JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 152, 470–487 (1990)

Return Times for Nonsingular Measurable Transformations

S. ALPERN

Department of Mathematics, London School of Economics, Houghton Street, London, England WC2A 2AE

AND

V. S. PRASAD*

Department of Mathematics, University of Lowell, Lowell, Massachusetts 01854

Submitted by Dorothy Maharam Stone

Received May 22, 1989

Let σ be a nonsingular antiperiodic automorphism of a Lebesgue probability space $(X, \mathscr{A}, \mu)$. Let $\pi = (\pi_1, \pi_2, ..., \pi_k, ...)$ be a probability vector with the property that the k's for which $\pi_k > 0$ is a relatively prime set of integers and $\sum_{i=1}^{\infty} i\pi_i < \infty$. Then there is a measurable set B of positive measure such that the relative distribution of return times under σ to B is the given distribution.

1. INTRODUCTION

In this paper we consider the following problem: Given a nonsingular automorphism of a Lebesgue space, find a measurable subset having a prescribed set of return times. We show that such a set can always be found for any nonsingular antiperiodic automorphism of a Lebesgue probability space $(X, \mathscr{A}, \mu)$ as long as the set of return times forms a relatively prime set of integers. Furthermore, we can say a good deal about the distribution of these return times.

These results imply Alpern's theorems [1, 2] on the return times of finite measure preserving transformations. The latter have proved to be very useful in ergodic theory, the study of measure preserving homeomorphisms [1], and in the coding of stationary stochastic processes (cf. Alpern, Prasad [3]). The results in this paper, in addition to implying Alpern's results for the measure preserving case, may also be used to infer analogous

* The authors gratefully acknowledge support from a NATO Collaborative Research grant. The second author was also supported by a grant from NSERC, Canada.

0022-247X/90 $3.00

properties for nonsingular homeomorphisms and nonstationary stochastic processes. We use our main result to prove several conjugacy theorems for nonsingular transformations. We point out that one such application, Corollary 1, is related to Theorem 6 of Choksi and Kakutani [5] concerning the conjugates of an ergodic infinite measure preserving transformation.

We let $G = G(X)$ denote the group of all μ-nonsingular bimeasurable bijections σ, of the Lebesgue probability space $(X, \mathscr{A}, \mu)$ (i.e., $\mu\sigma^{-1}(A) = 0$ if and only if $\mu(A) = 0$).

Our main results follow.

THEOREM 1. *Let σ be a nonsingular antiperiodic automorphism of a Lebesgue probability space $(X, \mathscr{A}, \mu)$. Let $\pi = (\pi_1, \pi_2, ..., \pi_k\ , ...)$ be a probability vector with the property that the k's for which $\pi_k > 0$ is a relatively prime set of integers and $\sum_{k=1}^{\infty} k\pi_k < \infty$. Then there is a measurable set B of positive measure such that the relative distribution of return times under σ to B is the given distribution.*

We note that without the finiteness of the mean return time (i.e., $\sum_{k=1}^{\infty} k\pi_k$) the theorem would not be true in general since, for a measure preserving transformation σ, Kac's Theorem (cf. [6]) says that $\mu(B) = (\sum_{k=1}^{\infty} k\pi_k)^{-1} = \pi^*$. We also note that Rohlin's Lemma (cf. [6]) is a special case of this theorem, since Rohlin's Lemma asserts the existence of a set with return times n and 1.

THEOREM 2. *Let σ and π be as above. There is a measurable partition $\mathbf{P} = \{P_{k,i}: k = 1, 2, ..., i = 1, 2, ..., k\}$ of X with base $B = \bigcup_{k=1}^{\infty} P_{k,1}$ satisfying*

(1) $\sigma(P_{k,i}) = P_{k,i+1}$, *for* $i < k$, $k = 1, 2, ...$, *and*

(2) $\mu(P_{k,1})/\mu(B) = \pi_k$ *for* $k = 1, 2, ...$

Furthermore, given any positive numbers γ and γ_k, $k = 1, 2, ...,$ the partition $\mathbf{P}$ can be made to satisfy the following, where $\pi^* = (\sum_k k\pi_k)^{-1}$:

(3) $(1+\gamma_k)^{-1} < \mu(P_{k,i})/\mu(P_{k,1}) < 1 + \gamma_k$, *for* $k = 1, 2, ...$ *and* $i = 1, ..., k$

(4) $\pi^*/(1+\gamma) < \mu(B) < (1+\gamma)\pi^*$ *and consequently*

(5) $(1+\gamma)^{-2} < \mu(P_{k,i})/(\pi_k\pi^*) < (1+\gamma)^2$.

