SOME CHARACTERIZATION PROBLEMS OF RANDOM VARIABLES WITH VALUES IN A LOCALLY COMPACT ABELIAN GROUP

By

PETER FLUSSER

Bachelor of Arts
Ottawa University
Ottawa, Kansas
1958

Master of Arts
The University of Kansas
Lawrence, Kansas
1960

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF EDUCATION
May, 1971

SOME CHARACTERIZATION PROBLEMS OF RANDOM VARIABLES WITH VALUES IN A LOCALLY COMPACT ABELIAN GROUP

Thesis Approved:

Thesis Adviser

Dean of the Graduate College

PREFACE

As early as 1963 the Committee on the Undergraduate Program in Mathematics recommended that a college curriculum for the pre-graduate preparation of research mathematicians include, as a bare minimum, courses in real analysis, complex analysis, abstract algebra, topology and probability; [3]. [1] While the committee admitted that at the time the report was written some of its goals were somewhat unrealistic, in the intervening years many undergraduate institutions have begun to implement a program along the lines suggested. The manner of implementation varies greatly with each school, and naturally, depends on its resources.

The purpose of this thesis is to discuss a rather specialized problem which requires for its solution a broad, but not very deep, understanding of all the branches of mathematics named in the report referred to above. As such it could serve to integrate material usually taught in different courses, and since some new results are also presented, it brings the more capable undergraduate to a point in mathematics where active research is still going on.

If $X_1 = R_1 e^{i\Theta_1}$ and $X_2 = R_2 e^{i\Theta_2}$ are complex random variables, their product $X_1 X_2 = R_1 R_2 e^{i(\Theta_1+\Theta_2)}$ where $\Theta_1 \dot{+} \Theta_2 = (\Theta_1 + \Theta_2) \bmod 2\pi$. Thus even in classical probability theory

---

[1] Numbers in parentheses refer to the bibliography at the end of the thesis.

one needs to consider operations other than the usual elementary ones of addition, multiplication and their inverses. This thesis is an attempt to illustrate how the concepts taught in the so-called "abstract" courses shed new light on the more classical material. Thus this material can be used by the college instructor who is looking for applications in his abstract courses in order to help motivate his students. Such an instructor might consider this material useful in deciding which topics to emphasize in other courses and in finding examples which integrate material from algebra, topology and probability. He will thus be able to give one more example which illustrates how the more "abstract" chapters in mathematics become tools for the solution of "concrete" problems, allow a unified treatment of many individual problems and yield new results in the classical theory.

I would like to take this opportunity to thank the members of my committee, Professors Jewett, Kotlarski, Ahmad, and Higgins for their generous help and advice. Special thanks are due to Professor Ignacy I. Kotlarski for many hours of fruitful discussions, his patience as a teacher and for showing me how mathematical research is really done. Thanks are also due to my wife Virginia and to my children for putting up with two years' privations while I returned to being a student, and for living with me during all the ups and downs which are entailed in the pursuit of an advanced degree. And lastly I must thank my father and his wife for their support, both moral and otherwise, without which this whole undertaking would have been impossible.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

We are interested in the following characterization problem:

Problem: For $n \geq 1$, let $X_1, \ldots, X_n$ be independent random variables and let f be a known measurable function from n-dimensional Euclidean space to a space of dimension less than or equal to n. If the distribution of $Y = f(X_1, \ldots, X_n)$ is known, what can be said about the distributions of the X_k's?

Example 1.1 Let X be a random variable. If X has the standard Cauchy distribution with probability density function

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2} ; \quad x \in R \tag{1.1}$$

then for every real number c, the random variable Y defined by

$$Y = \frac{X+c}{1-cX} \tag{1.2}$$

is also distributed like X. Conversely, if for one real c not equal to the tangent of a rational multiple of π, Y, given by (1.2), is distributed like X, then X has the standard Cauchy distribution with probability density function (1.1). (Williams, [26]).

To simplify the wording in the next two examples we shall adopt the following notation: If X_1 and X_2 are identically distributed random variables, we write: $X_1 \sim X_2$.

Example 1.2 Let X_1 and X_2 be independent random variables and assume that $X_1 \sim X_2$ are normally distributed with probability density function

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{x^2}{2\sigma^2}}, \quad x \in R. \tag{1.3}$$

Let

$$Y = \frac{X_1}{X_2}. \tag{1.4}$$

Then Y has the standard Cauchy distribution with probability density function (1.1). The converse, however, is false. For let $X_1 \sim X_2$ be independent, identically distributed random variables with probability density functions given by

$$g(x) = \frac{\sqrt{2}}{\pi} \cdot \frac{1}{1+x^4}, \quad x \in R. \tag{1.5}$$

Then if Y is defined by (1.4), it also has the standard Cauchy distribution with probability density function (1.1). (Mauldon, [20]).

Example 1.3 Let $X_0 \sim X_1 \sim X_2$ be three independent, normally distributed random variables with probability density function (1.3). Let

$$Y_1 = \frac{X_1}{X_0}\ ;\ Y_2 = \frac{X_2}{X_0}. \tag{1.6}$$

Then the joint distribution of (Y_1, Y_2) is the bivariate Cauchy distribution given by the density:

$$g(y_1, y_2) = \frac{1}{2\pi} \cdot \frac{1}{(1 + y_1^2 + y_2^2)^{3/2}}, \quad (y_1, y_2) \in R^2. \tag{1.7}$$

Conversely, let X_0, X_1, X_2 be independent random variables, symmetric about the origin and satisfying $P(X_k = 0) = 0$, $(k = 0, 1, 2)$. If (Y_1, Y_2) is defined by (1.6) and if the joint distribution of (Y_1, Y_2) is given by the probability density function (1.7) then $X_0 \sim X_1 \sim X_2$ and each is normally distributed with density (1.3). (Kotlarski, [14]).

A more thorough study of these and similar examples revealed that such characterizations can also be formulated for random variables with values in a locally compact Abelian group. Since the structure of such a group is simpler than that of the real line the basic difficulties inherent in such problems become more readily apparent and more amenable to solution. Furthermore, the insights gained studying the abstract case show that certain known properties of real random variables can easily be deduced from the more abstractly formulated characterization theorems while at the same time new results in the classical theory are obtained. The purpose of this thesis then is twofold: First, to give an expository account of group valued random variables and second, to present certain new results relating to the characterization of such random variables with applications to the classical theory. It is hoped that the expository part of the thesis will make all the material contained in it comprehensible to the well prepared college senior and thus place within his grasp material heretofore found only in journals or more advanced books.

Specifically, a well prepared college senior is one who has taken a one year course in probability with calculus as a prerequisite and the basic notions of measure theory as a tool. He has taken the standard course in linear algebra, has encountered most of those ideas of abstract algebra found for example in Herstein [8], Chapters II and IV, and those ideas of topology discussed, for instance, in Hocking and Young [9], Chapters I and II.

It is the intention of the writer to prove only those theorems which are either new or the proofs of which are not readily available in the literature. In either case the proofs will be the author's. The expository part of this thesis will contain, therefore, mainly definitions, examples and statements of the main results together with references to the literature where a more thorough discussion of these ideas can be found. We make no claim for completeness in this presentation, on the contrary, our aim here is to penetrate as quickly as possible to the results needed for an understanding of the new material. Thus, for example, we will only discuss the Haar integral on Abelian groups, and thereby avoid all the complications inherent in the non-Abelian case. Similarly we will assume that all our topological groups are second countable and Hausdorff; the reader meeting this material for the first time will be better able to use whatever intuition he has gained from his study of Lebesgue measure on the real line.

This thesis falls naturally into three sections. The first, consisting of Chapters II, III and IV is devoted to developing the basic machinery needed in the sequel. In Chapter II we discuss the basic properties of locally compact second countable Hausdorff Abelian groups, in Chapter III we describe the Haar measure on the σ-algebra

of Borel subsets of such groups, while Chapter IV is devoted to a discussion of the concept of duality. For the purposes of review and to standardize our notation we also present, in Chapter II, a brief discussion of those algebraic and topological concepts used subsequently.

The second part, consisting of Chapter V, deals with the notion of a random variable with values in a locally compact, second countable Abelian group. We discuss the distribution and characteristic function of such a random variable and present certain unpublished results of Professor Kotlarski's as an illustration of applications of these concepts.

In the third part, consisting of Chapters VI and VII we present some new results in the area of characterization problems. In Chapter VI we give a characterization of the uniform distribution on a compact, second countable Abelian group as well as some applications of this result to classical characterization problems. In Chapter VII we prove a theorem which enables us to characterize the marginal distributions of a random variable $X = (X_0, X_1, X_2)$ with values in a locally compact, second countable Abelian group G in terms of the joint distribution of $Y = (Y_1, Y_2)$ where $Y = T(X)$ and T is a homomorphism on G satisfying certain conditions. We also give some applications.

CHAPTER II

LOCALLY COMPACT SECOND COUNTABLE HAUSDORFF ABELIAN GROUPS

Algebra

All groups considered in this thesis will be Abelian. We shall use additive notation, thus "+" will denote the group operation, "0" the identity element, -g the inverse of the element g and for all integers n, we define:

$$ng = \begin{cases} g + \ldots + g \ (n \text{ times}) & \text{if } n > 0 \\ 0 & \text{if } n = 0 \\ (-g) + \ldots + (-g) \ (|n| \text{ times}) & \text{if } n < 0. \end{cases} \tag{2.1}$$

The only exceptions to this convention will arise if the elements of the group under consideration are real or complex numbers. In that case we will use "+" for ordinary addition, "·" for multiplication and symbols such as "$\dot{+}$", or "$\circ$" for specific operations which we may need to define in the sequel. From this point on we omit the word "Abelian" when talking about groups.

A subset H of G is called a <u>subgroup</u> of G if G is a group under the operation induced on H by G. If H and K are subsets of G we define:

$$-H = \{g \in G : -g \in H\} \tag{2.2}$$

and

$$H + K = \{h+k : h \in H,\ k \in K\}. \tag{2.3}$$

If $H = \{h\}$ we will write $h+K$ instead of $\{h\} + K$. If H is a subgroup of G we define the factor group G/H by:

$$G/H = \{g+H : g \in G\}. \tag{2.4}$$

G/H becomes a group under the operation:

$$(g_1 + H) + (g_2 + H) = (g_1 + g_2) + H. \tag{2.5}$$

If S is a subset of G, we say that S <u>generates</u> G, if the smallest subgroup of G which contains S is G itself. The elements of S, in this case, are called <u>generators</u> of G. G is called <u>cyclic</u> if there exists $g \in G$ such that $\{g\}$ generates G. The only infinite cyclic group is the group Z of integers under addition. If m_0 is an integer and $H_{m_0} = \{n m_0 : n \in Z\}$ then Z/H_{m_0} is a cyclic group of <u>order</u> m_0; i.e., Z/H_{m_0} contains exactly m_0 elements. We shall denote the group Z/H_{m_0} by the symbol Z_{m_0}. The only cyclic groups of finite order are the groups Z_{m_0}, $m_0 \in Z$. In general, if G is a group and $g_0 \in G$, we define $Z(g_0) = \{n g_0 : n \in Z\}$ and call $Z(g_0)$ the cyclic subgroup of G <u>generated</u> by g_0.

If G_1 and G_2 are subgroups of G we say that G is the <u>direct sum</u> of G_1 and G_2, in symbols:

$$G = G_1 \oplus G_2, \tag{2.6}$$

if $G = G_1 + G_2$ and $G_1 \cap G_2 = \{0\}$. It follows that in this case every element of G can be expressed uniquely as the sum of an

element of G_1 and an element of G_2. G_1 and G_2 are called direct summands of G. Alternately, if G_1 and G_2 are groups, we may define:

$$G = G_1 \times G_2 = \{(g_1, g_2) : g_1 \in G_1,\ g_2 \in G_2\}. \tag{2.7}$$

G becomes a group under the operation, also denoted by "+", and defined by:

$$(g_1, g_2) + (h_1, h_2) = (g_1 + h_1, g_2 + h_2). \tag{2.8}$$

The subsets $G_1' = \{(g_1, 0) : g_1 \in G_1\}$ and $G_2' = \{(0, g_2) : g_2 \in G_2\}$ are subgroups of G, isomorphic to (i.e. for all intents and purposes indistinguishable from) G_1 and G_2 respectively, and $G = G_1' \oplus G_2'$. Furthermore G/G_1' is isomorphic to G_2 and G/G_2' is isomorphic to G_1.

We shall use the notation $G_1' \oplus G_2'$ or $G_1 \times G_2$ indiscriminately, whichever is more convenient. There is an immediate generalization of these ideas to direct sums of a finite number of direct summands. We shall not need to consider the infinite case.

Topology

Let $(X, \mathfrak{J})$ be a topological space. We shall consider only Hausdorff spaces in this thesis; thus if $x_1, x_2 \in X$ there will always exist disjoint open sets N_1, N_2 containing x_1 and x_2 respectively. A subset $\mathfrak{U} \subset \mathfrak{J}$ is called a base for the topology $\mathfrak{J}$ if every set in $\mathfrak{J}$ is the union of sets in $\mathfrak{U}$. A topological space is called second countable if it has a countable base.

A family $\mathfrak{S}$ of subsets of X is called a cover of X if $X = \bigcup \{S : S \in \mathfrak{S}\}$. If moreover $\mathfrak{S} \subset \mathfrak{J}$ then $\mathfrak{S}$ is called an open cover of X. A cover $\mathfrak{S}_1$ of X which satisfies $\mathfrak{S}_1 \subset \mathfrak{S}$ is called a subcover of $\mathfrak{S}$. A topological space is called compact if every open cover has a finite subcover. A subset X' of X is compact if X' is a compact space under the topology induces on X' by $\mathfrak{J}$. A family $\mathfrak{S}$ of subsets of X is said to have the finite intersection property if every finite subfamily of $\mathfrak{S}$ has a non-empty intersection. We recall that a space X is compact if and only if every family $\mathfrak{S}$ of closed subsets of X with the finite intersection property has a non-empty intersection.

If $x \in X$, a neighborhood of x is an open set containing x. This definition is not the most general, (see, for example, Kelley [11]), but it suffices for the purposes of this presentation. A topological space X is called locally compact if every point has a neighborhood the closure of which is compact.

A topological space X is called connected if it is not the union of two disjoint open sets. It is called Hausdorff (or T_2) if for all $x_1, x_2 \in X$ there exist disjoint neighborhoods N_1 and N_2 of x_1 and x_2 respectively. A subset Y of X is called dense in X if $\overline{Y} = X$, where $\overline{Y}$ is the closure of Y, i.e., the smallest closed subset of X which contains Y. X is called separable if it contains a countable dense subset. Every second countable space is separable ; (Hocking and Young [9], p. 65).

If Y is a subset of the topological space X the family $\mathfrak{J}' = \{Y \cap T : T \in \mathfrak{J}\}$ is a topology for Y called the topology induced on Y by $\mathfrak{J}$. If X_1 and X_2 are topological spaces and we define:

$$X = X_1 \times X_2 = \{(x_1, x_2) : x_1 \in X_1, x_2 \in X_2\} \tag{2.9}$$

then X becomes a topological space if it is given the topology $\mathfrak{J}$ which has as a base the family $\{T_1 \times T_2 : T_1 \in \mathfrak{J}_1, T_2 \in \mathfrak{J}_2\}$. X is called the Cartesian product of X_1 and X_2 and $\mathfrak{J}$ is called the product topology, and denoted by $\mathfrak{J}_1 \times \mathfrak{J}_2$. Again there is an immediate generalization to the Cartesian product of a finite number of topological spaces; we will not need to consider the infinite case.

If X and Y are topological spaces, a function $f : X \to Y$ is called continuous if the inverse image of every open subset of Y is open in X. It is called open if f maps open sets onto open sets. If f is bijective, (one-one and onto), open and continuous, it is called a homeomorphism.

If X is a set, the set of all subsets of X is a topology for X, called the discrete topology. We shall denote the discrete topology by $\mathfrak{D}$.

The symbol R^1 will denote the set of real numbers, and R^n the set of n-tuples of real numbers. C will denote the set of complex numbers and T^1 the set of complex numbers of absolute value one. The symbol $\mathfrak{U}$ will be used to denote the usual topology on R^n and C; it will also be used to denote the topology induced on subsets of these spaces by the usual topology. This apparently inconsistent use of the symbol $\mathfrak{U}$ will not cause confusion.

Topological Groups

Definition 2.1 A topological group is a triple $(G, +, \mathfrak{J})$ where

$(G, +)$ is a group,[1] $(G, \mathfrak{I})$ is a Hausdorff space and

(a) the function $+ : G \times G \to G$ is continuous

(b) the mapping $g \mapsto -g$, $g \in G$, is also continuous.

Some authors, Husain [10], omit the condition that $\mathfrak{I}$ be Hausdorff. Others, Pontryagin [22], replace it with the apparently weaker condition that $\mathfrak{I}$ be T_1. We follow Pontryagin [22]; it can then be shown ([22] p. 97) that $\mathfrak{I}$ is Hausdorff. (As a matter of fact in this case $\mathfrak{I}$ is regular).

The following are examples of topological groups.

Example 2.1: Every group becomes a topological group under the discrete topology $\mathfrak{D}$. Such groups are called discrete groups. We shall have occasion to refer to the discrete groups $(Z, +, \mathfrak{U})$ and $(Z^n, +, \mathfrak{U})$ of integers and n-tuples of integers under addition. We shall denote these groups by Z and Z^n for short. Also, every finite group is discrete.

Example 2.2: $(R^1, +, \mathfrak{U})$, $(R^n, +, \mathfrak{U})$ and $(C, +, \mathfrak{U})$ are topological groups. Again we shall denote these by R^1, R^n and C respectively.

Example 2.3: Let R_+ be the set of positive real numbers. Then $(R_+, \cdot, \mathfrak{U})$ is a topological group.

Example 2.4: $(T^1, \cdot, \mathfrak{U})$ and $(C\setminus\{0\}, \cdot, \mathfrak{U})$ are topological groups. T^1 is called the torus or circle group.

---

[1] We recall that in this thesis we consider only Abelian groups.

The proofs of the following elementary properties of topological groups can be found in Pontryagin [22] pages 96 ff. We assume throughout that G is a topological group.

Proposition 2.1: If $a, b \in G$ then for every neighborhood W of $a+b$ there exist neighborhoods U of a and V of b such that $U + V \subset W$.

Proposition 2.2: If $a \in G$ then for every neighborhood V of $-a$ there exists a neighborhood U of a such that $-U \subset V$.

Proposition 2.3: If $(G, \mathfrak{J})$ is a Hausdorff space and + is an operation on G such that $(G, +)$ is a group and if moreover Propositions 2.1 and 2.2 are satisfied, then $(G, +, \mathfrak{J})$ is a topological group.

Proposition 2.4: If $a \in G$ the mapping $g \mapsto a+g$, $g \in G$, is a homeomorphism on G.

Proposition 2.5: The mapping $g \mapsto -g$, $g \in G$, is a homeomorphism on G.

