Z. Wahrscheinlichkeitstheorie verw. Geb. 5, 279–295 (1966)

On Infinite Products of Random Elements and Infinite Convolutions of Probability Distributions on Locally Compact Groups

I. Csiszár
Mathematical Institute of the Hungarian Academy of Sciences, Budapest
and
Birkbeck College, University of London

Received June 3, 1965

Introduction

Limiting distributions and stochastic or almost sure convergence of infinite products of independent — not necessarily equally distributed — random group elements as well as interconnections of the various types of convergence have been investigated in the last years by several authors, as Kloss [1], [2]; Tortrat [3]; Loynes [4]; Bártfai [5] and others. The aim of the present paper is to establish general theorems, containing some of the results of the above mentioned papers as special cases and, in particular, to find the proper generalization of the fact that for sums of independent real-valued random variables convergence in law, stochastic convergence and convergence with probability 1 are equivalent. The main results are theorems 3.1 and 3.2 of § 3. They will be obtained by applying considerations concerning topological semigroups in general to the semigroup of probability measures on a group; the basic tool will be the concept of the *tail idempotent*, introduced in § 2. In § 1 a summary of the necessary concepts and notations is given.

§ 1. Preliminaries

In this section we summarise the basic concepts and notations used in the sequel, and mention some simple lemmas, with or without proofs. Details and further references can be found, e. g., in the book of Grenander [6] or Heyer [7].

G will denote a *locally compact topological group;* commutativity will not be assumed. Saying "a *neighbourhood* N" we shall always mean a symmetric neighbourhood of the identity in G (i. e., $N = N^{-1}$).

By *separability* of G we shall mean the existence of a countable base for its topology. A locally compact group is separable if and only if it is σ-compact and has a countable base at the identity.

The smallest σ-algebra containing every compact subset of G will be denoted by $\mathscr{B}$.

A *probability distribution* μ on G is a regular measure on $\mathscr{B}$, i. e., a measure satisfying

$$\mu(A) = \sup_{K \subset A} \mu(K) = \inf_{U \supset A} \mu(U) \qquad (K \text{ compact}, U \text{ open}) \tag{1.1}$$

for every $A \in \mathscr{B}$, such that $\mu(G) = 1$.

The set of all probability distributions on G will be denoted by P_G. By the Riesz representation theorem, there is a one-to-one correspondence between the regular signed measures on $\mathscr{B}$ and the continuous linear functionals on the Banach space $\mathscr{C}_0(G)$; the probability distributions $\mu \in P_G$ correspond to the positive functionals with norm one. Here $\mathscr{C}_0(G)$ denotes the closure (with respect to the sup norm) of $\mathscr{K}(G)$, the linear normed space of all continuous functions on G vanishing outside some compact set. The *vague topology* in the set of all regular signed measures on B is defined as the weak* topology in the conjugate space $\mathscr{C}_0^*(G): \mu_n \to \mu$ if and only if[1]

$$\int f(x)\mu_n(dx) \to \int f(x)\mu(dx) \quad \text{for every } f \in \mathscr{C}_0(G)\,. \tag{1.2}$$

Actually, it would be enough to require (1.2) only for $f \in \mathscr{K}(G)$. Note, that P_G is closed in the vague topology if and only if G is compact. Otherwise, the closure of P_G is the set of all regular non-negative measures on $\mathscr{B}$ with $\mu(G) \leqq 1$, that will be denoted by Q_G.

Lemma 1.1. *If $\nu_n \to \nu_\infty \in Q_G$ then for any compact set K and open set U with $K \subset U \subset G$ we have $\liminf_{n\to\infty} \nu_n(U) \geqq \nu_\infty(K)$; $\limsup_{n\to\infty} \nu_n(K) \leqq \nu_\infty(U)$.*

Corollary. *If $\nu_\infty \in P_G$ then for any $\varepsilon > 0$ there exists a compact set K_ε such that $\nu_n(K_\varepsilon) > 1 - \varepsilon$ for n large enough.*

Proof. By Uryson's lemma, there exists $f \in \mathscr{K}(G)$ with $0 \leqq f(x) \leqq 1$, $f(x) = 1$ for $x \in K$ and $f(x) = 0$ for $x \notin U$. Applying (1.2) to such an f, we obtain the statement of the lemma. The corollary follows from the fact that — by regularity — a K with $\nu_\infty(K) > 1 - \varepsilon$ always exists, and there exists also an $U \supset K$ having compact closure (e. g., $U = KN$, where N is a neighbourhood having a compact closure).

Remark. In the corollary, the restriction "for n large enough" has been made in order to ensure the validity of the statement also for nets (Moore-Smith convergence).

The *convolution* $\mu = \mu_1 * \mu_2$ of two probability distributions (or, more generally, of two measures $\mu_1, \mu_2 \in Q_G$) is defined by

$$\mu(B) = \int \mu_1(Bx^{-1})\mu_2(dx) = \int \mu_2(x^{-1}B)\mu_1(dx) \qquad (B \in \mathscr{B}) \tag{1.3}$$

or, equivalently, by

$$\int f(z)\mu(dz) = \int\int f(xy)\mu_1(dx)\mu_2(dy) \qquad (f \in \mathscr{C}_0(G))\,. \tag{1.4}$$

The convolution of two probability distributions is again a probability distribution; moreover, P_G is a topological semigroup with respect to the convolution (i. e., the mapping $P_G \times P_G \to P_G: (\mu_1, \mu_2) \to \mu_1 * \mu_2$ is vaguely continuous). The latter statement can be deduced from (1.3), using lemma 1.1 (first one has to show, that the function $g^{(n)}(y) = \int f(xy)\mu^{(n)}(dx)$ converges to $g(y) = \int f(xy)\mu(dx)$ uniformly on compact sets, if $\mu^{(n)} \to \mu \in P_G$). Moreover, it can be shown in the same way that convolution is continuous even as $P_G \times Q_G \to Q_G$. Here P_G cannot be replaced by Q_G, i. e., Q_G is no topological semigroup, as the following simple

[1] Here and in the following, for the sake of simplicity, we use the notation $\mu_n \to \mu\,(n \to \infty)$, while it is understood that the same is true for Moore-Smith sequences too.

example shows: Let $\mu_1^{(n)}$ be the point mass at the point n of the real line and $\mu_2^{(n)}$ the point mass at $-n$; then $\mu_1^{(n)} \to 0$, $\mu_2^{(n)} \to 0$, but $\mu_1^{(n)} * \mu_2^{(n)}$ is equal to the point mass at 0, for $n = 1, 2, \ldots$.

If H is a compact subgroup of G, the Haar measure ω_H on H (extended in the obvious way to G) is an idempotent of the semigroup P_G (i. e., $\omega_H * \omega_H = \omega_H$). It is a very important fact that these measures are the only idempotents of P_G (for compact G this was proved first by WENDEL [8] and for the general case by HEYER [7]; actually, HEYER considered only separable groups but this restriction is not essential for the proof). A probability distribution $\mu \in P_G$ is called H-invariant (where H is a compact subgroup of G) if for every $x \in H$ and $B \in \mathscr{B}$ $\mu(Bx) = \mu(B)$. It is easy to see that μ is H-invariant if and only if $\mu = \mu * \omega_H$.

