

ERGODIC AND MIXING TRANSFORMATIONS

by

WILLIAM PARRY

Thesis submitted for the degree of Ph.D.
in the University of London, July 1960.

ABSTRACT

A space X of sequences $\dots x_{-1} x_0 x_1 \dots$ of "letters" chosen randomly from a finite "alphabet" together with a probability derived from a probability function on events $\{x: x = \dots x_{-1} x_0 x_1 \dots \in X; x_i \in \mathcal{A}\}$ constitutes the type of measure space studied in information theory.

This thesis investigates such measure spaces when sequences with the same "history" are identified, when the letters are not necessarily chosen randomly and when the alphabet is at most countably infinite.

Each space of this type (i.e. each partition system) is shown in Ch.2 to be a thick subset of a partition system isomorphic with the unit interval together with a transformation T ; where $Tx = (f^{-1}x)$ and f can be chosen either as a strictly increasing continuous function in its domain with range $[0, 1)$ or as a strictly decreasing continuous function in its domain with range $[0, 1)$. Moreover, if $x = x_0 x_1 \dots$ then the isomorph of x is $\bar{f}(x_0 + \bar{f}(x_1 + \dots))$ where $\bar{f}$ is the monotonic extension of f . These expansions include the f -expansions studied by A. Renyi and this latter class, examples of which are continued fractions and "decimals", are characterised in terms of the behaviour of sequences under the shift transformation. These expansions are "valid" if and only if $Tx = (f^{-1}x)$ is topologically transitive, (but not necessarily ergodic).

In Ch.3 four examples of f -expansions are considered. In each case T is ergodic, in fact weakly mixing, and the invariant measures are yielded by the solutions of a certain

functional equation. (Most of this work is well known)
 The frequency of digits in β -expansions can therefore be calculated.

Finally in ch.4 and ch.5 β -expansions (or generalised decimals) are studied more closely. Those β which have a recurrent tails in their β -expansions yield invariant measures of a particularly simple character. Such

β -transformations can be characterised by a matrix and the invariant measures equivalent to Lebesgue measure are integrals of step functions , obtained from eigen vectors of this matrix with value β .

ACKNOWLEDGEMENTS

I take this opportunity to thank Professor W.K. Hayman, Dr. T. Kovari and especially Mrs Y.N. Dowker, my supervisor, for their invaluable help mathematical and otherwise.

Acknowledgements are also due to the Department of Scientific and Industrial Research, and to the University of London (Arthur Jubber Studentship) , for the studentships awarded to me by these bodies during the course of my studies at Imperial College.

CONTENTS

ABSTRACT	3
ACKNOWLEDGEMENTS	5
<u>CHAPTER 0</u> <u>PRELIMINARIES</u>	8
A. MEASURE THEORY	8
B. ERGODIC THEORY	10
C. HILBERT AND BANACH SPACES	10
<u>CHAPTER 1.</u> <u>INTRODUCTION</u>	13
A. ABSOLUTE CONTINUITY AND SINGULARITY	13
B. MEASURABLE TRANSFORMATIONS	14
C. HILBERT AND BANACH SPACES	19
D. ERGODIC THEOREMS	24
<u>CHAPTER 2.</u> <u>ISOMORPHISMS AND REPRESENTATIONS</u>	28
A. ISOMORPHISMS	29
B. PARTITIONS	35
C. PARTITION SYSTEMS	41
D. REPRESENTATIONS	44
E. REPRESENTATION OF PARTITION SYSTEMS	49
F. FULL PARTITION SYSTEMS	53
G. TOPOLOGICAL TRANSITIVITY	56

<u>CHAPTER 3.</u>	<u>FREQUENCY, ERGODICITY, AND INVARIANCE</u>	67
A.	FREQUENCY	67
B.	ERGODICITY	69
C.	INVARIANCE	74
<u>CHAPTER 4.</u>	<u>β -TRANSFORMATIONS</u>	86
A.	β -NUMBERS AND THEIR CHARACTERISTIC EQUATIONS	86
B.	DISTRIBUTION OF β -NUMBERS	96
C.	THE NORMALISING FUNCTION	97
<u>CHAPTER 5.</u>	<u>β -NUMBERS AND MATRICES</u>	105
REFERENCES		109

CHAPTER 0PRELIMINARIESA.MEASURE THEORY

The elements of Measure Theory assumed in this thesis are to be found in [1].

(X, α, μ) is a measure space means :

- i) α is a σ -algebra of subsets of X i.e. α is closed under differences and countable unions and X is an element of α .
- ii) μ is a measure function i.e. μ is a non-negative function with domain α and $\mu \emptyset = 0$ and if $\{A_n\} \subset \alpha$ ($A_m \cap A_n = \emptyset$ when $m \neq n$) then $\mu \bigcup_n A_n = \sum_n \mu A_n$.
- iii) μ is σ -finite i.e. there exists $\{A_n\} \subset \alpha$ with $\mu A_n < \infty$ and $\bigcup_n A_n = X$.

(X, α, μ) is said to be a (finite) normalised measure space and μ a (finite) normalised measure function when $(\mu X < \infty)$ $\mu X = 1$. Evidently a finite measure space is normalised when the values of the measure function are divided by μX . A normalised measure (space) function is sometimes called a probability (space) function.

If (X, α, μ) is a measure space and $A \in \alpha$, A is said to be an atom when $\alpha \ni B \subset A$ implies either $\mu B = 0$ or $\mu(A-B) = 0$. An algebra with (without) atoms is said to be atomic (non-atomic). A measure (function)

space whose algebra (domain) is atomic is said to be atomic.

If (X, α, m) is a measure space there is an associated metric space (α, d) where $d(E, F) = m(E - F \cup F - E)$ ($E, F \in \alpha$) and (X, α, m) is said to be seperable if the associated metric space (α, d) is seperable.

If all subsets of sets of measure zero are measurable (i.e. belong to α) the measure space, measure function and algebra are said to be complete. If (X, α, m) is a measure space and α' is the set of all $(E - N \cup N - E)$ where $E \in \alpha$, $N \subset M \in \alpha$, $m M = 0$ and $m'(E - N \cup N - E) = m E$, then (X, α', m') is a complete measure space - the completion of (X, α, m) .

A real valued function f with domain X is said to be Borel (Lebesgue) measurable if $f^{-1}E \in \alpha$ for all Borel E (Lebesgue E) on the real line. A transformation T of X onto itself is said to be measurable if $TE, T^{-1}E \in \alpha$ for all $E \in \alpha$.

If μ, ν are two measure functions with the same domain μ is said to be absolutely continuous with respect to ν ($\mu \ll \nu$) if $\mu E = 0$ whenever $\nu E = 0$. μ, ν are said to be equivalent ^($\mu \sim \nu$) if $\mu \ll \nu$ and $\nu \ll \mu$.

μ, ν are said to be mutually singular ($\mu \perp \nu$) if there is a set $A \in \alpha$ with $\mu A = \nu A^c = 0$. ($A^c = X - A$ is the complement of A).

B.

ERGODIC THEORY

A measurable transformation T of (X, α, μ) onto itself is said to be non-singular if $\mu T A = 0$ and $\mu T^{-1} A = 0$ when $\mu A = 0$.

If T is a non-singular measurable transformation of (X, α, μ) onto itself a set $A \in \alpha$ is said to be a wandering set of positive measure if $\mu A > 0$ and $T^n A \cap T^m A = \emptyset$ for non-negative $n, m, n \neq m$. Evidently if A is a wandering set of positive measure $T^n A \cap T^m A = \emptyset$ for all integers $n, m, n \neq m$, for it suffices to note that if $n \geq 0$ and $k > 0$,

$$\begin{aligned} T^{-n} A \cap T^{-n+k} A &\subset T^{-n} A \cap T^{-n} A (T^k A) \\ &= T^{-n} (A \cap T^k A) \\ &= \emptyset. \end{aligned}$$

A measure function μ is invariant when $\mu E = \mu T E$ for all measurable E . A measurable transformation is ergodic means that invariant sets (i.e. sets A such that $A = T^{-1} A$) or their complements have measure zero. The orbit of a set E under T is the smallest invariant set $O(E)$ containing E . $(O(E) = \bigcup_{n \geq 0} T^{-n} (\bigcup_{m \geq 0} T^m E))$ If T is non-singular $O(E)$ has measure zero when E has measure zero. The orbit of a point x is the orbit of the one point set $\{x\}$ and is a countable set. Moreover, $O(x) = O(y)$ if $y \in O(x)$ and $O(x) \cap O(y) = \emptyset$ if $y \notin O(x)$.

C.

HILBERT AND BANACH SPACES

The elements of Hilbert and Banach Space Theory are to be found in [27].

A Hilbert space $\mathcal{H}$ is a set of elements satisfying :

- a.) $\mathcal{H}$ is a linear space i.e. multiplication of elements by complex scalars and addition of elements are defined and $\mathcal{H}$ is closed under these operations.
- b.) $\mathcal{H}$ is a metric space whose metric is derived from a scalar product i.e. to every pair $f, g \in \mathcal{H}$ there is associated a number (f, g) -- the scalar product of f and g , and (f, g) satisfies :
- i) $(af, g) = a(f, g)$ for every scalar a .
 - ii) $(f+g, h) = (f, h) + (g, h)$
 - iii) $(f, g) = \overline{(g, f)}$
 - iv) $(f, f) \geq 0$ for $f \neq 0$ and $(f, f) = 0$ for $f = 0$.
- The norm of f is defined as $\|f\| = (f, f)^{\frac{1}{2}}$ and the distance between f, g is defined as $\|f - g\|$.
- c.) $\mathcal{H}$ is complete in the sense that every Cauchy sequence converges to an element in $\mathcal{H}$.

A Banach space $\mathcal{B}$ is a set of elements satisfying conditions a.) and c.) of the definition of Hilbert space and instead of b.) satisfies :

- b!) There is associated with every element f of $\mathcal{B}$ a non-negative number $\|f\|$ -- the norm of f , such that $\|f+g\| \leq \|f\| + \|g\|$, $\|af\| = |a|\|f\|$ (a , scalar) and $\|f\| = 0$ if and only if $f = 0$.

A linear transformation T of a Banach space $\mathcal{B}$ into itself obeys $T(\alpha f + \beta g) = \alpha Tf + \beta Tg$ (α, β scalar, $f, g \in \mathcal{B}$) and is said to be bounded if $\|T\| < \infty$ where $\|T\| = \sup_{f \in \mathcal{B}} \frac{\|Tf\|}{\|f\|}$. $\|T\|$ is the norm of T .

If T is a bounded linear transformation of a Hilbert space H there is a unique bounded linear transformation T^* such that $(Tp, q) = (p, T^*q)$ for every $p, q \in H$, which is called the adjoint of T . Moreover $\|T^*\| = \|T\|$.

If T is a bounded linear transformation of a Banach space B and $Tf = \mu f$ where $0 \neq f \in B$ and μ is a complex scalar, f is said to be an eigen vector of T with eigen value μ .

The resolvent set of T is the set of all complex scalars μ for which $(T - \mu I)^{-1}$ exists and is bounded. The complement of the resolvent set in the complex plane is the spectrum of T . The radius of the smallest closed disc with centre the origin, which contains all the spectrum is called the spectral radius.

Imitating the definition of the radius of convergence of a sequence of complex numbers we define the radius of convergence R of a sequence $\{A_n\}$ in a Banach space as the least upper bound of $|\beta|$ taken over all those β (complex) such that $A_0 + A_1\beta + A_2\beta^2 + \dots$ converges. Similarly the Césaro radius of convergence R' of $\{A_n\}$ is the least upper bound of $|\beta|$ taken over all those β (complex) such that $\frac{1}{n} \sum_{r=0}^{n-1} A_r \beta^r$ converges.

We note that $A_0 + A_1\beta + \dots$ converges for all $|\beta| < R$ and $\frac{1}{n} \sum_{r=0}^{n-1} A_r \beta^r$ converges for all $|\beta| < R'$.

CHAPTER 1INTRODUCTION

This chapter is devoted to the exposition of basic measure and ergodic theorems and the proofs of some elementary consequences of these theorems.

A.ABSOLUTE CONTINUITY AND SINGULARITY

THEOREM . R.N. (Radon - Nikodym) [1.]

If μ, ν are two measure functions (σ -finite) with the same domain $\mathcal{X}$ and $\mu \ll \nu$ there exists a finite valued measurable ($\mathcal{X}$) function f such that

$$(1.1) \quad \mu E = \int_E f(x) d\nu \quad \text{for every } E \in \mathcal{X}$$

The function f is unique in the sense that if

$\mu E = \int_E g(x) d\nu$ for every $E \in \mathcal{X}$
 then $f = g$ ν -a.e. (i.e. $f = g$ except on a set of ν measure zero)

Any one of the functions defined by (1.1) is called a derivative of μ with respect to ν and written $\frac{d\mu}{d\nu}$.

THEOREM . L [1.]

If μ, ν are two measure functions (σ -finite) with the same domain $\mathcal{X}$, there exists two uniquely determined measure functions μ_1, μ_2 such that

$$(1.2) \quad \mu = \mu_1 + \mu_2 \quad \text{and } \mu_1 \ll \nu \text{ and } \mu_2 \perp \nu.$$

The decomposition (1.2) is called the Lebesgue decomposition.

B.

MEASURABLE TRANSFORMATIONS

LEMMA 1

If μ, ν are non-trivial (i.e. not identically zero) measure functions with the same domain X and T is a measurable ($\mathcal{A}$) and non-singular (μ) transformation and $\mu \perp \nu$, there exist T invariant sets A, B with $A \cup B = X$ (the whole space) and $A \cap B = \emptyset$ such that $\mu A = 0$ and $\nu B = 0$.

PROOF.

By the definition of singularity there exist two disjoint sets $E, F \in \mathcal{A}$ such that $E \cup F = X$ and $\mu E = 0$ and $\nu F = 0$. Let $A = O(E)$ and $B = A^c$, then $\mu A = 0$ and since $F \supset A^c = B$ $\nu B = 0$. A, B are obviously T invariant.

LEMMA 2

Let $(X, \mathcal{A}, \mu)$ be a non-trivial complete measure space and T measurable non-singular (μ) transformation and ν a non-trivial measure function with domain $\mathcal{A}$. If $\mu \perp \nu$ there are wandering sets of positive measure.

PROOF.

Let A, B be an invariant decomposition of X with $\mu A = 0$ and $\nu B = 0$. Every subset of A belongs to $\mathcal{A}$. Let V be a selection (admissible by the Axiom of Choice) of points, one from each of the distinct orbits of points in A . Obviously $T^m V \cap T^n V = \emptyset$ if $m \neq n$. Moreover $O(V) = A$. Should $\nu V = 0$ we would have $\nu A = 0$ and ν would be trivial. Therefore V is a wandering set of positive measures.

THEOREM 1.

If $(X, \mathcal{A}, \mu)$ is a complete non-trivial measure space and T is a μ preserving and μ ergodic transformation and if ν is a measure function with domain $\mathcal{A}$ invariant under T and admitting no wandering sets of positive ν measure (in fact, if the singular part of ν admits no wandering sets of positive measure) then $\nu = k\mu$ for some finite scalar k .

PROOF

By Theorem L there exist two measure functions ν_1, ν_2 with domain $\mathcal{A}$ such that $\nu = \nu_1 + \nu_2$ and $\nu_1 \ll \mu$ and $\nu_2 \perp \mu$. If there are no wandering sets of positive ν_2 measure (which happens in particular when there are no wandering sets of positive ν measure) by Lemma 2 $\nu_2 \equiv 0$ and therefore $\nu \ll \mu$.

By Theorem R.N. there exists a non-negative finite valued function f such that for all $E \in \mathcal{A}$

$$\nu E = \int_E f d\mu$$

and since $\nu T^{-1}E = \int_{T^{-1}E} f d\mu = \int_E f(Tx) d\mu$

and $\nu E = \nu T^{-1}E$

we have again by Theorem R.N. $f(x) = f(Tx)$ $[\mu]$

By altering the values of f on a set of μ measure zero we can write $\nu E = \int_E h(x) d\mu$ for all $E \in \mathcal{A}$

where $h(x) = h(Tx)$ everywhere and h is finite valued.

Let $S_a = \{x: h(x) < a\}$. For all x , $h(x) < a$ if and only if $h(Tx) < a$, consequently $T^{-1}S_a = S_a$, and by the μ ergodicity of T , $\mu S_a = 0$ or $\mu(X - S_a) = 0$. $\mu S_r > 0$ for some rational number r , otherwise h would not be finite valued. Let $S = \sup\{r: \mu S_r > 0\}$ then $\mu(X - S_r) = 0$ for $r > S$ and $\mu S_r = 0$ for $r < S$, r rational. $X - S_r = \{x: h(x) \geq r\}$

and $X - S_r ⊃ X - S_u$ if $r < u$

therefore $\{x: h(x) > s\} = \bigcup_{\substack{r > s \\ r \text{ rational}}} (X - S_r)$

and $\mu \{x: h(x) > s\} = 0$

Moreover $\{x: h(x) < s\} = \bigcup_{\substack{r < s \\ r \text{ rational}}} S_r$

therefore $\mu \{x: h(x) < s\} = 0$

Consequently $h(x) = s$ μ -a.e.

and

$$\forall E = \int_E h(x) d\mu = s \mu E.$$

Contained in the statement and proof of Theorem 1 is the following corollary.

COROLLARY 1.1

If T is μ ergodic and ν, ν are invariant under T and $\nu \ll \mu$ then $\nu = k\mu$ for some finite scalar k .

We consider in this thesis transformations which are not necessarily one - one. For the one-one case, however, it is well known [3] that the existence of a finite or σ -finite invariant measure ν equivalent to a given measure μ depends on the existence of a solution of a functional equation. We concern ourselves in the next theorem with an analogous functional equation, for the least complicated of many - one transformations.

If T is a one-one measurable non-singular transformation of $(X, \mathcal{A}, \mu)$, $T^{-1}E \in \mathcal{A}$, consequently by Theorem R.N. there exists a finite valued non-negative measurable function ω such that $\mu(T^{-1}E) = \int_E \omega(x) d\mu$ i.e. $\frac{d\mu \circ T^{-1}}{d\mu} = \omega(x)$

The measures ν invariant under T and equivalent to μ are given by $\nu E = \int_E h(x) d\mu$

where h satisfies

$$0 < h(x) = h(Tx) \omega(x) \quad [x]$$

THEOREM 2.

Let T be a non-singular measurable transformation of $(X, \mathcal{A}, \mu)$ onto itself such that $X = \bigcup_{n=0}^{\infty} X_n$, $(X_n \cap X_m = \emptyset, n \neq m)$ where $\{X_n\}$ are measurable and T sends each X_n one-one onto $T(X_n)$.

$$\nu E = \int_E h(x) d\mu = \nu T^{-1}E \quad \text{for all } E \in \mathcal{A}$$

if and only if

$$h(x) = \sum_{n=0}^{\infty} h(T_n x) \omega_n(x) \quad [x]$$

where $T_n(x)$ is the unique point $y \in X_n$ with $Ty = x$ if it exists and $h T_n(x) = 0$ otherwise and $\omega_n(x) = \frac{d\mu \circ T_n^{-1}}{d\mu}$ (i.e. $\mu T_n^{-1}E = \int_E \omega_n(x) d\mu$ all $E \in \mathcal{A}$)

PROOF

For any pair of measurable sets E, F
 $E \cap TF \supset \{x: x \in E \text{ and } T_n(x) \in F\}$

but if $x \in TF$ and $T_n(x) \notin F$ we have $T^{-1}(x) \cap X_n = \emptyset$

Consequently $T_n(E \cap TF) = T_n \{x: x \in E \text{ and } T_n(x) \in F\}$

$$\begin{aligned} \text{and } \mu(T_n E \cap F) &= \mu(T^{-1}E \cap X_n \cap F) \\ &= \mu T^{-1}(E \cap TF) \cap X_n = \mu T_n(E \cap TF) \\ &= \mu T_n \{x: x \in E \text{ and } T_n(x) \in F\} \end{aligned}$$

$$\text{i.e. } \int_{T_n E} \chi_F(x) d\mu = \int_E \chi_F(T_n(x)) d\mu_{T_n}$$

For simple functions (i.e. finite linear combinations of characteristic functions) f , therefore, we have

$$(1.3) \quad \int_{T_n E} f(x) d\mu = \int_E f(T_n(x)) d\mu_{T_n}$$

By an elementary continuity argument it follows that (1.3.) holds for all non-negative measurable functions and therefore

$$\begin{aligned} (1.4) \quad \left\{ \sum_{n=0}^{\infty} \int_{T_n E} h(x) d\mu \right. &= \int_E \left(\sum_{n=0}^{\infty} h(T_n(x)) \right) d\mu_{T_n} \\ &= \int_E \left(\sum_{n=0}^{\infty} h(T_n(x)) \omega_n(x) \right) d\mu \end{aligned}$$

$$\begin{aligned} \text{If } \nu E &= \nu T^{-1}E \text{ for all } E \in \mathcal{A} \text{ and} \\ \nu E &= \int_E h(x) d\mu \text{ then } \nu T^{-1}E = \sum_{n=0}^{\infty} \nu(T^{-1}E \cap X_n) \\ &= \sum_{n=0}^{\infty} \int_{T_n E} h(x) d\mu = \int_E \left(\sum_{n=0}^{\infty} h(T_n(x)) \omega_n(x) \right) d\mu \\ &= \nu E = \int_E h(x) d\mu \quad (\text{by (1.4.)}) \end{aligned}$$

$$\text{i.e. } h(x) = \sum_{n=0}^{\infty} h(T_n(x)) \cdot \omega_n(x) \quad [\mu]$$

On the other hand if $h(x) = \sum_{n=0}^{\infty} h T_n(x) \omega_n(x) \cdot [u]$

and $\forall E = \int_E h G_n d\mu$ for all $E \in \mathcal{A}$

$$\begin{aligned} \text{then } \forall E &= \int_E \left(\sum_{n=0}^{\infty} h T_n G_n \omega_n G_n \right) d\mu && (\text{by (1.4)}) \\ &= \sum_{n=0}^{\infty} \int_{T_n E} h G_n d\mu \\ &= \sum_{n=0}^{\infty} \nu(T_n E) = \underline{\nu T^{-1} E}. \end{aligned}$$

It is quite clear that in case T is one-one the above theorem reduces to the remark preceding it.

