On the limiting probability distribution on a compact topological group

by

K. Urbanik (Wrocław)

I. Let G be a compact (not necessarily commutative) topological group. A regular completely additive measure μ defined on the class of all Borel subsets of G, with $\mu(G)=1$, will be called a *probability distribution*. A sequence of probability distributions $\mu_1, \mu_2, \ldots$ is said to be *weakly convergent to a probability distribution* μ if

$$\lim_{n\to\infty} \int_G f(x)\,\mu_n(dx) = \int_G f(x)\,\mu(dx)$$

for any complex-valued continuous function f defined on G.

Let $X_1, X_2, \ldots$ be a sequence of independent G-valued random variables with the probability distributions $\mu_1, \mu_2, \ldots$ Put

$$Y_n = X_1 \cdot X_2 \cdot \ldots \cdot X_n \quad (n=1,2,\ldots), \tag{1}$$

where the product is taken in the sense of group multiplication in G. Let us denote by ν_n the probability distribution of the random variable Y_n. It is well known that

$$\nu_n = \mu_1 * \mu_2 * \ldots * \mu_n \quad (n=1,2,\ldots),$$

where the convolution $*$ is defined by the formula:

$$\nu * \lambda(E) = \int_G \nu(E\cdot x^{-1})\,\lambda(dx).$$

The limiting distribution of the sequence of random variables Y_n is the weak limit of the sequence of the probability distributions $\mu_1 * \mu_2 * \ldots * \mu_n$.

We say that the sequence of probability distributions $\mu_1, \mu_2, \ldots$ is *normal* (by an analogy with normal numbers in the sense of Borel), if for every integer k exists a sequence of integers $n_1 < n_2 < \ldots$ such that

$$\mu_s = \mu_{n_j+s} \quad (s=1,2,\ldots,k;\ j=1,2,\ldots).$$

In particular, the sequence $\mu_1 = \mu_2 = \ldots$ is normal.

The aim of this paper is to find the class of all possible limiting distributions of (1) and the conditions of convergence when the sequence of probability distributions $\mu_1, \mu_2, \ldots$ is normal. This is an answer to a problem raised by A. Rényi (see [6]).

For the case of the additive group of real numbers modulo 1 and the finite commutative group the results are known (cf. [1], [5], and [8]). The results of this paper are connected to some extent with the work of Kawada and Ito [4], who discussed the convergence of probability distributions and stable distributions for the case of a compact separable group.

II. Let K_μ, where μ is a probability distribution on G, denote the class of all compact subsets E of G such that $\mu(E)=1$. Put $A_\mu = \bigcap_{E \in K_\mu} E$. It is easy to see that A_μ is a compact subset of G. We shall prove that

$$\mu(A_\mu) = 1 . \tag{2}$$

Let V be an arbitrary open set containing A_μ. Then, since

$$V \cup \bigcup_{E \in K_\mu} (G \setminus E) = G$$

and G is a compact set, there exists a finite covering of G:

$$V \cup \bigcup_{j=1}^{n} (G \setminus E_j) = G,$$

where $E_j \in K_\mu$. Consequently, $\mu(V) = \mu(G) = 1$. Then, according to the regularity of μ, the equality

$$\mu(A_\mu) = \inf_V \mu(V) = 1$$

is true. The formula (2) is thus proved. It is easy to see that the values $\mu(A)$ for $A \subset A_\mu$ determine the measure μ.

$[E]$ will denote the smallest closed subgroup of G containing E. We shall use the following notation:

$$AB = \{xy\colon\ x \in A,\ y \in B\}, \qquad A^{-1} = \{x^{-1}\colon\ x \in A\}.$$

The following theorems will be proved:

THEOREM 1. *Let $\mu_1, \mu_2, \ldots$ be a normal sequence of probability distributions. If the sequence $\mu_1 * \mu_2 * \ldots * \mu_n$ weakly converges to ν, then ν is the Haar measure of the subgroup $A_\nu = [\bigcup_{n=1}^{\infty} A_{\mu_n}]$.*

THEOREM 2. *Let $\mu_1, \mu_2, \ldots$ be a normal sequence of probability distributions. The sequence $\mu_1 * \mu_2 * \ldots * \mu_n$ converges if and only if the equality*

$$[\bigcup_{n=1}^{\infty} A_{\mu_n}] = [\bigcup_{n=1}^{\infty} A_{\mu_1} A_{\mu_2} \ldots A_{\mu_n} A_{\mu_n}^{-1} \ldots A_{\mu_2}^{-1} A_{\mu_1}^{-1}] \tag{$*$}$$

holds.