Note that conclusions (1) and (2) above are simply a restatement of Theorem 1 in partition language. If σ preserves μ, then it follows from (1) and (2) that $\mu(B) = \pi^* = (\sum_{k=1}^{\infty} k\pi_k)^{-1}$ (i.e., Kac's Theorem), and that $\mu(P_{k,i}) = \mu(P_{k,1}) = \pi_k\mu(B) = \pi_k\pi^*$. The assertions of conclusions (3)–(5) are that these equations "hold approximately" for certain partitions, even in the more general case of nonsingular σ. In particular, we can guarantee that $\mu(\bigcup_{i=1}^{k} P_{k,i}) \approx k\pi_k\pi^* \equiv \phi_k$ (the ϕ_k's sum to 1). The next theorem shows that we can obtain this relation exactly.

THEOREM 3. *Let σ be a nonsingular antiperiodic automorphism of a Lebesgue probability space $(X, \mathscr{A}, \mu)$. Let $\phi = (\phi_1, \phi_2, ..., \phi_k, ...)$ be a probability vector with the property that the k's for which $\phi_k > 0$ is a relatively prime set of integers. Then there is a partition $\{P_{k,i}: k = 1, 2, ..., i = 1, 2, ..., k\}$ of X such that*

(1) $\sigma(P_{k,i}) = P_{k,i+1}$, *for* $i < k$, *and*

(2) $\mu(\bigcup_{i=1}^{k} P_{k,i}) = \phi_k$ *for all* k.

Theorem 3 is the extension of Alpern's result [2, Theorem 1] for measure preserving transformations. Note also that Theorems 2 and 3 extend Chacon and Friedman's [4] result that for nonsingular antiperiodic automorphisms one can always find a set having return times 2 and 3.

Before proceeding with the proofs of these theorems we state some consequences of these results. The first is a nonsingular extension of a result of Alpern [2, Theorem 4] for measure preserving automorphisms. The conclusion of the corollary below is also similar to the conclusion of Choksi and Kakutani's Theorem 6 in [5], concerning infinite measure preserving transformations. It is also related to a conjugacy result of Chacon and Friedman (cf. [6, Theorem 7.13]).

COROLLARY 1. *Let $\sigma, \tau \in G$ with σ antiperiodic and τ ergodic. Let $F \in \mathscr{A}$ with $\mu(F) < 1$. Assume either*

(1) $\mu(F \cup \tau F) < 1$, *or*

(2) τ^n *is ergodic for all* $n \geqslant 1$ (τ *is setwise aperiodic*): *for all* $A \in \mathscr{A}$, $0 < \mu(A) < 1$, $n \geqslant 1$, $\tau^n(A) \neq A$.

Then there is a $\theta \in G$ such that $\theta^{-1}\sigma\theta(x) = \tau(x)$ for $\mu -$ a.e. $x \in F$. Furthermore, if σ and τ preserve μ, then we can also choose θ so that it preserves μ.

