Z. Wahrscheinlichkeitstheorie 4, 78—102 (1965)

Probability Theory on Discrete Semigroups

By

PER MARTIN-LÖF

0. Introduction and summary

The study of the classical problems of probability theory, such as the determination of the distribution of the sum of a large number of independent random variables, on more general algebraic structures than the real line, attracts a rapidly increasing interest. A survey of the field, including the history and practical background of it, is given in the treatise by GRENANDER (1963).

This paper deals with discrete semigroups. It is the first paper abstaining from compactness assumptions in the non group case. The main tool is the theory of Markov chains and that is why the study is confined to the discrete case. It should be noted that, unless we put restrictions on the algebraic structure, assuming it to be a group for example, we do not have access (at present) to any sort of Fourier analysis, which is the basic tool in the classical studies.

In the following a survey of the paper is given. Section 1 deals with the measure theoretic (this sounds a bit too solemn since the topology is discrete) preliminaries and section 2 with the algebraic ones. This leads to the definition of the basic concept, the convolution operation, in section 3. The connection between the composition of independent random variables and the theory of Markov chains is established in section 4, permitting a complete description of the behaviour of the convolution iterates of an arbitrary probability distribution in section 5. We return to this problem in section 8, where necessary and sufficient convergence conditions are given. This is made possible by the study of the idempotent probability measures and the structure of a group of probability measures in section 6 and 7 respectively. Infinitely divisible distributions are the subject of section 9, where it is shown that the compound Poisson distributions are the only infinitely divisible ones on a finite semigroup. A discussion of the homogeneous processes is found in section 10. The bibliography should be reasonably complete up to and including 1963.

I am indebted to ULF GRENANDER, my teacher, for rousing my interest in the subject, and I wish to thank him for his encouragement. Also, I have benefited from discussions with ŠTEFAN SCHWARZ. HENRIK ERIKSSON read the manuscript, and his comments resulted in a number of improvements.

1. Measures on a discrete space

Let S be a set (possibly non denumerable) endowed with the discrete topology. By a *non negative measure* on S we understand an extended, non negative, countably additive set function, defined for all subsets of S, assuming finite values on finite sets and regular in the sense that the measure of an arbitrary set equals the supremum of the measures of all finite sets contained in it. See HALMOS (1950) for the general measure theoretic definitions.

A *probability measure* (or *probability distribution*) is a non negative measure such that the measure of the whole space equals one.

A non negative measure is said to be finite if it is not extended, i. e. if the measure of the whole space is finite. By a *finite (signed) measure* we mean the difference between two non negative finite measures.

The measure which attributes to each set the number of points in it is called *counting measure*. If E is a finite subset of S, the probability measure that gives equal masses to the points of E and mass zero to all other points will be referred to as the *uniform distribution* over E.

The regularity assumption we have imposed guarantees the one to one correspondence between a measure μ and its *density function* (with respect to counting measure) m given by

$$m(a) = \mu(\{a\}),$$

for all points a in S, and, for all subsets A of S,

$$\mu(A) = \sum_A {}_a\, m(a).$$

We shall consistently denote measures by small Greek letters and their densities by the corresponding small Latin ones.

It is trivial to give an example of an extended countably additive set function μ (defined for all subsets of the space) which is not regular. Let, for example, S be non denumerable and define $\mu(A)$ to be 0 when A is countable and $+\infty$ otherwise. However, no example seems to have been given of a finite valued μ with the required properties. We know even that when S is the real line no such function exists (assuming the continuum hypothesis). A discussion of the problem and further references are found in ULAM (1960).

Let μ be a finite measure. Then

$$\|\mu\| = \sum_S {}_a\, |m(a)|$$

is defined to be the *norm* of μ. It is evident that the sum converges precisely when μ is finite. For a non negative measure the norm is simply the measure of the whole space.

In the following the concept of the *support* of a measure μ, to be denoted $C(\mu)$, will be of fundamental importance. By definition

$$C(\mu) = \{a;\, m(a) \neq 0\}.$$

The support of a finite measure is denumerable.

Several topologies in the set of measures will concern us. In the *vague topology* a subbase neighbourhood of the (not necessarily finite) measure μ is given by

$$\{\nu;\, |\sum_S {}_a\, f(a)\,(n(a) - m(a))| < \varepsilon\},$$

where f vanishes outside a finite set. This is evidently the same as the pointwise topology for the densities.

The set of all finite measures may be viewed as the dual of the Banach space of all functions tending to zero at infinity. Thus, by letting the function f above belong to this space instead, we obtain a subbase neighbourhood of μ in the *weak star topology*. For a uniformly bounded (in norm) set of measures, in particular for the set of probability measures, the weak star topology evidently coincides with the vague topology. The set of all finite measures with norm $\leqq 1$ is compact as well as sequentially compact in the weak star topology.

The topology in the set of finite measures determined by the norm will be referred to as the *norm topology.*

The *weak topology* in the set of probability measures is determined by the requirement that a subbase neighbourhood of π is of the form

$$\{\varrho; \,|\sum_S {}_a f(a)\,(r(a) - p(a))\,| < \varepsilon\},$$

where f is a bounded function. This is the topology with which probability theory is mainly concerned. It is a consequence of the discreteness of the space that the weak topology coincides with the norm topology. We note that weak convergence of a sequence of probability measures is equivalent to vague convergence together with the fact that the limit measure has total mass one. A set of probability measures is conditionally compact in the weak topology if and only if, for every $\varepsilon > 0$, there exists a finite set C such that $\pi(C) > 1 - \varepsilon$ for all π in the set, i. e. if and only if the tails of the distributions are uniformly small.

It is evident what should be the definition of a *random variable* taking values in S, namely a function defined on a probability space and with range in S, satisfying the condition that the inverse image of every subset of S should be measurable (i. e. belong to the Borel algebra implicit in the definition of a probability space).

2. Semigroups

In this section the algebraic definitions and theorems, basic to the rest of this paper, are outlined. Everything is found in the monographs by LJAPIN (1960) and CLIFFORD and PRESTON (1961).

A *groupoid* is a set S together with a binary operation (i. e. a function from the Cartesian product $S \times S$ to S), which in the following will be denoted multiplicatively. If A and B are two subsets of S, $A\,B$ denotes the set of all products ab with a in A and b in B. We shall write $a\,B$ and $A\,b$ instead of $\{a\}\,B$ and $A\,\{b\}$.

A groupoid is a *semigroup* provided the multiplication is associative, i. e. if $a(bc) = (ab)c$ for all a, b and c in S. In a semigroup the product of an arbitrary finite number of elements and the power a^n of an element a are unambiguously defined.

A subset T of a groupoid is called a *subgroupoid* if $T^2 = T\,T \subset T$. If A is an arbitrary subset of a groupoid, the intersection of all subgroupoids containing A is called the *subgroupoid generated* by A. It sonsists of all finite products of elements of A. If S is a semigroup the subsemigroup generated by A is expressible as $\bigcup_{n=1}^{\infty} A^n$.

An element 1 of a groupoid S is called an *identity element if* $1a = a1 = a$ for all a in S. There can be at most one identity element. If there is none we can adjoin one and consider the enlarged set $S \cup \{1\}$ and extend the multiplication to it by defining $11 = 1$ and $1a = a1 = a$ for all a in S. If S is a semigroup, so is $S \cup \{1\}$.

An element e of a groupoid is called *idempotent* if $e^2 = e$. In particular, an identity element is idempotent.

By a *right (left) ideal* of a groupoid S we mean a subset $R(L)$ such that $RS \subset R(SL \subset L)$. A *twosided ideal* is a subset which is a right as well as a left ideal. A groupoid is *right (left) simple* if there is no right (left) ideal properly contained in it and *simple* provided S is itself the only twosided ideal. A right (left, twosided) ideal is *minimal* if there is no right (left, twosided) ideal properly contained in it. Two different minimal right (left) ideals are disjoint. There can be at most one minimal twosided ideal.

Let A be an arbitrary subset of a groupoid S. The intersection of all right (left, twosided) ideals containing A is called the *right (left, twosided) ideal generated by* A. If S is a semigroup it equals $A \cup AS(A \cup SA,\ A \cup AS \cup SA \cup SAS)$. In case $A = \{a\}$ we speak of the *principal right (left, twosided) ideal* generated by a.

Let S and T be groupoids. A mapping h of S into T is called a *homomorphism (antihomomorphism)* if $h(ab) = h(a)h(b)$ $(h(ab) = h(b)h(a))$ for all a and b in S. If, in addition, it is one to one, we speak of an *isomorphism (antiisomorphism)*.

Let S be a semigroup with identity element 1. The set of all elements a to which there exists an element a^{-1} such that $aa^{-1} = a^{-1}a = 1$ is a subgroup of S (containing 1). It is *maximal* in the sense that it contains any subgroup which meets it.

An element a of a semigroup S is called *regular* if $axa = a$ for some x in S.

The following equivalence relations, called Green relations after their discoverer, are important. Two elements a and b of a semigroup S are said to be *right (left) ideal equivalent* if they generate the same principal right (left) ideal. The equivalence classes modulo this equivalence relation are called *right (left) ideal layers*. Two different elements a and b are right (left) ideal equivalent if and only if each one of them is divisible on the left (right) by the other one, i. e. $b = ax(b = xa)$ and $a = by(a = yb)$ for some x and y in S.

