

PROBABILITY DISTRIBUTIONS ON LOCALLY COMPACT ABELIAN GROUPS

BY

K. R. PARTHASARATHY, R. RANGA RAO, AND S. R. S. VARADHAN

1. Introduction

For probability distributions on the real line there are three main theorems on which the entire study of limit theorems for sums of independent random variables is based. These are (1) the Lévy-Khinchin representation of an infinitely divisible distribution, (2) the criteria for weak convergence of such distributions, and (3) Khinchin's theorem on sums of infinitesimal summands stating that these converge weakly if and only if certain associated infinitely divisible laws converge. For a precise statement of these results we refer to Kolmogorov and Gnedenko [3].

During the last two decades or so these results have been extended by many authors to varying degrees of generality. We mention in particular the works of Lévy [12], Kawada and Itô [17], Takano [9], Bochner [1], [2], Hunt [4], Urbanik [13], [14], Kloss [16]. In this paper we study probability distributions on a locally compact abelian (separable) group and obtain definitive extensions of all the three main results mentioned above.

The preliminaries are developed in Section 2. We mention in particular the concept of shift-compactness introduced therein and the important role that Theorem 2.1 plays in our study. A slightly modified notion of an infinitely divisible law is given in this paper to take into account the fact that the group may not be divisible.

The main results of the paper are the following. Weak limits of sums of uniformly infinitesimal random variables (with values in a group) are infinitely divisible. These limits can be obtained from certain accompanying infinitely divisible distributions if they have no idempotent factors. If μ is any infinitely divisible distribution without an idempotent factor, then its characteristic functional $\hat{\mu}(y)$, defined on the character group, has the form

$$(x_0, y) \exp \left\{ \int [(x, y) - 1 - ig(x, y)] dF(x) - \phi(y) \right\}$$

where (x, y) is the value of the character y at x , x_0 is a fixed element of the group, $g(x, y)$ is a fixed function independent of μ , F is a σ -finite measure which integrates the function $(x, y) - 1 - ig(x, y)$ and has finite mass outside each neighborhood of the identity, and $\phi(y)$ is a nonnegative continuous function satisfying the equality

$$\phi(y_1 + y_2) + \phi(y_1 - y_2) = 2[\phi(y_1) + \phi(y_2)].$$

Received December 18, 1961.

Distributions μ with characteristic function of the form

$$(x_0, y) \exp(-\phi(y)),$$

where $\phi(y)$ has the above mentioned property, are precisely those characterized by the following algebraic property: $\mu = \alpha * \beta$, where α is a (generalized) Poisson distribution and β is infinitely divisible, implies that α is degenerate at the identity. We call them Gaussian since they are really so in the case of finite-dimensional Euclidean spaces.

One interesting feature of the representation of an infinitely divisible distribution without idempotent factors is its nonuniqueness. However, the Gaussian component is always unique. If F_1 and F_2 are two measures which occur in two different representations, their difference is always concentrated in a subgroup H of X characterized by the following property: H is the smallest closed subgroup containing all compact subgroups of X . Thus nonuniqueness is due to the prevalence of compact subgroups. If, however, we take a weakly continuous one-parameter convolution semigroup of measures $\{\mu_t\}$, then there exists a unique representation

$$\hat{\mu}_t(y) = (x_t, y) \exp \left[t \left(\int [(x, y) - 1 - iq(x, y)] dF - \phi(y) \right) \right].$$

In the last section we prove the following generalization of a result due to Khinchin [5] on the factorization of arbitrary probability distributions. Any distribution is the convolution of the normalized Haar measure of a compact subgroup, a finite or a countable number of indecomposable distributions, and an infinitely divisible distribution without indecomposable factors.

2. Preliminaries

2.1. All groups considered in this paper are locally compact abelian separable metric groups. Let X denote such a group, and let Y be its character group. For $x \in X$ and $y \in Y$, let (x, y) denote the value of the character y at x . By duality theory the relation between X and Y is perfectly symmetric, i.e., X is the character group of Y . Further, if G is a closed subgroup of X and H is the annihilator of G in Y , i.e., the set

$$H = [y: (x, y) = 1 \text{ for all } x \in G],$$

then G and Y/H are character groups of each other. These facts and some well-known results on the structure of locally compact abelian groups will be freely used in the sequel. For these details we refer to A. Weil [10], and Pontrjagin [18].

2.2. By a measure on X we shall mean a nonnegative completely additive set function defined on the Borel σ -field of subsets of X . We shall refer to probability measures as distributions. Let \mathfrak{M} denote the class of all dis-

tributions. For $\lambda, \mu \in \mathfrak{M}$ we write

$$(\lambda * \mu)(E) = \int \mu(E - x) d\lambda$$

for any Borel set E . $\lambda * \mu$ is a distribution obtained by the convolution operation. With this operation \mathfrak{M} becomes a commutative semigroup. If μ is the distribution degenerate at a point $x \in X$, then we write $\lambda * x$ for $\lambda * \mu$. We call $\lambda * x$ the shift or translate of λ by the element x . If $\lambda_1, \lambda_2, \dots, \lambda_n$ are some n distributions, then the distribution $\mu = \lambda_1 * \dots * \lambda_n$ has an obvious meaning. We denote μ by $\prod_{i=1}^n \lambda_i$. If all the distributions λ_i are identically equal to a single distribution λ , we write $\mu = \lambda^n$. These definitions can obviously be extended to all measures, and we shall have occasion to use them.

For any measure μ we write $\bar{\mu}(A) = \mu(-A)$ where $-A$ is the set of all inverses of elements in A . Then $\bar{\mu}$ is also a measure. We denote by $|\mu|^2$ the measure $\mu * \bar{\mu}$.

We introduce in \mathfrak{M} the weak topology, defined as follows: For $\mu_n, \mu \in \mathfrak{M}$, μ_n is said to converge to μ (or $\mu_n \Rightarrow \mu$) if $\int f d\mu_n \rightarrow \int f d\mu$ for each bounded continuous function f on X . We assume some familiarity with this convergence and in particular the description of compact subsets in this topology.

2.3. For each $\mu \in \mathfrak{M}$ its characteristic function $\hat{\mu}(y)$ is a function on the character group Y , defined as follows:

$$\hat{\mu}(y) = \int_X (x, y) d\mu(x).$$

Some of the basic properties of the characteristic function $\hat{\mu}(y)$ are given below.

- (1) $\hat{\mu}(y)$ is a uniformly continuous function of y ,
- (2) $\hat{\mu}(y)$ determines μ uniquely,
- (3) $(\mu * \lambda)^\wedge(y) = \hat{\mu}(y)\hat{\lambda}(y)$ for all $y \in Y$, and $\mu, \lambda \in \mathfrak{M}$,
- (4) $\widehat{\bar{\mu}}(y) = \overline{\hat{\mu}(y)}$,
- (5) $\mu_n \Rightarrow \mu$ if and only if $\hat{\mu}_n(y) \rightarrow \hat{\mu}(y)$ uniformly over compact subsets of Y ,
- (6) if $\hat{\mu}_n(y)$ converges to a limit uniformly on each compact subset of Y , then there is a $\mu \in \mathfrak{M}$ such that

$$(i) \hat{\mu}(y) = \lim_{n \rightarrow \infty} \hat{\mu}_n(y) \quad \text{and} \quad (ii) \mu_n \Rightarrow \mu.$$

2.4. A distribution μ is said to be idempotent if $\mu^2 = \mu * x$ for some $x \in X$. If we write $\lambda = \mu * (-x)$, then it is clear that $\lambda^2 = \lambda$, so that $\hat{\lambda}(y) = 0$ or 1 . From the inequality

$$(2.1) \quad 1 - R(x, y_1 + y_2) \leq 2[(1 - R(x, y_1)) + (1 - R(x, y_2))]$$

(R denoting the real part) it is clear that the set of all y for which $\hat{\lambda}(y) = 1$ is a both open and closed subgroup of Y . It is not difficult to see that the

annihilator G of this subgroup is compact and λ is the normalized Haar measure of G . Thus μ is the translate of the Haar distribution of a compact subgroup.

2.5. For $\alpha, \mu \in \mathfrak{M}$, α is said to be a factor of μ ($\alpha < \mu$ in symbols) if there is a $\beta \in \mathfrak{M}$ such that $\mu = \alpha * \beta$. A distribution μ whose only factors are either degenerate distributions or translates of μ is said to be indecomposable. We denote by $F(\mu)$ the collection of all factors of μ .

Two distributions α, β are said to be equivalent ($\alpha \sim \beta$ in symbols) if each is a shift of the other. “ \sim ” is an equivalence relation. For each α , $\tilde{\alpha}$ denotes the equivalence class containing α . The collection $\tilde{\mathfrak{M}}$ of equivalence classes forms a semigroup. $\tilde{\mathfrak{M}}$ will be endowed with the quotient topology.

DEFINITION 2.1. A subset $\mathfrak{N} \subset \mathfrak{M}$ is said to be *shift-compact* if its image $\tilde{\mathfrak{N}}$ in $\tilde{\mathfrak{M}}$ is conditionally compact.

The following theorem proved in [8] (see Section 3) plays a fundamental role in our study.

THEOREM 2.1. *Let \mathfrak{N} be a conditionally compact subset of \mathfrak{M} . Then $F(\mathfrak{N})$ consisting of the totality of all factors of elements belonging to \mathfrak{N} is shift-compact.*

COROLLARY 1. *A subset $\mathfrak{N} \subset \mathfrak{M}$ is shift-compact if and only if the set $|\mathfrak{N}|^2$ consisting of all elements $\alpha \in \mathfrak{M}$ of the form $|\mu|^2$ with $\mu \in \mathfrak{N}$, is conditionally compact.*

COROLLARY 2. *For any distribution μ , $F(\mu)$ is shift-compact.*

COROLLARY 3. *Suppose $\alpha_1 < \alpha_2 < \dots < \alpha_n < \dots$ and $\alpha_n < \mu$ for all n . Then there is a translate α'_n of α_n for every n such that α'_n is weakly convergent.*

Proof of Corollary 3. Since $\alpha_n \in F(\mu)$ for each n , $\{\alpha_n\}$ is shift-compact. Suppose that β and β' are any two limits of shifts of α_n . Then the fact that $\alpha_n < \alpha_{n+1}$ implies that β and β' are translates of each other. The corollary is an easy consequence of this fact.

The following corollary is proved in a similar manner.

COROLLARY 4. *Suppose $\alpha_1 > \alpha_2 \dots$. Then there exists a translate α'_n of α_n for each n such that α'_n weakly converges to a limit.*

Remark. Corollaries 2–4 have been obtained earlier by Itô [11] for distributions in the real line.

3. Two auxiliary lemmas

LEMMA 3.1. *For each compact set $C \subset Y$ there are a neighborhood N_C of the identity in X and a finite set $E \subset C$ such that*

$$\sup_{y \in C} [1 - R(x, y)] \leq M \cdot \sup_{y \in E} [1 - R(x, y)]$$

for all $x \in N_c$, where M is a finite constant depending on C . (Here $R(x, y)$ denotes the real part of (x, y) .)

Proof. From the inequality

$$1 - R(x_1 + x_2, y) \leq 2[(1 - R(x_1, y)) + (1 - R(x_2, y))]$$

it is clear that if the lemma is valid in two groups X_1 and X_2 , it is valid for their direct sum $X_1 \oplus X_2$. Let now Y' denote the closed subgroup generated by C , and Φ its annihilator in X . If τ denotes the canonical homomorphism from X onto $X' = X/\Phi$, it is obvious that $R(x, y) = R(\tau(x), y)$ for all $x \in X$ and $y \in Y'$. It is thus sufficient to prove the lemma when the groups concerned are X' and Y' instead of X and Y . Since Y' is compactly generated, it is of the form $V \oplus C \oplus I'$ where V is a finite-dimensional vector group, C is a compact group, and I' is the product of r copies of the integer group. Hence X' is of the form $V \oplus D \oplus K'$ where D is a discrete group and K' is the product of r copies of the circle group. Since the lemma is trivially valid in the case of the real line, discrete group, and compact group, the proof of the lemma is complete.

