NOT FOR QUOTATION
WITHOUT PERMISSION
OF THE AUTHOR

ON THE CONVERGENCE IN DISTRIBUTION OF MEASURABLE MULTIFUNCTIONS, NORMAL INTEGRANDS, STOCHASTIC PROCESSES AND STOCHASTIC INFIMA

Gabriella Salinetti
Roger J.-B. Wets

December 1982
CP-82-87

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

ABSTRACT

The concept of the distribution function of a closed-valued measurable multifunction is introduced and used to study the convergence in distribution of sequences of multifunctions and the epi-convergence in distribution of normal integrands; in particular various compactness criteria are exhibited. The connections with the convergence theory for stochastic processes is analayzed and for purposes of illustration we apply the theory to sketch out a modified derivation of Donsker's Theorem (Brownian motion as a limit of random walks). We also suggest the potential application of the theory to the study of the convergence of stochastic infima.

ON THE CONVERGENCE IN DISTRIBUTION OF MEASURABLE MULTIFUNCTIONS, NORMAL INTEGRANDS, STOCHASTIC PROCESSES AND STOCHASTIC INFIMA

Gabriella Salinetti, Universita di Roma[1)]
Roger J.-B. Wets[2)]

In [1] we have given various characterizations for the almost sure convergence and the convergence in probability of sequences of closed-valued measurable multifunctions, sometimes also called *random closed sets*. In this paper we study their convergence in distribution or equivalently the weak*-convergence of the induced probability measures. Actually, we derive the basic results by relying on the framework provided by the theory of weak*-convergence in metric spaces. As background to the study of the convergence of normal integrands, we exhibit the relationship between the convergence theory for measurable multifunctions and that for certain stochastic processes associated to measurable multifunctions, such as the processes determined by the distance and characteristic functions. After some general results about normal integrands we relate their epi-convergence in distribution to their convergence in the classical sense of stochastic processes. Next, we obtain compactness criteria. Finally, we derive convergence in distribution results for selections of measurable multifunctions and touch on the

1) Supported in part by CNR-GNAFA.
2) Supported in part by a Guggenheim Fellowship.

potential applications of these results to stochastic optimization and the convergence of stochastic processes.

There is an intimate relationship between normal integrands and stochastic process which can be exploited to devise a new approach to the convergence of stochastic processes. To illustrate this point, we give a modified derivation of Donsker's theorem. A related approach has been developed by W. Vervaat. His motivation comes from the study of extremal processes in statistics, whereas our work was originally motivated by the search for approximation schemes for stochastic optimization problems. However, we feel that there also many potential contributions that this approach could make to the study of "classical" stochastic processes.

The setting is the same as in [1]. Let (Ω, A, μ) be a probability space with A the class of measurable sets and μ a probability measure on A; (E,d) is an n-dimensional linear space equipped with a metric d. A map Γ with domain Ω and whose values are closed subsets of E, $\Gamma:\Omega \rightrightarrows E$, is said to be a *closed-valued measurable multifunction* if for all closed sets $F \subset E$,

$$\Gamma^{-1}(F) = \{\omega \in \Omega \mid \Gamma(\omega) \cap F = \emptyset\} \in A \quad .$$

Each such multifunction can be identified with a *measurable* function γ from Ω to F, the hyperspace of all closed subsets of E equipped with the topology T_F, a variant of the Vietoris topology, cf. [1, Proposition 1.1]. More precisely, T_F--or simply T if no confusion is possible--is generated by the subbase consisting of all families of sets of the type

$$\{F_K, K \in K\} \qquad \text{and} \qquad \{F^G, G \in G\}$$

where for any set $D \in E$,

$$F^D = \{F \in F \mid F \cap D = \emptyset\} \quad \text{and} \quad F_D = \{F \in F \mid F \cap D \neq \emptyset\} \quad ,$$

and G and K denote respectively the classes of open and compact subsets of E. The choice of this topology is motivated by the fact that sequences in F T-converge if and only if the corresponding sequence of subsets of E converge in the classical sense, see [1, Theorem 2.2]. The properties of E allow us also to generate this topology T from the subbase consisting of all sets of the type

$$\{F^{B_\varepsilon(x)}, \ \varepsilon > 0, x \in E\} \quad \text{and} \quad \{F_{B^\circ_\varepsilon(x)}, \ \varepsilon > 0, x \in E\}$$

where $B_\varepsilon(x)$ and $B^\circ_\varepsilon(x)$ are respectively the closed and open balls of radius ε and center x. In fact, it is sufficient to consider open and closed balls with $\varepsilon \in Q_+$ and $x \in D$, a countable dense subset of E, i.e., these sets (balls) determine a countable base for (the topology of) E. In turn, this yields a countable base for T, let

$$\mathcal{B} = \{B = B_\varepsilon(x) \mid \varepsilon \in Q_+ \ , \ x \in D \subset E\} \quad ,$$

then as can easily be verified

$$\{F^{B'_1,\dots,B'_r}_{B^\circ_1,\dots,B^\circ_s} \mid B'_j \in \mathcal{B} \ , \ j = 1,\dots,r \ ; \ B_i \in \mathcal{B} \ , \ i = 1,\dots,s\}$$

is a base for T, where

$$F^{B'_1,\dots,B'_r}_{B^\circ_1,\dots,B^\circ_s} = F^{B'_1 \cup \dots \cup B'_r} \cap F_{B^\circ_1} \cap \dots \cap F_{B^\circ_s} \quad .$$

It follows directly from the properties of E and the definition of T that the Borel sigma-field S_F--or simply S, when no confusion is possible--is generated by any one of the families $\{F^K, K \in K\}$, $\{F_G, G \in G\}$, $\{F^{B_\varepsilon(x)}, \varepsilon > 0, x \in E\}$, $\{F_{B^\circ_\varepsilon(x)}, \varepsilon > 0, x \in E\}$ and $\{F_B, B \in \mathcal{B}\}$ and their complements.

It is significant that the topological space (F, T) is compact, regular and has a countable base; hence is *metrizable*. The compactness follows from Alexander's theorem with the subbase $\{F^G, G \in G; F_K, K \in K\}$ for the closed subsets of F; note that $\emptyset \in F$. The local compactness of E yields regularity and a countable base has already been exhibited earlier; see [2], [1] for details and further developments.

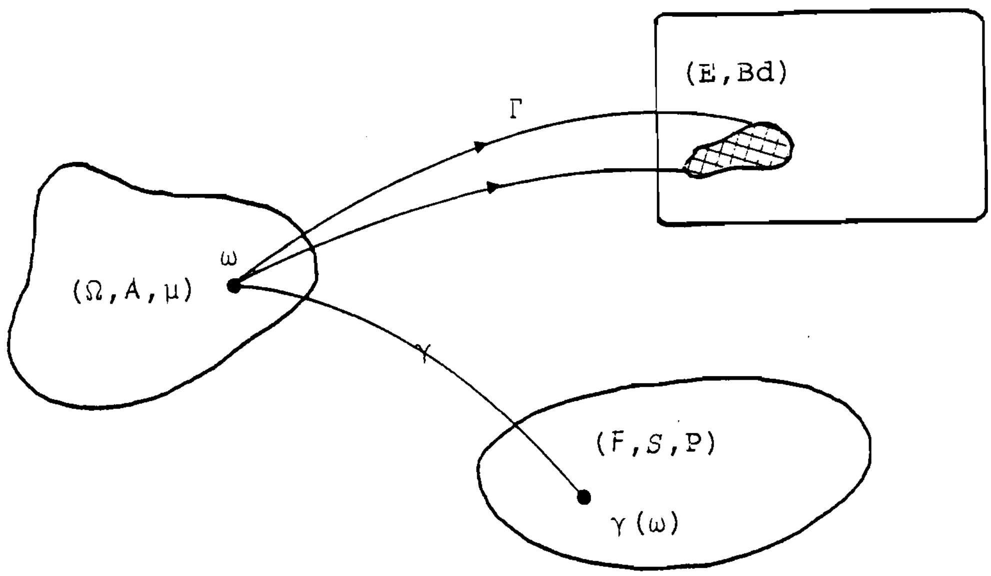


Figure 0.1. Multifunction Γ and associated function γ.

We can thus view γ as a measurable function (a random object) from Ω to the compact metrizable space F. The general theory of convergence in distribution on metric spaces is applicable and provides the underpinnings for the results on the convergence of closed-valued measurable multifunctions, as well as for the epi-convergence of normal integrands, cf. Section 3.

1. CONVERGENCE IN DISTRIBUTION

Since (F,T) is metrizable every probability measure P defined on S is *regular* [3, Theorem 1.1] and thus is determined by its value on the open (or closed) subsets of F. Since here every open set is the countable union of elements in the base, which in turn are obtained from the subbase by finite intersections, it will be sufficient to know the values of P on the subbase $\{F^K, K \in K; F_G, G \in G\}$ to completely determine P. We shall show that even a much smaller subclass of subsets of S will in fact suffice.

A function T from K into [0,1] with $T(\phi) = 0$ is called a *distribution function* if for any sequences of compact sets $\{K_\nu, \nu \in N\}$

(1.1) the sequence $\{T(K_\nu), \nu \in N\}$ decreases to $T(K)$ whenever the K_ν decrease to K;

(1.2) the functions $\{S_\nu, \nu = 0,1,\ldots\}$ defined recursively by

$$S_0(K_0) = 1 - T(K_0)$$

$$S_1(K_0;K_1) = S_0(K_0) - S_0(K_0 \cup K_1)$$

and for $\nu = 2,\ldots$

$$S_\nu(K_0;K_1,\ldots,K_\nu) = S_{\nu-1}(K_0;K_1,\ldots,K_{\nu-1}) - S_{\nu-1}(K_0 \cup K_\nu;K_1,\ldots,K_{\nu-1})$$

take on their values in [0,1].

A function on K with properties (1.1) and (1.2) is sometimes called a *Choquet capacity*. These properties of T on K are essentially the same as those of a distribution function defined on the real line. Property (1.1) can be viewed as an extension of the notion of right-continuity and property (1.2) as an extension of monotonicity.

1.3 THEOREM (Choquet). *Every probability measure* P *on* S *determines a distribution function* T *on* K *through the correspondence*

$$(1.4) \qquad T(K) = P(F_K) \quad .$$

$$(1.5) \qquad S_\nu(K_0;K_1,\ldots,K_\nu) = P(F^{K_0} \cap F_{K_1} \cap \ldots \cap F_{K_\nu}) \quad .$$

Conversely, every distribution function T *on* K *determines a probability measure* P *on* S *through* (1.4), *or alternatively* (1.5).

Matheron [4, p. 30-35] gives a proof of this theorem by relying on Choquet's Capacity Theorem. A new derivation relying exclusively on standard probabilistic tools is given in Appendix A.

Every closed-valued measurable multifunction $\Gamma:\Omega \rightrightarrows E$, or equivalently every measurable function (random element) $\gamma:\Omega \to F$, induces on S a probability measure, denoted by P, with

$$P(\mathcal{D}) = \mu\{\omega|\gamma(\omega) \in \mathcal{D}\}$$

for every $\mathcal{D} \in S$. For sets of the type F_D that belong to S, we have that

$$P(F_D) = \mu[\gamma^{-1}(F_D)] = \mu[\Gamma^{-1}(D)] \quad .$$

In particular, the values of T, the distribution function of Γ (i.e. the distribution function associated with P), are given by

$$T(K) = \mu[\Gamma^{-1}(K)] \quad .$$

The convergence in distribution of a sequence $\{\Gamma_\nu, \nu \in N\}$ of closed-valued measurable multifunctions can thus be studied in the framework of the weak*-convergence of the induced probability measures $\{P_\nu, \nu \in N\}$ defined on S - the Borel field of the meritizable space (F,T) - or equivalently, as we shall see, the convergence of the associated distribution functions $\{T_\nu, \nu \in N\}$ at every "continuity" point of the limit distribution function [3], [5]. This program is carried out in the remainder of this section.

A distribution function T defined on $\mathcal{K}$, is said to be *distribution-continuous* at K if for every sequence of sets $\{K_\nu \in \mathcal{K}, \nu \in N\}$ ($\mathcal{T}$-) increasing to K we have that the sequence $\{T(K_\nu), \nu \in N\}$ increases to T(K). Recall that $T(K) = \lim T(K_\nu)$, if $\{K_\nu, \nu \in N\}$ is any sequence decreasing to K. Note that

$$K_\nu \downarrow_{\mathcal{T}} K \quad \text{implies that} \quad K = \bigcap_{\nu \in N} K_\nu$$

and

$$K_\nu \uparrow_{\mathcal{T}} K \quad \text{implies that} \quad K = \mathrm{cl} \bigcup_{\nu \in N} K_\nu \qquad ,$$

see for instance [6, Prop. 1 and 2]. The *distribution-continuity set* C_T is the subset of $\mathcal{K}$ on which T is distribution-continuous. We shall show that the multifunctions $\{\Gamma_\nu \; \nu \in N\}$ converge in distribution to a limit multifunction Γ if and only if the associated distribution functions $\{T_\nu, \nu \in N\}$ converge to the distribution function T on C_T; in fact we shall go one step further and show that it is sufficient to demand convergence on $C_T^{ub} = C_T \cap \mathcal{K}^{ub}$, where $\mathcal{K}^{ub}$ *consists of all the compact subsets of* E *that can be obtained as the finite union of closed balls* (including ∅, the empty union of closed balls).

Recall that the random elements $\{\gamma_\nu, \nu \in N\}$ (respectively the multifunctions $\{\Gamma_\nu, \nu \in N\}$) converge in distribution to a random element γ (respectively, the multifunction Γ) if the induced probability measures $\{P_\nu, \nu \in N\}$ weak*-converge to P, in particular if

$$\text{(1.6)} \qquad \lim_\nu P_\nu(\mathcal{D}) = P(\mathcal{D})$$

for all P-continuity sets $\mathcal{D} \in \mathcal{S}$ [3, Theorem 2.1] where $\mathcal{D} \in \mathcal{S}$ is a P-continuity set if

$$P(\mathrm{bdy}\ \mathcal{D}) = 0$$

and bdy $\mathcal{D}$ is the boundary of $\mathcal{D}$ with respect to the topology $\mathcal{T}$.

Let V be a subclass of S endowed with the following properties:

(1.7) V is closed under finite intersections,

(1.8) given any $F \in \mathcal{F}$ and any ($\mathcal{T}$-)neighborhood N of F, there exists $W \in V$ such that

$$F \in \operatorname{int} W \subset W \subset N \quad .$$

Then V is a so-called *convergence-determining class*, i.e., convergence in distribution of the Γ_ν to Γ follows from having (1.6) satisfied on V rather than on S. The proof that V is indeed such a class of sets is the same as that of Theorem 2.2 and its Corollary 1 in [3], it suffices to recall that $(\mathcal{F},\mathcal{T})$ is separable. The next theorem shows that K^{ub} generates a convergence-determining class; this allows us to relate the convergence in distribution of multifunctions to the convergence of their distribution functions. A related result, for extremal processes on R, appears in [26, Theorem 8.3].

1.9 THEOREM. *The class of sets* $\{F_K, K \in K^{ub}\}$ *is a convergence-determining class, i.e., the* Γ_ν *converge in distribution to* Γ*, if for every* P*-continuity set* F_K*, with* $K \in K^{ub}$*, we have that*

$$\lim_\nu P_\nu(F_K) = P(F_K) \quad . \tag{1.10}$$

Moreover the Γ_ν *converge in distribution to* Γ *if and only if for every* $K \in C_T^{ub} \subset K^{ub}$ *we have that*

$$\lim_\nu T_\nu(K) = T(K) \quad . \tag{1.11}$$

The two following lemmas are needed in the proof of this theorem. They clarify the relationship between P-continuity and distribution-continuity. We define the ε-*enlargement* of K by $\varepsilon^\circ K = \{x | d(x,K) < \varepsilon\}$ and set $\varepsilon K = \operatorname{cl} \varepsilon^\circ K$, where $d(x,K) = \min_{y\in K} d(x,y)$.

1.12 LEMMA. *Suppose that* $K \in \mathcal{K}$ *is nonempty and* $\bar{\varepsilon} > 0$. *Then the family* $\{F_{\varepsilon K}, 0 < \varepsilon < \bar{\varepsilon}\}$ *contains at most countably many sets which are not* P-*continuous*.

PROOF. Formally, the argument is similar to the one used to show that the set of discontinuous points of a monotone function is at most countable.