In particular, for the case $\mu_n = \mu$ $(n = 1, 2, \ldots)$ we find that:

The sequence μ, $\mu * \mu$, $\mu * \mu * \mu$, ... converges if and only if the equality

$$[A_\mu] = [A_\mu A_\mu^{-1}]$$

holds. The limiting distribution is the Haar measure of the subgroup $[A_\mu]$.

III. Before proving the Theorems we shall give some elementary properties of the characteristic function of the probability distribution.

$\mathfrak{A}(G)$ will denote the class of all continuous finitely dimensional unitary representations of the group G (see [7], chapter IV). $\mathfrak{A}_0(G)$ will denote the subset of $\mathfrak{A}(G)$ containing all the irreducible representations $U \not\equiv 1$. The matrix-valued function

$$\varphi_\mu(U) = \int_G U(x)\mu(dx) \qquad \big(U \in \mathfrak{A}(G)\big)$$

is called the *characteristic function* of the probability distribution μ. If U is the unit representation: $U \equiv 1$, then $\varphi_\mu(U) = 1$. It is easy to prove that

$$\varphi_{\mu * \nu}(U) = \varphi_\mu(U) \cdot \varphi_\nu(U) .$$

Let $\mathcal{B}$ be the Banach space of all continuous complex-valued functions f in G with the norm $\|f\| = \max_{x \in G} |f(x)|$. The general form of linear functionals in $\mathcal{B}$ satisfying following conditions:

$$L(f) \geqslant 0 \quad \text{for} \quad f(x) \geqslant 0, \quad L(1) = 1,$$

is given by the formula

$$L(f) = L_\mu(f) = \int_G f(x)\mu(dx) ,$$

where μ is the uniquely determined probability distribution (see [3], p. 247, 248). The weak convergence of functionals L_{μ_n} is equivalent to the weak convergence of distributions μ_n.

$\mathcal{D}$ will denote the set of all linear combinations of matrix elements of $U \in \mathfrak{A}_0(G)$ and $U \equiv 1$. According to the theorem of Peter-Weyl (see [7], § 21, 22) $\mathcal{D}$ is a dense subset of $\mathcal{B}$. Hence the equality $\varphi_\mu(U) = \varphi_\nu(U)$ for $U \in \mathfrak{A}_0(G)$ implies $L_\mu(f) = L_\nu(f)$ for $f \in \mathcal{D}$, and, consequently, $\mu = \nu$. Thus the probability distribution is uniquely determined by the values of the characteristic function on $\mathfrak{A}_0(G)$. Obviously, if the sequence of distributions $\mu_1, \mu_2, \ldots$ weakly converges to the distribution μ, then the sequence of characteristic functions $\varphi_{\mu_1}(U), \varphi_{\mu_2}(U), \ldots$ converges to $\varphi_\mu(U)$. Conversely, if $\varphi_{\mu_1}(U), \varphi_{\mu_2}(U), \ldots$ converges to $\varphi_\mu(U)$ for $U \in \mathfrak{A}_0(G)$, then the sequence of linear functionals $L_{\mu_1}(f), L_{\mu_2}(f), \ldots$ converges to $L_\mu(f)$ for $f \in \mathcal{D}$ and, according to the density of $\mathcal{D}$ in $\mathcal{B}$, the sequence $L_{\mu_1}, L_{\mu_2}, \ldots$ weakly converges to L_μ. We see that the weak convergence of the sequence of probability distributions is equivalent to the convergence of the sequence of characteristic functions for $U \in \mathfrak{A}_0(G)$.

IV. The norm of matrix $B=(b_{ij})$ will be defined by the following well known formula:

$$\|B\| = \Big(\max_{1\leqslant j\leqslant n} \sum_{i=1}^{n} |b_{ij}|^2\Big)^{1/2}.$$

$\| \ \|$ is the submultiplicative norm. By e we shall denote the unit element of G.

