ON THE ERGODIC THEOREM

SHIGERU TSURUMI

(Received April 15, 1954)

1. Introduction. Let $(X, \mathfrak{B}, m)$ be a measure space such that X is a set, $\mathfrak{B}$ is a Borel field of subsets of X, and m is a σ-finite measure defined on $\mathfrak{B}$. A single valued transformation T of X into itself is called measurable if the inverse transformation T^{-1} sends every set of $\mathfrak{B}$ to a set of $\mathfrak{B}$. The measurable transformation T is called non-singular (with respect to m) if $A \in \mathfrak{B}$ and $m(A) = 0$ imply $m(T^{-1}A) = 0$. Throughout this paper it is assumed that all sets under consideration are in $\mathfrak{B}$ and the transformation T is measurablea nd non-singular. A measure μ defined on $\mathfrak{B}$ is said to be invariant under T (or T is said to be measure-preserving with respect to μ) if $\mu(T^{-1}A) = \mu(A)$ for every set A. Two measures λ and μ defined on $\mathfrak{B}$ are called equivalent if $\lambda(A) = 0$ implies $\mu(A) = 0$ and the converse. A set A is called an invariant set if $m(T^{-1}A - A) + m(A - T^{-1}A) = 0$.

We define the following statements.

(I) There exists a constant K such that

$$0 < \limsup_n \frac{1}{n} \sum_{i=0}^{n-1} m(T^{-i}A) \leqq K \cdot m(A)$$

for every set A of positive measure.

(I′) There exists a constant K such that

$$\limsup_n \frac{1}{n} \sum_{i=0}^{n-1} m(T^{-i}A) \leqq K \cdot m(A)$$

for every set A.

(II) There exists a sequence of sets $\{X_j\}$ and a constant K such that

$$X_1 \subset X_2 \subset \ldots., \quad X = \bigcup_{j=1}^{\infty} X_j, \quad m(X_j) < \infty \ (j = 1, 2, \ldots.),$$

and

$$0 < \sup_j \limsup_n \frac{1}{n} \sum_{i=0}^{n-1} m(X_j \cap T^{-i}A) \leqq K \cdot m(A)$$

for every set A of positive measure.

(II′) There exists a sequence of sets $\{X_j\}$ and a constant K such that

$$X_1 \subset X_2 \subset \ldots., \quad X = \bigcup_{j=1}^{\infty} X_j, \quad m(X_j) < \infty \ (j = 1, 2, \ldots.),$$

and

$$\sup_j \limsup_n \frac{1}{n} \sum_{i=0}^{n-1} m(X_j \cap T^{-i}A) \leqq K \cdot m(A)$$

for every set A.

(B) For any function $f \in L(X, \mathfrak{B}, m)$[1] the limit

$$\tilde{f}(x) = \lim_n \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x)$$

exists almost everywhere on X and $\tilde{f} \in L(X, \mathfrak{B}, m)$.

In the following, we use the notation $\tilde{f}$ ($\tilde{g}$ etc.) which denotes the limit function of the means $\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)$ $\left(\frac{1}{n}\sum_{i=0}^{n-1} g(T^i x)\text{ etc.}\right)$ in case the limit is well defined.

In case m is finite, N. Dunford and D. S. Miller [1][2] have given in their joint paper a necessary and sufficient condition[3] that Neumann's ergodic theorem holds, and as its consequence led the statement (B) from this condition. Hereafter F. Riesz [3] has given another proof of the latter and proved that, even if m is not finite, the above condition with a certain additional restriction implies (B) (see Corollary of Theorem 1 in § 2). Recently C. Ryll-Nardzewski [4] has shown that the statement (II′), which is weaker than (II) formulated by S. Hartman, is equivalent to (B) and that, in case m is finite, the statements (II), (II′) and (B) are equivalent to each other. However a part of the former is not quite right in case m is not finite. In fact we can construct a σ-finite (but not finite) measure space and a transformation for which (II′) holds and (B) does not hold (see Example 1 in § 3).

The main purpose of this paper is to work out that each of (I) and (II) implies (B), and (II′) does not necessarily imply (B).

2. Generalization of Birkhoff's ergodic theorem. Let $\Delta(A) = \Delta(A, \{A_k\}_{k=1,2,\ldots})$ denote a decomposition of the set A such that

$$A = \bigcup_{k=1}^{\infty} A_k, \quad A_k \cap A_l = 0 \ (k \neq l).$$

Let us put for every set A

$$\alpha(A) = \sup_{(A)} \sum_{(A:k)} \limsup_n \frac{1}{n} \sum_{i=0}^{n-1} m(T^{-i} A_k), \tag{1}$$

where $\sup_{(A)}$ denotes the supremum for all decompositions $\Delta(A)$ of A and $\sum_{(A:k)}$ means to sum up with respect to all sets A_k's of the decomposition

1) The notation $L(X, \mathfrak{B}, m)$ denotes the class of all integrable functions with respect to the measure space $(X, \mathfrak{B}, m)$.

2) Numbers in square brackets refer to the references at the end of this paper.

3) The condition reads as follows : there exists a constant K such that for any set A

$$\frac{1}{n} \sum_{i=0}^{n-1} m(T^{-i}A) \leqq K \cdot m(A) \qquad (n = 1, 2, \ldots\ldots).$$

$\Delta(A) = \Delta(A, \{A_k\}_{k=1,2,\ldots})$.

