Publ. RIMS, Kyoto Univ.
7 (1971/72), 541-580

Remarks on the Isomorphism Theorems for Weak Bernoulli Transformations in General Case

By

S. Ito*, H. Murata** and H. Totoki

§ 1. Introduction

Recently D. S. Ornstein [4] proved that entropy is a complete invariant for Bernoulli shifts with finite generators. M. Smorodinsky [7] generalized this important result to the case of Bernoulli shifts with countable generators. Ornstein [5] also proved the above fact for Bernoulli shifts with infinite entropy. On the other hand N.A. Friedman and Ornstein [2] obtained the result that a weak Bernoulli transformation with finite generator is isomorphic to a Bernoulli shift with the same entropy, so that entropy is also a complete invariant for these transformations.

Moreover Ornstein [5] proved that if T is a mixing (or only ergodic) transformation on a σ-field $\mathscr{F}$ and $\mathscr{F}$ is an increasing union of invariant sub-σ-fields $\mathscr{F}_i$ such that each $T|\mathscr{F}_i$ (restriction to $\mathscr{F}_i$) is a Bernoulli shift with finite entropy, then T is itself a (generalized) Bernoulli shift. Combining this theorem and the results mentioned above, the isomorphism theorem for generalized weak Bernoulli transformation (see Definition 1 in §2) is easily obtained, because these transformations restricted to each approximating σ-fields are weak Bernoulli transformations with finite generators and hence Bernoulli shifts. This argument, however, make use of the results of [2], [4], [5] and [7]. Therefore it seems to be significant to prove the isomorphism theorem in a straight way,

Received June 29, 1971.
* Department of Applied Mathematical Science, Tokyo University of Education, Otsuka Tokyo, 112, Japan.
** Department of Mathematics, Hiroshima University, Hiroshima, 730, Japan.

unifying the above mentioned results.

The purpose of this note is to give a direct proof of the isomorphism theorem for generalized weak Bernoulli transformations. Another proof of this theorem was also obtained by G. Maruyama, which appears in this issue, from another view point using a different way.

The proof of the theorem will be divided into two steps. Firstly we will prove the theorem for weak Bernoulli transformations with countable generators having finite entropy (Proposition 1) in §3. We do this without using the isomorphism theorem for Bernoulli shifts. So this includes the results of [4], [7] and [2], since Bernoulli shifts are weak Bernoulli as a matter of course. Next in §4 we will prove the theorem in general case, which includes the results of [5]. For this purpose we use the results in §3, especially Lemma 11 which is a stronger form of Lemma 7 of [4]. Although the method of our proof is analogous to the ones in [2] and [5], we will go into details to serve the purpose of this note.

In §5 we will give some examples of weak Bernoulli transformations which contain mixing Markov shifts with countable generators. §2 is the preliminary one where we will prepare notations and definitions, especially the definition of generalized weak Bernoulli transformations.

We want to express our thanks to Professor D.S. Ornstein who kindly sent us preprints of his papers.

§2. Notations and Definitions

Let $(X, \mathscr{F}, m)$ be a non-atomic Lebesgue probability space.[1] Transformation T is always invertible and measure preserving. Let $P, Q, R, \cdots$ denote measurable partitions of X, and when they are at most countable, we always assume they are ordered partitions. We will define some notations for (ordered) partitions. Given a partition $P=\{p_1, p_2 \cdots\}$, $N(P)$ denotes the number of the atoms $p_1, p_2, \cdots$ of P and

$$d(P)=\{m(p_1), m(p_2), \cdots\}$$

1) cf. [6] for the notions of the Lebesgue space and measurable partitions.

is the probability vector of the atoms of P. If $N(P)>N(Q)$ then we add ideal atoms $\emptyset$ to Q to make $N(P)=N(Q)$, and then we define

$$d(P, Q)=\sum_j |m(p_j)-m(q_j)|,$$

$$D(P, Q)=\sum_j m(p_j \triangle q_j).$$

Given a set F with positive measure and partitions P and Q, we will write

$$m(A|F)=m(A\cap F)/m(F), \quad A\in\mathscr{F},$$

$$P/F=\{p_1\cap F, p_2\cap F, \cdots\},$$

$$D(P, Q|F)=D(P/F, Q/F)=\sum_j m(p_j\triangle q_j|F).$$

Given partitions P and Q, $Q\overset{\varepsilon}{\subset}P$ denotes that there exists a partition $\bar{P}$ such that P is a refinement of $\bar{P}$ and $D(\bar{P}, Q)<\varepsilon$. We write $Q\subset P$ if $D(\bar{P}, Q)=0$. Note that $\bar{P}$ can be realized in the following way. Consider a partition $L=\{l_1, l_2, \cdots\}$ of the index set $\{1, 2, \cdots, N(P)\}$ of P. Then we can define a partition L_P of X as $L_P=\{\bigcup_{i\in l_j} p_i, j=1, 2, \cdots\}$; hence $N(L_P)=N(L)$. Thus there is a partition L such that $L_P=\bar{P}$. If Q is a subfamily of the sub-σ-field generated by P, then we will also write $Q\subset P$. If P and Q are partitions of X, $P\vee Q$, $\bigvee_0^n T^iP, \cdots$ will denote the ordered partitions with a canonical ordering.

Now we define several entropies as usual. Given countable partitions P and Q, we have the following

$$H(P)=-\sum_j m(p_j)\log m(p_j),$$

$$H(P|Q)=-\sum_{i,j} m(q_i)m(p_j|q_i)\log m(p_j|q_i),$$

$$h(P, T)=\lim_{n\to\infty}\frac{1}{n}H(P\vee\cdots\vee T^{n-1}P)$$

$$h(T)=\sup_P h(P, T)$$

where the supremum is taken over all finite partitions of X. Note that

$H(P\mid\bigvee_{-n}^{-1}T^iP)$ decreases to $h(P, T)$.

We will remark the following. Let $\mathscr{P}$ and $\mathscr{P}_k$ denote the collections of all countable partitions and partitions with at most k atoms respectively. Firstly, $(\mathscr{P}, d)$ and $(\mathscr{P}, D)$ are complete separable metric spaces. Next, the entropy $H(P)$ is continuous in $(\mathscr{P}_k, d)$ and $(\mathscr{P}_k, D)$. At last, the entropy $h(\cdot, T)$ is continuous in $(\mathscr{P}_k, D)$, and it is only lower semi-continuous in $(\mathscr{P}, D)$. These properties will be used in the proof of the theorem.

Given countable partitions P and Q, we say P and Q are ε-independent and denote $P\overset{\varepsilon}{\perp}Q$ if

$$\sum_{i,j}|m(p_i\cap q_j)-m(p_i)m(q_j)|<\varepsilon.$$

Although this definition is different from the one in [4], it is easy to see that they are equivalent, and our definition is convenient in the point that if $P\overset{\varepsilon}{\perp}Q$ and $\bar{P}\subset P$, $\bar{Q}\subset Q$ then $\bar{P}\overset{\varepsilon}{\perp}\bar{Q}$.

Definition 1. A countable partition P is called *weak Bernoulli* for T if for each $\varepsilon>0$ there exists $K=K(\varepsilon, P)$ such that $\bigvee_{-n}^{0}T^iP\overset{\varepsilon}{\perp}\bigvee_{K}^{K+n}T^iP$ for every $n\geqq 0$. A measurable partition P is called *generalized weak Bernoulli* for T if there exists an increasing sequence $\{P_n\}$ of finite partitions such that $\bigvee_{1}^{\infty}P_n=P$ and each P_n is weak Bernoulli for T. In this case the invariant sub-σ-fields generated by $\bigvee_{-\infty<i<\infty}T^iP_n$ $(n\geqq 1)$ are called *approximating σ-fields*. The transformation T is called (*generalized*) *weak Bernoulli* if T has a (generalized) weak Bernoulli generator.

§3. Countable Case

In this section we will prove the following

Proposition 1. *Two weak Bernoulli transformations with the same entropy whose weak Bernoulli generators are at most countable and have finite entropies are isomorphic.*

For the proof of this proposition, we prepare some lemmas in which we are concerned only with countable partitions.

Lemma 1. *Let* $\varepsilon>0$ *and let* Q_i, Q'_i, $0\leqq i\leqq n$, P_0, P'_0 *be partitions such that*

(1.1) $$Q_0 \overset{\varepsilon}{\perp} \bigvee_1^n Q_i,$$

(1.2) $$Q'_0 \overset{\varepsilon}{\perp} \bigvee_1^n Q'_i,$$

(1.3) $$d(\bigvee_1^n Q_i, \bigvee_1^n Q'_i)<\varepsilon,$$

(1.4) $$d(Q_0)=d(P_0),\ d(Q'_0)=d(P'_0).$$

Then there exist partitions P_i, P'_i, $1\leqq i\leqq n$, *such that*

(1.5) $$d(\bigvee_0^n Q_i)=d(\bigvee_0^n P_i), \quad d(\bigvee_0^n Q'_i)=d(\bigvee_0^n P'_i),$$

(1.6) $$\sum_{i=1}^n D(P_i, P'_i)<3n\varepsilon.$$

Noting the difference of the definitions of ε-independence, this lemma is the same as Lemma 3 in [2], so we may omit the proof.

Similarly the following lemma is the same as Lemma 1 in [7], so we also omit the proof.

Lemma 2. *For each* $\varepsilon>0$ *there exists* $\zeta=\zeta(\varepsilon)>0$ *such that if the partitions* P *and* Q *satisfy*

(2.1) $$H(P)-H(P\,|\,Q)<\zeta,$$

then $P \overset{\varepsilon}{\perp} Q$.

Using this lemma we can prove the following

Lemma 3. *Let* $\varepsilon>0$ *and* $\zeta=\zeta(\varepsilon)>0$ *as in Lemma 2. If* P_0, P, Q *and* R *are partitions such that*

(3.1) $$P_0\subset P,$$

(3.2) $$P_0 \overset{\varepsilon}{\perp} Q,$$

(3.3) $$H(P\,|\,Q)-H(P\,|\,Q\vee R)<\varepsilon\zeta,$$

then

(3.4) $$P_0 \overset{4\varepsilon}{\perp} Q\vee R.$$

Proof. Note that

$$H(P|Q)-H(P|Q\vee R)=\sum_{q\in Q} m(q)\,\{H(P/q)-H(P/q\,|\,R/q)\}.$$

Put $Q_1=\{q\in Q;\, H(P/q)-H(P/q\,|\,R/q)<\zeta\}$. Because of (3.3), we have $m(\cup Q_1)>1-\varepsilon$. Lemma 2 implies $P/q \overset{\varepsilon}{\perp} R/q$ for all $q\in Q_1$. Therefore

$$\begin{aligned}
&\sum_{p\in P_0,\, q\in Q,\, r\in R} |m(p\cap q\cap r)-m(p)m(q\cap r)| \\
&\leqq \sum_{p,q,r} m(q)\{|m(p\cap r\,|\,q)-m(p\,|\,q)m(r\,|\,q)| \\
&\qquad\qquad + m(r\,|\,q)\,|m(p\,|\,q)-m(p)|\} \\
&= \sum_{q\in Q_1} m(q) \sum_{p,r} |m(p\cap r\,|\,q)-m(p\,|\,q)m(r\,|\,q)| \\
&\qquad + \sum_{q\notin Q_1} m(q) \sum_{p,r} |m(p\cap r\,|\,q)-m(p\,|\,q)m(r\,|\,q)| \\
&\qquad\qquad + \sum_{p,q} |m(p\cap q)-m(p)m(q)| \sum_{r} m(r\,|\,q) \\
&< 4\varepsilon.
\end{aligned}$$

As we remarked in §2, the entropy of countable partitions is not continuous in d-metric, we have to add one more condition on the entropy to Lemmas 8, 9 and so on in [2].

