

NONCOMMUTING RANDOM PRODUCTS

BY
HARRY FURSTENBERG⁽¹⁾

Introduction. Let $X_1, X_2, \dots, X_n, \dots$ be a sequence of independent real valued random variables with a common distribution function $F(x)$, and consider the sums $X_1 + X_2 + \dots + X_n$. A fundamental theorem of classical probability theory is the strong law of large numbers which asserts that with probability one, $X_1 + X_2 + \dots + X_n \sim n \int x dF(x)$, provided that $\int |x| dF(x)$ is finite. It is natural to inquire whether there exist laws governing the asymptotic behavior of products $X_n X_{n-1} \dots X_1$, where the X_j are now identically distributed independent random variables with values in an arbitrary group. This type of question arises, for example, in considering solutions to differential or difference equations with random coefficients [13]. To illustrate this, consider the problem of determining the asymptotic behavior of a random sequence $\{\xi_n\}$ satisfying $\xi_n = u_n \xi_{n-1} + v_n \xi_{n-2}$, where (u_n, v_n) forms a sequence of independent identically distributed random vectors. In this case we can write

$$\begin{pmatrix} \xi_{n+1} \\ \xi_n \end{pmatrix} = X_n X_{n-1} \dots X_1 \begin{pmatrix} \xi_1 \\ \xi_0 \end{pmatrix}, \quad X_j = \begin{pmatrix} u_{j+1} & v_{j+1} \\ 1 & 0 \end{pmatrix},$$

and so the rate of growth of the ξ_n is governed by the behavior of the matrix product $X_n X_{n-1} \dots X_1$.

Bellman, who apparently was the first to consider questions of this kind, studied random products of 2×2 matrices with strictly positive entries [2]. He showed that in a certain sense the weak law of large numbers holds for the entries of the matrices $X_n X_{n-1} \dots X_1$. In [9] it was shown that a strong law of large numbers is valid. More precisely, if $y_{ij}^{(n)}$ is the typical entry of $X_n X_{n-1} \dots X_1$, then for a certain constant α , $n^{-1} \log y_{ij}^{(n)} \rightarrow \alpha$ with probability 1, provided the entries of X_j are positive and bounded away from ∞ and 0 in an appropriate sense. It should be emphasized that for matrices with arbitrary entries, this type of result does not hold and the analysis of [9] breaks down.

In the present investigation we shall consider a noncompact semi-simple Lie group G and independent G -valued random variables $\{X_n\}$ with a common distribution μ on G . The law of large numbers here can be given the following form. We shall exhibit a finite dimensional linear space \mathcal{V}_G of functions $\psi(g)$

Received by the editors August 14, 1962.

(1) This research was partially supported by the Air Force Office of Scientific Research.

on G , such that to every probability measure μ that is well behaved in a sense to be specified, there corresponds a linear functional α_μ on \mathcal{V}_G and, with probability 1,

$$\psi(X_n X_{n-1} \cdots X_1) \sim n\alpha_\mu(\psi)$$

as $n \rightarrow \infty$, for each $\psi \in \mathcal{V}_G$. This may be illustrated by an example. Let G be the group of conformal transformations of the unit disc onto itself. For every $g \in G$, set $\psi'(g)$ equal to the non-euclidean distance of $g(0)$ from 0. ψ' itself does not belong to \mathcal{V}_G but it differs from a function $\psi \in \mathcal{V}_G$ by a bounded function on G . Consequently, we will have $n^{-1} \psi'(X_n X_{n-1} \cdots X_1) \rightarrow \alpha_\mu(\psi)$.

In general, let G be any semi-simple group and Λ an irreducible representation of G by $m \times m$ matrices. Let u be a vector in R^m and let $\|\cdot\|$ denote a norm on R^m . There will be a $\psi \in \mathcal{V}_G$ such that for each distribution μ , $\log \|\Lambda(X_n X_{n-1} \cdots X_1)u\| \sim n\alpha_\mu(\psi)$. A similar relation holds, under more restricted conditions, for the angle between various column vectors of the product matrix $X_n X_{n-1} \cdots X_1$, where G is a matrix group. We find that the logarithm of the angle is asymptotic to $n\alpha_\mu(\psi_0)$, and in certain cases we can show that $\alpha_\mu(\psi_0) < 0$. This implies that the columns of the product $X_n X_{n-1} \cdots X_1$ are "ultimately" proportional.

In contrast to the classical situation, the computation of the limit $\alpha_\mu(\psi)$ of $n^{-1} \psi(X_n X_{n-1} \cdots X_1)$ cannot, in general, be made directly in terms of the measure μ . The expression for $\alpha_\mu(\psi)$ involves, in addition to μ , an auxiliary measure ν on a space $B(G)$ associated with the group G . It turns out that ν is determined by μ as the solution to a convolution equation $\mu * \nu = \nu$, the meaning of which will become clear presently.

The main step in the proof that the limit of $n^{-1} \psi(X_n X_{n-1} \cdots X_1)$ exists for $\psi \in \mathcal{V}_G$ consists of showing that the above convolution equation determines a unique probability measure ν . This equation is natural in any case in view of the appearance of ν in the expression for $\alpha_\mu(\psi)$. We shall study this question in some detail. If a group G acts on a space M , then the product gx is defined for $g \in G$, $x \in M$. It follows that if μ is a probability measure on G and ν a probability measure on M , then $\mu * \nu$ may be defined as the measure on M giving the distribution of gx where g and x independently have distributions μ and ν respectively. We now ask whether for given μ , there exists a probability measure ν on M satisfying $\mu * \nu = \nu$. We call such a measure a stationary measure for μ . Moreover, if ν exists, we may ask if it is unique for a given space M . These two questions are related, for if we let $M^{(2)}$ denote the space of pairs of distinct points of M , then a stationary measure exists on $M^{(2)}$ if and only if there exist distinct stationary measures on M .

As to the existence of a stationary measure for μ on M , our analysis shows that this is independent of the measure μ and depends only on the space M in

question (Theorem 2.1). More precisely, the necessary and sufficient condition is that a certain subgroup $H(G)$ leave fixed some probability measure on M . The subgroup $H(G)$ as well as the space $B(G)$ mentioned previously were introduced in our discussion of a generalized Poisson formula. The proof of Theorem 2.1 and much of our discussion depends strongly on [8] and we shall summarize these results in §0.

The question of the existence of a stationary measure for μ on a G -space M has independent interest. A probability measure on G induces on every G -space M a "random walk" $\{W_n\}$ where we let W_{n+1} arise by applying a random g to W_n , g having distribution μ . We may now inquire as to the nature of this random walk. There are at least three possibilities that may be distinguished. The walk may be *transient*; that is, with probability 1, W_n may ultimately leave every compact set. It may be *null-recurrent*; that is, W_n may return with probability 1 infinitely often to some compact set, but the expected time for return to the set from a point of M will be infinite. Finally, the expected time may be finite; this is sometimes called the *positively recurrent* case. (These definitions are informal and only serve to give an idea of the possibilities.) It may be seen that the latter case will occur if and only if μ possesses a stationary measure on M .

Our results imply that the question of the positive recurrence of $\{W_n\}$ depends only on the state space M and not on the measure μ generating the random walk. When M is compact, there will always be positive recurrence. For certain groups G , under the assumption that $M = G/L$ and L is a connected subgroup, we can say conversely that if M is not compact, the walk is either null-recurrent or transient.

For certain M we can also assert the transience of the random walk independent of μ . For example, if G is any noncompact semi-simple group the induced random walk on G itself is always transient. (This follows from Theorem 7.6.) However, there are homogeneous G -spaces for which the transience or recurrence of the W_n does depend on the measure μ .

The paper is divided into two parts, the first of which is concerned with the question of stationary measures and the second with laws of large numbers. Part II depends upon Part I primarily in the application of the uniqueness criterion for a stationary measure (Theorem 2.1).

0. Preliminaries.

0.1. G -SPACES. All the topological spaces to be considered in the sequel will be assumed to be locally compact, Hausdorff, and separable. If G is a group then M is a G -space if there is defined a continuous map $(g, x) \rightarrow gx$ of $G \times M \rightarrow M$, such that gx may be considered the result of applying the group element g to x . This means that $g_1(g_2x) = (g_1g_2)x$ and $ex = x$ where e is the identity of G . If for each pair $x, y \in M$ the equation $gx = y$ has a solution in M , then G is *transitive* on M and M is a homogeneous space of G . If $x_0 \in M$, the set of $g \in G$ with $gx_0 = x_0$

is called the *stability group* of x_0 . When G is transitive on M then M may be identified with the coset space G/L , where L is the stability group of a point of M .

An *equivariant* map ϕ of one G -space into another is one that commutes with the action of G : $g\phi(x) = \phi(gx)$. If L_1 and L_2 are subgroups of G , $L_1 \subset L_2$, there is a natural equivariant map of $G/L_1 \rightarrow G/L_2$ taking gL_1 into gL_2 .

0.2. MEASURES. The term measure will refer to a regular, borel measure. With the exception of Haar measure, all our measures will be bounded. We will be concerned almost exclusively with probability measures which are measures assigning to the whole space the value 1. The integral $\int f(x)d\mu(x)$, which is defined at least for borel measurable functions, will often be abbreviated to $\mu(f)$. On a locally compact space μ is determined by the induced functional $\mu(f)$ on the class of continuous functions vanishing at ∞ .

If M is a G -space and μ and ν are measures on G and M respectively, then $\mu * \nu$ is defined by

$$\mu * \nu(f) = \int f(gx)d\mu(g)d\nu(x).$$

Considering G itself as a G -space, this defines the convolution of measures on G . If x is a point of a space, δ_x will denote the probability measure whose support is the single point x . For a measure ν on a G -space M it will often be necessary to consider the measures $\delta_g * \nu$. We shall abbreviate this to $g\nu$. Note that $g\nu(f) = \int f(gx)d\nu(x)$ so that gx has the distribution $g\nu$ if x has distribution ν . It will sometimes be instructive to consider $\mu * \nu$ as a weak integral

$$\mu * \nu = \int g\nu d\mu(g) = \int \mu * \delta_x d\nu(x).$$

When we speak of convergence of a sequence of measures we will always mean convergence in the weak sense: $\mu_n \rightarrow \mu$ if $\mu_n(f) \rightarrow \mu(f)$ for every bounded continuous function f on the space in question.

0.3. RANDOM WALKS. We shall use the term random walk to refer to the assignment of transition probabilities to points of a space. Specifically, by a random walk on M we shall mean a continuous function μ_x assigning to each $x \in M$ a probability measure on M . The hypothesis of continuity implies that if we set $\tau f(x) = \mu_x(f)$ then τ takes continuous functions into continuous functions.

If in addition to the measures $\{\mu_x\}$ we are given an arbitrary distribution (= probability measure) ν_0 on M , there is determined a Markov process with stationary transition probabilities. Namely, we let W_0 have distribution ν_0 and the joint distribution of W_1, W_2, \dots, W_n is determined inductively by setting the conditional distribution of W_n , given W_1, W_2, \dots, W_{n-1} , equal to $\mu_{W_{n-1}}$. The measure ν_0 is called the *initial distribution* of the process.

0.4. BOUNDARIES AND POISSON SPACES. Most of our considerations are limited to noncompact semi-simple groups. By a *semi-simple* group we will always mean a semi-simple connected Lie group with finite center. It is for these groups that we established the results of [8] and aside from reference to these results we will not require very much of the general theory of these groups. In this class of groups are the so-called classical groups which include the groups $SL(m, R)$ and $SL(m, C)$ of $m \times m$ unimodular real or complex matrices.

The conclusions of [8] are as follows. If G is a noncompact semi-simple group it has nontrivial compact homogeneous spaces called *boundaries*. These are characterized by the property that for any probability measure π on such a space, there exists a sequence $\{g_n\}$ in G with $g_n\pi$ converging to a point measure. All boundaries of a group G are equivariant images of one of them, $B(G)$, called the *maximal boundary* of G (and conversely, every equivariant image of $B(G)$ is a boundary). $B(G)$ can be identified with a coset space $G/H(G)$. The subgroup $H(G)$ is not uniquely determined, but the conjugates $\{g^{-1}H(G)g\}$ exhaust this class of groups. They are referred to as *minimal B-subgroups* of G .

There are finitely many covering spaces $\{\Pi\}$ of $B(G)$ which are also G -spaces, and these correspond to the subgroups H_π of $H(G)$ containing the connected component $H_0(G)$ of the identity of $H(G)$. These spaces $\{\Pi\}$ are the *Poisson spaces* of G .

Now let μ be an absolutely continuous probability measure on G . The solutions to the equation

$$(0.1) \quad f(g) = \int f(gg')d\mu(g')$$

are called μ -harmonic functions on G . To each such measure μ there corresponds a Poisson space Π_μ (frequently the Poisson space of μ is the maximal boundary $B(G)$) and there exists a probability measure ν on Π_μ with the following property. Every bounded μ -harmonic function on G is given by the "Poisson" representation

$$(0.2) \quad f(g) = \int_{\Pi_\mu} \hat{f}(gx)d\nu(x)$$

for some bounded function \hat{f} on Π_μ ; every \hat{f} on Π_μ determines a μ -harmonic function on G , and the correspondence $f \leftrightarrow \hat{f}$ is one-to-one. The measure ν in formula (0.2) depends only on μ and satisfies $\mu * \nu = \nu$. The μ -harmonic functions corresponding to continuous boundary functions \hat{f} may be characterized by the property of being *left-uniformly continuous* on G . This means that as $g_n \rightarrow$ identity, $f(g_n g) \rightarrow f(g)$ uniformly in g .

If G is a semi-simple group, then G admits a decomposition $G = K \cdot A \cdot N$, referred to as the Iwasawa decomposition [13, Chapter VI, §3]. Here K is any maximal compact subgroup of G , A is an abelian group, N is nilpotent, and

the product $A \cdot N = S$ is a solvable, connected, and simply connected subgroup. This decomposition is meant to imply that each $g \in G$ can be expressed uniquely in the form $g = kan$, $k \in K$, $a \in A$, $n \in N$. In [8] it was shown that for any such decomposition, S is contained in a minimal B -subgroup $H(G)$. Moreover, the normalizer of S is also contained in $H(G)$. C. C. Moore has shown (unpublished) that $H(G)$ is identical with the normalizer of S . This means [5, Chapter III] that if M is the centralizer of A in K , then $H(G)$ can be written as $H(G) = M \cdot A \cdot N$, and $B(G) \approx K/M$. If M_0 is the connected component of the identity of M , then $H_0(G) = M_0 \cdot A \cdot N$, and the subgroups H_π correspond to the subgroups of the finite group M/M_0 . The group $M \cdot A \cdot N$ also plays an important role in the theory of representations of G [5].

Let μ be a probability measure on G . Associate with μ the random walk on G defined by setting $\mu_g = g\mu$ (§0.3); that is, from each $g \in G$ there is a transition to gg' where g' is distributed according to μ . A μ -harmonic function can then be seen to be harmonic with respect to this random walk: if $\{W_n\}$ denotes a Markov process associated with this random walk, then $E(f(W_{n+1}) | W_n) = f(W_n)$, if and only if f is μ -harmonic. For certain classes of Markov processes [15] there has been developed a notion of a Martin boundary of the state space induced by the process. The non-negative (an *a fortiori* the bounded) harmonic functions for the process are then given by a Poisson representation in terms of measures on the Martin boundary. Also the "minimal" non-negative harmonic functions correspond to a certain subset of the Martin boundary. In case the measure μ on G is sufficiently regular (μ of class B_∞ , see Definition 7.3 below), then a Poisson representation is valid for all non-negative μ -harmonic functions (in addition to the bounded ones) where the boundary space is taken as Π_μ [8, Chapter V, §5]. Moreover, each point $x \in \Pi_\mu$ determines a minimal, non-negative μ -harmonic function

$$f_x(g) = \frac{dg\nu}{dm}(x),$$

where ν is the measure in (0.2) and m is a fixed measure on Π_μ , and up to scalar multiples, these exhaust the class of minimal non-negative μ -harmonic functions. Finally, it is not difficult to show that f_x and f_y are proportional only if $x = y$. These properties suggest identifying Π_μ as the Martin boundary of G induced by the random walk on G corresponding to μ . It ought to be pointed out that using martingale convergence arguments, the existence of a Martin boundary with these properties may be established for probability measures on arbitrary locally compact groups G . It is also not hard to show that such a Martin boundary will necessarily be a G -space. The semi-simplicity of G is needed in showing that the space Π_μ is a homogeneous G -space, and in identifying it with one of a finite number of covering spaces of the maximal boundary $B(G)$ of G .

0.5. GROUPS OF MATRICES. A group G of $m \times m$ real matrices naturally acts on

an m -dimensional space R^m . The action is given by identifying R^m with real $m \times 1$ matrices and letting G act by left multiplication. This identification will be implicit whenever we speak of the action of a matrix group on a vector space. The group G is called *irreducible* if the only subspaces left fixed by the matrices of G are R^m and $\{0\}$. Otherwise it is called *reducible*.

The real $(m-1)$ -dimensional projective space P^{m-1} is obtained from $R^m - \{0\}$ by identifying two vectors if each is a scalar multiple of the other. If two vectors are so identified then their images under a nonsingular linear transformation will also be identified. It follows that if G is a group of matrices, then G acts on P^{m-1} as well as on R^m .

PART I. RANDOM WALKS ON G -SPACES.

1. Stationary μ -processes.

1.1. GENERALITIES. Let G be a locally compact topological group, μ a probability measure on G , and M a G -space. The measure μ determines a random walk on M in a natural way. Namely, to each point $x \in M$ we assign the probability measure $\mu * \delta_x$ on M as the transition probability measure from x . We speak of this as the random walk induced by μ .

DEFINITION 1.1. A Markov process $\{Z_n\}$ with state space M and with transition probability measures $\mu_x = \mu * \delta_x$ will be called a μ -process on M . (The index n ranges either over the non-negative integers or all integers.)

A μ -process $\{Z_n, n \geq 0\}$ may always be obtained in the following way. We consider a sequence X_1, X_2, \dots , of G -valued random variables which are independent and all have distribution μ . Let Z_0 be an M -valued random variable with arbitrary distribution and independent of the X_n , and set $Z_n = X_n X_{n-1} \cdots X_1 Z_0$. Then $Z_{n+1} = X_{n+1} Z_n$, where X_{n+1} is independent of Z_n . It follows that the conditional distribution of Z_{n+1} , given Z_n , is $\mu * \delta_{Z_n}$ which shows that $\{Z_n\}$ is a μ -process. Conversely, it is easily seen that every μ -process arises this way.

We are particularly interested in stationary μ -processes. Since a μ -process is determined by its initial distribution, its stationarity depends on the initial distribution.

LEMMA 1.1. Let $\{Z_n, n \geq 0\}$ be a μ -process on M and let v be the distribution of Z_0 . $\{Z_n\}$ is stationary if and only if $\mu * v = v$.

Proof. Since the distribution of Z_1 is $\mu * v$, it follows that the process can be stationary only if $\mu * v = v$. Conversely, if $\mu * v = v$, then all the Z_n have the same individual distributions. Now the joint distribution of a set of variables $(Z_h, Z_{h+1}, \dots, Z_{h+m})$ is determined by the distribution of Z_h and the transition probabilities between consecutive variables. Since all this is independent of h , it follows that the process is stationary.

DEFINITION 1.2. A probability measure v on M is called a *stationary measure* for μ if $\mu * v = v$.

There are two questions with which we shall be concerned: for a given G -space M , does there exist a stationary measure, and if it exists, is it unique? The first of these questions only arises for noncompact spaces, for we have

LEMMA 1.2. *If M is compact, there exists at least one stationary measure for μ on M .*

Proof. Let ν' be any probability measure on M and set $\nu_n = n^{-1} \sum_0^{n-1} \mu^k * \nu'$ where μ^k denotes the k -fold convolution of μ . Since M is compact, some subsequence of $\{\nu_n\}$ converges to a probability measure ν , and since $\mu * \nu_n - \nu_n \rightarrow 0$, we have $\mu * \nu = \nu$.