Proof. Given τ and F we construct a partition $\{F_{k,i}: k = 1, 2, ..., i = 1, 2, ..., k\}$ of X, based on the skyscraper with base $B = X \backslash \tau F$. The set $F_{1,1} = X \backslash (F \cup \tau F)$ has positive measure if hypothesis (1) holds. For $k \geqslant 2$, define $F_{k,1} = \{x \in F: \tau^{-1}(x) \notin F, \tau^i(x) \in F$, for $i = 0, ..., k-2$, and $\tau^{k-1}(x) \notin F\}$, and $F_{k,i} = \tau^{i-1}(F_{k,1})$ for $i = 2, ..., k$. Observe that $F = \bigcup_{k=2}^{\infty} \bigcup_{i=1}^{k-1} F_{k,i}$. Note that either of the hypotheses (1), (2) implies that $gcd\{k: \pi_k \equiv \mu(F_{k,1}) > 0\} = 1$. Given $\pi = (\pi_1, ..., \pi_k, ...)$ and σ we can apply part (1) of Theorem 2 to obtain a partition $\mathbf{P} = \{P_{k,i}: k = 1, 2, ..., i = 1, ..., k\}$ with $\sigma(P_{k,i}) = P_{k,i+1}$ for $i < k$, and $\mu(P_{k,1}) > 0 \Leftrightarrow \pi_k > 0 \Leftrightarrow \mu(F_{k,1}) > 0$. Hence we can define θ on $F_{k,1}$ to be any nonsingular bimeasurable bijection onto $P_{k,1}$. Extend θ to all of X by setting for $x \in \tau^{i-1} F_{k,1}$, $\theta(x) = \sigma^{i-1}\theta\tau^{-i+1}(x)$. Then θ has the required properties. If σ and τ are μ-preserving, then by the remarks after Theorem 2, we can choose $P_{k,i}$ so that $\mu(P_{k,1}) = \pi_k \pi^* = \mu(F_{k,1})$, $k = 1, ..., \infty$. Hence θ can be choosen to be μ-preserving.

The next corollary says that given any nonsingular antiperiodic automorphism σ and the transition probabilities of a positive recurrent aperiodic irreducible Markov chain with discrete state space, there is a partition $\mathbf{P}$ such that the transitions under σ of the partition elements are "approximately" that given by the Markov chain. This may be used to infer an approximate coding result for nonstationary processes analogous to that proved in [3]. See also the results of Alpern [1], Kieffer [8], and Grillenberger and Krengel [7] for the stationary case involving finite state Markov chains. All of these results follow from the next corollary if σ is measure preserving, since γ below can be taken to be 0.

COROLLARY 2. *Let $\sigma \in G$ be antiperiodic. Let $\mathbf{N}$ be the set of nonnegative integers. Suppose that $(p_{i,j}: i, j \in \mathbf{N})$ are the transition probabilities of a positive recurrent, aperiodic, irreducible Markov chain with state space $\mathbf{N}$. Then for each $\gamma > 0$, there is a partition $\mathbf{Q} = \{Q_i\}_{i \in \mathbf{N}}$ of X such that for all $i, j \in \mathbf{N}$,*

$$\frac{p_{i,j}}{(1+\gamma)^2} \leqslant \frac{\mu(Q_i \cap \sigma^{-1} Q_j)}{\mu(Q_i)} \leqslant (1+\gamma)^2 \, p_{i,j}.$$

If σ is μ-preserving the above holds with $\gamma = 0$.

Proof. Let $\pi = (\pi_k)_{k \in \mathbf{N}}$, be given by: π_k is the probability that the Markov chain starting at 0 first returns to 0 in exactly k steps. Then the aperiodicity of the chain implies that π satisfies the hypothesis of Theorem 2. Let $\gamma_i = \gamma$ and let $\mathbf{P}$ be the partition from Theorem 2. Attatch the label 0 to B the base of the partition. Then for each distinct loop $w = [0 = i_0, i_1, ..., i_k = 0]$ at 0, having probability α, find a subset F_w of $\bigcup_{i=1}^k P_{k,i}$ so that $\mu(P_{k,i} \cap (\bigcup_{j=1}^k \tau^{j-1} F_w)) = \alpha \mu(P_{k,i})$ for $1 \leqslant i \leqslant k$. This involves using the Erasing Lemma of the next section. The different F_w's can be chosen to be disjoint from each other. Attatch the labels $i_1, ..., i_{k-1}$ respectively to $\tau^{i-1} F_w$ for $i = 2, ..., k$. When this is done for each distinct loop at $w = [0 = i_0, i_1, ..., i_k = 0]$, each point in X has a single label. Let Q_i be the set consisting of all points with the label i. This is the required partition.