C. HILBERT AND BANACH SPACES

The following theorem implies and is a consequence of Theorem R.N. [1.] p. 178 Ex. 3. [2], [5]

THEOREM R. (The Riesz representation theorem.)

If Λ is a continuous positive linear functional on $L^2(\mu)$ where μ is a finite measure, then there exists $g \in L^2(\mu)$ depending only on Λ such that $\Lambda f = \int f g d\mu$ for all $f \in L^2$. Moreover, g is unique in the sense that if $\Lambda f = \int f h d\mu$ for all $f \in L^2$ then $g = h$ [u].

The following two theorems are statements about the spectral properties of a bounded linear operator and the iterates of that operator.

THEOREM 3.

Let A be a bounded linear transformation of a Banach space $\mathcal{B}$ and denote the radius of convergence of $\{A^n\}$ by R , the Cesàro radius of convergence of $\{A^n\}$ by R' , the least upper bound of the radii of convergence of $\{A^n f\}$ taken over all $f \in \mathcal{B}$ by R'' , and the spectral radius of A by S , then $R = R' = R'' = 1/S$.

PROOF

We note firstly that if $\{u_n\}$ is a sequence in a Banach space and $\frac{u_0 + \dots + u_n}{n+1}$ is bounded then $\sum_0^\infty \frac{u_n}{n+1}$ converges for $|u| < 1$. [4] p. 368 Ex.14. It follows easily, that the radius of convergence of a sequence $\{u_n\}$ is the same as the Cesàro radius of convergence of $\{u_n\}$.

If S is the spectral radius of A ,
 $\lim_{n \rightarrow \infty} \| (A/S)^n \|^{1/n} = 1$ [2.]
 and therefore $\| (A/\alpha)^n \|^{1/n} = \frac{S \varepsilon_n}{|\alpha|} < K < 1$ for all sufficiently large n . Consequently $\sum_0^\infty \| (A/\alpha)^n \|$ converges and $\frac{1}{n} \sum_0^{n-1} (A/\alpha)^m$ is bounded and $\frac{1}{n} \sum_0^{n-1} (A/\alpha)^m f$ is bounded ($f \in \mathcal{B}$).
 Therefore $\sum_0^\infty (A/\beta)^m$ and $\sum_0^\infty (A/\beta)^m f$ converges for $|\beta| > |\alpha| > S$.

If $\sum_0^\infty (A/\alpha)^m f$ converges for all $f \in \mathcal{B}$ converges. It remains, therefore, to show that $\sum_0^\infty (A/\alpha)^m$ does not converge for $|\alpha| < S$.

If $\sum_0^\infty (A/\alpha)^m$ converges, there is a constant C such that $\| \sum_0^n (A/\alpha)^m \| < C$ for all n , and therefore $(A - \alpha I)^{-1}$ exists and is bounded.

Consequently for every α in the spectrum of A , $\sum_{n=0}^{\infty} (A/\alpha)^n$ does not converge. This implies the existence of an β with absolute value arbitrarily close to α , such that $\sum_{n=0}^{\infty} (A/\beta)^n$ does not converge and since the convergence of $\sum_{n=0}^{\infty} (A/\beta)^n$ (for $|\alpha| > |\beta|$) entails the convergence of $\sum_{n=0}^{\infty} (A/\alpha)^n$, it follows that $\sum_{n=0}^{\infty} (A/\alpha)^n$ does not converge for $|\alpha| < 5$.

We restrict our attention in the following theorem to the Hilbert space $L^2(\mu)$ where μ is a finite measure.

THEOREM 4.

Let A be a bounded linear transformation of $L^2(\mu)$ such that $\overline{A}f = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} A^k f$ (mean convergence)

exists for all $f \in L^2(\mu)$.

If G is a bounded linear transformation such that $GAf = Gf$ $[\mu]$ for all $f \in L^2$ then

for a fixed $g \in L^2$ $\int_E Gf d\mu = \int_E gf d\mu$ depending only on G and

$A^*g = g$ $[\mu]$ (A^* is the adjoint of A).

If, moreover, $Af = f$ only when f is constant $[\mu]$, then $G \equiv G(1) \overline{A}$ $[\mu]$

and $\overline{A}f = \int g f d\mu / \int g d\mu$ $[\mu]$ for all $f \in L^2$
i.e. $G(f)/G(1) = \int Gf d\mu / \int G(1) d\mu$ $[\mu]$

On the other hand if $A^*g = g$ $[\mu]$ then $G \equiv g \overline{A}$ satisfies $GAf = Gf$ $[\mu]$ and $\int Gf d\mu = \int g f d\mu$ for all $f \in L^2$.

PROOF

By Theorem R. there exists $g \in L^2$ depending only on G such that $\int_E G f d\mu = \int_E g f d\mu$.

Therefore $\int A^* g h d\mu = \int g A h d\mu = \int G A h d\mu = \int G h d\mu$
 $= \int g h d\mu$ for all $h \in L^2$.

Therefore $A^* g = g$ $[E]$.

$A f = f$ $[E]$, implies f is constant $[E]$, is equivalent to the requirements that $\bar{A} f$ is constant $[E]$ for all f .

$$G f = G A f = G A^2 f = \dots \\ = \frac{1}{n} \sum_{i=0}^{n-1} G A^i f = G \bar{A} f = \bar{A} f \cdot G(1) \quad [E]$$

since $\bar{A} f$ is constant.

i.e. $G f = G(1) \bar{A} f$.

$\bar{A} f$ is constant $[E]$, i.e. $\bar{A} f = k(f) [E]$

and therefore k is a continuous positive linear functional

and by Theorem R, $\bar{A} f = \int h f d\mu [E]$

where h is independent of f .

$$\int G f d\mu = \int G(1) \bar{A} f d\mu = \int G(1) d\mu \cdot \int h f d\mu = \int h (\int G(1) d\mu) f d\mu$$

i.e. $g = h \int G(1) d\mu [E]$

But $\int G(1) d\mu = \int g d\mu$ therefore $h = g / \int g d\mu [E]$

i.e. $\bar{A} f = \int g f d\mu / \int g d\mu [E]$

$$\text{and } \frac{G(f)}{G(1)} = \frac{G(1) \bar{A} f}{G(1)} = \frac{\int G(f) d\mu}{\int G(1) d\mu}$$

$$\begin{aligned} \text{On the other hand, if } A^* g = g \text{ then } \int g \bar{A} f d\mu \\ = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \int G A^i f d\mu = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \int (A^*)^i g f d\mu \\ = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \int g f d\mu = \int g f d\mu. \end{aligned}$$

Under the conditions of the above theorem it is clear that all eigen vectors of A^* with value 1, are yielded by operators G such that $GAf = Gf$ [μ] for all $f \in L^2$.

Suppose T is a measurable non-singular point transformation of (X, α, μ) onto itself such that $\frac{1}{n} \sum_{k=0}^{n-1} f(T^k x)$ converges in L^2 for all $f \in L^2$ then (the induced transformation)

satisfies the initial condition of the theorem. The requirement that $Af = f$ imply f is constant is the ergodic condition and in this case the only G with $GAf = Gf$ for all $f \in L^2$, are the transformations of type $h\bar{A}$.

$$\int h\bar{A}f d\mu = \int h d\mu (\int f g d\mu) = k(h) \int f g d\mu$$

for some $g \in L^2$ depending only on $\bar{A}$, where k is a constant depending on h . Therefore the only eigen vector with value 1 of A^* is the unique function g such that $\bar{A}f = \int g f d\mu$ [μ]

If $A^*g = g$ [μ] then it is easy to show that $\nu E = \int_E g d\mu$ is an invariant measure equivalent to μ . Consequently we have the well known result :

COROLLARY 4.1

If $\frac{1}{n} \sum_{k=0}^{n-1} f(T^k x)$ converges in L^2 for every $f \in L^2$ then there exists an invariant measure equivalent to μ , $\nu E = \int_E g d\mu$ where g is a non-trivial, non-negative solution of $A^*g = g$ (where A is the induced linear transformation.) If T is ergodic the solution is unique.

D.

ERGODIC THEOREMS

The first ergodic theorem proved, the so-called mean-ergodic theorem, is due to Von-Neumann [6.], who made use of the fact, discovered by Koopmann [7.], that an operator A on $L^2(\mu)$ (where μ is a finite measure) induced by an invertible measure preserving transformation T is unitary i.e. $A^* = A^{-1}$ where $A f(x) = f(Tx)$ and $A^{-1} f = f(T^{-1}x)$.

Corollary 4.1 can be viewed as a converse to Von-Neumann's theorem.

THEOREM N. (The mean-ergodic theorem. Von-Neumann)

If $(X, \mathcal{A}, \mu)$ is a finite measure space and T is a one-one μ preserving transformation of X onto itself then for all $f \in L^2(\mu)$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) = f^*(x) \quad \text{exists (mean convergence)}$$

and $f^*(x) \in L^2(\mu)$.

Actually the above theorem is true when $L^2(\mu)$ and its norm is replaced by $L^p(\mu)$ and its norm for any $p \geq 1$ [8.]

Birkhoff's (individual) ergodic theorem is concerned with point-wise convergence in the case where T preserves the finite measure μ [9.]. Hopf generalised this to the case where T preserves a σ -finite measure μ [10.]. The most general result in this direction is due to Hurewicz [11.] whose theorem is concerned with the point-wise convergence of weighted averages, when the hypothesis of an invariant measure is dropped. We will

state the theorem in a form given it by Oxtoby [12] and Halmos [13] for one-one transformations and extended to the many-one case by Y.N. Dowker [14].

In the one-one case we can define measures $m \circ T^{-n}$, absolutely continuous with respect to m so that

$$m \circ T^{-n} E = \int_E \omega_n(x) dm \quad \text{for all } E \in \mathcal{A}$$

where $\omega_n(x)$ depends on n but not on E . It can be

shown that $\omega_n(x) = \omega(T^{-n}x) \cdots \omega(x)$ where

$$\omega(x) = \omega_1(x)$$

When T is not one-one, $m \circ T$ is not necessarily a measure function. It is a measure, however, when its domain is restricted to $T^{-1}(\mathcal{A})$ and in this case it is absolutely continuous with respect to m . Consequently there is a function $\omega(x)$ unique among $T^{-1}(\mathcal{A})$ measurable functions such that

$$m \circ T E = \int_E \omega(x) dm \quad \text{for all } E \in T^{-1}(\mathcal{A})$$

Defining formally $\omega_n(x) = \omega(T^{-n}x) \cdots \omega(x)$

Hurewicz's ergodic theorem in its most general form states:

THEOREM Hu. (Hurewicz's ergodic theorem)

If T is a measurable non-singular many-one transformation of $(X, \mathcal{A}, m)$ onto itself and if f is

integrable and g is non-negative and measurable and such

that $g(x) + g(Tx) + \cdots = \infty$ [m]

then $\lim_{n \rightarrow \infty} \frac{f(x) + \cdots + f(T^{n-1}x) \omega_{n-1}(x)}{g(x) + \cdots + g(T^{n-1}x) \omega_{n-1}(x)} = h(x) < \infty$ [m]

exists (point-wise convergence), and $g(x) h(x)$

is integrable.

$\int_E f(x) dm = \int_E g(x) h(x) dm$ for every invariant set E

such that

$$\int_E g(x) dm < \infty$$

The above theorem certainly reduces to Hurewicz's when T is one-one. If μ is invariant and σ -finite then $\omega(x) = 1$ and we have Hopf's ergodic theorem (which we refer to as Theorem Ho.) which in turn reduces to Birkhoff's ergodic theorem (which we refer to as Theorem B.) when μ is finite.

The following theorems ^{are useful} when proving the existence of a finite invariant measure equivalent to a given finite measure.

THEOREM D. M. (Dunford and Miller.) [15]

If (X, α, μ) is a finite measure space and T is a measurable non-singular many-one transformation the following two statements are equivalent :

- i) There exists a constant M such that

$$\frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}E) < M \mu E \quad \text{for all } E \in \alpha, n=1, 2, \dots$$
- ii) If f is integrable then so is $f(Tx)$ and

$$\frac{1}{n} \sum_{k=0}^{n-1} f(T^k x)$$
 converges in the mean $L(\mu)$.

Each of the above statements implies :

$$\frac{1}{n} \sum_{k=0}^{n-1} f(T^k x)$$
 converges point-wise $[\mu]$ for every integrable f .

THEOREM C. (Caldéron.) [16.]

If (X, α, μ) is a finite measure space and T is a many-one non-singular measurable transformation, a necessary and sufficient condition for the existence of a finite invariant measure equivalent to μ is that either

- i) $\lim_{n \rightarrow \infty} \mu T^{-n} A > 0$ for all A with $\mu A > 0$
 or ii) $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu T^{-i} A > 0$ for all A with $\mu A > 0$
-

Y.H. Dowker [17] states the following theorem for one-one transformations. It is true, however, for many-one transformations; the proof is exactly the same as in [17] and is based likewise on Theorem C., whilst noting that the latter theorem is true for many-one transformations.

THEOREM D. (Dowker.)

If $(X, \mathcal{A}, \mu)$ is a finite measure space and T is a many-one transformation of X onto itself, a necessary and sufficient condition for the existence of a finite invariant measure equivalent to μ is that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu T^{-i} A > 0 \quad \text{for all } A$$

with $\mu A > 0$.

CHAPTER 2

ISOMORPHISMS AND REPRESENTATIONS

Isomorphisms between two measure spaces are of two kinds ; set isomorphisms and point isomorphisms. A necessary and sufficient condition that a normalised measure space be set isomorphic with the unit interval is that it be seperable and non-atomic. [1], [18].

Halmos and Von-Neumann in [18], define a concept of normality for non-atomic seperable measure spaces, which, roughly speaking, requires measurable functions to have measurable ranges on the real line. Rohlin has also studied the nature of spaces isomorphic with the unit interval. [19], [20]. The Lebesgue spaces defined in [19], [20] (which are seperable) have the property that it is precisely measurable subsets which constitute Lebesgue sub-spaces. Consequently in order that a seperable space be a Lebesgue space, it is necessary and sufficient that it be absolutely measurable i.e. measurable in every seperable space containing it. Rohlin allows atoms in Lebesgue spaces , but since they are at most countable their union can be dealt with seperately. Rohlin's result states, essentially, that a normalised non-atomic Lebesgue space is isomorphic with the unit interval.

If (X_1, α_1, m_1) and (X_2, α_2, m_2) are measure spaces, a point isomorphism between X_1 and X_2 is a one-one mapping from almost all of X_1 onto almost all of X_2 such that $E \in \alpha_1$ if and only if $TE \in \alpha_2$, and then $m_1 E = m_2 TE$. If such a mapping exists X_1 and X_2 are point isomorphic.

An m -space is a complete measure space (X, α, m) on which a metric is defined so that topologically it is a complete separable space and which satisfies :

i) The measure of a non-empty open set is positive.

ii) For every measurable E ,

$$m E = \inf \{ m U : U \supset E, U \text{ open} \}$$

i.e. m is regular (with respect to open sets.)

A normalised m -space is isomorphic with the unit interval, in fact the isomorphism sends the entire space onto almost all of the unit interval. [18], [19], [20], [21] p. 269 , [22] p. 301 . The first proof of a theorem which implies this was due to Hausdorff, who also defined the notion of an absolute Borel set , which is the set or topological equivalent of an absolutely measurable set [19], [20] or a set measurable in itself [18].

A.

ISOMORPHISMS

Let (X, α, m) be a normalised measure space and $X' \subset X$ with $m^* X' = \inf \{ m E : E \supset X', E \in \alpha \} = 1$

(i.e. X' is a thick subset of X) . If α' is the set of all $E \Delta X'$ with $E \in \alpha$ and $m' E \Delta X = m E$, it is known that m' is unambiguously defined and (X', α', m') is a normalised measure space, (a thick sub-space of (X, α, m)) N.B. $m' E = m^* E$ for all $E \in \alpha'$. [18], [19], [23]

If (X, α, m) is a measurable sub-space of an m -space (i.e. (X, α, m) is an m -sub-space) then it is easy to see that m' is regular with respect to sets $X' \cap U$, U open.

We have remarked that normalised m -spaces are isomorphic with the unit interval ; our first result shows that normalised m -spaces essentially typify spaces isomorphic with the unit interval , in fact :

THEOREM 5.

A measure space (X, α, m) is isomorphic with a thick sub-set (I, λ, μ) of the unit interval if and only if (X, α, m) is a thick sub-space of a normalised m -space. (X, α, m) is isomorphic with the unit interval if and only if (X, α, m) is a measurable thick subset of a normalised m -space.

PROOF

Let T be an isomorphism of (X, α, m) onto a thick subset I of the unit interval I' . By reducing X and I each by a set of measure zero , if necessary, we suppose T sends the entire space X onto I . If μ' is the Lebesgue measure on I and $x, y \in X$, define $d(x, y) = \mu'(T_x, T_y)$.

(X, d) is a separable metric space whose completion we denote by (X', d') . If we extend T to T' by continuity so that the domain of T' is X' then the range of T' will be I' and T' will be one-one. Moreover, T' and the inverse of T' will be continuous. Therefore we can induce the order of I' on X' . Let β be the Borel sets generated by open sets in X' . Since T' is a homeomorphism $\beta \cap X \in \alpha$ if $\beta \in \beta$. We define $\gamma \beta = m \beta \cap X$ so that (X', β, γ) is

a measure space ; its measure completion we denote by (X', α', m') , and we show that (X', α', m') is an m -space, and (X, α, m) is a thick sub-space of (X', α', m') .

X' is a complete separable metric space. (X', α', m') is a normalised complete measure space. If $E \in \alpha'$, $E \cap X \in \alpha$ and $m'E = m(E \cap X)$; if $F \in \alpha$ there exists $E \in \alpha'$ such that $F = E \cap X$. If U is a non-empty open set in X' , $U \cap X$ is a non-empty open set in X , therefore $m'U = m(U \cap X)$. If $E \in \alpha'$

$$\begin{aligned} & \inf \{ m'U : U \supset E, U \text{ open in } X' \} \\ &= \inf \{ m(U \cap X) : U \supset E, U \text{ open in } X' \} \\ &= \inf \{ mT'U : U \supset E, U \text{ open in } X' \} \\ &= \inf \{ mV \cap I : V \supset T'E, V \text{ open in } I' \} \\ &= \inf \{ m'V : V \supset T'E, V \text{ open in } I' \} \\ &= m'(T'E) \end{aligned}$$

$$\begin{aligned} & \text{sub } \mu(T'E \cap I) \\ &= m(E \cap X) \\ &= m'E. \end{aligned}$$

(since $T'E$ is a Lebesgue measurable subset of I' .)

Therefore m' is regular with respect to open sets and (X', α', m') is an m -space . It is clear that if I is a Lebesgue set then X is a measurable thick subset of X' .

On the other hand suppose (X, α, μ) is a thick sub-space of a normalised μ -sub-space (X', α', μ') . A normalised μ -sub-space is isomorphic to the unit interval, since a measurable sub-set of a normal space is itself a normal space. [18]. It follows immediately that (X, α, μ) is isomorphic with a thick sub-space of the unit interval. Clearly if X is a measurable subset of X' then (X, α, μ) is isomorphic with the unit interval itself.

Incidental to the proof of Theorem 5 we have proved :
COROLLARY 5.1

A thick sub-space of a normalised μ -sub-space is a thick sub-space of a normalised μ -space.

In our next result we replace the complete separable metric space by a linearly ordered space.

THEOREM 6

If (X, α, μ) is a normalised non-atomic measure space where X is linearly ordered so that α is the σ -algebra generated by the intervals and μ is regular with respect to open sets and positive on non-empty open sets then (X, α, μ) is point isomorphic with a thick sub-set of the unit interval.

PROOF

Define $f(x) = \mu \{y : y < x\}$

$f(x)$ is strictly increasing (except on a countable set which we exclude for convenience) Since μ is positive on non-empty open intervals. If $x_n \uparrow x$, $f(x_n) \leq f(x)$

and since $\{y: y < x_n\} \uparrow \{y: y < x\}$
 $f(x_n) = m \{y: y < x_n\} \uparrow f(x) = m \{y: y < x\}$

If $x_n \downarrow x$, $-f(x_n) \geq -f(x)$,

and since $\{y: y \leq x_n\} \downarrow \{y: y \leq x\}$

and m is non-atomic, $-f(x_n) \downarrow -f(x)$

f maps X into $[0, 1]$.

The distribution of f is defined as

$$F(a) = m \{x: f(x) < a\}$$

If $a = f(y)$, $F(a) = m \{x: f(x) < f(y)\}$
 $= m \{x: x < y\} = f(y) = a.$

It is known, and it is easy to show, that $F(a)$ is continuous from the left. The points $f(y)$, $y \in X$ are everywhere dense in $[0, 1]$ for otherwise there would exist an interval (a, b) such that $f(y) < a$ or $f(y) > b$ for all

$y \in X$. If a' is the least upper bound of $f(y)$ such that $f(y) < a$ and b' is the greatest lower bound of $f(y)$ such that $f(y) > b$ then $a' \leq a < b \leq b'$. Let $x_n \uparrow$ and $y_n \downarrow$ so that $f(x_n) \uparrow a'$ and $f(y_n) \downarrow b'$ then

$$m \bigcap_n (x_n, y_n) = b' - a' > 0$$

Since m is non-atomic there exists t such that $x_n < t < x_{n+1}$ for all n ; $c = m \{y < t\} \cap \bigcap_n (x_n, y_n) > 0$ and

$$d = m \{y > t\} \cap \bigcap_n (x_n, y_n) > 0$$

$$f(t) = m \bigcup_n \{y < x_n\} + c = m \bigcap_n \{y < y_n\} - d$$

$$= a' + c = b' - d.$$

i.e. $a' < f(H) < b'$. This is evidently in contradiction with the definitions of a', b' on the assumption that there is no $y \in X$ with $a \leq f(y) \leq b$.

$F(a) = a$ in an everywhere dense set in $[0, 1]$ and since $F(a)$ is continuous from the left $F(a) = a$ for all $a \in [0, 1]$.