A simple semigroup S is *completely simple* if it contains at least one minimal right ideal and at least one minimal left ideal. Then

$$S = \bigcup_I {}_i R_i = \bigcup_J {}_j L_j .$$

Here $R_i(L_j)$ runs through all minimal right (left) ideals as $i(j)$ varies over the index set $I(J)$. For every i in I and j in J, $L_j R_i = S$ and $R_i L_j = R_i \cap L_j = G_{i,j}$ is a subgroup of S, all these groups being isomorphic. Moreover, $G_{i,j} G_{k,l} = G_{i,l}$ for every i, k in I and j, l in J. We refer to the decomposition

$$S = \bigcup_{I \times J} {}_{i,j}\, G_{i,j}$$

as the group decomposition of S.

A fundamental theorem of Rees tells us that a completely simple semigroup is isomorphic to a so called *Rees matrix semigroup*. Such a semigroup is of the form $G \times I \times J$, where G is a group and I and J arbitrary index sets, multiplication being defined in terms of a $J \times I$ matrix $(g_{j,i})$ of elements in G, called the sandwich matrix, by the relation

$$(g, i, j)\,(h, k, l) = (g\, g_{j,k}\, h, i, l)$$

for all g, h in G, i, k in I and j, l in J.

Let S be a semigroup. To each element c in S we make correspond the $S \times S$ matrix $T(c)$ whose a, b th element equals 1 if $ac = b\,(ca = b)$ and 0 otherwise. It is easily verified that $T(cd) = T(c)\,T(d)$ $(T(cd) = T(d)\,T(c))$, i. e. the mapping constructed is a homomorphism (antihomomorphism). It is a *representation (antirepresentation)* of the semigroup S by matrices over the field of real numbers.

3. Convolution of measures

Let S be a groupoid and μ and ν two finite measures on S. The *convolution* of μ and ν is the finite measure $\mu\nu$ defined by

$$(\mu\nu)(A) = (\mu \times \nu)(\{(b, c);\, b\,c \in A\})$$

for all $A \subset S$. The density of $\lambda = \mu\nu$ is

$$l(a) = \sum_{\{(b,c);\, b\,c = a\}} m(b)\, n(c) = \sum_{S} m(b)\, \nu(\{c;\, b\,c = a\}) = \sum_{S} \mu(\{b;\, b\,c = a\})\, n(c)\,.$$

If x and y are two random variables taking their values in S, so is their product xy, provided their probability distributions, say π and ϱ, are regular and the probability measure over the basic probability space is complete. The probabilistic importance of the concept of convolution is that if x and y are idependent, the probability distribution of xy is precisely the convolution $\pi\varrho$ of π and ϱ.

It is immediately verified that $\|\mu\nu\| \leqq \|\mu\|\,\|\nu\|$ and that $C(\mu\nu) \subset C(\mu)C(\nu)$. Equality holds in both cases provided μ and ν are non negative.

The convolution operation is associative or commutative if and only if the multiplication in S is. This follows immediately from the fact that the mapping which makes correspond to each point a of S the unit mass placed at a is an isomorphism from S into the set of probability measures on S.

If S is a semigroup it follows from what we have just said that the convolution of an arbitrary finite number of finite measures and the convolution power μ^n of a finite measure μ are unambiguously defined.

An important question is whether the convolution $\mu\nu$ depends in some way continuously on μ and ν. From the inequality $\|\mu\nu\| \leqq \|\mu\|\,\|\nu\|$ it follows that, in the norm topology, $\mu\nu$ is a jointly continuous function of μ and ν. Joint continuity holds also for probability measures in the weak topology (since, as we have remarked, it coincides with the norm topology).

In the vague or weak star topology, however, the situation is much worse. To exemplify this, let S be the set of non negative integers and define the multiplication by $ab = 0$ for all a and b in S. This makes S a commutative semigroup. Let π_n be the unit mass placed at n and ϱ a fixed probability measure. Then $\pi_n \to 0$ vaguely as $n \to \infty$. However, $\pi_n\varrho = \varrho\pi_n$ is the unit mass placed at 0 for every n. This shows that the convolution product may not be continuous in the vague or weak star topology even if one factor is held fixed.

The convolution $\mu\nu$ of two finite measures μ and ν depends continuously on $\nu(\mu)$ for arbitrary fixed $\mu(\nu)$ in the weak star topology if and only if the equation $ax = c$ $(xa = c)$ has at most finitely many solutions for every a and c. In particular, the mentioned continuity holds on a group.

To prove the sufficiency, let f be an arbitrary function tending to zero at infinity. Since

$$\sum_S {}_c f(c)\,(\mu\nu)(\{c\}) = \sum_S {}_b \Big(\sum_S {}_a f(a\,b)\,m(a)\Big)\,n(b)$$

it suffices to show that the function within the brackets tends to zero at infinity. Choose finite sets A and C such that $|\mu|(A') < \varepsilon$ and $|f(c)| < \varepsilon$ if $c \notin C$. On account of our assumption we can find a finite set B such that $A\,B'$ does not meet C. Then, for $b \notin B$

$$\Big|\sum_S {}_a f(a\,b)\,m(a)\Big| \leqq \max_A |f(a\,b)|\,|\mu|(A) + \max_{A'} |f(a\,b)|\,|\mu|(A') < \varepsilon(\|\mu\| + \|f\|),$$

where $\|f\|$ denotes the maximum of $|f|$.

Conversely, assume that there are infinitely many x such that $ax = c$ for some a and c. Let μ and ν be the unit masses placed at a and x respectively. Then $\nu \to 0$ in the weak star topology as x tends to infinity, while $\mu\nu$ identically equals the unit mass at c.

Let S be a semigroup and recall the matrix representation (antirepresentation) of the preceding section. We now extend this to a representation (antirepresentation) of the algebra of finite measures on S. To μ we make correspond

$$M = \sum_S {}_c m(c)\,T(c)\,.$$

The a, b th element of M equals $\mu(\{c;\, ac = b\})$ $(\mu(\{c;\, ca = b\}))$. It is immediately verified that kM corresponds to $k\mu$ for an arbitrary real constant k, $M + N$ to $\mu + \nu$ and $MN(NM)$ to $\mu\nu$. Here N denotes the matrix corresponding to ν.

4. Composition of independent random variables and Markov chains

Let $x_1, x_2, \ldots$ be a sequence of independent, identically distributed random variables taking their values in the (discrete) semigroup S and let π be their common probability distribution. (Given any probability distribution we can in the usual way construct a suitable sample space and define a sequence of independent random variables on it having the prescribed probability distribution.) Consider the partial products

$$y_n = x_1 x_2 \ldots x_n\,, \quad n = 1, 2, \ldots.$$

It is a fact, basic to this paper, that the sequence $y_1, y_2, \cdots$ forms a discrete parameter Markov chain with stationary transition probabilities. The Markov property is obvious from the relation

$$y_n = (x_1 x_2 \cdots x_{n-1})\,x_n = y_{n-1} x_n$$

and the independence of y_{n-1} and x_n. The one step transition probability, i. e. the probability that $y_n = b$ given that $y_{n-1} = a$, evidently equals

$$\pi(\{c;\, a\,c = b\})\,,$$

and hence the transition probabilities are stationary. The transition matrix is simply the matrix corresponding to π in the representation described in the previous section. The initial distribution of the Markov chain is $\{p(a);\, a \in S\}$ (remember that $p(a) = \pi(\{a\})$).

Changing left and right we can in an analogous way treat the partial products

$$z_n = x_n \cdots x_2 x_1, \quad n = 1, 2, \ldots.$$

The probability of going from a to b now equals

$$\pi(\{c;\, c\,a = b\}),$$

and the matrix of transition probabilities is identical with the matrix corresponding to π in the antirepresentation of the foregoing section.

It is evident that the case of non identically distributed random variables could be discussed in a similar way. However, the transition probabilities will then, in general, no longer be stationary.

The idea of considering a sequence of partial products of independent random variables as a Markov process is by no means new. In fact, many of the classical examples of Markov chains are of the above type, in particular the card mixing process and the unrestricted random walk in one or several dimensions (see the books by FELLER, DOOB and CHUNG). In the case of realvalued random variables (the law of composition being ordinary addition) CHUNG uses the term chain with independent increments. The double formulation of the problem has been exploited repeatedly with a purpose opposite to ours, i. e. limit theorems for products of independent random variables, derived mostly by means of Fourier analysis, have been interpreted as limit theorems for the corresponding Markov process. This is so in the papers by KAWADA and ITÔ (1940, separable compact groups), VOROBJOV (1954, finite Abelian groups) and KLOSS (1959, compact groups). ROSENBLATT (1960) and HEBLE and ROSENBLATT (1963) use the Markov process connection in deriving the complete description of the idempotent measures on a compact semigroup and, in his book, GRENANDER points out the possibility of deriving limit theorems on finite semigroups by borrowing results from the theory of Markov chains and considers a few examples.

Having constructed the above two Markov chains — for simplicity they will be called the right and left chain respectively — we can look upon an element (subset) of the semigroup also as a state (set of states) of any one of the two chains. The question arises as to what relations there are between the various concepts of the theory of Markov chains and those of the theory of semigroups. Further, if a state has a certain property in the right chain, does it necessarily have the same property as a state of the left chain? These two questions will be answered below in the form of a dictionary. A series of statements in Markov chain language are given in the left column and the algebraic reformulations in the right one. The Markov chain terminology is taken from FELLER's second edition.

4.1. *T is the minimal state space of the right (left) chain.*	*T is the semigroup generated by the support of π.*

Since the probability distribution of $y_n(z_n)$, $n = 1, 2, \ldots$, equals π^n the minimal state space is $\bigcup_{n=1}^{\infty} C(\pi^n) = \bigcup_{n=1}^{\infty} (C(\pi))^n$ and this is precisely the semigroup generated by the support of π.

From now on we will assume that the support of π generates the whole of S. This will simplify the formulations of our theorems a lot and evidently implies no essential restriction of generality. In Markov chain terms this means that if the state space is not minimal from the beginning we replace it by the minimal one.