1.2. A BASIC LEMMA. In this section we shall establish a lemma which is useful for showing the nonexistence of a stationary measure in certain cases. It is necessary first to introduce certain notation. If Z is a random variable with values in M , we shall denote the distribution of Z by $E^*(Z)$. That is to say, $E^*(Z)$ is a probability measure on M such that for any bounded continuous function f on M , $E^*(Z)(f) = E(f(Z))$. More generally, suppose that Z is defined over a sample space Ω and that \mathcal{F} is a σ -field on Ω . We define the conditional distribution $E^*(Z | \mathcal{F})$ by the equality

$$(1.1) \quad E^*(Z | \mathcal{F})(f) = E(f(Z) | \mathcal{F}).$$

Thus $E^*(Z | \mathcal{F})$ is itself a random variable with values in the space of probability measures on M . Because the space M is separable, as we shall assume, it may be shown that there is a uniquely determined (up to null sets) random variable $E^*(Z | \mathcal{F})$ satisfying (1.1) for bounded continuous f .

If $\{Y_n, n \geq 0\}$ is a stationary process, it may be imbedded in a doubly infinite stationary process in an essentially unique way, setting the joint distribution of $(Y_{-n}, Y_{-n+1}, \dots, Y_{-1})$ equal to that of $(Y_0, Y_1, \dots, Y_{n-1})$. Suppose now that $\{Z_n, n \geq 0\}$ is a stationary μ -process on a G -space M . We may suppose there are given variables $\{X_n, n \geq 1\}$ with values in G so that $Z_n = X_n X_{n-1} \cdots X_1 Z_0$, where the X_n are independent of each other and of Z_0 and where $E^*(X_n) = \mu$. The process $\{(X_{n+1}, Z_n), n \geq 0\}$ is stationary and may be extended so that the index n may also take on negative values. We then have $Z_{n+1} = X_{n+1} Z_n$ for all n , and X_{n+1} will be independent of all $Z_k, k \leq n$. We shall refer to the X_n as the associated G -valued random variables for the process $\{Z_n\}$.

LEMMA 1.3. *If $\{Z_n\}$ is a stationary μ -process and $\{X_n\}$ the associated G -valued random variables, then with probability 1,*

$$(1.2) \quad E^*(Z_n | X_n, X_{n-1}, \dots) = \lim_{k \rightarrow \infty} X_n X_{n-1} \cdots X_{n-k} E^*(Z_0).$$

Recall that since G acts on M , it acts on the measures on M so that the product on the right side of (1.2) is defined.

Proof. Let ψ be a bounded continuous function on M .

$$E^*(Z_n | X_n, X_{n-1}, \dots)(\psi) = E(\psi(Z_n) | X_n, X_{n-1}, \dots) = \lim_{k \rightarrow \infty} E(\psi(Z_n) | X_n, X_{n-1}, \dots, X_{n-k}),$$

by a martingale convergence theorem [6, Chapter VII, Theorem 4.3]. Now $Z_n = X_n X_{n-1} \dots X_{n-k} Z_{n-k-1}$, so that

$$\begin{aligned} E(\psi(Z_n) | X_n, X_{n-1}, \dots, X_{n-k}) &= E(\psi(X_n X_{n-1} \dots X_{n-k} Z_{n-k-1}) | X_n, X_{n-1}, \dots, X_{n-k}) \\ &= \int \psi(X_n X_{n-1} \dots X_{n-k} \zeta) d\nu(\zeta) \end{aligned}$$

where $\nu = E^*(Z_0)$, since $X_n, X_{n-1}, \dots, X_{n-k}$ are independent of Z_{n-k-1} , and the latter has distribution ν . It follows that

$$E(\psi(Z_n) | X_n, X_{n-1}, \dots, X_{n-k}) = X_n X_{n-1} \dots X_{n-k} \nu(\psi)$$

(see §0.2) and as $k \rightarrow \infty$,

$$\lim_{k \rightarrow \infty} E(\psi(Z_n) | X_n, X_{n-1}, \dots, X_{n-k}) = \lim_{k \rightarrow \infty} X_n X_{n-1} \dots X_{n-k} E^*(Z_0)(\psi).$$

A group G acts on itself by left multiplication. We may inquire whether G , as a G -space, possesses a stationary measure for μ .

THEOREM 1.1. *If G is any locally compact group, μ a probability measure on G , then G possesses a stationary measure for μ only if μ has its support in a compact subgroup of G .*

Proof. Suppose that $\{Z_n\}$ is a stationary μ -process on G and that $\{X_n\}$ are the associated independent variables. There is a compact set $\Delta_1 \subset G$ with $E^*(Z_0)(\Delta_1) > \frac{1}{2}$, and for some sample sequence for which (1.2) is valid we can find an open set Δ_2 with compact closure such that $E^*(Z_0 | X_0, X_{-1}, \dots)(\Delta_2) > \frac{1}{2}$. Then ultimately $X_0 X_{-1} \dots X_{-k} E^*(Z_0)(\Delta_2) > \frac{1}{2}$ by (1.2), and it follows that $X_0 X_{-1} \dots X_{-k} \Delta_1 \cap \Delta_2$ is nonempty. (Note that $\nu(\Delta) = g\nu(g\Delta)$.) Hence $X_0 X_{-1} \dots X_{-k} \in \Delta_2 \Delta_1^{-1} \subset \bar{\Delta}_2 \Delta_1^{-1} = \Delta_3$ where Δ_3 is compact. This holds for $k > k'$, say, so we have $X_{-n-1} \dots X_{-n-m} \in \Delta_3^{-1} \Delta_3$ for $k' \leq n < n + m$. Now by the ergodic theorem, any m -tuple (g_1, \dots, g_m) of elements of G in the support of μ can be approximated arbitrarily closely by some $(X_{-n-1}, \dots, X_{-n-m})$ occurring in almost any sample sequence. It follows that all products of group elements in the support of μ lie in the compact set $\Delta_4 = \Delta_3^{-1} \Delta_3$. But then the support of μ lies in a compact semigroup $\subset G$ and therefore in a compact subgroup of G .

We recall that a compact semigroup with cancellation (and therefore a compact sub-semigroup of a group) is a group.

We remark that if G is then taken to be the group of reals, then the standard device for proving this theorem would be the use of characteristic functions, a device which is not available for the general case.

Another application of Lemma 1.3 is the following.

THEOREM 1.2. *Let μ be a probability distribution on the set of $m \times m$ unimodular matrices and let G be the smallest closed subgroup of $SL(m, R)$ containing the support of μ . Then μ admits a stationary measure on the G -space $R^m - \{0\}$ only if G is either compact or reducible.*

Proof. Let $\{Z_n\}$ be a stationary μ -process on $R^m - \{0\}$, and let $\{X_n\}$ be the associated G -valued variables. Choose a sample sequence $\dots X_{-2}, X_{-1}, X_0$ so that

$$(1.3) \quad E^*(Z_0 | X_0, X_{-1}, \dots) = \lim_{k \rightarrow \infty} X_0 X_{-1} \dots X_{-k} E^*(Z_0),$$

and such that any m -tuple (g_1, \dots, g_m) of elements of G in the support of μ is approximated arbitrarily closely by m -tuples $(X_{n-1}, \dots, X_{n-m})$ of the sample sequence. If G is not compact, then as in the preceding proof, the set of matrices $\{X_0 X_{-1} \dots X_{-k}\}$ cannot be bounded in the group G . Since the matrices of G are unimodular, this implies that for some subsequence $g_n = X_0 X_{-1} \dots X_{-k_n}$, we have $\|g_n\| \rightarrow \infty$, where $\|g\| = \max |g_{ij}|$. Passing to a subsequence we may assume that $g_n / \|g_n\|$ converges elementwise to a matrix h . $\|h\| = 1$, and so h is clearly not the 0 matrix. It follows that for $u \in R^m$, hu can vanish only if u belongs to a proper subspace $V \subset R^m$. Now, if $hu \neq 0$, then $\|g_n u\| \rightarrow \infty$, where $\|\cdot\|$ now denotes a norm in R^m . On the other hand, by (1.3), $g_n E^*(Z_0)$ converges to a probability measure on R^m . This is therefore only possible if $E^*(Z_0)$ has its support in V . Let V' be the smallest subspace of R^m containing the support of $E^*(Z_0)$; we have $V' \subset V$. It is easily shown that since $\mu * E^*(Z_0) = E^*(Z_0)$, almost every g , but for a set of μ -measure 0, must take V' into V' . G is therefore generated by elements taking V' into V' and this says that G is reducible. This proves the theorem.

1.3. AN APPLICATION TO CONTINUED FRACTIONS. In this section we interrupt our main line of reasoning to present an application of Lemma 1.3 to a different kind of question. We will encounter this problem again in Part II in a more general setting.

THEOREM 1.3. *Let G be the group of fractional linear transformations taking the unit disc D of the complex plane onto itself:*

$$g \in G \Leftrightarrow gz = \rho \frac{z + \alpha}{1 + \bar{\alpha}z}, \quad |\rho| = 1, |\alpha| < 1.$$

Let μ be a probability measure on G whose support is not contained in a proper subgroup of G and let $\{X_n\}$ be a sequence of independent G -valued random variables with distribution μ . Then with probability 1, $X_1 X_2 \dots X_n(z)$ converges to a point of $\{|\zeta| = 1\}$ for each z in the interior of D , and the limit point is independent of z .

Proof. We can replace the variables $\{X_n, n \geq 1\}$ by variables $\{X_{-n}, n \leq 1\}$ and the assertion becomes that of the convergence of $X_{-1}X_{-2} \cdots X_{-n}(z)$. Since the group G is noncompact and the support of μ is not contained in a compact subgroup of G , we conclude as in the proof of Theorem 1.1 that with probability 1, the group elements $\{X_{-1}X_{-2} \cdots X_{-n}, n = 1, 2, \dots\}$ do not remain bounded. Now let us suppose that we have

$$g_n(z) = X_{-1}X_{-2} \cdots X_{-n}(z) = \rho_n(z + \alpha_n)(1 + \bar{\alpha}_nz)^{-1}.$$

It is clear that if we have a sequence $g_{n_k} \rightarrow \infty$, then $|\alpha_{n_k}| \rightarrow 1$, and since the g_n do not lie in a compact subset of G this must be true of some subsequence.

Now consider the action of G on the boundary B of D : $B = \{z : |z| = 1\}$. Since B is compact, there will be a stationary μ -process $\{Z_n\}$ on B (Lemma 1.2), and we can take the X_n as forming the associated G -valued variables. By Lemma 1.3,

$$\begin{aligned} E^*(Z_{-1} | X_{-1}, X_{-2}, \dots) &= \lim_{n \rightarrow \infty} X_{-1} \cdots X_{-n} E^*(Z_0) = \lim_{n \rightarrow \infty} g_n E^*(Z_0) \\ (1.4) \qquad \qquad \qquad &= \lim_{k \rightarrow \infty} g_{n_k} E^*(Z_0). \end{aligned}$$

Passing to a subsequence of $\{g_{n_k}\}$ we may suppose that $\rho_{n_k} \rightarrow \rho', \alpha_{n_k} \rightarrow \alpha'$, where necessarily $|\alpha'| = 1$. It follows that for all $\zeta \in B$,

$$g_{n_k}(\zeta) \rightarrow \rho'(\zeta + \alpha') (1 + \bar{\alpha}'\zeta)^{-1} = \rho'\alpha'$$

unless $1 + \bar{\alpha}'\zeta = 0$, or equivalently, unless $\zeta = -\alpha'$. We see from (1.4) that unless $E^*(Z_0)$ assigns positive measure to the point $-\alpha'$, $E^*(Z_{-1} | X_{-1}, X_{-2}, \dots)$ is concentrated on $\rho'\alpha'$. It is not difficult to show that the first alternative cannot occur. That is, if $E^*(Z_0)$ assigns positive measure to a single point of B , then since $\mu * E^*(Z_0) = E^*(Z_0)$ we can see that μ would be concentrated on a subgroup of G taking a finite point set of B into itself. Therefore, the measure $E^*(Z_{-1} | X_{-1}, X_{-2}, \dots)$ is concentrated on $\rho'\alpha'$, and by (1.4), $g_n E^*(Z_0)$ converges to the point measure at $\rho'\alpha'$. Since $E^*(Z_0)$ is not a point measure, this will happen only if for the entire sequence $\{g_n\}$, $|\alpha_n| \rightarrow 1$ and $\rho_n \alpha_n \rightarrow \rho'\alpha'$. But then we will have for every z with $|z| < 1$,

$$g_n(z) = \rho_n(z + \alpha_n)(1 + \bar{\alpha}_nz)^{-1} \rightarrow \rho'\alpha'.$$

This proves the theorem.

Theorem 1.3 may be reformulated for the half-plane $\text{Im } z > 0$. We then have,

COROLLARY. *Let $\{(u_n, v_n)\}$ be a sequence of independent pairs of random real numbers satisfying $u_n v_n < 0$, and such that all pairs have the same distribution which we assume is such that the matrices*

$$(-u_n v_n)^{-1/2} \begin{pmatrix} 0 & u_n \\ v_n & 1 \end{pmatrix}$$

are not concentrated in a proper subgroup of $SL(2, R)$. Then with probability 1, the continued fractions

$$\frac{u_1}{v_1 + \frac{u_2}{v_2 + \dots + \frac{u_n}{v_n + z}}}$$

converge whenever z is not real, and the limit is independent of z .

2. Application of the Poisson formula.

2.1. CRITERION FOR EXISTENCE OF A STATIONARY MEASURE. In this section we shall assume that G is a semi-simple group and that μ is an absolutely continuous probability measure on G . With these assumptions the conclusions of [8] are valid; these are summarized in §0.4. In particular, to the measure μ there corresponds a homogeneous G -space Π_μ which is one of a finite number of compact homogeneous covering spaces of the maximal boundary $B(G)$, and there is a measure ν on Π_μ satisfying $\mu * \nu = \nu$ which plays a distinguished role in the representation formula for μ -harmonic functions on G . In our present terminology, the measure ν is a stationary measure for μ , and it consequently generates a stationary μ -process.

DEFINITION 2.1. The stationary μ -process on Π_μ with initial distribution ν , where ν is the measure occurring in the Poisson representation formula (0.2), is called the *distinguished μ -process*.

We next observe that, conversely, beginning with a stationary measure for μ on any G -space, we can generate μ -harmonic functions. Namely, suppose that M is a G -space and π a stationary measure for μ on M . Set

$$(2.1) \quad f(g) = \int_M \phi(gx) d\pi(x)$$

where ϕ is a bounded borel measurable function on M . Then

$$(2.2) \quad \begin{aligned} \int_G f(gg') d\mu(g') &= \int_{G \times M} \phi(gg'x) d\mu(g') d\pi(x) = \int_M \phi(gy) d\mu * \pi(y) \\ &= \int_M \phi(gy) d\pi(y) = f(g). \end{aligned}$$

From this observation we can deduce

THEOREM 2.1. *Let M be a G -space. There is a one-to-one correspondence between the stationary measures of μ on M and the measures on M invariant*

under H_μ (recall $\Pi_\mu = G/H_\mu$). In particular, there exists a stationary μ -process on M if and only if H_μ has an invariant probability measure on M , and this occurs if and only if the subgroup $H(G)$ has an invariant probability measure on M .

Before proving this theorem we introduce some notation which will be useful. In general, let L be a closed subgroup of G , ω a bounded measure on G/L , let M be a G -space and suppose that λ is a measure on M that is invariant under L . If σ is the natural map $G \rightarrow G/L$ we can find some measure $\tilde{\omega}$ on G with $\sigma(\tilde{\omega}) = \omega$. Then the measure $\tilde{\omega} * \lambda$ is defined on M , and we claim that this measure on M is independent of the choice of $\tilde{\omega}$ for a given ω . Namely

$$(2.3) \quad \int_M f(x) d\tilde{\omega} * \lambda(x) = \int_{G \times M} f(gx) d\tilde{\omega}(g) d\lambda(x) = \int_G F(g) d\tilde{\omega}(g)$$

where $F(g) = \int f(gx) d\lambda(x)$. Since λ is invariant under L , it follows that $F(gl) = f(g)$ for $l \in L$, so that $F = F' \circ \sigma$ for some function F' on G/L . The last integral of (2.3) is then $\int_{G/L} F'(y) d\sigma(\tilde{\omega})(y) = \int_{G/L} F(y) d\omega(y)$. It follows that the measure $\tilde{\omega} * \lambda$ depends only on ω and λ whenever λ is invariant under L .

DEFINITION 2.2. We shall denote the measure $\tilde{\omega} * \lambda$, in case λ is invariant under L , simply as $\omega * \lambda$.

Theorem 2.1 will now follow from

LEMMA 2.1. If M is a G -space and λ is a probability measure on M invariant under H_μ , then $\nu * \lambda$ is a stationary measure for μ on M , where ν is the stationary measure on $\Pi_\mu = G/H_\mu$ of the distinguished μ -process. Conversely, every stationary measure ν' on M has the form $\nu * \lambda$ where λ is invariant under H_μ and the correspondence $\nu' \leftrightarrow \lambda$ is one-to-one.

Proof. For the first statement we observe that associativity holds for the convolutions of Definition 2.2 so that $\mu * (\nu * \lambda) = (\mu * \nu) * \lambda$. Since $\mu * \nu = \nu$ we have $\mu * (\nu * \lambda) = \nu * \lambda$, so that $\nu' = \nu * \lambda$ is stationary. To justify the associativity, we simply note that if $\sigma(\tilde{\omega}) = \omega$ then $\sigma(\mu * \tilde{\omega}) = \mu * \omega$ since σ is equivariant. We turn now to the converse statement, so let ν' be a stationary measure on M for μ . Let $M_c = M \cup \{\infty\}$ be the one point compactification of the locally compact M , if M is not compact already; otherwise, set $M_c = M$. In any case, letting $g \infty = \infty$ for $g \in G$, M_c will again be a G -space. Let $\phi(x)$ be a continuous function on M_c and form the μ -harmonic function (see (2.1))

$$(2.4) \quad \tilde{\phi}(g) = \int_{M_c} \phi(gx) d\nu'(x).$$

Since ϕ is continuous on M_c it follows that ϕ is left-uniformly continuous on G . By §0.4 we then have

$$(2.5) \quad \tilde{\phi}(g) = \int_{\Pi_\mu} \hat{\phi}(gy) dv(y)$$

where $\hat{\phi}$ is continuous on Π_μ . Each point $y \in \Pi_\mu$ induces a positive linear functional on the class of continuous functions on M_c by sending $\phi \rightarrow \hat{\phi}(y)$. This functional determines a probability measure λ_y on M_c and we will have $\hat{\phi}(y) = \lambda_y(\phi)$. In general, if ψ is a function on a G -space and $\gamma \in G$, let ψ^γ denote the function $\psi^\gamma(p) = \psi(\gamma p)$. Returning to the foregoing situation we observe that by (2.4), $(\phi^\gamma)^\sim = \tilde{\phi}^\gamma$. It follows by the uniqueness of the boundary function in (2.5), that $(\phi^\gamma)^\wedge = \hat{\phi}^\gamma$. From this we conclude that

$$(2.6) \quad \gamma \lambda_y = \lambda_{\gamma y}, \quad \gamma \in G, y \in \Pi_\mu.$$

Let y_0 be the point of Π_μ whose stability group is H_μ . (y_0 corresponds to the identity coset in G/H_μ .) Set $\lambda = \lambda_{y_0}$. From (2.6) we see that H_μ leaves λ invariant. Moreover, by (2.4) and (2.5),

$$v'(\phi) = \tilde{\phi}(e) = \int_{\Pi_\mu} \hat{\phi}(y) dv(y).$$

Now let \tilde{v} be a measure on G mapping onto v under $G \rightarrow G/H_\mu$. We then have

$$\begin{aligned} v'(\phi) &= \int_{\Pi_\mu} \hat{\phi}(y) dv(y) = \int_G \hat{\phi}(gy_0) d\tilde{v}(g) \\ &= \int_G \lambda_{gy_0}(\phi) d\tilde{v}(g) = \int_G g\lambda(\phi) d\tilde{v}(g) = \tilde{v} * \lambda(\phi). \end{aligned}$$

By Definition 2.2, $\tilde{v} * \lambda = v * \lambda$ and this proves that $v' = v * \lambda$. Note that if M is not compact, λ is nevertheless concentrated on M ($\lambda(\infty) = 0$), for if not, v' would have to assign positive measure to the point at ∞ and we are assuming that v' is a measure on M .

