KLIMKO, Eugene Martin, 1939-
CONTRIBUTIONS TO THE THEORY OF INFINITE
INVARIANT MEASURES.

The Ohio State University, Ph.D., 1967
Mathematics

University Microfilms, Inc., Ann Arbor, Michigan
CONTRIBUTIONS TO THE THEORY OF
INFINITE INVARIANT MEASURES

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Eugene Martin Klinko, B.S., M.S.

The Ohio State University
1967

Approved by
Adviser
Department of Mathematics
ACKNOWLEDGMENT

The author wishes to express his deep gratitude to his teacher, Professor Louis Sucheston, for the personal attention and interest given to the development of the author's mathematical career, secondly for suggesting the topics studied in this dissertation, and finally for endless discussions and encouragement.

The author is indebted to Professor P. V. Reichelderfer for valuable criticism and comments.

The author would also like to thank Mr. Lawrence A. Klimko for his assistance in reading this paper and Miss Loretta Klimko for typing the manuscript.
VITA

March 13, 1939 . . . . Born — Youngstown, Ohio

1961 . . . . B.S., The Ohio State University Columbus, Ohio

1961-1962 . . . . Assistant instructor, Department of Mathematics, The Ohio State University, Columbus, Ohio

1964 . . . . M.S., The Ohio State University Columbus, Ohio

1965-1967 . . . . Teaching Assistant, Department of Mathematics, The Ohio State University, Columbus, Ohio

PUBLICATION


MAJOR FIELDS OF STUDY

Probability Theory  Professor Louis Sucheston
Complex Variable Theory  Professor F. W. Carroll
Real Variable Theory  Professor P. V. Reichelderfer
## CONTENTS

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II.</td>
<td>ERGODIC THEOREMS</td>
<td>6</td>
</tr>
<tr>
<td>III.</td>
<td>THE GLIVENKO-CANTELLI THEOREM FOR INFINITE INVARIANT MEASURES</td>
<td>15</td>
</tr>
<tr>
<td>IV.</td>
<td>RATIO ERGODIC THEOREM IN INFORMATION THEORY</td>
<td>26</td>
</tr>
<tr>
<td>V.</td>
<td>WEAKLY WANDERING SETS OF INFINITE MEASURE</td>
<td>37</td>
</tr>
<tr>
<td>VI.</td>
<td>WEAKLY WANDERING FUNCTIONS</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>REFERENCES</td>
<td>54</td>
</tr>
</tbody>
</table>
I. INTRODUCTION

A *sigma-finite measure space* is a triple \((\Omega, \mathcal{F}, \mu)\) in which \(\Omega\) is an abstract set, \(\mathcal{F}\) is a *sigma-field* and \(\mu\) is a *sigma-finite* measure on \(\mathcal{F}\). That is: 
- \(\mathcal{F}\) is a family of subsets of \(\Omega\) containing \(\Omega\) and closed under the formation of complements and countable unions;
- \(\mu\) is a non-negative countably additive set-function on \(\mathcal{F}\);
- there is a countable decomposition \((\Omega_n)_{n=1}^\infty\) of \(\Omega\) such that \(\Omega_n \in \mathcal{F}\) and \(\mu(\Omega_n) < \infty\) for \(n \geq 1\).

For any set \(A\), \(\Omega - A\) is denoted by \(A^c\) and \(1_A\) is the indicator function of \(A\). All sets introduced in this paper are assumed measurable; that is, are elements of \(\mathcal{F}\). A *null set* has \(\mu\)-measure 0; all set relations in this paper are, unless otherwise stated, assumed to hold modulo null sets; that is, the relation holds except on a null set. However, the words a.e. (almost everywhere) will be sometimes stated for emphasis. A mapping \(\tau\) from \(\Omega\) into \(\Omega\) is measurable if \(\tau^{-1}\mathcal{F} \subseteq \mathcal{F}\); \(\tau\) is invertible if \(\tau\) is one to one and \(\tau^{-1}\) is measurable.

A wandering set \(W\) is disjoint from all of its preimages \(\tau^{-n}W\), \(n > 0\). A *conservative* transformation has no wandering sets of positive measure. \(\tau\) is ergodic if the
invariant sigma-field $\mathcal{Y} = \{ A : \tau^{-1} A = A \}$ is trivial; i.e. $A \in \mathcal{Y}$ implies $\mu(A) = 0$ or $\mu(A^c) = 0$. We shall only consider transformations which are null-preserving; i.e. $\mu(A) = 0$ implies $\mu(\tau^{-1} A) = 0$.

An invariant measure $\mu$ is a measure $\mu$ such that $\mu(A) = \mu(\tau^{-1} A)$ for each $A \in \mathcal{F}$, in which case we will also say that $\tau$ preserves $\mu$ or that $\tau$ is measure preserving (with respect to $\mu$).

A sequence of real-valued $\mathcal{F}$-measurable functions $(X_n)_{n=0}^{\infty}$ is stationary if $X_0$ is $\mu$-integrable and for every integer $n > C$, every $n$-dimensional interval $B$ (product of intervals of the real line), and every finite sequence of integers $i_1, \ldots, i_n$, the following relation holds:

$$
\mu[(X_{i_1} + h, \ldots, X_{i_n} + h) \in B] = \\
\mu[(X_{i_1}, \ldots, X_{i_n}) \in B].
$$

In a probabilistic setup ($\mu(\Omega) = 1$), measurable functions are random variables and a stationary sequence is called a strictly stationary stochastic process.

Given a $\mu$-integrable function $X_0$ and a measure-preserving point transformation $\tau$, we can easily generate a stationary sequence by means of the formula
\( x_n = x_0 \circ T^n \) for \( n = 1, 2, \ldots \). In the converse situation, given a stationary sequence \( (x_n)_{n=0}^\infty \) on a probability space \((\Omega, \mathcal{F}, \mu)\), it is possible to construct an auxiliary probability space \((\Omega', \mathcal{F}', \mu')\), called the sample space, on which is defined an invertible measure-preserving point transformation \( \tau' \) and a random variable \( X_0'\) such that the random variables \( x_n' = x_0' \circ T^n, n = 1, 2, \ldots \) are indistinguishable from the corresponding \( x_n \) in so far as their distribution properties are concerned. That is, for any \( n \)-dimensional Borel set \( B \), and any set of indices \( i_1, \ldots, i_n \), we have

\[
(1.2) \quad \mu[(x_{i_1}, \ldots, x_{i_n}) \in B] = \mu'[(x'_{i_1}, \ldots, x'_{i_n}) \in B].
\]

The measure thus introduced on \( \mathbb{R}^n \) is called the joint distribution of \( x_{i_1}, \ldots, x_{i_n} \). A detailed discussion of the sample space can be found in the book of Doob [8], p. 452 ff.; p. 617 ff.. We point out that null-recurrent Markov chains and Markov processes satisfying the Harris condition (see Harris [16]) give rise to point transformations on sample spaces on which the measure \( \mu \) is sigma-finite (see Harris and Robbins [17] and also Kakutani and Parry [21]).
The discussion of the preceding paragraph motivates the consideration of a general setting in which a stationary sequence is determined by an integrable function $X_0$ and a measure-preserving point transformation $\tau$ on a sigma-finite measure space. Chapter II contains various well known ergodic theorems which serve as basic tools in the succeeding two sections. We prove, in Chapter III, an extension of the Glivenko-Cantelli theorem to infinite measure spaces. In Chapter IV, a ratio version of the Shannon-McMillan-Breiman (see [1], [3], [20], [6], also [26]) theorem of information theory is proved. The results of the last two chapters are independent of those of Chapters III and IV and are described in the next paragraph.

Two measures $\mu$ and $\pi$ are equivalent if they both have the same null sets. If there exists an increasing sequence of positive integers $(n_i)_{i=0}^\infty$ with $n_0 = 0$ and such that the sets $W, \tau^{-n_1}W, \tau^{-n_2}W, \ldots$ are all mutually disjoint, then $W$ is weakly wandering under the sequence $(n_i)_{i=0}^\infty$. In Chapter V, we give a simple direct proof of a result of Hajian [10] concerning the existence of weakly wandering sets of infinite measure.
Finally, in Chapter VI, we obtain results analogous to those of Chapter V for operators determined by a certain class of matrices. Many of the results in the theory of point transformations are contained in general operator theory; however, so far as the author knows, the question remains open whether or not the results of Chapter V are contained in general operator theorems.
II. ERGODIC THEOREMS

We resume here ergodic theorems which are basic to the results of Chapters III and IV. The basic theorem for point transformations is the Hopf ergodic theorem which is contained in the Chacon-Ornstein operator ergodic theorem. We present the results in their historical order indicating how the various theorems contain their predecessors.

Given any sigma-field $\mathcal{G} \subseteq \mathcal{F}$ on which $\mu$ is sigma-finite, and any integrable function $f$, we ascribe to $E(f \mid \mathcal{G})$ its usual meaning of the Radon-Nikodym derivative of the measure $\gamma(A) = \int_A f \, d\mu$, taken with respect to the restriction of $\mu$ to $\mathcal{G}$. That is, $E(f \mid \mathcal{G})$ is a $\mathcal{G}$-measurable function having the property that for every $C \in \mathcal{G}$,

$$\int_C f \, d\mu = \int_C E(f \mid \mathcal{G}) \, d\mu;$$

$E(f \mid \mathcal{G})$ is unique modulo $\mathcal{G}$-null sets. Actually, to define conditional expectation, it suffices to have the measure $\gamma$ sigma-finite on $\mathcal{G}$ rather than have $f$ integrable.