LEMMA 1. *If the statement* (I′) *holds, the non-negative set function* α *defined by* (1) *has the following properties*:

(i) α *is finitely additive*;

(ii) $\alpha(A) \leqq \alpha(T^{-1}A)$ *for every set* A *of finite measure*;

(iii) $\alpha(A) \leqq K \cdot m(A)$ *for every set* A;

(iv) $\alpha(A) \geqq \limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m(T^{-i}A)$ *for every set* A.

PROOF. Proof of (i): Let us suppose that

$$A = \bigcup_{j=1}^{N} A^j, \qquad A^i \cap A^j = 0\,(i \neq j).$$

Let ε be any positive number. Then, for each j, there exists a decomposition $\Delta(A^j, \{A_k^j\}_{k=1,2,\ldots})$ such that

$$\alpha(A^j) - \frac{\varepsilon}{N} < \sum_{(A^j:k)} \limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m(T^{-i}A_k^j).$$

Combining all $\Delta(A^j)$'s we get a decomposition $\Delta(A, \{A_k^j\}_{k=1,2,\ldots}; j = 1, 2, \ldots, N)$, so that it follows

$$(2)\qquad \begin{aligned} \sum_{j=1}^{N} \alpha(A^j) - \varepsilon &< \sum_{j=1}^{N} \sum_{(A^j:k)} \limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m(T^{-i}A_k^j) \\ &= \sum_{(A:k,j)} \limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m(T^{-i}A_k^j) \leqq \alpha(A). \end{aligned}$$

On the other hand, there exists a decomposition $\Delta(A, \{A_k\}_{k=1,2,\ldots})$ such that

$$(3)\qquad \alpha(A) - \varepsilon < \sum_{(A:k)} \limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m(T^{-i}A_k).$$

Let us put

$$A_k^j = A_k \cap A^j (k = 1, 2, \ldots; j = 1, 2, \ldots, N).$$

Then, for each j, the collection of sets $\{A_k^j\}_{k=1,2,\ldots}$ gives a decomposition $\Delta(A^j, \{A_k^j\}_{k=1,2,\ldots})$, so that we have

$$(4)\qquad \begin{aligned} \sum_{(A:k)} \limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m(T^{-i}A_k) &\leqq \sum_{(A:k,j)} \limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m(T^{-i}A_k^j) \\ &= \sum_{j=1} \sum_{(A^j:k)} \limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m(T^{-i}A_k^j) \leqq \sum_{j=1}^{N} \alpha(A^j). \end{aligned}$$

Thus by (3) and (4) we have

$$(5)\qquad \alpha(A) - \varepsilon < \sum_{j=1}^{N} \alpha(A^j).$$

Since ε is arbitrary, (2) and (5) imply

$$\alpha(A) = \sum_{j=1}^{N} \alpha(A^j).$$

Hence it follows that α is finitely additive.

Proof of (ii): If A is a set of finite measure, then for each i the set $T^{-i}A$ is of finite measure. Further, to a decomposition $\Delta(A, \{A_k\}_{k=1,2,\ldots})$ there corresponds a decomposition $\Delta(T^{-1}A, \{T^{-1}A_k\}_{k=1,2,\ldots})$. Hence we can easily show the property (ii).

The properties (iii) and (iv) are the immediate consequences of the definition of α.

C Ryll-Nardzewski [4] has proved the following

LEMMA 2. *Let $(X', \mathfrak{B}', \mu)$ be a measure space, and let $\widetilde{T}$ be a transformation of $L(X', \mathfrak{B}', \mu)$ into itself which has the following properties:*

(i) *if $f(x) = g(x)$ almost everywhere (μ), $\widetilde{T}f(x) = \widetilde{T}g(x)$ almost everywhere (μ);*

(ii) *$\widetilde{T}$ is additive and homogeneous;*

(iii) *if $f(x)$ is positive almost everywhere (μ), $\widetilde{T}f(x)$ is also. Then $\widetilde{T}$ is a linear operator of $L(X', \mathfrak{B}', \mu)$ into itself.*

We shall now prove the following theorem which is a generalization of Birkhoff's ergodic theorem.

THEOREM 1. *The statement* (I) *implies the statement* (B), *but the converse is not true.*

If m is finite, three statements (I), (I′) and (B) are equivalent to each other.

PROOF. (I) $\rightarrow (B)$: We put for any set A of finite measure

$$\beta(A) = \lim_{n} \alpha(T^{-n}A).$$

This definition is justified by (ii) of Lemma 1.

Let B be a fixed set of finite measure, then the non-negative set function $\beta(A \cap B)$ of variable A has the following properties:

(i) $\beta(A \cap B) \leqq K^2 \cdot m(A \cap B)$ for every set A;

(ii) $\beta(A \cap B)$ is completely additive as the set function of A;

(iii) $\beta[T^{-1}(A \cap B)] = \beta(A \cap B)$ for every set A;

(iv) $\beta(A \cap B) \geqq \alpha(A \cap B) \geqq \limsup_{n} \frac{1}{n} \sum_{i=0}^{n-1} m[T^{-i}(A \cap B)]$ for every set A.

