The Annals of Probability
1980, Vol. 8, No. 1, 131–141

ON CODING A STATIONARY PROCESS TO ACHIEVE A GIVEN MARGINAL DISTRIBUTION

BY JOHN C. KIEFFER

University of Missouri-Rolla

The problem of coding a stationary process $\{X_i\}_{i=-\infty}^{\infty}$ onto a stationary process $\{Y_i\}_{i=-\infty}^{\infty}$ so that for some positive integer m, $(Y_0, Y_1, \cdots, Y_{m-1})$ has a given marginal distribution is considered. The problem is solved for $\{X_i\}$ nonergodic as well as ergodic. The associated universal coding problem is also solved, where one seeks to find a coding function which yields the desired marginal distribution for each member of a class of possible distributions for $\{X_i\}$.

Introduction. Let $(S, \mathbb{S})$ be a measurable space. Let $\{X_i\}_{i=-\infty}^{\infty}$ be a stationary $(S, \mathbb{S})$-valued process defined on some probability space. Let A be a finite set. Suppose it is desired to code $\{X_i\}$ onto a stationary A-valued process $\{Y_i\}_{i=-\infty}^{\infty}$. Letting $(S^\infty, \mathbb{S}^\infty)$ be the measurable space consisting of S^∞, the set of bilateral infinite sequences from S, and $\mathbb{S}^\infty$, the usual product σ-field, this means there is a measurable coding function $f: S^\infty \to A$ such that $Y_i = f(\{X_{i+j}\}_{j=-\infty}^{\infty})$, for each integer i. Suppose we wish to find f so that for some positive integer m, $(Y_0, \cdots, Y_{m-1})$ will have some previously specified marginal distribution π on A^m. For an ergodic process $\{X_i\}$, the solution to this problem follows easily from a result of Grillenberger and Krengel [1, Theorem 1.1]. In this paper we solve the problem for nonergodic stationary processes $\{X_i\}$. We also solve a related universal coding problem: given a family $\mathcal{P}$ of possible distributions for $\{X_i\}$, find f so that the desired marginal is attained for every member of $\mathcal{P}$. This type of problem would arise if the distribution of $\{X_i\}$ were unknown.

NOTATION. We fix for the remainder of the paper a finite set A, a positive integer $m \geqslant 2$, and a probability measure π on A^m. We assume π is an *invariant* distribution; that is, if $X_1^*, X_2^*, \cdots, X_m^*$ are the successive projections from $A^m \to A$, we assume that $(X_1^*, \cdots, X_{m-1}^*)$ and $(X_2^*, \cdots, X_m^*)$ have the same distribution on A^{m-1} under π. For $1 \leqslant k \leqslant m$, let π_k be the probability distribution on A^k which is the distribution of $(X_1^*, \cdots, X_k^*)$ under π.

Let Z be the set of all integers. Let $\mathcal{Q}$ be the set of all subsets of A. We let $(A^\infty, \mathcal{Q}^\infty)$ be the measurable space consisting of A^∞, the set of all bilateral sequences $x = (x_i)_{i=-\infty}^{\infty}$ from A, and $\mathcal{Q}^\infty$, the usual product σ-field. We let $\{X_i\}_{i=-\infty}^{\infty}$ be the family of all projections from $A^\infty \to A$; thus, $X_i(x) = x_i$, $i \in Z$. We let $T_A: A^\infty \to A^\infty$ be the shift transformation; thus, if $x = (x_i)$ then $T_A x = x' = (x_i')$, where $x_i' = x_{i+1}$, $i \in Z$.

Received January 1978.

AMS 1970 *subject classifications.* Primary 28A65, 60G10.

Key words and phrases. Stationary aperiodic process, ergodic process, stationary coding, mixing invariant marginal, ergodic decomposition.

If μ is a probability measure on $\mathcal{Q}^\infty$, for $n = 1, 2, \cdots$, let μ_n be the measure on A^n which is the distribution of $(X_1, \cdots, X_n)$ under μ. If $(i_1, \cdots, i_m) \in A^m$, define $\pi(i_m|i_1, \cdots, i_{m-1}) = \pi(i_1, \cdots, i_m)/\pi_{m-1}(i_1, \cdots, i_{m-1})$ if $\pi_{m-1}(i_1, \cdots, i_{m-1}) > 0$. Otherwise, define $\pi(i_m|i_1, \cdots, i_{m-1}) = 0$. For $n > m$, let π_n be the probability measure on A^n such that $\pi_n(i_1, \cdots, i_n) = \pi_{m-1}(i_1, \cdots, i_{m-1})\pi(i_m|i_1, \cdots, i_{m-1})\pi(i_{m+1}|i_2, \cdots, i_m) \cdots \pi(i_n|i_{n-m+1}, \cdots, i_{n-1})$. Let $\hat{\pi}$ be the T_A-stationary probability measure on $\mathcal{Q}^\infty$ such that $\hat{\pi}_n = \pi_n$, $n = 1, 2, \cdots$. $\hat{\pi}$ is called the *Markov extension* of π. (Markov extensions were previously considered in [1], [2], [3], and [4].) $\hat{\pi}$ is an $(m-1)$-Markovian measure.

If $n < m$, let $A_n(\pi) = \{(i_1, \cdots, i_n) \in A^n: \pi_n(i_1, \cdots, i_n) > 0\}$. If $n \geq m$, let $A_n(\pi) = \{(i_1, \cdots, i_n) \in A^n: \pi(i_1, \cdots, i_m) > 0,\ \pi(i_2, \cdots, i_{m+1}) > 0, \cdots, \pi(i_{n-m+1}, \cdots, i_n) > 0\}$. Let $A_\infty(\pi) = \{x \in A^\infty: \pi(x_i, \cdots, x_{i+m-1}) > 0, i \in Z\}$. Note that $A_\infty(\pi)$ is a set of $\hat{\pi}$ probability 1.

Let $\{Z_i\}_{i=-\infty}^{\infty}$ be the process where $Z_i = (X_i, \cdots, X_{i+m-2})$. Then $\{Z_i\}$ is a stationary Markov process under $\hat{\pi}$. Let $\pi^*: A_{m-1}(\pi) \times A_{m-1}(\pi) \to [0, 1]$ be the stochastic matrix such that

$$\begin{aligned}\pi^*((i_1, \cdots, i_{m-1}), (i'_1, \cdots, i'_{m-1})) &= 0 \text{ if } (i'_1, \cdots, i'_{m-2}) \neq (i_2, \cdots, i_{m-1}) \\ &= \pi(i'_{m-1}|i_1, \cdots, i_{m-1}), \qquad \text{otherwise.}\end{aligned}$$

Then, $\hat{\pi}[Z_{i+1} = x|Z_i = x'] = \pi^*(x', x)$, $x, x' \in A_{m-1}(\pi)$.

We define π to be *mixing* if T_A is mixing with respect to $\hat{\pi}$, and define π to be *ergodic* if T_A is ergodic with respect to $\hat{\pi}$. It is easily seen that π is mixing if and only if $\{Z_i\}$ is a mixing process under $\hat{\pi}$, and π is ergodic if and only if $\{Z_i\}$ is an ergodic process under $\hat{\pi}$. It is well-known what it means for a stationary Markov process to be mixing or ergodic. Thus π is mixing if and only if π^* is regular (that is, some power of π^* has all positive elements), or if and only if π^* is irreducible and aperiodic. π is ergodic if and only if π^* is irreducible. The following lemma gives another characterization of these concepts.

