F(\omega, x) := \begin{cases} |x| & \text{if } x \in F(\omega) \\ +\infty & \text{otherwise} \end{cases}$$

Then $I_{g_F}(\delta) < +\infty$ is obvious. By Theorem 9.3 there exists a sequence (f_k) in $\mathcal{L}^0_{\mathbf{R}^d}$ such that $\epsilon_{f_k} \Longrightarrow \delta$, $J_h(f_k) = I_h(\delta)$ and $J_{g_F}(f_k) = I_{g_F}(\delta)$ for all k. By the last identity we have $f_k \in \mathcal{L}^1_F$ for all k. Also, the first identity implies that (f_k) is uniformly integrable. So Example 4.9 implies the existence of a subsequence (f_n) of (f_k) with $f_n \rightharpoonup$ bar δ. □

Finally, we use Theorem 9.3 to take a fresh look at M. Valadier's improvement [Val2] of the main functional relaxation result in the textbook of I. Ekeland and R. Temam [ET, IX] (see also [E, BL]). Although the presentation is limited to the simple dynamical system of Example 5.11, the method clearly applies to the dynamical system used in [ET, Val2] as well.

Example 9.5 (functional relaxation in optimal control) We reconsider the original control problem $(\mathcal{P})$ of Example 5.11 and determine, under some extra assumptions, the precise nature of the lower semicontinuous hull of the objective functional J. These extra assumptions are as follows: we take $S := \mathbf{R}^m$, $c(\omega, x, y) := x$, $e = 0$ and we suppose that the constraint (5.3) is no longer explicitly present. Instead, we require the following: there exists $h' : \mathbf{R}_+ \to \mathbf{R}_+$, convex, lower semicontinuous and nondecreasing, such that $\lim_{\xi\to\infty} h'(\xi)/\xi = +\infty$ and

$$h(\omega, x) + \psi(\omega) \le g_0(\omega, x, y) \le \alpha h(\omega, x) + \chi(|y|) + \psi'(\omega) \text{ for all } (\omega, x, y) \in [0,1] \times \mathbf{R}^m \times \mathbf{R}^m,$$

where $h(\omega, x) := h'(|x|)$, ψ and ψ' belong to $\mathcal{L}^1_{\mathbf{R}}$, α is a constant and $\chi : \mathbf{R}_+ \to \mathbf{R}_+$ is a nondecreasing function. The objective functional J of $(\mathcal{P})$, defined on $\mathcal{L}^1_{\mathbf{R}^m}$, is given by

$$J(f) := \int^*_{[0,1]} g_0(\omega, f(\omega), y_f(\omega))d\omega.$$

As usual, let $J_h(f) := \int_{[0,1]} h(\omega, f(\omega))d\omega$; observe that $J_h(f) = +\infty$ implies $J(f) = +\infty$. For each $\omega \in \Omega$, $y \in \mathbf{R}^m$, let $g_0^{**}(\omega, \cdot, y)$ be the Fenchel biconjugate of the function $g_0(\omega, \cdot, y)$ on $\mathbf{R}^m$. Then we shall show that the lower semicontinuous hull $\bar{J}$ of $J : \mathcal{L}^1_{\mathbf{R}^m} \to (-\infty, +\infty]$, with respect to the weak topology $\sigma(\mathcal{L}^1_{\mathbf{R}^m}, \mathcal{L}^\infty_{\mathbf{R}^m})$, is given by

$$\bar{J}(f) = J'(f) := \int^*_{[0,1]} g_0^{**}(\omega, f(\omega), y_f(\omega))d\omega.$$

To see this, begin by observing that J' is weakly lower semicontinuous. This follows from Corollary 7.2 and the obvious continuity of the mapping $f \mapsto y_f(\omega)$ for each fixed $\omega \in [0,1]$ (observe that $y_f(\omega) = \int_0^\omega f(\omega')d\omega'$). Therefore, it follows immediately that $\bar{J} \ge J'$. To show the converse, let us fix an arbitrary $f_0 \in \mathcal{L}^1_{\mathbf{R}^m}$. Because of the inequality $\bar{J}(f_0) \ge J'(f_0)$ we may suppose without loss of generality that $J'(f_0)$ is finite. This implies $g_0^{**}(\omega, f_0(\omega), y_{f_0}(\omega)) < +\infty$ for a.e. ω. Also, by lower semicontinuity and convexity of h' we have

$$g_0^{**}(\omega, x, y_{f_0}(\omega)) \ge h'(|x|) + \psi(\omega). \tag{9.1}$$

Thus, for a.e. ω we may apply Corollary A.15 to $g_0(\omega, \cdot, y_{f_0}(\omega))$, which gives the existence of $\nu_\omega \in \mathcal{P}(\mathbf{R}^m)$ such that

$$g_0^{**}(\omega, f_0(\omega), y_{f_0}(\omega)) = \int_{\mathbf{R}^m} g_0(\omega, x, y_{f_0}(\omega))\nu_\omega(dx),$$

and such that bar ν_ω exists and equals $f_0(\omega)$. By Theorem A.13(ii) and Carathéodory's theorem (see the proof of Theorem 8.2) it follows that for a.e. ω there exist $m+2$ points $s_{i,\omega}$ and convex coefficients $\alpha_{i,\omega}$, $1 \le i \le m+2$, such that

$$g_0^{**}(\omega, f_0(\omega), y_{f_0}(\omega)) = \sum_{i=1}^{m+2} \alpha_{i,\omega} g_0(\omega, , s_{i,\omega}, y_{f_0}(\omega)), \;\; f_0(\omega) = \sum_{i=1}^{m+2} \alpha_{i,\omega} s_{i,\omega}.$$

By an application of Theorem A.4 we may suppose without loss of generality that the functions $s_i : \omega \mapsto s_{i,\omega}$ and $\alpha_i : \omega \mapsto \alpha_{i,\omega}$ are $\mathcal{A}$-measurable. Set $\delta' := \sum_i \alpha_i \epsilon_{s_i}$ (a.e.); then, just as in Example 8.3, the above implies $y_{\delta'} = y_{f_0}$ and $I_{g_{m+1}}(\delta') = J'(f_0) < +\infty$, where the integrand g_{m+1} on $[0,1] \times \mathbf{R}^m$ is given by $g_{m+1}(\omega, x) := g_0(\omega, x, y_{f_0}(\omega))$. Secondly, we also have trivially that $J_{\bar g}(\delta') = 0 < +\infty$, where $\bar g(\omega, x) := 0$ if $x \in \{s_{i,\omega} : 1 \le i \le m+2\}$ and $\bar g(\omega, x) := +\infty$ otherwise. Thirdly, it follows by (9.1) that $I_h(\delta')$ is finite. Since $I_h(\delta') = \int_{\mathbf{R}^m} h'(|x|)\nu'(dx)$ for $\nu' := \mu \otimes \delta'(\Omega \times \cdot)$, it follows by de la Vallée-Poussin's Theorem A.3 that there exists $h'' : \mathbf{R}_+ \to \mathbf{R}_+$, convex, continuous and nondecreasing, with $\lim_{\xi \to \infty} h''(\xi)/\xi = +\infty$, such that $I_{\tilde h}(\delta') := \int_{\mathbf{R}^m} h''(h'(|x|))\nu'(dx) < +\infty$. Here $\tilde h(\omega, x) := h''(h'(|x|))$. We now invoke the denseness Theorem 9.3, which yields the existence of a sequence (f_n) in $\mathcal{L}^0_{\mathbf{R}^m}$ such that $\epsilon_{f_n} \Longrightarrow \delta'$ and with $J_{g_{m+1}}(f_n) = I_{g_{m+1}}(\delta')$, $J_{\bar g} f_n) = I_{\bar g}(\delta')$, and $J_{\tilde h}(f_n) = I_{\tilde h}(\delta')$ for all $n \in \mathbf{N}$. Of course, the first of these identities implies $J_{g_{m+1}}(f_n) = J'(f_0)$ for all n. The second one implies that $f_n(\omega) \in \{s_{i,\omega} : 1 \le i \le m+2\}$ a.e. for each n, and by Theorem A.3 the third one implies that $(h'(|f_n(\cdot)|))$ is uniformly integrable. *A fortiori* this implies that (f_n) is uniformly integrable (by the reverse implication in Theorem A.3, which remains valid for h' lower semicontinuous), so by Example 4.9 we get existence of a subsequence $(f_{n'})$ with $f_{n'} \rightharpoonup$ bar δ'; in other words, $f_{n'} \rightharpoonup f_0$. To conclude the proof of the identity $\bar J = J'$ it is now enough to demonstrate the inequality $\liminf_{n'} J(f_{n'}) \le J'(f_0)$. By virtue of the identity $J_{g_{m+1}}(f_{n'}) = J'(f_0)$, already shown to hold for all n', it remains to prove that

$$\liminf_{n'} \int_{[0,1]} [g_0(\omega, f_{n'}(\omega), y_{f_{n'}}(\omega)) - g_0(\omega, f_{n'}(\omega), y_{f_0}(\omega))]d\omega \ge 0.$$