LEMMA 3.2. For any $y \in Y$ there is a continuous function $h_y(x)$ on X with the following properties:

- (1) $|h_y(x)| \leq \pi$ for all $x \in X$, and $h_y(-x) = -h_y(x)$,
- (2) $(x, y) = \exp(ih_y(x))$ for all $x \in N_y$ where

$$N_y = [x : |(x, y) - 1| \leq \frac{1}{2}].$$

Proof. Let $(x, y) = \exp(i\phi(x))$ where $-\pi \leq \phi(x) < \pi$. Then it is not difficult to verify that $\phi(x)$ is a continuous function of x in the closed set N_y . Now choose any continuous extension of $\phi(x)$ to X such that the first condition is fulfilled. This will serve our purpose.

LEMMA 3.3. There is a function $g(x, y)$ defined on the product space $X \times Y$ possessing the following properties:

- (1) $g(x, y)$ is a continuous function of both the variables x and y .
- (2) $\sup_{x \in X} \sup_{y \in C} |g(x, y)| < \infty$ for each compact set $C \subset Y$.
- (3) $g(x, y_1 + y_2) = g(x, y_1) + g(x, y_2)$ for each $x \in X$ and $y_1, y_2 \in Y$, and $g(-x, y) = -g(x, y)$.
- (4) If C is any compact subset of Y , then there is a neighborhood N_c of the identity in X such that $(x, y) = \exp[ig(x, y)]$ for all $x \in N_c$ and $y \in C$.
- (5) If C is any compact subset of Y , then $g(x, y)$ tends to zero uniformly in $y \in C$ as x tends to the identity of the group X .

Proof. We shall reduce the proof of the proposition to the case of certain simple groups by making use of the structure theory. Suppose that the proposition is true for an open subgroup G of X . Let H and Y be the character groups of G and X respectively. Since H can be obtained as a quotient group of Y by taking the quotient with respect to the annihilator of G in Y ,

there is a canonical homomorphism τ from Y to H . Suppose $g(x, h)$ has been defined for $x \in G$ and $h \in H$ with the required properties. We extend the definition of g as follows. For $x \in G$ and $y \in Y$ we define

$$g(x, y) = g(x, \tau(y)).$$

For $x \notin G$, we define

$$g(x, y) = 0 \quad \text{for all } y \in Y.$$

Since an open subgroup is closed, the continuity of $g(x, y)$ follows immediately. The rest of the properties of $g(x, y)$ are immediate consequences of their validity in $G \times H$.

In the case of a general group X we take G to be the group generated by a compact neighborhood of the identity. This is both open and closed in X . This group G has the simple structure, $V \oplus C \oplus I^r$ where V is a finite-dimensional vector group, C is a compact group, and I^r is the product of the integer group taken r times. We now observe that if functions $g_1(x, y)$ and $g_2(u, v)$ with the properties mentioned in the lemma exist in groups X and U with character groups Y and V respectively, then a function $g(\xi, \eta)$ with the same properties exists for $\xi \in X \oplus U$ and $\eta \in Y \oplus V$. We have only to define

$$g(\xi, \eta) = g_1(x, y) + g_2(u, v)$$

where x and u are projections of ξ into X and U respectively, and y and v are projections of η into Y and V respectively. Thus it is enough to construct $g(x, y)$ in the case of a real line, a compact group, and the integer group. In the case of the integer group we can take $g(x, y)$ to be identically zero. In the case of the real line we can take $g(x, y) = \theta(x)y$ where

$$\begin{aligned} \theta(x) &= x, & x &\in [-1, 1], \\ &= 1, & x &> 1, \\ &= -1, & x &< -1. \end{aligned}$$

(Note that the character group of the real line is itself.) Thus, in order to complete the proof of the lemma it is enough to consider the case of a compact group X .

Let X be a compact group with Y as character group. Let X_0 be the component of identity in X , Y_1 the annihilator of X_0 in Y , $X_1 = X/X_0$, and $Y_0 = Y/Y_1$. Then Y_0 is the character group of X_0 , and Y_1 is the character group of X_1 . Since X_0 is connected and compact, Y_0 is a discrete torsion-free group. Let $\{d_\alpha\}$ be a maximal family of mutually independent elements in Y_0 . Then for $d \in Y_0$ there exist elements $d_{\alpha_1}, \dots, d_{\alpha_k}$ from the maximal family and integers n, n_1, \dots, n_k ($n > 0$) such that

$$(3.1) \quad nd = n_1 d_{\alpha_1} + \dots + n_k d_{\alpha_k}.$$

This representation is unique except for multiplication by an integer on both sides.

Each element of Y_0 is a coset of Y_1 in Y . We take the coset d_α and pick

out an element y_α of Y from this coset. We fix the elements y_α . We define

$$g(x, y_\alpha) = h_{y_\alpha}(x)$$

for every α where $h_{y_\alpha}(x)$ is as in Lemma 3.2. Let now $y \in Y$ be arbitrary. Then y belongs to some coset of Y_1 which is an element of Y_0 . If this element is denoted by d , then there exist integers n, n_1, \dots, n_k ($n > 0$) and elements $d_{\alpha_1}, \dots, d_{\alpha_k}$ from the collection $\{d_\alpha\}$ such that equation (3.1) is satisfied. We define

$$g(x, y) = (n_1/n)g(x, y_{\alpha_1}) + \dots + (n_k/n)g(x, y_{\alpha_k}).$$

We shall now prove that the function $g(x, y)$ constructed in this way has all the required properties.

Since $g(x, y)$ is continuous in x for each fixed y and Y is discrete, the continuity of $g(x, y)$ in both the variables follows immediately. Properties (2) and (3) are obvious from the nature of the construction.

Since compact sets in Y are finite sets, it is enough to prove property (4) for each $y \in Y$. For any $y \in Y$, let $[y]$ denote the coset of Y_0 to which y belongs. Then $[y]$ is an element of Y_0 . If we write $[y]$ for d and $[y_\alpha]$ for d_α , then equation (3.1) can be written as

$$(3.2) \quad n[y] = n_1[y_{\alpha_1}] + \dots + n_k[y_{\alpha_k}].$$

For any two elements $y_1, y_2 \in [y]$ it is clear that $y_1 - y_2 \in Y_1$. Since Y_1 is the character group of X/X_0 which is totally disconnected, every element of Y_1 is of finite order. Hence, for any $y \in Y_1$ there exists a neighborhood of the identity in X where $(x, y) = 1$. Thus for any two elements $y_1, y_2 \in [y]$ there exists a neighborhood of the identity in X where $(x, y_1) = (x, y_2)$.

Making use of the remarks made in the previous paragraph we shall complete the proof of the lemma. From the construction of $g(x, y)$ and Lemma 3.2 it is clear that, for each y_α , there exists a neighborhood of the identity in X where $e^{ig(x, y_\alpha)} = (x, y_\alpha)$. Let now $y \in Y$ be arbitrary. From (3.2) it is clear that there exist elements $y_{\alpha_j}, \dots, y_{\alpha_j}$ in $[y_{\alpha_j}]$, for $j = 1, 2, \dots, k$, such that

$$(3.3) \quad ny = \sum_{j=1}^k (y_{\alpha_j} + \dots + y_{\alpha_j}).$$

From the remarks made in the previous paragraph it follows that there exists a neighborhood of the identity in X (depending on y_{α_j} and y_{α_j}) where

$$(x, y_{\alpha_j r}) = (x, y_{\alpha_j}).$$

Denoting by N the intersection of all the neighborhoods corresponding to $y_{\alpha_j r}$ ($r = 1, 2, \dots, n_j$ and $j = 1, 2, \dots, k$) we have

$$(3.4) \quad (x, y_{\alpha_j r}) = (x, y_{\alpha_j}) \quad \text{for } x \in N, r = 1, 2, \dots, n_j, j = 1, 2, \dots, k.$$

From (3.3) and (3.4) we have

$$(x, y)^n = (x, ny) = \prod_{j=1}^k (x, y_{\alpha_j})^{n_j} \quad \text{for } x \in N.$$

Since there are neighborhoods of the identity where $(x, y_{\alpha_i}) = \exp \{ig(x, y_{\alpha_i})\}$, it follows that there exists a neighborhood of the identity where

$$(x, y)^n = e^{ing(x,y)}.$$

Since (x, y) and $e^{ig(x,y)}$ are continuous functions nonvanishing at the identity of X , there exists a neighborhood of the identity where

$$(x, y) = e^{ig(x,y)}.$$

Property (5) is obvious from property (1) and the fact that $g(x, y)$ vanishes when either x or y is the identity of the corresponding group. This completes the proof of the lemma.

In the following paragraphs we give examples of the function $g(x, y)$ for some particular groups.

Example 1. Let $X = Y = R^n$. If $x = (x_1, \dots, x_n) \in X$ and $y = (y_1, y_2, \dots, y_n) \in Y$, then

$$g(x, y) = \sum_{i=1}^n \phi_i(x_i)y_i$$

where $\phi_i(t)$ ($i = 1, 2, \dots, n$) are bounded continuous functions on the real line such that $\phi_i(t) = t$ in a neighborhood of $t = 0$ and $\phi_i(-t) = -\phi_i(t)$.

Example 2. Let K denote the circle group, and let $X = K^n$ and $Y = I^n$, I being the integer group. Let

$$X = [(x_1, \dots, x_n) : -1 < x_i \leq 1 \text{ for } i = 1, 2, \dots, n]$$

addition being taken modulo 2. Then if $y = (y_1, \dots, y_n) \in I^n$,

$$g(x, y) = \sum_{i=1}^n \phi_i(x_i)y_i$$

where the functions $\phi_i(t)$ are as in Example 1.

Example 3. $Y =$ additive group of rationals, and X is its character group. Let $\phi(x)$ be a bounded continuous function on X such that $\exp(i\phi(x)) = (x, y_0)$ for x in a neighborhood of the identity in X , where y_0 is a fixed element of Y , other than the identity. Then

$$g(x, y) = \phi(x)y/y_0.$$

Example 4. If X is totally disconnected, then every homomorphism of Y into the real line is trivial, so that in this case $g(x, y) \equiv 0$ for all $x \in X$ and $y \in Y$.

4. Infinitely divisible distributions

We shall now introduce the definition of infinitely divisible distributions and study some of their elementary properties.

DEFINITION 4.1. A distribution μ is said to be *infinitely divisible* if, for each n , there are elements $x_n \in X$ and $\lambda_n \in \mathfrak{M}$ such that $\mu = \lambda_n^n * x_n$.

We remark that the definition given here is slightly different from the classical definition in the case of the real line. Such a modification is necessary if we want to avoid the role of divisibility of elements from the group. As yet it is not clear whether there exists a single element x of the group X with the property $\mu = \lambda_n^n * x$ for every n . That such is the case for every infinitely divisible distribution will be obvious from the representation which we shall give later in Section 7.

THEOREM 4.1. *The infinitely divisible distributions form a closed subsemigroup of \mathfrak{M} .*

Proof. If λ and μ are infinitely divisible, it is obvious from the definition that $\lambda * \mu$ is also infinitely divisible. Let now $\mu_k, k = 1, 2, \dots$, be a sequence of infinitely divisible distributions weakly converging to μ . For any fixed integer n , let

$$(4.1) \quad \mu_k = \lambda_{k_n}^n * x_{k_n} .$$

From Theorem 2.1 it is clear that there exists a subsequence of λ_{k_n} which after a suitable shift converges to a distribution λ_n . Since $\mu_k \Rightarrow \mu$, it is obvious from (4.1) that there exists an element x_n such that $\mu = \lambda_n^n * x_n$. This completes the proof.