Lemma 1.2. *Let $K \subset G$ be compact and assume $\sup_{x \in G} \mu_1(Kx) \leqq \alpha < 1$; then there exists a compact set $K^* \subset G$ (depending on μ_1) such that for any $\mu_2 \in P_G$ and any $x \in G$*

$$(\mu_1 * \mu_2)(Kx) \leqq \alpha - \frac{\alpha}{2}(1 - \mu_2(K^* x)).$$

Proof. Let $\tilde{K}$ be a compact set such that $\mu_1(\tilde{K}) > 1 - \alpha/2$. Then $u \notin K^{-1}\tilde{K}$ implies $Ku \cap \tilde{K} = \emptyset$ and thus $\mu_1(Ku) \leqq \alpha/2$. Write $\tilde{K}^{-1}K = K^*$; then for $y \notin K^* x$ we have $xy^{-1} \notin K^{*-1} = K^{-1}\tilde{K}$ and hence

$$(\mu_1 * \mu_2)(Kx) = \int_{K^* x} \mu_1(Kxy^{-1})\mu_2(dy) + \int_{G \setminus K^* x} \mu_1(Kxy^{-1})\mu_2(dy) \leqq$$
$$\leqq \alpha\mu_2(K^* x) + \frac{\alpha}{2}(1 - \mu_2(K^* x)) = \alpha - \frac{\alpha}{2}(1 - \mu_2(K^* x)).$$

The support $S(\mu)$ of a probability distribution $\mu \in P_G$ is defined as the set of those elements of G, all neighbourhoods of which have a positive μ measure. $S(\mu)$ is the smallest closed set with $\mu(S(\mu)) = 1$. In fact, if K is an arbitrary compact set and $K \cap S(\mu) = \emptyset$, then each point of K has some neighbourhood of μ-measure 0; as K is compact, already a finite number of these neighbourhoods cover K and thus $\mu(K) = 0$. This means, by the regularity of μ, that $\mu(G \setminus S(\mu)) = 0$, i.e., $\mu(S(\mu)) = 1$.

The following easily verifiable statement is very important:

Lemma 1.3. *If $\mu = \mu_1 * \mu_2$ then $S(\mu)$ is the closure of the set $S(\mu_1)S(\mu_2)$.*

(Here and in the following, a notation like $A\,B$ is to be understood as

$$A\,B = \{xy : x \in A, y \in B\}).$$

A *random element* ξ of G is a $\mathscr{B}$-measurable mapping of some probability space $(\Omega, \mathscr{F}, P)$ into G. When considering random elements, we shall always assume that G is separable. Thus the *distribution of* ξ, i. e., the measure

$$\mu_\xi(B) = P\{\xi^{-1}(B)\} \qquad (B \in \mathscr{B}) \tag{1.5}$$

is automatically a regular measure. If ξ_1 and ξ_2 are *independent random elements*, i. e.,

$$P\{\xi_1^{-1}(B_1) \cap \xi_2^{-1}(B_2)\} = P\{\xi_1^{-1}(B_1)\}P\{\xi_2^{-1}(B_2)\} \quad (B_1, B_2 \in \mathscr{B}), \tag{1.6}$$

then the distribution of $\xi_1\xi_2$ is equal to the convolution $\mu_{\xi_1} * \mu_{\xi_2}$. This fact suggests several simple but useful inequalities, that can be easily proved also directly from (1.3), e. g.,

$$(\mu_1 * \mu_2)(AB) \geqq \mu_1(A)\mu_2(B) \tag{1.7}$$

$$\mu_1(AB^{-1}) \geqq (\mu_1 * \mu_2)(A) + \mu_2(B) - 1 \tag{1.8}$$

$$\mu_2(A^{-1}B) \geqq (\mu_1 * \mu_2)(B) + \mu_1(A) - 1 \qquad \text{etc.} \tag{1.8'}$$

A sequence $\xi_n (n = 1, 2, \ldots)$ of random elements converges stochastically to a random element ξ iff for every neighbourhood N (i. e., symmetric neighbourhood of the identity) we have

$$\lim_{n\to\infty} P\{\xi_n \in N\xi\} = \lim_{n\to\infty} P\{\xi_n \xi^{-1} \in N\} = \lim_{n\to\infty} P\{\xi\xi_n^{-1} \in N\} = 1 \tag{1.9}$$

or equivalently (as for the equivalence of (1.9) and (1.10) see e. g. [*4*])

$$\lim_{n\to\infty} P\{\xi_n \in \xi N\} = \lim_{n\to\infty} P\{\xi^{-1}\xi_n \in N\} = \lim_{n\to\infty} P\{\xi_n^{-1}\xi \in N\} = 1\,. \tag{1.10}$$

If G is separable — what we always assume when dealing with random elements — the Cauchy criterion

$$P\{\xi_n \in N\xi_m\} = P\{\xi_n\xi_m^{-1} \in N\} \geqq 1 - \varepsilon \quad \text{for} \quad n, m \geqq n_0 = n_0(\varepsilon) \tag{1.11}$$

is necessary and sufficient for the stochastic convergence of the sequence ξ_n to some random element ξ (see e. g. [*6*], p. 108).

The concept of the *quotient space* G/H of the topological group G with respect to some compact subgroup H will play an important role in the sequel. G/H is the set of all (left, say) cosets of H, with the quotient topology with respect to the natural mapping π of G onto G/H (i.e., $\pi(x) = xH$). Thus we obtain a homogeneous space, which is Hausdorff, locally compact and separable if G is separable; G/H is a topological group, iff H is a normal subgroup of G. If N ranges over a base at the identity of G, the sets

$$N(x') = \pi(NxH) \tag{1.12}$$

form a base for the neighbourhoods of $x' = \pi(x) \in G/H$. Observe that if $N = N^{-1}$ — what will be always assumed — we have

$$y' \in N(x') \quad \text{iff} \quad x' \in N(y')\,. \tag{1.13}$$

A random element ξ' of G/H can be defined as a $\mathscr{B}'$-measurable mapping from $(\Omega, \mathscr{F}, P)$ into G/H, where $\mathscr{B}'$ denotes the σ-algebra generated by the compact subsets of G/H. The stochastic convergence on G/H can be defined in the same way as on G: $\lim\limits_{n\to\infty} \text{st}\, \xi_n' = \xi'$ iff

$$\lim_{n\to\infty} P\{\xi_n' \in N(\xi')\} = 1 \tag{1.9'}$$

for all neighbourhoods N (of the unity element of G). The Cauchy criterion can be written as

$$P\{\xi_n' \in N(\xi_m')\} \geqq 1 - \varepsilon \qquad \text{for} \quad n, m \geqq n_0 = n_0(\varepsilon) \tag{1.11'}$$

and one can see[2] in a similar way as for G itself that this is necessary and sufficient for the stochastic convergence of ξ_n' to some ξ'.

[2] As a matter of fact, this is a direct consequence of a general theorem due to Doss [*10*], since the homogeneous space G/H can be uniformized in an obvious way, yielding a separable complete uniform space.

Since the natural mapping $\pi: G \to G/H$ is continuous and hence measurable, the map of a random element of ξ is a random element ξ' of G/H; this will be referred to as $\xi \bmod H$. In particular, by saying that a sequence $\xi_1, \xi_2, \ldots$ of random elements of G is (stochastically or with probability one) convergent mod H we shall mean the convergence of the corresponding sequence $\pi(\xi_1), \pi(\xi_2), \ldots$ of random elements of G/H.