We do this by applying the Fatou-Vitali lemma. Observe that

$$\liminf_{n'} g_0(\omega, f_{n'}(\omega), y_{f_{n'}}(\omega)) - g_0(\omega, f_{n'}(\omega), y_{f_0}(\omega)) \ge 0 \text{ for a.e. } \omega,$$

since all values $f_{n'}(\omega)$ lie in $\{s_{i,\omega} : 1 \le i \le m+2\}$ and $g_0(\omega, s_{i,\omega}, \cdot)$ is lower semicontinuous for every $1 \le i \le m+2$. Also, $\epsilon_{f_{n'}} \Longrightarrow \delta'$ implies $\sup_{[0,1]} |y_{f_{n'}} - y_{\delta'}| \to 0$ by Example 5.11. Therefore, for n' sufficiently large $\chi(y_{n'}(\omega)) \le \chi(\sup_{[0,1]} |y_{f_0}| + 1)$. Finally, as we already observed, the sequence $(\alpha h'(|f_{n'}|))$ is uniformly integrable. So the Fatou-Vitali lemma can indeed be applied and we are done. □

In particular, the above example applies to the situation of Example 1.3, which explains our remarks in section 1 on the subject of functional relaxation.

10 Comments

Global comments

The fundamental choice to use K-convergence to extend classical narrow convergence results systematically from the domain of probability measures to that of Young measures, lends these lecture notes a unique form. This contrasts with earlier expositions, such as [Y2, Mc2, War, BL, Ba3, Val3, Val5], which are also considerably less general in several respects and leave several of the subjects and applications discussed here untreated. Not discussed in the text is the following functional-analytic interpretation of Young measures in terms of abstract L^∞-spaces: Let $\hat{S}$ be a compactification of S. One identifies the set $\mathcal{R}(\Omega, \mathcal{A}, \mu; S)$ of Young measures as a subset of the space $\mathcal{L}^\infty(\Omega, \mathcal{A}, \mu; \mathcal{M}(\hat{S}))$, viz., the abstract $\mathcal{L}^\infty$-space of $\mathcal{C}(\hat{S})$-scalarly measurable functions from Ω into $\mathcal{M}(\hat{S})$, the set of all bounded signed measures on $\hat{S}$. Since $\mathcal{L}^\infty_{\mathcal{M}(\hat{S})}$ is the dual space of $\mathcal{L}^1_{\mathcal{C}(\hat{S})}$, the Alaoglu-Bourbaki theorem, combined with tightness – serving to undo the effects of the compactification – easily implies the topological version of the Prohorov-type theorem. Theorem 4.7 itself, i.e., the sequential version, then follows by a standard conditional expectation argument. For compact S this approach was started by J. Warga [War] and for S locally compact and Polish it was pursued by C. Castaing, H. Berliocchi and J.M. Lasry, who used Alexandrov's one point compactification (see [BL, p. 179-180]) and who extended the scope of Young measure theory considerably by showing how integral functionals with lower semicontinuous (measurable) integrands could be included in the theory. In [Ba1, Ba2, Ba3] their treatment was expanded to a metrizable Lusin space S, using Hilbert cube compactification. Subsequently, this approach was further extended to the case of a completely regular Suslin space S [Ba14]. (See also [JM] for an independent treatment of the case where S is Polish.) These lecture notes continue this development by dealing with a separable metric space S and allowing stronger topologies on S in the principal results (as explained in the main text, this includes the case where S is completely regular and Suslin). As an aside, let it be observed that a considerable shortcut is available in case the underlying measure space is topological in nature (but direct contact with classical measure theory is then lost of course). The shortcut can be achieved by use of the Scorza-Dragoni theorem (a Lusin-type theorem). The direct objects of study then become the product measures on the topological space $\Omega \times S$ (cf. Theorem A.1), for which the classical narrow convergence theory is directly available. Results in this vein were obtained by L.C. Young [Y2], E.J. McShane [Mc2], and H. Berliocchi and J.-M. Lasry [BL].

Section 1

A particularly elegant and comparatively short proof of Komlós' Theorem 1.2 by martingale arguments can be found in [Cha1]. For extensive background information on Komlós' theorem, especially on the influence which it has had on the so-called 'subsequence principle', the survey paper [Cha2] should be consulted.

Section 2

The pertinent credits and references for the classical material in this section can be found in P. Billingsley's textbook [Bi1]. Definition 2.12 and Theorem 2.13 go back to Yu.V. Prohorov, with later extensions by V.S. Varadarayan and L.M. LeCam [L2]. Theorem 2.15 is due to L.M. LeCam [L2] (with input from S. Ulam and P.-A. Meyer); observe that without the Suslin assumption this result still holds for a sequence of *Radon* probability measures on the metric space S. The use of the Alexandrov-compactification $\hat{\mathbf{N}}$ in these considerations, which is a recurring technical trick in these notes, was introduced in [Ba4]. The precision reached for the support in the synthesizing Theorem 2.20, by using sequential closures, seems new (the same comment applies to Theorems 5.5 and 6.2).

Section 3

Definition 3.1, the classical definition of a transition probability, was not studied as such by L.C. Young [Y1, Y2]. In contrast, Young's *generalized curves* started out as linear continuous functionals, which correspond to product measures on $\Omega \times S$ – by Riesz's representation theorem – with one marginal fixed (i.e., equal to μ). In statistical decision theory, however, the notion of a *statistical decision function* [Wal, L1, Fe] had precisely the form of a transition probability, and also the narrow convergence Definition 3.3 arose in this area [Wal, L1]. Definition 3.10 of K-convergence was introduced in [Ba16] for quite general scalarly measurable functions, including transition probabilities. Outer integral functionals for Young measures, and the concomitant notion of measurable regularization were introduced in [Ba3, Ba7]. Theorem 3.13 was given in [Ba17]; a flaw in the proof there has been corrected by giving an entirely new proof here. Support properties of narrow limits of Young measures, as given in Corollary 3.18, were first studied in [Ba3], and [Ba17] contains the first version of Theorem 3.11. Theorem 3.15, a corollary of the relative K-compactness Theorem 4.7, was first given in [Ba16, Ba17] for S Polish. The present extension to the metrizable Suslin case is new; the same holds for Corollary 3.16. In view of the comment above about the significance of Theorem 2.15, it would have been possible to replace the Suslin condition for S systematically by a restriction to Young measures mapping into the *Radon* probability measures on the separable metric space S. However, to vary the assumptions on S seemed more natural than to vary the set of Young measures itself (besides, S has to be Suslin anyway when measurable selections are considered). The counterExample 3.17 is new. The Young measure portmanteau Theorem 3.19 has many fathers; it contains elements that can be traced back to at least H. Berliocchi and J.-M. Lasry [BL] and A.D. Ioffe [Iof]. Theorem 3.23 was first given in [Ba11]; the proof given here is particularly simple.

