ISRAEL JOURNAL OF MATHEMATICS, Vol. 42, No. 4, 1982

A SIMPLE PROOF OF THE ERGODIC THEOREM USING NONSTANDARD ANALYSIS

BY
TETURO KAMAE

ABSTRACT

A simple proof of the individual ergodic theorem is given. The essential tool is the nonstandard measure theory developed by P. Loeb. Any dynamical system on an abstract Lebesgue space can be represented as a factor of a "cyclic" system with a hyperfinite cycle. The ergodic theorem for such a "cyclic" system is almost trivial because of its simple structure. The general case follows from this special case.

We consider a countably saturated nonstandard model [1]. Let k be a nonstandard positive hyperinteger and K be the set of nonnegative hyperintegers less than k. Let $\boldsymbol{K}$ be the set of all internal subsets of K. The σ-field generated by $\boldsymbol{K}$ is denoted by $\sigma(\boldsymbol{K})$. It is known that there exists a unique probability measure P defined on $(K, \sigma(\boldsymbol{K}))$ satisfying

$$P(B) = {}^\circ\left(\frac{\# B}{k}\right)$$

for any $B \in \boldsymbol{K}$, where "$\#$" denotes the internal cardinality and "$^\circ$" denotes the standard part mapping. Let $\boldsymbol{B}$ be the P-completion of $\sigma(\boldsymbol{K})$. Thus $(K, \boldsymbol{B}, P)$ is a complete probability space.

LEMMA 1 (P. Loeb [2]). *For any real-valued integrable function f on $(K, \boldsymbol{B}, P)$ and $\varepsilon > 0$ (standard), there exist internal functions F and G such that*

(1) $G(x) \leqq f(x) \leqq F(x)$ *for all* $x \in K$, *and*

(2) $|\int_B f(x)dP(x) - (1/k)^*\Sigma_{x\in B} F(x)| < \varepsilon$, $|\int_B f(x)dP(x) - (1/k)^*\Sigma_{x\in B} G(x)| < \varepsilon$ *for any internal subset B of K.*

Let φ be the transformation on K such that

$$\varphi(x) = \begin{cases} x+1 & (x < k-1), \\ 0 & (x = k-1). \end{cases}$$

Received September 25, 1981 and in revised form April 22, 1982

Then it is clear that φ preserves P. Such a dynamical system $(K, \boldsymbol{B}, P, \varphi)$ will be called a *universal system* since, as is proved later, any measure preserving transformation on an abstract Lebesgue space is a factor of it.

THEOREM 1 (Ergodic Theorem for a universal system). *For any* $f \in L_1(K, \boldsymbol{B}, P)$,

$$\hat{f}(x) \equiv \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(\varphi^i x)$$

exists for almost all $x \in K$ *and is integrable* (w.r.t. P). *Moreover,*

$$\int \hat{f}(x)dP(x) = \int f(x)dP(x).$$

PROOF. We may assume that $f(x) \geqq 0$ for any $x \in K$ since the general case follows from this special case. Let

$$\bar{f}(x) = \overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{i=0}^{n-1} f(\varphi^i x),$$

$$\underline{f}(x) = \underline{\lim_{n\to\infty}} \frac{1}{n} \sum_{i=0}^{n-1} f(\varphi^i x).$$

To prove our theorem, it is sufficient to prove that $\bar{f}$ is integrable and

$$\int \bar{f}(x)dP(x) \leqq \int f(x)dP(x) \leqq \int \underline{f}(x)dP(x).$$

Take any $\varepsilon > 0$ and a large integer M (both standard). By Lemma 1, there exist internal functions F and G on K satisfying

(1) $f(x) \leqq F(x)$ and $G(x) \leqq \bar{f}(x) \wedge M$ for all $x \in K$, and