Proof of (i): From (ii), (iii) of Lemma 1 and (I), it follows that for any set A

$$\beta(A \cap B) = \lim_{n} \alpha[T^{-n}(A \cap B)]$$

$$\lim \frac{1}{n} \sum_{i=0}^{n-1} \alpha[T^{-i}(A \cap B)]$$

$$\leqq \limsup_n \frac{1}{n}\sum_{i=0}^{n-1} K\cdot m[T^{-i}(A\cap B)] \leqq K^2\cdot m(A\cap B),$$

which is the required.

Proof of (ii): From (i) of Lemma 1 it follows that $\beta(A\cap B)$ is finitely additive as the set function of A.

Let $\{A_n\}$ be any sequence of sets such that

$$A_1 \supset A_2 \supset \ldots,\quad \bigcap_{n=1}^{\infty} A_n = 0.$$

Since B is of finite measure, $m(A_n\cap B)$ tends to zero with $1/n$, so that, from (i), $\beta(A_n\cap B)$ tends to zero with $1/n$. Thus $\beta(A\cap B)$ is completely additive.

The properties (iii) and (iv) follows evidently from the definition of β and (iv) of Lemma 1.

Now we choose a sequence of sets $\{Y_k\}$ such that

$$X = \bigcup_{k=1}^{\infty} Y_k,\quad Y_k\cap Y_l = 0\ (k \neq l),\quad m(Y_k) < \infty\ (k = 1, 2, \ldots).$$

Let us put for each set A

$$\gamma(A) = \sum_{k=1}^{\infty} \beta(A\cap Y_k).$$

Then the non-negative set function γ has the following properties:

(v) $\gamma(A) \leqq K^2\cdot m(A)$ for every set A;

(vi) γ is an invariant measure on $\mathfrak{B}$;

(vii) γ is equivalent to m;

(viii) $\gamma(A) \geqq m(A)$ for any invariant set A.

The property (v) is the immediate consequence of (i) and definition of γ.

Proof of (vi): From (ii) it is evident that γ is a measure on $\mathfrak{B}$, so that we shall prove only that γ is invariant under T. Since Y_k's are of finite measure and hence, from (I), $T^{-1}Y_k$'s are also, it follows, from (iii), that for any set A

$$\begin{aligned}\gamma(A) &= \sum_{k=1}^{\infty}\beta(A\cap Y_k) = \sum_{k=1}^{\infty}\beta[T^{-1}(A\cap Y_k)]\\ &= \sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\beta(T^{-1}A\cap T^{-1}Y_k\cap Y_l)\\ &= \sum_{l=1}^{\infty}\sum_{k=1}^{\infty}\beta(T^{-1}A\cap T^{-1}Y_k\cap Y_l)\\ &= \sum_{l=1}^{\infty}\beta(T^{-1}A\cap Y_l) = \gamma(T^{-1}A).\end{aligned}$$

Proof of (vii): It is clear that $m(A) = 0$ implies $\gamma(A) = 0$, so that we shall prove the converse.

Let A be any set of positive measure. Then there is a set Y_{k_0} such that the set $A\cap Y_{k_0}$ is of positive measure. From (iv) and (I) it follows

$$\gamma(A)=\sum_{k=1}^{\infty}\beta(A\cap Y_k)\geqq\beta(A\cap Y_{k_0})$$
$$\geqq\limsup_n\frac{1}{n}\sum_{i=0}^{n-1}m[T^{-i}(A\cap Y_{k_0})]>0.$$

Proof of (viii): Let A be an arbitrary fixed invariant set. If A is not of finite γ-measure, it holds obviously

$$\gamma(A)\geqq m(A),$$

so that we shall consider the case where A is of finite γ-measure. Then from (vi) and (vii) it follows that γ is a finite invariant measure equivalent to m as the measure on $\mathfrak{B}_A$[4]. By Birkhoff's ergodic theorem, for each $f\in L(A, \mathfrak{B}_A, \gamma)$ the limit function $\tilde{f}(x)$ of the means $\frac{1}{n}\sum_{i=0}^{n-1}f(T^ix)$ is defined almost everywhere (γ) on A and $\tilde{f}\in L(A, \mathfrak{B}_A, \gamma)$. Since m is equivalent to γ, the function $\tilde{f}$ is also defined almost everywhere (m) on A. Let $\tilde{T}$ denotes the transformation of $L(A, \mathfrak{B}_A, \gamma)$ into itself defined by

$$\tilde{T}f=\tilde{f}$$

for each $\tilde{f}\in L(A, \mathfrak{B}_A, \gamma)$. Then $\tilde{T}$ has the properties (i), (ii) and (iii) of Lemma 2, so that $\tilde{T}$ is a linear operator of $L(A, \mathfrak{B}_A, \gamma)$ into itself. Let us now suppose that $\Delta(B, \{B_k\}_{k=1,2,\dots})$ is a decomposition of any subset B of A, and let φ_B and φ_{B_k}'s be the characteristic functions of B and B_k's, respectively[5]. Then it holds

$$\int_A\left|\varphi_B(x)-\sum_{k=1}^{N}\varphi_{B_k}(x)\right|d\gamma\to 0\ (N\to\infty),$$

so that from the linearity of $\tilde{T}$ it follows

$$\int_A\left|\tilde{\varphi}_B(x)-\sum_{k=1}^{N}\tilde{\varphi}_{B_k}(x)\right|d\gamma=\int_A\left|\tilde{T}\Big(\varphi_B(x)-\sum_{k=1}^{N}\varphi_{B_k}(x)\Big)\right|d\gamma\to 0\ (N\to\infty).$$

Since $\sum_{k=1}^{N}\tilde{\varphi}_{B_k}(x)$ is monotone-increasing as N increases, we have that

$$\sum_{k=1}^{\infty}\tilde{\varphi}_{B_k}(x)=\tilde{\varphi}_B(x)$$

almost everywhere (γ) on A and hence almost everywhere (m) on A.

