LINEAR FRACTIONAL TRANSFORMATIONS MOD ONE AND ERGODIC THEORY

by

Stephan Martin Rudolfer

A Thesis submitted for the degrees of Doctor of Philosophy in the University of London and Diploma of Membership of the Imperial College.

Department of Mathematics,
Imperial College,
London, S.W.7.

August 1968.

ABSTRACT

After an introductory chapter, we study characterisations of mixing, weak mixing and ergodicity of a finite measure-preserving transformation T due to N. Oishi [25]. These characterisations are in terms of convergence of suitably defined entropies of finite partitions. We show that the characterisations can be given in terms of (countable) partitions with finite entropy, extend the characterisation to mixing of degree r and give further characterisations in terms of convergence of the suitably defined measures of Jordan measurable sets and, in the case of a compact measure space, in terms of weak convergence of these measures. It is shown that these characterisations cannot be extended to convergence of the corresponding entropies of TxT nor to all measurable subsets, respectively.

Chapter III studies the ergodic properties of two classes of linear fractional transformation mod one, which turn out mostly to have similar properties to previously studied f-transformations [29], [32]. The main tool is a sufficient condition for ergodicity of non-singular, many-one transformations of a probability space, which, applied to f-transformations, generalises a similar theorem of A. Rényi [29]. Rényi's theorem states the existence of a finite invariant measure equivalent to Lebesgue measure. In some cases, using a result of W. Parry [27], we have succeeded in constructing this invariant measure. Throughout, results were only obtained for f-transformations with independent digits (in the sense of Rényi). The dependent digit case is much more delicate, and we were unable to obtain results in this direction.

Of particular interest is the ergodic transformation $T_{\frac{x}{1+x}}$ whose σ-finite invariant measure is exhibited. Its associated f-expansions have a striking distribution of digits. It is an open question whether $T_{\frac{x}{1+x}}$ is exact and what value its entropy [22] takes.

In chapter IV the isomorphism problem for irreducible, null recurrent and aperiodic Markov shifts is studied using a necessary and sufficient condition for ergodicity due to Kakutani and Parry [13] and the divergence properties of certain renewal sequences. The latter provide metric invariants which are then used to investigate three classes of Markov shift, it being shown that they each consist of a continuum of non-isomorphic transformations. Non-isomorphism between the three classes is also discussed. A generalised Hopf ergodic theorem is proved as a corollary to the methods developed.

ACKNOWLEDGEMENT

My thanks go to my supervisor, Dr. Y.N. Dowker, to Dr. W. Parry for their help, guidance and criticism, and to my wife for her constant encouragement, during the preparation of this thesis.

CONTENTS

CHAPTER I

Preliminary Definitions and Results

§1 Measure Spaces

If A is a set and x is an element of A then we write $x \in A$. If x is not an element of A then we write $x \notin A$. If A is a subset of B then we write $A \subset B$ or $B \supset A$. The union, intersection, difference and symmetric difference of two sets A and B are denoted by $A \cup B$, $A \cap B$, $A - B$ and $A + B$ respectively. If we are studying the subsets A of a set X then we write $\mathcal{C} A$ for $X - A$, the complement of A. The empty set is denoted by ϕ ; sets A and B are said to be <u>disjoint</u> if $A \cap B = \phi$. $\{x\}$ denotes the <u>point set</u> x.

Unless otherwise stated, the same notation as that relating to sets will also be used when dealing with classes or families of sets.

1.1 <u>Definition</u> If X is a non-empty set and $\mathcal{R}$ is a non-empty collection of subsets of X then $\mathcal{R}$ is a <u>ring</u> if $E \in \mathcal{R}$ and $F \in \mathcal{R}$ imply $E - F \in \mathcal{R}$ and $E \cup F \in \mathcal{R}$.

An <u>algebra</u> is a ring $\mathcal{R}$ such that $X \in \mathcal{R}$.

A <u>σ-ring</u> $\mathcal{S}$ is a ring such that $E_n \in \mathcal{S}$ $(n = 1, 2, \ldots)$ imply $\bigcup_{n=1}^{\infty} E_n \in \mathcal{S}$.

A <u>σ-algebra</u> $\mathcal{B}$ is a σ-ring such that $X \in \mathcal{B}$. If $\mathcal{A}, \mathcal{B}$ are σ-algebras we write $\mathcal{A} \subset \mathcal{B}$ if $A \in \mathcal{A}$ implies $A \in \mathcal{B}$.

Since the intersection of an arbitrary class of σ-algebras is again a σ-algebra, the σ-algebra <u>generated</u> by a class of subsets of X is uniquely defined as the intersection of all σ-algebras of subsets of X containing this class. It always exists since the σ-algebra of all subsets of X contains every class of subsets of X.

1.2 <u>Definition</u> The pair $(X, \mathcal{B})$, where X is a non-empty set and $\mathcal{B}$ is a σ-algebra of subsets of X, is a <u>measurable space</u>. Elements of $\mathcal{B}$ are <u>measurable</u> or <u>Borel</u> subsets of X.

A non-negative, possibly infinite-valued set function μ defined on $(X, \mathcal{B})$ is a <u>measure</u> if

(i) $\mu(\phi) = 0$

(ii) $E_n \in \mathcal{B}$ $(n = 1, 2, \ldots)$, $E_i \cap E_j = \phi$ $(i \neq j)$ imply

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n) .$$

$E \in \mathcal{B}$ is a <u>null set</u> of μ if $\mu(E) = 0$. An important principle of measure theory is that of neglecting null sets. This gives far more generality to results and definitions than would be possible if a purely set-theoretic approach were adopted. Thus, if P(x) is a proposition depending on x, we say P(x) holds <u>almost everywhere</u> if $\{x : P(x) \text{ does not hold}\}$ is a null set. Similarly, $A = B \pmod 0$ or $A = B$ modulo zero mean that $A + B$ is a null set. In particular

$A \subset B \pmod 0$ if $A - B$ is a null set. Applying this to σ-algebras, we say that $\mathcal{A} \subset \mathcal{B} \pmod 0$ if for every $A \in \mathcal{A}$ there exists $B \in \mathcal{B}$ with $A = B \pmod 0$. Hence $\mathcal{A} = \mathcal{B} \pmod 0$ means that $\mathcal{A} \subset \mathcal{B} \subset \mathcal{A} \pmod 0$.

μ is <u>finite</u> if $\mu(X) < \infty$; μ is <u>σ-finite</u> if $\mu(X) = \infty$,

$X = \bigcup_{n=1}^{\infty} X_n$, $X_i \cap X_j = \phi$ $(i \neq j)$ and $\mu(X_n) < \infty$ $(n = 1, 2, \ldots)$.

$\{X_n\}_o^\infty$ which is not unique, is called a <u>σ-finite partition</u> of X.

$(X, \mathcal{B}, \mu)$ is then a <u>finite</u> or <u>σ-finite</u> <u>measure space</u> respectively. It is a <u>probability space</u> (μ is a <u>probability measure</u>) if $\mu(X) = 1$.

A is an <u>atom</u> of $(X, \mathcal{B}, \mu)$ if $A \in \mathcal{B}$, $\mu(A) > 0$ and $B \subset A$, $B \in \mathcal{B}$ implies $\mu(B) = 0$ or $\mu(B) = \mu(A)$. If $(X, \mathcal{B}, \mu)$ has no atoms then it is <u>non-atomic</u>.

If $(X, \mathcal{B}, \mu)$ satisfies this condition:

$A \in \mathcal{B}$, $\mu(A) = 0$, $B \subset A$ implies $B \in \mathcal{B}$,

then $(X, \mathcal{B}, \mu)$ is a <u>complete</u> measure space. There is no loss of generality in assuming, as we do from now on, that $(X, \mathcal{B}, \mu)$ is complete; there exists a unique σ-algebra $\overline{\mathcal{B}} \supset \mathcal{B}$ of subsets of X and a unique measure $\overline{\mu}$ on $(X, \overline{\mathcal{B}})$ such that $(X, \overline{\mathcal{B}}, \overline{\mu})$ is complete, $\overline{\mu}(E) = \mu(E)$ for all $E \in \mathcal{B}$ and for each $\overline{A} \in \overline{B}$ there exists $A \in \mathcal{B}$, $A \supset \overline{A}$ such that $\overline{\mu}(A - \overline{A}) = 0$ [7, p. 55].

A countable collection $A \subset \mathcal{B}$ is a <u>basis</u> for $(X, \mathcal{B}, \mu)$ if the following conditions are satisfied:-

(i) For every $A \in \mathcal{B}$ there is a set B belonging to the σ-algebra generated by Λ such that $A \subset B$ and $\mu(B - A) = 0$.

(ii) For every pair x, y of distinct points in X there is a set $A \in \Lambda$ such that either $x \in A$, $y \notin A$ or $x \notin A$, $y \in A$.

$(X, \mathcal{B}, \mu)$ is <u>separable</u> if it has a basis. Note that the definition of separability applies equally to σ-finite measure spaces, since (i) depends only on null sets of μ, and that countable subsets of a separable measure space are measurable by (ii), since $\{x\} = \bigcap_{A \in \Lambda} A$. If $(X, \mathcal{B}, \mu)$ is separable and non-atomic then $p(\{x\}) = 0$ for all $x \in X$. $(X, \mathcal{B}, \mu)$ is <u>complete</u> with respect to $\Lambda = \{A_i : i \in I\}$ if all intersections

$$\bigcap_{i \in I} B_i \neq \phi ,$$

where $B_i = A_i$ or $X - A_i$. If a probability space $(X, \mathcal{B}, \mu)$ is separable and complete with respect to a basis then it is called a <u>Lebesgue</u> space.

For example, if $Y = [0, 1)$, $\mathcal{L}$ = σ-algebra generated by the ring of finite, disjoint unions of intervals of the form $[a, b)$ and λ is the measure determined by $\lambda([a, b)) = b - a$ then $(Y, \mathcal{L}, \lambda)$ is a Lebesgue space. In fact, all non-atomic Lebesgue spaces are essentially the same as $(Y, \mathcal{L}, \lambda)$ [31]. That measures on a ring can be extended to measures on the generated σ-algebra follows from the next result.

1.3 Theorem [7, p 54]. If ν is a measure on a ring R of subsets of X then there exists a unique measure μ defined on the σ-algebra generated by R such that $\nu(E) = \mu(E)$, $E \in R$.

1.4 Definition $\xi = \{E_n\}_1^N$, where $1 < N \leq \infty$, is a measurable partition (mod 0) of the Lebesgue space $(X, \mathcal{B}, \mu)$ if

(i) $E_n \in \mathcal{B}$, $\mu(E_n) > 0$ $(n = 1, \ldots, N)$

(ii) $X_i \cap X_j = \phi \ (i \neq j)$

(iii) $X = \bigcup_{n=1}^{N} X_n$ (mod 0) .

The set of integers $\{1, \ldots, N\}$ is the index set of ξ . The (at most) countability of ξ is essential: the existence of an uncountable, disjoint class of sets of positive measure whose union is X would contradict the finiteness of μ. We shall consider measurable partitions only of Lebesgue spaces, and refer to them simply as 'partitions'. Two frequently used partitions are ν and ε, where $\nu = \{X\}$ and $\varepsilon = \{\{x\} : x \in X\}$. The latter is not strictly a partition in the sense of 1.4, since all its elements except at most a countable number of atoms are null sets. However, we shall refer to ε as a partition, as it consists of a disjoint class of measurable subsets of X whose union is X.

If $A \subset X$, let $\xi_A = \{B \cap A : B \in \xi\}$ denote the partition of A induced by the partition ξ of X.

A partial ordering on the class of all partitions of $(X, \mathcal{B}, \mu)$ is given by $\leq$, where $\xi \leq \eta$ if η is a refinement of

ξ , i.e. if every element of η is a subset of some element of ξ. We say that $\xi \leq \eta$ (mod 0) if there exists a set $A \in \mathcal{B}$, $A = X$ (mod 0) such that $\xi_A \leq \eta_A$. $\xi = \eta$ (mod 0) is defined similarly. $\xi \leq \eta$ (mod 0) and $\eta \leq \xi$ (mod 0) imply $\xi = \eta$ (mod 0) since $\xi_A \leq \eta_A$ and $\xi_B \leq \eta_B$, $\mu(\mathscr{C}A) = \mu(\mathscr{C}B) = 0$, imply $\xi_{A\cap B} = \eta_{A\cap B}$, $\mu(\mathscr{C}(A\cap B)) = 0$. The smallest partition of X is ν , the largest is ε .

For any collection of partitions $\{\xi_i : i \in I\}$ of X there exists the join $\bigvee_{i\in I} \xi_i$ defined as a partition ξ of X having the properties:-

(i) $\xi_i \leq \xi$(mod 0) for all $i \in I$

(ii) if $\xi_i \leq \xi'$ (mod 0) for all $i \in I$ where ξ' is a partition then $\xi \leq \xi'$ (mod 0) .

Similarly there exists the meet $\bigwedge_{i\in I} \xi_i$ defined as a measurable partition ξ having the properties:-

(i) $\xi_i \geq \xi$ (mod 0) for all $i \in I$

(ii) if $\xi_i \geq \xi'$ (mod 0) for all i where ξ' is a partition then $\xi \geq \xi'$ (mod 0) .

The join and meet of two partitions ξ and η are written $\xi \vee \eta$ and $\xi \wedge \eta$ respectively. It is easy to verify that

$$\bigvee_{i=1}^{n} \xi_i = \{ \bigcap_{i=1}^{n} A_i : A_i \in \xi_i \} .$$

We write $\xi_n \nearrow \xi$ $(n \longrightarrow \infty)$ if $\xi_n \leq \xi_{n+1}$, $\xi = \bigvee_{n=1}^{\infty} \xi_n$ and $\xi_n \searrow \xi$ $(n \longrightarrow \infty)$ if $\xi_n \geq \xi_{n+1}$, $\xi = \bigwedge_{n=1}^{\infty} \xi_n$.

Let $\hat{\xi}$ denote the sub-σ-algebra of $\mathcal{B}$ generated by the partition ξ of $(X, \mathcal{B}, \mu)$. Then [32] $\hat{\xi} \subset \hat{\eta}$ if, and only if, $\xi \leq \eta$. Also $\hat{\varepsilon} = \mathcal{B}$.

For r = 1, 2, ..., $(X^{(r)}, \mathcal{B}^{(r)}, \mu^{(r)})$ denotes the <u>r-fold direct product</u> of the measure space $(X, \mathcal{B}, \mu)$ with itself, i.e.

$$X^{(r)} = X \times \ldots \times X \quad (r \text{ times})$$

$\mathcal{B}^{(r)}$ = σ-algebra generated by the ring of finite, disjoint unions of sets of the form $B_1 \times \ldots \times B_r$ $(B_i \in \mathcal{B})$

$\mu^{(r)}$ = measure uniquely determined (see 1.3) by

$$\mu^{(r)}(B_1 \times \ldots \times B_r) = \mu(B_1) \ldots \mu(B_r) . \quad (B_i \in \mathcal{B}) .$$

A <u>measurable rectangle</u> in $\mathcal{B}^{(r)}$ is a set of the form $E_1 \times \ldots \times E_r$, $E_i \in \mathcal{B}$ $(i = 1, \ldots, r)$.

$\mathcal{R}^{(r)}$ = ring of finite, disjoint unions of measurable rectangles in $\mathcal{B}^{(r)}$.

$E \in \mathcal{B}^{(r)}$ is <u>Jordan measurable</u> if for all $\delta > 0$ there exist R_δ, $S_\delta \in \mathcal{R}^{(r)}$, $R_\delta \subset E \subset S_\delta$, such that

$$\mu(E - R_\delta) < \delta \text{ and } \mu(S_\delta - E) < \delta \,\partial$$

If $\xi_1, \ldots, \xi_r$ are partitions of X then $\xi_1 \times \ldots \times \xi_r$ stands for the partition $\{E_1 \times \ldots \times E_r : E_i \in \xi_i\}$ of $X^{(r)}$. In particular $S(\xi, \xi')$ denotes the sub-σ-algebra of $\mathcal{B}^{(2)}$ generated by

$\xi \times \xi'$ and $\mu^{\xi,\xi'}$ denotes the restriction of $\mu^{(2)}$ to $S(\xi,\xi')$, i.e. the measure uniquely defined on $(X^{(2)}, S(\xi,\xi'))$ by

$$\mu^{\xi,\xi'}(E) = \mu(E) \ , \ E \in S(\xi,\xi') \ .$$

Note that $\mu^{\xi,\xi'}$ has atoms $E \times F$, $E \in \xi$, $F \in \xi'$.

If $E \in \mathcal{B}^{(r)}$ then

$$\pi(E) = \{x \in X : \text{for some } (x_2,\ldots, x_r) \in X^{(r-1)}, (x_1,x_2,\ldots,x_r) \in E\}$$

$$E_{x_1} = \{(x_2,\ldots, x_r) \in X^{(r-1)} : (x_1,x_2,\ldots,x_r) \in E\} \ .$$

§2 Measurable Functions and Absolute Continuity

Let $(X, \mathcal{B}, \mu)$ be a measure space.

2.1 Definition If $E \subset X$ the characteristic function $\chi_E(x)$ of E is defined as follows:-

$$\chi_E(x) = \begin{cases} 1 \ , & x \in E \\ 0 \ , & x \notin E. \end{cases}$$

f is a simple function if $f(x) = \sum_{i=1}^{n} a_i \chi_{E_i}(x)$, where a_i are real.

f is an elementary function if $f(x) = \sum_{i=1}^{\infty} a_i \chi_{E_i}(x)$, where a_i are real.

2.2 Definition A real-valued function f on X is measurable if for all Borel subsets H of $[-\infty, \infty]$ $f^{-1}(H) \in \mathcal{B}$.

$\chi_E(x)$ is measurable if, and only if, $E \in \mathcal{B}$. The sum of two measurable functions and multiples of a measurable function are measurable. For further facts about measurable functions, see [7] or [19].

2.3 <u>Definition</u> If ν is another measure on $(X, \mathcal{B})$ then μ is <u>absolutely continuous</u> with respect to ν, $\mu \ll \nu$, if $\mu(E) = 0$ whenever $\nu(E) = 0$. μ is <u>equivalent</u> to ν if $\mu \ll \nu \ll \mu$.

2.4 <u>Theorem</u> (Radon-Nikodym) [7] [19]. If the measures μ and ν are defined on $(X, \mathcal{B})$ and μ is absolutely continuous with respect to ν then there exists a finite valued measurable function f such that

$$\mu(E) = \int_E f(x)\, d\,\nu(x) \;, \quad E \in \mathcal{B} \;.$$

f is called the <u>Radon-Nikodym derivative</u> of μ with respect to ν.

Two points on notation:- $\{f \in E\}$ will sometimes be used instead of $\{x : f(x) \in E\}$; $\inf_X f(x)$ is to be understood as $\inf_{x \in X} f(x)$.

§3 <u>Integrable Functions</u>

Let $(X, \mathcal{B}, \mu)$ be a measure space.

3.1 <u>Definition</u> A measurable elementary function $f(x) = \sum_{i=1}^{\infty} a_i \chi_{E_i}(x) \geq 0$ is <u>integrable</u> if $\sum_{i=1}^{\infty} a_i \mu(E_i) < \infty$ and its integral is then written $\int_X f(x)\, d\,\mu(x) = \sum_{i=1}^{\infty} a_i \mu(E_i)$.
In writing integrals we sometimes adopt the convention that $\int f\, d\,\mu$ stands for $\int_X f(x)\, d\mu(x)$.

3.2 <u>Definition</u> An elementary function $f \geq 0$ is <u>integrable on E</u>, $E \in \mathcal{B}$ if $\chi_E(x)\, f(x)$ is integrable; we write $\int \chi_E\, f\, d\mu = \int_E f\, d\mu$.

3.3 <u>Definition</u> A measurable function $f(x) \geq 0$ is <u>integrable on E</u>, $\mu(E) < \infty$, if there exists a sequence of elementary functions $f_n(x) \geq 0$ such that $f_n(x) \nearrow f(x)$ $(n \longrightarrow \infty)$ uniformly on E and $f_n(x)$ is integrable on E $(n = 1, 2, \ldots)$. Its integral over E

$$\int_E f(x)\, d\mu = \lim_{n \to \infty} \int_E f_n(x)\, d\mu \ .$$

3.4 <u>Definition</u> A measurable function $f(x)$ is <u>integrable on E</u>, $\mu(E) < \infty$, if $f^+(x)$ and $f^-(x)$ are integrable on E, where

$$f^+(x) = \begin{cases} f(x) \ , & \text{if } f(x) \geq 0 \\ 0 \ , & \text{if } f(x) < 0 \end{cases}$$

$$f^-(x) = \begin{cases} -f(x) \ , & \text{if } f(x) \leq 0 \\ 0 \ , & \text{if } f(x) > 0 \ , \end{cases}$$

and then $\int f\, d\mu = \int f^+\, d\mu - \int f^-\, d\mu$.

3.5 <u>Definition</u> If $(X, \mathcal{B}, \mu)$ is σ-finite with $X = \bigcup_1^\infty X_n$, $\mu(X_n) < \infty$, then f is <u>integrable</u> if it is integrable on each X_n and $\sum_1^\infty \left| \int_{X_n} f\, d\mu \right| < \infty$ in which case we write

$$\int_X f \, d\mu = \sum_1^\infty \int_{X_n} f \, d\mu .$$

For more details of this approach to integration theory see [21] .

Write $L_p(\mu) = \{f : |f|^p$ is integrable on $(X, \mathcal{B}, \mu)\}$, $p \geq 1$.

$L_p(\mu)$ is a Banach algebra with norm

$$\|f\|_p = \left(\int |f|^p \, d\mu\right)^{1/p} ,$$

provided that functions which are equal almost everywhere are identified; $\int$ is a linear functional on $L_p(\mu)$, i.e. $f, g \in L_1(\mu)$ imply $\int(\alpha f + \beta g) \, d\mu = \alpha \int f \, d\mu + \beta \int g \, d\mu$ for all real α, β.

3.6 Theorem: Hölder's Inequality [19]. Let $p > 1$, $q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$, $f \in L_p(\mu)$, $g \in L_q(\mu)$. Then $f g \in L_1(\mu)$ and $\|f g\|_1 \leq \|f\|_p \, \|g\|_q$.

3.7 <u>Theorem</u> If $f \in L_1(\mu)$, $g \in L_1(\mu)$ and $\int_E f \, d\mu = \int_E g \, d\mu$ for all $E \in \mathcal{B}$ then $f = g$ almost everywhere.

<u>Proof</u> Since $\int$ is a linear functional on $L_1(\mu)$, it is sufficient to prove that $\int_E f \, d\mu = 0$, $E \in \mathcal{B}$, implies $f = 0$ almost everywhere.

Let $E_{k,n} = \{x : \frac{k}{2^n} < f(x) < \frac{k+1}{2^n}\}$, $k = 0, \pm 1, \ldots, n = 0, 1, \ldots$.

Then $E_{k,n} \in \mathcal{B}$ since f is measurable and

$$0 = \int_{E_{k,n}} f \, d\mu \geq \frac{k}{2^n} \mu(E_{k,n}) ,$$

i.e. $\mu(E_{k,n}) = 0 \quad k > 0, n = 0, 1, \ldots$. Similarly,

$\mu(E_{k,n}) = 0 \quad k < -1$. Since $\{x : f(x) \neq 0\} = \bigcup_{\substack{k=-\infty \\ k \neq 0,-1}}^{\infty} \bigcup_{n=0}^{\infty} E_{k,n}$

the result follows.

3.8 <u>Definition</u> If $\{f_n(x)\}_0^\infty$ is a sequence of integrable functions and $f(x) \in L_1(\mu)$ then we say that

$f_n \longrightarrow f \ (n \longrightarrow \infty)$ <u>(pointwise)</u> if for each $x \in X$ and $\delta > 0$ there exists $n_0(\delta, x)$ such that $n > n_0$ implies $|f_n(x) - f(x)| < \delta$.

$f_n \longrightarrow f \ (n \longrightarrow \infty)$ <u>almost everywhere</u> if $\{x : f_n(x) \not\longrightarrow f(x) \ (n \longrightarrow \infty)\}$ is a null set.

$f_n \longrightarrow f \ (n \longrightarrow \infty)$ <u>almost uniformly</u> if $\{x : f_n(x) \not\longrightarrow f(x) \ (n \longrightarrow \infty)$ uniformly$\}$ is a null set.

$f_n \longrightarrow f \ (n \longrightarrow \infty)$ <u>in measure</u> (<u>in probability</u> if $\mu(X) = 1$) if $\mu\{x : |f_n(x) - f(x)| \geq \delta\} \longrightarrow 0 \ (n \longrightarrow \infty)$ for each $\delta > 0$.

$f_n \longrightarrow f \ (n \longrightarrow \infty)$ <u>in $L_1(\mu)$ norm</u> if $\|f_n - f\|_1 \longrightarrow 0 \ (n \longrightarrow \infty)$.

The relationships between the various forms of convergence are thoroughly discussed in [19].

3.9 <u>Definition</u> If ξ is a partition of $(X, \mathcal{B}, \mu)$ and $\hat{\xi}$ denotes the generated σ-algebra, a function $f : X \longrightarrow [-\infty, \infty]$ is <u>measurable with respect to $\hat{\xi}$</u> if $f^{-1}(B) \in \hat{\xi}$ for all Borel subsets B of $[-\infty, \infty]$.

3.10 <u>Definition</u> Let $(X, \mathcal{B}, \mu)$ be a probability space, $\mathcal{B}'$ be a sub-σ-algebra of $\mathcal{B}$ and $0 \leq f \in L_1(\mu)$. Then the set function $\mu_f(B)$ defined by $\mu_f(B) = \int_B f \, d\mu$ $(B \in \mathcal{B}')$ is a finite measure on $\mathcal{B}'$ which is absolutely continuous with respect to the restriction of μ to $\mathcal{B}'$. Hence by 2.4 there exists $E(f|\mathcal{B}') \in L_1(\mu)$ which is measurable with respect to $\mathcal{B}'$ such that

$$\int_B f \, d\mu = \int_B E(f|\mathcal{B}') \, d\mu , \quad B \in \mathcal{B}' .$$

$E(f|\mathcal{B}')$, the <u>conditional expectation</u> of f with respect to $\mathcal{B}'$, is only determined up to a null-set, since any two versions of $E(f|\mathcal{B}')$ are equal almost everywhere (replacing $\mathcal{B}$ by $\mathcal{B}'$ in 3.7).

3.11 <u>Martingale Theorem</u> [2] If $\{\xi_n\}_0^\infty$ is a sequence of partitions of the probability space $(X, \mathcal{B}, \mu)$ such that $\xi_n \nearrow \xi$ $(n \longrightarrow \infty)$ and f is measurable with respect to $\hat{\xi}$ then $E(f|\hat{\xi}_n) \longrightarrow E(f|\hat{\xi})$ $(n \longrightarrow \infty)$.

If $(X_i, \mathcal{B}_i, \mu_i)$ $(i = 1, 2)$ are finite or σ-finite measure sequences then $(X_1 \times X_2, \mathcal{B}_1 \times \mathcal{B}_2, \mu_1 \times \mu_2)$ denotes the direct product of the two measure spaces, i.e.