Lemma 4. *Let P be weak Bernoulli for T with $H(P)<\infty$, and let $\varepsilon>0$. Let $K=K(\varepsilon/4, P)$ be as in Definition 1. Let H be a positive integer. Then there exist a positive integer n_* and $\eta>0$ such that if T' and P' satisfy*

(4.1) $$d\Big(\bigvee_{-n_*}^{K+H} T^i P,\ \bigvee_{-n_*}^{K+H} T'^i P'\Big)<\eta,$$

$$(4.2)\qquad |h(P,\ T)-h(P',\ T')|<\eta,$$

$$(4.3)\qquad |H(\bigvee_{-n_*}^{j} T^iP)-H(\bigvee_{-n_*}^{j} T'^iP')|<\eta,\quad -n_*\leqq j\leqq K+H,$$

then for each $m\geqq 0$,

$$(4.4)\qquad \bigvee_{K}^{K+H} T'^iP' \overset{\varepsilon}{\perp} \bigvee_{-m}^{0} T'^iP'.$$

Proof. Let $P_1=\bigvee_{1}^{K+H} T^iP$, $P_2=\bigvee_{K}^{K+H} T^iP$, $Q=\bigvee_{-n}^{0} T^iP$, and $R=\bigvee_{-n-m}^{-n-1} T^iP$. We define P_1', P_2', Q', and R' in the same way for T' and P'. Since $H(P_1|Q)$ decreases monotonically to $H(P_1|\bigvee_{-\infty}^{0} T^iP)=(K+H)h(P,\ T)$ as $n\to\infty$, we can take $n_*=n$ so large that

(1) $(K+H)h(P,\ T)+\varepsilon\zeta/4-H(P_1|Q)=\alpha>0$,

where $\zeta=\zeta(\varepsilon/4)$ is as in Lemma 2. The choice of K implies

(2) $P_2 \overset{\varepsilon/4}{\perp} Q$.

Now take $\eta>0$ so small that (2) and (4.1) imply

(3) $P_2' \overset{\varepsilon/4}{\perp} Q'$,

(4.2) implies

(4) $|(K+H)h(P,\ T)-(K+H)h(P',\ T')|<\alpha/2$,

and (4.3) implies

(5) $|H(P_1|Q)-H(P_1'|Q')|<\alpha/2$.

Then (1), (4) and (5) imply

(6) $H(P_1'|Q')<H(P_1'|Q'\vee R')+\varepsilon\zeta/4$.

Therefore (3) and (6) imply (4.4) by Lemma 3.

Lemma 5. *Let* P *be weak Bernoulli for* T *with* $H(P)<\infty$ *and let* $\varepsilon>0$. *There exist positive integers* n_1, n_2, *and* $\eta>0$ *such that if* T' *and* P' *satisfy*

(5.1) $$d(\bigvee_0^{n_1} T^i P, \bigvee_0^{n_1} T'^i P') < \eta,$$

(5.2) $$|h(P, T) - h(P', T')| < \eta,$$

(5.3) $$|H(\bigvee_0^{j} T^i P) - H(\bigvee_0^{j} T'^i P')| < \eta, \quad 0 \leqq j \leqq n_1,$$

then there exist partitions P_i *and* P'_i *of* X *such that*

(5.4) $$d(\bigvee_0^{n} P_i) = d(\bigvee_0^{n} T^i P), \qquad n = 0, 1, 2, \cdots,$$

(5.5) $$d(\bigvee_0^{n} P'_i) = d(\bigvee_0^{n} T'^i P'), \quad n = 0, 1, 2, \cdots,$$

(5.6) $$\sum_{i=0}^{n} D(P_i, P'_i) < \varepsilon n, \qquad n \geqq n_2.$$

Proof. Let $K = K(\varepsilon/50, P)$ be as in Definition 1. Choose H so that $K/K+H < \varepsilon/8$. Apply Lemma 4 to H, K and $\varepsilon/12$, then we have n_* and $\eta < \varepsilon/12$. Let $n_1 = n_* + K + H$. The choice of K and Lemma 4 implies

(1) $$\bigvee_{K+n}^{K+H+n} T^i P \overset{\varepsilon/12}{\perp} \bigvee_0^{n} T^i P, \quad \bigvee_{K+n}^{K+H+n} T'^i P' \overset{\varepsilon/12}{\perp} \bigvee_0^{n} T'^i P'$$

for all $n \geqq 0$.

We will define P_i, P'_i inductively. Let P_0 and P'_0 be partitions of X such that

(2) $$d(P_0) = d(P), \ d(P'_0) = d(P').$$

Let $Q_0 = P$, $Q'_0 = P'$, and

(3) $$Q_i = T^{i+K-1} P, \ Q'_i = T'^{i+K-1} P', \quad 1 \leqq i \leqq K+H.$$

Because of (1) (in case of $n = 0$), (2) and (5.1) we can apply Lemma 1 to get P_i, P'_i, $K \leqq i \leqq K+H$ such that

(4) $$d(P_0 \vee \bigvee_K^{K+H} P_i) = d(P \vee \bigvee_K^{K+H} T^i P),$$

(5) $$d(P'_0 \vee \bigvee_K^{K+H} P'_i) = d(P' \vee \bigvee_K^{K+H} T'^i P'),$$

(6) $$\sum_{i=K}^{K+H} D(P_i, P'_i) < 3(H+1)\varepsilon/12 = (H+1)\varepsilon/4.$$

We may now define P_i and P'_i, $1 \leqq i \leqq K-1$, so that (5.4) and (5.5) hold for $n \leqq K+H$. Now (6) implies

$$\sum_{i=0}^{K+H} D(P_i, P_i') < \sum_{i=0}^{K-1} D(P_i, P_i') + (H+1)\varepsilon/4$$
$$\leqq 2K + (H+1)\varepsilon/4 < (K+H)\varepsilon/2.$$

Let $n_2 = K+H$.

Assume we have already defined P_i and P_i', $0 \leqq i \leqq j(K+H)$, so that (5.4–6) hold. Assume also

$$(7) \quad \sum_{t(K+H)-H}^{t(K+H)} D(P_i, P_i') < (H+1)\varepsilon/4, \quad 1 \leqq t \leqq j.$$

By (5.1) we have

$$(8) \quad d\Big(\bigvee_{n+K}^{n+K+H} T^i P, \bigvee_{n+K}^{n+K+H} T'^i P'\Big) < \varepsilon/12.$$

Let $n_0 = j(K+H)$. Because of (1), (5) and (5.4–6) (for $n = j(K+H)$) we can apply Lemma 1 (P_0 and P_0' are replaced by $\bigvee_0^{n_0} P_i$ and $\bigvee_0^{n_0} P_i'$ respectively) to get P_i, P_i', $n_0 + K \leqq i \leqq n_0 + K + H$, such that

$$(9) \quad d\Big(\bigvee_0^{n_0} P_i \vee \bigvee_{n_0+K}^{n_0+K+H} P_i\Big) = d\Big(\bigvee_0^{n_0} T^i P \vee \bigvee_{n_0+K}^{n_0+K+H} T^i P\Big),$$

$$(10) \quad d\Big(\bigvee_0^{n_0} P_i' \vee \bigvee_{n_0+K}^{n_0+K+H} P_i'\Big) = d\Big(\bigvee_0^{n_0} T'^i P' \vee \bigvee_{n_0+K}^{n_0+K+H} T'^i P'\Big),$$

$$(11) \quad \sum_{i=n_0+K}^{n_0+K+H} D(P_i, P_i') < (H+1)\varepsilon/4.$$

We define P_i, P_i', $n_0 < i < n_0 + K$, so that (5.4) and (5.5) hold for $n = (j+1)(K+H)$. For $j(K+H) < n \leqq (j+1)(K+H)$, (7) and (11) imply

$$\sum_{i=0}^{n} D(P_i, P_i') < 2Kj + j(H+1)\varepsilon/4 + 2K + (H+1)\varepsilon/4$$
$$< n\varepsilon.$$

We will now prepare the following two lemmas which will be used to prove the approximation lemma (Lemma 8). Since they are exactly the same as Lemma 10 and 11 in [2], we will omit their proofs.

We consider a transformation T_1 and a finite partition $R = \{r_i;\ 1 \leqq$

$i \leqq k\}$. Let $Q=\bigvee_0^u T_1^i R$. Note that an atom $q \in Q$ has the form

$$q=\bigcap_0^u T_1^i r_{s_i}, \quad 1 \leqq s_i \leqq k, \quad 0 \leqq i \leqq u.$$

Let $l=\{l_0, l_1, \cdots, l_{n-1}\}$ be a sequence of length $n(>u)$, where $1 \leqq l_i \leqq k$, $0 \leqq i<n$. Let $N(l, q)$ be the number of times that $(s_u, s_{u-1}, \cdots, s_0)$ appears as a consecutive subsequence in l; hence $N(l, q) \leqq n-u$.

Definition 2. Let $\varepsilon>0$. l is called an *ε-sequence* for Q if

$$|N(l, q)/n-m(q)|<\varepsilon,$$

for all $q \in Q$.

Lemma 6. *Let R, T_1, and Q be as above and $\varepsilon>0$. Let T be a transformation and let B_j, $1 \leqq j \leqq J$, be measurable sets such that $T^i B_j$, $0 \leqq i<n$, $1 \leqq j \leqq J$, are disjoint. Assume $u/n<\varepsilon/4$ and $X_1=\bigcup_{i=0}^{n-1} \bigcup_{j=1}^{J} T^i B_j$ satisfies $m(X_1)>1-\varepsilon/4$. Let $l_j=\{l_{j,i}, 0 \leqq i<n\}$, $1 \leqq j \leqq J$, be $\varepsilon/4k^{u+1}$-sequences for Q. Let $P=\{p_1, \cdots, p_k\}$ be a partition such that $p_t/X_1=\bigcup_{l_{j\,i}=t} T^i B_j$ and p_t/X_1^c is arbitrary for $1 \leqq t \leqq k$. Then*

$$d\Big(\bigvee_0^u T^i P,\ Q\Big)<\varepsilon. \tag{6.1}$$

If the sequences l_j, $1 \leqq j \leqq J$, are distinct, then

$$\bigvee_{-n}^{n} T^i(P \vee\{B, B^c\}) \supset \{T^i B_j: 0 \leqq i<n,\ 1 \leqq j \leqq J\}, \tag{6.2}$$

where $B=\bigcup_{j=1}^{J} B_j$.

Lemma 7. *Let T_1 be ergodic. Let $R=\{r_1, \cdots, r_k\}$ be a finite partition and $Q=\bigvee_0^u T_1^i R$. Let $L_n=\bigvee_0^{n-1} T_1^{-i} R$ and write $l=\bigcap_0^{n-1} T_1^{-i} r_{l_i} \in L_n$ as $l=(l_0, l_1, \cdots, l_{n-1})$. Let $a>0$ and $b>0$ be arbitrary. Then there exists n_0 such that for all $n \geqq n_0$ we have a subfamily $L_n' \subset L_n$ such that $m(\cup L_n')>1-a$, and $l \in L_n'$ implies l is a b-sequence for Q.*

Now we can prove the following approximation lemma which is different from the Lemma 12 in [2]. Precisely speaking, the condition that T is "Bernoulli" in [2] is only used in changing the partition P to satisfy the condition $0<h(T_1)-h(P, T)<\delta$ in the case $h(T_1)=h(P, T)$. Noting this point, we can state this lemma in the following way.