LEMMA 1. (*a*) *π is ergodic if and only if there exists a probability measure μ on $\mathcal{Q}^\infty$, stationary and ergodic with respect to T_A, such that $\mu_m = \pi$.*

(*b*) *π is mixing if and only if there exists a probability measure μ on $\mathcal{Q}^\infty$, stationary and mixing with respect to T_A, such that $\mu_m = \pi$.*

PROOF. (a) If π is ergodic then $\hat{\pi}$ is ergodic. Conversely, suppose π is not ergodic. Then there exists $C \subset A^{m-1}$ such that

$$\begin{aligned}0 &< \pi\{(i_1, \cdots, i_m): (i_1, \cdots, i_{m-1}) \in C\} \\ &= \pi\{(i_1, \cdots, i_m): (i_1, \cdots, i_{m-1}) \in C, (i_2, \cdots, i_m) \in C\} < 1.\end{aligned}$$

Suppose μ is a T_A-stationary probability measure on $\mathcal{Q}^\infty$ such that $\mu_m = \pi$. Let $F = \cap_{i=-\infty}^{\infty}\{(X_i, \cdots, X_{i+m-2}) \in C\}$. Then $T_A F = F$ and $0 < \mu(F) < 1$, so μ is not ergodic.

(b) If π is mixing then $\hat{\pi}$ is mixing. Conversely, suppose π is not mixing. If π is not ergodic, by (a) there is no mixing μ on $\mathcal{Q}^\infty$ such that $\mu_m = \pi$. Thus we may suppose π is not mixing but ergodic. Let r be the period of π^*. There exist disjoint sets $C_0, \cdots, C_{r-1} \subset A^{m-1}$ such that $\Sigma_{i=0}^{r-1}\pi\{(i_1, \cdots, i_m): (i_1, \cdots, i_{m-1}) \in C_i, (i_2, \cdots, i_m) \in C_{i+1}\} = 1$. (In the preceding $C_r = C_0$.) Suppose μ on $\mathcal{Q}^\infty$ is stationary and $\mu_m = \pi$. Set $F = \cap_{i=-\infty}^{\infty} \cap_{j=0}^{r-1}\{(X_{ir+j}, \cdots, X_{ir+j+m-2}) \in C_j\}$. Then $F, T_A F, \cdots, T_A^{r-1}F$ are disjoint, $T_A^r F = F$, and $\mu(\cup_{j=0}^{r-1}T_A^j F) = 1$. Thus μ cannot be mixing since $\lim_{i\to\infty} \mu(T_A^i F \cap F)$ does not exist.

We fix now a standard measurable space $(\Omega, \mathcal{F})$ and a one-to-one bimeasurable mapping T of Ω onto itself. Let $\mathcal{P}_s$ be the family of all probability measures μ on $\mathcal{F}$ such that T is an aperiodic measure-preserving transformation on $(\Omega, \mathcal{F}, \mu)$. Let $\mathcal{P}_e$ be the family of all probability measures μ on $\mathcal{F}$ such that T is an aperiodic ergodic measure-preserving transformation on $(\Omega, \mathcal{F}, \mu)$. We assume $\mathcal{P}_s \neq \phi$. Then by ergodic decomposition theory, $\mathcal{P}_e \neq \phi$.

By a partition P of Ω we mean a finite measurable partition $P = \{P^j: j \in E\}$. (Thus the index set E is finite.) Given $P = \{P^j: j \in E\}$ and $i \in Z$, T^iP is the partition $T^iP = \{T^iP^j: j \in E\}$. If $m \leq n$, $\vee_{i=m}^{n} T^{-i}P$ is the partition indexed by E^{n-m+1} given by

$$\{T^{-m}P^{j_1} \cap T^{-(m+1)}P^{j_2} \cap \cdots \cap T^{-n}P^{j_{n-m+1}}: (j_1, \cdots, j_{n-m+1}) \in E^{n-m+1}\}.$$

If $\mu \in \mathcal{P}_s$, and $P = \{P^j: j \in E\}$ is a partition, then $\text{dist}_\mu P$ is the probability measure λ on E such that $\lambda(j) = \mu(P^j), j \in E$.

The following theorem is an easy consequence of Theorem 1.1 of [1].

THEOREM 1. *Let $\mu \in \mathcal{P}_e$. Let π be mixing. Then there exists a partition $P = \{P^j: j \in A\}$ of Ω such that* $\text{dist}_\mu(\vee_0^{m-1}T^{-i}P) = \pi$.

REMARK. We may translate this theorem into the context of coding one stationary process onto another to achieve a given marginal distribution. For, suppose $\{X_i\}$ is an aperiodic ergodic $(S, \mathbb{S})$-valued process defined on some probability space $(\Omega', \mathcal{F}', \lambda')$. In Theorem 1, take $(\Omega, \mathcal{F})$ to be $(S^\infty, \mathbb{S}^\infty)$ with T the shift on S^∞. Take μ on $\mathbb{S}^\infty$ to be the distribution of $\{X_i\}$. Let P be the partition of S^∞ given by Theorem 1. Define $f: S^\infty \to A$ so that $f(x) = j$ if and only if $x \in P^j$, where $x \in S^\infty, j \in A$. Then f codes $\{X_i\}$ onto an A-valued process $\{Y_i\}$ such that the distribution of $(Y_0, \cdots, Y_{m-1})$ is $\text{dist}_\mu(\vee_0^{m-1}T^{-i}P)$, which is π. In this paper we will describe our coding results in terms of the existence of a certain partition P rather than in terms of a certain coding function f, since it makes the notation simpler. But the reader may easily translate any theorem we give into a theorem about coding one stationary process onto another to achieve a desired marginal distribution.

PROOF OF THEOREM 1. Let $H(\hat{\pi})$ denote the entropy of $\hat{\pi}$. From [3], it may be calculated as follows: for each $i \in A_{m-1}(\pi)$, let p_i be the probability vector $p_i = (\pi^*(i,j): j \in A_{m-1}(\pi))$. Let $H(p_i)$ be the entropy of p_i. Then $H(\hat{\pi}) = \Sigma_i \pi_{m-1}(i)H(p_i)$. If $H(\hat{\pi}) = 0$ then $H(p_i) = 0$ for all i. This means every row of π^*

has only one nonzero element. Thus every power of π^* has this form. Since π^* is regular, this implies $A_{m-1}(\pi)$ has only one element. Then, for some $a^* \in A$, $\pi(a^*, \cdots, a^*) = 1$. In this case, Theorem 1 follows trivially. (Take $P^j = \Omega, j = a^*$; $P^j = \phi$, otherwise.) Thus, we can assume $H(\hat{\pi}) > 0$. Find a partition Q of Ω such that the entropy $H_\mu(Q, T)$ of Q with respect to T is less than $H(\hat{\pi})$ and Q contains a set of irrational measure. Let $\mathcal{F}_Q$ be the sub-σ-field of $\mathcal{F}$ generated by $\{T^iQ: i \in Z\}$. Let μ_Q be the restriction of μ to $\mathcal{F}_Q$. Let T_Q be the automorphism of $(\Omega, \mathcal{F}_Q, \mu_Q)$ induced by T. T_Q is ergodic and aperiodic. The entropy $H(T_Q)$ of T_Q is $H_\mu(Q, T) < H(\hat{\pi})$. Thus by Theorem 1.1 of [1], there exists a $\mathcal{F}_Q$-measurable partition $P = \{P^j: j \in A\}$ such that $\text{dist}_\mu(\vee_0^{m-1}T_Q^{-i}P) = \pi$ and $\{T^iP: i \in Z\}$ generate $\mathcal{F}_Q$.