We shall be interested in products $\xi_1 \xi_2 \ldots \xi_n$ of independent random elements of G. For such products, as LOYNES [4] has shown, stochastic convergence is equivalent to convergence with probability one. This result holds also for the mod H convergence, as it will be shown later (see the last step of the proof of theorem 3.2).

§ 2. The Concept of Tail Idempotents

The set P_G of all probability distributions on a locally compact group G is a topological semigroup with respect to the vague topology and the convolution as multiplication.

Since P_G is a subset of the unit sphere in the conjugate space of the Banach space $\mathscr{C}_0(G)$, the vague closure of P_G is (vaguely) compact. In particular, the compactness of G implies that of P_G, too.

When dealing with a sequence of powers (in the convolution sense) of a given distribution on a compact group G, the theorem asserting that the sequence of powers of an element of a compact semigroup has a (unique) idempotent accumulation point is very useful (cf. SCHWARZ [12], KLOSS [1]). The following theorem will play a similar role in our further investigations.

Theorem 2.1. *Let X be a compact semigroup, and $x_1, x_2, \ldots$ an arbitrary sequence of elements of X. Write*

$$y_k^l = x_{k+1} x_{k+2} \ldots x_l \qquad (0 \leqq k < l)$$

and assume that $x \in X$ is an accumulation point of the sequence $y_0^l = x_1 x_2 \ldots x_l$.

Then there exist a directed set D and an integer-valued function $n(d)$ on D with the property $n(d) \geqq n_0$ for $d \geqq d_0 = d_0(n_0)$ — i.e., a subnet of the sequence of non-negative integers — such that the limits (in the Moore-Smith sense, cf., e.g. [11])

$$\lim_d y_k^{n(d)} = \tilde{y}_k \qquad (k = 0, 1, 2, \cdots; \tilde{y}_0 = x) \tag{2.1}$$

and

$$\lim_d \tilde{y}_{n(d)} = y_\infty \tag{2.2}$$

exist; y_∞ is necessarily an idempotent, and

$$\tilde{y}_k = \tilde{y}_k y_\infty \qquad (k = 0, 1, 2, \ldots). \tag{2.3}$$

Furthermore if $x' \in X$ is another accumulation point of the sequence $y_0^l = x_1 x_2 \ldots x_l$ and $n(d')$ $(d' \in D')$ is a subnet of the sequence of non-negative integers such that the limits

$$\lim_{d'} y_k^{n(d')} = \tilde{y}'_k \quad (k = 0, 1, 2, \ldots; \tilde{y}'_0 = x') \tag{2.4}$$

$$\lim_{d'} \tilde{y}'_{n(d')} = y'_\infty \tag{2.5}$$

exist, then for any accumulation points y' and y'' of the nets $\tilde{y}'_{n(d)}$ and $\tilde{y}_{n(d')}$ respectively, we have

$$x' = xy'; \; y'y'' = y_\infty . \tag{2.6}$$

If, in particular, X satisfies the first axiom of countability, one can take as $n(d)$ some ordinary sequence of positive integers n_l $(l = 1, 2, \ldots)$.

The assertions of the theorem remain valid also for non-compact topological semigroups X, supposing that X is a subspace (in the topological sense) of a compact Hausdorff space X_0 and some further conditions ensure, that the limiting points in question, surely existing in X_0, belong also to X.

Proof. By assumption, we have $\lim_{d_1} y_0^{n(d_1)} = x$ for some subnet $n(d_1)$ $(d_1 \in D_1)$ of the sequence of non-negative integers. Consider the topological product of countable copies of $X: Y = \mathop{\mathbf{X}}_{l=0}^{\infty} X^l$, $X^l = X$. Tyhonov's theorem implies that Y is compact, thus for some subnet $n(d_2)$ $(d_2 \in D_2)$ of $n(d_1)$, the net $y^{n(d_2)}$ will be convergent, where y^l denotes the vector $y^l = (y_0^l, y_1^l, \ldots)$:

$$\lim_{d_2} y^{n(d_2)} = \tilde{y} = (\tilde{y}_0, \tilde{y}_1, \ldots) \qquad (\tilde{y}_0 = x) . \tag{2.7}$$

Now let $n(d)$ $(d \in D)$ be a subnet of $n(d_2)$ such that $\tilde{y}_{n(d)}$ $(d \in D)$ is convergent; the relation (2.7) obviously remains valid for $n(d)$, too. Thus for this $n(d)$, the relations (2.1) and (2.2) hold. To prove (2.3) and that y_∞ is an idempotent, take in equation

$$y_k^l = y_k^j y_j^l \qquad (k < j < l) \tag{2.8}$$

three times limits, first as $l = n(d)$, then as $j = n(d)$ and at last as $k = n(d)$. By continuity, we obtain

$$\tilde{y}_k = y_k^j \tilde{y}_j; \; \tilde{y}_k = \tilde{y}_k \, y_\infty; \; y_\infty = y_\infty \, y_\infty . \tag{2.9}$$

To prove (2.6), take a subnet $n(d')$ of the sequence of non-negative integers satisfying (2.4).

Choose further subnets $n(d_0)$ $(d_0 \in D_0)$ and $n(d_0')$ $(d_0' \in D_0')$ of $n(d)$ and $n(d')$, respectively, such that

$$\lim_{d_0} \tilde{y}'_{n(d_0)} = y' , \quad \lim_{d_0'} \tilde{y}_{n(d_0')} = y'' . \tag{2.10}$$

Then, taking limits in (2.8) first as $l = n(d')$, then as $j = n(d_0)$, we obtain, putting $k = 0$, $x' = xy'$; further taking limits in (2.8) first as $l = n(d)$, then as $j = n(d'_0)$ and at last as $k = n(d_0)$, we obtain $y_\infty = y'y''$.

If X satisfies the first axiom of countability, one can take everywhere ordinary sequences instead of nets.

The last statement of the theorem is obvious.

Definition. *An idempotent y_∞ associated with the infinite product $x_1 x_2 \ldots$ as in theorem 2.1 will be called a tail idempotent of the infinite product.*

Let us emphasize, that an infinite product may have several tail idempotents. The following theorem indicates a relation between them enabling us in certain cases (e. g., for commutative X) to establish uniqueness.

Theorem 2.2. *Let $n(d)$ and $n(d')$ be two subnets of the sequence of positive integers such that* (2.1), (2.2), (2.4), (2.5) *hold. Then we have*

$$y'_\infty = y'' y_\infty y' ; y' y'' = y_\infty , \tag{2.11}$$

where y' and y'' are the same as in theorem 2.1. *(Under the condition, that all limiting points in question belong to X.)*

Proof. Taking four times limits in the equation

$$y_k^l = y_k^i y_i^j y_j^l , \tag{2.12}$$

first as $l = n(d')$, next as $j = n(d_0)$, then as $i = n(d)$ and at last as $i = n(d'_0)$ — where $n(d_0)$ and $n(d'_0)$ are the same as in the proof of theorem 2.1 — we obtain by continuity equation (2.11).

Observe, that for commutative X the relations (2.11) imply

$$y'_\infty = y_\infty y'' y' = y_\infty y_\infty = y_\infty .$$

Now let us apply the above results to the semigroup $X = P_G$ of probability distributions μ on a locally compact group G. When dealing with this semigroup, we shall write μ and ν instead of x and y respectively.