We must still prove that the correspondence $v' \leftrightarrow \lambda$ is one-to-one. What must be shown is that $v * \lambda_1 = v * \lambda_2$ for two measures λ_1 and λ_2 on M , implies $\lambda_1 = \lambda_2$. For this we note that if $g_n v \rightarrow v_0$, then $g_n(v * \lambda) \rightarrow v_0 * \lambda$; this is an easy consequence of Definition 2.2. By Lemma 5.4 of [8], if $\{X_n\}$ is a sequence of independent G -valued random variables with distribution μ , then with probability 1, $X_1 X_2 \cdots X_n v$ converges to a point measure on Π_μ . All we need is that there is some sequence $\{g_n\}$ in G with $g_n v$ converging to a point measure, say $g_n v \rightarrow \delta_y$, $y \in \Pi_\mu$. Therefore if $v * \lambda_1 = v * \lambda_2$ it follows that $\delta_y * \lambda_1 = \delta_y * \lambda_2$. Now if $y = gH$ then $\delta_y * \lambda_i = g\lambda_i$ so that it follows that $g\lambda_1 = g\lambda_2$, whence $\lambda_1 = \lambda_2$. This proves the lemma.

Proof of Theorem 2.1. All that remains to the proof of Theorem 2.1 is showing that H_μ has an invariant measure on M , if and only if $H(G)$ leaves invariant a measure on M . This follows from the fact that H_μ is always a subgroup of

finite index in $H(G)$ (see [8, Chapter V]). Clearly, if $H(G)$ has an invariant probability measure on M so does $H_\mu \subset H(G)$. Suppose conversely that λ is a probability measure fixed under H_μ . Then the set of measures $\{h\lambda, h \in H(G)\}$ will be finite. The average of these measures is evidently an $H(G)$ -invariant measure and so this proves our assertion.

An important consequence of Theorem 2.1 is that the existence of a stationary measure on M for μ depends only on M and not on μ . For, the group $H(G)$ occurring in Theorem 2.1 does not depend on the measure μ .

2.2. APPLICATIONS OF THE CRITERION. In the remainder of this section we will apply the criterion of Theorem 2.1 to the existence and uniqueness problem for stationary measures.

LEMMA 2.2. *If $L_1 \subset L_2$ are subgroups of G with L_2/L_1 compact, then there is a stationary measure for μ on G/L_1 if and only if there is one on G/L_2 .*

Proof. Let σ denote the natural map $G/L_1 \rightarrow G/L_2$. Since σ is equivariant it follows that if ν_1 is a stationary measure on G/L_1 , then $\sigma(\nu_1)$ is a stationary measure on G/L_2 . For the converse, let ν_2 be a stationary measure on G/L_2 and let Q denote the set of all probability measures ν_1 on G/L_1 satisfying $\sigma(\nu_1) = \nu_2$. It is easy to see by the Hahn-Banach extension theorem [1], that Q is nonempty. Q is clearly convex; we claim it is also compact. Note that under the hypothesis that L_2/L_1 is compact, if Δ is a compact subset of G/L_2 , then $\sigma^{-1}(\Delta)$ is a compact subset of G/L_1 . Now suppose that $\nu_1^{(n)}$ is a sequence of measures in Q , and suppose $\varepsilon > 0$ given. Let Δ be a compact subset of G/L_2 with $\nu_2(\Delta) > 1 - \varepsilon$; then $\nu_1^{(n)}(\sigma^{-1}(\Delta)) > 1 - \varepsilon$. It follows that if $\nu_1^{(n)}$ converges to a measure on G/L_1 , that measure assigns a value $\geq 1 - \varepsilon$ to $\sigma^{-1}(\Delta)$. This being the case for every ε , we see that the limit measure is again a probability measure. Hence Q is compact. Define the mapping $\tau: Q \rightarrow Q$ by setting $\tau(\nu_1) = \mu * \nu_1$. Since $\sigma(\mu * \nu_1) = \mu * \sigma(\nu_1) = \mu * \nu_2 = \nu_2$ it follows that τ actually does take Q into itself. It is also an affine map and since Q is compact and convex we conclude [7] that there is a measure $\nu_1 \in Q$ with $\tau\nu_1 = \nu_1$ or $\mu * \nu_1 = \nu_1$. This completes the proof of the lemma.

COROLLARY. *Let M_1 be a compact G -space and let $\sigma: M_1 \rightarrow M_2$ be an equivariant map. If ν_2 is a stationary measure for μ on M_2 then $\nu_2 = \sigma(\nu_1)$ where ν_1 is a stationary measure for μ on M_1 . Here σ is onto.*

Proof. With Q defined as before, Q is again nonempty and convex. This time the compactness of Q follows from the compactness of M_1 . The remainder of the proof is then the same as for the lemma.

THEOREM 2.2. *If M is a boundary of G (see §0.4) then the G -space $M \times M$ supports a unique stationary measure for μ and this measure is concentrated on the diagonal of $M \times M$.*

($M \times M$ is the space of all ordered pairs of points of M ; it is a G -space by setting $g(x, y) = (gx, gy)$. It is not homogeneous because the diagonal which consists of pairs $\{(x, x)\}$ remains invariant under G . The theorem implies, in particular, that the space of distinct pairs of points of M does not carry a stationary measure.)

Proof. Suppose the theorem has been demonstrated in case M is the maximal boundary $B(G)$. According to §0.4, every boundary is an equivariant image of $B(G)$ and so $M \times M$ is an equivariant image of $B(G) \times B(G)$. Applying the corollary to Lemma 2.2 to a stationary measure on $M \times M$, it would then follow that such a measure is concentrated on the diagonal of $M \times M$, and is unique. We therefore restrict our attention to the case $M = B(G)$. By Lemma 2.1, it will suffice then to prove that there is a unique measure λ^* on $B(G) \times B(G)$ invariant under H_μ , and that this measure is concentrated on the diagonal. Let ϕ_1 and ϕ_2 denote the two projections from $B(G) \times B(G)$ to $B(G)$. The ϕ_i are equivariant maps. Since $H_\mu \subset H(G)$ and the latter is the stability group of some point of $B(G)$, it follows that H_μ leaves fixed a point $x_0 \in B(G)$, x_0 representing the identity coset in $G/H(G)$.

Let $\phi_i(\lambda^*) = \lambda_i$. λ_i is a measure on $B(G)$ invariant under H_μ . Since $[H(G):H_\mu]$ is finite, the set of measures $\{h\lambda_i, h \in H(G)\}$ is finite. Let $\bar{\lambda}$ be the average of the measures in this set. Then $\bar{\lambda}$ is invariant under $H(G)$. Since $G/H(G)$ is compact it follows that the set of measures $\{g\bar{\lambda}, g \in G\}$ is compact. But $B(G)$ is a boundary, and so some $g_n\bar{\lambda}$ converges to a point measure. Since a limit of these is again of the form $g\bar{\lambda}$, it follows that $\bar{\lambda}$ is a point measure, $\bar{\lambda} = \delta_x$. If $x \neq x_0$ then $\frac{1}{2}\delta_{x_0} + \frac{1}{2}\delta_x$ would be invariant under $H(G)$ and would not be a point measure. The preceding argument shows that this cannot happen since $B(G)$ is a boundary. Therefore $x = x_0$ and $\bar{\lambda}$ is concentrated at x_0 . By the definition of $\bar{\lambda}$ it follows that λ_i must also be concentrated at x_0 and so $\lambda_i = \delta_{x_0}$. The only measure on $B(G) \times B(G)$ satisfying $\phi_i(\lambda^*) = \delta_{x_0}$, $i = 1, 2$, is the measure concentrated at the diagonal point (x_0, x_0) , so this proves the theorem.

The next theorem will be very important for the derivation of the laws of large numbers to be discussed in Part II.

THEOREM 2.3. *If M is a boundary of G , there is one and only one stationary measure for μ on M .*

Proof. The existence of a stationary measure follows from Lemma 1.2. Suppose now that ν_1 and ν_2 are two stationary measures for μ on M . Let ϕ_1 and ϕ_2 be the two projections of $M \times M$ on M and observe that if π is a measure on $M \times M$ then $\mu * \phi_i(\pi) = \phi_i(\mu * \pi)$. If we take $\pi = \nu_1 \times \nu_2$ then $\phi_i(\pi) = \nu_i$ so we also have $\phi_i(\mu^n * \pi) = \mu^n * \nu_i = \nu_i$ by stationarity. Now some sequence of averages of $\mu^n * \pi$ converges to a stationary probability measure π' on $M \times M$ and the latter still satisfies $\phi_i(\pi') = \nu_i$. By Theorem 2.2, π' has support on the diagonal of $M \times M$. But then $\phi_1(\pi') = \phi_2(\pi')$ and $\nu_1 = \nu_2$ as was to be shown.

REMARK. We could also have demonstrated Theorem 2.3 without Theorem 2.2 by using Lemma 2.1. The proof given shows, in addition, the connection between the uniqueness of the stationary measure on M and the nonexistence of a stationary measure off the diagonal of $M \times M$.

2.3. EXISTENCE OF STATIONARY MEASURES. We shall next apply Theorem 2.1 to show that certain homogeneous spaces of G cannot support stationary measures.

DEFINITION 2.3. A topological group is said to have the *fixed-point property* if whenever it acts continuously on a compact convex set by affine transformation, then there necessarily exists a fixed point of the group in this set.

This notion was used frequently in [8]. We found there that a connected Lie group has the fixed-point property if and only if it has a normal solvable subgroup such that the quotient group is compact.

LEMMA 2.3. *If L is any connected Lie group then L possesses a subgroup H with the fixed-point property such that L/H is compact.*

Proof. This follows from Theorem 1.10 of [8] where it is shown that the minimal B -subgroups of a connected Lie group have the fixed-point property. The quotient of a Lie group by a B -subgroup is a boundary and this is compact by definition.

LEMMA 2.4. *Let H be a closed subgroup of G with the fixed-point property. Then μ admits a stationary measure on G/H only if H is contained in a minimal B -subgroup $H(G)$, and if $H(G)$ preserves a probability measure on $H(G)/H$.*

Proof. Suppose μ admits a stationary measure on G/H . Then by Theorem 2.1, $H(G)$ leaves fixed a measure λ on G/H . Now H has the fixed-point property, and so H leaves fixed a measure π on the compact space $G/H(G)$. In accordance with Definition 2.2 we may form the convolution $\lambda * \pi$, λ being a measure on G/H and the measure π being invariant under H . For $h \in H(G)$, $h\lambda = \lambda$, so $h(\lambda * \pi) = \lambda * \pi$, and $\lambda * \pi$ is an $H(G)$ -invariant measure on $G/H(G)$. Since the latter space is a boundary, it follows that the only such measure is the one concentrated on the identity coset, $H(G)$, of $G/H(G)$ (see the proof of Theorem 2.2). Let x_0 denote this point of $G/H(G)$. Since $\lambda * \pi$ is concentrated on one point, the same is, *a fortiori*, true of π . Thus H keeps some point, say γx_0 , fixed, or $H\gamma H(G) \subset \gamma H(G)$ so that $H \subset \gamma H(G)\gamma^{-1}$. Since $\gamma H(G)\gamma^{-1}$ is again a minimal B -subgroup of G we may suppose to begin with that $H(G)$ was chosen as this subgroup. We may thus suppose that $H \subset H(G)$.

Consider now the measure λ left fixed by $H(G)$ on G/H . Since $H \subset H(G)$ there is a natural equivariant map $G/H \rightarrow G/H(G)$. The measure λ is mapped into a measure on $G/H(G)$ invariant under $H(G)$. The image of λ is thus concentrated on the coset $H(G)$ of $G/H(G)$. This means that in G/H , the measure λ was concentrated on $H(G)/H \subset G/H$. It follows that $H(G)$ preserves a probability measure on $H(G)/H$. We have thus established the assertions of the lemma.

THEOREM 2.4. *In case $H(G)$ is solvable and L is a closed subgroup of G , then μ admits a stationary measure on G/L only if either G/L is compact or L has infinitely many connected components.*

Proof. By Lemma 2.3, if L has finitely many connected components, it contains a subgroup H with the fixed-point property such that L/H is compact. In that case, by Lemma 2.2, μ will admit a stationary measure on G/L only if it does so on G/H . We may now apply the foregoing lemma to conclude that, in this case, H is contained in $H(G)$, a minimal B -subgroup of G , and $H(G)$ preserves a finite measure on $H(G)/H$. But Mostow has shown that a solvable group will preserve a finite measure on a homogeneous space only if the space is compact [14]. Thus we would find $H(G)/H$ compact. Since $G/H(G)$ is compact, G/H would be compact and, a fortiori, G/L is compact. This proves the theorem.

REMARK 1. The hypothesis that $H(G)$ be solvable is verified in the cases that $G = \text{SL}(m, R)$ or $\text{SL}(m, C)$. In fact, in [8], the subgroups $H(G)$ are determined explicitly in these cases, and it is found that $H(G)$ may in both cases be taken as the subgroup of triangular matrices.

REMARK 2. If L has infinitely many connected components, G/L may carry a stationary measure without being compact. For example, let $G = \text{SL}(m, R)$ and let L be the subgroup of matrices with integer entries. It is known [3] that G/L has a finite invariant measure which will also be a stationary measure. On the other hand, G/L is not compact.

The next theorem shows the significance of the nonexistence of a stationary measure for the random walk induced by μ on the space. For the proof we require a preliminary lemma.

LEMMA 2.5. *Let M be a locally compact space and $\{\mu_x\}$ a random walk on M . Define*

$$(2.7) \quad \tau f(x) = \int f(y) d\mu_x(y) = \mu_x(f),$$

where $f(x)$ is a bounded borel measurable function on M , and set $F(x) = \tau f(x) - f(x)$. If $\{Z_n\}$ is a Markov process with transition probability measures $\{\mu_x\}$, then with probability 1,

$$(2.8) \quad \frac{1}{n} \sum_0^{n-1} F(Z_k) \rightarrow 0$$

as $n \rightarrow \infty$.

Proof. Set $W_{n+1} = \sum_0^n (\tau f(Z_k) - f(Z_{k+1})) / (k+1)$.

$$\begin{aligned}
 E(W_{n+1} | Z_1, \dots, Z_n) &= W_n + \frac{1}{n+1} E(\tau f(Z_n) - f(Z_{n+1}) | Z_1, \dots, Z_n) \\
 &= W_n + \frac{1}{n+1} \tau f(Z_n) - \frac{1}{n+1} E(f(Z_{n+1}) | Z_1, \dots, Z_n) = W_n,
 \end{aligned}$$

since it follows from (2.7) that $E(f(Z_{n+1}) | Z_1, \dots, Z_n) = \tau f(Z_n)$. Hence $\{W_n\}$ forms a martingale. Since $f(x)$ is bounded, say $|f| < c$, $E(W_{n+1}^2) \leq 4c^2 \sum_1^\infty (1/k^2)$. This implies that W_n converges, as $n \rightarrow \infty$, with probability 1 [5, Chapter VII, Theorem 4.1]. By the Kronecker's lemma, it follows that

$$\sum_0^n \{ \tau f(Z_k) - f(Z_{k+1}) \} = o(n),$$

and rearranging terms, we have the assertion of our lemma.

THEOREM 2.5. *Let M be a G -space that does not admit a stationary measure for μ . Let $\{Z_n\}$ be any μ -process on M , and let Δ be any compact subset of M . If $n(k)$ denotes the index n for which Z_n belongs to Δ for the k th time for $n \geq 0$ ($n(k) \leq \infty$), then $n(k)/k \rightarrow \infty$ with probability 1.*

In other words, the process drifts to ∞ in the sense that it returns to any fixed compact set less and less frequently.

Proof. The conclusion of the theorem is equivalent to the statement that if $\psi(x)$ is any continuous function with compact support on M then with probability 1,

$$(2.9) \quad \frac{1}{n} \sum_0^{n-1} \psi(Z_n) \rightarrow 0.$$

Consider the random walk on M induced by μ (§1.1) and apply Lemma 2.5. We will in fact show that (2.9) will be valid under our hypotheses whenever (2.8) is valid. So let $\{Z_n\}$ be a sample sequence for which (2.8) is valid for all continuous f vanishing at ∞ . (It would suffice to require this for a dense set of f and so this will be the case for almost all sample sequences.) Suppose that there is a ψ' with compact support for which (2.9) is not valid; so for some subsequence $\{n_i\}$, $n_i^{-1} \sum_0^{n_i-1} \psi'(Z_n) \rightarrow \alpha \neq 0$. By refining the sequence n_i sufficiently we may assume that $n_i^{-1} \sum_0^{n_i-1} \psi(Z_n)$ converges as $i \rightarrow \infty$ for any countable set of bounded functions ψ , and hence for a dense set of continuous functions vanishing at ∞ , and hence for all continuous functions vanishing at ∞ . This defines a linear functional on the class of continuous functions on M vanishing at ∞ and since the functional is positive and bounded, it corresponds to a non-negative bounded measure π on M . Since $\pi(\psi') \neq 0$, the measure π does not vanish. On the other hand, we have $\pi(F) = 0$ when $F = \tau f - f$. Thus

$$\int \tau f(x) d\pi(x) = \int f(x) d\pi(x)$$

or

$$\int \int f(gx) d\mu(g) d\pi(x) = \int f(x) d\pi(x)$$

or $\mu * \pi(f) = \pi(f)$, whence $\mu * \pi = \pi$. Since $\pi \neq 0$, this contradicts the hypothesis that μ does not admit a stationary measure on M . This proves the theorem.

In Theorem 2.2 we saw that if M is a boundary of G , then the only stationary measure on $M \times M$ is concentrated on the diagonal. Denoting the diagonal by D , we conclude that the G -space $M \times M - D$ does not support a stationary measure. Suppose \mathcal{U} is a neighborhood of D in $M \times M$; the complement of \mathcal{U} is then a compact subset of $M \times M - D$. Applying Theorem 2.5 we find that any μ -process on $M \times M$ stays in \mathcal{U} with the exception of a set of indices of density 0. More specifically, let $\{Z_n\}$ be a μ -process on $M \times M$ obtained by setting $Z_0 = (x_1, x_2)$ where the x_i are distinct points of M and letting $Z_n = X_n \cdots X_1 Z_0$, where the X_n are the associated G -valued random variables. We then have

COROLLARY. *Let $\{X_n\}$ be a sequence of G -valued independent random variables with distribution μ , let M be a boundary of G and $d(\cdot, \cdot)$ a metric on M . If x_1 and x_2 are any two points of M , the sequences $\{X_n X_{n-1} \cdots X_1 x_i\}$ for $i=1, 2$ approach one another with probability one in the sense that $d(X_n X_{n-1} \cdots X_1 x_1, X_n X_{n-1} \cdots X_1 x_2) < \varepsilon$ for any $\varepsilon > 0$, for all n outside a set of density 0.*

2.4. UNIQUENESS OF STATIONARY MEASURES. According to Theorem 2.1, the stationary measures for μ on a G -space are in correspondence with the measures on the space invariant under the group H_μ . Previously, we used this to draw conclusions about the stationary measures for μ using properties of the group H_μ , but we can also reverse the direction of this argument. Thus, suppose that μ_1 and μ_2 are two measures with the same Poisson space, $\Pi_{\mu_1} = \Pi_{\mu_2}$, and suppose that one can show that for μ_1 there exists a unique stationary measure on a certain G -space. Then $H_{\mu_1} = H_{\mu_2}$ has a unique invariant measure on the space, from which it follows that μ_2 also has a unique stationary measure. We will use this reasoning to show that if $H_\mu = H(G)$, then μ has a unique stationary measure on any compact homogeneous space of G .