4.2. *The state a leads to b in the right (left) chain.* — *b is divisible to the left (right) by a.*

$p^n(a, b) > 0$ if and only if there exists $x \in C(\pi^n)$ such that $ax = b$. Thus a leads to b if and only if $ax = b$ for some $x \in \bigcup_{n=1}^{\infty} C(\pi^n) = S$.

4.3. *R is a closed set of states in the right (left) chain.* — *R is a right (left) ideal.*

The set of states that can be reached from R in n steps equals $RC(\pi^n)$. Thus no state outside R can be reached from R if and only if

$$R \supset \bigcup_{n=1}^{\infty} RC(\pi^n) = R\left(\bigcup_{n=1}^{\infty} C(\pi^n)\right) = RS.$$

The following theorems are now obvious.

4.4. *R is a minimal closed set in the right (left) chain.* — *R is a minimal right (left) ideal.*

4.5. *R is the closure of E in the right (left) chain.* — *R is the right (left) ideal generated by E.*

4.6. *The closure of a set of states E in the right (left) chain is the set of all states that can be reached from it (including E).* — *The right ideal generated by E equals $E \cup ES$.*

A class in a Markov chain is an equivalence class with respect to the equivalence relation $\sim$, where $a \sim b$ if and only if a communicates with b or $a = b$ (see CHUNG's book). From what we have proved above it follows that, in the right (left) chain, this means precisely that a and b are right (left) ideal equivalent and hence the following correspondence is proved.

4.7. *R is a class in the right (left) chain.* — *R is a right (left) ideal layer.*

Since an essential class in the right (left) chain is the same as a minimal closed set, the algebraic translation is again minimal right (left) ideal, and a state is essential if and only if it is contained in a minimal right (left) ideal. Thus the following theorem holds (since its algebraic reformulation is true), making clear the relation between the properties of being essential in the right and left chain respectively

4.8. *If there exist essential states in both the right and the left chain, the essential states in the two chains are the same.* — *If there exist both minimal right and minimal left ideals, the union of the former equals the union of the latter (and is, in fact, the completely simple minimal twosided ideal).*

The following example, a modification of one given by ŠTEFAN SCHWARZ, shows that it may happen that there are essential states in one chain although there are none in the other. We have to construct a denumerable semigroup possessing a minimal right (left) ideal but no minimal left (right) ideal. Let $S = \{a, b, \ldots\}$ be the set of polygonal lines of the form

$$a(t) = \begin{cases} s_{m-1} + \dfrac{t - r_{m-1}}{r_m - r_{m-1}}(s_m - s_{m-1}) & \text{if} \quad r_{m-1} \leqq t \leqq r_m, \\ & \quad m = 1, 2, \ldots, n \\ s_n + t - r_n & \text{if} \quad t \geqq r_n \end{cases}$$

where $0 = r_0 < r_1 < \cdots < r_n$ and $0 < s_0 < s_1 < \cdots < s_n$ are all rational, $n = 0, 1, \ldots$, multiplication being the ordinary composition of functions. S is itself the unique minimal left ideal, since for every b and c the equation $xb = c$ is solvable. A suitable solution is

$$x(t) = \begin{cases} c(b^{-1}(t)) & \text{if} \quad b(0) \leqq t < +\infty \\ \dfrac{c(0)}{2}\left(1 + \dfrac{t}{b(0)}\right) & \text{if} \quad 0 \leqq t \leqq b(0) \end{cases}$$

However, S can not contain a minimal right ideal as that would imply complete simplicity and hence the existence of an idempotent. For $e^2 = e$ implies in particular $e(e(0)) = e(0)$, contradicting the fact that e is strictly increasing and $e(0) > 0$.

The next two theorems show that a state is persistent, persistent null or transient in one chain if and only if it is in the other so that we may unambiguously use these terms without reference to left and right.

4.9. *The state b is persistent (transient) in the right (left) chain.* $\qquad \sum_{n=1}^{\infty} \pi^n(\{b\}) = +\infty \, (< +\infty)$.

Find an m such that $b \in C(\pi^m)$, i. e. $p^m(b) > 0$. From the relation $(n \geqq m)$

$$\pi^n(\{b\}) = \sum_a p^m(a) p^{n-m}(a, b) \geqq p^m(b) p^{n-m}(b, b)$$

it is evident that the divergence of $\sum_{n=0}^{\infty} p^n(b, b)$ implies that of $\sum_{n=1}^{\infty} \pi^n(\{b\})$. (We define $p^0(a, b) = 1$ if $a = b$ and 0 if $a \neq b$.) Assume instead that $\sum_{n=0}^{\infty} p^n(b, b)$ converges. Using the fact that (see p. 21 in CHUNG's book)

$$\sum_{n=0}^{\infty} p^n(a, b) = h(a, b) \sum_{n=0}^{\infty} p^n(b, b)$$

where $h(a, b)$ is the probability that b will ever be reached from a, $a \neq b$, and $h(b, b) = 1$ we find

$$\begin{aligned} \sum_{n=1}^{\infty} \pi^n(\{b\}) &= \sum_{n=1}^{\infty} \sum_a p(a) p^{n-1}(a, b) \\ &= \sum_a p(a) \sum_{n=0}^{\infty} p^n(a, b) = \sum_{n=0}^{\infty} p^n(b, b) \sum_a p(a) h(a, b) \\ &\leqq \sum_{n=0}^{\infty} p^n(b, b). \end{aligned}$$

Hence $\sum_{n=1}^{\infty} \pi^n(\{b\})$ converges and we are finished.

4.10. *b is a persistent null state in the right (left) chain.* $\qquad \sum_{n=1}^{\infty} \pi^n(\{b\}) = +\infty$ *but* $\pi^n(\{b\}) \to 0$ *as* $n \to \infty$.

From the first formula of the preceding proof it follows that $\pi^n(\{b\}) \to 0$ implies that $p^n(b, b) \to 0$ as $n \to \infty$. Conversely, if b is null $p^n(a, b) \to 0$ for all a and thus so does

$$\pi^n(\{b\}) = \sum_a p(a) p^{n-1}(a, b).$$

4.11. *$\{s(c);\ c \in S\}$ is a stationary distribution of the right (left) chain.* $\qquad \sigma = \sigma\pi \, (\sigma = \pi\sigma)$.

This follows from the relations

$$\sigma(\{c\}) = s(c),$$

$$(\sigma\pi)(\{c\}) = \sum_{\{(a,b);\, ab=c\}} s(a)p(b) = \sum_a s(a) \sum_{\{b;\, ab=c\}} p(b)$$

$$= \sum_a s(a)p(a, c),$$

where $c \in S$ is arbitrary.

There remains the question whether a state necessarily has the same period in the two chains. It is trivially seen that it may happen that the period is undefined in both chains or defined in one of them but undefined in the other. The following example, due to HENRIK ERIKSSON, shows that the right and left period may be both defined and yet unequal. Take $S = \{0, a, b, c, 1\}$ with the multiplication table

	0	a	b	c	1
0	0	0	0	0	0
a	0	0	0	b	a
b	0	0	0	a	b
c	0	a	b	1	c
1	0	a	b	c	1

and choose a probability distribution with the support $\{a, c\}$. Then the support generates the whole semigroup and the period of a in the right chain is two, while it is one in the left chain. However, we shall mainly be interested in the period of persistent non null states, and the following theorem is accordingly more than sufficient for our purposes.

4.12. *Let a be a regular element. Then the period d of a in the right (left) chain is the largest integer with the property that the integers n such that $a \in C(\pi^n)$ differ only be multiples of d. In particular, a state which is essential in both chains, e. g. a persistent state, has the same period in the right chain as in the left one.*

We first show that the largest integer with the property that the values of n such that $a \in C(\pi^n)$ differ only by multiples of it, is a divisor of d. For this part of the proof we need not assume a to be regular. In fact, choose r such that $a \in C(\pi^r)$. Then $p^n(a, a) > 0$, i. e. $a \in aC(\pi^n)$, implies $a \in C(\pi^r)C(\pi^n) = C(\pi^{r+n})$ from which the assertion follows. Conversely, assume that a is regular and choose x such that $a = axa$. Since $x \in C(\pi^r)$ for some r, $a \in C(\pi^m) \cap C(\pi^n)$ implies $a \in aC(\pi^{r+m}) \cap aC(\pi^{r+n})$ and so $r + m$ and $r + n$ are both multiples of d. Hence, so is $n - m = (r + n) - (r + m)$. The last part of the theorem follows from the fact that a state which is essential in both chains, e. g. a persistent state, is contained in the completely simple semigroup of all essential states and hence regular.

The condition of regularity is, of course, not necessary as can be seen from the following example. Put $S = \{0, a, 1\}$, define multiplication by

	0	a	1
0	0	0	0
a	0	0	a
1	0	a	1

and choose a probability distribution with the whole of S as its support. Then the right and left period both equal one, although a is not regular.

The following theorem contains the basic facts we need for the proof of the main limit theorem of the following section.

4.13. *If there exist essential states in both chains, e. g. persistent states, all essential states have the same period. If there exist persistent states, they are either all null or all non null and there are no essential transient states.*

Suppose that there are essential states in both chains. Then, if not empty, the set of all essential states having a certain period, being closed in both the right and the left chain, is a twosided ideal contained in the minimal twosided ideal of all essential states and hence equals it. The same reasoning applies to the set of all persistent states and the set of all persistent null states.