Theorem 1 allows us to answer a question posed at the end of [3]. Suppose the invariant distribution π satisfies $\pi(x) > 0$, $x \in A^m$. It was asked at the end of [3] whether there is an ergodic μ on $\mathcal{Q}^\infty$ of entropy zero such that $\mu_m = \pi$. Since π is mixing (π^* is clearly regular), choose $(\Omega, \mathcal{F}, \mu)$ and $T: \Omega \to \Omega$ so that $\mu \in \mathcal{P}_e$ and $H_\mu(T) = 0$. Find the P given by Theorem 1. Using P define an ergodic A-valued process $\{X_i\}$ with the distribution of $(X_0, \cdots, X_{m-1})$ equal to π. $\{X_i\}$ has entropy 0, so let μ on $\mathcal{Q}^\infty$ be the distribution of $\{X_i\}$.

We note that Theorem 1 becomes false if we require only that π be ergodic. For by Lemma 1, if μ is mixing, so is π. So if we want to extend Theorem 1 to a π which is ergodic but not mixing, we will have to require that μ be ergodic but not mixing. This question is settled in Theorem 4. If π is not ergodic, then by Lemma 1 there is no ergodic μ which yields the marginal π. Since we want to code at least ergodic μ, there is no point in considering nonergodic π.

Another possible extension of Theorem 1 would be to find P so that $\text{dist}_\mu(\vee_0^{m-1}T^{-i}P) = \pi$ for *all* $\mu \in \mathcal{P}_s$. This result we will obtain in Theorem 3. To get Theorem 3, we will need a sharpening of Theorem 1, namely Theorem 2. Also, from Theorem 2 we can get Theorem 1 as a corollary without having to use Theorem 1.1 of [1] to prove Theorem 1 as we did. This is of interest because the proof of Theorem 1.1 of [1] is very difficult. The proof of Theorem 2 we will give will be an easy application of some lemmas from [1].

DEFINITIONS. If π_1 and π_2 are probability distributions on A^m define $|\pi_1 - \pi_2| = \sum_{i \in A^m}|\pi_1(i) - \pi_2(i)|$. If $P = \{P^j: j \in E\}$ and $Q = \{Q^j: j \in E\}$ are two partitions of Ω and $\mu \in \mathcal{P}_s$, let $|P - Q|_\mu = \frac{1}{2}\sum_j \mu(P^j \Delta Q^j)$. Define $\mathcal{M}(\pi)$ to be the set of all invariant probability measures $\tilde{\pi}$ on A^m such that $\tilde{\pi}$ is absolutely continuous with respect to π.

If $n > m$, and $Q = (i_1, \cdots, i_n) \in A^n$, define μ_Q to be the probability measure on A^m such that $\mu_Q(x) = (n - m + 1)^{-1}\sum_{j=1}^{n-m+1}\delta((i_j, \cdots, i_{j+m-1}), x)$, $x \in A^m$. ($\delta(i,j)$ denotes the Kronecker delta: $\delta(i,j) = 1$ if $i = j$; $\delta(i,j) = 0$, $i \neq j$.)

If π is mixing, then since π^* is regular we may find a positive integer $L > m$ and for each $j \geq L$ a map $\tau_j: A^{m-1} \times A^{m-1} \to A^j$ such that if $b, b' \in A_{m-1}(\pi)$ then $b\tau_j(b, b')b' \in A(\pi)_{2(m-1)+j}$.

The following lemma we will need later. It is a combination of Lemmas 2.1 and 2.3 of [1].

LEMMA 2. *Let* $0 < \varepsilon < 1$. *Let* π *be mixing. There exists* $h(\varepsilon) > 0$ *and a positive integer* $N(\varepsilon) > 2m$ *such that given* $n \geqslant N(\varepsilon)$ *and any* $Q^* \in A_{m-1}(\pi)$, *there exist finitely many sequences* $Q_1, Q_2, \cdots, Q_K$ *from* $A_n(\pi)$, *beginning and ending with* Q^*, *for which the following holds*:

If $\tilde{\pi} \in \mathfrak{M}(\pi)$ *and* $|\tilde{\pi} - \pi| < h(\varepsilon)$, *and if* $1 > \varepsilon' \geqslant \varepsilon$, *there exist nonnegative numbers* $\alpha_1, \alpha_2, \cdots, \alpha_K$, *such that* $\sum_{i=1}^{K}\alpha_i = \varepsilon'$ *and* $\pi = (1 - \varepsilon')\tilde{\pi} + \sum_{i=1}^{K}\alpha_i\mu_{Q_i}$.

LEMMA 3. *Let* π *be mixing. Let* $\mu \in \mathcal{P}_e$. *Let* $0 < \varepsilon < 1$. *Then if* $P = \{P^j: j \in A\}$ *is a partition of* Ω *satisfying* $\operatorname{dist}_\mu(\vee_0^{m-1}T^{-i}P) \in \mathfrak{M}(\pi)$ *and* $|\operatorname{dist}_\mu(\vee_0^{m-1}T^{-i}P) - \pi| < h(\varepsilon/2)$, *there exists a partition* $\tilde{P} = \{\tilde{P}^j: j \in A\}$ *such that* $\operatorname{dist}_\mu(\vee_0^{m-1}T^{-i}\tilde{P}) = \pi$ *and* $|P - \tilde{P}|_\mu < \varepsilon$.

PROOF. Let $\bar{\pi} = \operatorname{dist}_\mu(\vee_0^{m-1}T^{-i}P)$. By a strong form of the Rohlin-Kakutani Theorem [7, page 22], we may find $F \in \mathcal{F}$ and positive integers N_1 and N_2 such that:

(a) $\{T^iF: i = 0, \cdots, N_1 + N_2 + 2L - 1\}$ are disjoint.

(b) $\mu(E \cap F) = \mu(E)\mu(F)$, $E \in \vee_0^{(N_1+N_2+2L)m}T^{-i}P$.

(c) $$\frac{N_1 + N_2 + 2L - (N_1 + N_2 - 2m + 2)\mu\left(\cup_0^{N_1+N_2+2L-1}T^jF\right)}{N_1 + N_2 + 2L - (N_1 - m + 1)\mu\left(\cup_0^{N_1+N_2+2L-1}T^jF\right)} < \frac{h(\varepsilon/2) - |\bar{\pi} - \pi|}{2}.$$

(d) $\mu(\cup_0^{N_1+N_2+2L-1}T^jF) < 1$.