Suppose that $F_{\varepsilon_1 K}$ and $F_{\varepsilon_2 K}$ with $0 < \varepsilon_1 < \varepsilon_2 < \bar{\varepsilon}$ are not P-continuity sets, i.e., for $i = 1,2$, $\alpha_i' = P(\text{int } F_{\varepsilon_i K}) < P(F_{\varepsilon_i K}) = \alpha_i$. The two open intervals

$$I(\varepsilon_i) =]\alpha_i', \alpha_i[\subset [0,1] \qquad i = 1,2$$

are disjoint, since for any $\hat{\varepsilon} \in]\varepsilon_1, \varepsilon_2[$ we have that

$$\text{int } F_{\varepsilon_1 K} \subset F_{\varepsilon_1 K} \subset F_{\hat{\varepsilon} K} \subset \text{int } F_{\varepsilon_2 K} \subset F_{\varepsilon_2 K}$$

and thus $\alpha_1' < \alpha_1 \leq \hat{\alpha} = P(F_{\hat{\varepsilon} K}) \leq \alpha_2' < \alpha_2$. It follows that the class of sets $\{F_{\varepsilon K}, 0 < \varepsilon < \bar{\varepsilon}\}$ that are not P-continuity sets correspond to a class of disjoint subintervals of $[0,1]$, each one containing a distinct rational number. There can only be a countable number of such intervals and hence of sets $F_{\varepsilon K}$ that are not P-continuity sets. □

1.13 LEMMA. *Suppose that* T *is a distribution function on* $\mathcal{K}$. *Then*

$$C_T^{ub} = \{K \in \mathcal{K}^{ub} \mid F_K \text{ P-}\textit{continuity set}\} \quad .$$

PROOF. Suppose first that $K \in \mathcal{K}^{ub}$ and F_K is a P-continuity set. Let us consider $\{K_\nu, \nu \in \mathbb{N}\}$ an arbitrary collection of compact sets increasing to K. We need to show that

$$\lim_\nu T(K_\nu) = T(K) \quad .$$

Since $K_\nu \uparrow_{\mathcal{T}} K = \text{cl } \cup_{\nu=1}^\infty K_\nu$, we have that

$$F_K = \text{cl } F_{\cup K_\nu} \quad \text{and} \quad \text{int } F_K \subset F_{\cup K_\nu} \quad .$$

These relations and the P-continuity of F_K imply that

$$T(K) = P(F_K) = P(\text{int } F_K) \leq P(F_{\cup K_\nu}) = \lim_\nu P(F_{K_\nu})$$

$$= \lim_\nu T(K_\nu) \leq P(F_K) = T(K)$$

and thus $K \in C^{ub}$.

Next suppose that $K \in C^{ub}$. We suppose that K is nonempty, since otherwise the assertion holds trivially. Because K is a compact set with nonempty interior, we can find an (strictly) increasing sequence of open precompact sets $\{G_\nu, \nu \in N\}$ such that $\text{cl } G_\nu = K_\nu \uparrow_T K$, and $\text{cl } G_\nu \subset G_{\nu+1}$. We have that

$$F_{K_\nu} \subset F_{G_{\nu+1}} \subset F_{K_{\nu+1}} \subset F_K \quad .$$

The sets $F_{G_{\nu+1}}$ are T-open and hence

$$F_{K_\nu} \subset F_{G_{\nu+1}} \subset \text{int } F_K \quad .$$

Since T is distribution-continuous at K, we have that

$$P(F_K) = T(K) = \lim_\nu T(K_\nu) = \lim P(F_{K_\nu}) \leq P(\text{int } F_K) \leq P(F_K)$$

which implies that $P(\text{bdy } F_K) = 0$, i.e., F_K is a P-continuity set. □

PROOF OF THEOREM 1.9. Let

$$V = \{F_K | K \in K^{ub}, P(\text{bdy } F_K) = 0\}$$

and

$$U = \{F^{K_0}_{K_1,\ldots,K_r} = F^{K_0} \cap F_{K_1} \cap \ldots \cap F_{K_r} | F_{K_i} \in V,\ i = 0,\ldots,r,\ r \geq 0$$

We prove first that (1.10) implies that for every $\mathcal{D} \in U$

(1.14) $\qquad \lim P_\nu(\mathcal{D}) = P(\mathcal{D})$

and that U satisfies (1.7) and (1.8). This will yield the first assertion of the theorem.

Since K^{ub} is closed under finite unions and for any collection of subsets $D_1,\ldots,D_r$ of E

$$\text{bdy } F_{\cup D_i} \subset \cup \text{ bdy } F_{D_i} \quad ,$$

it follows that

$$(1.15) \qquad F_{\cup_{i=1}^r K_i} \in V \text{ whenever } F_{K_i} \in V,\ i = 1,\ldots,r.$$

And thus from (1.10), with $F_{K_i} \in V$, $i = 1,\ldots,r$, we get that

$$\lim P_\nu(F^{K_0}) = \lim [1 - P_\nu(F_{K_0})] = 1 - P(F_{K_0}) = P(F^{K_0}) \quad ,$$

$$\lim_\nu P_\nu(F_{K_1}^{K_0}) = \lim [P_\nu(F_{K_0 \cup K_1}) - P_\nu(F_{K_0})] = P(F_{K_0 \cup K_1}) - P(F_{K_0}) = P(F_{K_1}^{K_0})$$

from which (1.14) follows by induction.

That U is closed under finite intersections, i.e., satisfies (1.7), follows from (1.15) and the fact that

$$F^{K_0}_{K_1,\ldots,K_r} \cap F^{L_0}_{L_1,\ldots,L_s} = F^{K_0 \cup L_0}_{K_1,\ldots,K_r,L_1,\ldots,L_s} \quad ,$$

where for $i = 0,\ldots,r$, $K_i \in K^{ub}$ and for $j = 0,\ldots,s$, $L_j \in K^{ub}$.

Next, we show that U also satisfies (1.8). Take $F \in F$ and $N \in B$, a base neighborhood of F, say

$$(1.16) \qquad N = F^{B_1',\ldots,B_r'}_{B_1^\circ,\ldots,B_s^\circ} = F^{C}_{B_1^\circ,\ldots,B_s^\circ}$$

with $C = \cup_{i=1}^r B_i'$, and for $i = 1,\ldots,r$, $B_i' \in B$ and for $j = 1,\ldots,s$, $B_j \in B$. We assume that $F \neq \emptyset$, $C \neq \emptyset$ and $s \geq 1$, the other cases being straightforward. $F \in F_{B^\circ}$ with $B^\circ = B_\eta^\circ(x)$ means not only that $F \cap B_\eta^\circ(x) \neq \emptyset$ but also that

$$F \cap B_\varepsilon^\circ(x) = \emptyset$$

for all $\varepsilon \in]\hat{\eta},\eta[$ for some $\hat{\eta} < \eta$. From Lemma 1.12 it follows that then there exists $\hat{\varepsilon} = \hat{\eta} + \bar{\varepsilon}$ with $0 < \bar{\varepsilon} < \eta - \hat{\eta}$, such that $F_{\hat{B}}$, with $\hat{B} = B_{\hat{\varepsilon}}(x)$, is a P-continuity set. Hence to every B_i that appears in the definition (1.16) of N, there corresponds $\hat{B}_i \in \mathcal{B}$ such that $F_{\hat{B}_i}$ is a P-continuity set and $F \in F_{\hat{B}_i^\circ} \subset F_{\hat{B}_i} \subset F_{B_i^\circ}$. Thus

$$F \in \operatorname{int} \mathcal{W}' \subset \mathcal{W}' \subset F_{B_1^\circ} \cap \ldots \cap F_{B_s^\circ}$$

where $\mathcal{W}' = F_{\hat{B}_1} \cap \ldots \cap F_{\hat{B}_s}$ is a P-continuity set, and clearly

$$\operatorname{int} \mathcal{W}' \supset F_{\hat{B}_1^\circ} \cap \ldots \cap F_{\hat{B}_s^\circ}$$

as follows directly from the construction of $\mathcal{T}$.

On the other hand, since $C \in K^{ub}$ is compact and $F \in F^C$ is closed, there exists $\bar{\varepsilon} > 0$ such that for every $\varepsilon \in]0,\bar{\varepsilon}[$

$$\varepsilon C \cap F = (C + B_\varepsilon(0)) \cap F \neq \emptyset \quad .$$

Again by Lemma 1.12 there then exists $\hat{\varepsilon} \in]0,\bar{\varepsilon}[$ such that $F_{\hat{C}}$, with $\hat{\varepsilon}C = \hat{C}$, and necessarily its complement $F^{\hat{C}}$ are P-continuity sets. The set

$$\mathcal{W} = \mathcal{W}' \cap F^{\hat{C}}$$

is a continuity set and we have that

$$F \in \operatorname{int} \mathcal{W} \subset \mathcal{W} \subset F^C_{B_1^\circ,\ldots,B_x^\circ} = N$$

since $F \in F^{\hat{C}}_{\hat{B}_1^\circ,\ldots,\hat{B}_s^\circ} \subset \operatorname{int} \mathcal{W}$ by definition of $\mathcal{T}$. This completes the proof of the first assertion of the theorem.

By Lemma 1.13 for every $K \in C_T^{ub}$, F_K is a P-continuity set and hence the convergence in distribution of the Γ_ν to Γ implies via (1.6) that

$$\lim_\nu T_\nu(K) = \lim_\nu P_\nu(F_{K_\nu}) = P(F_K) = T(K)$$

which naturally yields (1.11).

On the other hand, if $\lim T_\nu(K) = T(K)$ on C_T^{ub} it follows from Lemma 1.13 and the relation (1.4) that

$$\lim_\nu P_\nu(F_K) = P(F_K)$$

for every $K \in K^{ub}$. But we just proved that this is a convergence-determining class. Hence the Γ_ν converge in distribution to Γ. □

In [1, Section 5] it was demonstrated that the almost sure convergence of a sequence of measurable multifunction $\{\Gamma_\nu, \nu \in N\}$ to the multifunction Γ implies their convergence in probability. We show that in turn, this implies their convergence in distribution. Recall that the Γ_ν *converge* to Γ *in probability* of for all $\varepsilon > 0$ and any compact $K \in K$,

$$\text{(1.17)} \qquad \lim_{\nu\to\infty} \mu[\Delta_{\varepsilon,\nu}^{-1}(K)] = 0$$

where

$$\Delta_{\varepsilon,\nu} = (\Gamma_\nu \setminus \varepsilon^\circ \Gamma) \cup (\Gamma \setminus \varepsilon^\circ \Gamma_\nu) \quad .$$

Note that $\Delta_{\varepsilon,\nu} : \Omega \rightrightarrows E$ is a measurable multifunction.

1.18 PROPOSITION. *Suppose* $\{\Gamma_\nu, \nu \in N\}$ *is a collection of closed-valued measurable multifunction converging in probability to the closed-valued measurable multifunction* Γ. *Then the* $\{\Gamma_\nu, \nu \in N\}$ *also converge in distribution to* Γ.

PROOF. We start with proving that for all $K \in K^{ub}$

$$\limsup_{\nu\to\infty} T_\nu(K) \leq T(K)$$

Suppose Λ_1 and Λ_2 are two closed-valued measurable multifunctions defined on Ω. Then for any $\varepsilon > 0$ and $K \in K$ we get

$$\Lambda_1^{-1}(\varepsilon K) = (\varepsilon\Lambda_1)^{-1}(K) \quad .$$

One also has that

$$(1.19)\qquad \Lambda_1^{-1}(K) \subset \Lambda_1^{-1}(\varepsilon K) \cup (\Lambda_1 \setminus \varepsilon^\circ \Lambda_2)^{-1}(K) \qquad ,$$

as follows from the following relations

$$\begin{aligned}\Lambda_1^{-1}(K) &= [\Lambda_1 \cap (\varepsilon^\circ \Lambda_2 \cup (E \setminus \varepsilon^\circ \Lambda_2))]^{-1}(K) \\ &\subset (\Lambda_1 \cap \varepsilon \Lambda_2)^{-1}(K) \cup (\Lambda_1 \setminus \varepsilon^\circ \Lambda_2)^{-1}(K) \\ &\subset (\varepsilon \Lambda_2)^{-1}(K) \cup (\Lambda_1 \setminus \varepsilon^\circ \Lambda_2)^{-1}(K) \\ &= \Lambda_2^{-1}(\varepsilon K) \cup (\Lambda_1 \setminus \varepsilon^\circ \Lambda_2)^{-1}(K) \qquad ,\end{aligned}$$

where the last equality follows from the preceeding identity. Thus

$$(1.20)\qquad \mu[\Lambda_1^{-1}(K)] \le \mu(\Lambda_2^{-1}(\varepsilon K)] + \mu[(\Lambda_1 \setminus \varepsilon^\circ \Lambda_2)^{-1}(K)]$$

since by definition of measurability for multifunctions all sets involved are measurable. In particular if we set $\Lambda_1 = \Gamma_\nu$ and $\Lambda_2 = \Gamma$ and take $K \in \mathcal{K}^{ub}$ this gives

$$\begin{aligned}T_\nu(K) = \mu[\Gamma_\nu^{-1}(K)] &\le \mu[\Gamma^{-1}(\varepsilon K)] + \mu[(\Gamma \setminus \varepsilon^\circ \Gamma_\nu)^{-1}(K)] \\ & T(\varepsilon K) + \mu(\Lambda_{\varepsilon,\nu}^{-1}(K)) \qquad .\end{aligned}$$

Since this holds for all ν, taking lim sup on both sides yields:

$$\limsup_{\nu\to\infty} T_\nu(K) \le T(\varepsilon K)$$

as follows from convergence in probability, in particular (1.11). The above holds for all $\varepsilon > 0$ and since T is a distribution function we have that $\lim_{\varepsilon\downarrow 0} T(\varepsilon K) = T(K)$ from which the relation $\limsup_{\nu\to\infty} T_\nu(K) \le T(K)$ follows directly.

There remains only to show that for $K \in C_T^{ub}$

$$\liminf_{\nu\to\infty} T_\nu(K) \ge T(K) \qquad .$$

Since $K \in K^{ub}$ we can write it as

$$K = B_{r_1}(x_1) \cup \ldots \cup B_{r_q}(x_q)$$

with q finite. Let

$$K_\varepsilon = B_{r_1-\varepsilon}(x_1) \cup \ldots \cup B_{r_q-\varepsilon}(x_q)$$

with $0 < \varepsilon < \min\{r_1, r_2, \ldots, r_q\}$. As $\varepsilon \downarrow 0$ we have that $\varepsilon K_\varepsilon \uparrow K$. If $T \in C_T^{ub}$ we have that

$$T(K) = \lim_{\varepsilon \downarrow 0} T(K_\varepsilon) = \lim_{\varepsilon \downarrow 0} T(\varepsilon K_\varepsilon) \quad .$$

Thus given any $\delta > 0$ there exists $\varepsilon\delta$ such that for all $0 < \varepsilon < \varepsilon\delta$

$$T(K) \leq T(K_\varepsilon) + \delta \quad .$$

We now again use (1.19) with $\Lambda_1 = \Gamma$ and $\Lambda_2 = \Gamma_\nu$ to get

$$\begin{aligned} T(K) &\leq T(K_\varepsilon) + \delta = \mu[\Gamma^{-1}(K_\varepsilon)] + \delta \\ &\leq \mu[\Gamma_\nu^{-1}(\varepsilon K_\varepsilon)] + [(\Gamma \backslash \varepsilon^\circ \Gamma_\nu)^{-1}(K_\varepsilon)] + \delta \\ &\leq T_\nu(\varepsilon K_\varepsilon) + \mu[\Delta_{\varepsilon,\nu}^{-1}(K_\varepsilon)] + \delta \\ &\leq T_\nu(K) + \mu[\Delta_{\varepsilon,\nu}^{-1}(K_\varepsilon)] + \delta \end{aligned}$$

Taking lim inf on both sides, using (1.17) and the fact that δ is arbitrary yields the desired relation. $\square$

2. CLOSED-VALUED MEASURABLE MULTIFUNCTIONS AND ASSOCIATED STOCHASTIC PROCESSES

A closed-valued multifunction Γ is completely described by any one of the following associated processes:

(2.1) the indicator (function) process 1_Γ defined on $E \times \Omega$ with

$$1_\Gamma(x,\omega) = \begin{cases} 1 & \text{if } x \in \Gamma(\omega) \quad , \\ 0 & \text{otherwise} \quad ; \end{cases}$$

(2.2) the extended indicator (function) process ψ_Γ defined on $E \times \Omega$ with

$$\psi_\Gamma(x,\omega) = \begin{cases} 0 & \text{if } x \in \Gamma(\omega) \quad , \\ +\infty & \text{otherwise} \quad ; \end{cases}$$

(2.3) the distance (function) process $d_\Gamma = d(\cdot,\Gamma)$ defined on $E \times \Omega$ with

$$d_\Gamma(x,\omega) = d(x,\Gamma(\omega)) = \min_{y \in \Gamma(\omega)} d(x,y) \quad .$$

In this section, we analyse the relationship between the convergence of a sequence of multifunctions and the convergence of the associated stochastic processes, by which is meant, as usual, the convergence in distribution of the finite dimensional sections of the stochastic process.

It is remarkable that the convergence in distribution of the Γ_ν to Γ is equivalent to the convergence in distribution of the corresponding distance processes d_{Γ_ν} to d_Γ, whereas it does not imply, nor does it follow from, the convergence of the indicator processes. We start with an example, involving nonempty convex sets, that illustrates some of the problems that may arise.