The unsymmetrical random walk shows that, in case there are no persistent states, there may very well exist essential transient states. In fact, let S be the group of integers under addition and define π by

$$p(a) = \begin{cases} p & \text{if } a = 1 \\ 1 - p & \text{if } a = -1 \\ 0 & \text{if } a \neq 1, -1 \end{cases}$$

where $p \neq 0, {}^1/_2, 1$. Then the support of π generates the whole of S and all states are both essential and transient.

We finally remark that, although it would be unnatural, we could prove well-known theorems on semigroups using the ideas of this section. To exemplify this let us prove that a finite semigroup has at least one minimal right (left) ideal. Choose a probability distribution the support of which generates the whole semigroup and consider the right (left) associated chain. Since there are only a finite number of states not all of them can be transient and the persistent ones may be divided into minimal closed sets, i. e. minimal right (left) ideals.

5. The main limit theorem

The object of this section is to give a complete description of the behaviour of π^n as $n \to \infty$. We begin with the following lemma of independent interest, proved for a compact semigroup by ROSENBLATT (1960) (cf. GRENANDER's book, where a heuristic version of the proof is given).

5.1. *If I is a twosided ideal*

$$\pi^n(I) \to 1$$

as $n \to \infty$.

Retaining the notations of the previous section, $y_n \in I$ evidently implies $y_{n+1} = y_n x_{n+1} \in I$ and hence $\pi^{n+1}(I) \geqq \pi^n(I)$, $n = 1, 2, \ldots$. Now choose m so that $C(\pi^m) \cap I \neq 0$, i. e. $\pi^m(I) > 0$. From the decomposition

$$y_{km} = (x_1 \cdots x_m) \cdots (x_{km-m+1} \cdots x_{km})$$

and the fact that as soon as one of the blocks belongs to I so does the whole product it follows that

$$1 \geqq \pi^{km}(I) \geqq 1 - (1 - \pi^m(I))^k \to 1$$

as $k \to \infty$.

Let a be an arbitrary element of a (discrete) semigroup and consider the sequence of its powers,

$$a, a^2, \ldots, a^n, \ldots.$$

Two things can happen, either they are all distinct or else there are $h + d - 1$ different ones,

$$a, \ldots, a^{h-1}, a^h, \ldots, a^{h+d-1},$$

where $a^{h+d} = a^h$ and $\{a^h, a^{h+1}, \ldots, a^{h+d-1}\}$ is (isomorphic to) a cyclic group of order d.

This may be interpreted as a limit theorem for the probability distribution that assigns mass one to the point a. We have seen that, in the limit, the probability mass either escapes to infinity or jumps round cyclically. The following theorem shows that the behaviour of an arbitrary probability distribution is essentially the same.

5.2. *As $n \to \infty$ either*

$$\pi^n \to 0$$

vaguely or else there is a natural number d such that

$$\pi^{nd+r} \to \sigma_r, \quad r = 1, 2, \ldots, d,$$

weakly where $\sigma_1, \sigma_2, \ldots, \sigma_d$ have disjoint supports and

$$\sigma_r \pi = \pi \sigma_r = \sigma_{r+1 \,(\mathrm{mod}\, d)}, \quad \sigma_r \sigma_s = \sigma_{r+s \,(\mathrm{mod}\, d)}.$$

Suppose first that all states of the associated Markov chain are transient or persistent null. Then, as we have shown in the previous section, $\pi^n(\{a\}) \to 0$ for every $a \in S$, i. e. $\pi^n \to 0$ vaguely.

There remains to see what happens when there are non null persistent states. In this case the theorems 4.13 and 4.8 tell us that the persistent states are all non null and all have the same period, d, and that they form the (completely) simple minimal twosided ideal, I. By the lemma which we have just proved, $\pi^n(I) \to 1$ as $n \to \infty$, i. e. the probability of remaining for ever in the transient states is zero. These facts show firstly that no probability mass can escape to infinity and secondly, since

$$\pi^{nd+r}(\{b\}) = \sum_a p(a)\, p^{nd+r-1}(a, b),$$

that π^{nd+r} converges vaguely as $n \to \infty$, $r = 1, 2, \ldots, d$. Hence the convergence is, as a matter of fact, weak. The last two formulas of the theorem are now obvious consequences of the weak continuity of the convolution operation. It remains to show that the limit distributions have disjoint supports. Suppose that $a \in C(\sigma_r) \cap C(\sigma_s)$. Then there exist m and n such that $a \in C(\pi^{md+r}) \cap C(\pi^{nd+s})$ and hence, according to theorem 4.12, $r = s$.

The theorem we have just proved has the following immediate corollary, proved for a compact semigroup by Rosenblatt (1960) and for a locally compact group by Grenander (p. 59 in his book).

5.3. *The Césaro mean*

$$\frac{1}{n}(\pi + \pi^2 + \cdots + \pi^n)$$

converges as $n \to \infty$, either vaguely to zero or weakly to an idempotent measure σ such that

$$\sigma \pi = \pi \sigma = \sigma^2 = \sigma.$$

The proof is contained in the following lines, letting k denote the integral part of n/d,

$$\begin{aligned}&\frac{1}{n}(\pi+\pi^2+\cdots+\pi^n)\\ &=\frac{kd}{n}\frac{1}{d}\left(\frac{1}{k}(\pi+\cdots+\pi^{(k-1)d+1})+\cdots+\frac{1}{k}(\pi^d+\cdots+\pi^{kd})\right)+\\ &\quad+\frac{1}{n}(\pi^{kd+1}+\cdots+\pi^n)\to\frac{1}{d}(\sigma_1+\sigma_2+\cdots+\sigma_d).\end{aligned}$$

Alternatively, we could have applied directly the corresponding Markov chain theorem.

6. Idempotent measures

In the present section we proceed to give an algebraic description of the idempotent probability measures. Maybe it is not out of place to recapitulate briefly the literature on this subject. LÉVY (1939) determined all idempotents (real and complex) on the circle group and his results have been rediscovered, in the case of probability measures, by KAKEHASHI (1949) and, in the general case, by HELSON (1953). This line of investigation was carried on by RUDIN (1959) and completed by COHEN (1960), who determined all idempotents on a general locally compact Abelian group. On a compact, not necessarily commutative, group the idempotent probability distributions were described by KAWADA and ITÔ (1940) and their result has been rederived (without their separability assumption) in almost all papers on the subject. In the case of probability measures on a compact semigroup partial results have been obtained (round 1960) by a large number of authors. The complete solution was found by PYM (1962) and HEBLE and ROSEERBLATT (1963) We are now able to do without the compactness condition, still assuming, however, that our semigroup is discrete.

6.1. *The support of an idempotent probability measure, π, is a completely simple semigroup such that the (isomorphic) groups $G_{i,j}=R_iL_j=R_i\cap L_j$ in its group decomposition are finite. Letting $\sigma_{i,j}$ denote the normed Haar measure over $G_{i,j}$ we have, with $\pi(R_i)=t_i$ and $\pi(L_j)=u_j$,*

$$\pi=\sum\nolimits_{i,j}t_i u_j\sigma_{i,j},\quad \sum\nolimits_i t_i=\sum\nolimits_j u_j=1.$$

Conversely, any measure of this type is idempotent.

Suppose that π is idempotent. Then $\pi^n=\pi$ for all natural n and, in particular, $\pi^n\to\pi$ as $n\to\infty$. Hence, all states of the associated chain are persistent (non null), so that the support of π is a completely simple semigroup. Let us restrict our attention to an arbitrary fixed minimal right ideal R_i, i. e. an irreducible closed set in the right chain. A theorem of DERMAN (see p. 50 in CHUNG's book) tells us that the stationary distribution equations

$$s(c)=\sum_{R_i}{}_a\, s(a)p(a,c)$$

admit of a unique (up to a multiplicative constant) non negative solution. Let us show that

$$s(a)=\pi(L_j)\quad\text{if}\quad a\in G_{i,j}$$

is one such solution. Suppose that $c \in G_{i,l}$. Then

$$\sum_{R_i} {}_a s(a)\, p(a,c) = \sum{}_j \pi(L_j) \sum_{G_{i,j}} {}_a \sum_{\{b;\, ab=c\}} {}_b\, p(b)$$

$$= \sum{}_j \pi(L_j) \sum_{L_l} {}_b\, p(b) \sum_{G_{i,j} \cap \{a;\, ab=c\}} {}_a 1 = \pi(L_l) = s(c)\,,$$

since for every $c \in G_{i,l}$ and $b \in L_l$ there exists exactly one $a \in G_{i,j}$ such that $ab = c$. On the other hand, due to the idempotence of π, the restriction of p to R_i also satisfies the equations above. From this we conclude, firstly, that the groups $G_{i,j}$ are finite and, secondly, that the restriction of π to R_i equals $\pi(R_i) \sum_j \pi(L_j)\, \sigma_{i,j}$.

The last assertion of the theorem follows from the relation

$$\sigma_{i,j}\, \sigma_{k,l} = \sigma_{i,l}\,,$$

which is an immediate corollary to the fact, used previously in the proof, that to every $c \in G_{i,l}$ and $b \in G_{k,l}$ there exists exactly one $a \in G_{i,j}$ such that $ab = c$.

We could just as well, in accordance with ROSENBLATT and PYM, formulate our theorem in terms of the REES structure theorem. In fact, we have shown that the support of an idempotent probability measure π is isomorphic with a REES matrix semigroup $G \times I \times J$ over a finite group G (I and J can of course be at most denumerable). Further

$$\pi = \sigma \times \tau \times v\,,$$

where σ is the normalized Haar measure over G and τ and v are probability measures on I and J respectively. Conversely, any such measure is idempotent.

In the commutative case the theorem takes the following form.