LEMMA 1. *Let μ be a probability distribution and*

$$e \in A_\mu . \tag{3}$$

If $U \in \mathfrak{A}([A_\mu])$ and

$$\|\varphi_\mu(U)\| = 1 , \tag{4}$$

then $U \notin \mathfrak{A}_0([A_\mu])$.

Proof. Let $U(x)=\big(u_{ij}(x)\big)$. There is an integer k such that the equality

$$\|\varphi_\mu(U)\|^2 = \sum_{i=1}^{n} \Big| \int_{A_\mu} u_{ik}(x)\, \mu(dx) \Big|^2 \tag{5}$$

is true. Since $U(x)$ is a unitary matrix, then

$$\sum_{i=1}^{n} |u_{ik}(x)|^2 = 1 \quad \text{for} \quad x \in [A_\mu].$$

Hence, according to (4) and (5),

$$\sum_{i=1}^{n} \Big(\Big| \int_{A_\mu} u_{ik}(x)\, \mu(dx) \Big|^2 - \int_{A_\mu} |u_{ik}(x)|^2 \mu(dx) \Big) = 0 .$$

Since $u_{ik}(x)$ are continuous functions, then the last equality implies $u_{ik}(x) = \text{const}$ for $x \in A_\mu$. From condition (3) it follows that the equalities

$$u_{ik}(x) = u_{ik}(e) = \delta_{ik} \quad \text{for} \quad x \in A_\mu,\ i=1,2,\ldots,n,$$

are satisfied. Then $\langle \delta_{1k}, \delta_{2k}, \ldots, \delta_{nk} \rangle$ is the invariant vector under the transformations $U(x)$ for $x \in A_\mu$, and consequently for $x \in [A_\mu]$. The lemma is thus proved.

LEMMA 2. *The formula*

$$A_{\mu * \nu} = A_\mu \cdot A_\nu$$

is true for every probability distributions μ and ν.

Proof. The formulas

$$\mu * \nu(A_\mu \cdot A_\nu) = \int_{A_\nu} \mu(A_\mu \cdot A_\nu x^{-1})\, \nu(dx)$$

and

$$A_\mu \cdot A_\nu x^{-1} \supset A_\mu \quad \text{for} \quad x \in A_\nu$$

imply the inequality

$$\mu * \nu(A_\mu \cdot A_\nu) \geqslant \int_{A_\nu} \mu(A_\mu)\, \nu(dx) = 1 .$$

Thus $\mu * \nu(A_\mu \cdot A_\nu) = 1$. Since $A_\mu \cdot A_\nu$ is the compact subset of G, therefore we obtain the following inclusion:

$$A_{\mu * \nu} \subset A_\mu \cdot A_\nu . \tag{6}$$

The equality

$$1 = \mu * \nu(A_{\mu*\nu}) = \int_{A_\nu} \mu(A_{\mu*\nu} \cdot x^{-1})\, \nu(dx)$$

implies

$$\mu(A_{\mu*\nu} \cdot x^{-1}) = 1 \quad \text{for} \quad x \in A_\nu \setminus N ,$$

where

$$\nu(N) = 0 . \tag{7}$$

Since $A_{\mu*\nu} x^{-1}$ is the compact subset of G, therefore we obtain

$$A_{\mu*\nu} x^{-1} \supset A_\mu \quad \text{for} \quad x \in A_\nu \setminus N .$$

This implies

$$A_{\mu*\nu} \supset A_\mu \cdot (A_\nu \setminus N) . \tag{8}$$

From the formula (7) and the definition of the set A_ν it follows that $A_\nu = \overline{A_\nu \setminus N}$. Consequently, in view of (8), we obtain

$$A_{\mu*\nu} \supset A_\mu \cdot A_\nu .$$

Then it follows from (6) that $A_{\mu*\nu} = A_\mu \cdot A_\nu$. The lemma is thus proved.

