On the convergence in measure.

By S. HARTMAN*) and E. MARCZEWSKI in Wrocław (Poland).

In 1909 F. RIESZ introduced the important notion of *convergence in measure* (called later by other authors *asymptotical convergence*) of sequences of functions [1]). The importance of this concept increased particularly since KOLMOGOROFF gave the measure-theoretical interpretation of probability theory [2]). In that interpretation the so-called *convergence in probability* (e. g. in Bernoulli's law of large numbers) is the same as the convergence in measure.

The present contribution to the study of the convergence in measure has been suggested by the following theorem, frequently used in the probability theory: If a sequence $\{f_n\}$ of measurable functions is convergent in measure to a function f, then the distribution functions F_n of f_n are convergent to the distribution function F of f in each continuity point of F [3]). Here we give a stronger condition which is not only necessary but also sufficient for the convergence in measure of a sequence of functions (Theorems 3 and 3') [4]). Our result gives a new relation between the properties of functions and those of the sets obtained from them by the operation of converse image.

The last paragraph contains some applications to the study of stochastically independent functions.

1. Preliminaries.

We consider an abstract space X, a σ-field $\mathbf{M}$ of subsets of X and a σ-measure (i. e. a σ-additive and non negative set function) $\mu(E)$ in $\mathbf{M}$ such that $\mu(X)=1$. Points of X will be denoted by x [5]).

---

*) Scholar of the Commission for the Restoration of Polish Science.

1) F. RIESZ, Sur les suites de fonctions mesurables, *Comptes Rendus Acad. Sci. Paris*, **148** (1909), pp. 1303–1305.

2) A. KOLMOGOROFF, *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Berlin, 1933).

3) See e. g. M. FRÉCHET, *Recherches théoriques modernes sur la théorie des probabilités*. I (Paris, 1937), p. 170 and S. MAZURKIEWICZ, Sur les espaces de variables aléatoires, *Fundamenta Math.*, **36** (1949), pp. 288—302, especially p. 291.

4) Another necessary and sufficient condition was given by W. KOZAKIEWICZ, Sur les conditions nécessaires et suffisantes pour la convergence stochastique, *Fundamenta Math.*, **31** (1938), pp. 160—178.

5) With indices if necessary.

We consider a metric space Y; we denote by y[5]) its points and by $\varrho(y_1, y_2)$ the distance of y_1 and y_2. For each subset E of Y we design by $\mathrm{Fr}E$ the boundary of E, i. e. the set $\overline{E}\cdot(\overline{Y-E})$. In general our set-theoretical terminology and notation is that of KURATOWSKI's *Topologie*[6]).

Each σ-measure defined in the class of Borel subsets of Y will be called a *Borel measure* in Y.

In the sequel the letter f (and g)[5]), if nothing to the contrary is explicitly stated, shall denote a mapping $y=f(x)$ of X into Y which is measurable with respect to μ, i. e. such that $f^{-1}(B)\in\mathbf{M}$ for each Borel subset B of Y. Obviously

(i) For each f the set function $\mu[f^{-1}(B)]$ is a Borel measure in Y.

We say, following F. RIESZ, that a sequence $\{f_n\}$ *converges in measure* to f or, else, *μ-converges* to f and we write $f_n \underset{\mu}{\longrightarrow} f$, if for each $\varepsilon>0$

$$\lim_{n=\infty} \mu \mathop{E}_{x} \{\varrho[f_n(x), f(x)] > \varepsilon\} = 0.$$

We say that a sequence of sets $E_n\in\mathbf{M}$ is *μ-convergent* to a set $E\in\mathbf{M}$ and we write $E_n \underset{\mu}{\longrightarrow} E$, if the sequence of characteristic functions of E_n is μ-convergent to the characteristic function of E, or, in other words, if $\lim_{n=\infty}\mu(E_n \dot{-} E)=0$ (where $A\dot{-}B$ denotes the *symmetric difference* of A and B, i. e. the set $(A-B)+(B-A)$)[7]).

It easily follows from elementary properties of the symmetric difference[8]) that

(ii) If $A_n \underset{\mu}{\longrightarrow} A$ and $B_n \underset{\mu}{\longrightarrow} B$, then

$$A_n+B_n \underset{\mu}{\longrightarrow} A+B, \quad A_nB_n \underset{\mu}{\longrightarrow} AB, \quad A_n-B_n \underset{\mu}{\longrightarrow} A-B.$$

(iii) If $A_n \underset{\mu}{\longrightarrow} A$, then $\lim_{n=\infty}\mu(A_n)=\mu(A)$.