Since $(X, \mathscr{A}, \mu)$ is a Lebesgue space and all Lebesgue spaces are (measure theoretically) the same, we specialize in this and the next corollary to the case when X is the open unit disk in the plane, and μ is the normalized planar Lebesgue measure. On $G(X)$ consider the following three topologies: *compact-equal*, *compact-open*, and the *uniform topology*.

Bases for $\tau \in G$ in these topologies are given respectively by setting for compact $K \subset X$, and $\varepsilon > 0$

(1) $V(\tau, K) = \{\sigma \in G : \tau(x) = \sigma(x)\ \mu - \text{a.e. } x \in K\}$.

(2) $N(\tau, K, \varepsilon) = \{\sigma \in G : |\tau(x) - \sigma(x)| < \varepsilon\ \mu - \text{a.e. } x \in K\}$.

(3) $U(\tau, \varepsilon) = \{\sigma \in G : \mu\{x : \tau(x) \neq \sigma(x)\} < \varepsilon\}$.

Since there are compact sets of arbitrarily large measure in X the compact-equal topology is finer than the compact-open topology and the uniform topology. It follows that:

COROLLARY 3. *For $\sigma \in G$ an antiperiodic automorphism the compact-equal closure of the conjugacy class of σ (i.e., $\{\theta^{-1}\sigma\theta : \theta \in G\}$) contains the setwise aperiodic automorphisms.*

The next result is the nonsingular analogue of one for measure preserving automorphisms which has proved to be very useful in proving approximation results for measure preserving homeomorphisms.

COROLLARY 4. *The compact-equal (and, therefore, the compact-open) closure of the conjugacy class of an antiperiodic $\sigma \in G(X)$ contains the ergodic nonsingular homeomorphisms of the open disk X.*

Proof. Let τ be an ergodic homeomorphism of X and let $F \subset X$ be compact. Then $F \cup \tau F$ is compact and so $\mu(F \cup \tau F) < \mu(X) = 1$. Hence Corollary 1 applies with condition (1). Consequently $\theta^{-1}\sigma\theta \in V(\tau, F)$.

Remark. In Corollary 4, X can be taken to be any noncompact sigma-compact, connected manifold, $\mathscr{A}$ the Borel subsets of X, and μ any non-atomic Borel probability measure strictly positive on nonempty open sets.

2. LYAPUNOV'S THEOREM AND SKYSCRAPERS

Since all our theorems apply to a single antiperiodic automorphism in G, fix an arbitrary one and call it σ. We say that a measurable partition $\{E_i : i \geqslant 0\}$ of X is a *skyscraper* if $\sigma(E_i) \subset E_0 \cup E_{i+1}$ for $i = 0, 1, \ldots$ and if almost all points of E_0 return under σ to E_0. The *top* of $\{E_i\}$ is defined as $\sigma^{-1}E_0$ and denoted by $T(\{E_i\})$. For any measurable subset F of E_0 define $F_0 = F$ and for $i \geqslant 1$, $F_i = \sigma(F_{i-1}) \cap E_i$. Define $C(F) = \bigcup_{i=0}^{\infty} F_i$, to be the column(s) over F. Since σ is antiperiodic, the following lemma follows from the existence of recurrent sweep-out sets of arbitrarily small measure

LEMMA. *For each positive real number ε, $0 < \varepsilon < 1$, and σ a nonsingular* *antiperiodic automorphism of a Lebesgue probability space* $(X, \mathscr{A}, \mu)$ *there is a skyscraper* $\{E_i\}$ *such that*

$$\mu(T(\{E_i\})) < \varepsilon.$$

In this language, Theorem 1 states that X can be partitioned as a skyscraper with only columns of specified heights as long as the set of heights is relatively prime; furthermore, we can specify the probability distribution of the base of these columns (Theorem 1) or we can specify the distribution of the measure of these columns to be any probability vector (Theorem 3).

To prove Theorem 1 we use the following well-known result due to Lyapunov, on the range of a vector measure (cf. [9]).