(2) $|\int_B f(x)dP(x) - (1/k)^*\Sigma_{x\in B} F(x)| < \varepsilon$ and
$|\int_B \bar{f}(x) \wedge M dP(x) - (1/k)^*\Sigma_{x\in B} G(x)| < \varepsilon$

for any internal subset B of K. For any $x \in K$, there exists a positive integer n (standard) such that

$$\bar{f}(x) \wedge M \leqq \frac{1}{n} \sum_{i=0}^{n-1} f(\varphi^i x) + \varepsilon.$$

Therefore, for $m = 0, 1, \cdots, n-1$, it holds that

$$\begin{aligned} G(\varphi^m x) &\leqq \bar{f}(\varphi^m x) \wedge M \\ &\leqq \frac{1}{n} \sum_{i=0}^{n-1} f(\varphi^i x) + \varepsilon \\ &\leqq \frac{1}{n} {}^*\sum_{i=0}^{n-1} F(\varphi^i x) + \varepsilon. \end{aligned}$$

Thus,

$$(*) \qquad {}^*\sum_{i=0}^{n-1} G(\varphi^i x) \leqq {}^*\sum_{i=0}^{n-1} F(\varphi^i x) + n\varepsilon.$$

For any $x \in K$, let $T(x)$ be the least positive hyperinteger n satisfying (*) or k if such n does not exist. Then, T is an internal function such that $T(x)$ is a standard integer for any $x \in K$. Therefore,

$$r \equiv \max_{x \in K} T(x)$$

exists and is standard. Let $T_0 = 0$ and for positive hyperintegers j, define T_j (hyper-)inductively by

$$T_j = T_{j-1} + T(T_{j-1}).$$

Let J be the first j such that $k - r \leqq T_j < k$. Then it holds that

$$\begin{aligned}
\frac{1}{k} {}^*\sum_{x=0}^{T_J - 1} G(x) &= \frac{1}{k} {}^*\sum_{j=0}^{J-1} {}^*\sum_{i=0}^{T(T_j)-1} G(\varphi^i T_j) \\
&\leqq \frac{1}{k} {}^*\sum_{j=0}^{J-1} \left({}^*\sum_{i=0}^{T(T_j)-1} F(\varphi^i T_j) + T(T_j)\varepsilon \right) \\
&= \frac{1}{k} {}^*\sum_{x=0}^{T_J-1} F(x) + \frac{T_J}{k}\varepsilon \\
&\leqq \frac{1}{k} {}^*\sum_{x=0}^{T_J-1} F(x) + \varepsilon.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\int \bar{f}(x) \wedge M dP(x) &= \int_{\{0 \leq x < T_J\}} \bar{f}(x) \wedge M dP(x) \\
&\leqq \frac{1}{k} {}^*\sum_{x=0}^{T_J-1} G(x) + \varepsilon \\
&\leqq \frac{1}{k} {}^*\sum_{x=0}^{T_J-1} F(x) + 2\varepsilon \\
&\leqq \int_{\{0 \leq x < T_J\}} f(x) dP(x) + 3\varepsilon \\
&= \int f(x) dP(x) + 3\varepsilon.
\end{aligned}$$

Letting $M \uparrow \infty$ and then $\varepsilon \downarrow 0$, we have

$$\int \bar{f}(x)dP(x) \leqq \int f(x)dP(x).$$

Since $\bar{f}$ is integrable, $\underline{f}$ is also integrable. By Lemma 1, for any $\varepsilon > 0$ (standard), there exist internal functions G and H such that

(1) $G(x) \leqq f(x)$ and $\underline{f}(x) \leqq H(x)$ for any $x \in K$, and

(2) $|\int_B f(x)dP(x) - (1/k)^* \Sigma_{x \in B} G(x)| < \varepsilon$ and
$|\int_B \underline{f}(x)dP(x) - (1/k)^* \Sigma_{x \in B} H(x)| < \varepsilon$

for any internal subset B of K. The same argument as above leads us to the conclusion that

$$\int f(x)dP(x) \leqq \int \underline{f}(x)dP(x).$$

Thus we have

$$\int \bar{f}(x)dP(x) \leqq \int \underline{f}(x)dP(x)$$

which completes the proof.