$\mathbf{Q}$ and $\mathbf{K}$ being two subclasses of $\mathbf{M}$ we say that $\mathbf{Q}$ is *μ-dense* on $\mathbf{K}$ if for each $E\in\mathbf{K}$ there is a sequence of sets $E_n\in\mathbf{Q}$ μ-convergent to E.

A class $\mathbf{V}$ of Borel subsets of Y is called a *basis* of Y, if each open set $G\subset Y$ is the sum of a denumerable subclass of $\mathbf{V}$. Obviously

(iv) If $\mathbf{V}$ is a denumerable basis of Y, then for each $\varepsilon>0$ the class of all sets $V\in\mathbf{V}$ with $\delta(V)<\varepsilon$ is again a basis of Y.

(v) If Y is a separable metric space (i. e. if there is a denumerable basis of Y) then each basis of Y contains a denumerable basis.

---

[6]) C. KURATOWSKI, *Topologie* I, deuxième édition (Warszawa-Wrocław, 1948).

[7]) The number $\mu(A\dot{-}B)$ may be considered as the distance of A and B See e. g. O. NIKODYM, Sur une généralisation des intégrales de M. J. Radon, *Fundamenta Math.*, **15** (1930), pp. 131—179, especially p. 137.

[8]) See e. g. F. HAUSDORFF, *Mengenlehre* (Berlin—Leipzig, 1935), Ergänzungen, pp. 276—278 and E. MARCZEWSKI, Concerning the symmetric difference in the Theory of Sets and in Boolean algebras, *Colloquium Math.*, 1 (1948), pp. 199-202, especially pp. 200-201.

Now let us prove that

(vi) For each finite Borel measure $\varphi(B)$ in Y the class $\mathbf{J}_\varphi$ of all open subsets G of Y with $\varphi(\operatorname{Fr} G)=0$ is an additive basis of Y.

The relation $\operatorname{Fr} A+\operatorname{Fr} B \supset \operatorname{Fr}(A+B)$ implies the additivity of $\mathbf{J}_\varphi$. Further, since for each open set $G \subset Y$ the set R of all $r>0$ such that

$$\varphi\{\underset{y}{E}[\varrho(y, Y-G)=r]\}>0$$

is denumerable, there exists a sequence of positive numbers $r_n \overline{\in} R$ convergent to 0. Putting

$$V_n=\underset{y}{E}[\varrho(y, Y-G)>r_n]$$

we have $V_n \in \mathbf{J}_\varphi$ and $G=V_1+V_2+\ldots$, whence $\mathbf{J}_\varphi$ is a basis of Y.

We call a *generalized basis* of Y each class $\mathbf{V}$ of Borel sets such that the smallest field $\mathbf{V}_0$ containing $\mathbf{V}$ is a basis of Y. Consequently

(vii) Each basis of Y is a generalized basis of Y.

We shall deal with an important case of the generalized basis for the Euclidean space R^k. Let L_a denote, for each $a=(a_1, a_2, \ldots, a_k) \in R^k$, the set

$$\underset{(\xi_1, \ldots, \xi_k)}{E}[\xi_1<a_1, \ldots, \xi_k<a_k].$$

It is obvious that

(viii) The set of all subsets L_a of R^k where a runs over a dense subset of R^k is a generalized basis of R^k.

(ix) For each finite Borel measure φ in R^k the set of the continuity points of the function $F(a)=\varphi(L_a)$ (where a runs over R^k) is dense in R^k.

(x) For each function f, the values of which belong to R^k, its distribution function, i. e. the function $F(a)=\mu[f^{-1}(L_a)]$, is continuous at $a_0 \in R^k$, if and only if $\mu[f^{-1}(\operatorname{Fr} L_{a_0})]=0$.

2. Convergence in measure of functions characterized by that of sets.

Theorem 1. *If $f_n \underset{\mu}{\longrightarrow} f$ and, if E is a Borel subset of Y such that*

(1) $$\mu[f^{-1}(\operatorname{Fr} E)]=0$$

then $f_n^{-1}(E) \underset{\mu}{\longrightarrow} f^{-1}(E)$.