Theorem 2.1 implies, in particular, that for a *compact* G every infinite convolution $\mu_1 * \mu_2 * \ldots$ has at least one tail idempotent ν_∞, and by theorem 2.2 we know that if G (and thus also P_G) is commutative, the tail idempotent is unique. In the non-commutative case there may exist several tail idempotents, as the following simple example shows: Let H be a non-normal (compact) subgroup of G and x an arbitrary element of G such that $xHx^{-1} \neq H$; let us define $\mu_{3k+1} = \delta_x$, $\mu_{3k+2} = \delta_{x^{-1}}$ and $\mu_{3k} = \omega_H$ (the Haar measure on H). Then both ω_H and $x^{-1}\omega_H x = \omega_{x^{-1}Hx} \neq \omega_H$ are tail idempotents of the infinite convolution $\mu_1 * \mu_2 * \ldots$. Recall that the idempotents of P_G are exactly the Haar measures on the compact subgroups of G.

If G is only locally compact, P_G is not compact, but it is embedded into the (vaguely) compact set Q_G. Thus theorems 2.1 and 2.2 apply for P_G also in case of a locally compact G, if some condition makes sure that the limiting measures in question belong not only to Q_G but also to P_G. Now if

$$\nu_0^n = \mu_1 * \cdots * \mu_n \qquad (\mu_k \in P_G, k = 1, 2, \ldots)$$

possesses at least one accumulation "point" $\mu \in P_G$, we obtain that for some subnet $n(d)$ $(d \in D)$ of the sequence $0, 1, 2, \ldots$ the limits

$$\lim_d \nu_k^{n(d)} = \tilde{\nu}_k \qquad (k = 0, 1, 2, \ldots; \tilde{\nu}_0 = \mu) \tag{2.1'}$$

and

$$\lim_d \tilde{\nu}_{n(d)} = \nu_\infty \tag{2.2'}$$

exist (in Q_G). Further, utilising that the convolution yields a continuous mapping $P_G \times Q_G \to Q_G$, the first equality in (2.9) is also true (replacing y by ν), and thus $\tilde{\nu}_0 = \mu \in P_G$ implies $\tilde{\nu}_j \in P_G$ for every j. Then, by the same argument, also the second equality of (2.9) holds true and $\nu_\infty \in P_G$, assuring the validity of the third equality, too. In particular, substituting in (2.3) $k = 0$, we have

$$\mu = \mu * \nu_\infty \tag{2.13}$$

where ν_∞ is an idempotent of P_G, i.e., the Haar measure on some compact subgroup H of G: $\nu_\infty = \omega_H$. This result contains a theorem of TORTRAT [3], asserting that if the sequence $\mu_1 * \cdots * \mu_n$ is convergent, then the limiting distribution μ satisfies the equation $\mu = \mu * \omega_H$, where ω_H is the Haar measure on some compact subgroup $H \subset G$ and it is an accumulation point of the double sequence

$$\{\nu_k^n;\ 0 \leqq k < n\}.$$

If $\mu' \in P_G$ is another accumulation point of the sequence $\nu_0^n = \mu_1 * \cdots * \mu_n$, we obtain just the same way as above that all the limiting measures $\tilde{\nu}_k', \nu_\infty', \nu', \nu''$ (playing the role of the elements $\tilde{y}_k', y_\infty', y', y''$ in theorem 2.1) belong to P_G. Thus theorem 2.1 applies for $X = P_G$ without any further condition.

We finish this section with a generalization of a theorem of KLOSS [1].

Theorem 2.3. *If the infinite convolution $\mu_1 * \mu_2 * \cdots$ ($\mu_k \in P_G, k = 1, 2, \ldots$) possesses accumulation "points" (measures) in P_G, then any two such accumulation points can be shifted into each other, i.e., if $\mu \in P_G$ and $\mu' \in P_G$ are accumulation points of the sequence $\nu_0^n (n = 1, 2, \ldots)$ then there exists $a \in G$ such that $\mu' = \mu * \delta_a$. (δ_a denotes the measure concentrated at the point a).*

Remark. KLOSS [1] proved this theorem for compact groups G (in which case all accumulation points belong to P_G). He claims that his result implies also the existence of a sequence $a_1, a_2, \ldots$ of elements of G, such that the sequence $\mu_1 * \cdots * \mu_n * \delta_{a_n}$ converges[3], but does not give the proof. If G is separable, the equivalence of the two statements can be easily shown; for the general case, however, I have not been able to find a proof.

For locally compact separable groups, the necessary and sufficient condition of the existence of elements $a_1, a_2, \ldots, a_n, \ldots$ of G such that $\mu_1 * \cdots * \mu_n * \delta_{a_n}$ converges is contained in theorem 3.1.

Proof of theorem 2.3. We have just shown that theorem 2.1 is valid for $X = P_G$ (replacing x by μ and y by ν) if μ and μ' both belong to P_G. Thus, for some ν', $\nu'' \in P_G$ we have

$$\mu' = \mu * \nu' \tag{2.14}$$

and

$$\nu' * \nu'' = \nu_\infty = \omega_H . \tag{2.15}$$

From (2.15) and lemma 1.3 follows that for any $x \in S(\nu'')$ we have

$$H \supset S(\nu') \cdot x , \qquad S(\nu') \subset H x^{-1} \tag{2.16}$$

and hence, writing $x^{-1} = a$,

$$\omega_H * \nu' = \omega_H * \delta_a . \tag{2.17}$$

Finally, (2.14), (2.17) and (2.13) imply $\mu' = \mu * \delta_a$.

---

[3] KLOSS formulated this statement in terms of random elements, and called it "the general principle of convergence".

§ 3. Limit Theorems on Infinite Products of Random Elements

In this section G will denote a locally compact *separable* group. Let $\xi_1, \xi_2, \ldots$ be independent random elements of G with distributions $\mu_1, \mu_2, \ldots$.

The results of the previous section now enable us to characterise the limiting properties of the infinite product $\xi_1 \xi_2 \ldots$ in some detail. Our main results are contained in theorems 3.1 and 3.2.

Theorem 3.1. *Either*

(3.1) $$\sup_{x \in G} P\{\xi_1 \ldots \xi_n \in Kx\} \to 0 \quad as \quad n \to \infty,$$

for every compact set $K \subset G$ *or there exists a sequence* $a_1, a_2, \ldots, a_n, \ldots$ *of elements of* G *such that all the products*

(3.2) $$\xi_{k+1} \ldots \xi_n a_n \qquad (k < n)$$

have limiting distributions as $n \to \infty$ $(k = 0, 1, 2, \ldots)$.

Remarks. 1. This theorem extends B.M. Kloss' "general principle of convergence", referred to in the remark to theorem 2.3, to locally compact (separable) groups. Another slight generalization, that will turn out also to be useful, is that we have a limiting distribution not only for $\xi_1 \ldots \xi_n a_n$ but for all the products (3.2).

2. If we introduce the one-point compactification G_∞ of G, (3.1) can be described equivalently as follows: For any sequence $a_1, a_2, \ldots$ of elements of G we have

(3.1′) $$\lim_{n \to \infty} \operatorname{st} \xi_1 \xi_2 \ldots \xi_n a_n = \infty .$$

Here the stochastic convergence cannot be replaced by convergence with probability one, as e. g., the example of the random walk in one dimension shows.