$\mathcal{B}_1 \times \mathcal{B}_2$ = σ-algebra generated by the ring of finite, disjoint unions of sets of the form $B_1 \times B_2$ $(B_i \in \mathcal{B}_i)$

$\mu_1 \times \mu_2$ = measure uniquely determined (see 1.3) by $\mu_1 \times \mu_2 (B_1 \times B_2) = \mu_1(B_1)\ \mu_2(B_2)$ $(B_i \in \mathcal{B}_i)$.

3.12 Fubini's Theorem [7, p. 148] If $h \in L_1(\mu_1 \times \mu_2)$ then

$\int_{X_2} h\ d\mu_2 \in L_1(\mu_1)$, $\int_{X_1} h\ d\mu_1 \in L_2(\mu_2)$ and

$$\int_{X_1 \times X_2} h\ d(\mu_1 \times \mu_2) = \int_{X_1} \left(\int_{X_2} h\ d\mu_2 \right) d\mu_1 = \int_{X_2} \left(\int_{X_1} h\ d\mu_1 \right) d\mu_2 .$$

§4 Topological Measure Spaces

Let X be a compact, Hausdorff topological space. Measurability and measure are connected with the topology as follows:-

4.1 Definition $\mathcal{B}$ = σ-algebra generated by the compact subsets of X.

μ = a regular measure on $\mathcal{B}$, i.e. for all $E \in \mathcal{B}$

$$\mu(E) = \inf_{\substack{E \subset U \\ U \text{ open}}} u(U)$$

$$= \sup_{\substack{C \subset E \\ C \text{ compact}}} \mu(C)$$

the r-fold direct product of $(X, \mathcal{B}, \mu)$ with itself, $(X^{(r)}, \mathcal{B}^{(r)}, \mu^{(r)})$ is also compact and Hausdorff with respect to the product topology [15], $r = 2, 3, \ldots$.

4.2 Definition $C^{(r)}$ = the Banach algebra of all continuous, real-valued functions f defined on $X^{(r)}$, provided with the uniform topology, i.e. the topology defined by the norm

$$\|f\| = \sup_{\underline{x} \in X^{(r)}} |f(\underline{x})| .$$

$C^{(r)}$ is a sub-algebra of the Banach algebra $L_1(\mu^{(r)})$.

$$\mathcal{A}^{(r)} = \{ f(\underline{x}) = \sum_1^n a_i \, f_i^1(x_1) \ldots f_i^r(x_r) \; : \; a_i \text{ real}, \; f_i^j \in C^{(r)} \text{ and } \underline{x} = (x_1, \ldots, x_r) \} .$$

$\mathcal{A}^{(r)}$ is a subalgebra of $C^{(r)}$.

§5 Transformations

Let $(X_1, \mathcal{B}_1, \mu_1)$ and $(X_2, \mathcal{B}_2, \mu_2)$ be measure spaces.

5.1 Definition The transformation $T : X_1 \longrightarrow X_2$ (X_1 is the domain of T) is measurable if $A \in \mathcal{B}_2$ implies $T^{-1} A \in \mathcal{B}_1$. Tx is the image of $x \in X_1$ under T.

T is non-singular if it is measurable and if $A \in \mathcal{B}_2$, $\mu_2(A) = 0$ implies $\mu_1(T^{-1}A) = 0$.

T is measure-preserving or a homomorphism if it is measurable and if $A \in \mathcal{B}_2$ implies $\mu_2(A) = \mu_1(T^{-1}A)$.

T is <u>one-one</u> if $Tx = Ty$ implies $x = y$. It is <u>many-one</u> if at most a countable number of distinct points can have the same image under T. If T is 1-1 onto and both T and T^{-1} are homomorphisms T is called an <u>isomorphism</u>. $(X_1, \mathcal{B}_1, \mu_1)$ and $(X_2, \mathcal{B}_2, \mu_2)$ are then said to be <u>isomorphic</u>. If the two spaces coincide, homomorphisms are called <u>endomorphisms</u> and isomorphisms, <u>automorphisms</u>. The endomorphism T_2 of $(X_2, \mathcal{B}_2, \mu_2)$ is <u>isomorphic</u> to the endomorphism T_1 of $(X_1, \mathcal{B}_1, \mu_1)$ if there is an isomorphism T from X_1 to X_2 such that $T\,T_1 = T_2\,T$.

If $T : (X_1, \mathcal{B}_1, \mu_1) \longrightarrow (X_2, \mathcal{B}_2, \mu_2)$ then $T^{(r)}$ denotes the transformation from $(X_1^{(r)}, \mathcal{B}_1^{(r)}, \mu_1^{(r)})$ to $(X_2^{(r)}, \mathcal{B}_2^{(r)}, \mu_2^{(r)})$ given by $T^{(r)}(x_1, \ldots x_r) = (Tx_1, \ldots, Tx_r)$. Similarly, if T_i is an endomorphism of $(X_i, \mathcal{B}_i, \mu_i)$ $(i = 1, 2)$ then $T_1 \times T_2$ defined by $T_1 \times T_2\,(x_1, x_2) = (T_1\,x_1, T_2\,x_2)$ is an endomorphism of $(X_1 \times X_2, \mathcal{B}_1 \times \mathcal{B}_2, \mu_1 \times \mu_2)$, with the latter defined analogously to $(X^{(r)}, \mathcal{B}^{(r)}, \mu^{(r)})$.

If $E \subset X_1$ the <u>restriction</u> of T to E, denoted by $T|_E$, is defined by $T|_E\,x = Tx$, $x \in E$. If $T : (X_1, \mathcal{B}_1, \mu_1) \longrightarrow (X_2, \mathcal{B}_2, \mu_2)$ is measurable and non-singular then so is $T|_E$:

$T|_E : (E, E \cap \mathcal{B}_1, \mu_{1E}) \longrightarrow (X_2, \mathcal{B}_2, \mu_2)$, where

$$E \cap \mathcal{B}_1 = \{E \cap B : B \in \mathcal{B}_1\}$$

$$\mu_{1E}(F) = \mu_1(F), \quad F \in E \cap \mathcal{B}_1 .$$

If $T^{-1}E = E \in \mathcal{B}_1$ and T is measure-preserving then so is $T|_E$.

If $(X_i, \mathcal{B}_i, \mu_i)$ are compact Hausdorff measure spaces $(i = 1, 2)$ then T is continuous if $T^{-1} 0_2$ is open for all open sets $0_2 \subset X_2$.

A measurable non-singular transformation T of $(X, \mathcal{B}, \mu)$ is ergodic if $T^{-1} E = E \in \mathcal{B}$ implies $\mu(E) = 0$ or $\mu(X-E) = 0$. Such sets E are invariant under T.

5.2 Theorem [8] T is ergodic if, and only if, $f(Tx) = f(x)$, where f is a measurable function, implies $f(x)$ = constant almost everywhere. Such functions f are invariant under T.

5.3 Theorem If T is an ergodic endomorphism of $(X, \mathcal{B}, \mu)$ which preserves another measure ν equivalent to μ then $\mu = c\,\nu$ where c is a positive constant.

Proof Denoting the Radon-Nikodym derivative of μ with respect to ν by $\frac{d\mu}{d\nu}$, we have $\mu(E) = \int_E \frac{d\mu}{d\nu}(x)\, d\nu$

$$= \int_{T^{-1}E} \frac{d\mu}{d\nu}(x)\, d\nu$$

$$= \int_E \frac{d\mu}{d\nu}(Tx)\, d\nu\ ,\ E \in \mathcal{B}\ .$$

If $\mu(X) < \infty$, i.e. $\frac{d\mu}{d\nu} \in L_1(\nu)$ then 3.7 can be applied, giving that $\frac{d\mu}{d\nu}(Tx) = \frac{d\mu}{d\nu}(x)$ almost everywhere.

In the general case, write $f(x) = \frac{d\mu}{d\nu}(x)$ and $C^+ = \{x : f(Tx) > f(x)\}$. Then $C^+ \in \mathcal{B}$, and for each n, $0 \leq \mu(C^+ \cap X_n) < \infty$ where $\{X_n\}_1^\infty$ is a σ-finite partition of X with respect to the measure μ. Thus

$$\int_{C^+ \cap X_n} f(Tx)\, d\nu = \int_{C^+ \cap X_n} f(x)\, d\nu < \infty ,$$

i.e. $$\int_{C^+ \cap X_n} \{f(Tx) - f(x)\}\, d\nu = 0 .$$

Putting $E_{k,m} = \{x : \frac{k}{2^m} \leq f(Tx) - f(x) < \frac{k+1}{2^m}\}$ and using the argument of 3.7 we see that $\nu(C^+ \cap X_n \cap E_{k,m}) = 0$, $k = 1, 2, \ldots$, $m = 0, 1, \ldots$. Hence $\nu(C^+) = 0$. Similarly, $C^- = \{x : f(Tx) < f(x)\} = \phi \pmod 0$.
Thus $\frac{d\mu}{d\nu}(Tx) = \frac{d\mu}{d\nu}(x)$ almost everywhere.

5.2 now implies that $\frac{d\mu}{d\nu}(x) = c$ almost everywhere. $c > 0$ since μ and ν are both non-negative set functions which take positive values for some sets. //

5.4 <u>Definition</u> A subset J of the positive integers has density $\delta(J) = 0$ if $\lim_{n\to\infty} \frac{\nu_n(J)}{n} = 0$, where $\nu_n(J)$ = number of integers between 1 and n inclusive which belong to J.

If T is an endomorphism of a probability space $(X, \mathcal{B}, \mu)$ then T is <u>weak mixing</u> if any of the following three conditions holds:-

$$\left.\begin{array}{l}\frac{1}{n}\sum_{0}^{n-1}|\mu(A\cap T^{-k}B)-\mu(A)\ \mu(B)|\longrightarrow 0\quad(n\longrightarrow\infty)\\[2ex] \frac{1}{n}\sum_{0}^{n-1}[\mu(A)\cap T^{-k}B)-\mu(A)\ \mu(B)]^{2}\longrightarrow\quad(n\longrightarrow\infty)\\[2ex] \mu(A\cap T^{-n}B)\longrightarrow\mu(A)\ \mu(B)\ (n\longrightarrow\infty,\ n\notin J,\ \delta(J)=0)\end{array}\right\}\ \begin{array}{l}\text{for all}\\ A,b\in\mathfrak{E}\end{array}$$

The three definitions of weak mixing are equivalent, since the forms of convergence to which they correspond are equivalent for bounded sequences of real numbers.

5.5 Definition T has ergodic index $ek(T) = r$ if

$$T^{(s)} \text{ is } \begin{cases} \text{ergodic}, & 1 \leqslant s \leqslant r \\ \text{not ergodic}, & r < s. \end{cases}$$

Clearly, $T^{(s)}$ ergodic implies $T^{(s')}$ ergodic for all $1 \leqslant s' < s$. If $T^{(s)}$ is ergodic for all positive integers s then we put $ek(T) = \infty$, while if T is not ergodic we put $ek(T) = 0$

5.6 Theorem [13] If T is an endomorphism of the finite measure space $(X, \mathcal{B}, \mu)$ then $ek(T) = 0, 1$ or ∞.

Proof It suffices to prove that $T^{(2)}$ ergodic implies $T^{(n)}$ ergodic for $n > 2$. Now $T^{(2)}$ is ergodic if, and only if, T is weak mixing [8, p. 39]. S weak mixing and T weak mixing imply $S \times T$ weak mixing, since for bounded sequences of real numbers $\{a_n\}$, $\{b_n\}$,

$a_n \longrightarrow a$ $(n \longrightarrow \infty, n \notin J_a, \delta(J_a) = 0)$ and

$b_n \longrightarrow b$ $(n \longrightarrow \infty, n \notin J_b, \delta(J_b) = 0)$ imply $a_n b_n \longrightarrow ab(n \longrightarrow \infty,$

$n \notin J_a \cup J_b$, $\delta(J_a \cup J_b) = 0)$: see 5.6. Putting $S = T^{(n)}$ and using induction gives the required result. //

Let $\mathcal{D}^{(r)}$ denote the class of sequences of positive integer r-tuples $\Delta_n^r = (k_n^1 , \ldots, k_n^r)$ such that

$$\lim_{n\to\infty} \inf_{1\leq i<j\leq r} |k_n^i - k_n^j| = \infty .$$

T is <u>mixing of degree r</u> if

$$\lim_{n\to\infty} \mu(E_0 \cap T^{-k_n^1} E_1 \cap \ldots \cap T^{-k_n^r} E_r) = \mu(E_0)\ldots\mu(E_r)$$

for all $\{\Delta_n^r\}_0^\infty \in \mathcal{D}^{(r)}$ and $E_i \in \mathcal{B}$ $(0 \leq i \leq r)$.

This definition is equivalent to the usual definition of mixing of degree r[32]. Also if T is mixing of degree r it is mixing of all degrees less than r.

When r = 1 we say that T is mixing, simply.

A partition ξ of $(X, \mathcal{B}, \mu)$ is a <u>generator</u> of T if $\bigvee_{n=0}^{\infty} T^{-n} \xi = \varepsilon$ (mod 0). T is said to be <u>exact</u> [32] if it has a generator ξ such that $\bigwedge_{n=0}^{\infty} T^{-n} \xi = \nu$ (mod 0) when $\nu = \{X\}$. Since the definition of exactness only depends on the null sets of μ we can, and do, extend the definition to σ-finite endomorphisms (see for example [28]).

5.7 <u>Ergodic Theorem</u> [24]

If T is a finite or σ-finite endomorphism of the measure space $(X, \mathcal{B}, \mu)$ and $f \in L_1(\mu)$ then

$$\frac{1}{n} \sum_{0}^{n-1} f(T^k x) \longrightarrow f^*(x) \ (n \longrightarrow \infty)$$ almost everywhere and in $L_1(\mu)$ norm

$f^* \in L_1(\mu)$ and $f^*(Tx) = f^*(x)$ almost everywhere.

If $\mu(X) < \infty$ then $\int f \, d\mu = \int f^* \, d\mu$.

5.8 Corollary If T is a finite endomorphism then the ergodicity of T is equivalent to the following condition :-

$$\frac{1}{n} \sum_{0}^{n-1} \mu(A \cap T^{-k} B) \longrightarrow \frac{\mu(A)\ \mu(B)}{\mu(X)} \quad (n \longrightarrow \infty) \text{ for all } A, B \in \mathcal{B} .$$

Proof If the Cesaro convergence of measures holds and $E = T^{-1} E \in \mathcal{B}$, put $A = B = E$ to give $\mu(E) = \frac{\mu(E)^2}{\mu(X)}$.

Conversely, the Ergodic Theorem implies

$\frac{1}{n} \sum_{0}^{n-1} f(x)\ g(T^k x) \longrightarrow f(x)\ g^*(x) \quad (n \longrightarrow \infty)$ in $L_1(\mu)$ norm if $f(x)$ is a bounded function. Put $f = \chi_A$, $g = \chi_B$ and integrate term by term (norm convergence implies convergence of the sequences of integrals. //

5.9 Corollary Mixing implies weak mixing implies ergodicity.

Proof Ordinary convergence implies strong Cesáro convergence implies weak Cesáro convergence. //

5.10 Theorem [12, p.405] Let T be an endomorphism of the σ-finite measure space $(X, \mathcal{B}, \mu)$. Then there exists an invariant set C, unique up to null sets, such that $0 \leq f \in L_1(\mu)$, $x \in C$ and $\sum_{0}^{\infty} f(T^n x) > 0$ implies that $\sum_{0}^{\infty} f(T^n x) = \infty$. C is the conservative part of T. $D = X - C$ is the dissipative part of T.

If $X = C \pmod 0$, T is conservative.

5.11 Definition $W \in \mathcal{B}$ is a wandering set of T if $W \cap T^{-n} W = \phi \quad n = 1, 2, \ldots$.

5.12 Theorem T as defined in 5.10 is conservative if, and only if, it has no wandering sets of positive measure.

Proof Let $\{X_n\}_1^\infty$ denote any σ-finite partition of X. For $E \in \mathcal{B}$, $\mu(E) < \infty$, put $f_E(x) = \sum_0^\infty \chi_E(T^n x)$

and
$$\overline{\lim_n}\, T^{-n} E = \bigcap_{m=0}^{\infty} \bigcup_{n=m}^{\infty} T^{-n} E$$
$$= \{x : T^n x \in E \text{ infinitely often}\} .$$

Clearly $f_E(x) = \infty$ if, and only if, $x \in \overline{\lim_n}\, T^{-n} E$. Thus $C = \overline{\lim_n}\, T^{-n} E \cup \{f_E = 0\}$.

Suppose T is conservative and that $E \in \mathcal{B}$ is a wandering set of positive measure. There is no loss of generality in assuming that $\mu(E) < \infty$, since $0 < \mu E \cap X_n) < \infty$ for some n and $E \cap X_n$ is a wandering set.

$$f_E(x) = \begin{cases} 0, & x \notin \bigcup_0^\infty T^{-n} E \\ 1, & x \in \bigcup_0^\infty T^{-n} E . \end{cases}$$

Thus $D \neq \phi \pmod 0$, which contradicts the assumption that T is conservative. Hence T has no wandering sets of positive measure.

Suppost T is not conservative, i.e. $\mu(D) > 0$. Then there exists $E \in \mathcal{B}$, $\mu(E) > 0$, $E \subset D$; for example, $D \cap X_n$ for some n. Put

$$F = \bigcup_{0}^{\infty} T^{-n} E - \bigcup_{1}^{\infty} T^{-n} E .$$

F is a wandering set, since T^{-m} commutes with set-theoretic difference and countable union and also $\bigcup_{m+1}^{\infty} T^{-n} E \subset \bigcup_{m}^{\infty} T^{-n} E \subset \bigcup_{0}^{\infty} T^{-n} E$ $(m = 1, 2, \ldots)$. If $\mu(F) = 0$ then $\bigcup_{m}^{\infty} T^{-n} E = \bigcup_{0}^{\infty} T^{-n} E$ (mod 0) $(m = 1, 2, \ldots)$ and $\overline{\lim_n}\ T^{-n} E = \bigcup_{0}^{\infty} T^{-n} E$ (mod 0). It follows that $E \subset C$ (mod 0) which contradicts $E \subset D$. Hence $\mu(F) > 0$. //

5.13 <u>Hopf Ergodic Theorem</u> [11] Let T be a conservative endomorphism of the σ-finite measure space $(X, \mathcal{B}, \mu)$.

Then for $f \in L_1(\mu)$, $0 \leq g \in L_1(\mu)$

$$\lim_{n \to \infty} \frac{\sum_{0}^{n} f(T^k x)}{\sum_{0}^{n} g(T^k x)} = h_{f,g}(x)$$ exists and is finite almost everywhere on $\{x : \sum_{0}^{\infty} g(T^n x) > 0\}$. $h_{f,g}(x)$ is invariant and $\int f \, d\mu = \int g h_{f,g} \, d\mu$.

Applying the principle of ignoring null sets, we say that endomorphisms T_i of $(X_i, \mathcal{B}_i, \mu_i)$ $(i = 1, 2)$ respectively are <u>almost isomorphic</u> if after discarding null sets from either or both of X_1, X_2 the resulting endomorphisms are isomorphic.

It follows easily that if T_3 is an endomorphism of $(X_3, \mathcal{B}_3, \mu_3)$ such that $T_1 \times T_3$ is not isomorphic to $T_2 \times T_3$ then T_1 cannot be isomorphic to T_2. In other words T_1 isomorphic to T_2 implies $T_1 \times T_3$

isomorphic to $T_2 \times T_3$. A quantitative or qualitative function ι (T) of the endomorphism T is said to be a metric invariant if S isomorphic to T implies $\iota(S) = \iota(T)$. Mixing and ergodicity are examples of qualitative invariants, while one of the most powerful quantitative metric invariants is entropy.

5.14 Definition Let T be an endomorphism of the probability space $(X, \mathcal{B}, \mu)$ and $\xi = \{E_n\}_1^\infty$ be a countable partition of X.
The entropy [2] of ξ with respect to μ is $H_\mu(\xi) = -\sum_{n=1}^{\infty} \mu(E_n)\log\mu(E_n)$.

If ξ is uncountable we define $H_\mu(\xi) = \infty$

$Z_\mu = \{$ partitions $\xi : H_\mu(\xi) < \infty\}$.

If $\xi, \eta \in Z_\mu$ the conditional entropy of ξ relative to η is $H_\mu(\xi|\eta) = \sum_{\substack{E\in\xi \\ F\in\eta}} \mu(E\cap F)\log \dfrac{\mu(E\cap F)}{\mu(F)}$. The main properties of partition entropy are as follows: -

1) $0 \leq H_\mu(\xi|\eta) \leq \infty$; $H_\mu(\xi|\eta) = 0$ if and only if $\xi \leq \eta$ (mod 0)

2) $H_\mu(\xi \vee \eta|\zeta) = H_\mu(\xi|\zeta) + H_\mu(\eta|\xi \vee \zeta)$

3) $H_\mu(\xi|\eta) \leq H_\mu(\zeta|\eta)$ if $\xi \leq \zeta$

4) $H_\mu(\xi|\eta) \geq H_\mu(\xi|\zeta)$ if $\eta \leq \zeta$.

5) $H_\mu(\xi \vee \eta|\zeta) \leq H_\mu(\xi|\zeta) + H_\mu(\eta|\zeta)$

6) $H_\mu(T^{-1}\xi|T^{-1}\eta) = H_\mu(\xi|\eta)$,
where $T^{-1}\xi = \{T^{-1}E : E \in \xi\}$,

7) $H_\mu(\xi) = H_\mu(\xi|\nu)$.

5.15 Theorem [2] $H_\mu(T, \xi) = \lim_{n \to \infty} \frac{1}{n} H_\mu \left(\bigvee_{i=o}^{n-1} T^{-i} \xi \right)$ exists for all $\xi \in Z_\mu$.

5.16 Definition [2] The entropy or Kolmogorov-Sinai invariant of T with respect to μ , $h_\mu(T) = \sup_{\xi \in Z_\mu} h_\mu(T, \xi)$.

5.17 Theorem [2] If T_i is an endomorphism of $(X_i, \mathcal{B}_i, \mu_i)$ $(i = 1, 2)$ then T_1 almost isomorphic to T_2 implies that

$$h_{\mu_1}(T_1) = h_{\mu_2}(T_2) .$$

5.18 Theorem [2] If $\xi \in Z_\mu$ is a generator of the endomorphism T of $(X, \mathcal{B}, \mu)$ then $h_\mu(T) = h_\mu(T, \xi)$.

For futher properties of $h_\mu(T)$,see [2] .

5.19 Definition An endomorphism T of the measure space $(X, \mathcal{B}, \mu)$ is periodic with period n if there exists a positive integer n such that $T^n x = x$, $x \in X$.

5.20 Theorem If the endomorphism T of the non-atomic Lebesgue space $(X, \mathcal{B}, \mu)$ is periodic then $h_\mu(T) = 0$ and T is not ergodic.

Proof $\bigvee_{i=o}^{kn-1} T^{-i}\xi = \bigvee_{i=o}^{k-1} T^{-i}\xi$ $(k = 1, 2, \ldots)$ and so $h_\mu(T, \xi) = 0$, where n is the period of T.

Since μ is non-atomic there exists a set $E \in \mathcal{B}$ with $0 < \mu(E) < \frac{1}{n}$. $E_n = \bigcup_{o}^{n-1} T^{-i} E$ is invariant as T has period n, and $0 < \mu(E_n) \leq n\, \mu(E) < 1$. Thus T is not ergodic. //

§6 Markov and Bernoulli Shifts

A measure-preserving transformation T on a finite or σ-finite measure space $(X, \mathcal{B}, \mu)$ is said to be a Markov shift or M-shift if T has a generator $\xi = \{X_n\}_0^N$, where $0 < N \leqslant \infty$, such that

$$\frac{\mu(X_{n_0} \cap \ldots \cap T^{-k} X_{n_k})}{\mu(X_{n_0} \cap \ldots \cap T^{-(k-1)} X_{n_{k-1}})} = \frac{\mu(X_{n_{k-1}} \cap T^{-1} X_{n_k})}{\mu(X_{n_{k-1}})}$$

for all $0 \leqslant n_r \leqslant N$, $r = 0, 1, \ldots, k$ provided all the measures involved are positive. The set $\{0, 1, \ldots, N\}$ is the state space of the M-shift.

Putting $\lambda_n = \mu(X_n)$ and $p_{ij} = \dfrac{\mu(X_i \cap T^{-1} X_j)}{\mu(X_i)}$,

we see that $\lambda_n \geqslant 0$, $\sum_j p_{ij} = 1$, $\lambda_j = \sum_i \lambda_i p_{ij}$ and that $(X, \mathcal{B}, \mu, T)$ is isomorphic to the shift $(\Omega, \mathfrak{M}, m, S)$ defined as follows:-

$$\Omega_n = \{0, \ldots, N\}, \quad n = 1, 2, \ldots$$

$$\Omega_n = \prod_1^\infty \Omega_n$$

$\mathfrak{M}$ = σ-algebra of measurable sets generated by the cylinders $\{\omega : \omega_n = i_0, \ldots, \omega_{n+k} = i_k\}$
$0 \leqslant i_r \leqslant N \quad (r = 0, 1, \ldots, k)$
and $n = 0, 1, \ldots$,

m = measure uniquely determined by the equations

$$m(\{\omega : \omega_n = i_o\}) = \lambda_{i_o} \quad ,$$

$$m(\{\omega : \omega_n = i_o, \ldots, \omega_{n+k} = i_k\}) = \lambda_{i_o} p_{i_o i_1} \cdots p_{i_{k-1} i_k}$$

(the uniqueness follows by the Kolmogorov Extension Theorem [19, p.159])

$$(S\omega)_n = \omega_{n+1} \;, \quad \text{where } \omega = (\omega_1, \omega_2, \ldots) \in \Omega \;.$$

We note that S preserves m, since the measure of a cylinder is independent of its initial co-ordinate, and that m is finite if, and only if $\sum_{n=0}^{N} \lambda_n < \infty$.

Writing $\underline{\lambda}$ for the N + 1 dimensional vector $(\lambda_o, \ldots, \lambda_N)$ and P for the $(N+1) \times (N+1)$ matrix (p_{ij}), the pair $(\underline{\lambda}, P)$ determines the Markov shift T up to isomorphism. When talking of M-shifts we shall sometimes refer to $(\underline{\lambda}, P)$ as the M-shift. In particular, a shift $(\Omega, \mathfrak{M}, m, S)$ defined as above in terms of $(\underline{\lambda}, P)$ is an M-shift if $\lambda_n \geq 0$, $p_{ij} \geq 0$, $\sum_j p_{ij} = 1$ and $\lambda_j = \sum_i \lambda_i p_{ij}$, i.e. $\underline{\lambda} = \underline{\lambda} P$ in matrix form ($X_n = \{\omega\colon \omega_1 = n\}$ form a generator for S).

$$\text{Let } p_{ij}(0) = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

$$p_{ij}(1) = p_{ij}$$

and $$p_{ij}(n) = \sum_{i_1, \ldots, i_{n-1}} p_{ii_1} \cdots p_{i_{n-1} j} \;, \quad n > 1 \;.$$

Then we see by induction that P^n, the n^{th} power of P, is the matrix $(p_{ij}(n))$. Also, $\mu(X_i \cap T^{-n} X_j) = \lambda_i \, p_{ij}(n)$.