LEMMA 3. *If the sequence of probability distributions $\nu_1, \nu_2, \ldots$ weakly converges to ν, then*

$$A_\nu \subset \Big[\bigcup_{n=1}^{\infty} A_{\nu_n}\Big].$$

If $\nu_n = \mu_1 * \mu_2 * \ldots * \mu_n$, where $\mu_1, \mu_2, \ldots$ is a normal sequence, then $\nu * \nu = \nu$ and

$$A_\nu = \Big[\bigcup_{n=1}^{\infty} A_{\mu_n}\Big].$$

Proof. ν is a regular measure and $[\bigcup_{n=1}^{\infty} A_{\nu_n}]$ is a compact subset of G. Then, for arbitrary $\varepsilon > 0$, there is a continuous function f satisfying the following conditions:

$$f(x) = 1 \quad \text{for} \quad x \in [\bigcup_{n=1}^{\infty} A_{\nu_n}],$$

$$\nu([\bigcup_{n=1}^{\infty} A_{\nu_n}]) \geqslant \int_G f(x)\nu(dx) - \varepsilon. \tag{9}$$

Since $A_{\nu_n} \subset [\bigcup_{n=1}^{\infty} A_{\nu_n}]$, we have $\int_G f(x)\nu_n(dx) = 1$, and consequently

$$\lim_{n\to\infty} \int_G f(x)\nu_n(dx) = \int_G f(x)\nu(dx) = 1.$$

According to (9),

$$\nu([\bigcup_{n=1}^{\infty} A_{\nu_n}]) = 1.$$

Then $A_\nu \subset [\bigcup_{n=1}^{\infty} A_{\nu_n}]$. The first part of the lemma is thus proved.

Let $\nu_n = \mu_1 * \mu_2 * \ldots * \mu_n$, where $\mu_1, \mu_2, \ldots$ is a normal sequence. Let $n_1 < n_2 < \ldots$ be a sequence of integers such that

$$\mu_1 * \mu_2 * \ldots * \mu_k = \mu_{n_j+1} * \mu_{n_j+2} * \ldots * \mu_{n_j+k} \quad (j = 1, 2, \ldots).$$

Then the equalities

$$\lim_{j\to\infty} \nu_{n_j+k} = \nu,$$

$$\lim_{j\to\infty} \nu_{n_j+k} = \lim_{j\to\infty} \nu_{n_j} * \mu_1 * \mu_2 * \ldots * \mu_k = \nu * \mu_1 * \mu_2 * \ldots * \mu_k$$

imply

$$\nu = \nu * \mu_1 * \mu_2 * \ldots * \mu_k. \tag{10}$$

Hence

$$\nu = \nu * \lim_{k\to\infty} \mu_1 * \mu_2 * \ldots * \mu_k = \nu * \nu.$$

From Lemma 2 it follows that $A_\nu = A_\nu \cdot A_\nu$. Hence A_ν is a compact semi-group. Thus in view of the theorem of Iwasawa (see [2]) A_ν is a sub-group of G. Equality (10) implies $\nu = \nu * \mu_n$ for every n. Consequently

$$A_\nu = A_\nu \cdot A_{\mu_n} \quad (n = 1, 2, \ldots).$$

Since A_ν is a subgroup of G, we have

$$A_\nu \supset A_{\mu_n} \quad (n = 1, 2, \ldots)$$

and, consequently,

$$A_\nu \supset [\bigcup_{n=1}^{\infty} A_{\mu_n}].$$

The last inclusion and

$$A_\nu \subset [\bigcup_{n=1}^{\infty} A_{\mu_1 * \mu_2 * \ldots * \mu_n}] \subset [\bigcup_{n=1}^{\infty} A_{\mu_n}]$$

imply

$$A_\nu = [\bigcup_{n=1}^{\infty} A_{\mu_n}].$$

The lemma is thus proved.

Proof of Theorem 1. Let ν be a limiting distribution of a sequence $\mu_1 * \mu_2 * \ldots * \mu_n$. From Lemma 3 it follows that $\nu * \nu = \nu$ and

$$A_\nu = [\bigcup_{n=1}^{\infty} A_{\mu_n}].$$

Hence

$$\varphi_\nu(U) \cdot \varphi_\nu(U) = \varphi_\nu(U) \quad \text{for} \quad U \in \mathfrak{A}(A_\nu),$$

and consequently

$$\|\varphi_\nu(U)\| \leqslant \|\varphi_\nu(U)\|^2 \quad \text{for} \quad U \in \mathfrak{A}(A_\nu). \tag{11}$$