Proof of theorem 3.1. The proof will be divided into several steps.

a) Write

(3.3) $$\sup_{x \in G} P\{\xi_1 \ldots \xi_n \in Kx\} = \sup_{x \in G} (\mu_1 * \cdots * \mu_n)(Kx) = \alpha_n(K).$$

As

(3.4) $$\mu_1 * \cdots * \mu_n = \nu_0^n = \nu_0^k * \nu_k^n \qquad (0 < k < n),$$

we have for every $x \in G$

(3.5) $$(\mu_1 * \cdots * \mu_n)(Kx) = \int \nu_0^k(Kxy^{-1}) \nu_k^n(dy) \leqq \alpha_k(K) \int \nu_k^n(dy) = \\ = \alpha_k(K) \qquad (1 < k < n).$$

Thus the sequence $\alpha_n(K)$ is non-increasing.

Now define

(3.6) $$\alpha(K) = \lim_{n \to \infty} \alpha_n(K)$$

and

(3.7) $$\alpha_0 = \sup_K \alpha(K) \qquad (K \subset G, K \text{ compact}).$$

First we prove, that either $\alpha_0 = 0$ or $\alpha_0 = 1$. Suppose namely $0 < \alpha_0 < 1$, and choose an α such that

$$\alpha_0 < \alpha < 1\,, \quad 0 < \alpha \frac{1+\alpha}{2} < \alpha_0\,. \tag{3.8}$$

Then, by assumption, for any compact set $K \subset G$ there exists such K that

$$\sup_{x \in G} \nu_0^k(Kx) < \alpha\,. \tag{3.9}$$

Now apply lemma 1.2 for $\nu_0^n = \nu_0^k * \nu_k^n (n > k)$. We obtain that for a suitable compact set $K^* \subset G$ (depending on K)

$$\nu_0^n(Kx) \leqq \alpha - \frac{\alpha}{2}\,(1 - \nu_k^n(K^*x)) \qquad (x \in G, n > k)\,. \tag{3.10}$$

Here for n large enough, ($n \geqq n_0 = n_0(K^*)$, say) we must have

$$\nu_k^n(K^*x) < \alpha\,, \tag{3.11}$$

otherwise, choosing a compact set K' with $\nu_0^k(K') > \frac{\alpha_0 + \alpha}{2\,\alpha}$, the inequality (1.7) would yield

$$\nu_0^n(K'K^*x) = (\nu_0^k * \nu_k^n)\,(K'K^*x) \geqq \nu_0^k(K')\,\nu_k^n\,(K^*x) > \frac{\alpha_0 + \alpha}{2} \qquad (n > k)$$

thus $\alpha(K'K^*) \geqq \frac{\alpha_0 + \alpha}{2} > \alpha_0$, contradicting (3.7).

But (3.10) and (3.11) imply

$$\nu_0^n(Kx) \leqq \alpha - \frac{\alpha}{2}\,(1 - \alpha) = \alpha\,\frac{1+\alpha}{2} \qquad (n \geqq n_0)$$

thus

$$\alpha(K) \leqq \frac{\alpha(1+\alpha)}{2}\,. \tag{3.12}$$

Since $K \subset G$ was an arbitrary compact set, (3.12) means that $\alpha_0 \leqq \frac{\alpha(1+\alpha)}{2}$, contradicting (3.8).

b) We have proved, that either $\alpha_0 = 0$, i. e., (3.1) holds, or $\alpha_0 = 1$, i. e., to every $\alpha < 1$ there exists a compact set $K_\alpha \subset G$ such that

$$\sup_{x \in G} \nu_0^n(K_\alpha x) = \sup_{x \in G}\,(\mu_1 * \cdots * \mu_n)\,(K_\alpha x) > \alpha \qquad (n = 1, 2, \ldots)\,. \tag{3.13}$$

Moreover, these compact sets can be chosen also in such a way that for a fixed sequence $x_1, x_2, \ldots, x_n, \ldots$ of elements of G

$$\nu_0^n(K_\alpha x_n) > \alpha \qquad (n = 1, 2, \ldots) \tag{3.14}$$

for every $\alpha < 1$. Indeed, let us choose to $\alpha = \frac{1}{2}$ a $K_{\frac{1}{2}}$, satisfying (3.13) and then a sequence $x_1, x_2, \ldots$ such that (3.14) be fulfilled for $\alpha = \frac{1}{2}$. Find now to every $\alpha > \frac{1}{2}$ some K_α, satisfying (3.13). Then for some $x_n^\alpha \in G$ we have

$$\nu_0^n(K_\alpha x_n^\alpha) > \alpha\,. \tag{3.15}$$

By (3.15) and $\nu_0^n(K_{\frac{1}{2}} x_n) > \frac{1}{2}$ the sets $K_\alpha x_n^\alpha$ and $K_{\frac{1}{2}} x_n$ cannot be disjoint, and therefore

$$x_n^\alpha \in K_\alpha^{-1} K_{\frac{1}{2}} x_n \quad \text{and} \quad K_\alpha x_n^\alpha \subset K_\alpha K_\alpha^{-1} K_{\frac{1}{2}} x_n\,.$$

This means that replacing each K_α by $K'_\alpha = K_\alpha K_\alpha^{-1} K_{\frac{1}{2}}$, the inequalities (3.14) will be valid for every $\alpha < 1$.

Introduce now

(3.16) $$\hat{\mu}_n = \delta_{x_{n-1}} * \mu_n * \delta_{x_n^{-1}}$$

where the x_n's are the same as in (3.14) and $x_0 = e$. Then

(3.17) $$\hat{\nu}_0^n = \hat{\mu}_1 * \cdots * \hat{\mu}_n = \mu_1 * \cdots * \mu_n * \delta_{x_n^{-1}}$$

and, by (3.14)

(3.18) $$\hat{\nu}_0^n(K_\alpha) = \nu_0^n(K_\alpha x_n) > \alpha$$

for every n and α.

c) (3.18) means that for every α there exists a compact set K such that $\hat{\nu}_0^n(K) > \alpha$; according to (1.8′), the same is true for the whole set $\{\hat{\nu}_k^n : 0 \leqq k < n\}$. This obviously implies that every accumulation "points" (measures) of this set belong to P_G (cf. lemma 1.1) and therefore theorem 2.1 can be applied[4] to P_G (replacing y everywhere by ν). Thus, being P_G now separable, for a suitable sequence $n_1 < n_2 < \cdots < n_k < \ldots$ of positive integers all the limits

(3.19) $$\lim_{l\to\infty} \hat{\nu}_k^{n_l} = \tilde{\hat{\nu}}_k \qquad (k = 0, 1, 2, \ldots)$$

(3.20) $$\lim_{k\to\infty} \tilde{\hat{\nu}}_{n_k} = \hat{\nu}_\infty = \omega_H$$

exist, where the tail idempotent $\hat{\nu}_\infty = \omega_H$ is the Haar measure on some compact subgroup $H \subset G$. Moreover,

(3.21) $$\tilde{\hat{\nu}}_k = \tilde{\hat{\nu}}_k * \omega_H \qquad (k = 0, 1, \ldots).$$

d) From (3.19) and (3.20) follows — utilising also the separability of P_G, implied by the separability of G — that the sequence $n_1 < n_2 < \cdots < n_k < \cdots$ can be chosen also in such a way (replacing the original sequence with a suitable subsequence) that also

(3.22) $$\lim_{k\to\infty} \hat{\nu}_{n_k}^{n_{k+1}} = \omega_H$$

holds.