Define $f_{ij}(0) = 0$

$$f_{ij}(1) = p_{ij}$$

$$f_{ij}(n) = \sum_{\substack{i_r \neq j \\ r=1,\ldots,n-1}} p_{ii_1} \cdots p_{i_{n-1}j} , \quad n > 1$$

The relation between $p_{ij}(n)$ and $f_{ij}(n)$ is $p_{ij}(n) = \sum_{r=0}^{n-1} f_{ij}(n-r) p_{jj}(r)$

which can be expressed in terms of generating functions

$F_{ij}(z) = \sum_{n=0}^{\infty} f_{ij}(n) z^n$ and $P_{ij}(z) = \sum_{n=0}^{\infty} p_{ij}(n) z^n$ by

$$P_{ij}(z) = \begin{cases} F_{ij}(z)\, P_{jj}(z) & , \quad i \neq j \\ F_{ii}(z)\, P_{ii}(z) + 1 & , \quad i = j . \end{cases}$$

For the following definitions, we follow [30].

The M-shift $(\underline{\lambda}, P)$ is <u>irreducible</u> if for any states i, j there exists a positive integer n such that $p_{ij}(n) > 0$. Unless otherwise stated, all M-shifts in the sequel will be irreducible.

The state i is <u>transient</u> if $\sum_{n=0}^{\infty} p_{ii}(n) < \infty$,

<u>positive recurrent</u> if $\sum_{n=0}^{\infty} p_{ii}(n) = \infty$ and

$$\sum_{n=1}^{\infty} n\, f_{ii}(n) < \infty .$$

<u>null-recurrent</u> if $\sum_{n=0}^{\infty} p_{ii}(n) = \infty$ and

$$\sum_{n=1}^{\infty} n\, f_{ii}(n) = \infty .$$

aperiodic if $p_{ii}(n) > 0$ for all large enough n.

We consider only aperiodic M-shifts.

6.1 Theorem If $(\underline{\lambda}, P)$ is irreducible and aperiodic, all states are of the same type, ie. transient, positive recurrent or null recurrent [5, p. 355].

For all states i, $\lim_{n\to\infty} \frac{1}{n} \sum_{0}^{n-1} p_{ii}(k)$ exists. i is positive recurrent if, and only if, $\lim_{n\to\infty} \frac{1}{n} \sum_{0}^{n-1} p_{ii}(k) > 0$ [30] .

One can therefore talk of a transient, positive recurrent or null recurrent M-shift.

6.2 Theorem Given two M-shifts $(\underline{\lambda}, P) = (\Omega, \mathfrak{M}, m, S)$ and $(\underline{\lambda}', P') = (\Omega', \mathfrak{M}', m', S')$ with $\Omega_n = \Omega'_n = \{0, 1, \ldots\}$, n = 1, 2, ...

let $\Omega'' = \prod_{1}^{\infty} (\Omega_n \times \Omega'_n)$,

$\mathfrak{M}''$ = σ-algebra generated by the cylinders

$\{\omega'' : \omega''_n = (i_o, j_o), \ldots, \omega''_{n+r} = (i_r, j_r)\}$

m" = measure uniquely determined by the equation

$$m''(\{\omega'' : \omega''_n = (i_o, j_o), \ldots, \omega''_{n+r} = (i_r, j_r)\})$$

$$= \lambda_{i_o} \lambda'_{j_o} p_{i_o i_1} p'_{j_o j_1} \cdots p_{i_{r-1} i_r} p'_{j_{r-1} j_r}$$

Then $(\Omega \times \Omega', \mathfrak{M} \times \mathfrak{M}', m \times m', S \times S')$ is isomorphic to $(\Omega'', \mathfrak{M}'', m'', S'')$.

Proof Define $\psi : \Omega \times \Omega' \longrightarrow \Omega''$ by $\psi(\omega, \omega') = \{(\omega_n, \omega'_n)\}_1^{\infty}$, where

$\omega = \{\omega_n\}_1^\infty$ and $\omega' = \{\omega'_n\}_1^\infty$. Clearly ψ is 1-1 onto. It is measurable since

$$\psi^{-1}\{\omega'' : \omega''_n = (i_o, j_o), \ldots, \omega''_{n+r} = (i_r, j_r)\}$$

$$= \{\omega : \omega_n = i_o, \ldots, \omega_{n+r} = i_r\} \times \{\omega' : \omega'_n = j_o, \ldots, \omega'_{n+r} = j_r\}$$

and the ring of finite disjoint unions of cylinders generates the measurable subsets in each respective sequence space. ψ is measure-preserving by the definitions of m". Finally,

$$\psi S = S'\psi . \quad //$$

6.3 <u>Theorem</u> [13] If T is an M-shift on the σ-finite measure space $(X, \mathcal{B}, \mu)$ then T is ergodic if, and only if, it is irreducible and recurrent (positive or null).

6.4 <u>Corollary</u> If $(\Omega, \mathfrak{M}, m, S)$ and $(\Omega', \mathfrak{M}', m', S')$ are M-shifts preserving σ-finite measure such that $S \times S'$ is irreducible then $S \times S'$ is ergodic if, and only if, it is recurrent.

<u>Proof</u> By 6.2, $S \times S'$ is isomorphic to S" as defined in 6.2. Clearly S" is an irreducible M-shift preserving σ-finite measure. The result follows from 6.3 noting that ergodicity is a metric invariant. //

The next result shows that the recurrence in 6.3 and 6.4 is null recurrence.

6.5 <u>Theorem</u> Let $T = (\underline{\lambda}, P)$ be an irreducible M-shift.

(i) If T is transient the only measure preserved by T is the trivial one $\mu(E) = 0$ for all $E \in$.

(ii) If T is positive recurrent is preserves a finite measure.

(iii) If T is null-recurrent it preserves a σ-finite measure.

In (ii) and (iii) the invariant measure is unique up to constant multiples.

<u>Proof</u> (i) $\sum_{0}^{\infty} p_{ii}(n) < \infty$ implies that $p_{ii}(n) \longrightarrow 0$ $(n \longrightarrow \infty)$. Since $p_{ii}(2n) \geq p_{ij}(n)\, p_{ji}(n)$, it follows that $P^n \longrightarrow 0$, the zero $(N+1) \times (N+1)$ matrix, $(n \longrightarrow \infty)$. For $\underline{\lambda}$ to satisfy $\underline{\lambda}P = \underline{\lambda}$, we must have $\underline{\lambda}P^n = \underline{\lambda}$ for all n, i.e. $\underline{\lambda} = \underline{0}$.

(ii) [30] There is a unique vector $\underline{\lambda} = \underline{\lambda}P$ with $\sum_{0}^{N} \lambda_n = 1$.

(iii) [30] There is a unique vector $\underline{\lambda} = \underline{\lambda}P$ with $\sum_{0}^{\infty} \lambda_n = \infty$.

The uniqueness of the invariant measure in (ii) and (iii) follows from the ergodicity of T (I. 5.3). //

6.6 <u>Definition</u> $(\underline{\lambda}, P)$ is a <u>Bernoulli shift</u> or <u>Bernoulli endomorphism</u> if $\lambda_j = p_{ij}$ for all i, j. Thus, with the notation at the beginning of §6, $m(\{\omega : \omega_n = i_o, \ldots, \omega_{n+k} = i_k\}) = \lambda_{i_o} \ldots \lambda_{i_k}$. We shall sometimes use $\underline{\lambda}$ to denote the associated Bernoulli shift.

6.7 <u>Theorem</u> (i) The entropy of the Bernoulli shift $\underline{\lambda}$ is $-\sum_{1}^{N} \lambda_n \log \lambda_n$.

(ii) Every Bernoulli shift is mixing of all degrees.

<u>Proof</u> (i) See for example [2]

(ii) Let r be a positive integer, $A_o, \ldots, A_r$ be cylinders and $\Delta_n^r = (k_n^1, \ldots k_n^r) \in \mathcal{D}^{(r)}$. Then A_o, $T^{-k_n^1} A_1, \ldots, T^{-k_n^1} A_r$

are cylinders depending on sets of co-ordinates which are pairwise disjoint if n is large enough. Hence for large enough n,

$$\mu(A_o \cap T^{-k_n^1} \cap \ldots \cap T^{-k_n^r} A_r) = \mu(A_o) \ldots \mu(A_r) .$$

The result follows, since the cylinders generate the measurable sets. //

§7 Renewal Sequences

7.1 Definition [18] A sequence $\underline{p} = \{p_n\}_o^\infty$ of real numbers is a renewal sequence if $p_o = 1$, $p_n = \sum_{k=1}^{n} \sum_{r_1+\ldots+r_k=n} f_{r_1} \ldots f_{r_k}$ $(n \geq 1)$,

where $0 \leq f_n$ and $\sum_{n=1}^{\infty} f_n \leq 1$.

This can be expressed in terms of generating functions: $P(z) = \sum_o^\infty p_n z^n$ and $F(z) = \sum_1^\infty f_n z^n$ satisfy $P(z) = \frac{1}{1-F(z)}$. The series $F(z)$ and $P(z)$ were called by T. Kaluza "reciprocal power series" [14], and we shall sometimes write $F_{\underline{p}}(z)$ for the power series reciprocal to $P(z)$. A very important subclass of renewal sequences was studied by him, although in a different context to the present one. It was rediscovered by J. Lamperti [21] and developed by J.F.C.Kingman [18].

7.2 Definition [18] A sequence $\{p_n\}_o^\infty$ of real numbers is a Kaluza sequence if $p_o = 1$, $0 \leq p_n \leq 1$ and $p_n^2 \leq p_{n-1} p_{n+1}$ for all n.

7.3 Theorem [18] If $\underline{p}$ is a Kaluza sequence then $p_n \geq p_{n+1}$ for all n.

<u>Proof</u> Let $u_n = \frac{p_n}{p_{n-1}}$. Then $u_n \leqslant u_{n+1}$. Also by induction $p_n = u_1 \dots u_n$. $p_n \geqslant p_{n+1}$ if, and only if, $u_{n+1} \leqslant 1$. Suppose $u_{n_o} > 1$ for some n_o. Then $u_n > 1$ for all $n > n_o$, in particular $u_n > 1 + \varepsilon$ for some $\varepsilon > 0$. Thus $p_{n_o+n} > u_1 \dots u_{n_o}(1 + \varepsilon)^n \longrightarrow \infty$ $(n \longrightarrow \infty)$. This contradicts the boundedness of $\underline{p}$. //

7.4 <u>Theorem</u> [24] If $\underline{p}$ is a Kaluza sequence then it is a renewal sequence.

<u>Proof</u> The proof of 7.3 implies that $\lim_{n\to\infty} \frac{p_{n+1}}{p_n}$ exists and does not exceed 1. Hence P(z) converges for $|z| < 1$. So does F(z), since it is the reciprocal power series of P(z). However, the relation between F and P holds irrespective of the value of z since it follows on equating coefficients of z^n in the identity $P(z) = F(z)\, P(z) + 1$. Now $1 \leqslant P(z) \leqslant \infty$ if $0 \leqslant z$ and so $\sum_1^\infty f_n \leqslant 1$.

An induction shows that $f_n \geqslant 0$ for all n:

$$0 = p_n - \sum_{\nu=1}^{n} p_{n-\nu}\, f_\nu \ ,$$

$$f_{n+1} = p_{n+1} - \sum_{\nu=1}^{n} p_{n-\nu+1}\, f_\nu$$

and hence $f_{n+1}\, p_n = \sum_{\nu=1}^{n} (p_{n-\nu}\, p_{n+1} - p_{n-\nu+1}\, p_n)\, f_\nu$; since

$\frac{p_{n+1}}{p_n} \geqslant \frac{p_n}{p_{n-1}} \geqslant \dots \geqslant \frac{p_{n-\nu+1}}{p_{n-\nu}}$, $f_1 \geqslant 0, \dots, f_n \geqslant 0$ imply $f_{n+1} \geqslant 0$.

The first step of the induction is given by $f_1 = p_1$. //

7.5 <u>Definition</u> [16] A renewal sequence $\{p_n\}_0^\infty$ is <u>infinitely divisible</u> if $\{p_n^t\}_0^\infty$, where p_n^t denotes p_n raised to the power t, is a renewal sequence for all $t > 0$.

The defining inequality for Kaluza sequences gives the following

7.6 <u>Theorem</u> [16] Every Kaluza sequence is infinitely divisible.

We mention the interesting converse to 7.6.

7.7 <u>Theorem</u> [16] Every zero-free infinitely divisible renewal sequence is a Kaluza sequence.

Anticipating Chapter IV, we shall use the same notation for renewal sequences as we have done for M-shifts.

7.8 <u>Definition</u> A renewal sequence $\underline{p} = \{p_n\}_0^\infty$ is said to be

<u>transient</u> if $\sum_0^\infty p_n < \infty$

<u>positive recurrent</u> if $\sum_0^\infty p_n = \infty$ and $\sum_1^\infty n\, f_n < \infty$

<u>null recurrent</u> if $\sum_0^\infty p_n = \infty$ and $\sum_1^\infty n\, f_n = \infty$

<u>aperiodic</u> if $p_n > 0$ for n large enough.

7.9 <u>Theorem</u> If $\underline{p}$ is a Kaluza sequence then either it has only a finite number of positive terms or it is aperiodic.

<u>Proof</u> By 7.3 any zero in $\underline{p}$ is followed by zeros. //

When $\underline{p}$ is a recurrent Kaluza sequence, f_n grows more slowly than p_n :

7.9 <u>Theorem</u> If $\underline{p}$ is Kaluza and recurrent then $\frac{f_n}{p_n} \longrightarrow 0 \ (n \longrightarrow \infty)$.

<u>Proof</u> $\frac{f_n}{p_n} = 1 - \sum_1^n f_s \frac{p_{n-s}}{p_n} \leqslant 1 - \sum_1^n f_s$.

Hence $0 \leqslant \overline{\lim_n} \frac{f_n}{p_n} \leqslant 0$. //

§8. <u>Continued Fractions</u>

8.1 <u>Definition</u> If x is a real number the <u>integer part</u> of x is defined as the greatest integer $[x] \leqslant x$. Thus $x - 1 < [x] \leqslant x$.

The <u>fractional part</u> of x is defined by $(x) = x - [x]$. Thus $0 \leqslant (x) < 1$.

If f is a 1-1 real-valued function of a real variable then f^{-1} denotes the <u>functional inverse</u> of $f : f(f^{-1} x) = x$.

8.2 <u>Definition</u> An <u>nth order continued fraction</u> is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \dots + \cfrac{1}{a_n}}} ,$$

where $a_0, \dots a_n$ are real numbers. We shall always write such a continued fraction as

$$a_0 + \frac{1}{a_1 +} \ \frac{1}{a_2 +} \ \cdots \ + \frac{1}{a_n} .$$

If $\lim_{n\to\infty} \left(a_o + \frac{1}{a_1+} \cdots + \frac{1}{a_n}\right)$ exists, we denote it by $a_o + \frac{1}{a_1+} \; \frac{1}{a_2+} \cdots$ and refer to it as an <u>infinite continued fraction</u>.

Since $a_o + \frac{1}{a_1+} \cdots + \frac{1}{a_n}$ is obtained by a finite number of rational operations on $a_o, \ldots, a_n$, it can be represented as the ratio of two polynomials.

$$\frac{P(a_o, \ldots, a_n)}{Q(a_o, \ldots, a_n)}$$

in $a_o, \ldots, a_n$ with integral coefficients. This representation is not unique since in evaluating the finite continued fraction a factor common to numerator and denominator may occur. To overcome the ambiguity we define $\frac{P}{Q}$ 'canonically' [17] :

$$\frac{P(a_o)}{Q(a_o)} = a_o$$

$$\frac{P(a_o, \ldots, a_n)}{Q(a_o, \ldots, a_n)} = a_o + \cfrac{1}{a_1 + \frac{1}{a_2+} \cdots + \frac{1}{a_n}} = a_o + \frac{1}{r_1}, \quad \text{say}$$

r_1 is an $(n - 1)$ st order continued fraction with canonical representation $r_1 = \frac{p'}{q'}$, say. We then define

$$\frac{P(a_o, \ldots, a_n)}{Q(a_o, \ldots, a_n)} = a_o + \frac{q'}{p'} = \frac{a_o p' + q'}{p'} ,$$

i.e. $P(a_o, \ldots, a_n) = a_o p' + q'$, $Q(a_o, \ldots, a_n) = p'$. The definition is now completed by induction.

$$\frac{P(a_o,\ldots,a_n)}{Q(a_o,\ldots,a_n)}$$ is the nth <u>convergent</u> or <u>partial quotient</u> of

$$a_o + \frac{1}{a_1 +} \cdots + \frac{1}{a_n} \quad .$$

For more details about continued fractions, see [17] .

CHAPTER II.

Characterisations of Mixing Properties for Measure-Preserving Transformations.

§1 Introduction

N. Oishi [25] characterised mixing, weak mixing and ergodicity of finite measure-preserving transformations in terms of convergence of suitably defined entropies. We have extended the convergence criterion from all finite partitions to all partitions with finite entropy and also characterised mixing of degree r in the same way. Further characterisations in terms of convergence of suitably defined measures and, in the case of a topological measure space, weak convergence of the same measures are given. It is shown that these results are the 'best possible'.

§2 A lemma

Throughout this chapter, T will denote a measure-preserving transformation of the non-atomic Lebesgue space $(X, \mathcal{B}, p)$.

If γ is another probability measure on $(X, \mathcal{B})$, we define

$$H_p(\gamma) = \begin{cases} \displaystyle\int_X \log \frac{d\gamma}{dp}\, d\gamma , & \text{if } \gamma \text{ is absolutely continuous with respect to } p \\ +\infty , & \text{otherwise} \end{cases}$$

where $\frac{d\gamma}{dp}$ denotes the Radon-Nikodym derivative of γ with respect to p.

2.1 Lemma [25] Let $\{\gamma_1, \gamma_2, \ldots\}$ be a sequence of probability measures on $(X, \mathcal{B})$ such that $\gamma_n \leq c\, p$ for all n , where $c \geq 1$ is a constant then

$$\gamma_n(E) \longrightarrow p(E) \quad (n \longrightarrow \infty) \quad \text{uniformly for } E \in \mathcal{B}$$

if, and only if

$$\lim_{n \to \infty} H_p(\gamma_n) = 0.$$

Proof. Since, by assumption, $\int_E (\frac{d\gamma_n}{dp} - c)dp \leq 0$ for all $E \in \mathcal{B}$, $\frac{d\gamma_n}{dp} \leq c$ p - almost everywhere for each n. $\gamma_n(E) \longrightarrow p(E)$ $(n \longrightarrow \infty)$ uniformly for $E \in \mathcal{B}$ if, and only if, $\frac{d\gamma_n}{dp} \longrightarrow 1$ $(n \longrightarrow \infty)$ in $L_1(p)$ norm, since, on the one hand,

$$|\gamma_n(E) - p(E)| \leq \int_X \left| \frac{d\gamma_n}{dp} - 1 \right| dp$$

while, on the other, if $S = \{ \frac{d\gamma_n}{dp} > 1 \}$, $S \in \mathcal{B}$ and

$$\int_X \left| \frac{d\gamma_n}{dp} - 1 \right| dp = \int_S (\frac{d\gamma_n}{dp} - 1)dp + \int_{\complement S} (1 - \frac{d\gamma_n}{dp})dp$$

$$= | \gamma_n(S) - p(S) | + | \gamma_n(\complement S) - p(\complement S) | .$$

Suppose $H_p(\gamma_n) \longrightarrow 0$ $(n \longrightarrow \infty)$. Since $x \log x \geq x - 1 + \frac{1}{2c}(x-1)^2$ for any x with $0 \leq x \leq c$,

$$H_p(\gamma_n) \geq \frac{1}{2c} \int_X (\frac{d\gamma_n}{dp} - 1)^2 dp \geq 0 \quad \text{for each } n .$$

Hence $\frac{d\gamma_n}{dp} \longrightarrow 1$ $(n \longrightarrow \infty)$ in $L_2(p)$ norm and so also in $L_1(p)$ norm (by Holder's inequality, $(\int |f| \, dp)^2 \leq \int |f|^2 dp$ for $f \in L_2(p)$)

Conversely, let $\frac{d\gamma_n}{dp} \longrightarrow 1$ $(n \longrightarrow \infty)$ in $L_1(p)$ norm.

$$\frac{d\gamma_n}{dp} \longrightarrow 1 \;(n \longrightarrow \infty) \text{ in probability, since for all } \varepsilon > 0$$

$$\varepsilon\, p(E_{n,\varepsilon}) \leqslant \int_{E_{n,\varepsilon}} \left| \frac{d\gamma_n}{dp} - 1 \right| dp \leqslant \int_X \left| \frac{d\gamma_n}{dp} - 1 \right| dp$$

where $E_{n,\varepsilon} = \{ \, | \frac{d\gamma_n}{dp} - 1 | \geqslant \varepsilon \, \}$.

Now $|x \log x| \leqslant |x - 1| + \frac{1}{2}(x - 1)^2$ for any $x \geqslant 0$, so that

$$\frac{d\gamma_n}{dp} \log \frac{d\gamma_n}{dp} \longrightarrow 0 \quad (n \longrightarrow \infty) \text{ in probability .}$$

$$\left|\frac{d\gamma_n}{dp} \log \frac{d\gamma_n}{dp}\right| \leqslant \text{Max}[c \log c, \frac{1}{e} \log \frac{1}{e}] = K, \text{ say, for all } n .$$

$$\left| \int_X \frac{d\gamma_n}{dp} \log \frac{d\gamma_n}{dp} \, dp \right| \leqslant \int_X \left| \frac{d\gamma_n}{dp} \log \frac{d\gamma_n}{dp} \right| dp$$

$$= \int_{E_{n,\varepsilon}} \left| \frac{d\gamma_n}{dp} \log \frac{d\gamma_n}{dp} \right| dp + \int_{\mathcal{C}E_{n,\varepsilon}} \left| \frac{d\gamma_n}{dp} \log \frac{d\gamma_n}{dp} \right| dp$$

$$\leqslant K\, p\, (E_{n,\varepsilon}) + \varepsilon .$$

Hence

$$\lim_{n \to \infty} H_p(\gamma_n) = 0 . \qquad //$$

§3 The Characterisations

Let
$$G(T^n) = \{(x, T^n x) : x \in X\},$$

μ_n denote the measure defined on $(X^{(2)}, \mathcal{B}^{(2)})$ by
$$\mu_n(E) = p[\pi \{E \cap G(T^n)\}], \quad E \in \mathcal{B}^{(2)}.$$

$\mu_n^{\theta,\theta'}$ denotes the restriction of μ_n to the σ-algebra $S(\theta\ \theta')$ generated by $\theta \times \theta'$, θ, θ' being finite partitions of X, for which we also define

$$I_{\mu_n}(\theta \times \theta') = \sum_{E\in\theta, F\in\theta'} \mu_n(E \times F)\log \frac{\mu_n(E \times F)}{\mu(E \times F)}$$

3.1 Theorem. The following are equivalent:-

(i) T is mixing

(ii) $H_{\mu_n}(\theta \times \theta') \longrightarrow H_\mu(\theta \times \theta')$ $(n \longrightarrow \infty)$ for all $\theta,\theta' \in Z_p$

(iii) $\mu_n(M) \longrightarrow \mu(M)$ $(n \longrightarrow \infty)$ for all Jordan measurable $M \in \mathcal{B}^{(2)}$.

Proof. (i) <=> (ii): We first consider the case of finite θ,θ' [2.5].

$$\frac{d\mu_n^{\theta,\theta'}}{d\mu^{\theta,\theta'}}(x, y) = \sum_{E\in\theta, F\in\theta'} \chi_{E\times F}(x, y) \frac{\mu_n(E \times F)}{\mu(E \times F)}$$

and $$\int_{X^{(2)}} \left| \frac{d\mu_n^{\theta,\theta'}}{d\mu^{\theta,\theta'}} - 1 \right| d\mu^{\theta,\theta'} = \sum_{E\in\theta, F\in\theta'} \left| \mu_n(E \times F) - \mu(E \times F) \right|$$

Hence T is mixing if, and only if,

$$\frac{d\mu_n^{\theta,\theta'}}{d\mu^{\theta,\theta'}} \longrightarrow 1 \ (n \longrightarrow \infty) \text{ in } L_1(\mu^{\theta,\theta'}) \text{ norm for all finite}$$

partitions θ,θ'. The proof of 2.1 now implies that T is mixing if, and only if, for all finite θ,θ', $\mu_n^{\theta,\theta'}(M) \longrightarrow \mu^{\theta,\theta'}(M) \quad (n \longrightarrow \infty)$ uniformly for $M \in S(\theta, \theta')$.

The result for finite θ,θ' follows from 2.1, noting that

$$\frac{d_{\mu_n}^{\theta,\theta'}}{d_{\mu}^{\theta,\theta'}} \leq \max_{\substack{F\in\theta' \\ p(F)\neq 0}} \frac{1}{p(F)} ,$$

$$I_{\mu_n}(\theta \times \theta') = H_{\mu^{\theta,\theta'}}(\mu_n^{\theta,\theta'})$$

and $H_{\mu_n}(\theta \times \theta') = H_{\mu}(\theta \times \theta') - I_{\mu_n}(\theta \times \theta')$.

Now let $\theta, \theta' \in Z_p$ be infinite. They can be at most countable. Thus $\theta \times \theta' = \{D_1, D_2, \dots\}$, where D_k are disjoint, measurable rectangles.

$$\begin{aligned} H_{\mu_n}(\theta \times \theta') &= H_p(\theta \vee T^{-n}\theta') \\ &\leq H_p(\theta) + H_p(\theta') \\ &= H_{\mu}(\theta \times \theta') , \text{ for each } n . \end{aligned}$$

Thus $\overline{\lim_n}\ H_{\mu_n}(\theta \times \theta') \leq H_{\mu}(\theta \times \theta')$. Let ξ_n be the partition of $X^{(2)}$ given by $\xi_n = \{D_1, D_2, \dots, D_n, \bigcup_{k=n+1}^{\infty} D_k\}$. Then $\xi_n \nearrow \theta \times \theta'$ $(n \longrightarrow \infty)$, since each set of ξ_n is a union of (at most two) sets of ξ_{n+1}, while if $\eta \geq \xi_n$ for all n, each D_n is a union of sets of η and hence $\theta \times \theta' \leq \eta$.