2.4 EXAMPLE. For all $\omega \in \Omega$, let

$$\Gamma_\nu(\omega) = \{(x_1,x_2) = \lambda(1,\nu^{-1}) \ , \ \lambda \geq 0\} \qquad \nu \in N$$

$$\Gamma(\omega) = \{(0,0)\}$$

and

$$\Gamma'(\omega) = \{(x_1,x_2) = \lambda(1,0) \ , \ \lambda \geq 0\} \qquad .$$

Note that for all $\omega \in \Omega$, $\lim_\nu \Gamma_\nu(\omega) = \Gamma'(\omega)$ and consequently the sequence of closed-valued measurable multifunctions $\{\Gamma_\nu \ \nu \in N\}$ converge in distribution to Γ'. However, the processes $\{1_{\Gamma_\nu}, \nu \in N\}$ and $\{\psi_{\Gamma_\nu}, \nu \in N\}$ converge, as processes, to the stochastic processes 1_Γ and ψ_Γ respectively, i.e., for every finite collection of R^2-vectors

$$x^1 = (x_1^1, x_2^1), x^2, \ldots, x^q \qquad ,$$

the random vectors

$$\{(1_{\Gamma_\nu}(x^1), \ldots, 1_{\Gamma_\nu}(x^q)), \nu \in N\}$$

converge in distribution to the (random) vector

$$1_\Gamma(x^1), \ldots, 1_\Gamma(x^q) \qquad ,$$

and similarly for ψ_{Γ_ν} and ψ_Γ. To see this, simply observe that (i) the variables $1_{\Gamma_\nu}(x^1), \ldots, 1_{\Gamma_\nu}(x^q)$ are independent, (ii) if $x \neq (0,0)$, then for ν sufficiently large the distribution function of $1_{\Gamma_\nu}(x)$ is given by

$$H_{x',\nu}(z) = \begin{cases} 0 & \text{if } z \leq 0 \\ 1 & \text{if } z > 0 \end{cases}$$

which is also the form of the distribution function of $1_\Gamma(x)$, and (iii) for all ν, the distribution function of $1_{\Gamma_\nu}(0,0)$ and that of $1_\Gamma(0,0)$ is the same, viz.,

$$H_{0,\nu}(z) = \begin{cases} 0 & \text{if } z \leq 1 \quad , \\ 1 & \text{if } z > 1 \quad . \end{cases}$$

But the Γ_ν do not converge in distribution to Γ. Simply let K be the ball with center (1,1) and of radius 1. Then

$$T_\nu(K) = \mu[\omega | \Gamma_\nu(\omega) \cap K \neq \emptyset] = 1 \quad ,$$

does not converge to

$$T(K) = \mu[\omega | \Gamma(\omega) \cap K \neq \emptyset] = 0 \quad ;$$

and moreover clearly K is a continuity set of T since for any $K' \subset K$ we have that $T(K') = 0$. In fact we already knew that the Γ_ν do not converge in distribution to Γ, since they actually converge to Γ'. It is also easy to verify that the processes d_{Γ_ν} converge to the process $d_{\Gamma'}$. For every $x = (x_1,x_2)$, the functions

$$d_{\Gamma_\nu}(x,\omega) = \begin{cases} d((x_1,x_2),(0,0)) & \text{if } \nu x_1 + x_2 \leq 0 \\ d[(x_1,x_2),[\frac{\nu x_1+x_2}{\nu+\nu^{-1}} , \frac{x_1+\nu^{-1}x_2}{\nu+\nu^{-1}}]] & \text{if } \nu x_1 + x_2 \geq 0 \end{cases}$$

converge, for all ω, to

$$d_{\Gamma'}(\omega,x) = \begin{cases} d((x_1,x_2),(0,0)) & \text{if } x_1 \leq 0 \\ d((x_1,x_2),(x_1,0)) & \text{if } x_1 \geq 0 \end{cases}$$

and from this it follows that for any $q \geq 1$, we have that the vectors

$$(d_{\Gamma_\nu}(x^1,\cdot),\ldots,d_{\Gamma_\nu}(x^q,\cdot))$$

converge in distribution to the vector

$$(d_\Gamma(x^1,\cdot),\ldots,d_\Gamma(x^q,\cdot)) \quad .$$

2.5 THEOREM. *Suppose that* $\{\Gamma;\Gamma_\nu,\nu \in N\}$ *are closed-valued measurable multifunctions. Then, the* Γ_ν *converge in distribution to* Γ *if and only if the stochastic processes* $\{(d_{\Gamma}(x,\cdot),x \in E),\nu \in N\}$ *converge in distribution to the stochastic process* $(d_\Gamma(x,\cdot),x \in E)$.

We give two separate proofs of this assertion, the first one providing a simple direct argument, whereas the second proof relies on a general result and yields some insight in the structural relation between the two types of convergence. The second proof is given in Section 3, it follows Example 3.17.

PROOF. We first show that the distribution function of a closed-valued measurable multifunction Λ determines and is completely determined by the distribution of the process d_Λ. For $a_1,\ldots,a_q$, any finite collection of positive real numbers and $x_1 \in E,\ldots,x_q \in E$, we have that

$$\mu\{\omega|d_\Lambda(x_i,\omega) \leq a_i, i = 1,\ldots,m\} = \mu\{\omega|\Lambda(\omega) \cap B_{a_i}(x_i) \neq \emptyset, i = 1,\ldots,m\} \quad .$$

This in turn can be (uniquely) expressed as the sum and difference of probabilities associated with sets of the type $\{\omega|\Lambda(\omega) \cap K \neq \emptyset\}$ with $K \in K^{ub}$. We do this for $q = 2$, the generalization is immediate

$$\begin{aligned}
&\mu\{\omega|\Lambda(\omega) \cap B_{a_i}(x_i) \neq \emptyset \ , \ i = 1,2\} \\
&= \sum_{i=1} \mu\{\omega|\Lambda(\omega) \cap B_{a_i}(x_i) \neq \emptyset\} - \mu\{\omega|\Lambda(\omega) \cap (B_{a_1}(x_1) \cup B_{a_2}(x_2)) \neq \emptyset\} \\
&= T_\Lambda(B_{a_1}(x_1)) + T_\Lambda(B_{a_2}(x_2)) - T_\Lambda(B_{a_1}(x_1) \cup B_{a_2}(x_2)) \quad .
\end{aligned}$$

The result now follows directly from the above, because if (z_1,z_2) is a continuity point of the distribution function of the random vector $(d(x_1,\cdot),d(x_2,\cdot))$, then the sets $B_{z_1}(x_1), B_{z_2}(x_2)$ and $B_{z_1}(x_1) \cup B_{z_2}(x_2)$ are continuity sets for T_Γ--the distribution function of Λ--and vice versa. □

3. CONVERGENCE OF INTEGRANDS

Any function $f:E \times \Omega \to \overline{R}$ will be called an *integrand*, with $\overline{R} = R \cup \{\pm\infty\}$ denoting the extended reals. The function f is completely determined by its *epigraph multifunction*

$$\omega \longmapsto \text{epi } f(\cdot,\omega) = \{(x,\alpha) \in E \times R \mid \alpha \geq f(x,\omega)\} \quad .$$

We say that f is a *normal integrand* if its epigraph multifunction is closed-valued and measurable. The theory of normal integrands was introduced and developed by Rockafellar [7], [8] to study variational problems involving constraints. It also provides the natural framework to analyze the convergence of stochastic optimization problems, as is sketched out in Section 8. Vervaat [26] who also saw the concept of normal integrands emerge in his study of extremal processes refers to them as *random lower semicontinuous functions*.

Every normal integrand can be viewed as a stochastic process with lower semicontinuous (l.sc.) realizations; the functions $x \longmapsto f(x,\omega)$ are lower semicontinuous since their epigraphs are closed. On the other hand, any finite-valued stochastic process with l.sc. realizations is a normal integrand. This follows immediately from [8, Corollary 2E]. We give a direct proof of this fact for the reader unfamiliar with the general theory.

3.1 PROPOSITION. *Suppose* $f:E \times \Omega \to R$ *is a stochastic process with lower semicontinuous realizations, in other words* f *is a finite-valued integrand with* $\omega \longmapsto f(\omega,x)$ *measurable for all* $x \in R^n$ *and* x $\quad$ $f(\omega,x)$ *l.sc. for all* $\omega \in \Omega$. *Then* f *is a normal integrand.*

PROOF. We need to show that the closed-valued multifunction $\omega \longmapsto \text{epi } f(\cdot,\omega) = \Lambda(\omega)$ is measurable. To do so it suffices to show that Λ admits a *Castaing representation*, i.e., that dom Λ $= \{\omega \mid \Lambda(\omega) \neq \emptyset\} \in A$ and there exists a countable collection $\{\nu_k, k \in N\}$ of measurable functions from dom Λ into E such that for all $\omega \in$ dom Λ,

$$\mathrm{cl}\ [\cup_{k=1}^{\infty} \nu_k(\omega)] = \Lambda(\omega) \quad ,$$

cf. [8, Theorem 1B]. Since f is finite, dom $\Lambda = \Omega \in A$. Now let

$$D = \{(a_k, \alpha_k) \in E \times R \ , \ k \in N\}$$

be a countable dense subset of $E \times R$, and for $k = 1,\ldots$ define the functions

$$\nu_k(\omega) = (a_k \ , \ \max\ [\alpha_k, f(a_k, \omega)])$$

They clearly determine a Castaing representation for Λ. □

At first it might appear that, at least in terms of the classical analysis of stochastic process, the requirement that the processes have l.sc. realizations is rather limiting. That is true, in some sense, and in Section 7 we suggest another approach which overcomes the difficulties one might have with this restriction. However, note that *any cád-lág process* (whose realizations are right continuous and have at all points left limits) admits trivially a *modification* with l.sc. realizations. Thus the class of normal integrands includes not only stochastic processes with extended real values, a form in which they arise in stochastic optimization for example, but also a very wide class of the "standard" real-valued processes.

Crucial to the development is the fact that for stochastic processes representable by normal integrands, we can introduce a notion of convergence which not only is the appropriate one if we are interested in extremal properties of the processes, but also in many situations provides us with a more satisfactory approach to convergence questions as the classical functional approach. We start with a short description of the epi-topology on the space of lower semicontinuous functions.

A collection of lower semicontinuous functions $\{f_\nu : E \to \overline{R}\ ,\ \nu \in N\}$ is said to epi-converge[†] to the function $f : R^n \to \overline{R}$ *at the point* x if

(3.2) given any subsequence of functions $\{f_{\nu_k}, k \in N\}$ and any sequence $\{x_k, k \in N\}$ converging to x, we have that

$$\liminf_{k\to\infty} f_{\nu_k}(x_k) \geq f(x) \qquad ,$$

and

(3.3) there exist a sequence $\{x_\nu, \nu \in N\}$ converging to x such that

$$\limsup_{\nu\to\infty} f_\nu(x_\nu) \leq f(x) \qquad .$$

If the above holds at every point x in R^n, we say that the collection *epi-converges* to f. This type of convergence, introduced by Wijsman [11], is closely related to the notion of pointwise convergence but it is neither implied nor does it imply pointwise convergence. Simply note that (3.2) implies but does not follow from

(3.4) $$\liminf_{\nu\to\infty} f_\nu(x) \leq f(x) \qquad ,$$

whereas (3.3) follows from but does not imply

(3.5) $$\limsup f_\nu(x) \geq f(x) \qquad .$$

Consider for example the sequence $\{f_\nu : R \to \overline{R}\ ,\ \nu \in N\}$ with $f_\nu(\nu^{-1}) = 0$ and 1 otherwise, which epi-converges to f with $f(0) = 0$ and 1 otherwise, but pointwise converges to the function identically 1. The terminology epi-convergence comes from the

---

[†]In the context of the Calculus of Variations, this type of convergence is sometimes called Γ-convergence, cf. [9], [10], for example; in [26] Vervaat refers to it as the inf-vague convergence.

fact that a collection of functions epi-converges if and only if the epigraphs of the functions converge as sets as was already observed by Mosco [12]. Further details can be found in [13], [14], [15], [16], [17].

Epi-convergence engenders a topology $\mathcal{E}$ on the *space* SC(E) *of lower semicontinuous functions* defined on E and with values in the extended reals [17]; note that SC(E) is a convex cone. In view of the preceding comments, it can be identified with the restriction of the topology $\mathcal{T}$ on the closed subsets of $E \times R$ to the space of epigraphs. It can be verified [17, Section IV], [18, Theorem 4] that this topology $\mathcal{E}$ on SC(E) can be generated by the subbase consisting of the sets of the type

$$\mathcal{E}^{K,a} = \{f \in SC(E) \,|\, \inf_K f > a\} \qquad K \in \mathcal{K}\ ,\ a \in R \quad ,$$

and

$$\mathcal{E}_{G,a} = \{f \in SC(E) \,|\, \inf_G f < a\} \qquad G \in \mathcal{G}\ ,\ a \in R \quad .$$

Recall that $\mathcal{K}$ and $\mathcal{G}$ denote the space of compact and open subsets of E respectively. For obvious reasons we refer to this topology $\mathcal{E}$ as the *epi-topology* on SC(E). The topological space $(SC(E),\mathcal{E})$ is *regular* and *compact* [17, Corollary 4.3]. In particular this means that every sequence of lower semicontinuous functions contains a convergent subsequence. Moreover, the properties of E allow us to replace, in the construction of a subbase for $\mathcal{E}$, the class of sets $\mathcal{K}$ and $\mathcal{G}$ by

$$\{B_\varepsilon(x) \,|\, \varepsilon \in Q_+\ ,\ x \in D \subset E\}$$

and

$$B^\circ_\varepsilon(x) \,|\, \varepsilon \in Q_+\ ,\ x \in D \subset E\}$$

respectively, and restrict a to Q and x to D, where D is a countable dense subset of E, cf. the Introduction. Thus the topological space $(SC(E),\mathcal{E})$ has a countable base, and hence is *metrizable* and *separable*.

As already pointed out at the beginning of this Section, although pointwise and epi-convergence are intimately related, in general they do not imply each other. In other words, this means that the epi-topology $\mathcal{E}$ and the product topology $\mathcal{P}$ (that corresponds to pointwise convergence) do not coincide on SC(E). However they do on epi-lower semicontinuous subsets of SC(E) [17, Theorem 2.18 and Theorem 4.6]. A set $\mathcal{Q} \subset SC(E)$ is *equi-lower semicontinuous*, if to all $x \in E$ and $\varepsilon > 0$ sufficiently small there corresponds $V \in \mathcal{N}(x)$, a neighborhood of x, such that for all $f \in \mathcal{Q}$.

$$\text{(3.6)} \qquad \inf_{y \in V} f(y) \geq \min\,[\varepsilon^{-1}, f(x) - \varepsilon] \quad .$$

Let us now return to normal integrands and their convergence.

3.7 DEFINITION. *We say that the sequence of normal integrands* $\{f_\nu : E \times \Omega \to \overline{R}\ ,\ \nu \in N\}$ *epi-converge almost surely [in probability, in distribution respectively] to the normal integrand* $f : E \times \Omega \to \overline{R}$, *if the corresponding epigraph multifunctions* $\{\text{epi } f_\nu : \Omega \rightrightarrows E \times R\ ,\ \nu \in N\}$ *converge almost surely [in probability, in distribution respectively] to the epigraph multifunction* epi $f : \Omega \rightrightarrows E \times R$.

We note that as a consequence of [1, Corollary 3.2] and the definition of almost sure convergence of normal integrands, it follows that if $\{f_\nu, \nu \in N\}$ is a sequence of normal integrands and they have almost surely an epi-limit, then *this limit is also a normal integrand*, ignoring possibly a set of measure 0.

3.8 PROPOSITION. *Suppose* $\{f_\nu : E \times \Omega \to \overline{R}\ ,\ \nu \in N\}$ *is a sequence of normal integrands. They epi-converge almost surely to the normal integrand* $f : E \times \Omega \to \overline{R}$ *if and only if for all* $\omega \in \Omega$, *except possibly on a set of measure* 0, *we have that for all* $x \in E$

(3.9) *for any subsequence* $\{f_{\nu_k}, k \in N\}$ *and any sequence* $\{x_k, k \in N\}$ *converging to* x *we have that*

$$\liminf_{k \to \infty} f_{\nu_k}(x_k, \omega) \geq f(x, \omega) \quad ,$$

and

(3.10) *there exist a sequence* $\{x_\nu, \nu \in N\}$ *converging to* x *such that*

$$\limsup_{\nu\to\infty} f_\nu(x_\nu,\omega) \leq f(x,\omega) \quad .$$

PROOF. It really suffices to observe that conditions (3.9) and (3.10) are nothing more than the conditions for epi-convergence and that they hold if and only if the epigraph converge [17, Proposition 1.9]. Thus the epigraph multifunctions converge almost surely if and only if (3.9) and (3.10) hold for almost all ω. □

To characterize the convergence in probability of normal integrands it is useful to introduce some perturbations of the given integrands. Let g denote an arbitrary function defined on E and with values in the extended reals. For any positive number a, we denote by g^a the function defined by

$$g^a(x) := \inf_{y\in aB} g(x - y) - a$$

where $aB = B_a(0)$ is the closed ball of radius a. It is not difficult to verify that

$$\text{epi } g^a = \text{epi } g + \{(x,\alpha) \mid x \in aB \ , \ |\alpha| \leq a\} \quad .$$

If epi g is closed, i.e., if g is l.sc., then so is epi g^a since the second term of the (Minkowski) sum is compact. Suppose $f:E \times \Omega \to \overline{R}$ is a normal integrand then so is f^a. To see this simply observe that the epigraph multifunction

$$\omega \longmapsto \text{epi } f^a(\cdot,\omega) = \text{epi } f(\cdot,\omega) + \{(x,\alpha) \mid x \in aB \ , \ |\alpha| \leq a\}$$

is closed-valued and measurable, since for each ω it is the sum of a closed-valued measurable multifunction and a compact-valued constant (measurable) multifunction [8, Proposition 1I].