Then according to lemma 1.1, for any neighbourhood N and any $\varepsilon > 0$ there exists a positive integer k_0 such that

(3.23) $$\hat{\nu}_{n_k}^{n_{k+1}}(NH) > 1 - \varepsilon \qquad (k \geqq k_0).$$

Since $\hat{\nu}_{n_k}^{n_{k+1}}(NH) = \int \hat{\nu}_{n_k}^m (NHx^{-1}) \hat{\nu}_m^{n_{k+1}}(dx)$ $(n_k < m < n_{k+1})$, (3.23) implies that for some $\hat{x} \in G$ we have

$$\hat{\nu}_{n_k}^m (NH\hat{x}^{-1}) > 1 - \varepsilon$$

i. e.,

(3.24) $$(\hat{\nu}_{n_k}^m * \delta_{\hat{x}})(NH) > 1 - \varepsilon.$$

[4] One could use theorem 2.3 as well, but this would not make the proof essentially shorter. We prefer therefore the direct derivation.

Now let $N_1 \supset N_2 \supset \cdots \supset N_i \supset \cdots$ be a base at the identity in G, and $\varepsilon_1 > \varepsilon_2 > \cdots$ $\cdots > \varepsilon_i > \cdots$ a sequence converging to 0. For every i, choose a k_i such that for $k \geqq k_i$ (3.23) holds with N_i and ε_i, where we may and do assume $k_1 < k_2 < \cdots$ $\cdots < k_j < \cdots$. Further, for every m, find an $\hat{x}_m \in G$ such that (3.24) holds with N_i and ε_i, where k and i (depending on m) are defined by the inequalities

$$n_k < m \leqq n_{k+1} \quad \text{and} \quad k_i < k \leqq k_{i+1}$$

(if $m = n_{k+1}$, we put $\hat{x}_m = e$).

e) Now put $a_m = x_m^{-1}\hat{x}_m$. We show that then

$$\lim_{m\to\infty} \mu_{k+1} * \cdots * \mu_n * \delta_{a_m} = \lim_{n\to\infty} \nu_k^n * \delta_{a_m} = \delta\, x_k^{-1} * \tilde{\hat{\nu}}_k \tag{3.25}$$

for every k.

Since, according to (3.16), $\nu_k^m = \delta_{x_k} * \nu_k^m * \delta_{x_m^{-1}}$, (3.25) is equivalent to

$$\lim_{m\to\infty} (\hat{\nu}_k^m * \delta_{\hat{x}_m}) = \tilde{\hat{\nu}}_k \qquad (k = 0, 1, \ldots). \tag{3.26}$$

If (3.26) did not hold, there would exist a positive integer l and a sequence $m_j \to \infty$ such that

$$\lim_{j\to\infty} (\hat{\nu}_l^{m_j} * \delta_{\hat{x}_{m_j}}) = \mu' \neq \tilde{\hat{\nu}}_l, \tag{3.27}$$

(where, as was pointed out in step c), necessarily $\mu' \in P_G$).

Hence, defining k_j by the inequality $n_{k_j} < m_j \leqq n_{k_{j+1}}$, and choosing a subsequence of m_j (which, for the sake of simplicity, will be denoted again by m_j) such that

$$\lim_{j\to\infty} \hat{\nu}_{n_{k_j}}^{m_j} * \delta_{\hat{x}_{m_j}} = \nu'$$

exists (and then necessarily belongs to P_G), we would obtain

$$\mu' = \lim_{j\to\infty} \hat{\nu}_l^{n_{k_j}} * \nu_{n_{k_j}}^{m_j} * \delta_{\hat{x}_{m_j}} = \tilde{\hat{\nu}}_l * \nu'. \tag{3.28}$$

But from the construction in step d) and lemma 1.1 clearly follows that $\nu'(NH) = 1$ for every neighbourhood N, i. e., $S(\nu') \subset H$ and thus $\omega_H * \nu' = \omega_H$.

Hence from (3.21) and (3.28) we obtain $\mu' = \tilde{\hat{\nu}}_l$, contradicting (3.27).

Thus (3.25) is proved and the proof of theorem 3.1 is complete.

In the proof of theorem 3.2 we shall need the following simple lemma:

Lemma 3.1. *Let N be a neighbourhood of the identity in an arbitrary topological group G, and $K \subset G$ a compact set. Then there exists a neighbourhood N^* such that $xN^* \subset Nx$ for every $x \in K$.*

Proof. Take a neighbourhood N' with $N'^3 \subset N$ and select a finite subset $(x_1, x_2, \ldots, x_n)$ of G such that

$$\bigcup_{i=1}^{n} N'x_i \supset K. \quad \text{Put} \quad N^* = \bigcap_{i=1}^{n} x_i^{-1}N'x_i.$$

Then for every $x \in K$ we have $x \in N'x_i$ (and thus, by symmetry, also $x_i \in N'x$) for some $1 \leqq i \leqq n$; hence, utilizing $x_iN^* \subset N'x_i$, we obtain

$$xN^* \subset N'x_iN^* \subset N'^2x_i \subset N'^3x \subset Nx.$$

Theorem 3.2. *Let $\xi_1, \xi_2, \ldots$ be a sequence of independent random elements of a separable locally compact group G. Assume that the product $\xi_1 \xi_2 \ldots \xi_n$, as well as all the products $\xi_{k+1}\xi_{k+2} \ldots \xi_n (k \geqq 1)$ have limiting distributions as $n \to \infty$.*

Then there exists a unique compact subgroup $H \subset G$ such that all the above limiting distributions are H-uniform and the product $\xi_1 \xi_2 \ldots \xi_n$ converges mod H *with probability* 1.

Proof. Denote the distribution of ξ_k by $\mu_k (k = 1, 2, \ldots)$ and write

(3.29) $$\nu_k^n = \mu_{k+1} * \cdots * \mu_n .$$

Then by assumption all the sequences $\nu_k^n (k = 0, 1, 2, \ldots)$ are convergent in P_G, as $n \to \infty$:

(3.30) $$\lim_{n \to \infty} \nu_k^n = \tilde{\nu}_k \in P_G \qquad (k = 0, 1, 2, \ldots) .$$

First we show, that this implies

(3.31) $$\lim_{k \to \infty} \tilde{\nu}_k = \nu_\infty = \omega_H$$

where ω_H is the Haar measure on some compact subgroup $H \subset G$. (And thus the tail idempotent is unique.) In fact, for any two accumulation "points" (measures) ν' and ν'' of the sequence $\tilde{\nu}_k$, taking limits in

(3.32) $$\tilde{\nu}_k = \nu_k^n * \tilde{\nu}_n \qquad (k < n)$$

first for $\tilde{\nu}_{n_i} \to \nu''$ then for $\tilde{\nu}_{k_j} \to \nu'$, we obtain $\nu' = \nu' * \nu''$.

In particular, for any accumulation point ν' of $\{\tilde{\nu}_k\}$, we have

(3.33) $$\omega_H = \omega_H * \nu' , \quad \nu' = \nu' * \omega_H$$

where ω_H denotes a tail idempotent of $\mu_1 * \mu_2 * \ldots$, being the Haar measure on some compact subgroup $H \subset G$. Now the first equality in (3.33) yields $S(\nu') \subset H$ and then from the second one follows $\nu' = \omega_H$, proving (3.31).