(e) $\left(\frac{N_1 - m + 1}{N_1 + N_2 + 2L}\right)\mu(\cup_0^{N_1+N_2+2L-1}T^jF) > \varepsilon/2$.

(f) $\frac{2L + N_1}{N_1 + N_2 + 2L} < \varepsilon$.

(g) $N_1 \geqslant N(\varepsilon/2)$.

Fix $Q^* \in A_{m-1}(\pi)$. By Lemma 2, pick $Q_1, \cdots, Q_K \in A_{N_1}(\pi)$ beginning and ending with Q^* so that $|\tilde{\pi} - \pi| < h(\varepsilon/2)$, $\tilde{\pi} \in \mathfrak{M}(\pi)$ and $1 > \varepsilon' \geqslant \varepsilon/2$ imply the existence of nonnegative $\alpha_1, \cdots, \alpha_K$ such that $\alpha_1 + \cdots + \alpha_K = \varepsilon'$ and $\pi = (1 - \varepsilon')\tilde{\pi} + \sum_{i=1}^{K}\alpha_i\mu_{Q_i}$.

Let $\{X_i'\}_{i=-\infty}^{\infty}$ be the collection of maps from Ω to A such that $X_i'(\omega) = j$ if and only if $T^i\omega \in P^j$. We will define a map $\Psi: \Omega \to A^\infty$ such that $T_A \circ \Psi = \Psi \circ T$. At the end of the proof we will have to adjust Ψ to obtain a map $\Psi': \Omega \to A^\infty$ such that $T_A \circ \Psi' = \Psi' \circ T$. Then $\tilde{P}$ will be the partition $\tilde{P}^j = \{\omega: \Psi'(\omega)_0 = j\}$.

Let $\omega \in \Omega$. We wish to define $\Psi(\omega)$. Say $\Psi(\omega) = (y_i)_{i=-\infty}^{\infty}$, where now we have to define each $y_i \in A$. Let $\{i_k\}_{k=-\infty}^{\infty}$ be a sequence of integers such that

(h) $\cdots < i_k < i_{k+1} < \cdots$

(i) $\{i_k\} = \{i \in Z: T^i\omega \in F\}$.

If $m \leqslant n$ are integers, we let $[m, n] = \{m, m + 1, \cdots, n\}$. We need to define

$(y_i: i \in [i_k, i_{k+1} - 1])$ for each $k \in Z$. To do this, for each k divide $[i_k, i_{k+1} - 1]$ into 4 parts: $A_k = [i_k, i_k + L - 1]$, $B_k = [i_k + L, i_k + L + N_1 - 1]$, $C_k = [i_k + L + N_1, i_k + 2L + N_1 - 1]$, $D_k = [i_k + 2L + N_1, i_{k+1} - 1]$. Note that A_k, B_k, C_k consist of L, N_1, L consecutive integers, respectively. D_k contains at least N_2 consecutive integers since $i_{k+1} - i_k \geq N_1 + N_2 + 2L$. First, for each k we define $\Psi(\omega)$ over B_k and D_k. For $k \in Z$, define $(y_i: i \in D_k)$ to be $(X_i'(\omega): i \in D_k)$, and define $(y_i: i \in B_k)$ to be Q_1. Now we need to define $\Psi(\omega)$ over each A_k and C_k. For each k, let $b_1^k, b_2^k, b_3^k, b_4^k$ be the following 4 blocks from A^{m-1}:

b_1^k = last $m - 1$ entries of $\Psi(\omega)$ in D_{k-1}.
b_2^k = first $m - 1$ entries of $\Psi(\omega)$ in B_k.
b_3^k = last $m - 1$ entries of $\Psi(\omega)$ in B_k.
b_4^k = first $m - 1$ entries of $\Psi(\omega)$ in D_k.

Since $\Psi(\omega)$ was defined over the B_k's and D_k's first, $b_1^k, \cdots, b_4^k$ are defined. (In fact, $b_2^k = b_3^k = Q^*$.) For each $k \in Z$, define $(y_i: i \in A_k)$ to be $\tau_L(b_1^k, b_2^k)$ and define $(y_i: i \in C_k)$ to be $\tau_L(b_3^k, b_4^k)$. This completes the definition of $\Psi(\omega) = (y_i)$.

Since $\bar{\pi} \in \mathfrak{M}(\pi)$, with probability one $(X_i'(\omega)) \in A_\infty(\pi)$ and so with probability one, $\Psi(\omega) \in A_\infty(\pi)$.

If $m \leq n$ are integers let $F_m^n = \cup_m^n T^i F$. Let $\{Y_i\}_{i=-\infty}^{\infty}$ be the collection of maps from $\Omega \to A$ such that $\{Y_i(\omega)\} = \Psi(\omega)$. We define probability distributions π_1, π_2, π_3 on A^m as follows:

$$\pi_1(x) = \mu\left[\{(Y_0, \cdots, Y_{m-1}) = x\} \cap F_{N_1+2L}^{N_1+2L+N_2-m}\right] / \mu\left[F_{N_1+2L}^{N_1+2L+N_2-m}\right], \quad x \in A^m.$$

$$\pi_2(x) = \mu\left[\{(Y_0, \cdots, Y_{m-1}) = x\} \cap \left(F_L^{L+N_1-m}\right)^c\right] / \mu\left[\left(F_L^{L+N_1-m}\right)^c\right], \quad x \in A^m.$$

$$\pi_3(x) = \mu\left[\{(Y_0, \cdots, Y_{m-1}) = x\} \cap F_L^{L+N_1-m}\right] / \mu\left[F_L^{L+N_1-m}\right], \quad x \in A^m.$$

($(F_L^{L+N_1-m})^c$ denotes the complement of $F_L^{L+N_1-m}$.) Note that on $F_{N_1+2L}^{N_1+2L+N_2-m}$, we have $(Y_0, \cdots, Y_{m-1}) = (X_0', \cdots, X_{m-1}')$. Using this fact and (b), we have $\pi_1 = \bar{\pi}$. Now $(F_L^{L+N_1-m})^c$ is the disjoint union of $F_{N_1+2L}^{N_1+2L+N_2-m}$ and some set G. We have

$$\mu(G) = 1 - (N_1 - m + 1)\mu(F) - (N_2 - m + 1)\mu(F)$$

$$= 1 - \left(\frac{N_1 + N_2 - 2m + 2}{N_1 + N_2 + 2L}\right)\mu\left(F_0^{N_1+N_2+2L-1}\right).$$

Also,

$$\mu\left[\left(F_L^{L+N_1-m}\right)^c\right] = 1 - \left(\frac{N_1 - m + 1}{N_1 + N_2 + 2L}\right)\mu\left(F_0^{N_1+N_2+2L-1}\right).$$

Let π_G be the probability measure on A^m such that

$$\pi_G(x) = \mu\left[\{(Y_0, \cdots, Y_{m-1}) = x\} \cap G\right] / \mu(G), \qquad x \in A^m.$$

We have $\pi_2 = (1-\alpha)\pi_1 + \alpha\pi_G$, where $\alpha = \mu(G)/\mu[(F_L^{L+N_1-m})^c]$. By (c), $\alpha < \frac{1}{2}[h(\varepsilon/2) - |\bar{\pi} - \pi|]$. Thus, $|\pi_2 - \pi_1| \leq \alpha|\pi_G - \pi_1| < h(\varepsilon/2) - |\bar{\pi} - \pi|$, and since $\pi_1 = \bar{\pi}$, we have $|\pi - \pi_2| < h(\varepsilon/2)$. Now π_2 is absolutely continuous with respect to π since for μ-almost all ω, $(Y_i(\omega), \cdots, Y_{i+m-1}(\omega)) \in A_m(\pi)$ for all $i \in Z$.