In 1931 it was observed by Koopman that by means of the formula

$$E(f \mid \mathcal{G}) = E(f \mid \mathcal{G}^c);$$

$$E(f \mid \mathcal{G}) = E(f \mid \mathcal{G}^c);$$

$$(Tf)(\omega) = f(\tau\omega),$$

6
a measure preserving point transformation \( \tau \) generates an isometry on the space \( L_2(\mu) \) of functions \( f \) for which \( f^2 \) is \( \mu \)-integrable. That is, \( \tau \) generates a linear mapping \( T: L_2 \to L_2 \) having the property that (i) \( T(af + bg) = aTf + bTg \) for any pair \( a, b \) of reals and any \( f, g \in L_2 \) and (ii) the \( L_2 \)-norm of \( f: \|f\|_2 = \left[ \int \Omega |f|^2 \, d\mu \right]^{1/2} \), is the same as the norm of its image \( Tf \). This observation by Koopman led von Neumann to prove the \( L_2 \) mean convergence of

\[
(2.3) \quad \frac{1}{n} \sum_{i=0}^{n-1} f(\tau^i \omega) .
\]

In 1932, Birkhoff proved the pointwise convergence of (2.3) for \( f \in L_1 \). The theorem which bears his name, also called the Individual Ergodic Theorem (see [53] p. 18), is now stated:

**Theorem 2.1 (Birkhoff).** Let \( \tau \) be a measure preserving conservative point transformation and let \( f \in L_1 \). Then the limit

\[
(2.4) \quad f^*(\omega) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(\tau^i \omega)
\]

exists almost everywhere. Moreover, \( f^* \) is \( \mathcal{M} \)-mea-
surable and belongs to $L_1$. Moreover, if $\mu$ is finite, then $\int_C f \, d\mu = \int_C f^* \, d\mu$ for every $C \in \mathcal{C}$.

If $\tau$ is ergodic, the only $\mathcal{C}$-measurable functions are constants, which implies that on infinite sigma-finite measure spaces, $f^* = 0$ a.e. ($0$ being the only integrable constant). This relatively uninformative situation was remedied by a ratio limit theorem proved in 1937 by E. Hopf ([18] p. 49), which we now state.

**Theorem 2.2 (Hopf).** Let $\tau$ be a measure-preserving conservative point transformation, $f, g \in L_1$ and $g \geq 0$. If $\sum_{i=0}^{\infty} g(\tau^i \omega) > 0$ a.e., then the limit

$$ (2.5) \quad h(f, g)(\omega) = \lim_{n \to \infty} \frac{\sum_{i=0}^{n-1} f(\tau^i \omega)}{\sum_{i=0}^{n-1} g(\tau^i \omega)} $$

exists a.e., is $\mathcal{C}$-measurable and $\int_C f \, d\mu = \int_C h(f, g)g \, d\mu$ for every $C \in \mathcal{C}$.

These theorems were originally proved under the assumption that $\tau$ is invertible; however, this assumption is not necessary. Halmos [12] and
Hurewicz [19], for invertible transformations, relaxed the requirement that \( \tau \) be measure preserving and obtained a more general theorem than that of Hopf (see also [25]). Let \( \phi_n \) be the unique (modulo null sets) function such that for each \( A \in \mathcal{F} \), the following relation holds

\[
\mu(\tau^n A) = \int_A \phi_n \, d\mu.
\]

We now state the Halmos-Hurewicz ergodic theorem.

**Theorem 2.3 (Halmos-Hurewicz).** Let \( \tau \) be invertible, \( f \in L_1 \) and let \( g \geq 0 \) be measurable and such that

\[
\sum_{i=0}^{\infty} \phi_i(\omega)g(\tau^i \omega) = \infty \quad \text{a.e.}
\]

Then the limit

\[
h(f, g)(\omega) = \lim_{n \to \infty} \frac{\sum_{i=0}^{n-1} \phi_i(\omega)f(\tau^i \omega)}{\sum_{i=0}^{n-1} \phi_i(\omega)g(\tau^i \omega)}
\]

exists a.e.. Furthermore, \( h(f, g) \) is \( J \)-measurable, \( h(f, g)g \) is integrable and

\[
\int_C f \, d\mu = \int_C h(f, g)g \, d\mu
\]

for every \( C \in \mathcal{F} \) for which \( \int_C g \, d\mu < \infty \).

When \( \tau \) is invertible, the Halmos-Hurewicz theorem implies the Hopf theorem, since \( \phi_i = 1 \) for \( i = 0, 1, \ldots \) when \( \tau \) is measure preserving. The Halmos-
Hurewicz theorem also shows that the assumption $g \in L^1$ in the Hopf theorem is not necessary; it suffices to assume that $g \geq 0$ and measurable. The last assertion of the Hopf theorem must then be replaced by the last assertion of the Hurewicz–Halmos theorem. The preceding extension is also valid when $\tau$ is not invertible. It is now seen that with $g(\omega) = 1$ for all $\omega \in \Omega$, the Hopf theorem becomes the Birkhoff theorem.

We now turn our attention to operator ergodic theory. We will state without proof the Chacon–Ornstein ergodic theorem which contains the Hopf and the Halmos-Hurewicz theorems. We also will use the Chacon-Ornstein theorem to obtain an alternate identification of the limit in the preceding theorems.

A mapping $T : L^1 \to L^1$ is **positive** if $f \geq 0$ implies $Tf \geq 0$; $T$ is a **contraction** if for each $f \in L^1$, $|Tf|_1 \leq |f|_1$. A contraction $T$ is **conservative** if for every $L^1$-function $f \geq 0$, the set on which $T_\infty f \neq 0$ or $\infty$ is $\mu$ null, where $T_\infty f \overset{\text{def}}{=} \sum_{i=0}^{\infty} T^i f$. We have already pointed out that measure preserving point transformations generate isometries on $L^2$; (2.2) also generates an isometry on $L^1$ which
is also clearly a positive contraction. In order to relate the notions of conservative operators and conservative point transformations, we shall require the following recurrence theorem due to Halmos [13] and to Sucheston [30]. For each set $A$, let $A^\tau = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} \tau^{-i}A$; $A^\tau$ is the set of points which return "infinitely often" to $A$ under $\tau$.

**Theorem 2.4 (Recurrence Theorem).** $\tau$ is conservative if and only if $\mu(A - A^\tau) = 0$.

**Lemma 2.1.** Let $\tau$ generate an operator $T$ by means of (2.2). Then $T$ is conservative if and only if $\tau$ is.

**Proof.** Let $W$ be a wandering set of positive measure. Then on the set $\bigcup_{i=0}^{\infty} \tau^{-i}W$, $T^\infty W = 1$ and thus $T$ cannot be conservative. On the other hand, it follows by the Recurrence Theorem that for any set $B$ of positive measure, $T^\infty 1_B = \infty$ on $B$. Consider any $L_1$-function $f \geq 0$. Suppose now that $T^\infty f > 0$ on $A$, a set of positive measure. Then there is a non-null set $B \subset A$, a $\delta > 0$, and an integer $i_0 \geq 0$ such that $T^{i_0}f \geq \delta 1_B$. This implies however,
that $T^\infty f \geq \delta T^\infty T^0 1_B = \infty$ a.e. on $B$. Therefore, on every set of positive measure, $T^\infty f = 0$ or $\infty$ which implies that $T$ is conservative.

If $A$ is such that $f = 0$ on $A^c$ implies that $Tf = 0$ on $A^c$, then $A$ is an invariant set for $T$. When $T$ is conservative, the preceding definition is one of several equivalent definitions of an invariant set for $T$ (see [24] p. 196). We next relate the notions of $T$-invariant sets and $\tau$-invariant sets.

**Lemma 2.2.** When $T$ is generated by a conservative point transformation $\tau$, the class of $T$-invariant sets coincides with the class of $\tau$-invariant sets.

**Proof.** Let $A$ be $\tau$-invariant. If $f \in L_1$ and $f = 0$ on $A^c$, then $f = 1_A \cdot f$ and $(Tf)(\omega) = 1_A \cdot f(\tau \omega) = 0$ on $A^c$. Conversely, let $A$ be $T$-invariant and let $f > 0$ a.e. and $f \in L_1$. Then $1_A \cdot f \in L_1$ and $1_A \cdot f = 0$ on $A^c$ hence $T(1_A \cdot f) = 0$ on $A^c$. Since $f > 0$ and $T(1_A \cdot f) = 1_A \cdot (f \circ \tau)$, $1_A \cdot f \geq 1_A$. Since $\tau$ is conservative, the wandering set $A - \tau^{-1}A$ has measure zero and $\tau^{-1}A = A$.

We now state the Chacon-Ornstein operator ergodic theorem [5].
Theorem 2.5 (Chacon-Crnstein). Let $T$ be a conservative positive contraction on $L^1$. Let $f, g \in L^1$ and $g \geq 0$. Then

\[
\lim_{n \to \infty} \sum_{i=0}^{n-1} T^i f
\]

exists on the set $C$ where $T^\infty g > 0$. Denote the limit of (2.8) by $h(f, g)$. $h(f, g)$ is $C \cap J$-measurable and $\int_C f \, du = \int_C h(f, g) g \, du$ for every $C \in \mathcal{C} \cap C$.

Remarks. 1. The preceding theorem contains the Hopf theorem without the invertibility assumption since the operator $T$, where $Tf = f \circ \tau$, is a positive isometry on $L^1$ and the conservativity of $\tau$ implies that of $T$.

2. When $\tau$ is conservative, the Chacon-Crnstein theorem implies the Halmos-Hurewicz theorem provided $g \in L^1$. Indeed, the mapping $T$, where $(Tf)(\omega) = \phi_1(\omega)f(\tau \omega)$ is a positive isometry on $L^1$; the relation

$\phi_{i+j}(\omega) = \phi_i(\tau^j \omega) \phi_j(\omega)$

is proved in [12]; it implies that $(T^n f)(\omega) = \phi_n(\omega)f(\tau^n \omega)$ for $n = 0, 1, \ldots$. 
3. The assumption \( \sum_{i=0}^{\infty} g(\tau^i \omega) > 0 \) a.e. in the Hopf theorem is not necessary if the assertion is weakened to: the convergence holds on the set where the sum \( \sum_{i=0}^{\infty} g(\tau^i \omega) \) is greater than zero. The corresponding statement is valid for the Halmos-Hurewicz theorem.

We now obtain a rewording of the identification of the limit in Theorem 2.5. The proof of Corollary 2.1 is taken from [4].