6.2. *An idempotent probability measure on a commutative semigroup is the normed Haar measure on a finite subgroup and conversely.*

PYM (1962) and LOYNES (1963) have proved that the idempotent probability measures on a locally compact group are precisely the Haar measures on compact subgroups. In the discrete case a simple direct proof of this fact has been given by RUDIN (1963). We deduce it as a corollary to the theorem above.

6.3. *An idempotent probability measure on a group is the normed Haar measure on a finite subgroup and conversely.*

We know already that the support is a completely simple semigroup such that the groups in its group decomposition are finite. However, since a group does not contain more than one idempotent (namely the identity element), there is only one group in the group decomposition.

7. The semigroup of probability measures and the structure of its subgroups

We have already proved that the convolution of two probability measures is again a probability measure and that the convolution operation is associative (provided the underlying algebraic structure is a semigroup) and jointly continuous in the weak topology. This can be expressed by saying that the set of all probability measures is a topological semigroup. It has been studied from an algebraic point of view in a series of papers (mostly assuming that the underlying structure is a compact semigroup). Typical exponents of this line of investigation are WENDEL (1954), SCHWARZ (1957, 1963), GLICKSBERG (1959), COLLINS (1960, 1961) and COLLINS and KOCH (1962). Results in this spirit may be obtained as reformulations and corollaries of many of the theorems of this paper.

One problem that has attracted attention is to determine the structure of a group of measures, see GLICKSBERG (1959, compact groups and compact Abelian semigroups, and 1961, compact Abelian separately continuous semigroups), COLLINS (1962, compact semigroups), KLOSS (1962, locally compact groups representable in compact ones) and SCHWARZ (1963, finite semigroups). We have already met one such group, namely the set of all weak limit points of the sequence $\pi, \pi^2, \ldots$. In fact, in theorem 5.2 we proved that (if non empty) it was isomorphic to a finite cyclic group. The problem will now be solved in the discrete case, assuming neither compactness nor the group property. First we introduce the following concept. If Γ is a set of measures (on a discrete space), the support of Γ, $C(\Gamma)$, is the union of the supports of all measures contained in Γ.

7.1. *Let Γ be a group of probability measures on a discrete semigroup and let ε be its identity element. Then $C(\Gamma) = G \times I \times J$ is a completely simple semigroup and $C(\varepsilon) = H \times I \times J$ a completely simple subsemigroup. Here H is a finite normal subgroup of G. The elements of Γ are of the form $\nu \times \tau \times v$, where τ and v are fixed probability measures over I and J respectively and ν is the uniform distribution over a coset of H, i. e. the Haar measure over H translated.*

ε is idempotent and hence

$$C(\varepsilon) = \bigcup_i R_i = \bigcup_j L_j = \bigcup_{i,j} H_{i,j} = H \times I \times J$$

is a completely simple semigroup.

$$(C(\Gamma))^2 = C(\Gamma^2) = C(\Gamma),$$

so that $C(\Gamma)$ is a semigroup, and

$$C(\Gamma) = C(\varepsilon\Gamma) = C(\varepsilon)\, C(\Gamma) = \bigcup_i R_i\, C(\Gamma).$$

$R_i C(\Gamma)$ is obviously a right ideal of $C(\Gamma)$. If we can prove that it is minimal, it will follow (since we can prove analogously that $C(\Gamma) L_j$ is a minimal left ideal) that $C(\Gamma)$ is completely simple.

Take $a \in R_i C(\Gamma)$. Then $a \in R_i C(\pi)$ for some π in Γ. Consequently

$$a\, C(\pi^{-1}) \subset R_i\, C(\pi)\, C(\pi^{-1}) = R_i\, C(\pi\pi^{-1}) = R_i\, C(\varepsilon) = R_i,$$

and

$$a\, C(\pi^{-1})\, C(\varepsilon) = a\, C(\pi^{-1}\varepsilon) = a\, C(\pi^{-1}),$$

so that $a C(\pi^{-1})$ is a right ideal in $C(\varepsilon)$ contained in the minimal right ideal R_i, whence $a C(\pi^{-1}) = R_i$. Finally,

$$a C(\Gamma) = a\, C(\pi^{-1}\Gamma) = a\, C(\pi^{-1})\, C(\Gamma) = R_i\, C(\Gamma),$$

and the minimality of $R_i C(\Gamma)$ is proved.

The groups in the group decomposition of $C(\Gamma)$ are $G_{i,j} = R_i C(\Gamma) C(\Gamma) L_j \supset R_i C(\varepsilon) L_j = R_i L_j = H_{i,j}$ with i in I and j in J. Recalling how the representation of a completely simple semigroup as a REES matrix semigroup is constructed (see the book by LJAPIN, p. 281), we find that

$$C(\Gamma) = G \times I \times J,$$

where G is a group containing H as a subgroup and the sandwich matrix defining the multiplication in $G \times I \times J$ may be chosen as the one corresponding to $C(\varepsilon) = H \times I \times J$. In particular the elements of it belong to H, which is also necessary since $H \times I \times J$ is a subsemigroup of $G \times I \times J$.

Let π be an arbitrary element of Γ and choose (g, i, j) in $C(\pi)$. Then

$$HgH \times I \times J = (H \times I \times J)\,(g, i, j)\,(H \times I \times J) \subset C(\varepsilon)\,C(\pi)\,C(\varepsilon) = C(\pi)$$

and hence

$$C(\pi) = A \times I \times J$$

where $AH = HA = A$. Likewise

$$C(\pi^{-1}) = B \times I \times J$$

with $BH = HB = B$, and

$$AB \times I \times J = C(\pi)\,C(\pi^{-1}) = C(\varepsilon) = H \times I \times J\,.$$

This can not hold unless A is a right and $B = A^{-1}$ a left coset of H. Changing the order of multiplication we find that A is also a left coset of H, and, since any point of G is contained in the projection of $C(\pi)$ for some π in Γ, it follows that H is normal.

It remains to show that $\pi = \nu \times \tau \times \upsilon$ where ν is the uniform distribution over a coset of H. We know already that this is true for ε. Let m be the mass that the normed Haar measure over H attributes to each of its points. From $\pi = \varepsilon\pi$ it follows that, letting $c = (g, i, j)$ be an arbitrary point of $C(\pi)$,

$$\begin{aligned} p(c) &= \sum\nolimits_{\substack{a,b \\ \{(a,b);\, ab = c\}}} e(a)\,p(b) = \sum\nolimits_k m\,t_i\,u_k \sum\nolimits_b p(b) \sum\nolimits_{\substack{a \\ H_{i,k} \cap \{a;\, ab = c\}}} 1 \\ &= \sum\nolimits_k m\,t_i\,u_k\,\pi(G \times I \times \{j\}) = m\,t_i\,\pi(G \times I \times \{j\})\,. \end{aligned}$$

Dually,

$$p(c) = m\,\pi(G \times \{i\} \times J)\,u_j\,,$$

wherefrom we conclude that

$$p(c) = m\,t_i\,u_j\,.$$

The proof is finished.

8. Necessary and sufficient convergence conditions

The theorem of the previous section allows us to give necessary and sufficient conditions for the convergence (weak and vague) of π^n as $n \to \infty$. We shall see that the limit behaviour is completely determined by the support and the algebraic properties of the semigroup. In the case of a compact group the solution was essentially given already by Kawada and Itô (1940) and since then it has been rederived in almost all papers on the subject, using Fourier analytic as well as algebraic methods. Recently, Rosenblatt (1964) has settled the problem on an arbitrary compact semigroup.

8.1. $\pi^n \to 0$ *vaguely as* $n \to \infty$ *if and only if the semigroup generated by the support of* π *does not posess a completely simple minimal twosided ideal with finite groups in its group decomposition.*

We know already from section 5 and 6 that the condition is sufficient.

Conversely, assume that the semigroup generated by the support of π contains the completely simple minimal twosided ideal $K = \bigcup_i R_i = \bigcup_j L_j = \bigcup_{i,j} G_{i,j}$, where $G_{i,j}$ is finite. According to theorem 5.1. we can find m such that $\pi^m(K) > 1 - \varepsilon$ and then

$$\pi^m\left(\left(\bigcup_{i=1}^{r} R_i\right) \cap \left(\bigcup_{j=1}^{s} L_j\right)\right) > 1 - \varepsilon$$

for sufficiently large r and s. Hence, if $n \geqq m$,

$$\pi^n\left(\bigcup_{i=1}^{r} R_i\right) \geqq \pi^m\left(\bigcup_{i=1}^{r} R_i\right) > 1 - \varepsilon$$

and

$$\pi^n\left(\bigcup_{j=1}^{s} L_j\right) \geqq \pi^m\left(\bigcup_{j=1}^{s} L_j\right) > 1 - \varepsilon,$$

so that

$$\pi^n\left(\left(\bigcup_{i=1}^{r} R_i\right) \cap \left(\bigcup_{j=1}^{s} L_j\right)\right) > 1 - 2\varepsilon.$$

Since

$$\left(\bigcup_{i=1}^{r} R_i\right) \cap \left(\bigcup_{j=1}^{s} L_j\right) = \bigcup_{i=1}^{r} \bigcup_{j=1}^{s} G_{i,j}$$

is finite, this means that the sequence $\pi, \pi^2, \ldots$ is weakly conditionally compact, i. e. no probability mass escapes to infinity.

8.2. *π^n converges weakly as $n \to \infty$ if and only if the semigroup generated by the support of π posesses a completely simple minimal twosided ideal $G \times I \times J$ with G finite, which contains no subsemigroup $H \times I \times J$, H being a proper normal subgroup of G, such that*

$$C(\pi)\,(H \times I \times J) \subset g H \times I \times J$$

for some g outside H.