We now argue that $\pi_3 = \mu_{Q_1}$. For $\omega \in \Omega$, let $\{i_k\}_{k=-\infty}^{\infty}$ be the sequence chosen earlier satisfying $(h) - (i)$. Let $B_k^* = [i_k + L, i_k + L + N_1 - m]$, $k \in Z$. Then, $\cup_{k=-\infty}^{\infty} B_k^* = \{i \in Z: T^i\omega \in F_L^{L+N_1-m}\}$. Thus, for μ-almost all ω, and each $x \in A^m$,

$$\pi_3(x) = \lim_{N\to\infty} \frac{\Sigma_{k=-N}^{N}\Sigma_{i\in B_k^*}\delta((Y_i(\omega), \cdots, Y_{i+m-1}(\omega)), x)}{(2N+1)(N_1 - m + 1)} = \mu_{Q_1}(x),$$

since $\Psi(\omega)$ is Q_1 over each B_k, $k \in Z$. Now μ_{Q_1} is invariant because Q_1 starts and ends with the same sequence in A^{m-1}. We have $\operatorname{dist}(Y_0, \cdots, Y_{m-1})$, the probability distribution of $(Y_0, \cdots, Y_{m-1})$ on A^m, given as follows:

$$\operatorname{dist}(Y_0, \cdots, Y_{m-1}) = \mu\left[F_L^{L+N_1-m}\right]\pi_3 + \mu\left[\left(F_L^{L+N_1-m}\right)^c\right]\pi_2.$$

Now $\operatorname{dist}(Y_0, \cdots, Y_{m-1})$ and π_3 are invariant and $\mu[(F_L^{L+N_1-m})^c] > 0$ by (d), so π_2 is invariant. Thus, $\pi_2 \in \mathfrak{M}(\pi)$ and $|\pi_2 - \pi| < h(\varepsilon/2)$. Setting $\varepsilon' = \mu[F_L^{L+N_1-m}]$, we have $1 > \varepsilon' \geq \varepsilon/2$, by (e) and (d). Thus, pick non-negative $\alpha_1, \cdots, \alpha_K$ such that $\alpha_1 + \cdots \alpha_K = \varepsilon'$ and $\pi = (1-\varepsilon')\pi_2 + \Sigma_{i=1}^K \alpha_i \mu_{Q_i}$. Since T is aperiodic on $(\Omega, \mathfrak{F}, \mu)$, μ must be nonatomic. Thus, we may pick $\{F_1, \cdots, F_K\}$, a partition of F, such that $\mu[\cup_{j=L}^{L+N_1-m} T^j F_i] = \alpha_i$, $i = 1, \cdots, K$. Let $(F_i)_L^{L+N_1-m} = \cup_{j=L}^{L+N_1-m} T^j F_i$, $i = 1, \cdots, K$. We now define $\Psi': \Omega \to A^\infty$. Let $\omega \in \Omega$; we define $(y_i')_{i=-\infty}^{\infty}$ where $(y_i') = \Psi'(\omega)$. Let $\{i_k\}_{k=-\infty}^{\infty}$ be given by (h) − (i). For each k, define $(y_i': i \in B_k)$ to be Q_j where j is the unique integer in the range $1 \leq j \leq K$ such that $T^{i_k}\omega \in F_j$. Define $y_i' = Y_i(\omega)$, $i \notin \cup_k B_k$.

Let $\{Y_i'\}_{i=-\infty}^{\infty}$ be the collection of maps from Ω to A such that $\{Y_i'(\omega)\} = \Psi'(\omega)$, $\omega \in \Omega$. Define the following probability distributions on A^m:

$$\pi_i'(x) = \mu\left[\{(Y_0', \cdots, Y_{m-1}') = x\} \cap (F_i)_L^{L+N_1-m}\right]/\alpha_i, \quad i = 1, \cdots, K, x \in A^m;$$

$$\pi'(x) = \mu\left[\{(Y_0', \cdots, Y_{m-1}') = x\} \cap \left(F_L^{L+N_1-m}\right)^c\right]/(1-\varepsilon'), \ x \in A^m.$$

Now if $T^i\omega \notin F_L^{L+N_1-m}$, then $(Y_i'(\omega), \cdots, Y_{i+m-1}'(\omega)) = (Y_i(\omega), \cdots, Y_{i+m-1}(\omega))$. Thus $\pi' = \pi_2$. Also $\pi_i' = \mu_{Q_i}$, $i = 1, \cdots, K$. (Argue as we did to show $\pi_3 = \mu_{Q_1}$.)

Let $\tilde{P}$ be the partition such that $\tilde{P}^j = \{Y_0' = j\}, j \in A$. Then $\operatorname{dist}_\mu(\vee_0^{m-1} T^{-i}\tilde{P}) = \operatorname{dist}(Y_0', \cdots, Y_{m-1}') = (1-\varepsilon')\pi' + \Sigma_{i=1}^K \alpha_i \pi_i' = (1-\varepsilon')\pi_2 + \Sigma_{i=1}^K \alpha_i \mu_{Q_i} = \pi$.

We now estimate $|\tilde{P} - P|_\mu$. For μ-almost all ω, $|\tilde{P} - P|_\mu = \mu[Y_0' \neq X_0'] =$

$$\lim_{N\to\infty} \frac{\Sigma_{k=-N}^{N}\Sigma_{j=i_k}^{i_{k+1}-1}\left[1 - \delta(Y_j'(\omega), X_j'(\omega))\right]}{\Sigma_{k=-N}^{N}(i_{k+1} - i_k)}.$$

Since $Y_j'(\omega) = X_j'(\omega)$ for $j \in D_k$, this limit is no bigger than $(2L + N_1)/(2L + N_1 + N_2)$, which is less than ε by (f).