Let $U \in \mathfrak{A}_0(A_\nu)$. Then, in view of Lemma 1, $\|\varphi_\nu(U)\| < 1$. Hence, according to (11),

$$\varphi_\nu(U) = 0 \quad \text{for} \quad U \in \mathfrak{A}_0(A_\nu). \tag{12}$$

Let λ be the Haar measure of the group A_ν, with $\lambda(A_\nu) = 1$. It is well known that the equality

$$\varphi_\lambda(U) = \int_{A_\nu} U(x)\lambda(dx) = 0 \quad \text{for} \quad U \in \mathfrak{A}_0(A_\nu)$$

holds (see [7]). From (12) it follows that

$$\varphi_\nu(U) = \varphi_\lambda(U) \quad \text{for} \quad U \in \mathfrak{A}_0(A_\nu),$$

and consequently $\nu = \lambda$. Theorem 1 is thus proved.

Proof of Theorem 2. Let

$$y_n \in A_{\mu_1} \cdot A_{\mu_2} \cdot \ldots \cdot A_{\mu_n} \quad (n = 1, 2, \ldots).$$

Then following formula

$$A_{\mu_1} A_{\mu_2} \cdot \ldots \cdot A_{\mu_n} y_n^{-1} \subset A_{\mu_1} A_{\mu_2} \ldots A_{\mu_n} A_{\mu_n}^{-1} \ldots A_{\mu_2}^{-1} A_{\mu_1}^{-1}$$
$$\subset [A_{\mu_1} A_{\mu_2} \ldots A_{\mu_n} y_n^{-1}] \quad (n = 1, 2, \ldots)$$

is satisfied. Hence

$$[\bigcup_{n=1}^{\infty} A_{\mu_1} A_{\mu_2} \ldots A_{\mu_n} y_n^{-1}] = [\bigcup_{n=1}^{\infty} A_{\mu_1} A_{\mu_2} \ldots A_{\mu_n} A_{\mu_n}^{-1} \ldots A_{\mu_2}^{-1} A_{\mu_1}^{-1}].$$

Put

$$\lambda_n(E)=\begin{cases}1 & \text{if} \quad y_n^{-1}\in E,\\ 0 & \text{if} \quad y_n^{-1}\notin E,\end{cases}$$

$$\pi_n=\mu_1*\mu_2*\ldots*\mu_n*\lambda_n \quad (n=1,2,\ldots).$$

Then

$$A_{\pi_n}=A_{\mu_1}A_{\mu_2}\ldots A_{\mu_n}y_n^{-1},$$

and consequently

(13) $$[\bigcup_{n=1}^{\infty}A_{\pi_n}]=[\bigcup_{n=1}^{\infty}A_{\mu_1}A_{\mu_2}\ldots A_{\mu_n}A_{\mu_n}^{-1}\ldots A_{\mu_2}^{-1}A_{\mu_1}^{-1}].$$

Moreover,

(14) $$\varphi_{\pi_n}(U)=\varphi_{\mu_1*\mu_2*\ldots*\mu_n}(U)\,U(y_n^{-1}) \quad (n=1,2,\ldots).$$

The necessity of the condition (*). Let a sequence $\mu_1*\mu_2*\ldots*\mu_n$ weakly converge to ν. From Theorem 1 it follows that ν is the Haar measure of the subgroup

(15) $$A_\nu=[\bigcup_{n=1}^{\infty}A_{\mu_n}].$$

Moreover,

$$\varphi_{\mu_1*\mu_2*\ldots*\mu_n}(U)\to\varphi_\nu(U)=0 \quad \text{for} \quad U\in\mathfrak{A}_0(A_\nu).$$

Then formula (14) implies

$$\|\varphi_{\pi_n}(U)\|\leqslant\|\varphi_{\mu_1*\mu_2*\ldots*\mu_n}(U)\|\to 0 \quad \text{for} \quad U\in\mathfrak{A}_0(A_\nu).$$

Hence

$$\varphi_{\pi_n}(U)\to 0=\varphi_\nu(U) \quad \text{for} \quad U\in\mathfrak{A}_0(A_\nu),$$

and consequently the sequence π_n weakly converges to ν. Then in view of Lemma 3 it follows that

(16) $$A_\nu\subset[\bigcup_{n=1}^{\infty}A_{\pi_n}].$$

The formulas (13), (15) and (16) imply the condition (*).