It is clear, that all the $\tilde{\nu}_k$'s are H-uniform, as $\tilde{\nu}_k = \tilde{\nu}_k * \omega_H$ $(k = 0, 1, 2, \ldots)$. It is also clear that if all $\tilde{\nu}_k$'s are H'-uniform ($H' \neq H$) then $\nu_\infty = \omega_H$ is also H'-uniform and thus $H' \subset H$.

We are going to prove that the sequence $\xi_1 \xi_2 \ldots \xi_n$ converges with probability one mod H. From this and the preceding paragraph also the uniqueness part of the theorem will follow, since if $\xi_1 \xi_2 \ldots \xi_n$ converges (at least stochastically) mod H', then the support of the tail idempotent must be contained in H', i. e., $H \subset H'$.

First we prove the stochastic convergence mod H. To this end, according to (1.11′), we have to show, that for every neighbourhood N of the identity in G and every $\varepsilon > 0$ we have

(3.34) $$P\{\pi(\xi_1 \ldots \xi_n) \in N(\pi(\xi_1 \ldots \xi_m))\} > 1 - \varepsilon \quad \text{for} \quad n, m \geqq n_0 = n_0(N, \varepsilon)$$

(where π denotes the natural mapping $G \to G/H$, i. e., $\pi(x) = xH$). (3.34) can be written also as

(3.34a) $$P\{\xi_1 \ldots \xi_n \in N \xi_1 \ldots \xi_m H\} > 1 - \varepsilon \qquad (n, m \geqq n_0) .$$

Let N be an arbitrary (symmetric) neighbourhood of the identity in G, and $\varepsilon > 0$. Find a compact set K such that

(3.35) $$P\{\xi_1 \ldots \xi_n \in K\} = (\mu_1 * \cdots * \mu_n)(K) > 1 - \varepsilon/2$$

for every n; the existence of such K follows from the corollary of lemma 1.1. Take further a neighbourhood N^* of the identity in G such that $xN^* \subset Nx$ for every $x \in K$; this is possible by lemma 3.1. At last, let N_1 be a neighbourhood with $N_1^2 \subset N^*$ and $N_1^* \subset N_1$ such a neighbourhood that $xN_1^* \subset N_1 x$ for every $x \in H$. Then we have

$$N_1^* H N_1^* \subset N_1^* N_1 H \subset N^* H . \tag{3.36}$$

As $\tilde{\nu}_k \to \omega_H$, by lemma 1.1 we can choose such k_0 that

$$\hat{\nu}_k(N_1^* H) > 1 - \frac{\varepsilon}{4} \qquad (k \geqq k_0) . \tag{3.37}$$

The (3.36), (3.37) and (1.8) imply

$$\nu_k^n(N^* H) \geqq \nu_k^n(N_1^* H N_1^*) > 1 - \frac{\varepsilon}{2} \qquad (n > k \geqq k_0) . \tag{3.38}$$

From (3.38) and (3.35) now follows

$$\begin{aligned} &P\{\xi_1 \dots \xi_n \in N \xi_1 \dots \xi_m H\} \geqq \\ &\geqq P\{\xi_1 \dots \xi_n \in \xi_1 \dots \xi_m N^* H; \xi_1 \dots \xi_m \in K\} \geqq \\ &\geqq P\{\xi_{m+1} \dots \xi_n \in N^* H\} - P\{\xi_1 \dots \xi_m \notin K\} = \\ &= \mu_m^n(N^* H) - P\{\xi_1 \dots \xi_m \notin K\} > 1 - \varepsilon \qquad (n > m \geqq k_0) . \end{aligned}$$

Thus (3.34) is proved (we need no further proof for the case $m > n$, since by the symmetry of N the event $\xi_1 \dots \xi_n H \subset N \xi_1 \dots \xi_m H$ is identical with $\xi_1 \dots \xi_m H \subset N \xi_1 \dots \xi_n H$).

We remark that if G is Abelian or anyway if H happens to be a normal subgroup of G, the stochastic convergence mod H of $\xi_1 \dots \xi_n$ implies convergence with probability 1 mod H, by a theorem of LOYNES [4] (applied to the group G/H). In the general case we can complete the proof by a reasoning similar to the one of Loynes, although with more computational difficulties. Let us denote the stochastic limit of the sequence $\pi(\xi_1 \dots \xi_n)$ in G/H by η'.

Let N be an arbitrary neighbourhood of the identity in G, $K \subset G$ a compact set, N_1 a neighbourhood with $N_1^2 \subset N$ and N_1^* a neighbourhood (depending on K) such that $xN_1^* \subset N_1 x$ for every $x \in K$ (cf. lemma 3.1). At last, take neighbourhoods N_2 and N_2^* such that the closure of N_2 is contained in N_1 and $N_2^* H N_2^* \subset N_1^* H$.

According to (3.34) and (3.38) we may select a sequence

$$n_1 < n_2 < \dots < n_k < \dots$$

such that for every k

$$P\{\pi(\xi_1 \dots \xi_n) \notin N_2(\pi(\xi_1 \dots \xi_{n_k}))\} < 2^{-k} \quad \text{for} \quad n \geqq n_k \tag{3.39}$$

and

$$P\{\xi_{n_k+1} \dots \xi_{n_{k+1}} \notin N_2^* H\} < 2^{-k} . \tag{3.40}$$

From (3.39), in particular, (utilising that N_1 contains the closure of N_2) follows

$$P\{\eta' \notin N_1(\pi(\xi_1 \dots \xi_{n_k}))\} < 2^{-k} \tag{3.41}$$

i. e.

$$P\{\pi(\xi_1 \dots \xi_{n_k}) \notin N_1(\eta')\} < 2^{-k} . \tag{3.42}$$

This means that with probability 1 we have

$$\pi(\xi_1 \dots \xi_{n_k}) \in N_1(\eta') \tag{3.43}$$

for all but a finite number of k's.

Now we make use of a lemma of Loève, asserting that if A_k and B_k $(k = 1, \dots, m)$ are arbitrary random events such that for each fixed k A_k and B_k are independent, then

$$P(\bigcup_{i=1}^{m} A_i B_i) \geqq \inf_{1 \leqq i \leqq m} P(B_i) \cdot P\{\bigcup_{i=1}^{m} A_i\}. \tag{3.44}$$

Applying (3.44) for the events

$$A_i = \{\xi_{n_k+1} \dots \xi_{n_k+i} \notin N_1^* H\}, \quad B_i = \{\xi_{n_k+i+1} \dots \xi_{n_{k+1}} \in N_2^* H\}$$

we obtain for every k — utilising $N_2^* H N_2^* \subset N_1^* H$ —

$$\begin{aligned} &P\{\xi_{n_k+1} \dots \xi_{n_{k+1}} \notin N_2^* H\} \geqq \\ &\geqq \inf_{n_k < n \leqq n_{k+1}} P\{\xi_{n+1} \dots \xi_{n_{k+1}} \in N_2^* H\}\, P\{\bigcup_{n=n_k+1}^{n_{k+1}} \xi_{n_{k+1}} \dots \xi_n \notin N_1^* H\}. \end{aligned} \tag{3.45}$$

As for k large enough we obviously have

$\inf_{n_k < n \leqq n_{k+1}} P\{\xi_{n+1} \dots \xi_{n_{k+1}} \in N_2^* H\} \geqq \frac{1}{2}$ (say), (3.45) and (3.40) imply

$$\sum_{k=1}^{\infty} P\{\bigcup_{n=n_k+1}^{n_{k+1}} \xi_{n_{k+1}} \dots \xi_n \notin N_1^* H\} < \infty.$$

Thus, by the Borel-Cantelli lemma, we have with probability one

$$\xi_{n_k+1} \dots \xi_n \in N_1^* H \qquad (n = n_k + 1, \dots, n_{k+1}) \tag{3.46}$$

except for a finite number of k's.