Proposition 2.6: Let F be a closed subset, U an open subset, P an arbitrary subset and K a compact subset of G. Then $K+F$ and $-F$ are closed and $P+U$ and $-U$ are open subsets of G.

Proposition 2.7: If P and Q are compact subsets of G, so is $P+Q$.

Proposition 2.8: If $g_1, g_2 \in G$ then there exists a homeomorphism φ on G such that $\varphi(g_1) = g_2$.

Definition 2.2: A family $\mathfrak{G}$ of neighborhoods of 0 is called a complete system of neighborhoods of the identity if for every neighborhood N of 0 there exists an element $M \in \mathfrak{G}$ such that $M \subset N$.

Proposition 2.9: If $\mathfrak{G}$ is a complete system of neighborhoods of the identity then the family $\{g+M : g \in G, M \in \mathfrak{G}\}$ is a basis of $\mathfrak{I}$.

Homomorphisms

Let G_1 and G_2 be topological groups and $f: G_1 \to G_2$ a function. Consider the following four conditions which f may or may not satisfy:

A. $f(x_1 + x_2) = f(x_1) + f(x_2)$ for all $x_1, x_2 \in G_1$.

B. f is continuous.

C. f is open.

D. f is bijective.

Definition 2.3: If f satisfies (A) it is called an algebraic homomorphism. If f satisfies (A) and (B) it is called a homomorphism. If f satisfies (A) and (D) it is called an algebraic isomorphism. If f satisfies (A), (B), (C) and (D) it is called an isomorphism.

Note that f is a homeomorphism if and only if it satisfies (B), (C) and (D). Thus every isomorphism is a homeomorphism.

Example 2.5: The logarithm is an isomorphism between $(R_+, \cdot, \mathfrak{U})$ and $(R, +, \mathfrak{U})$.

Definition 2.4: Let $\varphi: G_1 \to G_2$ be an algebraic homomorphism. Then the kernel of φ is the set of all elements of G_1 which

are mapped into the identity of G_2 : $\ker \varphi = \{g \in G_1 : \varphi(g) = 0\}$.

We recall that $\ker \varphi$ is an (algebraic) subgroup of G_1 and that $\ker \varphi = \{0\}$ if and only if φ is an algebraic isomorphism.

Important Convention: From now on we shall consider only locally compact second countable topological groups. To simplify our phraseology, the word "group" will mean "second countable, locally compact, Hausdorff, Abelian topological group." If G is a group and if for some reason we need to ignore its topology, we shall refer to the algebraic group G.

We call attention to the fact that all groups mentioned in our examples are groups under the new meaning of the term. In order to give counter-examples, we recall that $(R^1, +, \mathfrak{D})$ is not second countable and that the Hilbert space ℓ^2, considered as a topological group under vector addition is not locally compact.

There is one small difficulty with our new convention however. A subgroup of a locally compact group need not be locally compact. This is illustrated by the subgroup Q (the rational numbers) of R. However, apart from the fact that the closure $\overline{H}$ of a subgroup H of G is also a subgroup of G we do have the following additional results.

Proposition 2.10: If $f: G_1 \to G_2$ is an open homomorphism from the group G_1 into the group G_2, then $f(G_1)$ is locally compact. (Hocking and Young [9], p. 72).

Proposition 2.11: A subgroup H of a group G is locally compact if and only if it is closed. (Husain [10], p. 69).

Proposition 2.12: A topological group is locally compact if and only if there exist a neighborhood U of 0 such that $\overline{U}$ is compact. (Husain [10], p. 67).

Since most subgroups which we shall need to consider will be closed, Proposition 2.11 overcomes the difficulty inherent in our convention. Non-closed subgroups will be used mainly as counter-examples. We need one more result from the general theory.

Proposition 2.13: Every surjective (onto) homomorphism is open. Hence every bijective homomorphism is an isomorphism. (Pontryagin [22], p. 115).

Many applications of results obtained in the abstract theory of random variables with values in a topological group are based on the following simple considerations:

Let G be a group, S a set and let $f: G \to S$ be bijective. For all $s_1, s_2 \in S$ define:

$$s_1 \circ s_2 = f(f^{-1}(s_1) + f^{-1}(s_2)), \tag{2.10}$$

and let $\mathfrak{F}$ be the set of all subsets B of S such that $f^{-1}(B)$ is open in G. Then $(S, \circ, \mathfrak{F})$ is a topological group and f is an isomorphism between G and S.

Example 2.6: Let $G = (T^1, \cdot, \mathfrak{U})$, $S = [0, 1)$ and define $f: G \to S$ by:

for all $t = e^{i\theta} \in T^1$, $f(t) = f(e^{i\theta}) = \frac{1}{2\pi}\theta$, $0 \leq \theta < 2\pi$.

For this function f, we define:

$$x \dotplus y = f(f^{-1}(x) \cdot f^{-1}(y)), \quad x, y \in S \tag{2.11}$$

and the topology $\mathfrak{T}^1$ on S as in the previous paragraph. Then $(S, \dotplus, \mathfrak{T}^1)$ is a topological group isomorphic to $(T^1, \cdot, \mathfrak{U})$. The operation $\dotplus$ is called addition modulo $[0, 1)$, we shall sometimes write $(x+y) \bmod [0, 1)$ instead of $x \dotplus y$. Furthermore we shall use the symbol T^1 to stand for $(S, \dotplus, \mathfrak{T}^1)$ as well as for $(T^1, \cdot, \mathfrak{U})$, depending on whether the additive or multiplicative notation is more convenient.

We remark that the topological space $(S, \mathfrak{T}^1)$ is not homeomorphic to $(S, \mathfrak{U})$ since the former is compact while the latter is not. The following example illustrates this point further.

Example 2.7: Let S and $\dotplus$ be as in Example 2.6. Then $(S, \dotplus, \mathfrak{U})$ is not a topological group since $\dotplus$ is not continuous in both variables, (Husain [10], pp. 28 and 44), even though it is continuous in each variable separately.

For future reference we state formally as definitions and theorems the following simple applications of the above considerations. For proofs and further details see Aczél [1], pp. 253 ff.

Theorem 2.1: Let $-\infty \leq a < b \leq \infty$ and let $f: [a, b) \to [0, 1)$ be strictly increasing, continuous, surjective and satisfy $f(b - 0) = 1$. For $x, y \in [a, b)$ define:

$$x \circ y = f^{-1}(f(x) \dotplus f(y)) \tag{2.12}$$

and endow $[a,b)$ with the topology $\mathfrak{J}$ induced on it by f^{-1} considered as a function on T^1. Then $([a,b), \circ, \mathfrak{J})$ is a topological group isomorphic to T^1 and f is an isomorphism.

Definition 2.5: Let f satisfy the conditions of Theorem 1. The group $([a,b), \circ, \mathfrak{J})$ will be denoted by $([a,b), f)$.

Theorem 2.2: Let $f: R_+ \to R^1$ be strictly increasing, surjective and continuous. For $x, y \in R$ define:

$$x \circ y = f(f^{-1}(x) \cdot f^{-1}(y)). \tag{2.13}$$

Then $(R, \circ, \mathfrak{U})$ is a topological group isomorphic to R_+ (and hence to R^1) and f is an isomorphism.

Definition 2.6: Let f satisfy the conditions of Theorem 2.2 and let $\circ$ be defined by (2.13). Then $(R, \circ, \mathfrak{U})$ will be denoted by R_f.

Subgroups and Factor Groups

A subgroup H of the topological group $(G, +, \mathfrak{J})$ becomes a topological group under the topology induced on it by $\mathfrak{J}$. Furthermore if H is a subgroup of G so is $\overline{H}$. For reasons discussed in the previous section we shall, for the moment, confine our attention to closed subgroups of G. We note also that every open subgroup of a topological group is closed. The converse, clearly, is false.

Definition 2.7: Let H be a closed subgroup of G. The mapping $\nu: G \to G/H$ defined by

$$\nu(g) = g + H, \quad g \in G \tag{2.14}$$

is an algebraic homomorphism from G onto G/H called the natural homomorphism. The quotient topology on G/H is defined by declaring a subset of G/H to be open if and only if its inverse image under ν is open in G.

Proposition 2.14: Let H be a closed subgroup of G. Then G/H is a locally compact, second countable Hausdorff Abelian group and ν is both open and continuous. (Pontryagin [22], p. 112).

Proposition 2.15: Let H be a closed subgroup of G and $\nu : G \to G/H$ the natural homomorphism. Then M is a closed subgroup of G/H if and only if $\nu^{-1}(M)$ is a closed subgroup of G containing H. (Pontryagin [22], p. 114).

Proposition 2.16: Let G_1 and G_2 be two topological groups, not necessarily locally compact or second countable. If $\varphi : G_1 \to G_2$ is a homomorphism, then $\ker \varphi$ is a closed subgroup of G_1. If φ is open and surjective then G_2 is isomorphic to $G_1/\ker \varphi$ under the mapping $f : G_1/\ker \varphi \to G_2$ defined by

$$f(g + \ker \varphi) = \varphi(g), \quad g \in G_1. \tag{2.15}$$

If φ is not open then f is a continuous algebraic isomorphism between $G_1/\ker \varphi$ and G_2, however it will fail to be open. If G_1 and G_2 are second countable and locally compact and φ is surjective then f is an isomorphism. (Pontryagin [22], p. 113).

Example 2.8: Consider R/Z. The natural homomorphism $\nu : R \to R/Z$ is given by $\nu(x)$ = fractional part $(x) + Z$, and hence R/Z is isomorphic to $(T^1, +)$.

The Subgroups of R^1 and T^1

We give now a full description of the subgroups of R^1 and T^1. These results are both elementary and well known; we provide proofs because we have not been able to find elementary proofs in the literature. Most authors, Hocking and Young [9], Husain [10], Kelley [11], Pontryagin [22], either take the following theorems for granted, or state them as exercises.

We assume that the reader is familiar with the axioms of a complete ordered field (Kelley [11], pp. 19 ff.), is aware of the fact that these axioms are categorical, and that R^1 with the usual operations of addition and multiplication is a complete ordered field. We recall that a complete ordered field is Archimedean; i.e., for all $0 < x < y < R^1$ there exists a unique non-negative integer n such that $(n-1)x < y \leq nx$.

Lemma 2.1: Let G be a subgroup of R^1 and assume that there exists a non-zero sequence $\{x_n\}$ of elements of G converging to zero. Then G is dense in R^1.

Proof: Let $a \in G$ and choose $\varepsilon > 0$. We shall show that there exists an element $g \in G$ such that $|a - g| < \varepsilon$.

Without loss of generality we may assume that $\varepsilon < a - \varepsilon$.

Since $\{x_n\} \to 0$ there exists a natural number N such that $|x_N| < \varepsilon$. Let $y = |x_N|$. Then $y \in G$ and $0 < y < \varepsilon$.

Since the ordering on R^1 is Archimedean there exists a positive integer m such that

$$(m-1)y < a \leq my. \tag{2.16}$$

Since $y < \varepsilon$, we obtain from (2.16)

$$my - y = (m-1)y > a - \varepsilon. \tag{2.17}$$

Putting $g = (m-1)y$, and noting that $g \in G$, (2.16) and (2.17) imply

$$|a - g| < \varepsilon \tag{2.18}$$

and the lemma is proved. //.

Lemma 2.2: Every non-trivial subgroup of R^1 is infinite.

Proof: If H is a non-trivial subgroup of R^1 and $g \in H$, $g \neq 0$, then $Z(g) \subset H$. //.

Corollary: R^1 has no non-trivial compact subgroups.

Theorem 2.3: Every closed proper subgroup of R^1 is cyclic, and conversely, every cyclic subgroup of R^1 is closed.

Proof: Let G be a closed non-trivial subgroup of R^1 and let $P = G \cap R_+$. Let $x_0 = \inf P$. Note that $0 \notin P$, and that, since G is closed, $x_0 \in G$.

If $x_0 = 0$ then there exists a non-zero sequence $\{x_n\}$ in G converging to 0. By Lemma 2.1, G is then dense in R^1, and since G is closed, $G = R$, a contradiction. Hence $x_0 \neq 0$. Thus x_0 is the smallest positive element of G.

We claim that $G = Z(x_0)$.

If not, then there exists $y \in G$, $y > 0$, such that $y \neq nx_0$, $n = 0, 1, 2, \ldots$. Clearly $x_0 < y$. Hence we can find an integer $m \geq 2$ such that

$$(m-1)x_0 < y < mx_0 . \tag{2.19}$$

Hence

$$0 < y - (m-1)x_0 < mx_0 - (m-1)x_0 = x_0 \tag{2.20}$$

and since $y - (m-1)x_0 \in G$ and $m \geq 2$, (2.20) contradicts the fact x_0 is the smallest positive element of G.

The converse is trivial, since every cyclic subgroup of R^1 is discrete and hence closed. //.

Corollary: A subgroup G of R^1 is closed if and only if it is discrete.

Theorem 2.4: If G is a non-trivial subgroup of R^1 and G is not cyclic (or not closed) then G is dense in R^1.

Proof: By Lemma 2.2, G is infinite. Since G is not cyclic, neither is $\overline{G}$, (the closure of G). Hence, by Lemma 2.1, $\overline{G} = R^1$ and the theorem is proved. //.

In connection with Propositions 2.6 and 2.7, the following example may be of interest since it illustrates the fact that the sum of two closed subsets of a topological group need not be closed. As a matter of fact we show that the direct sum of two closed subgroups need not be closed.

Example 2.9: Let $H_1 = Z$, $H_2 = Z(\sqrt{2})$. Then clearly H_1 and H_2 are closed subgroups of R^1, and

$$H_1 + H_2 = H_1 \oplus H_2 = \{m + \sqrt{2}\, n : m, n \in Z\}.$$

To show that $H_1 \oplus H_2$ is not closed, it suffices to exhibit a sequence of non-zero elements of $H_1 \oplus H_2$ converging to zero.

We first note that if $m + n\sqrt{2} \in H_1 \oplus H_2$, then

$$(m + n\sqrt{2})^2 = m^2 + 2n^2 + 2mn\sqrt{2} \in H_1 \oplus H_2 .$$

Furthermore $0 < -1 + \sqrt{2} < 1$.

Thus $(-1 + \sqrt{2})^{(2^n)} \in H_1 \oplus H_2$ for all positive integers n and since $\lim_{n\to\infty} (-1 + \sqrt{2})^{(2^n)} = 0$ our assertion is proved. //.

We now examine the subgroups of T^1. We recall (Example 2.8), that $T^1 = R^1/Z$. In what follows, we let $\nu : R^1 \to T^1$ be the natural homomorphism.

Theorem 2.5: Let M be a subgroup of T^1. Then M is closed if and only if M is a finite cyclic group.

Proof: Assume that M is closed. By Proposition 2.15, $M = \nu(H)$, where H is a closed (and therefore cyclic) subgroup of R^1 which contains Z. Hence H is generated by a rational number n/m, and if n and m are so chosen that $m > 0$ and n and m are relatively prime, then M is isomorphic to Z_m.

The converse is trivial. //.

Theorem 2.6: Let D be a subgroup of T^1 which is not closed (or not finite cyclic). Then D is dense in T^1.

Proof: The proof is completely analogous to the proof of Theorem 2.4. //.

Theorem 2.7: Let $\alpha \in [0, 1)$ be irrational. Then $Z(\alpha)$ is dense in T^1.

Proof: If $Z(\alpha)$ is not dense in T^1 then by Theorem 2.6, $Z(\alpha)$ is closed in T^1. Hence (by Proposition 2.15), $\nu^{-1}(Z(\alpha))$ is a cyclic subgroup of R^1 which contains Z. But $Z \cap \nu^{-1}(Z(\alpha)) = \{0\}$, and this contradiction proves the theorem. //.

Direct Sums and Cartesian Products

Assume that $G = G_1 \oplus G_2$. Since G is assumed to be locally compact and second countable, G_1 and G_2 with the topologies induced on them by G are closed (and hence locally compact) subgroups of G. Furthermore the product topology on $G_1 \times G_2$ is the same as the original topology on G; (Pontryagin [22], pp. 121 ff.). Also G_1 is isomorphic to G/G_2 and G_2 is isomorphic to G/G_1.

Similarly if G_1 and G_2 are groups and $G = G_1 \times G_2$ is given the product topology then G decomposes into the direct sum $G_1' \oplus G_2'$ where $G_1' = \{(g_1, 0) : g_1 \in G_1\}$ and $G_2' = \{(0, g_2) : g_2 \in G_2\}$, and G_k' is isomorphic to G_k for $k = 1, 2$. For $k = 1, 2$, the maps $p_k : G \to G_k'$ defined by

$$\left.\begin{aligned} P_1(g_1, g_2) &= (g_1, 0) \\ P_2(g_1, g_2) &= (0, g_2) \end{aligned}\right\} (g_1, g_2) \in G \qquad (2.21)$$

are open homomorphisms.

We will thus use the notation $G = G_1 \times G_2$ or $G = G_1' \oplus G_2'$ indiscriminately, which ever is more convenient. The generalization

to an arbitrary finite number of direct summands is immediate; we shall not need to consider infinite direct products.

Example 2.10: R^n is the direct sum of n copies of R^1; (n = 2, 3, ...).

By analogy with Example 2.10, if G is a group, we shall denote by G^n the direct sum of n copies of G.

Structure Theorems

In this section we state certain results which will not be needed in the sequel but which are interesting in their own right and which may give the reader a stronger intuition about the structure of the groups under consideration. We recall that the word "group" means locally compact, second countable Hausdorff Abelian group. Some of the theorems stated below are valid under less restrictive conditions. For proofs see the references listed.

Theorem 2.8: Every group is homeomorphic to a complete metric space. (Husain [10], p. 69).

Definition 2.8: A group G is compactly generated if there exists a neighborhoof V of 0 such that $\overline{V}$ is compact and V generates G.

Theorem 2.9: A connected group is compactly generated. (Pontryagin [22], p. 129).

Theorem 2.10: If G is compactly generated, then $G = R^n \oplus K + Z^m$ where K is a compact subgroup of G. (Pontryagin [22], p. 269).

Theorem 2.11: Let G be a group and F a compact subset of G. Then there exists a compactly generated subgroup H of G such that $F \subset H$. (Pontryagin [22], p. 265).

CHAPTER III

MEASURE AND INTEGRATION ON GROUPS

Haar Measure

Let Ω be a nonvoid set. A family $\mathfrak{S}$ of subsets of Ω is called a σ-algebra if

(a) $\Omega \in \mathfrak{S}$

(b) $A \in \mathfrak{S} \Rightarrow \Omega \setminus A \in \mathfrak{S}$

(c) $\{A_k\}_{k=1}^{\infty} \subset \mathfrak{S} \Rightarrow \bigcup_{k=1}^{\infty} A_k \in \mathfrak{S}.$

It follows that if $\mathfrak{S}$ is a σ-algebra on Ω, then

(d) $\emptyset \in \mathfrak{S}$

(e) $A, B \in \mathfrak{S} \Rightarrow A \setminus B \in \mathfrak{S}$

(f) $\{A_k\}_{k=1}^{\infty} \subset \mathfrak{S} \Rightarrow \bigcap_{k=1}^{\infty} A_k \in \mathfrak{S}.$

If $\mathfrak{F}$ is a family of subsets of Ω, the smallest σ-algebra containing $\mathfrak{F}$ is called the σ-algebra generated by $\mathfrak{F}$ and denoted by gen $\mathfrak{F}$.