The normalized Haar measure of a compact subgroup is an example of an infinitely divisible distribution. We shall now prove a result concerning the absence of zeros for the characteristic function of an infinitely divisible distribution without idempotent factors.

THEOREM 4.2. *Let $\hat{\mu}(y)$ be the characteristic function of an infinitely divisible distribution μ . If $\hat{\mu}(y_0) = 0$ for some character y_0 , then μ has an idempotent factor.*

Proof. From the definition of infinite divisibility it follows that, for each n , there exist an element $x_n \in X$ and a distribution λ_n such that $\mu = \lambda_n^n * x_n$. Since $\hat{\mu}(y_0) = 0$, $\hat{\lambda}_n(y_0)$ also vanishes for every n . By Theorem 2.1, λ_n is shift-compact. Let λ be a limit of shifts of λ_n . Then $\hat{\lambda}(y_0) = 0$ and is hence a nondegenerate distribution. It is also clear that every power of λ is a factor of μ . Thus the sequence λ^n is shift-compact, and any limit of shifts is a nondegenerate idempotent factor of μ . This completes the proof.

DEFINITION 4.2. If F is any totally finite measure on X the *distribution $e(F)$ associated with F* is defined as follows:

$$e(F) = e^{-F(X)} [1 + F + F^2/2! + \dots + F^n/n! + \dots]$$

where 1 is used to denote the measure with unit mass and degenerate at the identity.

$e(F)$ is obviously an infinitely divisible distribution since $e(F) = [e(F/n)]^n$.

Its characteristic function is given by

$$(e(F))^{\wedge}(y) = \exp \left[\int ((x, y) - 1) dF(x) \right].$$

Suppose F_n is a sequence of totally finite measures and we form the sequence $e(F_n)$. We shall now obtain a necessary condition (which will be shown to be sufficient in Section 9) for the shift-compactness of $e(F_n)$.

THEOREM 4.3. *Let $\mu_n = e(F_n)$ where F_n is a sequence of totally finite measures. Then, in order that*

(a) μ_n be shift-compact,
 (b) if μ is any limit of shifts of $\{\mu_n\}$, then μ have no idempotent factor,
 the following conditions are necessary:

- (i) For each neighborhood N of the identity the family $\{F_n\}$ restricted to $X - N$ is weakly conditionally compact.
- (ii) For each $y \in Y$

$$\sup_n \int [1 - R(x, y)] dF_n < \infty,$$

$R(x, y)$ denoting the real part of (x, y) .

Before proceeding to the proof of this theorem we shall prove the following.

LEMMA 4.1. *Let F_n be as in Theorem 4.3, and suppose $\sup_n F_n(N') \leq k$ where N' denotes the complement of a symmetric neighborhood N of the identity. If the sequence $e(F_n)$ is shift-compact, then the sequence of measures F_n is tight when restricted to N' .*

Proof. Let G_n denote the restriction of F_n to N' . Then $e(G_n)$ is a factor of $e(F_n)$. Since $\{e(F_n)\}$ is shift-compact, so is the sequence $\{e(G_n)\}$ by Theorem 2.1. Let $H_n = G_n + \bar{G}_n$. Then, by Corollary 1 of Theorem 2.1, the sequence $\{e(H_n)\}$ is compact. Hence, for any $\varepsilon > 0$ there exists a compact set C such that $e(H_n)(C') < \varepsilon$. Since $e(H_n) = e^{-H_n(x)} [\sum_{r=0}^{\infty} H_n^r / r!]$, we have

$$\begin{aligned} \varepsilon > e(H_n)(C') &\geq e^{-H_n(x)} H_n(C') \geq e^{-2k} H_n(C') \\ &\geq e^{-2k} G_n(C') \end{aligned}$$

for all n . Since k is a constant not depending on n and ε is arbitrary, it follows that the family $\{G_n\}$ is tight.

Proof of Theorem 4.3. Since any neighborhood of the identity contains a symmetric neighborhood, we can assume that N is symmetric. Suppose (a) and (b) are valid. Let, if possible, $\sup_n F_n(N') = \infty$. We can then choose a subsequence for which

$$(4.2) \quad F_{n_k}(N') \geq 2k \quad \text{for } k = 1, 2, \dots$$

Let $L_k, k = 1, 2, \dots$, be measures such that

$$(4.3) \quad \begin{aligned} L_k(A) &\leq (1/k)F_{n_k}(A) && \text{for every Borel set } A, \\ L_k(N) &= 0, \\ L_k(N') &= 1. \end{aligned}$$

The distribution $\lambda_k = e(L_k)$ is a factor of $e(F_{n_k})$, and the shift-compactness of $e(F_n)$ implies the shift-compactness of $\{\lambda_k\}$ by Theorem 2.1. Let λ be any limit of shifts of λ_k . From (4.2) and (4.3) it follows that any power of λ is a factor of μ . Thus the sequence λ^n is shift-compact, and any limit of shifts of λ^n is a factor of μ . Since these limits will be idempotent and μ has no idempotent factors, it follows that any such limit must be degenerate. Since λ^n is an increasing sequence (in the order $<$), it follows that λ itself must be degenerate. Thus the sequence $|\lambda_k|^2$ converges to the distribution degenerate at the identity. Hence

$$e(L_k + \bar{L}_k)(N') \rightarrow 0 \quad \text{as } k \rightarrow \infty.$$

But

$$e(L_k + \bar{L}_k)(N') = e^{-(L_k + \bar{L}_k)(x)} \sum \frac{(L_k + \bar{L}_k)^r}{r!} (N') \geq e^{-2} L_k(N') = e^{-2},$$

which is a contradiction. Thus we have $\sup_n F_n(N') < \infty$. Now an application of Lemma 4.1 shows that condition (i) of the theorem is necessary.

In order to prove the necessity of (ii) we observe that $e(F_n + \bar{F}_n) = |e(F_n)|^2$ is a compact sequence, and an application of Theorem 4.2 shows that any limit of $|e(F_n)|^2$ has a nonvanishing characteristic function. Thus, for any y ,

$$\lim_n \exp \left\{ \int [R(x, y) - 1] d(F_n + \bar{F}_n) \right\} \neq 0,$$

which implies condition (ii). The proof is thus complete.

5. General limit theorems for sums of infinitesimal summands

In the case of the real line a well-known result due to Bawly and Khinchin (see [3, Chapter 4]) asserts that the limit of sums of infinitesimal random variables is infinitely divisible, and it can be obtained as the limit of a certain accompanying sequence of infinitely divisible distributions. The purpose of this section is to introduce the notion of infinitesimal distributions in a group and prove a generalized version of the above-mentioned result in the case when the limiting distribution has no idempotent factor.

DEFINITION 5.1. A triangular sequence $\{\alpha_{nj}\}, j = 1, 2, \dots, k_n$ of distributions is said to be *uniformly infinitesimal* if

$$\lim_{n \rightarrow \infty} \sup_{1 \leq j \leq k} \sup_{y \in K} |\hat{\alpha}_{nj}(y) - 1| = 0$$

for each compact set $K \subset Y$.

Before going to the statement of the main result of this section we shall prove a lemma which will be often used in the sequel.

LEMMA 5.1. *Let $\mu_n = \prod_{j=1}^{k_n} \alpha_{nj}$ where the sequence $\{\alpha_{nj}\}$ is uniformly infinitesimal. If the distribution μ is a limit of shifts of μ_n , then the set of characters $[y: \hat{\mu}(y) \neq 0]$ is an open subgroup of Y , and consequently the normalized Haar measure of the annihilator of this subgroup in X is a factor of μ .*

Proof. Since μ is a limit of shifts of μ_n , it is clear that for a subsequence (which we shall denote by μ_n itself) $|\mu_n|^2 \Rightarrow |\mu|^2$, and hence

$$(5.1) \quad \lim_{n \rightarrow \infty} \prod_{j=1}^{k_n} |\hat{\alpha}_{nj}(y)|^2 = |\hat{\mu}(y)|^2.$$

If $\hat{\mu}(y) \neq 0$, it is obvious that $\hat{\mu}(-y) \neq 0$. (5.1) implies that a necessary and sufficient condition that $\hat{\mu}(y) \neq 0$ is that

$$(5.2) \quad \sup_n \sum_{j=1}^{k_n} (1 - |\hat{\alpha}_{nj}(y)|^2) < \infty.$$

Thus if we make use of the inequality

$$1 - \phi(y_1 + y_2) \leq 2[(1 - \phi(y_1)) + (1 - \phi(y_2))]$$

for any real-valued characteristic function ϕ , it is clear that the validity of (5.2) for y_1 and y_2 implies its validity for $y_1 + y_2$. The continuity of $\hat{\mu}(y)$ now implies that the set $[y: \hat{\mu}(y) \neq 0]$ is an open subgroup of Y .

We choose and fix a function $g(x, y)$ defined on $X \times Y$ and satisfying all the properties mentioned in Lemma 3.3. The main theorem of this section can now be stated as follows:

THEOREM 5.1. *Let $\{\alpha_{nj}\}$ be a uniformly infinitesimal sequence of distributions, and let*

$$\mu_n = \prod_{j=1}^{k_n} \alpha_{nj}.$$

Suppose that $\{\mu_n\}$ is shift-compact such that no limit of shifts of μ_n has an idempotent factor. Let

$$\beta_{nj} = e(\alpha_{nj} * x_{nj})$$

where x_{nj} is that element of the group X defined by the equality

$$(x_{nj}, y) = \exp \left[-i \int g(x, y) d\alpha_{nj} \right].$$

*If $\lambda_n = (\prod_{j=1}^{k_n} \beta_{nj}) * x_n$, where $x_n = -\sum_j x_{nj}$, then*

$$\lim_{n \rightarrow \infty} \sup_{y \in K} |\hat{\lambda}_n(y) - \hat{\mu}_n(y)| = 0$$

for each compact set K of Y .

Proof. During the course of the proof of the theorem we shall adopt the following conventions: We denote by C_1, C_2, \dots constants depending only on the compact set K (and not on n). All the statements that we make

are for sufficiently large n . By N we denote any arbitrarily small neighborhood of the identity in X .