3.11 PROPOSITION. *Suppose* $\{f_\nu : E \times \Omega \to \overline{R}\ ,\ \nu \in N\}$ *is a sequence of normal integrands. They epi-converge in probability to the normal integrand* $f: E \times \Omega \to \overline{R}$ *if and only if for all* $\varepsilon > 0$, $r > 0$ *and* $x \in E$

$$\lim_{\nu\to\infty} \mu\{\omega \mid \ y \text{ such that } d(x,y) \leq r, \text{ and either}$$

$$-r \leq f_\nu(y,\omega) < f^\varepsilon(y,\omega) \leq r \text{ or}$$

$$-r \leq f(y,\omega) < f_\nu^\varepsilon(y,\omega) \leq r\} = 0 \quad .$$

PROOF. It really suffices to observe that there exist y satisfying the conditions laid out if and only if

$$[(\text{epi } f_\nu(\cdot,\omega)\setminus \varepsilon\text{-epi } f(\cdot,\omega)) \cup (\text{epi } f(\cdot,\omega)\setminus \varepsilon\text{-epi } f_\nu(\cdot,\omega)] \cap B_r(x) \neq \emptyset \quad .$$

This is exactly the definition of convergence in probability of the multifunctions $\{\text{epi } f_\nu(\cdot,\omega), \nu \in N\}$ to epi $f(\cdot,\omega)$, [1, Section 5].

In parallel to the development in Section 1, it is possible to associate to each normal integrand a distribution function defined on $K^{ub} \times R$, cf. Section 1 and the expressions given for the subbase of the epi-topology. The epi-convergence in distribution of normal integrands can thus be given a characterization similar to that given by Theorem 1.2 for closed-valued measurable multifunctions. It is an excellent exercise that the conscientious reader would not want to bypass. As a consequence of [1, Section 5] and Proposition 1.18 we can conclude the following:

3.12 PROPOSITION. *Suppose* $\{f; f_\nu : E \times \Omega \to \overline{R}, \nu \in N\}$ *is a collection of normal integrands and* μ *is a probability measure. Then the* f_ν *epi-converge a.s. to* f *if and only if they epi-converge* μ*-almost uniformly to* f. *Moreover almost sure epi-convergence implies epi-convergence in probability which in turn implies epi-convergence in distribution.*

As already indicated at the beginning of this Section normal integrands can be viewed as stochastic processes with l.sc. realizations albeit with values in the extended reals.

The main result of the Section is a characterization of the normal integrands for which epi-convergence in distribution and convergence in the classical sense (Kolmogorov) for stochastic processes coincide. The key ingredient is the concept of equi-lower semicontinuity defined by relation (3.6).

3.13 THEOREM. *Suppose that* $\{f_\nu : E \times \Omega \to \overline{R}, \nu \in \Omega\}$ *is a collection of a.s. equi-lower semicontinuous, normal integrands, i.e., there exists* $\Omega' \in A$ *with* $\mu(\Omega') = 1$ *such that for all* $\omega \in \Omega'$, *the collection of functions*

$$\{x \longmapsto f_\nu(x,\omega) : E \to \overline{R} \ , \ \nu = 1,\dots\}$$

is equi-lower semicontinuous. Then the f_ν *epi-converge in distribution to a normal integrand* $f : E \times \Omega \to \overline{R}$ *if and only if the stochastic processes*

$$\{f_\nu(x,\cdot) \ , \ x \in E \ ; \ \nu = 1,\dots\}$$

converge in (the classical sense) to the process $(f(x,\cdot), x \in E)$.

PROOF. The product topology P on SC(E) can be generated by the base of open neighborhoods of the type

$$V_{\varepsilon;x_1,\dots,x_q}(g) = \{f \in SC(E) \,|\, |f(x_i) - g(x_i)| < \varepsilon \ , \ i = 1,\dots,q\}$$

with $\varepsilon > 0$, $q \in N$ and $x_1,\dots,x_q$ a collection of points in E. The topological space $(SC(E),P)$ is compact and regular but not separable. However any equi-lower semicontinuous subset $\mathcal{Q}$ of SC(E) equipped with the relative P-topology is a separable metric space since on $\mathcal{Q}$ the epi-topology and the product topology coincide [17, Theorem 2.18 and Theorem 4.6]. This also implies that the Borel fields on $\mathcal{Q}$ generated by the E-open or P-open sets are the same.

Let $\pi_{x_1,\ldots,x_q}$ denote the natural projection with respect to $x_1,\ldots,x_q$ (all in E) from $\mathcal{Q}$ to R^q, where

$$\pi_{x_1,\ldots,x_q} f = (f(x_1),\ldots,f(x_q))$$

The *finite dimensional sets*

$$\mathcal{U} = \{\pi^{-1}_{x_1,\ldots,x_q} H \mid H \in S^q; x_i \in E \ , \ i = 1,\ldots,q \text{ and } q \in N\}$$

where S^q denotes the Borel sigma-field on R^q, contain the open neighborhoods $V_{\varepsilon;x_1,\ldots,x_q}(g)$ that determine a base for $\mathcal{P}$. Moreover, since $(\mathcal{Q},\mathcal{P} = \mathcal{E})$ is separable it follows that $\mathcal{U}$ generates the Borel field on $\mathcal{Q}$. This is of crucial importance to the proof of this theorem, since it follows that for probability measures defined on $(\mathcal{Q},S_{\mathcal{Q}})$, the class $\mathcal{U}$ is a *convergence-determining class* [3, p. 15]. To see this we rely on a theorem of Kolmogorov and Prohorov [3, Theorem 2.2] and observe that $\mathcal{U}$ is closed under finite intersections and every open set is a countable union of elements of $\mathcal{U}$; recall that the open sets $V_{\varepsilon;x_1,\ldots,x_q}$ belong to $\mathcal{U}$ and that $(\mathcal{Q},\mathcal{P} = \mathcal{E})$ is Lindelöf.

The projections $\pi_{x_1,\ldots,x_q}$ from $(SC(E),\mathcal{E})$ to R^q are in general *not* continuous, epi-convergence would then imply pointwise convergence. But on $\mathcal{Q}$, an equi-lower semicontinuous subset, these projections are continuous. In turn, this implies that weak*-convergence of probability measures defined on $(\mathcal{Q},S_{\mathcal{Q}})$ yields the weak*-convergence of their projections [3, Section 5]. The next lemma encapsulates what we have shown so far.

3.14 LEMMA. *Suppose* $\mathcal{Q}$ *is an equi-lower semicontinuous subset of* $(SC(E),\mathcal{E})$ *and* $\{P_\nu, \nu = 1,\ldots\}$ *is a sequence of probability measures defined on the Borel field generated by the* $\mathcal{E}$*-open sets of* $\mathcal{Q}$. *Then the* P_ν *weak***-converge to a probability measure* P *(defined on the Borel field generated by the* $\mathcal{E}$*-open sets of* SC(E)) *if and only if for all finite dimensional sets* $A \in \mathcal{U}$.

$$P(A) = \lim_{\nu\to\infty} P_\nu(A)$$

PROOF OF THEOREM 3.13 (Continued). We now consider $\{Q^{\nu}, \nu = 1,\ldots\}$ and Q the probability measures induced by the normal integrands $\{f_{\nu}, \nu = 1,\ldots\}$ and f on the measure space $(SC(E), \mathcal{S}_E)$ where $\mathcal{S}_E$ denotes the Borel sigma-field generated by the $\mathcal{E}$-open sets. The Theorem will be proved if we show that the Q_{ν} weak*-convergence to Q if and only if they converge on $\mathcal{U}$--the class of finite dimensional sets--when the $\{f_{\nu}(\cdot,\omega); \nu = 1,\ldots,\omega \in \Omega'\}$ determine an equi-lower semicontinuous subset of $SC(E)$. But this is precisely the content of Lemma 3.14. □

Before we continue, we note that all the results obtained for epi-convergence, the space of lower semicontinuous functions, and normal integrands with lower semicontinuous sections have their counterpart in the mirror-setting: hypo-convergence, the space of upper semicontinuous functions and normal integrands with upper semicontinuous sections. Recall that the *hypograph* of a function f is defined by

$$\text{hypo } f = \{(x,\alpha) \in E \times R \mid \alpha \leq f(x)\}$$

For the record, we give here the definition of equi-upper semicontinuity and the corresponding version of Theorem 3.13. Let $-SC(E)$ denote the space of upper semicontinuous functions. A subset $\mathcal{Q} \subset -SC(E)$ is *equi-upper semicontinuous*, if to all $x \in E$ and $\varepsilon > 0$ sufficiently small there corresponds $V \in \mathcal{N}(x)$ such that for all $f \in \mathcal{Q}$

$$\sup_{y \in V} f(y) \leq \max\,[-\varepsilon^{-1}\,,\ f(x) + \varepsilon] \tag{3.15}$$

3.16 THEOREM. *Suppose that* $\{f_{\nu}: E \times \Omega \to \overline{R}, \nu \in \Omega\}$ *is a sequence of a.s. equi-upper semicontinuous normal integrands (with upper semicontinuous sections). Then the* f_{ν} *hypo-converge in distribution to a normal integrand (with upper semicontinuous sections)* $f: E \times \Omega \to \overline{R}$ *if and only if the stochastic processes*

$$\{f_{\nu}(x,\cdot)\ ,\ x \in E\ ;\ \nu = 1,\ldots\}$$

converge (as stochastic processes) to $(f(x,\cdot)\ ,\ x \in E)$.

We could now define a concept of a.s. equi-continuity and combining Theorems 3.13 and 3.16, obtain limit results for processes with continuous paths but defined on E not as usual on a compact subset of E. We do not work out those results, cf. [17, Section 4] for how this program could be carried out.

The equi-semicontinuity with probability 1 is, as we have seen sufficient to ensure the equivalence of these two types of convergence for sequences of normal integrands. But in fact it is *nearly necessary*. From [17, Theorem 2.18] it follows that if for all $\omega \in \Omega'$ with $\mu(\Omega') = 1$, the functions $\{x \longrightarrow f_\nu(x,\omega), \nu = 1,\dots\}$ both epi- and pointwise converge to $x \longrightarrow f(x,\omega) > -\infty$, then the collection is equi-lower semicontinuous with probability 1. On the other hand, if for some x, there is some subset $A \subset \Omega$ of positive probability such that for all $\omega \in A$, either the $f_\nu(\cdot,\omega)$ epi-converge to $f(\cdot,\omega)$ at x but do not pointwise converge at x, or vice-versa, then convergence in distribution of the finite dimensional sections never implies epi-convergence in distribution, as can easily be verified. However, the sequence may fail to be equi-lower semicontinuous with probability 1 and still it may converge in distribution, in both the epi- and the pointwise sense, to limit processes that are distribution-equivalent. The following example illustrates such a situation.

3.17 EXAMPLE. Consider the sequence of normal integrands, for $\nu = 1,\dots$, let

$$f_\nu(x,\omega) = \begin{cases} -\nu(x-\omega) & \text{if } \omega \le x \le \omega + \nu^{-1} \\ +\infty & \text{otherwise} \end{cases}$$

with $(\Omega, A, \mu) = ([0,1], \mathcal{B}_{[0,1]}$, uniform measure). Then for all ω, the $f_\nu(\cdot,\omega)$ epi-converge to

$$f(\cdot,\omega) = \begin{cases} -1 & \text{if } x = \omega, \\ +\infty & \text{otherwise,} \end{cases}$$

and they pointwise converge to

$$f'(\cdot,\omega) = \begin{cases} 0 & \text{if } x = \omega, \\ +\infty & \text{otherwise.} \end{cases}$$

Let $F_\nu(x,\cdot):\bar{R} \to [0,1]$ be the distribution functions associated to the random variables $f_\nu(x,\cdot)$. They are given by the expression: for $x \in)0,1]$ and $\nu^{-1} \le x$

$$F_\nu(x,z) = \begin{cases} 0 & \text{if} \quad z \le 1 \\ \nu^{-1}(z+1) & \text{if} \quad -1 \le z \le 0 \\ \nu^{-1} & \text{if} \quad 0 \le z < \infty \\ 1 & \text{if} \quad z = \infty \end{cases} .$$

that converge (as distribution functions) to $F(x,\cdot)$ with

$$F(x,z) = \begin{cases} 0 & \text{if} \quad z < \infty \\ 1 & \text{if} \quad z = \infty \end{cases}$$

This is the distribution function (on $\bar{R}$) of both $f(x,\cdot)$ and $f'(x,\cdot)$, again taking $x \in)0,1]$. For $x \notin)0,1]$ the expressions for $F_\nu(x,\cdot)$ are a little more involved but the limit distribution function is again the same. Since for all ω, there is some x at which the pointwise limit and the epi-limit are different, we have

$$\mu\{\omega | f_\nu(\cdot,\omega) \ , \ \nu = 1,\ldots \text{ are equi-l.sc.}\} = 0$$

but the normal integrands both epi-converge in distribution and in the (classical) sense of stochastic processes to f, or f' for that matter. □

Theorem 3.13 and the comments that follow also enable us to return with gained insight to the questions raised in Section 2 in connection with the convergence of stochastic processes associated to multifunctions. We start with a new proof of Theorem 2.5, which shows that it is actually a Corollary to Theorem 3.13.

SECOND PROOF OF THEOREM 2.5. Let $\mathcal{D} = \{d_F : E \to \bar{R} | F \in \mathcal{F}\}$ be the space of distance functions. There is a natural bijection relating the elements of $\mathcal{F}$ and $\mathcal{D}$. Let us observe that the topology $\mathcal{T}$ on $\mathcal{F}$ corresponds to that generated on $\mathcal{D}$ by the subbase consisting of the sets of type

$$\mathcal{D}^K = \{d_F | 0 \notin d_F(K)\} \qquad K \in \mathcal{K} \quad ,$$

$$\mathcal{D}_G = \{d_F | 0 \in d_F(G)\} \qquad G \in \mathcal{G} \quad ,$$

where for any set Q, $d_F(Q) = \{a = d_F(x), x \in Q\}$. This is in fact the epi-topology on $\mathcal{D}$. To see this, recall that the epi-topology is generated by the subbase

$$\mathcal{E}^{K,a} = \{d_F | a < d_F(K)\} = \{d_F | 0 \in d_F(\text{cl } aK)\} \quad , \quad K \in \mathcal{K} \ , \ a \in R_+$$

and

$$\mathcal{E}_{G,a} = \{d_F | a \in d_F(G)\} = \{d_F | 0 \in d_F(a^\circ G)\} \quad , \quad G \in \mathcal{G} \ , \ a \in R_+$$

it is sufficient to consider $a \in R_+$ because the functions d_F are nonnegative. Since for every $a \in R_+$--the nonnegative reals--the sets aK are compact and $a^\circ G$ are open, it follows that $\mathcal{D}^K = \mathcal{E}^{K,a}$ and $\mathcal{D}_G = \mathcal{E}_{G,a}$.

The set $\mathcal{D} \subset SC(E)$ is equi-lower semicontinuous. To prove this it certainly suffices to show that:

to every $x \in E$ and $\varepsilon > 0$, there corresponds $V \in \mathcal{N}(x)$ such that

$$\inf_{y \in V} d_F(y) \geq d_F(x) - \varepsilon$$

for all $F \in \mathcal{F}$.

Simply take $V = \{y | d(x,y) < \varepsilon\} = B^\circ_\varepsilon(x)$, then for any $y \in B^\circ_\varepsilon(x)$ we have that

$$d_F(x) \leq d_F(y) + d(x,y) \leq d_F(y) + \varepsilon \quad .$$

Now consider $\{\Gamma; \Gamma_\nu, \nu \in N\}$ a collection of closed-valued measurable multifunctions. The multifunctions $\{\Gamma_\nu, \nu \in N\}$ converge in distribution to Γ if and only if the distance processes $\{d_{\Gamma_\nu}, \nu \in N\}$ epi-converge in distribution to the distance process d_Γ, as follows directly from the relationship between $(\mathcal{F},\mathcal{T})$ and $(\mathcal{D},\mathcal{E})$. Since $\mathcal{D}$ is equi-lowersemicontinuous, we simply appeal to Theorem 3.13 to complete the proof. □

Contrary to the set of distance functions, the space of indicator functions of closed subsets of E is not an equi-lower semicontinuous subset of SC(E). Thus it is possible to find sequences whose epi-limit and pointwise limit differ on a subset with positive measure. This observation was certainly instrumental in the construction of Example 2.4. We give here a necessary and sufficient condition for equi-lower semicontinuity for a sequence of indicator functions which can then be applied to sequences of multifunctions.

3.18 PROPOSITION. *Suppose that* $\{F;F_\nu, \nu \in N\}$ *is a collection of closed subsets of* E *with* $F = \lim_\nu F_\nu$. *Then the collection of indicator functions* $\{\psi_{F_\nu}, \nu \in N\}$ *is equi-l.sc. if and only if to every* $x \in F$ *there corresponds* ν_x *such that* $x \in F_\nu$ *for all* $\nu \geq \nu_x$.