**Corollary 2.1.** If \( \mu \) is sigma finite on \( \mathcal{C} \), then

\[
(2.9) \quad h(f, g) = \frac{E(f | \mathcal{C})}{E(g | \mathcal{C})}.
\]

**Proof.** From the relation \( \int_C f \, d\mu = \int_C h(f, g)g \, d\mu \); it follows that \( \int_C E(f | \mathcal{C}) \, d\mu = \int_C E(h(f, g)g | \mathcal{C}) \, d\mu \) for every \( C \in \mathcal{C} \). This implies that \( E(f | \mathcal{C}) = E(h(f, g)g | \mathcal{C}) = h(f, g)E(g | \mathcal{C}) \). The last inequality holding because the \( \mathcal{C} \)-measurable function \( h(f, g) \) can be factored from the conditional expectation (see Theorem [22] p. 350 where it is proved in case \( \mu \) is a probability; the theorem carries over to the sigma-finite case.).
III. THE GLIVENKO-CANTELLI THEOREM FOR INFINITE INVARIANT MEASURES

Let \( \mu \) be a probability measure on \( \mathcal{F} \) and let \( X_0, X_1, \ldots \) be a sequence of random variables. The sequence is \textit{independent} if for every sequence \( B_1, B_2, \ldots, B_n \) of Borel subsets of the real line and for every sequence of integers \( i_1, \ldots, i_n \)

\[
\mu\left[ \bigcap_{j=1}^{n} X^{-1}_{i_j}(B_j) \right] = \prod_{j=1}^{n} \mu[X^{-1}_{i_j}(B_j)].
\]

The distribution function \( F(x) \) of a random variable \( X \) is defined for each real \( x \) as \( F(x) = \mu[X < x] \). A sequence of independent random variables is \textit{identically-distributed} if all of the random variables in the sequence have the same distribution function. Let

\[
F_n(x) = \mathbb{1}_{(-\infty, x]} X_n \quad \text{for } n = 0, 1, \ldots.
\]

The experimental distribution function of a sample of size \( n \)

\[
\frac{1}{n} \sum_{i=0}^{n-1} F_1(x)
\]

is the average number of times \( X_i < x \) for \( i = 0, \ldots, n-1 \). The relationship between the experimental distribution function and the distribution function is given by the Glivenko-Cantelli theorem which follows
(see [22] p. 20 and [28] p. 335).

**Theorem 3.1 (Glivenko-Cantelli).** Let $X_0, X_1, ...$ be a sequence of independent identically distributed random variables, $X_0 \in L_1$. Then $\lim_{n \to \infty} \Delta_n = 0$ a.e.,

where

$$
(3.3) \quad \Delta_n = \sup_{-\infty < x < \infty} \left| \frac{1}{n} \sum_{i=0}^{n-1} F_i(x) - F(x) \right|.
$$

The Glivenko-Cantelli theorem has been generalized by various authors. Fortet and Mourier [93] considered a Banach space setting and allowed more general functions than indicator functions in the expression for $F_n(x)$. They also note that the sequence $X_0, X_1, ...$ need only be a strictly stationary stochastic process. Wolfowitz [34] obtained generalizations in several directions (see also Blum [2]); his results, however, are in part contained in those of Fortet and Mourier. Tucker [32] proved the theorem for strictly stationary stochastic processes in the non-ergodic case. Various forms of the theorem have also been proved in a metric space setting by Varadarajan [33] and Ranga Rao [27].

There are two difficulties encountered in formulating the Glivenko-Cantelli theorem for infinite invariant measures: (1) there is no distribution
function, and (2) the experimental distribution

\[
\frac{1}{n} \sum_{i=0}^{n-1} 1_B \circ X_i \to 0 \quad \text{for every Borel set } B \text{ for which } \mu(X_0^{-1}B) < \infty.
\]

The last assertion follows from the Birkhoff ergodic theorem. The first of these difficulties is easily overcome (see (3.5) below) while the second is eliminated by consideration of ratios.

We now allow \( \mu \) to be sigma-finite and assume that \( \tau \) is ergodic and conservative. Let \( X_0, Y_0 \), be fixed real-valued measurable functions on \( \Omega \) and let \( X_n = X_0 \circ \tau^n, \ Y_n = Y_0 \circ \tau^n, \ n = 1, 2, \ldots \). If \( s, x, t, y \) are extended real numbers, for \( n = 0, 1, \ldots \), let

\[
F_n(x) = 1_{(s,x)} \circ X_n, \ G_n(y) = 1_{(t,y)} \circ Y_n,
\]

and let

\[
F(x) = \int_{\Omega} F_0(x) \mu(d\omega), \ G_t(y) = \int_{\Omega} G_0(y) \mu(d\omega).
\]

Our theorem asserts that the ratio

\[
\frac{\sum_{k=0}^{n-1} F_k(x)}{\sum_{k=0}^{n} G_k(y)}
\]

converges almost everywhere uniformly in \((x, y)\), which is however restricted to a set on which \( F_s \), \( G^t \) behave with some moderation.
Theorem 3.2. Let \( s, t \in \mathbb{R} \) (extended real line). Let \( C \) and \( D \) be sets in \( \mathbb{R} \) such that for some positive constants \( c, d \)

\[
C = \{ x : F^s(x) \leq c \}, \quad D = \{ y : G^t(y) \geq d \}.
\]

Let \( B = C \times D \) and

\[
(3.7) \quad \triangle_n = \sup_{(x,y) \in B} \left| \frac{\sum_{i=0}^{n-1} F_i^s(x)}{n} - \frac{F^s(x)}{n} \right|.
\]

Then for almost all \( \omega \in \Omega \)

\[
(3.8) \quad \lim_{n \to \infty} \triangle_n = 0.
\]

We note that Theorem 3.2 implies the Glivenko-Cantelli theorem. Let \( \mu \) be a probability measure and let \( X_0 = Y_0 \). Further set \( s = t = -\infty \) and \( c = d = 1 \). Then the denominator in the first ratio in (3.7) is simply \( n \) and Theorem 3.2 asserts the uniform convergence a.e. of the experimental distribution function \( \frac{1}{n} \sum_{i=0}^{n-1} F_i^\infty(x) \) of a strictly stationary ergodic process \((X_n)_{n=0}^\infty\), to the distribution function \( F^\infty(x) \) of \( X_0 \). Indeed, a stationary process on a probability space gives rise to a measure-
preserving (hence conservative) point-transformation on the sample probability space. The uniform convergence a.e. of experimental distribution functions carries over from the second space to the first one.

We first prove Theorem 3.2. Then we show how Theorem 3.2 extends to the non-ergodic case: the second ratio in (3.7) is then replaced by a ratio of "conditional distribution functions".

Proof of Theorem 3.2. Since $s$ and $t$ remain fixed throughout the proof, we omit the superscripts from $F_s, F_n, G^t,$ and $G^t_n$. We may assume that $c = \sup_{x \in C} F(x)$ and $d = \inf_{y \in D} G(y)$. We next let $M$ and $j$ be positive integers with $0 < j < M$ and we form a partition of $B$ by letting $x_{Mj}$ and $y_{Mj}$ be the smallest real numbers such that

$$F(x_{Mj}) \leq \frac{j}{M} \leq F(x_{Mj} + 0),$$

$$\frac{1}{G(y_{Mj} + 0)} \leq \frac{M - j}{dM} \leq \frac{1}{G(y_{Mj})},$$

and set $x_{Mo} = \inf\{x \in C\}, x_{MM} = \sup\{x \in C\}, y_{Mo} = \inf\{y \in D\}, y_{MM} = \sup\{y \in D\}$. For each pair $(x, y) \in B$, we define
\[(3.10) \quad \delta_n(x,y) = \begin{vmatrix} \sum_{i=0}^{n-1} F_i(x) & - & F(x) \\ \sum_{i=0}^{n-1} G_i(y) & G(y) \end{vmatrix} \]

From (3.4) the definition of \( F_n \) and \( G_n \), it follows that

\[(3.11) \quad F_n(x) = 1_{(s,x)} X_o \circ \tau^n, \quad G_n(y) = 1_{(t,y)} Y_o \circ \tau^n.\]

Considering \( F_n \), \( G_n \) as functions of \( \omega \), we have for fixed \( x, y \)

\[(3.12) \quad F_n(x) = F_0(x) \circ \tau^n, \quad G_n(y) = G_0(y) \circ \tau^n.\]

It follows from (3.4) and Hopf's ergodic theorem (Theorem 2.2 and Corollary 2.1) that for fixed \((x, y) \in B\)

\[(3.13) \quad \lim_{n \to \infty} \delta_n(x, y) = 0 \quad \text{a.e.} \]

The preceding argument is also valid with \( x \) or \( y \)
in (3.13) replaced by \( x + 0 \) or \( y + 0 \), respectively, provided that \( x \neq x_Mi \) or \( y \neq y_Mi \) . (3.13) also holds when \( x = x_Mo + 0 \) or \( y = y_Mo + 0 \). Let \((x, y) \in B\)
with \( x \neq x_Mo \), \( y \neq y_Mo \). Then there is an \( i \) and a \( j \) such that \( x_{M,i-1} < x < x_Mi \) and \( y_{M,j-1} < y < y_Mj \).

The monotonicity of \( F, F_n, G, G_n \) implies that

\[(3.14) \quad \frac{F(x_{M,i-1} + 0)}{G(y_{Mj})} \leq \frac{F(x)}{G(y)} \leq \frac{F(x_{Mi})}{G(y_{M,j-1} + 0)}\]

and
\[
\sum_{k=0}^{n} F_k(x_{M,i-1} + 0) \leq \sum_{k=0}^{n} F_k(x) \leq \sum_{k=0}^{n} F_k(x_{M,i}) ;
\]
\[
\sum_{k=0}^{n} G_k(y_{M,j}) \leq \sum_{k=0}^{n} G_k(y) \leq \sum_{k=0}^{n} G_k(y_{M,j-1} + 0)
\]

hence

\[
\sum_{k=0}^{n} \frac{F_k(x)}{G_k(y)} \leq \sum_{k=0}^{n} \frac{F_k(x_{M,i})}{G_k(y_{M,j-1} + 0)} - \frac{F(x_{M,i})}{G(y_{M,j-1} + 0)} + \frac{F(x_{M,i-1} + 0)}{G(y_{M,j})}.
\]

By \((3.9)\), the last difference in \((3.16)\) is bounded by

\[
\frac{ic \cdot M - j + 1}{M \cdot dM} - \frac{(i-1)c \cdot M - j}{M \cdot dM} \leq \frac{2c}{dM}.
\]

An inequality similar to \((3.16)\), giving a lower bound for the left side of \((3.16)\) is obtained, and the two inequalities together yield:

\[
\delta_n(x, y) \leq 2c/dM +
\]
\[
+ \max[\delta_n(x_{M,i}, y_{M,j-1} + 0), \delta_n(x_{M,i-1} + 0, y_{M,j})].
\]

In case \(x = x_{M_0}\), a computation similar to the previous one shows that \((3.17)\) holds with \(x_{M,i}\) and \(x_{M,i-1}\) both replaced by \(x_{M_0}\); similarly if \(y = y_{M_0}\). Hence if \(y_{M_0} \notin D\), an upper bound for \(\Delta_n = \sup_{(x,y)B} \delta_n(x, y)\) is given by:

\[
\max_{i=1,2} \left[ \Delta_n^{(i)} + \frac{2c}{dM} \right]
\]

where
\[ \Delta_{n,M}^{(1)} \overset{\text{def}}{=} \max_{0 \leq i \leq M} \delta_n(x_{Mi}, y_{Mj} + 0) \]

(3.19)

\[ \Delta_{n,M}^{(2)} \overset{\text{def}}{=} \max_{0 \leq i \leq M} \delta_n(x_{Mi} + 0, y_{Mj}) \]

while, if \( y_{Mo} \in \mathcal{D} \), we allow \( j = 0 \) in the definition of \( \Delta_{n,M}^{(2)} \). It follows from (3.13) that for \( i = 1, 2 \), each \( M \) and almost every \( \omega \in \Omega \), \( \lim_{n} \Delta_{n,M}^{(i)} = 0 \), and therefore \( \limsup \Delta_n \leq 2c/dM \). Since \( M \) is arbitrary, it follows that \( \lim \Delta_n = 0 \) almost everywhere, which completes the proof of the theorem.