If $(G, +, \mathfrak{I})$ is a group we shall write $\mathfrak{B} = \text{gen } \mathfrak{I}$. An element of $\mathfrak{B}$ is called a Borel set; (Rudin, [25]). We note that $\mathfrak{B}$ is also generated by the closed subsets of G, and since G is locally compact and second countable, $\mathfrak{B}$ is also generated by the compact subsets of G.

If Ω is a set and $\mathfrak{S}$ a σ-algebra of subsets of Ω, the pair $(\Omega, \mathfrak{S})$ is called a measurable space, and the elements of $\mathfrak{S}$ are called measurable sets.

A real valued set function on $\mathfrak{S}$ is called a measure on the measurable space $(\Omega, \mathfrak{S})$ if it satisfies the following conditions:

(A) $\mu(A) \geq 0$ for all $A \in \mathfrak{S}$

(B) $\mu(\emptyset) = 0$

(C) If $A_1, A_2, \ldots$ is a sequence of disjoint measurable sets, then $\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu\left(A_k\right)$.

The triple $(\Omega, \mathfrak{S}, \mu)$ is called a measure space. If in addition μ satisfies

(D) $\mu(\Omega) = 1$

then μ is called a probability measure and $(\Omega, \mathfrak{S}, \mu)$ is called a probability space.

A measure μ on $(\Omega, \mathfrak{S})$ is called regular if for all $A \in \mathfrak{S}$,

$$\mu(A) = \sup\{\mu(K) : K \text{ is compact and } K \subset A\}. \qquad (3.1)$$

The support of μ is the smallest closed subset F of Ω such that $\mu(\Omega \setminus F) = 0$. We write $F = \operatorname{supp} \mu$. If μ is a measure on the Borel sets of a group G then the support of μ is defined to be the smallest closed subgroup H of G such that $\mu(G \setminus H) = 0$. In that case we write $H = c(\mu)$. Note, that in general, if μ is a measure on the group G, $\operatorname{supp} \mu \neq c(\mu)$. It has become traditional, in the literature, to use the same term for these different concepts.

Let $(G, +, \mathfrak{I})$ be a group and $\mathfrak{B} = \text{gen}\,\mathfrak{I}$. A measure μ on $(G, \mathfrak{B})$ is called a Haar measure if it satisfies the following conditions:

(α) $\mu(G) \neq 0$

(β) For all $A \in \mathfrak{B}$ and $y \in G$, $\mu(A) = \mu(g + A)$

(γ) For all $A \in \mathfrak{B}$, $\mu(A) = \mu(-A)$

Conditions (α) and (β) are really the defining properties of Haar measure. Condition (β) is called the translation invariant property of Haar measure. Since G is Abelian, it can be shown that (α) and (β) imply (γ); Halmos [7].

Example 3.1: If $G = R^n (n = 1, 2, \ldots)$ then Lebesgue measure restricted to the Borel subsets of R^n is a Haar measure.

Example 3.2: If G is a finite or countable group then $\mathfrak{B}$ is the set of all subsets of G. If $A \subset G$, we shall let $\#(A)$ denote the number of elements of A. Then $\#$ is a Haar measure on G. We note that $\#(A) = \infty$ for all infinite subsets of G. The measure $\#$ is also called counting measure.

The following is the basic result of this chapter.

Theorem 3.1: If G is a group then there exists a Haar measure on G. If μ_1 and μ_2 are two Haar measures on G then there exists a constant $k > 0$ such that $\mu_1(A) = k\mu_2(A)$ for all $A \in \mathfrak{B}$.

A proof of Theorem 3.1 together with a description of the actual construction of Haar measure can be found in Halmos [7], Chapters XI

and XII. The proofs of the following basic properties of Haar measure can also be found there.

Theorem 3.2: A Haar measure is regular. Moreover if μ is a Haar measure on G then for all $A \in \mathfrak{B}$

$$\mu(A) = \inf\{\mu(U) : A \subset U, \ U \text{ is open}\}.$$

If U is an open set, $\mu(U) > 0$.

Theorem 3.3: $\mu(G) < \infty$ if and only if G is compact.

If G is compact we may choose a Haar measure μ on G such that $\mu(G) = 1$. This measure is called the normalized Haar measure or the Haar measure on G.

Example 3.3: The normalized Haar measure on $(T^1, \dot{+})$ is the Lebesgue measure restricted to the Borel subsets of $[0, 1)$. (Pontryagin [22], p. 201).

Theorem 3.4: Let μ be a Haar measure on G. Then G is discrete if and only if there exists an element $g \in G$ such that $\mu(g) > 0$. (Halmos [7]).

The Haar Integral

Let $(G, +, \mathfrak{I})$ be a group, $\mathfrak{B} = \text{gen } \mathfrak{I}$ and let μ be a Haar measure on G.

A function $f : G \to X$ where $(X, \mathfrak{I}_X)$ is a topological space is called a measurable function if the inverse image of every open subset of X belongs to $\mathfrak{B}$. If $\mathfrak{B}_X = \text{gen } \mathfrak{I}_X$, then f is called a Borel function

if the inverse image of every element of $\mathfrak{B}_X$ belongs to $\mathfrak{B}$. Since we will be interested only in functions whose ranges are groups these two concepts will coincide in this thesis and we shall refer to measurable functions for short. We recall that all continuous functions are measurable, as are all (pointwise) limits of sequences of measurable functions.

Since $(G, \mathfrak{B}, \mu)$ is a measure space, we may, by standard methods (Rudin [25], pp. 19 ff.) define the Lebesgue integral of a measurable function $f: G \to R_+$. We shall denote by $L^1(G)$ the set of all complex valued Borel functions on G satisfying:

$$\int_G |f| d\mu < \infty \tag{3.2}$$

where $\int_G |f| d\mu$ is the Lebesgue integral of the positive function $|f|$.

Denoting the positive and negative parts of a real-valued function f on G by f^+ and f^-, we define the Haar integral of a function $f = g + hi$ by:

$$\int_G f d\mu = \int_G f^+ d\mu - \int_G f^- d\mu + i\left[\int_G h^+ d\mu - \int_G h^- d\mu\right]. \tag{3.3}$$

Since $g^+ \leq |g| \leq |f|$ etc., the Haar integral is defined for every complex measurable function satisfying (3.2); i.e., for all $f \in L^1(G)$. Finally, we define:

$$||f|| = \int_G |f| d\mu, \quad f \in L^1(G). \tag{3.4}$$

The material that follows is intended for the more advanced reader. The beginner need only study Theorem 3.6 below.

Let $f, g \in L^1(G)$. The convolution of f and g is defined by the formula:

$$(f * g)(x) = \int_G f(x - y)\, g(y)\, d\mu(y) \tag{3.5}$$

provided that

$$\int_G |f(x - y)\, g(y)|\, d\mu(y) < \infty . \tag{3.6}$$

If $f, g \in L^1(G)$, (3.6) holds for almost all $x \in G$; (Rudin [24], p. 4).

If $f, g \in L^1(G)$ and if $\mu\{x \in G : f(x) \neq g(x)\} = 0$ we shall write $f = g$ a.e. $[\mu]$. Writing temporarily $f \sim g$ instead of $f = g$ a.e. $[\mu]$, it can easily be seen that $\sim$ is an equivalence relation on $L^1(G)$. We define:

$$\mathfrak{L}^1(G) = L^1(G)/\sim, \tag{3.7}$$

i.e., $\mathfrak{L}^1(G)$ is the set $L^1(G)$ with functions which differ on a set of μ-measure zero identified.

The basic properties of the Haar integral can now be summarized in the following theorem. (For proofs see Rudin [24], p. 6).

Theorem 3.5: If G is a group and μ a Haar measure on G, then $\mathfrak{L}^1(G)$ is a commutative Banach Algebra under the usual vector operations, convolution and norm defined by (3.4). The Haar integral is a bounded translation invariant linear functional on $\mathfrak{L}^1(G)$.

We shall find it more convenient to restate the most pertinent parts of this theorem as follows:

Theorem 3.6: The Haar integral satisfies the following equations for all $f, g \in \mathfrak{L}^1(G)$:

$$\int_G \alpha f \, d\mu = \alpha \int_G f \, d\mu, \quad \alpha \in C \tag{3.8}$$

$$\int_G (f+g) d\mu = \int_G f \, d\mu + \int_G g \, d\mu \tag{3.9}$$

$$\int_G |f| d\mu \geq 0 \tag{3.10}$$

$$\int_G f(x+a) d\mu(x) = \int_G f \, d\mu \text{ for all } a \in G \tag{3.11}$$

$$\int_G f(-x) d\mu(x) = \int_G f \, d\mu \tag{3.12}$$

$$\left| \int_G f \, d\mu \right| \leq \int_G |f| d\mu \tag{3.13}$$

$$\int_E d\mu = \mu(E) = \int_G \mathcal{J}_E d\mu \tag{3.14}$$

where $\mathcal{J}_E$ is the indicator function of E and $E \in \mathfrak{B}$.

Product Measures

Let $(\Omega_1, \mathfrak{S}_1, \mu_1)$ and $(\Omega_2, \mathfrak{S}_2, \mu_2)$ be measures spaces. Then $\Omega_1 \times \Omega_2$ becomes a measurable space under the σ-algebra $\mathfrak{S}$ generated

by the family $\{S_1 \times S_2 : S_1 \in \mathfrak{S}_1, S_2 \in \mathfrak{S}_2\}$. We shall write:

$$\mathfrak{S} = \mathfrak{S}_1 \times \mathfrak{S}_2 . \tag{3.15}$$

Let $E \subset \Omega_1 \times \Omega_2$, $x \in \Omega_1$, $y \in \Omega_2$. The x-section E_x of E is that subset of Ω_2 defined by

$$E_x = \{t : (x, t) \in E\} \tag{3.16}$$

and the y-section, E^y of E is that subset of Ω_1 defined by

$$E^y = \{s : (s, y) \in E\} \tag{3.17}$$

If $E \in \mathfrak{S}$, then $E_x \in \mathfrak{S}_2$ and $E^y \in \mathfrak{S}_1$ for all $x \in \Omega_1$, respectively $y \in \Omega_2$.

Now let $(G_1, +, \mathfrak{I}_1)$ and $(G, +, \mathfrak{I}_2)$ be groups, $\mathfrak{B}_1 = \text{gen } \mathfrak{I}_1$, $\mathfrak{B}_2 = \text{gen } \mathfrak{I}_2$ and μ_1 and μ_2 measures, not necessarily Haar measures, on $(G_1, \mathfrak{B}_1)$ and $(G_2, \mathfrak{B}_2)$ respectively. Let $G = G_1 \times G_2$ and $\mathfrak{B} = \mathfrak{B}_1 \times \mathfrak{B}_2$. Since G_1 and G_2 are locally compact and second countable, $\mathfrak{B} = \text{gen } (\mathfrak{I}_1 \times \mathfrak{I}_2)$. Let $E \in \mathfrak{B}$. Define the functions f and g on G_1 and G_2 respectively by:

$$\begin{aligned} f(x) &= \mu_2(E_x), \quad x \in G_1 \\ g(y) &= \mu_1(E^y), \quad y \in G_2 . \end{aligned} \tag{3.18}$$

Then f and g are non-negative measurable functions and

$$\int_{G_1} f \, d\mu_1 = \int_{G_2} g \, d\mu_2 . \tag{3.19}$$

For proofs of the above statements see Halmos [7], p. 143.

For all sets $E \subset \mathfrak{B}$ we now define

$$\mu(E) = \int_{G_1} \mu_2(E_x)d\mu_1(x) = \int_{G_2} \mu_1(E^y)d\mu_2(y) \tag{3.20}$$

Then μ becomes a measure on $(G,\mathfrak{B})$ satisfying:

$$\mu(S_1 \times S_2) = \mu_1(S_1)\mu_2(S_2), \quad S_1 \in \mathfrak{S}_1, \quad S_2 \in \mathfrak{S}_2, \tag{3.21}$$

and (3.21) determines μ uniquely; (Halmos [7], p. 143).

The measure μ defined in this manner is called the product of μ_1 and μ_2; in symbols:

$$\mu = \mu_1 \times \mu_2. \tag{3.22}$$

We shall need the following theorems in the sequel; the first is a special case of Fubini's Theorem.

Theorem 3.7: If f and g are integrable functions on the groups G_1 and G_2 respectively, then the function h defined by $h(x, y) = f(x)\,g(y)$ is an integrable function on $G_1 \times G_2$ and

$$\int_{G_1 \times G_2} h\,d(\mu_1 \times \mu_2) = \int_{G_1} f\,d\mu_1 \int_{G_2} g\,d\mu_2. \tag{3.23}$$

(Halmos [7], p. 149).

Theorem 3.8: Let $(\Omega_1, \mathfrak{B}_1, \mu_1)$ be a measure space and $(\Omega_2, \mathfrak{B}_2)$ be a measurable space. Let $T : \Omega_1 \to \Omega_2$ be measurable and define a set function μ_2 on $\mathfrak{B}_2$ by

$$\mu_2(\mathfrak{B}) = \mu_1(T^{-1}(B)), \quad B \in \mathfrak{B}_2 \tag{3.24}$$

Then μ_2 is a measure on $(\Omega_2, \mathfrak{B}_2)$. Moreover, for every real or complex-valued measurable function g on Ω_2 we have:

$$\int_{\Omega_2} g(y)d\mu_2(y) = \int_{\Omega_1} g(T(x))d\mu_1(x) \qquad (3.25)$$

in the sense that if either integral exists so does the other and the two are equal. (Halmos [7], p. 163).

Haar Measure on Subgroups, Factor Groups and Direct Sums of Compact Groups

Since we are primarily interested in probability measures on groups, we will, in view of Theorem 3.3, restrict our attention to compact groups in this section. Although some of the results mentioned will be valid in a more general setting, the general theory is considerably more difficult. For a more comprehensive account see Loomis, [18].

Let $(G, +, \mathfrak{J})$ be a compact group, $\mathfrak{B} = \text{gen } \mathfrak{J}$, and μ the normalized Haar measure on G. Thus $\mu(G) = 1$, that is $(G, \mathfrak{B}, \mu)$ is a probability space. Let H be a closed subgroup of G. If $\mathfrak{J}'$ is the topology induced on H by $\mathfrak{J}$, $\mathfrak{B}' = \text{gen } \mathfrak{J}'$ and $\mu' = \mu\big|_{\mathfrak{B}'}$, then $(H, \mathfrak{B}', \mu')$ is a measure space. However μ' need not be a Haar measure, as the following example illustrates:

Example 3.4: Consider T^1. The only closed proper subgroups of T^1 are finite cyclic groups. (Theorem 2.5). Haar measure on T^1 is Lebesgue measure, m. (Example 3.3). Thus if H is a closed subgroup of T^1, $\mu'(H) = 0$ and so μ' is not a Haar measure on H.

We do, however, have the following results:

Theorem 3.9: Let $G = G_1 \oplus G_2$ where G_1 and G_2 need not be compact. Let μ_1 and μ_2 be Haar measures on G_1 and G_2 respectively. If μ is a Haar measure on G, then μ is a constant multiple of $\mu_1 \times \mu_2$. (Halmos [7], p. 263).

Theorem 3.10: Let H be a compact subgroup of G, and μ be a Haar measure on G. Let $\nu : G \to G/H$ be the natural homomorphism defined by (2.14). Then $\tilde{\mu} = \mu \nu^{-1}$ is a Haar measure on G/H. (Halmos [7], p. 279).

The following is a partial converse of Theorem 3.9 and is an immediate Corollary of Theorem 3.10.

Theorem 3.11: Let G_1 and G_2 be compact groups and let $G = G_1 \times G_2$. Let μ be the normalized Haar measure on G. For every Borel subset B_1 of G_1 define

$$\mu_1(B_1) = \mu(B_1 \times G_2) \tag{3.26}$$

and similarly, for every Borel subset B_2 of G_2 define

$$\mu_2(B_2) = \mu(G_1 \times B_2). \tag{3.27}$$

Then u_1 and μ_2 are the normalized Haar measures on G_1 and G_2 respectively.

The Metric Space of Borel Subsets of a Compact Group

We conclude this chapter with some technical results needed subsequently in Chapter VI.

Definition 3.1: A pseudo-metric space is a pair (X, d) where X is a set and d is a function; $d : X \times X \to R$ satisfying:

(A) $d(x, y) \geq 0$; $\quad x, y \in X$

(B) $d(x, x) = 0$; $\quad x \in X$

(C) $d(x, y) = d(y, x)$; $\quad x, y \in X$

(D) $d(x, y) + d(y, z) \geq d(x, z)$; $\quad x, y, z \in X$.

The function d is called a pseudo-metric on X. If in addition d satisfies:

(E) $d(x, y) = 0 \Leftrightarrow x = y$; $\quad x, y \in X$

then (X, d) is called a metric space and d is called a metric on X.

Let (X, d) be a pseudo-metric space. Define a relation $\sim$ on X by:

$$x \sim y \Leftrightarrow d(x, y) = 0. \tag{3.28}$$

It is readily seen that $\sim$ is an equivalence relation on X. Hence we may consider $X/\sim$; the set of equivalence classes induced by $\sim$ on X. Writing $X' = X/\sim$ and denoting the element of X' to which $x \in X$ belongs by x', the function $d' : X' \times X' \to R$ defined by

$$d'(x', y') = d(x, y); \quad x, y \in X \tag{3.29}$$

becomes a metric on X'.

If (X, d) is a metric space, $x \in X$ and $r > 0$, we define the open disk $S(x, r)$ by:

$$S(x, r) = \{y \in X : d(x, y) < r\}. \tag{3.30}$$

The topology $\mathfrak{J}$ induced by d on X is the topology with base $S(x, r)$ for all $x \in X$, and $r > 0$. $(X, \mathfrak{J})$ is a Hausdorff space. If the metric space (X, d) is separable, it is second countable; (Hocking and Young [9], p. 11).

A sequence in the metric space (X, d) is a function on the set of natural numbers with range in X. A sequence $\{x_n\}$ is said to converge if there exists a point $x_0 \in X$ such that every neighborhood of x_0 contains all except at most a finite number of points of $\{x_n\}$. The point x_0 is called the limit of the sequence $\{x_n\}$. A limit, if it exists, is unique. A sequence $\{x_n\}$ is called a Cauchy sequence if for every $\varepsilon > 0$ there exists a natural number N such that if $n, m > N$ then $d(x_n, x_m) < \varepsilon$. Every convergent sequence is a Cauchy sequence. If a metric space (X, d) has the property that every Cauchy sequence converges, it is called complete.

Now let G be a compact group, $\mathfrak{B}$ the family of Borel subsets of G, and let π be a probability measure on $(G, \mathfrak{B})$. For $E, F \in \mathfrak{B}$, define the symmetric difference on E and F by:

$$E \Delta F = (E \cup F) \setminus (E \cap F) \tag{3.31}$$

We note that $E \Delta F = (E \setminus F) \dot{\cup} (F \setminus E)$ where the dot over the symbol $\cup$ indicates that the sets are disjoint.