LYAPUNOV'S THEOREM. *Let* v_j, $j = 1, ..., n$, *be finite nonatomic measures on a measurable space* $(E, \mathscr{A}_E)$. *Let* $\mathbf{v}: \mathscr{A}_E \to \mathbf{R}^n$ *be defined by*

$$\mathbf{v}(F) = (v_1(F), ..., v_n(F)).$$

Then the range $\mathbf{v}(\mathscr{A}_E)$ *is closed and convex. In particular we have*:

(1) *Given any number* α, $0 < \alpha < 1$, *there is a set* $F_\alpha \in \mathscr{A}_E$ *such that* $v_j(F_\alpha) = \alpha v_j(E)$, $j = 1, ..., n$.

(2) *Given any positive integer* k *there is a measurable partition* $\{F_i\}_{i=0}^{k-1}$ *of* E *such that* $v_j(F_i) = v_j(E)/k$, $j = 1, ..., n$.

The following two lemmas concern an arbitrary skyscraper $\{E_i\}$ and are simple consequences of Lyapunov's Theorem:

ERASING LEMMA. *Let* $\{E_i\}$ *be a skyscraper. Let* $\mathbf{P} = \{P_j\}_{j=1}^n$ *be any partition of* $(X, \mathscr{A})$ *and let* α, $0 < \alpha < 1$, *be any number. Then there is a measurable subset* F *of* E_0 *with*

$$\mu(P_j \cap C(F)) = \alpha\mu(P_j) \qquad \text{for} \quad j = 1, ..., n$$

and consequently

$$\mu(C(F)) = \alpha.$$

Proof. Define nonatomic measures v_j, $j = 1, ..., n$ on $(E_0, \mathscr{A}_{E_0})$ by

$$v_j(A) = \mu(P_j \cap C(A)).$$

Let F be the set F_α given by part (1) of Lyapunov's Theorem. Then for all j,

$$\mu(P_j \cap C(F)) = v_j(F) = \alpha v_j(E_0) = \alpha\mu(P_j)$$

since $C(E_0) = X$.

LABELLING LEMMA. *For any skyscraper $\{E_i\}$, any measurable subset F of E_0 and any positive integer k there is a measurable partition $\{S_j\}_{j=1}^k$ of $C(F)$, with the following properties:*

(1) $\sigma(S_j) \subset S_{j+1} \cup E_0$, *where addition on the indices is done* mod k.

(2) $\mu(S_j \cap C(F)) = \mu(C(F))/k$, $j = 1, ..., k$.

Proof. For $j = 1, ..., k$, $A \subset F$, let

$$C_j(A) = \bigcup_{j' = j (\mathrm{mod}\, k)} \sigma^{j'}(A).$$

Thus $C_j(A)$ is the j mod k subtower over A. For $j = 1, ..., k$ define measures λ_j on $(F, \mathscr{A}_F)$ by

$$\lambda_j(A) = \mu(C_j(A)).$$

Since the k subtowers $C_j(F)$ have $C(F)$ as their disjoint union, we have for all measurable subsets A of F

$$\sum_{j=1}^{k} \lambda_j(A) = \mu(C(A)).$$

By part two of Lyapunov's Theorem there is a measurable partition $\{F_i\}_{i=0}^{k-1}$ of F with $\lambda_j(F_i) = \lambda_j(F)/k$, $i = 0, ..., k-1$, $j = 1, ..., k$. Define S_j by

$$S_j = \bigcup_{i=0}^{k-1} C_{j+i}(F_i).$$

We claim that $\{S_j\}_{j=1}^k$ has the required properties. To establish (1), observe that a typical subset of S_j has the form $\sigma^{j'}(F_i)$, where $j' = j + i$ (mod k), and so there is some nonnegative integer n such that $j' = j + i + nk$. Thus, a typical subset of the left side of (1) has the form $\sigma(\sigma^{j'}(F_i)) = \sigma^{(j+1)+i+nk}(F_i) \subset C_{j+1+i}(F_i) \cup E_0 \subset S_{j+1} \cup E_0$.