The non-ergodic case. We now show that Theorem 3.2, with suitable modifications, remains true even though the invariant sigma-field \( \mathcal{J} \) is not trivial. We assume that \( \mu \) is sigma-finite on \( \mathcal{J} \). For each \( s, x, t, y \in \mathbb{R} \), we define

\[ F^s(x|\mathcal{J}) = E(l(s,x)^o X_0|\mathcal{J}) \]

(3.20)

\[ G^t(y|\mathcal{J}) = E(l(t,y)^o Y_0|\mathcal{J}) \]  

As in the case of conditional probability distributions, we may and do assume that for every \( \omega \in \Omega \), \( F^s(x|\mathcal{J}) \) and \( G^t(y|\mathcal{J}) \) are: (i) nondecreasing in \( x \) \( (y) \), (ii) left-continuous, and (iii) \( F^s(s|\mathcal{J}) = G^t(t|\mathcal{J}) = 0 \). The existence of such functions follows from the fact that we may compute the conditional expectations in (3.20) with respect to a probability measure \( \pi \).
Indeed, let \( \pi \) be a \( \mu \)-equivalent probability measure for which the Radon-Nikodým derivative \( \alpha = d\mu / d\pi \) is \( \mathcal{F} \)-measurable and \( \alpha > 0 \) for every \( \omega \in \Omega \). We let subscripts denote the measure with respect to which conditional expectation is computed and obtain

\[
E_{\mu}(1_B \circ X_0 | \mathcal{F}) = E_{\pi}(\alpha 1_B \circ X_0 | \mathcal{F}) / \alpha = E_{\pi}(1_B \circ X_0 | \mathcal{F})
\]

for every Borel set \( B \). By Theorem B [22] p. 363, there is a regular version \( Q(\omega, B) \) of \( E_{\pi}(1_B \circ X_0 | \mathcal{F}) \); that is, (i) for every \( \omega \in \Omega \), \( Q(\omega, B) \) is a probability measure on the Borel sets of the real line, (ii) for every Borel set \( B \), \( Q(\omega, B) \) is \( \mathcal{F} \)-measurable and (iii) for every Borel set \( B \), \( Q(\omega, B) = E_{\pi}(1_B \circ X_0 | \mathcal{F}) \) \( \text{a.e. } \pi \). The required functions in (3.20) may then be taken to be

\[
F^s(x | \mathcal{F}) = Q(\omega, (s, x)) \\
G^t(y | \mathcal{F}) = Q(\omega, (t, y)).
\]

In the sequel, \( F^s(x) \) and \( G^t(y) \) are assumed to be replaced in \( \Delta_n \), \( \delta_n(x, y) \) etc. by \( F^s(x | \mathcal{F}) \) and \( G^t(y | \mathcal{F}) \) respectively. The proof of the next theorem uses an idea of Tucker [32] and is an extension of his result.

**Theorem 3.3.** Let \( s, t \in \bar{\mathbb{R}} \) and let \( C \) and \( D \) be sets in \( \bar{\mathbb{R}} \) such that for some positive \( \text{a.e. finite-valued } \mathcal{F} \)-measurable functions \( c(\omega) \) and \( d(\omega) > 0 \),

\[
C = \{ x : F^s(x | \mathcal{F}) \leq c(\omega) \} \quad D = \{ y : G^t(y | \mathcal{F}) \geq d(\omega) \},
\]
the inequalities holding for all $\omega$ outside of a null set $N$ independent of $x, y$. Let $B = C \times D$. Then for almost all $\omega \in \Omega$,

$$
\lim_{n \to \infty} \Delta_n = 0.
$$

Proof. The proof is similar to that of Theorem 3.2 and we merely sketch it, indicating the essential changes. We may and do assume that for every $\omega$,

$$
c(\omega) = \sup_{x \in C} F(x \mid \mathcal{I})(\omega), \quad d(\omega) = \inf_{y \in D} G(y \mid \mathcal{I})(\omega).
$$

Let $M$ and $j$ be integers with $0 < j < M$. We define $\mathcal{I}$-measurable functions $X_{Mj}$ and $Y_{Mj}$ by letting for each fixed $\omega$, $X_{Mj}$ and $Y_{Mj}$ be the smallest real numbers for which

$$
F(X_{Mj} \mid \mathcal{I}) \leq \frac{\mathcal{J}c(\omega)}{M} \leq F(X_{Mj} + 0 \mid \mathcal{I})
$$

(3.25)

$$
\frac{1}{G(Y_{Mj} + 0 \mid \mathcal{I})} \leq \frac{M - j}{Md(\omega)} \leq \frac{1}{G(Y_{Mj} \mid \mathcal{I})},
$$

and set $X_{M0} = \inf \{ x \in C \}$, $X_{MM} = \sup \{ x \in C \}$, $Y_{M0} = \inf \{ y \in D \}$, $Y_{MM} = \sup \{ y \in D \}$. $\mathcal{I}$-measurable functions are shift invariant; since $X_{Mj}$, $Y_{Mj}$ are $\mathcal{I}$-measurable,

$$
\tau^{-1}[s < X_n < X_{Mj}] = [s < X_{n+1} < X_{Mj}]
$$

(3.26)

$$
\tau^{-1}[t < Y_n < Y_{Mj}] = [t < Y_{n+1} < Y_{Mj}],
$$

and therefore we can write

(3.27) $F_n(X_{Mj}) = F_\circ(X_{Mj}) \circ \tau^n$, $G_n(Y_{Mj}) = G_\circ(Y_{Mj}) \circ \tau^n$.

From (3.27) we can conclude by Hopf's ergodic theorem
(Theorem 2.2 and Corollary 2.1) that the following holds

\[(3.28) \quad \lim_{n \to \infty} d_n(X_{M_i}, Y_{M_j}) = 0 .\]

For each fixed \( \omega \in \Omega \), we apply the argument used in the proof of Theorem 3.2 and obtain

\[(3.29) \quad \Delta_n \leq \max_{i=1,2} \Delta(i)_{n,M} + \frac{2c(\omega)}{Kd(\omega)}.\]

The theorem then follows by noting that \( c(\omega)/d(\omega) \) is finite valued and \( M \) is arbitrary.

**Remark.** One may ask whether the stationarity of the sequence \( (X_n)_{n=0}^{\infty} \) is essential. If \( \tau \) is invertible, we may drop the assumption that \( \tau \) is measure preserving provided that \( F_n^S \) and \( G_n^t \) are suitably weighted. Theorem 3.2 remains valid with \( (3.7) \) replaced by

\[(3.30) \quad \Delta_n = \sup_{(x,y) \in B} \left| \begin{array}{c}
\sum_{i=0}^{n-1} \phi_i F_i^S(x) - F^S(x) \\
\sum_{i=0}^{n-1} \phi_i G_i^t(y) - G^t(y)
\end{array} \right| ,\]

where \( \phi_i \) is defined by \( (2.6) \). Theorem 3.3 also remains valid when \( F^S(x) \) and \( G^t(y) \) in \( (3.30) \) are replaced by \( F^S(x|\mathcal{F}) \) and \( G^t(y|\mathcal{F}) \). Indeed, the role of Hopf's ergodic theorem in the proof of theorems 3.2 and 3.3 may be played by the Halmos-Hurewicz ergodic theorem (Theorem 2.3).
IV. RATIO ERGODIC THEOREM IN INFORMATION THEORY

The object of this chapter is to prove a ratio version of the Shannon-Breiman-McMillan ergodic theorem of information theory.

We begin by introducing some new terminology. Throughout this chapter, \( \tau \) will be assumed to be ergodic, conservative, invertible and measure preserving. For any sigma-field \( \mathcal{G} \subset \mathcal{F} \), on which \( \mu \) is sigma-finite, and for any set \( A \in \mathcal{F}, \mu(A) < \infty \), we define the conditional measure given \( \mathcal{G} \) by

\[
\mu(A \mid \mathcal{G}) = E(1_A \mid \mathcal{G}).
\]

If \( \mathcal{A} \) is a countable measurable partition of \( \Omega \) into sets of positive finite measure, \( I(\mathcal{A} \mid \mathcal{G}) \) the conditional information of \( \mathcal{A} \) with respect to \( \mathcal{G} \) is (logarithm has base 2)

\[
I(\mathcal{A} \mid \mathcal{G}) = -\sum_{A \in \mathcal{A}} 1_A \cdot \log \mu(A \mid \mathcal{G}).
\]

The absolute information \( I(\mathcal{A}) \) is

\[
I(\mathcal{A}) = -\sum_{A \in \mathcal{A}} 1_A \cdot \log \mu(A).
\]

We define the conditional entropy as

\[
H(\mathcal{A} \mid \mathcal{G}) = \int_{\Omega} I(\mathcal{A} \mid \mathcal{G}) \, d\mu.
\]
and similarly the **absolute entropy** is the integral of the absolute information. For any partition $\mathcal{A}$ and any set $B \in \mathcal{F}$, the entropy in $B$ is

\[
(4.5) \quad H(\mathcal{A} \cap B) = - \sum_{A \in \mathcal{A}} \mu(A \cap B) \log \mu(A \cap B).
\]

For any set $B \in \mathcal{F}$, the **conditional information in** $B$ is

\[
(4.6) \quad I_B(\mathcal{A} | \mathcal{C}) = - \sum_{A \in \mathcal{A}} 1_{A \cap B} \log \mu(A | \mathcal{C}) = 1_B \cdot I(\mathcal{A} | \mathcal{C}),
\]

and the **conditional entropy in** $B$ is

\[
(4.7) \quad H_B(\mathcal{A} | \mathcal{C}) = \int_{\mathcal{C}} I_B(\mathcal{A} | \mathcal{C}) \, d\mu.
\]

Similar relations define the **absolute information** and **absolute entropy in** $B$. If $(\mathcal{A}_n)_{n=0}^{\infty}$ is a sequence of countable measurable partitions of $\Omega$, we denote by $\bigvee_{i=0}^{m} \mathcal{A}_i$ the partition consisting of sets of the form $\bigcap_{i=0}^{m} A_i$ where $A_i \in \mathcal{A}_i$. The partition $\bigvee_{i=0}^{m} \mathcal{A}_i$ generates a sigma-field denoted by $\mathcal{F}(\bigvee_{i=0}^{m} A_i)$; this sigma field will, by a customary
abuse of notation, also be denoted by \( \bigcup_{i=0}^{m} A_i \).