THEOREM 2.6. *Let M be any compact homogeneous space of G . Then $H(G)$ leaves fixed a unique probability measure on M .*

Proof. To prove the theorem we shall show the existence of an absolutely continuous probability measure μ on G with $H_\mu = H(G)$ and such that μ has a unique stationary measure on M . Suppose momentarily that μ has been chosen. Let Q denote the compact, convex set of probability measures on M which are stationary for μ . To show that Q consists of a single member it will suffice by the Krein-Milman theorem, to show that Q contains only one extremal. Let us

first observe that if ν is an extremal of \mathcal{Q} , the corresponding stationary μ -process on M is ergodic. For if the μ -process $\{Z_n\}$ were not ergodic, then by [6, p. 460], there would be a nontrivial stationary variable which is a function of Z_0 ; say $f(Z_0) = f(Z_1)$. We can take f between 0 and 1, and form the measure $d\nu' = f d\nu$. We claim that ν' is again stationary:

$$\begin{aligned} \mu * \nu'(\phi) &= \iint \phi(gx)f(x)d\mu(g)d\nu(x) = E(\phi(Z_1)f(Z_0)) \\ &= E(\phi(Z_1)f(Z_1)) = \int \phi(x)f(x)d\nu(x) = \nu'(\phi). \end{aligned}$$

But then $\nu = \nu' + (\nu - \nu')$ is a nontrivial decomposition of ν and it cannot be extremal.

From this it follows that if ν' and ν'' are distinct extremal measures in \mathcal{Q} , they must be mutually singular. For, let $\{X_n\}$ be a sequence of G -valued random variables, independent and with distribution μ . By the ergodic theorem (for the stationary μ -process determined by ν'), we have for almost all x with respect to the measure ν' :

$$\lim_{N \rightarrow \infty} N^{-1} \sum_1^N f(X_n \cdots X_1 x) = \nu'(f)$$

with probability 1. The same assertion holds for $\nu''(f)$ for almost all x with respect to ν'' . Hence, unless $\nu' = \nu''$, these two must be mutually singular. As a consequence we find that if we wish to show that μ has a unique stationary measure on M , it suffices to show that no two stationary measures for μ can be mutually singular. We now proceed to construct a measure μ for which this will be the case.

If $S \subset G$, let m_S denote the Haar measure of G restricted to S : $m_S(\Delta) = m(\Delta \cap S)$ where m denotes the Haar measure on G , (which we suppose fixed once and for all). Let x_0 be a point of M . Since M is compact we can find a compact set $J \subset G$ such that $Jx_0 = M$. Now let $A \subset G$ be a compact set with $m(A) > 0$. Set $B = AJ^{-1}$ so that for any $g \in J$, $Bg \supset A$. We can suppose that A is large enough that B contains an open neighborhood of the identity of G . Now set $\mu = m(B)^{-1}m_B$. We claim that if x is any point in M then the measure $\mu * \delta_x$ dominates a fixed measure, namely, $m(B)^{-1}m_A * \delta_{x_0}$. For if x is any point in M , there is a $g \in J$ with $x = gx_0$. Then $m_B * \delta_x = m_{Bg} * \delta_{x_0}$. Now since $Bg \supset A$, for $g \in J$, $m_{Bg} \geq m_A$ and $m_B * \delta_x \geq m_A * \delta_{x_0}$. We thus have

$$(2.10) \quad \mu * \delta_x \geq m(B)^{-1}m_A * \delta_{x_0} = \pi_0.$$

Since $m(A) > 0$, the measure on the right is not trivial. Since (2.10) holds for all $x \in M$, we find by integration that

$$(2.11) \quad \mu * \pi \geq \pi_0$$

for any probability measure π . Consequently, if π is a stationary measure for μ , $\pi \geq \pi_0$. It follows that no two stationary measures for this μ can be mutually singular; hence, μ has a unique stationary measure on M . To complete the proof of the theorem, we remark that by Theorem 5.3 of [8], since the support of μ contains a neighborhood of the identity, the Poisson space of μ coincides with $B(G)$.

Combining the foregoing with Theorem 2.1 we have

COROLLARY. *If μ is a measure whose Poisson space coincides with $B(G)$, then on every compact homogeneous space of G , there exists a unique stationary measure for μ .*

We might mention a connection between Theorem 2.6 and some problems of ergodic theory. Let $G = \text{SL}(2, \mathbb{R})$ and let Γ be a discrete subgroup with G/Γ compact. The "horocycle flow" on G/Γ is the flow obtained by left multiplication on G/Γ with the group $\{g_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}\}$. It is known that this flow is ergodic with respect to the unique G -invariant probability measure on G/Γ , and that the flow is minimal. (See *Flows on homogeneous spaces*, by L. Auslander, L. Green, and F. Hahn, Annals of Mathematics Studies Number 53.)

Both of these would result simultaneously if it were known that the G -invariant measure on G/Γ is the only probability measure invariant under the subgroup $\{g_t\}$. Theorem 2.6 implies a weaker result, namely, that this measure is the only one left invariant by the subgroup $H(G)$ which, in this case, can be taken as

$$H(G) = \begin{pmatrix} s & t \\ 0 & s^{-1} \end{pmatrix},$$

a group containing the subgroup $\{g_t\}$.

3. Proper μ -processes.

3.1 DEFINITION 3.1. Let M be a G -space, $\{Z_n, -\infty < n < \infty\}$ a stationary μ -process and let $\{X_n, -\infty < n < \infty\}$ be the associated G -valued random variables, so that $Z_{n+1} = X_{n+1}Z_n$. We say that $\{Z_n\}$ is a *proper μ -process* if each Z_n is measurable over the σ -field generated by $\{X_n, X_{n-1}, \dots\}$.

We have $Z_n = X_n X_{n-1} \cdots X_{n-k+1} Z_{n-k}$, so that Z_n is a function of the "past" X_m and any Z_{n-k} . If the effect of Z_{n-k} becomes negligible as $k \rightarrow \infty$ so that Z_n actually is determined by the X_m , $m \leq n$, then we say that the process is proper. For example, the corollary to Theorem 2.5 suggests that this might be the case when Z_n is a stationary μ -process on a boundary of G . We shall see that in fact this is so, and we will moreover be able to determine all the proper μ -processes for a given μ .

Let Π_μ be the Poisson space of μ , and ν the measure on Π_μ for which the Poisson formula was valid. This measure induces the *distinguished μ -process* (Definition 2.1).

LEMMA 3.1. *The distinguished μ -process is a proper process.*

Proof. According to [8], Lemma 5.4, $X_0X_{-1} \cdots X_{-k}v$ converges to a point measure on Π_μ with probability 1 as $k \rightarrow \infty$. By Lemma 1.3, denoting the distinguished μ -process by $\{Z_n\}$, we have $E^*(Z_0 | X_0, X_{-1} \cdots) = \lim_{k \rightarrow \infty} X_0X_{-1} \cdots X_{-k}v$ which is a point measure. It follows that Z_0 is measurable with respect to $\{X_0, X_{-1}, \dots\}$ and so $\{Z_n\}$ is a proper process.

If M is a boundary of G then by Theorem 2.3, M carries a unique stationary measure for μ . In other words there is a unique stationary μ -process on M .

THEOREM 3.1. *The unique stationary μ -process on a boundary of G is a proper μ -process.*

Proof. Every boundary is an equivariant image of the maximal boundary $B(G)$. On the other hand, each Π_μ is a covering space of $B(G)$ and the projection $\Pi_\mu \rightarrow B(G)$ is equivariant [8, Chapter V]. Thus every boundary is an equivariant image of Π_μ , say $\sigma: \Pi_\mu \rightarrow M$. If $\{Z_n\}$ is a stationary μ -process on Π_μ it follows that $\sigma(Z_n)$ is a stationary μ -process on M ; the latter being unique, we conclude that the stationary μ -process on M is an equivariant image of the distinguished μ -process. Since the latter is a proper μ -process, it follows readily that any equivariant image is one as well.

THEOREM 3.2. *Every proper μ -process is an equivariant image of the distinguished μ -process.*

Proof. Let $\{Z_n\}$ denote the distinguished μ -process and $\{W_n\}$ an arbitrary proper μ -process, and let M be the state space of $\{W_n\}$. If the stationary measure for $\{W_n\}$ is v' , then by Lemma 2.1, $v' = v * \lambda$ where λ is an H_μ -invariant measure on M . Now $X_0X_{-1} \cdots X_{-n}v = X_0X_{-1} \cdots X_{-n}E^*(Z_0)$ converges to $E^*(Z_0 | X_0, X_{-1} \cdots)$ by Lemma 1.3, and the latter is a point measure since $\{Z_n\}$ is proper. We may thus write $X_0X_{-1} \cdots X_{-n}v \rightarrow \delta_{Z_0}$. It follows that $X_0X_{-1} \cdots X_{-n}v' \rightarrow \delta_{Z_0} * \lambda = g\lambda$ for some $g \in G$. But, if $\{W_n\}$ is a proper μ -process, then

$$X_0X_{-1} \cdots X_{-n}v' = X_0X_{-1} \cdots X_{-n}E^*(W_0) \rightarrow E^*(W_0 | X_0, X_{-1}, \dots) = \delta_{W_0}$$

so that $g\lambda = \delta_{W_0}$ and λ must be a point measure. We thus find that H_μ leaves fixed a point $x \in M$, $\lambda = \delta_x$, and $v' = v * \delta_x$. From the expression for v' we see that its support is contained in the set $Gx \subset M$. Now the set Gx is compact since $H_\mu x = x$, and so Gx is a continuous image of $G/H_\mu = \Pi_\mu$ which is compact. Thus $\{W_n\}$ takes place on an equivariant image of Π_μ . Finally we have

$$\delta_{W_0} = \delta_{Z_0} * \lambda$$

from which it follows that $\delta_{W_n} = \delta_{Z_n} * \lambda$ and so $\{W_n\}$ is an equivariant image

of $\{Z_n\}$. (For the definition of the convolution used here, see Definition 2.2.) This proves the theorem.

In particular, if μ is such that $\Pi_\mu = B(G)$, then a stationary μ -process is proper if and only if the state space is a boundary of G . In general, this notion provides a probabilistic characterization of the Poisson space Π_μ . It is the state space for a "maximal" μ -process which is a subprocess of $\{X_n\}$.

An illustration of a nonproper stationary μ -process is the following. Let Γ be a discrete subgroup of G such that G/Γ has finite invariant measure. The G -invariant probability measures π on G/Γ is clearly a stationary measure for any μ on G . The resulting process $\{W_n\}$ cannot be a proper process since

$$E^*(W_0 | X_0, X_{-1}, \dots) = \lim_{k \rightarrow \infty} X_0 X_{-1} \cdots X_{-k} E^*(W_0) = \lim_{k \rightarrow \infty} X_0 X_{-1} \cdots X_{-k} \pi = \pi.$$

4. Comparison with solvable groups.

4.1. We will show by a single example that a number of the foregoing results are not valid for Lie groups in general. In the case to be considered, G will be solvable. By [8], Corollary to Theorem 1.2, the boundary of a solvable group is trivial. For semi-simple groups the μ -harmonic functions on G and the proper μ -processes are related to a covering space Π_μ of the maximal boundary. In the case to be considered we will find that proper μ -processes exist and nontrivial μ -harmonic functions exist even though the maximal boundary is trivial.

Let variables X_n be given by

$$X_n = \begin{pmatrix} \xi_n & 1 \\ 0 & 1 \end{pmatrix}$$

where $\{\xi_n\}$ is a sequence of independent, identically distributed random variables satisfying $\xi_n > 0$ and $E(\log \xi_n) < 0$. The group G is the two-dimensional solvable group $G = \{ \begin{pmatrix} u & v \\ 0 & 1 \end{pmatrix}, u > 0 \}$. Consider the random power series $\sum_{k=0}^\infty \xi_n \xi_{n-1} \cdots \xi_{n-k} z^k$. Its radius of convergence is given by $\liminf_{k \rightarrow \infty} (\xi_n \xi_{n-1} \cdots \xi_{n-k})^{-1/k}$, and by the law of large numbers, since $E(\log \xi_n) < 0$, this is seen to be larger than 1 with probability 1. We may thus define variables u_n by

$$(4.1) \quad u_n = 1 + \xi_n + \xi_n \xi_{n-1} + \cdots + \xi_n \xi_{n-1} \cdots \xi_{n-k} + \cdots.$$

Then $u_{n+1} = 1 + \xi_{n+1} u_n$, or

$$\begin{pmatrix} u_{n+1} \\ 1 \end{pmatrix} = \begin{pmatrix} \xi_{n+1} & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_n \\ 1 \end{pmatrix}$$

so that if we set $Z_n = \begin{pmatrix} u_n \\ 1 \end{pmatrix}$, $\{Z_n\}$ forms a μ -process. Here the G -space M is the set of vectors $\begin{pmatrix} u \\ 1 \end{pmatrix}$. This shows that our criterion for the existence of a stationary measure cannot be carried over verbatim from the semi-simple case, since in this case $H(G) = G$ and there is no measure on M preserved by G . (Compare with Theorem 2.1.)

Moreover, it is immediately evident from (4.1) that $\{Z_n\}$ is a proper μ -process. Thus, for solvable groups, a proper μ -process may have a noncompact state space.

Finally, by means of the process Z_n , we can construct nontrivial μ -harmonic functions on G . Namely, let $\nu = E^*(Z_0)$ and for any bounded continuous function ϕ on M , set $\tilde{\phi}(g) = g\nu(\phi)$.

$$\begin{aligned} \int \tilde{\phi}(gg')d\mu(g') &= \int gg'\nu(\phi)d\mu(g') = g(\mu*\nu)(\phi) \\ &= g\nu(\phi) = \tilde{\phi}(g) \end{aligned}$$

so $\tilde{\phi}(g)$ is a μ -harmonic function. To see that $\tilde{\phi}(g)$ will in general be nontrivial observe that

$$\begin{aligned} \tilde{\phi}(X_0X_{-1}\cdots X_{-k}) &= X_0X_{-1}\cdots X_{-k}E^*(Z_0)(\phi) \rightarrow E^*(Z_0 | X_0, X_{-1}, \dots)(\phi) \\ &= E(\phi(Z_0) | X_0, X_{-1}, \dots). \end{aligned}$$

Since $\{Z_n\}$ is a proper μ -process, $\phi(Z_0)$ is measurable with respect to $\{X_n, n \leq 0\}$ so that $\tilde{\phi}(X_0X_{-1}\cdots X_{-k}) \rightarrow \phi(Z_0)$. Hence $\tilde{\phi}(g)$ will not always be constant, and there exist nontrivial μ -harmonic functions on G although it has a trivial boundary.

5. Brownian motion in symmetric spaces.

5.1. DEFINITION OF BROWNIAN MOTION. When the measure μ on G is properly restricted it is possible to associate with it a class of Markov processes other than the μ -processes considered up till now. Namely, let G be a noncompact semi-simple group (with finite center) and K a maximal compact subgroup of G . We will suppose that the measure μ is left-invariant under $K : k\mu = \mu$ for $k \in K$. If m_K denotes the normalized Haar measure on K we can write $\mu = m_K * \mu$; conversely, if $\mu = m_K * \mu'$, then μ will have this property. The space G/K is a symmetric Riemannian space and any noncompact symmetric Riemannian space has this form. We will denote the natural map $G \rightarrow G/K$ by σ . As usual, σ is an equivariant map. Now consider a "transposed" μ -process on G ; that is, we consider the random walk that assigns to each $g \in G$ the measure $g\mu$ on G . More precisely, if $\{X_n\}$ are G -valued independent random variables with distribution μ and W_0 is G -valued and independent of all the X_n , we set $W_n = W_0X_1X_2\cdots X_n$, so that we now multiply on the right rather than on the left. Here W_0 may have any distribution.

DEFINITION 5.1. For any absolutely continuous probability measure μ on G satisfying $k\mu = \mu$, $k \in K$, the process $U_n = \sigma(W_n)$ on G/K is called a *discrete brownian motion*.

Since σ does not commute with right multiplication by elements of G , it is not evident that $\{U_n\}$ is again a Markov process.

THEOREM 5.1. *A discrete brownian motion is a Markov process.*

Proof. Suppose $U_n = \sigma(W_n)$ is the brownian motion in question. To show that $\{U_n\}$ is a Markov process, it will suffice to show that if f is a bounded borel measurable function on G/K then $E(f(U_n) | W_{n-1}, W_{n-2}, \dots)$ is a function of U_{n-1} , since the σ -field generated by $\{W_j, j < n\}$ contains that generated by $\{U_j, j < n\}$. Let \tilde{f} denote the composition of f with $\sigma: f(U_n) = \tilde{f}(W_n)$. Then

$$\begin{aligned} E(f(U_n) | W_{n-2}, W_{n-1}, \dots) &= E(\tilde{f}(W_n) | W_{n-1}, W_{n-2}, \dots) \\ &= E(\tilde{f}(W_{n-1}X_n) | W_{n-1}, W_{n-2}, \dots) \end{aligned}$$

and since X_n is independent of $\{W_j, j < n\}$, the right side can be written

$$\int_G f(W_{n-1}g)d\mu(g) = \int_K \int_G f(W_{n-1}kg)dkd\mu(g)$$

with dk denoting the element of Haar measure on K . But $\int_K f(W_{n-1}kg)dk$ depends only on the coset $W_{n-1}K = \sigma(W_{n-1}) = U_{n-1}$ so that $E(f(U_n) | W_{n-1}, W_{n-2}, \dots)$ is a function of U_{n-1} as was to be shown.

The space G/K is a Riemannian space and one may speak of harmonic functions on G/K as those annihilated by the Laplace-Beltrami operator on G/K . In [8, Chapter IV], we showed that a bounded f is harmonic if and only if the corresponding function $\tilde{f}(g)$ on G satisfies

$$\int_K \tilde{f}(gkg')dk = \tilde{f}(g)$$

for all g' . Comparing this with the expression above for $E(f(U_n) | W_{n-1}, W_{n-2}, \dots)$ we conclude that

COROLLARY. *If $f(g)$ is a harmonic function on G/K and $\{U_n\}$ is a discrete brownian motion on G/K , then $E(f(U_n) | U_{n-1}, U_{n-2}, \dots) = f(U_{n-1})$ and $\{f(U_n)\}$ is a martingale.*

The continuous parameter brownian motions in G/K may be defined as processes $\{U(t)\}$ with continuous paths for which $\{U(n\tau)\}$ defines a discrete brownian motion for each $\tau > 0$. These are much more restricted than the discrete brownian motions and correspond to a choice of an invariant Laplacian operator on G/K . On the other hand, every measure of the form $m_k * \mu'$ generates a discrete brownian motion. One of the conclusions of [8] is that all these processes define the same set of harmonic functions.

5.2. LIMIT BEHAVIOR OF BROWNIAN MOTION. Let $B(G)$ be the maximal boundary of G ; it is shown in [8] that the maximal compact subgroup K is transitive on $B(G)$. It follows that there is a unique probability measure m on $B(G)$ which is K -invariant. Consider the set of measures $Gm = \{gm, g \in G\}$. Since $gkm = gm$

for $k \in K$ it can be seen that there is a map from G/K onto Gm . In Chapter II of [8] we showed that this map is actually one-to-one, so that the symmetric space G/K may be realized as the space Gm of measures on $B(G)$. We let $D = Gm$ and let \bar{D} denote the closure of D as a set of probability measures. The fact that $B(G)$ is a boundary implies that the point measures of $B(G)$ are members of \bar{D} , and in this way we can identify $B(G)$ with a subset of \bar{D} . We refer to this subset as the *distinguished boundary* of D .