LEMMA 4. *Let π be mixing. Let $\mu \in \mathcal{P}_e$. Let $k = 2(m-1) + L$. Let $P = \{P^j: j \in A\}$ be a partition of Ω. Let $\varepsilon > 0$. If $|\text{dist}_\mu(\vee_0^{m-1}T^{-i}P) - \pi| < \varepsilon/(2k+1)$, there exists a partition $\tilde{P} = \{\tilde{P}^j: j \in A\}$ such that $\text{dist}_\mu(\vee_0^{m-1}T^{-i}\tilde{P}) \in \mathfrak{M}(\pi)$ and $|(\vee_0^{m-1}T^{-i}\tilde{P}) - (\vee_0^{m-1}T^{-i}P)|_\mu < \varepsilon$.*

PROOF. Let $\{Y_i\}_{i=-\infty}^{\infty}$ be the collection of mappings from $\Omega \to A$ such that $Y_i(\omega) = j$ if and only if $T^i\omega \in P^j$. Let $S = \{(Y_0, \cdots, Y_{m-1}) \in A_m(\pi)\}$. Let $F = \cap_{j=-k}^{k} T^{-j}S$. We will define a certain map $\Phi: \Omega \to A^\infty$ such that $\Phi \circ T = T_A \circ \Phi$. Then $\tilde{P}$ will be the partition $\tilde{P}^j = \{\omega: \Phi(\omega)_0 = j\}$. Let $\omega \in \Omega$. We will define a sequence $(x_i)_{i=-\infty}^{\infty}$ to be $\Phi(\omega)$. Let $B(\omega) = \{i \in Z: T^i\omega \notin F\}$. Let $D(\omega) = \{i \in B(\omega): i - 1 \notin B(\omega)\}$. For each $j \in D(\omega)$, let $C_j(\omega) = [j, n]$, where $C_j(\omega) \subset B(\omega)$ and $n + 1 \notin B(\omega)$. We have $B(\omega) = \cup_{j \in D(\omega)} C_j(\omega)$, a disjoint union. Define $x_i = Y_i(\omega)$, $i \notin B(\omega)$. For each $j \in D(\omega)$, we define $(x_i: i \in C_j(\omega))$. Note that by choice of F, $C_j(\omega)$ has at least k elements. Define $(x_j, \cdots, x_{j+m-2}) = (Y_j(\omega), \cdots, Y_{j+m-2}(\omega)) = b$ and $(x_{n-m+2}, \cdots, x_n) = (Y_{n-m+2}(\omega), \cdots, Y_n(\omega)) = b'$. Define $(x_{j+m-1}, \cdots, x_{n-m+1})$ to be $\tau_r(b, b')$, where $r = (n - m + 1) - (j + m - 1) + 1$. This completes the definition of $\Phi(\omega)$.

Let $\{X_i'\}_{i=-\infty}^{\infty}$ be the collection of mappings from $\Omega \to A$ such that $\{X_i'(\omega)\} = \Phi(\omega)$, $\omega \in \Omega$. Let $\tilde{P}$ be the partition such that $\tilde{P}^j = \{X_0' = j\}$. By construction, $\{X_i'(\omega)\} \in A_\infty(\pi)$ for every ω. Therefore $\text{dist}_\mu(\vee_0^{m-1}T^{-i}\tilde{P}) \in \mathfrak{M}(\pi)$. Note that if $T^i\omega \in F$, then $(X_i', \cdots, X_{i+m-1}') = (Y_i, \cdots, Y_{i+m-1})$. Thus $|(\vee_0^{m-1}T^{-i}P) - (\vee_0^{m-1}T^{-i}\tilde{P})|_\mu = \mu[(X_0', \cdots, X_{m-1}') \neq (Y_0, \cdots, Y_{m-1})] \leq \mu(F^c) \leq (2k+1)\mu[(Y_0, \cdots, Y_{m-1}) \notin A_m(\pi)] < (2k+1)[\varepsilon/(2k+1) + \pi[(A_m(\pi))^c]] = \varepsilon$.

Lemmas 3 and 4 give us the following.

THEOREM 2. *Let π be mixing. Let $\mu \in \mathcal{P}_e$. Given $\varepsilon > 0$, there exists $\delta > 0$ such that if $P = \{P^j: j \in A\}$ is a partititon of Ω and $|\text{dist}_\mu(\vee_0^{m-1}T^{-i}P) - \pi| < \delta$, then there exists a partition $\tilde{P} = \{\tilde{P}^j: j \in A\}$ such that $|\tilde{P} - P|_\mu < \varepsilon$ and $\text{dist}_\mu(\vee_0^{m-1}T^{-i}\tilde{P}) = \pi$. δ depends only on ε and π and not on P, T, or $(\Omega, \mathcal{F}, \mu)$.*

Now, given any $\delta > 0$, and any invariant π on A^m, there exists P such that $|\text{dist}_\mu(\vee_0^{m-1}T^{-i}P) - \pi| < \delta$ [5, Lemma 5, page 22]. Thus Theorem 2 implies Theorem 1.

We are now in a position to find a P in Theorem 1 which will yield the marginal distribution π simultaneously for all $\mu \in \mathcal{P}_s$. To do this we need the following ergodic decomposition theorem.

LEMMA 5. *There exists a family $\{\mu_\omega: \omega \in \Omega\} \subset \mathcal{P}_e$ such that*

(a) *For any $E \in \mathcal{F}$, the map $\omega \to \mu_\omega(E)$ from Ω to the real line is measurable.*

(b) *For any $\mu \in \mathcal{P}_s$ and $E \in \mathcal{F}$, $\mu(E) = \int_\Omega \mu_\omega(E)d\mu(\omega)$.*

(c) *For any $\mu \in \mathcal{P}_e$, $\mu\{\omega: \mu_\omega = \mu\} = 1$.*

SKETCH OF PROOF. Let T_Ω be the shift on $(\Omega^\infty, \mathcal{F}^\infty)$. We have a measurable map α: $(\Omega, \mathcal{F}) \to (\Omega^\infty, \mathcal{F}^\infty)$ such that $\alpha(\omega) = (T^i\omega)_{i=-\infty}^{\infty}$. Let $\mathcal{P}_e(\Omega^\infty)$ and $\mathcal{P}_s(\Omega^\infty)$ be respectively the sets of T_Ω-ergodic, stationary, aperiodic probability measures and T_Ω-aperiodic, stationary probability measures on $\mathcal{F}^\infty$. α carries $\mathcal{P}_e$ into (but not onto) $\mathcal{P}_e(\Omega^\infty)$ and $\mathcal{P}_s$ into $\mathcal{P}_s(\Omega^\infty)$. By [6], there is a family $\{\mu_x: x \in \Omega^\infty\} \subset \mathcal{P}_e(\Omega^\infty)$ which satisfies the obvious analogues of (a)–(c) for the space $(\Omega^\infty, \mathcal{F}^\infty)$. Let $\mathcal{C}$ be a countable subfield of $\mathcal{F}$ which generates $\mathcal{F}$. Let $F = \{\omega \in \Omega: \mu_{\alpha(\omega)}[x \in \Omega^\infty: x_0 \in C$, some $x_i \notin T^iC] = 0, C \in \mathcal{C}\}$. Then $F \in \mathcal{F}$. Define $\{\mu_\omega: \omega \in \Omega\}$ as follows: if $\omega \notin F$, define μ_ω to be some fixed element of $\mathcal{P}_e$. If $\omega \in F$, define μ_ω so that $\mu_\omega(E) = \mu_{\alpha(\omega)}\{x \in \Omega^\infty: x_0 \in E\}$, $E \in \mathcal{F}$. We omit the tedious verification that $\{\mu_\omega: \omega \in \Omega\} \subset \mathcal{P}_e$ and (a)–(c) hold.