Turning to the proof of the theorem, we observe that the elements x_{n_j} are well defined, since, from the properties of $g(x, y)$, it follows that $\exp [-i \int g(x, y) d\alpha]$ is a character on Y for any distribution α . Further, for any neighborhood N of the identity all the points x_{n_j} are in N for sufficiently large n . Therefore the uniform infinitesimality of $\{\alpha_{n_j}\}$ implies the uniform infinitesimality of the sequence $\{\beta_{n_j}\}$. Thus $\hat{\lambda}_n(y)$ and $\hat{\mu}_n(y)$ are nonvanishing in K , and hence we use the logarithmic notation freely. Since no limit of shifts of μ_n has an idempotent factor, it follows from Lemma 5.1 that the sequence $\hat{\mu}_n(y)$ is uniformly bounded away from zero for all $y \in K$. Thus it is enough to prove that

$$\lim_{n \rightarrow \infty} \sup_{y \in K} | \log \hat{\lambda}_n(y) - \log \hat{\mu}_n(y) | = 0.$$

We have

$$\begin{aligned} \log \hat{\lambda}_n(y) &= \sum_j \log \hat{\beta}_{n_j}(y) - \sum_j \log (x_{n_j}, y) \\ &= \sum_j \log \hat{\beta}_{n_j}(y) + i \sum_j \int g(x, y) d\alpha_{n_j}(x) \\ &= \sum_j [(\alpha_{n_j} * x_{n_j})^\wedge(y) - 1] + i \sum_j \int g(x, y) d\alpha_{n_j}(x), \end{aligned}$$

and

$$\log \hat{\mu}_n(y) = \sum_j \log \hat{\alpha}_{n_j}(y).$$

Writing $\theta_{n_j} = \alpha_{n_j} * x_{n_j}$, we obtain

$$\begin{aligned} | \log \hat{\lambda}_n(y) - \log \hat{\mu}_n(y) | &= \left| \sum_j (\hat{\theta}_{n_j}(y) - 1) + i \sum_j \int g(x, y) d\alpha_{n_j} \right. \\ &\quad \left. - \sum_j \log \hat{\theta}_{n_j}(y) + \sum_j \log (x_{n_j}, y) \right| \\ &= | \sum_j (\hat{\theta}_{n_j}(y) - 1) - \sum_j \log \hat{\theta}_{n_j}(y) | \\ &\leq C_1 (\sum_j | 1 - \hat{\theta}_{n_j}(y) |) \sup_j | 1 - \hat{\theta}_{n_j}(y) |. \end{aligned}$$

Since $\{\theta_{n_j}\}$ is uniformly infinitesimal, it is clear from the above inequality that it is enough to prove that

$$(5.3) \quad \sup_n \sup_{y \in K} [\sum | 1 - \hat{\theta}_{n_j}(y) |] < \infty.$$

We have, for any neighborhood N of the identity in X

$$\begin{aligned} (5.4) \quad | 1 - \hat{\theta}_{n_j}(y) | &\leq \left| \int_N (1 - (x, y)) d\theta_{n_j} \right| + \left| \int_{N'} (1 - (x, y)) d\theta_{n_j} \right| \\ &\leq \left| \int_N (1 - (x, y)) d\theta_{n_j} \right| + 2\theta_{n_j}(N'). \end{aligned}$$

From property (4) of $g(x, y)$ in Lemma 3.3 it follows that there exists a neighborhood of the identity in X where

$$(x, y) = e^{ig(x,y)} \quad \text{for } y \in K.$$

In such a neighborhood we have, for $y \in K$,

$$(5.5) \quad |1 - (x, y) + ig(x, y)| \leq C_2 g^2(x, y).$$

(5.4) and (5.5) imply

$$(5.6) \quad |1 - \hat{\theta}_{n_j}(y)| \leq \left| \int_N g(x, y) d\theta_{n_j} \right| + C_2 \int_N g^2(x, y) d\theta_{n_j} + 2\theta_{n_j}(N')$$

for all $y \in K$. By property (2) of $g(x, y)$ in Lemma 3.3

$$(5.7) \quad \left| \int_X g(x, y) d\theta_{n_j} \right| = \left| \int_X g(x + x_{n_j}, y) d\alpha_{n_j}(x) \right| \\ \leq \left| \int_N g(x + x_{n_j}, y) d\alpha_{n_j}(x) \right| + C_3 \alpha_{n_j}(N').$$

Since all the x_{n_j} will be in any small neighborhood of the identity after a certain stage, and since $e^{ig(x,y)} = (x, y)$ for $x \in N$ and $y \in K$, we conclude, by making use of property (5) of $g(x, y)$ in Lemma 3.3, that

$$(5.8) \quad g(x + x_{n_j}, y) = g(x, y) + g(x_{n_j}, y)$$

for $x \in N$ and $y \in K$. Further

$$e^{ig(x_{n_j}, y)} = (x_{n_j}, y) = \exp \left\{ -i \int g(x, y) d\alpha_{n_j} \right\}$$

for all $y \in K$ and sufficiently large n . By property (5) of $g(x, y)$ in Lemma 3.3, we get

$$(5.9) \quad g(x_{n_j}, y) = - \int g(x, y) d\alpha_{n_j}.$$

(5.8) and (5.9) imply

$$\left| \int_N g(x + x_{n_j}, y) d\alpha_{n_j} \right| = \left| \int_N (g(x, y) + g(x_{n_j}, y)) d\alpha_{n_j} \right| \\ = \left| \int_N g(x, y) d\alpha_{n_j} - \alpha_{n_j}(N) \int g(x, y) d\alpha_{n_j} \right| \\ = \left| \alpha_{n_j}(N') \int_N g(x, y) d\alpha_{n_j} - \alpha_{n_j}(N) \int_{N'} g(x, y) d\alpha_{n_j} \right| \\ \leq C_4 \alpha_{n_j}(N').$$

The above inequality together with (5.7) implies that

$$(5.10) \quad \left| \int g(x, y) d\theta_{nj} \right| \leq C_5 \alpha_{nj}(N') \quad \text{for } y \in K.$$

(5.6), (5.10), and property (2) of Lemma 3.3 give

$$|1 - \hat{\theta}_{nj}(y)| \leq C_2 \int g^2(x, y) d\theta_{nj} + C_6 \theta_{nj}(N') + C_7 \alpha_{nj}(N')$$

for $y \in K$. Thus, in order to complete the proof of the theorem we have only to show that

$$(5.11) \quad \limsup_n \sum_j \theta_{nj}(N') < \infty,$$

$$(5.12) \quad \limsup_n \sum_j \alpha_{nj}(N') < \infty,$$

$$(5.13) \quad \limsup_n \sup_{y \in K} \sum_j \int_N g^2(x, y) d\theta_{nj} < \infty.$$

To this end we consider the distribution

$$|\mu_n|^2 = \prod_{j=1}^{k_n} |\alpha_{nj}|^2.$$

Since $|\mu_n|^2$ is compact and no limit of $|\mu_n|^2$ has an idempotent factor, according to Lemma 5.1, $|\hat{\mu}_n(y)|^2$ is bounded away from zero uniformly for $y \in K$ and in n . Thus

$$\limsup_n \sup_{y \in K} \sum_j (1 - |\hat{\alpha}_{nj}(y)|^2) < \infty.$$

This is the same as (5.3) with $|\alpha_{nj}|^2$ replacing θ_{nj} , and hence

$$\lim_{n \rightarrow \infty} \sup_{y \in K} [\exp(\sum_j (|\hat{\alpha}_{nj}(y)|^2 - 1)) - |\hat{\mu}_n(y)|^2] = 0.$$

Thus the sequence $e(\sum_j |\alpha_{nj}|^2)$ is compact. We now appeal to Theorem 4.3. Then

$$(5.14) \quad \limsup_n \sum_j |\alpha_{nj}|^2(N') < \infty,$$

$$(5.15) \quad \limsup_n \sum_j \int (1 - R(x, y)) d|\alpha_{nj}|^2 < \infty.$$

We now choose a neighborhood V of the identity such that $V + V \subset N$. Then

$$\begin{aligned} \sum_j \alpha_{nj}(N') &\leq \sum_j \alpha_{nj}((V + V)') \\ &\leq \sum_j \inf_{x \in V} \alpha_{nj}((V + x)') = \sum_j \inf_{x \in V} \alpha_{nj}(V' + x) \\ &\leq \sum_j [\alpha_{nj}(V)]^{-1} \int_V \alpha_{nj}(V' + x) d\alpha_{nj} \\ &\leq \sum_j [\alpha_{nj}(V)]^{-1} \int \alpha_{nj}(V' + x) d\alpha_{nj} \\ &\leq (\sup_j [\alpha_{nj}(V)]^{-1}) \sum_j |\alpha_{nj}|^2(V'). \end{aligned}$$

Since $\{\alpha_{nj}\}$ is uniformly infinitesimal, $\sup_j [\alpha_{nj}(V)]^{-1} < 1 + \varepsilon$, for any given $\varepsilon > 0$ and all sufficiently large n depending on ε . The above inequality and the validity of (5.14) for any neighborhood N of the identity imply (5.12). Since $|\alpha_{nj}|^2 = |\theta_{nj}|^2$ and $\{\theta_{nj}\}$ is uniformly infinitesimal, the same argument leads to (5.11).

From (5.15) we have, for any neighborhood V of the identity in X ,

$$(5.16) \quad \limsup_n \sum_j \iint_{V \times V} [1 - R(x_1 - x_2, y)] d\theta_{nj}(x_1) d\theta_{nj}(x_2) < \infty.$$

We now choose V such that $V - V \subset N$. Then

$$R(x_1 - x_2, y) = \cos g(x_1 - x_2, y).$$

Since $1 - \cos \theta > \theta^2/4$ for sufficiently small θ , we have from property (5) of $g(x, y)$ in Lemma 3.3,

$$1 - R(x_1 - x_2, y) \geq \frac{1}{4}g^2(x_1 - x_2, y)$$

for $y \in K$. Since $e^{ig(x,y)} = (x, y)$ for $x \in N$ and $y \in K$, the same property of $g(x, y)$ gives

$$g(x_1 - x_2, y) = g(x_1, y) - g(x_2, y), \quad x_1, x_2 \in V, \quad y \in K.$$

Thus, for $x_1, x_2 \in V, y \in K$,

$$(5.17) \quad 1 - R(x_1 - x_2, y) \geq \frac{1}{4}[g^2(x_1, y) + g^2(x_2, y) - 2g(x_1, y)g(x_2, y)].$$

(5.16) and (5.17) imply

$$(5.18) \quad \limsup_n \sup_{y \in K} \left\{ \sum_j \int_V g^2(x, y) d\theta_{nj}(x) - \left(\int_V g(x, y) d\theta_{nj} \right)^2 \right\} < \infty.$$

(5.10), (5.12), and (5.18) imply (5.13). This completes the proof of the theorem.

THEOREM 5.2. *If $\{\alpha_{nj}\}$ is uniformly infinitesimal, $\mu_n = \prod_j \alpha_{nj}$, and $\mu_n \Rightarrow \mu$, then μ is infinitely divisible.*

Proof. If μ has no idempotent factor, then it is also a limit of the sequence λ_n where λ_n is constructed as in Theorem 5.1. Since λ_n is infinitely divisible for each n , μ is also infinitely divisible.

Now let us consider the general case. By Lemma 5.1 the set $[y: \hat{\mu}(y) \neq 0]$ is an open subgroup H . If G is the annihilator of this subgroup in X , then the normalized Haar measure of the compact group G is a factor of μ . Let τ be the canonical homomorphism from X to X/G . Then the sequence $\{\alpha_{nj} \tau^{-1}\}$ is uniformly infinitesimal in X/G , and $\mu_n \tau^{-1} \Rightarrow \mu \tau^{-1}$. $\mu \tau^{-1}$ has no idempotent factor and hence is infinitely divisible. For $y \in H, \mu(y) = \hat{\mu} \tau^{-1}(y)$ and $\hat{\mu}(y) = 0$ for $y \notin H$. Thus μ itself is infinitely divisible.

Remark. In the statement of Theorem 5.2 we have assumed that $\{\alpha_{nj}\}$ is

uniformly infinitesimal. However, it is enough to assume the existence of a sequence $\{x_{nj}\}$ of elements from the group with the property that $\{\alpha_{nj} * x_{nj}\}$ is uniformly infinitesimal. This is equivalent to the statement that any limit of shifts of α_{nj} is degenerate.

6. Gaussian distributions

In this section we give an algebraic definition of a Gaussian distribution and obtain its representation. This definition is also consistent with the classical definition of Gaussian laws in the case of the finite-dimensional vector spaces. Another definition of a Gaussian law is due to Urbanik [14] who also discussed some of its properties.

From the point of view of limit theorems, the Gaussian laws arise very naturally as follows. Suppose F_n is a sequence of finite measures on the group X such that (1) outside each neighborhood of the identity $F_n \rightarrow 0$ as $n \rightarrow \infty$, and (2) $e(F_n)$ converges to a limit after a suitable shift. If the total mass of F_n is not uniformly bounded, $e(F_n)$ may actually converge to a nondegenerate distribution. These are precisely the Gaussian laws.