Suppose that π^n does not converge weakly. Then either $\pi^n \to 0$ vaguely, in which case we use the previous theorem, or else there is an integer $d > 1$ such that

$$\pi^{nd+r} \to \sigma_r, \quad r = 1, \ldots, d,$$

where

$$\bigcup_{r=1}^{d} C(\sigma_r) = G \times I \times J,$$

G being finite, is the completely simple minimal twosided ideal of the semigroup generated by the support of π. σ_d is the identity element of the group of limit distributions. Hence, by theorem 7.1, $G \times I \times J$ has the subsemigroup

$$C(\sigma_d) = H \times I \times J,$$

where H is a proper normal subgroup of G. From theorem 5.2 and 7.1 we conclude that

$$C(\pi)\,(H \times I \times J) = C(\pi)\,C(\sigma_d) = C(\pi\,\sigma_d) = C(\sigma_1) = g H \times I \times J$$

for a suitable g outside H.

Conversely,

$$C(\pi)\,(H \times I \times J) \subset g H \times I \times J$$

implies

$$C(\pi^n) \cap (G \times I \times J) \subset C(\pi^n)\,(H \times I \times J) \subset C(\pi^{n-1})\,(g H \times I \times J) \subset \cdots \subset g^n H \times I \times J,$$

which, together with the fact that $\pi^n(G \times I \times J) \to 1$, shows that π^n does not converge weakly as $n \to \infty$.

9. Infinitely divisible distributions

A probability distribution π is said to be infinitely divisible if, for every natural number n, there exists a probability distribution π_n such that

$$\pi = \pi_n^n.$$

Particularly important infinitely divisible distributions are the compound Poisson distributions. Such a distribution is of the form

$$e^{-c}\left(\varepsilon + \frac{c}{1!}\varrho + \frac{c^2}{2!}\varrho^2 + \cdots\right) = \sum_{k=0}^{\infty} e^{-c}\frac{c^k}{k!}\varrho^k,$$

where c is a positive number, $\varepsilon = \varrho^0$ an idempotent probability measure and ϱ a probability distribution such that

$$\varrho\,\varepsilon = \varepsilon\,\varrho = \varrho\,.$$

The last condition assures us that

$$\varrho^m \varrho^n = \varrho^{m+n}$$

for all $m, n \geqq 0$. Using this it is immediate that the distribution above is infinitely divisible. In fact, an nth root is obtained by replacing c by c/n.

A basic question is whether, conversely, an infinitely divisible distribution is necessarily a compound Poisson one. This is known to be so in the classical cases when the semigroup under consideration is the set of non negative integers under addition (see FELLER's book) or the whole integer group. VOROBJOV (1954) obtained the same result by means of Fourier analysis for an arbitrary finite Abelian group and BÖGE (1959), using quite different methods, removed the commutativity assumption.

One might think that that the same result would hold for an arbitrary discrete semigroup. That this is not so even for a commutative group is shown by the following example.

Consider the set of all rational numbers under addition and let π be the point mass at 1. Then, for every natural number n, the point mass at $1/n$ is the unique probability measure π_n satisfying the definition of infinite divisibility. However, π is not a compound Poisson distribution since such a distribution can not be degenerate except at 0.

It does not seem to be known at present what the appropriate conditions are that should be imposed to avoid the pathologies exhibited in this example.

In the rest of this section we restrict our attention to finite (i. e. compact discrete) semigroups. We shall show that an infinitely divisible distribution is then necessarily compound Poisson, thus generalising the result of BÖGE to the non group case. The following lemma will be needed.

9.1. *An infinitely divisible distribution on a finite semigroup has, for each n, at least one infinitely divisible nth root.*

The proof is wellknown, see VOROBJOV (1954) and KLOSS (1961, 1962), and works just as well on an arbitrary compact semigroup. Let π_m be an mth root, $m = 1, 2, \ldots$. An infinitely divisible nth root may then be chosen as a limit point of the sequence

$$\pi_{m!\,n}^{m!}, \quad m = 1, 2, \ldots.$$

Following KLOSS (1961) we now imbed our infinitely divisible distribution π in a rational parameter homogeneous process, i. e. we construct a family of probability measures π_t, $t > 0$ rational, such that

$$\pi_s\,\pi_t = \pi_{s+t}$$

for all $s, t > 0$, and $\pi_1 = \pi$.

According to the lemma we can find a sequence of infinitely divisible distributions $\varrho_1, \varrho_2, \ldots$ such that $\varrho_2^2 = \pi$, $\varrho_3^3 = \varrho_2, \ldots$. We put for $t = p/q$ with p and q natural,

$$\pi_t = \varrho_k^{(p/q)k!}, \quad k \geqq q,$$

which is a permissable definition. The equation above is easily seen to be satisfied and $\pi_1 = \varrho_k^{k!} = \pi$.

We shall now (and this is the clue) show that π_t converges as $t \to 0$. Let $\mu = \lim_{i\to\infty} \pi_{s_i}$ and $\nu = \lim_{i\to\infty} \pi_{t_i}$ be two arbitrary limit points. We may suppose that $s_i < t_i$ for all i. From the relation

$$\pi_{s_i}\pi_{t_i - s_i} = \pi_{t_i - s_i}\pi_{s_i} = \pi_{t_i}$$

it follows that the equations

$$\mu\xi = \nu, \quad \xi\mu = \nu$$

are solvable. In fact, ξ may be chosen as a limit point of the sequence $\pi_{t_i - s_i}$, $i = 1, 2, \ldots$. This means that the set of all limit distributions is a (commutative) group. By theorem 7.1 it is even finite (because of the finiteness of the semigroup under consideration). Let n be the order of it and η a limit point of the sequence $\pi_{s_i/n}$, $i = 1, 2, \ldots$. Since $\pi_{s_i} = \pi_{s_i/n}^n$ for every i,

$$\mu = \eta^n = \varepsilon,$$

where ε denotes the unit element of the group. Consequently ε is the only limit distribution, i. e. $\pi_t \to \varepsilon$ as $t \to 0$.

We could now, following BÖGE, directly finish the proof that π is a compound Poisson distribution. Alternatively, we can first show that π_t, now defined merely for positive rational t, may be continuously extended to all non negative real t and then apply the results of the following section.

Let t and h denote positive rational numbers. We know that $\lim_{h\to 0} \pi_h$ exists. But then (remember that all types of convergence are equivalent on a finite set)

$$\begin{aligned}\|\pi_{t+h} - \pi_t\| &= \|\pi_{t-h}\pi_{2h} - \pi_{t-h}\pi_h\| \\ &\leqq \|\pi_{t-h}\|\,\|\pi_{2h} - \pi_h\| = \|\pi_{2h} - \pi_h\| \to 0\end{aligned}$$

and

$$\begin{aligned}\|\pi_{t-h} - \pi_t\| &= \|\pi_{t-2h}\pi_h - \pi_{t-2h}\pi_{2h}\| \\ &\leqq \|\pi_{t-2h}\|\,\|\pi_h - \pi_{2h}\| = \|\pi_{2h} - \pi_h\| \to 0\end{aligned}$$

as $h \to 0$. Hence, by continuity, π_t may be extended to all non negative real t, and the extension is continuous. The proof of the following theorem is finished.

9.2. *A probability measure on a finite semigroup is infinitely divisible if and only if it is a compound Poisson distribution.*

10. Homogeneous stochastic processes

In this section we shall study the continuous analogue of the composition of independent random variables, i. e. the homogeneous stochastic processes. By such a process we mean a family of probability distributions π_t, $t > 0$, with the property

$$\pi_s\pi_t = \pi_{s+t}$$

for all real $s, t > 0$. It is said to be continuous if $\pi_s \to \pi_t$ as $s \to t$ for all $t > 0$. π_t is physically interpreted as the probability distribution of the increment of the process over a time interval of length t. The equation above then states that the increments of the process over successive time intervals are independent.

Just as we in the foregoing sections used the theory of discrete parameter Markov chains, we can now borrow results from the continuous parameter case. Let P_t be the matrix corresponding to π_t in the representation of section 3, i. e.

$$p_t(a, b) = \pi_t(\{c; ac = b\}) = \sum_{\{c;\, ac = b\}} {}_c p_t(c).$$

Being stochastic and satisfying the Chapman-Kolmogorov equation

$$P_s P_t = P_{s+t}, \quad s, t > 0,$$

P_t, $t > 0$, is a transition matrix.

10.1. *Let π_t, $t > 0$, be a continuous homogeneous process. Then the support of π_t is independent of t.*

We note firstly that $\bigcup_t C(\pi_t)$ is denumerable since it suffices to extend the union over the denumerable set of positive rational numbers. Secondly, being an at most denumerable sum of continuous functions, $p_t(a, b)$ is a measurable (in fact, lower semicontinuous) function of t for each a and b. Hence $p_t(a, b)$ is either identically zero or never zero (see CHUNG's book, p. 121). In particular, so is $p_t(a) = p_t(1, a)$, where 1 is the identity element (if there was no identity element from the beginning we could have adjoined one).

In case the semigroup under consideration is denumerable the continuity assumption may be replaced by the seemingly weaker condition that $p_t(a)$ is a measurable function of t for each a (see p. 121 in CHUNG's book). In the non denumerable case this is no longer true. Consider for example the real line with the discrete topology and let π_t be the point mass at t. Then

$$p_t(a) = \begin{cases} 1 & \text{if} \quad t = a \\ 0 & \text{if} \quad t \neq a \end{cases}$$

so that measurability though not continuity holds.

When the semigroup is denumerable more effort is needed to construct a non continuous process since it is then not even measurable. However, consider the commutative group of rational numbers under addition. It is known that there exists a non trivial rational valued function f such that

$$f(s) + f(t) = f(s + t), \quad -\infty < s, \quad t < +\infty.$$

This result goes back to HAMEL (for a proof see p. 259 in CHUNG's book). The desired non measurable process is obtained by letting π_t be the point mass at $f(t)$ for each $t > 0$.