Proof. Put for positive integers j and n

$$A_j^n=\underset{x}{E}\{\varrho[f(x), f_n(x)] \geqq 1/j\},$$

$$H_j=\underset{y}{E}[\varrho(y, E)<1/j] \cdot \underset{y}{E}[\varrho(y, Y-E)<1/j].$$

Consequently we have

(2) $$H_1 \supset H_2 \supset \ldots, \qquad H_1 \cdot H_2 \ldots=\operatorname{Fr} E$$

(3) $$f^{-1}(E) \dot{-} f_n^{-1}(E) \subset A_j^n+f^{-1}(H_j) \qquad \text{for } j=1,2, \ldots, \ n=1,2, \ldots.$$

By (1) and (2) we have

$$\lim_{j=\infty} \mu[f^{-1}(H_j)] = \mu[f^{-1}(\mathrm{Fr}\, E)] = 0.$$

Thus, for any $\varepsilon > 0$ there is a positive integer j_0 such that

(4) $$\mu[f^{-1}(H_{j_0})] < \varepsilon/2.$$

The sequence $\{f_n\}$ being μ-convergent to f, there is a positive integer N such that

(5) $$\mu(A_{j_0}^n) < \varepsilon/2 \qquad \text{for } n > N.$$

From (3), (4) and (5) we obtain $\mu[f^{-1}(E) \dot{-} f_n^{-1}(E)] < \varepsilon$ for $n > N$, q. e. d.

Theorem 1 and propositions 1(iii) and 1(x) imply the well-known theorem on the convergence of distribution functions, quoted in the Introduction above.

Theorem 2. *Let* $\mathbf{V}$ *be a denumerable generalized basis of* Y. *If*

(6) $$f_n^{-1}(V) \underset{\mu}{\longrightarrow} f^{-1}(V)$$

for each $V \in \mathbf{V}$, *then* $f_n \underset{\mu}{\longrightarrow} f$.

Proof. Since (6) holds for each $V \in \mathbf{V}$, it also holds by 1(ii) for each V belonging to the smallest field $\mathbf{V}_0$ containing $\mathbf{V}$.

Let us suppose that the sequence $\{f_n\}$ is not μ-convergent to f, or, in other words, that there are two numbers $\varepsilon > 0$ and $\eta > 0$, and an increasing sequence $\{k_n\}$ of positive integers such that

(7) $$\mu(E_n) > \eta \qquad \text{for } n = 1, 2, \ldots,$$

where

$$E_n = \underset{x}{E}\{\varrho[f_{k_n}(x), f(x)] > \varepsilon\}.$$

By definition of generalized basis and 1(iv), there is a sequence $\{V_n\}$ of sets belonging to $\mathbf{V}_0$ such that $Y = V_1 + V_2 + \ldots$ and $\delta(V_n) < \varepsilon$.

Further, by the σ-additivity of the measure μ there is a positive integer s such that putting $Y_0 = V_1 + V_2 + \ldots + V_s$ we have $\mu[f^{-1}(Y - Y_0)] < \eta/2$.

From this and (7) we obtain

(8) $$\mu[E_n \cdot f^{-1}(Y_0)] = \mu[E_n - f^{-1}(Y - Y_0)] > \eta/2.$$

Since $E_n \cdot f^{-1}(Y_0) = E_n[f^{-1}(V_1) + f^{-1}(V_2) + \ldots + f^{-1}(V_s)]$, there is by (8) for each positive integer n a positive integer j_n such that

$$\mu[E_n \cdot f^{-1}(V_{j_n})] > \eta/2s \qquad 1 \leq j_n \leq s.$$

Thus, there is a positive integer j_0 and an increasing sequence $\{l_n\}$ of positive integers such that $j_{l_n} = j_0$ for $n = 1, 2, \ldots$. Denoting by V the set V_{j_0} we obtain

(9) $$\mu[E_{l_n} \cdot f^{-1}(V)] > \eta/2s \qquad \text{for } n = 1, 2, \ldots.$$

If $x \in E_{l_n} \cdot f^{-1}(V)$, then

$$f(x) \in V \quad \text{and} \quad \varrho[f(x), f_{k_{l_n}}(x)] > \varepsilon,$$

whence, in virtue of the inequality $\delta(V) < \varepsilon$, we have $f_{k_{l_n}}(x) \overline{\in} V$. Consequently

$$f_{k_{l_n}}^{-1}(V) \cdot f^{-1}(V) \cdot E_{l_n} = 0,$$

whence

$$f^{-1}(V) \dot{-} f_{k_{l_n}}^{-1}(V) \supset f^{-1}(V) - f_{k_{l_n}}^{-1}(V) \supset E_{l_n} \cdot f^{-1}(V)$$

and finally

$$\mu[f^{-1}(V) \dot{-} f_{k_{l_n}}^{-1}(V)] > \eta/2s \qquad \text{for } n = 1, 2, \ldots.$$

Therefore the sequence $f_{k_{l_n}}^{-1}(V)$ is not μ-convergent to $f^{-1}(V)$, which contradicts the hypothesis. The proof is thus complete.