DEFINITION 6.1. A distribution μ is said to be *Gaussian* if it has the following properties: (i) μ is infinitely divisible, and (ii) if $\mu = e(F) * \alpha$ where α is infinitely divisible, then F is degenerate at the identity.

THEOREM 6.1. A function on Y is the characteristic function of a Gaussian distribution on X if and only if it has the form

$$(x, y) \exp [-\phi(y)]$$

where x is a fixed point of X and $\phi(y)$ is a continuous, nonnegative function on Y satisfying the equality

$$(6.1) \quad \phi(y_1 + y_2) + \phi(y_1 - y_2) = 2[\phi(y_1) + \phi(y_2)]$$

for all y_1, y_2 in Y .

Proof. Let μ be Gaussian. Then μ cannot have a nondegenerate idempotent factor. For, otherwise, the Haar measure of some compact subgroup will be a factor of μ , and hence if F is any measure concentrated in that subgroup, then $\mu = e(F) * \mu$. This contradicts property (ii) of Definition 6.1. From the definition of infinite divisibility it follows that, for each n , there exist a distribution α_n and an element g_n of X such that $\mu = \alpha_n^n * g_n$.

Since μ has no idempotent factors, any limit of shifts of α_n is degenerate, and hence α_n 's can be shifted so as to converge to the distribution degenerate at the identity. All these shifts may be absorbed into g_n , so that α_n itself can be assumed to converge to the distribution degenerate at the identity. We now write (as in the proof of Theorem 5.1)

$$\theta_n = \alpha_n * x_n, \quad \beta_n = e(\theta_n), \quad \lambda_n = \beta_n^n * (-nx_n) * g_n,$$

where the element x_n is determined by the identity

$$(x_n, y) = \exp \left[-i \int g(x, y) d\alpha_n \right].$$

The absence of idempotent factors for μ implies

$$\limsup_n \sup_{y \in K} |\hat{\lambda}_n(y) - \hat{\mu}(y)| = 0$$

by Theorem 5.1.

Thus

$$|\hat{\mu}(y)| = \lim_{n \rightarrow \infty} \exp \left(n \int [R(x, y) - 1] d\theta_n \right).$$

We shall first show that the function

$$\phi(y) = \lim_{n \rightarrow \infty} \left(n \int [1 - R(x, y)] d\theta_n \right)$$

satisfies (6.1). We write $P_n = n\theta_n$. Then $e(P_n)$ is a shift of λ_n , and hence $e(P_n)$ is shift-compact. Now Theorem 4.3 implies that P_n restricted to N' is tight for every neighborhood N of the identity. But any limit P of P_n restricted to N' will be such that $\mu = e(P) * \alpha$ where α is also infinitely divisible. From condition (ii) of Definition 6.1 it follows that the mass of P_n outside every neighborhood of the identity tends to zero. Thus

$$(6.2) \quad \phi(y) = \lim_{n \rightarrow \infty} \int_N [1 - R(x, y)] dP_n$$

for every neighborhood N of the identity. (6.2) and the following identity¹

$$\lim_{x \rightarrow e} \frac{[1 - R(x, y_1 + y_2)] + [1 - R(x, y_1 - y_2)]}{2[1 - R(x, y_1)] + 2[1 - R(x, y_2)]} = 1$$

(e denoting the identity element of the group X) imply

$$\phi(y_1 + y_2) + \phi(y_1 - y_2) = 2[\phi(y_1) + \phi(y_2)].$$

Thus, in order to complete the proof of the theorem it suffices to show that $\hat{\mu}(y)/|\hat{\mu}(y)|$ is a character on Y . Let us denote this by $\chi(y)$. It is not difficult to verify that, for every neighborhood N of the identity,

$$(6.3) \quad \begin{aligned} & \chi(y_1 + y_2)[\chi(y_1)\chi(y_2)]^{-1} \\ &= \exp \left(\lim_{n \rightarrow \infty} \int [I(x, y_1 + y_2) - I(x, y_1) - I(x, y_2)] dP_n \right) \end{aligned}$$

where $I(x, y)$ denotes the imaginary part of (x, y) . For any given $\varepsilon > 0$, we

¹ Since both the numerator and denominator can vanish, this limiting relation is to be interpreted as follows: the numerator lies between $(1 - \varepsilon)$ and $(1 + \varepsilon)$ times the denominator, if x is near enough to e .

choose a neighborhood N of the identity such that

$$|I(x, y_1)| < \varepsilon, \quad |I(x, y_2)| < \varepsilon$$

for $x \in N$. Since

$$|I(x, y_1 + y_2) - I(x, y_1) - I(x, y_2)| \leq |I(x, y_1)| |1 - R(x, y_2)| + |I(x, y_2)| |1 - R(x, y_1)|$$

and

$$\limsup \int [1 - R(x, y)] dP_n < \infty$$

(by Theorem 4.3), we have

$$\left| \int_N [I(x, y_1 + y_2) - I(x, y_1) - I(x, y_2)] dP_n \right| < C \cdot \varepsilon$$

where C is a constant depending only on y_1 and y_2 . Since ε is arbitrary, the right side of (6.3) is equal to unity. The continuity of $\chi(y)$ is obvious. This shows that χ is a character on Y . Thus there exists an element $x \in X$ such that

$$\hat{\mu}(y) = (x, y) \exp(-\phi(y)).$$

This proves the necessity part.

Conversely, let $\hat{\mu}(y) = (x, y) \exp(-\phi(y))$ where $\phi(y)$ is a nonnegative continuous function of y satisfying (6.1). Let y_1, \dots, y_k be some k characters. Then it is easily verified that $\exp[-\phi(n_1 y_1 + \dots + n_k y_k)]$, considered as a function of integers n_1, \dots, n_k , is positive-definite in the product of integer group taken k times. This implies the positive-definiteness of $\exp(-\phi(y))$, and hence this function is the characteristic function of a measure. Since $\hat{\mu}(y) = 1$ at the identity of Y , the measure is a distribution. The infinite divisibility of μ is obvious. We shall now prove property (ii) of Definition 6.1. Let, if possible, $\hat{\mu}(y) = \hat{\mu}_1(y)\hat{\mu}_2(y)$, where $\mu_1 = e(F)$ and μ_2 is infinitely divisible. Since $\hat{\mu}(y)$ does not vanish for any y , $\hat{\mu}_2(y)$ also does not vanish, and hence by Theorem 5.1, μ_2 is a limit of distributions of the type $e(H)$. From (2.1) it is clear that for any finite measure H

$$(6.4) \quad -\log |(e(H))^{\wedge}(y_1 + y_2)| - \log |(e(H))^{\wedge}(y_1 - y_2)| \leq 2[-\log |(e(H))^{\wedge}(y_1)| - \log |(e(H))^{\wedge}(y_2)|].$$

Thus (6.4) is also valid when $e(H)$ is replaced by either μ_1 or μ_2 . Substituting μ_1 and μ_2 for $e(H)$ in (6.4) and adding, we get $\phi(y_1 + y_2) + \phi(y_1 - y_2) \leq 2[\phi(y_1) + \phi(y_2)]$. Since equality holds good in this case, we must have

$$\begin{aligned} \int ([1 - R(x, y_1 + y_2)] + [1 - R(x, y_1 - y_2)]) dF \\ = 2 \int ([1 - R(x, y_1)] + [1 - R(x, y_2)]) dF \end{aligned}$$

for each $y_1, y_2 \in Y$, i.e.,

$$\int [1 - R(x, y_1)] [1 - R(x, y_2)] dF = 0.$$

Since F is a measure, this implies that F must be degenerate at the origin. This completes the proof.

Remark 1. Consider real-valued continuous functions $\psi(y_1, y_2)$ defined for $y_1, y_2 \in Y$ possessing the following properties:

- (i) $\psi(y_1, y_2) = \psi(y_2, y_1)$,
- (ii) $\psi(y_1 + y_2, y_3) = \psi(y_1, y_3) + \psi(y_2, y_3)$,
- (iii) $\psi(y, y) \geq 0$.

Clearly for any such function ψ , $\phi(y) = \psi(y, y)$ satisfies the identity (6.1). Conversely, any nonnegative continuous function satisfying (6.1) can be obtained in this way. In fact, $\psi(y_1, y_2)$ can be recovered from $\phi(y)$ by the relation

$$\psi(y_1, y_2) = \frac{1}{2}[\phi(y_1 + y_2) - \phi(y_1) - \phi(y_2)].$$

Remark 2. If X_0 is the component of the identity in X , then its annihilator Y_0 is the smallest closed subgroup containing all compact subgroups of Y . Consequently any $\psi(y_1, y_2)$ with the properties stated above vanishes identically in Y_0 , i.e., $\psi(y_1, y_2) = 0$ if y_1 or $y_2 \in Y_0$. Thus if μ is a symmetric Gaussian distribution on X , then $\hat{\mu}(y) = 1$ for $y \in Y_0$. In other words μ is necessarily concentrated in X_0 . Theorem 6.1 and Remark 1 above also show that every connected locally compact group has nontrivial Gaussian measures defined on it.

7. Representation of infinitely divisible distributions

As we have mentioned in the introduction, the characteristic function of any infinitely divisible distribution on the real line possesses the famous Lévy-Khinchin representation [3]. The purpose of this section is to obtain such a canonical representation in the case of a general locally compact abelian group.

DEFINITION 7.1. An infinitely divisible distribution λ is said to be a *proper factor* of another infinitely divisible distribution μ , if $\mu = \lambda * \alpha$ and α is infinitely divisible.

LEMMA 7.1. *The set of proper factors of an infinitely divisible distribution is closed.*

LEMMA 7.2. *If $e(F_n)$ converges to the distribution degenerate at the identity, then $F_n(N') \rightarrow 0$ as $n \rightarrow \infty$ for every neighborhood N' of the identity.*

The proof of both the lemmas is quite elementary and is left to the reader.

THEOREM 7.1. *If μ is an infinitely divisible distribution without idempotent factors, then $\hat{\mu}(y)$ has a representation*

$$(7.0) \quad \hat{\mu}(y) = (x_0, y) \exp \left[\int [(x, y) - 1 - ig(x, y)] dF(x) - \phi(y) \right]$$

where x_0 is a fixed element of X , $g(x, y)$ is a function on $X \times Y$ which is independent of μ and has the properties mentioned in Lemma 3.3, F is a σ -finite measure with finite mass outside every neighborhood of the identity in X which satisfies

$$\int [1 - R(x, y)] dF < \infty \quad \text{for every } y,$$

and $\phi(y)$ is a nonnegative continuous function satisfying

$$\phi(y_1 + y_2) + \phi(y_1 - y_2) = 2[\phi(y_1) + \phi(y_2)]$$

for each $y_1, y_2 \in Y$. Conversely, any function of the type (7.0) is the characteristic function of an infinitely divisible distribution.