Since the points $f(y), y \in X$ are dense in $[0, 1]$, $J = f^{-1}K$ is a sub-interval of X , for every sub-interval K of $[0, 1]$. Moreover, if J is a sub-interval of X , $J = f(f(J)) = f^{-1}K$, where K is a sub-interval of $[0, 1]$, (viz. the closure of $f(J)$ save perhaps at end points.) This is sufficient to establish $f(\beta) \supset \alpha$ and $f(\alpha) \supset \beta \cap f(X)$ where β is the class of Borel subsets of $[0, 1]$, i.e. $\alpha = f^{-1}\beta$ and $\beta \cap f(X) = f(\alpha)$.

It follows from the regularity of m and the properties of the distribution function $F(a) = a$, that if $A \in \alpha$ then $f(A) = \beta \cap f(X)$ for some $\beta \in \beta$ and $m' \beta \cap f(X) = m \beta$ (m Borel measure) and $m A = m' \beta \cap f(X)$

The proof is complete , since $(f(X), \beta \cap f(X), m')$ is a thick subset of $[0, 1]$, by virtue of the fact that if U is open in $[0, 1]$ and $U \supset f(X)$ then $f^{-1}U \supset X$ and $m f^{-1}U \geq 1$.

B.

PARTITIONS

If (X, α, m) is a finite measure space and $\{a^n\}$ is a sequence of measurable partitions ($a^n = A_1^n, A_2^n, \dots$) we call sets of the form $A_{n_0}^0 \cap A_{n_1}^1 \cap \dots \cap A_{n_m}^m$ cylinders. m is said to be regular with respect to cylinders if when $E \in \alpha$ and $\varepsilon > 0$ we can find a sequence of cylinders $\{C_n\}$ such that $UC_n \supset E$ and

$$m(UC_n - E) < \varepsilon$$

LEMMA 3

If α is generated by the sets of $\{a^n\}$ and (X, α, m) is a normalised measure space so that m is regular with respect to cylinders then m is non-atomic if and only if

$$m(A_{n_0}^0 \cap A_{n_1}^1 \cap \dots) = 0$$

for every

sequence $n_0, n_1, \dots$.PROOF

If there exists an atom E , then either $C \cap E = \emptyset$ $[m]$ or $C \supset E$ $[m]$ for all cylinders C . Since E is approximable from outside by unions of cylinders, E is approximable $[m]$ from outside by single cylinders. Therefore there is a sequence $n_0, n_1, \dots$ such that

$$A_{n_0}^0 \cap A_{n_1}^1 \cap \dots = E \quad [m]$$

and

$$m(A_{n_0}^0 \cap A_{n_1}^1 \cap \dots) > 0$$

On the other hand if (X, α, m) is non-atomic we have
 $\lim_{m \rightarrow \infty} m(A_{n_0}^0 \cap \dots \cap A_{n_m}^m) = 0$ [1.] Th. A. p.172.

If $x \in A_{n_0}^0 \cap A_{n_1}^1 \cap \dots$ we say that $n_0, n_1, \dots$ is the representation of x . (It is easy to see that each x has one and only one representation by virtue of the A^n being partitions.) We write $x \sim n_0 n_1 \dots$.

We say $x \sim y$ (x is equivalent to y) when x, y have the same representation; and say that the representation of a point x is valid if the set of points with the same representation, (which is measurable) , has measure zero. Evidently by Lemma 3, all representations are valid if and only if (X, α, m) is non-atomic.

The equivalence classes we denote by $\tilde{x}$ and $\tilde{X}$ denotes the collection of all equivalence classes.

LEMMA 4

Let (X, α, m) be a measure space where α is generated by a sequence of partitions. If $x \in E \in \alpha$, $\tilde{x} \subset E$.

PROOF

Let β be the collection of sets $F \in \alpha$ such that either $F \supset \tilde{x}$ or $F \cap \tilde{x} = \emptyset$. If $A, B \in \beta$ then $A \cap B \in \beta$. $X \in \beta$. If $A_n \in \beta$ and one of

$A_n \supset \tilde{\mathcal{X}}$ then $\cup A_n \supset \tilde{\mathcal{X}}$ and $\cup A_n \in \beta$. If $A_n \cap \tilde{\mathcal{X}} = \emptyset$ for all A_n , then $\tilde{\mathcal{X}} \cap \cup A_n = \emptyset$ and $\cup A_n \in \beta$. Therefore β is a σ -algebra. Moreover, if $\mathcal{E} = A_{n_0}^0 \cap A_{n_1}^1 \cap \dots$ and $C = A_{m_0}^0 \cap \dots \cap A_{m_k}^k$ then $C \supset \tilde{\mathcal{X}}$ if and only if $m_i \neq n_i, i \leq k$ and if $m_i \neq n_i$ for some $i \leq k, C \cap \tilde{\mathcal{X}} = \emptyset$. Therefore β contains all cylinders and hence $\beta = \alpha$.

If we write $\tilde{E}$ for the set of $\tilde{\mathcal{X}}$ such that $\mathcal{X} \in E$ and $\tilde{\alpha}$ for the set of all $\tilde{E}$ such that $E \in \alpha$ and $\tilde{\omega} \tilde{E} = \omega E$ for $E \in \alpha$ (i.e. $\tilde{E} \in \tilde{\alpha}$), $(\tilde{\mathcal{X}}, \tilde{\alpha}, \tilde{\omega})$ is a measure space which is the quotient with respect to $\sim$, and we call it the representation quotient. Evidently the representation quotient is non-atomic if and only if the original space is non-atomic.

In the remainder of this section we consider a normalised non-atomic measure space (X, α, ω) where α is generated by a sequence of partitions $\{\alpha^n\}$ where each partition has the same countable or finite number of components and such that ω is regular with respect to cylinders and positive on non-empty cylinders, and no two points have the same representation.

It is possible to lineally order X in two distinct ways;
D.O. The direct ordering.

We say $n_0 n_1 \dots < m_0 m_1 \dots$ when for the first i such that $n_i \neq m_i$ we have $n_i < m_i$.
 $x < y$ if the representation of x is less than the representation of y .

A.O. The alternating ordering.

We say $n_0 n_1 \dots < m_0 m_1 \dots$ when for the first i such that $n_i \neq m_i$ we have $n_i > m_i$ if i is even, $n_i < m_i$ if i is odd. $x < y$ is the

representation of x less than the representation of y .

Let $x \sim u_0 u_1 \dots$ then $y \geq x$ if and only if

$$y \in \left(\bigcup_{t_0 \geq u_0} A_{t_0}^0 \right) \cap \left(\bigcup_{t_0 t_1 \geq u_0 u_1} A_{t_0 t_1}^1 \right) \cap \dots \in \alpha$$

therefore $\{y : y \geq x\} \in \alpha$

Similarly $\{y : y \leq x\} \in \alpha$

Therefore open, closed, and half-open half-closed intervals are measurable.

Since (X, α, m) is non-atomic, if C is any non-empty cylinder

$$\inf_{x \in C} m\{y : y < x\} \cap C = 0$$

$$\inf_{x \in C} m\{y : y > x\} \cap C = 0$$

Let $x_n, z_n \in C$ so that

$$m\{y : y < x_n\} \cap C < \frac{1}{2^n}$$

$$m\{y : y > z_n\} \cap C < \frac{1}{2^n}$$

then $(x_n, y_n) \subset C$ and $m(x_n, y_n) > mC - 1/n$

Therefore $\bigcup (x_n, y_n) = C$ (except perhaps at two points).

Therefore every cylinder is the union of intervals.

LEMMA 5.

α is the class of Borel sets generated by the open sets, and m is regular with respect to open sets.

LEMMA 6.

$mI > 0$ for every non-empty open interval I .

PROOF.

If $I = (x, y) \ni z, x < y$
 and $x \sim x_0, x_1, \dots$
 $y \sim y_0, y_1, \dots$
 $z \sim z_0, z_1, \dots$ for some m
 $x_0, x_1, \dots, x_m < z_0 \dots z_m < y_0 \dots y_m$
 therefore $(x, y) \supset A_{z_0}^0 \cap \dots \cap A_{z_m}^m \ni z$
 and since $mC > 0$ for every non-empty cylinder $C, mI > 0$

We say that (X, m) has an associated transformation if
 when $A_{n_0}^0 \cap A_{n_1}^1 \cap \dots$ is non-empty, $A_{n_1}^0 \cap A_{n_2}^1 \cap \dots$
 is non-empty and $A_{n_0}^1 \cap A_{n_1}^2 \cap \dots$ is non-empty.

The associated transformation is defined as $Tx = y$
 where $y \sim n_1, n_2, \dots$, if $x \sim n_0, n_1, \dots$. Obviously T is a
 transformation of X onto itself; we note that if $x, y \in A_n^0$
 then

$$(2.1) \quad \begin{cases} Tx < Ty \text{ (D.O.) if } x < y \text{ (D.O.)} \\ Tx > Ty \text{ (A.O.) if } x < y \text{ (A.O.)} \end{cases}$$

$T^{-m} A_\theta^0$ is the set of all x with representation of
 the form $x \sim n_0, n_1, \dots, n_{m-1}, \theta, n_{m+1}, \dots$
 i.e. $T^{-m} A_\theta^0 = A_\theta^m$

and therefore

$$(2.2) \quad T^{-m} a^0 = a^m$$

$$(2.3) \quad \left\{ \begin{array}{l} T^m \text{ is one-one from } A_{n_0}^0 \cap \dots \cap A_{n_{m-1}}^{m-1} (= A_{n_0}^0 \cap \dots \cap T^{-(m-1)} A_{n_{m-1}}^0) \\ \text{onto } T^m (A_{n_0}^0 \cap \dots \cap T^{-(m-1)} A_{n_{m-1}}^0) (= T^m A_{n_0}^0 \cap \dots \cap T A_{n_{m-1}}^0) \\ \text{for if} \end{array} \right.$$

$$x \sim n_0, n_1, \dots, n_{m-1}, r_m, \dots$$

$$y \sim n_0, \dots, n_{m-1}, s_m, \dots$$

$$\text{then } T^m x = \tau_m \dots$$

$$T^m y = s_m \dots$$

and $T^m x = T^m y$ if and only if $\tau_m \dots = s_m \dots$
 i.e. $x = y$

LEMMA 7.

T is measurable.

PROOF.

It suffices to show that TI and $T^{-1}I$ are measurable when I is an interval, for if I_n is an interval contained in A_n and $\mathcal{B}(I_n)$ is the class of all measurable subsets E of I_n such that TE and $T^{-1}E$ are measurable, then $\mathcal{B}(I_n)$ is the σ -algebra of subsets of I_n containing all sub-intervals of I_n , (utilising the fact that T and T^{-1} both distribute over unions and differences since T is one-one on I_n .)

We show that $T^{-1}I$ and TI are measurable for every interval I .

$I = \bigcup_n (A_n \cap I)$ therefore $T^{-1}I = \bigcup_n T^{-1}(A_n \cap I)$
 and $TI = \bigcup_n T(A_n \cap I)$. $A_n \cap I$ is the union of an increasing sequence of intervals contained in A_n and therefore we need only show that $TI, T^{-1}I$ are measurable when $I \subset A_n$ and I is an interval. By (2.1) TI is an interval and using (2.1) it is easy to see that $T^{-1}I \cap A_n$ is a union of an increasing sequence of intervals or empty or a single point for every A_n . For if $xy \in T^{-1}(I \cap A_n)$ Tx, Ty and all points in between belong to I .

C.

PARTITION SYSTEMS.

In this section we suppose (X, α, μ) is a normalised non-atomic measure space and T is a measurable transformation of X onto itself, where α is generated by a partition α and its inverse iterates such that no two points have the same representation (with respect to the sequence of partitions $\alpha, T^{-1}\alpha, \dots$), μ is regular with respect to cylinders and μ is positive on non-empty cylinders. The space together with the transformation are said to form a partition system.

By our previous construction of an associated transformation it is clear that (X, α, μ) has an associated transformation and that this transformation is T . We remark also that in the above had we not supposed distinct points to have distinct representation our considerations would apply to the representation quotient.

Define $\phi(x) = \mu \{y : y < x\}$ where the ordering is direct or alternating. By Lemma 5 and Lemma 6, (X, α, μ) satisfies conditions of Theorem 6, consequently:

THEOREM 7.

(X, α, μ) is point isomorphic under ϕ with a thick subset of the unit interval. This subset can be taken to be the ~~entire~~ unit interval when and only when the range of ϕ is measurable.

$\phi(x) = \phi(y)$ If and only if there are no points between x, y . There are only a countable number of such pairs (x, y) . If N is a selection of one point from each such pair and $O(N)$ is the smallest T invariant set containing N , then $O(N)$ is a countable set. Excluding this set from X we may assume T sends X onto itself and $\phi(x)$ is strictly increasing. We can also assume that X has no maximal element, by excluding it and its iterates if it exists. In fact we exclude the larger countable sets viz. the smallest invariant set $O(M)$ containing end points of cylinders M . We define a transformation of $\phi(X) \subset [0, 1]$ onto itself using the same symbol, without confusion, as for the transformation of X onto itself:

$$T\phi(x) = \phi(Tx)$$

Two partition systems $(X_1, \alpha_1, m_1, T_1), (X_2, \alpha_2, m_2, T_2)$ are said to be point isomorphic under ϕ (where ϕ is the isomorphism) if (X_1, α_1, m_1) and (X_2, α_2, m_2) are point isomorphic under ϕ and $\phi^{-1}T_2\phi = T_1$ $[m_1]$

From theorem 7 it follows :

COROLLARY 7.1.

(X, α, m, T) is point isomorphic under ϕ with a thick subset of the unit interval together with the transformation $T\phi(x) = \phi(Tx)$.

If $x_n \downarrow x \notin O(M)$ where $x_n \in \phi(A_n)$ define $Tx = \lim_{n \rightarrow \infty} Tx_n$ which exists since $Tx_n \downarrow$ if (D.O.) and $Tx_n \uparrow$ if (A.O.). (if $T(0)$ is not defined, which can only happen (A.O.) case

when $\{A_n\}$ is an infinite sequence, define $T(0) = 0$.)
 Let $I = [0, 1 - \phi(0, \infty))$. T sends I onto itself,
 moreover, one can show that T is continuous from the left
 and continuous from the right everywhere in I , and
 consequently that T is Borel measurable.

Let $\overline{A}$ denote the topological closure of A .

D.O. Direct Ordering.

$$\begin{aligned} \text{Let } I_0 &= \overline{\phi(A_0)} \cap I = (0, x_1) \cap I \\ I_1 &= \overline{\phi(A_1)} \cap I = (x_1, x_2) \cap I \\ &\vdots \end{aligned}$$

then $0 < x_1 < x_2 \dots \leq 1$

Define $\psi(x) = n + T(x)$ if $x \in I_n$

then ψ sends I one-one ~~onto~~ $\cup J_n'$ where
 $J_n' = J_n \cap (I + n)$ and J_n' is an interval contained in
 $[n, n+1)$. If $\phi'(x) = \psi^{-1}(x)$ then ϕ' is a strictly
 increasing function sending $\cup J_n'$ onto I . Furthermore
 ϕ' is continuous on each J_n' . We may extend the
 domain of ϕ' so that the extended function —

A. ϕ' is a continuous, strictly increasing function
 with domain $\cup J_n$ and range $[0, 1)$ where
 $J_n = [a_n, b_n) \subset [n, n+1)$ and $Tx = (\phi'^{-1}x)$ for $x \in I$ and
 $\cup J_n \bmod 1 = [0, 1)$ ($\{z\}$ is the fractional part of z)

A.O. Alternating Ordering.

Similarly there is a function ϕ such that

B. ϕ is a strictly decreasing continuous function with
 domain $\cup J_n$, where $J_n = [a_n, b_n] \subset (n, n+1]$ and range
 $[0, 1)$ such that $Tx = (\phi^{-1}x)$ for $x \in I$ and $\cup J_n \bmod 1 = [0, 1)$
 (Should this be ambiguous for $x = 0$; $T(0) = 0$ when

$$\varphi^{-1}(0) = \infty).$$

Conversely it is easy to show that:

LEMMA 8.

If φ is a function satisfying A. or B. then $T_x = (\varphi^{-1}x)$ for $x \in [0, 1)$ is a measurable transformation of $[0, 1)$ onto itself.

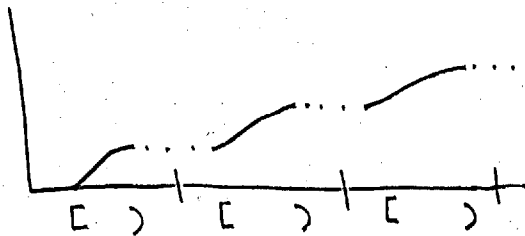
D.

REPRESENTATIONS.

D.O.

Let φ satisfy A. We note that if $x \in [0, 1)$
 $T^i(x) = (\varphi^{-1} T^{i-1}(x))$
 and we define $\varepsilon_{i+1}(x) = [\varphi^{-1} T^i x]$

$\bar{\varphi}$ denotes the monotone continuous extension of φ to the domain $[0, \infty]$.



$$\begin{aligned} \varphi &: \text{---} \\ \bar{\varphi} &: \text{---} \dots \end{aligned}$$

Let $\psi_n(x) = \bar{\varphi}(\varepsilon_1(x) + \bar{\varphi}(\varepsilon_2(x) + \dots + \bar{\varphi}(\varepsilon_n(x))))$
 then $\psi_n(x)$ is non-decreasing as n increases for every $x \in [0, 1)$
 and $\psi_n(x)$ is non-decreasing for a fixed n as x increases, and since $x = \varphi(\varepsilon_1 + \dots + \varphi(\varepsilon_n + T^n(x)))$

$$\psi_n(x) \leq x \quad \text{for every } x$$

i.e.

$$\psi(x) = \lim_{n \rightarrow \infty} \psi_n(x) \leq x.$$

Suppose $y = \psi_n(x) < x < 1$
 then $y = \varphi(\varepsilon_1 + \overline{\varphi}(\varepsilon_2 + \dots) + \delta_1) < x$ for some $\delta_1 > 0$
 and $\varepsilon_1 + \overline{\varphi}(\varepsilon_2 + \dots) < \varphi^{-1}(y) < \varphi^{-1}(x)$
 therefore $\overline{\varphi}(\varepsilon_2 + \dots) < Ty < Tx$ and $\varepsilon_1(x) = \varepsilon_1(y)$
 Again $\varphi(\varepsilon_2 + \overline{\varphi}(\dots) + \delta_2) < Ty < Tx$ for some $\delta_2 > 0$
 and $\varepsilon_2 + \overline{\varphi}(\varepsilon_3 + \dots) < \varphi^{-1}Ty < \varphi^{-1}Tx$
 and $\overline{\varphi}(\varepsilon_3 + \dots) < T^2y < T^2x$ and $\varepsilon_2(x) = \varepsilon_2(y)$.
 Proceeding in this manner we get $\varepsilon_i(x) = \varepsilon_i(y)$ for
 $i \in \mathbb{N}$.

Consequently

$$(2.4.) \quad \psi_m(\psi_n(x)) = \psi_m(x) \quad \text{for } m \leq n$$

$$\text{and } \varepsilon_m(\psi_n(x)) = \varepsilon_m(x) \quad \text{for } m \leq n$$

Since $\psi_m(x)$ is non-decreasing for a fixed m as x increases

$$\psi_m(x) = \psi_m(\psi_n(x)) \leq \psi_m(\psi(x)) \leq \psi_m(x).$$

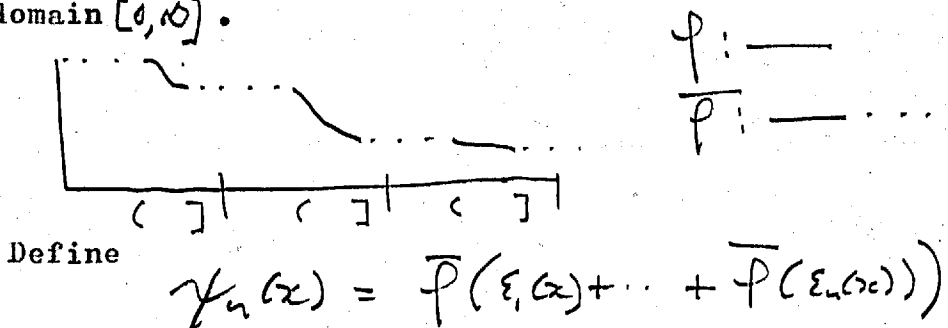
$$(2.5.) \quad \text{Therefore } \psi_m(x) = \psi_m(\psi(x)) \quad \text{and } \varepsilon_m \psi = \varepsilon_m$$

and

$$(2.6.) \quad \psi(x) = \psi(\psi(x)).$$

A.O.

Let φ satisfy B. We note that $T^i x = (\varphi^{-1} T^{i-1}(\varphi(x)))$
 for $x \in [0, 1)$ and define $\varepsilon_{i+1}(x) = [\varphi^{-1}(T^i x)]$
 $\overline{\varphi}$ denotes the continuous monotone extension of
 the domain $[0, \infty)$.



$\psi_{2n}(x)$ is non-decreasing as n increases for every x ; $\psi_{2n+1}(x)$ is non-increasing as n decreases for every x ; $\psi_n(x)$ is non-decreasing as x increases for a fixed n .

Since $x = \varphi(\varepsilon_1 + \dots + \varphi(\varepsilon_n + T^n(x)))$
 $\psi_{2n}(x) \leq x \leq \psi_{2n+1}(x)$ and

$$\underline{\psi}(x) = \lim_{n \rightarrow \infty} \psi_{2n}(x) \leq x \leq \overline{\psi}(x) = \lim_{n \rightarrow \infty} \psi_{2n+1}(x)$$

Suppose $y = \psi_{2n}(x) < x < 1$
 then* $y = \varphi(\varepsilon_1 + \overline{\varphi}(\varepsilon_2 + \dots) - \delta_1) < x$ for some $\delta_1 > 0$

i.e. $\varepsilon_1 + \overline{\varphi}(\varepsilon_2 + \dots) - \delta_1 = \varphi^{-1}y > \varphi^{-1}x$

since $\overline{\varphi}(\varepsilon_2 + \dots) - \delta_1 < 1$

$\overline{\varphi}(\varepsilon_2 + \dots) > \overline{\varphi}(\varepsilon_2 + \dots) - \delta_1 = T y > T x$

and $\varepsilon_1(x) = \varepsilon_1(y)$

Therefore, for some $\delta_2 > 0$

$$\varphi(\varepsilon_1 + \overline{\varphi}(\dots) + \delta_2) > T y > T x$$

i.e. $\varepsilon_2 + \overline{\varphi}(\varepsilon_3 + \dots) + \delta_2 < \varphi^{-1}T y < \varphi^{-1}T x$

and $\overline{\varphi}(\varepsilon_3 + \dots) < T^2 y < T^2 x$

and $\varepsilon_2(x) = \varepsilon_2(y)$.