Finally for $E, F \in \mathfrak{B}$ define:

$$\rho(E, F) = \pi(E \Delta F). \tag{3.32}$$

It can easily be verified that $(\mathfrak{B}, \rho)$ is a pseudo-metric space.

If we now identify Borel sets whose symmetric difference is of π-measure zero, that is if we carry out the construction described at the beginning of this section, we obtain a metric space $(\mathfrak{B}', \rho')$. As before, if $B \in \mathfrak{B}$, we denote the equivalence class to which B belongs by B'. We shall need the results stated in the following two Theorems.

Theorem 3.12: $(\mathfrak{B}', \rho')$ is a complete separable metric space. (Halmos [7], p. 168).

Theorem 3.13: Let $B_0 \in \mathfrak{B}$. Then the function $f: G \to \mathfrak{B}'$ given by

$$f(g) = (g + B_0)' \tag{3.33}$$

is continuous. (Halmos [7], p. 268).

CHAPTER IV

DUALITY

The Dual of a Group

Let $(G, +, \mathfrak{I})$ be a group. A (continuous) homomorphism $\gamma : G \to (T^1, \cdot, \mathfrak{U})$ is called a <u>character</u> on G. Thus a character is a continuous function γ from G into C satisfying:

$$\gamma(g_1 + g_2) = \gamma(g_1)\gamma(g_2), \qquad g_1, g_2 \in G$$

$$|\gamma(g)| = 1, \qquad g \in G.$$

The set of all characters on G is called the <u>dual</u> of G and is denoted by G^*. If $g \in G$ and $\gamma \in G^*$ we shall write $< g, \gamma >$ instead of $\gamma(g)$. If G is not trivial, neither is G^*; as a matter of fact for all $g \in G$, $g \neq 0$, there exists $\gamma \in G^*$ such that $< g, \gamma > \neq 1$. (Pontryagin [22], p. 274).

If for all $\gamma_1, \gamma_2 \in G^*$ we define:

$$< g, \gamma_1 + \gamma_2 > = < g, \gamma_1 > < g, \gamma_2 >; \qquad g \in G \tag{4.1}$$

G^* becomes an algebraic group with the following identity: the constant map of G into the complex number 1.

We note that for all $g \in G$, $\gamma \in G^*$:

$$< -g, \gamma > = < g, -\gamma > = (< g, \gamma >)^{-1} = \overline{< g, \gamma >}. \tag{4.2}$$

Definition 4.1, Theorem 4.1 and the remark in the succeeding paragraph enable us to give an elegant definition of the topology of G^*. It is intended for the more advanced reader, since it is beyond the scope of this thesis to define the concepts of "a maximal ideal space of a Banach algebra" and "Gelfand topology." A more elementary definition of the topology of G^* is given in Theorem 4.2. The beginner may omit this material if he is willing to accept the fact that there exists a topology on G^* satisfying the conclusions of Theorems 4.3 to 4.9.

Definition 4.1: Let $f \in \mathfrak{L}^1(G)$, and let μ be a Haar measure on G. The Fourier transform of f is the complex valued function $\hat{f}$ on G^* defined by:

$$\hat{f}(\gamma) = \int_G f(x) < x, \gamma > d\mu(x); \qquad \gamma \in G^*. \tag{4.3}$$

Theorem 4.1: Let $\gamma \in G^*$. Then the map $f \mapsto \hat{f}(\gamma)$ is a complex homomorphism on $\mathfrak{L}^1(G)$ and is not identically zero. Conversely, every non-zero complex homomorphism of $\mathfrak{L}^1(G)$ is obtained in this way, and distinct characters induce distinct homomorphisms. (Rudin [24], p. 7).

In view of Theorem 4.1 we may identify G^* with the maximal ideal space of $\mathfrak{L}^1(G)$. Thus G^* becomes a locally compact Hausdorff space when endowed with the Gelfand topology.

A simpler way to characterize the topology of G^* and at the same time to show that with this topology G^* becomes a locally compact Hausdorff Abelian group is given in the next theorem.

Theorem 4.2: Let G and G^* be as above. Then

(a) $< g, \gamma >$ is a continuous function on $G \times G^*$.

(b) Let K and K^0 be compact subsets of G and G^* respectively, and let $r > 0$. Let

$$U_r = \{z \in C : |1 - z| < r\}$$

and put

$$N(K, r) = \{\gamma \in G^* : < g, \gamma > \in U_r \text{ for all } g \in K\}$$

and

$$N(K^0, r) = \{g \in F : < g, \gamma > \in U_r \text{ for all } \gamma \in K^0\}.$$

Then $N(K, r)$ and $N(K^0, r)$ are open subsets of G^* and G respectively.

(c) The family $\{N(K, r) : r > 0\}$ and their translates is a base for the topology of G^*.

(d) G^* is a locally compact Hausdorff Abelian group. (Rudin [24], p. 10).

Finally to show that G^* is second countable we have, as a simple Corollary of Theorem 57 in Pontryagin [22], p. 276:

Theorem 4.3: If G is second countable, so is G^*.

Example 4.1: If $G = R$, $G^* = R$ and $< x, y > = e^{ixy}$, $x, y \in R$

Example 4.2: If $G = T^1$, $G^* = Z$ and $< x, n > = e^{inx}$, $x \in T^1$, $n \in Z$.

Example 4.3: If $G = Z$, $G^* = T^1$ and $< n, x > = e^{inx}$, $n \in Z$, $n \in T^1$.

For proofs of the above statements see Rudin [24], p. 12.

Theorem 4.4 (Pontryagin's Theorem): Let G be a group, G^* its dual and G^{**} the dual of G^*. Define $\varphi : G \to G^{**}$ by:

$$< \gamma, \varphi(g) > = < g, \gamma > \text{ for all } \gamma \in G^*. \tag{4.4}$$

Then φ is an isomorphism from G onto G^{**}.[1] (Pontryagin [22], p. 273).

In view of Pontryagin's Theorem, we shall identify G and G^{**} and write $G^{**} = G$.

Theorem 4.5: The dual of a discrete group is compact and the dual of a compact group is discrete. (Pontryagin [22], p. 237).

Definition 4.2: Let G be a group, G^* its dual and H a subset of G. The annihilator $H^\perp$ of H is defined by

$$H^\perp = \{\gamma \in G^* : < g, \gamma > = 1 \text{ for all } g \in H\} \tag{4.5}$$

Clearly $H^\perp$ is a closed subgroup of G^*.

Theorem 4.6: Let H be a closed subgroup of G. Then $(G/H)^*$ is isomorphic to $H^\perp$. (Pontryagin [22], p. 243).

---

[1] The reader who finds the use of the symbol $< , >$ in (4.4) confusing, should read this formula as follows: $[\varphi(g)](\gamma) = \gamma(g)$ for all $\gamma \in G^*$.

Theorem 4.7: Let H be a closed subgroup of G. Then H^* is isomorphic to $G^*/H^\perp$. (Pontryagin [22], p. 274).

Theorem 4.8: Let $G = G_1 \oplus G_2$. Then $G^* = G_1^* \oplus G_2^*$. (Rudin [24], p. 36).

Example 4.4: $(R^n)^* = R^n$

Example 4.5: T^n and Z^n are duals of each other

Example 4.6: For all positive integers n and m $(Z_m^n)^* = Z_m^n$.

Theorem 4.9: Let H be a closed subgroup of G, and let $g \in G \setminus H$ and let $\zeta \in H^*$. Then there exists a character $\gamma \in G^*$ such that

(a) $< g, \gamma > \neq 0$

and

(b) $< h, \gamma > = < h, \zeta >$ for all $h \in H$.

(Pontryagin, [22], p. 275).

The Adjoint of a Homomorphism

The following definition is analogous to the well known concept of the adjoint of a linear operator on a reflexive Banach space (Dunford and Schwartz [4], p. 478).

Definition 4.3: Let $T : G \to H$ be a homomorphism. The adjoint T^* of T is the function from H^* to G^* given by

$$< g, T^*(\zeta) > = < T(g), \zeta >, \quad g \in G, \quad \zeta \in H^*. \tag{4.6}$$

We shall occasionally write (4.3) in the form:

$$T^*(\zeta) = \zeta T, \quad \zeta \in H^*. \tag{4.7}$$

We state in Theorem 4.10 below several easily established properties of the adjoint of a homomorphism. For proofs see Pontryagin [22], pp. 242 ff. Theorem 4.11 is a simple consequence of Theorem 4.9; we furnish a proof for the purpose of completeness and because we could not find a proof in the literature.

Theorem 4.10: Let $T, T_1, T_2 : G \to H$ and $S : H \to K$ be homomorphisms. Then:

(1) $T^* : H^* \to G^*$ is a homomorphism.

(2) T^* is continuous and if T is open so is T^*.

(3) $(T_1 + T_2)^* = T_1^* + T_2^*$.

(4) $(ST)^* = T^*S^*$.

(5) If I_G is the identity map on G, then $(I_G)^* = I_{G^*}$ is the identity map on G^*.

(6) If T is invertible, so is T^* and $(T^{-1})^* = (T^*)^{-1}$.

(7) $T^{**} = T$; more precisely, if for every $g \in G$, g^{**} denotes the image of g under the identification map (Theorem 4.4) of G and G^{**} then $< \zeta, T^{**}(g^{**}) > = < T(g), \zeta >$ for all $\zeta \in H^*$.

Theorem 4.11: Let $T : G \to H$ be an open and injective homomorphism. Then for all $\gamma \in G^*$ there exists $\zeta \in H^*$ such that $\gamma = T^*(\zeta)$.

Proof: Let $\gamma \in G^*$ and let $K = T(G)$. Since T is open, K is locally compact (Proposition 2.10) and hence closed (Proposition 2.11). Since $T : G \to K$ is invertible, we may let T_0^* be the restriction of T^* to K^* and use (6) of Theorem 4.11 to define:

$$\xi = (T^*)^{-1}(\gamma) \in K^*. \tag{4.8}$$

But then we may extend ξ to a character ζ on all of H; (Theorem 4.9).

Now consider the function $\zeta T : G \to C$. Since $T = T^*(\zeta)$ (by 4.4) we have for all $g \in G$;

$$\begin{aligned} < g, T^*(\zeta) > &= < T(g), \zeta > \\ &= < T(g), \xi > \text{ (since } T(g) \in K) \\ &= < g, T_0^*(\xi) > \\ &= < g, \gamma > \text{ (by 4.7).} \end{aligned}$$

Thus $\gamma = T^*(\zeta)$ and the proposition is proved. //.

CHAPTER V

RANDOM VARIABLES WITH VALUES IN A GROUP

Basic Notions

In classical Probability Theory, a random variable is defined as a measurable function from a probability space to R^n for some positive integer n. In this chapter we generalize this concept by replacing R^n by a locally compact, second countable Hausdorff Abelian group. Again, we make no claim for completeness in this presentation we only develop sufficient machinery to prove certain characterization theorems in this and the next two chapters. Actually the main thrust of most books and papers on this subject is the derivation of limit theorems analogous to the limit theorems of classical probability theory, and matters treated here are mostly ignored in the literature. For the more standard development of the subject of group-valued random variables see Grenander [6], Parthasarathy [21], and Sazonov and Tutubalin [27].

We assume throughout that $(\Omega, \mathfrak{S}, P)$ is a probability space, $(G, +, \mathfrak{J})$ a group,[1] $\mathfrak{B} = \text{gen}\,\mathfrak{J}$, and μ a Haar measure on $(G, \mathfrak{B})$

---

[1] We recall that we use the work "group" to denote a locally compact, second countable, Hausdorff, Abelian, topological group. For convenience, we display explicitly the group operation "+" and the topology "$\mathfrak{J}$".

which is normalized if G is compact. If more than one group is considered, we will use appropriate subscripts; the underlying probability space $(\Omega, \mathfrak{S}, P)$ remains fixed throughout.

Definition 5.1: A random variable X is a measurable function $X : \Omega \to G$, i.e., a function X which satisfies the condition

$$X^{-1}(B) \in \mathfrak{S} \text{ for all } B \in \mathfrak{S} \tag{5.1}$$

Definition 5.2: The distribution of X is the measure μ_X on $(G, \mathfrak{B})$ defined by

$$\begin{aligned} \mu_X(B) &= P\{\omega \in \Omega: X(\omega) \in B\} \\ &= PX^{-1}(B) \text{ for all } B \in \mathfrak{B}. \end{aligned} \tag{5.2}$$

It is clear that if X is a random variable then $(G, \mathfrak{B}, \mu_X)$ is a probability space.

Definition 5.3: If X and Y are random variables with values in G, we shall write $X \sim Y$ if $\mu_X = \mu_Y$, i.e., if $P\{\omega: X(\omega) \neq Y(\omega)\} = 0$.

Definition 5.4: The random variable X is said to be uniformly distributed on G if $\mu_X = \mu$, (the normalized Haar measure on G).

It follows from Theorem 3.3 that a necessary (but not sufficient) condition for X to be uniformly distributed on G is that G be compact. Furthermore, since $\mu(g+B) = \mu(B)$ for all $g \in B$ and $B \in \mathfrak{B}$, the uniqueness of the normalized Haar measure (Theorem 3.1) implies the following simple Proposition:

Proposition 5.1: X is uniformly distributed on G if and only if $X \sim g + X$ for all $g \in G$.

Since neither R nor Z are compact none of the random variables studied in elementary probability are uniformly distributed in the sense of Definition 5.4. We shall see later how the uniform continuous distribution on $[0, 1]$ may be considered to be a uniform distribution on $(T^1, +)$.

Definition 5.5: Let $g \in G$ and let the random variable X have the distribution given by

$$\mu_X(B) = \begin{cases} 0 & \text{if } g \notin B \\ 1 & \text{if } g \in B \end{cases} \quad \text{for all } B \in \mathfrak{B}. \tag{5.3}$$

Then X is called degenerate at g and denoted by g. Its distribution is called a degenerate distribution and denoted by δ_g.

Definition 5.6: If $X_1, \ldots, X_n$ are random variables with values in $G_1, \ldots, G_n$ respectively, their joint distribution is a measure $\mu_{X_1,\ldots,X_n}$ on $(\Pi\{G_k : k = 1, \ldots, n\}, \Pi\{\mathfrak{B}_k : k = 1, \ldots, n\})$ generated by the set function:

$$\mu_{X_1,\ldots,X_n}(B_1 \times \ldots \times B_n) = P\{\omega : X_k(\omega) \in B_k,\ k = 1, \ldots, n\} \tag{5.4}$$

where $B_k \in \mathfrak{B}_k$.

Definition 5.7: The random variables $X_1, \ldots, X_n$ are said to be independent if their joint distribution is the product (see 3.22) of their distributions; i.e., if

$$\mu_{X_1,\ldots,X_n} = \mu_{X_1} \times \mu_{X_2} \times \ldots \times \mu_{X_n} \tag{5.5}$$

Convolutions

As a motivation for the concepts to be discussed now, we translate Theorem 3.8 into probabilistic language.

Theorem 5.1: Let G_1, G_2 be groups, X a random variable with values in G_1 and let $T : G_1 \to G_2$ be a measurable function. Let $Z = T(X)$. Then Z is a random variable with values in G_2 and with distribution given by:

$$\mu_Z(B_2) = \mu_X(T^{-1}(B_2)), \quad B_2 \in \mathfrak{B}_2 \tag{5.6}$$

Moreover if f is a measurable real or complex valued function on G_2, then

$$\int_{G_2} f(z)d\mu_Z(z) = \int_{G_1} f(T(x))d\mu_X(x) \tag{5.7}$$

We can generalize Theorem 5.1 to the case where T is a function of two variables.

Theorem 5.2: Let G_1, G_2 be groups, X, Y random variables with values in G_1 and let $T : G_1 \times G_1 \to G_2$ be a measurable function. Let $Z : \Omega \to G_2$ be the mapping defined by

$$Z(\omega) = T(X(\omega), Y(\omega)), \quad \omega \in \Omega. \tag{5.8}$$

Then Z is a random variable with values in G_2. (Rudin [25], p. 11). The distribution of Z is given by

$$\mu_Z(B_2) = \mu_{(X, Y)}(T^{-1}(B_2)), \qquad B_2 \in \mathfrak{B}_2. \tag{5.9}$$

If X and Y are independent, then

$$\mu_Z(B_2) = (\mu_X \times \mu_Y)(T^{-1}(B_2)), \qquad B_2 \in \mathfrak{B}_2. \tag{5.10}$$

A generalization of Theorem 5.2 to functions T of any finite number of variables is immediate. We shall, however, be interested in the following special case of that theorem:

Let $G_1 = G_2 = G$, let X and Y be independent and let $T(x, y) = x + y$, $x, y \in G$. In this case it is natural to write $Z = X + Y$. We will compute μ_Z.

Let $B \in \mathfrak{B}$. We recall that since G is locally compact and second countable the σ-algebra $\mathfrak{B} \times \mathfrak{B}$ is precisely the family of Borel subsets of $G \times G$. Now:

$$\mu_Z(B) = P\{\omega: X(\omega) + Y(\omega) \in B\}, \qquad B \in \mathfrak{B}. \tag{5.11}$$

We now consider the set

$$B^{(2)} = \{(x, y) \in G \times G: x + y \in B\} \tag{5.12}$$

and note that

$$T^{-1}(B^{(2)}) = B. \tag{5.13}$$

Hence by (5.9) and (5.11)

$$\mu_Z(B) = \mu_{(X, Y)}(B^{(2)}), \tag{5.14}$$

and since X and Y are independent, (5.14) becomes

$$\mu_Z(B) = (\mu_X \times \mu_Y)(B^{(2)}). \tag{5.15}$$

Formula (5.15) enables us to calculate μ_Z, given μ_X and μ_Y. We are interested in other expressions for μ_Z and to simplify our further considerations we introduce the following definition:

Definition 5.8: Let μ_X and μ_Y be the distributions of the G-valued random variables X and Y respectively. For all $B \in \mathfrak{B}$, let $B^{(2)}$ be defined by (5.12). The convolution of μ_X and μ_Y is defined by:

$$\mu_X * \mu_Y(B) = (\mu_X \times \mu_Y)(B^{(2)}), \quad B \in \mathfrak{B}. \tag{5.16}$$

We see immediately that if X and Y are independent, then $\mu_X * \mu_Y$ is the distribution of $X+Y$. The following theorem is also immediate:

Theorem 5.3: $\mu_X * \mu_Y$ is a probability distribution. $\mu_X * \mu_Y = \mu_Y * \mu_X$ and $(\mu_X * \mu_Y) * \mu_Z = \mu_X * (\mu_Y * \mu_Z)$.

The following theorem summarizes the basic properties of the convolution.