Comparing Theorems 1 and 2, we obtain the

Theorem 3. *If Y is separable, then the following conditions are equivalent:*

I. $f_n \xrightarrow[\mu]{} f$.

II. *There is a denumerable generalized basis* $\mathbf{V}$ *of Y such that* $f_n^{-1}(V) \xrightarrow[\mu]{} f^{-1}(V)$ *for each* $V \in \mathbf{V}$.

III. *There is a denumerable basis* $\mathbf{V}$ *of Y such that* $f_n^{-1}(V) \xrightarrow[\mu]{} f^{-1}(V)$ *for each* $V \in \mathbf{V}$.

IV. $f_n^{-1}(B) \xrightarrow[\mu]{} f^{-1}(B)$ *for each Borel set* $B \subset Y$ *such that* $\mu[f^{-1}(\mathrm{Fr}\, B)] = 0$.

In fact: the implication I → IV is stated by Theorem 1; IV → III in virtue of 1 (vi) and 1 (v); III → II in view of 1 (vii), and finally Theorem 2 gives the implication II → I.

From Theorem 3 and the propositions 1 (i) and 1 (viii—x) we obtain

Theorem 3′. *If the values of* f_n *and* f *are points of the Euclidean space* R^k, *the following conditions are equivalent:*

I. $f_n \xrightarrow[\mu]{} f$.

V. *There is a dense subset D of* R^k *such that* $f_n^{-1}(L_a) \xrightarrow[\mu]{} f^{-1}(L_a)$ *for each* $a \in D$.

VI. $f_n^{-1}(L_a) \xrightarrow[\mu]{} f^{-1}(L_a)$ *for each continuity point a of the distribution function of f.*

3. Convergence in measure and stochastic independence.

Two sets $A, B \in \mathbf{M}$ are called *stochastically independent* or simply *independent* if $\mu(AB) = \mu(A) \cdot \mu(B)$. Two classes $\mathbf{A}$ and $\mathbf{B}$ contained in $\mathbf{M}$ are called independent, if any two sets $A \in \mathbf{A}$ and $B \in \mathbf{B}$ are independent. Two functions f and g are called independent, if for any two Borel sets A and B in Y we have $\mu[f^{-1}(A) \cdot g^{-1}(B)] = \mu[f^{-1}(A)] \cdot \mu[g^{-1}(B)]$ or, in other words, if the classes $\mathbf{B}_f$ and $\mathbf{B}_g$ of all converse images of Borel sets under f and g respectively are independent.

In virtue of 1 (ii) and 1 (iii):

(i) If for each positive integer n the sets $A_n \in \mathbf{M}$ and $B_n \in \mathbf{M}$ are independent, and if $A_n \underset{\mu}{\rightarrow} A$ and $B_n \underset{\mu}{\rightarrow} B$, then A and B are independent.

Consequently:

(ii) If the classes $\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{K}_1, \mathbf{K}_2$ are contained in $\mathbf{M}$, and if $\mathbf{Q}_j$ is μ-dense on $\mathbf{K}_j$ for $j = 1, 2$, then the independence of $\mathbf{Q}_1$ and $\mathbf{Q}_2$ implies that of $\mathbf{K}_1$ and $\mathbf{K}_2$.

Now we prove the following lemma:

(iii) If $\mathbf{V}$ is an additive basis of Y, and if φ is a finite Borel measure in Y, then $\mathbf{V}$ is φ-dense on the class $\mathbf{B}$ of all Borel sets in Y.

It is well known that for each Borel set B and each $\eta > 0$ there is an open set $G \supset B$ such that $\varphi(G - B) < \eta$[9]. The class $\mathbf{V}$ being an additive basis of Y, there is an increasing sequence of sets $V_n \in \mathbf{V}$ such that $G = V_1 + V_2 + \ldots$. Consequently, by the σ-additivity of μ there is a positive integer s such that $\varphi(G - V_s) < \eta$. Since

$$B \dot{-} V_s \subset (B \dot{-} G) + (G \dot{-} V_s) = (G - B) + (G - V_s),$$

we obtain finally $\varphi(B \dot{-} V_s) < 2\eta$, q. e. d.