Proof. Let μ be any infinitely divisible distribution without an idempotent factor. Choose and fix a sequence $\{N_k\}$ of neighborhoods of the identity in X descending to the identity. Let μ_1 be that proper factor of μ which is of the type $e(F)$ and for which $F(N_1) = 0$ and $F(N'_1) = 0$ and $F(N'_1)$ is maximum. Such an $e(F)$ exists because of Theorems 2.1 and 4.3 and Lemma 7.1. Let the F at which the maximum is attained be F_1 , and let $\mu = \mu_1 * \lambda_1$ and $\mu_1 = e(F_1)$. Since λ_1 is infinitely divisible and without idempotent factors, the same argument can be applied to λ_1 and the neighborhood N_2 . Thus there exists a measure F_2 for which $F_2(N_2) = 0$, $F_2(N'_2)$ is a maximum, $\mu_2 = e(F_2)$, $\lambda_1 = \mu_2 * \lambda_2$, and λ_2 is infinitely divisible. Repeating this procedure we can write

$$(7.1) \quad \mu = \mu_1 * \mu_2 * \cdots * \mu_n * \lambda_n,$$

$$(7.2) \quad \lambda_{n-1} = \mu_n * \lambda_n,$$

$$(7.3) \quad \mu_n = e(F_n), \quad F_n(N_n) = 0.$$

λ_n is infinitely divisible, and $F_n(N'_n)$ is a maximum in the sense explained earlier. Thus by Theorem 2.1 there exist shifts of $\mu_1 * \cdots * \mu_n$ and λ_n converging to ν and λ respectively, and $\mu = \nu * \lambda$. We now assert that λ cannot have a proper factor of the type $e(F)$. Suppose, on the contrary, $e(F)$ is a proper factor of λ . Then it will have a positive mass outside some N_k . Further, since the sequence λ_n is descending (in the order $<$), $e(F)$ is a proper factor of λ_k . $\lambda_k = e(F) * \theta$ where θ is infinitely divisible. If F' is the restriction of F to N'_k , then (7.2) and (7.3) imply that

$$\lambda_{k-1} = e(F_k + F') * e(F - F') * \theta.$$

This is a contradiction since the total mass of $F_k + F'$ exceeds that of F_k . Thus λ has no proper factor of the type $e(F)$ and is therefore a Gaussian distribution. An application of Theorem 6.1 leads to the existence of a function $\phi(y)$ and an element $x \in X$ for which

$$(7.4) \quad \hat{\lambda}(y) = (x, y) \exp [-\phi(y)]$$

and

$$\phi(y_1 + y_2) + \phi(y_1 - y_2) = 2[\phi(y_1) + \phi(y_2)], \quad y_1, y_2 \in Y.$$

Now we write $H_n = F_1 + F_2 + \dots + F_n$. From the construction of the distributions $\nu_n = \mu_1 * \dots * \mu_n$, it is clear that

$$\nu_n(y) = (x_n, y) \exp \left[\int [(x, y) - 1] dH_n(x) \right]$$

for some element $x_n \in X$. Since $\exp [i \int g(x, y) dH_n(x)]$ is a character on Y , it can be considered as an element of X . Thus there exists an element $z_n \in X$ such that

$$(7.5) \quad \nu_n(y) = (z_n, y) \exp \left[\int [(x, y) - 1 - ig(x, y)] dH_n(x) \right].$$

Since $e(H_n)$ is a factor of μ and H_n increases as $n \rightarrow \infty$, it follows from the shift-compactness of $e(H_n)$, Theorem 4.3, and Lemma 2.1 that H_n increases to a σ -finite measure H for which $H(N') < \infty$ for every neighborhood N of the identity and

$$\int \sup_{y \in K} [1 - R(x, y)] dH < \infty,$$

for every compact $K \subset Y$. Since $[(x, y) - 1 - ig(x, y)]$ is bounded uniformly in $y \in K$, by property (2) of $g(x, y)$ in Lemma 3.3, we have, for every neighborhood N of the identity,

$$\lim_{n \rightarrow \infty} \int_{N'} [(x, y) - 1 - ig(x, y)] dH_n = \int_{N'} [(x, y) - 1 - ig(x, y)] dH,$$

uniformly in $y \in K$. When N is sufficiently small, we have, by the properties (4) and (5) of $g(x, y)$ in Lemma 3.3,

$$\begin{aligned} (x, y) &= e^{ig(x, y)} && \text{for } x \in N, \quad y \in K, \\ g^2(x, y) &\leq C_1[1 - R(x, y)], && x \in N, \quad y \in K, \end{aligned}$$

where C_1 is a constant depending on K only. Thus

$$(7.6) \quad |(x, y) - 1 - ig(x, y)| \leq C_2[1 - R(x, y)], \quad x \in N, \quad y \in K,$$

where C_2 is a constant depending on K only. Thus

$$\int_N \sup_{y \in K} |(x, y) - 1 - ig(x, y)| dH_n \leq C_2 \int_N \sup_{y \in K} [1 - R(x, y)] dH.$$

The above inequality implies the convergence of

$$\int [(x, y) - 1 - ig(x, y)] dH_n \quad \text{to} \quad \int [(x, y) - 1 - ig(x, y)] dH$$

uniformly in $y \in K$. Now (7.5) implies that $\nu_n(y)$, after a suitable shift, converges uniformly over compact sets to $\exp \int [(x, y) - 1 - ig(x, y)] dH$. This completes the proof of the first part since ν_n converges to ν after a suitable shift, $\mu = \nu * \lambda$, and λ satisfies (7.4).

To prove the converse, we first observe that if F is a totally finite measure, then $\exp[i \int g(x, y) dF]$ is a character, and hence $\hat{\mu}(y)$ given by (7.0) is the characteristic function of an infinitely divisible distribution. In the general case we consider a sequence F_n of totally finite measures increasing to F . If K is any compact subset of Y , then, according to Lemma 2.1, $1 - R(x, y)$ is uniformly integrable with respect to F for $y \in K$. (7.6) implies the uniform integrability of $[(x, y) - 1 - ig(x, y)]$ for $y \in K$. This shows that the function

$$(x_0, y) \exp \left(\int [(x, y) - 1 - ig(x, y)] dF_n - \phi(y) \right)$$

converges uniformly over compact sets to $\hat{\mu}(y)$. Thus, by Theorem 5.2' $\hat{\mu}(y)$ is infinitely divisible.

Remark. If the group X is totally disconnected, then the representation takes a simpler form. For in such groups $\phi(y) = 0$ (see Remark 2 following Theorem 6.1), and $g(x, y) = 0$ (see Example 4 at the end of Section 3). Thus every infinitely divisible distribution without idempotent factors has the representation

$$\hat{\mu}(y) = (x_0, y) \exp \left[\int [(x, y) - 1] dF \right]$$

where F is a σ -finite measure which has finite mass outside each neighborhood of the identity.

8. Uniqueness of the representation

In the case of the real line it is well known that the canonical representation of an infinitely divisible distribution is unique. However, this is not true in the case of a general group. We shall show that the nonuniqueness is essentially due to the presence of compact subgroups in the original group or equivalently due to the disconnectedness of the character group.

Before proceeding to the statement of the main result of this section we shall explain a few conventions and prove an elementary lemma. If μ is any infinitely divisible distribution without idempotent factors, we say that μ has the representation (x_0, F, ϕ) where x_0, F , and ϕ are as in Theorem 7.1. If F is any signed measure, we denote by F_y the measure given by

$$F_y(A) = \int_A [1 - R(x, y)] dF.$$

LEMMA 8.1. *Let μ be a totally finite signed measure. If $\hat{\mu}(y)$ is constant on the cosets of a closed subgroup Y_0 of Y , then μ vanishes identically on the complement of the annihilator of Y_0 in X .*

Proof. Let $Y_1 = Y/Y_0$, and let X_1 be the annihilator of Y_0 in X . $\hat{\mu}(y)$, being constant on cosets of Y_0 , can be considered as a function on Y_1 . Then $\hat{\mu}(y_1)$, $y_1 \in Y_1$ is the characteristic function of a signed measure on X_1 . Since, for $x \in X$, (x, y) remains constant on cosets of Y_1 and $\hat{\mu}(y)$ has the same property, we can write

$$\hat{\mu}(y) = \int_{x_1} (x, y_1) d\lambda = \int_{x_1} (x, y) d\lambda$$

where y_1 denotes that coset of Y_1 to which y belongs. This shows that the signed measures μ and λ are identical, and hence μ vanishes identically on the complement of X_1 .

THEOREM 8.1. *If (x_1, F_1, ϕ_1) and (x_2, F_2, ϕ_2) are two representations of the same infinitely divisible distribution without idempotent factors, then (i) $\phi_1 = \phi_2$, and (ii) the signed measure $F_1 - F_2$ vanishes identically on the complement of the annihilator of the component of identity of the character group Y .*

Proof. Writing $F = F_1 - F_2$, $\phi = \phi_1 - \phi_2$, and $x_0 = x_2 - x_1$, we have

$$(8.1) \quad \exp \int [(x, y) - 1 - ig(x, y)] dF = (x_0, y) \exp \phi(y),$$

$$(8.2) \quad \phi(y_1 + y_2) + \phi(y_1 - y_2) - 2[\phi(y_1) + \phi(y_2)] = 0, \quad y_1, y_2 \in Y.$$

Equating the logarithm of the absolute value on both sides of (8.1) we obtain

$$(8.3) \quad \phi(y) = \int [R(x, y) - 1] dF.$$

Substituting the values of the above expression at $y_1 + y_2, y_1 - y_2, y_1$, and y_2 in (8.2), we get

$$(8.4) \quad \int (1 - R(x, y_1))(1 - R(x, y_2)) dF = 0.$$

(8.4) can be rewritten as

$$\int [1 - R(x, y_1)] dF_{y_2} = 0, \quad \text{for } y_1, y_2 \in Y.$$

Since F_{y_2} is totally finite, we have

$$(8.5) \quad \int (x, y) d(F_{y_2} + \bar{F}_{y_2}) = 2F_{y_2}(X).$$

Since the right side of (8.5) is constant when y_2 is fixed, we conclude that the

signed measure $F_{y_2} + \bar{F}_{y_2}$ is degenerate at the identity. But the mass of F_{y_2} at the identity is zero. Thus $F_{y_2} + \bar{F}_{y_2} = 0$. In particular $F_{y_2}(X) = 0$, i.e.,

$$(8.6) \quad \int [1 - R(x, y_2)] dF = 0 \quad \text{for every } y_2 .$$

(8.3) and (8.6) imply the equality of ϕ_1 and ϕ_2 .

In order to prove the second part of the theorem we make use of the equality of ϕ_1 and ϕ_2 and rewrite (8.1) as

$$(8.7) \quad \exp \int [(x, y) - 1 - ig(x, y)] dF = (x_0, y), \quad y \in Y .$$

Substituting $y = y_1 + y_2$, $y_1 - y_2$, and y_1 successively in (8.7) and dividing the product of the first two by the square of the third, we obtain

$$(8.8) \quad \exp \int (x, y_1)[1 - R(x, y_2)] dF = 1, \quad y_1, y_2 \in Y ,$$

or equivalently

$$\int (x, y_1)[1 - R(x, y_2)] dF = 2\pi n(y_1, y_2)$$

where $n(y_1, y_2)$ is an integer-valued continuous function of y_1 and y_2 . We fix y_2 for the present. Then $n(y_1, y_2)$ remains constant on every connected subset of Y and, in particular, on the cosets of the component of identity in Y . This implies, by Lemma 8.1, that the signed measure F_{y_2} vanishes identically on the complement of the annihilator in X of the component of the identity of Y . Since this is true for each y_2 , it follows that F itself vanishes identically outside this annihilator. This completes the proof of the theorem.

Remark 1. It is not difficult to show that the annihilator of the component of the identity of Y is the smallest closed subgroup containing all compact subgroups of X . This reflects the role of compact subgroups in making the representation nonunique. In particular, if the group X has no compact subgroups, then the representation is unique.

Remark 2. It was shown in the course of the proof of Theorem 8.1 that the measure F_y is antisymmetric for each character y , i.e., $F_y(A) = -F_y(-A)$ for every Borel set. But if every element in the group were of order two, then such a measure would be identically zero. Coupling this with Remark 1 we can say that if the group X is such that every compact subgroup of X consists only of elements of order two, then the representation is unique.