The sufficiency of the condition (*). Suppose that the condition (*) is satisfied. Then, according to (13),

(17) $$[\bigcup_{n=1}^{\infty}A_{\pi_n}]=[\bigcup_{n=1}^{\infty}A_{\mu_n}].$$

Let $U\in\mathfrak{A}([\bigcup_{n=1}^{\infty}A_{\pi_n}])$. If $\|\varphi_{\pi_n}(U)\|=1$ for every n, then, according to Lemma 1, $U\notin\mathfrak{A}_0([A_{\pi_n}])$ for every n. Since $[A_{\pi_{n+1}}]\supset[A_{\pi_n}]$, we have $U\notin\mathfrak{A}_0([\bigcup_{j=1}^{n}A_{\pi_j}])$ for every n, and consequently $U\notin\mathfrak{A}_0([\bigcup_{j=1}^{\infty}A_{\pi_j}])$. Hence,

according to (17), $U\notin\mathfrak{A}_0([\bigcup_{n=1}^{\infty}A_{\mu_n}])$. Therefore, if $U\in\mathfrak{A}_0([\bigcup_{n+1}^{\infty}A_{\mu_n}])$, then there is such an integer k that $\|\varphi_{\pi_k}(U)\|<1$ and by (14)

(18) $$\|\varphi_{\mu_1*\mu_2*\ldots*\mu_k}(U)\|<1.$$

Let $n_1,n_2,\ldots$ be a sequence of integers such that

$$\mu_1*\mu_2*\ldots*\mu_k=\mu_{n_j+1}*\mu_{n_j+2}*\ldots*\mu_{n_j+k} \quad (j=1,2,\ldots),$$

$$n_j+k<n_{j+1} \quad (j=1,2,\ldots).$$

It is easy to see that

$$\|\varphi_{\mu_1*\mu_2*\ldots*\mu_n}(U)\|\leqslant\|\varphi_{\mu_1*\mu_2*\ldots*\mu_k}(U)\|^{n_j} \quad \text{for} \quad n_j+k\leqslant n.$$

Then, according to (18), $\varphi_{\mu_1*\mu_2*\ldots*\mu_n}(U)\to 0$ for $U\in\mathfrak{A}_0([\bigcup_{n=1}^{\infty}A_{\mu_n}])$. The convergence of the sequence of characteristic functions

$$\varphi_{\mu_1*\mu_2*\ldots*\mu_n}(U) \quad \text{for} \quad U\in\mathfrak{A}_0([\bigcup_{n=1}^{\infty}A_{\mu_n}])$$

is equivalent to the weak convergence of the sequence of probability distributions $\mu_1*\mu_2*\ldots*\mu_n$. Theorem 2 is thus proved.

Note added in proof. The results of this paper were obtained, by a different method, at the same time by B. M. Kloss (see Доклады Акад. Наук. СССР 109, No 3 (1956), p. 453-455).

References

[1] A. Dvoretzky and I. Wolfowitz, *Sums of random integers reduced modulo m*, Duke Math. J. 18 (1951), p. 501-507.

[2] B. Gelbaum, G. K. Kalish and J. M. H. Olmsted, *On the embedding of topological semigroups and integral domains*, Proc. Amer. Math. Soc. 2 (1951), p. 807-821.

[3] P. R. Halmos, *Measure theory*, New York 1950.

[4] K. Ito and Y. Kawada, *On the probability distribution on a compact groups*, Proc. Phys.-Math. Soc. of Japan 22 (1940), p. 977-998.

[5] P. Lévy, *L'addition des variables aléatoires définies sur une circonférence*, Bull. Soc. Math. de France 67 (1939), p. 1-41.

[6] A. Prékopa, A. Rényi and K. Urbanik, *О предельном распределении для сумм независимых случайных величин на бикомпактных коммутативных топологических группах*, to appear in Acta Math. Acad. Sci. Hungaricae.

[7] A. Weil, *L'intégration dans les groupes topologiques et ses applications*, Paris 1940.

[8] Н. Н. Воробьев, *Сложение независимых случайных величин на конечных абелевых группах*, Мат. Сбор. 34(1) (1954), p. 89-126.

MATHEMATICAL INSTITUTE OF THE POLISH ACADEMY OF SCIENCES

Reçu par la Rédaction le 3.5.1956