Theorem 5.4: For every $B \in \mathfrak{B}$ let $\mathcal{J}_B$ be the indicator function of G; i.e.,

$$\mathcal{J}_B(x) = \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{if } x \notin B \end{cases} \tag{5.17}$$

Then:

$$(\mu_X * \mu_Y)(B) = \int_G \mathcal{J}_B d(\mu_X * \mu_Y) = \int_B d(\mu_X * \mu_Y), \quad B \in \mathfrak{B}; \tag{5.18}$$

$$(\mu_X * \mu_Y)(B) = \iint_{GG} \mathcal{J}_B(x+y)\, d\mu_X\, d\mu_Y ; \quad B \in \mathfrak{B}, \tag{5.19}$$

$$(\mu_X * \mu_Y)(B) = \int_G \mu_X(B - y)\, d\mu_Y(y) ; \quad B \in \mathfrak{B}. \tag{5.20}$$

For proofs see Rudin [24].

Characteristic Functions

In classical Probability Theory the characteristic function $\varphi(t)$ of the real random variable X is defined by:

$$\varphi(t) = E\, e^{itx}, \quad t \in R. \tag{5.21}$$

We define, by analogy, the characteristic function of a random variable X with values in G to be the Fourier-Stieltjes transform of μ_X.

Definition 5.9: The characteristic function of the random variable X is the complex valued function μ_X defined on G^* by:

$$\hat{\mu}_X(\gamma) = \int_G < g, \gamma > d\mu_X(g) \quad \gamma \in G^*. \tag{5.22}$$

The following theorem summarizes the basic properties of μ_X which will be needed in the sequel. For proofs see Rudin [24], Chapter I.

Theorem 5.5: Let X and Y be independent random variables with values in G, μ_X and μ_Y their distributions and $\hat{\mu}_X$ and $\hat{\mu}_Y$ their characteristic functions. Then:

(1) $\hat{\mu}_X$ is uniformly continuous on G^*.

(2) $\hat{\mu}_X$ determines μ_X uniquely.

(3) $\hat{\mu}_X(0) = 1$ where 0 is the identity of G^*

($< g, 0 > = 1$ for all $g \in G$).

(4) $|\mu_X(\gamma)| \leq 1$ for all $\gamma \in G^*$.

(5) $\widehat{\mu_X * \mu_Y}(\gamma) = \hat{\mu}_X(\gamma)\hat{\mu}_Y(\gamma)$ for all $\gamma \in G^*$.

(6) If $Y = g + X$, $\hat{\mu}_Y(\gamma) = < g, \gamma > \hat{\mu}_X(\gamma)$

and conversely,

(7) $< g, \gamma > \hat{\mu}_X(\gamma)$ is the characteristic function of $g + X$.

Example 5.1: Let X be degenerate at $g_0 \in G$. Then $\hat{\mu}_X(\gamma) = < g_0, \gamma >$ for all $\gamma \in G^*$.

For a description of the characteristic function of the uniform distribution, see Theorem 5.6 in the next section.

Uniform Distributions

It is well known that if X is a real random variable then its characteristic function $\varphi_X(t) = 1$ for all $t \in R^1$ if and only if X is degenerate at 0. For a random variable with values in a group G we have a much richer theory. Theorem 5.6 is our basic tool as we develop that theory.

Theorem 5.6: The G-valued random variable X is uniformly distributed on a closed subgroup H of G if and only if $\hat{\mu}_X = \mathcal{J}_{H^\perp}$, i.e., if

$$\hat{\mu}_X(\gamma) = \begin{cases} 1 & \text{if } \gamma \in H \\ 0 & \text{if } \gamma \notin H \end{cases} \tag{5.23}$$

Proof: By (2) of Theorem 5.5 it suffices to prove necessity.

Let X be uniformly distributed on the compact subgroup H of G.

If $\gamma \in H^{\perp}$ then,

$$\hat{\mu}_X(\gamma) = \int_G < g, \gamma > d\mu_X(g) = \int_H < g, \gamma > d\mu_X(g)$$

$$= \int_H d\mu_X(g)$$

$$= 1$$

If $\gamma \notin H^{\perp}$, then there exists $g_0 \in H$ such that $< \gamma, g_0 > \neq 1$. Hence

$$\hat{\mu}_X(\gamma) = \int_G < g, \gamma > d\mu_X(g) = \int_H < g, \gamma > d\mu_X(g)$$

$$= \int_G < g - g_0 + g_0, \gamma > d\mu_X(g)$$

$$= \int_G < g - g_0, \gamma > < g_0, \gamma > d\mu_X(g)$$

$$= < g_0, \gamma > \int_G < g - g_0, \gamma > d\mu_X(g)$$

$$= < g_0, \gamma > \int_G < g, \gamma > d\mu_X(g)$$

$$= \langle g_0, \gamma \rangle \mu_X(\gamma).$$

Since

$$\langle g_0, \gamma \rangle \neq 1, \ \hat{\mu}_X(\gamma) = 0. \qquad //.$$

Corollary: X is uniformly distributed on G if and only if

$$\hat{\mu}_X(\gamma) = \begin{cases} 1 & \text{if } \gamma = 0 \\ 0 & \text{if } \gamma \neq 0 \end{cases} \tag{5.24}$$

Proof: $G^{\perp} = 0$. //.

We note that the apparently surprising result of this corollary does not contradict the fact ((1) of Theorem 5.5) that $\hat{\mu}_X$ is uniformly continuous. We recall that if X is uniformly distributed on G then G is compact (Theorem 3.3) and that the dual of a compact group is discrete. (Theorem 4.5).

The following theorem describes an interesting property of the uniform distribution.

Theorem 5.7: Let G be compact, let X and Y be independent random variables with values in G and let X be uniformly distributed on G. Then $X+Y$ is uniformly distributed on G.

Proof: For all $\gamma \in G^*$

$$\hat{\mu}_{X+Y}(\gamma) = \hat{\mu}_X(\gamma)\hat{\mu}_Y(\gamma) = \begin{cases} 0 & \text{if } \gamma \neq 0 \\ 1 & \text{if } \gamma = 0. \end{cases} \qquad //.$$

It is very important to be careful when applying Theorem 5.7 to real random variables. Since the only compact subgroup of R^1 is $\{0\}$

the only uniform distribution on R^1 is the distribution which is degenerate at 0. We recall the following well-known example:

Example 5.2: If X_1 and X_2 are independent real random variables, both uniformly distributed on $[0, 1]$ then $X_1 + X_2$ has the probability density function f given by

$$f(x) = \begin{cases} x & \text{if } x \in [0, 1] \\ -x+2 & \text{if } x \in [1, 2] \\ 0 & \text{if } x \notin [0, 2]. \end{cases}$$

This apparent contradiction to Theorem 5.7 arises from the fact that $[0, 1]$ is not a subgroup of R^1. We also remark that for the random variable X_1 of Example 5.2 $\text{supp } X_1 = [0, 1] \neq c(X_1) = R^1$.

The converse of Theorem 5.7 is false as the following examples illustrate.

Example 5.3: Let $G = T^2$, $H_1 = \{(g, 0) \in T^2\}$ and $H_2 = \{(0, g) \in T^2\}$. Let X_1 and X_2 be independent and uniformly distributed on H_1 and H_2 respectively. Since $G^* = Z^2$, we may represent an element $\gamma \in G^*$ by the ordered pair (m, n) of integers. Writing $z = (x, y)$ for $z \in G$, we note that

$$< z, \gamma > = e^{i(mx + ny)}.$$

Hence

$$\hat{\mu}_{X_1}(m, n) = \begin{cases} 0 & \text{if } n \neq 0 \\ 1 & \text{if } n = 0 \end{cases}$$

and

$$\hat{\mu}_{X_2}(m,n) = \begin{cases} 0 & \text{if } m \neq 0 \\ 1 & \text{if } m = 0 \end{cases}$$

and hence

$$\hat{\mu}_{X_1 + X_2}(m,n) = \begin{cases} 0 & \text{if } m \neq 0 \text{ or } n \neq 0 \\ 1 & \text{if } (m,n) = (0,0) \end{cases}$$

Thus $X_1 + X_2$ is uniformly distributed on G while neither X_1 nor X_2 is uniformly distributed on G.

In the following example we compute $X+Y$ without recourse to characteristic function.

Example 5.4: Let G be the Klein 4 group. We denote the elements of G by $\{0,a,b,c\}$, with 0 the identity element of G, and recall that $x+x = 0$ for all $x \in G$ and that if x,y,z are distinct non-zero elements of G, then $x+y = z$. (See Herstein [8] for more details).

Let X meet the values	0	a	b	c
with probabilities	$\frac{1}{2}$	$\frac{1}{2}$	0	0
and let Y meet the values	0	a	b	c
with probabilities	$\frac{1}{2}$	0	$\frac{1}{2}$	0

Then

$$P(X+Y=0) = P(X=0, Y=0) + P(X=a, Y=a) + P(X=b, Y=b) + P(X=c, Y=c)$$

$$P(X+Y=a) = P(X=0, Y=a) + P(X=a, Y=0) + P(X=b, Y=c) + P(X=c, Y=b)$$

with similar expressions for $P(X+Y=b)$ and $P(X+Y=c)$.

A simple substitution then yields the result that the random variable $X+Y$ meets the values $0 \quad a \quad b \quad c$ with probabilities $\frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4}$

Thus $X+Y$ is uniformly distributed on G while neither X nor Y is.

Example 5.4 is really a rewording of Example 5.3, for the Klein 4 group is really Z_2^2, and writing $H_1 = \{0,a\}$, $H_2 = \{0,b\}$ we note that X and Y are uniformly distributed on H_1 and H_2 respectively and that $G = H_1 \oplus H_2$; (Herstein, [8]). However, the Klein 4 group illustrates the necessity in Theorem 5.7 of the condition that X and Y be independent. An easy computation shows that no matter what the distribution of X on Z_2^2, $X+X$ is degenerate at 0.

The following theorem is an attempt at a partial converse to Theorem 5.7. We do not know whether the condition in Theorem 5.8 that $\hat{\mu}_Y$ do not vanish could be weakened.

Theorem 5.8: Let X and Y be independent random variables with values in G. If $\hat{\mu}_Y$ does not vanish and $X+Y$ is uniformly distributed on G, then so is X.

Proof: For all $\gamma \in G^*$

$$\hat{\mu}_{X+Y}(\gamma) = \hat{\mu}_X(\gamma)\hat{\mu}_Y(\gamma) = \begin{cases} 0 & \text{if } \gamma \neq 0 \\ 1 & \text{if } \gamma = 0. \end{cases}$$

If $\gamma = 0$, this equation is trivially satisfied. If $\gamma \neq 0$ then $\hat{\mu}_X(\gamma) = 0$ since $\hat{\mu}_Y(\gamma) \neq 0$. Hence X is uniformly distributed on G.

//.

We note that Theorem 5.7 could be restated as follows:

Theorem 5.7': Let G be compact, let X and Y be independent random variables with values in G and let X be uniformly distributed on G. Then $X+Y \sim X$.

Theorem 5.9 below is a useful statement of a partial converse of Theorem 5.7'.

Theorem 5.9: Let X and Y be independent G-valued random variables with $c(X) = G$. Let Y satisfy the condition:

$$\text{For all } \gamma \in G^*,\ \hat{\mu}_Y(\gamma) = 1 \Rightarrow \gamma = 0. \qquad (*)$$

Then if $X+Y \sim X$, X is uniformly distributed on G.

Proof: For all $\gamma \in G^*$

$$\hat{\mu}_{X+Y}(\gamma) = \hat{\mu}_X(\gamma)\hat{\mu}_Y(\gamma) = \hat{\mu}_X(\gamma). \qquad (5.25)$$

If $\gamma = 0$, (5.25) is trivially satisfied.

If $\gamma \neq 0$, $\hat{\mu}_Y(\gamma) \neq 1$ and hence (5.25) implies that $\hat{\mu}_X(\gamma) = 0$. Hence

$$\hat{\mu}_X(\gamma) = \begin{cases} 0 & \text{if } \gamma \neq 0 \\ 1 & \text{if } \gamma = 0 \end{cases}$$

and the theorem is proved. //.

We do not know if the condition (*) of Theorem 5.9 can be weakened. That it cannot be dispensed with altogether, is illustrated by the following example.

Example 5.5: Let G be the Klein 4 group, and adopt the notation of Example 5.4.

Let X meet the values	0	a	b	c
with probabilities	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
and let Y meet the values	0	a	b	c
with probabilities	$\frac{2}{3}$	0	$\frac{1}{3}$	0

A simple calculation now shows that $X+Y$ meets the values

	0	a	b	c
with probabilities	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{3}{8}$

Hence $X+Y \sim X$ even though X is not uniformly distributed on G. Recalling that $G^* = G$, we note that $\hat{\mu}_Y(0) = \hat{\mu}_Y(a) = 1$ while $\hat{\mu}_Y(b) = \hat{\mu}_Y(c) = \frac{1}{3}$.

It turns out, however, that if $X+Y \sim X$ then Y and $X+Y$ are independent even if (*) of Theorem 5.9 is not satisfied. This fact is proved in the following theorem.

Theorem 5.10: Let X and Y be independent random variables with values in G, and let $X \sim X+Y$. Then Y and $X+Y$ are independent.

Proof: For all $B_1, B_2 \in \mathfrak{B}$, we have:

$$\begin{aligned}
\mu_{(Y, X+Y)}(B_1 \times B_2) &= P\{\omega : Y(\omega) \in B_1,\ (X+Y)(\omega) \in B_2\} \\
&= P\{\omega : Y(\omega) \in B_1,\ X(\omega) \in B_2\} \\
&= \mu_{(Y,X)}(B_1 \times B_2) \\
&= \mu_Y(B_1)\,\mu_X(B_2) \\
&= \mu_Y(B_1)\,\mu_{X+Y}(B_2) .
\end{aligned}$$

Since the sets $B_1 \times B_2$, $B_1, B_2 \in \mathfrak{B}$, generate the Borel sets of $G \times G$, the theorem is proved. //.

To give a final characterization of the uniform distribution on a group we need the following definition:

Definition 5.10: A probability distribution π is called idempotent, if $\pi * \pi = \pi$.

The following theorem shows the connection between idempotent and uniform distributions. For a proof see Grenander [6], p. 106. For further details on the structure of idempotent measures on groups see Rudin [24], Chapter III.

Theorem 5.11: The G-valued random variable X has an idempotent distribution μ_X if and only if X is uniformly distributed on a compact subgroup H of G.

It follows from Theorem 5.10 that the only idempotent measure on R^1 is δ_0. The indempotent measures on T^1 are Haar measure and counting measures on cyclic groups generated by rational numbers.

Applications to Real Random Variables

In this section we distinguish between the distribution of a real random variable as given in Definition 5.2 and its distribution function as defined in Ash [2], p. 52. We recall that the distribution function of a real random variable X is defined by

$$F_X(x) = P\{\omega : X(\omega) < x\}, \quad x \in R^1, \tag{5.26}$$

and that its characteristic function is given by

$$\varphi_X(t) = E\, e^{itX}, \qquad t \in R^1. \tag{5.27}$$

We wish to establish a connection between these ideas and the concepts introduced in the previous section. We assume throughout that a and b are fixed (extended) real numbers satisfying $-\infty \leq a < b \leq \infty$.

Definition 5.11: The class $\mathfrak{F}(a, b)$ is the set of all real-valued function F on $[a, b)$ to $[0, 1)$ satisfying:

(1) $F(a) = 0$

(2) $F(b - 0) = 1$

(3) F is continuous

(4) F is strictly increasing.

It follows that if $F \in \mathfrak{F}(a, b)$ then F is the distribution function of a random variable X. We shall write $X \sim F_X$ to indicate that F_X is the distribution function of X. We also note that all functions $F \in \mathfrak{F}(a, b)$ satisfy the conditions of Theorem 2.1 and that we may thus consider the group $([a, b), F)$, (see Definition 2.5).

Now let $X \sim F_X \in \mathfrak{F}(a, b)$ and consider X to be a random variable with values in $([a, b), F_X)$. We wish to determine μ_X. For this purpose it suffices to consider half open intervals $[c, d)$ with $a \leq c < d \leq b$. We note that

$$\mu_X([c, d)) = P\{\omega: c \leq X(\omega) < d\} = F_X(d) - F_X(c) \tag{5.28}$$

Next we recall that F_X is an isomorphism between $([a, b), F_X)$ and T^1, and that the normalized Haar measure on T^1 is Lebesgue measure m. Since $\mu_X([c, d)) = m(F_X([c, d)))$ it follows that μ_X is

translation invariant on $([a, b), F_X)$, and is thus the normalized Haar measure on $([a, b), F_X)$. We have thus proved the following theorem:

Theorem 5.12: Let $X \sim F_X \in \mathfrak{F}(a, b)$. Then X may be considered to be a uniformly distributed random variable on the group $([a, b), F_X)$.

The converse of Theorem 5.12 follows immediately from the uniqueness of the Haar measure. We state it in the next theorem.

Theorem 5.13: Let $\Psi \in \mathfrak{F}(a, b)$ and let $X \sim F_X$. If X is uniformly distributed on $([a, b), \Psi)$, then $F_X = \Psi$.

Theorems 5.12 and 5.13 are well known. If $X \sim F_X$ it is uniformly distributed on $([a, b), F_X)$, it follows that $F_X(X)$ is uniformly distributed on T^1. Theorems 5.12 and 5.13 can be found in the classical literature stated as follows: (see Lindgren [17]p. 274 ff.).

Theorem 5.14: The real random variable X has the distribution function F if and only if $F(X)$ is the uniform continuous distribution on $[0, 1]$.

Formula (5.28) exhibits the relationship between the distribution function of a real random variable X and its distribution when X is considered to be group-valued on $([a, b), F_X)$. We now derive a similar relationship between its characteristic functions as given by (5.27) and (5.22).

Theorem 5.15: Let $\psi \in \mathfrak{F}(a,b)$ and let $X \sim F_X \in \mathfrak{F}(a,b)$. Let $(G,\circ) = ([a,b), \psi)$. Then X considered as a G-valued random variable has a distribution μ_X defined by (5.28). Furthermore, every element in G^* is of the form $\psi^*(n)$, $n \in Z$. Finally, for all $n \in Z$,

$$\varphi_X(\psi^*(n)) = \hat{\mu}_X(\psi^*(n)). \tag{5.29}$$

Proof: The fact that μ_X is given by (5.28) has already been shown.

Since ψ is an isomorphism from G to T^1, ψ^* is an isomorphism from T^{1*} to G^*; (Theorem 4.10). Since $T^{1*} = Z$, every element of G^* is of the form $\psi^*(n)$, $n \in Z$.

