ADVANCES IN MATHEMATICS 12, 171–177 (1974)

On Pointwise Convergence, Compactness, and Equicontinuity. II

A. IONESCU TULCEA*

Department of Mathematics, Northwestern University, Evanston, Illinois 60201

This note is a sequel and a completion to [5]. Its purpose is to underline the important role played by the "separation property" ($h_1 \in H$, $h_2 \in H$, $h_1 \neq h_2 \Rightarrow \tilde{h}_1 \neq \tilde{h}_2$) in measurability questions (such as weak versus strong). The separation property is characteristic of liftings (see [4] for the definition of a lifting).

The basic notation and terminology used below is as follows:

We denote by $(E, \mathscr{E}, \mu)$ the underlying probability space and by $\mathscr{L} = \mathscr{L}(E, \mathscr{E}, \mu)$ the algebra of all $\mathscr{E}$-measurable mappings $f: E \to R$. For $f \in \mathscr{L}$, $g \in \mathscr{L}$ we write

$$f = g \qquad \text{if} \quad f(t) = g(t) \quad \text{for all } t \in E$$

and

$$f \equiv g \ (\mu) \qquad \text{if} \quad f(t) = g(t) \quad \mu\text{-almost surely.}$$

The latter defines the usual equivalence relation in $\mathscr{L}$. We denote by $\tilde{f}$ the equivalence class of each $f \in \mathscr{L}$ with respect to this equivalence relation.

We say that a set F *carries* μ if $F \in \mathscr{E}$ and $\mu(E - F) = 0$.

We begin with the following result, which in a certain sense generalizes Proposition 1 of [5]:

THEOREM 1. *Let $H \subset \mathscr{L}$ be a set with the following properties.*

(a) *The relations $h_1 \in H$, $h_2 \in H$, $h_1 \neq h_2$ imply $\tilde{h}_1 \neq \tilde{h}_2$.*

(b) *H is convex.*

* Research supported by the U. S. Army Research Office (Durham) under contract No. DAHCO 68 C005.

607/12/2-3

Then the following assertions about H are equivalent:

(i) *H is sequentially compact for the topology of pointwise convergence*[1] *on E.*

(ii) *H is compact metrizable for the topology of pointwise convergence on E.*

(iii) *H is compact for the topology of pointwise convergence on E.*

Proof. We consider on H the topology $\mathcal{T}$ of pointwise convergence on E and the topology $\mathcal{T}^{(\mu)}$ of convergence in μ-probability.[2]

Since (ii) $\Rightarrow$ (iii) obviously, it remains to prove (i) $\Rightarrow$ (ii) and (iii) $\Rightarrow$ (i).

Let e be the identity mapping of H into H. We begin by showing that:

(*) Under *either one* of the assumptions (i) or (iii),

$$e\colon (H, \mathcal{T}^{(\mu)}) \to (H, \mathcal{T}) \text{ is continuous.}$$

Let $f \in H$ and $(f_\alpha)_{\alpha \in N}$ a sequence of elements of H such that

$$(f_\alpha)_{\alpha \in N} \xrightarrow{\mathcal{T}^{(\mu)}} f.$$

Suppose that $(f_\alpha)_{\alpha \in N}$ does *not* converge pointwise to f. There is then some $t_0 \in E$ and an $\epsilon_0 > 0$ such that for *any* $\alpha \in N$, we can find an $\alpha' \geqslant \alpha$ with

$$|f_{\alpha'}(t_0) - f(t_0)| > \epsilon_0 .$$

Now for each positive integer k, choose $\alpha(k) \in N$ such that

$$\beta \geqslant \alpha(k) \Rightarrow d(f_\beta, f) \leqslant 1/2^k.$$

Then for $\alpha'(k) \geqslant \alpha(k)$ (chosen as above) we have

$$\begin{gathered} |f_{\alpha'(k)}(t_0) - f(t_0)| > \epsilon_0 \\ d(f_{\alpha'(k)}, f) = \int \frac{|f_{\alpha'(k)} - f|}{1 + |f_{\alpha'(k)} - f|}\, d\mu \leqslant \frac{1}{2^k}. \end{gathered} \tag{1}$$

[1] That is, for any sequence (h_n) of elements of H, there is a subsequence (h_{n_k}) and an $h \in H$ such that $h_{n_k} \to h$ pointwise on E.