COROLLARY (Ergodic Theorem). *Let T be a measure preserving transformation on a probability measure space $(\Omega, \mathbf{A}, \mu)$. Then for any $f \in L_1(\Omega, \mathbf{A}, \mu)$,*

$$\hat{f}(\omega) \equiv \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i \omega)$$

exists for almost all $\omega \in \Omega$ (w.r.t. μ) and $\int \hat{f} d\mu = \int f d\mu$.

PROOF. Let $\mathbf{N}$ be the set of nonnegative integers. Let σ be the shift on the product space $\mathbf{R}^{\mathbf{N}}$; i.e. $(\sigma\alpha)(n) = \alpha(n+1)$ for any $\alpha \in \mathbf{R}^{\mathbf{N}}$ and $n \in \mathbf{N}$. Define $\tau : \Omega \to \mathbf{R}^{\mathbf{N}}$ by $\tau(\omega)(n) = f(T^n\omega)$ for any $\omega \in \Omega$ and $n \in \mathbf{N}$. Let ν be the probability measure on $(\mathbf{R}^{\mathbf{N}}, C)$ induced by μ through τ, where C is the Borel field on $\mathbf{R}^{\mathbf{N}}$. Let $\pi : \mathbf{R}^{\mathbf{N}} \to \mathbf{R}$ be the 0-th coordinate mapping; i.e. $\pi(\alpha) = \alpha(0)$. Then, our corollary is equivalent to the statement that

$$\hat{\pi}(\alpha) \equiv \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \pi(\sigma^i \alpha)$$

exists for almost all $\alpha \in \mathbf{R}^{\mathbf{N}}$ (w.r.t. ν) and $\int \hat{\pi} d\nu = \int \pi d\nu$. Therefore, it is sufficient to prove that $(\mathbf{R}^{\mathbf{N}}, C, \nu, \sigma)$ is a *factor* of some universal system in virtue of Theorem 1; i.e. there exists a universal system $(\boldsymbol{K}, \boldsymbol{B}, P, \varphi)$ and a measure preserving transformation $g:(K, \boldsymbol{B}, P) \rightarrow (\mathbf{R}^{\mathbf{N}}, C, \nu)$ such that $g(\varphi(x)) = \sigma(g(x))$ for almost all $x \in K$ (w.r.t. P). Thus, our corollary follows from the following theorem.

THEOREM 2. *Let $(\Omega, \boldsymbol{A}, \mu)$ be a normalized abstract Lebesgue space and T be a measure preserving transformation on it. Then, $(\Omega, \boldsymbol{A}, \mu, T)$ is a factor of any universal system $(K, \boldsymbol{B}, P, \varphi)$.*

PROOF. There exists an isomorphism h from the probability space $(\Omega, \boldsymbol{A}, \mu)$ to $([0,1], \boldsymbol{D}, \lambda)$, where $\boldsymbol{D}$ is the Borel field on $[0,1]$ and λ is a probability Borel measure on $[0,1]$. Define $r: \Omega \rightarrow [0,1]^{\mathbf{N}}$ by $r(\omega)(n) = h(T^n\omega)$ for any $\omega \in \Omega$ and $n \in \mathbf{N}$. Let ν be the measure on $([0,1]^{\mathbf{N}}, \boldsymbol{C})$ induced by μ through r, where $\boldsymbol{C}$ is the Borel field on $[0,1]^{\mathbf{N}}$. It is clear that the dynamical system $(\Omega, \boldsymbol{A}, \mu, T)$ is isomorphic to $([0,1]^{\mathbf{N}}, \boldsymbol{C}, \nu, \sigma)$, where σ is the shift on $[0,1]^{\mathbf{N}}$. We recall that $\alpha \in [0,1]^{\mathbf{N}}$ is *typical* for ν, if