PROOF. By hypothesis, $x \in F$ implies that for all $\nu \geq \nu_x$, $x \in F_\nu$ and thus for all $y \in E$.

$$\psi_{F_\nu}(x) = 0 \leq \psi_{F_\nu}(y) + \varepsilon \quad .$$

For all $\nu < \nu_x$, the l.sc. of the ψ_{F_ν} implies that for all $\varepsilon > 0$ there exist a neighborhood V such that for all $y \in V$

$$\psi_{F_\nu}(y) \geq \min\,[\varepsilon^{-1}\ ,\ \psi_{F_\nu}(x) - \varepsilon]$$

for all $\nu > \nu_x$, since there are only a finite number of ν. The two preceding inequalities yield (3.6).

Now suppose that the collection $\{\psi_{F_\nu}, \nu \in N\}$ is equi-l.sc. and that $F = \lim F_\nu$, or equivalently that the $\{\psi_{F_\nu}, \nu \in N\}$ epi-converge to ψ_F. If for some x, there is no ν_x such that $x \in F_\nu$ for all $\nu \geq \nu_x$, it means that there exists a subsequence $\{F_\mu, \mu \in M \subset N\}$ of sets such that $x \notin F_\mu$. Conditions (3.6) yields for all $\varepsilon > 0$ sufficiently small, a neighborhood V of x such that for all $y \in V$

$$\psi_{F_\mu}(y) \geq \varepsilon^{-1} \qquad ,$$

for μ sufficiently large, or equivalently there is a neighborhood V of x such that for all μ sufficiently large

$$F_\mu \cap V = \emptyset$$

This implies that $x \notin \lim F_\mu$ and thus contradicts the assumption that $F = \lim F_\nu$, see for example, [1, Theorem 2.2.i_a]. □

3.19 COROLLARY. *Suppose that* $\{\Gamma;\Gamma_\nu,\nu \in N\}$ *is a collection of closed-valued measurable multifunctions with domain* Ω. *Suppose moreover that for every* $\omega \in \Omega\backslash A$ *with* $\mu(A) = 0$, *we have*

$$x \in \Gamma(\omega) \textit{ implies } x \in \Gamma_\nu(\omega) \textit{ for } \nu \textit{ sufficiently large.}$$

Then the multifunctions Γ_ν *converge in distribution to* Γ *if and only if the stochastic processes* $\{\psi_{\Gamma_\nu}(x,\cdot),x \in E;\ \nu = 1,\ldots\}$ *converge to the stochastic process* $(\psi_\Gamma(x,\cdot),x \in E)$.

A similar result also holds for the indicator processes 1_F. One passes this time through hypo-converge and equi-upper semi-continuity. In fact, we can obtain the results directly from the above if one observes that for any closed set F

$$1_F = 1 - \frac{2}{\pi} \arctan \psi_F \qquad ,$$

and now all the arguments can just be repeated with the obvious adjustments for signs.

We close this Section with an observation concerning continuous convergence and lower semicontinuous convergence, convergence concepts that have surfaced in the context of approximation schemes for stochastic dynamic programming problems [19], [20]. A collection of functions $\{f_\nu:E \to \overline{R},\nu \in N\}$ is said to *continuously converge* to $f:E \to \overline{R}$ if at all $x \in E$ conditions (3.2) and (3.20) are satisfied, where

(3.20) given any sequence $\{x_\nu, \nu \in N\}$ converging to x we have

$$\limsup_{\nu\to\infty} f_\nu(x_\nu) \leq f(x) \quad .$$

They *lower-semicontinuously converge* to f if at all $x \in E$ conditions (3.2) and (3.5) are satisfied.

Continuous convergence implies lower semicontinuous convergence and clearly both types of convergence imply pointwise *and* epi-convergence. Thus, any sequence of normal integrands that either continuously or lower semicontinuously converge forms an equi-lower semicontinuous set. Hence Theorem 3.13 informs us that for normal integrands lower semicontinuous-convergence in distribution takes place only if their finite dimensional sections converge in distribution.

4. EQUI-LOWER SEMICONTINUITY IN PROBABILITY

Theorem 3.13 was proved by relying almost exclusively on topological arguments involving the relationship between the epi-topology and the product topology on the space of lower semi-continuous functions. Here we actually refine Theorem 3.13 by paying more attention to probabilistic structures, in particular we rely on the results obtained in Section 1. As in Section 3, we shall be working with sequences of normal integrands and will be concerned with the relationship between their epi-convergence in distribution and their convergence when viewed as stochastic processes.

Let $\{f; f_\nu : E \times \Omega \to \overline{R}, \nu \in N\}$ be a collection of normal integrands, $\{Q; Q_\nu, \nu \in N\}$ the induced probability measures on $(SC(E), S_E)$ and $\{T; T_\nu, \nu \in N\}$ the associated distribution functions, cf. Section 1. Relying on the characterization of the epi-topology given in Section 3, we see that the distribution function T associated to a normal integrand f can be defined on $K^{ub} \times R$ by

$$(4.1) \quad T(K,a) = Q(f \mid \inf_K f \leq a) = \mu\{\omega \mid \inf_{y \in K} f(y,\omega) \leq a\}$$

where $K \in K^{ub}$, $a \in R$ and as before K^{ub} denotes the compact subsets of E obtained as the finite union of closed balls. By abuse of notation we shall use T to denote a distribution function on $K^{ub} \times R$ as well as on arbitrary compact subsets of $E \times R$.

In Section 2 and 3 we have stressed the fact that for lower semicontinuous functions epi-convergence and pointwise convergence are not comparable. In view of this the next theorem is somewhat remarkable. At the conceptual level, the validity of this theorem rests on the fact that the Borel sigma-field S_E contains that generated by the subsets of SC(E) with respect to which the projection maps of $f \longmapsto \pi_x f = f(x)$ are measurable.

4.2 THEOREM. *Suppose* $\{f; f_\nu : E \times \Omega \to R, \nu \in N\}$ *is a collection of normal integrands. Then epi-convergence in distribution of the*

f_ν *to* f *implies the convergence in distribution of the stochastic processes* $\{f_\nu(\cdot,x), x \in E; \nu \in N\}$ *to the process* $\{f(\cdot,x), x \in E\}$.

PROOF. From Theorem 1.9, it follows that epi-convergence in distribution implies that for any finite collection $x_1,\ldots,x_p$ of points in E and reals $a_1,\ldots,a_p$

$$\mu\{\omega | f_\nu(x_1,\omega) > a_1,\ldots,f_\nu(x_p,\omega), > a_p\}$$

$$= \mu\{\omega | (x_1,a_1) \notin \text{epi } f_\nu(\cdot,\omega),\ldots,(x_p,a_p) \in \text{epi } f_\nu(\cdot,\omega)\}$$

$$= 1 - \mu\{\omega | \{(x_1,a_1),\ldots,(x_p,a_p)\} \cap \text{epi } f_\nu(\cdot,\omega) \neq 0\}$$

$$= 1 - T_\nu(\{(x_1,a_1),\ldots,(x_p,a_p)\})$$

converge to

$$1 - T(\{(x_1,a_1),\ldots,(x_p,a_p)\}) = \mu\{\omega | f(x_1,\omega) > a_1,\ldots,f(x_p,\omega) > a_p\}$$

provided that $\{(x_1,a_1),\ldots,(x_p,a_p)\}$ is a $\mathcal{E}$-continuity set. But here that means that

$$\mu\{\omega | f(x_l,\omega) > a_l \ , \ l = 1,\ldots,p\}$$

$$= \lim_{\varepsilon_l \to 0} \mu\{\omega | f(x_l,\omega) > a_l - \varepsilon_l \ , \ l = 1,\ldots,p\}$$

where $\varepsilon_1 > 0,\ldots,\varepsilon_p > 0$, which fails only if the point $(a_1,\ldots,a_p)$ is not a continuity point of the p-dimensional distribution function of the random vector

$$\omega \longmapsto (f(x_1,\omega),\ldots,f(x_p,\omega)) \quad . \quad \square$$

In order to prepare the tools necessary to derive a converse of the preceding theorem, we obtain a sufficient condition that allows us to determine the distribution function of a normal integrand from the distribution of its finite dimensional sections. To do so it is useful to introduce a class of normal integrands whose sections $x \longmapsto f(x,\omega)$ have uniform semicontinuity properties.

4.3 DEFINITION. *A function* $f:E \to R$ *is totally lower semicontinuous (totally l.sc.) if given any bounded set* $D \subset E$ *and* $\varepsilon > 0$ *sufficiently small, one can find a finite collection* $x_1,\dots,x_p$ *in* D *and associated neighborhoods* $V_1 \in N(x_1),\dots,V_p \in N(x_p)$ *such that*

$$(4.4) \qquad \cup_{l=1}^{p} V_l \text{ covers } D \quad ,$$

and for $l = 1,\dots,k$,

$$(4.5) \qquad \inf_{y \in V_l} f(y) \geq \min\,[\varepsilon^{-1}, f(x_l) - \varepsilon] \quad .$$

A normal integrand $f:E \times \Omega\ \overline{R}$ *is in probability totally l.sc. if given any* $\eta > 0$, *any bounded set* $D \subset E$ *and* $V_1 \in N(x_1),\dots,V_p \in N(x_p)$ *such that the* V_l *cover* D *and*

$$(4.6) \quad \mu\{\omega \in \Omega \,|\, \inf_{y \in V_l} f(y,\omega) \geq \min\,[\varepsilon^{-1}, f(x_l,\omega) - \varepsilon], l = 1,\dots,p\} > 1-\eta$$

A stochastic process $(X_t(\cdot), t \in T)$ with continuous paths $t \longmapsto X_t(\omega)$ is totally l.sc. in probability as can easily be verified. This is also the case if $(X_t(\cdot), t \in T)$ is a l.sc. modification of a stochastic process with cád-lág paths, i.e., in the space $D[T]$.

4.7 PROPOSITION. *Suppose* $f:E \times \Omega \to \overline{R}$ *is in probability a totally l.sc. normal integrand. Then the distribution function* T *of* f *is completely determined by its values on finite sets, i.e.,*

$$T(\{x_1,\dots,x_p\} \ , \ a)$$

with $x_l \in E$, $l = 1,\dots,p$ *and* $a \in R$.

PROOF. Let $a \in R$ and $K \in E$ be compact. Now, since f is a totally l.sc. in probability integrand, for any $\eta > 0$ and $\varepsilon < 0$ sufficiently small, in particular with $\varepsilon^{-1} > a + \varepsilon$, we can find $x_1,\dots,x_p$ and $V_1 \in N(x_1),\dots,V_p \in N(x_p)$ such that conditions (4.4) and (4.6) are satisfied. This means that $\mu(\Omega_\varepsilon) > 1 - \eta$ where

$$\Omega_\varepsilon = \{\omega \,|\, \inf_{y \in V_l} f(y,\omega) \geq \min\,[\varepsilon^{-1}, f(x_l,\omega) - \varepsilon] \ , \ l = 1,\dots,p\} \quad .$$

We have

$$\mu\{\omega \in \Omega | f(x_l,\omega) > a + \varepsilon \ , \ l = 1,\dots,p\}$$

$$= \mu\{\omega \in \Omega_\varepsilon | f(x_l,\omega) > a + \varepsilon \ , \ l = 1,\dots,p\}$$

$$+ \mu\{\omega \in \Omega\backslash\Omega_\varepsilon | f(x_l, \ ) > a \ , \ l = 1,\dots,p\}$$

$$< \mu\{\omega \in \Omega | f(y,\omega) > a \quad \text{for all } y \in \cup_l V_l\} + \eta \quad .$$

$$\leq \mu\{\omega \in \Omega | \inf_{y\in K} f(y,\omega) > a\} + \eta = 1 - T(K,a) + \eta \quad ,$$

where the last inequality follows from the lower semicontinuity of f and the compactness of K.
On the other hand for any choice $x_1,\dots,x_p$ in K, we always have

$$\mu\{\omega \in \Omega | f(x_l,\omega) > a + \varepsilon \ , \ l = 1,\dots,p\} \geq \mu\{\omega \in \Omega | \inf_{y\in K} f(y,\omega) > a+\varepsilon$$

Thus for any $\eta > 0$ and any $\varepsilon > 0$ sufficiently small, we can find $x_1,\dots,x_p$ such that

$$(4.8) \qquad T(K,a) - \eta < T(\{x_1,\dots,x_p\},a + \varepsilon) \leq T(K,a + \varepsilon)$$

from which the assertion directly follows because as ε tends to 0, $T(K,a + \varepsilon)$ goes to $T(K,a)$ as follows from (1.1). □

We now consider sequences of normal integrands and we are led to the following concept:

4.9 DEFINITION. *Let $\mathcal{D} = \{f:E \times \Omega \to \bar{R}\}$ be a family of normal integrands. We say that they are equi-totally lower semicontinuous in probability if given any $\eta > 0$, any bounded set $D \subset E$ and $\varepsilon > 0$ sufficiently small, there exist $x_1,\dots,x_p$ in D and $V_1 \in N(x_1),\dots,V_p \in N(x_p)$ such that the V_l cover D and (4.6) is satisfied for all $\nu \in N$.*

4.10 THEOREM. *Suppose $\{f;f_\nu:E \times \Omega \to R,\nu \in N\}$ is a collection of equi-totally l.sc. normal integrands in probability. Then they* *epi-converge in distribution if and only if the stochastic process* $\{f_\nu(x,\cdot), x \in E; \nu \in N\}$ *converge (in distribution) to the process* $\{f(x,\cdot), x \in E\}$.

PROOF. Theorem 4.2 yields the assertion in the "only if" direction, we only need to prove the "if" part. First note that

$$T(\{x_1,\ldots,x_p\},a) = 1 - \mu\{\omega | f(x_1,\omega) > a,\ldots,f(x_p,\omega) > a\}$$

and thus if $[\{x_1,\ldots,x_p\},a]$ is a continuity point of T, the point $(a_1 = a\ ,\ a_2 = a,\ldots,a_p = a)$ is a continuity point of the distribution function of the random vector

$$(f(x_1,\cdot)\ ,\ f(x_2,\cdot),\ldots,f(x_p,\cdot)) \quad .$$

This means that convergence of the normal integrands in the classical sense of stochastic processes implies that for any finite set $[\{x_1,\ldots,x_p\},a]$ we have

$$\lim_{\nu\to\infty} T_\nu(\{x_1,\ldots,x_p\},a) = T(\{x_1,\ldots,x_p\},a) \qquad ,$$

whenever $[\{x_1,\ldots,x_p\},a]$ belongs to C_T. But since the normal integrands are totally lower semicontinuous the distribution functions on $\mathcal{K}^{ub}$ or $\mathcal{K}$ are completely determined by the values of the T or T on the subsets of type $[\{x_1,\ldots,x_p\},a]$ cf. Proposition 4.7. Moreover the *equi*(-totally lower semicontinuity in probability) condition guarantees the convergence of $\{T_\nu(K), \nu \in N\}$ to T(K) for every $K \in \mathcal{K} \cap C_T$. This is more than what we need, since in view of Theorem 1.9, it would have been sufficient to prove the convergence of T_ν to T on $\mathcal{K}^{ub} \cap C_T$. □

5. COMPACTNESS THEOREMS

The fact that every sequence of closed-valued measurable multifunctions contains a subsequence converging in distribution has many consequences, in particular in the study of limit problems associated to a sequence of stochastic optimization problems but also in the convergence theory for stochastic processes, see Section 6. This section is devoted to proving this assertion and deriving a few elementary implications.

5.1 THEOREM. *Every collection of closed-valued measurable multifunctions is pre-compact with respect to the topology induced by convergence in distribution. In particular this implies that every sequence of closed-valued measurable multifunctions contains a subsequence converging in distribution to a closed-valued measurable multifunction (possibly the empty-valued multifunction).*

PROOF. Recall that (F,T) is a compact separable metric space. Let $\{Q_\alpha, \alpha \in I\}$ be the distributions defined on the sigma-field S_T induced by the multifunctions $\{\Gamma_\alpha, \alpha \in I\}$ where I is an arbitrary index set. Since F is separable it follows that the space of distributions on S_T is separable and also metrizable by the Prohorov metric [3, Appendix III]. Since F is compact, the collection of probability measures $\{Q_\alpha, \alpha \in I\}$ is tight and hence it follows from the Theorem of Prohorov and Varadarajan [3, Theorems 6.1 and 6.2] that they form a pre-compact set of probability measures on S_T. To complete the proof it now suffices to appeal to the Theorem of Engl and Wakolbinger [21]; they prove that the set of distributions of multifunctions defined on Ω and with values in F (a Polish space) is weak* closed. □

5.2 COROLLARY. *Any collection of normal integrands is pre-compact with respect to epi-convergence in distribution. In particular this implies that any sequence of normal integrands contains a subsequence epi-converging in distribution to a normal integrand.*

PROOF. Simply apply the Theorem to the epigraph multifunction. □

To know which multifunction (normal integrand respectively), defined on (Ω, G, μ) corresponds to this limit must usually be determined through other means. In the special case when $\Omega = [0,1]$ and μ is the uniform measure there is always the Prohorov construction [22] that yields a multifunction (normal integrand respectively) with the given probability. In the case of normal integrands the a.s. equi-lower semicontinuity condition or its equivalent often plays a useful role in characterizing a limit normal integrand, as can easily be surmised.