Section 4

Definition 4.1 was introduced in [Ba2, Ba3]; this extends earlier versions of tightness, given by L.M. LeCam [L1] and H. Berliocchi and J.-M. Lasry [BL] for instance. Definition 4.2 and the equivalence Proposition 4.3 are due to M. Valadier and A. Jawhar [Jaw]. The Prohorov-type Theorem 4.7 for K-convergence was first given in [Ba16] (for S completely regular Suslin); Prohorov-type theorems for narrow convergence (both sequential and nonsequential) can be found in [L1, BL] (for S locally compact Polish), in [JM] (for S Polish), and in [Ba1, Ba3, Ba14] (for S metrizable Lusin and S completely regular Suslin respectively). Metrizability of S in Theorem 4.7 might seem retrogressive, in view of [Ba16]; yet this is not the case, because of the novel observation made in Theorem 5.5. To see the relation with [Ba16, Ba14] precisely, observe that Theorem 5.5 applies to a completely regular Suslin space (S, τ), since such a space always has a weaker *metrizable* topology (quite related to the weak metric d_w used in section 6) for which it is still Suslin. A different, somewhat related tightness criterion has been given by C. Castaing [Ca1]. Proposition 4.10 is due to Z. Artstein [Ar2].

Section 5

The growth dominance property (γ) has its roots in the classical notion of coercivity in the calculus of variations; the form presented is as in [BL] and owes much to R.J. Aumann, M. Perles and L. Shapley [AP]. The more general property (γ') is introduced here for the first time. Theorem 5.5 forms a synthesis of material whose history has been described in the above comments regarding section 4; e.g., see [Ba14]. Proposition 5.8, with origins in [AP], comes from H. Berliocchi and J.-M. Lasry [BL] (in first instance, for locally compact Polish S – but already the result for Euclidean S happens to be equivalent to the one with S Suslin; see [Ba4] and section 8).

Section 6

Definitions 6.1 and 6.7 are classical [DU]. Theorem 6.2 is a specialization of Theorem 5.5 to an L^1-setting; forerunners can be found in [BL, Ba3, Ba8]. Compactness criteria like Theorem 6.8 go back to N. Dunford and B.J. Pettis. The connection with Prohorov-type narrow compactness criteria for Young measures was first made by H. Berliocchi and J.-M. Lasry [BL] (for finite-dimensional E) and later in [Ba8] (for a reflexive Banach space E) and [BH]. Of course, extensions of the Dunford-Pettis criterion have also been obtained without Young measures (and much earlier); without any attempt to be fair or exhaustive (and disregarding the related work in Orlicz spaces), I would mention [DU, CV, Ca3, ACV, Ba16] and their references. As indicated in the main text, Theorem 6.9 originated with V. Jalby [Jal], who proved s-relative compactness of the sequence $(\int_A f_{n'})$ for each A. The present formulation refines this result, and its proof is completely new Definition 6.10 is from [BC]; here the main idea, as expressed by Theorem 6.11, is due to V.F. Gaposhkin [Gap]. Definition 6.14 of the Bocce criterion is due to M. Girardi [Gi]. Definition 6.17 of limited convergence is was first given in [Ba9]. Theorems 6.19 (and 6.21 as well) originated with A. Visintin, and was subsequently improved in [Ba9, Ba17]; the present version continues these improvements further. See [Val4] for more details.

Section 7

The lower semicontinuity Theorem 7.1 has a long history, which goes back to L. Tonelli [ET, Ce]; versions of this result can be discerned in the work of L.C. Young and E.J. McShane as well. Also H. Berliocchi and J.-M. Lasry [BL] proved similar results by means of Young measures. All the above authors suppose that the Banach space E is finite-dimensional. The first version of the lower semicontinuity result Theorem 7.1 in infinite dimensions appears to have been given in [BO] (for a reflexive Banach space E), and the first lower closure result of this type seems to have been given in [Ba8] (also with E reflexive). The use of w^2-convergence to unify lower closure and lower semicontinuity results, as in Corollaries 7.2 and 7.4, was suggested in [Ba3, p. 588] and worked out in [Ba13].

Section 8

The gist of Theorem 8.2 can be found in many works on optimal control (e.g. [Ce]). The Fatou-Vitali Theorem 8.6 in several dimensions is essentially as in [Ba4, Ba3] (see also Corollary 8.8); the present version is from [Ba21]. Lemma 8.7, the main tool section 8, is from [Ba21], where one can find a more complicated proof, based on using Choquet-type representations of Young measures. The original Fatou result in several dimensions of D. Schmeidler [Schm] is subsumed by Corollary 8.8. Corollary 8.9 is due to Z. Artstein [Ar2].

Section 9

The existence Theorem 9.1 was originally given by R.J. Aumann and M. Perles [AP] (without using Young measures). Their result inspired H. Berliocchi and J.-M. Lasry [BL] and the present author [Ba1, Ba4] into proofs which use Young measures. A different approach, which avoids Young measures and is more in the spirit of [AP], was given by Z. Artstein [Ar1]. The denseness Theorem 9.3 is classical in Young measure theory and goes back to approximation arguments of L.C. Young. See J. Warga's textbook [War] for several applications (with S compact) and see [Ba6] for a very general topological version. The limiting bang-bang Theorem 9.4 generalizes a result of M. Valadier [Val1]. The method of proof is new. Although only stated for a concrete example, the relaxation Example 9.5, which goes back to work by I. Ekeland [E] and H. Berliocchi and J.-M. Lasry [BL], improves upon [Val2] in one significant aspect: it only needs lower semicontinuity in the state variable. The proof contains several new elements, not in the least its use of Theorem A.14 and its application of Theorem A.3 to certain marginal measures (this simplifies earlier proofs given in [BL, ET, Val2, Val1]).

Appendix A

All results are known, except for Theorem A.14 and its corollary. To my own great surprise I have not been able to locate references for Corollary A.11, apart from [Ba21, Proposition 3.2] and an unproven statement in [Fl, Proposition 2.4] (for finite-dimensional Ω). The proof given in [Ba21] deduces Corollary A.11 from the analogue of Lemma 8.7 (which in turn is proven by a Choquet-type representation for Young measures that is not discussed here). This is much more complicated than the the proof given here. Theorem A.14 appears not to have appeared before (I have left it unpublished for a long time – the proof uses a duality/perturbation scheme with compactifications in the spirit of [Ba10]). It is remarkable that the connection of such a result with duality was apparently not made in [ET] (a textbook on duality theory).

Appendix B

The results in this appendix ought to be well-known; I was simply unable to find a suitable reference for all the results on outer integration used in the main text.

E.J.B.

A Auxiliary results

We recall and derive some results from measure theory and convex analysis which play a role in the main text. Our first result is a Fubini-type theorem from [N, III.2] (see also [As, 2.6]). As in the main text, $(\Omega, \mathcal{A}, \mu)$ is a finite measure space, S a separable metric space and $(E, \|\cdot\|)$ a separable Banach space.

Theorem A.1 (Fubini-Tonelli) *For any $\delta \in \mathcal{R}(\Omega, \mathcal{A}, \mu; S)$ the formula*

$$[\mu \otimes \delta](A \times B) := \int_A \delta(\omega)(B)\, \mu(d\omega)$$

defines a unique product measure $\mu \otimes \delta$ on $(\Omega \times S, \mathcal{A} \times \mathcal{B}(S))$. Moreover, for every $\mathcal{A} \times \mathcal{B}(S)$-measurable function $g : \Omega \times S \to [0, +\infty]$

$$\omega \mapsto \int_S g(\omega, x)\delta(\omega)(dx) \text{ is } \mathcal{A}\text{-measurable}$$

and

$$\int_{\Omega \times S} g\, d(\mu \otimes \delta) = \int_\Omega [\int_S g(\omega, x)\delta(\omega)(dx)]\mu(d\omega).$$

Proposition A.2 *Let $\delta : \Omega \to \mathcal{P}(S)$. The following are equivalent:*
(a) $\delta \in \mathcal{R}(\Omega, \mathcal{A}, \mu; S)$.
(b) δ is measurable with respect to $\mathcal{A}$ and the narrow Borel σ-algebra on $\mathcal{P}(S)$.

Proof. $(a) \Rightarrow (b)$: For every $c \in \mathcal{C}_b(S)$ the mapping $\omega \mapsto \int_S c(x)\, \delta(\omega)(dx)$ is $\mathcal{A}$-measurable by Theorem A.1. Since $\mathcal{P}(S)$ is separable and metrizable for the narrow convergence topology [DM, III.60], (b) follows elementarily.