10.2. *A homogeneous process on a finite semigroup is continuous.*

This follows from the corresponding Markov chain theorem due to DOEBLIN (1938). (The reference is found in CHUNG's book. I am indebted to him for pointing it out to me.) Alternatively, we could have proved the theorem directly, repeating almost word for word the argument in the previous section.

We turn to the study of the behaviour of π_t as $t \to \infty$. The situation is somewhat simpler than the discrete time case since the probability mass can no longer jump round cyclically.

10.3. *Let π_t, $t > 0$, be a continuous homogeneous process. As $t \to \infty$ π_t converges, either vaguely to zero or weakly to an idempotent measure σ such that*

$$\sigma \pi_t = \pi_t \sigma = \sigma .$$

The corresponding Markov chain theorem tells us (after having adjoined an identity element if necessary) that π_t converges vaguely as $t \to \infty$. But then the limit distribution coincides with the limit distribution of $\pi_n = \pi_1^n$ as $n \to \infty$, which is either zero or an idempotent probability measure. The last relation follow from the weak continuity of the convolution operation.

Theorem 8.1 tells us precisely when the probability mass escapes to infinity.

10.4. *A continuous homogeneous process π_t, $t > 0$, converges weakly as $t \to \infty$ if and only if the support of the process contains a completely simple minimal twosided ideal with finite groups in its group decomposition.*

We now turn to study the infinitesimal properties of the process.

10.5. *A continuous homogeneous process converges vaguely as $t \to 0$.*

After having, if necessary, adjoined an identity element this follows from the corresponding Markov chain theorem (p. 118 in CHUNG's book). As is shown by the first example above, the converse is, at least in the non denumerable case, false.

A homogeneous process is said to be a compound Poisson process if it is of the form

$$\pi_t = e^{-ct}\left(\varepsilon + \frac{ct}{1!}\varrho + \frac{(ct)^2}{2!}\varrho^2 + \cdots\right) = \sum_{n=0}^{\infty} e^{-ct}\frac{(ct)^n}{n!}\varrho^n$$

where c is a non negative constant, $\varepsilon = \varrho^0$ an idempotent probability measure and ϱ a probability distribution such that

$$\varrho\,\varepsilon = \varepsilon\,\varrho = \varrho .$$

It is immediately verified that $\pi_s \pi_t = \pi_{s+t}$, $s, t > 0$, and that π_t is continuous.

Since

$$\lim_{t\to 0} \pi_t = \varepsilon$$

and

$$\lim_{t\to 0} 1/t(\pi_t - \varepsilon) = c(\varrho - \varepsilon)$$

(with convergence in the norm topology), ε and $c(\varrho - \varepsilon)$ are uniquely determined by the process. Conversely, as we shall see in a while, π_t may be expressed solely in terms of ε and $c(\varrho - \varepsilon)$. It will be proved below that the restriction of ϱ to $C(\varepsilon)$ is proportional to ε and hence we can just as well assume that ϱ and ε have disjoint supports. With this convention not only $c(\varrho - \varepsilon)$ but c and ϱ themselves are unique (unless $c = 0$, in which case ϱ may be arbitrary).

Suppose now that $ab = c$ with a, c in $C(\varepsilon)$ and b in $C(\varrho)$. Since $\varrho = \varepsilon\varrho$, $C(\varrho) = C(\varepsilon)C(\varrho)$, there is an idempotent e in $C(\varepsilon)$ such that $b = eb$. Consequently $b = (eae)^{-1}ec \in C(\varepsilon)$, where the inverse is taken in that one of the groups in the group decomposition of $C(\varepsilon)$, whose identity element is e. Letting ν be the restriction of ϱ to $C(\varepsilon)$ we conclude that

$$\nu\,\varepsilon = \varepsilon\,\nu = \nu .$$

Hence, remembering the proof of theorem 6.1, ν is proportional to ε and we have proved what we promised.

10.6. *A homogeneous process is a compound Poisson process if and only if it converges weakly as $t \to 0$.*

We have already proved that a compound Poisson process converges weakly as $t \to 0$. The proof of the converse is wellknown from the analytical theory of operator semigroups.

First we note that the existence of $\lim_{t\to 0} \pi_t = \varepsilon$ (in the weak or, equivalently, the norm topology) implies continuity for all t. A direct proof was given at the end of section 9. We could also argue as follows. Adjoin an identity element if necessary. The continuity of $p_t(a, b) = \pi_t(\{c;\, ac = b\})$ at zero implies continuity for all t (p. 118 in CHUNG's book). In particular, continuity holds for $p_t(a) = p_t(1,a)$.

Now find $\gamma > 0$ such that $\|\pi_t - \varepsilon\| \leqq \delta$ for $t \leqq \gamma$. If δ is sufficiently small,

$$\log \pi_t = -\frac{(\varepsilon - \pi_t)}{1} - \frac{(\varepsilon - \pi_t)^2}{2} - \cdots$$

is welldefined and continuous for $t \leqq \gamma$,

$$\log \pi_\gamma = \log \pi_{\gamma/n}^n = n \log \pi_{\gamma/n}$$

and

$$\log \pi_{m\gamma/n} = \log \pi_{\gamma/n}^m = m \log \pi_{\gamma/n} = \frac{m}{n} \log \pi_\gamma$$

for an arbitrary rational number $m/n \leqq 1$. By continuity

$$\log \pi_t = t\,\frac{1}{\gamma} \log \pi_\gamma = t\,\mu\,,$$

$$\pi_t = e^{t\mu} = \varepsilon + \frac{t}{1!}\,\mu + \frac{t^2}{2!}\,\mu^2 + \cdots$$

for $t \leqq \gamma$. If $t > 0$ is arbitrary, we can choose n so large that $t/n \leqq \gamma$ and conclude that

$$\pi_t = \pi_{t/n}^n = (e^{(t/n)\mu})^n = e^{t\mu}\,.$$

Since

$$\mu = \lim_{t\to 0} \frac{\pi_t - \varepsilon}{t}$$

and

$$\mu\,\varepsilon = \varepsilon\,\mu = \mu$$

we see, firstly, that the total variation of μ is zero and that μ is non negative outside $C(\varepsilon)$ and, secondly, that the restriction of μ to $C(\varepsilon)$ is proportional to ε. Hence $\mu = c(\varrho - \varepsilon)$ for a suitable $c \geqq 0$ and probability measure ϱ with $\varrho\,\varepsilon = \varepsilon\,\varrho = \varrho$, wherefrom we conclude that

$$\pi_t = \sum_{n=0}^{\infty} \frac{t^n}{n!}\, c^n (\varrho - \varepsilon)^n = \sum_{n=0}^{\infty} \frac{(ct)^n}{n!} \sum_{k=0}^{n} \binom{n}{k} \varrho^k (-1)^{n-k}$$

$$= \sum_{k=0}^{\infty} \frac{(ct)^k}{k!}\, \varrho^k \sum_{n=k}^{\infty} \frac{(-ct)^{n-k}}{(n-k)!} = \sum_{k=0}^{\infty} \frac{(ct)^k}{k!}\, e^{-ct} \varrho^k\,,$$

i. e. π_t is a compound Poisson process.

It is now easy to see that the only continuous homogeneous processes on a group are the compound Poisson ones. In the commutative case this has been proved, using Fourier analysis, by BOCHNER (1955).

10.7. *A continuous homogeneous process on a group is a compound Poisson process and conversely.*

Let π_t, $t > 0$, be a continuous homogeneous process on a group. We know already that π_t converges vaguely as $t \to 0$, say towards ε. Using the vague continuity of the convolution when one factor is held fixed, established in section 3, we conclude that

$$\pi_t = \lim_{s\to 0} \pi_{s+t} = \lim_{s\to 0} \pi_s\,\pi_t = \varepsilon\,\pi_t$$

(the limit is in the vague topology). Hence

$$\|\pi_t\| = \|\varepsilon\|\,\|\pi_t\|\,, \quad \|\varepsilon\| = 1\,,$$

so that, in fact, π_t converges weakly as $t \to 0$.

Without the group property the above theorem is no longer true. However, just as the asymptotic behaviour of π_t as $t \to \infty$ proved to be completely determined by the algebraic properties of the support of the process, one might hope that the condition that π_t converges weakly as $t \to 0$ is also equivalent to some algebraic property of the support. At least in the commutative case this is indeed so.

10.8. *A continuous homogeneous process on a commutative semigroup is a compound Poisson process if and only if the support of the process has an identity element.*

Let $C = C(\pi_t)$ be the support of the process. Suppose first that π_t converges weakly as $t \to 0$. The limit distribution, say ε, is then the uniform distribution over a finite subgroup G. From

$$\pi_t \varepsilon = \varepsilon \pi_t = \pi_t$$

it follows that

$$C G = G C = C .$$

Hence the identity of G is an identity element for the whole of C.