THEOREM 5.2. *If $\{U_n\}$ is a discrete brownian motion on $D = G/K$ then with probability 1, U_n converges to a point of the distinguished boundary of D .*

Proof. We may suppose that $U_0 = \sigma(g_0)$, then $U_n = \sigma(g_0 X_1 X_2 \cdots X_n)$. Now, in identifying G/K with Gm , we map $\sigma(g)$ onto gm . We thus may take $U_n = g_0 X_1 X_2 \cdots X_n m$. Next, observe that the K -invariant measure m on $B(G)$ is stationary for μ since $k(\mu * m) = (k\mu) * m = \mu * m$, so that $\mu * m$ is K -invariant, hence equal to m . Therefore, m is the unique stationary measure for μ on $B(G)$. Thus, if $\{Z_n\}$ is the unique stationary μ -process on $B(G)$, $E^*(Z_n) = m$. Since $B(G)$ is a boundary we have, by Theorem 3.1, that $\{Z_n\}$ is a proper process. Then Z_0 is measurable over $\{X_0, X_{-1}, \dots\}$ and so $E^*(Z_0 | X_0, X_{-1}, \dots) = \delta_{Z_0}$ with probability 1. By Lemma 1.3, $X_0 X_{-1} \cdots X_{-k} E^*(Z_0) \rightarrow \delta_{Z_0}$ or $X_0 X_{-1} \cdots X_{-k} m \rightarrow \delta_{Z_0}$ with probability 1. The X_n being independent, we can change indices and infer that with probability 1, $X_1 X_2 \cdots X_k m$ converges to a point measure on $B(G)$, and as we have seen, this is a restatement of our theorem.

COROLLARY. *Let $U(t)$ be a brownian motion in a Cartan domain \mathcal{D} . With probability 1, $\lim_{t \rightarrow \infty} U(t)$ exists and is a point of the Bergman-Šilov boundary of \mathcal{D} .*

Proof. A Cartan domain is a symmetric space which is imbedded in some complex C^n as a bounded set and whose isometries are analytic transformations. We have shown [8, Theorem 4.5] that there is an equivariant map of D into C^n taking D onto \mathcal{D} and the distinguished boundary $B(G)$ of D onto the Bergman-Šilov boundary of \mathcal{D} . The corollary then follows readily from Theorem 5.2.

PART II. LAWS OF LARGE NUMBERS.

6. Spherical functions on a semi-simple group.

6.1. SPHERICAL FUNCTIONS AND COCYCLES. If $\{X_n\}$ is a sequence of independent, identically distributed variables with values in a multiplicative group of positive reals, then the law of large numbers can be stated: $n^{-1} \log X_n X_{n-1} \cdots X_1 \rightarrow E(\log X_n)$. The function $\log t$ on this group is characterized up to a constant multiple by the functional equation: $\log st = \log s + \log t$, and the requirement of continuity. In extending this law to group-valued random variables, we shall consider a family of functions defined by analogous conditions. On a semi-simple group it would be without interest to consider functions

$f(g)$ satisfying $f(g_1g_2) = f(g_1) + f(g_2)$ since any such function would vanish identically. (A homomorphic image of a semi-simple group is semi-simple and if f did not vanish identically it would map the group onto the reals which do not form a semi-simple group.) There is, however, an analogous functional equation which has nontrivial solutions. In this section we shall define these functions and study how they may be represented.

As before, G is a semi-simple group (with finite center). K will denote a fixed maximal compact subgroup of G . If M is a boundary of G then K is transitive on M [8, Theorem 1.4] and there exists a unique K -invariant probability measure on M . We denote this measure in each case by m . M_K denotes the normalized Haar measure on K and we shall usually abbreviate $dm_K(k)$ to dk . A function $f(g)$ on G will be called *left-uniformly continuous* (l.u.c.) if $f(g'g) - f(g)$ is uniformly small in g for g' sufficiently close to the identity of G . A function may be l.u.c. without being bounded on G . However, if $f(g)$ is l.u.c., all the differences $f(g'g) - f(g)$ are bounded functions of g , as follows readily from the connectedness of G .

DEFINITION 6.1. An l.u.c. real valued function $\psi(g)$ is called an *A-spherical function* (*A-s.f.*) if for every $g_1, g_2 \in G$,

$$(6.1) \quad \int_K \psi(g_1kg_2)dk = \psi(g_1) + \psi(g_2).$$

From (6.1) we find that $\psi(kg_2) = \psi(g_2)$ and $\psi(g_1k) = \psi(g_1)$, so that in general

$$(6.2) \quad \psi(k_1gk_2) = \psi(g) \quad k_1, k_2 \in K.$$

Setting $g_2 = e$ in (6.1), we obtain

$$(6.3) \quad \psi(e) = 0.$$

DEFINITION 6.2. A real valued function $\phi(g)$ is called an *M-spherical function* (*M-s.f.*) if $\phi(g) > 0$, $\log \phi(g)$ is l.u.c., and if

$$(6.4) \quad \int_K \phi(g_1kg_2)dk = \phi(g_1)\phi(g_2)$$

for every $g_1, g_2 \in G$.

An *M-spherical function* is *zonal spherical* in the sense of Gel'fand [10; 13, Chapter X]. From (6.4) we deduce

$$(6.5) \quad \phi(k_1gk_2) = \phi(g) \quad k_1, k_2 \in K; \phi(e) = 1.$$

The set of *A-spherical functions* form a linear space, whereas the set of *M-spherical functions* have no such structure. As we will ascertain, the linear space of *A-spherical functions* is finite dimensional.

DEFINITION 6.3. If M is a G -space, an A -cocycle on M is a real valued continuous function on $G \times M$ satisfying $\rho(g_1g_2, x) = \rho(g_1, g_2x) + \rho(g_2, x)$ for $g_1, g_2 \in G, x \in M$, and $\rho(k, x) = 0$ for $k \in K$.

DEFINITION 6.4. If M is a G -space, an M -cocycle on M is a positive continuous function on $G \times M$ satisfying $\sigma(g_1g_2, x) = \sigma(g_1, g_2x)\sigma(g_2, x)$ for $g_1, g_2 \in G, x \in M$, and $\sigma(k, x) = 1$ for $k \in K$.

Clearly the M -cocycles are just the exponentials of the A -cocycles.

DEFINITION 6.5. The real vector space of A -spherical functions will be denoted by \mathcal{V}_G ; the space of A -cocycles on the maximal boundary $B(G)$ will be noted by \mathcal{W}_G .

6.2. REPRESENTATION OF A -SPHERICAL AND M -SPHERICAL FUNCTIONS. If M is a G -space on which K is transitive, then every A -cocycle on M leads to an A -s.f. and every M -cocycle to an M -s.f. Suppose ρ is an A -cocycle on M ; set

$$(6.6) \quad \psi_\rho(g) = \int_M \rho(g, x) dm(x)$$

where m is the unique K -invariant probability measure on M . Then

$$(6.7) \quad \begin{aligned} \psi_\rho(g'g) &= \int_M (\rho(g', gx) + \rho(g, x)) dm(x) \\ &= \int_M \rho(g', gx) dm(x) + \psi_\rho(g). \end{aligned}$$

Since $\rho(e, x) = 0$ for all $x \in M$ it follows that as $g' \rightarrow e$, the integral in (6.7) goes to 0 uniformly in g . This shows that ψ_ρ is l.u.c. Moreover,

$$\begin{aligned} \int_K \psi_\rho(g_1kg_2)dk &= \int_K \int_M \rho(g_1kg_2, x)dm(x)dk \\ &= \int_K \int_M \rho(g_1, kg_2x) dm(x)dk + \int_K \int_M \rho(k, g_2x)dm(x)dk \\ &\quad + \int_K \int_M \rho(g_2, x)dm(x)dk. \end{aligned}$$

If, in the first integral on the right, we integrate with respect to k first, we obtain $\int_M \rho(g_1, y)d\pi(y)$ where π is a K -invariant measure. Hence $\pi = m$. The second integral vanishes, so we have

$$\begin{aligned} \int_K \psi_\rho(g_1kg_2)dk &= \int_M \rho(g_1, y)dm(y) + \int_M \rho(g_2, x)dm(x) \\ &= \psi_\rho(g_1) + \psi_\rho(g_2). \end{aligned}$$

In a similar way it may be shown that an M -cocycle determines an M -s.f. by

$$(6.8) \quad \phi_\sigma(g) = \int_M \sigma(g, x) dm(x).$$

Since the M -cocycles are exponentials of A -cocycles, it will be convenient to use the notation ϕ_ρ for the M -s.f. determined by e^ρ where ρ is an A -cocycle. Thus

$$(6.9) \quad \phi_\rho(g) = \int_M e^{\rho(g, x)} dm(x).$$

THEOREM 6.1. *Every A -spherical function is given by (6.6) with $M = B(G)$. Moreover, the correspondence $\rho \rightarrow \psi_\rho$ is an isomorphism of \mathcal{W}_G onto \mathcal{V}_G .*

Proof. Let $\psi(g)$ be an A -s.f. We shall consider A -cocycles $R(g_1, g)$ on G itself. Let Q be the set of these cocycles satisfying (a):

$$|R(g_1, g)| \leq \sup_{g \in G} |\psi(g_1g) - \psi(g)|$$

(this is finite because ψ is l.u.c.) and (b): for every $g' \in G$,

$$\psi(g) = \int_K R(g, kg') dk.$$

The set Q forms a convex set in the linear space of real valued A -cocycles on G . We introduce a topology in this linear space by taking as neighborhoods of 0, sets of A -cocycles $R(g_1, g)$ which are uniformly close to 0 on compact subsets of $G \times G$. We shall show that Q is a compact set with respect to this topology.

Since ψ is l.u.c., we have by (a) that $R(\eta, g) \rightarrow 0$ uniformly in g and R , as $\eta \rightarrow e$. We shall show that the functions of Q are equicontinuous in both variables; in other words, that $R(\eta_1g_1, \eta_2g_2) \rightarrow R(g_1, g_2)$ uniformly in $R \in Q$ as $\eta_1 \rightarrow e$, $\eta_2 \rightarrow e$ for fixed g_1, g_2 . By the functional equation for A -cocycles we have

$$R(\eta_1g_1, \eta_2g_2) = R(\eta_1, g_1\eta_2g_2) + R(g_1, \eta_2g_2),$$

and the first term on the right is small if η_1 is close to e , uniformly in R . It suffices to consider $R(g_1, \eta g_2)$ with $\eta \rightarrow e$.

$$\begin{aligned} R(g_1, \eta g_2) &= R(g_1\eta, g_2) - R(\eta, g_2) \\ &= R(g_1\eta g_1^{-1}\eta^{-1}\eta g_1, g_2) - R(\eta, g_2) \\ &= R(g_1\eta g_1^{-1}\eta^{-1}, \eta g_1g_2) + R(\eta g_1, g_2) - R(\eta, g_2) \\ &= R(g_1\eta g_1^{-1}\eta^{-1}, \eta g_1g_2) + R(\eta, g_1g_2) + R(g_1, g_2) - R(\eta, g_2). \end{aligned}$$

Since $g_1\eta g_1^{-1}\eta^{-1} \rightarrow e$ as $\eta \rightarrow e$, the first, second, and fourth terms on the right $\rightarrow 0$ uniformly in $R \in Q$ and the result follows. From the equicontinuity we conclude in the usual way that Q is compact; hence a compact, convex subset of a locally convex topological linear space.

We next show that G acts on Q by affine transformations. Namely, set

$$\tau_\gamma R(g_1, g) = R(g_1, g\gamma);$$

both (a) and (b) are preserved, and $\tau_\gamma R$ is again a cocycle on G , so $\tau_\gamma : Q \rightarrow Q$.

The set Q is nonempty. In fact, setting $R(g_1, g) = \psi(g_1g) - \psi(g)$, we find $R \in Q$. To begin with, R is an A -cocycle on G as is readily verified. (a) is trivially satisfied and (b) becomes

$$\begin{aligned} \psi(g) &= \int_K (\psi(gkg') - \psi(kg')) dk \\ &= \int_K \psi(gkg') dk - \psi(g') \end{aligned}$$

which is (6.1).

Finally we recall from [8, Theorem 1.4], that the minimal B -subgroup $H(G)$ has the fixed-point property. This implies that $H(G)$ has a fixed point in Q . Equivalently, there exists an $R(g_1, g) \in Q$ with $R(g_1, gh) = R(g_1, g)$ for $h \in H(G)$. We may then identify R with an A -cocycle ρ on $G/H(G)$ or $B(G)$: $R(g_1, g) = \rho(g_1, gH)$. Then (b) becomes

$$\psi(g) = \int_K \rho(g, kx) dk = \int_{B(G)} \rho(g, y) d(m_K * \delta_x)(y),$$

where $x \in B(G)$. Now $m_K * \delta_x$ is a K -invariant measure on $B(G)$; hence $m_K * \delta_x = m$ and

$$\psi(g) = \int_{B(G)} \rho(g, y) dm(y) = \psi_\rho(g).$$

This shows that the mapping $\rho \rightarrow \psi_\rho$ of \mathcal{W}_G into \mathcal{V}_G is onto.

To complete the proof of the theorem we must show that if $\psi_\rho = 0$ then $\rho = 0$. Suppose $\psi_\rho = 0$, then

$$\begin{aligned} 0 &= \int \rho(gg', x) dm(x) = \int \rho(g, g'x) dm(x) + \int \rho(g', x) dm(x) \\ &= \int \rho(g, g'x) dm(x) + \psi_\rho(g') = \int \rho(g, g'x) dm(x) = \int \rho(g, y) dg'm(y). \end{aligned}$$

But $B(G)$ is a boundary and so a sequence $\{g'_n\}$ may be chosen so that $g'_n m$ converges to any preassigned point measure δ_x on $B(G)$. It follows that $\rho(g, x) = 0$, and this proves the theorem.

Let $G = K \cdot A \cdot N$ be an Iwasawa decomposition of G relative to K [13, Chapter VI]. $A \cdot N$ is a solvable group, N its commutator and A is an abelian group. We set $r = \dim A$.

THEOREM 6.2. $\dim \mathcal{V}_G = \dim \mathcal{W}_G = r$.

Proof. Let $S = A \cdot N$; by [8, Theorem 1.4], S may be taken as a subgroup

of $H(G)$. Let $x_0 \in B(G)$ be the point left fixed by $H(G)$; then S leaves x_0 fixed. Since K is transitive on $B(G)$, each $x \in B(G)$ can be written $x = kx_0$, so that $\rho(g, x) = \rho(g, kx_0) = \rho(gk, x_0) - \rho(k, x_0)$ or $\rho(g, x) = \rho(gk, x_0)$. Hence each A -cocycle on $B(G)$ is determined by the function $\rho(g, x_0)$ on G . Now $\rho(kg, x_0) = \rho(k, gx_0) + \rho(g, x_0) = \rho(g, x_0)$ so that this function is left invariant under K . Hence ρ is determined by the function $\rho(s, x_0)$ on S . But since S leaves x_0 fixed, we have $\rho(s_1s_2, x) = \rho(s_1, x_0) + \rho(s_2, x_0)$ so that $\rho(s, x_0)$ is a homomorphism from S to the reals. Such a homomorphism takes $N \rightarrow 0$ and it follows that ρ is determined by a character on $A = S/N$.

This shows that $\dim \mathcal{V}_G = \dim \mathcal{W}_G \leq r$. The reverse inequality depends upon the decomposition $H(G) = M \cdot A \cdot N$, where M is the centralizer of A in K (§0.4). Suppose that $\chi(s)$ is a homomorphism from S to the additive group of reals. We define a cocycle on $B(G)$ by setting $\rho(ks, x_0) = \chi(s)$, and $\rho(k_1s, k_2x_0) = \rho(sk_2, x_0)$. It may be verified that if this definition is unambiguous, then ρ is actually an A -cocycle. To show that the definition is unambiguous, it is necessary to show that if $k_1x_0 = k_2x_0$, then $\rho(sk_1, x_0) = \rho(sk_2, x_0)$ for all $s \in S$. To evaluate $\rho(sk_i, x_0)$, we set $sk_i = \tilde{k}_i s_i$, so that $\rho(sk_i, x_0) = \chi(s_i)$. We proceed now to show that for any χ , $\chi(s_1) = \chi(s_2)$ under these circumstances. We have

$$s_2 = \tilde{k}_2^{-1} s \tilde{k}_2 = \tilde{k}_2^{-1} \tilde{k}_1 s_1 k_1^{-1} k_2 = k' s_1 k$$

where $k = k_1^{-1} k_2$, $k' = \tilde{k}_2^{-1} \tilde{k}_1$. Since $k_1^{-1} k_2 x_0 = x_0$, $k = k_1^{-1} k_2 \in H(G)$. So $k \in H(G) \cap K = M$. Since M is in the normalizer of S , $k^{-1} s_1 k \in S$. We then find $k^{-1} (k')^{-1} s_2 \in S$, so $k^{-1} (k')^{-1} \in S \cap K$ or $k' = k^{-1}$. Hence $s_2 = k^{-1} s_1 k$. Now the automorphism $s \rightarrow k^{-1} s k$ takes N into itself (M is also in the normalizer of N). Writing $s_1 = an$, we find $k^{-1} s_1 k = k^{-1} a k k^{-1} n k = a(k^{-1} n k)$ since $k \in M$. Thus $\chi(k^{-1} s_1 k) = \chi(a) = \chi(s_1)$ since χ vanishes on N . This proves that $\chi(s_1) = \chi(s_2)$ which shows that ρ is well defined. This argument shows that a nonvanishing character on S leads to a nonvanishing A -cocycle on $B(G)$ and so $\dim \mathcal{V}_G = r$. This completes the proof of the theorem.

REMARK. It is natural to conjecture that in a like manner, all M -spherical functions are given by (6.9) with $M = B(G)$. In fact, using Harish-Chandra's spherical function formula [13, Chapter X, Theorem 6.16], this can be shown to be the case, but with ρ a complex-valued A -cocycle. In the latter representation, a spherical function is determined by a homomorphism from A to the complex numbers. The arguments of the foregoing theorem show that such a homomorphism corresponds to a complex-valued A -cocycle on $B(G)$.

6.3. AN EXAMPLE. Let G be the group of fractional linear transformations of the complex plane leaving fixed the unit circle:

$G = \{g : gz = \rho(z + \alpha)(1 + \bar{\alpha}z)^{-1}, |\rho| = 1, |\alpha| < 1\}$. It may be checked that $B(G)$ is the unit circle $\Gamma = \{\zeta : |\zeta| = 1\}$. We obtain an M -cocycle on $B(G)$ by taking $\sigma(g, \zeta) = |g'(\zeta)|$, since $(g_1 g_2)'(\zeta) = g_1'(g_2 \zeta) g_2'(\zeta)$ and for K the compact

group of rotations in G ($\alpha = 0$), we have $|k'(\zeta)| = 1$, $k \in K$. $\rho = \log \sigma$ will then be an A -cocycle on Γ .

The K -invariant probability measure on Γ is given by $(1/2\pi i)\zeta^{-1}d\zeta$, so

$$\begin{aligned} \psi_\rho(g) &= \frac{1}{2\pi i} \int_\Gamma \log |g'(\zeta)| \frac{d\zeta}{\zeta} \\ &= \mathcal{R} \left\{ \frac{1}{2\pi i} \int_\Gamma \log g'(\zeta) \frac{d\zeta}{\zeta} \right\} \\ &= \mathcal{R} \{ \log g'(0) \} = \log |g'(0)|. \end{aligned}$$

If $gz = \rho(z + \alpha)(1 + \bar{\alpha}z)^{-1}$, then $g'(0) = \rho(1 - |\alpha|^2)$ and $\psi_\rho(g) = \log(1 - |\alpha|^2) = \log(1 - |g(0)|^2)$. In this case, $r = 1$, and so every A -spherical function on G is a multiple of this ψ_ρ . The M -spherical functions are more complicated but they have been treated more extensively (see e.g. [11]).

7. Asymptotic behavior of $\psi(X_n X_{n-1} \cdots X_1)$.

7.1. UNIQUENESS OF STATIONARY MEASURE AND LAWS OF LARGE NUMBERS. The next lemma is essentially contained in [4].

LEMMA 7.1. *Let M be a compact space and $\{\mu_x\}$ a random walk on M (see §0.3). A probability measure ν on M will be called stationary for this random walk if for each f , $\nu(f) = \int \mu_x(f) d\nu(x)$. Suppose that there exists a unique stationary measure ν for the random walk, and let $\{Z_n\}$ be any Markov process with transition probability measures $\{\mu_x\}$. Then with probability 1*

$$(7.1) \quad \frac{1}{n} \sum_0^{n-1} f(Z_k) \rightarrow \nu(f)$$

for any continuous f on M .