$(b) \Rightarrow (a)$: For any open set $G \subset S$ there exists a nondecreasing sequence (c_n) in $\mathcal{C}_b(S)$ such that $\lim_n c_n(x) = 1_G(x)$ for every $x \in S$ [As, A6]. Hence, $\delta(\cdot)(G)$ is $\mathcal{A}$-measurable by an application of the monotone convergence theorem. Since finite intersections of open sets are open, (a) follows by an application of a σ-additive class result [As, 4.1.2]. □

We frequently use the following classical criterion for uniform integrability [DM, II.22]:

Theorem A.3 (de la Vallée-Poussin's criterion) *Let (ϕ_n) in $\mathcal{L}^1_{\mathbf{R}}$. The following are equivalent:*
(a) (ϕ_n) is uniformly integrable.
(b) There exists a function $h' : \mathbf{R}_+ \to \mathbf{R}_+$, convex, continuous and nondecreasing, such that $\lim_{\xi\to\infty} h'(\xi)/\xi = +\infty$ and

$$\sup_n \int_\Omega h'(|\phi_n(\omega)|)\mu(d\omega) < +\infty.$$

The following implicit measurable function result is taken from [CV, Theorem III.38].

Theorem A.4 (implicit measurable selection) *Let $(U, \mathcal{U})$ be a measurable space, S a Suslin space, and $\Theta : \Omega \to 2^U$ a multifunction whose graph*

$$\text{gph}\, \Theta := \{(\omega, u) \in \Omega \times U : u \in \Theta(\omega)\}$$

belongs to $\mathcal{A} \times \mathcal{B}(U)$. Let $g : \Omega \times S \to U$ be measurable with respect to $\mathcal{A} \times \mathcal{B}(S)$ and $\mathcal{U}$ such that $g(\omega, S) \cap \Theta(\omega) \neq \emptyset$ for a.e. ω. Then there exists $f \in \mathcal{L}^0_S$ such that $g(\omega, f(\omega)) \in \Theta(\omega)$ for a.e. ω in Ω.

Ky Fan's inequality is as follows [AE, Theorem 5, p. 330]:

Theorem A.5 *Let K be a compact convex subset of a topological vector space and let $\chi : K \times K \to \mathbf{R}$ be such that (i) χ is lower semicontinuous, (ii) $\chi(\cdot, z)$ is lower semicontinuous for every $z \in K$ and (iii) $\chi(z', \cdot)$ is concave for every $z' \in K$. Then there exists $z_* \in K$ such that*

$$\sup_{z \in K} \chi(z_*, z) \leq \sup_{z \in K} \chi(z, z).$$

Next, we give some Lyapunov-type results which lead up to the instrumental Corollary A.11.

Definition A.6 (atom) An *atom* of $(\Omega, \mathcal{A}, \mu)$ is a set $A \in \mathcal{A}$, $\mu(A) > 0$, for which there exists no $B \in \mathcal{A}$, $B \subset A$, such that $0 < \mu(B) < \mu(A)$. □

Note that as atoms we only accept nonnull sets. It is elementary to check that any $\mathcal{A}$-measurable function must be a.e. *constant* on any atom of $(\Omega, \mathcal{A}, \mu)$.

Proposition A.7 (decomposition property) *There exists an at most countable collection $(\bar{A}_j)$ of atoms of $(\Omega, \mathcal{A}, \mu)$, such that $\Omega^{na} := \Omega \setminus \cup_j \bar{A}_j$ contains no atoms.*

PROOF. For each $i \in \mathbf{N}$ there can be at most i atoms whose μ-measure is at least $\mu(\Omega)/i$. This gives the desired collection $(\bar{A}_j)$. □

If $\Omega = \Omega^{na}$ then $(\Omega, \mathcal{A}, \mu)$ is said to be *nonatomic*. The most important property of nonatomic measure spaces is as follows [CV, p. 118 ff.].

Theorem A.8 (Lyapunov's theorem) *Let $q \in \mathbf{N}$ and let $f \in \mathcal{L}^1_{\mathbf{R}^q}$. If Ω is nonatomic, then*

$$C := \{\int_A f \, d\mu : A \in \mathcal{A}\} = \{\int_\Omega f\alpha \, d\mu : \alpha \in \mathcal{L}^\infty_{\mathbf{R}}, 0 \leq \alpha \leq 1\}.$$

Corollary A.9 *Let $m, r \in \mathbf{N}$, let $f_1, \cdots, f_r$ be functions in $\mathcal{L}^1_{\mathbf{R}^m}$ and let $\alpha_1, \cdots, \alpha_r$ be nonnegative functions in $\mathcal{L}^\infty_{\mathbf{R}}$, with $\sum_{i=1}^r \alpha_i(\omega) = 1$ for all ω. If Ω is nonatomic, then there exists a measurable partition $B_1, \cdots, B_r$ of Ω such that*

$$\int_\Omega \sum_{i=1}^r \alpha_i f_i \, d\mu = \sum_{i=1}^r \int_{B_i} f_i \, d\mu.$$

PROOF. We use induction for r. For $r = 1$ the result holds trivially. Suppose it is true for $r = k - 1$. Denote $\sum_1^{k-1} \alpha_i' f_i$ by g, where $\alpha_i'(\omega) := \alpha_i(\omega)/(1 - \alpha_k(\omega))$ if $\alpha_k(\omega) < 1$ and $\alpha_i'(\omega) := 0$ if $\alpha_k(\omega) = 1$. By Theorem A.8, there exists $A \in \mathcal{A}$ for which $\int_A (g, f_k) = \int_\Omega \alpha_k (g, f_k)$. This gives $\int_{\Omega \setminus A} g = \int_\Omega (1 - \alpha_k) g$, so now the result follows by the induction step applied to the functions $f_i 1_{\Omega \setminus A}$. □

Consequently, the following result is immediate:

Corollary A.10 (Richter's theorem) *Let $F : \Omega \to 2^{\mathbf{R}^d}$ be an arbitrary multifunction. If Ω is nonatomic, then the set $\int_\Omega F \, d\mu$ is convex.*

The next result is [Ba21, Proposition 3.2], which extends Corollary A.9: the important fact to observe is that the participating functions are no longer supposed integrable.

Corollary A.11 (extended Lyapunov theorem) *Let $m, r \in \mathbf{N}$, let $f_1, \cdots, f_r$ be functions in $\mathcal{L}^0_{\mathbf{R}^m}$ and let $\alpha_1, \cdots, \alpha_r$ be nonnegative functions in $\mathcal{L}^\infty_{\mathbf{R}}$, with $\sum_{i=1}^r \alpha_i(\omega) = 1$ for all ω, such that*

$$\int_\Omega \sum_{i=1}^r \alpha_i |f_i| \, d\mu < +\infty.$$

If Ω is nonatomic, then there exists a measurable partition $B_1, \cdots, B_r$ of Ω such that

$$\int_\Omega \sum_{i=1}^r \alpha_i f_i \, d\mu = \sum_{i=1}^r \int_{B_i} f_i \, d\mu.$$

Proof. Define for every $p \in \mathbf{N}$ the set Ω_p to consist of all ω for which $\max_i |f_i(\omega)|$ belongs to $(p, p-1]$. Then the Ω_p are disjoint and on each Ω_p we can apply Corollary A.9. For each p this gives the existence of a measurable partition $B_{1,p}, \cdots, B_{r,p}$ such that

$$\int_{\Omega_p} \sum_{i=1}^r \alpha_i f_i = \sum_{i=1}^r \int_{B_{i,p}} f_i.$$

Hence, the result follows by

$$\int_\Omega \sum_{i=1}^r \alpha_i f_i = \sum_p \int_{\Omega_p} \sum_{i=1}^r \alpha_i f_i = \sum_{i=1}^r \int_{\cup_p B_{i,p}} f_i,$$

since for each i the $B_{i,p}$ are disjoint. □

The following result forms a trick to reduce arguments involving at most countably many product measurable sets or functions to a marginally countably generated situation; cf. [CV, p. 78] and [Val1, Appendix].