To establish (2) observe that

$$\begin{aligned}
\mu(S_j \cap C(F)) &= \mu\left(\bigcup_{i=0}^{k-1} C_{j+i}(F_i)\right) \\
&= \sum_{i=0}^{k-1} \mu(C_{j+i}(F_i)) \\
&= \sum_{i=0}^{k-1} \lambda_{i+j}(F_i) \\
&= \sum_{i=0}^{k-1} (\lambda_{i+j}(F)/k) \\
&= \frac{1}{k} \sum_{i=0}^{k-1} \lambda_{i+j}(F) \\
&= \mu(C(F))/k.
\end{aligned}$$

3. Proof of Theorem 1

A measurable partition $\mathbf{P} = \{P_{k,i}\}$ of $(X, \mathscr{A})$ indexed by pairs (k, i), $k = 1, ..., \infty$, $i = 1, ..., k$, is called a *triangular partition.* A singly indexed set P_k from a triangular partition $\mathbf{P}$ is $\bigcup_{i=1}^{k} P_{k,i}$. The base measure $b(\mathbf{P})$ of a triangular partition is defined by

$$b(\mathbf{P}) = \sum_{k=1}^{\infty} \mu(P_{k,1}).$$

The partition metric is given by

$$\|\mathbf{R} - \mathbf{Q}\| = \sum_{k=1}^{\infty} \sum_{i=1}^{k} \mu(R_{k,i} \Delta Q_{k,i}),$$

where Δ denotes the symmetric difference between two sets.

There are two distributions of a triangular partition $\mathbf{P}$ which we consider. The first is the base distribution given by

$$d_b(\mathbf{P}) = (\mu(P_{1,1})/b(\mathbf{P}), \mu(P_{2,1})/b(\mathbf{P}), ...).$$

The second is the column distribution

$$d_c(\mathbf{P}) = (\mu(P_1), \mu(P_2), ...).$$

Note that

$$|d_c(\mathbf{P}) - d_c(\mathbf{Q})| \leqslant \|\mathbf{P} - \mathbf{Q}\|.$$

The metric on the space of distributions is given by their l^1 distance

$$|X-Y| = \sum_{k=1}^{\infty} |X_k - Y_k|.$$

The triangular partition $\mathbf{P}$ is an n-partition if $\mu(P_{k,i})=0$ for all $k>n$. The partition of $\mathbf{P}$ of Theorem 1 is obtained as a limit of partitions $\mathbf{P}^n$, where each $\mathbf{P}^n$ is an n-partition. For any distribution π, set $\pi^n = (\pi_1/c_n, ..., \pi_n/c_n, 0, 0, ...)$, where $c_n = \sum_1^n \pi_k$, and observe that $\lim_n \pi^n = \pi$.

The partition $\mathbf{P}$ of Theorem 2 is a triangular partition which additionally satisfies $\sigma(P_{k,i}) = P_{k,i+1}$ for $i<k$. This means that if x belongs to $P_{k,i}$ then $\sigma(x)$ belongs to $P_{k,i+1}$, unless $i=k$, in which case $\sigma(x)$ belongs to $B = \bigcup_k P_{k,1}$. With this property in mind we define a transition $(k, i) \to (k', i')$ to be "legal" if $k=k'$ and $i'=i+1$, or if $i=k$ and $i'=1$. Define the "wrong set" of a partition $\mathbf{P}$ to be

$$W(\mathbf{P}) = \{x \in X: x \in P_{k,i}, \sigma(x) \in P_{k',i'} \text{ and } (k, i) \to (k', i') \text{ is not legal}\}.$$

Note that this may also be described as follows:

$$W(\mathbf{P}) = \bigcup_{k=1}^{\infty} \left[\bigcup_{i=1}^{k-1} (P_{k,i} \backslash \sigma^{-1} P_{k,i+1}) \cup \left(P_{k,k} \backslash \sigma^{-1} \left(\bigcup_{j=1}^{\infty} P_{j,1} \right) \right) \right].$$

Note that conclusion (1) of Theorem 2, $\sigma(P_{k,i}) = P_{k,i+1}$, $i<k$, is equivalent to $\mu(W(\mathbf{P}))=0$ for a triangular partition $\mathbf{P}$. To say that $\mu(W(\mathbf{P}))$ is small is an approximate form of conclusion (1), useful for our limiting argument.