By the first part of the proof,

$H_{\mu_k}(\xi_n) \longrightarrow H_\mu(\xi_n)$ ($k \longrightarrow \infty$) for each n. Also

$H_{\mu_k}(\xi_n) \leqslant H_{\mu_k}(\theta \times \theta')$ which implies that

$H_\mu(\xi_n) \leqslant \underline{\lim}_k H_{\mu_k}(\theta \times \theta')$ for each n. Let $n \longrightarrow \infty$:

$H_\mu(\theta \times \theta') \leqslant \underline{\lim}_k H_{\mu_K}(\theta \times \theta')$; hence

$$H_\mu(\theta \times \theta') \leqslant \underline{\lim}_k H_{\mu_k}(\theta \times \theta') \leqslant \overline{\lim}_k H_{\mu_k}(\theta \times \theta') \leqslant H_\mu(\theta \times \theta') ,$$

i.e. $H_\mu(\theta \times \theta') = \lim_{n\to\infty} H_{\mu_n}(\theta \times \theta')$.

(ii) => (i) follows trivially from the first part of the proof, since every finite θ is in Z_p .

(i) <=> (iii) : T is mixing if, and only if, $\mu_n(M) \longrightarrow \mu(M)$ $(n \longrightarrow \infty)$ for all $M \in \mathcal{R}$, by the finite additivity of measures. (iii) => (i) follows at once, since every set in $\mathcal{R}$ is Jordan measurable.

Let T be mixing and M, Jordan measurable. For all positive integers n, there exist R_n and S_n such that

$$M \supset R_n \in \mathcal{R} \quad , \qquad \mu(M - R_n) < \frac{1}{n}$$

$$M \subset S_n \in \mathcal{R} \quad , \qquad \mu(S_n - M) < \frac{1}{n}$$

Hence $\mu(M - \bigcup_1^\infty R_n) = 0 = \mu(\bigcap_1^\infty S_n - M)$

i.e. $\mu(\bigcup_1^\infty R_n) = \mu(M) = \mu(\bigcap_1^\infty S_n)$

Let $I_N = \bigcup_1^N R_n$. Then $I_N \in \mathcal{R}$ and $I_N \nearrow \bigcup_1^\infty R_n \quad (N \longrightarrow \infty)$.

Hence $\mu_k(I_N) \nearrow \mu_k(\bigcup_1^\infty R_n) \quad (N \longrightarrow \infty)$ for each k

$\mu(I_N) \nearrow \mu(\bigcup_1^\infty R_n) \quad (N \longrightarrow \infty)$

and $\mu_k(I_N) \longrightarrow \mu(I_N) \quad (k \longrightarrow \infty)$ for each N.

Let $J_N = \bigcap_1^N S_n$. Then $J_N \in \mathcal{R}$ and $J_N \searrow \bigcap_1^\infty S_n \quad (N \longrightarrow \infty)$.

Hence $\mu_k(J_N) \searrow \mu_k(\bigcap_1^\infty S_n) \quad (N \longrightarrow \infty)$ for each k

$\mu(J_N) \searrow \mu(\bigcap_1^\infty S_n) \quad (N \longrightarrow \infty)$

and $\mu_k(J_N) \longrightarrow \mu(J_N) \quad (k \longrightarrow \infty)$ for each N.

$\mu_k(I_N) \leq \mu_k(M) \leq \mu_k(J_N)$ for each k and N. Keeping N fixed, let $k \longrightarrow \infty$:

$\mu(I_N) \leq \underline{\lim}_k \mu_k(M) \leq \overline{\lim}_k \mu_k(M) \leq \mu(J_N)$. Let $N \longrightarrow \infty$.

$\mu(M) \leq \underline{\lim}_k \mu_k(M) \leq \overline{\lim}_k \mu_k(M) \leq \mu(M)$,

i.e. $\mu(M) = \lim_k \mu_k(M)$. //

<u>3.2 Lemma.</u> $T^{(2)}$ preserves μ_n for each n .

<u>Proof</u>

$$\pi\{G(T^n) \cap (T^{(2)})^{-1} E\} = T^{-1} \pi\{G(T^n) \cap E\} \quad \text{for all} \quad E \in \mathcal{B}^{(2)} .$$ //

Since $T^{(2)}$ also preserves μ , one might conjecture that

$$T \text{ is mixing if, and only if, } h_{\mu_n}(T^{(2)}) \longrightarrow h_\mu(T^{(2)}) \ (n \longrightarrow \infty)$$

where $h_{\mu_n}(T^{(2)})$ is the Kolmogorov-Sinai invariant of $T^{(2)}$ with respect to μ_n . That the conjecture is false is shown by

<u>3.3 Theorem</u> $(X^{(2)}, \mathcal{B}^{(2)}, \mu_n , T^{(2)})$ is almost isomorphic to $(X, \mathcal{B} , p, T)$

<u>Proof</u> Since μ_n is concentrated on $G(T^n)$, $X^{(2)} - G(T^n)$ is a μ_n-null set.

Let $\psi_n(x) = (x, T^n x)$, $x \in X$. ψ_n is one-to-one onto $G(T^n)$. It is measurable and measure-preserving, since

$$\psi_n^{-1}(E) = \psi_n^{-1}(E \cap G(T^n)) = \pi [E \cap G(T^n)] .$$

Finally $T^{(2)}\psi_n(x) = \psi_n(Tx)$ for all $x \in X$. //

<u>3.4 Corollary</u> $h_{\mu_n}(T^{(2)}) = h_p(T)$

<u>Proof</u> Almost isomorphic transformations have the same entropy. //

<u>3.5 Corollary</u> $h_{\mu_n}(T^{(2)}) \longrightarrow h_\mu(T^{(2)})$ $(n \longrightarrow \infty)$ if, and only if ,

$$h_p(T) = 0 \text{ or } \infty .$$

<u>Proof</u> $h_\mu(T^{(2)}) = 2h_p(T)$.

The conjecture is disproved since, on the one hand, Bernoulli endomorphisms with countable state space are mixing yet have finite, positive entropy, while on the other hand, periodic endomorphisms of non-atomic Lebesgue spaces are not ergodic yet have zero entropy.

One might also ask whether (iii) of 3.1 could not be replaced by

(iii)' $\mu_n(M) \longrightarrow \mu(M) \quad (n \longrightarrow \infty)$ for all $M \in \mathcal{B}^{(2)}$.

This is answered in the negative by

<u>3.6 Theorem</u> $\mu(G(T^n)) = 0$ for all n .

<u>Proof</u> $\mu(G(T^n)) = \int_X p[G(T^n)_x]dp = \int_X p(\{T^n x\})dp$

where $\{T^n x\}$ denotes the point set $T^n x$,

$= 0$ since p is non-atomic . //

(iii)' is false for $\bigcup_1^\infty G(T^n) \in \mathcal{B}^{(2)}$, since

$\mu_m(\bigcup_1^\infty G(T^n)) = 1$ for each m, but $\mu(\bigcup_1^\infty G(T^n)) = 0$.

<u>3.7 Theorem</u> The following are equivalent:-

(i) T is weak mixing

(ii) $\frac{1}{n}\sum_0^{n-1} H_{\mu_k}(\theta \times \theta') \longrightarrow H_\mu(\theta \times \theta') \quad (n \longrightarrow \infty)$ for all $\theta, \theta' \in Z_p$

$$\text{(iii)}\quad \left.\begin{array}{l} \frac{1}{n}\sum_{0}^{n-1}[\mu_k(M)-\mu(M)]^2 \longrightarrow 0 \quad (n\longrightarrow\infty) \\[2ex] \frac{1}{n}\sum_{0}^{n-1}|\mu_k(M)-\mu(M)| \longrightarrow 0 \quad (n\longrightarrow\infty) \\[2ex] \mu_n(M)\longrightarrow\mu(M) \quad (n\longrightarrow\infty,\ n\notin J,\ \delta(J)=0) \end{array}\right\} \text{for all Jordan measurable } M\in\mathcal{B} \quad (2)$$

<u>Proof</u> (i) <=> (ii) : For finite θ, θ' [2.5], the result follows from

$$\frac{1}{n}\sum_{0}^{n-1} H_{\mu_k}(\theta\times\theta') = H_{\mu}(\theta\times\theta') - \frac{1}{n}\sum_{0}^{n-1} I_{\mu_k}(\theta\times\theta')$$

and $\frac{1}{2c}\sum_{\substack{E\in\theta,\ F\in\theta'\\ \mu(E\times F)\neq 0}} \frac{1}{\mu(E\times F)}\frac{1}{n}\sum_{0}^{n-1}[\mu_k(E\times F)-\mu(E\times F)]^2$

$$\leq \sum_{\substack{E\in\theta,\ F\in\theta'\\ \mu(E\times F)\neq 0}} \frac{1}{n}\sum_{0}^{n-1} I_{\mu_k}(\theta\times\theta')$$

$$\leq \frac{1}{n}\sum_{0}^{n-1}|\mu_k(E\times F)-\mu(E\times F)| + \frac{1}{2}\sum_{\substack{E\in\theta,F\in\theta'\\ \mu(E\times F)\neq 0}}\frac{1}{\mu(E\times F)}\frac{1}{n}\sum_{0}^{n-1}[\mu_k(E\times F)-\mu(E\times F)]^2 ,$$

where $c = \max_{\substack{F\in\theta'\\ p(F)\neq 0}} \frac{1}{p(F)}$. The latter inequalities are obtained by putting

$x = \dfrac{d\mu_k^{\theta,\theta'}}{d\mu^{\theta,\theta'}}$ in the inequalities $x - 1 + \frac{1}{2c}(x-1)^2 \leq x\log x \leq |x-1| + \frac{1}{2}(x-1)^2$,

for $0\leq x\leq c$, integrating with respect to $\mu^{\theta,\theta'}$ and taking Cesaro sums.

(ii) => (i) now follows for $\theta,\theta'\in Z_p$, as in 3.1 . The rest of the proof is entirely analogous to that of 3.1.

(i) <=> (iii) : Since the three forms of convergence corresponding to weak mixing are equivalent for bounded sequences of real numbers, the proof need only consider one of them.

T is weak mixing if, and only if,

$$\frac{1}{n} \sum_{0}^{n-1} | \mu_k(M) - \mu(M) | \longrightarrow 0 \ (n \longrightarrow \infty) \quad \text{for all } M \in \mathcal{R}$$

by the finite additivity of measures and the triangle inequality for moduli. (iii) => (i) follows at once, since every set in $\mathcal{R}$ is Jordan measurable

Let T be weak mixing and M, Jordan measurable. With the notation of 3.1,

$$\frac{1}{n} \sum_{k=o}^{n-1} |\mu_k(J_N) - \mu(M)| \leqslant \frac{1}{n} \sum_{k=o}^{n-1} |\mu_k(J_N) - \mu(J_N)| + |\mu(J_N) - \mu(M)|$$

Therefore

$$\overline{\lim_n} \frac{1}{n} \sum_{k=o}^{n-1} |\mu_k(J_N) - \mu(M)| \leqslant |\mu(J_N) - \mu(M)| \quad \text{for each } N .$$

Similarly,

$$\overline{\lim_n} \frac{1}{n} \sum_{k=o}^{n-1} |\mu_k(I_N) - \mu(M)| \leqslant |\mu(I_N) - \mu(M)| \quad \text{for each } N .$$

Now $|\mu_k(M) - \mu(M)| \leqslant |\mu_k(I_N) - \mu(M)| + |\mu_k(J_N) - \mu(M)|$, since $a \leqslant b \leqslant c$ implies $|b| \leqslant |a| + |c|$, whatever the values of a, b, c . Thus

$$\overline{\lim_n} \frac{1}{n} \sum_{k=o}^{n-1} |\mu_k(M) - \mu(M)| \leqslant \overline{\lim_n} \frac{1}{n} \sum_{k=o}^{n-1} |\mu_k(I_N) - \mu(M)| + \overline{\lim_n} \frac{1}{n} \sum_{k=o}^{n-1} |\mu_k(J_N)-\mu(M)|$$

$$\leq |\mu(I_N) - \mu(M)| + |\mu(J_N) - \mu(M)| \quad \text{for all } N .$$

Let $N \longrightarrow \infty$: $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\mu_k(M) - \mu(M)| = 0$. //

The conjecture that T is weak mixing if, and only if,

$$\frac{1}{n} \sum_{0}^{n-1} h_{\mu_k}(T^{(2)}) \longrightarrow h_\mu(T^{(2)}) \quad (n \longrightarrow \infty)$$

if, and only if ,

$$\frac{1}{n} \sum_{0}^{n-1} |\mu_k(M) - \mu(M)| \longrightarrow 0 \quad (n \longrightarrow \infty) \quad \text{for all } M \in \mathcal{B}^{(2)}$$

is disproved by 3.5 and 3.6.

Let $\overline{\mu_n}$ be the measure defined on $\mathcal{B}^{(2)}$ by

$$\overline{\mu_n}(E) = \frac{1}{n} \sum_{0}^{n-1} \mu_k(E) , \quad E \in \mathcal{B}^{(2)} .$$

<u>3.8 Theorem</u> The following are equivalent

(i) T is ergodic

(ii) $H_{\overline{\mu_n}}(\theta \times \theta') \longrightarrow H_\mu(\theta \times \theta')$ $(n \longrightarrow \infty)$ for all $\theta, \theta' \in Z_p$

(iii) $\overline{\mu}_n(M) \longrightarrow \mu(M)$ $(n \longrightarrow \infty)$ for all Jordan measurable $M \in \mathcal{B}^{(2)}$

<u>Proof</u> Since by the ergodic theorem T is ergodic if, and only if,

$$\overline{\mu}_n(E \times F) \longrightarrow \mu(E \times F) \quad (n \longrightarrow \infty) \quad \text{for all } E, F \in \mathcal{B} ,$$

we can replace μ_n by $\overline{\mu}_n$ in 3.1 . //

That 3.8 represents the best possible results in these directions follows from 3.5 and 3.6.

For each sequence $\Delta_n^r \in \mathfrak{D}^{(r)}$ $(n = 1, 2, \ldots)$ let $\mu_{\Delta_n^r}$ be the measure defined on $(X^{(r+1)}, \mathcal{B}^{(r+1)})$ by

$$\mu_{\Delta_n^r}(E) = p[\pi\{E \cap G_{\Delta_n^r}(T)\}] , \quad E \in \mathcal{B}^{(r+1)} ,$$

where $G_{\Delta_n^r}(T) = \{(x, T^{k_n^1} x , \ldots, T^{k_n^r} x) : x \in X\}$.

<u>3.9 Theorem</u> The following are equivalent:-

(i) T is mixing of degree r

(ii) $H_{\mu_{\Delta_n^r}}(\theta_0 \times \ldots \times \theta_r) \longrightarrow H_{\mu^{(r+1)}}(\theta_0 \times \ldots \times \theta_r)$ $(n \longrightarrow \infty)$ for all $\Delta_n^r \in \mathfrak{D}^{(r)}$ and $\theta_i \in Z_p$ $i = 0, \ldots, r$.

(iii) $\mu_{\Delta_n^r}(M) \longrightarrow \mu^{(r+1)}(M)$ $(n \longrightarrow \infty)$ for all $\Delta_n^r \in \mathfrak{D}^{(r)}$ and Jordan measurable $M \in \mathcal{B}^{(r+1)}$.

<u>Proof</u>. follows that of 3.1. //

<u>3.10 Lemma</u> $T^{(r+1)}$ preserves $\mu_{\Delta_n^r}$ for each n and $\Delta_n^r \in \mathfrak{D}^{(r)}$

<u>Proof</u> $\pi\{G_{\Delta_n^r}(T) \cap (T^{(r+1)})^{-1} E\} = T^{-1} \pi\{G_{\Delta_n^r}(T) \cap E\}$

for all $E \in \mathcal{B}^{(r+1)}$. //

3.11 Theorem $\mu^{(r+1)}(G_{\Delta_n^r}(T)) = 0$ for all n and $\Delta_n^r \in \mathfrak{D}^{(r)}$.

Proof $\mu^{(r+1)}(G_{\Delta_n^r}(T)) = \int_X \mu^{(r)}\{(G_{\Delta_n^r}(T)]_x) dp$

$$= \int_X \mu^{(r)}(\{(T^{k_n^1} x, \ldots T^{k_n^r} x)\}) dp$$

$= 0$, since p is non-atomic. //

3.12 Theorem $(X^{(r+1)}, \mathcal{B}^{(r+1)}, \mu_{\Delta_n^r}, T^{(r+1)})$ is almost isomorphic to $(X, \mathcal{B}, p, T)$.

Proof $\psi_{\Delta_n^r}(x) = (x, T^{k_n^1} x, \ldots, T^{k_n^r} x) : X \longrightarrow G_{\Delta_n^r}(T)$ gives the required isomorphism. //

As before, 3.9 gives the best possible results in these directions: Bernoulli endomorphisms with countable state space are mixing if all degrees, yet have positive, finite entropy, while periodic endomorphisms of non-atomic Lebesgue spaces are not mixing of any degree $r \geq 1$, yet have zero entropy; for any given sequence $\Delta_n^r \in \mathfrak{D}^{(r)}$, $\bigcup_{n=1}^{\infty} G_{\Delta_n^r}(T) \in \mathcal{B}^{(r+1)}$,

$$\mu_{\Delta_n^r}\left(\bigcup_{m=1}^{\infty} G_{\Delta_m^r}(T)\right) = 1 \text{ for each } n \text{ , but } \mu^{(r+1)}\left(\bigcup_{m=1}^{\infty} G_{\Delta_m^r}(T)\right) = 0$$

§4 The Topological Case

Let $(X, \mathcal{B}, p)$ now be a compact, Hausdorff probability space, and T be a continuous measure-preserving transformation of $(X, \mathcal{B}, p)$. $(X^{(r+1)}, \mathcal{B}^{(r+1)}, \mu^{(r+1)}$ is also compact (by Tychonoff's theorem) and Hausdorff, with respect to the product topology, $r \geq 0$.

4.1 Theorem $\mathcal{A}^{(r+1)}$ is dense in $C^{(r+1)}$ with respect to the uniform topology on $C^{(r+1)}$, $r \geq 0$.

Proof $\mathcal{A}^{(r+1)}$ contains the constant functions, since $C^{(r+1)}$ does.

$\mathcal{A}^{(r+1)}$ separates points of $X^{(r+1)}$:

Let $(x_0, \dots, x_r) \neq (y_0, \dots, y_r)$. Then at least one $x_i \neq y_i$. Since $X^{(r+1)}$ is Hausdorff, there exist disjoint open sets U_i, V_i such that

$$x_i \in U_i, \; y_i \in V_i.$$

$X^{(r+1)}$, being also compact, is completely regular. Hence there exist $f_i, g_i \in C^{(1)}$ such that

$$f_i(x_i) = 0 \qquad f_i(\complement V_i) = 1$$

$$g_i(y_i) = 0 \qquad g_i(\complement V_i) = 1 \quad .$$

$$f(z_0, \dots, z_r) = \prod_i f_i(z_i) + 2 \prod_i g_i(z_i) ,$$

where the products are taken over indices i for which $x_i \neq y_i$, is in $C^{(r+1)}$ and separates $(x_0, \dots, x_r)$ from $(y_0, \dots, y_r)$. Hence by the Stone Weierstrass Theorem [15, p 244], $\mathcal{A}^{(r+1)}$ is dense in $C^{(r+1)}$. //

4.2 Theorem   T is mixing of degree r if, and only if,

$$\int_{X^{(r+1)}} F \, d\mu_{\Delta_n^r} \longrightarrow \int_{X^{(r+1)}} F \, d\mu^{(r+1)} \quad (n \longrightarrow \infty)$$

for all $\Delta_n^r \in \mathfrak{D}^{(r)}$ and $F \in C^{(r+1)}$.

Proof   T is mixing of degree r if, and only if ,

$$\int_{X^{(r+1)}} f_o(x_o) \dots f_r(x_r) \, d\mu_{\Delta_n^r} (x_o, \dots, x_r) = \int_X f_o(x) \dots f_r(T^{k_n^r} x) \, dp$$

$$\longrightarrow \int_{X^{(r+1)}} f_o(x_o) \dots f_r(x_r) d\mu^{(r+1)} (x_o, \dots, x_r) = \int_X f_o(x) dp \dots \int_X f_r(x) dp$$

$(n \longrightarrow \infty)$ for all $\Delta_n^r \in \mathfrak{D}^{(r)}$ and $f_i \in L_1(p)$, and in particular for all $f_i \in C^{(1)}$. To see this, consider characteristic functions and use $L_1(p)$ approximation. By linearity of integrals, it follows that T is mixing of degree r if, and only if,

$$\int_{X^{(r+1)}} F \, d\mu_{\Delta_n^r} \longrightarrow \int_{X^{(r+1)}} F \, d\mu^{(r+1)} \quad (n \longrightarrow \infty)$$

for all $\Delta_n^r \in \mathfrak{D}^{(r)}$ and $F \in \mathcal{a}^{(r+1)}$. The proof is completed by 4.1. //

This theorem includes the case when T is mixing, i.e. mixing of degree one. Putting r = 1 and replacing μ_n by $\bar{\mu}_n$ in 4.2, we get

<u>4.3 Theorem</u> T is ergodic if, and only if,

$$\int_{X^{(2)}} F d\bar{\mu}_n \longrightarrow \int_{X^{(2)}} F \, d\mu \quad (n \longrightarrow \infty)$$

for all $F \in C^{(2)}$.

<u>4.4 Theorem</u> T is weak mixing if, and only if,

$$\left.\begin{array}{l} \frac{1}{n} \sum_{0}^{n-1} \left[\int_{X^{(2)}} F d\mu_k - \int_{X^{(2)}} F \, d\mu \right]^2 \longrightarrow 0 \quad (n \longrightarrow \infty) \\ \\ \frac{1}{n} \sum_{0}^{n-1} \left| \int_{X^{(2)}} F \, d\mu_k - \int_{X^{(2)}} F \, d\mu \right| \longrightarrow 0 \quad (n \longrightarrow \infty) \\ \\ \int_{X^{(2)}} F d\mu_n \longrightarrow \int_{X^{(2)}} F \, d\mu \quad (n \longrightarrow \infty, \ n \notin J, \ \delta(J) = 0) \end{array}\right\} \text{for all } F \in C^{(2)} .$$

<u>Proof</u> Analogous to that of 4.2, replacing ordinary convergence by strong Cesaro convergence and its two equivalent forms of convergence. //

4.2 cannot be extended to all $F \in L_1(\mu^{(r+1)})$, as the function $\chi_{\bigcup_{n=1}^{\infty} G_{\Delta_n^r}(T)}$, for any sequence $\Delta_n^r \in \mathcal{D}$, shows. 4.3 and 4.4 similarly cannot be extended to all $F \in L_1(\mu)$ (consider $F = \chi_{\bigcup_1^{\infty} G(T^n)}$).

CHAPTER III

Linear Fractional Transformations Mod One.

§1 Introduction.

This chapter is concerned with the ergodic properties of f-transformations, T_f, which will be introduced, together with f-expansions, in §3. These transformations of the unit interval onto itself, which in general do not preserve Lebesgue measure, have been investigated by several authors. A. Rényi [29] gave sufficient conditions for the validity of f-expansions and a sufficient condition for T_f to be ergodic and have a finite invariant measure equivalent to Lebesgue measure. V. A. Rohlin [32] showed that this condition implies also that T_f is exact, and gave a formula for the entropy of T_f with respect to the invariant measure. W. Parry [27] gave an explicit formula for the invariant measure of one class of f-transformations, namely the linear mod one transformations $Tx = (\beta x + \alpha)$, $\beta > 1$, $0 \leq \alpha < 1$, where (y) denotes the fractional part of y.

Sufficient conditions for ergodicity and infinite ergodic index of a general many-one transformation of a probability space with a generator are given in §2. These conditions, when applied to f-transformations, generalise Rényi's condition for ergodicity and invariant measure, and are used in §§4, 5 to study two classes of linear fractional transformation mod one, some of which also satisfy Rényi's condition. In some cases, the invariant measure could be found, using a result of W. Parry [27] which is proved in §2, while in others we did not succeed in doing this. For the former, the entropy is computed [32] and also the frequency with which the digits occur in the f-expansion. Throughout the study of f-expansions, the distinction between dependent and independent digits plays an essential part. §6 lists those questions which we were unable to resolve.

§2 Ergodicity and Invariant Measure

Throughout this section, unless otherwise stated, T will denote a many-one, measurable and non-singular transformation of the probability space $(X, \mathcal{B}, p)$, where p is non-atomic.

2.1 Lemma. For each $n = 1, 2, \ldots$ let $\xi_n = \{E_n(y) : y \in X\}$ be a countable measurable partition of X such that $E_n(y) \searrow \{y\}$ $(n \longrightarrow \infty)$ for each $y \in X$, i.e. $\xi_n \nearrow \varepsilon$ $(n \longrightarrow \infty)$.

Then for all $E \in \mathcal{B}$,

$$\frac{p(E_n(y) \cap E)}{p(E_n(y))} \longrightarrow \chi_E(y) \quad (n \to \infty)$$ for almost all $y \in X$ and in $L_1(p)$ mean.

Proof Let $\hat{\xi}_n$ denote the sub-σ-algebra of $\mathcal{B}$ generated by ξ_n, $f(x)$ be any integrable function and $E(f|\hat{\xi}_n)$ be the conditional expectation of f with respect to $\hat{\xi}_n$. Then by the Martingale theorem,

$$E(f|\hat{\xi}_n)(x) \to E(f|\mathcal{B})(x) = f(x) \qquad (n \to \infty)$$

almost everywhere and in $L_1(p)$ mean.

Putting

$$F_n(x) = \sum_{y \in X} \chi_{E_n(y)}(x) \frac{\int_{E_n(y)} f(z)\, d\mu(z)}{\mu(E_n(y))}$$

we have that $E(f|\hat{\xi}_n)(x) = F_n(x)$, since

$F_n(x)$ is measurable with respect to $\hat{\xi}_n$ and

$$\int_{Y_n} F_n(x)\, d\mu(x) = \int_{Y_n} f(x)\, d\mu(x)$$

for all $Y_n \in \hat{\xi}_n$, each such Y_n being a disjoint union of sets $E_n(y)$.

Putting $y = x$ and $f(z) = \chi_E(z)$ gives the required result. //

If T has a finite or countable generator $\xi = \{X_n\}_0^N$, $0 < N \leq \infty$, let $\varepsilon_n(y) = y_n$ = the unique integer such that $T^{n-1} y \in X_{y_n}$, $n = 1, 2, \ldots$,

and $C_n(y) = X_{y_1} \cap T^{-1} X_{y_2} \cap \ldots \cap T^{-(n-1)} X_{y_n}$. Clearly, either $C_n(y) \cap C_n(y') = \phi$ or $C_n(y) = C_n(y')$, and

$$\{C_n(y) : y \in X\} = \bigvee_{i=0}^{n-1} T^{-i} \xi$$

Since ξ is a generator of T, $C_n(y) \searrow \{y\}$ $(n \longrightarrow \infty)$ for each $y \in X$.

$\varepsilon_n(x)$ is a measurable function of x, since for every Borel set B in $[0, \infty)$, $\varepsilon_n^{-1}(B) = \bigcup_{k \in B} T^{-(n-1)} X_k$.

$$\varepsilon_n(x) \in L_1(p) \text{ if, and only if, } \sum_{k=1}^{N} k\, p(X_k) < \infty .$$

$$\varepsilon_n(x) = \varepsilon_1(T^{n-1} x) .$$

For each $y \in X$ and $n = 1, 2, \ldots$, the probability measure

$$\nu_n(E; y) = p(C_n(y) \cap T^{-n} E)$$

is absolutely continuous with respect to p. Thus, by the Radon-Nikodym theorem there exists a positive, integrable function $\omega^n(x, y)$ defined almost everywhere such that

$$p(C_n(y) \cap T^{-n} E) = \int_E \omega^n(x, y)\, dp(x) \quad , \quad E \in \mathcal{B} .$$

In fact, for each y and n, $0 \leq \omega^n(x, y) \leq 1$ for almost all x.