Proposition A.12 *Let $(U, \mathcal{U})$ be a measurable space and let $g : \Omega \times U \to [-\infty, +\infty]$ be $\mathcal{A} \times \mathcal{U}$-measurable. Then there exists a countably generated sub-σ-algebra $\mathcal{A}_0$ of $\mathcal{A}$ such that g is also $\mathcal{A}_0 \times \mathcal{U}$-measurable. Moreover, if $(\Omega, \mathcal{A}, \mu)$ is nonatomic, then $\mathcal{A}_0$ can be chosen in such a way as to make $(\Omega, \mathcal{A}_0, \mu)$ nonatomic.*

Proof. If $g = 1_G$, $G \in \mathcal{A} \times \mathcal{U}$, then it suffices to observe that the union of all σ-algebra's $\mathcal{A}_0 \times \mathcal{U}$, $\mathcal{A}_0$ a countably generated sub-σ-algebra of $\mathcal{A}$, is a σ-algebra which must coincide with $\mathcal{A} \times \mathcal{U}$. The usual approximation by a sequence of simple functions then finishes the argument for general g. In addition, if $\mathcal{A}$ is nonatomic, then let $(\bar{A}_j)$ be an enumeration of the atoms of $\mathcal{A}_0$, just as in Proposition A.7. By nonatomicity of $\mathcal{A}$, for each $m \in \mathbf{N}$ each $\mathcal{A}_0$-atom $\bar{A}_j$ can be partitioned as $\bar{A}_j = \cup_{i=1}^m B_i^{m,j}$, with $\mu(B_i^{m,j}) \leq \mu(\bar{A}_j)/m$, $1 \leq i \leq m$. Now let $\mathcal{A}_1$ be the σ-algebra generated by $\mathcal{A}_0$ and all $B_i^{m,j}$. Suppose that A is an atom of $\mathcal{A}_1$. Of course, we can only have $\mu(A \cap (\Omega \backslash \cup_j \bar{A}_j)) > 0$ if $\mu(A) = \mu(A \cap (\Omega \backslash \cup_j \bar{A}_j))$. But this implies that, modulo a null set, the $\mathcal{A}_1$- and $\mathcal{A}_0$-atom A is contained in $\Omega \backslash \cup_j \bar{A}_j$, which is the nonatomic part of $(\Omega, \mathcal{A}_0, \mu)$ (cf. Proposition A.7). Therefore, it follows that $\mu(A \cap (\Omega \backslash \cup_j \bar{A}_j)) = 0$, i.e., A is essentially contained in $\cup_j \bar{A}_j$. Hence, for every $m \in \mathbf{N}$ there must be j and i, $1 \leq i \leq m$, with $\mu(A \cap B_i^{m,j}) > 0$. But since A is an atom this implies then $\mu(A) = \mu(B_i^{m,j}) \leq \mu(\bar{A}_j)/m \leq \mu(\Omega)/m$. So $\mu(A) = 0$, in contradiction to our Definition A.6. □

The following well-known property of barycenters is frequently used in the main text.

Theorem A.13 (barycenters and convexity) *Let $\nu \in \mathcal{P}(E)$ be such that $\int_E \|x\|\nu(dx) < +\infty$.*

(i) If $C \subset E$ is closed and convex with $\nu(C) = 1$ then

$$\text{bar } \nu := \int_E x\, \nu(dx) \text{ lies in } C.$$

(ii) If $E = \mathbf{R}^d$ and if $C \subset \mathbf{R}^d$ is convex – possibly nonmeasurable – with outer measure $\nu^*(C) = 1$ then

$$\text{bar } \nu := \int_{\mathbf{R}^d} x\, \nu(dx) \text{ lies in } C.$$

PROOF. From elementary facts about Bochner integration it follows that bar ν is well-defined. Part (a) follows directly from the Hahn-Banach theorem [War, I.6.13], and part (b) is proven by induction for the dimension of E (see [Pf, Lemma] and [Fe, Lemma 3, p. 74]). □

The next series of results generalizes a well-known result from Choquet analysis; the main proof, which depends upon a duality-perturbation scheme using compactifications, is quite in the spirit of [Ba10].

Theorem A.14 (barycenters and convexification) *Let $q : E \to (-\infty, +\infty]$ be lower semicontinuous and such that* dom $q^* := \{q^* < +\infty\}$ *is nonempty. Let q^{**} be the Fenchel biconjugate of q. Then for every $x_0 \in E$ with $q^{**}(x_0) < +\infty$ there exists $\nu \in \mathcal{P}(E)$ such that*

$$q^{**}(x_0) = \sup_{x^* \in \text{dom } q^*} \int_E [q(x) + < x_0 - x, x^* >] \nu(dx).$$

PROOF. It follows immediately from the definition of biconjugate that

$$-q^{**}(x_0) = \inf_{x^* \in \text{dom } q^*, \alpha \in \mathbf{R}, < \cdot, x^* > + \alpha \leq q} [- < x_0, x^* > - \alpha].$$

Let $\hat{E}$ be the Hilbert cube compactification of E [Cho, 6.3]; then we may identify E as a subset of $\hat{E}$. Since $(E, \|\cdot\|)$ is a Polish space, E is a G_δ-set in $\hat{E}$; hence we get $E \in \mathcal{B}(\hat{E})$ after identification; cf. [DM, p. 118]. Consider for $\hat{c} \in \mathcal{C}(\hat{E})$ the perturbed optimization problem

$$(\mathcal{P}_{\hat{c}}) : \inf_{x^* \in \text{dom } q^*, \alpha \in \mathbf{R}, < \cdot, x^* > + \alpha \leq q - \hat{c}|_E} [- < x_0, x^* > - \alpha].$$

Define $p(\hat{c}) := \inf(\mathcal{P}_{\hat{c}})$. Since dom q^* is supposed nonempty, there exist $\bar{x}^* \in E^*$ and $\bar{\alpha} \in \mathbf{R}$ such that $q \geq < \cdot, \bar{x}^* > + \bar{\alpha}$ on E. This gives $p(\hat{c}) \leq - < x_0, \bar{x}^* > - \bar{\alpha} + \sup_{\hat{E}} |\hat{c}|$ for all $\hat{c} \in \mathcal{C}(\hat{E})$, so p is certainly continuous and finite at the origin (note that $p(0) = -q^{**}(x_0) > -\infty$). In the terminology of [ET], this means that $(\mathcal{P})$ is *stable* for the chosen perturbation scheme. The corresponding dual problem is

$$(\mathcal{P}^*) : \sup_{\hat{\nu} \in \mathcal{M}(\hat{E})} -p^*(\hat{\nu}),$$

by the Riesz representation theorem [DM, III.35]. Here $\mathcal{M}(\hat{E})$ stands for the set of all bounded signed measures on the compact and metrizable space $\hat{E}$. Let us calculate p^*; we have

$$p^*(\hat{\nu}) = \sup_{\hat{c} \in \mathcal{C}(\hat{E}), x^* \in \text{dom } q^*, \alpha \in \mathbf{R}, < \cdot, x^* > + \alpha \leq q - \hat{c}|_E} [\int_{\hat{E}} \hat{c}\, d\hat{\nu} + < x_0, x^* > + \alpha].$$

We now claim that $p^*(\hat{\nu}) = +\infty$, unless there is a probability measure $\nu \in \mathcal{P}(E)$ such that $\hat{\nu}(\cdot) = \nu(E \cap \cdot)$. If this is true, then $(\mathcal{P}^*)$ is equivalent to an optimization problem over $\mathcal{P}(E)$. Indeed, if $\hat{\nu}(\hat{E} \backslash E) \neq 0$ then $p^*(\hat{\nu}) = +\infty$, and if $\hat{\nu}$ is negative, then there exists a nonpositive $\hat{c} \in \mathcal{C}(\hat{E})$ with $\int_{\hat{E}} \hat{c}\, d\hat{\nu} > 0$. Then taking positive multiples of $\hat{c}$ shows $p^*(\hat{\nu}) = +\infty$. Continuing, if $\hat{\nu}(\hat{E} \backslash E) = 0$ and if $\hat{\nu}$ is nonnegative, then

$$p^*(\hat{\nu}) = \sup_{x^* \in \text{dom } q^*, \alpha \in \mathbf{R}} [\int_E (q(x) - < x, x^* > -\alpha)\, \hat{\nu}(dx) + < x_0, x^* > +\alpha]. \tag{A.1}$$