Therefore, by Fatou's lemma we have

4) The notation $\mathfrak{B}_A$ denotes the Borel field relative to the set A (that is, $\mathfrak{B}_A$ is consisted of all subsets of A).

5) In the following we employ the notations φ_A, φ_B etc. for the characteristic functions.

$$\begin{aligned}\gamma(A) &= \sum_{l=1}^{\infty}\beta(A\cap Y_l) \geqq \sum_{l=1}^{\infty}\alpha(A\cap Y_l)\\ &= \sum_{l=1}^{\infty}\sup_{(A\cap Y_l)}\sum_{(A\cap Y_l;k)}\limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m_i[T^{-i}(A\cap Y_l)_k]\\ &= \sum_{l=1}^{\infty}\sup_{(A\cap Y_l)}\sum_{(A\cap Y_l;k)}\limsup_n \frac{1}{n}\sum_{i=0}^{n-1}\int_A \varphi_{(A\cap Y_l)_k}(T^i x)dm\\ &\geqq \sum_{l=1}^{\infty}\sup_{(A\cap Y_l)}\sum_{(A\cap Y_l;k)}\int_A \liminf_n \frac{1}{n}\sum_{i=0}^{n-1}\varphi_{(A\cap Y_l)_k}(T^i x)dm.\\ &= \sum_{l=1}^{\infty}\sup_{(A\cap Y_l)}\sum_{(A\cap Y_l;k)}\int_A \widetilde{\varphi}_{(A\cap Y_l)_k}(x)dm\\ &= \sum_{l=1}^{\infty}\int_A \widetilde{\varphi}_{A\cap Y_l}(x)dm = \int_A \widetilde{\varphi}_A(x)dm = \int_A 1dm = m(A).\end{aligned}$$

Thus the proof of (viii) is complete.

Next let us suppose $f\in L(X,\mathfrak{B},m)$, then by (v) we have $f\in L(X,\mathfrak{B},\gamma)$. Hence from Birkhoff's ergodic theorem it follows that the limit

$$\widetilde{f}(x) = \lim_n \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)$$

exists almost everywhere (γ) and then almost everywhere (m) on account of (vii), and furthermore $\widetilde{f}\in L(X,\mathfrak{B},\gamma)$. Since the limit function $\widetilde{f}$ is an invariant function, it follows that, for any pair of real numbers a and b, the set $\{x;\ a\leqq\widetilde{f}(x)<b\}$ is an invariant set. Hence, from $\widetilde{f}\in L(X,\mathfrak{B},\gamma)$, (viii), and the definition of the integral, we have easily $\widetilde{f}\in L(X,\mathfrak{B},m)$. Thus it was proved that (I) implies (B).

It will be shown by Example 2 in § 3 that (B) does not necessarily imply (I).

Next we shall prove that, in case m is finite, the statements (I), (I′) and (B) are equivalent to each other. We have already proved that (I) implies (B), so that it remains to prove that (B) implies (I′) and (I′) implies (I).

(B) → (I′)[6]: Since m is finite, we have, by Lebesgue's convergence theorem, (B) and Lemma 2, that there exists a constant K such that for any set A,

$$\limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m(T^{-i}A) = \limsup_n \frac{1}{n}\sum_{i=0}^{n-1}\int_X \varphi_A(T^i x)\,dm$$

6) The proof of this part is due to Ryll-Nardzewski [4].

$$\leqq \int_X \limsup_n \frac{1}{n}\sum_{i=0}^{n-1}\varphi_A(T^i x)dm = \int_X \widetilde{\varphi_A}(x)dm \leqq K\int_X \varphi_A(x)dm = K\cdot m(A),$$

which is the required.

(I′) → (I) : Since m is finite, we can show that for any set A

(6) $$\gamma(A) = \beta(A) = \lim_n \alpha\,(T^{-n}A).$$

Let us now suppose that there exists a set A such that

$$m(A) > 0,\ \limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m(T^{-i}A) = 0.$$

Then it is easy to see that for each positive integer j

$$\limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m[T^{-i}(T^{-j}A)] = \limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m(T^{-i}A) = 0.$$

Hence from the definition of α it follows

(7) $$\alpha(T^{-n}A) = 0\ (n = 0, 1, 2, \dots).$$

Let us put

$$\widetilde{A} = \bigcup_{i=0}^{\infty} T^{-i}A.$$

Since, in the present case, γ is a finite invariant measure and $\widetilde{A}\supset T^{-1}\widetilde{A}$, it follows that $\widetilde{A}$ is an invariant set. Thus we have

$$\gamma(\widetilde{A}) \geqq m(\widetilde{A}) \geqq m(A) > 0.$$

On the other hand, from (6) and (7), it follows

$$\gamma(\widetilde{A}) \leqq \sum_{i=0}^{\infty}\gamma(T^{-i}A) = \sum_{i=0}^{\infty}\gamma(A) = 0.$$

This contradiction shows that if A is any set of positive measure, it follows

$$\limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m(T^{-i}A) > 0.$$

Hence, from (I′) and the above inequality we get (I).

Thus Theorem 1 is now completely proved.