Now let $Y = \psi(X)$. Then Y is a T^1-valued random variable and for all Borel subsets B of G, $\mu_X(B) = \mu_Y(\psi(B))$. Hence

$$\mu_Y = \mu_X \psi^{-1} \tag{5.30}$$

Furthermore, for all $n \in Z$

$$\hat{\mu}_Y(n) = \int_{T^1} < y, n > d\mu_Y(y)$$

$$= \int_0^1 < y, n > d\mu_Y(y) \quad \text{(Pontryagin [22], p. 201)}$$

$$= \int_0^1 < y, n > d(\mu_X \psi^{-1})(y) \quad \text{(by (5.30))}$$

$$= \int_a^b < \psi(x), n > d\mu_X(x) \quad \text{(Theorem 3.8)}$$

$$= \int_a^b < x, \Psi^*(n) > d\mu_X(X)$$

$$= \hat{\mu}_X(\Psi^*(n)) .$$

Thus:

$$\hat{\mu}_Y(n) = \hat{\mu}_X(\Psi^*(n)) \tag{5.31}$$

Repeating the above steps, we note that for all $n \in Z$,

$$\hat{\mu}_Y(n) = \int_a^b < x, \Psi^*(n) > d\mu_X(x)$$

$$= \int_a^b e^{i\Psi^*(n)x} \, dF_X(x) \quad \text{(by (5.28) \& Ex. 4.2)}$$

$$= \varphi_X(\Psi^*(n)) \quad \text{(by 5.27)}$$

Thus:

$$\hat{\mu}_Y(n) = \varphi_X(\Psi^*(n)) \tag{5.32}$$

(5.31) and (5.32) now yield (5.29) and the theorem is proved. //.

We note that the well-known fact that the distribution of a random variable on $[-\pi, \pi]$ is completely characterized by its Fourier coefficients is an immediate corollary of Theorem 5.14.

We are now able to prove the following characterization theorems, which are unpublished results of Professor Kotlarski's.

Theorem 5.16: Let $\Psi \in \mathfrak{F}(a, b)$ and $(G, \circ) = ([a, b), \Psi)$. Let X and Y be independent and have distribution functions

F_X, $F_Y \in \mathfrak{D}(a,b)$. Then $X \circ Y \sim X$ if and only if $F_X = \Psi$.

Proof: If $F_X = \Psi$ then X is uniformly distributed on G. (Theorem 5.11), and hence $X \circ Y$ is uniformly distributed on G; (Theorem 5.7). Hence $X \circ Y \sim X$.

Conversely, assume that $X \circ Y \sim X$. Since $Y \sim F_Y \in \mathfrak{D}(a,b)$, $\varphi_Y(t) = 1$ implies that $t = 0$. (Lukacs [19], p. 25). Hence, by Theorem 5.14, $\mu_Y(\gamma) = 1$ implies that $\gamma = 0$. Hence, by Theorem 5.9, X is uniformly distributed on G, and so, by Theorem 5.12, $F_X = \Psi$. //.

Corollary: Let X and Y be independent random variables, with distribution functions F_X, $F_Y \sim \mathfrak{D}(a,b)$. Let $(G,\circ) = ([a,b), F_Y)$. If $X \circ Y \sim X$ then $X \sim Y$.

Proof: In Theorem 5.15, put $\Psi = F_Y$, and the rest is obvious. //.

Theorem 5.17: Let $\Psi \in \mathfrak{D}(a,b)$ and let $(G,\circ) = ([a,b), \Psi)$. Let X and Y be independent real random variables with support in $[a,b]$ and let $X \sim X \circ Y$. Then Y and $X \circ Y$ are independent.

Proof: Apply Theorem 5.10 to the G-valued random variables X and Y. //.

CHAPTER VI

A PROPERTY OF THE UNIFORM DISTRIBUTION ON COMPACT ABELIAN GROUPS

A Characterization Theorem

In Chapter V we have shown how a real random variable X with distribution function $F_X \in \mathfrak{F}(a, b)$ may be considered to be uniformly distributed on the group $([a, b), F_X)$. Theorems 5.7, 5.8 and 5.9 describe some of the properties of uniformly distributed group-valued random variables and Theorem 5.15 shows how these ideas could be applied to classical characterization problems. The basic theoretical result in Chapter V, namely Theorem 5.9, is somewhat unsatisfactory; however, since we do not know whether the condition that: $\mu_Y(\gamma) = 1$ implies $\gamma = 0$, could be weakened, nor do we know of a group theoretic analogue of the result, used in the proof of Theorem 5.15, which states that a characteristic function $\varphi_X(t)$ is the characteristic function of a lattice distribution if and only if there exists a real $t_0 \neq 0$ such that $|\varphi_X(t_0)| = 1$; (Lukacs [19], p. 25).

The purpose of this chapter is to prove a result similar to Theorem 5.9 under the assumption that Y is a degenerate random variable. We already know (Proposition 5.1) that the random variable X with values in the group G is uniformly distributed on G if and only

if $X \sim g+X$ for all $g \in G$. We show in Theorem 6.1 below that if $X \sim g+X$ for one $g \in G$ satisfying certain conditions then X is uniformly distributed on G. It turns out that classical characterization theorems such as the one described in Example 1.1 follow immediately from this result. We were also able to derive some new characterization theorems from Theorem 6.1.

We first need to prove a lemma.

Lemma 6.1: Let X be a random variable with values in G, and let H be a dense subgroup of G. If for all $h \in H$, $X \sim h+X$ then X is uniformly distributed on G.

Proof: Let $\mathfrak{B} = \text{gen}\,\mathfrak{J}$, where $\mathfrak{J}$ is the topology on G. For all $E, F \in \mathfrak{B}$ define:

$$\rho(E, F) = \mu_X(E \Delta F) \tag{6.1}$$

and recall (Theorem 3.12) that $(\mathfrak{B}', \rho')$ is a complete metric space, where $(\mathfrak{B}', \rho')$ is obtained from the pseudo-metric space $(\mathfrak{B}, \rho)$ by identifying sets whose symmetric difference is of μ_X-measure zero.

Now let $g_0 \in G$ and $B_0 \in \mathfrak{B}$ be fixed. We shall show that for all $\varepsilon > 0$,

$$|\mu_X(B_0) - \mu_X(g_0 + B_0)| < \varepsilon. \tag{6.2}$$

It will follow that μ_X is translation invariant, and hence that μ_X is the normalized Haar measure on G. This means that X is uniformly distributed on G.

By Theorem 3.13, the function $f: G \to \mathfrak{B}'$ given by

$$f(g) = (g + B_0)' \tag{6.3}$$

is continuous.

Let $\varepsilon > 0$. Since H is dense in G, we can find $h_0 \in H$ such that

$$\rho'(f(g_0), f(h_0)) < \varepsilon . \tag{6.4}$$

But

$$\begin{aligned} \rho'(f(g_0), f(h_0)) &= \rho'((g_0 + B_0)', (h_0 + B_0)') \\ &= \rho(g_0 + B_0, h_0 + B_0) \\ &= \mu_X((g_0 + B_0) \Delta (h_0 + B_0)) \end{aligned}$$

Hence:

$$\mu_X((g_0 + B_0) \Delta (h_0 + B_0)) < \varepsilon . \tag{6.5}$$

We intend to show that:

$$\begin{aligned} \mu_X[(g_0 + B_0) \cap (h_0 + B_0)] &\leq \mu_X(g_0 + B_0) \\ &\leq \mu_X[(g_0 + B_0) \cap (h_0 + B_0)] + \varepsilon . \end{aligned} \tag{6.6}$$

The first inequality is obvious since

$$(g_0 + B_0) \cap (h_0 + B_0) \subset g_0 + B_0$$

and the second follows from:

$$\mu_X(g_0 + B_0) = \mu_X[\{(g_0 + B_0) \cap (h_0 + B_0)\} \dot{\cup} \{(g_0 + B_0) \cap (h_0 + B_0)\}]$$

$$= \mu_X[(g_0 + B_0) \cap (h_0 + B_0)] + \mu_X[(g_0 + B_0) \cap (h_0 + B_0)]$$

$$\leq \mu_X[(g_0 + B_0) \cap (h_0 + B_0)] + \mu_X[(g_0 + B_0) \Delta (h_0 + B_0)]$$

$$\leq \mu_X[(g_0 + B_0) \cap (h_0 + B_0)] + \varepsilon . \qquad \text{(by 6.5)}$$

Similarly we can show that

$$\mu_X[(g_0 + B_0) \cap (h_0 + B_0)] \leq \mu_X(h_0 + B_0)$$

$$\leq \mu_X[(g_0 + B_0) \cap (h_0 + B_0)] + \varepsilon , \quad (6.7)$$

and combining (6.6) and (6.7) we obtain

$$|\mu_X(h_0 + B_0) - \mu_X(g_0 + B_0)| < \varepsilon . \qquad (6.9)$$

Now since $X \sim h_0 + X$, $X \sim (-h_0) + X$. Hence

$$P\{\omega : X(\omega) \in B_0\} = P\{\omega : (-h_0) + X(\omega) \in B_0\} = P\{\omega : X(\omega) \in h_0 + B_0\}$$

which means that

$$\mu_X(B_0) = \mu_X(h_0 + B_0) . \qquad (6.10)$$

Substituting (6.10) in (6.9) we obtain (6.2) and the lemma is proved.//.

We are now ready to state the main theorem of this chapter.

Theorem 6.1: Let X be a random variable with values in G and let g_0 generate a dense subgroup H of G. Let $X \sim g_0 + X$. Then X is uniformly distributed on G.

Proof: Since $X \sim g_0 + X$, $X \sim (-g_0) + X$, and hence a simple induction argument yields: $X \sim n g_0 + X$ for all $n \in Z$. Thus $X \sim h + X$ for all $h \in H$, and since H is dense in G the theorem is proved. //.

Corollary: Let $\alpha \in T^1$ be an irrational number. Then the random variable X with values in T^1 is uniformly distributed if and only if $X \sim \alpha + X$.

Proof: The cyclic subgroup of T^1 generated by an irrational number is dense in T^1. (Theorem 2.7). //.

The requirement, in Theorem 6.1, that g_0 generate a dense subgroup of G is essential. In Example 6.1 we exhibit a non-uniformly distributed random variable X on the group G, and an element $g_0 \in G$ such that $X \sim g_0 + X$.

Example 6.1: Consider Z_4. Let the random variable X meet

the values	0	1	2	3
with probabilities	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{3}{8}$

Then $X \sim 2 + X$ but X is not uniformly distributed on Z_4.

The proof of Theorem 6.1 is the one given in Flusser [5]. After that paper had been accepted for publication we discovered the following simpler proof.

Alternate Proof of Theorem 6.1: Let $H = \{n g_0 : n \in Z\}$. Since $X \sim g_0 + X$, $X \sim h + X$ for all $h \in H$. Let $g \in G$. Since H is dense in G, there exists a sequence $\{h_n\}$ of elements of H converging to G. Hence for all $\gamma \in G^*$ and $n = 1, 2, \ldots$

$$\hat{\mu}_X(\gamma) = \hat{\mu}_{h_n + X}(\gamma)$$
$$= < h_n, \gamma > \hat{\mu}_X(\gamma) \to < g, \gamma > \hat{\mu}_X(\gamma)$$

by the continuity of γ and $\hat{\mu}_X$.

Hence $\hat{\mu}_X(\gamma) = < g, \gamma > \hat{\mu}_X(\gamma) = \hat{\mu}_{g+X}(\gamma)$ for all $g \in G$ and $\gamma \in G^*$.

Thus $X \sim g + X$ for all $g \in G$ and the theorem follows from Proposition 5.1. //.

Applications to Real Random Variables

In this section we derive four characterization theorems from Theorem 6.1. The first is a result of Williams' [26], (see Example 1.1) the others are new but are similar in spirit to the first.

Theorem 6.2: (A Characterization of the Cauchy Distribution: Williams [26]).

Let X be a real random variable, and let c be a real number not the tangent of a rational multiple of π. Let

$$Y = \frac{X + c}{1 - cX} . \tag{6.11}$$

Then $X \sim Y$ if and only if X has the standard Cauchy distribution with probability density function (1.1).

Proof: Let

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \arctan x ; \quad x \in R. \tag{6.12}$$

Then $F \in \mathfrak{F}(-\infty, \infty)$; (see Definition 2.6). Let $(G, \circ) = ([-\infty, \infty), F)$.

$$\gamma = \frac{1}{\pi} \text{ arc tan } c . \tag{6.13}$$

Recalling that $0 \leq F(Y) < 1$ we obtain:

$$\begin{aligned} F(Y) &= \frac{1}{2} + \frac{1}{\pi} \text{ arc tan } \frac{X+c}{1-cX} \\ &= (\frac{1}{2} + \frac{1}{\pi} \text{ arc tan } X + \frac{1}{\pi} \text{ arc tan } c) \bmod [0, 1) \\ &= F(X) \dotplus \gamma . \end{aligned} \tag{6.14}$$

By our assumption on c and (6.13), γ is irrational, and hence by Theorem 6.1, $F(X)$ is uniformly distributed on T^1 if and only if $F(X) \sim F(Y)$, which is true if and only if X is distributed like Y on $(G, \circ)$. Furthermore, by Theorems 5.12 and 5.13, $F(X)$ is uniformly distributed on T^1 if and only if F is the distribution function of X, i.e., if and only if X has the standard Cauchy distribution with probability density function (1.1).

Theorem 6.3: [A Characterization of the Uniform Continuous Distribution on $[0, 1]$].

Let X be a real random variable, and for some irrational $c \in [0, 1]$ let

$$Y = (X+c) \bmod [0, 1) . \tag{6.15}$$

Then $X \sim Y$ if and only if X is uniformly distributed on $[0, 1]$.

Proof: Let $F(x) = x$, $x \in [0, 1)$. The theorem now follows from the Corollary to Theorem 6.1 and an argument similar to the one used in Theorem 6.2. //.

Theorem 6.4: [A Characterization of the Exponential Distribution].

Let X be a real random variable, and let c be a positive number which is not the logarithm of a rational number. Define

$$Y = \begin{cases} -\log(e^{-X} + e^{-c} - 1) & \text{if } X \leq -\log(1 - e^{-c}) \\ -\log(e^{-X} + e^{-c}) & \text{if } X > -\log(1 - e^{-c}) \end{cases} . \tag{6.16}$$

Then $X \sim Y$ if and only if $X \sim \text{Exp}(1)$; i.e., if and only if X has the probability density function

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ e^{-x} & \text{if } x > 0 \end{cases} . \tag{6.17}$$

Proof: Let

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 - e^{-x} & \text{if } x \geq 0 \end{cases} . \tag{6.18}$$

Then $F \in \mathfrak{F}(0, \infty)$. Let $(G, o) = ([0, \infty), F)$.

Since

$$F^{-1}(y) = -\log(y - 1) \tag{6.19}$$

we get

$$\begin{aligned} c \circ X &= F^{-1}[F(X) + F(c)] \\ &= F^{-1}[1 - e^{-X} + 1 - e^{-c}] \\ &= \begin{cases} -\log(e^{-X} + e^{-c} - 1) & \text{if } X \leq -\log(1 - e^{c}) \\ -\log(e^{-X} + e^{-c}) & \text{if } X > -\log(1 - e^{-c}) \end{cases} \end{aligned}$$

Hence: $c \circ X = Y$. (6.20)

To show that c satisfies the condition of Theorem 6.1 we need only verify that $F(c)$ is irrational. But $F(c) = 1 - e^{-c}$ and c is not the logarithm of a rational number. Hence c generates a dense subgroup of G and hence $X \sim Y$ if and only if X is uniformly distributed on G which is true if and only if F is the distribution of X. This proves the theorem. //.

Theorem 6.5: [A Characterization of the Hyperbolic Secant Distribution].

Let X be a real random variable, and let c be a real number, not the logarithm of the tangent of a rational multiple of π. For all non-zero x, let $L(x)$ be the odd function which is equal to $\log x$ for positive x. Define:

$$Y = L\left(\frac{e^X + e^c}{1 - e^c e^X}\right) . \tag{6.21}$$

Then $X \sim Y$ if and only if X is the absolutely continuous random variable with probability density function

$$f(x) = \frac{1}{\pi} \operatorname{sech} x, \quad x \in R. \tag{6.22}$$

Proof: Let

$$F(x) = \frac{2}{\pi} \arctan e^x, \quad x \in R. \tag{6.23}$$

Then $F \in \mathfrak{F}(-\infty, \infty)$. Let $(G, \circ) = (-\infty, \infty), F)$. We note that since $F(c)$ is irrational, c generates a dense subgroup of G. We show that

$$c \circ X = Y, \tag{6.24}$$

the required result then follows from an argument like that used in Theorem 6.4.

$$F(X) \dotplus F(c) = \frac{2}{\pi} [\arctan e^X + \arctan e^c] \bmod [0, 1)$$

$$= (\frac{2}{\pi} \arctan \frac{e^X + e^c}{1 - e^X e^c}) \bmod [0, 1)$$

$$= \frac{2}{\pi} [\arctan \frac{e^X + e^c}{1 - e^X e^c} + \frac{\pi}{2} \delta]$$

where

$$\delta = \begin{cases} 0 & \text{if } 1 - e^X e^c \geq 0, \text{ i.e., if and only if } X \leq -c \\ 1 & \text{if } 1 - e^X e^c < 0, \text{ i.e., if and only if } X > c \end{cases}$$

Hence

$$F(X) \dotplus F(c) = \begin{cases} \frac{2}{\pi} \arctan \frac{e^X + e^c}{1 - e^X e^c} & \text{if } X < -c \\ \frac{2}{\pi} \arctan \frac{e^X + e^c}{1 - e^X e^c} & \text{if } X \geq -c \end{cases}$$

noting that

$$F^{-1}(y) = \log \tan \frac{\pi}{2} y, \qquad y \in [0, 1)$$

we obtain

$$F^{-1}[F(X) \dotplus F(c)] = \begin{cases} \log \frac{e^X + e^c}{1 - e^X e^c} & \text{if } X < -c \\ \log \tan \left[\arctan \left(\frac{e^X + e^c}{1 - e^X e^c} + \frac{\pi}{2} \right) \right] & \text{if } X \geq -c \end{cases} \tag{6.25}$$

Recalling that

$$c \circ X = F^{-1}[F(X) \dotplus F(c)] \tag{6.26}$$

and that:

$$\log \tan\left[\arctan\left(\frac{e^X + e^c}{1 - e^X e^c} + \frac{\pi}{2}\right)\right] = \log \tan\left[\operatorname{arc\,cot}\left(\frac{e^X + e^c}{1 - e^X e^c}\right)\right]$$

$$= \log \frac{1 - e^X e^c}{e^X + e^c}, \tag{6.27}$$

we obtain from (6.25), (6.26) and (6.27)

$$c \circ X = \begin{cases} \log \dfrac{e^X + e^c}{1 - e^c e^X} & \text{if } X < -c \\ \log \dfrac{1 - e^X e^c}{e^X + e^c} & \text{if } X \geq -c \end{cases} \tag{6.28}$$

and hence (6.24) follows from the definition of the function L and (6.21). This concludes the proof. //.