Proof. Recall Lemma 2.5 where it was shown that for functions of the form $\tau f - f$, τ defined by (2.7), the averages in question tend to 0 with probability 1. Let $\{Z_k\}$ be a sample sequence for which this is the case and suppose that for some subsequence n_i we have $n_i^{-1} \sum_0^{n_i-1} f(Z_k) \rightarrow c$. By sufficiently refining the subsequence, we may assume that $\lim_{i \rightarrow \infty} n_i^{-1} \sum_0^{n_i-1} F(Z_k)$ exists for all continuous F on M .

This limit defines a functional on the space of continuous functions on M , and we may check that this functional corresponds to a probability measure on M . Let ν' be this measure; then $\nu'(f) = c$. Now $\nu'(\tau f - F) = 0$ or $\nu'(\tau F) = \nu'(F)$ and

$$\int \mu_x(F) d\nu'(x) = \nu'(F)$$

for all continuous F . But this implies that ν' is a stationary measure for $\{\mu_x\}$,

and so $v' = v$. Therefore $c = v'(f) = v(f)$. Thus any convergent subsequence of the left side of (7.1) converges to $v(f)$, and since the sequence is bounded, this establishes the desired result.

The same method of proof yields a stronger result:

LEMMA 7.2. *Let N be the set of stationary measures on M for $\{\mu_x\}$. With the notation of Lemma 7.1, we now have that (7.1) is valid in general, for functions f with the property that $v_1(f) = v_2(f)$ for any $v_1, v_2 \in N$. The measure on the right side of (7.1) is then any measure in N .*

These results apply in particular to G -spaces where the random walk $\{\mu_x\}$ is that induced by the measure μ on G . A stationary measure is then one that satisfies for all continuous f on M :

$$\begin{aligned} \int f(x)dv(x) &= \int \mu * \delta_x(f)dv(x) \\ &= \int f(gx)d\mu(g)dv(x) \\ &= \int f(y)d\mu * v(y) \end{aligned}$$

or, equivalently, $\mu * v = v$. A stationary measure for the random walk induced by μ is then a stationary measure for μ . Lemmas 7.1 and 7.2 may then be applied to compact G -spaces for which we have information regarding the stationary measures.

We have in mind a slightly more involved application. Let M be a compact G -space and $f(g, x)$ a continuous function on $G \times M$. Suppose that $|f(g, x)| < F(g)$ where $F(g)$ is integrable on G with respect to the measure μ . Let $\{Z_n\}$ be a μ -process on M and let $\{X_n\}$ be the associated G -valued random variables.

LEMMA 7.3. *Under the foregoing assumptions we will have with probability 1,*

$$(7.2) \quad \frac{1}{n} \sum_0^{n-1} f(X_{k+1}, Z_k) \rightarrow \int_G \int_M f(g, x)d\mu(g)dv(x),$$

where v is a stationary measure for μ on M , provided that the expression on the right side of (7.1) is independent of which stationary measure v is chosen.

Proof. Compactify the space $G \times M$ by forming $\tilde{M} = (G \times M) \cup M$ and assigning a topology to \tilde{M} as follows: $G \times M$ will be open in \tilde{M} and $(g_n, x_n) \rightarrow y \in M$ if $g_n \rightarrow \infty$ (i.e. g_n goes outside of every compact subset of G) and $g_n x_n \rightarrow y$. On M we define a random walk by assigning to the point $(g, x) \in \tilde{M}$ the product measure $\mu \times \delta_{gx}$ on $G \times M \subset \tilde{M}$ and to the point $y \in M$, the product measure $\mu \times \delta_y$ on $G \times M$. We can see that this is a continuous assignment of

measures. Let us determine the stationary measures for this random walk. If $\tilde{\nu}$ is such a measure, since $\tilde{\nu} = \int_{\tilde{M}} \mu_{\xi} d\tilde{\nu}(\xi)$, and each μ_{ξ} assigns measure 0 to $M \subset \tilde{M}$, the same is true of $\tilde{\nu}$. So $\tilde{\nu}$ will be a measure on the product space $G \times M$. Let us show that $\tilde{\nu}$ is a product measure. Suppose f_1 is a function of g alone, and f_2 a function of x alone. Then

$$(7.3) \quad \begin{aligned} \tilde{\nu}(f_1 f_2) &= \int \mu_{(g,x)}(f_1 f_2) d\tilde{\nu}(g, x) \\ &= \int \mu(f_1) f_2(gx) d\tilde{\nu}(g, x) = \mu(f_1) \int f_2(gx) d\tilde{\nu}(g, x). \end{aligned}$$

Take $f_2 = 1$; we see that $\tilde{\nu}(f_1) = \mu(f_1)$. Take $f_1 = 1$; then $\tilde{\nu}(f_2) = \int f_2(gx) d\tilde{\nu}(g, x)$. (7.3) then becomes

$$\tilde{\nu}(f_1 f_2) = \tilde{\nu}(f_1) \tilde{\nu}(f_2),$$

which implies that $\tilde{\nu}$ is a product measure. Moreover, we have $\tilde{\nu}(f_1) = \mu(f_1)$ which implies that the G component of $\tilde{\nu}$ is μ . From the expression for $\tilde{\nu}(f_2)$ it follows that if ν is the M component, then $\mu * \nu(f_2) = \nu(f_2)$ whence $\mu * \nu = \nu$. Thus we conclude that a stationary measure for the given random walk has the form $\mu \times \nu$ where ν is a stationary measure on M for μ .

Consider now the composite process $\{(X_{k+1}, Z_k)\}$ occurring in our lemma. This can be considered a Markov process on \tilde{M} since $G \times M \subset \tilde{M}$. Moreover, the transition probability measures agree with those under consideration since $(X_{k+1}, Z_k) \rightarrow (X_{k+2}, Z_{k+1}) = (X_{k+2}, X_{k+1} Z_k)$ and X_{k+2} is independent of all the preceding variables. The assertion of Lemma 7.3 now follows immediately from Lemma 7.2 for those functions $f(g, x)$ on $G \times M$ which can be extended to continuous functions on \tilde{M} . Clearly, functions $f(g, x)$ vanishing outside a compact set of $G \times M$ have this property. To establish Lemma 7.3 for the class of functions under consideration, it suffices to approximate f by an f' of compact support in $G \times M$ such that both sides of (7.2) are small for $f - f'$. Choose $f'(g, x) = f(g, x)$ for g inside a large compact set of G , and $f'(g, x) = 0$ for g outside a larger compact set, retaining the inequality $|f'(g, x)| < F(g)$ everywhere. Then

$$|(f - f')(g, x)| < F'(g)$$

where $\int F'(g) d\mu(g)$ can be made arbitrarily small. The right side of (7.2) is then clearly small for $f - f'$. The same, however, is true for the left side by the law of large numbers for $F'(X_k)$. This proves the lemma.

7.2. LAWS OF LARGE NUMBERS. Even in the classical case, some restriction must be placed on the distribution of a random variable for the validity of the law of large numbers. In our case we have the following restriction:

DEFINITION 7.1. A probability measure μ on G is of class B_1 if it is absolutely continuous, and if for each A -cocycle $\rho(g, x)$ on $B(G)$ we have

$$(7.4) \quad \int_G \sup_{x \in B(G)} |\rho(g, x)| d\mu(g) < \infty.$$

DEFINITION 7.2. Let μ be of class B_1 , and let ν be its unique stationary measure on $B(G)$ (see Theorem 2.3). If $\rho \in \mathcal{W}_G$ we set

$$(7.5) \quad \alpha_\mu(\rho) = \int_G \int_M \rho(g, x) d\mu(g) d\nu(x).$$

(7.5) assigns to each μ of class B_1 a linear functional α_μ on \mathcal{W}_G . This functional will play a very significant role in the sequel.

THEOREM 7.1. Let μ be of class B_1 and let $\{X_n\}$ be a sequence of independent G -valued random variables with distribution μ . If $\rho \in \mathcal{W}_G$ then with probability 1,

$$(7.6) \quad n^{-1} \rho(X_n X_{n-1} \cdots X_1, x) \rightarrow \alpha_\mu(\rho), \quad n \rightarrow \infty$$

for any $x \in B(G)$.

Proof. $\rho(X_n X_{n-1} \cdots X_1, x) = \sum_{k=0}^{n-1} \rho(X_{k+1}, X_k \cdots X_1, x) = \sum_{k=0}^{n-1} \rho(X_{k+1}, Z_k)$ where $\{Z_k\}$ is a μ -process on $B(G)$. Our theorem will now follow from Lemma 7.3 if we can verify the hypotheses of that lemma. For one thing, $|\rho(g, x)| < \sup_x |\rho(g, x)|$ which we assume is integrable with respect to μ by (7.4). Moreover, since μ has a unique stationary measure on $B(G)$ (Theorem 2.3), there is no ambiguity on the right hand side of (7.2) in our case. This completes the proof of the theorem.

The next theorem expresses our formulation of the law of large numbers for the group G .

THEOREM 7.2. Let ψ_ρ be the A -spherical function corresponding to ρ (by (6.6)). Under the hypotheses of Theorem 7.1 we will have, with probability one,

$$(7.7) \quad n^{-1} \psi_\rho(X_n X_{n-1} \cdots X_1) \rightarrow \alpha_\mu(\rho) \quad n \rightarrow \infty.$$

Proof. By (6.6), this will follow from (7.6) by integrating over $B(G)$, if we can justify the passage to the limit under the integral. Let $F(g) = \sup_{x \in B(G)} |\rho(g, x)|$; $F(g)$ is integrable with respect to μ by (7.4). Now

$$|\rho(X_n X_{n-1} \cdots X_1, x)| = \left| \sum_{k=0}^{n-1} \rho(X_{k+1}, X_k \cdots X_1, x) \right| \leq \sum_{k=0}^{n-1} F(X_{k+1}).$$

Therefore

$$|n^{-1} \rho(X_n X_{n-1} \cdots X_1, x)| \leq n^{-1} \sum_{k=0}^{n-1} F(X_{k+1}),$$

and by the law of large numbers for $\{F(X_k)\}$ this is bounded as $n \rightarrow \infty$, with probability 1. Thus, with probability one, the expressions $n^{-1} \rho(X_n X_{n-1} \cdots X_1, x)$ are uniformly bounded in x and we may pass to the limit under the integral sign.

Let us consider the example of §6.3. If $gz = \rho(z + \alpha)(1 + \alpha z)^{-1}$, $\rho(g, \zeta) = \log |g'(\zeta)|$, $\psi_\rho(g) = \log(1 - |g(0)|^2)$. We find that

$$\sup_{\zeta \in \Gamma} \rho(g, \zeta) = \log \frac{1 + |g(0)|}{1 - |g(0)|}$$

and this function must be integrable with respect to μ for μ to be of class B_1 . Rewriting the law of large numbers, (7.7), for that case, we find

$$(7.8) \quad n^{-1} \log(1 - |X_n X_{n-1} \cdots X_1(0)|^2) \rightarrow \alpha_\mu(\rho).$$

This expression, incidentally, gives information concerning the μ -process in the disc $\{|z| < 1\}$ which starts at the origin. As we will see later, $\alpha_\mu(\rho)$ is in this case strictly negative, and (7.8) shows the rate with which the μ -process in the disc converges to the boundary.

As was pointed out in the Introduction, the limits $\alpha_\mu(\rho)$ in (7.6) and (7.7) apparently cannot be computed directly in terms of μ and ρ , but their expression requires knowledge of the stationary measure ν on $B(G)$. There is an exception to this which we point out in the next theorem.

THEOREM 7.3. *If μ has the form $\mu = \mu_1 * m_K * \mu_2$, then $\nu = \mu_1 * m$, where m is the unique K -invariant measure on $B(G)$, and*

$$(7.9) \quad \alpha_\mu(\rho) = \int_K \int_K \int_G \int_G \int_G \rho(g_1 k g_2, g_1' k' x) d\mu_1(g_1) d\mu_1(g_1') d\mu_2(g_2) dk dk'$$

for any $x \in B(G)$.

Proof. It suffices to show that $\mu_1 * m$ is a stationary measure for μ . Now since m is the unique K -invariant probability measure on $B(G)$, it follows that if π is any probability measure on $B(G)$, $m_K * \pi = m$. Hence $(m_K * \mu_2) * (\mu_1 * m) = m$, and $(\mu_1 * m_K * \mu_2) * (\mu_1 * m) = \mu_1 * m$. This proves the theorem.

In general, however, we cannot determine ν and α_μ explicitly. In the remainder of this section we shall attempt to obtain information regarding $\alpha_\mu(\rho)$ without evaluating it explicitly. Of particular interest is the question of whether $\alpha_\mu(\rho)$ is > 0 or < 0 , since these would imply in particular that $\psi \rightarrow \infty$ or $\psi \rightarrow -\infty$ respectively. For example, it is not entirely trivial that the μ -process in the disc $\{|z| < 1\}$ referred to previously is “transient”, i.e., that the process does not return infinitely often to a compact subset of the disc. This, however, follows from (7.8) together with the information that $\alpha_\mu(\rho) < 0$ in that case. Notice that the strict inequality is essential here.

7.3. NEGATIVE A -COCYCLES. We will now further restrict the class of probability measures on G .

DEFINITION 7.3. An absolutely continuous probability measure μ on G cor-

responds to a non-negative function $p(g)$ in $L^1(G)$. If $p(g)$ is bounded and of compact support we shall say that μ is of class B_∞ .

It turns out that for certain A -cocycles ρ we can make the assertion $\alpha_\mu(\rho) < 0$ or $\alpha_\mu(\rho) > 0$ independent of the measure μ , so long as μ is of class B_∞ . It turns out to be more convenient to consider the case $\alpha_\mu(\rho) < 0$.

DEFINITION 7.4. An A -cocycle $\rho \in \mathcal{W}_G$ is called a negative cocycle if for every μ of class B_∞ , $\alpha_\mu(\rho) < 0$.

LEMMA 7.4. If μ is of class B_∞ and ν is the stationary measure for μ on $B(G)$, then the Radon-Nikodym derivative $d\nu/dm$ is bounded, where m is the K -invariant probability measure on $B(G)$.

Proof. Let $d\mu = p(g)dg$ where dg is the element of Haar measure on G and set $\Lambda(g) = \sup_{k \in K} p(gk)$. Since $p(g)$ is bounded and of compact support and K is compact, the same holds for $\Lambda(g)$. Then

$$\begin{aligned} \int_{B(G)} f(x)d\nu(x) &= \int_{B(G)} \int_G f(gx)d\mu(g)d\nu(x) = \int_{B(G)} \int_G f(gx)p(g)dg d\nu(x) \\ &= \int_{B(G)} \int_K \int_G f(gkx)p(gk)dgdkd\nu(x) \end{aligned}$$

since the Haar measure on G is right invariant (G is semi-simple). Since $p(gk) \leq \Lambda(g)$, if $f \geq 0$,

$$\begin{aligned} \int_{B(G)} f(x)d\nu(x) &\leq \int_{B(G)} \int_K \int_G f(gkx) \Lambda(g)dgdkd\nu(x) \\ &= \int_{B(G)} \int_G f(gy)\Lambda(g)dg d(m_K * \nu)(y) \\ &= \int_{B(G)} \int_G f(gy)\Lambda(g)dg dm(y) \end{aligned}$$

since $m_K * \nu = m$. We can replace gy by x if we replace $dm(g)$ by $dgm(x)$. Then

$$\int_{B(G)} f(x)d\nu(x) \leq \int_G \left\{ \int_{B(G)} f(x) \frac{dgm}{dm}(x)dm(x) \right\} \Lambda(g)dg.$$

It follows that

$$\frac{d\nu}{dm}(x) \leq \int_G \frac{dgm}{dm}(x)\Lambda(g)dg$$

which is finite by virtue of the fact that Λ is bounded and has compact support. Here dgm/dm denotes the Radon-Nikodym derivative of the two measures, and we have used the fact that $(dgm/dm)(x)$ is a continuous function on $G \times B(G)$.

We can prove now

THEOREM 7.4. *Let ρ be an A -cocycle on $B(G)$ such that the corresponding M -spherical function satisfies $\phi_\rho(g) \leq 1$ for all $g \in G$. Then unless $\rho \equiv 0$, ρ is a negative cocycle and $\alpha_\mu(\rho) < 0$ whenever μ is of class B_∞ .*

Proof. We have $\psi_\rho(g) = \int_{B(G)} \rho(g, x) dm(x)$ and we define $\psi'_\rho(g) = \int_{B(G)} \rho(g, x) dv(x)$, where v is the stationary measure for μ which we assume to be of class B_∞ . The contention of the theorem is that $\alpha_\mu(\rho) = \int_G \psi'_\rho(g) d\mu(g) < 0$. Now

$$\psi_\rho(g) \leq \log \int_{B(G)} e^{\rho(g, x)} dm(x) = \log \phi_\rho(g) \leq 0.$$

By Theorem 7.2, $n^{-1} \psi_\rho(X_n X_{n-1} \cdots X_1) \rightarrow \alpha_\mu(\rho)$ so that $\alpha_\mu(\rho)$ is, in any case, ≤ 0 . Thus it remains to show that we cannot have $\int \psi'_\rho(g) d\mu(g) = 0$. Suppose, to the contrary, that this integral vanishes. Then

$$\begin{aligned} \int_G \psi'_\rho(gg') d\mu(g') &= \int_G \int_{B(G)} \rho(gg', x) dv(x) d\mu(g') \\ &= \int_G \int_{B(G)} \rho(g, g'x) dv(x) d\mu(g') + \int_{B(G)} \int_G \rho(g', x) d\mu(g') dv(x) \\ &= \int_{B(G)} \rho(g, y) d(\mu * v)(y) + \int_G \psi'_\rho(g') d\mu(g') \\ &= \int \rho(g, y) dv(y) = \psi'_\rho(g). \end{aligned}$$

It follows that ψ'_ρ is a μ -harmonic function (§0.4). We shall show that if μ is of class B_∞ , ψ'_ρ is bounded away from $+\infty$. Namely, by Lemma 7.4, dv/dm is a bounded function on $B(G)$; say $dv/dm < c$. Then

$$\begin{aligned} \psi'_\rho(g) &= \int_{B(G)} \rho(g, x) \frac{dv}{dm}(x) dm(x) \\ &\leq \log \int_{B(G)} e^{\rho(g, x)} \frac{dv}{dm}(x) dm(x) \\ &\leq \log c + \log \int_{B(G)} e^{\rho(g, x)} dm(x) = \log c + \log \phi_\rho(g) \\ &\leq \log c. \end{aligned}$$

Added in proof. Let $f(g)$ be a μ -harmonic function, bounded on one side and satisfying $f(kg) = f(g)$, $k \in K$. If $\{X_n\}$ is a sequence of independent G -valued random variables with distribution μ , then $U_n(g) = f(gX_1 \cdots X_n)$ forms a martingale and converges to some $F(g; X_1, X_2, \dots)$. In Theorem 3.1 of [8] it was shown (for $|f|$ bounded, but boundedness on one side is sufficient) that

$F(g; x_1, x_2, \dots)$ is independent of g . The function $f(g) = \psi'_\rho(g)$ satisfies these conditions, so we may conclude that with probability 1,

$$\psi'_\rho(hgX_1 \cdots X_n) - \psi'_\rho(gX_1 \cdots X_n) \rightarrow 0$$

as $n \rightarrow \infty$, for all $h, g \in G$. This gives

$$\int \rho(h, gX_1 \cdots X_n x) \, dv(x) \rightarrow 0.$$

By [8, Lemma 5.4], $X_1 \cdots X_n v$ converges to a point measure on $B(G)$, with probability 1. It follows that $\rho \equiv 0$ and the theorem is proven.