Conversely, suppose that C has an identity element, say 1, and let G be the maximal group containing it. From $ab = g \in G$ we conclude that $a(bg^{-1}) = (ab)g^{-1} = gg^{-1} = 1$, i. e. a is in G. Similarly, b is contained in G. Consequently, the restriction of π_t to G, μ_t, satisfies

$$\mu_{s+t} = \mu_s \mu_t, \quad s, t > 0 .$$

From

$$\| \mu_{s+t} \| = \| \mu_s \| \, \| \mu_t \| , \quad s, t > 0 ,$$

it follows, since μ_t is non zero, that $\|\mu_t\| = e^{-ct}$ for some $c \geqq 0$. But then $e^{ct} \mu_t$ is a continuous homogeneous process on G and the previous theorem tells us that it converges weakly as $t \to 0$. Hence, so does

$$\pi_t = e^{ct} \mu_t + (1 - e^{ct}) \mu_t + (\pi_t - \mu_t) .$$

We finally give an example of a continuous homogeneous process on a commutative semigroup which is not a compound Poisson process. Let S be the set of integers with the multiplication rule

$$a b = \max(a, b) ,$$

and let π be an arbitrary probability distribution on S. It may be imbedded in the continuous homogeneous process π_t, $t > 0$, defined by

$$\pi_t(\{\cdots, a - 1, a\}) = (\pi(\{\cdots, a - 1, a\}))^t ,$$

$a = \cdots, -1, 0, +1, \ldots$. If the support of π is bounded below, the smallest element of it is an identity element and all probability mass concentrates at it as $t \to 0$. However, if the support has no lower bound,

$$\pi_t \to 0$$

vaguely as $t \to 0$.

It is interesting to note that the derivative at zero still exists, although now merely in the vague topology, and that it is unbounded. Indeed,

$$\frac{\pi_t - 0}{t} \to \mu$$

vaguely as $t \to 0$, where

$$m(a) = \log \frac{\pi(\{\cdots, a-1, a\})}{\pi(\{\cdots, a-2, a-1\})}, \quad a = \cdots, -1, 0, +1, \ldots.$$

References

BOCHNER, S.: Harmonic Analysis and the Theory of Probability. Berkeley: Univ. of Calif. Press 1955.

BÖGE, W.: Über die Charakterisierung unendlich teilbarer Wahrscheinlichkeitsverteilungen. J. reine angew. Math. **201**, 150—156 (1959).

CHUNG, K. L.: Markov Chains with Stationary Transition Probabilities. Berlin-Göttingen-Heidelberg: Springer 1960.

CIGLER, J.: Folgen normierter Masse auf kompakten Gruppen. Z. Wahrscheinlichkeitstheorie **1**, 3—13 (1962).

— Über die Grenzverteilung von Summen Markowscher Ketten auf endlichen Gruppen. Z. Wahrscheinlichkeitstheorie **1**, 415—432 (1962).

CLIFFORD, A. H., and G. B. PRESTON: The Algebraic Theory of Semigroups, Vol. 1. Math. Surveys, Number 7. Providence: Amer. Math. Soc. 1961.

COHEN, P. J.: On a conjecture of Littlewood and idempotent measures. Amer. J. Math. **82**, 191—212 (1960).

COLLINS, H. S.: Primitive idempotents in the semigroup of measures. Duke math. J. **27**, 397—400 (1960).

— The kernel of a semigroup of measures. Duke math. J. **28**, 387—392 (1961).

— Convergence of convolution iterates of measures. Duke math. J. **29**, 259—264 (1962).

— Idempotent measures on compact semigroups. Proc. Amer. math. Soc. **13**, 442—446 (1962).

—, and R. J. KOCH: Regular $\mathscr{D}$-classes in measure semigroups. Trans. Amer. math. Soc. **105**, 21—31 (1962).

DOOB, J. L.: Stochastic Processes. New York: Wiley 1953.

DVORETZKY, A., and J. WOLFOWITZ: Sums of random integers reduced modulo m. Duke math. J. **18**, 501—507 (1951).

FELLER, W.: An Introduction to Probability Theory and its Applications, Vol. 1, Second Edition. New York: Wiley 1957.

GLICKSBERG, I.: Convolution semigroups of measures. Pacific J. Math. **9**, 51—67 (1959).

— Weak compactness and separate continuity. Pacific J. Math. **11**, 205—214 (1961).

GRENANDER, U.: Probabilities on Algebraic Structures. Stockholm: Almqvist & Wiksell 1963.

GYIRES, B.: Die Grenzwertaufgabe der Summenverteilungen. Comptes Rendus du Premier Congrès des Mathématiciens Hongrois, 741—758. Budapest: Akadémiai Kiadó 1952.

HALMOS, P. R.: Measure Theory. New York: Van Nostrand 1950.

HEBLE, M., and M. ROSENBLATT: Idempotent measures on a compact topological semigroup. Proc. Amer. math. Soc. **14**, 177—184 (1963).

HELSON, H.: Note on harmonic functions. Proc. Amer. math. Soc. **4**, 686—691 (1953).

HEWITT, E., and H. S. ZUCKERMAN: Arithmetic and limit theorems for a class of random variables. Duke math. J. **22**, 595—615 (1955).

HLAWKA, E.: Statistik auf kompakten Gruppen I. Österreich. Akad. Wiss. math.-naturw. Kl. Anzeiger **96**, 64—76 (1959).

HORTON, H. B.: A method for obtaining random numbers. Ann. math. Statistics **19**, 81—85 (1948).

—, and R. T. SMITH III: A direct method for producing random digits in any number system. Ann. math. Statistics **20**, 82—90 (1949).

KAKEHASHI, T.: Stationary periodic distributions. J. Osaka Inst. Sci. Technology **1**, 21—25 (1949).

KAWADA, Y., and K. ITÔ: On the probability distribution on a compact group. I. Proc. phys.-math. Soc. Japan **22**, 977—998 (1940).

KLOSS, B. M.: Grenzwertverteilungen für Summen unabhängiger zufälliger Größen, welche Werte aus einer bikompakten Gruppe annehmen (Russisch). Doklady Akad. Nauk SSSR, n. Ser. **109**, 453—455 (1956).

KLOSS, B. M.: Über Wahrscheinlichkeitsverteilungen auf bikompakten topologischen Gruppen (Russisch). Teor. Verojatn. Primen. **4**, 255—290 (1959).
— Grenzverteilungen auf bikompakten abelschen Gruppen (Russisch). Teor. Verojatn. Primen. **6**, 392—421 (1961).
— Über stabile Verteilungen auf einer Klasse lokalkompakter Gruppen (Russisch). Teor. Verojatn. Primen. **7**, 249—270 (1962).
KOUTSKÝ, Z.: Einige Eigenschaften der modulo k addierten Markowschen Ketten. Trans. second Prague Conf., 263—278. Prague: Publ. House of the Czechoslovak Acad. of Sciences 1959.
LÉVY, P.: L'addition des variables aléatoires définies sur une circonférence. Bull. Soc. math. France **67**, 1—41 (1939).
LJAPIN, E. S.: Semigroups. Transl. of Math. Monographs, Vol. 3. Providence: Amer. Math. Soc. 1963.
LOYNES, R. M.: Products of independent random elements in a topological group. Z. Wahrscheinlichkeitstheorie **1**, 446—455 (1963).
— Fourier transforms and probability theory on a non-commutative locally compact topological group. Ark. Mat. **5**, 37—42 (1963).
PRÉKOPA, A., A. RÉNYI und K. URBANIK: Über die Grenzverteilung von Summen unabhängiger zufälliger Variablen auf bikompakten abelschen topologischen Gruppen (Russisch). Acta math. Acad. Sci. Hungar. **7**, 11—16 (1956).
PYM, J. S.: Idempotent measures on semigroups. Pacific J. Math. **12**, 685—698 (1962).
RIVKIND, JA. I.: Ein Grenzwertsatz der Wahrscheinlichkeitstheorie auf kompakten topologischen Gruppen (Russisch). Grosnenskij Gos. Ped. Inst. Uč. Zap. **1**, 51—58 (1955).
ROSENBLATT, M.: Limits of convolution sequences of measures on a compact topological semigroup. J. Math. Mech. **9**, 293—306 (1960).
— Equicontinuous Markov operators. Teor. Verojatn. Primen. **9**, 205—222 (1964).
RUDIN, W.: Idempotent measures on abelian groups. Pacific J. Math. **9**, 195—209 (1959).
— Idempotents in group algebras. Bull. Amer. math. Soc. **69**, 224—227 (1963).
SCHWARZ, Š.: On the structure of the semigroup of measures on a finite semigroup. Czechosl. math. J. **7** (**82**), 358—373 (1957).
— Probability measures on non-commutative semigroups. General Topology and its Relations to Modern Analysis and Algebra, 307—310. Prague: Publishing House of the Czechoslovak Academy of Sciences 1962.
— Probabilities on non-commutative semigroups. Czechosl. math. J. **13** (88), 372—426 (1963).
— Convolution semigroup of measures on compact non-commutative semigroups. Czechosl. math. J. **14** (**89**), 95—115 (1964).
— Product decomposition of idempotent measures on compact semigroups. Czechosl. math. J. **14** (**89**), 121—124 (1964).
STROMBERG, K.: Probabilities on a compact group. Trans. Amer. math. Soc. **94**, 295—309 (1960).
ULAM, S. M.: A Collection of Mathematical Problems. New York: Interscience 1960.
ULLRICH, M., and K. URBANIK: A limit theorem for random variables in compact topological groups. Colloquium math. **7**, 191—198 (1960).
URBANIK, K.: On the limiting probability distribution on a compact topological group. Fundamenta Math. **44**, 253—261 (1957).
— Poisson distributions on compact Abelian topological groups. Colloquium math. **6**, 13—24 (1958).
VOROBJOV, N. N.: Addition unabhängiger zufälliger Größen auf endlichen abelschen Gruppen (Russisch). Mat. Sbornik, n. Ser. **34**, 89—126 (1954).
WENDEL, J. G.: Haar measure and the semigroup of measures on a compact group. Proc. Amer. math. Soc. **5**, 923—929 (1954).

Inst. för Försäkringsmatematik
och Matematisk Statistik
Box 23015
Stockholm 23 (Schweden)

(Received May 25, 1964)