As a trivial application of this Theorem, we prove Helly's Theorem for random vectors. Let

$$\xi = (\xi_1, \xi_2, \ldots, \xi_p)$$

be a random vector defined on (Ω, A, μ) and with values in R^p. To each value $\xi(\omega)$ we associate the set

$$\Gamma(\omega) = \{(1, \xi_1(\omega)) \ , \ (2, \xi_2(\omega), \ldots, (p, \xi_p(\omega))\} \subset I \times R^p$$

where $I = \{1, 2, \ldots, n\}$. The multifunction $\Gamma: \Omega \rightrightarrows I \times R^p$ is closed-valued and measurable. Let P denote the distribution induced by Γ and F the (usual) distribution function of the random vector ξ, i.e.,

$$F(a_1, \ldots, a_p) = \mu\{\omega \mid \xi_1(\omega) < a_1, \ldots, \xi_p(\omega) < a_p\}$$

We note that there is a one-to-one correspondence between F and P defined through the relation

$$(5.3) \qquad F(a_1, \ldots, a_p) = P[\bigcap_{i=1}^{p} \{i\} \times)-\infty, a_i(]$$

on a class of open sets that determine P. Moreover, the continuity points of F correspond to P-continuity sets in this class. Let T be the distribution function of Γ defined as in Theorem 1.3.

5.4 COROLLARY (Helly's Theorem). *If* $\{F_\nu: R^p \to [0,1], \nu \in N\}$ *is a sequence of distribution functions associated with a sequence of* *random vectors* $\{\xi^\nu\ \Omega \to R^p, \nu \in N\}$, *then there exists a subsequence* $\{F_{\nu_k}, k = 1,\dots\}$ *and a function* $F: R^p \to [0,1]$ *that has all the properties of a distribution function except possibly that* $F(a)$ *does not necessarily go to* 1 *as* a_l *tends to* $+\infty$ *for* $l = 1,\dots,p$ *and such that*

$$\lim_{k\to\infty} F_{\nu_k}(a_1,\dots,a_p) = F(a_1,\dots,a_p)$$

for all continuity points of F.

PROOF. Let P_ν be the distribution and the T_ν be the distribution functions associated to the multifunctions Γ_ν defined as before by $\Gamma_\nu(\omega) = \times_{i=1}^{p}\{(i,\xi_i^\nu(\omega)\}$. Theorem 5.1 yields a subsequence $\{T_{\nu_k}, k = 1,\dots\}$ converging to a distribution function T on C_T or equivalently the subsequence $\{P_{\nu_k}, k = 1,\dots\}$ converge to a probability measure P with

$$\lim_k P_k[\bigcap_{i=1}^{p}\{i\} \times)-\infty\ ,\ a_1(] = P[\bigcap_{i=1}^{p}\{i\} \times)-\infty\ ,\ a_i(]$$

for all P-continuity set of this type. This completes the proof. The function F defined through relation (5.3) may fail to tend to 1 as the a_i tend to ∞; this comes from the fact that $P(I \times R^p)$ could be strictly less than 1, since positive probability may be assigned to the empty set. This happens when the $T_\nu(\omega)$ tend to the empty set on a set of positive probability which occurs if and only if the $\{\xi^\nu(\omega), \nu = 1,\dots\}$ are unbounded on a set of positive probability [1, Lemma 2.1]. □

In connection with Helly's Theorem it is worth noting that it can be derived directly from the compactness of the space of epigraphs. We sketch out the argument. Let us accept the convention that the distribution functions on R^k are everywhere continuous from below. Thus these are l.sc. functions defined on R^k and with values in [0,1]. Given any sequence of distribution functions, a subsequence epi-converges to an everywhere continuous from below function bounded by 0 and 1 [17, Theorem 4.6]. One then shows that this limit possesses the desired "monotonicity" property on each rectangle.

Compactness for processes whose paths are in C[0,1] or $D_{sc}[0,1]$ can be derived using Corollary 5.2; here $D_{sc}[0,1]$ is the space of l.sc. real-valued functions defined on [0,1] which we view as a l.sc. modification of D[0,1], the space of real-valued right-continuous functions with left-hand limits (cád-lág). In general, however, there is no guarantee that the limit probability measure, or equivalently the limit distribution function, is that of a process whose paths are of type C[0,1] or $D_{sc}[0,1]$.

We like to point out another approach to obtaining compactness criteria very much in keeping with the ideas developed here. It is possible to associate to every measure on (F, B_T) induced by closed-valued measurable multifunction Γ, a function $V: F \to [0,1]$, which has the same properties as a distribution function plus some "finite additive property, and which is upper semicontinuous on F with respect to the topology T. The weak*-convergence of measures on (F, B_T) can be identified with the hypo-convergence of functions of type V on F [23]. (For distribution functions defined on R^1 Vervaat [25] has also observed that weak convergence of the distribution function corresponded to their hypo-convergence.)

6. CONVERGENCE OF SELECTIONS

A measurable function v from Ω to E is called a *measurable selection* of Γ if $v(\omega) \in \Gamma(\omega)$ for all $\omega \in \mathrm{dom}\,\Gamma = \{\omega | \Gamma(\omega) \neq \emptyset\}$. The basic theorem on measurable selections asserts that a closed-valued multifunction is measurable if and only if it admits a *Castaing representation*, i.e., dom $\Gamma \in$ Q and there exists a countable collection of measurable selections $\{v_k, k \in N\}$ such that for all $\omega \in \mathrm{dom}\,\Gamma$

$$\mathrm{cl}\,\{\cup_k v_k(\omega)\} = \Gamma(\omega) \quad .$$

6.1 THEOREM. *Suppose* $\{\Gamma_\nu, \Omega \rightrightarrows E, \nu \in N\}$ *is a collection of closed-convex-valued measurable multifunctions converging in distribution to the closed-convex-valued measurable multifunction* $\Gamma: \Omega \rightrightarrows E$ *with* $\mu(\mathrm{dom}\,\Gamma_\nu = \mathrm{dom}\,\Gamma) = 1$. *Then there exist Castaing representations* $\{v_{\nu k}, \nu \in N\}$ *of the* Γ_ν *such that for every* k, *the* $\{v_{\nu k}, \nu \in N\}$ *converge in norm in distribution to a measurable function* v_k *such that* $\{v_k, k \in N\}$ *is a Castaing representation of* Γ.

PROOF. First observe that if the $\{\Gamma_\nu, \nu = 1,\ldots\}$ converge in distribution to Γ, then for all $z \in E$ the multifunctions $\{z + \Gamma_\nu, \nu = 1,\ldots\}$ converge in distribution to Γ. Let d be the Euclidean distance on E and fix D a countable dense subset of E whose elements are denoted by $\{a_k, k \in N\}$. For each $k \in N$ and $\omega \in \Omega$, let $x_{\nu k}(\omega) \in \Gamma_\nu(\omega)$ [$x_k(\omega) \in \Gamma(\omega)$ resp] be such that for $\nu = 1,\ldots$

$$d(a_k, x_{\nu_k}(\omega)) = d(a_k, \Gamma_\nu(\omega)) \quad .$$

and

$$d(a_k, x_k(\omega)) = d(a_k, \Gamma(\omega)) \quad .$$

Since the multifunctions are convex-valued, the $x_{\nu k}(\omega)$ and $x_k(\omega)$ are unique. Moreover the functions

$$\omega \longmapsto x_{\nu_k}(\omega) \qquad \text{and} \qquad \omega \longmapsto x_k(\omega)$$

are measurable [8]. The assertion now follow directly from Theorem 2.5 applied to the multifunctions $\Gamma_\nu - a_k$ (that converge in distribution to $\Gamma - a_k$) and the distance functions $d(0,\Gamma_\nu - a_k)$ which converge in distribution to $d(0,\Gamma - a_k)$. □

6.2 COROLLARY. *Suppose* $\{\Gamma_\nu : \Omega \rightrightarrows E, \nu = 1,\ldots\}$ *is a sequence of closed-convex-valued measurable multifunctions converging in distribution to a closed-convex-valued measurable multifunction* $\Gamma : \Omega \rightrightarrows E$. *Then there exist measurable selections of the* Γ_ν *that in norm converge in distribution to a measurable selection of* Γ.

It is easy to find a sequence of random vectors $\xi_\nu, \nu = 1,\ldots$ that in norm converge in distribution to a vector ξ but such that the vectors do not converge in distribution. Simply for each ν let $\xi_\nu(\omega) = -1$ for all ω and $\xi(\omega) = 1$ for all ω. But such a situation cannot arise if the vectors ξ_ν are viewed as multifunctions and when these multifunctions $\xi_\nu : \Omega \rightrightarrows R$ converge in distribution.

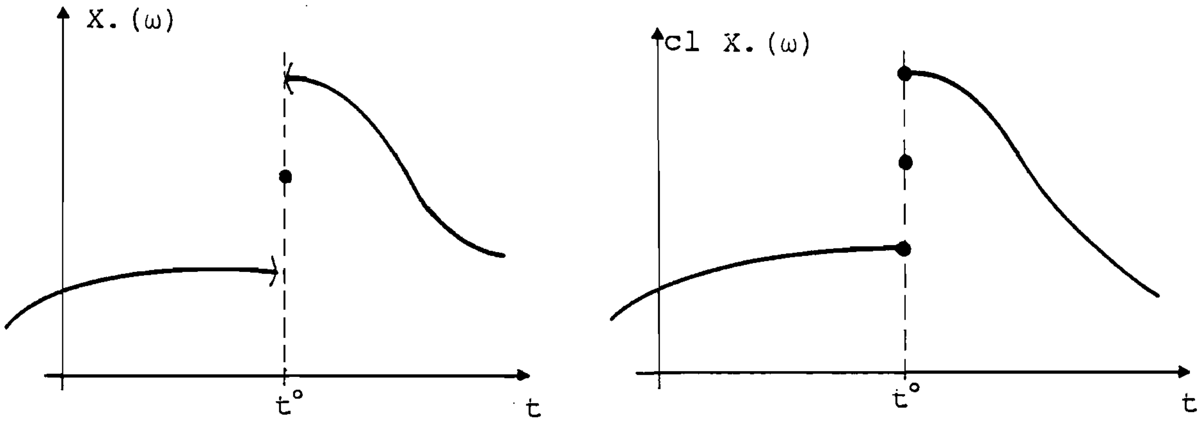
7. STOCHASTIC PROCESSES - THE WIENER PROCESS

We consider stochastic processes $(X_t(\cdot), t \in T)$ with values in R^p. The limit associated with a sequence of stochastic processes is usually obtained in the following manner: a function space $\mathcal{Y}$ is selected in terms of the expected properties of the paths of the limit process, (say the space of continuous functions), the approximating processes are redefined so that their paths are in $\mathcal{Y}$ (for example, random walk paths are "filled in" by interpolation), and some compactness criterion (the Arzela-Ascoli equi-continuity condition, for example) is invoked to obtain the relative compactness of the probability measures induced on $\mathcal{Y}$ by the stochastic processes of the sequence. In Section 3 we have already suggested another approach for processes that are (extended) real-valued. It identifies the paths of a process with normal integrands and consider convergence in terms of their associated epigraphs; this is particular useful if one is interested in extremal properties as we shall see in Section 8. In that setting we always have a "limit" in distribution (Corollary 5.2) and in favorable circumstances, such as under equi-semicontinuity conditions, the limit can be calculated explicitly. In this section we suggest another approach which more naturally fits the study of vector-valued processes. To illustrate this approach, we conclude with a derivation of Donsker's Theorem.

By the *closure of a* (vector-valued) *stochastic process* we mean the process obtained by taking the upper semicontinuous regularization of its paths (viewed as multifunctions). More precisely the closure of $(X_t(\cdot), t \in T)$ denoted by $((\mathrm{cl}\, X)_t(\cdot), t \in T)$ is defined as follows:

$$\mathrm{cl}\, X_t(\omega) = \{y = \lim y_k \mid y_k = X_{t_k}(\omega) \quad \text{for some } t_k \to t\} \quad .$$

This is a multivalued process, in general. The next figure illustrates this closure operation.



7.1 Figure: Closure of $(X_t(\cdot), t \in T)$

To each path $t \longmapsto \text{cl } X_t(\omega)$ we associate the closed set $\Gamma(\omega)$ defined by

$$\Gamma(\omega) = \{(t,y) \in T \times R^p \mid y \in \text{cl } X_t(\omega)\} = \text{graph cl } X_\bullet(\omega)$$

Clearly Γ is closed-valued, it is also measurable as follows from [1, Theorem 3.1] since $\Gamma = \text{ls } \Gamma_k'$, where for all $\nu, \Gamma_k(\omega) =$ graph $X_\bullet(\omega)$. Now let $\{(X_t^\nu, t \in T), \nu \in N\}$ be a sequence of stochastic processes, and for all $\nu, \Gamma_\nu : \Omega \rightrightarrows R^p$ the multifunction defined by

$$\Gamma_\nu(\omega) = \text{graph cl } X_\bullet^\nu(\omega) \quad .$$

We can now study the limit process associated to the sequence $\{(X_t^\nu), \nu \in N\}$ through the convergence in distribution of the multifunctions $\{\Gamma_\nu, \nu \in N\}$. One of the advantages of proceeding in this fashion, in preference to the more traditional functional approach, is the automatic existence of a limit distribution as guaranteed by Theorem 5.1 even when the (index) set T is not compact. But also, the fact that the space of closed subsets of $T \times R^q$ is richer than the class of subsets corresponding to the graph of functions with prescribed properties (continuity, ...) might enable us to study processes whose paths do not conveniently fit in a neat class, such as when the paths of the limit process do not have the same properties as those of the processes $\{(X_t^\nu), \nu \in N\}$.

We have used cl $X_\bullet$, the closure of $X_\bullet$, rather than $(X_t(\cdot), t \in T)$ to define the associated multifunction Γ because the theory of convergence in distribution for multifunction has been developed for closed-valued measurable multifunctions and thus for purely technical reasons we need the graph of the paths of the random variables to be closed. It would be possible to work out a theory for multifunctions defined by the graphs of $X_\bullet$, not necessarily closed-valued, however, since limits of sequences of sets are always closed, the multifunctions associated with limit processes will always be closed-valued and thus the associated processes will also be closed. It appears that nothing of substance would be gained by taking such an approach and thus will not be pursued here.

W. Vervaat [24, 25, 26], whose work has progressed independently of ours, has suggested a closely related approach for the convergence of stochastic processes. He was also led to this by his penetrating analysis of extremal processes, although his motivation is not quite of the same nature as ours. Briefly, to each path

$$t \longmapsto X_t(\omega) : T \to \overline{R}$$

of a process, he associates a closed subset $\Gamma'(\omega)$ of $\Gamma \times R^1$ obtained as follows: let $X_\bullet^u(\omega)$ and $X_\bullet^l(\omega)$ be the smallest upper semicontinuous function majoring $X_\bullet(\omega)$ and the largest lower semicontinuous function majorized by $X_\bullet(\omega)$ respectively. Then

$$\Gamma'(\omega) = \{(t,x) \mid X_t^u(\omega) \leq x \leq X_t^l(\omega)\}$$

Convergence of a sequence of processes can thus be studied in the framework of the convergence of the associated multifunctions of type Γ'. Theorem 5.1 again providing the needed compactness of the associated probability measures. His approach works well for $\bar{R}$-valued processes, but it is not clear how to handle in such a framework R^q-valued processes. In fact, he is more immediately concerned with classes of processes that we would handle in the structure provided by normal integrands.

As indicated at the outset of this section we conclude with a derivation of Donsker's Theorem. We do not seek to obtain the results in the fullest generality, since our aim is to illustrate the use of multifunction techniques. Let $\xi_j, j = 1,\ldots$ be a sequence of i.i.d. (independent identically distributed) real-valued random variables with mean 0 and variance σ^2 and define by

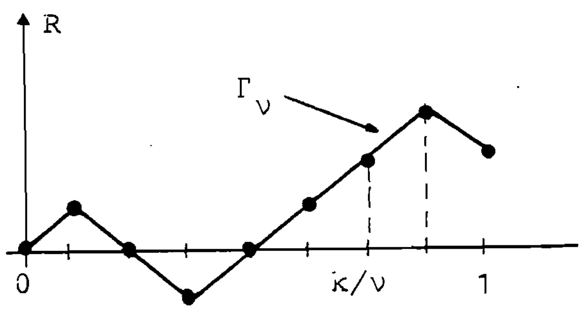
$$S_k(\omega) = \sum_{j=1}^{k} \zeta_j(\omega) \quad , \ k = 1,\ldots \quad ,$$

the associated random walk, with $S_0(\omega) = 0$. We give an argument leading to Brownian motion on bounded intervals. First, let us define for $\nu = 1,\ldots,$ the processes $(X_\nu(t,\cdot), t \in [0,1])$ as follows

$$X_\nu(t,\omega) = (\sigma\sqrt{\nu})^{-1}[S_{[\nu t]} + (\nu t - [\nu t])\xi_{[\nu t]+1}(\omega)] \tag{7.2}$$

where $[\nu t]$ designates the largest integer less than or equal to νt.

Associated to each $X_\nu(\cdot,\omega)$ we have the closed set $\Gamma_\nu(\omega)$ determined by their graphs:



7.3 Figure. $\Gamma_\nu(\cdot)$ associated to $t \longmapsto X_\nu(t,\cdot)$.

Clearly the Γ_ν are closed-valued measurable multifunctions and each one induces a probability measure on the closed subsets of $[0,1] \times R$. Theorem 5.1 guarantees of a weak*-convergent subsequence of the induced probability measures, however, possibly to a measure corresponding to the empty-valued multifunction. Actually, this does not occur as is argued here below.