Lemma 8. *Let R be weak Bernoulli for T_1 with $h(R, T_1)=h(T_1)$ and $H(R)<\infty$. Let n_1, n_2, and η be as in Lemma 5 for R, T_1, and $\varepsilon^2/3$. Let T be an ergodic transformation with $h(T)=h(T_1)$. Let P' be a partition such that*

$$d(\bigvee_0^{n_1} T_1^i R, \bigvee_0^{n_1} T^i P')<\eta, \tag{8.1}$$

$$h(T_1)-h(P', T)<\eta, \tag{8.2}$$

$$|H(\bigvee_0^{j} T_1^i R)-H(\bigvee_0^{j} T^i P')|<\eta, \quad 0\leqq j\leqq n_1. \tag{8.3}$$

Then given $\delta>0$ and a positive integer u, there exists a (finite) partition P such that

$$d(\bigvee_0^{u} T_1^i R, \bigvee_0^{u} T^i P)<\delta, \tag{8.4}$$

$$h(T_1)-h(P, T)<\delta, \tag{8.5}$$

$$|H(\bigvee_0^{j} T_1^i R)-H(\bigvee_0^{j} T^i P)|<\delta, \quad 0\leqq j\leqq u, \tag{8.6}$$

$$D(P', P)<6\varepsilon. \tag{8.7}$$

Proof. We first assume $h(T_1)=h(R, T_1)>h(P', T)$. Let $R=\{r_1, r_2,\cdots\}$ and define the k-section $R(k)=\{r_1, r_2,\cdots, r_{k-1}, \bigcup_{j\geqq k} r_j\}$ of R. $P'(k)$ is also difined in the same way. Then we can take k so large that

(1) $d(\bigvee_0^{u} T_1^i R, \bigvee_0^{u} T_1^i R(k))<\delta/2$,

(2) $H(\bigvee_0^{j} T_1^i R)-H(\bigvee_0^{j} T_1^i R(k))<\delta/2, \quad 0\leqq j\leqq u$,

(3) $h(T_1)-h(R(k), T_1)<\delta/4$,

(4) $h(P', T)<h(R(k), T_1)$,

(5) $D(P', P'(k))<\varepsilon$.

Denote $\bar{R}=R(k)$, and also $\bar{P}=P'(k)$. Take $\delta_1<\delta/2$ such that if Q_i, $0\leqq i\leqq u$ are partitions with k atoms and if $d(\bigvee_0^u T_1^i\bar{R}, \bigvee_0^u Q_i)<\delta_1$ then $|H(\bigvee_0^j T_1^i\bar{R})-H(\bigvee_0^j Q_i)|<\delta/2$, $0\leqq j\leqq u$. Because of $h(T_1)=h(T)$, there exists a finite refinement W of $\bar{P}$ such that

(6) $0<h(\bar{R}, T_1)-h(W, T)=\beta<\delta/4$.

Choose $\gamma<\min(\delta_{1/2}, \varepsilon/3, 5/2)$ such that $D(W', W)<\gamma$ implies

$$h(W', T)>h(W, T)-\delta/4.$$

We will now choose n so large that the following conditions (7)–(10) are satisfied.

(7) Applying Lemma 7 to T_1, $\bar{R}$, and $Q=\bigvee_0^n T_1^i\bar{R}$, we have a subfamily $\mathscr{L}'\subset L_n=\bigvee_0^{n-1} T_1^{-i}\bar{R}$ such that $m(\cup\mathscr{L}')>1-\gamma/10$, and $l\in\mathscr{L}'$ implies l is a $\delta_{1/4k^{u+1}}$-sequence for Q.

(8) Applying the Shannon-McMillan-Breiman theorem to T_1 and $\bar{R}$, we have a subfamily $\mathscr{L}''\subset L_n$ such that $m(\cup\mathscr{L}'')>1-\gamma/10$, and $l\in\mathscr{L}''$ implies $m(l)$ is between $2^{-[h(\bar{R},T_1)\pm\beta/3]n}$.

(9) Applying the Shannon-McMillan-Breiman theorem to T and W, we have a subfamily $\mathscr{W}_n\subset\bigvee_0^{n-1} T^{-i}W$ such that $m(\cup\mathscr{W}_n)>1-\gamma/10$, and $w\in\mathscr{W}_n$ implies $m(w)$ is between $2^{-[h(W,T)\pm\beta/3]n}$.

(10) $n>\max(n_1, n_2)$ such that $u/n<\delta_{1/4}$, $n\beta>3$, and $m(A)\leqq 1/n$ implies $H(\{A, A,^c\})<\beta$.

Apply Lemma 5 to obtain partitions R_i', P_i', $0\leqq i\leqq n-1$, such that

$$d(\bigvee_0^{n-1} R_i')=d(\bigvee_0^{n-1} T_1^{-i}R),\ d(\bigvee_0^{n-1} P_i')=d(\bigvee_0^{n-1} T^{-i}P'),$$
$$\sum_{i=0}^{n-1} D(R_i', P_i')<n\varepsilon^2/3.$$

Let φ be an automorphism of X such that

$$\varphi(\bigvee_0^{n-1} P_i')=\bigvee_0^{n-1} T^{-i}P'.$$

Putting $R_i''=\varphi(R_i')$, $0\leqq i\leqq n-1$, we have

$$d(\bigvee_0^{n-1} R_i'')=d(\bigvee_0^{n-1} T_1^{-i}R),\quad \sum_{i=0}^{n-1} D(R_i'',\ T^{-i}P')<n\varepsilon^2/3.$$

Hence there exist partitions $R_i=\{r_{i,1,\dots,}\ r_{i,k}\}$, $0\leqq i\leqq n-1$, such that

(11) $d(\bigvee_0^{n-1} R_i)=d(\bigvee_0^{n-1} T_1^{-i}\bar{R})$,

(12) $\sum_{i=0}^{n-1} D(R_i,\ T^{-i}\bar{P})<n\varepsilon^2/3$.

Let $L_n^*=\bigvee_0^{n-1} R_i$ and $\bar{P}_n=\bigvee_0^{n-1} T^{-i}\bar{P}$. A sequence $(s_0, s_1,\cdots, s_{n-1})$ is called the L_n^*-name of x if $x\in\bigcap_0^{n-1} r_{i,s_i}$, and we denote this by $s^*(x)=(s_0^*(x),\cdots, s_{n-1}^*(x))$. Similarly $(s_0, s_1,\cdots, s_{n-1})$ is called the $\bar{P}_n$-name of x if $x\in\bigcap_0^{n-1} T^{-i}\bar{p}_{s_i}$ where $\bar{P}=\{\bar{p}_1,\cdots,\ \bar{p}_k\}$, and we denote this by $\bar{s}(x)=(\bar{s}_0(x),\cdots,\ \bar{s}_{n-1}(x))$. Note that for each atom $l\in L_n^*$ $s^*(x)$ is the same for all $x\in l$. Hence we can talk about the L_n^*-name $s^*(l)$ of $l\in L_n^*$. Analogously we denote the $\bar{P}_n$-name of $w\in\bigvee_0^{n-1} T^{-i}W$ by $\bar{s}(w)$. Define

$$\rho(s^*(x),\ \bar{s}(x))=\text{ the number of } \{i\,;\ s_i^*(x)\neq\bar{s}_i(x)\},$$

$$G=\{x\,;\ \rho(s^*(x),\ \bar{s}(x))>n\varepsilon\}.$$

Then

$$n\varepsilon m(G)\leqq E\{\rho(s^*(x),\ \bar{s}(x))\}=\sum_{i=0}^{n-1} m\{x\,;\ s_i^*(x)\neq\bar{s}_i(x)\}$$
$$=\frac{1}{2}\sum_{i=0}^{n-1} D(R_i,\ T^{-i}\bar{P})<n\varepsilon^2/6$$

and so

(13) $m(G)<\varepsilon/6$.

Let $\bar{\mathscr{L}}^*=\{l\in L_n^*;\ s^*(l)$ is a $\delta_{1/4k^{u+1}}$-sequence for $Q\}$, $\tilde{\mathscr{L}}^*=\{l\in L_n^*;\ m(l)$ is between $2^{-[h(\bar{R},T_1)\pm\beta/3]n}\}$ and $\mathscr{L}=\bar{\mathscr{L}}^*\cap\tilde{\mathscr{L}}^*$. Then (7), (8) and (11) imply

(14) $m(\cup\mathscr{L})>1-\gamma/5$.

Let $\mathscr{C}=\{w\in\mathscr{W}_n;\ m(w\cap(\cup\mathscr{L})\cap G^c)\geqq m(w)/2\}$, then

$$m(X\setminus\cup\mathscr{C})=\sum_{w\notin\mathscr{C}}m(w)$$

$$=\sum_{w\notin\mathscr{C}}\{m(w\cap(\cup\mathscr{L})\cap G^c)+m(w\cap((\cup\mathscr{L})\cap G^c)^c)\}$$

$$<\frac{1}{2}\sum_{w\notin\mathscr{C}}m(w)+m(X\setminus\cup\mathscr{L})+m(G)$$

and so

$$(15)\quad m(X\setminus\cup\mathscr{C})<2\{m(X\setminus\cup\mathscr{L})+m(G)\}<\varepsilon/2.$$

Note that (6) and (10) imply

$$\max_{l\in\mathscr{L}}m(l)\leqq 2^{-[h(\bar{R},T_1)-\beta/3]n}$$

$$<\frac{1}{2}\cdot 2^{-[h(W,T)+\beta/3]n}\leqq\frac{1}{2}\min_{w\in\mathscr{W}_n}m(w).$$

To each $w\in\mathscr{C}$ we assign $l\in\mathscr{L}$ such that $l\cap w\cap G^c\neq\emptyset$. Then any t elements in $\mathscr{C}$ correspond to at least t elements in $\mathscr{L}$. Hence, applying the marriage lemma, each $w\in\mathscr{C}$ can be assigned an $l(w)\in\mathscr{L}$ such that $\rho(s^*(l(w)),\ \bar{s}(w))\leqq n\varepsilon$ and the mapping l is one-to-one. Since

the number of elements of $\mathscr{W}_n$

$$\leqq 2^{[h(W,T)+\beta/3]n}$$

$$<(1-\gamma/5)2^{[h(\bar{R},T_1)-\beta/3]n}$$

$<$ the number of elements of $\mathscr{L}$,

we can extend the map l from $\mathscr{W}_n$ into $\mathscr{L}$ in the one-to-one way. Thus we obtain

(16) for any $w\in\mathscr{W}_n$, $s^*(l(w))$ is $\delta_{1/4k^{u+1}}$-sequence for Q,

(17) if $w\in\mathscr{C}$ then $\rho(s^*(l(w)),\ \bar{s}(w))\leqq n\varepsilon$.

Rohlin's theorem asserts that there exists a set F' such that T^iF', $0\leqq i\leqq n-1$, are disjoint and

(18) $m(\bigcup_0^{n-1} T^i F')>1-\gamma/10.$

Becaue of (15) we have

$$\sum_{i=0}^{n-1} m(T^i F' \cap (X \setminus \cup \mathscr{C})) < \varepsilon/2.$$

Therefore for more than half of indices i,

(19) $m(T^i F' \cap (X \setminus \cup \mathscr{C})) < \varepsilon/n.$

Analogously using (9) we have

(20) $m(T^i F' \cap (X \setminus \cup \mathscr{W}_n)) < \gamma/5n$

for more than half of indices i. Hence there exists $i=i_0$ such that (19) and (20) hold together. Put $F=T^{i_0}F'$. Then T^iF, $0 \leqq i \leqq n-1$, are disjoint and

(21) $m(\bigcup_0^{n-1} T^i(F \cap \cup \mathscr{C})) > 1-\gamma/10-\varepsilon,$

(22) $m(\bigcup_0^{n-1} T^i(F \cap \cup \mathscr{W}_n)) > 1-3\gamma/10.$

We will now define the partition $P=\{p_1, p_2, \cdots, p_k\}$ as follows. Let $X_1 = \bigcup_0^{n-1} T^i(F \cap \cup \mathscr{W}_n)$ and $A = F \cap \cup \mathscr{W}_n$, then

(23) $m(A) \leqq 1/n, \ m(X_1) > 1-3\gamma/10 > 1-\delta_{1/4}.$

Define P on X_1 as

$$p_j/X_1 = \cup \{T^i w;\ w \in \mathscr{W}_n/F,\ s_1^*(l(w))=j\}, \quad 1 \leqq j \leqq k.$$

Obviously $P/X_1 \subset W^* = \bigvee_{-n}^{n} T^i(W \vee \{F, F^c\})$. Extend P to X_1^c so that $P \subset W^*$. Then Lemma 6 implies

(24) $d(\bigvee_0^u T^i P, \bigvee_0^u T_1^i \bar{R}) < \delta_1,$

(25) $\{T^i w;\ w \in \mathscr{W}_n/F, \quad 0 \leqq i \leqq n-1\} \subset P^* = \bigvee_{-n}^{n} T^i(P \vee \{A, A^c\})$

because $s^*(l(w))$, $w \in \mathscr{W}_n/F$ are distinct.