[2] $\mathcal{T}^{(\mu)}$ is clearly Hausdorff, metrizable being given by the metric

$$d(h, g) = \int \frac{|h - g|}{1 + |h - g|}\, d\mu$$

for $g \in H$, $h \in H$.

It is clear that the sequence $(f_{\alpha'(k)})_k$ converges to f, a.s. Under *either one* of the assumptions (i) or (iii), there is an element $g \in H$ which is a *cluster value* of the sequence $(f_{\alpha'(k)})_k$ for the topology of pointwise convergence on E. At a point $t \in E$ where $(f_{\alpha'(k)}(t))_k$ converges to a limit, we must have

$$g(t) = \lim_k f_{\alpha'(k)}(t).$$

Hence $g \equiv f(\mu)$, and using hypothesis (a), $g = f$. But at t_0 we have (by the first inequalities in relations (1)):

$$|f_{\alpha'(k)}(t_0) - g(t_0)| > \epsilon_0 \qquad \text{for } all\ k.$$

This contradicts the fact that $g(t_0)$ is a cluster value for the sequence $(f_{\alpha'(k)}(t_0))_k$. Hence the continuity assertion (*) is proved. From (*) we easily deduce that:

(**) Under *either one* of the assumptions (i) or (iii), for a sequence (u_n) of elements of H and u an element of H, the following assertions are equivalent.

(j) $u_n \to u$ μ-almost surely.

(jj) $u_n \to u$ in μ-probability.

(jjj) $u_n \to u$ pointwise on E.

(i) $\Rightarrow$ (ii). To prove this implication it is enough to note that under the assumption (i), $(H, \mathscr{T}^{(\mu)})$ is compact metrizable, that $\mathscr{T}$ is Hausdorff and weaker than $\mathscr{T}^{(\mu)}$ (by (*)), whence $\mathscr{T} = \mathscr{T}^{(\mu)}$.

(iii) $\Rightarrow$ (i). We divide the proof of this implication into several steps:

(I) For each $t \in E$ denote by ϵ_t the mapping $h \to h(t)$ of the compact space $(H, \mathscr{T})$ into R. It is clear that $\epsilon_t \in C_R(H)$ and hence that $\epsilon_t(H) = \{h(t) \mid h \in H\}$ is a *compact* subset of R. It follows that if (u_n) is an arbitrary sequence of elements of H, then $\sup_n |u_n(t)| < \infty$ for each $t \in E$, and thus

$$\sup_n |u_n| \in \mathscr{L}.$$

(II) Let (h_n) be an arbitrary sequence of elements of H. Let $h^* = \sup_n |h_n|$ and define the measure ν on $(E, \mathscr{E})$ by

$$d\nu = (1/1 + h^*)\, d\mu.$$

Then ν is a *finite* measure on $(E, \mathscr{E})$, ν is *equivalent* with μ (that is ν and μ admit the same sets of measure zero), and

$$\int h^* \, d\nu = \int \frac{h^*}{1 + h^*} \, d\mu < \infty,$$

that is $h^* \in \mathscr{L}^1(E, \mathscr{E}, \nu)$.

From (**) above we then deduce:

(***) For a sequence (u_n) of elements of H and u an element of H we have:

$$u_n \to u \ \nu\text{-almost surely} \Leftrightarrow u_n \to u \text{ pointwise on } E.$$

(III) We next recall a remarkable theorem due to Komlós (see [6]; see also [2]) of which we shall make use below.

Let $(E, \mathscr{E}, \nu)$ be a finite measure space.

THEOREM (Komlós). *Let (f_n) be a sequence of elements of $\mathscr{L}^1(E, \mathscr{E}, \nu)$ with $\sup_n \|f_n\|_1 < \infty$. Then one can find a subsequence $(f_{n_k})_k$ and an element $f \in \mathscr{L}^1(E, \mathscr{E}, \nu)$ such that $(f_{n_k})_k$, as well as any further subsequence extracted from $(f_{n_k})_k$, converges Cesaro to f, ν-almost surely.*

(IV) We finally show that H is sequentially compact for the topology of pointwise convergence on E.