This follows from defining $(\hat{c}_n) \subset \mathcal{C}(\hat{E})$ by

$$\hat{c}_n(\hat{x}) := \inf_{x \in E} [n \hat{\rho}(x, \hat{x}) + q'(x)],$$

where $q' := q - < \cdot, x^* > -\alpha$ is lower semicontinuous and bounded from below and where $\hat{\rho}$ stands for any choice of a metric on $\hat{E}$. For then the proof of [As, A6.6] can be mimicked completely to give $\hat{c}_n \uparrow q$ on E, so (A.1) follows by the monotone convergence theorem. Clearly, (A.1) implies that if $\hat{\nu}(\hat{E} \backslash E) = 0$ and if $\hat{\nu}$ is nonnegative, then $p^*(\hat{\nu}) = +\infty$, unless $\hat{\nu}(E) = 1$. Now the claim follows from the resulting simplification of (A.1). Finally, it remains to apply the duality-stability result of [ET, Proposition 2.2, p. 51]; this gives $\inf(\mathcal{P}) = \max(\mathcal{P}^*)$, so the result follows. □

Corollary A.15 *Let $q : E \to (-\infty, +\infty]$ be such that $q + < \cdot, x^* >$ is inf-compact for every $x^* \in E^*$. Then for every $x_0 \in E$ with $q^{**}(x_0) < +\infty$ there exists $\nu \in \mathcal{P}(E)$ such that*

$$q^{**}(x_0) = \int_E q(x) \nu(dx)$$

and such that $\int_E | < x, x^* > | \nu(dx) < +\infty$ and $\int_E < x, x^* > \nu(dx) =< x_0, x^* >$ for every $x^* \in E^*$.

Proof. By the given inf-compactness property, $q + < \cdot, x^* >$ attains its minimum over E for each $x^* \in E^*$. Hence, dom $q^* = E^*$. Therefore, the probability measure ν furnished by Theorem A.14 satisfies for each $x^* \in E^*$

$$\int_E [-q^*(0) + < x_0 - x, x^* >] \nu(dx) \leq q^{**}(x_0) < +\infty;$$

observe that this implies $\int_E \max(- < x^*, x >, 0)\, \nu(dx) < +\infty$.[14] Replacing x^* by $-x^*$, we see that $\int_E | < x, x^* > | \nu(dx) < +\infty$ for every $x^* \in E^*$; the stated property of ν now follows easily from the identity in Theorem A.14 by taking scalar multiples. □

[14] As everywhere else, the integration convention from Lemma B.2 is in force.

B Outer integration

Definition B.1 (outer integration) Let $\psi : \Omega \to [-\infty, +\infty]$ be arbitrary (possibly nonmeasurable). Then the *outer integral* $\int_\Omega^* \psi \, d\mu$ is defined by

$$\int_\Omega^* \psi \, d\mu := \inf\{\int_\Omega \phi \, d\mu : \phi \in \mathcal{L}^1_{\mathbf{R}}, \phi \geq \psi \text{ on } \Omega\},$$

where the infimum over the empty set is set equal to $+\infty$. □

Lemma B.2 (outer and ordinary integration) *Let $\psi : \Omega \to [-\infty, +\infty]$ be $\mathcal{A}$-measurable. Then*

$$\int_\Omega^* \psi \, d\mu = \int_\Omega \psi \, d\mu := \int_\Omega \psi^+ \, d\mu \dot{-} \int_\Omega \psi^- \, d\mu,$$

where $\psi^+ := \max(\psi, 0)$, $\psi^- := \max(-\psi, 0)$ and $\dot{-}$ is as ordinary subtraction, but with the additional convention $(+\infty) \dot{-} (+\infty) := +\infty$.

Proof. If $\int_\Omega \psi^+ = +\infty$, the result is immediate (the infimum in Definition B.1 is then taken over the empty set).

So suppose $\int_\Omega \psi^+ < +\infty$. Note that $\int_\Omega \phi \geq \int_\Omega \psi$ for every ϕ participating in the infimum in Definition B.1. Hence, $\int_\Omega^* \psi \geq \int_\Omega \psi$. Now if $\int_\Omega \psi^- < +\infty$, then $\psi \in \mathcal{L}^1_{\mathbf{R}}$, so Definition B.1 implies that $\int_\Omega^* \psi \leq \int_\Omega \psi$, which finishes the argument. And if $\int_\Omega \psi^- = +\infty$, then an obvious argument with the sequence $\phi_n := \psi^+ - \min(\psi^-, n)$ shows that $\int_\Omega^* \psi = -\infty = \int_\Omega \psi$. □

Lemma B.3 (attainment outer integral) *Let $\psi : \Omega \to [-\infty, +\infty]$ (possibly nonmeasurable) and $\phi \in \mathcal{L}^1_{\mathbf{R}}$ be such that $\psi \geq \phi$ on Ω and $\int_\Omega^* \psi \, d\mu < +\infty$. Then there exists $\tilde{\phi} \in \mathcal{L}^1_{\mathbf{R}}$, $\tilde{\phi} \geq \phi$, such that $\int_\Omega \tilde{\phi} \, d\mu = \int_\Omega^* \psi \, d\mu$.*

Proof. By Definition B.1 there exists a sequence (ϕ_k) in $\mathcal{L}^1_{\mathbf{R}}$ such that $\phi_k \geq \psi$ and $\int_\Omega \phi_k \leq \int_\Omega^* \psi + k^{-1}$ for all k. Define $\tilde{\phi} := \inf_k \phi_k$; then $\phi_1 \geq \tilde{\phi} \geq \psi \geq \phi$ (hence $\tilde{\phi} \in \mathcal{L}^1_{\mathbf{R}}$) and $\int_\Omega \tilde{\phi} \leq \int_\Omega^* \psi$. The converse inequality is trivial. □

Proposition B.4 (Fatou-Vitali for outer integrals) *Let (ψ_n) be a sequence of (possibly nonmeasurable) functions $\psi_n : \Omega \to [-\infty, +\infty]$ such that there exists a uniformly integrable sequence (ϕ_n) in $\mathcal{L}^1_{\mathbf{R}}$ for which for every $n \in \mathbf{N}$*

$$\psi_n(\omega) \geq \phi_n(\omega) \text{ for all } \omega \in \Omega.$$

Then

$$\liminf_n \int_\Omega^* \psi_n(\omega)\mu(d\omega) \geq \int_\Omega^* \liminf_n \psi_n(\omega)\mu(d\omega).$$

Proof. By Lemma B.3, for each n there exists $\tilde{\phi}_n \in \mathcal{L}^1_{\mathbf{R}}$ such that $\tilde{\phi}_n \geq \psi_n \geq \phi_n$ and $\int_\Omega \tilde{\phi}_n = \int_\Omega^* \psi_n$. By uniform integrability of (ϕ_n), the classical Fatou-Vitali lemma [As, 7.5.2] applies. This gives

$$\liminf_n \int_\Omega^* \psi_n = \liminf_n \int_\Omega \tilde{\phi}_n \geq \int_\Omega \liminf_n \tilde{\phi}_n.$$

Since $\liminf_n \tilde{\phi}_n \geq \liminf_n \psi_n$, Definition B.1 gives $\int_\Omega^* \liminf_n \tilde{\phi}_n \geq \int_\Omega^* \liminf_n \psi_n$. Since $\liminf_n \tilde{\phi}_n$ is $\mathcal{A}$-measurable, Lemma B.2 applies, and the result follows. □

Lemma B.5 (subadditivity outer integration) *Let $\psi, \psi' : \Omega \to [-\infty, +\infty]$ be arbitrary (possibly nonmeasurable). Then*

$$\int_\Omega^* \psi \, d\mu \dotplus \int_\Omega^* \psi' \, d\mu \geq \int_\Omega^* (\psi \dotplus \psi') d\mu,$$

where $\dotplus$ is defined just as ordinary addition, but with $(-\infty) \dotplus (+\infty) := +\infty$ as an additional convention.