We will now check the conclusions (8.4–7). By (1) and (24)

$$d(\bigvee_0^u T_1^i R, \bigvee_0^u T^i P) < \delta.$$

The choice of δ_1 and (24) imply

$$|H(\bigvee_0^j T_1^i \bar{R}) - H(\bigvee_0^j T^i P)| < \delta/2, \quad 0 \leqq j \leqq u,$$

which implies (8.6) by (2). Because of (25) and (23), there exists a partition $W' \subset P^*$ such that

$$D(W, W') = D(W, W' | X_1^c)\, m(X_1^c) < 6\gamma/10$$

which implies

$$h(W', T) > h(W, T) - \delta/4$$

by the choice of γ. Hence using (10), (23), (6) and (3) in turn

$$\begin{aligned} h(P, T) &\geqq h(P \vee \{A, A^c\}, T) - H(\{A, A^c\}) \\ &> h(P^*, T) - \delta/4 \\ &\geqq h(W', T) - \delta/4 \\ &> h(W, T) - \delta/2 \\ &> h(\bar{R}, T_1) - 3\delta/4 > h(T_1) - \delta. \end{aligned}$$

On the other hand (10) and (6) imply

$$\begin{aligned} h(P, T) &\leqq h(W^*, T) \\ &\leqq h(W, T) + H(\{F, F^c\}) \\ &< h(W, T) + \beta = h(\bar{R}, T_1) \leqq h(T_1). \end{aligned}$$

Thus (8.5) holds. Next, let $X_2 = \bigcup_0^{n-1} T^i(F \cap \cup \mathscr{C})$; hence $m(X_2^c) < \varepsilon + \gamma/10$. Since $p_j \cap X_2 = \cup \{T^i w;\ w \in \mathscr{C}/F,\ s_i^*(l(w)) = j\}$ and $\bar{p}_j \cap X_2 = \cup \{T^i w;\ w \in \mathscr{C}/F,\ \bar{s}_i(w) = j\}$, (17) implies

$$D(P, \bar{P} | X_2) m(X_2) = \sum_{j=1}^k m((p_j \cap X_2) \triangle (\bar{p}_j \cap X_2))$$

$$=2\sum_{w\in\mathscr{C}\cap F} m(w)\rho(s^*(l(w)),\ \bar{s}(w))$$

$$\leqq 2n\varepsilon\, m(F)<2\varepsilon.$$

Therefore

$$D(P,\ \bar{P})=D(P,\ \bar{P}\,|\,X_2)m(X_2)+D(P,\ \bar{P}\,|\,X_2^c)m(X_2^c)$$

$$<2\varepsilon+2\varepsilon+\gamma/5<5\varepsilon.$$

Thus (5) implies (8.7).

Finally, although we have assumed $h(T_1)>h(P',\ T)$ at the beginning of this proof, we can deduce to this case even if $h(T_1)=h(P',\ T)$. Indeed, in this case consider a product transformation $S_1=T_1\times T'$ on the product space $X\times Y$ where T' is a Bernoulli shift with generator $R'=\{r_1',\ r_2'\}$ on Y. Consider partitions $\tilde{R}=R\times R'$ on $X\times Y$, i.e. $\tilde{R}=(R\times\nu_Y)\vee(\nu_X\times R')$, and $\hat{P}=P'\vee\nu_X$ and $\hat{R}=R\vee\nu_X$ on X, where $\nu_X=\{X\}$ and $\nu_Y=\{Y\}$ are trivial partitions. Then by choosing the measure of r_2' so small, we have

(a) $d(\bigvee_0^v T_1^i\hat{R},\ \bigvee_0^v S_1^i\tilde{R})<\eta'$,

(b) $h(T')<\eta'$,

(c) $|H(\bigvee_0^j T_1^i\hat{R})-H(\bigvee_0^j S_1^i\tilde{R})|<\eta',\quad 0\leqq j\leqq v$,

where $v=\max(u,\ n_1)$ and $\eta'<\min(\delta/2,\ \eta)$. By choosing η' so small, we have

(8.1)′ $$d(\bigvee_0^{n_1} S_1^i\tilde{R},\ \bigvee_0^{n_1} T^i\hat{P})<\eta,$$

(8.2)′ $$0<h(S_1)-h(\hat{P},\ T)=h(T')<\eta,$$

(8.3)′ $$|H(\bigvee_0^j S_1^i\tilde{R})-H(\bigvee_0^j T^i\hat{P})|<\eta,\quad 0\leqq j\leqq n_1,$$

and $h(S_1)=h(\tilde{R},\ S_1)=h(T_1)+h(T')=h(T)+h(T')$.

Now we must remark the following three facts. Firstly the proof of this lemma in the case $0<h(T_1)-h(P',\ T)$ (and $h(T_1)=h(T)$) is also

available even if the space on which T_1 acts is different from the space of T. Secondly the conditions $h(T_1)=h(T)$ and $0<h(T_1)-h(P', T)$ are only used in taking a partition W satisfying (6). Lastly n_1, n_2, and η which are defined in Lemma 5 for R, T_1 and $\varepsilon^2/3$ remain unaltered even if we replace these by $\tilde{R}$, S_1 and $\varepsilon^2/3$. Noting these facts, we can apply the previous argument to $\tilde{R}$, S_1, n_1, n_2, η, T and $\hat{P}$, because we can take a partition W which satisfies a relation similar to (6) by choosing η' small enough even when $h(S_1)>h(T)=h(P', T)$. Hence for every $\delta/2$ we have a finite partition $\tilde{P}$ such that

$$d(\bigvee_0^u S_1^i\tilde{R}, \bigvee_0^u T^i\tilde{P})<\delta/2, \tag{8.4$'$}$$

$$0<h(S_1)-h(\tilde{P}, T)<\delta/2, \tag{8.5$'$}$$

$$|H(\bigvee_0^j S_1^i\tilde{R})-H(\bigvee_0^j T^i\tilde{P})|<\delta/2, \quad 0\leqq j\leqq u, \tag{8.6$'$}$$

$$D(\hat{P}, \tilde{P})<6\varepsilon. \tag{8.7$'$}$$

By changing the order of $\tilde{P}$ we have a partition P with the same atoms as $\tilde{P}$ such that (8.4–7) are valid.

Remark. In Lemma 8, without assuming (8.1–3) there exists a partition P which satisfies only (8.4–6).

In the following, we denote $(P, T)\sim(R, T_1)$ if $d(\bigvee_0^u T_1^iR)=d(\bigvee_0^u T^iP)$ for all $u\geqq 0$.

Lemma 9. *Let R be weak Bernoulli for T_1 with $h(R, T_1)=h(T_1)$ and $H(R)<\infty$. Given $\varepsilon>0$, there exist $\delta>0$ and a positive integer u such that if T is ergodic with $h(T)=h(T_1)$ and P' satisfies*

$$d(\bigvee_0^u T_1^iR, \bigvee_0^u T^iP')<\delta, \tag{9.1}$$

$$h(T_1)-h(P', T)<\delta, \tag{9.2}$$

$$|H(\bigvee_0^j T_1^iR)-H(\bigvee_0^j T^iP')|<\delta, \quad 0\leqq j\leqq u, \tag{9.3}$$

then there exists P such that

$$(9.4) \qquad (P, T) \sim (R, T_1),$$

$$(9.5) \qquad D(P', P) < \varepsilon.$$

Proof. Let $n_1(\varepsilon)$ and $\eta(\varepsilon)$ be as in Lemma 5 for R, T_1 and $\varepsilon^2/3$. Let $\eta'(\varepsilon) = \min(\eta(\varepsilon), \varepsilon)$ and $\varepsilon' = \varepsilon/12$. Let $\delta = \eta'(\varepsilon')$ and $u = n_1(\varepsilon')$. Applying Lemma 8 for $\eta'(\varepsilon'/2)$ and $u_1 > \max(u, n_1(\varepsilon'/2))$ we get P_1 such that

$$d(\bigvee_0^{u_1} T_1^i R, \bigvee_0^{u_1} T^i P_1) < \eta'(\varepsilon'/2),$$

$$h(T_1) - h(P_1, T) < \eta'(\varepsilon'/2),$$

$$|H(\bigvee_0^{j} T_1^i R) - H(\bigvee_0^{j} T^i P_1)| < \eta'(\varepsilon'/2), \quad 0 \leqq j \leqq u_1,$$

$$D(P', P_1) < 6\varepsilon'.$$

Assume we have P_k, $1 \leqq k \leqq n$, such that

(1) $d(\bigvee_0^{u_k} T_1^i R, \bigvee_0^{u_k} T^i P_k) < \eta'(\varepsilon'/2^k)$,

(2) $h(T_1) - h(P_k, T) < \eta'(\varepsilon'/2^k)$,

(3) $|H(\bigvee_0^{j} T_1^i R) - H(\bigvee_0^{j} T^i P_k)| < \eta'(\varepsilon'/2^k), \quad 0 \leqq j \leqq u_k$,

(4) $D(P_{k-1}, P_k) < 6\varepsilon'/2^{k-1}$,

where $u_k > \max(u_{k-1}, n_1(\varepsilon'/2^k), k)$.

Applying Lemma 8 for $\eta'(\varepsilon'/2^{n+1})$ and $u_{n+1} > \max(u_n, n_1(\varepsilon'/2^{n+1}), n+1)$ we get P_{n+1} which satisfies (1–4) for $k = n+1$. Thus we have a sequence of partitions P_k, $k = 1, 2, \cdots$, each of which satisfies (1–4). Since the space $(\mathscr{P}, D)$ is complete, (4) implies there exists a partition P such that

$$D(P_k, P) \to 0 \text{ as } k \to \infty.$$

For each pair k, n such that $u_k \geqq n$, (1) implies

$$d(\bigvee_0^{n} T_1^i R, \bigvee_0^{n} T^i P_k) < \eta'(\varepsilon'/2^k).$$

Letting $k \to \infty$ we obtain (9.4). By (4)

$$D(P', P) \leqq \sum_{k=1}^{\infty} D(P_k, P_{k+1}) + D(P', P_1) < 12\varepsilon' = \varepsilon.$$

Remark. In Lemma 9, without assuming (9.1–3) we have a partition P which satisfies only (9.4).

Thus we have the following

Extended Sinai's Theorem. *Let R be a weak Bernoulli generator for T_1 with $H(R) < \infty$ and T be ergodic with $h(T_1) \leqq h(T)$. Then there exists a partition P such that*

$$(P, T) \sim (R, T_1).$$

Proof. Choose a partition Q such that $h(Q, T) = h(T_1)$ and $H(Q) < \infty$. Then we can apply Lemma 9 to T acting on $\bigvee_{-\infty}^{\infty} T^i Q$ and T_1, so obtain a partition P which satisfies the conclusion.

Using Lemma 8 we have

Lemma 10. *Let P and Q be weak Bernoulli for T respectively and $H(P) < \infty$, $H(Q) < \infty$. Assume there exists $\gamma > 0$ such that*

(10.1) $$h(P, T) = h(Q, T),$$

(10.2) $$Q \subset \bigvee_{-\infty}^{\infty} T^i P,$$

(10.3) $$P \overset{\gamma}{\subset} \bigvee_{-\infty}^{\infty} T^i Q.$$

Then given $\varepsilon > 0$ and $\varepsilon' > 0$ there exists a partition $P_1 \subset \bigvee_{-\infty}^{\infty} T^i Q$ and a positive integer K such that

(10.4) $$(P_1, T) \sim (P, T),$$

(10.5) $$Q \overset{\varepsilon}{\subset} \bigvee_{-K}^{K} T^i P_1,$$

(10.6) $$D(P, P_1) < 2\gamma + \varepsilon',$$

(10.7) $$D(Q, \psi(Q)) < \varepsilon,$$

where ψ is the canonical map from (P, T) to (P_1, T).