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(\sigma^i \alpha) = \int f d\nu$$

holds for any $f \in C([0,1]^{\mathbf{N}})$. We will prove later that a typical element exists for any shift invariant probability measure on $([0,1]^{\mathbf{N}}, \boldsymbol{C})$. Take a typical element α for ν. Define $g: K \rightarrow [0,1]^{\mathbf{N}}$ by $g(x) = {}^\circ(\sigma^x\alpha)$ for any $x \in K$, where $\sigma^x\alpha \in {}^*([0,1]^{\mathbf{N}})$ is the value of the natural extension of the mapping $n \mapsto \sigma^n\alpha$ from $\mathbf{N}$ to $[0,1]^{\mathbf{N}}$ at $x \in {}^*\mathbf{N}$. Take any $f \in C([0,1]^{\mathbf{N}})$. Then, since ${}^\circ f(\sigma^x\alpha) = f(g(x))$ for any $x \in K$, it holds [1], [2] that

$$\int f d\nu = \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(\sigma^i\alpha) = {}^\circ\left(\frac{1}{k} {}^*\sum_{x=0}^{k-1} f(\sigma^x\alpha)\right) = \int f(g(x)) dP(x).$$

Thus, g is measure preserving. Since σ is continuous, it holds that

$$\sigma(g(x)) = \sigma({}^\circ\sigma^x\alpha) = {}^\circ\sigma(\sigma^x\alpha) = {}^\circ\sigma^{x+1}\alpha = g(x+1) = g(\varphi(x))$$

for any $x \in K$ except for $x = k-1$. Thus, $([0,1]^{\mathbf{N}}, C, \nu, \sigma)$ and $(\Omega, \boldsymbol{A}, \mu, T)$ are factors of $(K, \boldsymbol{B}, P, \varphi)$.

LEMMA 2. *Let ν be a σ-invariant probability measure on $([0,1]^{\mathbf{N}}, \boldsymbol{C})$, where σ is the shift on $[0,1]^{\mathbf{N}}$ and $\boldsymbol{C}$ is the Borel field on it. Then, there exists a typical element for ν.*

PROOF. Various proofs are known for this lemma. Here, we reproduce J. Ville's proof [3] which does not use the ergodic theorem. We take a sequence of periodic elements α_n $(n = 1, 2, \cdots)$ in $[0,1]^{\mathbf{N}}$ such that μ_{α_n} converges weakly to ν as $n \rightarrow \infty$, where

$$\mu_{\alpha_n} = \frac{1}{c_n}(\delta_{\alpha_n} + \delta_{\sigma\alpha_n} + \cdots + \delta_{\sigma^{c_n-1}\alpha_n}),$$

c_n being a period of α_n and δ_α denoting the Dirac measure at $\alpha \in [0,1]^{\mathbf{N}}$. This can be done as is shown later. Then, the following α is typical for ν provided that the sequence t_n $(n = 1,2,\cdots)$ of positive integers is selected so as to increase sufficiently fast:

$$\alpha(n) = \alpha_m(n - T_m)$$

for any $n \in \mathbf{N}$ with $T_m \leqq n < T_{m+1}$, where $T_0 = 0$ and $T_i = T_{i-1} + c_i t_i$ $(i = 1,2,\cdots)$. All we have to prove is that for any neighbourhood U of ν in the weak topology of the space of Borel measures on $[0,1]^{\mathbf{N}}$, there exists a periodic element $\beta \in [0,1]^{\mathbf{N}}$ such that $\mu_\beta \in U$. There exist positive integers m, n and a positive number δ such that if a Borel measure μ on $[0,1]^{\mathbf{N}}$ satisfies $|\mu(B) - \nu(B)| < \delta$ for any subset B of the form

$$\left[\frac{j_0}{m}, \frac{j_0+1}{m}\right] \times \cdots \times \left[\frac{j_{n-1}}{m}, \frac{j_{n-1}+1}{m}\right] \times [0,1] \times [0,1] \times \cdots,$$

where $j_0, j_1, \cdots, j_{n-1}$ are integers from 0 through $m-1$, then $\mu \in U$. Let