Proof. If either term on the left is equal to $+\infty$, the result is trivially true. So suppose that $\int_\Omega^* \psi \, d\mu < +\infty$ and $\int_\Omega^* \psi' \, d\mu < +\infty$ (hence both ψ and ψ' are a.e. not equal to $+\infty$). By Definition B.1, there exist sequences (ϕ_n) and (ϕ'_n) in $\mathcal{L}^1_{\mathbf{R}}$ such that $\int_\Omega \phi_n \to \int_\Omega^* \psi$ and $\int_\Omega \phi'_n \to \int_\Omega^* \psi'$, with $\phi_n \geq \psi$ and $\phi'_n \geq \psi'$. But then simple work with $(\phi_n + \phi'_n)$ gives the inequality immediately. □

Lemma B.6 *Let $\psi : \Omega \to [-\infty, +\infty]$ be arbitrary and let $\phi \in \mathcal{L}^1_{\mathbf{R}}$. Then*

$$\int_\Omega^* (\psi + \phi) d\mu = \int_\Omega^* \psi \, d\mu + \int_\Omega \phi \, d\mu.$$

Proof. An elementary consequence of Definition B.1. □

References

[ACV] Amrani, A., Castaing, C. and Valadier, M. Méthodes de troncature appliquées à des problèmes de convergence faible ou forte dans L^1. *Arch. Rational Mech. Anal.* **117** (1992), 167-191.

[Ar1] Artstein, Z. On a variational problem. *J. Math. Anal. Appl.* **45** (1974), 404-415.

[Ar2] — A note on Fatou's lemma in several dimensions. *J. Math. Econom.* **6** (1979), 277-282.

[As] Ash, R.B. *Real Analysis and Probability.* Academic Press, New York, 1972.

[AE] Aubin, J.-P. and Ekeland, I. *Applied Nonlinear Analysis.* Wiley, New York, 1984.

[Au] Aumann, R.J. Integrals of set-valued functions. *J. Math. Anal. Appl.* **12** (1965), 1-12.

[AP] — and Perles, M. A variational problem arising in economics. *J. Math. Anal. Appl.* **11** (1965), 488-503.

[Ba1] Balder, E.J. On a useful compactification for optimal control problems. *J. Math. Anal. Appl.* **72** (1979), 391-398.

[Ba2] — Mathematical foundations of statistical decision theory: A modern viewpoint I, Preprint No. 199, Mathematical Institute, University of Utrecht, 1981.

[Ba3] — A general approach to lower semicontinuity and lower closure in optimal control theory. *SIAM J. Control Optim.* **22** (1984), 570-598.

[Ba4] — A unifying note on Fatou's lemma in several dimensions. *Math. Oper. Res.* **9** (1984), 267-275.

[Ba5] — Existence results without convexity conditions for general problems of optimal control with singular components. *J. Math. Anal. Appl.* **101** (1984), 527-539.

[Ba6] — A general denseness result for relaxed control theory. *Bull. Austral. Math. Soc.* **30** (1984), 463-475.

[Ba7] — Lower closure for orientor fields by lower semicontinuity of outer integral functionals. *Ann. Mat. Pura Appl.* (IV) **139** (1985), 349-360.

[Ba8] — An extension of Prohorov's theorem for transition probabilities with applications to infinite-dimensional lower closure problems. *Rend. Circ. Mat. Palermo* **34** (1985), 427-447.

[Ba9] — On weak convergence implying strong convergence in L_1-spaces. *Bull. Austral. Math. Soc.* **33** (1986), 363-368.

[Ba10] — On a general duality result for moment problems. *Optimization* **18** (1987), 17-21.

[Ba11] — Generalized equilibrium results for games with incomplete information. *Math. Oper. Res.* **13** (1988), 265-276.

[Ba12] — Fatou's lemma in infinite dimensions. *J. Math. Anal. Appl.* **136** (1988), 450-465.

[Ba13] — On infinite-horizon lower closure results for optimal control. *Ann. Mat. Pura Appl.* (IV) **151** (1988), 239-246.

[Ba14] — On Prohorov's theorem for transition probabilities. *Sém. Analyse Convexe* **19** (1989), 9.1-9.11.

[Ba15] — Unusual applications of a.e. convergence. In: *Almost Everywhere Convergence.* (G.A. Edgar and L. Sucheston, eds.) Academic Press, New York, 1989, pp. 31-53.

[Ba16] — New sequential compactness results for spaces of scalarly integrable functions. *J. Math. Anal. Appl.* **151** (1990), 1-16.

[Ba17] — On equivalence of strong and weak convergence in L_1-spaces under extreme point conditions. *Israel J. Math.* **75** (1991), 21-48.

[Ba18] — On Cournot-Nash equilibrium distributions for games with differential information and discontinuous payoffs. *Econ. Theory* **1** (1991), 339-354.

[Ba19] — Existence of optimal solutions for control and variational problems with recursive objectives. *J. Math. Anal. Appl.* **178** (1993), 418-437.

[Ba20] — New existence results for optimal controls in the absence of convexity: the importance of extremality. *SIAM J. Control Optim.* **32** (1994), 890-916.

[Ba21] — A unified approach to several results involving integrals of multifunctions. *Set-valued Anal.* **2** (1994), 63-75.

[Ba22] — A unifying approach to existence of Nash equilibria. *Intern. J. Game Theory*, in press.

[Ba23] — On the existence of optimal contract mechanisms for incomplete information principal-agent models. *J. Econ. Theory*, to appear.

[BGJ] — , Girardi, M. and Jalby, V. From weak to strong types of $\mathcal{L}^1_E$-convergence by the Bocce-criterion. *Studia Math.* **111** (1994), 241-262.

[BH] — and Hess, C. Fatou's lemma for multifunctions with unbounded values. *Math. Oper. Res.* **20** (1995), in press.

[BR] — and Rustichini, A. An equilibrium result for games with private information and infinitely many players. *J. Econ. Theory* **62** (1994), 385-393.

[BJ] Ball, J.M. and James, R.D. Finite phase mixtures as minimizers of energy. *Arch. Rat. Mech. Anal.* **100** (1987), 13-52.

[BK] — and Knowles, G. Young measures and minimization problems of mechanics. in : *Elasticity, Mathematical Methods and Applications* (G. Eason and R.W. Ogden, eds.) Ellis Horwood, Chichester, UK, 1990, pp. 1-20.

[BL] Berliocchi, H. and Lasry, J.-M. Intégrandes normales et mesures paramétrées en calcul des variations. *Bull. Soc. Math. France* **101** (1973), 129-184.

[BS] Bertsekas, D.P. and Shreve, S.E. *Stochastic Optimal Control: the Discrete Time Case.* Academic Press, New York, 1978.

[Bi1] Billingsley, P. *Convergence of Probability Measures.* Wiley, New York, 1968.

[Bi2] — *Probability and Measure.* Wiley, New York, 1986.

[BO] Bottaro, G. and Opezzi, G. Semicontinuità inferiore di un funzionale integrale dipendente da funzionali di classe L^p a valori in un spazio di Banach. *Ann. Mat. Pura Appl.* (IV) **123** (1980), 16-26.

[Bo] Bourbaki, N. *Intégration*, Chap. IX. Hermann, Paris, 1969.

[BC] Brooks, J. and Chacon, R.V. Continuity and compactness of measures. *Adv. in Math.* **37** (1980), 16-26.

[BD] — and Dinculeanu, N. Weak compactness in spaces of Bochner integrable functions and applications. *Adv. in Math.* **24** (1977), 172-188.

[Bu] Buttazzo, G. *Semicontinuity, Relaxation and Integral Representation in the Calculus of Variations.* Pitman Research Notes in Math. **207**, Longman, Harlow (Essex), 1989.

[Ca1] Castaing, C. Compacité dans l'espace des mesures de probabilité de transition. *Atti Sem. Mat. Fis. Univ. Modena* **34** (1985-86), 337-351.

[Ca2] — Convergence faible et sections extrémales. *Sém. Analyse Convexe* **18** (1988), 2.1-2.18.