Proof. Let $0 < \varepsilon_1 < \min(\varepsilon/6, \varepsilon')$. By (10.2) there exists K_1 such that $Q \overset{\varepsilon_1}{\subset} \bigvee_{-K_1}^{K_1} T^i P$. Lemma 9 implies there exists $\delta > 0$ and u such that if $P' \subset \bigvee_{-\infty}^{\infty} T^i Q$ satisfies

(1) $d(\bigvee_0^u T^i P, \bigvee_0^u T^i P') < \delta$,

(2) $h(P, T) - h(P', T) < \delta$,

(3) $|H(\bigvee_0^j T^i P) - H(\bigvee_0^j T^i P')| < \delta, \quad 0 \leqq j \leqq u$,

then there exists $P_1 \subset \bigvee_{-\infty}^{\infty} T^i Q$ such that

(4) $(P_1, T) \sim (P, T)$,

(5) $D(P, P_1) < \varepsilon_{1/2K_1+1}$.

Let $0 < \delta_1 < \delta/2$. Choose $0 < \varepsilon_2 < \varepsilon_1$ so that $D(Q, Q') < 3\varepsilon_2$ implies $h(Q, T) - \delta_1 < h(Q', T)$. Choose $K_2 > \max(K_1, u)$ such that $Q \overset{\varepsilon_2}{\subset} \bigvee_{-K_2}^{K_2} T^i P$. Define the k-section $P(k) = \{p_1, \cdots, p_{k-1}, \bigcup_{j \geqq k} p_j\}$ of P. Then there exists k such that $\bar{P} = P(k)$ satisfies

(6) $D(P, \bar{P}) < \varepsilon_1$,

(7) $Q \overset{2\varepsilon_j}{\subset} \bigvee_{-K_j}^{K_j} T^i \bar{P}, \quad j = 1, 2$,

(8) $d(\bigvee_0^u T^i P, \bigvee_0^u T^i \bar{P}) < \delta_1$,

(9) $|H(\bigvee_0^j T^i P) - H(\bigvee_0^j T^i \bar{P})| < \delta_1, \quad 0 \leqq j \leqq u$.

Choose $0 < \delta_2 < \min(\delta_1, \varepsilon_2)$ so that if P' is a partition with k atoms and satisfies $d(\bigvee_0^u T^i \bar{P}, \bigvee_0^u T^i P') < \delta_2$ then $|H(\bigvee_0^j T^i \bar{P}) - H(\bigvee_0^j T^i P')| < \delta_1$, $0 \leqq j \leqq u$.

Choose n so large that $K_2/n<\delta_2/10$. Rohlin's theorem says that there exists a set $F\in\bigvee_{-\infty}^{\infty}T^iQ$ such that T^iF, $-n\leqq i\leqq n$, are disjoint and $X_1=\bigcup_{-n}^{n}T^iF$ satisfies $m(X_1)>1-\delta_2/10$. Define the partition $R=\{T^{-n}F, T^{-n+1}F,\cdots, T^nF, X_1^c\}$. By (10.3) we have a partition $\tilde{P}\subset\bigvee_{-\infty}^{\infty}T^iQ$ such that $D(P, \tilde{P})<\gamma$.

Consider now the partitions $\bigvee_{-n}^{n}T^i(P\vee Q)\vee\tilde{P}\vee R(\subset\bigvee_{-\infty}^{\infty}T^iP)$ and $\bigvee_{-n}^{n}T^iQ\vee\tilde{P}\bigvee_{-\infty}^{\infty}R(\subset\bigvee_{-\infty}^{\infty}T^iQ)$. Since $\bigvee_{-\infty}^{\infty}T^iQ$ has no atom, there exist partitions $\hat{P}_i\subset\bigvee_{-\infty}^{\infty}T^iQ$, $-n\leqq i\leqq n$, satisfying

$$(10)\quad d(\bigvee_{-n}^{n}T^i(P\vee Q)\vee\tilde{P}\vee R)=d(\bigvee_{-n}^{n}(\hat{P}_i\vee T^iQ)\vee\tilde{P}\vee R),$$

$$T^i\hat{P}_0/T^{k+i}F=\hat{P}_i/T^{k+i}F,$$

for $-n\leqq i\leqq n$, $-n\leqq k+i\leqq n$. Hence putting $X_2=\bigcup_{-n+K_2}^{n-K_2}T^iF$ we have

$$(11)\quad \bigvee_{-K_2}^{K_2}T^i\hat{P}_0/X_2=\bigvee_{-K_2}^{K_2}\hat{P}_i/X_2,$$

$$(12)\quad m(X_2)>1-\delta_2/5.$$

Thus we have

$$(13)\quad d(\bigvee_0^u T^iP, \bigvee_0^u T^i\hat{P}_0)=d(\bigvee_0^u\hat{P}_i, \bigvee_0^u T^i\hat{P}_0)$$

$$\leqq 2m(X_2^c)<2\delta_2/5.$$

Let $\bar{P}_i=\hat{P}_i(k)$ be the k-section of $\hat{P}_i$, $-n\leqq i\leqq n$. Then (13) implies

$$(14)\quad d(\bigvee_0^u T^i\bar{P}, \bigvee_0^u T^iP_0)<\delta_2,$$

and so the choice of δ_2 implies

$$(15)\quad |H(\bigvee_0^j T^i\bar{P})-H(\bigvee_0^j T^i\bar{P}_0)|<\delta_1,\quad 0\leqq j\leqq u.$$

Because of (7) we have a partition $L^{(j)}$ of the set of all sequences $(n_{-K_j}, n_{-K_j+1},\cdots, n_{K_j})$, $1\leqq n_i\leqq k$, $-K_j\leqq i\leqq K_j$, such that $D(L_P^{(j)}, Q)<2\varepsilon_j$, $j=1, 2$ (see §2). Let $L_{\bar{P}}^{(j)}$ and $L_{\bar{P}_0}^{(j)}(j=1, 2)$ be defined as in §2 by $\bigvee_{-K_j}^{K_j}\bar{P}_i$

and $\bigvee_{-K_j}^{K_j} T^i \bar{P}_0$ respectively and $L^{(j)}$. Then (10) implies $D(L_{\bar{P}}^{(j)}, Q) < 2\varepsilon_j$, $j=1, 2$, and (11) implies $L_{\bar{P}_0}^{(j)}/X_2 = L_{\bar{P}}^{(j)}/X_2$, $j=1, 2$. Thus (12) implies

$$(16) \quad D(L_{\bar{P}_0}^{(j)}, Q) < 2\varepsilon_j + 2m(X_2^c) < 3\varepsilon_j, \quad j=1, 2.$$

In particular $D(L_{\bar{P}_0}^{(2)}, Q) < 3\varepsilon_2$ which implies

$$(17) \quad h(Q, T) - \delta_1 < h(L_{\bar{P}_0}^{(2)}, T) \leqq h(\bar{P}_0, T).$$

Put $P' = \bar{P}_0$, then P' satisfies (1–3). Indeed (8) and (14) imply (1); (17) and (10.1) imply (2); (9) and (15) imply (3). Hence Lemma 9 implies there exists a partition $P_1 \subset \bigvee_{-\infty}^{\infty} T^i Q$ satisfying (4) and (5). Let $P^* = P_1(k)$ the k-section of P_1 and let $L_{P^*}^{(1)}$ be defined in the way of §2 by $\bigvee_{-K_1}^{K_1} T^i P^*$ and $L^{(1)}$. Then by (5)

$$(18) \quad D(L_{P^*}^{(1)}, L_{\bar{P}_0}^{(1)}) \leqq D(\bigvee_{-K_1}^{K_1} T^i P^*, \bigvee_{-K_1}^{K_1} T^i \bar{P}_0)$$

$$\leqq (2K_1+1) D(P^*, \bar{P}_0) < \varepsilon_1,$$

and so (16) implies

$$D(Q, L_{P^*}^{(1)}) < 4\varepsilon_1 < \varepsilon$$

which implies (10.5) for $K=K_1$. Since (10) and (6) we have

$$D(P, P_1) \leqq D(P, \tilde{P}) + D(\tilde{P}, \hat{P}_0) + D(\hat{P}_0, \bar{P}_0)$$

$$= 2D(P, \tilde{P}) + D(P, \bar{P}) < 2\gamma + \varepsilon_1,$$

which implies (10.6). Finally note $\psi(L_{\bar{P}}^{(1)}) = L_{P^*}^{(1)}$. Then (16), (18) and the choice of $L^{(1)}$ imply

$$D(Q, \psi(Q)) \leqq D(Q, L_{\bar{P}_0}^{(1)}) + D(L_{\bar{P}_0}^{(1)}, L_{P^*}^{(1)}) + D(\psi(L_{\bar{P}}^{(1)}), \psi(Q))$$

$$< 3\varepsilon_1 + \varepsilon_1 + 2\varepsilon_1 < \varepsilon.$$

Lemma 11. *Let R and P be weak Bernoulli generators for T_1 and T respectively and $h(T_1) = h(T)$ and $H(R) < \infty$, $H(P) < \infty$. Let Q be a* *partition such that*

(11.1) $$(Q, T) \sim (R, T_1).$$

Given $\varepsilon > 0$, *there exist a partition* Q_1 *and a positive integer* K *such that*

(11.2) $$(Q_1, T) \sim (R, T_1),$$

(11.3) $$P \overset{\varepsilon}{\subset} \bigvee_{-K}^{K} T^i Q_1,$$

(11.4) $$D(Q, Q_1) < \varepsilon.$$

Proof. Let $0 < \varepsilon_1 < \varepsilon/4$. Because of (11.1) we can apply Lemma 10 to obtain $P_1 \subset \bigvee_{-\infty}^{\infty} T^i Q$ such that

(1) $(P_1, T) \sim (P, T)$,

(2) $Q \overset{\varepsilon_1}{\subset} \bigvee_{-\infty}^{\infty} T^i P_1$,

(3) $D(Q, \psi(Q)) < \varepsilon_1$,

where ψ is the canonical map from (P, T) to (P_1, T). Now (1) implies $h(Q, T) = h(P, T) = h(P_1, T)$. We can again apply Lemma 10 to obtain $Q' \subset \bigvee_{-\infty}^{\infty} T^i P_1$ and $K > 0$ such that

(4) $(Q', T) \sim (Q, T) \sim (R, T_1)$,

(5) $P_1 \overset{\varepsilon_1}{\subset} \bigvee_{-K}^{K} T^i Q'$,

(6) $D(Q, Q') < 3\varepsilon_1$.

Let $Q_1 = \psi^{-1}(Q')$. Then (4) implies (11.2), (5) implies (11.3), and (3) and (6) imply

$$D(Q, Q_1) = D(\psi(Q), Q') \leqq D(\psi(Q), Q) + D(Q, Q') < 4\varepsilon_1 < \varepsilon.$$

Now we can prove Proposition 1.

Proof of Proposition 1. Let R and P be weak Bernoulli generators for T_1 and T respectively with $h(T_1)=h(T)$ and $H(R)<\infty$, $H(P)<\infty$. By the Extended Sinai's Theorem there exists a partition Q such that $(Q, T)\sim(R, T_1)$. Then Lemma 11 implies that for $\varepsilon_1<2^{-1}$ we can find a partition Q_1 and K_1 such that

(1) $(Q_1, T)\sim(R, T_1)$,

(2) $D(Q, Q_1)<\varepsilon_1$,

(3) $P\overset{2^{-1}}{\subset}\bigvee_{-K_1}^{K_1} T^iQ_1$.

Choose $\varepsilon_2<2^{-2}$ so small that (3) and

(4) $D(Q_1, Q_2)<\varepsilon_2$

imply

(5) $P\overset{2^{-1}+2^{-2}}{\subset}\bigvee_{-K_1}^{K_1} T^iQ_2$.

Apply Lemma 11 to obtain Q_2 and K_2 satisfying (4),

(6) $(Q_2, T)\sim(R, T_1)$,

(7) $P\overset{2^{-2}}{\subset}\bigvee_{-K_2}^{K_2} T^iQ_2$.