The Shannon-Breiman-McMillan theorem asserts the mean and a.e. convergence of \( \frac{1}{n} \sum_{i=0}^{n-1} I(\bigcup_{i=0}^{n-1} A_i) \) when \( \mu \) is a probability measure on \( \mathcal{F} \). If \( \mu \) is allowed to be infinite, a difficulty arises in that the limit generally will be 0. We attempt to overcome this difficulty by using a ratio ergodic theorem.

We shall require the following version of the Martingale Convergence Theorem.

**Proposition 4.1.** Let \( G_n \) be an increasing sequence of sigma-fields. Let \( (\Omega_i)_{i=1}^{\infty} = C_1 \) be a decomposition of \( \Omega \) into sets of finite measure. If \( f \) is such that \( \int_{\Omega_i} |f| \, d\mu < \infty \) for \( i \geq 1 \), then

\[
\mathbb{E}(f \| G_n) \rightarrow \mathbb{E}(f \| G) \text{ a.e.}
\]

where \( G \) is the sigma-field generated by \( \bigcup_{i=1}^{\infty} C_i \).

**Proof.** For each fixed \( i \), let \( f_i, \mu_i, C_{in} \) and \( C_i \) be respectively, the restrictions of \( f, \mu, C_n \) and \( C \) to \( \Omega_i \). Then \( \mathbb{E}(f_i \| C_{in}) \) is a martingale on the finite measure space \( (\Omega_i, \Omega_i \cap \mathcal{F}, \mu_i) \).
and it follows from the Martingale Convergence Theorem ([24] p. 42) that

\[
(4.9) \quad E(f_i | \mathcal{G}_{in}) \xrightarrow{n} E(f_i | \mathcal{G}_i) \quad a.e. .
\]

It now will suffice to show that for each \( i \) and each \( n \),

\[
(4.10) \quad E(f_i | \mathcal{G}_{in}) = E(f|\mathcal{G}_n) \quad a.e. \text{ on } \Omega_i .
\]

Since \( \Omega_i \in \mathcal{G}_1 \subset \mathcal{G}_n \), it follows that \( \mathcal{G}_{in} \subset \mathcal{G}_n \) for \( i = 1, 2, \ldots ; n = 1, 2, \ldots \) and therefore for all \( B \in \mathcal{G}_{in} \)

\[
(4.11) \quad \int_B E(f_i | \mathcal{G}_{in}) \, d\mu_i = \int_B f \, d\mu = \int_B E(f|\mathcal{G}_n) \, d\mu .
\]

Since (4.11) holds for every subset \( B \in \mathcal{G}_{in} \), it follows that on \( \Omega_i \), \( E(f|\mathcal{G}_{in}) = E(f|\mathcal{G}_n) \) and similarly that \( E(f|\mathcal{G}_i) = E(f|\mathcal{G}) \).

**Remark.** Since the logarithm is a continuous function, it follows that \( I(A|\mathcal{G}_n) \xrightarrow{n} I(A|\mathcal{G}) \) pointwise.

The next theorem is a somewhat generalized version of Hopf's ergodic theorem; the generalization is analogous to the one obtained by Breiman [3] for Birkhoff's ergodic theorem.
Theorem 4.1. Let \( f_n, g_n \) be sequences of measurable functions such that \( \sup_n |f_n| \in L_1 \) and \( \sup_n |g_n| \in L_1 \). Further let \( f_n \to f \) a.e., \( g_n \to g \) a.e. with \( g \geq 0 \). Then if \( \int_{\Omega} g \, d\mu > 0 \),

\[
\lim_{n \to \infty} \frac{\sum_{i=0}^{n} f_i(\tau^{n-i} \omega)}{\sum_{i=0}^{n} g_i(\tau^{n-i} \omega)} = \frac{\int_{\Omega} f \, d\mu}{\int_{\Omega} g \, d\mu}.
\]

Proof. The assumptions on \( f_n \) and \( g_n \) imply that \( f, g \in L_1 \) and therefore it follows from Hopf's theorem that

\[
\lim_{n \to \infty} \frac{\sum_{i=0}^{n} f_i(\tau^{n-i} \omega)}{\sum_{i=0}^{n} g_i(\tau^{n-i} \omega)} = \frac{\int_{\Omega} f \, d\mu}{\int_{\Omega} g \, d\mu}.
\]

We next show that the following holds a.e.:

\[
\lim_{n \to \infty} \frac{\sum_{i=0}^{n} f_i(\tau^{n-i} \omega)}{\sum_{i=0}^{n} g_i(\tau^{n-i} \omega)} = 0.
\]

For each \( N \geq 0 \), we define \( F_N(\omega) = \sup_{n \geq N} |f_n(\omega) - f(\omega)| \).

Since \( F_N \geq F_{N+1} \) for each \( N \geq 0 \) and \( F_0 \leq \sup_{n \geq 0} |f_n| + |f| \), it follows that \( F_N \in L_1 \). We have
The last limit is, by Hopf's theorem, \( \int_{\Omega} F_N \, d\mu / \int_{\Omega} g \, d\mu \). Since \( F_N \to 0 \), (4.14) follows by an application of the monotone convergence theorem [24] p. 42.

By the previous argument with \( g_i \) replacing \( f_i \) we have

\[
\lim_{n \to \infty} \frac{\sum_{i=0}^{n} g_i(\tau^{n-i} \omega)}{\sum_{i=0}^{n} g(\tau^{n-i} \omega)} = 1 \text{ a.e.}.
\]

The theorem now follows by writing

\[
\frac{\sum_{i=0}^{n} f_i(\tau^{n-i} \omega)}{\sum_{i=0}^{n} g(\tau^{n-i} \omega)} = \frac{\sum_{i=0}^{n} f_i(\tau^{n-i} \omega)}{\sum_{i=0}^{n} g_i(\tau^{n-i} \omega)} \cdot \frac{\sum_{i=0}^{n} g(\tau^{n-i} \omega)}{\sum_{i=0}^{n} g(\tau^{n-i} \omega)}
\]

The following theorem is essentially due to Ionescu Tulcea [20], Breiman [3] and Chung [6].
Theorem 4.2. Let $\mathcal{C}_n$, $n = 0, 1, \ldots$ be an
increasing sequence of sigma-fields on which $\mu$ is
sigma finite. Let $\mathcal{A}$ be a countable partition of $\Omega$.
Let $C \in \mathcal{C}_1$ and let $H(\mathcal{A} \cap C) < \infty$ and $\mu(C) < \infty$.
Then
\[(4.17) \quad \sup \frac{I_{\mathcal{C}(\mathcal{A} \mid \mathcal{C}_n)}}{n} \in L_1.\]

Proof. We enumerate the sets $A_i \in \mathcal{A}$ in such a
way that the sequence $u_i = \mu(A_i \cap C)$, $i = 1, 2, \ldots$
is non-increasing. We further set
\[(4.18) \quad \varepsilon_k = I_{\mathcal{C}(\mathcal{A} \mid \mathcal{C}_k)}.\]
To prove the theorem, it will suffice to show that

\[(4.19) \quad \sum_{m=0}^{\infty} \mu[\sup_k \varepsilon_k > m] < \infty,\]
because

\[0 \leq \int_\Omega \sup_k \varepsilon_k \, d\mu \leq \sum_{m=0}^{\infty} (m+1)\mu[m < \sup_k \varepsilon_k \leq m+1] =
\sum_{m=0}^{\infty} (m+1)(\mu[m < \sup_k \varepsilon_k] - \mu[m+1 < \sup_k \varepsilon_k]) =
\sum_{m=0}^{\infty} \mu[m < \sup_k \varepsilon_k].\]

For each fixed $m$, $i = 1, 2, \ldots$ we define the sets
The following relations follow from the above definitions:

(4.23) \( F_{ij}(m) \cap F_{ik}(m) = \emptyset \) if \( j \neq k \)

(4.24) \( F_{ik}(m) \in \mathcal{C}_k \)

(4.25) \( F_i(m) = \bigcup_{k=0}^{\infty} F_{ik}(m) \cap A_i \cap C \).

Since \( C \in \mathcal{C}_k \) it follows from (4.22) and (4.24) that

(4.26) \( \mu(F_{ik}(m) \cap A_i \cap C) = \\
\int_{C \cap F_{ik}(m)} \mu(A_i | \mathcal{C}_k) \, d\mu \leq 2^{-m} \mu(F_{ik}(m) \cap C) \).

From (4.23) and (4.25) we obtain

(4.27) \( \mu(F_i(m)) \leq 2^{-m} \mu(C) \).