We proceed in the same manner and have $\varepsilon_i(x) = \varepsilon_i(y)$, $i \leq 2n$
 Therefore $\psi_m(\psi_n(x)) = \psi_m(x)$ for all $m \leq 2n$
 (this is trivially true for the case $T^i x = 0$ for some $i \leq 2n$.)

* By $\varphi(\varepsilon_1 + \overline{\varphi}(\varepsilon_2 + \dots))$ we mean $\lim_{n \rightarrow \infty} \varphi(\varepsilon_1 + \overline{\varphi}(\dots + \overline{\varphi}(\varepsilon_{2n})))$

Suppose $y = \psi_{2n+1}(x) > x$ (when $\psi^{-1}(0) = \infty$ we suppose $T_i x > 0$ for all $i \leq 2n+1$) then for some $\delta_1 > 0$

$$* \quad y = \psi(\varepsilon_1 + \bar{\psi}(\cdot) + \delta_1) > x$$

$$\text{i.e.} \quad \varepsilon_1 + \bar{\psi}(\varepsilon_2 + \dots) + \delta_1 = \psi^{-1}y < \psi^{-1}x$$

$$\text{and} \quad \bar{\psi}(\varepsilon_2 + \dots) + \delta_1 = T_1 y < T_1 x$$

$$\text{and } \varepsilon_1(x) = \varepsilon_1(y)$$

Again we continue in the same manner and have $\varepsilon_i(x) = \varepsilon_i(y)$

$$\text{for all } i \leq 2n+1. \quad \text{Consequently } \psi_m(\psi_{2n+1}(x)) = \psi_m(x) \quad \text{for all } m \leq 2n+1$$

(This is trivially true when $T_i x = 0$ for some $i \leq 2n+1$.)

Quite generally, therefore, for all x and for all

$$(2.7) \quad \begin{cases} \psi_m(\psi_{2n}(x)) = \psi_m(x) \\ \text{and} \end{cases}$$

and

$$\varepsilon_m(\psi_{2n}(x)) = \varepsilon_m(x)$$

$\psi_m(x)$ is non-decreasing as x increases for a fixed m therefore,

$$\psi_m(x) = \psi_m(\psi_{2n}(x)) \leq \psi_m(\psi(x)) \leq \psi_m(x)$$

Therefore

$$(2.8) \quad \psi_m(\psi(x)) = \psi_m(x)$$

$$\text{and } \varepsilon_m \psi = \varepsilon_m$$

and

$$(2.9) \quad \bar{\psi}(\psi(x)) = \bar{\psi}(x)$$

and

$$(2.10) \quad \psi(\psi(x)) = \psi(x)$$

Similarly $\psi_m(\bar{\psi}(x)) = \psi_m(x)$

$$(2.11)$$

$$\text{and } \varepsilon_m \bar{\psi} = \varepsilon_m$$

and

$$(2.12) \quad \psi(\bar{\psi}(x)) = \psi(x)$$

and

$$(2.13) \quad \bar{\psi}(\psi(x)) = \bar{\psi}(x)$$

)))--

*By $\psi(\varepsilon_i + \bar{\psi}(\varepsilon_{i+1} + \dots))$ we mean $\lim_{n \rightarrow \infty} \psi(\varepsilon_i + \bar{\psi}(\dots + \bar{\psi}(\varepsilon_{2n+1})))$

According to Lemma 8; if f satisfies A, T is a measurable transformation of $[0, 1)$ onto itself. In this case let $\mathcal{A}$ denote the partition $\{f(J_n)\}$ of $[0, 1)$. If f satisfies B, T is a measurable transformation of $[0, 1) - P$ onto itself, where P is the set of x such that $T^i x = 0$ for some i , when $f^{-1}(0) = \infty$ and $P = \emptyset$ when $f^{-1}(0) < \infty$. In this case let $\mathcal{A}$ denote the partition $\{f(J_n) - P\}$ of $[0, 1) - P$.

Recalling our definition of representation, $x \sim n_1 n_2 \dots$ means $x \in f(J_{n_1}), T x \in f(J_{n_2}), \dots$ when f satisfies A. and $x \in f(J_{n_1}) - P, T x \in f(J_{n_2}) - P, \dots$ when f satisfies B. In either case $x \sim n_1 n_2 \dots$ means

$$\begin{aligned} \varepsilon_1(x) &= n_1 \\ \varepsilon_2(x) &= n_2 \\ &\vdots \end{aligned}$$

THEOREM 8.

If f satisfies A, the representations are valid for all $x \in [0, 1)$ if and only if

$$x = \overline{f}(\varepsilon_1(x) + \overline{f}(\varepsilon_2(x) + \dots))$$

for every $x \in [0, 1)$.

If f satisfies B, the representations are valid for all $x \in [0, 1)$ such that $T^i x \neq 0$ for all i , if and only if

$$x = \overline{f}(\varepsilon_1(x) + \overline{f}(\dots))$$

for every $x \in [0, 1)$.

PROOF.

If f satisfies A for all n , $\varepsilon_n \psi(x) = \varepsilon_n(x)$.
(If f satisfies B for all n , $\varepsilon_n \overline{\psi} = \varepsilon_n = \varepsilon_n \underline{\psi}$.)

Therefore $\psi(x) = x$ ($\overline{\psi}(x) = x = \underline{\psi}(x)$) if and

only if the set of points with representation $\xi_1, \xi_2, \dots$ is of measure zero, i.e. if and only if representations are valid for all $x \in [0, 1)$ (all $x \in [0, 1) - P$.)

We can associate, as indicated above, with any f for which $x = \bar{f}(\xi_1 + \dots)$ for all x , a partition system whose representations are evidently valid.

E.

REPRESENTATION OF PARTITION SYSTEMS.

In C. we defined a partition system and showed that every partition system is point isomorphic with a thick subset of the unit interval together with a transformation. Corresponding to each of the two isomorphisms, derived from the direct ordering and the alternating ordering, there are two transformations of a thick subset of the unit interval onto itself. Each of these transformations were altered on a countable set and extended to the whole of $[0, 1)$. Moreover, these extensions were shown to be representable as $(f_1^{-1}x)$ and $(f_2^{-1}x)$ where f_1 and f_2 satisfied conditions A. and B. respectively.

$$\bar{f}_1(\xi_1(x) + \bar{f}_1(\xi_2(x) + \dots)) = x \quad \text{for all } x \in [0, 1)$$

if and only if $\bar{f}_1(\xi_1(x) + \bar{f}_1(\xi_2(x) + \dots)) = x$ for all x

belonging to $\phi(X)$, since $\psi(x)$ is non-decreasing.

$$\text{Similarly } \bar{f}_2(\xi_1 + \dots) = x \quad \text{for all } x \in [0, 1)$$

if and only if $\bar{f}_2(\xi_1 + \dots) = x$ for all $x \in \phi(X)$

By Theorem 8, therefore, and since the representations i

a partition system are always valid,

$$x = \bar{f}_1(\varepsilon_1 + \dots)$$

and

$$x = \bar{f}_2(\varepsilon_1 + \dots)$$

for all $x \in \phi(X)$.

Therefore

THEOREM 9.

i) A partition system (X, α, m, T) is isomorphic with a thick subset of unit interval together with the transformation $(\bar{f}_1^{-1}x)$ and for every $x \in \phi(X)$, where ϕ is the isomorphism, $x = \bar{f}_1(\varepsilon_1 + \bar{f}_1(\varepsilon_2 + \dots))$, where $\bar{f}_1$ is a strictly increasing continuous function with domain $\bigcup J_n$, $J_n = [a_n, b_n) \subset [n, n+1)$ and range $[0, 1)$ and $\bar{f}_1$ is the continuous monotone extension of $\bar{f}_1$ to the domain $[0, \infty]$ and where

$$x \sim \varepsilon_1, \varepsilon_2, \dots$$

ii) A partition system (X, α, m, T) is isomorphic with a thick subset of unit interval together with a transformation $(\bar{f}_2^{-1}x)$ and for every $x \in \phi(X)$, where ϕ is the isomorphism, $x = \bar{f}_2(\varepsilon_1 + \bar{f}_2(\varepsilon_2 + \dots))$ where $\bar{f}_2$ is a strictly decreasing continuous function with domain $\bigcup J_n$, $J_n = (a_n, b_n] \subset (n, n+1]$ and range $[0, 1)$ and $\bar{f}_2$ is a continuous monotone extension of $\bar{f}_2$ to the domain $[0, \infty]$ and where

$$x \sim \varepsilon_1, \varepsilon_2, \dots$$

The question arises as to whether the thick subset of $[0, 1)$ in Theorem 9 is in fact almost all of $[0, 1)$. Following the Proof of Halmos and Von-Neumann [18], who proved the existence of an unmeasurable thick subset of $[0, 1)$ with inner measure zero, invariant under all one-one measure

preserving transformation, we have:

THEOREM 10.

If T is a transformation of $[0, 1)$ onto itself, there exists a thick unmeasurable subset invariant under T with inner measure zero.

PROOF.

The set of Borel sets of positive measure may be well ordered, so that to each ordinal $\beta < \aleph_1$ (the first uncountable ordinal) there corresponds a Borel set of positive measure B_β . Let $O(x)$ be the orbit of x under T . If $y \in O(x)$, $O(y) = O(x)$ if $y \notin O(x)$, $O(x) \cap O(y) = \emptyset$.

By transfinite induction we define x_β, y_β . x_1 is chosen in B_1 , y_1 is chosen in B_1 but not in $O(x_1)$. The choice is possible since B_1 is an uncountable set and $O(x_1)$ is countable.

If x_β, y_β are defined for all $\beta < \gamma < \aleph_1$ we define x_γ as follows:

$\bigcup_{\beta < \gamma} (O(x_\beta) \cup O(y_\beta))$ is countable and we may choose $x_\gamma \in B_\gamma$ so that x_γ is not in this set. Add $O(x_\gamma)$ to this set and choose y_γ so that $y_\gamma \in B_\gamma$ but y_γ is not in the larger set.

Points x_β, y_β are now defined for all $\beta < \aleph_1$. It is clear that $O(x_\alpha)$ and $O(y_\beta)$ are disjoint for all α, β .

Define

$$A = \bigcup_{\alpha < \Omega} O(x_\alpha)$$

$$B = \bigcup_{\beta < \Omega} O(y_\beta)$$

A, B are disjoint and since $A \ni x_\beta$ and $B \ni y_\beta$ for every $\beta < \Omega$. A and B each have at least one point in common with every Borel set of positive measure. Therefore A^c and B^c each contain no Borel set of positive measure.

Therefore $\mu^* A^c = \mu^* B^c = 0$
 and $\mu_* A = \mu_* B = 1$
 and $\mu_* A = \mu_* B = 0$

(μ_* , μ^* Borel inner and outer measure respectively)

It is obvious that both A and B are invariant under T .

Combining theorem 9 and theorem 10 and the method of extension adopted in C. and D. we have:

THEOREM 11.

Every partition system isomorphic with an unmeasurable thick subset of unit interval together with a transformation of type $(f^{-1}x)$ can be embedded in a partition system isomorphic with the entire unit interval together with a transformation $(g^{-1}x)$ where g is an extension of f .

Every partition system isomorphic with the entire unit interval together with a transformation of type $(g^{-1}x)$ contains a partition system isomorphic with a thick unmeasurable subset of the unit interval together with a transformation $(f^{-1}x)$ where f is a restriction of g .

F.FULL PARTITION SYSTEMS.

Evidently to study partition systems is equivalent to studying representations, since the correspondence between points of the partition system and their representations is obviously an isomorphism.

Let " $<$ " denote the direct ordering and " $<$ " the alternating ordering. If $\mathcal{C}$ is a collection of sequences of integers we define: $\sup_{<} \{a_1, a_2, \dots \in \mathcal{C}\} = \alpha_1, \alpha_2, \dots$

where

$$\alpha_1 = \sup \{a_1 : a_1, a_2, \dots \in \mathcal{C}\}$$

$$\alpha_2 = \sup \{a_2 : \alpha_1, a_2, \dots \in \mathcal{C}\}$$

$$\alpha_3 = \sup \{a_3 : \alpha_1, \alpha_2, a_3, \dots \in \mathcal{C}\}$$

$$\inf_{<} \{a_1, a_2, \dots \in \mathcal{C}\} = \beta_1, \beta_2, \dots$$

where

$$\beta_1 = \inf \{a_1 : a_1, a_2, \dots \in \mathcal{C}\}$$

$$\beta_2 = \inf \{a_2 : \beta_1, a_2, \dots \in \mathcal{C}\}$$

$$\sup \{a_1, a_2, \dots \in \mathcal{C}\} = \sigma_1, \sigma_2, \dots$$

where

$$\sigma_1 = \inf \{a_1 : a_1, \dots \in \mathcal{C}\}$$

$$\sigma_2 = \sup \{a_2 : \sigma_1, a_2, \dots \in \mathcal{C}\}$$

$$\inf \{a_1, a_2, \dots \in \mathcal{C}\} = \tau_1, \tau_2, \dots$$

where

$$\tau_1 = \sup \{a_1 : a_1, \dots \in \mathcal{C}\}$$

$$\tau_2 = \inf \{a_2 : \tau_1, a_2, \dots \in \mathcal{C}\}$$

(N.B. ∞ may appear among the infima and suprema.)

If χ is a space of representations and (χ, α, m, T) is a partition system where T is the shift transformation and the partition is $\{a_n\}$ (where $A_n = \{x_1, x_2, \dots, x_n : x_i = a_i\}$) let

$$\gamma_1, \gamma_2, \dots = \inf \{a_1, a_2, \dots, x\}$$

$$\delta_1, \delta_2, \dots = \sup \{a_1, a_2, \dots, x\}$$

with respect to the order ' $<$ '. $a_1, a_2, \dots$ is said to be allowable with respect to $<$ if

$$\gamma_1, \gamma_2, \dots \leq a_1, a_2, \dots < \delta_1, \delta_2, \dots$$

for every n . χ is said to be full with respect to $<$ if χ contains all allowable sequences.

We have definitions of allowable and full with respect to ' $<$ ', if we replace $<$ by $\leq$ in the above definitions.

If ϕ is the isomorphism* of a full partition system χ into $[0, 1]$ and $x \in [0, 1]$ corresponds to no member of χ under ϕ , then

$$\beta_1, \beta_2, \dots = \sup \{a_1, a_2, \dots : \phi(a_1, a_2, \dots) < x\} \notin \chi$$

$$\alpha_1, \alpha_2, \dots = \inf \{a_1, a_2, \dots : \phi(a_1, a_2, \dots) > x\} \notin \chi$$

It follows easily from this that $T_m x = 0$ or 1 , for some m , where T is the extension of the isomorph of T to $[0, 1]$. Consequently ϕ sends χ one-one onto the whole unit interval with the exception of a countable set.

We suppose χ is full. If $n+1$ is the number of components of the partition generating α ;

$$\gamma_1 = 0, \delta_1 = n \quad \text{if } \underline{D.O.}$$

$$\gamma_1 = n, \delta_1 = 0 \quad \text{if } \underline{A.O.}$$

and $n \geq 1$. (It may happen that $n = \infty$.)

* i.e. derived from $<$ or $\leq$.

D.O.

If $\{A_m\}$ is the partition generating α where

$$A_m = \{a_1, a_2, \dots \in X : a_1 = m\}$$

and if $0 \leq m \leq n-1$ there is a sequence $a_{r+1}, a_{r+2}, \dots$ with r large enough so that

$$\gamma_1, \gamma_2, \dots < m, \delta_1, \dots, \delta_r, a_{r+1}, a_{r+2}, \dots < \delta_1, \delta_2, \dots$$

and $\delta_1, \dots, \delta_r, a_{r+1}, a_{r+2}, \dots \in X$

therefore $m, \delta_1, \dots, \delta_r, a_{r+1}, \dots \in X$ (by definition of $\delta_1, \delta_2, \dots$ and $\gamma_1, \gamma_2, \dots$.)

If $1 \leq m \leq n$ there is a sequence $b_{r+1}, b_{r+2}, \dots$ with r large enough so that

$$\gamma_1, \gamma_2, \dots < m, \gamma_1, \dots, \gamma_r, b_{r+1}, \dots < \delta_1, \delta_2, \dots$$

and $\gamma_1, \dots, b_{r+1}, \dots \in X$

therefore $m, \gamma_1, \dots, b_{r+1}, \dots \in X$

Inducing the partition structure of X onto $\phi(X) = [0, 1] - N$ where N is a countable set, and retaining without confusion the same notation, we see that: T sends A_0 one-one onto an interval with right end point 1 , (except for a countable set); T sends A_n one-one onto an interval with left end point 0 , (except for a countable set) and T sends A_m ($0 < m < n$) one-one onto $[0, 1]$ (except for a countable set).

Defining $\psi(x) = m + Tx$ for $x \in A_m$, ψ is a strictly increasing continuous function sending $[0, 1]$ onto $[T(0), n + T(1)]$ (except for a countable set).

A.O.

Similarly if X is full with respect to the alternating order and we induce on $[0, 1]$ the partition

structure of X by the isomorphism derived from the alternating order, then

$$\psi(x) = m + Tx$$

for $x \in A_m$

is a strictly decreasing continuous function sending $[0, 1)$ onto $(T(1), m + T(0)]$ (except for a countable set.)

In either case $f = \psi^{-1}$ is (except for a countable set) a strictly monotone function (decreasing or increasing) sending an interval $((a, b]$ or $[a, b)$ onto $[0, 1)$ (where $a < 1$ and $b - a \geq 1$.)

On the other hand if f is a strictly monotone function (decreasing or increasing) sending an interval $((a, b]$ or $[a, b)$ onto $[0, 1)$ (where $a < 1$ and $b - a \geq 1$) and such that

$$x = \overline{f}(\varepsilon_1(x) + \overline{f}(\varepsilon_2(x) + \dots)) \quad \text{for every } x$$

then $X = \{\varepsilon_1 \varepsilon_2 \dots\}$ is full.

The above statements are simple consequences of the definition of a full partition system and the continuity of f .

G.

TOPOLOGICAL TRANSITIVITY.

We consider in this section transformations T of $[0, 1)$ onto itself where $Tx = (f^{-1}x)$ and f is either
 a.) an increasing function sending $[a, b)$ onto $[0, 1)$ where $a < 1$ and $b - a \geq 1$, or
 b.) a decreasing function sending $(a, b]$ onto $[0, 1)$ where $a < 1$ and $b - a \geq 1$.

A study of such transformations amounts to a study of

full partition systems.

T is said to be topologically transitive when the smallest invariant set containing an open set is essentially $[0, 1)$ for ϕ .

I $b-a=1$ and T is measure preserving.

a.) ϕ increasing.

$$\begin{aligned} \text{Since } Tx &= \mu[0, Tx) = \mu[0, T(0)) + \mu[T(0), Tx) \\ &= \mu[0, T_0) + \mu[T_0, 1) + \mu[0, Tx) - 1 \quad \begin{array}{l} \text{if } T(0) \leq Tx \\ \text{otherwise} \end{array} \\ Tx &= \mu[0, T_0) + \mu T[0, x) \bmod 1 \\ &= \alpha + \mu[0, x) \bmod 1 \\ &= \alpha + x \bmod 1 \end{aligned}$$

where $\alpha = \mu[0, T_0)$

If α is rational there exists an interval I such that $T^n I$ lies entirely in A_0 or entirely in A_1 for all n . Therefore the representations are not valid and T is not topologically transitive. On the other hand if x, y have the same representation then $I = (x, y)$ has the property that $T^n I = I$ or $T^n I \cap I = \emptyset$ for all n and so α is rational. Moreover, it is well-known that if α is irrational then T is ergodic.

b.) ϕ decreasing.

We can show similarly that $Tx = 1 + \alpha - x \bmod 1$, and representations are valid if and only if T is ergodic and/or α is irrational. Consequently

THEOREM 12*

A full partition system generated by a one-one transformation is isomorphic with the unit interval together with an ergodic transformation $Tx = x + \alpha \bmod 1$ where α is irrational; and is isomorphic with the unit interval together with an ergodic transformation $Tx = 1/\beta - x$ where β is irrational.

II. $a = 0, b > 1$ in which case T is not one-one.

If $\mathcal{P}$ generates a partition system then by (2.4) and (2.7) of D., $x \sim \varepsilon_1, \varepsilon_2, \dots$ implies

$$\overline{\mathcal{P}}(x) = \overline{\mathcal{P}}(\varepsilon_1 + \overline{\mathcal{P}}(\varepsilon_2 + \dots))$$

and $\varepsilon_m(y_n) = \varepsilon_m(x)$ for all $m \leq n$ where

$$y_n = \overline{\mathcal{P}}(\varepsilon_1 + \dots + \overline{\mathcal{P}}(\varepsilon_n))$$

If $\mathcal{P}$ is increasing and $T^i x \neq 0$ for all i , and $x \sim \varepsilon_1, \varepsilon_2, \dots$

then for infinitely many n ,

$$\overline{\mathcal{P}}(\varepsilon_n + \dots + \overline{\mathcal{P}}(\varepsilon_{n+1})) < 1$$

for all $m \leq n$. For otherwise there exists N such that for all $n \geq N$ there is an $m \leq n$ and therefore a least $m \leq n$ with

$$\overline{\mathcal{P}}(\varepsilon_m + \dots + \overline{\mathcal{P}}(\varepsilon_{n+1})) = 1$$

Consequently

$$\overline{\mathcal{P}}(\varepsilon_r + \dots + \overline{\mathcal{P}}(\varepsilon_m + \dots + \overline{\mathcal{P}}(\varepsilon_{n+1}))) = \overline{\mathcal{P}}(\varepsilon_r + \dots + \overline{\mathcal{P}}(\varepsilon_{m+1})) < 1$$

*Actually without the assumption of fullness, such a partition system is isomorphic with a thick subset of the unit interval together with $Tx = x + \alpha \bmod 1$; and with a thick subset of the unit interval together with $Tx = 1/\beta - x \bmod 1$.

for $r \leq m-1$ and it follows that $m-1 \leq N$ in which case

$\lim_{n \rightarrow \infty} \bar{f}(\varepsilon_1 + \dots + \bar{f}(\varepsilon_n))$
has a finite expansion and so $T'x = 0$ for some i , which contradicts our hypothesis.