7.4. *A-CYCLES ON ARBITRARY BOUNDARIES.* We have been treating primarily cocycles on the maximal boundary $B(G)$. However, we saw in §6.2 that an A -cocycle on any G -space on which K acts transitively determines an A -spherical function. In particular, this will be the case for boundaries.

If M is a boundary then there is an equivariant map $\sigma : B(G) \rightarrow M$. Let ρ be an A -cocycle on M . Form $\rho^*(g, x) = \rho(g, \sigma(x))$ for $x \in B(G)$. Then

$$\begin{aligned} \rho^*(g_1 g_2, x) &= \rho(g_1 g_2, \sigma(x)) = \rho(g_1, g_2 \sigma(x)) + \rho(g_2, \sigma(x)) \\ &= \rho(g_1, \sigma(g_2 x)) + \rho(g_2, \sigma(x)) = \rho^*(g_1, g_2 x) + \rho^*(g_2, x) \end{aligned}$$

so that ρ^* is an A -cocycle on $B(G)$. We claim that ρ^* determines the same A -s.f. and M -s.f. as ρ .

For

$$\begin{aligned} \psi_{\rho^*}(g) &= \int_{B(G)} \rho^*(g, x) dm(x) = \int_{B(G)} \rho(g, \sigma(x)) dm(x) \\ &= \int_M \rho(g, y) d\sigma m(y) = \int_M \rho(g, y) dm(y) = \psi_\rho(g) \end{aligned}$$

since σm is again K -invariant if m is. Similarly we find $\phi_{\rho^*}(g) = \phi_\rho(g)$. Finally, we notice that $\alpha_\mu(\rho^*)$ can be obtained in terms of ρ . For if ν is the unique stationary measure for μ on $B(G)$, then $\sigma(\nu)$ will be the unique stationary measure for μ on M . Then

$$\begin{aligned} \alpha_\mu(\rho^*) &= \int_G \int_{B(G)} \rho(g, \sigma(x)) d\mu(g) d\nu(x) \\ &= \int_G \int_M \rho(g, y) d\mu(g) d\sigma(\nu)(y) \end{aligned}$$

and the latter integral may be denoted by $\alpha_\mu(\rho)$. We can now extend Theorem 7.4 to the case of cocycles on an arbitrary boundary.

THEOREM 7.5. *Let ρ be an A -cocycle on a boundary M such that $\phi_\rho(g) \leq 1$ for all $g \in G$. Then unless $\rho \equiv 0$, we will have $\alpha_\mu(\rho) < 0$, whenever μ is of class B_∞ .*

Proof. $\phi_{\rho^*}(g) = \phi_{\rho}(g)$ and $\alpha_{\mu}(\rho^*) = \alpha_{\mu}(\rho)$.

To illustrate this, let M be a nondegenerate boundary of G and m its K -invariant probability measure. Form

$$\sigma(g, x) = \frac{dg^{-1}m}{dm}(x).$$

Since m is K -invariant we have clearly $\sigma(k, x) = 1$. Also

$$\begin{aligned} \sigma(g_1g_2, x) &= \frac{dg_2^{-1}g_1^{-1}m}{dm}(x) = \frac{dg_2^{-1}g_1^{-1}m}{dg_2^{-1}m}(x) \frac{dg_2^{-1}m}{dm}(x) \\ &= \frac{dg_1^{-1}m}{dm}(g_2x) \frac{dg_2^{-1}m}{dm}(x) = \sigma(g_1, g_2x)\sigma(g_2, x) \end{aligned}$$

by the rules for the Radon-Nikodym derivative. Thus σ is an M -cocycle on M and $\rho = \log \sigma$ is an A -cocycle. Evaluating ϕ_{ρ} we find

$$\begin{aligned} \phi_{\rho}(g) &= \int_M \sigma(g, x)dm(x) = \int_M \frac{dg^{-1}m}{dm}(x)dm(x) \\ &= \int_M dg^{-1}m(x) = 1. \end{aligned}$$

Moreover, ρ is not identically 0 since otherwise $\sigma \equiv 1$ and $g^{-1}m \equiv m$ for all $g \in G$. But this is impossible since M is a boundary and there is a sequence $g_n m$ converging to a point measure. We can therefore apply Theorem 7.5 to deduce that $\alpha_{\mu}(\rho) < 0$. Incidentally, this proves

THEOREM 7.6. *There is a negative cocycle for any compact semi-simple group.*

An interesting consequence is

COROLLARY. *If G is any noncompact semi-simple group and μ a measure of class B_{∞} then the random walk induced by μ on G is transient.*

Proof. Let ρ be a negative cocycle on G ; then $\psi_{\rho}(X_n X_{n-1} \cdots X_1) \rightarrow -\infty$ with probability 1.

The set of negative cocycles forms a cone in the vector space \mathscr{W}_G . We would conjecture that this cone has a nonempty interior, i.e., that there always exist r linearly independent negative cocycles, where $r = \dim \mathscr{W}_G$.

8. Random products of unimodular matrices. In this section we shall describe the asymptotic behavior of $X_n X_{n-1} \cdots X_1$ where the X_j are independent, identically distributed, random unimodular $m \times m$ matrices. We are interested in the rate of growth of the column vectors of the product $X_n X_{n-1} \cdots X_1$ (§8.3) and the

behavior of the angles between the various row vectors and column vectors (§8.1). In §8.2 we shall discuss another type of limit behavior that occurs here; namely, if the underlying distribution μ is of class B_∞ , then the rows of the product matrix all tend, as $n \rightarrow \infty$, to a fixed direction.

We shall denote by R^m the m -dimensional real vector space of $m \times 1$ real matrices. R_t^m will denote the space of $1 \times m$ real matrices. If $\alpha \in \text{SL}(m, R)$, $u \in R^m$, $v \in R_t^m$ then the products αu , $v\alpha$, and $v\alpha u$ are all defined. If in $R^m - \{0\}$ we identify two vectors that are scalar multiples of one another we obtain the space P^{m-1} . We denote by P_t^{m-1} the space obtained similarly from $R_t^m - \{0\}$. If $u \in R^m$ (resp. R_t^m) we will let \bar{u} denote the corresponding element in P^{m-1} (resp. P_t^{m-1}). Given two vectors u_1, u_2 , we let $\langle u_1, u_2 \rangle$ denote the acute angle between the lines through u_1 and u_2 , $0 \leq \langle u_1, u_2 \rangle \leq \pi/2$. $\langle u_1, u_2 \rangle$ depends only on \bar{u}_1 and \bar{u}_2 and we shall also use the notation $\langle \bar{u}_1, \bar{u}_2 \rangle$.

8.1. ANGLES BETWEEN COLUMN VECTORS. We take $G = \text{SL}(m, R)$ and μ will be a probability measure of class B_∞ on G . K will be the group of orthogonal matrices of G . In [8, Chapter I], it was shown that the maximal boundary of G is the flag manifold F_m consisting of all $m - 1$ -tuples of subspaces $(V_1, V_2, \dots, V_{m-1})$ where the V_i are i -dimensional subspaces of R^m and $V_1 \subset V_2 \subset \dots \subset V_{m-1}$. The action of G on F_m is that induced naturally by the action of $\text{SL}(m, R)$ on R^m . We will not be concerned directly with the maximal boundary F_m , but with two of the boundaries onto which F_m maps. Namely, we take M_1 to be the set of 1-dimensional subspaces of R^m and M_2 to be the set of pairs (V_1, V_2) , $V_1 \subset V_2$ and $\dim V_i = i$. Notice that M_1 is the same as P^{m-1} . M_1 and M_2 are both boundaries of G ; both have natural topologies with respect to which they are compact, and both are G -spaces.

We introduce a third G -space $M_1^{(2)}$ which is not compact. $M_1^{(2)}$ is defined as the set of ordered pairs of distinct points in M_1 . The space $M_1^{(2)}$ may be compactified by adjoining M_2 as follows. For every pair of distinct 1-dimensional subspaces V_1' and V_1'' of R^m let $\beta(V_1', V_1'')$ denote the element of M_2 given by $V_1 = V_1'$ and $V_2 = 2$ -dimensional space spanned by V_1' and V_1'' . Now set $M = M_1^{(2)} \cup M_2$ with the following topology: $M_1^{(2)}$ is open in M and a sequence $(V_{1,n}', V_{1,n}'') \rightarrow (V_1, V_2) \in M_2$ if the angle between $V_{1,n}'$ and $V_{1,n}''$ goes to 0 and $\beta(V_{1,n}', V_{1,n}'') \rightarrow (V_1, V_2)$. With this topology M is a compact G -space.

We can describe this setup geometrically as follows. A point of $M_1^{(2)}$ consists of a pair of points of M_1 . This point will tend to infinity on $M_1^{(2)}$ if the pair of points on M_1 tend to one another. In the limit, assuming that one of the pair converges on M_1 , the pair of points determine not just a single point of M_1 but also a line element at that limit point. The set of these (points of M_1 and line elements through them) constitutes M_2 .

If $x \in M_1^{(2)}$, $x = (u', u'')$ with u' and $u'' \in P^{m-1}$, set $\omega(x) = \langle u', u'' \rangle$. Define

$$(8.1) \quad \sigma(g, x) = \omega(gx) / \omega(x).$$

Since K acts on R^m by rotations $\omega(kx) = \omega(x)$ and $\sigma(k, x) = 1$. From the form of (8.1) it is immediate that $\sigma(g_1 g_2, x) = \sigma(g_1, g_2 x) \sigma(g_2, x)$. Hence $\sigma(g, x)$ is an M -cocycle on $M_1^{(2)}$. Now $M_1^{(2)}$ is dense in M , and it is possible to show that $\sigma(g, x)$ extends continuously to $x \in M$. The extension is then again an M -cocycle and its restriction to M_2 is also an M -cocycle. The latter, which we shall denote by σ_2 , may be described in this way. K acts on the manifold M_1 in such a way that there is a unique Riemmanian metric (up to scalar multiple) on M_1 which is invariant under K . Each $g \in G$ sends a tangent vector at $p \in M_1$ onto a tangent vector at gp . The ratio of the lengths of these vectors depends on p , on the direction of the tangent vector at p , and on g . The first two of these correspond to a point ξ of M_2 and the ratio in question is just $\sigma_2(g, \xi)$.

LEMMA 8.1. μ has a unique stationary measure ν on M which has its support on M_2 .

Proof. $M = M_1^{(2)} \cup M_2$ and $M_1^{(2)} = M_1 \times M_1 - D$ where D is the diagonal of $M_1 \times M_1$. By Theorem 2.2, since M_1 is a boundary of G , there is no stationary measure for μ on $M_1 \times M_1 - D$. Thus any stationary measure for μ on M must have its support on M_2 . But M_2 is also a boundary and so it possesses a unique stationary measure.

Let $\rho_2(g, \xi) = \log \sigma_2(g, \xi)$ be the A -cocycle on M_2 determined by σ_2 . We then have

LEMMA 8.2. If $\{X_n\}$ is the G -valued process of independent random variables with distribution μ , then with probability 1, for any $x \in M$,

$$(8.2) \quad \lim_{n \rightarrow \infty} n^{-1} \log \sigma(X_n X_{n-1} \cdots X_1, x) = \alpha_\mu(\rho_2).$$

Proof. We have

$$\begin{aligned} n^{-1} \log \sigma(X_n X_{n-1} \cdots X_1, x) &= n^{-1} \sum_{k=0}^{n-1} \log \sigma(X_{k+1}, X_k \cdots X_1 x) \\ &= n^{-1} \sum_{k=0}^{n-1} \log \sigma(X_{k+1}, Z_k) \end{aligned}$$

where $\{Z_k\}$ is a μ -process on M and $\{X_k\}$ are the associated G -valued random variables. We may thus apply Lemma 7.3 to the function $f(g, x) = \log \sigma(g, x)$. Since μ is of compact support the hypothesis that $f(g, x)$ be bounded by an integrable function of g is satisfied. Moreover, there is a unique stationary measure on M for μ so that the right side of (7.2) is unambiguous. Finally, the integral over M in (7.2) reduces to an integral on M_2 since ν is supported on M_2 . The resulting expression is then just $\alpha_\mu(\rho_2)$.

It should be observed that though the right side of (8.2) depends on the G -space M_2 , the limit is the same for all $x \in M$, and thus for $x \in M_1^{(2)}$. Using (8.1) we obtain

THEOREM 8.1. *Let μ be a measure of class B_∞ on $SL(m, R)$ and let $\{X_n\}$ be the corresponding process. If u' and u'' are any distinct nonzero vectors in R^m , then with probability 1,*

$$(8.3) \quad \langle X_n X_{n-1} \cdots X_1 u', X_n X_{n-1} \cdots X_1 u'' \rangle^{1/n} \rightarrow e^{\alpha_\mu(\rho_2)}$$

We shall show next that $\alpha_\mu(\rho_2) < 0$.

LEMMA 8.3. $\int_{M_2} \sigma_2(g, x) dm(x) = 1$ for all $g \in G$.

Proof. Each $x \in M_2$ determines a 2-dimensional subspace of R^m . Let K_x be the subgroup of K leaving fixed the orthogonal complement of this subspace. K_x is then isomorphic to the circle group. Since m is K -invariant we may write

$$\int_{M_2} \sigma_2(g, x) dm(x) = \int_{M_2} \left\{ \int_{K_x} \sigma_2(g, kx) dk \right\} dm(x)$$

where dk now represents the element of Haar measure on K_x . To prove the lemma it will suffice to show that for each g and x

$$(8.4) \quad \int_{K_x} \sigma_2(g, kx) dk = 1.$$

We may parametrize the elements of K_x by $\{k(t)\}$ with $0 \leq t < 1$, so that dk is replaced by dt . Let $u(0)$ denote a unit vector in R^m lying in the 1-dimensional subspace determined by x . Then let $u(t) = k(t)u(0)$. $u(t)$ describes a circle as t ranges over the unit interval, and the vector $du(t)/dt$ has constant magnitude 2π . Now let $v(t)$ be the unit vector in the direction of $gu(t)$. We then have

$$(8.5) \quad \sigma_2(g, k(t)x) = \left\| \frac{dv(t)}{dt} \right\| / \left\| \frac{du(t)}{dt} \right\|,$$

where $\|\cdot\|$ is the euclidean metric. Then

$$\begin{aligned} \int_{K_x} \sigma_2(g, kx) dk &= \int_0^1 \sigma_2(g, k(t)x) dt = \frac{1}{2\pi} \int_0^1 \left\| \frac{dv(t)}{dt} \right\| dt \\ &= \frac{1}{2\pi} \int ds \end{aligned}$$

where ds is the element of length on the curve described by $v(t)$. But the total length is 2π since the curve is a circle. This proves the lemma.

Applying Theorem 7.5 we find

THEOREM 8.2. *The exponent $\alpha_\mu(\rho_2)$ in (8.3) is strictly negative.*

8.2. THE ROWS OF $X_n X_{n-1} \cdots X_1$. In this section we shall give a generalization of Theorem 1.3 regarding the convergence of continued fractions. The assertion will be that if $u \in P_t^{m-1}$, then $u X_n X_{n-1} \cdots X_1$ converges with probability 1. In particular, the rows of $X_n X_{n-1} \cdots X_1$ converge as $n \rightarrow \infty$. For 2×2 matrices this yields Theorem 1.3, except that in the present section, we have to impose greater restrictions on the measure μ . We shall assume throughout this section that μ is of class B_∞ , a condition that was not needed in Theorem 1.3.

Let $u', u'' \in R_t^m$ and $v', v'' \in R^m$. We shall set

$$(8.6) \quad (u', u''; v', v'') = \frac{(u'v')(u''v'')}{(u'v'')(u''v')}$$

where the products $(u'v')$, etc., represent the usual scalar product.

If $v', v'' \in R^m$ or R_t^m , we shall denote by $v' \wedge v''$ the $(m(m-1)/2)$ -tuple of determinants

$$\left| \begin{array}{cc} v'_i & v''_i \\ v'_j & v''_j \end{array} \right|, \quad i < j.$$

We shall consider this as an element in the euclidean space of $m(m-1)/2$ dimensions. If v' and v'' are vectors of fixed magnitude then the magnitude of $v' \wedge v''$ is of the same order as the angle between v' and v'' , as this angle goes to 0. The following identity is well known:

$$(8.7) \quad (u'v')(u''v'') - (u'v'')(u''v') = (u' \wedge u'') \cdot (v' \wedge v'').$$

From (8.7) we obtain

LEMMA 8.4.

$$(u', u''; v', v'') - 1 = \frac{(u' \wedge u'') \cdot (v' \wedge v'')}{(u'v'')(u''v')}.$$

We see that if either v' and v'' are close together or u' and u'' are close, and neither v'' is close to being orthogonal to u' nor v' to u'' , then $(u', u''; v', v'')$ is close to 1.

LEMMA 8.5. *If μ is of class B_∞ and M is a boundary of G , then there is a constant c such that if m is the K -invariant probability measure on M and $\pi = \mu * \pi_0$ for any probability measure π_0 on M , then $d\pi/dm \leq c$.*

Proof. The proof is identical to that of Lemma 7.4, where this was proved for the stationary measure on M . The only property of a stationary measure used was the fact that it had the form $\mu * \nu$ for a probability measure ν .

Let $\{e_i, i = 1, \dots, m\}$ represent a basis in R^m and $\{f_j, j = 1, \dots, m\}$ a basis in R_t^m . We shall denote by Y_n the random product $X_n X_{n-1} \cdots X_1$.

LEMMA 8.6. *With probability 1,*

$$(8.8) \quad (f_i, f_j; Y_n e_l, Y_n e_k) \rightarrow 1$$

for all i, j, l, k as $n \rightarrow \infty$.

Proof. By Theorem 8.1 and 8.2 the angle between $Y_n e_l$ and $Y_n e_k$ goes to 0 at a strictly exponential rate: say $\langle Y_n e_l, Y_n e_k \rangle < c_1 \exp(-n\beta)$. Let $Q_n^i \subset P^{m-1}$ consist of the points whose corresponding unit vectors u satisfy $(f_i u) < \exp(-n\beta/3)$. If $\overline{Y_n e_l} \notin Q_n^j$ and $\overline{Y_n e_k} \notin Q_n^i$ then we have from Lemma 8.4 that

$$|1 - (f_i, f_j; Y_n e_l, Y_n e_k)| < c_2 \exp(-n\beta/3)$$

for some constant c_2 . If these conditions are fulfilled for all but finitely many n , then (8.8) is valid. The lemma will thus follow if we can show that the probabilities of $\overline{Y_n e_l} \in Q_n^j$ occurring infinitely often or of $\overline{Y_n e_k} \in Q_n^i$ occurring infinitely often vanish. It suffices to show (Borel-Cantelli) that $\sum P(\overline{Y_n e_l} \in Q_n^j) < \infty$. Since the distribution of $\overline{Y_n e_l} = \overline{Y_n} e_l$ has the form $\mu * \pi_0$ on P^{m-1} , and P^{m-1} is a boundary of G , it follows from Lemma 8.5 that it suffices to show $\sum_n m(Q_n^j) < \infty$. But clearly

$$m(Q_n^j) < c_3 \exp(-n\beta/3)$$

for some constant c_3 . This proves the lemma.