Let \( \gamma(m) = 2^m/(m+1)^2 \) and let \( \gamma^{-1}(i) \) be the number of integers \( m \) for which \( \gamma(m) < i \). Then,

(4.28) \( \gamma^{-1}(i) < 1 + \max \{ m : \gamma(m) < i \} \leq 18 + 2 \log i \) the last inequality following from the fact that if \( m_o = \max \{ m : \gamma(m) < i \} \) and \( m_o > 15 \), then \( \log i > m_o - 2 \log(m_o + 1) > m_o - 2(m_o + 1)/4 = m_o/2 - 1/2 \).
We sum $\mu(F_i(m))$ over $i$ and $m$ as follows

\[(4.29) \sum_{m=0}^{\infty} \sum_{i=1}^{\infty} \mu(F_i(m)) = \sum_{m=0}^{\infty} \sum_{i \leq \gamma(m)}^{\infty} \mu(F_i(m)) + \sum_{m=0}^{\infty} \sum_{i > \gamma(m)}^{\infty} \mu(F_i(m)).\]

From (4.27) we obtain an estimate for the first sum in the right side of (4.29)

\[(4.30) \sum_{m=0}^{\infty} \sum_{i \leq \gamma(m)}^{\infty} \mu(F_i(m)) \leq \sum_{m=0}^{\infty} 2^{-m} \gamma(m) \mu(C) \leq \pi^2 \mu(C)/6.\]

To estimate the second sum, we note that $F_i(m) \subset C \cap A_i$, and since $F_i(m) \subset C \cap A_i$, $\mu(F_i(m)) \leq \mu(C \cap A_i) = u_i$. Interchanging the order of summation yields

\[(4.31) \sum_{m=0}^{\infty} \sum_{i > \gamma(m)}^{\infty} \mu(F_i(m)) \leq \sum_{i=1}^{\infty} \sum_{m=0}^{\infty} \gamma^{-1}(i)-1 \mu(F_i(m)) \leq \sum_{i=1}^{\infty} \gamma^{-1}(i) u_i.\]

Since $\sum_{i=1}^{u_i} \leq \mu(C)$ and the sequence $(u_i)_{i=1}^{\infty}$ is non-increasing, we have $iu_i \leq \mu(C)$; hence

\[(4.32) \sum_{i=1}^{\infty} \gamma^{-1}(i) u_i \leq \sum_{i=1}^{\infty} (18 + 2 \log[\mu(C)/u_i]) u_i = \]
\[
\sum_{i=1}^{\infty} (18 + 2 \log \mu(C) - 2 \log u_i)u_i \leq \\
\mu(C)(18 + 2 \log \mu(C)) + 2H(\mathcal{R} \cap C).
\]

From (4.30) and (4.32) it follows that the sum on the left in (4.29) is finite which proves the theorem.

We may now state the main result of this chapter.

**Theorem 4.3.** Let \( \mathcal{A}, \mathcal{B} \) be partitions of \( \Omega \) into countably many sets of finite measure. Let \( C, D \in \mathcal{H}(\mathcal{R}^{-1}\mathcal{A}), D \in \mathcal{H}(\mathcal{R}^{-1}\mathcal{B}) \) be such that \( \mu(C), \mu(D) \), \( H(\mathcal{R} \cap C) \) and \( H(\mathcal{B} \cap D) \) are all finite. Further assume that

\[
(4.33) \quad \sum_{i=1}^{\infty} I_D(B) \bigvee_{i=1}^{\infty} \tau^{-i}B \geq 0.
\]

Then

\[
(4.34) \quad \lim_{n \to \infty} \sum_{i=0}^{n} I_C(A) \bigvee_{j=1}^{i+1} \tau^{-j}A \tau^{n-i} \mu(A) = \\
\sum_{i=0}^{n} I_D(B) \bigvee_{j=1}^{i+1} \tau^{-j}B \tau^{n-i} \mu(B)
\]

\[
\sum_{i=1}^{\infty} I_C(A) \bigvee_{j=1}^{\infty} \tau^{-j}A \mu(A) \\
= \sum_{i=1}^{\infty} I_D(B) \bigvee_{j=1}^{\infty} \tau^{-j}B \mu(B)
\]

**Proof.** We set
and we note that from Theorem 4.2 it follows that
\[ \sup_n f_n \in L_1, \quad \sup_n g_n \in L_1. \]
Furthermore, Proposition 4.1 and the remark following it imply that
\[ f_n \to f \quad \text{and} \quad g_n \to g \quad \text{a.e.} \]
Also, \[ \mu(B \upharpoonright \tau^{-j} \beta) \leq 1 \]
for all \( B \in \beta \) and all \( n \geq 0 \), implies that \( g \geq 0 \).
Since \[ \int g \, d\mu > C \], it follows from Theorem 4.1 that
\( (4.34) \) holds and the proof is complete.
V. WEAKLY WANDERING SETS OF INFINITE MEASURE

A basic question in the theory of point transformations is that of the existence of an equivalent finite invariant measure. Hajian and Kakutani [11] introduced the condition:

(HK) All weakly wandering sets are null sets.

and proved the following theorem under the assumption that $\tau$ is invertible; see also Sucheston [30].

**Theorem 5.1 (Hajian-Kakutani).** Condition (HK) is equivalent with the existence of a finite equivalent invariant measure.

Hajian [10] has shown that, in a sense, the opposite of condition (HK) is true when $\tau$ preserves an infinite sigma-finite measure: if $\tau$ is invertible, ergodic, conservative, and measure preserving, then there exists a weakly wandering set of infinite measure. Our purpose here is to present a simple direct proof of Hajian's result without the invertibility assumption. The proof presented here is a modified version of Sucheston's proof of Theorem 5.1, while Hajian's proof, by different methods, uses the Individual Ergodic Theorem.

A Banach limit is a linear functional $L$ on the Banach space of bounded sequences $(x_n)_{n=0}^{\infty}$ of real numbers, which satisfies the conditions...
i) \( L(x_n) = L(x_{n+1}) \)

(5.1) ii) \( L(x_n) \geq 0 \) if \( x_n \geq 0 \) for \( n = 0, 1, \ldots \)

iii) \( L(1) = 1 \).

The existence, basic properties, and extreme values of Banach limits are established in [30] and [31]. It is known that for any Banach limit \( L \), \( L(x_n) \geq \lim \inf x_n \).

Let \( \pi \) be a probability measure on \( \mathcal{F} \); a Banach limit \( L \) determines an invariant finitely additive set-function \( \lambda \) by means of the formula

(5.2) \[ \lambda(A) = L[\pi(r^{-n}A)] \quad A \in \mathcal{F}. \]

Because \( r \) is null preserving, \( \lambda \) is \( \pi \)-continuous; that is, \( \lambda \) vanishes on \( \pi \)-null sets. The next two lemmas are contained in Theorems 3, 4, and 6 of [30]. Lemma 5.1 permits the construction of invariant measures via Banach limits, while Lemma 5.2 is basic in constructing weakly wandering sets. In these two lemmas, we consider a partition \( \mathcal{P} \) of a set \( A \) to mean a class of sets whose union is \( A \) and which are disjoint in the set theoretic sense rather than modulo null sets.

**Lemma 5.1.** The set function \( \lambda \) admits a decomposition \( \lambda = \lambda_c + \lambda_m \) where \( \lambda_c \geq 0 \) and \( \lambda_m \) is a \( \pi \)-continuous invariant measure. \( \lambda_m \) is given by

(5.3) \[ \lambda_m(A) = \inf \sum_{i=1}^{\infty} \lambda(A_i), \]

where infimum is taken over all countable (including finite) partitions \( (A_i)_{i=1}^{\infty} \) of \( A \).
Proof. Clearly $\lambda_m$ is $n$-continuous and $\lambda_m(A) \leq \lambda(A)$. We next show that $\lambda_m$ is a measure. Indeed, let $A_i$, $i = 1, 2, \ldots$ be a countable partition of $A$ and for each $i = 1, 2, \ldots$ let $A_{in}$, $n = 1, 2, \ldots$ be a partition of $A_i$ such that

$$\sum_{n=1}^{\infty} \lambda(A_{in}) \leq \lambda_m(A_i) + \varepsilon 2^{-i}.$$ 

Now $[A_{in}, i = 1, 2, \ldots; n = 1, 2, \ldots]$ is a countable partition of $A$, hence

$$\lambda_m(A) \leq \sum_{i,n=1}^{\infty} \lambda(A_{in}) \leq \sum_{i=1}^{\infty} \lambda_m(A_i) + \varepsilon 2^{-i} \leq \sum_{i=1}^{\infty} \lambda_m(A_i) + \varepsilon,$$

which establishes the countable subadditivity of $\lambda_m$. On the other hand, the finite additivity of $\lambda$ together with the superadditive property of infimum, implies that $\lambda_m$ is finitely superadditive. Finite superadditivity in general implies countable superadditivity; thus $\lambda_m$ is a measure. The invariance of $\lambda_m$ is proved by noting that if $A_i$, $i = 1, 2, \ldots$ is a partition of $A$, for which

$$\lambda_m(A) + \varepsilon > \sum_{i=1}^{\infty} \lambda(A_i),$$

then $\tau^{-1}A_i$, $i = 1, 2, \ldots$ is a partition of $\tau^{-1}A$ and
\[ \lambda_m(A) + \varepsilon \geq \sum_{i=1}^{\infty} \lambda(A_i) = \sum_{i=1}^{\infty} \lambda(\tau^{-1}A_i) \geq \lambda_m(\tau^{-1}A) . \] Thus, 
\[ \lambda_m(A) \geq \lambda_m(\tau^{-1}A) . \] Strict inequality cannot hold, since 
\[ \tau^{-1}\Omega = \Omega \] implies that for any set \( A \), 
\[ \lambda_m(A) + \lambda_m(A^c) = \lambda_m(\tau^{-1}A) + \lambda_m(\tau^{-1}A^c) . \]

**Lemma 5.2.** If there exists no finite \( \pi \)-equivalent invariant measure, then there is a set \( A \) with \( \pi(A) > 0 \) and such that for each \( \varepsilon > 0 \), there is a set \( B \subset A \) with \( \pi(A - B) < \varepsilon \) and \( \lambda(B) = 0 \).

**Proof.** The measure \( \lambda_m \) of Lemma 5.1 is \( \pi \)-continuous and since there is no equivalent invariant measure, there is a set \( A \) for which \( \pi(A) > 0 \) and \( \lambda_m(A) = 0 \). Since \( \lambda_m(A) = 0 \), there is a sequence \( \mathcal{P}_n = (A_{in})_{i=1}^{\infty}, \ n = 1, 2, \ldots \) of partitions of \( A \) such that

\[ \lim_{n \to \infty} \sum_{i=1}^{\infty} \lambda(A_{in}) = 0 . \]

Let \( k(n) \) be the least integer for which

\[ \sum_{i=k(n)}^{\infty} \pi(A_{in}) < \varepsilon 2^{-n} . \]

Set

\[ C = \bigcup_{n=1}^{\infty} \bigcup_{i=k(n)}^{\infty} A_{in} . \]

Then \( \pi(C) < \sum_{n=1}^{\infty} \varepsilon 2^{-n} = \varepsilon \) and
(5.10) \[ \lambda(A - C) = \lambda\left(\bigcap_{n=1}^{\infty} \bigcup_{i=1}^{k(n)-1} A_{in}\right) < \lambda\left(\bigcup_{i=1}^{k(n)-1} A_{in}\right) = \sum_{i=1}^{k(n)-1} \lambda(A_{in}) \rightarrow 0. \]

The conclusion of the lemma now follows with \( B = A - C \).