Let π be the denumerable probability distribution given in Theorem 1. Define $\pi^* = (\sum_{k=1}^{\infty} k\pi_k)^{-1}$, the reciprocal of the expected return time to the base set B. There are numbers N such that the set $\{1 \leqslant k \leqslant N: \pi_k > 0\}$ is relatively prime. Fix one such N. Consequently, for any $n \geqslant N$ there is an integer $v = v(n)$ such that for any pairs (k, i) and (k', i') with $k, k' \leqslant n$, there is a legal sequence of transitions $(k, i) = (k_0, i_0) \to (k_1, i_1) \to \cdots \to (k_j, i_j) \to \cdots \to (k_v, i_v) = (k', i')$ from (k, i) to (k', i') in v transitions, with $k_j \leqslant n$ for all $j = 0, ..., v$. Let $s_n = \min\{\pi_k: \pi_k > 0, k \leqslant n\}$.

The next three lemmas concern a partition of a skyscraper and how it can be successively modified to another of the same skyscraper—the modifications done so that the changed partition is "closer" to the required partition of Theorem 1. In the proof of Theorem 1 (and 2) these lemmas are used consecutively on the same skyscraper. Lemma 1 says that the "wrong set" of an n-partition ($\mathbf{P}$) can be moved to the top of the skyscraper (which is chosen to have a small top). Lemma 2 takes the resulting partition ($\mathbf{Q}$) and modifies it to ($\mathbf{R}$) to have the base distribution π^n while keeping the wrong set at the top of the skyscraper. So far, all of these partitions are n-partitions. Lemma 3 then adds new labels $(n+1, 1), \ldots, (n+1, n+1)$ so that the resulting partition (**S**) has base distribution π^{n+1} and still has its wrong set at the top of the skyscraper. If the first of these partitions **P** is called $\mathbf{P}^n$, and the last **S** is called $\mathbf{P}^{n+1}$, then we may expect that if $\|\mathbf{P}^n - \bar{\mathbf{P}}\| \to 0$ for some limiting triangular partition $\bar{\mathbf{P}}$, $\bar{\mathbf{P}}$ satisfies $d_b(\bar{\mathbf{P}}) = \lim d_b(\mathbf{P}^{n+1}) = \lim \pi^{n+1} = \pi$. If these three lemmas are applied successively to skyscrapers having tops with measure decreasing to 0, we obtain $\mu(W(\bar{\mathbf{P}})) = \lim \mu(W(\mathbf{P}^n)) \to 0$.

LEMMA 1 (Moving the wrong set to top of skyscraper). *For any $n > N$ and $\varepsilon > 0$ there is a $\delta = \delta(n) > 0$, such that the following holds. Given any triangular n-partition* **P** *with $\mu(W(\mathbf{P})) < \delta$ and any skyscraper $\{E_i\}$, there is a triangular n-partition* **Q**, *with $W(\mathbf{Q}) \subset T(\{E_i\})$ and $\|\mathbf{Q} - \mathbf{P}\| \leqslant \varepsilon$.*

Proof. As we discussed earlier, there is a positive integer $v = v(n)$, such that all pairs (k, i) and (k', i') with $k, k' \leqslant n$, are connected by a sequence of exactly v legal transitions. Since the measure $\nu = \sum_{i=0}^{v} \mu\sigma^i$ is equivalent to the measure μ, there is a positive number $\delta = \delta(n)$, such that if $\mu(A) < \delta$ then $\nu(A) < \varepsilon/2$.

We can partition the base of the skyscraper into *columns*, given by the different **P**-m names for the points $x \in E_0$, where m is the number of floors in the skyscraper directly above x. By the **P**-m name of a point $x \in X$, we mean the m-tuple of indices $((k_1, i_1), \ldots, (k_m, i_m))$ such that $\sigma^{j-1}x \in P_{k_j, i_j}$, $j = 1, \ldots, m$. Partition the base E_0 into sets with the same **P**-names. If $F \subset E_0$ is one such set corresponding to a **P**-m name then $\bigcup_{i=0}^{m-1} \sigma^i(F) = C(F)$ is called a *column of $\{E_i\}$ with respect to* **P**. The sets $\sigma^i(F)$, $i = 0, \ldots, m-1$, are called column levels.