Suppose now we have got Q_n and $K_1,\cdots, K_n$ such that

(8) $(Q_n, T)\sim(R, T_1)$,

(9) $P\overset{2^{-j}+\cdots+2^{-n}}{\subset}\bigvee_{-K_j}^{K_j} T^iQ_n$, $1\leqq j\leqq n$.

Choose $\varepsilon_{n+1}<2^{-n-1}$ so small that (9) and

(10) $D(Q_n, Q_{n+1})<\varepsilon_{n+1}$

imply

(11) $P\overset{2^{-j}+\cdots+2^{-n-1}}{\subset}\bigvee_{-K_j}^{K_j} T^iQ_{n+1}$, $1\leqq j\leqq n$.

Apply Lemma 11 to obtain Q_{n+1} and K_{n+1} satisfying (10),

(12) $(Q_{n+1}, T) \sim (R, T_1)$,

(13) $P \overset{2^{-n-1}}{\subset} \bigvee_{-K_{n+1}}^{K_{n+1}} T^i Q_{n+1}$.

Thus by induction we have the sequence of partitions Q_n, $n=1, 2, \cdots$, which satisfy (8), (9) and (10). Then (10) implies there exists a countable partition Q_∞ such that

$$D(Q_n, Q_\infty) \to 0 \text{ as } n \to \infty.$$

Then (8) implies

(14) $(Q_\infty, T) \sim (R, T_1)$,

and (9) implies

(15) $P \subset \bigvee_{-\infty}^{\infty} T^i Q_\infty$.

Since P is a generator, the proposition follows from (14) and (15).

§4. General Case

In this section, using the results of the preceding section, we will prove the following theorem of Ornstein.

Theorem. *Two generalized weak Bernoulli transformations with the same entropy* (*including the case* ∞) *are isomorphic.*

To prove this theorem, we must prepare some lemmas. The essential one is Lemma 14 which is similar to Lemma 4′ in [5] and is proved by some modified argument. Using this lemma and Lemma 11 we can get Lemma 15 and 16 respectively, which prove our theorem. For completeness we will state the proofs of these lemmas.

Lemma 12. *Let* $\varepsilon>0$ *and* $\zeta=\zeta(\varepsilon)$ *as in Lemma* 2. *If countable partitions* P and R *satisfy*

$$H(P)+H(R)-h(P \vee R, T)<\zeta,$$

then

$$T^iP \overset{\varepsilon}{\perp} \bigvee_0^n T^jR \vee \bigvee_0^{i-1} T^jP, \qquad 0 \leqq i \leqq n.$$

Remark. This conclusion implies $P \overset{\varepsilon}{\perp} \bigvee_0^n T^jR$ when $i=0$.

Proof. Noting the following relations

$$h(P\vee R,\ T)=\lim_{n\to\infty}\frac{1}{n+1}\left[H(\bigvee_0^n T^jR)+\sum_0^n H(T^iP|\bigvee_0^n T^jR\vee\bigvee_0^{i-1}T^jP)\right],$$

$$h(R,\ T)=\lim_{n\to\infty}\frac{1}{n+1}H(\bigvee_0^n T^jR),$$

$$H(T^kP\mid \bigvee_0^{m+n} T^jR\vee\bigvee_0^{i-1}T^jP)\leqq H(T^iP|\bigvee_0^n T^jR\vee\bigvee_0^{i-1}T^jP),$$

for $m>0$, $i\leqq k\leqq i+m$, we have

$$h(P\vee R,\ T)\leqq h(R,\ T)+H(T^iP\mid \bigvee_0^n T^jR\vee\bigvee_0^{i-1}T^jP)$$

$$\leqq H(R)+H(T^iP\mid \bigvee_0^n T^jR\vee\bigvee_0^{i-1}T^jP),$$

for any $n\geqq 0$, $0\leqq i\leqq n$. Then the assumption and the above inequality imply

$$H(P)-H(T^iP|\bigvee_0^n T^jR\vee\bigvee_0^{i-1}T^jP)<\zeta,\quad 0\leqq i\leqq n.$$

Therefore the lemma follows from Lemma 2.

Lemma 13. *Let* $\varepsilon>0$ *and* $\varepsilon'>0$. *Let* $\{P_i\}$, $\{R_i\}$, $0\leqq i\leqq n$, *be the sequences of countable partitions and* π *be a probability vector such that*

(13.1) $$P_i \overset{\varepsilon}{\perp} \bigvee_0^n R_j\vee\bigvee_0^{i-1}P_j,\quad 0\leqq i\leqq n,$$

(13.2) $\{R_i\}$, $0\leqq i\leqq n$, *are independent*,

(13.3) $$d(P_i,\ \pi)<\varepsilon',\quad 0\leqq i\leqq n.$$

Then there exists a sequence of partitions $\{\bar{P}_i\}$, $0 \leqq i \leqq n$, *satisfying*

(13.4) $$\{\bar{P}_i, R_j\},\ 0 \leqq i, j \leqq n, \text{ are independent},$$

(13.5) $$d(\bar{P}_i) = \pi, \quad 0 \leqq i \leqq n,$$

(13.6) $$D(P_i, \bar{P}_i) < \varepsilon + \varepsilon', \quad 0 \leqq i \leqq n.$$

Proof. Note that the following statement is useful for the proof; if $P \overset{\varepsilon}{\perp} Q$ and $d(P, \pi) < \varepsilon'$, then there exists a partition $\bar{P}$ such that $\bar{P}$ and Q are independent, $d(\bar{P}) = \pi$, and $D(P, \bar{P}) < \varepsilon + \varepsilon'$. Using this fact we can easily prove this lemma by induction.

Lemma 14. *Let* T *be an ergodic transformation and* $h(T) < \infty$. *Let* $\varepsilon > 0$, $\zeta = \zeta(\varepsilon)$ *be as in Lemma* 2, *and* $\theta(\varepsilon) = \min(\zeta(\varepsilon^2/2), \varepsilon^2/2)/2$. *If a probability vector* $\pi = \{\pi_i; i \geqq 1\}$ *and countable partitions* P *and* R *satisfy*

(14.1) $$\{T^i R\},\ -\infty < i < \infty, \text{ are independent},$$

(14.2) $$H(\pi) = h(T) - H(R),$$

(14,3) $$d(P, \pi) < \theta(\varepsilon),$$

(14.4) $$H(P) + H(R) - h(P \vee R, T) < \theta(\varepsilon),$$

then given $\delta > 0$ *there exists a (finite) partition* $\tilde{P}$ *such that*

(14.5) $$d(\tilde{P}, \pi) < \delta,$$

(14.6) $$H(\tilde{P}) + H(R) - h(\tilde{P} \vee R, T) < \delta,$$

(14.7) $$D(P, \tilde{P}) < 15\varepsilon.$$

Proof. We may assume $\delta < \theta(\varepsilon)$. Let $\pi' = \{\pi'_i; 1 \leqq i \leqq l\}$ be a (finite) probability vector such that the following two conditions are satisfied;

(1) $d(\pi, \pi') = \sum_i |\pi_i - \pi'_i| < \dfrac{\delta}{4}$,

(2) $H(\pi')=H(\pi)+\alpha, \quad 0<\alpha<\frac{\delta}{4}.$

By the assumption (14.4) we can apply Lemma 12 and we get, for every $n\geqq 0$,

(3) $T^iP \overset{\varepsilon^2/2}{\perp} \bigvee_0^n T^jR \vee \bigvee_0^{i-1} T^jP, \quad 0\leqq i\leqq n,$

and using (14.3) and (1), we get

(4) $d(T^iP, \pi')<\varepsilon^2/2.$

So by (3), (14.1) and (4) we can apply Lemma 13 to get a sequence of partitions $\{\bar{P}_i;\, 0\leqq i\leqq n\}$ such that

(5) $\{\bar{P}_i, T^jR;\ 0\leqq i, j\leqq n\}$ are independent,

(6) $d(\bar{P}_i)=\pi', \quad 0\leqq i\leqq n,$

(7) $D(\bar{P}_i, T^iP)<\varepsilon^2, \quad 0\leqq i\leqq n.$

By the ergodicity, T has a countable generator, so that we have a refinement Q of $P\vee R$ satisfying

(8) $h(Q, T)=h(T).$

Choose $0<\varepsilon'<\min(1/2, \varepsilon/2, \delta/12)$ so small that the following two conditions are valid;

(9) $D(Q, Q')<10\varepsilon'$ implies $h(Q, T)-\delta/4<h(Q', T),$

(10) $d(\tilde{P}, \pi')<10\varepsilon'$ and $N(\tilde{P})\leqq l$ imply

$$|H(\tilde{P})-H(\pi')|<\delta/4.$$

Choose $n>3/\alpha$ so large that the following three statements are valid: (A) By Shannon-McMillan-Breiman theorem setting $\bar{Q}=\bigvee_0^{n-1} T^kQ$ there exists a $\bar{Q}$-measurable set Y such that

(11) $m(Y)>1-\varepsilon',$

(12) $2^{-n(h(T)+\alpha/3)}<m(B)<2^{-n(h(T)-\alpha/3)}, \quad B\in\bar{Q}/Y,$

(B) we will denote the $\bar{P}\vee R$-name of $x\in X$ as $\bar{s}(x)=\{(\bar{s}_k(x), t_k(x)); 0\leqq k\leqq n-1\}$ when $x\in\bigcap_0^{n-1}\bar{p}^k_{s_k}(x)\cap T^k r_{t_k}(x)$ for $\bar{P}_k=\{\bar{p}^k_i;\ 1\leqq i\leqq l\}$ and $R=\{r_i;\ i\geqq 1\}$. Let $N(j;\bar{P}, x)$ be the number of k which satisfies $\bar{s}_k(x)=j$. Then $N(j;\bar{P}, G)$ is defined naturally for $G\in\bigvee_0^{n-1}(\bar{P}_k\vee T^kR)$. Using the law of large numbers, by (5) and (6), setting $\hat{P}=\bigvee_0^{n-1}(\bar{P}_k\vee T^kR)$, there exists a $\hat{P}$-measurable set $\bar{Y}$ such that

(13) $m(\bar{Y})>1-\varepsilon'$,

(14) $2^{-n(H(\pi')+H(R)+\alpha/3)}<m(G)<2^{-n(H(\pi')+H(R)-\alpha/3)}, \quad G\in\hat{P}/\bar{Y},$

(15) $|N(j;\bar{P}, G)/n-\pi'_j|<\varepsilon'/l, \quad G\in\hat{P}/\bar{Y},$

(C) $m(A)<1/n$ implies $H(\{A, A^c\})<\delta/4$. By the choice of n in (A) and (B), we can estimate

(16) $N(\bar{Q}/Y)<2^{n(h(T)+\alpha/3)}$,

(17) $N(\hat{P}/\bar{Y})>2^{n(h(T)+\alpha/3)}$.

Now we will also define the $P\vee R$-name $s(x)=\{(s_k(x), t_k(x));\ 0\leqq k\leqq n-1\}$ of $x\in X$ by the same way as above. Note that this name is the same in every atom of $\bigvee_0^{n-1}T^kQ$. Let E_k denote the set of all points x such that $\bar{s}_k(x)\neq s_k(x)$ and D_e denote the set of all points x for which $\bar{s}_k(x)$ and $s_k(x)$ are distinct in at least e places, then we have $\sum_0^{n-1}m(E_k)\geqq em(D_e)$. On the other hand, for $P=\{p_i;\ i\geqq 1\}$ and $\bar{P}_k=\{\bar{p}^k_i;\ 1\leqq i\leqq l\}$, we have $D(T^k(P\vee R), \bar{P}_k\vee T^kR)=2m(E_k)$. Then setting $W=D^c_{n\varepsilon}$ we have

(18) $m(W)>1-\varepsilon$.

Let $\mathscr{B}$ be the collection of atoms B in $\bar{Q}/Y$ which satisfy $m(B\cap W\cap\bar{Y})\geqq m(B)/2$ and let $Z=\cup\mathscr{B}$ the union of sets in $\mathscr{B}$. Then we can estimate

(19) $m(Z^c)<2(\varepsilon+\varepsilon')<3\varepsilon$.