2.2 <u>Lemma</u> For each $m = 1, 2, \ldots$ let $\xi_m = \{I_m(x) : x \in X\}$ be a countable or finite measurable partition of $(X, \mathcal{B}, p)$ such that $I_m(x) \searrow \{x\}$ $(m \longrightarrow \infty)$ for each $x \in X$.

Then for $n = 1, 2, \ldots$ and $y \in X$,

$$\omega^n(x, y) = \lim_{m \to \infty} \frac{p(C_n(y) \cap T^{-n} I_m(x))}{p(I_m(x))}$$

for almost all x and $L_1(p)$ mean .

<u>Proof</u> As in 2.1,

$$\sum_{z \in X} \chi_{I_m(z)}(x) \frac{\int_{I_m(z)} \omega^n(w,y)\, dp(w)}{p(I_m(z))} \to \omega^n(x,y) \quad (m \to \infty)$$

for almost all x and in $L_1(p)$ mean.//

2.3 <u>Theorem</u> If

$$\overline{\lim_n} \frac{\inf_x \omega^n(x,y)}{p(C_n(y))} > 0 \qquad \text{for almost all } y ,$$

then T is ergodic.

If further

$$\underline{\lim_n} \frac{\inf_x \omega^n(x,y)}{p(C_n(y))} > 0 \qquad \text{for almost all } y ,$$

then T has infinite ergodic index.

<u>Proof</u>. Suppose $T^{-1} E = E \in \mathcal{B}$ and $0 < p(E) < 1$. Then

$$\frac{p(C_n(y) \cap E)}{p(C_n(y))} = \frac{p(C_n(y) \cap T^{-n} E)}{p(C_n(y))}$$

$$= \int_E \frac{\omega^n(x,y)}{p(C_n(y))}\, dp(x)$$

$$\geq \frac{\inf_x \omega^n(x,y)}{p(C_n(y))} \cdot p(E)$$

Since E is a generator of T , 2.1 implies that

$$\chi_E(y) = \lim_{n\to\infty} \frac{p(C_n(y) \cap E)}{p(C_n(y))}$$

$$\geq \overline{\lim_n} \frac{\inf_x \omega^n(x,y)}{p(C_n(y))} \; p(E)$$

$$\geq \underline{\lim_n} \frac{\inf_x \omega^n(x,y)}{p(C_n(y))} \cdot p(E) ,$$

from which it follows that $\chi_E(y) > 0$ for almost all $y \in X$, i.e. that $E = X \pmod 0$. Note that the theorem remains true if $\inf_x \omega^n(x, y)$ is taken, not over all x, but over almost all x.

To prove the second assertion, let S be another non-singular, measurable and many-one transformation on $(X, \mathcal{B}, p)$ with generator $\eta = \{Y_m\}_0^M$ where $0 < M \leq \infty$. We sharpen the previous notation as follows:-

$$C_n^S(y) = Y_{y_1} \cap S^{-1} Y_{y_2} \cap \ldots \cap S^{-(n-1)} Y_{y_n}$$

$$p(C_n^S(y) \cap S^{-n}E) = \int_E \omega_S^n(x,y)\, dp(x) , \quad E \in \mathcal{B}$$

and similarly for T. It is further assumed that

$$\underline{\lim_n} \frac{\inf_x \omega_S^n(x,y)}{p(C_n^S(y))} > 0 \qquad \text{for almost all } y .$$

T x S is a non-singular, measurable and many-one transformation on

$(X \times X, \mathcal{B} \times \mathcal{B}, p \times p)$ with generator $\xi \times \eta$. If $\underline{z}_i = (x_i, y_i)$ $i = 1, 2$, then

$$C_n^{T\times S}(\underline{z}_2) = [X_{\xi_1(x_2)} \times Y_{\xi_1(y_2)}] \cap \ldots \cap [(T\times S)^{-(n-1)} X_{\xi_n(x_2)} \times Y_{\xi_n(y_2)}]$$

$$= [X_{\xi_1(x_2)} \cap \ldots \cap T^{-(n-1)} X_{\xi_n(x_2)}] \times [Y_{\xi_1(y_2)} \cap \ldots \cap S^{-(n-1)} Y_{\xi_n(y_2)}].$$

Thus $(p \times p)(C_n^{T\times S}(\underline{z}_2)) = p(C_n^T(x_2))\, p(C_n^S(y_2))$.

Also

$$C_n^{T\times S}(\underline{z}_2) \cap (T\times S)^{-n}(E\times F) = [C_n^T(x_2) \cap T^{-n}E] \times [C_n^S(y_2) \cap S^{-n} F]$$

for all $E, F \in \mathcal{B}$ and $\underline{z}_2 \in X \times X$.

Therefore,

$$\iint_{E\times F} \omega_{T\times S}^n(\underline{z}_1, \underline{z}_2)\, d(p\times p)(\underline{z}_1) = \int_E \omega_T^n(x_1, x_2)\, dp(x_1) \int_F \omega_S^n(y_1, y_2)\, dp(y_1)$$

$$= \iint_{E\times F} \omega_T^n(x_1, x_2)\, \omega_S^n(y_1, y_2)\, d(p\times p), \text{ by Fubini's theorem.}$$

Thus $\omega_{T\times S}^n(\underline{z}_1, \underline{z}_2) = \omega_T^n(x_1, x_2)\, \omega_S^n(y_1, y_2)$, for almost all $\underline{z}_1$, since $\{E \times F : E \in \mathcal{B}, F \in \mathcal{B}\}$ generates $\mathcal{B} \times \mathcal{B}$. Finally, throwing out a $p \times p$-null set from $X \times X$, we have that

$$\lim_n \frac{\inf_{\underline{z}_1} \omega_{T\times S}^n(\underline{z}_1, \underline{z}_2)}{(p\times p)(C_n^{T\times S}(\underline{z}_2))} = \lim_n \left[\frac{\inf_{\underline{z}_1} \omega_T^n(x_1, x_2)\, \omega_S^n(y_1, y_2)}{p(C_n^T(x_2))\, p(C_n^S(y_2))} \right]$$

$$= \underline{\lim_n} \left[\frac{\inf_{x_1} \omega_T^n(x_1, x_2)}{p(C_n^T(x_2))} \cdot \frac{\inf_{y_1} \omega_S^n(y_1, y_2)}{p(C_n^S(y_2))} \right]$$

$$\geq \underline{\lim_n} \left[\frac{\inf_{x_1} \omega_T^n(x_1, x_2)}{p(C_n^T(x_2))} \right] \cdot \underline{\lim_n} \left[\frac{\inf_{y_1} \omega_S^n(y_1, y_2)}{p(C_n^S(y_2))} \right]$$

> 0 for almost all $\underline{z}_2$.

Putting $S = T^{n-1}$ $(n = 1, 2, \ldots)$ and using induction give that T has infinite ergodic index. //

We note that this result is independent of whether T preserves a finite or σ-finite measure. In the former case, T could only have ergodic index 0, 1 or $+\infty$, while in the latter case the ergodic index of T could take any non-negative integer or $+\infty$ as value [13].

Assume for the rest of this section that T is one-one on each set X_n of the generator ξ $(0 \leq n \leq N)$ and that $TF \in \mathcal{B}$ for all $F \in \mathcal{B}$. It follows that $X_n \cap T^{-1}\{x\}$ consists of a single point, where $\{x\}$ denotes the point set x. Let $\gamma_n(E) = p(X_n \cap T^{-1}E)$. The measure γ_n is absolutely continuous with respect to p, and hence has a Radon-Nikodym derivative $\frac{d\gamma_n}{dp}(x)$.

2.4 <u>Lemma</u> For all measurable functions h , for all $E \in \mathcal{B}$ and each n = 0, 1, ..., N,

$$\int_{X_n \cap T^{-1}E} h(y)\, dp(y) = \int_E h(X_n \cap T^{-1}\{x\})\, d\gamma_n(x)$$

in the sense that if one side is finite, then so is the other and they are equal.

<u>Proof</u> Let $F \in \mathcal{B}$. Since $p(X) = 1$, $\chi_F(x) \in L_1(p) \cap L_1(\gamma_n)$, $0 \leq n \leq N$. Noting that $\chi_{A \cup B}(x) = \chi_A(x) + \chi_B(x)$, whenever $A \cap B = \phi$, and that $F \cap X_n = \phi$ implies the vanishing of both the integrals in the Lemma, it is sufficient to consider $F \subset X_n$, for which

$$F = X_n \cap T^{-1} T F .$$

The inclusion $F \subset X_n \cap T^{-1} T F$ follows from $F \subset T^{-1} T F$, while if $x \in X_n \cap T^{-1} T F$, then $T x \in T F$, say $Tx = Ty$, $y \in F$; but T is one-one on each X_n and so $x = y$. Also

$$\chi_F(X_n \cap T^{-1}\{x\}) = \chi_{TF}(x) ,$$

and so

$$\int_E \chi_F(X_n \cap T^{-1}\{x\})\, d\gamma_n(x) = p(X_n \cap T^{-1}(E \cap TF))$$

$$= p(X_n \cap T^{-1}E \cap T^{-1}TF)$$

$$= p(T^{-1}E \cap F)$$

$$= p(T^{-1}E \cap X_n \cap F)$$

$$= \int_{X_n \cap T^{-1}E} \chi_F(x)\, dp(x) .$$

Since $\int_E f(x)dp(x)$ is a linear functional on $L_1(p)$, it follows that 2.4 is true for arbitrary measurable, simple functions $\sum_{\nu=1}^{k} a_\nu \chi_{A_\nu}(x)$.

$$\int_{X_n \cap T^{-1}E} h(y)\,dp(y) < \infty \qquad \text{if, and only if,}$$

$h(y)$ is integrable on $X_n \cap T^{-1} E$ if, and only if,

there exist measurable, elementary functions $h_m(x)$ such that

$h_m(x) \nearrow h(x)$ $(m \longrightarrow \infty)$ uniformly on $X_n \cap T^{-1} E$ and then

$$\int_{X_n \cap T^{-1}E} h(y)\,dp(y) = \lim_{m\to\infty} \int_{X_n \cap T^{-1}E} h_m(y)\,dp(y)$$

$$= \lim_{m\to\infty} \int_E h_m(X_n \cap T^{-1}\{x\})\,d\gamma_n(x)$$

$$= \int_E h(X_n \cap T^{-1}\{x\})\,d\gamma_n(x) \qquad \text{if, and only if,}$$

there exist measurable, elementary functions $h_m(x)$ such that

$h_m(X_n \cap T^{-1}\{x\}) \nearrow h(X_n \cap T^{-1}\{x\})$ $(m \longrightarrow \infty)$ uniformly on E

$[X_n \cap T^{-1}\{x\} \in X_n \cap T^{-1} E$ if, and only if, $x \in E]$

if, and only if, $h(X_n \cap T^{-1}\{x\})$ is integrable on E . //

2.5 Theorem T has an invariant measure ν equivalent to p if, and only if, there exists a measurable function $h(x)$, $0 < h(x) < \infty$ almost everywhere, such that

$$h(x) = \sum_{n=0}^{N} h(X_n \cap T^{-1}\{x\}) \frac{d\gamma_n}{dp}(x) \qquad \text{almost everywhere,}$$

and then $\quad h(x) = \dfrac{d\nu}{dp}(x)$

<u>Proof</u> Suppose T has an invariant measure ν. Put $h(x) = \frac{d\nu}{dp}(x)$. Then $0 < h(x) < \infty$ almost everywhere.

$$\int_E h(x)\,dp(x) = \sum_n \int_{X_n \cap T^{-1}E} h(y)\,dp(y)$$

$$= \sum_n \int_E h(X_n \cap T^{-1}\{x\}) \frac{d\gamma_n}{dp}(x)\,dp(x) \text{ , by 2.4.}$$

If $N < \infty$, the integration and summation commute, while if $N = \infty$,

$$\sum_{n=0}^{m} h(X_n \cap T^{-1}\{x\}) \frac{d\gamma_n}{dp}(x) \nearrow \sum_{n=0}^{\infty} h(X_n \cap T^{-1}\{x\}) \frac{d\gamma_n}{dp}(x) \quad (m \to \infty)$$

and hence by [7 , theorem 27.B],

$$\sum_{n=0}^{\infty} \int_E h(X_n \cap T^{-1}\{x\}) \frac{d\gamma_n}{dp}(x)\,dp(x) = \int_E \sum_{n=0}^{\infty} h(X_n \cap T^{-1}\{x\}) \frac{d\gamma_n}{dp}(x)\,dp(x)$$

for all $E \in \mathcal{B}$.

Conversely, let $\nu(E) = \int_E h(x)\,dp(x)$, $E \in \mathcal{B}$.

Clearly, ν is a measure equivalent to p .

$$\nu(T^{-1}E) = \int_{T^{-1}E} h(x)\,dp(x)$$

$$= \sum_n \int_{X_n \cap T^{-1}E} h(y)\,dp(y)$$

$$= \sum_n \int_E h(X_n \cap T^{-1}\{x\}) \frac{d\gamma_n}{dp}(x)\,dp(x)$$

$$= \int_E \sum_n h(X_n \cap T^{-1}\{x\}) \frac{d\gamma_n}{dp}(x)\,dp(x)$$

$$= \nu(E) \quad //$$

§3 f-expansions and f-transformations

Let $f : [0, \infty) \longrightarrow [0, \infty)$ be a differentiable function such that $f' : [0, \infty) \longrightarrow (-\infty, \infty)$ is continuous. We distinguish two cases:

A) $f(x)$ strictly decreasing, $x \in (f^{-1}(1), \infty)$

$$\lim_{x \longrightarrow \infty} f(x) = 0$$

$$|f'(x)| < 1 \ , \quad x \in (f^{-1}(1), \infty)$$

B) $f(x)$ strictly increasing, $x \in [0, f^{-1}(1))$

$$f(0) = 0$$

$$|f'(x)| < 1 \ , \quad x \in [0, f^{-1}(1)) \ .$$

In both cases a further distinction is necessary, namely

(1) $f^{-1}(1)$ is zero or a non-negative integer or $+\infty$

(2) $f^{-1}(1)$ is a finite, positive non-integer.

Such a function f can be associated with a measurable, non-singular transformation $T_f : (I, \mathcal{B}, p) \longrightarrow (I, \mathcal{B}, p)$,

where
$$I = \begin{cases} (0, 1) & \text{case A)} \\ [0, 1) & \text{case B)} \end{cases}$$

$\mathcal{B}$ = Borel subsets of I

p = Lebesgue measure on I .

Let

$$T_f(x) = (f^{-1}(x)) \ , \quad x \in I$$

and $\varepsilon_n(x) = [f^{-1}\{T_f^{n-1}(x)\}]$, $x \in I$, $n \geqslant 1$,
where (y) and $[y]$ denote the fractional and integer parts of y, respectively. T_f is called an <u>f-transformation</u>.

For any given function f , f-expansions are said to be valid if for all $x \in I$ either

$$T_f^n(x) = 0 \text{ for some } n ,$$

in which case x has the finite f-expansion

$$\begin{aligned} x &= f(\varepsilon_1(x) + f(\varepsilon_2(x) + \dots + f(\varepsilon_n(x))\dots)) \\ &\equiv f_n(\varepsilon_1(x), \dots, \varepsilon_n(x)) , \end{aligned}$$

or

$$\lim_{n\to\infty} f_n(\varepsilon_1(x), \dots, \varepsilon_n(x)) \text{ exists and equals } x .$$

$\varepsilon_1(x)$, $\varepsilon_2(x)$, ... are called the <u>digits</u> of x in its f-expansion. They take non-negative, integral values. The values they can take, or their <u>admissible</u> values, depend on f, as will be seen later. A finite sequence of non-negative integers $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$ is called <u>canonical</u> if there is a $y \in X$ such that

$$\varepsilon_i(y) = \varepsilon_i \quad (i = 1, \dots n) .$$

f-expansions with <u>independent digits</u> occur if every sequence of admissible digits is canonical. In the opposite case, the digits are <u>dependent</u>. This terminology was introduced by Renyi [29], and should not be confused with stochastic independence. This distinction between these two kinds of independence is discussed later (3.9). $\varepsilon_n(x)$ are measurable functions of x, since for all Borel sets B of $[0, \infty)$, $\{x : \varepsilon_n(x) \in B\} = \bigcup_{k\in B} T_f^{-(n-1)}X_k$.

There is a natural partition $\xi = \{X_n\}$ associated with T_f, namely that for which $X_n = \{x : \varepsilon_1(x) = n\}$. For the four cases considered, we have

A1) $X_n = (f(n+1), f(n))$, $n = f^{-1}(1),\ f^{-1}(1) + 1, \ldots$

A2) $$X_n = \begin{cases} (f([f^{-1}(1)]+ 1), 1)\ , & n = [f^{-1}(1)] \\ (f(n+1), f(n)) & ,\ n = [f^{-1}(1)]+1,\ [f^{-1}(1)]+2, \ldots . \end{cases}$$

Note that ξ is a partition mod 0 of I, since the countable, and hence null, set of subdivision points of ξ are omitted. The admissible digits here are $[f^{-1}(1)]$, $[f^{-1}(1)] + 1, \ldots$.

B1) $X_n = [f(n), f(n+1))$, $n = 0, 1, \ldots, f^{-1}(1) - 1$.

B2) $$X_n = \begin{cases} [f(n), f(n+1))\ , & n = 0, 1, \ldots, [f^{-1}(1)] - 1 \\ [f([f^{-1}(1)]), 1), & n = [f^{-1}(1)]\ . \end{cases}$$

The admissible digits now are

$0, 1, \ldots, f^{-1}(1) - 1$ case B1)

$0, 1, \ldots, [f^{-1}(1)]$ case B2)

Each $C_n(y)$ is an interval, being a finite intersection of intervals.

T_f is measurable, since for each $[a, b) \subset I$,

$X_n \cap T_f^{-1}[a, b)$ is an interval,

$T_f^{-1}[a, b) = \bigcup_n X_n \cap T_f^{-1}[a, b)$, the union being taken over the index set of ξ, the ring (algebra in case B)) of finite, disjoint unions of half-open intervals, i.e. of the form $[a, b)$, generates $\mathcal{B}$ and T_f^{-1} commutes with set-theoretic union, intersection and difference. Also $T_f E \in \mathcal{B}$ for all $E \in \mathcal{B}$ since for all $[a, b) \subset I$ $T_f[a, b)$ is a finite or countable disjoint union of half-open intervals,

$$T_f A = \bigcup_n T_f (A \cap X_n)$$

and the restriction of T_f to each X_n commutes with set-theoretic operations of union, intersection and difference.

$[f^{-1}(x)]$ is a step function with countable number of discontinuities and so $T_f(x)$ is an almost everywhere differentiable function of x, with

$$T_f'(x) = \frac{1}{f'(f^{-1}(x))} \quad \text{almost everywhere.}$$

Since f' is assumed continuous and strictly positive (or negative), for each $E \in \mathcal{B}$, $X_n \cap T_f^{-1} E$ is a 'continuously shrunk' ('continuously shrunk and reversed') version of E. Thus T_f is non-singular.

One would expect valid f-expansions to distinguish between different points, although two distinct f-expansions may represent the same point. A simple example is afforded by any r-adic expansion ($f(x) = \frac{x}{r}$, r an integer), for which

$$0.1 \quad \text{and} \quad 0.0\,(r-1)(r-1)\ldots$$

represent the same real number $\frac{1}{r}$. In fact the following result is true.

3.1 <u>Theorem</u> [27] Let $f : [0, \infty) \longrightarrow [0, 1)$ be either strictly increasing or strictly decreasing throughout its domain of definition. Then f-expansions are valid if, and only if,

$$\varepsilon_n(x) = \varepsilon_n(y) \qquad n = 1, 2, \ldots$$

implies $x = y$.

<u>Proof</u> <u>f increasing</u>:

Let $\rho_n(x) = f_n(\varepsilon_1(x), \ldots, \varepsilon_n(x))$. Then for each $x \in [0, 1)$

$$\rho_n(x) \leq \rho_{n+1}(x) \leq (x) = f_n(\varepsilon_1(x), \ldots, \varepsilon_n(x) + T_f^n(x)).$$

Thus $\rho(x) = \lim\limits_{n \to \infty} \rho_n(x)$ exists and $\rho(x) \leq x$.

(=>) If $\varepsilon_n(x) = \varepsilon_n(y) \quad n = 1, 2, \ldots$

then $\rho_n(x) = \rho_n(y) \quad n = 1, 2, \ldots$

and so $x = \rho(x) = \rho(y) = y$.

(<=) Since for each x and for each n, $\rho_n(x) \leq \rho(x) \leq x$, the result will follow from $[\rho_n(x), x] \subset C_n(x)$ for each n:

Suppose

$$\rho_n(x) \leq y \leq x .$$

Then $\varepsilon_1(x) + T_f(\rho_n(x)) \leq T_f(y) + \varepsilon_1(y) \leq T_f(x) + \varepsilon_1(x)$,

from which it follows that $\varepsilon_1(y) = \varepsilon_1(x)$ and $T_f(\rho_n(x)) \leq T_f(y) \leq T_f(x)$.

So by induction $\varepsilon_i(y) = \varepsilon_i(x) \quad i = 1, \ldots, n$.

f decreasing : the proof is analogous, noting that for each $x \in (0, 1)$

$$\rho_{2n}(x) \leq \rho_{2n+2}(x) \leq x \leq \rho_{2n+1}(x) \leq \rho_{2n-1}(x) \quad n = 1, 2, \ldots \quad //$$

3.2 Corollary f -expansions are valid if, and only if, ξ is a generator of T_f .

Proof f-expansions are valid if, and only if, $c_n(y) \searrow \{y\}$ $(n \to \infty)$ for each y.

$$\bigvee_{i=0}^{n-1} T^{-i}\xi = \{c_n(y) : y \in X\} \quad //$$

3.3 Theorem [27] If f satisfies A) or B) , then f-expansions are valid for all $x \in I$.

Proof The Mean Value theorem implies that for $x \neq y$

$$\left| \frac{f(x) - f(y)}{x - y} \right| < 1$$

This is equivalent to the following: if $\delta > 0$, there exists $\varepsilon > 0$ such that for $|x - y| > \delta$,

$$\left| \frac{f(x) - f(y)}{x - y} \right| < 1 - \varepsilon .$$

The second condition trivially implies the first, while if for some $\delta > 0, |x - y| > \delta$ implies $\left| \frac{f(x) - f(y)}{x - y} \right| \geq 1 - \varepsilon$ for all $\varepsilon > 0$ we get a contradiction of the first condition.

Suppose $x, y \in I, x \neq y, |x - y| > \delta$, say, yet

$\varepsilon_n(x) = \varepsilon_n(y) \quad n = 1, 2, \ldots$.

Then there is $\varepsilon > 0$ such that

$$\left| \frac{T_f(x) - T_f(y)}{x - y} \right| = \left| \frac{f^{-1}(x) - f^{-1}(y)}{x - y} \right| > \frac{1}{1 - \varepsilon} .$$

By induction,

$$\left| \frac{T_f^n(x) - T_f^n(y)}{x - y} \right| > \frac{\delta}{(1 - \varepsilon)^n} , \quad n = 1, 2, \ldots ,$$

which is impossible since $|T_f^n(x) - T_f^n(y)| \leq 1$. Hence the result, by 3.1. //

3.4 <u>Corollary</u> If $|T_f'(x)| > 1$ almost everywhere, then f-expansions are valid.

<u>Proof</u> $T_f(x) = f^{-1}(x) - [f^{-1}(x)]$, and the set of discontinuities of the step function $[f^{-1}(x)]$ is at most countable. //

Regarding the dependence of the digits in f-expansions where f satisfies A) or B), case 1) corresponds to independent digits, since then $T_f X_n = I$ for all n.

Case 2) gives rise to dependent digits:

By the assumption on the domain of f, $f^{-1}(1) \in [0, \infty)$, i.e. $f^{-1}(1) < \infty$.

A2) Let $M = [f^{-1}(1)]$. Although we have only considered $(0, 1)$, $T_f(1)$ is well-defined by $T_f(1) = (f^{-1}(1))$. Thus $0 < T_f(1)$.

Consider $(0, T_f(1))$. Since $T_f(X_M) = (T_f(1), 1)$,

$$X_M \cap T_f^{-1}(0, T_f(1)) = \phi .$$

Now $\lim_{x \to \infty} f(x) = 0$ implies that

$$X_n \cap (T_f(1), 1) = \phi$$

if, and only if, $f(n) < T_f(1)$, i.e. $n > f^{-1}(T_f(1))$. Thus

$$X_n \subset (0, T_f(1))$$

for all $n > \varepsilon_2(1)$. No assumption is made about the validity of the f-expansion for 1. Any sequence of digits containing the subsequence $M, \varepsilon_2(1) + 1$, for example, is not canonical.

B2) Let $N = [f^{-1}(1)]$,

$$\delta_1 = N$$

$$\delta_n = \text{Max}\,\{\delta : N\delta_1 \ldots \delta_{n-1} \text{ is canonical}\}, \quad n > 1 .$$

Then $\varepsilon_n(1) = \delta_n$, $n = 1, 2, \ldots$, since $f_n(\delta_1, \ldots, \delta_n)$ increases with n, is not greater than one, and $x < \lim_{n \to \infty} f_n(\delta_1, \ldots, \delta_n)$ for all $x \in [0, 1)$.

If $\delta_n = N$ for all n, then $T_f(1) = 1$, which is impossible since $f^{-1}(1)$ is not an integer. Let $\bar{n}$ be the least n for which $\delta_n < N$. Then the sequence

$$\delta_1, \delta_2, \ldots, \delta_{\bar{n}} + 1$$

is not canonical, yet consists of admissible digits.

We now obtain an explicit formula for $\omega^n(x, y)$. For each $y \in I$ and $n \geq 1$, let

$$S_n^y(x) = C_n(y) \cap T_f^{-n}(x) \ : \quad I \longrightarrow C_n(y)$$

where $C_n(y)$ is defined in terms of the natural partition ξ associated with f. S_n^y is one-valued since T^n is 1 - 1 on each $C_n(y)$. Since f-expansions are assumed to be valid,

$$C_m(y) \searrow \{y\} \quad (m \longrightarrow \infty) ;$$

also $C_m(x)$ is an interval, with end points $a_m(x) < b_m(x)$, say. Then

$$p(C_n(y) \cap T^{-n} C_m(x)) = | S_n^y(b_m(x)) - S_n^y(a_m(x)) | ,$$

and so by 2.2

$$\omega^n(x, y) = \left| \frac{d}{dt} S_n^y(t) \Big|_{t = x} \right| ,$$

since in case A) or case B) $S_n^y(b_m(x)) - S_n^y(a_m(x))$ is of constant sign as $m \longrightarrow \infty$.