Now

$$(f_i, f_j; Y_n e_l, Y_n e_k) = (f_i Y_n, f_j Y_n; e_l, e_k).$$

Apply Lemma 8.4 to the latter cross-ratio. Since (8.8) is valid for all i, j, l, k we conclude that for any i, j the angle between $f_i Y_n$ and $f_j Y_n$ goes to 0 at an exponential rate. Combining Theorems 8.1 and 8.2 with the foregoing we find

THEOREM 8.3. *Under the hypothesis that μ is of class B_∞ , the angles between the various rows and the angles between the various columns of $X_n X_{n-1} \cdots X_1$ go to 0 at an exponential rate as $n \rightarrow \infty$, with probability 1.*

Let us now consider P_t^{m-1} as a G -space with G acting on the right. P_t^{m-1} then carries a unique stationary measure π for μ . We will then have $\pi * \mu = \pi$ and from this it follows (since μ is absolutely continuous) that π is not concentrated on a linear subvariety of P_t^{m-1} . Consider the corresponding stationary process on P_t^{m-1} . Since P_t^{m-1} is a boundary, this is a proper μ -process (§3). Combining this with Lemma 1.3 we conclude that, with probability 1, $\pi X_n X_{n-1} \cdots X_1$ converges to a point measure δ_W where W is a random variable on P_t^{m-1} . Now the rows of $Y_n = X_n X_{n-1} \cdots X_1$ are almost proportional (Theorem 8.3) and so for all elements $u \in P_t^{m-1}$ but for those close to a linear subvariety, the direction of $u Y_n$ will be close to that of any row of Y_n . Since π is not concentrated on a linear subvariety and $\pi Y_n \rightarrow \delta_W$ we conclude that for any e_i , $e_i Y_n \rightarrow W$. Finally since the basis $\{e_i\}$ is arbitrary we may state

THEOREM 8.4. *Let μ be of class B_∞ and let $u \in P_t^{m-1}$. There is a random variable W defined in terms of X_1, X_2, X_3, \dots such that with probability 1,*

$$(8.9) \quad uX_nX_{n-1} \cdots X_1 \rightarrow W \in P_t^{m-1}.$$

Moreover, W does not depend upon u .

REMARK. It is obvious that for $v \in P^{m-1}$, $X_nX_{n-1} \cdots X_1v$ will not converge in general. For if $\{Z_n\}$ is the stationary μ -process on P^{m-1} , then $X_nX_{n-1} \cdots X_1Z_0 = Z_n$. Thus the sequence $X_nX_{n-1} \cdots X_1v$ in general oscillates in P^{m-1} .

8.3. RATE OF GROWTH OF THE VECTORS $X_nX_{n-1} \cdots X_1u$. It follows from Theorem 8.3 and the fact that the matrices $X_nX_{n-1} \cdots X_1$ are unimodular that the norms of at least some of the column vectors in $X_nX_{n-1} \cdots X_1$ must be growing at an exponential rate. We shall find that this is the case quite generally and for a general distribution μ on $SL(m, R)$, the μ -processes on $R^m - \{0\}$ tend to ∞ at an exponential rate.

Again we let μ be a measure on $SL(m, R)$. This time, however, we let G denote the smallest closed subgroup of $SL(m, R)$ containing the support of μ . μ will not be required to be absolutely continuous. As usual $\{X_n\}$ denotes the G -valued variables with distribution μ . We are interested in the action of G on R^m .

Consider the M -cocycle on $R^m - \{0\}$ defined by

$$(8.10) \quad \sigma(g, u) = \|gu\| / \|u\|.$$

Note that $\sigma(g, u)$ depends only on g and $\bar{u} \in P^{m-1}$. We can therefore define an M -cocycle on P^{m-1} by $\sigma_1(g, \bar{u}) = \sigma(g, u)$. We set $\rho = \log \sigma$, $\rho_1 = \log \sigma_1$.

Let us consider the stationary measures on P^{m-1} for μ . These form a convex set whose extremals correspond to the ergodic stationary μ -processes on P^{m-1} (see the proof of Theorem 2.6). More precisely we can say if $\{Z_n\}$ is the corresponding stationary μ -process and $\{X_n\}$ the associated G -valued variables, that $\{(X_{n+1}, Z_n)\}$ is an ergodic process. We shall call the extremal measures ergodic.

LEMMA 8.7. *Suppose that G is irreducible and that $\int \log \|g\| d\mu(g) < \infty$ and let π be an ergodic stationary measure for μ on P_{m-1} . Then there exist m linearly independent vectors u_1, \dots, u_m in R^m such that, with probability 1,*

$$(8.11) \quad n^{-1} \log \|X_n \cdots X_1 u_i\| \rightarrow \int_G \int_{P^{m-1}} \rho_1(g, \xi) d\mu(g) d\pi(\xi)$$

as $n \rightarrow \infty$, $i = 1, \dots, m$.

Proof. Denote the right hand side of (8.11) by α . If $\{X_n, Z_n\}$ are as before, we have $X_{n+1}Z_n = Z_{n+1}$. Then

$$(8.12) \quad n^{-1} \rho_1(X_n \cdots X_1, Z_0) = n^{-1} \sum_{k=0}^{n-1} \rho_1(X_{k+1}, Z_k).$$

Now $|\rho_1(X_{k+1}, Z_k)| \leq \log \|X_{k+1}\|$ so that we may apply the ergodic theorem and we find that the expression in (8.12) converges to α , with probability 1. This means that for almost all $\xi \in P^{m-1}$ (with respect to π), $n^{-1} \rho_1(X_n \cdots X_1, \xi) \rightarrow \alpha$. Now G is irreducible and we have $\mu * \pi = \pi$. If π had its support in a linear subvariety of P^{m-1} , the support of μ would leave fixed a linear subvariety and G could not be irreducible. Hence the support of π is not contained in a proper linear subvariety and so there exist independent ξ_1, \dots, ξ_m with

$$n^{-1} \rho_1(X_n \cdots X_1, \xi_i) \rightarrow \alpha.$$

Now let $\xi_i = \bar{u}_i$, $u_i \in R^m$; then $\rho_1(g, \xi_i) = \log(\|gu_i\|/\|u_i\|)$. We thus have $n^{-1} \log \|X_n \cdots X_1 u_i\| \rightarrow \alpha$.

THEOREM 8.5. *If G is irreducible and $\int \|g\| d\mu(g) < \infty$, the expression $\int \int \rho_1(g, \xi) d\mu(g) d\pi(\xi)$ for a stationary measure π for μ is independent of the stationary measure π . Denoting the common value by $\alpha_\mu(\rho_1)$, we have, with probability 1,*

$$(8.13) \quad n^{-1} \log \|X_n \cdots X_1 u\| \rightarrow \alpha_\mu(\rho_1)$$

or all nonzero vectors $u \in R^m$.

Proof. The second assertion is equivalent to the statement that $n^{-1} \rho_1(X_n \cdots X_1, \xi) \rightarrow \alpha_\mu(\rho_1)$ for all $\xi \in P^{m-1}$. In this form it follows from Lemma 7.3 once it is established that the expression denoted by $\alpha_\mu(\rho_1)$ is independent of the stationary measure π . Now let π_1 be one ergodic stationary measure for μ on P^{m-1} . By Lemma 8.7,

$$n^{-1} \log \|X_n \cdots X_1 u\| \rightarrow \iint \rho_1(g, \xi) d\mu(g) d\pi_1(\xi)$$

for a basis of R^m . But then it follows that for all $u \in R^m$,

$$(8.14) \quad \limsup_{n \rightarrow \infty} n^{-1} \log \|X_n \cdots X_1 u\| \leq \iint \rho_1(g, \xi) d\mu(g) d\pi_1(\xi).$$

Now let π_2 be any other ergodic stationary measure, and apply Lemma 8.7. We then see that

$$(8.15) \quad \iint \rho_1(g, \xi) d\mu(g) d\pi_2(\xi) \leq \iint \rho_1(g, \xi) d\mu(g) d\pi_1(\xi).$$

Interchanging the roles of π_1 and π_2 we find that in (8.15), equality must hold. Then $\int \int \rho_1(g, \xi) d\mu(g) d\pi(\xi)$ is the same for all ergodic—or extremal—stationary measures, and hence for all stationary measures.

The rest of our discussion will be devoted to showing that if G is noncompact, $\alpha_\mu(\rho_1)$ is strictly positive. This will show that the vectors $X_n \cdots X_1$ grow exponentially under our hypotheses. Note that the condition of noncompactness of G alone is not sufficient. For example, if $G = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$ then $X_n \cdots X_1 u$ will not be of exponential growth because G is not irreducible.

Let ω denote the K -invariant probability measure on P^{m-1} . In §7.4 we observed that $(dg^{-1}\omega/d\omega)(\xi)$ was an M -cocycle of $SL(m, R)$ on P^{m-1} . The next lemma gives the relationship between this cocycle and the cocycle $\sigma_1(g, \xi)$ defined above.

LEMMA 8.8. $\sigma_1(g, \xi)^m (dg^{-1}\omega/d\omega)(\xi) = 1$.

The proof is straightforward and we omit it.

LEMMA 8.9. Suppose that a stationary measure π for μ on P^{m-1} is equivalent to ω in the sense of absolute continuity ($\pi < \omega, \omega < \pi$). Then

$$(8.16) \quad \alpha_\mu(\rho_1) = -\frac{1}{m} \iint \log \frac{dg^{-1}\pi}{d\pi}(\xi) d\mu(g) d\pi(\xi).$$

Proof. By Lemma 8.8, $\rho_1(g, \xi) = -(1/m) \log (dg^{-1}\omega/d\omega)(\xi)$. (8.16) will follow if we can replace $dg^{-1}\omega/d\omega$ by $dg^{-1}\pi/d\pi$ in the integral. The difference between the corresponding expressions is

$$\begin{aligned} & \iint \log \frac{dg^{-1}\omega}{d\omega}(\xi) \frac{d\pi}{dg^{-1}\pi}(\xi) d\mu(g) d\pi(\xi) \\ &= \iint \log \frac{dg^{-1}\omega}{dg^{-1}\pi}(\xi) d\mu(g) d\pi(\xi) - \iint \log \frac{d\omega}{d\pi}(\xi) d\mu(g) d\pi(\xi) \\ &= \iint \log \frac{d\omega}{d\pi}(g\xi) d\mu(g) d\pi(\xi) - \iint \log \frac{d\omega}{d\pi}(\xi) d\mu(g) d\pi(\xi) \\ &= \iint \log \frac{d\omega}{d\pi}(\eta) d\mu(g) d(\mu * \pi)(\eta) - \iint \log \frac{d\omega}{d\pi}(\xi) d\mu(g) d\pi(\xi) = 0 \end{aligned}$$

since $\mu * \pi = \pi$.

COROLLARY. If a stationary measure π for μ on P^{m-1} is equivalent to ω , then $\alpha_\mu(\rho_1) > 0$ unless G leaves fixed the measure π .

Proof. For

$$\int \log \frac{dg^{-1}\pi}{d\pi}(\xi) d\pi(\xi) \leq \log \int \frac{dg^{-1}\pi}{d\pi}(\xi) d\pi(\xi) = 0$$

by Jensen's inequality, and equality can only hold if the integrand of the second integral is constant almost everywhere with respect to π . But in this case

$g^{-1}\pi = \pi$. If this holds for almost all g in the support of μ it must hold for all of G and we see that G leaves π fixed.

Let $\|\pi\|$ denote the norm of the measure π on P^{m-1} as an element of the dual to the Banach space of continuous functions on P^{m-1} .

LEMMA 8.10. *If π_1 and π_2 are probability measures on P^{m-1} then there is a constant $c > 0$ such that*

$$(8.17) \quad \|\pi_1 - \pi_2\| < c \left\{ - \int \log \frac{d\pi_2}{d\pi_1} d\pi_1 \right\}^{1/2}.$$

(If π_2 is not absolutely continuous with respect to π_1 the right hand side is ∞ taking $\log 0 = -\infty$.)

Proof. $\|\pi_1 - \pi_2\| = \sup | \int (\phi d\pi_1 - \phi d\pi_2) |$ where $|\phi| \leq 1$. We may assume that $\pi_2 < \pi_1$, so we have

$$\begin{aligned} \|\pi_1 - \pi_2\| &= \sup \left| \int \phi \left(1 - \frac{d\pi_2}{d\pi_1} \right) d\pi_1 \right| = \int \left| 1 - \frac{d\pi_2}{d\pi_1} \right| d\pi_1 \\ &= \int \left| 1 - \sqrt{\frac{d\pi_2}{d\pi_1}} \right| \left| 1 + \sqrt{\frac{d\pi_2}{d\pi_1}} \right| d\pi_1 \\ &\leq \left\{ \int \left(1 - \sqrt{\frac{d\pi_2}{d\pi_1}} \right)^2 d\pi_1 \int \left(1 + \sqrt{\frac{d\pi_2}{d\pi_1}} \right)^2 d\pi_1 \right\}^{1/2} \\ &= \left\{ \left(2 - 2 \int \sqrt{\frac{d\pi_2}{d\pi_1}} d\pi_1 \right) \left(2 + 2 \int \sqrt{\frac{d\pi_2}{d\pi_1}} d\pi_1 \right) \right\}^{1/2} \\ &= 2 \left\{ 1 - \left[\int \sqrt{\frac{d\pi_2}{d\pi_1}} d\pi_1 \right]^2 \right\}^{1/2}. \end{aligned}$$

Since, in any case, $\|\pi_1 - \pi_2\| \leq 2$, it is only necessary to establish that (8.17) holds as $\int \log(d\pi_2/d\pi_1) d\pi_1 \rightarrow 0$. Now

$$\begin{aligned} \int \sqrt{\frac{d\pi_2}{d\pi_1}} d\pi_1 &\geq \exp \int \log \sqrt{\frac{d\pi_2}{d\pi_1}} d\pi_1, \\ 1 - \left[\int \sqrt{\frac{d\pi_2}{d\pi_1}} d\pi_1 \right]^2 &\leq 1 - \left[\exp \frac{1}{2} \int \log \frac{d\pi_2}{d\pi_1} d\pi_1 \right]^2 \sim - \int \log \frac{d\pi_2}{d\pi_1} d\pi_1, \end{aligned}$$

and this establishes the lemma.

THEOREM 8.6. *Let G be a noncompact subgroup of $SL(m, R)$ such that no subgroup of G of finite index is reducible. Then $\alpha_\mu(\rho_1) > 0$ and with probability 1, $\|X_n \cdots X_1 u\|$ grows exponentially as $n \rightarrow \infty$ for all $u \in R^m - \{0\}$.*

Proof. Let us first show that under the present hypotheses, G cannot leave fixed a measure on P^{m-1} . For since G is noncompact we can find a sequence

$\{g_n\}$ in G and two linear subvarieties $V_1, V_2 \subset P^{m-1}$ such that for all $\xi \notin V_1, g_n \xi \rightarrow V_2$. This implies that if π is an invariant measure on P^{m-1} for G then π must be concentrated on $V_1 \cup V_2$. Consider finite unions of proper linear subvarieties of P^{m-1} with π -measure unity. $V_1 \cup V_2$ is such a union. Let $W_1 \cup W_2 \cup \dots \cup W_k$ be a minimal such finite union in the sense of set-theoretic inclusion. Since G preserves π each $\bigcup_j g W_j$ again has this property. The intersection $\bigcup W_i \cap g W_j$ would comprise a smaller set with this property unless the two are identical which is only possible if each g permutes the varieties W_i . Then the subgroup of g taking $W_i \rightarrow W_i$ is of finite index in G and is reducible contrary to hypothesis. Thus G has no fixed measure on P^{m-1} .

Now choose a sequence of measures $\{\mu_n\}$ on $SL(m, R)$ with $\mu_n \rightarrow \mu$ weakly and with the property that each μ_n has a stationary measure π_n on P^{n-1} such that π_n is equivalent to ω . For this it is only necessary to choose μ_n as a measure given by a nowhere vanishing density in $SL(m, R)$. We may assume that π_n converges on P^{m-1} and the limit measure π will be a stationary measure for μ since one verifies easily that $\mu_n \rightarrow \mu, \pi_n \rightarrow \pi$ implies $\mu_n * \pi_n \rightarrow \mu * \pi$. We have, moreover, $\alpha_{\mu_n}(\rho_1) \rightarrow \alpha_{\mu}(\rho_1)$. To prove the theorem it will therefore be sufficient to prove that

$$\liminf_{n \rightarrow \infty} \alpha_{\mu_n}(\rho_1) > 0.$$

Since π_n is equivalent to ω we may apply Lemma 8.9, and what we shall show is that

$$(8.17) \quad \liminf_{n \rightarrow \infty} \left\{ - \iint \log \frac{dg^{-1} \pi_n}{d\pi_n}(\xi) d\mu_n(g) d\pi_n(\xi) \right\} > 0.$$

Finally, by Lemma 8.10, this will be the case unless for some subsequence, which we take to be the original sequence, we have

$$(8.18) \quad \int \|g^{-1} \pi_n - \pi_n\|^2 d\mu_n(g) \rightarrow 0.$$

Suppose now that (8.18) were valid. Since $\mu_n \rightarrow \mu$, it follows from (8.18) that if g is in the support of μ , there is a sequence $g_n \rightarrow g$ such that $\|g_n \pi_n - \pi_n\| = \|\pi_n - g_n^{-1} \pi_n\| \rightarrow 0$ (since any neighborhood of g has μ_n -measure $> \varepsilon > 0$ for some ε and for n larger than some n_0). Now let ϕ be a continuous function on P^{m-1} with $|\phi| \leq 1$. Then

$$|g \pi_n(\phi) - \pi_n(\phi)| \leq |g \pi_n(\phi) - g_n \pi_n(\phi)| + \|g_n \pi_n - \pi_n\|.$$

Now $g \pi_n(\phi) = \pi_n(L_g \phi)$ where $L_g \phi(\xi) = \phi(g\xi)$. Then

$$(8.19) \quad |g \pi_n(\phi) - \pi_n(\phi)| \leq \|L_g \phi - L_{g_n} \phi\|_{\infty} + \|g_n \pi_n - \pi_n\|.$$

Since $g_n \rightarrow g, L_{g_n} \phi \rightarrow L_g \phi$ uniformly, and both parts of the right side of (8.19)

go to 0. We conclude that $|g\pi(\phi) - \pi(\phi)| = \lim |g\pi_n(\phi) - \pi_n(\phi)| = 0$ and $g\pi = \pi$. Thus if (8.18) were true, then every g in the support of μ , and hence all of G , would leave π fixed. We have already shown that this will not happen. This completes the proof of the theorem.

COROLLARY. *If G is a semi-simple group with no compact factors and Λ is any irreducible representation of G , then for any vector v in the representation space of Λ , $\|\Lambda(X_n X_{n-1} \cdots X_1)v\|$ grows exponentially with probability 1, as $n \rightarrow \infty$.*

BIBLIOGRAPHY

1. S. Banach, *Théorie des opérations linéaires*, Monogr. Mat., Tom I, Warsaw, 1932.
2. R. Bellman, *Limit theorem for non-commutative operations*. I, Duke Math. J. **21** (1954), 491–500.
3. A. Borel and Harish-Chandra, *Arithmetic subgroups of algebraic groups*, Ann. of Math. (2) **75** (1962), 485–535.
4. L. Breiman, *The strong law of large numbers for a class of Markov chains*, Ann. Math. Statist. **31** (1960), 801–803.
5. F. Bruhat, *Sur les représentations induites des groupes de Lie*, Bull. Soc. Math. France **84** (1956), 97–205.
6. J. L. Doob, *Stochastic processes*, Wiley, New York, 1953.
7. N. Dunford and J. T. Schwartz, *Linear operators*, Interscience, New York, 1958.
8. H. Furstenberg, *A Poisson formula for semi-simple Lie groups*, Ann. of Math. (2) **77** (1963), 335–386.
9. H. Furstenberg and H. Kesten, *Products of random matrices*, Ann. Math. Statist. **31** (1960), 457–469.
10. I. Gel'fand, *Spherical functions on symmetric Riemann spaces*, Dokl. Akad. Nauk SSSR **70** (1950), 5–8. (Russian).
11. R. Godement, *A theory of spherical functions*. I, Trans. Amer. Math. Soc. **73** (1952), 496–556.
12. U. Grenander, *Some non-linear problems in probability theory*, Probability and Statistics, The Harold Cramer Volume, pp. 108–129, Wiley, New York, 1959.
13. S. Helgason, *Differential geometry and symmetric spaces*, Academic Press, New York, 1962.
14. G. D. Mostow, *Homogeneous spaces with finite invariant measure*, Ann. of Math. (2) **75** (1962), 17–31.
15. T. Watanabe, *On the theory of Martin boundaries induced by countable Markov processes*, Mem. Coll. Sci. Univ. Kyoto Ser. A Math. **33** (1960/61), 39–108.

UNIVERSITY OF MINNESOTA,
MINNEAPOLIS, MINNESOTA