Fix for the rest of the paper the family $\{\mu_\omega: \omega \in \Omega\} \subset \mathcal{P}_e$ such that (a)–(c) of Lemma 5 hold. For each $\mu \in \mathcal{P}_e$, let $A_\mu = \{\omega: \mu_\omega = \mu\}$. We have $\mu(A_\mu) = 1$ and $\mu'(A_\mu) = 0$ for $\mu' \in \mathcal{P}_e$, $\mu' \neq \mu$. $\{A_\mu: \mu \in \mathcal{P}_e\}$ is an uncountable partition of Ω.

LEMMA 6. *Let π be mixing. Let $\varepsilon > 0$. Let $\delta > 0$ be given by Theorem 2. Suppose $|\mathrm{dist}_\mu(\vee_0^{m-1}T^{-i}P) - \pi| < \delta$ for all $\mu \in \mathcal{P}_e$. Then for any $\eta > 0$, there exists $\tilde{P}$ such that $|\tilde{P} - P|_\mu < \varepsilon$ and $|\mathrm{dist}_\mu(\vee_0^{m-1}T^{-i}\tilde{P}) - \pi| < \eta$ for all $\mu \in \mathcal{P}_e$.*

PROOF. Let $\mathcal{C} \subset \mathcal{F}$ be a countable field which generates $\mathcal{F}$. Let $P_1, P_2, \cdots$ be an enumeration of all A-indexed partitions whose sets come from $\mathcal{C}$. Fix $\mu \in \mathcal{P}_e$. By Theorem 2, find P' such that $|P' - P|_\mu < \varepsilon$ and $\mathrm{dist}_\mu(\vee_0^{m-1}T^{-i}P') = \pi$. We may pick P_j so that $|P^j - P'|_\mu$ is so small that $|P_j - P|_\mu < \varepsilon$ and $|\mathrm{dist}_\mu(\vee_0^{m-1}T^{-i}P_j) - \pi| < \eta$. Let $A_n = \{\omega \in \Omega: |P_n - P|_{\mu_\omega} < \varepsilon, |\mathrm{dist}_{\mu_\omega}(\vee_0^{m-1}T^{-i}P_n) - \pi| < \eta\}$. Let $B_n = \{\omega: \omega \in A_n, \omega \notin \cup_{i=0}^{n-1}A_i\}$, $n = 1, 2, \cdots$. The B_n's partition Ω. If for $n = 1, 2, \cdots$, $P_n = \{P_n^j: j \in A\}$, define $\tilde{P} = \{\tilde{P}^j: j \in A\}$ so that $\tilde{P}^j = \cup_{n=1}^{\infty}(P_n^j \cap B_n)$. Let $\mu \in \mathcal{P}_e$. Then $A_\mu \subset B_n \subset A_n$ for a unique n. This implies that for each $j \in A$, $\tilde{P}^j \cap A_\mu = P_n^j \cap A_\mu$, for that unique n. Consequently, since $\mu(A_\mu) = 1$, we have $|\tilde{P} - P|_\mu = |P_n - P|_\mu < \varepsilon$ and $|\mathrm{dist}_\mu(\vee_0^{m-1}T^{-i}\tilde{P}) - \pi| = |\mathrm{dist}_\mu(\vee_0^{m-1}T^{-i}P_n) - \pi| < \eta$.

LEMMA 7. *Given $\delta > 0$, there exists P such that $|\mathrm{dist}_\mu(\vee_0^{m-1}T^{-i}P) - \pi| < \delta$ for all $\mu \in \mathcal{P}_e$.*

PROOF. A method of proof analogous to that of Lemma 6 will work here.

THEOREM 3. *Let π be mixing. There exists P such that $\mathrm{dist}_\mu(\vee_0^{m-1}T^{-i}P) = \pi$ for all $\mu \in \mathcal{P}_s$.*

PROOF. Pick positive numbers $\{\varepsilon_i\}_{i=1}^{\infty}$ so that $\Sigma_i\varepsilon_i < \infty$. For each ε_i pick the corresponding δ_i given by Theorem 2. We can assume $\delta_i \to 0$. By Lemmas 6 and 7 find a sequence of partitions $\{P_i\}_{i=1}^{\infty}$ such that for all $i = 1, 2, \cdots$, and all $\mu \in \mathcal{P}_e$, $|\mathrm{dist}_\mu(\vee_{j=0}^{m-1}T^{-j}P_i) - \pi| < \delta_i$ and $|P_{i+1} - P_i|_\mu < \varepsilon_i$. Letting $P_i = \{P_i^j: j \in A\}$, we have $\Sigma_{i=1}^{\infty}\mu(P_{i+1}^j \Delta P_i^j) < \infty$ for all $j \in A$, $\mu \in \mathcal{P}_e$. Therefore if we set $E^j = \limsup_{i\to\infty} P_i^j$, it follows that $\lim_{i\to\infty}\mu(P_i^j \Delta E^j) = 0$, $j \in A$, $\mu \in \mathcal{P}_e$. Letting

$A = \{a_1, \cdots, a_k\}$ and defining $P = \{P^j : j \in A\}$ to be the partition such that

$$P^{a_s} = \{\omega : \omega \in E^{a_s}, \omega \notin E^{a_t}, 1 \leq t \leq s-1\}, 1 \leq s \leq k-1,$$

we have $|P_i - P|_\mu \to 0$ for all $\mu \in \mathcal{P}_e$. We must have $\text{dist}_\mu(\vee_0^{m-1} T^{-j}P) = \pi$ for all $\mu \in \mathcal{P}_e$. By Lemma 5, this must hold also for all $\mu \in \mathcal{P}_s$.

DEFINITION. Let N be a positive integer. We say $\mu \in \mathcal{P}_s$ is *N-decomposable* if there exists $F \in \mathcal{F}$ such that $F, TF, \cdots, T^{N-1}F$ are disjoint, $T^N F = F$, and $\mu(\cup_0^{N-1} T^i F) = 1$.

LEMMA 8. *Let π be ergodic but not mixing. Let r be the period of π^*. Let $\mu \in \mathcal{P}_s$. Let* $\text{dist}_\mu(\vee_0^{m-1} T^{-i}P) = \pi$. *Then μ is r-decomposable.*

PROOF. This is clear from the proof of Lemma 1(b).

DEFINITION. If $P = \{P^j : j \in E\}$ is a partition of Ω, and if $F \in \mathcal{F}$, then $P \cap F$ is the partition $\{P^j \cap F : j \in E\}$ of F.

THEOREM 4. *Let π be ergodic but not mixing. Let r be the period of π^*. Then there exists P such that* $\text{dist}_\mu(\vee_0^{m-1} T^{-i}P) = \pi$ *for all r-decomposable $\mu \in \mathcal{P}_s$.*

PROOF. Find disjoint $C_0, \cdots, C_{r-1} \subset A^{m-1}$ such that $\Sigma_{i=0}^{r-1} \pi\{(i_1, \cdots, i_m): (i_1, \cdots, i_{m-1}) \in C_i, (i_2, \cdots, i_m) \in C_{i+1}\} = 1$, where $C_r = C_0$. Let $F = \cap_{i=-\infty}^{\infty} \cap_{j=0}^{r-1} \{(X_{ir+j}, \cdots, X_{ir+j+m-2}) \in C_j\}$. Let $\nu = \hat{\pi}$. Let ν' be the measure on $\mathcal{Q}^\infty$ such that $\nu'(E) = \nu(E \cap F)/\nu(F)$, $E \in \mathcal{Q}^\infty$. Recall that $\{Z_i\}_{-\infty}^{\infty}$ is a stationary ergodic Markov process under ν, where $Z_i = (X_i, \cdots, X_{i+m-2})$. It follows that $\{U_i\}_{-\infty}^{\infty}$ is a stationary mixing Markov process under ν', where $U_i = (Z_{ir}, \cdots, Z_{ir+r-1})$. It follows that $\{(X_{ir}, \cdots, X_{ir+r-1})\}$ is stationary and mixing under ν' since $(X_{ir}, \cdots, X_{ir+r-1})$ is a function of U_i. Thus ν' is stationary and mixing relative to T_A^r.