For the independent digit case,

$$S_n^y(t) = f_n(\varepsilon_1(y), \ldots, \varepsilon_n(y) + t) ,$$

since $T^n f_n(\varepsilon_1(y), \ldots, \varepsilon_n(y) + t) = t$

and

$$\varepsilon_r(f_n(\varepsilon_1(y), \ldots, \varepsilon_n(y)+t)) = \begin{cases} [f^{-1} f_{n-r+1}(\varepsilon_r(y), \ldots, \varepsilon_n(y)+t)] = \varepsilon_r(y) , & r = 1, \ldots, n-1 \\ [f^{-1} f(\varepsilon_n(y)+t)] = \varepsilon_n(y) , & r = n . \end{cases}$$

So for case 1),

$$\omega^n(x, y) = \left| \frac{d}{dt} f_n(\varepsilon_1(y), \ldots, \varepsilon_n(y) + t) \Big|_{t = x} \right| .$$

We have shown incidentally that for case 1) $C_n(y)$ is the interval with endpoints $f_n(\varepsilon_1(y), \ldots, \varepsilon_n(y))$ and $f_n(\varepsilon_1(y), \ldots, \varepsilon_n(y) + 1)$.

3.5 Theorem For the independent digit case T_f has invariant measure μ equivalent to p if, and only if, there exists a measurable function $h(x)$, $0 < h(x) < \infty$ almost everywhere, such that

$$h(x) = \sum_n h(f(x + n)) \, |f'(x + n)| \quad \text{almost everywhere,}$$

where the summation is taken over the index set of ξ, and then $h(x) = \frac{d\mu}{dp}(x)$.

Proof Using the notation of 2.5, $X_n \cap T^{-1}\{x\} = f(x + n)$ and

$$\frac{d\gamma_n}{dp}(x) = \omega^1(x, y), \text{ for any } y \text{ with } \varepsilon_1(y) = n$$

$$= f'(x + n) . \quad //$$

3.6 Theorem [29] If f-expansions are valid, with independent digits, and further

condition C)
$$\frac{\sup_{0<x<1} \omega^n(x, y)}{\inf_{0<x<1} \omega^n(x, y)} \leq C, \quad \text{for almost all } y$$

where $C \geq 1$ is independent of y and n,

then T_f is ergodic and has a finite, invariant measure ν equivalent to p such that

$$\frac{1}{C} \leq \frac{d\nu}{dp}(x) \leq C$$

2.3 is a genuine extension of 3.6, since for b = c = 1, the f-transformation studied in §5 satisfies 2.3 and has a σ-finite invariant measure equivalent to p. The first condition of 2.3 generalises condition C), because

$$\underline{\lim_{n}} \frac{\sup_x \omega^n(x,y)}{p(C_n(y))} \geq 1 \qquad \text{for all } y$$

and if C) holds,

$$\frac{1}{C} \leq \overline{\lim_{n}} \left[\frac{\inf_x \omega^n(x,y)/p(C_n(y))}{\sup_x \omega^n(x,y)/p(C_n(y))} \right]$$

$$\leq \frac{\overline{\lim_{n}} \inf_x \omega^n(x,y)/p(C_n(y))}{\underline{\lim_{n}} \sup_x \omega^n(x,y)/p(C_n(y))} \qquad \text{for all } y$$

and so $$\overline{\lim_{n}} \frac{\inf_x \omega^n(x,y)}{p(C_n(y))} \geq \frac{1}{C}$$

For the second condition of 2.3 suppose

$$\underline{\lim_{n}} \frac{\inf_x \omega^n(x,y)}{p(C_n(y))} = 0 \qquad \text{for } y \in E,\ p(E) > 0$$

and that C) holds. There exist $n_i = n_i(y)$, i = 1, 2, ... , such that

$$\frac{\inf_x \omega^{n_i}(x,y)}{p(C_{n_i}(y))} < \frac{1}{3C} \qquad i = 1, 2, \ldots \quad y \in E .$$

On the other hand,

$$\frac{\sup_x \omega^{n}(x,y)}{p(C_n(y))} \geq 1 \qquad \text{for all } y \text{ and } n .$$

Thus, for each $y \in E$ there exists a positive integer n_y such that

$$\frac{\sup_x \omega^{n_y}(x,y)/p(C_{n_y}(y))}{\inf_x \omega^{n_y}(x,y)/p(C_{n_y}(y))} > 3C > C ,$$

a contradiction.

3.7 Theorem. If f-expansions are valid, T_f is ergodic and has finite, invariant measure μ equivalent to p, then the (asymptotic) frequency of occurrence of the sequence of admissible digits $i_1 \ldots i_k$ in f-expansions is well-defined by

$$\phi_{i_1 \ldots i_k} = \lim_{n\to\infty} \frac{1}{n-k+1} \sum_{\nu=0}^{n-k} \chi_{X_{i_1} \cap \cdots \cap T_f^{-(k-1)} X_{i_k}} (T_f^{\nu}(x))$$

$$= \frac{\mu(X_{i_1} \cap \cdots \cap T_f^{-(k-1)} X_{i_k})}{\mu(I)} .$$

Proof The existence of $\phi_{i_1 \ldots i_k}$ follows from the Pointwise Ergodic Theorem, while its independence of x is implied by the ergodicity of T_f. //

3.8 Corollary If $i_1 \ldots i_k$ is a non-canonical sequence of digits, then $\phi_{i_1 \ldots i_k} = 0$.

Proof $i_1 \ldots i_k$ being non-canonical implies that

$$X_{i_1} \cap \ldots \cap T_f^{-(k-1)} X_{i_k} = \phi \ . \quad //$$

Normalising the measure μ, we have the following

3.9 Corollary If the random variables $\{\varepsilon_n(x)\}_{n=1}^{\infty}$ are independent [19, p.295], then the digits in f-expansions are independent.

Proof The independence implies that for any admissible digits $i_1 \ldots i_k$

$$\phi_{i_1 \ldots i_k} = \phi_{i_1} \ldots \phi_{i_k} \ .$$

Since $\phi_{i_r} > 0$, $1 \leq r \leq k$, 3.8 gives that $i_1 \ldots i_k$ is a canonical sequence of digits. //

The converse of 3.9 is false, as can be seen with $T_f(x) = (\frac{1}{x})$, 4.7, which has independent digits 1, 2, ...

$$\mu(X_1 \cap T_{\frac{1}{x}}^{-1} X_1) = \mu\left(\tfrac{1}{2}, \tfrac{2}{3}\right)$$

$$= \log \frac{10}{9}$$

$$\neq \left(\log \frac{4}{3}\right)^2$$

$$= \mu(X_1)^2 .$$

3.10 Theorem [32]. Under the conditions of 3.6, T_f is exact. T_f has finite entropy if, and only if, $\int_0^1 \log\left|\frac{d}{dx} f^{-1}(x)\right| d\mu < \infty$, in which case $h_\mu(T_f) = \int_0^1 \log\left|\frac{d}{dx} f^{-1}(x)\right| d\mu$.

§4 The transformation $T_f(x) = (\frac{1}{ax} - \frac{b}{a})$, when $f(x) = \frac{1}{ax+b}$

4.1 Theorem f-expansions are valid for all $a > 0$ and $0 \leqslant b \leqslant 1$.

Proof $f'(x)$ must be negative : $a > 0$

$f^{-1}(1)$ must be in $[0, \infty)$: $b \leqslant 1$.

If $b < 0$, dependent digits could give rise to negative $f_n(\varepsilon_1(x), \ldots, \varepsilon_n(x))$, e.g. $b = -\frac{1}{2}$, $a = 1$ when the admissible digits are 0, 1, ... and $f(0) = -2$. Independent digits never do this, however, since $a\,\varepsilon_n(x) + b \geqslant a\,(\frac{1-b}{a}) + b = 1$. Even in the latter case, where there do not seem to be a priori objections to $b < 0$, complications arise in the proofs. We thus take $b \geqslant 0$.

$$T_f(x) = \frac{1}{ax} - \frac{b}{a} - \left[\frac{1}{ax} - \frac{b}{a}\right]$$

is a piecewise continuous and differentiable function whose points of discontinuity are $\frac{1}{an+b}$, $n > [\frac{1-b}{a}]$. Therefore, provided x is not a point of discontinuity of T_f, T_f' exists and

$$T_f'(x) = -\frac{1}{ax^2} .$$

Hence

$$\infty > \frac{d}{dx} T_f^2(x) = T_f'(T_f(x))\, T_f'(x)$$

$$= \frac{1}{a\left\{\frac{1-bx}{ax} - \left[\frac{1-bx}{ax}\right]\right\}^2} \cdot \frac{1}{ax^2}$$

$$> 1 ,$$

since $0 < 1 - bx - ax\,[\frac{1-bx}{ax}] < 1$ for $x \neq \frac{1}{an+b}$, $n > [\frac{1-b}{a}]$.

From now on in the study of $T_{\frac{1}{ax+b}}$, we consider $I = \bigcup_{n=[\frac{1-b}{a}]+1}^{\infty} \{\frac{1}{an+b}\}$ $= I(\text{mod } 0)$.

Suppose $x \neq y$, yet $\varepsilon_n(x) = \varepsilon_n(y)$ for all n. Then $T_f^2 x$ is continuous in $[x,y]$ and differentiable in (x,y), so that

$$\left|\frac{T_f^2(x)-T_f^2(y)}{x-y}\right| = \frac{d}{dx} T_f^2(\xi) \quad , \quad \xi \in [x,y]$$

$$> 1.$$

This is equivalent to the following: if $\delta > 0$, there exists $\varepsilon > 0$ such that for $|x - y| > \delta$

$$\left|\frac{T_f^2(x)-T_f^2(y)}{x-y}\right| > 1+\varepsilon .$$

So $|x - y| > \delta$, $\varepsilon_n(x) = \varepsilon_n(y)$ for all n imply that

$$\left| T_f^{2n}(x) - T_f^{2n}(y)\right| > (1+\varepsilon)^n \delta ,$$

which is impossible for all n. Hence $\varepsilon_n(x) \neq \varepsilon_n(y)$ for some n, and f-expansions are valid by 3.1. //

4.2 Theorem Let

$$\frac{P_n(y,t)}{Q_n(y,t)} = f_n(\varepsilon_1(y),\ldots,\varepsilon_n(y)+t)$$

$$\equiv \frac{1}{b+a\varepsilon_1(y)+}\,\frac{a}{b+a\varepsilon_2(y)+}\cdots\frac{a}{+\,b+a(\varepsilon_n(y)+t)} .$$

Then

$$\frac{P_n(y,t)}{Q_n(y,t)} = \frac{\{b+a(\varepsilon_n(y)+t)\}P_{n-1}(y,0) + a\,P_{n-2}(y,0)}{\{b+a(\varepsilon_n(y)+t)\}Q_{n-1}(y,0) + a\,Q_{n-2}(y,0)} \quad , \; n \geqslant 3 .$$

Proof. Writing $\varepsilon_n(y) = y_n$, $P_n(t) = P_n(y,t)$ and $P_n = P_n(0)$ we have

$$\frac{P_n(t)}{Q_n(t)} = \frac{P_{n-1}\left(\frac{1}{b+a(y_n+t)}\right)}{Q_{n-1}\left(\frac{1}{b+a(y_n+t)}\right)}$$

$$= \frac{\left\{b+a\left(y_{n-1} + \frac{1}{b+a(y_n+t)}\right)\right\}P_{n-2} + a\,P_{n-3}}{\left\{b+a\left(y_{n-1} + \frac{1}{b+a(y_n+t)}\right)\right\}Q_{n-2} + a\,Q_{n-3}}$$

$$= \frac{\{b+a(y_n+t)\}\left[(b+ay_{n-1})P_{n-2} + a\,P_{n-3}\right] + a\,P_{n-2}}{\{b+a(y_n+t)\}\left[(b+ay_{n-1})Q_{n-2} + a\,Q_{n-3}\right] + a\,Q_{n-2}}$$

$$= \frac{\{b+a(y_n+t)\}P_{n-1} + a\,P_{n-2}}{\{b+a(y_n+t)\}Q_{n-1} + a\,Q_{n-2}} \quad , \qquad \text{assuming relation}$$

holds for n - 1; but

$$P_3(t) = \{b+a(y_3+t)\}(b+ay_2) + a$$
$$= \{b+a(y_3+t)\}P_2 + a\,P_1$$

and

$$Q_3(t) = \{b+a(y_3+t)\}\{(b+ay_1)(b+ay_2)+a\} + a(b+ay_1)$$

$$= \{b+a(y_3+t)\}Q_2 + a\,Q_1 . \quad //$$

4.2 implies that

$$\frac{P_n(y,t)}{Q_n(y,t)} = \frac{P_n(y,0) + at\,P_{n-1}(y,0)}{Q_n(y,0) + at\,Q_{n-1}(y,0)} .$$

4.3 <u>Theorem</u> For $a > 0$, $0 < b \leqslant 1$ and $f^{-1}(1) = \frac{1-b}{a} = 1, 2, \ldots,$ T_f is exact and has a finite invariant measure equivalent to p.

<u>Proof</u>

$$\omega^n(x,y) = \frac{a\,|P_{n-1}Q_n - P_nQ_{n-1}|}{(Q_n + axQ_{n-1})^2} .$$

Therefore,

$$\frac{\sup_{0<x<1} \omega^n(x,y)}{\inf_{0<x<1} \omega^n(x,y)} = \left(1 + \frac{a\,Q_{n-1}(y,0)}{Q_n(y,0)}\right)^2 .$$

But $Q_n(y,0) = (b+ay_n)\,Q_{n-1}(y,0) + a\,Q_{n-2}(y,0) > (b+ay_n)Q_{n-1}(y,0)$,

thus

$$\frac{a\,Q_{n-1}(y,0)}{Q_n(y,0)} < \frac{a}{b+ay_n} \leqslant \frac{a}{b} .$$

So 3.4 holds with $C = (1 + \frac{a}{b})^2$. //

4.4 <u>Theorem</u> For $b = 1$, T_f has finite, invariant measure μ given by

$$\frac{d\mu}{dp}(x) = \frac{1}{ax+1} .$$

<u>Proof</u>

$$\sum_{n=0}^{\infty} h\left(\frac{1}{a(x+n)+1}\right)\frac{a}{[a(x+n)+1]^2} = \sum_{n=0}^{\infty} \frac{1/a}{[x+\frac{1}{a}+n][x+\frac{1}{a}+n+1]}$$

$$= \frac{1}{ax+1} .$$

So by 3.4 , $\frac{d\mu}{dp}(x) = \frac{1}{ax+1}$. $\mu(I) = \frac{1}{a}\log(1+a) < \infty$. //

4.5 <u>Corollary</u>

$$\phi_i = \frac{\log\left(\frac{[a(i+1)+1]^2}{[ai+1][a(i+2)+1]}\right)}{\log(a+1)} , \quad i = 0, 1, \ldots .$$

4.6 <u>Proposition</u> $\varepsilon_n(x) \notin L_1(\mu)$ for each n.

<u>Proof</u> $\varepsilon_n(x) \in L_1(\mu)$ if, and only if, $\sum_{n=1}^{\infty} n \log\left(\frac{[a(n+1)+1]^2}{[an+1][a(n+2)+1]}\right) < \infty$;

but

$$\sum_{n=1}^{N} n\log\left(\frac{[a(n+1)+1]^2}{[an+1][a(n+2)+1]}\right) = \log[a(N+2)+1] + (N+1)\log\left[1-\frac{a}{a(N+2)+1}\right] - \log(a+1)$$

$\longrightarrow \infty$ ($N \longrightarrow \infty$) . //

4.7 <u>Corollary</u> $\frac{1}{n} \sum_{k=1}^{n} \varepsilon_k(x) \longrightarrow \infty$ $(n \longrightarrow \infty)$ almost everywhere.

<u>Proof</u> Note that $\varepsilon_k(x) = \varepsilon_1(T^{k-1}x)$ and, for N a positive integer, put

$$f_N(x) = \begin{cases} \varepsilon_1(x), & \text{if } \varepsilon_1(x) \le N \\ N, & \text{if } \varepsilon_1(x) > N. \end{cases}$$

Then $f_N(x) \in L_1(\mu)$ for each N and so

$$\frac{1}{n}\sum_{1}^{n} \varepsilon_k(x) \geqslant \frac{1}{n}\sum_{0}^{n-1} f_N(T^k x)$$

$$\to \sum_{1}^{N} k \log\left(\frac{[a(k+1)+1]^2}{[ak+1][a(k+2)+1]}\right) + N\sum_{N+1}^{\infty} \log\left(\frac{[a(k+1)+1]^2}{[ak+1][a(k+2)+1]}\right) \quad (n \to \infty)$$

$\longrightarrow \infty \; (N \longrightarrow \infty)$. //

4.8 <u>Corollary</u> The digits in f-expansions are almost everywhere unbounded.

4.9 <u>Corollary</u> $h_\mu(T_f) < \infty$; in fact

$$h_\mu(T_f) = -\frac{\log a \log(a+1)}{a} - 2\int_0^1 \frac{\log x}{ax+1}\,dx .$$

<u>Proof</u>

$$\int_0^1 \log\left|\frac{d}{dx} f^{-1}(x)\right| dp = \int_0^1 \log\left(\frac{1}{ax^2}\right) dx$$

$$= 2 - \log a < \infty . //$$

$I(a) = \int_0^1 \frac{\log x}{ax+1}\,dx$ does not appear to have a closed form expression when $a > 1$, although for all $a > 0$ it satisfies the differential equation

$$I'(a) + \frac{I(a)}{a} + \frac{\log(a+1)}{a^2} = 0 .$$

$I'(a) = \int_0^1 \frac{\partial}{\partial a} \frac{\log x}{ax+1}\,dx = a\int_0^1 \frac{-\log x}{(ax+1)^2}\,dx$, since

$-\frac{\log x}{(ax+1)^2} < -\log x$, $\int_0^1 -\log x\,dx = 1$ and thus $\int_0^1 -\frac{\log x}{(ax+1)^2}\,dx$ converges uniformly for $a > 0$; $\int_0^1 \frac{\log x}{ax+1}\,dx$ converges pointwise by 4.9;

$f(x, a) = -\frac{a \log x}{(ax+1)^2}$ is continuous on $(0,1] \times (0,\infty]$ and hence by [1 , p.443] integration and partial differentiation can be reversed.

For $a = 1$, [6 , p.563],

$$\int_0^1 \frac{\log x}{x+1}\,dx = -\frac{\pi^2}{12} ,$$

while for $0 < a < 1$,

$$\int_0^1 \frac{\log x}{ax+1}\,dx = -\sum_1^\infty \frac{(-a)^{n-1}}{n^2} :$$

$$\int_0^1 \frac{\log x}{ax+1}\,dx = \int_0^1 \log x\,dx + \int_0^1 \sum_1^\infty (-ax)^n \log x\,dx$$

$$= -1 + \sum_1^\infty \int_0^1 (-ax)^n \log x\,dx\ ,$$

since $|(-ax)^n \log x| \leq \frac{a^n}{ne}$ for $0 \leq x \leq 1$ and thus the series is uniformly convergent.

4.10 Theorem If $b = 0$, $a = \frac{1}{N}$ $(N = 1, 2, \ldots)$, then T_f is exact and has a finite, invariant measure equivalent to p.

Proof The admissible digits are $\frac{1}{a}, \frac{1}{a} + 1, \ldots$, i.e. zero is not an admissible digit. Thus

$$Q_n = ay_n\, Q_{n-1} + a\, Q_{n-2} > a\, Q_{n-1}$$

and so

$$\frac{\sup_{0<x<1} \omega^n(x,y)}{\inf_{0<x<1} \omega^n(x,y)} = \left(1 + a\,\frac{Q_{n-1}}{Q_n}\right)^2 < 4 :$$

3.4 holds, with $C = 4$. //

4.11 Theorem For $b = 0$, $a = \frac{1}{N}$ $(N = 1, 2, \ldots)$, the invariant measure μ is given by

$$\mu(E) = \int_E \frac{dx}{x + \frac{1}{a}}\ .$$

Proof

$$\sum_{n=N}^\infty \frac{1}{N + \frac{N}{x+n}} \cdot \frac{N}{(x+n)^2} = \frac{1}{x+N}\ . \quad //$$

$\mu(0, 1) = \log\left(\frac{N+1}{N}\right) < \infty$,

and $\quad \phi_n = \frac{\mu(X_n)}{\mu(0,1)} = \frac{\int_{\frac{N}{n+1}}^{\frac{N}{n}} \frac{dx}{N+x}}{\log\left(\frac{N+1}{N}\right)}$

$$= \frac{\log\left[\frac{(n+1)^2}{n(n+2)}\right]}{\log\left(\frac{N+1}{N}\right)} , \quad n = N, N+1, \ldots .$$

$h_\mu(T_f) < \infty$, since $\int_0^1 \log \left| \frac{d}{dx} f^{-1}(x) \right| dx = \int_0^1 \log\left(\frac{N}{x^2}\right) dx$

$$= \log N + 2 .$$

$$h_\mu(T_f) = \int_0^1 \log\left(\frac{N}{x^2}\right) \frac{dx}{N+x}$$

$$= (\log N)\left[\log\left(\frac{N+1}{N}\right)\right] - 2 \int_0^1 \frac{\log x}{N+x}\, dx$$

$$= (\log N)\left[\log\left(\frac{N+1}{N}\right)\right] - 2 \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2 N^n} .$$

4.12 <u>Proposition</u> $\quad \varepsilon_n(x) \notin L_1(\mu)$ for each n.

<u>Proof</u> $\quad \sum_{n=N}^{M} n \log\left[\frac{(n+1)^2}{n(n+2)}\right] = \log(M+2) + (M+1)\log\left(1 - \frac{1}{M+2}\right)$

$- N \log N$. //

4.13 <u>Corollary</u> $\quad \frac{1}{n} \sum_{1}^{n} \varepsilon_k(x) \longrightarrow \infty \ (n \longrightarrow \infty)$ almost everywhere.

<u>Proof</u> Apply the truncation argument of 4.7. //

4.14 Proposition $\log \varepsilon_n(x) \in L_1(\mu)$ for each n.

Proof. $\log n \log\left(\frac{(n+1)^2}{n(n+2)}\right) = \log n \log\left(1 + \frac{1}{n(n+2)}\right)$

$$< (n^{\frac{1}{2}} + \log 2)\frac{1}{n^2} .$$

4.14 follows since $\sum_{n=N}^{\infty} \frac{1}{n^p} < \infty$ if $p > 1$. //

4.15 Corollary $\sqrt{\varepsilon_1(x) \ldots \varepsilon_n(x)} \longrightarrow \left(\prod_{n=N}^{\infty} \left(1 + \frac{1}{n(n+2)}\right)^{\frac{\log n}{\log(1+\frac{1}{N})}}\right) (n \to \infty)$

almost everywhere.

Proof By 4.14 and the Individual Ergodic Theorem,

$$\frac{1}{n}\sum_{1}^{n} \log \varepsilon_k(x) \longrightarrow \frac{\int_I \log \varepsilon_1(x)\, d\mu}{\mu(I)} \quad (n \longrightarrow \infty) \text{ almost everywhere.}$$

The result follows, noting that for $a_n > 1$ and $\sum_1^{\infty} a_n < \infty$

$$\sum_1^{\infty} \log a_n = \log\left(\prod_1^{\infty} a_n\right) . \quad //$$

4.6 - 4.8 and 4.12 - 4.15 generalise corresponding results for $Tx = \left(\frac{1}{x}\right)$ [2 , p. 45].

§5 The transformation $T_f(x) = (\frac{bx}{1-cx})$, where $f(x) = \frac{x}{b+cx}$

5.1 Theorem. f-expansions are valid for $b \geqslant 1$ and $0 \leqslant c \leqslant 1$.

Proof $f'(x)$ must be positive : $b > 0$

$f^{-1}(1)$ must be in $(0, \infty]$: $c \leqslant 1$.

$c \geqslant 1 - b$, since if $c < 1 - b$, $\frac{b}{1-c} < 1$ and $f^{-1}[0,1) \subsetneq [0,1)$,

$\varepsilon_n(x) = 0$ for all $x \in [0,1)$ and all n, and f-expansions are not valid, by 3.1.

c must be non-negative, since otherwise negative $f_n(\varepsilon_1(x), \ldots, \varepsilon_n(x))$ can occur, e.g.

$b = 4,\ c = -1$: $c > 1 - b$ and $\frac{b}{1-c} = 2$, i.e. independent digits.

The admissible digits are 0, 1 .

$$f_2(0, 1) = -\frac{1}{13} .$$

$b = 4,\ c = -\frac{1}{2}$: $c > 1 - b$, $[\frac{b}{1-c}] = 2$, i.e. dependent digits.

Admissible digits are 0, 1 .

$$f_2(0, 1) = -\frac{2}{113}$$

and the sequence (0, 1) is canonical since $T_f X_o = I$.

Now suppose $b \geqslant 1$ and $c > 0$

$$f'(x) = \frac{b}{(b+cx)^2} < 1$$

$$\Leftrightarrow \quad g(x) = c^2 x^2 + 2bcx + b(b-1) > 0 .$$

Zeros of g are $\frac{\pm\sqrt{b} - b}{c}$ and if $b \geq 1$, $c > 0$,

$$\frac{\pm\sqrt{b} - b}{c} \leq 0 .$$

Hence $g(x) > 0$ for all $x \in [0, \frac{b}{1-c})$, i.e. by 3.3 f-expansions are valid.

Finally, f-expansions are not valid for $0 < b < 1$ and $0 < c \leq 1$: if $c < 1 - b$, f-expansions are not valid (see first part of this proof) while if $x \in X_0$ is an invariant point, then

$$x = \frac{bx}{1-cx} ,$$

i.e. $x = 0$ or $\frac{1-b}{c}$; but if $0 < b < 1$, $1 - b \leq c$ and $0 < c \leq 1$,

$$\frac{1}{b+c} - \frac{1-b}{c} = \frac{b(b+c)-b}{c(b+c)} = \frac{b(b+c-1)}{c(b+c)} \geq 0 .$$

Thus $\frac{1-b}{c} \in X_0$ and so $\varepsilon_n(y) = 0$ for all $y \in [0, \frac{1-b}{c})$ giving no valid f-expansion by 3.3 again.