CHAPTER VII

A CHARACTERIZATION OF THREE INDEPENDENT RANDOM VARIABLES

Introduction

While characterizing the normal, gamma and chisquare distributions, I. I. Kotlarski showed that the distributions of three independent random variables X_0, X_1 and X_2 are determined by the joint distribution of a two-dimensional random variable (Y_1, Y_2), where Y_1 and Y_2 depend on X_0, X_1 and X_2; Kotlarski [12], Kotlarski [13], Kotlarski [14]. In the special case where $Y_1 = X_0 + X_1$ and $Y_2 = X_0 + X_2$, and X_0, X_1, X_2, Y_1 and Y_2 are random variables with values in Hilbert space, Kotlarski [15] has shown that the distribution of (Y_1, Y_2) also determines the distributions of X_0, X_1 and X_2 up to a shift. Under the assumption that X_0, X_1, X_2, Y_1 and Y_2 have their values in a locally compact Abelian group, B. L. S. Prakasa Rao [23] obtained a similar result. More recently, Kotlarski [16] proved a similar characterization theorem under the assumption that X_0, X_1 and X_2 are random vectors of different dimension, T is a linear transformation satisfying certain conditions and $(Y_1, Y_2) = T(X_0, X_1, X_2)$.

In this chapter we prove a result similar to Kotlarski's in [16]. We assume that X_0, X_1 and X_2 are independent random variables with values in a group G, that T is a homomorphism from G onto a group H which is isomorphic to a proper subgroup of G and show that the joint distribution of $(Y_1, Y_2) = T(X_0, X_1, X_2)$ determines the distributions of X_0, X_1 and X_2 up to a shift which depends on only one parameter. We then show how the results of Kotlarski [16] and Prakasa Rao [23] follow immediately from these considerations, and finally, we give some applications to classical characterization problems.

The Main Theorem

Theorem 7.1: Let the locally compact, second countable, Hausdorff Abelian group G be the direct sum of three of its subgroups, G_0, G_1 and G_2:

$$G = G_0 \oplus G_1 \oplus G_2 . \tag{7.1}$$

For $k = 0, 1, 2,$ let p_k be the projection of G onto its k^{th} direct summand. Let X be a random variable with values in G and let

$$X_k = p_k X, \quad k = 0, 1, 2 . \tag{7.2}$$

Assume:

(A1) X_0, X_1 and X_2 are independent random variables (with values in G_0, G_1 and G_2 respectively).

(A2) The characteristic functions of X_0, X_1 and X_2 do not vanish.

Let H be another locally compact, second countable, Hausdorff Abelian group and let $T : G \to H$ be a (continuous) homomorphism from G onto H. Let

$$T_k = Tp_k, \quad k = 0, 1, 2 . \tag{7.3}$$

Assume further:

(A3) $T_0\big|_{G_0}$ is injective.

(A4) $(T_1 + T_2)\big|_{G_1 \oplus G_2}$ is bijective

(A5) $T(G_0) \cap T(G_1) = \{0\}$ and $T(G_0) \cap T(G_2) = \{0\}$, where 0 is the identity of H.

Let

$$Y = T(X) . \tag{7.4}$$

Then the distribution of Y determines the distributions of X_0, X_1 and X_2 up to a shift. The shift for X_0 is given by an arbitrary element $g_0 \in G$, those for X_1 and X_2 are determined by g_0.

Before proceeding with the proof of Theorem 7.1 we introduce some notation and prove two lemmas. Throughout this section we assume the hypothesis of Theorem 7.1. We also denote the distributions of X, X_0, X_1, X_2 and Y by μ, μ_0, μ_1, μ_2 and μ_Y respectively, with a corresponding notation for their characteristic functions. We denote the inverse of the restriction of $T_1 + T_2$ to $G_1 \oplus G_2$ by $(T_1 + T_2)^{-1}$; this mapping exists and is continuous by (A4) and Proposition 2.13.

We now define the following homomorphisms:

$$U = (T_1 + T_2)^{-1} T_0 : G \to G_1 \oplus G_2 , \tag{7.5}$$

$$U_0 = U\big|_{G_0} : G_0 \to G_1 \oplus G_2 , \tag{7.6}$$

$$U_1 = p_1 U_0 : G_0 \to G_1 , \tag{7.7}$$

$$U_2 = p_2 U_0 : G_0 \to G_2 . \tag{7.8}$$

Lemma 7.1: U, U_0, U_1 and U_2 are open and injective, and

$$U_0 = U_1 + U_2 . \tag{7.9}$$

Proof: The fact that U and U_0 are injective follows from (A5) and (A4). Equation (7.9) is obvious.

T is open by Proposition 2.13. Since projections are open homomorphisms (2.16), and the composition of open maps is open, U is open. Since the restriction of an open homomorphism is open, U_0 is open, and hence so are U_1 and U_2.

To show that U_1 is injective, consider an element $g \in G_0$, $g \neq 0$. We shall show that $g \notin \ker U_1$.

Since $g \neq 0$, $T_0(g) = h_1 + h_2 \neq 0$ (by (A3)), where $h_1 \in T_1(G)$, $h_2 \in T_2(G)$ and this representation is unique by (A4). If $h_1 = 0$, $T_0(g) = h_2 \in T_2(G)$, and hence $T_0(G) \cap T_1(G) \neq \{0\}$, contradicting (A5). So $h_1 \neq 0$ and similarly $h_2 \neq 0$.

Hence, by (A4) we can find a unique pair (g_1, g_2) with $0 \neq g_1 \in G_1$, $0 \neq g_2 \in G_2$ such that $T_1(g_1) = h_1$ and $T_2(g_2) = h_2$.

Now

$$U_1(g) = p_1(T_1 + T_2)^{-1}T_0(g)$$

$$= p_1(T_1 + T_2)^{-1}(h_1 + h_2)$$

$$= p_1(g_1 + g_2)$$

$$= g_1 \neq 0.$$

Hence $U_1(g) \neq 0$, thus $g \notin \ker U_1$, and hence U_1 is injective.

Similarly we can show that U_2 is injective, and the lemma is proved. //.

Corollary 1: $U_1^* \neq 0$ and $U_2^* \neq 0$.

Corollary 2: Let $\gamma_0 \in G_0^*$. Then there exist $\gamma_1 \in G_1^*$ and $\gamma_2 \in G_2^*$ such that

$$\gamma_0 = U_1^*(\gamma_1) = U_2^*(\gamma_2). \tag{7.10}$$

Proof: Apply Theorem 4.11 to U_1 and U_2. //.

Lemma 7.2: For all $\zeta \in H^*$

$$\hat{\mu}_Y(\zeta) = \hat{\mu}_0(T_0^*(\zeta))\hat{\mu}_1(T_1^*(\zeta))\hat{\mu}_2(T_2^*(\zeta)) \tag{7.11}$$

Proof: For every Borel subset $A \subset G$, $\mu(A) = \mu_Y(T(A))$. Hence $\hat{\mu}_Y = \mu T^{-1}$, and therefore, for all $\zeta \in H^*$:

$$\hat{\mu}_Y(\zeta) = \int_H < h, \zeta > d\mu_Y(h), \qquad h \in H$$

$$= \int_H < h, \zeta > d\mu T^{-1}(h)$$

$$= \int_G \langle T(g), \zeta \rangle \, d\mu(g), \qquad g \in G \qquad \text{(by Thm. 3.8)}$$

$$= \int_G \langle g, T^*(\zeta) \rangle \, d\mu(g)$$

$$= \int_G \langle g, [T_0^*(\zeta) + T_1^*(\zeta) + T_2^*(\zeta)] \rangle \, d\mu(g) \qquad \text{(by Thm. 4.10)}$$

$$= \int_G \langle g, T_0^*(\zeta) \rangle \langle g, T_1^*(\zeta) \rangle \langle g, T_2^*(\zeta) \rangle \, d\mu(g)$$

$$= \int_G \langle g, T_0^*(\zeta) \rangle \langle g, T_1^*(\zeta) \rangle \langle g, T_2^*(\zeta) \rangle \, d[\mu_0 \times \mu_1 \times \mu_2](g) \quad \text{(by (A1))}$$

$$= \int_G \langle g, T_0^*(\zeta) \rangle \, d\mu_0(g) \int_G \langle g, T_1^*(\zeta) \rangle \, d\mu_1(g) \int_G \langle g, T_2^*(\zeta) \rangle \, d\mu_2(g)$$

(by Thm. 3.7)

$$= \mu_0(T_0^*(\zeta)) \mu_1(T_1^*(\zeta)) \mu_2(T_2^*(\zeta)) . \qquad //.$$

Corollary: If the characteristic function of Y does not vanish, the characteristic functions of X_0, X_1 and X_2 do not vanish, and conversely.

Proof: This follows from (7.11), (4) of Theorem 5.5 and the fact that (A2) was not used in the proof of Lemma 2. //.

We are now ready to prove the main theorem.

Proof of Theorem 7.1: Let X' be another random variable with values in G satisfying (A1) and (A2). Let $X'_k = p_k X'$, $k = 0, 1, 2$, and let $Y' = T(X')$. Let the distributions of X', X'_0, X'_1, X'_2 and Y'

be ν, ν_0, ν_1, ν_2 and $\nu_{Y'}$ respectively, with a corresponding notation for their characteristic functions.

We assume that $\hat{\mu}_Y = \hat{\nu}_{Y'}$, and shall show that X_k and X'_k have the same distributions up to a shift.

By Lemma 2 we have for all $\zeta \in H^*$:

$$\hat{\nu}_0(T_0^*(\zeta))\hat{\nu}_1(T_1^*(\zeta))\hat{\nu}_2(T_2^*(\zeta)) = \hat{\mu}_0(T_0^*(\zeta))\hat{\mu}_1(T_1^*(\zeta))\hat{\mu}_2(T_2^*(\zeta)). \quad (7.12)$$

Let:

$$\gamma_0 = T_0^*(\zeta) \quad (7.13)$$

$$\gamma_1 = T_1^*(\zeta) \quad (7.14)$$

$$\gamma_2 = T_2^*(\zeta). \quad (7.15)$$

Then by (7.14), (7.15) and Theorem 4.10

$$\zeta = [(T_1 + T_2)^{-1}]^* (\gamma_1 + \gamma_2) \quad (7.16)$$

hence, again by Theorem 4.10 and (7.5) - (7.8),

$$\begin{aligned} \gamma_0 &= [(T_1 + T_2)^{-1} T_0]^* (\gamma_1 + \gamma_2) \\ &= U^*(\gamma_1 + \gamma_2) \\ &= U_0^*(\gamma_1 + \gamma_2) \end{aligned}$$

and since $U_1^*(\gamma_2) = U_2^*(\gamma_1) = 0$,

$$\gamma_0 = U_1^*(\gamma_1) + U_2^*(\gamma_2). \quad (7.17)$$

Moreover, since $T_1 + T_2$ is an isomorphism between $G_1 \oplus G_2$ and H, and since $(G_1 \oplus G_2)^* = G_1^* \oplus G_2^*$ (by Theorem 4.8) we may first substitute (7.17) in (7.13) and then substitute (7.13) - (7.15) in (7.12) and obtain:

$$\hat{\nu}_0(U_1^*(\gamma_1) + U_2^*(\gamma_2))\hat{\nu}_1(\gamma_1)\hat{\nu}_2(\gamma_2) = \hat{\mu}_0(U_1^*(\gamma_1) + U_2^*(\gamma_2))\hat{\mu}_1(\gamma_1)\hat{\mu}_2(\gamma_2)$$
$$\text{for all } \gamma_1 \in G_1^*,\ \gamma_2 \in G_2^*. \quad (7.18)$$

Now define:

$$\Psi_0 = \hat{\nu}_0/\hat{\mu}_0\ ;\ \Psi_1 = \hat{\mu}_1/\hat{\nu}_1\ ;\ \Psi_2 = \hat{\mu}_2/\hat{\nu}_2 \quad (7.19)$$

By (A2); (7.18) then becomes:

$$\Psi_0(U_1^*(\gamma_1) + U_2^*(\gamma_2) = \Psi_1(\gamma_1)\Psi_2(\gamma_2), \text{ for all } \gamma_1 \in G_1^*,\ \Psi_2 \in G_2^* \quad (7.20)$$

By Corollary 1 of Lemma 7.1 we may successively set $\gamma_2 = 0$ and $\gamma_1 = 0$, and, since $\Psi_1(0) = \Psi_2(0) = 1$ we obtain:

$$\Psi_0(U_1^*(\gamma_1)) = \Psi_1(\gamma_1) \text{ for all } \gamma_1 \in G_1^* \quad (7.21)$$

and

$$\Psi_0(U_2^*(\gamma_2)) = \Psi_2(\gamma_2) \text{ for all } \gamma_2 \in G_2^*. \quad (7.22)$$

Substituting (7.21) and (7.22) in (7.20) we obtain:

$$\Psi_0(U_1^*(\gamma_1) + U_2^*(\gamma_2)) = \Psi_0(U_1^*(\gamma_1))\Psi_0(U_2^*(\gamma_2))$$
$$\text{for all } \gamma_1 \in G_1^*,\ \gamma_2 \in G_2^*. \quad (7.23)$$

Now let $\gamma_0', \gamma_0'' \in G_0^*$. By Corollary 2 of Lemma 7.1 there exist $\gamma_1 \in G_1^*$, $\gamma_2 \in G_2^*$ such that $\gamma_0' = U_1^*(\gamma_1)$ and $\gamma_0'' = U_2^*(\gamma_2)$.

Hence (7.23) implies:

$$\Psi_0(\gamma_0' + \gamma_0'') = \Psi_0(\gamma_0')\,\Psi_0(\gamma_0'') \quad \text{for all} \quad \gamma_0', \gamma_0'' \in G_0^* . \tag{7.24}$$

Now Ψ_0 is continuous and since $\Psi_0(-\gamma_0) = \overline{\Psi_0(\gamma_0)}$, $|\Psi_0(\gamma_0)| = 1$ for all $\gamma_0 \in G_0^*$. Hence (7.24) implies that $\Psi_0 \in G_0^{**}$. Applying Pontryagin's Theorem, we conclude that there exists an element $g_0 \in G_0$ such that

$$\Psi_0(\gamma_0) = < g_0, \gamma_0 > \quad \text{for all} \quad \gamma_0 \in G_0^* . \tag{7.25}$$

Setting $\gamma_0 = U_1^*(\gamma_1)$ in (7.21) we obtain

$$\Psi_1(\gamma_1) = \Psi_0(U_1^*(\gamma_1)) = < g_0, U_1^*(\gamma_1) > = < U_1(g_0), \gamma_1 > \quad \text{for all} \quad \gamma_1 \in G_1^* , \tag{7.26}$$

and similarly, setting $\gamma_0 = U_1^*(\gamma_2)$ in (7.22) we obtain:

$$\Psi_2(\gamma_2) = < U_2(g_0), \gamma_2 > \quad \text{for all} \quad \gamma_2 \in G_2^* . \tag{7.27}$$

Hence, by (7.19) we obtain:

$$\hat{\nu}_0(\gamma_0) = < g_0, \gamma_0 > \hat{\mu}_0(\gamma_0) \quad \text{for all} \quad \gamma_0 \in G_0^* \tag{7.28}$$

$$\hat{\nu}_1(\gamma_1) = (< U_1(g_0), \gamma_1 >)^{-1} \hat{\mu}_1(\gamma_1) \quad \text{for all} \quad \gamma_1 \in G_1^* \tag{7.29}$$

$$\hat{\nu}_2(\gamma_2) = (< U_2(g_0), \gamma_2 >)^{-1} \hat{\mu}_2(\gamma_2) \quad \text{for all} \quad \gamma_2 \in G_2^* \tag{7.30}$$

which means that, (by Theorem 5.5, (6) and (7)), X_0' is distributed like $g_0 + X_0$, X_1' is distributed like $-U_1(g_0) + X_1$ and X_2' is distributed like $-U_2(g_0) + X_2$. This completes the proof. //.

Derivation of Kotlarski's and Prakasa Rao's Theorems

The theorem stated here is a generalization of Kotlarski's result in [16]. In that theorem X_0 had to be a one-dimensional random variable.

Theorem 7.2: Let $X_0 = [X_{0,1}, X_{0,2}, \ldots, X_{0,n_0}]$, $X_1 = [X_{1,1}, X_{1,2}, \ldots, X_{1,n_1}]$ and $X_2 = [X_{2,1}, X_{2,2}, \ldots, X_{2,n_2}]$ be three independent random vectors with non-vanishing characteristic functionals and values in the real vector spaces $\mathfrak{x}_0$, $\mathfrak{x}_1$ and $\mathfrak{x}_2$ respectively. Let:

$$\dim \mathfrak{x}_k = n_k, \qquad k = 0, 1, 2, \tag{7.31}$$

and denote

$$n = n_0 + n_1 + n_2 \tag{7.32}$$

$$m = n_1 + n_2 . \tag{7.33}$$

Assume:

$$1 \leq n_0 \leq \min (n_1, n_2) . \tag{7.34}$$

Define:

$$Y_\ell = \sum_{k=1}^{n_0} a_{\ell k} X_{0,k} + \sum_{k=n_0+1}^{n_0+n_1} a_{\ell k} X_{1,k} + \sum_{k=n_0+n_1+1}^{n} a_{\ell k} X_{2,k},$$

$$\ell = 1, \ldots, m \tag{7.36}$$

where $A = [a_{\ell k}]$ is an $m \times n$ matrix satisfying the conditions stated below. For any set $\{i_1, \ldots, i_{n_0}\}$ of n_0 natural numbers with $1 \leq i_1 < \ldots < i_{n_0} \leq n$, let $A(i_1, \ldots, i_{n_0})$ be the (square) matrix obtained from A by deleting the i_1-th, i_2-th, ..., i_{n_0}-th columns. Assume that:

(K1) $|A(1, 2, \ldots, n_0)| \neq 0$

(K2) There exists at least one set of n_0 natural numbers with $n_0 + 1 \leq i_1 < \ldots < i_{n_0} \leq n_0 + n_1$ such that $|A(i_1, \ldots, i_{n_0})| \neq 0$.

(K3) There exists at least one set of n_0 natural numbers with $n_0 + n_1 + 1 \leq i_1 < \ldots < i_{n_0} \leq n$ such that $|A(i_1, \ldots, i_{n_0})| \neq 0$.

Then the joint distribution of $(Y_1, \ldots, Y_m)$ determines the distributions of X_0, X_1 and X_2 up to a shift. The shift for X_0 is determined by an arbitrary element $x_0 \in \mathfrak{X}_0$, while the shifts for X_1 and X_2 depend on x_0.

Proof: Let $G = \mathfrak{X}_0 \times \mathfrak{X}_1 \times \mathfrak{X}_2 = R^n$ and let $H = \mathfrak{X}_1 \times \mathfrak{X}_2 = R^m$.

Since all the $\mathfrak{X}_k$'s are finite dimensional, G and H are locally compact, second countable, Hausdorff Abelian groups under the operation of vector addition.

Let $T : G \to H$ be the linear transformations the matrix of which, relative to the standard bases of R^n and R^m, is A. Then T is a (continuous) group homomorphism from G onto H.

Since (A1) and (A2) are stated explicitly in the hypothesis of the Theorem, it suffices to verify that assumptions (A3) - (A5) of

Theorem 7.1 hold.