REMARK 1. The statement (I) implies that for any set A of positive measure

$$\liminf_n \frac{1}{n}\sum_{i=0}^{n-1} m(T^{-i}A) > 0.$$

In fact, let A be any set of positive measure, then we can choose a set Y such that $0 < m(A \cap Y) < \infty$. Then it is easy to see that for each positive integer j

$$\limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m[T^{-i}(A \cap Y)]$$

$$= \limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m[T^{-i}(T^{-j}(A\cap Y))] \leqq K\cdot m[T^{-j}(A\cap Y)].$$

Thus we have

$$\limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m[T^{-i}(A\cap Y)] \leqq K\cdot \inf_j m[T^{-j}(A\cap Y)],$$

so that

$$\liminf_n \frac{1}{n}\sum_{i=0}^{n-1} m(T^{-i}A) \geqq \inf_j m[(T^{-j}(A\cap Y)]$$

$$\geqq \frac{1}{K}\cdot \limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m[T^{-i}(A\cap Y)] > 0.$$

By the above remark, the proof of Theorem 1 may be considerably simplified, but our method of proof of Theorem 1 enables us to prove Theorem 2.

REMARK. 2. The question of whether (I′) implies (B) or not is still open.

The following result due to F. Riesz [3] follows immediately from Theorem 1.

COROLLARY. *If there exist two positive constants K_1 and K_2 such that for any set A*

$$K_1\cdot m(A) \leqq \frac{1}{n}\sum_{i=0}^{n-1} m(T^{-i}A) \leqq K_2\cdot m(A) \quad (n = 1, 2, \ldots.),$$

then the statement (B) *holds.*

Next we shall state the theorem which is a modefication of Ryll-Nardzewski's theorem [4].

THEOREM 2. *The statements* (II) *and* (B) *imply the statements* (B) *and* (II′) *respectively, but the converses are not true.*

PROOF. (II′) $\to (B)$: This implication is proved similarly as Lemma 1 and Theorem 1, so that we shall sketch the proof.

With respect to the sets X_j's in (II), we put

$$\alpha^*(A) = \sup_j \sup_{(A)} \sum_{(A\,k)} \limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m(X_j\cap T^{-i}A_k)$$

for any set A. Then the non-negative set function α^* has the following properties:

(i) α^* is finitely additive;

(ii) $\alpha^*(A) \leqq \alpha^*(T^{-1}A)$ for every set A;

(iii) $\alpha^*(A) \leqq K\cdot m(A)$ for every set A;

(vi) $\alpha^*(A) \geqq \sup_j \limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m(X_j\cap T^{-i}A)$ for every set A.

We put further for any set A

$$\beta^*(A) = \lim_n \alpha^*(T^{-n}A).$$

For any fixed set B of finite measure, the non-negative set function $\beta^*(A \cap B)$ of variable A has the following properties:

(v) $\beta^*(A\cap B) \leqq K^2 \cdot m(A \cap B)$ for every set A;

(vi) $\beta^*(A\cap B)$ is completely additive as the set function of A;

(vii) $\beta^*[T^{-1}(A\cap B)] = \beta^*(A\cap B)$ for every set A;

(viii) $\beta^*(A\cap B) \geqq \sup_j \limsup_n \frac{1}{n}\sum_{i=0}^{n-1} m[X_j \cap T^{-i}(A\cap B)]$ for every set A.

Finally we put for any set A

$$\gamma^*(A) = \lim_j \beta^*(A\cap X_j).$$

Then the non-negative set function γ^* has the following properties:

(ix) $\gamma^*(A) \leqq K^2 \cdot m(A)$ for every set A;

(x) γ^* is an invariant measure on $\mathfrak{B}$;

(xi) γ^* is equivalent to m;

(xii) $\gamma^*(A) \geqq m(A)$ for any invariant set A.

From the properties (ix)-(xii) of γ^*, it is easy to see that (II) implies (B).

$(B) \to (\mathrm{II}')$[7]: From (B) and Lemma 2 there exists a constant K such that for any sets Y and A of finite measure

$$\limsup_n \frac{1}{n}\sum_{i=1}^{n-1} m(Y\cap T^{-i}A) = \limsup_n \frac{1}{n}\sum_{i=0}^{n-1}\int_Y \varphi_A(T^i x)dm$$

$$\leqq \int_Y \limsup_n \frac{1}{n}\sum_{i=0}^{n-1} \varphi_A(T^i x)dm = \int_Y \tilde{\varphi}_A(x)dm \leqq \int_X \tilde{\varphi}_A(x)dm$$

$$\leqq K\int_X \varphi_A(x)dm = K\cdot m(A),$$

which implies (II′).

It will be shown by Example 1 in § 3 that (B) and (II′) do not necessarily imply (II) and (B), respectively.

By use of Theorem 2 we shall prove the following two theorems.

THEOREM 3. *If there exists a finite invariant measure equivalent to m, then the statements* (II), (II′) *and* (B) *are equivalent to each other.*

The assumption of the theorem cannot be omitted.

PROOF. Theorem 2 shows that (II) implies (B) and (B) Implies (II′), so that for the present purpose it is sufficient to prove that (II′) implies (II).