$$\Sigma = \left\{\frac{1}{2m}, \frac{3}{2m}, \cdots, \frac{2m-1}{2m}\right\}.$$

Let λ be the measure on Σ^n such that

$$\lambda\left(\left(\frac{j_0+\frac{1}{2}}{m}, \cdots, \frac{j_{n-1}+\frac{1}{2}}{m}\right)\right) = \nu\left(\left[\frac{j_0}{m}, \frac{j_0+1}{m}\right] \times \cdots \times \left[\frac{j_{n-1}}{m}, \frac{j_{n-1}+1}{m}\right] \times [0,1] \times \cdots\right)$$

for any integers $j_0, j_1, \cdots, j_{n-1}$ from 0 through $m-1$. Since ν is σ-invariant, λ satisfies

$$(\#) \qquad \sum_{\xi_0 \in \Sigma} \lambda((\xi_0, \xi_1, \cdots, \xi_{n-1})) = \sum_{\xi_0 \in \Sigma} \lambda((\xi_1, \cdots, \xi_{n-1}, \xi_0))$$

for any $\xi_1, \cdots, \xi_{n-1} \in \Sigma$. Let N be a sufficiently large integer. We modify λ a little to get such a probability measure $\tilde{\lambda}$ on Σ^n that

(1) for any $\xi \in \Sigma^n$, $N\tilde{\lambda}(\xi)$ is a positive integer such that $|\lambda(\xi) - \tilde{\lambda}(\xi)| < \delta$, and

(2) $(\#)$ holds for $\tilde{\lambda}$ in place of λ.

This can be done because

(i) rational points are dense in the set of points $(\lambda(\xi_0, \cdots, \xi_{n-1});\ \xi_0, \cdots, \xi_{n-1} \in \Sigma)$ satisfying $(\#)$ together with

$$(*) \qquad \sum_{\xi_0, \cdots, \xi_{n-1} \in \Sigma} \lambda(\xi_0, \cdots, \xi_{n-1}) = 1, \qquad \text{and}$$

(ii) there exists a solution of (#) together with (*) satisfying $\lambda(\xi_0,\cdots,\xi_{n-1})=m^{-n}>0$ for any $\xi_0,\cdots,\xi_{n-1}\in\Sigma$.

Take a longest sequence $\xi^0,\xi^1,\cdots,\xi^{r-1}$ of elements in Σ^n such that

(1) $(\xi^i_1,\cdots,\xi^i_{n-1})=(\xi^{i+1}_0,\cdots,\xi^{i+1}_{n-2})$ for any $i=0,1,\cdots,r-1$, and

(2) $\#\{i;0\leqq i<r,\xi^i=\xi\}\leqq N\lambda(\xi)$ for any $\xi\in\Sigma^n$,

where $\xi^i=(\xi^i_0,\cdots,\xi^i_{n-1})$ $(i=0,1,\cdots,r)$ and $\xi^r\equiv\xi^0$.

Then, by a graph-theoretical consideration, it is not difficult to prove that the equality holds in the above inequality in (2) for any $\xi\in\Sigma^n$. Let β be the periodic element in $[0,1]^{\mathbb{N}}$ with period $n+r-1$ defined by

$$(\beta(0),\beta(1),\cdots,\beta(n+r-2))=(\xi^0_0,\xi^0_1,\cdots,\xi^0_{n-1},\xi^1_{n-1},\xi^2_{n-1},\cdots,\xi^{r-1}_{n-1}).$$

Then, it holds that $\mu_\beta\in U$.

Acknowledgement

The author would like to thank Prof. Benjamin Weiss (The Hebrew University of Jerusalem) for his useful advice.

References

1. M. Davis, *Applied Nonstandard Analysis*, John Wiley & Sons, New York, 1977.

2. P. Loeb, *Conversion from non-standard to standard measure space and applications to probability theory*, Trans. Am. Math. Soc. **211** (1975).

3. J. Ville, *Etude Critique de la Notion de Collectif*, Paris, 1939.

Department of Mathematics
Osaka City University
Sugimoto-cho, Osaka, Japan