Now we can assign to each atom $B\in\mathscr{B}$ an atom $G=\psi(B)\in\hat{P}/\bar{Y}$

such that $B\cap G\cap W\neq\emptyset$ in a 1–1 manner. This can be done by the same way as in the proof of Lemma 8 using the marriage lemma because of (16) and (17). Note that two names $s(B)$ and $\bar{s}(\psi(B))$ are the same in the second component and are distinct in at most $n\varepsilon$ places in the first component. We can extend ψ to a 1–1 mapping from $\bar{Q}/Y$ into $\hat{P}/\bar{Y}$ by (16) and (17).

Using Rohlin's theorem and the same argument as in the proof of Lemma 8, we get a measurable set F such that $T^{-k}F$, $0\leqq k\leqq n-1$, are disjoint, $m(\bigcup_0^{n-1} T^{-k}F)>1-\varepsilon'$ and

$$(20)\quad m(\bigcup_0^{n-1} T^{-k}(F\cap Z))>1-\varepsilon'-6\varepsilon,$$

$$(21)\quad m(\bigcup_0^{n-1} T^{-k}(F\cap Y))>1-3\varepsilon'.$$

We will now define the partition $\tilde{P}=\{p_1, p_2,\cdots, p_l\}$ as follows. Let $X_1=\bigcup_0^{n-1} T^{-k}(F\cap Y)$ and $A=F\cap Y$. For simplicity we denote the $\bar{P}\vee R$-name of $\psi(B)$, $B\in\bar{Q}/Y$ as $\bar{s}(B)=\{(\bar{s}_k(B), t_k(B));\ 0\leqq k\leqq n-1\}$. Define $\tilde{P}$ on $T^{-k}A$, $0\leqq k\leqq n-1$, by

$$\tilde{p}_j\cap T^{-k}A=\cup\{T^k(F\cap B);\ B\in\bar{Q}/Y,\ s_k(B)=j\},\quad 1\leqq j\leqq l.$$

And extend $\tilde{P}$ to X_1^c so that $X_1^c\subset\tilde{p}_1$. Then this partition $\tilde{P}$ satisfies (14.5–7). Indeed

$$\begin{aligned} m(\tilde{p}_j)&=\sum_{k=0}^{n-1}[m(\tilde{p}_j\cap T^{-k}A)+m(\tilde{p}_j\cap X_1^c)]\\ &=\sum_{k=1}^{n-1}[\sum_{\substack{B\in\bar{Q}/Y\\ \bar{s}_k(B)=j}} m(F\cap B)+m(\tilde{p}_j\cap X_1^c)]\\ &=\sum_{B\in\bar{Q}/Y} N(j:\bar{P},\ B)m(F\cap B)+m(X_1^c), \end{aligned}$$

by (15) and (21) we have

$$\begin{aligned} (22)\quad d(\tilde{P},\pi')&\leqq\sum_{j=1}^{l}|m(\tilde{p}_j)-\pi'_j|\\ &<\sum_{B\in\bar{Q}/Y} nm(F\cap B)\sum_{j=1}^{l}|N(j;\bar{P},\ B)/n-\pi'_j|+2m(X_1^c) \end{aligned}$$

$$< 7\varepsilon' < \frac{3}{4}\delta,$$

so by (1) we have (14.5);

$$d(\tilde{P}, \pi) \leqq d(\tilde{P}, \pi') + d(\pi', \pi) < \delta.$$

Secondly the choice (C) of n implies

$$H(\{A, A^c\}) < \delta/4.$$

On the other hand putting $P^* = \bigvee_{-n}^{n} T^k(\tilde{P} \vee \{A, A^c\} \vee R)$ we have

$$\{T^{-k}(F \cap B);\ 0 \leqq k \leqq n-1,\ B \in \bar{Q}/Y\} \subset P^*,$$

so we have a subpartition Q' of P^* which coincides with Q on X_1 so that $D(Q, Q') \leqq 2m(X_1^c) < 8\varepsilon'$. Therefore by (8) and (9), these relations imply

$$\begin{aligned} h(T) = h(Q, T) &< h(Q', T) + \delta/4 \\ &\leqq h(P^*, T) + \delta/4 \\ &\leqq h(\tilde{P} \vee R, T) + H(\{A, A^c\}) + \delta/4 \\ &< h(\tilde{P} \vee R, T) + \delta/2, \end{aligned}$$

and (22) implies

$$|H(\tilde{P}) - H(\pi')| < \delta/4,$$

so by (2) we have (14.6);

$$H(\tilde{P}) + H(R) - h(\tilde{P} \vee R, T) < \delta.$$

Finally noting the similarity of the following two relations

$$\tilde{p}_j \cap T^{-k}(B \cap F) = \begin{cases} T^{-k}(B \cap F), & \bar{s}_k(B) = j \\ \emptyset \quad , & \bar{s}_k(B) \neq j, \end{cases}$$

$$p_j \cap T^{-k}(B \cap F) = \begin{cases} T^{-k}(B \cap F), & s_k(B) = j \\ \emptyset, & s_k(B) \neq j \end{cases}$$

and noting the estimation by (20)

$$m(X_2) < \varepsilon' + 6\varepsilon \quad \text{where} \quad X_2 = X \setminus \bigcup_{k=0}^{n-1} T^{-k}(F \cap Z),$$

we have

$$\begin{aligned} D(P, \tilde{P}) &= \sum_j \sum_{B \in \mathscr{B}} \sum_{k=0}^{n-1} m[(p_j \cap T^{-k}(B \cap F)) \triangle (\tilde{p}_j \cap T^{-k}(B \cap F))] \\ &\qquad + D(P, \tilde{P} \mid X_2) m(X_2) \\ &\leqq \sum_B \sum_{k;\bar{s}_k(B) \neq s_k(B)} 2m(T^{-k}(B \cap F)) + 13\varepsilon \\ &\leqq 2n\varepsilon \sum_{B \in \mathscr{B}} m(B \cap F) + 13\varepsilon < 15\varepsilon. \end{aligned}$$

Remark. If we only assume (14.1) and (14.2), we obtain $\tilde{P}$ which satisfies (14.5) and (14.6).

Using this lemma we can prove the following

Lemma 15. *Let T be an ergodic transformation and $h(T) < \infty$. If a partition R satisfies*

(15.1) $\{T^i R\}$, $-\infty < i < \infty$, *are independent,*

(15.2) $H(R) \leqq h(T)$,

then there is a (finite) partition $\tilde{P}$ such that

(15.3) $H(\tilde{P}) + H(R) = h(T)$,

(15.4) $\{T^i(\tilde{P} \vee R)\}$, $-\infty < i < \infty$, *are independent and $\tilde{P}$ and R are independent.*

Proof. Take a (finite) probability vector $\pi = \{\pi_i;\ 1 \leqq i \leqq l\}$ such that $H(\pi) = h(T) - H(R)$. Then, by the previous remark we can apply Lemma

14 to get a (fiinite) partition P_0 for every δ_0 such that

(1) $d(P_0, \pi)<\theta(\delta_0)$,

(2) $H(P_0)+H(R)-h(P_0\vee R, T)<\theta(\delta_0)$,

where $\theta(\delta_0)$ is defined in Lemma 14. Choose $\delta_n\searrow 0$, $n\geqq 1$, we can apply Lemma 14 to get a partition P_1 for $\theta(\delta_1)$ such that

(3) $d(P_1, \pi)<\theta(\delta_1)$,

(4) $H(P_1)+H(R)-h(P_1\vee R, T)<\theta(\delta_1)$,

(5) $D(P_0, P_1)<15\delta_0$.

Then we obtain a sequence of partitions $\{P_i\}$, $i\geqq 0$, by induction such that

(6) $d(P_j, \pi)<\theta(\delta_j)$, $j=0, 1,\cdots, n,\cdots$

(7) $H(P_j)+H(R)-h(P_j\vee R, T)<\theta(\delta_j)$, $j=0, 1,\cdots, n,\cdots$

(8) $D(P_{j-1}, P_j)<15\delta_{j-1}$, $j=1, 2,\cdots, n,\cdots$

Choosing δ_n so small that (8) implies the existence of a partition $\tilde{P}$ such that $D(P_n, \tilde{P})\to 0$ as $n\to\infty$. Then $d(\tilde{P})=\pi$ and $\tilde{P}$ is a finite partition and satisfies (15.3). On the other hand for every $n\geqq 0$ we will take $A\in T^n(\tilde{P}\vee R)$ and $B\in \bigvee_0^{n-1} T^i(\tilde{P}\vee R)$, and take $A_k\in T^n(P_k\vee R)$ and $B_k\in \bigvee_0^{n-1} T^i(P_k\vee R)$ corresponding to A and B respectively for every $k\geqq 0$. Then

$$|m(A\cap B)-m(A_k\cap B_k)|\leqq(n+1)D(\tilde{P}, P_k)\to 0, \quad k\to\infty,$$

$$|m(A)-m(A_k)|\to 0, \qquad |m(B)-m(B_k)|\to 0, \qquad k\to\infty,$$

and by (7) Lemma 12 says

$$T^nP_k \overset{\delta_k^2/2}{\perp} \bigvee_0^n T^jR\vee \bigvee_0^{n-1} T^jP_k, \quad T^nR \overset{\delta_k^2/2}{\perp} \bigvee_0^n T^jP_k\vee \bigvee_0^{n-1} T^jR.$$

Therefore we have

$$T^n(P_k \vee R) \overset{3\delta_k^2/2}{\perp} \bigvee_0^{n-1} T^j(P_k \vee R)$$

and then

$$|m(A_k \cap B_k) - m(A_k)m(B_k)| < \frac{3}{2}\delta_k^2 \to 0 \text{ as } k \to \infty.$$

Hence $T^n(\tilde{P} \vee R)$ and $\bigvee_0^{n-1} T^i(\tilde{P} \vee R)$ are independent. Similarly we can prove the independence of $\tilde{P}$ and R. Thus we have (15.4).

Remark. If $H(R) = h(T)$, then we take the trivial partition as $\tilde{P}$.

For proving our theorem it is convenient to state a special case of Lemma 11 in the following

Lemma 16. *Let P be a weak Bernoulli generator for T and $H(P) < \infty$. Let Q be weak Bernoulli for T such that $h(Q, T) = h(T)$ and $H(Q) < \infty$. Given $\varepsilon > 0$, there exist a partition Q_1 and a positive integer K such that*

(16.1) $$(Q_1, T) \sim (Q, T),$$

(16.2) $$P \overset{\varepsilon}{\subset} \bigvee_{-K}^{K} T^i Q_1,$$

(16.3) $$D(Q, Q_1) < \varepsilon.$$

Proof of the theorem. Let P be a generalized weak Bernoulli generator for T. Take an increasing sequence of finite partitions $\{P_n\}$, $n \geqq 1$, such that each P_n is weak Bernoulli for T and $\bigvee_{n=1}^{\infty} P_n = P$ (hence $h(P_n, T) \nearrow h(P, T) = h(T)$).

Since P_1 is weak Bernoulli for T and $H(P_1) < \infty$, there exists a finite partition ${}_1Q_1$ by Proposition 1 such that

(1) ${}_1Q_1 \subset \bigvee_{-\infty}^{\infty} T^i P_1$

(2) $\{T_1^i Q_1\}$, $-\infty < i < \infty$, are independent

(3) $\bigvee_{-\infty}^{\infty} T_1^i Q_1 = \bigvee_{-\infty}^{\infty} T^i P_1$. (Hence $H({}_1Q_1) = h(P_1, T)$.)

Applying Lemma 15 for P_2 and T, we have a finite partition ${}_1Q_2 \subset \bigvee_{-\infty}^{\infty} T^i P_2$ such that

(4) $H({}_1Q_1) + H({}_1Q_2) = h(P_2, T)$

(5) $\{T^i({}_1Q_1 \vee {}_1Q_2)\}$, $-\infty < i < \infty$, are independent and ${}_1Q_1$ and ${}_1Q_2$ are independent.