Now from (3.43) and (3.46) we obtain — utilising that $x N_1^* \subset N_1 x$ for $x \in K$ — that almost surely

$$\xi_1 \dots \xi_n \in \pi^{-1}(N_1(\eta')) \cdot N_1^* H = N_1 \pi^{-1}(\eta') N_1^* H \subset N_1^2 \pi^{-1}(\eta') H \subset N \pi^{-1}(\eta')$$

i. e.,

$$\pi(\xi_1 \dots \xi_n) \in N(\eta') \tag{3.47}$$

except for a finite number of n's, if only $\pi^{-1}(\eta') \in K$. As this is true for any compact $K \subset G$, and G is σ-compact, the last restriction obviously can be omitted. Now if N ranges over a countable base at the identity of G, we obtain that with probability one $\pi(\xi_1 \dots \xi_n) \to \eta'$, and the proof is complete.

If G is the additive group of all real numbers, the only compact subgroup of G is $H = \{0\}$. Further in this case convergence in law of $\xi_1 + \xi_2 + \cdots + \xi_n$ implies the same also for $\xi_{k+1} + \xi_{k+2} + \cdots + \xi_n$, for all k. Thus theorem 3.2 reduces to the well-known theorem that for sums $\xi_1 + \xi_2 + \cdots + \xi_n$ of independent random variables convergence in law and convergence with probability one are equivalent. For an arbitrary group G, however, there are some trivial cases when $\lim_{n \to \infty} \mu_1 * \mu_2 * \cdots * \mu_n$ surely exists; e. g., if μ_1 is the Haar measure ω_H on some compact subgroup $H \subset G$, and $S(\mu_n) \subset H$ for every n, then $\mu_1 * \mu_2 * \cdots * \mu_n = \mu_1 = \omega_H$, not depending on the μ_n's at all. In order to exclude such trivial cases, it is quite natural to require the convergence of distributions $\mu_{k+1} * \mu_{k+2} * * \cdots * \mu_n (n \to \infty)$ for all k.

Thus theorem 3.2 is the proper generalization of the above mentioned theorem we were looking for.

In view of theorem 3.2 the statement of theorem 3.1 can be strengthened as follows:

If for a sequence $\xi_1, \xi_2, \ldots$ of independent random elements of a locally compact separable group G there exists a sequence $a_1, a_2, \ldots$ of (non-random) elements of G such that for some compact set K $\lim_{n\to\infty} P\{\xi_1 \xi_2 \ldots \xi_n a_n \in K\} > 0$ then this sequence can be chosen also in such a way that for some compact subgroup $H \subset G$ the sequence $\xi_1 \xi_2 \ldots \xi_n a_n$ is convergent mod H *with probability one and it has a H-uniform limiting distribution.*

In particular, if G is compact and $\xi_1 \xi_2 \ldots \xi_n a_n$ does not converge mod H with probability one for any choice of the a_n's and of the compact sub-group $H \neq G$, then the H in the above statement must be equal to G; hence follows that in this case $\xi_1 \xi_2 \ldots \xi_n$ as well as any $\xi_{k+1} \xi_{k+2} \ldots \xi_n$ have uniform limiting distribution ω_G as $n \to \infty$. In fact, since for some sequence $a_1, a_2, \ldots, a_n$ all $\xi_{k+1} \xi_{k+2} \ldots \xi_n a_n$ have some limiting distribution as $n \to \infty$, which has to be G-uniform, i. e., uniform, the same remains true also when deleting the a_n's. From the uniqueness statement of theorem 3.2 obviously follows that the above condition is also necessary for the limiting uniformity in the described sense. This result generalizes a theorem of Bártfai [5], obtained for compact Abelian groups[5], to the non-commutative case.

Addendum

After having submitted this paper for publication, I was informed by Prof. A. Tortrat that some of the results presented here are closely related to some recent results of his (see A. Tortrat: "Lois de probabilité sur un espace topologique completement régulier et produits infinis a termes indépendants dans un groupe topologique", Ann. Inst. Henri Poincaré, 1, 217—237 (1965) and „Lois et convolutions denombrables dans un groupe topologique", Seminar de calcul des probabilités, Séance du Mardi 2, février 1965). In view of this the statements of our theorems 2.3 and 3.1 are not new (except for the slight generalization, referred to in remark 1 to Theorem 3.1), but the proofs — based on the concept of the tail idempotent — seem to be simpler[6]. Our second main theorem 3.2 seems to be new.

(Added September 17, 1965)

References

[1] Kloss, B. M.: Probability distributions on bicompact topological groups (in Russian). Teor. Veroyatn. Primen. **4**, 255—290 (1959).

[2] — Limiting distributions on bicompact abelian groups (in Russian). Teor. Veroyatn. Primen. **6**, 392—421 (1961).

[5] In the paper of Bártfai separability was not assumed. Since convergence in P_G implies and is implied by the convergence of projections of the corresponding measures to all separable factor-groups of G, it is easy to see, that our above result also remains valid for non-separable groups, too.

[6] It should be noted that Tortrat considered the problem in a somewhat more general context, involving not locally compact groups, too. The method presented here can be extended to this more general case with almost no modifications.

[3] TORTRAT, A.: Lois tendues, convergence en probabilité et équation $P * P' = P$. C. R. Acad. Sci. Paris **258**, 3813—3816 (1964).
[4] LOYNES, R. M.: Products of independent random elements in a topological group. Z. Wahrscheinlichkeitstheorie verw. Geb. **1**, 446—455 (1963).
[5] BÁRTFAI, P.: Grenzverteilungssätze auf der Kreisperipherie und auf kompakten Abelschen Gruppen. Studia Math. Acad. Sci. Hung. **1**, (1966) to appear.
[6] GRENANDER, U.: Probabilities on algebraic structures. Stockholm: Almquist and Wiksell 1963.
[7] HEYER, H.: Untersuchungen zur Theorie der Wahrscheinlichkeitsverteilungen auf lokalkompakten Gruppen. Dissertation, Hamburg, 1963.
[8] WENDEL, J. C.: Haar measure and the semigroup of measures on a compact group. Proc. Amer. math. Soc. **5**, 923—929 (1954).
[9] PONTRIAGIN, J. S.: Topological groups (translation from Russian). Princeton: Princeton Univ. Press 1939.
[10] DOSS, S.: Sur la convergence stochastique dans les espaces uniformes. An. sci. École norm. sup., III. Sér. **71**, 87—100 (1954).
[11] KELLEY, J.: General topology. Princeton: Van Nostrand 1955.
[12] SCHWARZ, S.: To the theory of compact Hausdorff semigroups (in Russian). Czechoslov. math. J. **5**, 2—23 (1955).

Mathematical Institute
of the Hungarian Academy
of Sciences
Budapest, Hungary