For every class $\mathbf{Q}$ of subsets of Y and for every f let $\mathbf{Q}_f$ denote the class of the converse images of all $E \in \mathbf{Q}$ under f. Then, from (ii) and (iii) we obtain:

(iv) For each additive basis $\mathbf{V}$ of Y the independence of $\mathbf{V}_f$ and $\mathbf{V}_g$ is a necessary and sufficient condition for the independence of f and g[10].

The necessity of the condition is obvious. For the sufficiency it is enough 1^0 to apply (iii) for two Borel measures in Y:

$$\varphi(B) = \mu[f^{-1}(B)] \text{ and } \psi(B) = \mu[g^{-1}(B)],$$

2^0 to remark that the φ-density and the ψ-density of $\mathbf{V}$ on the class $\mathbf{B}$ of all Borel subsets of Y imply the μ-density of $\mathbf{V}_f$ on $\mathbf{B}_f$ and the μ-density of $\mathbf{V}_g$ on $\mathbf{B}_g$, 3^0 to apply (ii).

Theorem 4. *If f_n and g_n are independent for each positive integer n, if, further, $f_n \underset{\mu}{\rightarrow} f$ and $g_n \underset{\mu}{\rightarrow} g$, then f and g are independent.*

Proof. By Theorem 1

$$(1) \qquad f_n^{-1}(G) \underset{\mu}{\rightarrow} f^{-1}(G) \text{ and } g_n^{-1}(G) \underset{\mu}{\rightarrow} g^{-1}(G)$$

[9]) Or, which is the same for finite measures, B is contained in a set G_δ the measure of which is equal to $\mu(B)$. See e. g. E. Marczewski and R. Sikorski, Remark on measure and category, *Colloquium Math.*, **2** (1949), pp. 13—19, especially p. 14.

[10]) This theorem is a special case of a theorem proved (in a more complicated way) by S. Hartman, Sur l'indépendance stochastique de familles d'ensembles, *Comptes Rendus de la Société des Sciences et des Lettres de Wrocław*, **1** (in print), Théorème 5.

for every G such that

(2) $$\mu[f^{-1}(\mathrm{Fr}\,G)] = 0 = \mu[g^{-1}(\mathrm{Fr}\,G)].$$

In virtue of 1(i) the set function

$$\gamma(B) = \mu[f^{-1}(B)] + \mu[g^{-1}(B)]$$

is a Borel measure in Y and, consequently, according to 1(vi), the class $\mathbf{V}$ of all open sets G satisfying (2) is an additive basis of Y. The functions f_n and g_n being independent, it follows from (1) and (i) that $\mathbf{V}_f$ and $\mathbf{V}_g$ are independent and from (iv) that f and g are independent. Theorem 4 is thus proved.

We shall prove the following lemma:

(v) If Y is of the power of the continuum[11]), then f is independent with respect to itself, if and only if f is constant almost everywhere (i. e. if there is an y_0 such that $\mu\{\underset{x}{E}[f(x) \neq y_0]\} = 0$).

The sufficiency is obvious. To prove the necessity it is enough 1° to remark that if f and f are independent, then the Borel measure $\mu[f^{-1}(B)]$ in Y assumes only the values 0 and 1, and 2° to make use of the fact that for each Borel measure in Y, which assumes only the values 0 and 1, there is an $y_0 \in Y$ such that $\varphi[(y_0)] = 1$.[12])

From Theorem 4 and (v) we obtain

Theorem 5. *If Y is of the power of the continuum*[11]), *$f_n \underset{\mu}{\rightarrow} f$ and for each N there exist $k > l > N$ such that f_k and f_l are independent, then f is constant almost everywhere.*

(Received November 25, 1949.)

[11]) And more generally, if the power of Y is less than the first aleph inaccessible in the strong sense.

[12]) The proof is easy in case Y separable. For the non-separable case see E. Marczewski and R. Sikorski, Measures in non-separable metric spaces, *Colloquium Math.*, 1 (1948), pp. 132—139, especially p. 139, Theorem VI. We remark besides that Theorem III in the same paper (p. 137) permits to replace the hypothesis of separability of Y in Theorems 2 and 3 in the present paper by the weaker hypothesis that the separability character $\mathfrak{m}$ of Y is $\aleph_0$, $\aleph_1$ or, more generally, that $\mathfrak{m}$ is less than the first aleph inaccessible in the weak sense.