Let $0 < \varepsilon_1 < 2^{-1}$ and take $K_1 > 0$ such that

(6) $P_1 \overset{\varepsilon_1}{\subset} \bigvee_{-K_1}^{K_1} T^i_1 Q_1$.

Choosing $0 < \varepsilon_2 < 2^{-2}$ so small that the following statement holds.

(7) $D({}_1Q_1, {}_2Q_1) < \varepsilon_2$ implies $P_1 \overset{2^{-1}}{\subset} \bigvee_{-K_1}^{K_1} T^i_2 Q_1$.

Applying Lemma 16 for P_2, T and ${}_1Q_1 \vee {}_1Q_2$, we have partitions ${}_2Q_1$, ${}_2Q_2 \subset \bigvee_{-\infty}^{\infty} T^i P_2$ and $K_2 > 0$ such that

(8) $\{T^i({}_2Q_1 \vee {}_2Q_2)\}$, $-\infty < i < \infty$, are independent and ${}_2Q_1$ and ${}_2Q_2$ are independent,

(9) $d({}_2Q_1 \vee {}_2Q_2) = d({}_1Q_1 \vee {}_1Q_2)$,

(10) $P_2 \overset{\varepsilon_2}{\subset} \bigvee_{-K_2}^{K_2} T^i ({}_2Q_1 \vee {}_2Q_2)$,

(11) $D({}_2Q_1 \vee {}_2Q_2, {}_1Q_1 \vee {}_1Q_2) < \varepsilon_2$.

Hence by the choice of ε_2, (11) implies

(12) $P_1 \overset{2^{-1}}{\subset} \bigvee_{-K_1}^{K_1} T^i {}_2Q_1$.

Applying Lemma 15 for P_3, T and ${}_2Q_1 \vee {}_2Q_2$, we have a partition ${}_2Q_3 \subset \bigvee_{-\infty}^{\infty} T^i P_3$ such that

(13) $H({}_2Q_1) + H({}_2Q_2) + H({}_2Q_3) = h(P_3, T)$,

(14) $\{T^i(\bigvee_1^3 {}_2Q_j)\}$, $-\infty < i < \infty$, are independent and $\{{}_2Q_j\}$, $1 \leqq j \leqq 3$, are independent.

Similarly we get a sequence of partitions $\{{}_iQ_j;\ 1 \leqq j \leqq i+1,\ i \geqq 1\}$ and a sequence of positive integers $\{K_i\}$, $i \geqq 1$, by induction such that

$$(15)\quad \bigvee_1^n {}_nQ_j \subset \bigvee_{-\infty}^{\infty} T^i P_n;\ {}_nQ_{n+1} \subset \bigvee_{-\infty}^{\infty} T^i P_{n+1},$$

$$(16)\quad \{T^i(\bigvee_1^{n+1} {}_nQ_j)\},\ -\infty < i < \infty, \text{ are independent,}$$

$$(17)\quad \{{}_nQ_j\},\ 1 \leqq j \leqq n+1, \text{ are independent,}$$

$$(18)\quad d(\bigvee_1^n {}_nQ_j) = d(\bigvee_1^n {}_{n-1}Q_j),$$

$$(19)\quad P_m \overset{2^{-m}+\cdots+2^{-n}}{\subset} \bigvee_{-K_m}^{K_m} T^i(\bigvee_1^m {}_nQ_j),\quad 1 \leqq m \leqq n,$$

$$(20)\quad D(\bigvee_1^n {}_nQ_j, \bigvee_1^n {}_{n-1}Q_j) < 2^{-n},$$

$$(21)\quad \sum_1^{n+1} H({}_nQ_j) = h(P_{n+1}, T).$$

By (20) we obtain partitions ${}_\infty Q_j$, $j \geqq 1$, such that

$$D({}_nQ_j, {}_\infty Q_j) \to 0, \quad \text{as} \quad n \to \infty.$$

Then $d({}_\infty Q_j) = d({}_jQ_j) = d({}_{j-1}Q_j)$, $\{{}_\infty Q_j\}$, $j \geqq 1$, are independent by (17), $\{T^i(\bigvee_1^\infty {}_\infty Q_j)\}$, $-\infty < i < \infty$, are independent by (16), and by (19) we have

$$P_k \overset{2^{-k+1}}{\subset} \bigvee_{-\infty}^{\infty} T^i(\bigvee_1^k {}_\infty Q_j),\quad k = 1, 2, \cdots,$$

and so

$$\bigvee_{-\infty}^{\infty} T^i(\bigvee_1^\infty {}_\infty Q_j) = \bigvee_{-\infty}^{\infty} T^i P.$$

Thus $R = \bigvee_1^\infty {}_\infty Q_j$ is a Bernoulli generator for T and so (R, T) is a Bernoulli shift.

Now we will prove the isomorphism between the generalized weak Bernoulli transformations T_1 and T_2 with $h(T_1) = h(T_2) \leqq \infty$. Using the above argument there exist two partitions R_1 and R_2 such that (R_1, T_1) and (R_2, T_2) are Bernoulli shifts. If $h(T_1) = h(T_2) < \infty$, then $H(R_1) = H(R_2) < \infty$ and therefore R_1 and R_2 are at most countable partitions. So that T_1 and T_2 are isomorphic by Proposition 1.

When $h(T_1)=h(T_2)=\infty$, we will prove that they are isomorphic to the Bernoulli shift on the product space $\prod_{-\infty<i<\infty} X_i$, $X_i=[0, 1]$. Let P be a generalized weak Bernoulli generator for T and $h(T)=h(P, T)=\infty$. Then we can also assume in the above argument that $h(P_n, T)<h(P_{n+1}, T)-\log 2$, $n\geqq 1$ (by choosing subsequence of P_n). Noting that the condition for the choice of π in the proof of Lemma 15 is only $H(\pi)=h(T)-H(R)$, we can assume that all atoms of ${}_nQ_{n-1}$ have measures less than 1/2. Hence for the partition R taken for P and T in the above argument, the factor space X/R is isomorphic to the space $[0, 1]$ with the ordinary Lebesgue measure. Thus the proof of the theorem is complete.

§5. Examples

We will now give some examples of weak Bernoulli transformations. Firstly we will discuss *Markov shifts*. Let X_0 be $\{1, 2,\cdots, N\}$ or $\{1, 2,\cdots\}$ the set of all positive integers. Let $X=\prod_{-\infty}^{\infty} X_n$ where $X_n=X_0$ for all n. Let $M=(m_{ij})_{i,j\in X_0}$ be a transition matrix with stationary probability $\{m_i;\ i\in X_0\}$;

$$\sum_{i\in X_0} m_i m_{ij}=m_j, \qquad j\in X_0.$$

We assume $m_i>0$ for all $i\in X_0$. The pair of M and $\{m_i\}$ gives a Markov measure m on $\mathscr{F}$, where $\mathscr{F}$ is the complete σ-field generated by cylinders. The shift T on X defined by

$$(Tx)_n=x_{n-1}, \quad n=0, \pm 1,\cdots, \quad x=(\cdots, x_0, x_1,\cdots)\in X,$$

preserves the measure m; T on $(X, \mathscr{F}, m)$ is called a Markov shift. We assume T is ergodic. Thus M is assumed to be irreducible, recurrent and of positive type. Let $M^k=(m_{ij}^{(k)})$ be the k-step transition matrix. It is easy to see that T is mixing if and only if

$$(1) \quad \lim_{k\to\infty} m_{ij}^{(k)}=m_j, \quad i, j\in X_0.$$

Note (1) holds if and only if M is aperiodic.

Let $P=\{p_j;\, j\in X_0\}$ be the partition of X defined as $p_j=\{x;\, x_0=j\}$, $j\in X_0$. Denote $P_j^n=\bigvee_j^n T^iP$ for $j\leqq n$. We will prove the following

Proposition 2. *P is weak Bernoulli for T if and only if* (1) *holds.*

First we will see that P is a K-partition for T, i.e. $\bigcap_n P_{-\infty}^n$ is trivial, if and only if (1) holds. Indeed P is a K-partition for T if and only if

$$(2)\quad \lim_{k\to\infty}\sum_i\Big|\sum_j f(j)m_{ij}^{(k)}-\sum_j f(j)m_j\Big|m_i=0$$

for every bounded function f on X_0 (cf. [8]), and it is easy to see that (2) is equivalent to (1). On the other hand the Markov property implies for each $n\geqq 0$

$$(3)\qquad \sum_{p\in P_{-n}^0,\, q\in P_k^{k+n}}|m(p\cap q)-m(p)m(q)|$$

$$=\sum_{i,j\in X_0}|m_im_{ij}^{(k)}-m_j|$$

$$=2\sum_{i\in X_0}m_i\sup_{J\subset X_0}\Big|\sum_{j\in J}m_{ij}^{(k)}-\sum_{j\in J}m_j\Big|.$$

Since k-step transition probabilities of the time reversed chain are $\hat{m}_{ij}^{(k)}=m_jm_{ji}^{(k)}/m_i$, $i, j\in X_0$, if (1) holds then P is also a K-partition for T^{-1} and hence

$$(4)\quad \lim_{k\to\infty}\sup_{B\in P_k^\infty}|m(A\cap B)-m(A)m(B)|=0$$

for each $A\in\mathscr{F}$. Take $A=\{x_0=i\}$ and $B=\{x_k\in J\}$, then (4) implies (3) converges to 0 uniformly in n as $k\to\infty$. Thus mixing Markov shifts with countable states are weak Bernoulli transformations.

We will now consider the *continued-fraction transformation* T on $X=[0, 1]$ defined by

$$Tx=\begin{cases}\{1/x\}, & x\neq 0,\\ 0, & x=0,\end{cases}$$

with the invariant measure

$$m(A)=\frac{1}{\log 2}\int_A \frac{dx}{1+x}, \qquad A\in\mathscr{F}.$$

We will assert that T is weak Bernoulli, although it is not invertible. Consider the partition $P=\{[1/(j+1),\ 1/j];\ j=1, 2,\cdots\}$ of X. Let $P_j^n=\bigvee_j^n T^{-i}P$ for $0\leqq j\leqq n$. It is easy to see that P is a generator. Let Δ_n be the generic atom of P_0^n. Then it is known that

$$m(T^{-(k+n)}A\,|\,\Delta_n)=m(A)\,(1+\theta\rho^k), \quad A\in\mathscr{F},$$

where $|\theta|\leqq M$, $\rho<1$ and M and ρ are constants independent of A, k, n, and Δ_n (cf. [1], p. 50). This implies

$$\sum_{\Delta_n\in P_0^n,\, \Delta\in P_{n+k}^{2n+k}} |m(\Delta_n\cap\Delta)-m(\Delta_n)m(\Delta)|\leqq M\rho^k.$$

Thus T is weak Bernoulli with the generator P, and hence each natural extension of T is a Bernoulli transformation.

We remark that Y. Takahashi and one of the authors [3] showed that β-expansion transformations are also weak Bernoulli, while they are not invertible.

References

[1] Billingsley, P., *Ergodic Theory and Information*, Wiley, New York, 1965.
[2] Friedman N. A. and D. S. Ornstein, On isomorphism of weak Bernoulli transformations, *Advances in Math.* **5** (1971), 365-394.
[3] Ito, S. and Y. Takahashi, Markov subshifts, to appear.
[4] Ornstein, D. S., Bernoulli shifts with the same entropy are isomorphic, *Advances in Math.* **4** (1970), 337-352.
[5] ———, Two Bernoulli shifts with infinite entropy are isomorphic, *Advances in Math.* **5** (1971), 339-348.
[6] Rohlin, V. A., On the fundamental ideas of measure theory, *Mat. Sb.* **25** (67) (1949), 107-150. [*Amer. Math. Soc. Transl. I. Ser.* 10 (1961), 1-54.]
[7] Smorodinsky, M., An exposition of Ornstein isomorphism theorem, to appear.
[8] Totoki, H., On a class of special flows, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **15** (1970), 157-167.