Let (h_n) be a sequence of elements of H. As in part (II) of the proof let $h^* = \sup_n |h_n|$ and let ν be the measure on $(E, \mathscr{E})$ with density $1/(1+h^*)$ with respect to μ. Consider the *measure space* $(E, \mathscr{E}, \nu)$: The sequence (h_n), as a sequence of elements of $\mathscr{L}^1(E, \mathscr{E}, \nu)$, satisfies the hypothesis of Komlós' Theorem. We can then extract a subsequence satisfying the conclusion of Komlós' Theorem. To simplify the notation we shall assume that the sequence (h_n) itself satisfies the conclusion of Komlós' Theorem, i.e., *there is* $h \in \mathscr{L}^1(E, \mathscr{E}, \nu)$ such that for *any* subsequence (h_{n_p}) extracted from (h_n) we have:

$$\lim_p((h_{n_1} + h_{n_2} + \cdots + h_{n_p})/p) = h \ \nu\text{-almost surely.} \tag{2}$$

For each n define

$$g_n = \frac{h_1 + h_2 + \cdots + h_n}{n}.$$

Then (g_n) is a sequence of elements of H (use hypothesis (b)) and

$$\lim_n g_n(t) = h(t), \qquad \nu\text{-almost surely.} \tag{3}$$

Let now $g \in H$ be a *cluster value* of the sequence $(g_n)_n$ for the topology of pointwise convergence on E (use assumption (iii)). Then clearly (3) implies that

$$g(t) = h(t), \qquad \nu\text{-almost surely.} \tag{4}$$

Since H is convex, we deduce from (2), (4), and (***) in part (II) of the proof, that for *any* subsequence (h_{n_p}) extracted from (h_n)

$$(h_{n_1} + h_{n_2} + \cdots + h_{n_p})/p \to g \text{ pointwise on } E.$$

This of course means that the sequence (h_n) itself converges to g pointwise on E and hence the proof of Theorem 1 is complete.

Remark 1. Let $H \subset \mathscr{L}$ be a set with the properties (a) and (b) of Theorem 1 and satisfying one (and hence all) of the equivalent conditions (i), (ii), and (iii) of Theorem 1. Let $B \subset H$. Let u be the *upper envelope* of B and v the *lower envelope* of B (the mappings u and v are defined by the equations

$$u(t) = \sup_{h\in B} h(t), \qquad v(t) = \inf_{h\in B} h(t), \qquad \text{for} \quad t \in E).$$

Then $u\colon E \to R$, $v\colon E \to R$ and u, v are *$\mathscr{E}$-measurable*, i.e., $u \in \mathscr{L}$ and $v \in \mathscr{L}$.

In fact, it is enough to remark that B, as a subspace of the compact metric space H, is separable, to consider an at most countable set $B_0 \subset B$ *dense* in B for the topology of pointwise convergence on E and to note that $u = \sup_{h\in B_0} h$, $v = \inf_{h\in B_0} h$.

Remark 2. In Theorem 1 above, hypothesis (b) was used only in the proof of the implication (iii) $\Rightarrow$ (i). The equivalence (i) $\Leftrightarrow$ (ii) holds without assuming hypothesis (b).

Using Theorem 1 above in conjection with the beautiful generalization of Egorov's theorem due to P. A. Meyer (see [7, p. 199, Proposition 2]), and the classical criterion of relative compactness for a bounded set in $B(S, \Sigma)$ (see [3, p. 260, Theorem 6]) we obtain the following result—much in the same way that we derived Theorem 1 in [5]:

THEOREM 2. *Let $H \subset \mathscr{L}$ be a set with the following properties*:

(a) *The relations $h_1 \in H$, $h_2 \in H$ and $h_1 \neq h_2$ imply $\tilde{h}_1 \neq \tilde{h}_2$;*

(b) *H is convex.*

Consider the following assertions about H:

(α) *H is compact for the topology of pointwise convergence on E.*

(β) *There is a set $E_0 \in \mathscr{E}$ which carries μ, with the following property: For every $\epsilon > 0$ there is a countable partition $(E_n^\epsilon)_n$ of E_0 into sets belonging to $\mathscr{E}$ with $\mu(E_n^\epsilon) > 0$ such that*

$$s \in E_n^\epsilon,\ t \in E_n^\epsilon \Rightarrow |\, h(s) - h(t)| \leqslant \epsilon, \qquad \textit{for all} \quad h \in H.$$

Then (α) $\Rightarrow$ (β).

We shall not formulate the "corresponding converse" to Theorem 2 above (see Theorem 2 in [5]). Instead we shall make the following remark which suffices for our purposes.