Remark 3. Conversely, if X is a compact group such that not all elements of X are of order two, then the representation is not unique as can be seen from the following example. We take an element y_0 in the character group which is not of order two and consider the function

$$f(x) = 2\pi i[\overline{(x, y_0)} - (x, y_0)] .$$

$f(x)$ is real and not identically zero. If h denotes the normalized Haar measure of X , we have

$$\begin{aligned} \int (x, y)f(x) dh(x) &= 2\pi i && \text{if } y = y_0, \\ &= -2\pi i && \text{if } y = -y_0, \\ &= 0 && \text{otherwise.} \end{aligned}$$

Writing f^+ and f^- to denote the positive and negative parts of $f(x)$, we define the two measures

$$F_1(A) = \int_A f^+(x) dh, \quad F_2(A) = \int_A f^-(x) dh.$$

Then $F_1 \neq F_2$, but

$$\exp \int [(x, y) - 1 - ig(x, y)] dF_1 = (x_0, y) \exp \int [(x, y) - 1 - ig(x, y)] dF_2$$

where $(x_0, y) = \exp [i \int g(x, y) d(F_2 - F_1)]$ is a character on Y and hence an element of X . Thus $(e, F_1, 0)$ and $(x_0, F_2, 0)$ are two representations of the same infinitely divisible distribution.

9. Compactness criteria

When the group is the real line, necessary and sufficient conditions that a sequence of infinitely divisible distributions may converge to a given infinitely divisible distribution have been obtained in terms of their representations (cf. Gnedenko and Kolmogorov [3]). Such a result fails to be valid in the general case because of the nonuniqueness of the representations. However, it is possible to obtain conditions for the compactness of a family of infinitely divisible distributions in terms of their representations.

Before proceeding to state the main result of this section we shall investigate what happens to the representation when we pass over from a group to its quotient group. Let $G \subset X$ be some closed subgroup of X , and $X' = X/G$ the quotient group. Let τ denote the canonical homomorphism from X to X' . If Y' is the character group of X' , we choose and fix a function $g'(x', y')$ defined on $X' \times Y'$ and satisfying all the properties of Lemma 3.3. We observe that Y' is the annihilator of G in Y and hence a subgroup of Y . Any infinitely divisible distribution μ' on X' without idempotent factors has a representation (x', F', ϕ') (with g replaced by g') according to Theorem 7.1.

LEMMA 9.1. *Let μ be an infinitely divisible distribution on X with a representation (x, F, ϕ) . If $\mu' = \mu\tau^{-1}$, $x' = \tau x$, $F' = F\tau^{-1}$, and ϕ' is the restriction of ϕ to Y' , then μ' is an infinitely divisible distribution on X' and is a shift of the distribution represented by (x', F', ϕ') .*

Proof. Let $\mu = \mu_1 * \mu_2$ where μ_2 is the unique Gaussian component of μ . Clearly $\mu' = \mu\tau^{-1} = \mu_1\tau^{-1} * \mu_2\tau^{-1}$. Since $e(\alpha)\tau^{-1} = e(\alpha\tau^{-1})$ for every finite measure α , it follows easily that $\mu_1\tau^{-1}$ and the distribution represented by $(x', F', 0)$ are shifts of each other. Thus $\mu_2\tau^{-1}$ is the Gaussian component of μ' . This completes the proof.

We shall now prove the following.

THEOREM 9.1. *Let $\{\mu_\alpha\}$ be a family of infinitely divisible distributions without idempotent factors and with representations $\{(x_\alpha, F_\alpha, \phi_\alpha)\}$. Necessary and sufficient conditions that $\{\mu_\alpha\}$ be shift-compact and any limit of shifts of $\{\mu_\alpha\}$ be devoid of idempotent factors, are*

(1) *The family $\{F_\alpha\}$ of measures is compact when restricted to N' for every neighborhood N of the identity.*

$$(2) \sup_\alpha \left[\int [1 - R(x, y)] dF_\alpha + \phi_\alpha(y) \right] < \infty \text{ for all } y.$$

Proof. The necessity of the above two conditions is obvious in view of Theorem 4.3. Regarding sufficiency we first observe that if $\{\mu_\alpha\}$ is a shift-compact family, then condition (2) is sufficient to ensure the absence of idempotent factors in any limit of shifts of μ_α . We shall now prove shift-compactness.

We have only to prove the compactness of the family $\{|\mu_\alpha|^2\}$. We now observe that if (x, F, ϕ) is a representation of an infinitely divisible distribution μ in a group X and τ is a continuous homomorphism of X onto another group X' , then $\mu\tau^{-1}$ is a shift of the distribution represented by $(\tau x, F\tau^{-1}, \phi')$ where ϕ' is the restriction of ϕ to the character group of X' (which is a subgroup of Y). Further if $\{(x_\alpha, F_\alpha, \phi_\alpha)\}$ satisfies conditions (1) and (2), so does the family $\{(\tau x_\alpha, F_\alpha\tau^{-1}, \phi'_\alpha)\}$. Making use of these remarks we shall reduce the proof of the general case to that of certain simple groups.

In order that a family of measures be compact, it is necessary and sufficient that the family be tight. If C is a compact subgroup of X , and if the family of measures induced by the canonical homomorphism on X/C is tight, then the original family itself is tight. We now choose the group C in such a manner that X/C has the structure $V \oplus D \oplus K^r$ where V is a vector group, D a discrete group, and K^r the r -dimensional torus. The existence of such a compact subgroup is well known. But a family of measures in the product of two topological spaces is tight as soon as the two marginal families are so. Thus it is enough to prove the sufficiency of (1) and (2) in the case of the real line, discrete groups, and compact groups. In the case of the real line the boundedness of $\int [1 - R(x, y)] dF_\alpha$ implies the boundedness of $\int_{|x| < \varepsilon} x^2 dF_\alpha$ for a suitable ε , which together with condition (1) implies the equicontinuity of the family of functions $\exp\{-\int (1 - R(x, y)) dF_\alpha\}$. Since $\phi_\alpha(y)$ assumes the form $\sigma_\alpha^2 y^2$ and hence σ_α^2 is bounded, it is clear that $\{|\mu_\alpha(y)|^2\}$ is equicon-

tinuous. But equicontinuity implies compactness. In the case of a discrete group, identity itself is an open set, and hence the family $\{F_\alpha\}$ is compact outside the identity. This, together with the fact that every infinitely divisible distribution without idempotent factors is a shift of $e(F)$ for some F with zero mass at the identity, implies the required result. In the case of compact groups any family of measures is compact. This completes the proof of the theorem.

COROLLARY 9.1. *In addition to the conditions (1) and (2) of Theorem 9.1 the condition that $\{x_\alpha\}$ be a conditionally compact set is necessary and sufficient to ensure the conditional compactness of $\{\mu_\alpha\}$ with representations*

$$(x_\alpha, F_\alpha, \phi_\alpha).$$

Proof. From (7.6) we have

$$|(x, y) - 1 - ig(x, y)| \leq C[1 - R(x, y)]$$

for all $x \in N, y \in K$ where C is a constant depending on K only, N is a sufficiently small neighborhood of the identity in X , and K is a compact subset of Y . This implies the equicontinuity of the family of functions

$$\exp \left[\int [(x, y) - 1 - ig(x, y)] dF_\alpha - \phi_\alpha(y) \right].$$

Hence $\{x_\alpha\}$ is conditionally compact if and only if $\{\mu_\alpha\}$ is so.

For any σ -finite measure F which has finite mass outside every neighborhood of the identity and integrates the function $1 - R(x, y)$ for every y , we write

$$(9.1) \quad (E(F))^\wedge(y) = \exp \int [(x, y) - 1 - ig(x, y)] dF.$$

Theorem 7.1 implies that $(E(F))^\wedge(y)$ is the characteristic function of an infinitely divisible distribution. By proceeding along the same lines as in the proof of Theorem 9.1 it is possible to prove the following.

THEOREM 9.2. *Let $\{\mu_n\}$ be a sequence of infinitely divisible distributions with representations (x_n, F_n, ϕ_n) . Let μ_n converge to μ after a suitable shift, and F_n to F outside each neighborhood of the identity. Then μ has a representation (x, F, ϕ) for a suitable choice of x and ϕ .*

10. Representation of convolution semigroups

We have observed earlier that the representation of an infinitely divisible distribution is not unique. We shall now consider the representation problem for a one-parameter convolution semigroup of distributions. By such a semigroup we mean a family $\{\mu_t\}$ of distributions indexed by $t \geq 0$ such that $\mu_t * \mu_s = \mu_{t+s}$. We shall further assume that μ_t converges weakly to the

distribution degenerate at the identity when $t \rightarrow 0$. Obviously, for such semigroups $\hat{\mu}_t(y) \neq 0$ for any $t > 0$ and $y \in Y$. That such a semigroup has a unique canonical representation is the content of the following.

THEOREM 10.1. *Let $\{\mu_i\}$ be a one-parameter convolution semigroup of distributions such that μ_t converges weakly to the distribution degenerate at the identity as $t \rightarrow 0$. Then $\hat{\mu}_t(y)$ has the canonical representation*

$$\hat{\mu}_t(y) = (x_t, y) \exp \left(t \int [(x, y) - 1 - ig(x, y)] dF - t\phi(y) \right)$$

where F and ϕ are as in Theorem 7.1 and $\{x_t\}$ is a continuous one-parameter semigroup in X . Moreover, $\{x_t\}$, F , and ϕ are uniquely determined by $\{\mu_t\}$.

Proof. Since $\mu_t = (\mu_{t/n})^n$, μ_t is infinitely divisible. As remarked at the beginning, $\hat{\mu}_t(y)$ is nonvanishing at any point and hence has no idempotent factor. Thus by Theorem 7.1, μ_t has a representation (z_t, F_t, ϕ_t) . The uniqueness of ϕ_t implies that $\phi_t = t\phi_1$. We write $\phi = \phi_1$ and $\hat{\lambda}_t(y) = \hat{\mu}_t(y)e^{t\phi}$. Then $\{\lambda_t\}$ is a weakly continuous convolution semigroup, and λ_t has neither idempotent nor Gaussian factors. For any σ -finite measure F which has finite mass outside every neighborhood of the identity and integrates the function $1 - R(x, y)$, we define the distribution $E(F)$ as in (9.1). We now observe that the distribution $E(n!F_{1/n!})$ is a shift of λ_1 for every n . By Theorem 4.3 the sequence of measures $n!F_{1/n!}$ is compact outside every neighborhood of the identity. Thus we can choose a subsequence $n_k!F_{1/n_k!}$ which converges to a measure F outside every neighborhood of the identity. Thus by Theorem 9.2, λ_1 has a representation $(z, F, 0)$ for some $z \in X$. We now define

$$\hat{P}_t(y) = \exp \left(t \int [(x, y) - 1 - ig(x, y)] dF \right).$$

If $t = p/q$ is rational, then $E((p/q)n!F_{1/n!})$ is a shift of $\lambda_{p/q}$ for all sufficiently large n . Since $(p/q)n_k!F_{1/n_k!}$ converges to $(p/q)F$ outside every neighborhood of the identity, another application of Theorem 9.2 shows that $\lambda_{p/q}$ is a shift of $P_{p/q}$. By the continuity of the semigroups it is clear that, for every t , λ_t is a shift of P_t . Thus $\hat{\lambda}_t(y)$ can be written as $(x_t, y)\hat{P}_t(y)$. Then x_t automatically becomes a continuous one-parameter semigroup in X .

If now $(x_t, tF, t\phi)$ and $(x'_t, tF', t\phi)$ are two representations of the semigroup $\{\mu_i\}$, then by proceeding in the same way as in the proof Theorem 8.1 we obtain

$$\exp \left[t \int (x, y_1)[1 - R(x, y)] d(F - F') \right] = 1$$

for every t . But this can happen only if $F - F' = 0$. Thus the representation of the semigroup is always unique.