For $c = 0$ we have $T_f(x) = (bx)$, which has been studied in [26],[29]. //

5.2 Theorem

$$\frac{P_n(y,t)}{Q_n(y,t)} = \frac{1}{c}\left[1 - \frac{b}{b+1+c\varepsilon_1(y) -} \; \frac{b}{b+1+c\varepsilon_2(y) -} \cdots - \frac{b}{b+c(\varepsilon_n(y)+t)}\right]$$

$$= \frac{b\,P_{n-1}(y,0) + c(\varepsilon_n(y)+t)(P_{n-1}(y,0)+\ldots+P_1(y,0)+1)}{b\,Q_{n-1}(y,0) + c(\varepsilon_n(y)+t)(Q_{n-1}(y,0)+\ldots+Q_1(y,0)+c)}, \; n \geq 2.$$

<u>Proof</u> Defining the partial quotients 'canonically' [17] , we have, using the notion of 4.2 ,

$$\frac{P_1}{Q_1} = \frac{cy_1}{c(1+by_1)}$$

$$\frac{P_2(t)}{Q_2(t)} = \frac{bcy_1 + c(y_2+t) + c^2 y_1(y_2+t)}{c[b^2 + bcy_1 + (b+1)c(y_2+t) + c^2 y_1(y_2+t)]}$$

$$= \frac{bP_1 + c(y_2+t)(P_1+1)}{bQ_1 + c(y_2+t)(Q_1+c)} .$$

Assuming relation true for n - 1 ,

$$\frac{P_n(t)}{Q_n(t)} = \frac{P_{n-1}\left(\frac{y_n+t}{b+c(y_n+t)}\right)}{Q_{n-1}\left(\frac{y_n+t}{b+c(y_n+t)}\right)}$$

$$= \frac{bP_{n-2} + c\left(y_{n-1} + \frac{y_n+t}{b+c(y_n+t)}\right)(P_{n-2}+\dots+1)}{bQ_{n-2} + c\left(y_{n-1} + \frac{y_n+t}{b+c(y_n+t)}\right)(Q_{n-2}+\dots+c)}$$

$$= \frac{bP_{n-1} + c(y_n+t)(P_{n-1}+\dots+1)}{bQ_{n-1} + c(y_n+t)(Q_{n-1}+\dots+c)} . \quad //$$

5.3 <u>Theorem</u> If $\frac{b}{1-c} = 1, 2, \dots, \infty$,

$b \geq 1$, $0 < c \leq 1$ implies T_f is ergodic ;

$b > 1$ $0 < c \leq 1$ implies T_f is exact and has a finite, invariant measure equivalent to p.

Proof
$$\omega^n(x,y) = \frac{c\,|Q_n(P_{n-1}+\ldots+1) - P_n(Q_{n-1}+\ldots+c)|}{[Q_n + cx(Q_{n-1}+\ldots+c)]^2}$$

$$p(C_n(y)) = \frac{c\,|Q_n(P_{n-1}+\ldots+1) - P_n(Q_{n-1}+\ldots+c)|}{Q_n[Q_n + c(Q_{n-1}+\ldots+c)]}$$

$$\frac{\inf_x \omega^n(x,y)}{p(C_n(y))} = \frac{Q_n}{Q_n + c(Q_{n-1}+\ldots+c)}$$

Now $Q_n = b\,Q_{n-1} + cy_n(Q_{n-1} + \ldots + c)$

$> cy_n(Q_{n-1} + \ldots + c)$

$\geqslant c(Q_{n-1} + \ldots + c)$ if $y_n \neq 0$.

Hence
$$\frac{Q_n}{Q_n + c(Q_{n-1} + \ldots + c)} \geqslant \frac{1}{2} \text{ if } y_n \neq 0 .$$

$Z = \mathscr{C}\{y : y_n \neq 0$ infinitely often$\} = \{y : y_n = 0$, n large enough$\}$

$= \bigcup_{n=1}^{\infty} \{y : y_i = 0 , \ i \geqslant n\}$.

Each $\{y : y_i = 0 , \ i \geqslant n\}$ is countable and hence null by the non-atomicity of p, i.e. $p(Z) = 0$. Hence

$$\overline{\lim_n} \frac{\inf_x \omega^n(x,y)}{p(C_n(y))} \geqslant \frac{1}{2}$$ for almost all y . The first part follows by 2.3.

If $y_n \neq 0$, then $\dfrac{c(Q_{n-1} + \dots + c)}{Q_n} \leqslant 1$,

while if $y_n = y_{n-1} = \dots = y_{n-r+1} = 0$, $y_{n-r} \neq 0$, then

$$\frac{c(Q_{n-1} + \dots + c)}{Q_n} \leqslant \frac{1}{b}\left(c + \frac{1}{b}\left(c + \dots + \frac{1}{b}(c+1)\dots\right)\right) = \sum_{i=1}^{r-1} \frac{c}{b^i} + \frac{c+1}{b^r} .$$

If $b > 1$, suppose firstly that $0 < c \leqslant b - 1$. Then

$$\frac{1}{b}(c+1) \geqslant \frac{1}{b}\left(c + \frac{1}{b}(c+1)\right) \geqslant \dots ,$$

i.e.

$$\frac{c(Q_{n-1} + \dots + c)}{Q_n} \leqslant \frac{c+1}{b} \leqslant 1 .$$

Thus 3.5 holds with $C = 4$.

If $b - 1 < c \leqslant 1$, then

$$\frac{1}{b}(c+1) < \frac{1}{b}\left(c + \frac{1}{b}(c+1)\right) < \dots .$$

$$\sum_{\nu=1}^{i} \frac{c}{b^\nu} + \frac{c+1}{b^{i+1}} - \left[\sum_{\nu=1}^{i-1} \frac{c}{b^\nu} + \frac{c+1}{b^i}\right] = \frac{c+1-b}{b^{i+1}} \qquad (i = 1, 2, \dots) .$$

Hence

$$\begin{aligned}
\sum_{\nu=1}^{r-1} \frac{c}{b^\nu} + \frac{c+1}{b^r} &= \frac{c+1}{b} + \sum_{i=2}^{r} \frac{c+1-b}{b^i} \\
&= \frac{c+1}{b} + \frac{(c+1-b)\left(1 - \frac{1}{b^{r-1}}\right)}{b(b-1)} \\
&\nearrow \frac{c+1}{b} + \frac{c+1-b}{b(b-1)} \qquad (r \to \infty) \\
&= \frac{c}{b-1} .
\end{aligned}$$

Thus 3.5 holds with $C = \left(1 + \frac{c}{b-1}\right)^2$.

Finally, we show 3.5 breaks down for b = 1 :

if $y_n = y_{n-1} = \dots = y_{n-r+1} = 0$ and $y_{n-r} \neq 0$, then

$$\frac{c(Q_{n-1}+\dots+c)}{Q_n} = rc + \frac{c(Q_{n-r-1}+\dots+c)}{Q_{n-r}}$$

$$> r\,c\ ;$$

but for all n and r $\left\{ y : \varepsilon_i(y) = \begin{array}{ll} 1 & 1 \leq i \leq n-r \\ 0 & n-r < i \leq n \end{array} \right\}$

is non-trivial, thereby contradicting 3.5. //

5.4 Theorem For c = 1, the invariant measure μ equivalent to p is given by

$$\mu(E) = \int_E \frac{b\,dx}{b+x-1} .$$

Proof $\sum_{n=0}^{\infty} \frac{b}{b-1+\frac{x+n}{(b+x+n)}} \cdot \frac{b}{(b+x+n)^2} = \frac{b}{b+x-1}$. //

$\mu[0, 1) = b \log \frac{b}{b-1}$, i.e. μ if finite for $b > 1$, σ-finite for $b = 1$.

5.5 For $b > 1$,

$$\phi_n = \int_{\frac{n}{b+n}}^{\frac{n+1}{b+n+1}} \frac{b}{b+x-1}\,dx \Big/ b \log \frac{b}{b-1}$$

$$= \frac{\log\left[\frac{(b+n)^2}{(b+n-1)(b+n+1)}\right]}{\log\left(\frac{b}{b-1}\right)} , \qquad n = 0, 1, \dots .$$

5.6 <u>Proposition</u> $\varepsilon_n(x) \notin L_1(\mu)$ for each n.

<u>Proof</u>

$$\sum_{n=1}^{N} n \log\left(\frac{(b+n)^2}{(b+n-1)(b+n+1)}\right) = \log(b+N+1) + (N+1)\log\left(1-\frac{1}{b+N+1}\right) - \log b . \quad //$$

5.7 <u>Corollary</u> $\frac{1}{n}\sum_{k=1}^{n} \varepsilon_k(x) \longrightarrow \infty \quad (n \longrightarrow \infty)$ almost everywhere.

<u>Proof</u> Apply the truncation argument of 4.7. //

5.8 <u>Corollary</u> The digits in f-expansions are almost everywhere unbounded.

5.6 - 5.8 generalise corresponding results for $Tx = (\frac{1}{x})$ [2 , p.45].

$h_\mu(T_{\frac{x}{b+x}}) < \infty$, since

$$\int_0^1 \log \frac{b}{(1-x)^2}\, dx = \log b - 2\int_0^1 \log(1-x)\, dx$$
$$= \log b + 2 .$$

$$h_\mu\left(T_{\frac{x}{b+x}}\right) = \int_0^1 \left[\log \frac{b}{(1-x)^2}\right] \frac{b}{b+x-1}\, dx$$
$$= [b\log b]\left[\log\left(\frac{b}{b-1}\right)\right] - 2b\int_0^1 \frac{\log x}{b-x}\, dx$$

$$= [b\log b]\left[\log\left(\frac{b}{b-1}\right)\right] + 2\sum_{n=1}^{\infty} \frac{1}{n^2 b^{n-1}} .$$

5.9 For $b = 1$, $T_{\frac{x}{1+x}}$ is conservative : $\chi_{\bigcup_1^\infty X_n}(x) \in L_1(\mu)$

and $\sum_0^\infty \chi_{\bigcup X_m}(T^n x) = \infty$ almost everywhere, since

$$\{x : \sum_0^\infty \chi_{\bigcup_1^\infty X_m}(T^n x) < \infty\} = \bigcup_{i=1}^\infty \{x : \varepsilon_n(x) = 0 , \ n \geq i\}$$

$$= \phi \ (\text{mod } 0) .$$

Let $1_\varepsilon(x) = \begin{cases} 1 , & \varepsilon \leq x < 1 \\ \frac{x}{\varepsilon} , & 0 \leq x < \varepsilon \end{cases}$,

where $0 < \varepsilon < \frac{1}{2}$. Then $1_\varepsilon(x) \in L_1(\mu)$. By the Hopf Ergodic Theorem, for each admissible digit $K \neq 0$,

$$\frac{1}{n} \sum_{i=0}^{n-1} \chi_{X_K}(T_f^i(x)) \leq \frac{\sum_{i=0}^{n-1} \chi_{X_K}(T_f^i(x))}{\sum_{i=0}^{n-1} 1_\varepsilon(T_f^i(x))}$$

$$\longrightarrow h_{k,\varepsilon}(x) \quad (n \longrightarrow \infty) \text{ almost everywhere.}$$

$h_{k,\varepsilon}(x)$ is invariant and hence constant by the ergodicity of T_f :

$$h_{k,\varepsilon}(x) = \frac{\mu(X_k)}{\int 1_\varepsilon(x)\,d\mu} = \frac{\mu(X_k)}{1-\log\varepsilon}$$

$$\to 0 \quad (\varepsilon\to 0) .$$

Hence

$$\overline{\lim}\ \frac{1}{n}\sum_{i=0}^{n-1} \chi_{X_k}(T_f^i(x)) \leq \lim_{\varepsilon\to 0} h_{k,\varepsilon} = 0 ,$$

i.e. $\phi_k = 0$ for $k \neq 0$. $\phi_0 = 1$ since

$$\frac{1}{n}\sum_{i=0}^{n-1} \chi_{X_0}(T_f^i(x)) = 1 - \frac{1}{n}\sum_{i=0}^{n-1} \chi_{\bigcup_{k=1}^{\infty} X_k}(T_f^i(x))$$

and

$$\overline{\lim_n}\ \frac{1}{n}\sum_{0}^{n-1} \chi_{\bigcup_{k=1}^{\infty} X_k}(T_f^i(x)) \leq \lim_n \frac{\sum_0^{n-1} \chi_{\bigcup_1^\infty X_k}(T_f^i(x))}{\sum_0^{n-1} 1_\varepsilon(T_f^i(x))}$$

$$= \frac{\mu(\bigcup_1^\infty X_k)}{\int 1_\varepsilon(x)\,d\mu}$$

$$\to 0 \quad (\varepsilon\to 0) .$$

Finally, for any non-negative integers $i_1, \dots, i_k$,

$$\phi_{i_1 \dots i_k} = \begin{cases} 0, & \text{if } i_r \neq 0 \text{ for some } r,\ 1 \leq r \leq k \\ 1, & \text{if } i_r = 0,\ 1 \leq r \leq k. \end{cases}$$

For by the Hopf Ergodic Theorem,

$$\overline{\lim_{n}} \frac{1}{n-k+1} \sum_{\nu=0}^{n-k} \chi_{X_{i_0} \cap \dots \cap T_f^{-k} X_{i_k}} (T_f^{\nu}(x))$$

$$\leq \frac{\mu(X_{i_0} \cap \dots \cap T_f^{-k} X_{i_k})}{\int 1_\varepsilon(x)\, d\mu}$$

$$\leq \frac{\mu(X_{i_r})}{\int 1_\varepsilon(x)\, d\mu}$$

$$\rightarrow 0 \quad (\varepsilon \rightarrow 0), \text{ if } i_r \neq 0,$$

while

$$\frac{1}{n-k+1} \sum_{0}^{n-k} \chi_{X_0 \cap \dots \cap T_f^{-k} X_0} (T_f^{\nu}(x)) = 1 - \frac{1}{n-k+1} \sum_{0}^{n-k} \chi_{\bigcup_1^\infty X_i \cap T_f^{-1} X_0 \cap \dots \cap T_f^{-k} X_0} (T_f^{\nu}(x))$$

$\longrightarrow 1 \ (n \longrightarrow \infty)$, since $\mu(\bigcup_1^\infty X_i) < \infty$.

§6 Unresolved Questions

6.1 The preceeding discussion makes the restrictive assumption that the f-expansion digits are independent. It would be of great interest to know what results hold for the dependent digit case.

A. Renyi [29] and V. A. Rohlin [32], after proving their theorems for indpendent digits, applied them to f-transformations with dependent digits by observing the behaviour of the particular f-expansions. It seems probable that similar results could be obtained with linear fractional mod one transformations.

6.2 It would be interesting to know whether 2.3, in addition to implying ergodicity, also implied the existence of an invariant measure equivalent to p. This measure need not be finite, as $T_{\frac{x}{1+x}}$ shows.

6.3 In the majority of cases considered it was not possible to compute the invariant measure, even when its existence was known by 3.5. The generalisation of the exhibited cases is by no means clear.

6.4 Is $T_{\frac{x}{1+x}}$ an infinite exact endomorphism and has it got finite entropy in the sense of Krengel [22] ? What value does its ergodic index take? $T_{\frac{x}{1+x}}$ does not satisfy the stronger condition of 2.3, which would imply that it has infinite ergodic index.

CHAPTER IV

Metric Invariants for M-shifts

§1 Introduction

In the isomorphism problem, invariants play an important part in the negative sense of exhibiting non-isomorphism. Indeed, it is generally much harder to prove two transformations isomorphic than non-isomorphic. For M-shifts this is just the case. Metric invariants for null-recurrent, irreducible M-shifts are introduced and studied in §3. They are based on a certain class of Kaluza sequence which was mentioned by J.F.C. Kingman [18]. They depend for their effectiveness on the criterion for ergodicity for null-recurrent M-shifts given by S. Kakutani and W. Parry [13]. The duality between M-shifts and renewal sequences, which is well-known, is studied in §2. Indeed, isomorphism of M-shifts is studied entirely in terms of their associated renewal sequences. As a consequence of the methods of §3, a generalised Hopf ergodic theorem is proved. In §4 three classes of M-shifts are studied using the invariants of §3. One of these classes was introduced by U. Krengel [22] and shown to consist of a continuum of non-isomorphic M-shifts using basically the Kakutani-Parry theorem.

§2 The Relation between M-shifts and Renewal Sequences.

In this chapter we shall only consider null-recurrent, irreducible M-shifts. If T is the M-shift $(\underline{\lambda}, P)$ then for any state i, $\{p_{ii}(n)\}_0^\infty$ is a renewal sequence. This follows on putting $f_n = f_{ii}(n)$, the probability of returning to state i for the first time after n steps, starting at state i.

Conversely, given a renewal sequence $\{p_n\}_0^\infty$, the following theorem gives an M-shift $(\underline{\lambda}, P)$ for which $p_{00}(n) = p_n$. We shall see in §3 that the choice of state 0 is immaterial for our purposes. The construction in 2.1 is well-known [3 , p. 40] ; attempts were made to construct other M-shifts having $\{p_n\}_0^\infty$ as renewal sequence in the above sense, but no results were obtained in this direction.

2.1 Theorem Given a null-recurrent renewal sequence $\{p_n\}_0^\infty$, let $\{f_n\}_1^\infty$ be the sequence in terms of which it is defined and put $F_n = \sum_{n+1}^{\infty} f_n$. Then the null-recurrent, irreducible M-shift $T = (\underline{\lambda}, P)$ with

$$\lambda_n = F_n$$

and for i = 0, 1, ...

$$p_{ij} = \begin{cases} f_{i+1} / F_i & , \ j = 0 \\ F_{i+1} / F_i & , \ j = i + 1 \end{cases}$$

has $\quad p_{oo}(n) = p_n$

and $\quad h(T) = - \sum_{1}^{\infty} f_n \log f_n$.

<u>Proof</u> T preserves the measure generated by $(\underline{\lambda}, P)$ since

$$\sum_{i=0}^{\infty} \lambda_i p_{ij} = \begin{cases} F_{j-1} F_j / F_{j-1} , & j > 0 \\ \sum_{i=0}^{\infty} f_{i+1} , & j = 0 \end{cases}$$

$$= \lambda_j \quad .$$

$$p_{oo}(n) = \sum_{i_1, \ldots, i_{n-1}} p_{oi_1} \cdots p_{i_{n-1}o} \quad .$$

This sum contains two types of terms, namely,

$$\prod_{\nu=1}^{s} p_{o1} \cdots p_{i_\nu o} = \prod_{\nu=1}^{s} F_1 \frac{F_2}{F_1} \cdots \frac{f_{i_\nu+1}}{F_{i_\nu}}$$

$$= \prod_{\nu=1}^{s} f_{i_\nu} \quad ,$$

where $1 \leq s \leq [\frac{n-1}{2}]$, $0 \leq i_\nu \leq n-1$, and

$$(p_{00})^k = f_1^k ,$$

where $0 \leq k \leq n$. Thus, $p_{00}(n) = \sum_{k=1}^{n} \sum_{i_1+\dots+i_k = n} f_{i_1} \cdots f_{i_k}$

$$= p_n , \quad n = 1, 2, \dots$$

$p_{00}(0) = 1 = p_0$, by definition.

$$h(T) = -\sum_{i=0}^{\infty} \lambda_i \sum_{j=0}^{\infty} p_{ij} \log p_{ij} \qquad [22]$$

$$= -\sum_{i=0}^{\infty} F_i \left\{ \frac{f_{i+1}}{F_i} \log\left(\frac{f_{i+1}}{F_i}\right) + \frac{F_{i+1}}{F_i} \log\left(\frac{F_{i+1}}{F_i}\right) \right\}$$

$$= -\sum_{i=0}^{\infty} \{ f_{i+1} \log f_{i+1} + F_{i+1} \log F_{i+1} - F_i \log F_i \}$$

$$= -\sum_{i=1}^{\infty} f_i \log f_i ,$$

since $F_i \longrightarrow 0$ $(i \longrightarrow \infty)$ and hence $F_i \log F_i \longrightarrow 0$ $(i \longrightarrow \infty)$. T is irreducible since any state can be reached from any other via state 0. That T preserves a σ-finite measure follows from I.65. It can also be verified directly that $(\underline{\lambda}, P)$ gives rise to a σ-finite measure noting that

$$\sum_{n=0}^{\infty} F_n = \sum_{n=1}^{\infty} n f_n \, . \quad //$$

Write $S_{\underline{p}}$ for the M-shift constructed in 2.1 to have renewal sequence $\underline{p}$. Although we shall only use the notation $S_{\underline{p}}$ in this sense, the results of §3 would also hold if $S_{\underline{p}}$ were any other irreducible M-shift having $\underline{p}$ as renewal sequence.

2.2 Theorem. For renewal sequences $\underline{p}$ and $\underline{q}$,

$$S_{\underline{p}} \times S_{\underline{q}} \text{ is isomorphic to } S_{\underline{pq}} ;$$

where $\underline{p}\,\underline{q} = \{p_n q_n\}_0^\infty$.

Proof. For simplicity of notation we assume without loss of generality that the state spaces of $S_{\underline{p}}$ and $S_{\underline{q}}$ are $\{0, 1, \ldots\}$. $S_{\underline{p}} \times S_{\underline{q}}$ is isomorphic to the shift on $\prod_{n=1}^{\infty} X_n \times Y_n$ where $X_n = Y_n = \{0, 1, \ldots\}$. If $P_{(i_1,j_1)(i_2,j_2)}$ are the transition probabilities for the latter shift, then (I.6.2) $P_{(i_1,j_1)(i_2,j_2)}(n) = p_{i_1 i_2}(n) \; q_{j_1 j_2}(n)$. Thus if $p_n = p_{ii}(n)$ and $q_n = q_{jj}(n)$, then $S_{\underline{p}} \times S_{\underline{q}}$ is isomorphic to $S_{\underline{p}\,\underline{q}}$. //

§3 Some Metric Invariants for M-shifts.

This section studies the isomorphism problem for null-recurrent, irreducible M-shifts $S_{\underline{p}}$ in terms of the divergence properties of $\underline{p}$.

For a null-recurrent renewal sequence $\underline{p}$, let $k(\underline{p})$ be the unique number such that

$$\sum_{n=0}^{\infty} (p_n)^{\iota} \begin{cases} =\infty & , \quad \iota < k(\underline{p}) \\ <\infty & , \quad \iota > k(\underline{p}) . \end{cases}$$

If $\sum_{n=0}^{\infty} (p_n)^{\iota} = \infty$ for all $\iota > 0$, as for example when $p_n = \dfrac{1}{\log(n+e)}$, put $k(\underline{p}) = \infty$. Otherwise $\sum_{n=0}^{\infty} (p_n)^{\iota'} < \infty$ for some $\iota' > 0$. By the comparison test, [20, p.111] $\sum_{n=0}^{\infty} (p_n)^{\iota} < \infty$ for all $\iota > \iota'$. Also, $\sum_{n=0}^{\infty} p_n = \infty$ and $\sum_{n=0}^{\infty} (p_n)^{\iota} = \infty$ for all $0 < \iota < 1$. Thus

$$L = \{ \iota : \sum_{0}^{\infty} (p_n)^{\iota} = \infty \} \neq \phi$$

and

$$R = \{ \iota : \sum_{0}^{\infty} (p_n)^{\iota} < \infty \} \neq \phi .$$

By the comparison test, every element of L is less than every element of R. Dedekind's theorem [10 ,p.30] now implies the uniqueness asserted above. $k(\underline{p})$ generalises the ergodic index, since the powers ι are not restricted to integer values.

If $\underline{p}^{\lambda}$ denotes $\{p_n^{\lambda}\}_0^{\infty}$ and $\underline{p}\ \underline{q}$ denotes $\{p_n\, q_n\}_0^{\infty}$, the following is true:-

3.1 <u>Theorem</u> (i) $\lambda\, k(\underline{p}^{\lambda}) = k(\underline{p})$

$$\text{(ii)} \quad \frac{1}{k(\underline{pq})} \geqslant \frac{1}{k(\underline{p})} + \frac{1}{k(\underline{q})} .$$

<u>Proof</u> (i)

$$\sum_0^{\infty} (p_n^{\lambda})^{\iota} < \infty \quad \text{if} \quad \iota > k(\underline{p}^{\lambda})$$

$$= \infty \quad \text{if} \quad \iota < k(\underline{p}^{\lambda})$$

$$\sum_0^{\infty} p_n^{\lambda\iota} < \infty \quad \text{if} \quad \iota > \frac{k(\underline{p})}{\lambda}$$

$$= \infty \quad \text{if} \quad \iota < \frac{k(\underline{p})}{\lambda} .$$

If $\lambda\, k(\underline{p}^{\lambda}) < k(\underline{p})$, say, then $\sum_0^{\infty} (p_n^{\lambda})^{\iota} < \infty$ while $\sum_0^{\infty} p_n^{\lambda\iota} = \infty$ for some ι. This is a contradiction since $(p_n^{\lambda})^{\iota} = p_n^{\lambda\iota}$.

(ii) Let $\lambda = \frac{1}{k(\underline{p})} + \frac{1}{k(\underline{q})}$, $P = \lambda\, k(\underline{p})$ and $Q = \lambda\, k(\underline{q})$.

Then $\frac{1}{P} + \frac{1}{Q} = 1$ and so by Hölder's inequality [19, p.186]

$$\sum_{n=0}^{\infty} p_n^{\iota} q_n^{\iota} \leq \left(\sum_{n=0}^{\infty} p_n^{\iota P}\right)^{1/P} \left(\sum_{n=0}^{\infty} q_n^{\iota Q}\right)^{1/Q} .$$

If $\iota P > k(\underline{p})$ and $\iota Q > k(\underline{q})$, i.e. if $\iota > \frac{1}{\lambda}$, then $\sum_0^{\infty} p_n^{\iota} q_n^{\iota} < \infty$.

Hence $k(\underline{pq}) \leq \frac{1}{\lambda}$. //

Define

$$\log^r x = \begin{cases} x & , \quad r = 0 \\ \log \log^{r-1} x, & r = 1, 2, \ldots \end{cases}$$

and

$$e_r = \begin{cases} 1 & , \quad r = 0 \\ e^{e_{r-1}} & , \quad r = 1, 2, \ldots \end{cases} .$$

3.2 <u>Theorem</u> $\left\{\frac{1}{\log^r(n+e_r)}\right\}_0^{\infty}$ is a Kaluza sequence, $r = 0, 1, \ldots$.

Proof For $r = 0$, we have to show $\frac{1}{(n+1)^2} \leq \frac{1}{n(n+2)}$, $n = 0, 1, \ldots$.
which is equivalent to $0 \leq 1$; while for $r > 0$,

$$\frac{d}{dx} \frac{\log^r(x+e_r)}{\log^r(x+1+e_r)} > \frac{\log^r(x+1+e_r) - \log^r(x+e_r)}{[\log^r(x+1+e_r)]^2[\log^{r-1}(x+1+e_r)]\ldots(x+1+e_r)}$$

$$> 0 ,$$

since $\log^r x$ is an increasing function of x. //

If $\underline{p}$ is a renewal sequence, let $\alpha_o(\underline{p})$ be the unique number such that

$$\sum_0^\infty \frac{p_n}{(n+1)^\alpha} \begin{cases} = \infty , & \alpha < \alpha_o(\underline{p}) \\ < \infty , & \alpha > \alpha_o(\underline{p}) \end{cases} ,$$

again with the convention that $\alpha_o(\underline{p}) = \infty$ if $\sum_0^\infty \frac{p_n}{(n+1)^\alpha} = \infty$ for all $\alpha > 0$. For $r > 0$, let $\alpha_r(\underline{p})$ be the unique number such that

$$\sum_0^\infty \frac{p_n}{(n+1)^{\alpha_o(\underline{p})}\ldots[\log^{r-1}(n+e_{r-1})]^{\alpha_{r-1}(\underline{p})}[\log^r(n+e_r)]^\alpha} \begin{cases} = \infty, & \alpha < \alpha_r(\underline{p}) \\ < \infty, & \alpha > \alpha_r(\underline{p}) . \end{cases}$$

3.3 Theorem If $\underline{p}$ is a Kaluza sequence,

$$\alpha_o(\underline{p}) = 1 - \frac{1}{k(\underline{p})} .$$

Proof Since $\underline{p}$ is monotone decreasing, by the Cauchy Condensation Test [20, p.120] $\sum_0^\infty (p_n)^t < \infty$ if, and only if, $\sum_0^\infty 2^n (p_{2^n})^t < \infty$. The latter series converges if $\overline{\lim_n} \sqrt[n]{2^n (p_{2^n})^t} < 1$, i.e. if $\lambda^t < \frac{1}{2}$, where

$\lambda = \overline{\lim_n} \sqrt[n]{p_{2^n}}$. It diverges if $\lambda^t > \frac{1}{2}$.