In the following lemma, the existence of a special type of weakly wandering set is shown. Under the assumption that \( \tau \) is invertible, the lemma is due to Hajian [10].

**Lemma 5.3.** If \( \pi \) is a probability measure on \( \mathcal{F} \) which admits of no finite \( \pi \)-equivalent invariant measure, then there exists a sequence of integers \( 0 = r_0 < r_1 < \ldots \) and a set \( W \) of positive \( \pi \)-measure such that the following holds:

(5.11) \[ \tau^{-r_j-r_k} W \cap \tau^{-r_i-r_h} W = \emptyset \]

for any \( j > k, i > h, \) with \( j \neq i \) or \( h \neq k \).

**Proof.** Let \( A \) be the set obtained from Lemma 5.2. Let \( \varepsilon < \pi(A) \) and let \( B \) be the corresponding set for which \( \lambda(B) = 0 \) and \( \pi(A - B) < \varepsilon/2 \). We construct a sequence \( (r_i)_{i=0}^{\infty} \) and a set \( W \subset B \) which has the property (5.11). Set \( r_0 = 0 \). Let \( \varepsilon_j = \varepsilon/2^{j+1} \) for \( j = 1, 2, \ldots \). Since \( \lambda(B) = 0 \), \( \liminf \pi(\tau^{-n}B) = 0 \) and we may choose \( r_1 \) such that \( \pi(\tau^{-r_1}B) < \varepsilon_1 \). Set \( s_1 = r_1 \), \( N_1 = \{ n: 0 \leq n \leq 2s_1 \} \) and \( B_1 = \bigcup_{i \in N_1} \tau^{-i}B \).
Since $\lambda$ is finitely additive and invariant, $\lambda(B_1) \leq \sum_{i \in N_1} \lambda(\tau^{-i} B) = 0$, hence $\lim \inf \pi(\tau^{-n} B_1) = 0$. We can therefore choose $r_2 > s_1$ in such a way that $\pi(\tau^{-r_2+s_1} B_1) < \varepsilon_2$. Assume that we have already defined $s_n, N_n, B_n$ and chosen $r_{n+1}$. We define

$$s_{n+1} = \sum_{j=1}^{n+1} r_j, \quad N_{n+1} = \{m : 0 \leq m \leq 2s_{n+1}\},$$

(5.12)

$$B_{n+1} = \bigcup_{i \in N_{n+1}} \tau^{-i} B.$$ 

Since $\lambda$ is invariant, $\lambda(B_{n+1}) = 0$ and

$$\lim \inf \pi(\tau^{-i} B_{n+1}) = 0.$$ 

We can choose $r_{n+2} > s_{n+1}$ so large that

$$\pi(\tau^{-r_{n+2}+s_{n+1}} B_{n+1}) < \varepsilon_{n+1},$$

etc.

Now let

(5.13)

$$C = \bigcup_{j=1}^{\infty} \bigcup_{i \in N_{j}} \tau^{-r_{j+1}+s_{j}^{-i}} B = \bigcup_{j=1}^{\infty} \tau^{-r_{j+1}+s_{j}^{-i}} \bigcup_{i \in N_{j}} \tau^{-i} B,$$

then $\pi(C) < \sum_{j=1}^{\infty} \pi(\tau^{-r_{j+1}+s_{j}^{-i}} B_j) < \varepsilon/2$. Letting $W = B - C$ we obtain that $\pi(W) \geq \pi(B) - \varepsilon/2 \geq \pi(A) - \varepsilon$.

We now prove that (5.11) holds. It is sufficient to consider only two cases.
Case I. \( j > k \) and \( j > i > h \). Then, 
\[
- s_{j-1} < r_k - r_i - r_h < s_{j-1} \quad \text{and} \quad s_{j-1} + r_k - r_i - r_h \in N_{j-1}.
\]
Since \( W \subset B \) and 
\[
- r_j + s_{j-1} - (s_{j-1} + r_k - r_i - r_h) = - r_j - r_k + r_i + r_h,
\]
it follows that
\[
(5.14) \quad W \cap \tau^{-r_j - r_k + r_i + r_h} W \subset W \cap \tau^{-r_j - r_k + r_i + r_h} B \subset W \cap C = \emptyset.
\]

Case II. \( j = i > k > h \). Then \( s_{k-1} - r_h \in N_{k-1} \)
and (5.14) still holds. On applying \( \tau^{r_i - r_h} \) to (5.14), we obtain (5.11).

Remark. The special property (5.11) is not unusual in the sense that if \( \tau \) is ergodic, conservative and preserves no finite \( \pi \)-equivalent measure, and if \( D \) is a set of positive measure, then for each \( \varepsilon > 0 \) there exist \( W' \subset D \) and a sequence \( (r_i)_{i=0}^{\infty} \) such that
\[
\pi(D - W') < \varepsilon \quad \text{and} \quad (5.11) \text{ holds for } W' \text{ and } (r_i)_{i=0}^{\infty}.
\]
Indeed, in the proof of Lemma 5.2, we considered a set \( A \) for which \( \lambda_m(A) = 0 \) and \( \pi(A) > 0 \). Then
\[
\lambda_m(A^\tau) = 0, \quad \pi(A^\tau) = 1 \quad \text{and the arguments of Lemma 5.2 may be applied with } A \text{ replaced by } A^\tau \text{ to obtain a set } B \text{ satisfying the conditions of the conclusion of Lemma 5.2.}
\]
We then obtain, by Lemma 5.3, a weakly wandering set \( W \) such that \( \pi(W) > 1 - \varepsilon \) and \( W \) has
property (5.11). The set \( W' = D \cap \tilde{W} \) has the desired property.

**Lemma 5.4.** Let \( \tau \) be ergodic and preserve \( \mu \), a sigma-finite measure for which \( \mu(\Omega) = \infty \). Then there is no finite \( \mu \)-equivalent invariant measure.

**Proof.** *Ab contrario,* suppose that \( \pi \) is a finite equivalent invariant measure. By the Radon-Nikodým theorem, there is a function \( f \) such that

\[
(5.15) \quad \pi(A) = \int_A f \, d\mu \quad A \in \mathcal{F}.
\]

we have

\[
(5.16) \quad \pi(A) = \pi(\tau^{-1}A) = \int_{\tau^{-1}A} f(\omega) \, d\mu = \int_A f(\tau\omega) \, d\mu,
\]

the last inequality following by Theorem C p. 163 of [14] and the fact that \( \tau \) preserves \( \mu \). From (5.16) we conclude that \( f(\omega) = f(\tau\omega) \) a.e. which implies that \( f \) is \( \mathcal{J} \)-measurable, a contradiction, since the only \( \mathcal{J} \)-measurable functions are constants and \( f \) cannot be a constant.

We now show the existence of a weakly wandering set of infinite measure.

**Theorem 5.2.** Let \( \tau \) be an ergodic measure preserving point transformation defined on an infinite sigma-finite measure space \((\Omega, \mathcal{F}, \mu)\). Then there is a weakly wandering set of infinite measure.

**Proof.** By Lemma 5.4, there is no finite equivalent invariant measure. Since \( \mu \) is sigma-finite, we
may obtain an equivalent probability measure \( \pi \) by means of the formula

\[
(5.17) \quad \pi(A) = \sum_{n=1}^{\infty} \mu(A \cap \Omega_n) / 2^n \mu(\Omega_n)
\]

where \( (\Omega_n)_{n=1}^{\infty} \) is a decomposition of \( \Omega \) into sets of finite positive measure. By Lemma 5.3, we can find a set \( W \) of positive \( \pi \)-measure and a sequence \( (r_i)_{i=0}^{\infty} \) for which (5.11) holds. Since \( \pi \) and \( \mu \) are equivalent, \( \mu(W) > 0 \). Let \( N \) be the set of odd integers and set \( C = \bigcup_{i \in N} \tau^{-r_i} W \). If \( h, k \notin N \) and \( h \neq k \), then

\[
(5.18) \quad \tau^{-r_h} C \cap \tau^{-r_k} C = \bigcup_{i \in N} \bigcup_{j \in N} \tau^{-r_i-r_h} W \cap \tau^{-r_j-r_k} W.
\]

Each set of the double union falls into one of the following four categories:

(i) \( h > i, k > j \) \hspace{1cm} (iii) \( i > h, k > j \)

(ii) \( h > i, j > k \) \hspace{1cm} (iv) \( i > h, j > k \).

In cases (i) and (iv), we have \( h \neq k \) while in (ii) and (iii), we have \( h \neq j \); hence each element is empty by (5.11) and \( C \) is weakly wandering under the sequence \( 0 = r_0, r_2, r_4, \ldots \). Since \( \tau \) preserves \( \mu \),

\[
(5.19) \quad \mu(C) = \sum_{i \in N} \mu(\tau^{-r_i} W) = \sum_{i \in N} \mu(W) = \infty.
\]
VI. WEAKLY WANDERING FUNCTIONS

Many results which are valid for point transformations extend to general $L_1$ operators although the extension is sometimes very difficult. A point transformation $\tau$ generates an $L_1$ operator $T$ via the adjoint operator $T^*$ acting on $L_\infty$; $T^*$ is defined by the duality relationship

\[(6.1) \quad \int_{\Omega} fT^*g \, du = \int_{\Omega} Tf \cdot g \, du \quad f \in L_1, \; g \in L_\infty.\]

$\tau$ generates $T^*$ by means of the formula

\[(6.2) \quad T^*g = g \circ \tau \quad \text{for} \; g \in L_\infty.\]

A function $h \in L_\infty$, with $|h|_\infty \overset{\text{def}}{=} \text{ess sup} \; |h| \leq 1$, is weakly wandering under a sequence $0 = n_0 < n_1 < \ldots$ if $\left| \sum_{i=0}^{\infty} T^{n_i} h \right|_\infty < 2$. The result of Hajian and Kakutani (Theorem 5.1) has been extended to general $L_1$ operators by Dean and Sucheston [7] and independently by Neveu [23]. Here we extend the results of the preceding chapter to a class of infinite matrices. Let $\Omega = 1, 2, \ldots$, let $\mathcal{F}$ be the family of all subsets of $\Omega$, and let $\mu$ be the counting measure ascribing to each set in $\mathcal{F}$ the number of its points. An
infinite matrix \([t_{ij}]\) transforms a function \(g = (g_i)_{i=1}^{\infty}\) into a function \(T^*g\) whose \(i\)-th coordinate is given by

\[(6.3) \quad (T^*g)_i = \sum_{j=1}^{\infty} t_{ij} g_j.
\]

When \(t_{ij} \geq 0\) for \(i, j = 1, 2, \ldots\) and \(\sum_{j=1}^{\infty} t_{ij} \leq 1\) for each \(i = 1, 2, \ldots\), the matrix \([t_{ij}]\) is sub-stochastic; if moreover, \(\sum_{i=1}^{\infty} t_{ij} \leq 1\) for each \(j = 1, 2, \ldots\), then \([t_{ij}]\) is doubly sub-stochastic.