Let $\mathcal{F}_0$ be a countable subfield of $\mathcal{F}$ which generates $\mathcal{F}$. Let $E_1, E_2, \cdots$ be an enumeration of the sets in $\mathcal{F}_0$. Letting I_{E_i} denote the indicator function of E_i, define $F^* \in \mathcal{F}$ to be the set of all $\omega \in \Omega$ such that

(a) $\lim_{n\to\infty} n^{-1} \Sigma_{k=1}^n I_{E_i}(T^{kr+j}\omega)$ exists, $\quad 0 \leq j \leq r-1, i = 1, 2, \cdots$

(b) There is a least $i \geq 1$ such that the r numbers $\{\lim_{n\to\infty} n^{-1}\Sigma_{k=1}^n I_{E_i}(T^{kr+j}\omega): 0 \leq j \leq r-1\}$ are distinct.

(c) For the i satisfying (b), $\lim_{n\to\infty} n^{-1}\Sigma_{k=1}^n I_{E_i}(T^{kr}\omega) > \lim_{n\to\infty} n^{-1}\Sigma_{k=1}^n I_{E_i}(T^{kr+j}\omega)$, $\quad 1 \leq j \leq r-1$.

It is not hard to see that $F^*, TF^*, \cdots, T^{r-1}F^*$ are disjoint, $T^rF^* = F^*$, and $\mu[\cup_0^{r-1} T^i F^*] = 1$ for every r-decomposable $\mu \in \mathcal{P}_e$. By Lemma 5, $\mu[\cup_0^{r-1} T^i F^*] = 1$ for all r-decomposable $\mu \in \mathcal{P}_s$.

Let $\hat{Q} = \{\hat{Q}^j : j \in A^r\}$ be the partition of A^∞ such that $\hat{Q}^j = \{(X_0, \cdots, X_{r-1}) = j\}$. Choose a positive integer k such that $kr \geq m + r - 1$. By Lemma 1, since ν' is T_A^r stationary and mixing, $\text{dist}_{\nu'}(\vee_{i=0}^{k-1} (T_A^r)^{-i}\hat{Q})$ is a mixing distribution on $(A^r)^k$. By Theorem 3, pick $\hat{P} = \{\hat{P}^j : j \in A^r\}$, a partition of Ω, such that $\text{dist}_{\mu'}(\vee_0^{k-1} (T^r)^{-i}\hat{P}) = \text{dist}_{\nu'}(\vee_0^{k-1} (T_A^r)^{-i}\hat{Q})$, for every μ' on $\mathcal{F}$ such that T^r is an aperiodic automorphism of $(\Omega, \mathcal{F}, \mu')$. Let $P = \{P^j : j \in A\}$ be a partition of Ω such that

$$P^j \cap T^i F^* = T^i\big[\cup \{\hat{P}^{(k_1, \cdots, k_r)} \cap F^* : k_i = j\}\big], j \in A,$$
$$0 \leqslant i \leqslant r - 1.$$

Then $(\vee_0^{r-1} T^{-i} P) \cap F^* = \hat{P} \cap F^*$. Let $\mu \in \mathcal{P}_s$ be r-decomposable. We show that $\text{dist}_\mu(\vee_0^{m-1} T^{-i} P) = \pi$. Define μ' on $\mathcal{F}$ so that $\mu'(E) = \mu(E \cap F^*)/\mu(F^*)$, $E \in \mathcal{F}$. Since μ' is concentrated on F^* and $T^r F^* = F^*$, we obtain $\text{dist}_{\mu'}(\vee_0^{rk-1} T^{-j} P) = \text{dist}_{\mu'}(\vee_0^{k-1} (T^r)^{-i} \hat{P})$. Since T^r is an aperiodic automorphism of $(\Omega, \mathcal{F}, \mu')$, we have $\text{dist}_{\mu'}(\vee_0^{rk-1} T^{-j} P) = \text{dist}_{\nu'}(\vee_0^{k-1} (T_A^r)^{-i} \hat{Q})$. Now $\hat{Q} = \vee_0^{r-1} (T_A)^{-i} Q$, where $Q = \{Q^j : j \in A\}$ is the partition of A^∞ such that $Q^j = \{X_0 = j\}$. Thus, $\text{dist}_{\mu'}(\vee_0^{rk-1} T^{-j} P) = \text{dist}_{\nu'}(\vee_0^{rk-1} T_A^{-j} Q)$. Now $\text{dist}_\mu(\vee_0^{m-1} T^{-i} P) = r^{-1} \Sigma_{j=0}^{r-1} \text{dist}_{\mu'}(\vee_j^{j+m-1} T^{-i} P)$ and $\pi = \text{dist}_\nu(\vee_0^{m-1} T^{-i} Q) = r^{-1} \Sigma_{j=0}^{r-1} \text{dist}_{\nu'}(\vee_j^{j+m-1} T_A^{-i} Q)$. Since $[j, j + m - 1] \subset [0, kr - 1]$, $0 \leqslant j \leqslant r - 1$, we have $\pi = \text{dist}_\mu(\vee_0^{m-1} T^{-i} P)$.

Acknowledgment. I would like to thank Professor U. Krengel for pointing out that Theorem 1.1 of [1] answers the question raised at the end of [3], and also Professor N. D. Ylvisaker for bringing [2] to my attention.

REFERENCES

[1] Grillenberger, C. and Krengel, U. (1976). On marginal distributions and isomorphisms of stationary processes. *Math. Z.* **149** 131–154.

[2] Hobby, C. and Ylvisaker, N. D. (1964). Some structure theorems for stationary probability measures on finite state sequences. *Ann. Math. Statist.* **35** 550–556.

[3] Kieffer, J. C. (1974). On the approximation of stationary measures by periodic and ergodic measures. *Ann. Probability* **2** 530–534.

[4] Krieger, W. (1970). On entropy and generators of measure-preserving transformations. *Trans. Amer. Math. Soc.* **149** 453–464.

[5] Ornstein, D. S. (1974). *Ergodic Theory, Randomness, and Dynamical Systems*. Yale Univ. Press.

[6] Oxtoby, J. C. (1952). Ergodic sets. *Bull. Amer. Math Soc.* **58** 116–136.

[7] Shields, P. C. (1973). *The Theory of Bernoulli Shifts*. Univ. of Chicago Press.

Department of Mathematics
University of Missouri-Rolla
Rolla, Missouri 65401