Let A be any set of positive measure, μ a finite invariant measure equivalent to m, and $\{X_j\}$ the sequence of sets in (II′). Then from Birkhoff's ergodic theorem it follows that the limit function $\tilde{\varphi}_A$ of the means of φ_A is

7) The proof of this part is due to Ryll-Nardzewski [4].

$$\int_X \widetilde{\varphi}_A(x)d\mu = \mu(A). \tag{8}$$

defined almost everywhere (μ) and hence almost everywhere (m), and that Since μ is equivalent to m and $m(A) > 0$, it follows $\mu(A) > 0$. Let B be the set $\{x; \widetilde{\varphi}_A(x) > 0\}$, then, from $\mu(A) > 0$ and (8), it follows $\mu(B) > 0$. Hence we can choose a set $X_{j_0} \in \{X_j\}$ such that $\mu(X_{j_0} \cap B) > 0$. Since m is equivalent to μ, it holds $m(X_{j_0} \cap B) > 0$. Thus we have

$$\sup_j \limsup_n \frac{1}{n} \sum_{i=0}^{n-1} m(X_j \cap T^{-i}A) \geqq \limsup_n \frac{1}{n} \sum_{i=0}^{n-1} m(X_{j_0} \cap T^{-i}A)$$

$$= \limsup_n \frac{1}{n} \sum_{i=0}^{n-1} \int_{X_{j_0}} \varphi_A(T^i x)dm \geqq \int_{X_{j_0}} \liminf_n \frac{1}{n} \sum_{i=0}^{n-1} \varphi_A(T^i x)dm$$

$$= \int_{X_{j_0}} \widetilde{\varphi}_A(x)dm > 0.$$

Combining (II′) and the above inequality we get (II).

From Theorem 2 and the fact proved above, it follows that the assumption of the theorem cannot be omitted (cf. Example 1 in § 3).

THEOREM 4. *If X is the union of countable invariant subsets of finite measure, then the statements* (II), (II′) *and* (B) *are equivalent to each other.*

The assumption of the theorem cannot be omitted.

PROOF. Theorem 2 shows that (II) implies (B) and (B) implies (II′), so that for the purpose it is sufficient to prove that (II′) implies (II).

From the assumption there exist the invariant sets Y_j's such that

$$X = \bigcup_{j=1}^{\infty} Y_j, \quad Y_i \cap Y_j = 0 \ (i \neq j), \quad 0 < m(Y_j) < \infty \ (j = 1\ 2, \dots).$$

If we define the set function β^* as in the proof of Theorem 2, then it is easy to see that, for each j, β^* is a finite invariant measure on $\mathfrak{B}_{Y_j}$ and is equivalent to m as the measure on $\mathfrak{B}Y_j$.

Let us now put for any set A

$$\mu(A) = \sum_{j=1}^{\infty} \beta^*(A \cap Y_j)/2^j \beta^*(Y_j).$$

Then it is clear that μ is a finite invariant measure on $\mathfrak{B}$ equivalent to m. Hence we obtain the conclusion by Theorem 3.

It follows, from Theorem 2 and the fact proved above, that the assumption of the theorem cannot be omitted (cf. Example 1 in § 3).

REMARK. Theorem 3 is essentially equivalent to Theorem 4.

In the Proof of Theorem 4 we have shown that the assumption of Theorem 3 follows from the assumption of Theorem 4 and the statement (II′), so that for the present purpose it is sufficient to show that the assumption of Theorem 3 and the statement (B) imply the assumption of Theorem 4.

Let $\mathfrak{B}_0$ denote the class such that sets of $\mathfrak{B}_0$ are the invariant sets of finite positive measure and are mutually equivalent. Since m is σ-finite, the class $\mathfrak{B}_0$ is at most countable. Let us denote by Y the union of all sets of $\mathfrak{B}_0$, then the set $X-Y$ must be a set of measure zero or an invariant set which has no invariant subsets of finite positive measure.

Now let us suppose that $X-Y$ is not the set of measure zero. Then there is a set A such that

$$A \subset X-Y,\ 0<m(A)<\infty.$$

Since $\varphi_A \in L(X, \mathfrak{B}, m)$ and (B) holds, the limit $\tilde{\varphi}_A$ of the means of φ_A is defined almost everywhere (m) and $\tilde{\varphi}_A \in L(X, \mathfrak{B}, m)$.

On the other hand, let μ be a finite invariant measure equivalent to m, then from Birkhoff's ergodic theorem it follows that $\tilde{\varphi}_A$ is dsfined almost everywhere (μ) and that

$$\int_X \tilde{\varphi}_A(x) d\mu = \mu(A). \tag{9}$$

From $m(A)>0$ we get $\mu(A)>0$. Hence, if we put $B=\{x;\ \tilde{\varphi}_A(x)>0\}$, then from (9) and $\mu(A)>0$ it follows $\mu(B)>0$, so that $m(B)>0$. Then there exists a positive number ε such that the set $\{x; \tilde{\varphi}_A(x)>\varepsilon\}$ is of positive measure. Since $\tilde{\varphi}_A$ is the invariant function and $\tilde{\varphi}_A \in L(X, \mathfrak{B}, m)$, the set $\{x;\ \tilde{\varphi}_A(x)>\varepsilon\}$ is an invariant set of finite positive measure. This contradicts the assumption for the set $X-Y$. Hence $X-Y$ is the set of measure zero.

Thus we can conclude that X is the union of countable invariant subsets of finite measure.

3. Counter examples. We shall now show by example that, in case m is not finite, (B) does not necessarily imply (II) and that (II′) does not necessarily imply (B).