We have therefore infinitely many for which

$$z_n = \bar{f}(\varepsilon_1 + \dots + \bar{f}(\varepsilon_n + 1))$$

is defined, and clearly $\varepsilon_i(z_n) = \varepsilon_i(x)$

for all $i < n$.

If $\bar{f}$ is decreasing and $x \sim \varepsilon_1, \varepsilon_2, \dots$ then
 $\bar{f}(\varepsilon_1 + \dots + \bar{f}(\varepsilon_n)) \leq \bar{f}(\varepsilon_1 + \dots + \bar{f}(\varepsilon_n + 1)) \leq \bar{f}(\varepsilon_1 + \dots + \bar{f}(\varepsilon_{n+1}))$
 and $\bar{f}(\varepsilon_1 + \dots + \bar{f}(\varepsilon_{n+1})) \geq \bar{f}(\varepsilon_1 + \dots + \bar{f}(\varepsilon_n + 1)) \geq \bar{f}(\varepsilon_1 + \dots + \bar{f}(\varepsilon_n))$

therefore $\varepsilon_i(x) = \varepsilon_i(z_n)$ if $z_n = \bar{f}(\varepsilon_1 + \dots + \bar{f}(\varepsilon_n + 1))$
 $(i < n)$

LEMMA 9.

If $\bar{f}$ (increasing or decreasing) generates a partition system and $a=0, b>1$ (where the domain of $\bar{f}$ is either $[a, b)$ or $(a, b]$) and $x \sim \varepsilon_1, \varepsilon_2, \dots$ ($\varepsilon_n \in [b]$ for infinitely many n) and $T'x \neq 0$ for all i then for infinitely many n , $T^n I_n = (0, 1)$ where I_n is the interval with endpoints $\bar{f}(\varepsilon_1 + \dots + \bar{f}(\varepsilon_n))$, $\bar{f}(\varepsilon_1 + \dots + \bar{f}(\varepsilon_n + 1))$ which contains x .

PROOF.

We have shown that for infinitely many n , $\varepsilon_i(x) = \varepsilon_i(z_n)$ for $i < n$, if $z_n = \bar{f}(\varepsilon_1 + \dots + \bar{f}(\varepsilon_n + 1))$ and for all $i \leq n$, $\varepsilon_i(x) = \varepsilon_i(y_n)$ if $y_n = \bar{f}(\varepsilon_1 + \dots + \bar{f}(\varepsilon_n))$

$T^{n-1} I_n$ is therefore an interval with end points $\bar{f}(\varepsilon_n)$

and $\bar{f}(\xi_{n+1})$. If f is increasing we can write for these n , $\bar{f}(\xi_{n+1}) = f(\xi_{n+1})$

If f is decreasing for infinitely many n , $\bar{f}(\xi_{n+1})$ can be written $f(\xi_{n+1})$, for otherwise $\xi_n = [b]$ for all $n > N$ which contradicts the hypothesis. Therefore $T^n I_n = [0, 1)$

LEMMA 10.

If f (increasing or decreasing) generates a partition system and $a = 0, b > 1$ and U is open then $T^n U = [0, 1)$ for some n .

PROOF.

For all but a countable number of $x \in U$, an interval I_n with end points $\bar{f}(\xi_1 + \dots + \bar{f}(\xi_n))$, $\bar{f}(\xi_1 + \dots + \bar{f}(\xi_{n+1}))$, (where $\xi_n < [b]$) is contained in U .

THEOREM 13.

f - representations are valid if and only if T is topologically transitive.

PROOF.

By virtue of Lemma 10 it suffices to show that topological transitivity implies the validity of representations.

In fact if the representation of a point x is not valid, $(x \sim \xi_0 \xi_1 \dots)$

then

$$\mu(A_{\xi_0} \cap T^{-1} A_{\xi_1} \cap \dots) > 0$$

Let $J = A_{\varepsilon_0} \cap T^{-1}A_{\varepsilon_1} \cap \dots$

then $T^n J \subset A_{\varepsilon_n} \cap T A_{\varepsilon_{n+1}} \cap \dots$

and therefore either

i) $T^{n+1} J \subset J$ for some $n \geq 0$, or

ii) $T^{m+n} J \cap T^n J = T^n (J \cap T^m J) = \emptyset$

for all n and all $m \geq 0$.

i) If n is the least integer for which $T^{n+1} J \subset J$ and $T^{n+1} J = J$ then $J \cup \dots \cup T^n J$

is not the whole space, for if it were $A_0 = \bigcup T^{n_i} J$ for some sequence of n_i , $0 \leq n_i \leq n$ and therefore

$$T^{n+1} A_0 = [0, 1) = \bigcup T^{n_i+n+1} J = \bigcup T^{n_i} J = A_0$$

which is impossible, since a partition must have at least two components. Therefore if $J \cup \dots \cup T^n J = B$

and $T^{-1}B = B \setminus L$, T is not topologically transitive. On the other hand $T^{-1}B \supset B$ and $T^{-1}B - B \supset L$

where L is some interval. In this case $T^m L \cap T^{m+k} L = \emptyset$, $k \geq 0$ for all m . This case we consider with;

ii) When there exists an interval J such that $T^m J \cap T^{m+k} J = \emptyset$ for all m , $k \geq 0$.

Let U, V be two disjoint open intervals contained in J . If

$$G = U \cup T U \cup \dots$$

$$H = V \cup T V \cup \dots, \quad G \cap H = \emptyset$$

Evidently if

$$E = \bigcup_{n \geq 0} T^{-n} G$$

and

$$F = \bigcup_{n \geq 0} T^{-n} H$$

$T^{-1}E = E, T^{-1}F = F$. Should $x \in E \cap F$ there would exist an n such that $x \in T^{-n}G \cap T^{-n}H$
 $= T^{-n}(G \cap H)$
 $= \emptyset$

Consequently T is not topologically transitive.

COROLLARY 13.1.

If T is ergodic then f - representations are valid.

If f satisfies either A. or B. and $Tx = (f^{-1}x) \dagger$ if $P = \{x : T^i x = 0 \text{ some } i\}$
 when f satisfies B. and $f^{-1}0 = \emptyset$ and $P = \emptyset$
 otherwise, and if for all $x \in [0, 1) - P$
 $x = -\bar{P}(\epsilon_1 + \bar{P}(\epsilon_2 + \dots))$
 $[0, 1) - P$ together with T is a partition system which we say is generated by f .

Our next result is a counter example:

THEOREM 14.

There exists a partition system generated by a function f , which is strictly increasing and continuous and sends $[0, 2)$ onto $[0, 1)$, such that $Tx = (f^{-1}x)$ is not ergodic, (but which is topologically transitive).

PROOF.

We first show there exists a Borel set $B \subset [0, 1)$ such that $\mu B \cap I > 0$ and $\mu B^c \cap I > 0$ for every interval I , where μ is Lebesgue measure, and $\mu B = \frac{1}{2}$.

Let α'_0 be the length of an interval $I(\frac{1}{2})$ symmetric about $\frac{1}{2}$; α'_1 the length of each interval $I(\frac{1}{4}), I(\frac{3}{4})$ where $I(\frac{1}{4}), I(\frac{3}{4})$ are symmetric about centres of the intervals complementary to $I(\frac{1}{2})$; let α'_2 be the length that each interval $I(\frac{1}{8}), I(\frac{3}{8}), I(\frac{5}{8}), I(\frac{7}{8})$ where $I(\frac{1}{8}), I(\frac{3}{8}), I(\frac{5}{8}), I(\frac{7}{8})$ are symmetric about the centres of the intervals complementary to $I(\frac{1}{4}) \cup I(\frac{1}{2}) \cup I(\frac{3}{4})$; and so on.

The union of these intervals is a Borel set B_0 with measure $\alpha'_0 + 2\alpha'_1 + 4\alpha'_2 + \dots$

We proceed again in the same manner with each of the intervals $I(\frac{n}{2^m})$ replacing $(0, 1)$, the proportions, instead of $\alpha'_0, \alpha'_1, \alpha'_2, \dots$, this time being $\alpha_0^2, \alpha_1^2, \alpha_2^2, \dots$.

If n is odd, the union of these sub-intervals of $I(\frac{n}{2^m})$ is a Borel set of measure $\alpha_{n-1}' (\alpha_0^2 + 2\alpha_1^2 + 4\alpha_2^2 + \dots)$ and consequently the total union taken over all the sub-intervals of all intervals $I(\frac{n}{2^m})$ is a Borel set B_1 of measure $(\alpha'_0 + 2\alpha'_1 + \dots)(\alpha_0^2 + 2\alpha_1^2 + \dots)$

We label these sub-intervals in an analogous manner by $I(\frac{n_1}{2^{m_1}}, \frac{n_2}{2^{m_2}})$.

Proceeding again as before with proportions
with sub-intervals $I(\frac{n_1}{2^{m_1}}, \frac{n_2}{2^{m_2}}, \frac{n_3}{2^{m_3}})$ of $I(\frac{n_1}{2^{m_1}}, \frac{n_2}{2^{m_2}})$

we arrive at a Borel set B_2 of measure

$$(\alpha_0' + 2\alpha_1' + \dots)(\alpha_0'^2 + 2\alpha_1'^2 + \dots)(\alpha_0'^3 + 2\alpha_1'^3 + \dots)$$

Inductively we arrive at a sequence of Borel sets B_n
with $B_{n+1} \subset B_n$ and

$$\mu B_n = (\alpha_0' + 2\alpha_1' + \dots) \dots (\alpha_0'^{n+1} + 2\alpha_1'^{n+1} + \dots)$$

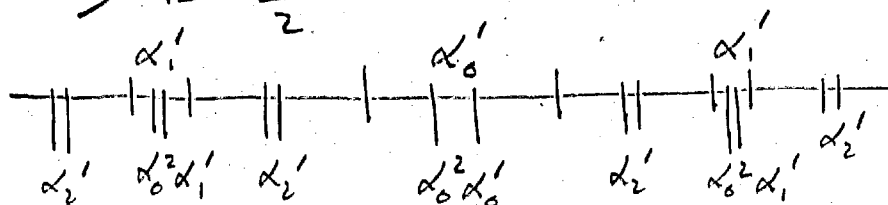
Let $B = \bigcap_n B_n$
then $\mu B = p_1 p_2 \dots$ where $p_i = (\alpha_0'^i + 2\alpha_1'^i + \dots)$

Let $\alpha_j^i = (\alpha_0^i)^{j+1}$ then $p_i = \alpha_0^i / (1 - 2\alpha_0^i)$

Let $\alpha_0^i = 1 / (e^{(\frac{1}{2^{i-1}} - \frac{1}{2^i})} + 2)$ then $p_i = e^{-(\frac{1}{2^{i-1}} - \frac{1}{2^i})}$

and $p_1 p_2 \dots = e^{-(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots)} = e^{-\log 2} = \frac{1}{2}$

and $\mu B = \frac{1}{2}$



Let $\mathcal{C}_q$ be the class of sub-intervals $I(\frac{n_1}{2^{m_1}}, \dots, \frac{n_{q+1}}{2^{m_{q+1}}})$.

Choose an interval J arbitrarily. If $\mu(I \cap J) > 0$
and $I \in \mathcal{C}_q$ there is a sub-interval of I , say $K \in \mathcal{C}_{q+1}$
such that $J \supset K$, and if the length of K is l ,

$$l > \mu B \cap K = l p_{q+2} p_{q+3} \dots > 0$$

Therefore B and B^c intersect K in sets of positive measure and consequently B and B^c intersect J in sets of positive measure.

By the construction, $\mu_{J \cap I} > 0$ for some $I \in \mathcal{C}_q$.
Therefore B and B^c intersect each interval in sets of positive measure.

$$\text{Let } B_1 = \{ \tfrac{1}{2}x : x \in B \}, B_2 = \{ x + \tfrac{1}{2} : x \in B_1 \}$$

then $B_1 \cup B_2$ and its complement $(B_1 \cup B_2)^c$ intersect every interval in a set of positive measure. Moreover

$$\mu B_1 = \mu \{ [0, \tfrac{1}{2}) - B_1 \} = \mu B_2 = \mu \{ [\tfrac{1}{2}, 1) - B_2 \} = 1/4$$

$\phi(x) = \mu [0, x) \cap (B_1 \cup B_2)$ sends $B_1 \cup B_2$ onto almost all of $[0, \tfrac{1}{2})$.

Let N be the zero subset not in the range of ϕ and let $O(N)$ be the smallest $\mathcal{S}$ invariant set containing N , where $Sx = 2x \bmod \tfrac{1}{2}$

It is well known that $O(N)$ is a set of measure zero and $I = [0, \tfrac{1}{2}) - O(N)$ is S invariant, and is contained in the range of ϕ , and $\mu M = 0$ where $M = (B_1 \cup B_2 - \phi^{-1}I)$ since ϕ preserves measure by the isomorphism theorems.

If $C_1 = B_1 - M$ and $C_2 = B_2 - M$ then $\phi'(x) = \mu [0, x) \cap (C_1 \cup C_2)$

has the same value on $C_1 \cup C_2$ as ϕ and ϕ sends $C_1 \cup C_2$

one-one onto I , ϕ' sends C_1 one-one onto $[0, \frac{1}{2}) \cap I$,
and ϕ sends C_2 one-one onto $[\frac{1}{4}, \frac{1}{2}) \cap I$.

Define $Tx = \phi'^{-1} \circ S \circ \phi'(x)$
for $x \in C_1 \cup C_2$ then T sends C_1 onto $C_1 \cup C_2$ and sends
 C_2 onto $C_1 \cup C_2$. Moreover if $x, y \in C_1$ and $x < y$
 $Tx < Ty$ and if $x, y \in C_2$ and $x < y$, $Tx < Ty$. If
 $E \subset C_1$ is measurable

$$\mu TE = 2\mu E$$

and if $E \subset C_2$ is measurable $\mu TE = 2\mu E$.

It is easy to extend the definition of T to points of
 $[0, 1) - C_1 \cup C_2$ so that the extended T is continuous from
the left and from the right in each point of $[0, \frac{1}{2})$, $[\frac{1}{2}, 1)$.

$$\text{Let } \psi(x) = \begin{cases} Tx & \text{if } x \in [0, \frac{1}{2}) \\ 1 + Tx & \text{if } x \in [\frac{1}{2}, 1) \end{cases}$$

$\psi(x) = \psi^{-1}(x)$ is a strictly increasing continuous
function (since necessarily $T(\frac{1}{2}) = T(0) = 0$) and sends
 $[0, 2)$ onto $[0, 1)$ and $Tx = (\psi^{-1}x)$.

If U is open and invariant then U contains the
smallest invariant set containing $U \cap (C_1 \cup C_2)$ which is
 $C_1 \cup C_2$ by virtue of the ergodicity of $Sx = 2x \bmod \frac{1}{2}$
on $[0, \frac{1}{2})$ and U contains the smallest invariant set
containing $U \cap (C_1 \cup C_2)^c$ (which is $(C_1 \cup C_2)^c$) for the
same reason. Therefore $U = [0, 1)$ and T is
topologically transitive. T therefore generates a
partition system. On the other hand T is not ergodic
since $C_1 \cup C_2$ is invariant and $\mu(C_1 \cup C_2) = \mu(B_1 \cup B_2)$
 $= \frac{1}{2}$

CHAPTER 3. FREQUENCY, ERGODICITY, AND INVARIANCE.

A.

FREQUENCY.

If $\varepsilon_1 \varepsilon_2 \dots$ is the representation of x in a partition system generated by a function f satisfying A. or B. $f(\varepsilon_1 + f(\varepsilon_2 + \dots))$ is called the f -expansion of x .

It is important to find a measure invariant under T and equivalent to Lebesgue measure, (if one exists) so that the frequency of digits in f -expansions may be calculated.

Let $\phi_n^m(x)$ be the number of m 's appearing in the first n places of the f -expansion of $x \in [0,1) - P$. We define $\phi^m(x) = \lim_{n \rightarrow \infty} \frac{\phi_n^m(x)}{n}$ (if it exists)

and the frequency of m as the f -expansion of x .

If the representation of x is $x \sim \varepsilon_1 \varepsilon_2 \dots$ let $\chi_m(x) = \begin{cases} 1 & \text{if } \varepsilon_0(x) = m \text{ (i.e. } \chi_m(x) = \chi(x) \text{)} \\ 0 & \text{otherwise} \end{cases}$ $\{\varepsilon_0 = m\}$

then $\chi_m(T^n x) = 1$ if and only if $\varepsilon_n(x) = m$

Therefore $\phi_n^m(x) = \sum_{i=0}^{n-1} \chi_m(T^i x)$

and $\phi^m(x) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \chi_m(T^i x)$

which exists almost everywhere if there exists a finite invariant measure μ equivalent to Lebesgue measure, by virtue of Theorem B.

Furthermore if μ is normalised and T is ergodic
 $\phi^m(x) = \int \chi_m(x) d\mu$ almost everywhere.

If T is topologically transitive (which happens when the domain of f is of the form $[0, t)$, $t > 1$ when f is increasing, and when the domain of f is of the form $(0, t]$, $t > 1$, when f is decreasing),

$$\sum_{i=0}^{\infty} \chi_m(T^i x) = \infty \quad \text{almost everywhere, (if } \mu(A) > 0 \text{)}$$

since if we put $E_n = \bigcap_{k=1}^n \bigcup_{j=k}^{\infty} T^{-j} A = \bigcup_{j=n}^{\infty} T^{-j} A$

therefore $\mu(E_n) = 1$ and $E_n \downarrow E = \bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} T^{-j} A$ (Lebesgue measure) and

This latter set is the set of points which belong to infinitely many $T^{-j} A$. Consequently for almost all $x \in [0, 1)$:

$$\sum_{j=0}^{\infty} \chi_{T^{-j} A}(x) = \infty \quad \text{i.e.} \quad \sum_{i=0}^{\infty} \chi_m(T^i x) = \infty$$

The relative frequency of digits m to n is defined as
 $\lim_{k \rightarrow \infty} \frac{\phi_k^m(x)}{\phi_k^n(x)}$ if it exists.

By Theorem H, if μ is a σ -finite invariant measure equivalent to Lebesgue measure and ψ_n is μ integrable,

$$\lim_{k \rightarrow \infty} \frac{\sum_{i=0}^{k-1} \psi_n(T^i x)}{\sum_{i=0}^{\infty} \psi_n(T^i x)} = \phi_{\mu, \psi_n}(x) \quad \text{exists almost everywhere, when}$$

$\sum_{i=0}^{\infty} \psi_n(T^i x)$ almost everywhere which happens when the domain of f is $[0, t)$ or $(0, t]$, $t > 1$.

Moreover if $\psi_n(x)$ is μ integrable

$$\int \psi_n(x) d\mu = \int \psi_n(x) \phi_{\mu, \psi_n}(x) d\mu$$

In the ergodic case, $\phi_{\mu, \psi_n}(x)$ is constant almost everywhere and therefore

$$\phi_{\mu, \psi_n}(x) = \int \psi_n(x) d\mu / \int \psi_n(x) d\mu \quad \text{almost everywhere.}$$

B.

ERGODICITY.

We remark that there is no mere generality in assuming f to have domain $[n, n+t)$ or $(n, n+t]$, $t > 1$, than in assuming f to have domain $[0, t)$ or $(0, t]$, $t > 1$, since the difference is only in the choice of alphabet.

A. Renyi [24] has formulated conditions for the existence of f -expansions the existence of finite measures invariant under $Tx = (f^{-1}x)$

and equivalent to Lebesgue measure and for the ergodicity of $T^t x = (\varphi^{-t} x)$, when φ has domain $[0, t)$, $t > 1$ when φ is increasing and domain $(1, 1+t]$ when φ is decreasing. His results generalise those of Bissenger [25] (who studied the case where φ is decreasing) and Everett [26] (who studied the case where φ is increasing).

With the exception of a particular class of φ ($\varphi(x) = x/\beta$ for $x < \beta$ which we will deal with later) Renyi considered the ergodic and invariance problems only for the cases where t is an integer or infinity. In these cases he defined

$$H_n(x, t) = \frac{d}{dt} (\varphi(\varepsilon_1(x) + t) + \dots + \varphi(\varepsilon_n(x) + t)) \quad (0 < t < 1)$$

and imposed the condition:

C. There exists a content C independent of x and n such that

$$\frac{\max_{0 < t < 1} |H_n(x, t)|}{\min_{0 < t < 1} |H_n(x, t)|} < C \quad \text{for all } x, n$$

This condition, however, implies the existence of a finite invariant measure equivalent to Lebesgue measure, and the ergodicity of T_t simultaneously.

In the following we generalise C, to the case where t is not necessarily an integer or infinity, and at the same time separate the problem of invariance from that of ergodicity.

We assume $\mathcal{G}$ to generate a partition process.

From Lemma 9, except for a countable set N of x , for infinitely many n the interval $I_n(x)$ with end parts $\bar{p}(\varepsilon_1(x) + \dots + \bar{p}(\varepsilon_n))$, $\bar{p}(\varepsilon_1 + \dots + \bar{p}(\varepsilon_{n+1}))$ has the property that $T^{-n} I_n(x) = [0, 1) \cap x \in I_n(x)$. Let $\mathcal{G}$ be the class of such intervals $I_n(x)$. Clearly $\mathcal{G}$ generates the class of Borel sets and for all $I_n(x) \in \mathcal{G}$ there is an essentially unique function $\omega_x^n(y)$ (by Theorem R.M.) such that

$$\mu T^{-n} A \cap I_n(x) = \int_A \omega_x^n(y) d\mu$$

for all Borel sets A .

In the case considered by Renyi we note that

$$\mu T^{-n} A \cap I_n(x) = \int_A \left| \frac{d(p(\varepsilon_1 + \dots + p(\varepsilon_n + t)))}{dt} \right| dt$$

Renyi's condition C therefore amounts to

$$\underline{C}. \quad \frac{\text{Max}_y |\omega_x^n(y)|}{\text{Min}_y |\omega_x^n(y)|} < C$$

Evidently C implies

$$\frac{\text{Min}_y |\omega_x^n(y)|}{\mu I_n(x)} > \frac{1}{C}$$

Since $\mu I_n(x) = \int \omega_x^n(y) d\mu \leq \text{Max}_y \omega_x^n(y)$.