COROLLARY. *If μ_t is symmetric for each t , i.e., $\mu_t(A) = \mu_t(-A)$ for each Borel set A , then the representation stated in Theorem 10.1 takes the simpler form*

$$\hat{\mu}_t(y) = \exp \left[t \int [(x, y) - 1] dF - t\phi(y) \right]$$

where F is a symmetric measure which integrates $[1 - (x, y)]$ for each y , and $\phi(y)$ is continuous, nonnegative, and satisfies (6.1).

Remark. The above corollary could be used to give a representation for the “negative definite” functions of Herz [15, p. 198]. (The problem of obtaining such a representation for “negative definite” functions is raised by Herz on page 207 of [15].) A function $\nu(y)$ on Y is said to be “negative definite” if

- (i) $\nu(y)$ is continuous, $\nu(y) = \nu(-y)$, and $\nu(y) \geq \nu(0) = 0$, and
- (ii) if y_1, \dots, y_n are arbitrary elements of Y and c_1, c_2, \dots, c_n are real numbers such that $\sum c_j = 0$, then

$$\sum \sum c_i c_j \nu(y_i - y_j) \leq 0.$$

From the corollary above and Theorem 3.1 of [15, p. 198], it follows that $\nu(y)$ is negative definite if and only if it is of the form

$$\nu(y) = \int_x [1 - (x, y)] dF + \phi(y)$$

where F and ϕ have properties stated in the above corollary.²

11. A decomposition theorem

According to a theorem of Khinchin [5] any distribution on the real line can be written as the convolution of two distributions one of which is the convolution of a finite or countable number of indecomposable distributions, and the other of which is a distribution without indecomposable factors. Further, any distribution which is not infinitely divisible has an indecomposable factor. The object of this section is to extend this result to the general case with a slight modification to counteract the existence of idempotent factors.

THEOREM 11.1. *Let μ be any distribution on X . Then it can be written as $\lambda_\sigma * \lambda$ where λ is a distribution without any idempotent factor and λ_σ is the maximal idempotent factor of μ .*

Proof. Let H be the group generated by the set of all characters y at which $\hat{\mu}(y) \neq 0$. H is open, and hence its annihilator G in X is compact. The normalized Haar measure λ_σ of G is the required maximal idempotent factor. If we denote by τ the canonical homomorphism from X to X/G , then the distribution $\mu\tau^{-1}$ in X/G has no idempotent factors. We now choose a Borel

² The authors wish to thank the referee for pointing out that this representation could be derived from the results of the paper.

set $A \subset X$ such that τ restricted to A maps A onto X/G in a one-to-one manner. The existence of such a Borel set follows from a result due to Mackey [7, Lemma 1.1, p. 102]. By a result of Kuratowski [6, p. 251] it follows that the inverse of τ from X/G to A is a measurable map. Hence $\mu\tau^{-1}$ induces a measure λ on A . This λ satisfies the requirements of the theorem.

We shall now introduce a function $\theta(\alpha)$ defined for all factors of a distribution μ and similar to the function introduced by Khinchin. Let μ be a distribution without any idempotent factor. Then there is a sequence y_1, y_2, \dots of characters in Y such that $\hat{\mu}(y_i) \neq 0$ for $i = 1, 2, \dots$ and the smallest closed subgroup generated by this sequence is Y . Since $-\log |\hat{\mu}(y_i)|$ is well defined, we can choose a sequence $\varepsilon_n > 0$ such that

$$-\sum_{n=1}^{\infty} \varepsilon_n \log |\hat{\mu}(y_n)| < \infty.$$

This implies at once that, for every factor α of μ , the function

$$\theta(\alpha) = -\sum_{n=1}^{\infty} \varepsilon_n \log |\hat{\alpha}(y_n)|$$

is well defined and has the following obvious properties:

- (i) $\theta(\alpha) \geq 0$.
- (ii) $\theta(\alpha) = 0$ if and only if α is degenerate.
- (iii) $\theta(\alpha_1 * \alpha_2) = \theta(\alpha_1) + \theta(\alpha_2)$.
- (iv) If $\alpha_n \Rightarrow \alpha$, then $\theta(\alpha_n) \rightarrow \theta(\alpha)$.
- (v) $\theta(\alpha) = \theta(\beta)$ if α is a shift of β .

THEOREM 11.2. *Let μ be a distribution without any indecomposable or idempotent factors. Then μ is infinitely divisible.*

Proof. In view of the remark made under Theorem 5.2 it is enough to factorize μ in the form

$$(11.2) \quad \mu = \alpha_{n_1} * \dots * \alpha_{n_{2^n}}$$

where $\{\alpha_{n_j}\}$ is such that any limit of their shifts is degenerate. From the properties (i)–(v) of the θ function it is clear that it is sufficient to factorize μ in the form (11.2) with $\theta(\alpha_{n_j}) = 2^{-n}\theta(\mu)$. If μ satisfies the conditions stated in the theorem, then any factor of μ also satisfies them. Thus it suffices to prove that μ can be written as $\mu_1 * \mu_2$ with $\theta(\mu_1) = \theta(\mu_2) = \frac{1}{2}\theta(\mu)$. A repetition of this argument will then complete the proof. In order to do this we first observe that

$$\inf_{\alpha \in F(\mu), \theta(\alpha) \neq 0} \theta(\alpha) = 0$$

where $F(\mu)$ is the class of all factors of μ . For, otherwise, the class $F(\mu)$ being shift-compact, the infimum would be attained at an indecomposable distribution. But this is a contradiction. Thus there are factors of μ with arbitrarily small θ -values. We now take two distributions μ_1, μ_2 for which $\mu =$

$\mu_1 * \mu_2$ and $|\theta(\mu_1) - \theta(\mu_2)|$ is minimum. From the shift-compactness of $F(\mu)$ it follows that the minimum is attained. This minimum has to be zero. For, otherwise, by transferring a factor of μ_1 or μ_2 with an arbitrarily small θ -value to μ_2 or μ_1 we can make $|\theta(\mu_1) - \theta(\mu_2)|$ smaller. This completes the proof of the theorem.

THEOREM 11.3. *Any distribution μ on X can be written as $\lambda_H * \lambda_1 * \lambda_2$ where λ_H is the maximal idempotent factor of μ , λ_1 is the convolution of a finite or a countable number of indecomposable distributions, and λ_2 is an infinitely divisible distribution without indecomposable or idempotent factors.*

Proof. An application of Theorem 11.1 shows that μ can be written as $\lambda_H * \lambda$ where λ_H is the maximal idempotent factor and λ has no idempotent factors. Thus we can define a θ function on $F(\lambda)$ satisfying the properties (i)–(v) of (11.1). Let now δ_1 be the maximum of $\theta(\alpha)$ as α varies over the indecomposable factors of λ . If δ_1 is greater than zero, we write $\lambda = \alpha_1 * P_1$ where α_1 is indecomposable and $\theta(\alpha_1) \geq \delta_1/2$. We now denote by δ_2 the maximum of $\theta(\alpha)$ as α varies over the indecomposable factors of P_1 . If $\delta_2 > 0$, then we write $P_1 = \alpha_2 * P_2$ where α_2 is indecomposable and $\theta(\alpha_2) \geq \delta_2/2$. We repeat this argument. If the process terminates at the n^{th} stage, then $\lambda = \alpha_1 * \cdots * P_n$ and $\delta_{n+1} = 0$, which means that P_n has no indecomposable factors. Otherwise the process continues ad infinitum. Since $\sum_1^\infty \theta(\alpha_i)$ is convergent, $\theta(\alpha_n) \rightarrow 0$ as $n \rightarrow \infty$. The sequence $\alpha_1 * \cdots * \alpha_n$ being monotonic (in the order $<$) converges after a suitable shift. Absorbing this shift in α_n we can assume that $\alpha_1 * \cdots * \alpha_n$ converges to a distribution λ_1 . Automatically P_n will converge to a distribution λ_2 . If λ_2 has an indecomposable factor α , then it is a factor of P_n . Thus $\theta(\alpha) \leq \delta_n$ for each n . But $\delta_n \leq 2\theta(\alpha_n)$ and hence tends to zero. Therefore $\theta(\alpha) = 0$, or equivalently λ_2 has no indecomposable factors.

12. Concluding remarks

In obtaining the representation for an infinitely divisible distribution as well as in the proof of the theorem on accompanying laws (Theorems 5.1 and 7.1), we have assumed that the distributions under consideration do not have any idempotent factors. If a distribution μ has an idempotent factor, then, as in the proof of Theorem 11.1, we can construct a compact subgroup G such that the measure induced by the canonical homomorphism onto X/G has no idempotent factors. Since uniform infinitesimality and infinite divisibility are preserved by this map, these results can now be discussed in the quotient group X/G .

Another assumption we have made in the paper is that the group is separable (i.e., satisfies the second axiom of countability). All the results of Sections 3–10 are easily extended, with little essential modification in the proofs, in place of this restriction on the group, we suppose that the measures under consideration have supports contained in a σ -compact subset.

REFERENCES

1. S. BOCHNER, *Harmonic analysis and the theory of probability*, Berkeley, University of California Press, 1955.
2. ———, *General analytical setting for the central limit theory of probability*, Calcutta Mathematical Society Golden Jubilee Commemoration Volume (1958), Part I, pp. 111–128.
3. B. V. GNEDENKO AND A. N. KOLMOGOROV, *Limit distributions for sums of independent random variables*, trans. from the Russian, Cambridge, Mass., Addison-Wesley Pub. Co., 1954.
4. G. A. HUNT, *Semi-groups of measures on Lie groups*, Trans. Amer. Math. Soc., vol. 81 (1956), pp. 264–293.
5. A. KHINCHIN, *Contribution à l'arithmétique des lois de distribution*, Bull. Univ. État Moscou, Sér. Int., Sect. A: Math. et Mécan., vol. 1 (1937), pp. 6–17.
6. C. KURATOWSKI, *Topologie I*, 2nd ed., Warsaw, 1948.
7. G. W. MACKEY, *Induced representations of locally compact groups. I*, Ann. of Math. (2), vol. 55 (1952), pp. 101–139.
8. K. R. PARTHASARATHY, R. RANGA RAO, AND S. R. S. VARADHAN, *On the category of indecomposable distributions on topological groups*, Trans. Amer. Math. Soc., vol. 102 (1962), pp. 200–217.
9. K. TAKANO, *Central convergence criterion in the multidimensional case*, Ann. Inst. Statist. Math., Tokyo, vol. 7 (1956), pp. 95–102.
10. A. WEILL, *L'intégration dans les groupes topologiques et ses applications*, 2 éd., Paris, Hermann, 1951.
11. K. ITÔ, *On stochastic processes (I)*, Japanese J. Math., vol. 18 (1942), pp. 261–301.
12. P. LÉVY, *L'addition des variables aléatoires définies sur une circonférence*, Bull. Soc. Math. France, vol. 67 (1939), pp. 1–41.
13. K. URBANIK, *Poisson distributions on compact Abelian topological groups*, Colloq. Math., vol. 6 (1958), pp. 13–24.
14. ———, *Gaussian measures on locally compact Abelian topological groups*, Studia Math., vol. 19 (1960), pp. 77–88.
15. C. HERZ, *The spectral theory of bounded functions*, Trans. Amer. Math. Soc., vol. 94 (1960), pp. 181–232.
16. B. M. KLOSS, *Limiting distributions on bicomact Abelian groups*, Teor. Veroyatnost. i Primenen., vol. 6 (1961), pp. 392–421 (in Russian with English summary).
17. Y. KAWADA AND K. ITÔ, *On the probability distribution on a compact group. I*, Proc. Phys.-Math. Soc. Japan (3), vol. 22 (1940), pp. 977–998.
18. L. PONTRJAGIN, *Topological groups*, trans. from the Russian, Princeton University Press, 1946.

INDIAN STATISTICAL INSTITUTE
CALCUTTA, INDIA