EXAMPLE 1. We shall start from the measure space $(X, \mathfrak{B}, \mu)$ and the transformation T constructed by P. R. Halmos [2; pp. 743-744].

Let us define the collection of the linear intervals $J_{n,k}$'s in the (s,t)-plane by

$$J_{n,k}=\left\{(s,k);\ \frac{1}{2^{n+1}} \leqq s < \frac{1}{2^n}\right\} \begin{pmatrix} k=0,1,\dots,2^{n+1}-1; \\ n=0,1,2,\dots \end{pmatrix}.$$

Let $(X, \mathfrak{B}, \mu)$ be the measure space such that X is the union of all $J_{n,k}$'s, $\mathfrak{B}$ is the class of the Lebesgue measurable subsets of X, and μ is the ordinary linear Lebesgue measure on $\mathfrak{B}$. Let T_0 be any one to one, measurable, measure-preserving (with respect to μ), and ergodic transformation of $\bigcup_{n=0}^{\infty} J_{n,0}$ onto itself (for example, the transformation T_0 is defined by

$$T_0(s,0)=(\{s+\theta\},0),\ \ (s,0)\in\bigcup_{n=0}^{\infty}J_{n,0},$$

where θ is a fixed irrational number and $\{s+\theta\}$ denotes the fractional part of $s+\theta$). Further, let us define a transformation T by

$$T(s,t)=(s,t+1),\ \text{if } (s,t)\in J_{n,k}\begin{pmatrix}k=0,1,\ldots,2^{n+1}-2;\\ n=0,1,2,\ldots\end{pmatrix}$$
$$=T_0(s,0),\ \text{if } (s,t)\in J_{n,2^{n+1}-1}\ (n=0,1,2,\ldots).$$

Then it can be proved that T is a one to one, measurable, measure-preserving (with respect to μ), and ergodic transformation of X onto itself (see [2]).

In the following, by x we denote the point (s,t) of X simply.

From Birkhoff's ergodic theorem it follows that for any function $f\in L(X,\mathfrak{B},\mu)$ the limit function $\tilde{f}$ of the means of f is defined almost everywhere (μ) and $\tilde{f}\in L(X,\mathfrak{B},\mu)$.

Since $\tilde{f}$ is an invariant function and T is ergodic, the function $\tilde{f}(x)$ is constant almost everywhere (μ), so that $\tilde{f}\in L(X,\mathfrak{B},\mu)$ implies that $\tilde{f}(x)$ vanishes almost everywhere (μ). Hence, for any sets Y and A of finite measure, we have

$$\limsup_n\frac{1}{n}\sum_{i=0}^{n-1}\mu(Y\cap T^{-i}A)=\limsup_n\frac{1}{n}\sum_{i=0}^{n-1}\int_Y\varphi_A(T^ix)d\mu$$

(1)
$$\leqq\int_Y\limsup_n\frac{1}{n}\sum_{i=0}^{n-1}\varphi_A(T^ix)d\mu=\int_Y\tilde{\varphi}_A(x)d\mu=0,$$

so that the statement (II) does not holds.

On the other hand, since μ is invariant under T, the statement (B) is the immediate consequence of Birkhoff's ergodic theorem.

Thus we conclude that (B) does not necessarily imply (II) (see Theorem 2).

Next we define a new measure m on $\mathfrak{B}$ as follows:

$$m(A)=\mu(A)/[2(n+1)^2-1],\ \text{if } A\subset J_{n,2^{n+1}-1}(n=0,1,\ldots),$$
$$=\mu(A),\qquad \text{if } A\subset J_{n,k}\begin{pmatrix}k=0,1,\ldots,2^{n+1}-2;\\ n=0,1,2,\ldots\end{pmatrix}.$$

Then it is obvious that m is a σ-finite (but not finite) measure equivalent to μ, and T is measurable, non-singular (with respect to m) and ergodic. We shall now show that the measure space $(X,\mathfrak{B},m)$ and the transformation T have the following properties:

(i) the statement (B) does not hold;

(ii) the statement (II′) holds.

Proof of (i): Let us put

$$f(x)=2^{n+1}[2(n+1)^2-1]/2(n+1)^2,\ \text{if } x\in J_{n,2^{n+1}-1}\ (n=0,1,2,\cdots),$$

$$= 2^{n+1}/2(n+1)^2(2^{n+1}-1), \text{ if } x \in J_{n,k} \begin{pmatrix} k = 0, 1, \dots, 2^{n+1}-2; \\ n = 0, 1, 2, \dots \end{pmatrix}.$$

Then we have

$$\int_X f(x)dm = \sum_{n=0}^{\infty} \int_{J_{n,2^{n+1}-1}} f(x)dm + \sum_{n=0}^{\infty} \sum_{k=0}^{2^{n+1}-2} \int_{J_{n,k}} f(x)dm$$

$$= \sum_{n=0}^{\infty} \left(\frac{2^{n+1}[2(n+1)^2-1]}{2(n+1)^2} \right) \cdot \left(\frac{1}{2^{n+1}} \cdot \frac{1}{2(n+1)^2-1} \right)$$

$$+ \sum_{n=0}^{\infty} \sum_{k=0}^{2^{n+1}-2} \left(\frac{2^{n+1}}{2(n+1)^2(2^{n+1}-1)} \right) \cdot \left(\frac{1}{2^{n+1}} \right)$$

$$= \sum_{n=0}^{\infty} \frac{1}{(n+1)^2} < \infty,$$

so that $f \in L(X, \mathfrak{B}, m)$.