We replace this latter condition by a weaker one;

$$\underline{C}'. \quad \inf_{x \notin N} \frac{\overline{\lim}_n \omega_x^n(y)}{\mu I_n(x)} > 0$$

THEOREM 15.

If $Tx = (\varphi^{-1}x)$ and φ satisfies C',
then T is ergodic.

PROOF.

$$\underline{C'} \text{ implies for all } x \notin N \quad \lim_n \frac{\omega_x^n(y)}{\mu I_n(x)} \geq h(y) \quad \text{where } h(y)$$

is positive almost everywhere.

For each $x \notin N$, $\frac{\omega_x^n(y)}{\mu I_n(x)} \geq h(y)$ for infinitely many n .

If E is invariant and $\varepsilon > 0$, then E^c may be covered by a countable union U of disjoint sets $I_n(x) \notin \mathcal{F}$ satisfying

$$\frac{\omega_x^n(y)}{\mu I_n(x)} \geq h(y) \text{ and so that } \mu(U - E^c) < \varepsilon$$

Since

$$\begin{aligned} \mu E \cap I_n(x) &= \mu T^{-n} E \cap I_n(x) \\ &= \int_E \omega_x^n(y) d\mu \\ &\geq \int_E h(y) d\mu \cdot \mu I_n(x). \end{aligned}$$

we have

$$\varepsilon > \mu E \cap U \geq \int_E h(y) d\mu \cdot \mu U.$$

Therefore either $\mu E = 0$ or $\mu E^c = 0$.

REMARK : The same method, which imitates Renyi's use of a Theorem of Knopp [27], can be used to show that $\overline{T \times T}$ is ergodic and hence that $\overline{T}$ is weakly mixing [3].

If we strengthen $\underline{C}'$ but weaken $\underline{C}$ so that for all $x \notin N$ and all n

$$\frac{\underline{C}''}{\mu I_n(x)} \omega_n^h(y) > h(y) \text{ where } h(y) \text{ is positive almost everywhere,}$$

then :

THEOREM 16

$\underline{C}''$ implies the existence of an invariant measure equivalent to Lebesgue measure.

PROOF

For every Borel A with $\mu A > 0$

$$\int_A \omega_n^h(y) d\mu \geq \int_A h(y) d\mu \cdot \mu I_n(x)$$

$$\text{i.e. } \mu T^{-n}A \cap I_n(x) \geq \int_A h(y) d\mu \cdot \mu I_n(x),$$

and summing over disjoint $I_n(x)$ for a fixed n we have

$$\mu T^{-n}A \geq \int_A h(y) d\mu$$

and therefore

$$\liminf \mu T^{-n}A > 0$$

and by Theorem C, there exists a finite invariant measure equivalent to Lebesgue measure.

C.INVARIANCE

Let f be a function satisfying A. or B.
 The partition $\{f(J_n)\}$ of $[0, 1)$ has the property that
 T is one-one on each $f(J_n)$. We can write
 $X_n = f(J_n)$ and use Theorem 2. It is easily seen
 that for every n ,

$$h(T_n x) \omega_n(x) = h(f(x+n)) |f'(x+n)|$$

(see Theorem 2 for notation.) As a corollary to
 Theorem 2 we have :

THEOREM 17

If ν is a measure equivalent to Lebesgue
 measure then

$$\nu E = \int_E h(x) dx \text{ and } \nu T^{-1}E = \nu E$$

for all Lebesgue sets E , if and only if

$$(3.1) \quad h(x) = \sum_{0 < f(x+n) < 1} h(f(x+n)) |f'(x+n)| \cdot [n].$$

It is easy to see that $h(x) = 1$ satisfies (3.1) when
 $f(x) = x - \alpha, \alpha \leq x < 1 + \alpha$ or $f(x) = 1 + \alpha - x, \alpha < x \leq 1 + \alpha$

i.e. when $Tx = x + \alpha \bmod 1$ or $Tx = 1 + \alpha - x \bmod 1$.

In the remainder of this section we find solutions
 of the equation (3.1) for the following four cases :

$$\text{I.} \quad f(x) = \frac{1}{x+1}$$

$$\text{II.} \quad f(x) = 1 - \frac{1}{x+1}$$

$$\text{III.} \quad f(x) = \frac{x}{\beta} \quad \text{for } x < \beta$$

$$\text{IV.} \quad f(x) = 1 - \frac{x}{\beta} \quad \text{for } x \leq \beta.$$

$$\text{I.} \quad f(x) = \frac{1}{x+1} \quad \text{Continued fractions.}$$

The representation of real numbers by f -expansions ($f(x) = \frac{1}{x+1}$, $x > 0$) are the continued fraction representations, previously studied by P. Lévy [28], [29], [30], A. Khintchine [31], [32], [33], C. Ryll-Nardzewsky [34], [35], E. Marczewski [35], and S. Hartman [35], [36].

An invariant measure equivalent to Lebesgue measure is yielded by any positive function h satisfying,

$$h(x) = \sum_{n=0}^{\infty} h\left(\frac{1}{x+n+1}\right) \frac{1}{(x+n+1)^2}$$

$$\text{If } h(x) = \frac{1}{x+1},$$

$$\sum_{n=0}^{\infty} h\left(\frac{1}{x+n+1}\right) \frac{1}{(x+n+1)^2} = \sum_{n=0}^{\infty} \left(\frac{1}{x+n+1} + 1\right) \frac{1}{(x+n+1)^2}$$

$$= \sum_{n=0}^{\infty} \left(\frac{1}{x+n+1} - \frac{1}{x+n+2}\right) = \frac{1}{x+1} = h(x).$$

Consequently $\nu E = \int_E \frac{1}{x+1} d\mu / \log 2$.

is a normalised invariant measure equivalent to Lebesgue measure. The ergodicity of $Tx = (\varphi^{-1}x)$ has been proved by various authors and therefore ν is (to within a constant factor of multiplication) the only σ -finite invariant measure equivalent to Lebesgue measure.

II. $\rho(x) = 1 - \frac{1}{x+1}$ Alternating Continued Fractions

We consider φ -expansions (where $\varphi(x) = 1 - \frac{1}{x+1}$ for $x \geq 0$) of real numbers x in the form :

$$x = 1 - \frac{1}{\varepsilon_1 + 2 - \frac{1}{\varepsilon_2 + 2 - \frac{1}{\varepsilon_3 + \dots}}}$$

where $\varepsilon_1(x) = [\varphi^{-1}x] = \left[\frac{x}{1-x} \right]$
 $\varepsilon_2(x) = [\varphi^{-1}(\varphi^{-1}x)]$,

etc .

Any invariant measure equivalent to Lebesgue measure will be yielded by positive solutions of ,

$$(3.2.) \quad h(x) = \sum_{n=0}^{\infty} h\left(1 - \frac{1}{x+n+1}\right) \left(\frac{1}{x+n+1}\right)^2$$

Let $h(x) = \frac{1}{x}$ then

$$\begin{aligned} \sum_{n=0}^{\infty} h\left(1 - \frac{1}{x+n+1}\right) \left(\frac{1}{x+n+1}\right)^2 &= \sum_{n=0}^{\infty} \frac{1}{(x+n)(x+n+1)} \\ &= \sum_{n=0}^{\infty} \left(\frac{1}{x+n} - \frac{1}{x+n+1} \right) = \frac{1}{x} = h(x). \end{aligned}$$

Since $\int_0^1 h(x) dx = \infty$ and $h(x) < \infty$ for $x > 0$
 $\nu E = \int_E \frac{1}{x} dx$ is a σ -finite infinite

invariant measure equivalent to Lebesgue measure. Moreover
 ν is (to within a constant factor of multiplication)
the only σ -finite invariant measure equivalent to
Lebesgue measure and $h(x) = 1/x$ is an essentially
unique positive solution of equation (3.2) as we will
show by proving that T is ergodic.

If $x \in [0, 1)$ let P_n, Q_n be the polynomials
in $\varepsilon_1 \dots \varepsilon_n$ such that,

$$P_n / Q_n = 1 - \frac{1}{\varepsilon_1 + 2 - \frac{1}{\varepsilon_2 + \dots - \frac{1}{\varepsilon_n + 1}}}$$

$$\frac{P_1}{Q_1} = 1 - \frac{1}{\varepsilon_1 + 1} = \frac{\varepsilon_1}{\varepsilon_1 + 1}$$

Suppose
$$\frac{P_n}{Q_n} = \frac{P_{n-1} + \varepsilon_n(1 + \dots + P_{n-1})}{Q_{n-1} + \varepsilon_n(1 + \dots + Q_{n-1})}$$

then

$$\begin{aligned} \frac{P_{n+1}}{Q_{n+1}} &= \frac{P_{n-1} + (\varepsilon_n + 1 - \frac{1}{\varepsilon_{n+1} + 1})(1 + \dots + P_{n-1})}{Q_{n-1} + (\varepsilon_n + 1 - \frac{1}{\varepsilon_{n+1} + 1})(1 + \dots + Q_{n-1})} \\ &= \frac{P_n + \varepsilon_{n+1}(1 + \dots + P_n)}{Q_n + \varepsilon_{n+1}(1 + \dots + Q_n)} \end{aligned}$$

By the principle of induction

$$P_n = P_{n-1} + \varepsilon_n (1 + \dots + P_{n-1})$$

$$Q_n = Q_{n-1} + \varepsilon_n (1 + \dots + Q_{n-1})$$

for all integers n .

$$\begin{aligned} f(\varepsilon_1 + \dots + f(\varepsilon_n + t)) &= \frac{P_{n-1} + (\varepsilon_n + t)(1 + \dots + P_{n-1})}{Q_{n-1} + (\varepsilon_n + t)(1 + \dots + Q_{n-1})} \\ &= \frac{P_n + t(1 + \dots + P_{n-1})}{Q_n + t(1 + \dots + Q_{n-1})} \end{aligned}$$

Therefore $H_n(x, t) = \frac{Q_n(1 + \dots + P_{n-1}) - P_n(1 + \dots + Q_{n-1})}{[Q_n + t(1 + \dots + Q_{n-1})]^2}$

and

$$\begin{aligned} \min_{0 < t < 1} H_n(x, t) &= \frac{Q_n(1 + \dots + P_{n-1}) - P_n(1 + \dots + Q_{n-1})}{(1 + \dots + Q_{n-1} + Q_n)^2} \\ &= \frac{Q_n(1 + \dots + P_n) - P_n(1 + \dots + Q_n)}{(1 + \dots + Q_n)^2} \\ &= \frac{Q_n}{1 + \dots + Q_n} \left(\frac{1 + \dots + P_n}{1 + \dots + Q_n} - \frac{P_n}{Q_n} \right) \\ &= \frac{Q_n}{1 + \dots + Q_n} \sim I_n(x) \end{aligned}$$

where $I_n(x)$ is the interval $[f(\varepsilon_1 + \dots + f(\varepsilon_n)), f(\varepsilon_1 + \dots + f(\varepsilon_n + 1))]$

Similarly $\max_{0 < t < 1} H_n(x, t) = \frac{1 + \dots + Q_n}{Q_n} \sim I_n(x).$

Therefore $\frac{\max H_n(x, t)}{\min H_n(x, t)} = \left(\frac{1 + \dots + Q_n}{Q_n} \right)^2$

There exists x such that $\left(\frac{1 + \dots + Q_n}{Q_n} \right)^2$

is arbitrarily large, for large n and therefore C is not satisfied.

In fact $Q_n = Q_{n-1} + \varepsilon_n (1 + \dots + Q_{n-1})$

and if $\varepsilon_n = 0$ then $Q_n = Q_{n-1}$

Suppose $\varepsilon_n = \varepsilon_{n-1} = \dots = \varepsilon_1 = 0$,

since

$P_1 = \varepsilon_1, Q_1 = \varepsilon_1 + 1; Q_n = Q_{n-1} = \dots = Q_1 = 1$

i.e. $\left(\frac{1 + \dots + Q_n}{Q_n} \right)^2 = (n+1)^2$

However, $\underline{C}'$ is satisfied since

$$\frac{\text{Max } H_n(x, t)}{\text{Min } H_n(x, t)} = \left(\frac{1 + \dots + Q_n}{Q_n} \right)^2 = \left(\frac{Q_n - Q_{n-1} + 1}{\varepsilon_n Q_n} \right)^2$$

if $\varepsilon_n \neq 0$

$$\leq \left(\frac{1}{\varepsilon_n} + 1 \right)^2 = \left(\frac{\varepsilon_n + 1}{\varepsilon_n} \right)^2 \leq 2^2 = 4$$

i.e. $\inf_{n \in \mathbb{N}} \frac{H_n(x, t)}{I_n(x)} \geq \frac{1}{4}$ if $\varepsilon_n \neq 0$.

For each irrational $x \in [0, 1)$ this is so for an infinity of n . Therefore $\lim_{n \rightarrow \infty} \frac{H_n(x, t)}{I_n(x)} \geq \frac{1}{4}$ if x is irrational and $\underline{C}'$ is satisfied i.e. T is ergodic.

The alternating fraction expansion has been noted previously by A. Renyi [37.], who proved that the relative frequency of m to n is

$$\frac{\log \frac{(m+1)^2}{m(m+2)}}{-\log \frac{(n+1)^2}{n(n+2)}} \quad \text{if } m, n > 0$$

and the relative frequency of 0 to m for $m > 0$ is infinite. These results are obtained by a simple application of Theorem Ho.

III.

 β -expansions.

$$f(x) = x/\beta \quad \text{for } x < \beta > 1$$

When $\beta = q$ is an integer, β -expansions are the well known q -adic expansions or "decimal" expansions. [38].

$$f(\varepsilon_1 + \dots + f(\varepsilon_n + t)) = \frac{\varepsilon_1}{\beta} + \dots + \frac{\varepsilon_n + t}{\beta^n}$$

therefore $H_n(x, t) = 1/\beta^n$

and $\frac{\max H_n(x, t)}{\min H_n(x, t)} = 1$ and consequently T

is ergodic. There is, therefore, (to within a constant factor of multiplication) at most one invariant measure equivalent to Lebesgue measure and essentially at most one positive solution $h(x)$ of

$$(3.3.) \quad h(x) = \sum_{0 \leq \frac{x+y}{\beta} < 1} h\left(\frac{x+y}{\beta}\right) \frac{1}{\beta} \equiv \sum_{y=0}^{\lfloor \beta-x \rfloor} h\left(\frac{x+y}{\beta}\right) \frac{1}{\beta}$$

We show that * $h(x) = \sum_{x < T^n(1)} \frac{1}{\beta^n}$

(where $T^0(1) = 1$, $T(1) = (\beta)$) satisfies (3.3).

In fact, more generally :

THEOREM 17

If the β -expansion of $\beta > 1$ is
 $\beta = \varepsilon_0 + \frac{\varepsilon_1}{\beta} + \dots$ $\varepsilon_0 = \lfloor \beta \rfloor$
 and α is any complex solution of

* This is a Saltus function. For properties see [2].

$$z = \varepsilon_0 + \frac{\varepsilon_1}{\alpha} + \dots$$

$$\text{and } h(x) = \sum_{x < T^n(1)} \frac{1}{\alpha^n}$$

$$\text{then } \sum_{n=0}^{\lfloor \beta-x \rfloor} h\left(\frac{x+n}{\beta}\right) \frac{1}{\alpha} = h(x).$$

PROOF

$$h(x) = \sum_{x < T^n(1)} \frac{1}{\alpha^n} = \sum_{n=0}^{\infty} \frac{a_n(x)}{\alpha^n}$$

$$\text{where } a_n(x) = \begin{cases} 1 & \text{if } x < T^n(1) \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Let } a_{n,m}(x) = \begin{cases} 1 & \text{if } \frac{x+n}{\beta} < T^m(1) \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} \sum_{n=0}^{\lfloor \beta-x \rfloor} h\left(\frac{x+n}{\beta}\right) \frac{1}{\alpha} &= \sum_{n=0}^{\lfloor \beta-x \rfloor} \sum_{m=0}^{\infty} \frac{a_{n,m}(x)}{\alpha^m} \frac{1}{\alpha} \\ &= \sum_{m=0}^{\infty} \frac{1}{\alpha^{m+1}} \left(\sum_{n=0}^{\lfloor \beta-x \rfloor} a_{n,m}(x) \right) \end{aligned}$$

$$\sum_{n=0}^{\lfloor \beta-x \rfloor} a_{n,m}(x) \quad \text{is the number of } n \text{ such that } \frac{x+n}{\beta} < T^m(1) \quad (\text{i.e. such that } n < \beta T^m(1) - x)$$

$$\text{and this number is } \begin{cases} 1 + \lfloor \beta T^m(1) \rfloor & \text{if } x < (\beta T^m(1)) = T^{m+1}(1) \\ \lfloor \beta T^m(1) \rfloor & \text{otherwise} \end{cases}$$

$$\text{Therefore } \sum_{n=0}^{\lfloor \beta-x \rfloor} a_{n,m}(x) = a_{m+1}(x) + \varepsilon_m$$

$$\text{since } \varepsilon_m = \lfloor \beta T^m(1) \rfloor$$

$$\text{and } \sum_{n=0}^{\lfloor \beta-x \rfloor} h\left(\frac{x+n}{\beta}\right) \frac{1}{\alpha} = \sum_{m=0}^{\infty} \frac{a_{m+1}(x) + \varepsilon_m}{\alpha^{m+1}}$$

$$= h(x) - a_0 + \frac{1}{\alpha} \sum_{m=0}^{\infty} \frac{\varepsilon_m}{\alpha^m}$$

$$= h(x) - a_0 + \alpha/\alpha = h(x)$$

$$\text{since } a_0 = 1$$

It is easy to see that $h(x) = \sum_{x < T^n(1)} 1/\beta^n$

is bounded and therefore

$\nu_E = \int_E h d\mu / \int_0^1 h d\mu$ is a normalised invariant measure equivalent to Lebesgue measure.

REMARK : β has a recurrent tail in its β -expansion if and only if $T^m(1) = T^n(1)$ for some $m \neq n$. Consequently $h(x)$ is a step function with a finite number of steps if and only if β has a recurrent tail.

If β has a recurrent tail in its β -expansion then $\beta^{n+1} - (\varepsilon_0 \beta^n + \dots + \varepsilon_n) = \beta^{m+1} - (\varepsilon_0 \beta^m + \dots + \varepsilon_m)$ for some $m \neq n$
[or = 0]

COROLLARY 17.1

If α is any solution of $z^{n+1} - (\varepsilon_0 z^n + \dots + \varepsilon_n) = z^{m+1} - (\varepsilon_0 z^m + \dots + \varepsilon_m)$

then $h(x) = \sum_{x < T^n(1)} 1/\alpha^n$ satisfies

$$\sum_{m=0}^{[\beta-x]} h\left(\frac{x+m}{\beta}\right) \frac{1}{\alpha} = h(x),$$

We omit the proof as it imitates the proof of Theorem 17.

IV.

Alternating β -expansions

$$f(x) = 1 - x/\beta, \quad x \leq \beta > 1$$

$$f(\varepsilon_1 + \dots + f(\varepsilon_n + \varepsilon)) = 1 - \frac{(\varepsilon_1 + 1)}{\beta} + \dots + (-1)^n \frac{(\varepsilon_n + \varepsilon)}{\beta^n}$$

and $|H_n(x, t)| = 1/\beta^n$ and therefore
 $\frac{\text{Max } |H_n(x, t)|}{\text{Min } |H_n(x, t)|} = 1$. Consequently T is ergodic.

Let
$$h(x) = \sum_{x \geq T^n(0)} \left(-\frac{1}{\beta}\right)^n$$

$h(x)$ is a Saltus function [2.] and is therefore continuous from the right everywhere and continuous from the left except at points $T^n(0)$.

$h(0) = \sum_{0 \geq T^n(0)} \left(-\frac{1}{\beta}\right)^n = 1$ if $T^n(0) \geq 0$
 for all $n > 0$. If $T^n(0) = 0$ for some positive integer n then there is a least $m > 0$ such that $T^m(0) = 0$. The only positive l for which $T^l(0) = 0$ are multiples of m consequently

$$h(0) = \sum_{k=0}^{\infty} \left(-\frac{1}{\beta}\right)^{km} = \frac{1}{1 - \left(-\frac{1}{\beta}\right)^m} > 0$$

An invariant measure equivalent to Lebesgue measure (if it exists) is yielded by a positive solution

$h(x)$ of ,
 (3.4)
$$h(x) = \sum_{n=0}^{\lfloor \beta - x \rfloor} h\left(1 - \frac{x+n}{\beta}\right) \frac{1}{\beta}$$

Real solutions of (3.4) are either non-negative or non-positive, by the ergodicity of T , and if we show that

$$h(x) = \sum_{x \geq T^n(0)} \left(-\frac{1}{\beta}\right)^n$$

is a solution of (3.4), it will follow that $h(x)$ is non-negative since $h(0) > 0$ and $h(x)$ is continuous from the right.

It is clear that $h(x)$ is bounded and therefore

$$\nu E = \int_E h(x) d\mu / \int_0^1 h(x) d\mu$$

will be (except for a constant factor of multiplication) the unique invariant measure equivalent to Lebesgue measure and ν will be normalised.