Let

$$A = [A_0, A_1, A_2] \tag{7.37}$$

where

$$\begin{aligned} A_0 &= [a_{ij}]: i = 1, \ldots, n_0 ; j = 1, \ldots, m \\ A_1 &= [a_{ij}]: i = n_0 + 1, \ldots, n_0 + n_1 ; j = 1, \ldots, m \\ A_2 &= [a_{ij}]: i = n_0 + n_1 + 1, \ldots, m ; j = 1, \ldots, m \end{aligned} \tag{7.38}$$

Then the matrix of T_0 is $[A_0, 0, 0]$, that of T_1 is $[0, A_1, 0]$ and that of T_2 is $[0, 0, A_2]$. Hence either (K2) or (K3) implies that rank $A_0 = n_0$. Thus T_0 is injective and (A3) is satisfied. Also (K1) implies (A4).

Now (K3) implies that rank $(T_0 + T_1) = n_0 + n_1$ while (K2) implies that rank $(T_0 + T_2) = n_0 + n_2$. Thus if $d = \dim [T(\mathfrak{X}_0) \cap T(\mathfrak{X}_1)]$, then rank $(T_0 + T_1) = n_0 + n_1 - d$, and hence $d = 0$. Thus $T(\mathfrak{X}_0) \cap T(\mathfrak{X}_1) = \{0\}$. Similarly $T(\mathfrak{X}_0) \cap T(\mathfrak{X}_2) = \{0\}$ and hence (A5) is satisfied.

According to the main theorem, the shift of X_0 is determined by an arbitrary element $x_0 \in \mathfrak{X}_0$. The shifts of X_1 and X_2 are then determined by $-U_1(x_0)$ and $-U_2(x_0)$ respectively.

Now the matrix of U_0 is $[A_1, A_2]^{-1} A_0$.

Putting $B = [A_1, A_2]^{-1} A_0$, and writing the $m \times n_0$ matrix B in the form $B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$ where B_k has dimension $n_k \times n_0$, $k = 1, 2$, we see that if the shifts of X_1 and X_2 are determined by x_1 and x_2 respectively, then

$$x_1 = -\begin{bmatrix} B_1 \\ 0 \end{bmatrix} x_0 \quad \text{and} \quad x_2 = -\begin{bmatrix} 0 \\ B_2 \end{bmatrix} x_0 .$$

This concludes the proof of the Theorem. //.

Theorem 7.3 (Prakasa Rao's Theorem): [23]. Let X_0, X_1 and X_2 be three independent random variables with non-vanishing characteristic functions and values in a group G'. Let $Y_1 = X_0 + X_1$ and $Y_2 = X_0 + X_2$. Then the joint distribution of (Y_1, Y_2) determines the distributions of X_0, X_1 and X_2 up to a shift. The shift for X_0 is determined by an arbitrary element $g_0 \in G'$, those for X_1 and X_2 are then determined by the element $-g_0$.

Proof: Let $G_0 = G_1 = G_2 = G'$, $G = G_0 \times G_1 \times G_2$, $H = G_1 \times G_2$ and $X = (X_0, X_1, X_2)$. For $(g_0, g_1, g_2) \in G$ define $T : G \to H$ by:

$$T(g_0, g_1, g_2) = (g_0 + g_1, g_0 + g_2).$$

Then

$$T_0(g_0, g_1, g_2) = (g_0, g_0)$$

$$T_1(g_0, g_1, g_2) = (g_1, 0)$$

$$T_2(g_0, g_1, g_2) = (0, g_2)$$

and

$$T(X) = (Y_1, Y_2).$$

Clearly (A1) - (A5) are satisfied. Moreover for all $g_0 \in G_0$:

$$U_1(g_0) = p_1 U_0(g_0) = p_1(T_1 + T_2)^{-1} T_0(g_0)$$

$$= p_1(T_1 + T_2)^{-1}(g_0, g_0)$$

$$= p_1(g_0, g_0)$$

$$= g_0$$

and similarly $U_2(g_0) = g_0$.

Hence the Theorem is proved. //.

Two Characterizations of the Gamma Distribution

We now show how Theorem 7.1 could be applied to prove characterization theorems of real random variables. We have chosen the two characterizations of the gamma distribution given below because these two well-known theorems provided the basic intuition needed to formulate Theorem 7.2 which in turn led to Theorem 7.1. Theorem 7.4 is a rather straightforward application of the main theorem of this chapter, the proof of Theorem 7.5 requires some of the machinery developed in Chapter II.

Theorem 7.4: Let X_0, X_1 and X_2 be three independent positive random variables. Let

$$Y_1 = X_1/X_0; \quad Y_2 = X_2/X_0. \tag{7.39}$$

A necessary and sufficient condition for X_k to be gamma distributed with parameters $p_k > 0$ and $a > 0$ (a-common, $k = 0, 1, 2$) is that the joint distribution of (Y_1, Y_2) be the bivariate beta distribution of

the second kind given by the density

$$g(y_1, y_2) = \begin{cases} \dfrac{\Gamma(p_0 + p_1 + p_2)}{\Gamma(p_0)\Gamma(p_1)\Gamma)p_2)} \cdot \dfrac{y_1^{p_1 - 1} y_2^{p_2 - 1}}{(1 + y_1 + y_2)^{p_0 + p_1 + p_2}}; & y_1, y_2 > 0 \\ 0 \text{ elsewhere} \end{cases} \tag{7.40}$$

(Kotlarski [14]).

Proof: A straightforward computation (Kotlarski [14]) shows that if X_0, X_1 and X_2 are gamma distributed with parameters $p_k > 0$ and $a > 0$ (a-common, $k = 0, 1, 2$) then (Y_1, Y_2) given (7.39) is distributed with density (7.40).

Conversely, let $G = (R_+^3, \cdot)$, $H = (R_+^2, \cdot)$, $X = [X_0, X_1, X_2]$ and define $T : G \to H$ by

$$T(g_0, g_1, g_2) = (g_1/g_0, g_2/g_0). \tag{7.41}$$

Then for all $g_0, g_1, g_2 \in R_+$

$$T_0(g_0, g_1, g_2) = (1/g_0, 1/g_0)$$

$$T_1(g_0, g_1, g_2) = (g_1, 1)$$

$$T_2(g_0, g_1, g_2) = (1, g_2)$$

and so clearly (A3) - (A5) of the main theorem are satisfied.

(A1) is satisfied by assumption.

To see that (A2) holds, let $Z = (Z_1, Z_2) = (\log Y_1, \log Y_2)$. Then μ_Z does not vanish, Kotlarski [14], and so (A2) follows from the Corollary to Lemma 7.2.

Finally a simple calculation shows that for all $x \in R_+$, $U_1(x) = U_2(x) = \frac{1}{x}$.

Theorem 7.1 then implies that the joint distribution of (Y_1, Y_2) determines the distributions of X_0, X_1 and X_2 up to a change of scale. The required results then follow from the first part of this Theorem and the fact that if X is gamma distributed with parameters p and a, then cX is gamma distributed with parameters p and ca for all $c \in R_+$.

Theorem 7.5: Let X_0, X_1 and X_2 be three independent positive random variables. Let:

$$Y_1 = \sqrt{\frac{p}{2}} \cdot \frac{X_1 - X_0}{\sqrt{X_1 \cdot X_0}} \; ; \; Y_2 = \sqrt{\frac{p}{2}} \cdot \frac{X_2 - X_0}{\sqrt{X}} . \tag{7.43}$$

A necessary and sufficient condition for X_0, X_1 and X_2 to be identically gamma distributed with parameters $p > 0$ fixed, and $a > 0$ free is that (Y_1, Y_2) have joint distribution given by

$$g(y_1, y_2) = \frac{2\Gamma(3p)}{p[\Gamma(p)]^3} \cdot \frac{\left(\frac{y_1}{\sqrt{2p}} + \sqrt{1 + \frac{y_1^2}{2p}}\right)^{2p} \left(\frac{y_2}{\sqrt{2p}} + \sqrt{1 + \frac{y_2^2}{2p}}\right)^{2p}}{\left[1 + \left(\frac{y_1}{\sqrt{2p}} + \sqrt{1 + \frac{y_1^2}{2p}}\right)^2 + \left(\frac{y_2}{\sqrt{2p}} + \sqrt{1 + \frac{y_2^2}{2p}}\right)^2\right]^{3p}} \cdot$$

$$\cdot \frac{1}{\sqrt{1 + \frac{y_1^2}{2p}} \sqrt{1 + \frac{y_2^2}{2p}}} \; ; \; y_1, y_2 \in R. \tag{7.4}$$

(Kotlarski [12]).

Proof: Again a straightforward computation (Kotlarski [12]) shows that if X_0, X_1 and X_2 are gamma distributed with parameters $p > 0$ fixed and $a > 0$ free, then (Y_1, Y_2) given by (7.43) is distributed with density (7.44).

Conversely, let $G = (R_+^3, \cdot)$. Then (A1) of the main theorem is satisfied by assumption.

To see that (A2) holds, let $Z = (Z_1, Z_2) = (\log Y_1, \log Y_2)$. Then μ_Z does not vanish, Kotlarski [12], and so (A2) follows from the Corollary to Lemma 7.2.

Now define $f: R_+ \to R$ by

$$f(x) = \sqrt{\frac{p}{2}}\ (x - \frac{1}{x}) . \tag{7.45}$$

Clearly f satisfies the conditions of Theorem 2.2 and so we may set $H = (R_f^2, o)$, (see Definition 2.6). Now define $T: G \to H$ by

$$T(g_0, g_1, g_2) = \left(\sqrt{\frac{p}{2}}\ \frac{g_1 - g_0}{\sqrt{g_1 g_0}}, \sqrt{\frac{p}{2}}\ \frac{g_2 - g_0}{\sqrt{g_2 g_0}} \right) \tag{7.46}$$

To see that T is a homomorphism satisfying the conditions of Theorem 7.1, consider the following diagram:

$$G \xrightarrow{T'} (R_+^2, \cdot) \xrightarrow{f'} (R_f^2, 0) \tag{7.47}$$

where

$$T'(g_0, g_1, g_2) = (\sqrt{g_1/g_0}\ , \sqrt{g_2/g_0}) \tag{7.48}$$

and

$$f'(x_1, x_2) = (f(x_1), f(x_2)), \ (x_1, x_2) \in R_+^2 . \tag{7.49}$$

Now:

$$\begin{aligned} T_0'(g_o, g_1, g_2) &= (\sqrt{1/g_0} \ , \ \sqrt{1/g_0}) \\ T_1'(g_0, g_1, g_2) &= (\sqrt{g_1} \ , \ 1) \\ T_2'(g_0, g_1, g_2) &= (1, \sqrt{g_2}) , \end{aligned} \tag{7.50}$$

and hence T' satisfies (A3) - (A5) of Theorem 7.1. But $T = f'T'$ and hence by Theorem 2.2, T also satisfies (A3) - (A5) of Theorem 7.1.

Finally, a straightforward calculation shows that for all $x \in R_+$, $U_1(x) = U_2(x) = 1/x$.

Theorem 7.1 then implies that the joint distribution of (Y_1, Y_2) determines the distributions of X_0, X_1 and X_2 up to a change of scale. The required results then follow from the first part of this Theorem and the fact that if X is gamma distributed with parameters p and a, then cX is gamma distributed with parameters p and ca for all $c \in R_+$. //.

CHAPTER VIII

CONCLUSIONS

Attempts to characterize the distributions of certain real random variables lead, in a natural way, to a consideration of random variables with values in a topological group. One of the more interesting insights which the author gained as a result of this study was the fact that a real random variable, with distribution function F in the class $\mathfrak{F}(a, b)$ may be considered to be uniformly distributed on the group $([a, b), F)$. As a result of this insight, it was possible to prove the main characterization theorems of Chapters V and VI, and to apply these, in turn, to the derivation of some classical characterization theorems.

Since it is possible to put a group structure on the support $[a, b]$ of a real random variable X by choosing functions which do not belong to the class $\mathfrak{F}(a, b)$, the question arises whether such a procedure might yield some fruitful results. It is suggested that further investigations in this area might be of interest.

The main theorem of Chapter VII is a straightforward generalization of Professor Kotlarski's result [16]. It was possible to obtain this generalization rather easily because there is basically only one interesting topology on the dual G^* of a locally compact Abelian group G. The Fourier-Stieltjes transform of a probability measure μ on G

is uniformly continuous on G^*, and we do have Pontryagin's theorem available as a basic tool.

The problem becomes considerably more difficult when attempts are made to generalize Theorem 7.1 to random variables with values in a topological vector space $\mathfrak{X}$. In that case we obtain a whole host of topologies on $\mathfrak{X}^*$, and the Fourier-Stieltjes transform need not be continuous on $\mathfrak{X}^*$ if the wrong topology is chosen. Although a theorem of Banach's could play the same role as Pontryagin's Theorem played in the proof of Theorem 7.1, this theorem cannot be as readily applied because of the aforementioned topological complications.

Still the problem of generalizing Theorem 7.1 to vector spaces remains intriguing, and it is felt that perhaps the theory of duality for topological vector spaces might hold the key to the solution.

In addition to proving the new results in Chapters VI and VII, it has been the author's intention to provide a road map which will lead the beginner along the easiest route to that point where he can participate to some extent in active mathematical research.

The basic notions needed to understand probability on abstract spaces have been developed in Chapters II, III and IV. Although most of the subjects touched upon lead to deep and beautiful theories in their own right, we have developed them only to the point where we could use them as tools in our later studies. An undergraduate might find it a novel experience to realize that certain topics, which he had filed in his mind under the heading of "abstract mathematics" can be used as tools in other branches of his subject and lead to some very "concrete" results.

One last point needs to be made. In our study of group-valued random variables in Chapter V, we have not mentioned any of the limit theorems, even though these are the cap-stones of the whole theory. We have avoided them because they are more difficult and because they are only indirectly related to the type of characterization problems that formed the main subject of this study.

A SELECTED BIBLIOGRAPHY

[1] Aczél, J. Lectures on Functional Equations and their Applications, Academic Press, New York, (1966).

[2] Ash, R. B. Basic Probability Theory, John Wiley and Sons, New York, (1970).

[3] Committee on the Undergraduate Program in Mathematics: Pregraduate Preparation of Research Mathematicians, Mathematical Association of America, Berkeley, California, (1963).

[4] Dunford, N. and Schwartz, J. T. Linear Operators, Part I, General Theory, Interscience Publishers, New York, (1964).

[5] Flusser, P. "A Property of the Uniform Distribution on Compact Abelian Groups with Applications to Characterization Problems in Probability," Rendiconti del' Accademia Nazionale dei Lincei, to appear (1971).

[6] Grenander, U. Probabilities on Algebraic Structures, John Wiley and Sons, New York, (1963).

[7] Halmos, P. R. Measure Theory, D. Van Nostrand Co., New York, (1950).

[8] Herstein, I. N. Topics in Algebra, Blaisdell Publishing Co., New York, (1964).

[9] Hocking, J. G. and Young, G. S. Topology, Addison Wesley Publishing Co., Reading, Mass., (1961).

[10] Husain, T. Introduction to Topological Groups, W. B. Saunders Co., Philadelphia, (1966).

[11] Kelley, J. L. General Topology, D. Van Nostrand Co., New York, (1955).

[12] Kotlarski, I. I. "On Characterizing the Chi-Square Distribution by the Student Law," Journal of the American Statistical Association, Vol. 61, (Dec. 1966), pp. 976-981.

[13] Kotlarski, I. I. "On Characterizing the Normal Distribution by Student's Law," Biometrika, 53, 324, (1966) pp. 603-606.

[14] Kotlarski, I. I. "On Characterizing the Gamma and the Normal Distribution," Pacific Journal of Mathematics, 20, (1967).

[15] Kotlarski, I. I. "On Some Characterization of Probability Distributions in Hilbert Space," Annali di Matematica Pura ed Applicata, IV, Ser. 74, (1966), pp. 129-134.

[16] Kotlarski, I. I. "On a Characterization of Probability Distributions by the Joint Distribution of Some of their Linear Forms," Sankhya, (to appear 1971).

[17] Lindgren, B. W. Statistical Theory, Macmillan Co., New York, (1960).

[18] Loomis, L. H. An Introduction to Abstract Harmonic Analysis, D. Van Nostrand Co., New York, (1953).

[19] Lukacs, E. Characteristic Function, Hafner Publishing Co., New York, (1960).

[20] Mauldon, J. G. "Characterizing Properties of Statistical Distributions," Quarterly Journal of Mathematics, Oxford Series, 2.7, (1956), pp. 155-160.

[21] Parthasarathy, K. R. Probability Measures on Metric Spaces, Academic Press, New York, (1967).

[22] Pontryagin, L. S. Topological Groups, Gordon and Breach, New York, (1966).

[23] Prakasa Rao, B. L. S. "On a Characterization of Probability Distributions on Locally Compact Abelian Groups," Zeitschrift der Wahrscheinlichkeits-theorie, Feb. 9, (1968), pp. 98-100.

[24] Rudin, W. Fourier Analysis on Groups, Interscience Publishers, New York, (1962).

[25] Rudin, W. Real and Complex Analysis, McGraw-Hill Book Co., New York, (1966).

[26] Williams, E. J. "Cauchy-Distributed Functions and a Characterization of the Cauchy Distribution," Annals of Mathematical Statistics, (1969), Vol. 40, No. 3, pp. 1083-1085.

[27] Sazonov, V. V. and Tutubalin, V. N. "Probability Distributions on Topological Groups," (translated by B. Seckler), Theory of Probability and its Applications, XI, No. 1, (1966).

VITA

Peter R. Flusser

Candidate for the Degree of

Doctor of Education

Thesis: SOME CHARACTERIZATION PROBLEMS OF RANDOM VARIABLES WITH VALUES IN A LOCALLY COMPACT ABELIAN GROUP

Major Field: Higher Education

Biographical:

Personal Data: Born in Vienna, Austria, July 3, 1930, the son of Mr. and Mrs. Rudolf Flusser.

Education: Elementary and secondary education received in Vienna, Austria, Prague, Czechoslovakia and Shanghai, China; graduated from Rhodes School, New York, New York. Attended Columbia University, New York, New York, 1948-1950 and 1952-1953. Received the Bachelor of Arts degree from Ottawa University, Ottawa, Kansas in 1958 with a major in Mathematics, received the Master of Arts degree from the University of Kansas, Lawrence, Kansas in 1960. Attended the University of Kansas part time 1960-1967. Completed requirements for the Doctor of Education degree at Oklahoma State University, Stillwater, Oklahoma, in May 1971.

Professional Experience: Graduate Assistant, the University of Kansas, Lawrence, Kansas 1958-1960 and Oklahoma State University, Stillwater, Oklahoma, 1969-1971. Assistant Professor of Mathematics, Ottawa University, Ottawa, Kansas 1960-1967, promoted to associate professor, 1967. Currently on leave from Ottawa University, Ottawa, Kansas. Associate health physicist, Oak Ridge National Laboratory, Oak Ridge, Tennessee, summers 1958, 1959, 1960, 1961 and 1963. Research mathematician, Tech/Ops. Burlington, Massachusetts, summers 1964 and 1965. Research assistant, the University of Kansas, Lawrence Kansas, summer 1966 and academic year 1966-1967, while on sabbatical leave from Ottawa University. Member of the Mathematical Association of America and the American Mathematical Society.