Remark. Let $H \subset \mathscr{L}$ be a set satisfying the following condition:

(β) There is a set $E_0 \in \mathscr{E}$ which carries μ, with the following property: For every $\epsilon > 0$ there is a countable partition (E_n^ϵ) of E_0 into sets belonging to $\mathscr{E}$ with $\mu(E_n^\epsilon) > 0$ such that

$$s \in E_n^\epsilon,\ t \in E_n^\epsilon \Rightarrow |\, h(s) - h(t)| \leqslant \epsilon, \qquad \text{for all} \quad h \in H.$$

Then the relations $h_1 \in H$, $h_2 \in H$ and $h_{1|E_0} \neq h_{2|E_0}$ imply $\tilde{h}_1 \neq \tilde{h}_2$, that is $H_{|E_0}$ has the "separation property."

We now turn our attention to vector-valued mappings. Let X be a *Banach space*, X' its *Banach space dual*. For the duality between X and X' we use the notation $\langle x', x \rangle = x'(x)$, for $x \in X$, $x' \in X'$. If $g\colon E \to X$ and $x' \in X'$, we denote by $\langle x', g \rangle$ the mapping $t \to \langle x', g(t) \rangle$ of E into R. For the sake of completeness we recall the terminology concerning weakly measurable and strongly measurable mappings of E into X, as used in [5]:

We say that $g\colon E \to X$ is *weakly measurable* ("scalairement mesurable" in Bourbaki's terminology; see [1]) if $\langle x', g \rangle \in \mathscr{L}$ for each $x' \in X'$. We say that $g\colon E \to X$ is *strongly* (Bochner) *measurable* if there is a sequence (s_n) of simple functions (that is $s_n\colon E \to X$ is countably valued and each value is assumed on a measurable set), such that $\lim_n s_n(t) = g(t)$, μ-almost surely.

We may now extend Theorem 3 of [5] to *arbitrary* weakly measurable mappings, as follows:

THEOREM 3. *Let $g: E \to X$ be a weakly measurable mapping. We have:*

(1) *Suppose that the relations $x' \in X'$, $y' \in X'$ and $\langle x', g\rangle \neq \langle y', g\rangle$ imply $\langle x', g\rangle \not\equiv \langle y', g\rangle(\mu)$. Then $g: E \to X$ is strongly measurable.*

(2) *Conversely, if $g: E \to X$ is strongly measurable, there is a set $E_0 \in \mathscr{E}$ carrying μ such that the relations $x' \in X'$, $y' \in X'$ and $\langle x', g\rangle_{|E_0} \neq \langle y', g\rangle_{|E_0}$ imply $\langle x', g\rangle \not\equiv \langle y', g\rangle(\mu)$.*

Proof. The proof is similar to that of Theorem 3 in [5]. In fact, let $X_1' = \{x' \in X' \mid \| x' \| \leqslant 1\}$ and define

$$H = \{\langle x', g\rangle \mid x' \in X_1'\}.$$

Clearly $H \subset \mathscr{L}$, H is convex and H is compact for the topology of pointwise convergence on E (Alaoglu's theorem; see [3, p. 424]).

Part (1) follows by applying Theorem 2 above. Part (2) follows by approximating g with simple functions and making use of the Remark at the end of Theorem 2 above.

Remark. It seems likely (although this is somewhat loosely stated) that in general, for abstract-valued functions, the "separation property" is what makes the difference between weak measurability and strong measurability.

REFERENCES

1. N. BOURBAKI, Intégration, Hermann, Paris, 1959–1968.
2. S. D. CHATTERJI, Un principe de sous-suites dans la théorie des probabilités, "Séminaire de Probabilités VI, Université de Strasbourg," pp. 72–89, Springer-Verlag, New York, 1972.
3. N. DUNFORD AND J. T. SCHWARTZ, "Linear Operators," Part I, Interscience, New York, 1958.
4. A. IONESCU TULCEA AND C. IONESCU TULCEA, "Topics in the Theory of Lifting," Springer-Verlag, New York, 1969.
5. A. IONESCU TULCEA, On pointwise convergence, compactness and equicontinuity in the lifting topology (I), *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **26** (1973), 197–205.
6. J. KOMLÓS, A generalization of a problem of Steinhaus, *Acta Math. Acad. Sci. Hungary* **18** (1967), 217–229.
7. P. A. MEYER, Représentation intégrale des fonctions excessives. Résultats de Mokobodzki, "Séminaire de Probabilités V, Université de Strasbourg," pp. 196–208, Springer-Verlag, Berlin, 1971.