We show more generally :

THEOREM 18

If the alternating β -expansion of β is

$$\beta = (\varepsilon_0 + 1) - \frac{(\varepsilon_1 + 1)}{\beta} + \dots, \quad \varepsilon_0 = [\beta]$$

and α is any complex solution of

$$z = (\varepsilon_0 + 1) - \frac{(\varepsilon_1 + 1)}{z} + \dots$$

and
$$h(x) = \sum_{x \geq T^n(0)} \left(-\frac{1}{\alpha}\right)^n$$

then
$$\sum_{n=0}^{[\beta-x]} h\left(1 - \frac{x+n}{\beta}\right) \frac{1}{\alpha} = h(x) \quad \text{(except perhaps at points } T^n(0) \text{)}$$

PROOF

We note that $\varepsilon_n(\beta) = [\beta(1 - T^n(0))]$

and suppose

$$\alpha = (\varepsilon_0 + 1) - \frac{(\varepsilon_1 + 1)}{\alpha} + \dots$$

Write
$$a_{n,m}(x) = \begin{cases} 1 & \text{if } x \geq T^m(0) \\ 0 & \text{otherwise} \end{cases}$$

and
$$a_{n,m}(x) = \begin{cases} 1 & \text{if } 1 - \frac{(x+n)}{\beta} \geq T^m(0) \\ 0 & \text{otherwise.} \end{cases}$$

$$\begin{aligned} \sum_{n=0}^{[\beta-x]} h\left(1 - \frac{x+n}{\beta}\right) \frac{1}{\alpha} &= \sum_{n=0}^{[\beta-x]} \frac{1}{\alpha} \sum_{m=0}^{\infty} a_{n,m}(x) \left(-\frac{1}{\alpha}\right)^m \\ &= \sum_{m=0}^{\infty} \frac{1}{\alpha} \left(-\frac{1}{\alpha}\right)^m \left(\sum_{n=0}^{[\beta-x]} a_{n,m}(x) \right) \end{aligned}$$

But $\sum_{n=0}^{[\beta-x]} a_{n,m}(x)$ is the number of n such that

$$1 - \frac{x+n}{\beta} \geq T^m(0) \quad (\text{i.e. such that } \beta(1 - T^m(0)) - x \geq n)$$

This number is $\varepsilon_m = [\beta(1 - T^m(0))]$ if $x > (\beta(1 - T^m(0)))$
 $= T^{m+1}(0)$ and $\varepsilon_m + 1$ otherwise.

Therefore $\sum_{n=0}^{[\beta-x]} a_{n,m}(x) = \varepsilon_m + 1 - a_{m+1}(x)$

if $x \notin \{T^i(0)\}$ and at such x

$$\begin{aligned}
\sum_{n=0}^{\lfloor \beta-\gamma \rfloor} h\left(1 - \frac{\gamma+n}{\beta}\right) \frac{1}{\alpha} &= \sum_{m=0}^{\infty} \frac{1}{\alpha} \left(-\frac{1}{\alpha}\right)^m \left(\varepsilon_{m+1} - a_{m+1}(\gamma)\right) \\
&= \frac{1}{\alpha} \sum_{m=0}^{\infty} (\varepsilon_{m+1}) \left(-\frac{1}{\alpha}\right)^m + \sum_{m=0}^{\infty} \left(-\frac{1}{\alpha}\right)^{m+1} a_{m+1}(\gamma) \\
&= 1 + h(\gamma) - a_0(\gamma) \\
&= h(\gamma) \quad \text{since } a_0(\gamma) = 1.
\end{aligned}$$

REMARK : In the above four cases we were able to assume the validity of representations, since in each case T is ergodic and hence topologically transitive.

CHAPTER 4

 β - TRANSFORMATIONS

In this chapter we make a special study of β -expansions. We omit the parallel study which could be made of alternating β -expansions.

Those $\beta > 1$ which have recurrent tails in their β -expansions we call β -numbers and those $\beta > 1$ which have zero tails in their β -expansions (which happens when $T^m(1) = 0$ for some $m > 0$) we call simple β -numbers.

Renyi has studied the class of simple β -numbers of the form $\beta^{n+1} = \beta^n + 1$ [39]. The results in Chapter 3. concerning β -expansions together with the results in this chapter have been submitted for publication [40]

A. β -NUMBERS AND THEIR CHARACTERISTIC EQUATIONS

If $(a_0, a_1, \dots)$ and $(b_0, b_1, \dots)$ are finite or infinite sequences of non-negative integers with the same number of terms, we write

$$(a_0, a_1, \dots) < (b_0, b_1, \dots)$$

when $a_n < b_n$ for the first n for which $a_n \neq b_n$.

We also write $(a_0, a_1, \dots)(\beta)$ for $a_0 + \frac{a_1}{\beta} + \dots$.

LEMMA 11

If $\beta = a_0 + \frac{a_1}{\beta} + \dots$ where $(a_0, a_1, \dots)$ is a sequence of

non-negative integers and if $(b_0, b_1, \dots)$ is a sequence of non-negative integers with $(b_n, b_{n+1}, \dots) < (a_0, a_1, \dots)$ for all $n \geq 1$, then

$$(4.1) \quad \begin{cases} (b_n, b_{n+1}, \dots)(\beta) < (a_m, \dots)(\beta) \\ (b_n, \dots) \not\leq (a_m, \dots) \end{cases} \quad \text{whenever}$$

unless $\beta = a_0 + \frac{a_1}{\beta} + \dots + \frac{a_{q-1}}{\beta^{q-1}}$ and the tail of $\{b_n\}$ coincides with $\{c_n\}$. $\left(\begin{array}{ll} c_i = a_i & \text{if } i \not\equiv 0 \pmod{q} \\ = a_{q-1} & \text{if } i \equiv 0 \pmod{q} \end{array} \right).$

PROOF

We first prove that $(b_n, \dots)(\beta) \leq (a_m, \dots)(\beta)$ whenever $(b_n, \dots) < (a_m, \dots)$ by showing

$$(4.2) \quad \begin{cases} (b_n, \dots, b_{n+r})(\beta) < (a_m, \dots, a_{m+r})(\beta) \\ (b_n, \dots, b_{n+r}) < (a_m, \dots, a_{m+r}) \end{cases} \quad \text{whenever}$$

In the case $r=0$, $(b_n) < (a_m)$ implies $(b_n)(\beta) = b_n < a_m = (a_m)(\beta)$. Therefore (4.2) holds for all m, n where $r=0$. We suppose (4.2) holds for all m, n when $r < k$.

If $(b_n, \dots, b_{n+k}) < (a_m, \dots, a_{m+k})$ either $b_n = a_m$ and $(b_n, \dots, b_{n+k}) = (a_m, \dots, a_{m+k})$ (in which case $(b_{n+1}, \dots, b_{n+k})(\beta) < (a_{m+1}, \dots, a_{m+k})(\beta)$ by hypothesis) or $b_n < a_m$. In the latter case, since $(b_{n+1}, \dots, b_{n+k}) \leq (a_0, \dots, a_{k-1})$ and again by hypothesis we have

$$(4.3) \quad (b_n, \dots, b_{n+k})(\beta) \leq (b_n, a_0, \dots, a_{k-1})(\beta) \leq (a_m, 0, \dots, 0)(\beta) \leq (a_m, \dots, a_{m+k})(\beta)$$

The central inequality in (4.3) holds by virtue of the relation

$$\frac{a_0}{\beta} + \dots + \frac{a_{k-1}}{\beta^k} \leq 1 \leq a_m - a_n$$

We note that equality holds throughout (4.3) when and only when

$$\beta = a_0 + \frac{a_1}{\beta} + \dots + \frac{a_{k-1}}{\beta^k} \text{ and } (b_{n+1}, \dots, b_{n+k}) = (a_0, \dots, a_{k-1})$$

However this is excluded by the condition $(b_n, \dots) < (a_n, \dots)$ for all $n \geq 1$.

Consequently (4.2) is proved and it follows immediately that

$$(b_n, \dots)(\beta) \leq (a_n, \dots)(\beta) \quad \text{whenever} \\ (b_n, \dots) \leq (a_n, \dots)$$

Suppose $(b_n, \dots) < (a_n, \dots)$ then for some $p, q, p-n = q-m \geq 0$
 $(b_0, \dots)(\beta) \leq (b_p, a_0, a_1, \dots)(\beta) \leq (a_q, 0, \dots)(\beta) \leq^{(3)} (a_q, a_{q+1}, \dots)(\beta)$
 and strict equality holds throughout only if

$$a_{q+1} = a_{q+2} = \dots = 0$$

In this case
$$\beta = a_0 + \frac{a_1}{\beta} + \dots + \frac{a_{q-1}}{\beta^q}$$

$$= c_0 + \frac{c_1}{\beta} + \dots, \text{ where } \begin{cases} c_i = a_i & \text{if } i \not\equiv 0 \pmod{q} \\ c_i = a_{q-1} & \text{if } i \equiv 0 \pmod{q} \end{cases}$$

and consequently $(b_n, \dots)(\beta) < (c_n, \dots)(\beta) = (a_n, \dots)(\beta)$

whenever $(b_n, \dots) < (c_n, \dots)$

But $(b_n, \dots) < (a_n, \dots)$ implies either the tail of $\{b_n\}$ coincides with $\{c_n\}$, or $(b_n, \dots) < (c_n, \dots)$.

This completes the proof.

THEOREM 19

If $\beta > 1$ and the β -expansion of β is

$$\beta = a_0 + \frac{a_1}{\beta} + \dots$$
and if $(b_0, b_1, \dots)$ is a sequence of non-negative numbers whose tail, in case the β -expansion of β is finite, does not coincide with $\{a_n\}$ (see Lemma 11), a necessary and sufficient condition for the existence of x with β -expansion

$$x = b_0 + \frac{b_1}{\beta} + \dots$$

is that $(b_n, \dots) < (a_0, \dots)$ for all $n \geq 1$.

In particular $(a_n, \dots) < (a_0, \dots)$ for all $n \geq 1$.

PROOF

If there exists x with β -expansion $x = b_0 + \frac{b_1}{\beta} + \dots$, for every $n \geq 1$,

$$\frac{b_n}{\beta} + \frac{b_{n+1}}{\beta^2} + \dots < 1 = \frac{a_0}{\beta} + \frac{a_1}{\beta^2} + \dots$$

Consequently for every $n \geq 1$,

$$(b_n, b_{n+1}, \dots) \neq (a_0, a_1, \dots)$$

Let k be the first integer with $b_{n+k} \neq a_k$ then

$$b_{n+k} + \frac{b_{n+k+1}}{\beta} + \dots < a_k + \frac{a_{k+1}}{\beta} + \dots$$

and since $\frac{a_{k+1}}{\beta} + \frac{a_{k+2}}{\beta^2} + \dots < 1$ we have $b_{n+k} < a_k + 1$

i.e. $b_{n+k} < a_k$.

By Lemma 11 for all $n \geq 1$, $(b_n, b_{n+1}, \dots) < (a_0, a_1, \dots)$ implies $\frac{b_n}{\beta^n} + \frac{b_{n+1}}{\beta^{n+1}} + \dots < \frac{a_0}{\beta^n} + \frac{a_1}{\beta^{n+1}} + \dots = 1$ and there exists x with β -expansion

$$x = b_0 + \frac{b_1}{\beta} + \frac{b_2}{\beta^2} + \dots$$

In the following Corollaries $(a_0, a_1, \dots)$ is a sequence of non-negative integers with $a_n \leq a_0 \geq 1$ for all $n \geq 0$.

It should be noted that under this condition there is a unique solution $\beta > 1$ of the equation

$$x = a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \dots$$

and this solution lies between a_0 and $a_0 + 1$.

By Lemma 11 and Theorem 19 we have :

COROLLARY 19.1

The unique solution $\beta > 1$ of

$$x = a_0 + \frac{a_1}{x} + \dots$$

has for its β -expansion

$$\beta = a_0 + \frac{a_1}{\beta} + \dots$$

if and only if $(a_n, a_{n+1}, \dots) < (a_0, a_1, \dots)$ for all $n \geq 1$.

We omit the proofs of the following corollaries.

COROLLARY 19.2

The unique solution $\beta > 1$ of

$$(4.4) \quad x = \sum_{k=0}^{\infty} \frac{a_k}{x^k} \quad \left(\text{where } a_{n+r(m)+k} = a_{n+k}, \right. \\ \left. r \geq 0, 0 \leq k < m \right)$$

has $\beta = \sum_{k=0}^{\infty} \frac{a_k}{\beta^k}$ for its β -expansion if and

only if

$$(4.5) \quad \begin{cases} (a_0, \dots, a_{n+m-1}) > (a_1, \dots, a_{n+m-1}, a_{n+m}) \\ (a_0, \dots, a_{n+m-1}) > (a_2, \dots, a_{n+m-1}, a_{n+1}, a_{n+2}) \\ (a_0, \dots, a_{n+m-1}) > (a_{n+m-1}, \dots, a_{n+1}, a_{n+m-1}) \end{cases}$$

where

$$r(k) \equiv k \pmod{m}.$$

COROLLARY 19.3

The unique solution $\beta > 1$ of

$$x^{n+m} - (a_0 x^{n+m-1} + \dots + a_{n+m-1}) = x^n - (a_0 x^{n-1} + \dots + a_{n-1})$$

has $\beta = \sum_{k=0}^{\infty} \frac{a_k}{\beta^k}$ (where $a_{n+r(m)+k} = a_{n+k}$, $r \geq 0, 0 \leq k < m$.)

for its β -expansion if and only if (4.5) holds.

COROLLARY 19.4

The unique solution $\beta > 1$ of

$x = a_0 + \frac{a_1}{x} + \dots + \frac{a_n}{x^n}$ has for its β -expansion

$\beta = a_0 + \frac{a_1}{\beta} + \dots + \frac{a_n}{\beta^n}$ if and only if

$$(4.6) \quad \begin{cases} (a_0, \dots, \underline{a_n}) > (\underline{a_1}, \dots, a_n, 0) \\ (a_0, \dots, \underline{a_n}) > (\underline{a_n}, \dots, a_n, 0 \dots 0) \\ (a_0, \dots, \underline{a_n}) > (\underline{a_n}, 0 \dots 0) \end{cases}$$

(4.6) is equivalent to

$$(4.7) \quad \begin{cases} (a_0, \dots, \underline{a_{n-1}}) \geq (\underline{a_1}, \dots, a_n) \\ (a_0, \dots, \underline{a_{n-k}}) \geq (\underline{a_k}, \dots, a_n) \\ (\underline{a_0}) \geq (\underline{a_n}) \end{cases}$$

If β is a non-simple β -number with β -expansion

$$\beta = \sum_{k=0}^{\infty} \frac{a_k}{\beta^k} \quad \left(\text{where } a_{n+r m+k} = a_{n+k}, \right. \\ \left. r \geq 0, 0 \leq k < m \right)$$

the equation $z^{n+m} = (a_0 z^{n+m-1} + \dots + a_{n+m-1}) = z^n (a_0 z^{m-1} + \dots + a_{n-1})$

is termed the characteristic equation of β .

Likewise, if β is a simple β -number with β -expansion

$$\beta = a_0 + \frac{a_1}{\beta} + \dots + \frac{a_n}{\beta^n}$$

the equation

$$z^{n+1} - (a_0 z^n + \dots + a_n) = 0$$

is termed the characteristic equation of β .

LEMMA 12

If β is a simple [non-simple] β -number with characteristic equation

$$z^{n+m} - (\varepsilon_0 z^{n+m-1} + \dots + \varepsilon_{n+m-1}) = 0$$

$$[= z^n - (\varepsilon_0 z^{n-1} + \dots + \varepsilon_{n-1})]$$

the conjugates $\beta_1, \dots, \beta_{n+m-1}$ of β with respect to the characteristic equation, satisfy

$$(4.8) \quad z^{n+m-1} + T(1)z^{n+m-2} + \dots + T^{n+m-1}(1) = 0$$

$$[= z^{n-1} + T(1)z^{n-2} + \dots + T^{n-1}(1)]$$

PROOF

This is easily verified by multiplying both sides of (4.8) by $(z - \beta)$ and adding $T^{n+m}(1) = 0$

$$[= T^n(1)]$$

to both sides. This gives the characteristic equation of β .

THEOREM 20.

The conjugates of a β -number with respect to its characteristic equation have absolute value less than 2 *

PROOF

Using the notation employed in Lemma 12 the conjugates of β with respect to its characteristic equation satisfy

$$z^{n+m-1} + T(1)z^{n+m-2} + \dots + T^{n+m-1}(1) = 0$$

$$[= z^{n-1} + \dots + T^{n-1}(1)]$$

$$\text{If } |z| > 1, (|z|^{n-1}) |z^{n-1} + T(1)z^{n-2} + \dots + T^{n-1}(1)|$$

$$< |z|^{n-1} + \dots + 1$$

$$\text{i.e. } |z^{n-1} + T(1)z^{n-2} + \dots + T^{n-1}(1)| < \frac{1}{|z| - 1}$$

$$\text{Therefore } |z|^{n-1} < \frac{1}{|z| - 1} + |T(1)z^{n-2} + \dots + T^{n-1}(1)|$$

$$< \frac{1}{|z| - 1} + \frac{|z|^{n-1} - 1}{|z| - 1} = \frac{|z|^{n-1}}{|z| - 1}$$

$$\text{i.e. } |z| < 2$$

* In this respect β -numbers resemble Pisot-Vijayaraghavan numbers ([41.] p.p. 133 - 144) defined as those algebraic integers greater than 0, with conjugates of maximum absolute value less than one.

If we define analogously the characteristic function of a non- β - number with β -expansion

$$\beta = \varepsilon_0 + \frac{\varepsilon_1}{\beta} + \dots$$

as
$$z = \varepsilon_0 + \frac{\varepsilon_1}{z} + \dots$$

its solutions must have absolute value greater than one.
(By the properties of convergence.)

On the other hand there are β -numbers whose conjugates have absolute value less than one :

i) The simple β -number given by $\beta^2 = \beta + 1$

The conjugate of $\beta > 1$ is $\frac{1-\sqrt{5}}{2}$ and

$$-1 < \frac{1-\sqrt{5}}{2} < 0$$

ii) The non-simple β -number given by $\beta^2 - 2\beta - 1 = \beta - 2$

The conjugate of $\beta > 1$ is $\frac{3-\sqrt{5}}{2}$

and $0 < \frac{3-\sqrt{5}}{2} < 1$.

Lemma 12 ♀

It can be proved as in Theorem 20 that the conjugated of the non- β -number β satisfy

$$1 + \frac{T(1)}{z} + \frac{T^2(1)}{z^2} + \dots = 0$$

And consequently $1 < \frac{1}{|z|} \left(1 + \frac{1}{|z|} + \frac{1}{|z|^2} + \dots \right)$

i.e. $1 < \frac{1}{|z|-1}$ and $|z| < 2$.

B.

DISTRIBUTION OF β -NUMBERSTHEOREM 21

The set of simple β -numbers is everywhere dense in $(1, \infty)$.

PROOF

Every $\beta > 1$ has a β -expansion

$$\beta = \varepsilon_0 + \frac{\varepsilon_1}{\beta} + \dots \quad \text{where } \varepsilon_n(\beta) = \varepsilon_n, n = 0, 1, \dots$$

and according to Lemma 12,

$$(\varepsilon_0, \varepsilon_1, \dots) > (\varepsilon_n, \varepsilon_{n+1}, \dots) \quad \text{for } n \geq 1,$$

$$\begin{aligned} \text{and therefore } (\varepsilon_0, \dots, \varepsilon_n) &> (\varepsilon_1, \dots, \varepsilon_n, 0) \\ (\varepsilon_0, \dots, \varepsilon_n) &> (\varepsilon_n, 0, \dots, 0) \end{aligned}$$

Consequently by Corollary 19.4 there is a simple β -number $\beta(n)$ with $\beta(n)$ -expansion

$$\beta(n) = \varepsilon_0 + \frac{\varepsilon_1}{\beta(n)} + \dots + \frac{\varepsilon_n}{\beta(n)^n}$$

It suffices therefore to show that $\lim_{n \rightarrow \infty} \beta(n) = \beta$. $\beta(n)$ is non-decreasing as n increases for otherwise

$$\beta(n+1) < \beta(n) < \varepsilon_0 + \frac{\varepsilon_1}{\beta(n+1)} + \dots + \frac{\varepsilon_n}{\beta(n+1)^n} \leq \beta(n+1).$$

Suppose there exists K with $\beta(n) < K < \beta$ for all n then $\beta(n) > \varepsilon_0 + \frac{\varepsilon_1}{K} + \dots + \frac{\varepsilon_n}{K^n}$ for all n and therefore

$$K \geq \lim_{n \rightarrow \infty} \beta(n) \geq \varepsilon_0 + \frac{\varepsilon_1}{K} + \dots > \varepsilon_0 + \frac{\varepsilon_1}{\beta} + \dots = \beta$$

This contradiction shows that $\lim_{n \rightarrow \infty} \beta(n) = \beta$ and this completes the proof.

C.

THE NORMALISING FUNCTION

The frequency of each digit exists and is the same for almost all x and is expressed by the function $\phi(n)$:

$$(4.9) \quad \phi(n) = \nu\left[\frac{n}{\beta}, \frac{n+1}{\beta}\right) \quad \text{if } n < [\beta]$$

$$= \nu\left[\frac{[\beta]}{\beta}, 1\right) \quad \text{if } n = [\beta]$$

where $\nu E = \int_0^1 h_\beta(x) dx / F(\beta)$ and $F(\beta)$ is the normalising factor $\int_0^1 h_\beta(x) dx$.

We note the following formulae

$$F(\beta) = \int_0^1 h_\beta(x) dx = \int_0^1 \left(\sum_{x < T^n(1)} \frac{1}{\beta^n} \right) dx$$

$$= \int_0^1 \left(\sum_{n=0}^{\infty} \frac{a_n(x)}{\beta^n} \right) dx \quad (\text{where } a_n(x) = 1 \text{ if } x < T^n(1))$$

$$= \sum_{n=0}^{\infty} \frac{1}{\beta^n} \int_0^1 a_n(x) dx \quad = 0 \text{ otherwise)}$$

$$(4.10) \quad \begin{cases} = \sum_{n=0}^{\infty} T^n(1) / \beta^n \\ = \sum_{n=0}^{\infty} \frac{(n+1) \varepsilon_n(\beta)}{\beta^n} \end{cases} \quad \left(\text{since } T^n(1) = \frac{\varepsilon_n(\beta)}{\beta} + \frac{\varepsilon_{n+1}(\beta)}{\beta^2} + \dots \right.$$

for all $n \geq 0$)

We have therefore $F(n) = 1$, for all $n \geq 1$

and $1 \leq F(\beta) \leq 1 + \frac{1}{\beta} + \frac{1}{\beta^2} + \dots = \beta / (\beta - 1)$

and consequently

$$(4.11) \quad \lim_{\beta \rightarrow \infty} F(\beta) = 1$$

LEMMA 13

If the α -expansion of α is

$$\alpha = a_0 + \frac{a_1}{\alpha} + \dots$$

and the β -expansion of β is

$$\beta = b_0 + \frac{b_1}{\beta} + \dots$$

then $\alpha > \beta$ if and only if $(a_0, a_1, \dots) > (b_0, b_1, \dots)$.