7.4 PROPOSITION. *Let $K' = K\backslash\{\emptyset\}$ be the space of nonempty compact subsets of* E *and $\mathcal{D}$ a subset of K'. Suppose*

$$D = \cup_{K\in\mathcal{D}} K \subset E$$

is bounded. Then

$$\mathrm{cl}_{T}\mathcal{D} \subset K'$$

PROOF. Suppose $F \in \mathrm{cl}_T\mathcal{D}$, then since every sequence of (nonempty) compact sets converging to F is contained in D it follows that $F \subset \mathrm{cl}\ D$ and hence $F \in K$. Now $F \neq \emptyset$ because any sequence contained in D can never satisfy the criterion for convergence to the empty set given by Lemma 2.1 of [1]. □

There is also a converse to this Proposition but not with cl_T but with T replaced by a topology related to T by finer, namely by T_h the topology on K' generated by the Hausdorff distance. The above statement then becomes: $\mathcal{D}$ is T_h-precompact if and only if D is precompact.

7.5 PROPOSITION. *The probability measures $P_\nu, \nu \in N$ induced by the measurable multifunctions $\Gamma_\nu:\Omega \rightrightarrows [0,1] \times R$ on K' are tight.*

PROOF. The measures $P_\nu, \nu = 1,\ldots$ are tight [3] on K' if to every $\varepsilon > 0$ there corresponds a compact $\mathcal{D}_\varepsilon \in K'$ such that for all ν

$$P_\nu(\mathcal{D}_\varepsilon) > 1 - \varepsilon \quad .$$

In view of the characterization provided by Proposition 7.4 and the definition of P_ν, it suffices that to every $\varepsilon > 0$, there corresponds $\eta > 0$ such that for all ν,

$$\mu\{\varepsilon | \Gamma_\nu(\omega) \subset [0,1] \times [-\eta,\eta]\} \geq 1 - \varepsilon \quad ,$$

since certainly $[0,1] \times [-\eta,\eta]$ is a bounded subset of $[0,1] \times R$. Using the definition (7.2) of Γ_ν we see that this is equivalent to having for all $\varepsilon > 0$ a corresponding η such that

$$\mu\{\omega | \max_{0 \leq k \leq \nu} |S_k(\omega)| \leq \eta\sigma\sqrt{\nu}\} \geq 1 - \varepsilon$$

for all $\nu \in N$. But this inequality follows directly from Kolmogorov's inequality [27, p. 247] if we choose $\eta = 1/\sqrt{\varepsilon}$. □

We have thus shown that the $\{P_\nu, \nu \in N\}$ are relatively weak*-compact on K'. We now need to exhibit the limit to complete the argument. Using the properties of the symmetric random walk ($\sigma = 1$), it appears possible to calculate explicitly the distribution functions $\{T^\nu : K^{ub} \to [0,1], \nu \in N\}$ of the multifunctions $\{\Lambda^\nu : \Omega \rightrightarrows [0,1] \times R', \nu \in N\}$ where

$$\Lambda^\nu(\omega) = \{(k/\nu \ , \ (\sqrt{\nu})^{-1} S_k) \ , \ k = 0,\ldots,\nu\}$$

and then obtain the distribution function $T : K^{ub} \to [0,1]$ of the multifunction corresponding to the graph of the Wiener process as the limit of the sequence of the distribution functions $\{T_\nu, \nu \in N\}$. The carrying out of those calculations are beyond the scope of the article, but it is easy to see what needs to be done. Since the closed balls in $[0,1] \times R$ can be taken to be closed rectangles, the union of finite balls is then the union of a finite number of rectangles. Let $R = [\alpha,\beta] \times [a,b]$ be such a rectangle. To find

$$T_\nu([\alpha,\beta] \times [a,b]) = T_\nu(R)$$

we need to find the probability that the random walk $\{S_k \ , \ k = 0,\ldots,\nu\}$ will pass through the bounds $[a\sqrt{\nu} \ , \ b\sqrt{\nu}]$ when $k \in [\alpha\nu, \beta\nu]$, i.e.

$$T_\nu(R) = \mu\{\text{for some } k \in [\alpha\nu, \beta\nu] \ , \ S_k \in [a\sqrt{\nu}, b\sqrt{\nu}]\} \quad .$$

Thus we need to count the number of (symmetric) random walk paths that pass through given a "window", or more generally through a finite number of such "windows", obtain limiting expressions and show that they yield the formulas for the distribution function T of the Wiener process.

There is an alternative approach which is closer to the standard proof of Donsker's Theorem. It consists in identifying the random walks through interpolation such as done in (7.2), with normal integrands and show that the collection $\{X_\nu(\cdot,\cdot), \nu \in N\}$ is equi-totally lower semicontinuous. This allows us to apply Theorem 4.10 and complete the proof by showing that the finite dimensional distribution of the $\{X_\nu, \nu \in N\}$ converge to those of the Wiener process, such as done in the first part of the proof of Donsker's Theorem [3, Theorem 10.1].

8. CONVERGENCE OF STOCHASTIC INFIMA

In this section, our purpose is to suggest the potential applications of the preceding results to stochastic optimization problems. Here we deal only with the inf-compact case that allows for a self-contained analysis. A more comprehensive treatment is foreseen [28].

Many stochastic optimization problems can be cast in the following (abstract) form:

(8.1) find $x_1 \in E_1$ that minimize $\int u[Q(x,\omega)]\mu(d\omega)$

where u is a scaling, e.g., utility function, and

$$Q(x_1,\omega) = \inf_{x_2 \in E_2} g(x_1,x_2,\omega) \quad ,$$

with $g:(E_1 \times E_2) \times \Omega \to \overline{R}$ is a normal integrand and μ, as before, is a probability measure. Models that fit (8.1) are stochastic programs with recourse, certain classes of Markov decision processes, stochastic control problems in discrete time, statistical estimation problems, and so on. Typically g is defined as follows:

$$(8.2) \qquad g(x_1,x_2,\omega) = \begin{cases} f_0(x_1,x_2,\omega) & \text{if for } i = 1,\dots,m \text{ , } f_i(x_1,x_2,\omega) \le 0 \\ +\infty & \text{otherwise,} \end{cases}$$

where for $i = 0,1,\dots,m$, the functions $f_i:(E_1 \times E_2) \times \Omega \to R$ are continuous in (x_1,x_2) and measurable in x. The proof that g, as well as Q by the way, is again a normal integrand follows from basic properties [8]. This is worked out in detail in [29,30] for stochastic programs with recourse and [31] for stochastic control problems in discrete time, cf. also [32] for stochastic optimization problems of Bolza type.

If $u[\cdot]$ is linear, we may restrict our attention to comparing the expectations of the random variables $\{Q(x_1,\cdot), x_1 \in E_1\}$, but more generally it is the whole distribution of $Q(x_1,\cdot)$ which is of interest. Unless g is very particular, the only possible approach to solving (8.1) is via approximations, in which case we are interested in the convergence in distribution of a sequence of the type

$$\{Q^{\nu}(x_1,\cdot) = \inf_{x_2 \in E_2} g^{\nu}(x_1,x_2,\omega) \ , \ \nu = 1,\ldots\}$$

This and related questions [33, Chapter II], [34], [35] lead to the study of the following question, known as *the distribution problem in stochastic optimization:*

(8.3) *Given* $\{f^{\nu}: E \times \Omega \to \overline{R} \ , \ \nu = 1,\ldots\}$ *a sequence of normal integrands and an associated limit normal integrand* f. *Find minimal conditions that guarantee the convergence in distribution of the random variables* $\{z^{\nu}(\cdot) = \inf_{x \in E} f^{\nu}(x,\cdot) \ , \ \nu = 1,\ldots\}$ *to* $z(\cdot) = \inf_{x \in E} f(x,\cdot)$.

The remainder of this section deals with this question, mainly when the normal integrands are also inf-compact, a restrictive situation that already covers a large number of applications. However, we start with a general result that does not require any additional assumptions about the f^{ν}.

Let $H^{\nu}: \overline{R} \to [0,1], \nu = 1,\ldots\}$ and $H: \overline{R} \to [0,1]$ denote the distribution functions of the random variables $\{z^{\nu}, \nu = 1,\ldots\}$ and z respectively, i.e.

(8.4) $$H(\zeta) = \mu[\omega | z(\omega) < \zeta]$$

and similarly for the $H^{\nu}, \nu = 1,\ldots$. To show the convergence in distribution of the $\{z^{\nu}, \nu = 1,\ldots\}$ to z, we need to show that

(8.5) $$H(\zeta) = \lim_{\nu \to \infty} H^{\nu}(\zeta) \qquad \text{for all } \zeta \in C_H$$

where C_H is the set of continuity point of H. In general every continuity point of the distribution function of the random (measurable) multifunction $\omega \longmapsto \text{epi } f(\cdot,\omega)$ is a continuity point of H but not conversely. This is at the crux of the difficulties. We always have the following:

8.6 THEOREM. *Suppose* $\{f^{\nu}: E \times \Omega \to \bar{R}, \nu = 1, \ldots\}$ *is a sequence of normal integrands epi-converging in distribution to (the normal integrand)* $f: E \times \Omega \to \bar{R}$. *Then, for all* $\zeta \in R$

$$(8.7) \qquad H(\zeta) \leq \liminf_{\nu \to \infty} H^{\nu}(\zeta)$$

PROOF. As in the proof of Theorem 3.13, let $\{Q^{\nu}, \nu = 1, \ldots\}$ and Q be the probability measures induced by the normal integrands on $(SC(E), S_E)$. We have that

$$H(\zeta) = \mu[\omega \in \Omega | \text{epi } f(\cdot,\omega) \cap G_\zeta \neq \emptyset] = Q(E_\zeta)$$

and

$$H^{\nu}(\zeta) = \mu[\omega \in \Omega | \text{epi } f^{\nu}(\cdot,\omega) \cap G_\zeta \neq \emptyset] = Q^{\nu}(E_\zeta)$$

where

$$G_\zeta = \{(x,\alpha) | \alpha < \zeta\}$$

and thus $E_\zeta = \{g \in SC(E) | \inf g < \zeta\}$ is an open subset of SC(E). Since the epi-convergence in distribution of the f^{ν} to f is the weak*-convergence of the Q^{ν} to Q, it follows that for every open subset O of SC(E), we have [3, Theorem 2.1]

$$\liminf Q^{\nu}(O) \geq Q(O).$$

The inequality (8.7) is obtained with $O = E_\zeta$. □

Unfortunately, the inequality

$$\limsup_{\nu \to \infty} H^{\nu}(\zeta) \leq H(\zeta)$$

does not hold for every $\zeta \in R$. To obtain this inequality we need to impose a certain type of compactness conditions, sufficient conditions are provided by Theorems 8.8 and 8.11.

8.8 THEOREM. *Suppose* $\{f^\nu : E \times \Omega \to \overline{R}\ ,\ \nu = 1,\ldots\}$ *is a sequence of normal integrands epi-converging in distribution to the normal integrand* $f : E \times \Omega \to \overline{R}$. *Suppose moreover that for all* $\zeta \in R$ *and* $\varepsilon > 0$ *there exist compact sets* K *and* K' *such that for all* $\nu \in N$

$$(8.9)\qquad \mu\{\omega | \inf f^\nu(\cdot,\omega) < \zeta\} \leq \varepsilon + \mu\{\omega | \inf_K f^\nu(\cdot,\omega) \leq \zeta + \varepsilon\}\quad ,$$

and

$$(8.10)\qquad \mu\{\omega | \inf_{K'} f(\cdot,\omega) \leq \zeta\} \leq \varepsilon + \mu\{\omega | \inf f(\cdot,\omega) \leq \zeta\}\quad .$$

Then $\{z^\nu, \nu = 1,\ldots\}$ *converge in distribution to* z.

PROOF. Hypothesis (8.9) implies that for all $\varepsilon > 0$

$$\limsup_{x\to\infty} H^\nu(\zeta) \leq \limsup_{\nu\to\infty} \mu\{\omega | \inf_K f^\nu(\cdot,\omega) \leq \zeta + \varepsilon\} + \varepsilon$$

$$= \varepsilon + \limsup_{\nu\to\infty} Q^\nu(E_{K,\zeta+\varepsilon})$$

where

$$E_{K,\zeta+\varepsilon} = \{g \in SC(E) | \text{epi } g \cap (K \times [\zeta + \varepsilon, -\infty)) \neq \emptyset\}$$

is a closed subset of $SC(E)$, as follows from the definition of the epi-topology. Again $\{Q^\nu, \nu = 1,\ldots\}$ and Q will denote the probability measures induced on $SC(E)$ by the normal integrands $\{f^\nu, \nu = 1,\ldots\}$ and f respectively. The convergence in distribution implies [3, Theorem 2.1] that

$$\limsup_{\nu\to\infty} Q^\nu(E_{K,\zeta+\varepsilon}) \leq Q(E_{K,\zeta+\varepsilon})$$

from which it follows that for all $\varepsilon > 0$

$$\limsup_{\nu\to\infty} H^\nu(\zeta) \leq \varepsilon + \mu\{\omega | \inf_K f(\cdot,\omega) \leq \zeta + \varepsilon\}$$

Nothing is lost if K is enlarged to contain the set K' used to obtain (8.10), thus we have

$$\limsup_{\nu\to\infty} H^\nu(\zeta) \leq 2\varepsilon + \mu\{\omega | \inf f(\cdot,\omega) \leq \zeta\}\quad .$$

At every continuity point $\zeta \in C_H$ we have that

$$\mu\{\omega | \inf f(\cdot,\omega) = z(\omega) \leq \zeta\} = H(\zeta) \quad .$$

This with what precedes, implies that for all $\zeta \in C_H$

$$\limsup_{\nu\to\infty} H^{\nu}(\zeta) \leq H(\zeta) \quad .$$

Combining this with the lim inf inequality (8.7) gives us the convergence in distribution of the random variables $\{z^{\nu}, \nu = 1,\ldots\}$ to z. □

It might not always be easy to verify the hypotheses of Theorem 8.8, a more approachable set of conditions is given by the next theorem.

A normal integrand $f:E \times \Omega \to \overline{R}$ will be called *inf-compact* if for every $\omega \in \Omega$, the function $x \longmapsto f(x,\omega)$ is inf-compact, i.e., for all $\alpha \in R$ the level sets

$$\text{lev}_{\alpha}\, f(\cdot,\omega) = \{x | f(x,\omega) \leq \alpha \}$$

are compact. For example, if g is defined by (8.2) it is inf-compact if the function $(x_1,x_2) \longmapsto f_0(x_1,x_2,\omega)$ is inf-compact or if the set

$$\{(x_1,x_2) | f_i(x_1,x_2,\omega) \leq 0 \quad , \quad i = 1,\ldots,m\} = S(\omega)$$

is compact, or the function $f_0(\cdot,\cdot,\omega)$ tends to ∞ on every unbounded arc contained in $S(\omega)$. A sequence $\{f^{\nu}:E \times \Omega \to \overline{R}, \nu \in N\}$ is *equi-inf-compact* if they are inf-compact and for every $\omega \in \Omega$, the sets $\text{lev}_{\alpha}\, f^{\nu}(\cdot,\omega)$ are *equi-bounded* for all $\alpha \in R$, i.e., to all $\omega \in \Omega$ and $\alpha \in R$ there corresponds $D \subset E$ such that $D \supset \text{lev}_{\alpha}\, f^{\nu}(\cdot,\omega)$ for all $\nu \in \mathbb{N}$.

8.11 THEOREM. *Suppose* $\{f^{\nu}:E \times \Omega \to \overline{R}\ , \nu = 1,\ldots\}$ *is a sequence of equi-inf-compact normal integrands that almost surely epi-converge to the normal integrand* $f:E \times \Omega \to R$. *Then the sequence of random variables* $\{z^{\nu}, \nu = 1,\ldots\}$ *converges almost surely to* z, *in particular this implies that on* C_H

$$\lim_{\nu\to\infty} H^\nu(\zeta) = H(\zeta)$$

PROOF. Ignoring a set of measure 0, we need to prove that for ever ω,

$$\lim_{\nu\to\infty} (\inf f_\nu(\cdot,\omega)) = \inf f(\cdot,\omega)$$

We are thus essentially in the deterministic case and can rely on [36, Proposition 12] to obtain the preceding equality for sequences of equi-inf-compact functions $\{f^\nu(\cdot,\omega), \nu = 1,\ldots\}$. The last assertion simply follows from the fact that a.s. convergence implies convergence in distribution. □

One application of this theorem is to random linear programs [33], [34], more detailed applications appear in [28].