The operator \(T^*\), determined by a sub-stochastic matrix, maps \(l^\infty\) into itself and is the adjoint of the operator \(T\) which maps \(l_1\) into itself by taking a function \(f = (f_i)_{i=1}^{\infty}\) into \(Tf\) whose \(j\)-th coordinate is given by

\[(6.4) \quad (Tf)_j = \sum_{i=1}^{\infty} t_{ij} f_i.
\]

For \(n = 0, 1, \ldots\) let \(t_{ij}^{(n)}\) \(i, j = 1, 2, \ldots\) be the entries of the matrix \([t_{ij}]^n\).

The function \(h\) constructed in the following theorem is, in two ways, analogous to a weakly wandering
set of infinite measure (see previous chapter):
(1) the set on which \( h > 0 \) has infinite measure and (2) \( \int h = \Sigma h_i = \infty \). It would seem that the result should extend to general operators, but the author has not been able to obtain such an extension.

Theorem 6.1. Let \( [t_{ij}] \) be a doubly sub-stochastic matrix. Suppose further that \( \lim \inf \sup_{n \to \infty} t_{ij}^{(n)} = 0 \). Then for any \( \delta > 0 \), there exists a function \( h \) whose coordinates are 0's and 1's, and a sequence of integers \( 0 = n_0 < n_1 < \ldots \) such that
\[
\| \sum_{k=0}^{\infty} T^k h \|_{\infty} \leq 1 + \delta.
\]

Proof. Let \( a_n = \sup_{i,j} t_{ij}^{(n)} \), then since \( \lim \inf_{i,j} a_n = 0 \), there is a subsequence \( 0 = n_0 < n_1 < \ldots \) such that \( \sum_{k=0}^{\infty} a_{n_k} \leq 1 + \delta/2 \). Let \( \epsilon_i, i = 1, 2, \ldots \) be a sequence such that \( \epsilon_i > 0 \) for all \( i \geq 1 \) and \( \sum_{i=1}^{\infty} \epsilon_i \leq \delta/2 \). Since both the row and column sums of \( [t_{ij}^n] \) are convergent series, the following conditions hold: (A) \( \lim_{i \to \infty} t_{ij}^{(n)} = 0 \) for all fixed \( n, j \) and
(B) \( \lim_{j \to \infty} t^{(n)}_{ij} = 0 \) for all fixed \( n, i \). We shall construct an infinite sequence of vectors \( h(i), \ i = 1, 2, \ldots \) for which only one coordinate is one and all others are zero. The vectors \( h(i) \) will have the further property that no two of them have a one in the same position. The required function will then be \( h = \sum_{i=1}^{\infty} h(i) \). Set \( N_0 = 1 \) and define \( h(1) = 1_{\{N_0\}} \) (the indicator function of the singleton set \( \{N_0\} \)). Then

\[
(6.5) \quad \left| \sum_{k=0}^{\infty} t^{\ast n_k} h(1) \right|_\infty = \sup_{i} \sum_{k=0}^{\infty} t^{(n_k)}_{il} \leq \sum_{k=0}^{\infty} a_{n_k} \leq 1 + \delta/2.
\]

Let \( K_1 \) be such that \( \sum_{k=K_1}^{\infty} a_{n_k} \leq \varepsilon_1/2 \). Since for each \( n \), \( \lim_{i \to \infty} t^{(n)}_{il} = 0 \), we can find an \( M_1 \) such that \( i > M_1 \) implies that \( t^{(n_k)}_{il} \leq \varepsilon_1/2K_1 \) for \( k = 0, 1, \ldots, K_1 - 1 \). Such an \( M_1 \) is found by determining \( M_{1k} \) such that \( t^{(n_k)}_{il} \leq \varepsilon_1/2K_1 \) for \( i > M_{1k} \) and setting \( M_1 = \max_{0 \leq k \leq K_1 - 1} M_{1k} \). Combining
estimates we find that for $i > M_1$

$$(6.6) \sum_{k=0}^{\infty} t(n_k)_{il} = \sum_{k=0}^{K_1-1} t(n_k)_{il} + \sum_{k=K_1}^{\infty} t(n_k)_{il} \leq \epsilon_1.$$ 

Choose $N_1$ so large that $j \geq N_1$ implies that for each

$i = 1, 2, \ldots, M_1$, $\sum_{k=0}^{K_1-1} t(n_k)_{ij} \leq \epsilon_1/2$. Such an $N_1$
can be chosen in the following way: for each fixed

$i = 1, 2, \ldots, M_1$ select $N_{li}$ such that $j \geq N_{li}$
implies that $t(n_k)_{ij} \leq \epsilon_1/2K_1$ for $k = 0, 1, \ldots, K_1 - 1$, then set $N_1 = \max_{0 \leq i \leq M_1} N_{li}$. Next, $h^{(2)}_{\langle K_1 \rangle}$
is defined. We note that for $0 < i \leq M_1$ we have that

$$(6.7) \left( \sum_{k=0}^{\infty} T^* n_k h^{(2)} \right)_{i} = \sum_{k=0}^{K_1-1} t(n_k)_{i, N_1} +$$

$$+ \sum_{k=K_1}^{\infty} t(n_k)_{i, N_1} \leq \epsilon_1$$

while for $i > N_1$ we have that

$$(6.8) \left( \sum_{k=0}^{\infty} T^* n_k h^{(2)} \right)_{i} \leq \sum_{k=0}^{\infty} a_n \leq 1 + \delta/2.$$
Combining the above with (6.5) and (6.6) we see that

for all $i$, \( \left( \sum_{k=0}^{\infty} T^k_n (h(1) + h(2)) \right)_i \leq 1 + \delta/2 + \epsilon_1 \)

(there are two cases: $i \leq M_1$ and $i > M_1$). Assuming that $K_{r-1}, M_{r-1}, N_{r-1}$, and $h^{(r-1)}$ have been defined, we define $K_r, M_r, N_r$ and $h^{(r)}$ in the following way: $K_r$ is such that $K_r > K_{r-1}$, and

\[ \sum_{k=K_r}^{\infty} a_{nk} \leq \epsilon_r/2. \]

In analogy to the case $r = 1$, $M_r$ is defined in such a way that $M_r > M_{r-1}$ and $i > M_r$ implies that $t^{(n_k)}_{i,N_{r-1}} < \epsilon_r/2K_r$ for

$k = 0, 1, \ldots, K_r - 1$. It immediately follows that for $i > M_r$,

\[ \sum_{k=0}^{\infty} t^{(n_k)}_{i,N_{r-1}} = \sum_{k=0}^{K_r-1} t^{(n_k)}_{i,N_{r-1}} + \sum_{k=K_r}^{\infty} t^{(n_k)}_{i,N_{r-1}} \leq \epsilon_r. \]

We next define $N_r$ in such a way that $j \geq N_r$ implies

\[ \sum_{k=1}^{K_r-1} t^{(n_k)}_{ij} \leq \epsilon_r/2. \]
Indeed, for each fixed \( i = 1, 2, \ldots, M_r \), select \( N_{ri} \) such that \( j \geq N_{ri} \) implies that \( t_{ij}^{(nk)} \leq \varepsilon_r / 2K_r \) for \( k = 0, 1, \ldots, K_r - 1 \), then set \( N_r = \max_{0 \leq i \leq M_r} N_{ri} \).

The function \( h(r) \) is defined by means of the formula

\[ h(r) = 1 \{ N_r \} . \]

For \( i \leq M_r \) it follows that

\[
\left( \sum_{k=0}^{\infty} t_{i}^{(nk)} h_{i}(r) \right)_{i} = \sum_{k=0}^{\infty} t_{i}^{(nk)} i^{N_{r}} \leq \varepsilon_r ,
\]

while for \( i > M_r \);

\[
\left( \sum_{k=0}^{\infty} t_{i}^{(nk)} h_{i}(r) \right)_{i} = \sum_{k=0}^{\infty} t_{i}^{(nk)} i^{N_{r}} \leq \sum_{k=0}^{\infty} a_n^{(nk)} \leq 1 + \delta/2 .
\]

We assert that the function \( h = \sum_{r=1}^{\infty} h(r) \) is the required function. It is only necessary to check that

\[
\left| \sum_{k=0}^{\infty} t_{i}^{(nk)} h \right|_{\infty} \leq 1 + \delta . \]

Let \( r \) be a given positive integer. For convenience in notation, set \( g^{(r)} = \sum_{n=0}^{r} h(n) \). For each \( i \), it is clear that
The following tabulation indicates a bound on the terms appearing in the inner sum of the last double sum. These bounds are a consequence of the construction employed.

\[
\sum_{k=0}^{\infty} t^{n_k} \varepsilon(r)_i = \sum_{k=0}^{\infty} t^{(n_k)} \varepsilon_i = \sum_{j=0}^{r} \sum_{k=0}^{\infty} t^{n_k} \varepsilon_j
\]

From the above array, it is clear that for any \( i > 0 \),

\[
\left( \sum_{k=0}^{\infty} T^{n_k} g(r) \right)_i \leq 1 + \delta/2 + \sum_{j=1}^{r+1} \varepsilon_j \leq 1 + \delta.
\]

Letting \( r \to \infty \) in (6.9) and interchanging the order of summation which is justified because all terms are positive, we obtain that for all \( i > 0 \),

\[
\left( \sum_{k=0}^{\infty} T^{n_k} h \right)_i \leq 1 + \delta,
\]

which completes the proof of the theorem.
REFERENCES


54