[Ca3] — Méthode de compacité et de décomposition, applications: minimisation, convergence des martingales, lemma de Fatou multivoque. *Ann. Mat. Pura Appl.* (IV) **164** (1993), 51-75.

[CV] — and Valadier, M. *Convex Analysis and Measurable Multifunctions*, Lecture Notes in Math. **580**, Springer-Verlag, Berlin, 1977.

[CC] Cellina, A. and Colombo, G. On a classical problem of the calculus of variations without convexity assumptions. *Ann. Inst. H. Poincaré, Analyse nonlinéaire* **7** (1989), 97-106.

[Ce] Cesari, L. *Optimization – Theory and Applications.* Springer-Verlag, Berlin, 1983.

[Cha1] Chatterji, S.D. Les martingales et leurs applications analytiques. In: *Ecole d'Eté de Probabilités: Processus Stochastiques* (J.L. Bretagnolle et al., eds.). Lecture Notes in Math. **307**, Springer-Verlag, Berlin, 1973, pp. 27-164.

[Cha2] — A subsequence principle theory. *Jber. d. Dt. Math.-Verein.* **87** (1985), 91-107.

[Cho] Choquet, G. *Lectures on Analysis.* Benjamin, Reading, Mass., 1969.

[DM] Dellacherie, C. and Meyer, P.-A. *Probabilités et Potentiel.* Hermann, Paris, 1975 (English translation: North-Holland, Amsterdam, 1978).

[DU] Diestel, J. and Uhl, J. J. *Vector Measures.* Math. Surveys **15**, American Mathematical Society, Providence, RI, 1977.

[E] Ekeland, I. Sur le contrôle optimal de systèmes gouvernés par des equations elliptiques. *J. Functional Anal.* **9** (1972), 1-62.

[ET] — and Temam, R. *Convex Analysis and Variational Problems.* North-Holland, Amsterdam, 1977.

[Fe] Ferguson, T. *Mathematical Statistics : a Decision Theoretic Approach.* Academic Press, New York, 1967.

[Fl] Flores-Bazán, F. Existence theorems for a class of nonconvex problems in the calculus of variations. *J. Optim. Theory Appl.* **78** (1993), 31-48.

[Gam] Gamkrelidze, R.V. On sliding optimal states. *Soviet Math. Dokl.* **3** (1962), 559-561.

[Gap] Gaposhkin, V.F. Convergence and limit theorems for sequences of random variables. *Theory Probab. Appl.* **17** No. 3 (1972), 379-400.

[Gh] Ghouila-Houri, A. Sur la généralisation de la notion de commande d'un système guidable. *Rev. Francaise Automat. Informat. Recherche Opérationelle* **4** (1967), 7-32.

[Gi] Girardi, M. Weak vs. norm compactness in L_1: the Bocce criterion. *Studia Math.* **98** (1991), 95-97.

[Ha] Harsanyi, J.C. Games with incomplete information played by 'Bayesian' players. *Management Sci.* **14**, 159-182.

[He] Hess, C. Measurability and integrability of the weak upper limit of a sequence of multifunctions. *J. Math. Anal. Appl.* **153** (1990), 226-249.

[Iof] Ioffe, A.D. On lower semicontinuity of integral functionals, I, II. *SIAM J. Control Optim.* **15** (1977), 521-538, 991-1000.

[Ion] Ionescu Tulcea, A. and C. *Topics in the Theory of Lifting.* Springer-Verlag, Berlin, 1969.

[JM] Jacod, J. and Mémin, J. Sur un type de convergence intermédiaire entre la convergence en loi et la convergence en probabilité. *Séminaire de Probabilités* XV (J. Azéma and M. Yor, eds.) Lecture Notes in Math. **986**, Springer-Verlag, Berlin, 1983, pp. 529-546.

[Jal] Jalby, V. *Contribution aux Problèmes de Convergence des Fonctions Vectorielles et des Intégrales Fonctionelles.* Thèse, Universit'e Montpellier II, Montpellier, 1993.

[Jaw] Jawhar, A. Mesures de transition et applications. *Sém. Analyse Convexe* **14** (1984), 13.1-13.62.

[K] Komlós, J. A generalisation of a problem of Steinhaus. *Acta Math. Acad. Sci. Hungar.* **18** (1967), 217-229.

[L1] — An extension of Wald's theory of statistical decision functions. *Ann. Math. Stat.* **26** (1955), 69-81.

[L2] LeCam, L. Convergence in distribution of stochastic processes. *Univ. California Publ. Statist.* **2**, No. 11 (1957), 207-236.

[Ma] Mas-Colell, A. On a theorem of Schmeidler. *J. Math. Econom.* **13** (1984), 201-206.

[Mc1] McShane, E.J. Generalized curves. *Duke Math. J.* **6** (1940), 513-536.

[Mc2] McShane, E.J. Relaxed controls and variational problems. *SIAM J. Control* **5** (1967), 438-485.

[NM] von Neumann, J. and Morgenstern, O. *Theory of Games and Economic Behavior*, Princeton Univ. Press, Princeton, 1944.

[N] Neveu, J. *Mathematical Foundations of the Calculus of Probability*, Holden-Day, San Fransisco, 1965.

[Pa] Page, F.H. Incentive compatible strategies for general Stackelberg games with incomplete information. *Intern. J. Game Theory* **18** (1989), 409-421.

[Pf] Pfanzagl, J. Convexity and conditional expectations. *Ann. Prob.* **2** (1974), 490-494.

[PV] Piccinini, L. and Valadier, M. Uniform integrability and Young measures. *J. Math. Anal. Appl.*, to appear.

[Pr] Prohorov, Yu. V. Probability distributions in functional spaces. *Uspehi Matem. Nauk* **55** (1953), 167.

[Ra] Raymond, J.-P. Existence theorems in optimal control theory without convexity assumptions. *J. Optim. Theory Appl.* **67** (1990), 109-132.

[Ro] Rockafellar, R.T. Integrals which are convex functions II. *Pacific J. Math.* **39** (1971), 439-469.

[SV] aadoune, M. and Valadier, M. Convergence in measure: the Fréchet criterion from local to global. Preprint, Département de Mathématiques, Université Montpellier II.

[Schm] Schmeidler, D. Fatou's lemma in several dimensions. *Proc. Amer. Math. Soc.* **24** (1970), 300-306.

[Schw] Schwartz, L. *Radon Measures.* Oxford University Press, 1975.

[Sl] Slaby, M. Strong convergence of vector valued pramarts and subpramarts. *Prob. Math. Stat.* **5**, 187-196.

[Val1] Valadier, M. Some bang-bang theorems. in: *Multifunctions and Integrands* (G. Salinetti, ed.) Lecture Notes in Math. **1091**, Springer-Verlag, Berlin, 1984, pp. 225-234.

[Val2] — Régularisation sci, relaxation et théorèmes bang-bang. *C.R. Acad. Sci. Paris Sér. I* **293** (1981), 115-116.

[Val3] — Young measures. *Methods of Nonconvex Analysis* (A. Cellina, ed.), Lecture Notes in Math. **1446**, Springer-Verlag, Berlin, 1990, pp. 152-188.

[Val4] — Young measures, weak and strong convergence and the Visintin–Balder theorem. *Set-Valued Anal.* **2** (1994), 357-367.

[Val5] — A course on Young measures. *Rend. Istit. Mat. Univ. Trieste*, to appear.

[Var] Varadarajan, V.S. Weak convergence of measures on separable metric spaces. *Sankhya* **19** (1958), 15-22.

[Vi] Visintin, A. Strong convergence results related to strict convexity. *Comm. Partial Differential Equations* **9** (1984), 439-466.

[Wal] Wald, A. *Statistical Decision Functions.* Wiley, New York, 1950.

[War] Warga, J. *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

[Y1] Young, L.C. Generalized curves and the existence of an attained absolute minimum in the calculus of variations, *C. R. Sci. Lettres Varsovie*, C III **30** (1937), 212-234.

[Y2] — *Lectures on the Calculus of Variations and Optimal Control Theory*, Saunders, Philadelphia, PA, 1969.