8.12 PROPOSITION. *Suppose* $\{A^\nu, b^\nu, c^\nu\ , \nu = 1,\ldots\}$ *is a sequence of random matrices* (m x n) *and vectors* (m x 1) *and* (1 x n) *respectively, such that*

(i) $A^- \leq A^\nu(\cdot) \leq A^+$ *and* $b^- \leq b^\nu(\cdot) \leq b^+$ *for all* ν

(ii) *the interior of* $\{x \geq 0 | A^- x \geq b^+\}$ *is nonempty and a.s.* $(A_i(\ ), b_i(\ )) \neq 0$ *for* $i = 1,\ldots,m$,

(iii) $\{x \geq 0 | A^+ x \geq 0\} = \{0\}$,

and almost surely

$$\lim_{\nu\to\infty} A^\nu(\cdot) = A(\cdot)\ , \lim_{\nu\to\infty} b^\nu(\cdot) = b(\cdot)\ , \lim_{\nu\to\infty} c^\nu(\cdot) = c(\cdot)\ .$$

Then the

$$z^\nu(\cdot) = \inf_x [c^\nu(\cdot)x | A^\nu(\cdot)x \geq b^\nu(\cdot)\ , x \geq 0] ,$$

converge almost surely to

$$z(\cdot) = \inf_x \, [c(\cdot)x \; A(\cdot)x \geq b(\cdot) \; , \; x \geq 0] \quad .$$

PROOF. Clearly for all ν, and for all ω

$$\{x \geq 0 | A^\nu(\omega)x \geq b^\nu(\omega)\} \subset \{x \geq 0 | A^+x \geq b^-\}$$

and this later set is compact as follows from conditions (iii). For $\nu = 1,\ldots,$ let

$$f^\nu(x,\omega) = \begin{cases} c^\nu(\omega)x & \text{if} \quad A^\nu(\omega)x \geq b^\nu(\omega) \; , \; x \geq 0 \quad , \\ +\infty & \text{otherwise,} \end{cases}$$

and

$$f(x,\omega) = \begin{cases} c(\omega)x & \text{if} \quad A(\omega)x \geq b(\omega) \; , \; x \geq 0 \quad , \\ +\infty & \text{otherwise.} \end{cases}$$

It is easy to verify that these functions are normal integrands and in view of the above the sequence $\{f^\nu : E \times \Omega \to R \cup \{+\infty\}, \nu = 1,\ldots\}$ is equi-inf-compact. The almost sure epi-convergence of the $f_\nu(\cdot,\omega)$ to $f(\cdot,\omega)$ follows from condition (ii). To see this simply observe that (i) and (ii) imply that int $\{x \geq 0 | A(\omega)x \geq b(\omega)\}$ is nonempty for almost all ω. Thus to prove epi-convergence it suffices to prove pointwise convergence of the f^ν to f as follows from [37, Corollary 2A]. In turn, this follows from the almost sure convergence of the parameters of the random linear program and the fact that for almost all ω, no row of $(A(\omega), b(\omega))$ is identically 0. □

APPENDIX: PROOF OF THEOREM 1.3

This new proof of Choquet's Theorem helps clarify its relationship to the classical Correspondence theorem which shows that there is a natural bijection between probability measures and distribution functions defined on R^1 (or more generally R^n). It shows that multifunctions can and should be viewed as "thick" functions and all that is required is an appropriate adaptation of the definitions.

We start by showing that every probability measure P on S determines a distribution T on K with $T(K) = P(F_K)$ for all $K \in K$. Clearly $T(\emptyset) = 0$ and $T(K) \in [0,1]$, thus we need to show that T satisfies conditions (1.1) and (1.2).

Let $\{K_\nu \in K, \nu \in N\}$ be a sequence that decreases monotonically to $K \in K$. If K is empty, then the K_ν must be empty for all but finitely many ν; this follows from [1, Lemma 2.1] and the fact that $\{K_\nu, \nu \in N\}$ is a decreasing sequence. The corresponding sets F_{K_ν} are themselves empty and thus obviously

$$0 = \lim_\nu T(K_\nu) = \lim_\nu P(F_{K_\nu}) = P(F_\emptyset) = T(\emptyset) = 0 \quad .$$

If $K \neq \emptyset$, then

$$F_K = \bigcap_{\nu=1}^{\infty} F_{K_\nu} \quad .$$

To see this, first note that $K = \cap K_\nu$, because the K_ν are decreasing to K, cf. [6, Proposition 2] and thus $F_K = F_{\cap K_\nu} \subset \cap F_{K_\nu}$. On the other hand if $F \in \cap F_{K_\nu}$, then for all $\nu \in N$, the elements of the decreasing sequence $\{F \cap K_\nu, \nu \in N\}$ of compact sets are nonempty. From this it follows that $F \in F_K$ i.e. $F \cap K \neq \emptyset$, since $F \cap K = \emptyset$ would imply that $F \cap K_\nu = \emptyset$ for all ν sufficiently large [1, Lemma 2.1] in view of the fact that $F \cap K \supset \cap_\nu (F \cap K_\nu)$. And thus

$$\lim_\nu T(K_\nu) = \lim_\nu P(F_{K_\nu}) = P \bigcap_{\nu=1}^{\infty} F_{K_\nu} = P(F_K) = T(K) \quad .$$

As far as (1.2) is concerned we have immediately that

$$S_0(K_0) = 1 - T(K_0) = 1 - P(F_{K_0}) = P(F^{K^0}) \quad ,$$

and recursively, for $\nu = 1,\ldots$

$$\begin{aligned} S_\nu(K_0;K_1,\ldots,K_\nu) &= S_{\nu-1}(K_0;K_1,\ldots,K_{\nu-1}) - S_{\nu-1}(K_0 \cup K_\nu;K_1,\ldots,K_{\nu-1}) \\ &= P\, F^{K_0}_{K_1,\ldots,K_{\nu-1}} - P\, F^{K_0 \cup K_\nu}_{K_1,\ldots,K_{\nu-1}} \\ &= P\, F^{K_0}_{K_1,\ldots,K_\nu} = P(F^{K_0} \cap F_{K_1} \cap \ldots \cap F_{K_\nu}) \quad , \end{aligned}$$

because for any collection of sets $\{A_0,\ldots,A_\nu\}$ we have that

$$F^{A_0}_{A_1,\ldots,A_\nu} = F^{A_0}_{A_1,\ldots,A_{\nu-1}} \setminus F^{A_0 \cup A_\nu}_{A_1,\ldots,A_{\nu-1}} \quad .$$

Clearly for all $\nu = 1,\ldots, 0 \le S_\nu(K_0;K_1,\ldots,K_\nu) \le 1$.

We now prove the converse, namely that a distribution function on K determines through the correspondence: for $\nu = 1,\ldots$

$$(1.4) \qquad P\, F^{K_0}_{K_1,\ldots,K_\nu} = S_\nu(K_0;K_1,\ldots,K_\nu)$$

(or equivalently $P(F_K) = T(K)$), a probability measure on S. Let S_σ be the algebra (field) consisting of the finite unions of pairwise disjoint sets of the type $T' = F^{K_0}_{K_1,\ldots,K_\nu}$ for some collection $K_0,\ldots,K_\nu$ in K. On S_σ, P is defined by the relation

$$P\left(\bigcup_{i=1}^{h} T_i\right) = \sum_{i=1}^{h} P(T_i)$$

where the sets T_i are sets whose measure have been defined by (1.4) above. P is a nonnegative finitely additive measure on S_σ bounded above by 1. The proof can be found in [4, p.33].

In face P is countably additive on S_σ. To prove this, it is sufficient to show that given any sequence of sets $\{\mathcal{D}_\nu \in S_\sigma, \nu \in N\}$ decreasing to $\emptyset$ - i.e. such that

$$\bigcup_{\nu=1}^{\infty} \bigcap_{\mu=\nu}^{\infty} \mathcal{D}_\mu = \bigcap_{\nu=1}^{\infty} \bigcup_{\mu=\nu}^{\infty} \mathcal{D}_\mu = \emptyset$$

with $\mathcal{D}_\nu \supset \mathcal{D}_{\nu+1}$ for $\nu = 1,\ldots,$ we have that

$$\lim_\nu P(\mathcal{D}_\nu) = 0 \quad ,$$

cf. for example [38, Theorem 9.F]. This will follow, if we can exhibit for every $\varepsilon > 0$ a collection of sets $\{\mathcal{D}'_\nu \in S_\sigma, \nu \in N\}$ such that for every $\nu \in N$,

(A.1) $$\mathcal{D}'_\nu \subset \text{cl } \mathcal{D}'_\nu \subset \mathcal{D}_\nu \quad ,$$

and

(A.2) $$P(\mathcal{D}_\nu) < P(\mathcal{D}'_\nu) + \varepsilon \cdot 2^{-\nu}$$

The sets cl $\mathcal{D}'_\nu$ are compact, F being compact. Also since $\cap_{\nu=1}^{\infty} \mathcal{D}_\nu = \emptyset$ and (A.1) holds then also

$$\bigcap_{\nu=1}^{\infty} \mathcal{D}'_\nu = \emptyset$$

from which it follows, by the finite intersection property, that for $\bar{\nu}$ sufficiently large $\cap_{\nu=1}^{\bar{\nu}}$ cl $\mathcal{D}'_\nu = \emptyset$. In view of (A.1) we also have that

$$\bigcap_{\nu=1}^{\bar{\nu}} \mathcal{D}'_\nu = \emptyset \quad .$$

Since the $\mathcal{D}_\nu$ are monotonically decreasing, for every $\mu \geq \bar{\nu}$, we have that

$$\mathcal{D}_\mu = \bigcap_{\nu\leq\mu} \mathcal{D}_\nu \subset \bigcup_{\nu\leq\mu} (\mathcal{D}_\nu \backslash \mathcal{D}'_\nu)$$

where $\mathcal{D}_\nu \backslash \mathcal{D}'_\nu \in S_\sigma$. Thus

$$P(\mathcal{D}_\mu) \leq \sum_{n=1}^{m} P(\mathcal{D}_\nu \backslash \mathcal{D}'_\nu) \quad .$$

By (A.2) this implies that lim $P(\mathcal{D}_\nu) = 0$.

To complete the proof it thus suffices to produce the sets $\{\mathcal{D}'_\nu \in \mathcal{S}_\sigma, \nu \in \mathbb{N}\}$ that satisfy (A.1) and (A.2). Recall that $\mathcal{S}_\sigma$ is the field consisting of the finite unions of pairwise disjoint sets of the type $F^{K_0}_{K_1,\ldots,K_h}$ and thus in particular the $\mathcal{D}_\nu$ are unions of such sets. Hence it will be sufficient to show that each set of the type $F^{K_0}_{K_1,\ldots,K_h}$ - used to construct $\mathcal{D}_\nu$ - admits an approximation that satisfies (A.1) and (A.2). Given any compact K_0 there always exists a sequence of open relatively compact sets $\{G'_\sigma, \sigma \in \mathbb{N}\}$ decreasing strictly to K_0 such that

$$K'_{\sigma+1} \subset G'_\sigma \subset K'_\sigma$$

with $K'_\sigma = \mathrm{cl}\, G'_\sigma$. Hence

$$\text{(A.3)} \qquad F^{K'_\sigma}_{K_1,\ldots,K_h} \subset F^{G'_\sigma}_{K_1,\ldots,K_h} \subset F^{K'_{\sigma+1}}_{K_1,\ldots,K_h} \quad .$$

The sets

$$F^{G'_\sigma}_{K_1,\ldots,K_h}$$

are closed and thus also compact, since F is compact. Since (1.1) implies that $P(F^{K'_\sigma}) \uparrow P(F^{K_0})$, it follows that

$$\lim P\left(F^{K'_\sigma}_{K_1,\ldots,K_h}\right) = P\left(F^{K_0}_{K_1,\ldots,K_h}\right) \quad .$$

Thus given any $\varepsilon > 0$ we can find σ sufficiently large so that

$$P\left(F^{K_0}_{K_1,\ldots,K_h}\right) < P\left(F^{K'_\sigma}_{K_1,\ldots,K_h}\right) + \varepsilon \cdot 2^{-\nu} \quad .$$

This combined with (A.3) yields (A.1) and (A.2), and thus completes the proof. □

REFERENCES

[1] G. Salinetti and R. Wets, On the convergence of closed-valued measurable multifunctions, *Trans.Amer.Math.Soc.*, 266(1981), 275-289.

[2] G. Choquet, Convergences, *Annales de l'Univ. de Grenoble*, 23(1947-1948), 55-112.

[3] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, New York, 1968.

[4] G. Matheron, *Random Sets and Integral Geometry*, John Wiley & Sons, New York, 1975.

[5] K. Parthasaraty, *Probability Measures on Metric Spaces*, Academic Press, New York, 1967.

[6] G. Salinetti and R. Wets, On the convergence of sequences of convex sets in finite dimension, *SIAM Review*, 21(1979), 18-33.

[7] R.T. Rockafellar, Integrals which are convex functionals, I & II. *Pacific J. Math.* 24(1968), 525-539 and 39(1971), 439-469.

[8] R.T. Rockafellar, Integral functionals, normal integrands and measurable selections, in *Nonlinear Operators and Calculus of Variations*, Lecture Notes in Math. 543, Springer Verlag, Berlin 1977.

[9] E. De Giorgi, Convergence problems for functionals and operators, in *Recent Methods in Nonlinear Analysis*, eds. E. De Giorgi, E. Magenes and U. Mosco, Pittagora Editrice, Bologna, 1979, 131-188.

[10] H. Attouch, Sur la Γ-convergence, Seminaire Brezis-Lions, College de France (February 1979).

[11] R. Wijsman, Convergence of sequences of convex sets, cones and functions, II, *Trans.Amer.Math.Soc.*, 123(1966), 32-45.

[12] U. Mosco, Convergence of convex sets and of solutions of variational inequalities, *Advance Math.* 3(1969), 273-299.

[13] H. Attouch, *Variational Convergence for Functions and Operators*, Research Notes in Mathematics, Pittman, London (forthcoming).

[14] R. Wets, Convergence of convex functions, variational inequalities and convex optimization problems, in *Variational Inequalities and Complementarity Problems*, eds. R. Cottle, F. Giannessi and J.-L. Lions, Wiley & Sons, Chichester (UK), 1980, 375-403.

[15] E. De Giorgi and T. Franzoni, Su un tipo di convergenza variazionale, *Atti.Acc.Naz.Licei* (8), 58(1975), 842-850. See also: Fascicolo delle demostrazioni, Scuola Normale Superiore di Pisa, May 1976.

[16] G. Buttazzo, su una definizione generale dei Γ-limiti, *Bolletino U.M.I.*, 14-B(1977), 722-744.

[17] S. Dolecki, G. Salinetti and R. Wets, Convergence of functions: equi-semicontinuity, *Trans.Amer.Math.Soc.* (1983), forthcoming.

[18] R. Wets, On a compactness theorem for epi-convergent sequences of functions, in *Proceed. of Rio de Janeiro Congress on Mathematical Programming*, ed. M. Kelmanson, North-Holland, 1983 (forthcoming).

[19] W. Whitt, Continuity of Markov processes and dynamic programs, Tech. Report, Yale University, 1975.

[20] H.-J. Langen, Convergence of dynamic programming models, *Mathematics of Operations Research* 6(1981), 493-512.

[21] H. Engl and A. Wakolbinger, On weak limits of probability distributions on Polish spaces, Tech. Report 218, Universitaet Linz, 1982.

[22] Y. Prohorov, Convergence of random processes and limit theorems in probability theory, *Theor. Probability Applic.* 1(1956), 157-214.

[23] G. Salinetti, On the convergence of multifunctions and stochastic processes, European Congress of Statisticians, Palermo, September 1982.

[24] W, Vervaat, Une compactification des espaces fonctionnels C et D; une alternative pour la démonstration des theorèmes limites fonctionnels, *C.R. Acad. Sc. Paris* 292(1981), 441-444.

[25] W. Vervaat, The structure of limit theorems in probability, Lecture Notes, Spring 1981, Nijmegen and Leuven.

[26] W. Vervaat, Random upper semicontinuous functions and extremal processes, Manuscript, University of Nijmegen, 1982.

[27] M. Loève, *Probability Theory I* (Fourth edition), Springer Verlag, New York 1977.

[28] G. Salinetti and R. Wets, On the convergence of infima, especially stochastic infima, in preparation.

[29] R. Wets, Duality relations in stochastic programming, in *Symposia Mathematica XIX*, Academic Press, New York 1976, 341-355.

[30] P. Olsen, Multistage stochastic programming with recourse: the equivalent deterministic problem, *SIAM J. Control Optim.*, 14(1976), 495-517.

[31] R.T. Rockafellar and R. Wets, The optimal recourse problem in discrete time: L^1-multipliers for inequality constraints, *SIAM J. Control Optim.*, 16(1978), 16-36.

[32] R.T. Rockafellar and R. Wets, Deterministic and stochastic optimization problems of Bolza type in discrete time, *Stochastics* (1983).

[33] P. Kall, *Stochastic Linear Programming*, Springer Verlag, Berlin 1975.

[34] G. Salinetti, Convergence for measurable multifunctions: an application to stochastic optimization problems, in *Mathematical Programming and its Economic Applications*, eds. G. Castellani and P. Mazzoleni, Angeli, Milano 1981.

[35] A. Prékopa, Large scale stochastic programming problems, Lecture at the XI. International Symposium on Mathematical Programming, Bonn 1982.

[36] R. Wets, A formula for the level sets of epi-limits and some applications, in *Mathematical Theory of Optimization*, eds. P. Cecconi and T. Zolezzi, Springer Verlag Lecture Notes in Mathematics, 1982 (forthcoming).

[37] G. Salinetti and R. Wets, On the relation between two types of convergence for convex functions, *J. Math.Anal.Appl.* 60(1977), 211-226.

[38] P. Halmos, *Measure Theory*, Van Norstand, New York, 1950.