Similarly, $\sum_{0}^{\infty} \frac{p_n}{(n+1)^{\alpha}}$ converges if $\lambda^{\frac{1}{1-\alpha}} < \frac{1}{2}$, diverges if $\lambda^{\frac{1}{1-\alpha}} > \frac{1}{2}$. Hence $\lambda^{k(\underline{p})} = \frac{1}{2} = \lambda^{\frac{1}{1-\alpha_0(\underline{p})}}$. //

3.4 Corollary (i) $\alpha_0(\underline{p}^{\lambda}) = 1 - \lambda + \lambda\, \alpha_0(\underline{p})$

(ii) $1 + \alpha_0(\underline{p}\ \underline{q}) \leq \alpha_0(\underline{p}) + \alpha_0(\underline{q})$.

Proof Use 3.1 and 3.3 . //

3.5 Theorem. If $\underline{p}$ and $\underline{q}$ are renewal sequences and $\alpha_n(\underline{p}) \neq \alpha_n(\underline{q})$ for some n, then $S_{\underline{p}}$ is not isomorphic to $S_{\underline{q}}$.

Proof Taking the above n to be the least such n, let $\bar{\alpha} = \frac{1}{2}\{\alpha_n(\underline{p}) + \alpha_n(\underline{q})\}$ and suppose $\alpha_n(\underline{p}) < \alpha_n(\underline{q})$

Then $$\sum_{K=0}^{\infty} \frac{p_K}{(K+1)^{\alpha_0(\underline{p})} \cdots [\log^n(K+e_n)]^{\bar{\alpha}}} < \infty$$

while $$\sum_{K=0}^{\infty} \frac{}{(K+1)^{\alpha_0(\underline{q})} \cdots [\log^n(K+e_n)]^{\bar{\alpha}}} = \infty .$$

Thus [I.6.4] the irreducible N-shift $S_{\underline{p}} \times T$ is not ergodic, although $S_{\underline{q}} \times T$ being irreducible is ergodic, where T is the irreducible, aperiodic M-shift associated with the Kaluza sequence

$$\left\{ \frac{1}{(k+1)^{\alpha_0(p)} \ldots [\log^r(k+e_n)]^{\bar{\alpha}}} \right\}_{k=0}^{\infty} .$$ It follows that $S_{\underline{p}} \times T$ is not isomorphic to $S_{\underline{q}} \times T$ and hence that $S_{\underline{p}}$ is not isomorphic to $S_{\underline{q}}$. //

3.6 __Corollary__ If for Kaluza sequences $\underline{p}$ and $\underline{q}$, $k(\underline{p}) \neq k(\underline{q})$, then $S_{\underline{p}}$ is not isomorphic to $S_{\underline{q}}$.

__Proof__ By 3.3 , $\alpha_0(\underline{p}) = 1 - \frac{1}{k(\underline{p})}$. //

3.7 __Lemma__ For n = 0, 1, ... let $p_n > 0$, $q_n > 0$,

$0 < w_{n+1} \leq w_n$

and $\sum_0^{\infty} p_n = \sum_0^{\infty} q_n = 0$. Then

$$\frac{\sum_0^N p_n}{\sum_0^N q_n} \longrightarrow c < \infty \quad (N \longrightarrow \infty)$$ implies the existence and finiteness of $\breve{w}(\underline{p}, \underline{q}) = \lim_{N \to \infty} \frac{\sum_0^N w_n p_n}{\sum_0^N w_n q_n}$.

$w(\underline{p}, \underline{q}) = c$ if, and only if, $\sum_0^{\infty} w_n q_n = \infty$.

<u>Proof</u> We follow the proof of [9 , theorem 1 4]. For $N = 0, 1, \ldots$ let

$$s_N = \frac{\sum_0^N p_n}{\sum_0^N q_n} , \qquad t_N = \frac{\sum_0^N w_n p_n}{\sum_0^N w_n q_n} ,$$

$$P_N = \sum_0^N p_n , \quad Q_N = \sum_0^N q_n , \quad R_N = \sum_0^N w_n q_n .$$

Then $p_o = s_o Q_o$ and $p_N = s_N Q_N - s_{N-1} Q_{N-1}$ $(N > 0)$.

$$t_m = \frac{\sum_0^m w_n p_n}{R_m}$$

$$= \sum_0^\infty c_{m,n} s_n ,$$

where

$$c_{m,n} = \begin{cases} \dfrac{(w_n - w_{n+1}) Q_n}{R_m}, & n < m \\ \dfrac{w_m Q_m}{R_m} , & n = m \\ 0 , & n > m . \end{cases}$$

Putting $s_n = 1$ for all n we see that $p_n = Q_n - Q_{n-1} = q_n$ and $t_n = 1$ for all n. Hence $\sum_0^\infty c_{m,n} = 1$ for all m. Since $w_{n+1} \leq w_n$ for all n, $c_{m,n} > 0$ and so $\sum_0^\infty |c_{m,n}| = 1$ for all m. Finally, either $R_m \longrightarrow \infty$ $(m \longrightarrow \infty)$ or $R_m \longrightarrow R < \infty$ $(m \longrightarrow \infty)$. These correspond respectively to $c_{m,n} \longrightarrow 0$ and $c_{m,n} \longrightarrow \dfrac{(w_n - w_{n+1}) Q_n}{R}$ $(m \longrightarrow \infty)$.

The first part of the lemma now follows from [9 , theorem 1] and the second part, from [9 , theorem 2] . //

3.8 <u>Theorem</u> If $\underline{p}$ and $\underline{q}$ are renewal sequences such that

$$\frac{\sum_0^N p_n}{\sum_0^N q_n} \longrightarrow c \quad (N \longrightarrow \infty) , \qquad 0 < c < \infty ,$$

then $\alpha_r(\underline{p}) = \alpha_r(\underline{q})$, $r = 0, 1, \ldots$.

<u>Proof</u> Write

$$w_n(p, \alpha, r) = \frac{1}{(n+1)^{\alpha_0(p)} \cdots [\log^{r-1}(n+e_{r-1})]^{\alpha_{r-1}(p)} [\log^r(n+e_r)]^{\alpha}}$$

and similarly for $\underline{q}$.

We must show that $\sum_0^\infty p_n w_n(\underline{p}, \alpha, r) < \infty$ if, and only if ,

$$\sum_0^\infty q_n w_n(\underline{q}, \alpha, r) < \infty .$$

Suppose that $\sum_0^\infty p_n w_n(\underline{p}, \alpha, r) < \infty$ but that

$$\sum_0^\infty q_n w_n(\underline{q}, \alpha, r) = \infty .$$

This implies that

$$\frac{\sum_0^N p_n w_n(p, \alpha, r)}{\sum_0^N q_n w_n(q, \alpha, r)} \to 0 \quad (N \to \infty) ,$$

which contradicts 3.7. The converse assertion follows on interchanging $\underline{p}$ and $\underline{q}$. //

3.9 Corollary If $(\underline{\lambda}\ P)$ is a null-recurrent M-shift and $\underline{p}_i = \{\underline{p}_{ii}(n)\}_0^\infty$, then

$$\alpha_n(\underline{p}_i) = \alpha_n(\underline{p}_j) \quad n = 0, 1, \ldots$$

for all states i and j.

Proof

$$\frac{\sum_0^N p_{ii}(n)}{\sum_0^N p_{jj}(n)} \to \frac{\lambda_i}{\lambda_j} \quad (N \to \infty) \quad [30] .$$

//

We now prove a generalised Hopf Ergodic Theorem. If $\underline{w} = \{w_n\}_0^\infty$ where $w_n \geq w_{n+1} > 0$, T is a conservative, infinite measure-preserving transformation on $(X, \mathcal{B}, \mu)$ and $0 < f(x) \in L_1(\mu)$, put

$$C_{f,\underline{w}} = \{x : \sum_0^\infty w_n f(T^n x) = \infty\} .$$

3.10 Lemma If f, $g \in L_1(\mu)$ and $g(x) \geq 0$ almost everywhere, then

$$\lim_{n\to\infty} \frac{\sum_0^n w_k f(T^k x)}{\sum_0^n w_k g(T^k x)} = h_{f,g,\underline{w}}(x)$$

exists and is finite on $\{x : \sum_0^\infty g(T^n x) > 0\}$.

Proof By the Hopf Ergodic Theorem, $h_{f,g}(x) = \lim_{n\to\infty} \frac{\sum_0^n f(T^k x)}{\sum_0^n g(T^k x)}$ exists and is finite on $\{x : \sum_0^\infty g(T^n x) > 0\}$. There is no loss of generality in assuming that $f(x) \geq 0$ almost everywhere, since in the general case apply 3.10 to f^+, g and f^-, g respectively and note that $h_{f,g,\underline{w}}(x) = h_{f^+,g,\underline{w}}(x) - h_{f^-,g,\underline{w}}(x)$. 3.7 now gives the required result. //

3.11 Corollary $C_{f,\underline{w}}$ is invariant and independent of f.

Proof $T^{-1} C_{f,\underline{w}} = \{x : w_k f(T^{k+1} x) = \infty\}$. Since $w_{n+1} \leq w_n$, $\sum_0^n w_k f(T^k x) \leq \sum_0^n w_k f(T^{k+1} x)$. Conversely, by 3.10

$$\frac{\sum_0^n w_k f(T^{k+1} x)}{\sum_0^n w_k f(T^k x)} \to c < \infty \quad (n \to \infty) .$$

Hence $\sum_0^\infty w_k f(T^k x) = \infty$ if, and only if, $\sum_0^\infty w_k f(T^{k+1} x) = \infty$, i.e. $C_{f,\underline{w}}$ is invariant.

Again by 3.10, $h_{f,g}(x) < \infty$ and $h_{g,f}(x) < \infty$ for $0 < f, g \in L_1(\mu)$. Hence $C_{f,\underline{w}}$ is independent of f. //

Write $C_{\underline{w}} = C_{f,\underline{w}}$ if there exists $0 < f \in L_1(\mu)$ such that $C_{f,\underline{w}} \neq \phi$. U. Krengel [23], working with more general positive contractions T on $L_1(\mu)$, calls $C_{\underline{w}}$ the $\underline{w}$-conservative part of T.

3.12 Theorem Under the conditions of 3.10, $h_{f,g,\underline{W}}(x)$ is invariant on $C_{\underline{W}}$ and

$$\int_{C_{\underline{W}}} g(x)\, h_{f,g,\underline{W}}(x)\, d\mu = \int_{C_{\underline{W}}} f(x)\, d\mu .$$

Proof. Apply the Hopf Ergodic Theorem to $\chi_{C_{\underline{W}}}(x) f(x)$ and $g(x)$, noting that by the invariance of $C_{\underline{W}}$,

$$h_{\chi_{C_{\underline{W}}} f, g}(x) = \chi_{C_{\underline{W}}}(x)\, h_{f,g}(x) . \quad //$$

§4 Non-Isomorphism of Certain M-Shifts

Put
$$u_n(\alpha,r) = \frac{1}{[\log^r(n+e_r)]^\alpha} , \quad \alpha > 0 , \quad r = 0, 1, \ldots$$

$$v_n(\alpha) = \frac{n!}{(1+\alpha)\ldots(n+\alpha)} , \quad \alpha > 0 ,$$

$$w_n(p) = \frac{\Gamma(n+p)}{\Gamma(p)\Gamma(n+1)} , \quad 0 < p < 1 .$$

4.1 Theorem $\underline{u}(\alpha,r)$ is a null-recurrent Kaluza sequence for $0 < \alpha \leq 1$ $(r = 0)$ and $0 < \alpha$ $(r > 0)$

$$\alpha_n(\underline{u}(\alpha,r)) = \begin{cases} 1-\alpha & , \ n = r \\ 1 & , \ n \neq r . \end{cases}$$

Proof 3.2 states that $\underline{u}(\alpha, r)$ is a Kaluza sequence.

$u_n(\alpha,0) = \dfrac{1}{(n+1)^\alpha}$ is a null-recurrent if, and only if, $\alpha \leq 1$.

[20, p. 120]. $\underline{u}(\alpha, r)$ is null-recurrent for all $\alpha > 0$ by the comparison test, since $\dfrac{n+1}{[\log^r(n+e_r)]^\alpha} \longrightarrow \infty (n \longrightarrow \infty)$ for all

$\alpha > 0$ and $r > 0$. The last assertion follows from the rates of convergence of the logarithmic scale [20, p. 123]. //

Note that when $r > 0$, $\alpha_n(\underline{u}(\alpha,r))$ can take negative values.

4.2 Corollary $S_{\underline{u}(\alpha,r)}$ form a continuum of non-isomorphic M-shifts.

Proof Apply 3.5. //

4.3 Corollary $S_{\prod_{\nu=1}^{k} \underline{u}(\alpha_\nu, r_\nu)}$ from a further continuum of non-isomorphic M-shifts, where $\prod_{\nu=1}^{k} \underline{u}(\alpha_\nu, r_\nu) = \left\{ \prod_{\nu=1}^{k} u_n(\alpha_\nu, r_\nu) \right\}_0^\infty$

Proof $$\alpha_n\left(\prod_{\nu=1}^{k} \underline{u}(\alpha_\nu, r_\nu)\right) = \begin{cases} 1 & , \quad n \neq r_\nu , \quad 1 \leq \nu \leq k \\ 1-\alpha_\nu & , \quad n = r_\nu . \end{cases}$$ //

4.4 Theorem $\underline{v}(\alpha)$ is a null-recurrent Kaluza sequence for $0 < \alpha \leq 1$.

$$\alpha_n(\underline{v}(\alpha)) = \begin{cases} 1 - \alpha & , \quad n = 0 \\ 1 & , \quad n > 0 . \end{cases}$$

<u>Proof</u> $v_n^2(\alpha) \leq v_{n-1}(\alpha)\, v_{n+1}(\alpha)$ if, and only if,

$$\frac{n}{n+\alpha} \leq \frac{n+1}{n+\alpha+1} \quad \text{if, and only if,}$$

$$\frac{d}{dx}\ \frac{x+1}{x+\alpha+1} = \frac{\alpha}{(x+\alpha+1)^2} > 0 \ .$$

$\underline{v}(\alpha)$ is recurrent if, and only if, $0 < \alpha \leq 1$ by the Gauss test [20, p.288], since

$$\frac{v_{n+1}(\alpha)}{v_n(\alpha)} = \left(1 + \frac{\alpha}{n+1}\right)^{-1}$$

$$= 1 - \frac{\alpha}{n} + 0\left(\frac{1}{n^2}\right) \ .$$

$\underline{v}(\alpha)$ is either transient or null-recurrent since the convergence of the infinite product $\prod_{n=0}^{\infty} \left(1 - \frac{\alpha}{n+\alpha+1}\right) = \lim_{n \to \infty} v_n(\alpha)$ is equivalent to that of $\sum_{n=0}^{\infty} \frac{\alpha}{n+\alpha+1}$; but $\frac{\alpha}{n+\alpha+1} \sim \frac{\alpha}{n}$, i.e. $\sum_{n=0}^{\infty} \frac{\alpha}{n+\alpha+1} = \infty$. Since $v_n(\alpha) \searrow$ $(n \longrightarrow \infty)$, $\underline{v}(\alpha)$ being a Kaluza sequence, $\lim_{n \to \infty} v_n(\alpha)$ exists. If it were positive, the infinite product would converge; thus $v_n(\alpha) \searrow 0$ $(n \longrightarrow \infty)$.

$$\left(\frac{v_{n+1}(\alpha)}{v_n(\alpha)}\right)^{\iota} = \left(1 + \frac{\alpha}{n+1}\right)^{-\iota}$$

$$= 1 - \frac{\iota\alpha}{n} + 0\left(\frac{1}{n^2}\right) \ .$$

Hence, again by Gauss' test, $k(\underline{v}(\alpha)) = \frac{1}{\alpha}$. It follows by 3.3 that $\alpha_o(\underline{v}(\alpha)) = 1 - \alpha$.

That $\alpha_n(\underline{v}(\alpha)) = 1$, $n > 0$, follows from $v_n(\alpha) \sim \Gamma(\alpha+1)u_n(\alpha,0)$ [33 , p. 11]. Indeed, 4.4 follows from 4.1 using this result. A different proof of the first part of 4.4 is given since it seems to have interest of its own. //

4.5 Corollary $S_{\underline{v}(\alpha)}$ is not isomorphic to $S_{\underline{u}(\beta,r)}$ for all $0 < \alpha , \beta < 1$ when $r > 0$ and for all $0 < \alpha \neq \beta < 1$ when $r = 0$.

Proof When $r = 0$, for $\alpha \neq \beta$ $\alpha_o(\underline{v}(\alpha)) \neq \alpha_o(\underline{u}(\beta,r))$. When $r > 0$ $\alpha_r(\underline{v}(\alpha)) \neq \alpha_r(\underline{u}(\beta,r))$ for all $0 < \alpha , \beta < 1$. //

We have no information when $r = 0$ and $\alpha = \beta$, since then $\alpha_n(\underline{v}(\alpha)) = \alpha_n(\underline{u}(\beta,r))$ for all n. This is to be expected, since $\underline{v}(\alpha)$ and $\underline{u}(\alpha,0)$ are esentially the same renewal sequence (in terms of convergence properties).

4.6 Theorem $\underline{w}(p)$ is a null-recurrent Kaluza sequence for $0 < p < 1$.

$$\alpha_n(\underline{w}(p)) = \begin{cases} p , & n = 0 \\ 1 , & n > 0 . \end{cases}$$

Proof $w_n(p) \sim \frac{1}{\Gamma(p)\, n^{1-p}}$ [34 , p. 58] . Since $\underline{u}(1 - p, 0)$ is null-recurrent, so is $\underline{w}(p)$. Also $\alpha_n(\underline{w}(p)) = \alpha_n(\underline{u}(1 - p,0))$. As $\underline{u}(1 - p,0)$ is a Kaluza sequence, the asymptotic relation implies that $\underline{w}(p)$ is Kaluza too. Alternatively,

$$\frac{w_{n+1}(p)}{w_n(p)} = \frac{n+p}{n+1} \nearrow 1 \ (n \longrightarrow \infty) . \quad //$$

4.7 Theorem $S_{\underline{w}(p)}$ is not isomorphic to $S_{\underline{u}(\alpha,r)}$ for all $0 < \alpha$, $p < 1$ and $r = 0, 1, \ldots$.

Proof $\alpha_o(\underline{w}(p)) \neq \alpha_o(\underline{u}(\alpha,r))$. //

Attempts were made to compute the entropy of the three classes of M-shifts studied above, the problem being to find an exact or asymptotic expression for f_n . For $\underline{u}(\alpha,r)$ and $\underline{v}(\alpha)$ there appear to be no useful closed form expressions for the generating functions of the former or latter. However, we did obtain the following

4.8 Theorem $\sum_{n=o}^{\infty} u_n(\alpha,0)\, z^n = \Phi(z, \alpha, 1)$,

where $\Phi(z, \alpha, \nu) = \sum_{n=o}^{\infty} \frac{z^n}{(n+\nu)^\alpha}$ [4 , pp. 27-31] .

Proof This identity is none other than the definition of Φ . //

4.9 Corollary $F_{\underline{u}(\alpha,0)}(z) = z \frac{\Phi(z,\alpha,2)}{\Phi(z,\alpha,1)}$.

Proof Follows from the definitions of F and Φ . //

4.10 <u>Theorem</u> $\sum_{n=0}^{\infty} v_n(\alpha)\, z^n = H(1,\ 1;\ \alpha + 1\ ;\ z)$,

where H denotes the hypergeometric function [33, p. 19]

<u>Proof</u> $H(a,\ b;\ c\ ;\ z) = \sum_{n=0}^{\infty} \frac{\Gamma(a+n)\Gamma(b+n)\Gamma(c)}{\Gamma(a)\Gamma(b)\Gamma(c+n)\Gamma(n+1)}\, z^n$. //

4.11 <u>Corollary</u> $F_{\underline{v}(\alpha)}(z) = \frac{H(1,2;\ \alpha + 2;\ z)}{H(1,1;\ \alpha = 1;\ z)}$.

Since $\underline{w}(p)$ is generated by a closed form expression, namely $(1 - z)^{-p}$, $F_{\underline{w}(p)}(z) = 1 - (1 - z)^p$ and

$$f_n = -\frac{\Gamma(n - p)}{\Gamma(-p)\Gamma(n+1)}$$

$$\sim \frac{-1}{\Gamma(-p)\, n^{p+1}} \quad .$$ Hence

4.12 <u>Theorem</u> [22] $h(S_{\underline{w}(p)}) < \infty$ for all $0 < p < 1$.

<u>Proof</u> $\sum_{n=2}^{\infty} \frac{1}{n^{p+1}} \log \frac{1}{n^{p+1}} < \infty$ if, and only if,

$$\sum_{n=2}^{\infty} \frac{\log n}{n^{p+1}} < \infty$$

But

$$\left(\frac{n}{n+1}\right)^{p+1} \frac{\log(n+1)}{\log n} = \left(1 + \frac{1}{n}\right)^{-(p+1)} \left\{1 + \frac{\log\left(1 + \frac{1}{n}\right)}{\log n}\right\}$$

$$= 1 - \frac{(p+1)}{n} + O\left(\frac{1}{n^2}\right)$$

and so by Gauss' test the latter series converges. The proof is completed by noting that series of asymptotically equal terms converge or diverge together. //

Although the renewal sequences studied in this section throw some light on the effectiveness of $\alpha_n(\underline{p})$ as metric invariants, they are too closely connected with the test sequences $\underline{u}(\alpha, r)$ to indicate whether, for example $\{\alpha_n(\underline{p})\}_0^\infty$ might be a complete invarient for $S_{\underline{p}}$ where $\underline{p}$ is Kaluza. The answer to this question is clearly connected with the universality or otherwise of the logarithmic scale of ratio tests. However, as K. Knopp [20, p. 304] points out, no "boundary" exists such that all monotonic series on one side of it converge, while those on the other side all diverge, irrespective of the manner of definition of the boundary.

B I B L I O G R A P H Y

1. Apostol, T.M., Mathematical Analysis, Addison-Wesley Publishing Co.Inc., 1957.

2. Billingsley, P., Ergodic Theory and Information, John Wiley and Sons, Inc., 1965.

3. Chung, K.L., Markov Chains with Stationary Transition Probabilities, Springer Verlag, 2nd edition, 1967.

4. Erdélyi, A., Editor, Higher Transcendental Functions, volume I, McGraw-Hill Book Co.Inc., 1953.

5. Feller, W., An Introduction to Probability Theory and its Applications, volume I, John Wiley and Sons,Inc., 2nd edition, 1965.

6. Gradšteĭn, I.S., and Ryžik, I.M., Tables of Integrals, Summation, Series and Products, Moscow State Publishing House, 1963 (in Russian).

7. Halmos, P.R., Measure Theory, van Nostrand Co.Inc., 1950.

8. Halmos, P.R., Lectures on Ergodic Theory, Chelsea Publishing Co., 1956.

9. Hardy, G.H., Divergent Series, Oxford University Press, 1956.

10. Hardy, G.H., Pure Mathematics, Cambridge University Press, 10th edition, 1952.

11. Hopf, E., Ergodentheorie, Chelsea Publishing Co., 1948.

12. Jacobs, K., Lecture Notes on Ergodic Theory, Matematisk Institut, Aarhus Universitet, 1962/63.

13. Kakutani, S., and Parry, W., Infinite Measure Preserving Transformations with "Mixing", Bull. Amer. Math. Soc. 69, 752-756 (1963).

14. Kaluza, T., Über die Koeffizienten reziproker Potenzreihen, Math. Zeitschrift, 28, 161-170 (1928).

15. Kelley, J.L., General Topology, van Nostrand Co.Inc., 1955.

16. Kendall, D.G., Renewal Sequences and their Arithmetic, Lecture Notes in Mathematics no. 31, 147-175, Springer, 1967.

17. Khinchin, A.Ya., Continued Fractions, Chicago University Press, 1964.

18. Kingman, J.F.C., An Approach to the Study of Markov Processes, Journal Roy. Stat. Soc. (B), 28, 417-447, (1966).

19. Kingman, J.F.C., and Taylor, S.J., Introduction to Measure and Probability, Cambridge University Press, 1966.

20. Knopp, K., Theory and Application of Infinite Series, Blackie and Son ltd., 1951.

21. Kolmogorov, A.N., and Fomin, S.V., Measure, Lebesgue Integrals, and Hilbert Space, Academic Press, 1962.

22. Krengel, U., Entropy of Conservative Transformations, Z. Wahrscheinlichkeitstheorie verw. Geb., 7, 161-181 (1967).

23. Krengel, U., Classification of States for Operators, Proc. Fifth Berkeley Symposium Math. Statist. Probability, 1966.

24. Lamperti, J., On the Coefficients of Reciprocal Power Series, Amer. Math. Monthly 65, 90-94 (1958).

25. Oishi, N., Notes on Ergodicity and Mixing Property, Proc. Jap. Acad., 41, 767-770 (1965).

26. Parry, W., On the β-expansions of Real Numbers, Acta Math. Acad. Sci. Hung. 11, 401-416 (1960).

27. Parry, W., Representations for Real Numbers, Acta Math. Acad. Sci. Hung., 15, 95-105 (1964).

28. Parry, W., Ergodic and Spectral Analysis of Certain Infinite Measure Preserving Transformations, Proc. Amer. Math. Soc. 16, 960-966 (1965).

29. Rényi, A., Representations for Real Numbers and their Ergodic Properties, Acta Math. Acad. Sci. Hung., 8, 477-493 (1957).

30. Reuter, G.E.H., Markov Chains, mimeographed lecture notes, Durham Probability Conference, 1963.

31. Rohlin, V.A., On the Fundamental Ideas of Measure Theory, Mat. Sbornik (N.S.) 25 (67), 107-150 (1949) [English translation in Amer. Math. Soc. Transl. no. 71 (1952)].

32. Rohlin, V.A., Exact Endomorphisms of a Lebesgue Space, Izv. Akad. Nauk SSSR 25, 499-530 (1961) [English translation in Amer. Math. Soc. Transl. (2) 39, 1-36 (1964)].

33. Sneddon, I.N., Special Functions of Mathematical Physics and Chemistry, Oliver and Boyd, 1961.

34. Titchmarsh, E.C., The Theory of Functions, Oxford University Press, 2nd edition, 1952.

INDEX OF NOTATION AND DEFINITIONS

Numbers refer to the page(s) where the notation and terms are defined.