On the other hand, it holds that for each n

$$f(x) + f(Tx) + \cdots + f(T^{2^{n+1}-1}x)$$

$$= \underbrace{\frac{2^{n+1}}{2(n+1)^2(2^{n+1}-1)} + \frac{2^{n+1}}{2(n+1)^2(2^{n+1}-1)} + \cdots + \frac{2^{n+1}}{2(n+1)^2(2^{n+1}-1)}}_{(2^{n+1}-1) \text{ terms}}$$

$$+ \frac{2^{n+1}[2(n+1)^2-1]}{2(n+1)^2} = 2^{n+1}, \qquad x \in J_{n,0},$$

and further to any point x of X there corresponds a positive integer $p(x)$ such that

$$T^{p(x)}x \in \bigcup_{n=0}^{\infty} J_{n,0}.$$

Hence we get for any x of X

$$\limsup_n \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = 1.$$

If we suppose that (B) holds, then we have

$$\tilde{f}(x) = \lim_n \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = 1$$

almost everywhere (m) and $\tilde{f} \in L(X, \mathfrak{B}, m)$. This contradicts the fact that m is not finite. Hence (B) does not hold.

Proof of (ii): It is easy to see that for any set A of finite measure

(2) $$m(A) \leqq \mu(A) < \infty.$$

By (1) and (2) we get that for any sets A and Y of finite measure

$$\limsup_n \frac{1}{n} \sum_{i=0}^{n-1} m(Y \cap T^{-i}A) = 0.$$

Since the inequality in (II′) holds evidently for any set A of infinite measure, the statement (II′) holds.

Thus we conclude that (II′) does not necessarily imply (B) (see Theorem 2).

In the connection to Theorem 3 and 4 we note that the following properties can be easily shown:

(iii) there exists no finite invariant measure equivalent to m;

(iv) X has no invariant subset of finite measure.

Finally we shall show by example that, in case m is not finite, (B) does not necessarily imply (I′).

EXAMPLE 2. We define $X, \mathfrak{B}, \mu, T$ and $J_{n,k}$ ($k = 0, 1, \ldots, 2^{n+1}-1$; $n = 0, 1, 2, \ldots$) as in Example 1. We introduce a new σ-finite measure m on $\mathfrak{B}$ such that m is equivalent to μ, and for any set A

$$m(A \cap J_{n,k}) = \mu(A \cap J_{n,k}) \begin{pmatrix} k = 1, 2, \ldots, 2^{n+1}-1; \\ n = 0, 1, 2, \ldots \end{pmatrix} \tag{3}$$

and

$$m(J_{n,0}) = \infty, \quad m(A \cap J_{n,0}) \geqq \mu(A \cap J_{n,0}) \ (n = 0, 1, 2, \ldots). \tag{4}$$

In fact, it is easy to construct such measure.

Let us suppose $f \in L(X, \mathfrak{B}, m)$, then by (3) and (4) we get $f \in L(X, \mathfrak{B}, \mu)$. Hence we get similarly as in Example 1 that the limit function $\tilde{f}$ of the means of f is defined almost everywhere (m) and $\tilde{f}(x)$ vanishes almost everywhere (m). Thus the statement (B) holds.

On the other hand, if we put

$$A = \bigcup_{n=0}^{\infty} J_{n, 2^{n+1}-1},$$

we have

$$m(A) = \mu(A) = 1. \tag{5}$$

Then from (4) it follows that for each n

$$\frac{1}{2^{n+1}} \sum_{i=0}^{2^{n+1}-1} m(T^{-i}A) \geqq \frac{1}{2^{n+1}} \sum_{i=0}^{2^{n+1}-1} m(T^{-i}J_{n,2^{n+1}-1})$$
$$\geqq \frac{1}{2^{n+1}} m(T^{-(2^{n+1}-1)}J_{n,2^{n+1}-1}) = \frac{1}{2^{n+1}} m(J_{n,0}) = \infty,$$

so that

$$\limsup_{n} \frac{1}{n} \sum_{i=0}^{n-1} m(T^{-i}A) = \infty. \tag{6}$$

By (5) and (6) we have that (I′) does not hold.

Thus we conclude that (B) does not necessarily imply (I′) and then (I) (see Theorem 1).

Finally I have to express my cordial thanks to Mr. S. Yano who gave me valuable remarks and advices.

Supplement

The results of the present paper were sketched in the preliminary report, S. Tsurumi, On ergodic theorems, Proc. Japan Acad., 30(1954) pp. 331–334 in which the sentence misinserted in lines 16–17 of page 333 should be omitted.

References

[1] N. DUNFORD and D. S. MILLER, On the ergodic theorems Trans. Amer. Math. Soc., 60(1946), pp. 538-549.
[2] P. R. HALMOS, Invariant measures, Ann. of Math., 48(1947), pp. 735-754.
[3] F. RIESZ, On a recent generalisation of G. D. Birkhoff's ergodic theorem, Acta de Szeged, 11(1946-1948), pp. 193-200.
[4] C. RYLL-NARDZEWSKI, On the ergodic theorems (I) (Generalized ergodic theorems), Studia Math., 12(1951), pp. 65-73.

MATHEMATICAL INSTITUTE, TOKYO METROPOLITAN UNIVERSITY.