PROOF

Suppose on the contrary $\alpha < \beta$ and $(a_0, a_1, \dots) > (b_0, b_1, \dots)$
(The case of equality is trivially dismissed.)

If n is the first integer with $a_n > b_n$,

$$b_n + \frac{b_{n+1}}{\beta} + \dots < b_n + 1 \leq a_n < a_n + \frac{a_{n+1}}{\alpha} + \dots$$

therefore

$$\beta = b_0 + \frac{b_1}{\beta} + \dots < a_0 + \frac{a_1}{\beta} + \dots < a_0 + \frac{a_1}{\alpha} + \frac{a_2}{\alpha^2} + \dots = \alpha$$

This contradiction completes the proof.

LEMMA 14

If the β -expansion of β is

$$\beta = \varepsilon_0 + \frac{\varepsilon_1}{\beta} + \dots + \frac{\varepsilon_{n-1}}{\beta^{n-1}}$$

for every r there exists $\alpha = \alpha(r)$ with α -expansion

$$(4.12) \quad \alpha = \sum_{m=0}^{(r+1)n-1} \frac{a_m}{\alpha^m}$$

(where $a_{sn+t} = \varepsilon_t$ for $0 \leq s \leq r$
 $0 \leq t \leq n-1$)

and $a_{sn+n-1} = \varepsilon_{n-1}$ for $0 \leq s \leq r$)

and $\alpha(r) \uparrow \beta$ as $r \rightarrow \infty$.

PROOF

$$(\varepsilon_0, \dots, \varepsilon_{n-1}) > (\varepsilon_k, \dots, \varepsilon_{n-1}, 0, \dots, 0)$$

$$\text{i.e. } (\varepsilon_0, \dots, \varepsilon_{n-k-1}) > (\varepsilon_k, \dots, \varepsilon_{n-1}, 0, \dots, 0)$$

Therefore

$$(a_0, \dots, a_{(r+1)n-1}) > (a_k, \dots, a_{(r+1)n-1}, 0, \dots, 0)$$

where the a_i 's are defined in (4.4)

By Corollary 19.4 there exists $\alpha = \alpha(r)$ with α -expansion (4.12). Evidently

$$\begin{aligned} \alpha &= \left(\varepsilon_0 + \frac{\varepsilon_1}{\alpha} + \dots + \frac{\varepsilon_{n-1}-1}{\alpha^{n-1}} \right) \left(1 + \frac{1}{\alpha^n} + \dots + \frac{1}{\alpha^{rn}} \right) \\ &= \left(\varepsilon_0 + \frac{\varepsilon_1}{\alpha} + \dots + \frac{\varepsilon_{n-1}-1}{\alpha^{n-1}} \right) \left(\frac{1 - 1/\alpha^{(r+1)n}}{1 - 1/\alpha^n} \right) \end{aligned}$$

According to Lemma 13, $\alpha(r)$ is non-decreasing as r increases. Let $\gamma = \sup_r \alpha(r)$ then

$$\gamma = \left(\varepsilon_0 + \frac{\varepsilon_1}{\gamma} + \dots + \frac{\varepsilon_{n-1}-1}{\gamma^{n-1}} \right) \left(\frac{1}{1 - 1/\gamma^n} \right)$$

i.e.

$$\gamma = \varepsilon_0 + \frac{\varepsilon_1}{\gamma} + \dots + \frac{\varepsilon_{n-1}-1}{\gamma^{n-1}}$$

and therefore

$$\gamma = \beta.$$

THEOREM 22

If β is a simple β -number with β -expansion

$$\beta = \varepsilon_0 + \frac{\varepsilon_1}{\beta} + \dots + \frac{\varepsilon_{n-1}}{\beta^{n-1}}$$

then

$$\lim_{\gamma \uparrow \beta} F(\gamma) = F(\beta) \left(\frac{\beta^n}{\beta^n - 1} \right).$$

PROOF

If $\alpha = \alpha(r) \leq \gamma < \beta$ ($\alpha(r)$ defined in (4.12)) γ will have in its γ -expansion the same $(r+1)_n$ initial digits as $\alpha(s)$ for all $s \geq r$. (By Lemma 14.,

Therefore

$$\gamma = \left(\varepsilon_0 + \frac{\varepsilon_1}{\gamma} + \dots + \frac{\varepsilon_{n-1}-1}{\gamma^{n-1}} \right) + \dots + \left(\frac{\varepsilon_0}{\gamma^n} + \dots + \frac{\varepsilon_{n-1}-1}{\gamma^{(r+1)_n-1}} \right) + \frac{\varepsilon_{(r+1)_n}(\gamma)}{\gamma^{(r+1)_n}} + \dots$$

$$\text{and } \frac{\varepsilon_0}{\beta} + \frac{2\varepsilon_1}{\beta^2} + \dots + \frac{(r+1)_n(\varepsilon_{n-1}-1)}{\beta^{(r+1)_n}} \leq F(\gamma)$$

$$= \frac{\varepsilon_0}{\gamma} + \frac{2\varepsilon_1}{\gamma^2} + \dots + \frac{(r+1)_n(\varepsilon_{n-1}-1)}{\gamma^{(r+1)_n}} \quad (\text{where } \lim_{r \rightarrow \infty} \Theta_r = 0)$$

$$\leq \frac{\varepsilon_0}{\alpha} + \frac{2\varepsilon_1}{\alpha^2} + \dots + \frac{(r+1)_n(\varepsilon_{n-1}-1)}{\alpha^{(r+1)_n}} + \Theta_r$$

It suffices to show * :

$$\lim_{r \rightarrow \infty} F(\alpha_r) = \lim_{r \rightarrow \infty} \left(\frac{\varepsilon_0}{\alpha} + \dots + \frac{(r+1)_n(\varepsilon_{n-1}-1)}{\alpha^{(r+1)_n}} \right) = F(\beta) \frac{\beta^n}{(\beta^n - 1)}$$

* In what follows the dependence of α on r should be noted. $T_r(n)$ denotes $(r+1)_n$.

Put $x(r) = \varepsilon_0 + \frac{\varepsilon_1}{\alpha} + \dots + \frac{(\varepsilon_{n-1}-1)}{\alpha^{n-1}}$

and $x'(r) = \frac{\varepsilon_0}{\alpha} + \frac{2\varepsilon_1}{\alpha^2} + \dots + \frac{n(\varepsilon_{n-1}-1)}{\alpha^n}$

Since $F(x(r)) - x'(r) = \frac{(n+1)\varepsilon_0}{\alpha^n} + \dots + \frac{(r+1)n(\varepsilon_{n-1}-1)}{\alpha^{(r+1)n}}$

$$= \frac{1}{\alpha^n} \left[n \left\{ \frac{\varepsilon_0}{\alpha} + \dots + \frac{\varepsilon_{n-1}-1}{\alpha^{n-1}} \right\} + \frac{\varepsilon_0}{\alpha} + \dots + \frac{r n (\varepsilon_{n-1}-1)}{\alpha^{rn}} \right]$$

$$= \frac{1}{\alpha^n} \left[n T_r^n(1) + F(x(r)) - \left\{ \frac{(r+1)\varepsilon_0}{\alpha^{r+1}} + \dots + \frac{(r+1)n(\varepsilon_{n-1}-1)}{\alpha^{(r+1)n}} \right\} \right]$$

$$= \frac{1}{\alpha^n} \left[n T_r^n(1) + F(x(r)) - \frac{1}{\alpha^{rn}} \left\{ r n x(r) + x'(r) \right\} \right]$$

we have

$$\begin{aligned} \lim_{r \rightarrow \infty} T_r^n(1) &= \lim_{r \rightarrow \infty} \left\{ \frac{1}{\alpha} \left(\varepsilon_0 + \frac{\varepsilon_1}{\alpha} + \dots + \frac{\varepsilon_{n-1}-1}{\alpha^{n-1}} \right) \left(\frac{1 - 1/\alpha^{rn}}{1 - 1/\alpha^n} \right) \right\} \\ &= \frac{1}{\beta} \left(\varepsilon_0 + \frac{\varepsilon_1}{\beta} + \dots + \frac{\varepsilon_{n-1}-1}{\beta^{n-1}} \right) \frac{1}{(1 - 1/\beta^n)} = 1 \end{aligned}$$

and

$$\lim_{r \rightarrow \infty} x'(r) = \frac{\varepsilon_0}{\beta} + \dots + \frac{n(\varepsilon_{n-1}-1)}{\beta^n}$$

Therefore

$$\lim_{r \rightarrow \infty} F(x(r)) \left(1 - \frac{1}{\alpha^n} \right) = \frac{\varepsilon_0}{\beta} + \dots + \frac{n \varepsilon_{n-1}}{\beta^n}$$

i.e.

$$\lim_{r \rightarrow \infty} F(x(r)) = \frac{F(\beta) \beta^n}{(\beta^n - 1)}$$

THEOREM 23

If $\beta > 1$ then $\lim_{\gamma \downarrow \beta} F(\gamma) = F(\beta)$

and if β is not a simple β -number

$$\lim_{\gamma \uparrow \beta} F(\gamma) = F(\beta)$$

PROOF

In either case suppose that $|\gamma - \beta| < \delta$ and let $n = n(\delta)$ be the number of initial consecutive digits common to both the γ -expansion of γ and the β -expansion of β . By Lemma 13 it is easy to see that $\lim_{\delta \rightarrow 0} n(\delta) = \infty$,

$$\begin{aligned} F(\beta) &= \frac{\varepsilon_0}{\beta} + \frac{2\varepsilon_1}{\beta^2} + \dots + \frac{n\varepsilon_{n-1}}{\beta^n} + \theta_n \quad \text{where } \lim_{n \rightarrow \infty} \theta_n = 0 \\ F(\gamma) &= \frac{\varepsilon_0}{\gamma} + \frac{2\varepsilon_1}{\gamma^2} + \dots + \frac{n\varepsilon_{n-1}}{\gamma^n} + \phi_n \quad \text{where } \lim_{n \rightarrow \infty} \phi_n = 0 \\ |F(\beta) - F(\gamma)| &\leq \left| \frac{1}{\beta} - \frac{1}{\gamma} \right| \left(\varepsilon_0 + 2\varepsilon_1 \left(\frac{1}{\beta} + \frac{1}{\gamma} \right) + \dots + n\varepsilon_{n-1} \left(\frac{1}{\beta^{n-1}} + \dots + \frac{1}{\gamma^{n-1}} \right) \right) \\ &\quad + \theta_n + \phi_n \\ &\leq \left| \frac{1}{\beta} - \frac{1}{\gamma} \right| \left(\varepsilon_0 + \frac{2^2\varepsilon_0}{\alpha} + \dots + \frac{n^2\varepsilon_0}{\alpha^{n-1}} \right) + \theta_n + \phi_n \\ &\leq K \left| \frac{1}{\beta} - \frac{1}{\gamma} \right| + \theta_n + \phi_n. \quad \text{where } \alpha = \min(\beta, \gamma) \end{aligned}$$

Therefore $\lim_{\gamma \uparrow \beta} F(\gamma) = F(\beta)$ if β is not a simple β -number and $\lim_{\gamma \downarrow \beta} F(\gamma) = F(\beta)$ for all $\beta > 1$.

THEOREM 24

$$\lim_{\gamma \downarrow 1} F(\gamma) = \infty$$

PROOF

Let us write $\beta = \beta(n)$ for the positive solution of

$$\beta = 1 + 1/\beta^n$$

then $\beta(n)$ decreases as n increases and $\lim_{n \rightarrow \infty} \beta(n) = 1$.
For if

$$\beta(n) \leq \beta(n+1) ; \beta(n+1) \leq 1 + \frac{1}{\beta(n)^{n+1}} < 1 + \frac{1}{\beta(n)^n} = \beta(n).$$

And if $\beta(n) \geq K > 1$ for all n ,

$$\lim_{n \rightarrow \infty} \beta(n) \leq \lim_{n \rightarrow \infty} 1 + \frac{1}{K^n} = 1.$$

Therefore, by virtue of the continuity from the right of $F(\beta)$ at points $\beta > 1$, it suffices to show that

$$\lim_{n \rightarrow \infty} F(\beta(n)) = \infty.$$

$$F(\beta(n)) = \frac{1}{\beta(n)} + \frac{n+1}{\beta(n)^{n+1}} \text{ we show that } \lim_{n \rightarrow \infty} \frac{n}{\beta(n)^n} = \infty.$$

Suppose $\frac{n}{\beta(n)^n} < K$ for infinitely many n then $\beta(n)^n$

$$= \left(1 + \frac{1}{\beta(n)^n}\right)^n < \left(1 + \frac{1}{\beta(n)^n}\right)^{K/\beta(n)^n}$$

for infinitely many n . But

$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{\beta(n)^n}\right)^{K/\beta(n)^n}$
 converges to e^K since $\lim_{n \rightarrow \infty} \beta(n)^n = \infty$,

by virtue of the relation $\beta(n)^n = \frac{1}{\beta(n) - 1}$.

Therefore there exists arbitrarily large n with

$\beta(n)^n < e^K$ and yet $\lim_{n \rightarrow \infty} \beta(n)^n = \infty$.

This contradiction establishes the Theorem.

CHAPTER 5 β -NUMBERS AND MATRICES.

Throughout this Chapter β is assumed to be a β -number i.e. $\beta > 1$ and satisfies an equation of the form

$$(5.1) \quad \beta^{n+1} - (\varepsilon_0 \beta^n + \dots + \varepsilon_n) = 0 \quad (m \leq n)$$

Let n be the least k so that $T^{k+1}(1)$ appears among $0, T(1), \dots, T^k(1)$ and let $I_0, \dots, I_n$ be the lengths of $n+1$ disjoint, consecutive intervals $I_0, \dots, I_n$ with end points $0, 1, T(1), \dots, T^n(1)$.

Write $I_k = \frac{n_k}{\beta} + r_k$ where n_k is a non-negative integer and $r_k < \frac{1}{\beta}$. If the end points of I_k are x_k, x_{k+1} ($x_k < x_{k+1}$) then

$$[x_k, x_{k+1}) = [x_k, x_k + \frac{1}{\beta}) \cup \dots \cup [x_k + \frac{n_k-1}{\beta}, x_k + \frac{n_k}{\beta}) \cup [x_k + \frac{n_k}{\beta}, x_{k+1})$$

T maps each of $[x_k, x_k + \frac{1}{\beta}), \dots, [x_k + \frac{n_k-1}{\beta}, x_k + \frac{n_k}{\beta})$

one-one onto $[0, 1)$ and maps $[x_k + \frac{n_k}{\beta}, x_{k+1})$ one-one onto $S_k = [T(x_k), T(x_{k+1}))$ if $\beta T(x_k) < T(x_{k+1})$

$$= [0, T(x_{k+1})) \cup [T(x_k), 1) \text{ otherwise.}$$

If $x \in I_k$, ε_{0k} of the points in $T^{-1}(x)$ are in I_0 .

ε_{nk} of the points in $T^{-1}(x)$ are in I_n

Therefore if

$$(5.2) \quad \varepsilon_{0k} h_0 + \dots + \varepsilon_{nk} h_n = \sum_{T y = x} h(y) = \sum_{m=0}^{[s-x]} h\left(\frac{x+m}{s}\right)$$

Let us denote the operator on functions

$$\sum_{m=0}^{[s-x]} g\left(\frac{x+m}{s}\right) \quad \text{by } A.$$

There is a one-one correspondence between the step functions described above and the vectors $\begin{pmatrix} h_0 \\ \vdots \\ h_n \end{pmatrix}$ in V_{n+1}

(the space of vectors $\begin{pmatrix} h_0 \\ \vdots \\ h_n \end{pmatrix}$) and from (5.2) if

$\begin{pmatrix} h_0 \\ \vdots \\ h_n \end{pmatrix}$ corresponds to $h(x)$ then $\bar{A} \begin{pmatrix} h_0 \\ \vdots \\ h_n \end{pmatrix}$ corresponds to $\bar{A} h(x)$

With the natural interpretation of the word "isomorphism", we have :

THEOREM 25

The set of step functions whose steps rest on I_j together with the operator A is isomorphic with V_{n+1} together with the matrix $\bar{A}$.

By Corollary 17.1 all solutions of the characteristic equation of β , are eigen values of A and therefore eigen values of $\bar{A}$. This is some justification for using the term "characteristic equation"

It follows from Theorem 25 :

THEOREM 26

ν is an invariant measure equivalent to Lebesgue measure if and only if

$$\nu E = \int_E h(x) dx \quad \text{where } h(x) \text{ is a step function}$$

with values $h_j > 0$ on I_j and

$$\bar{A} \begin{pmatrix} h_0 \\ \vdots \\ h_n \end{pmatrix} = \beta \begin{pmatrix} h_0 \\ \vdots \\ h_n \end{pmatrix}$$

We have therefore a simple method of evaluating invariant measures for the case of β -numbers.

REFERENCES

1.) P.R. Halmos. Measure Theory. D. Van Nostrand.
New York (1950)
2.) F. Riesz & B. Sz. Nagy . Functional Analysis. Blackie & Son Ltd
London and Glasgow (1956)
- 3.) P.R. Halmos . Lectures on Ergodic Theory. Mathematical Society of Japan (1956)
4.) A.C Zaanen. Linear Analysis. Amsterdam-Groningen
(1953)
5.) L.H Loomis. An Introduction to Abstract Harmonic Analysis. D.Van Nostrand. New York
(1953)
6.) J. Von Neumann. Proof of the quasi-ergodic hypothesis. Proc. Nat. Acad. Sci. , U.S.A. Vol 18.
(1932.) 78 - 82.
7.) B.O.Koopman. Hamiltonian systems and transformation in Hilbert space. Proc.Nat.Acad.Sci.,
U.S.A. Vol. 17 (1931) 315 -318.
- 8) Y.N. Dowker. A note on the ergodic theorems. Bull. on the Am.M.Soc. Vol 55 No. 4 (1949)
379 - 383.
9.) G.B. Birkhoff. Proof of the ergodic theorem. Proc. Nat.Acad.Sci., U.S.A. Vol 17 (1931)
656 - 660.
- 10) E. Hopf . Ergodentheorie. Berlin (1937).
- 11) W. Hurewicz . Ergodic Theorem without invariant measure. Ann. Math. Vol 45 (1944)
192 - 206.
- 12) J.C.Oxtoby. On the ergodic theorem of Hurewicz. Ann.Math. Vol 49. (1948) 872 - 884.
- 13) P.R. Halmos An ergodic theorem. Proc.Nat Acad. Sci U.S.A. Vol 32 (1946) 156 - 161.

- 14.) Y.N. Dowker. A new proof of the general ergodic theorem. Acta. Sci. Math. Szeged. 12 (1950) 162 - 166.
- 15.) N. Dunford & D.S. Miller. On the ergodic theorem. Trans. Amer. Math. Soc. 60. (1946) 538 - 549.
- 16.) A.P. Calderón. Sur les Mesures Invariantes. Compte Rendus. 246. (1955) 1960- 1962.
- 17.) Y.N. Dowker. Sur les Applications Mesurables. Comptes Rendus. 242. (1956) 329-331.
- 18.) P.R. Halmos & J. Von Neumann. Operator methods in classical mechanics. II. Ann. Math. 43 (1942) 332 - 350.
- 19.) V.A. Rohlin. On the fundamental ideas of measure theory. Mat. Sbornik. (N.S.) 25. (67) (1949). 107 - 150.
- 20.) V.A. Rohlin. On the fundamental ideas of measure theory. Am. Math. Soc. Translation No. 71.
- 21.) F. Hausdorff. Mengenlehre. Berlin. (1935).
- 22.) F. Hausdorff. Set Theory. Chelsea Publishing Co. New York. (1957)
- 23.) J.L. Doob. Stochastic Processes depending on a continuous parameter. Trans. Amer. Math. Soc. 42 (1937) 109 - 110.
- 24.) A. Renyi. Representations for real numbers and their ergodic properties. Acta. Math. Acad. Sci. Hung. 8. (1957) 477-493.
- 25.) B.H. Bissenger. A generalisation of continued fractions Bull. Amer. Math. Soc. 50. (1944) 868 - 876.
- 26.) C.I. Everett. Representations for real numbers. Bull. Amer. Math. Soc. 52. (1946) 861 - 869.

- 27.) K. Knopp. Mengentheoretische Behandlung einiger Problem der Diophantischen Approximationen und der Transfiniten Wahrscheinlichkeiten. Math. Annalen. 95. (1926) 409 - 426.
- 28.) P. Lévy. Sur les lois de probabilité dont dependent les quotients complets et incomplets d'une fraction continue. Bull. Soc. Math 57. (1929) 178 - 194.
- 29.) P. Lévy. Sur le développement en fraction continue d'un nombre choise en hazard. Comp. Math 3. (1936) 286 - 303.
- 30.) P. Lévy. Theorie de l'addition des variables aleatoires. (Paris) 1954.
- 31.) A. Khintchine. Metrische Kettenbruchprobleme. Comp. Math. 1. (1935) 359 - 382.
- 32.) A. Khintchine. Zur Metrischen Kettenbruchtheorie. Comp. Math. 3. (1936) 276-285.
- 33.) A. Khintchine. Kettenbrüche. (Leipzig.) (1956)
- 34.) C. Ryll - On the ergodic theorem II. Ergodic Theory of continued fractions. Studia Math. 12½ (1951) 74 - 79.
- Nardzewski .
- 35.) S. Hartman , Théorèmes ergodiques et leurs applicat-
E. Marczewski & ions . Coll. Math. 2 (1951)
C. Ryll-Nardzewski , 109 - 123.
- 36.) S. Hartman. Quelques propriétés ergodiques des fractions continues. Studia Math. 12 (1951) 271 - 278.
- 37.) A. Rényi Valós számok előállítására szolgáló algoritmusokról. M.T.A. Mat. es Fiz. Oszt. Kozl. 7. (1957) 265 - 293.
- 38.) É. Borel Les probabilités dénombrables et leurs applications arithmétique. Rendiconti del Circ. Math. di Palermo. 27. (1909) 247 - 271.

- 39.) A. Renyi, Probability Methods in number theory.
Proc. Int. Cong. Math. (1958) 529 - 539.
- 40.) W. Parry, Measures invariant under β -transformations. To appear in Acta Math. Acad. Sci. Hung.
- 41.) J.W.S. Cassels, An introduction to Diophantine Approximation. Cambridge Tracts. (1957).
-