We note that $W(\mathbf{P}) \cap \{E_i\}$ is the union of certain column levels of $\{E_i\}$ with respect to **P**. We obtain the partition **Q** by changing the label of certain column levels corresponding to illegal transitions. Choose any column $C(F)$ of $\{E_i\}$ with respect to **P** which has a column level in $W(\mathbf{P}) \cap \{E_i\}$. The set F corresponds to some **P**-m name

$$(k_1, i_1), \ldots, (k_j, i_j), \ldots, (k_m, i_m).$$

Let j_1 be the smallest value of $j < m$ such that the transition from (k_j, i_j) to (k_{j+1}, i_{j+1}) is illegal. This means that $\sigma^{j_1}(F)$ is the lowest column level belonging to $W(\mathbf{P}) \cap C(F)$. Choose (k'_{j_1+l}, i'_{j_1+l}), $l = 1, \ldots, v(n) - 1$, so that the transitions

$$(k_{j_1}, i_{j_1}) \to (k'_{j_1+1}, i'_{j_1+1})$$

$$(k'_{j_1+l}, i'_{j_1+l}) \to (k'_{j_1+l+1}, i'_{j_1+l+1}), \qquad l = 1, \ldots, v(n) - 2$$

$$(k'_{j_1+v(n)-1}, i'_{j_1+v(n)-1}) \to (k_{j_1+v(n)}, i_{j_1+v(n)})$$

are all legal transitions. If $j_1 + v(n) > m$, simply ignore the last $j_1 + v(n) - m$ conditions. Define $\mathbf{Q}^1$ equal to $\mathbf{P}$ except on the column levels $\sigma^{j_1+l}(F)$, $l = 1, ..., \min(v(n) - 1, m - j)$, where the $\mathbf{Q}^1$ labels are respectively (k'_{j_1+l}, i'_{j_1+l}). Observe that $W(\mathbf{Q}^1)$ contains at least one fewer column level, that is, $\sigma^{j_1}(F)$, than did $W(\mathbf{P})$. Now apply the same process to $\mathbf{Q}^1$ instead of to the partition $\mathbf{P}$ and call the resulting partition $\mathbf{Q}^2$. After repeating this process a finite number of times for each column name, we arrive at a partition $\mathbf{Q}$ with $W(\mathbf{Q}) \subset T(\{E_i\})$. To estimate the distance $\|\mathbf{Q} - \mathbf{P}\|$ between the two partitions, observe that for each column level in $W(\mathbf{P}) \cap \{E_i\}$, $\mathbf{Q}$ differs from $\mathbf{P}$ at most on the next $v(n) - 1$ images under σ of that level. Since $\| - \|$ measures such changes twice, we obtain the estimate

$$\|\mathbf{Q} - \mathbf{P}\| \leqslant 2 \sum_{i=0}^{v-1} \mu\sigma^i(W(\mathbf{P})) = 2v(W(\mathbf{P})) < \varepsilon$$

since $\mu(W(\mathbf{P})) < \delta$.

LEMMA 2 (Correcting the distribution). *Let $\{E_i\}$ be any skyscraper and let $\mathbf{Q}$ be an n-partition with $W(\mathbf{Q}) \subset T(\{E_i\})$. Then there is a triangular n-partition $\mathbf{R}$ with $W(\mathbf{R}) \subset T(\{E_i\})$, $d_b(\mathbf{R}) = \pi^n$ and*

$$\|\mathbf{R} - \mathbf{Q}\| \leqslant 2n\, |d_b(\mathbf{Q}) - \pi^n| / s_n,$$

where $s_n = \min\{\pi_k : 1 \leqslant k \leqslant n, \pi_k \neq 0\}$.

Proof. Choose α, $0 < \alpha < 1$, so that

$$\frac{\alpha}{1-\alpha} = \frac{s_n}{n\, |d_b(\mathbf{Q}) - \pi^n|}.$$

Let $\beta = 1 - \alpha$. Using the Erasing Lemma applied to the partition $\mathbf{Q}$ we find a subset F_α of E_0 with

$$\mu(Q_{k,1} \cap C(F_\alpha)) = \alpha\mu(Q_{k,1}), \qquad k = 1, ..., n.$$

Define $F_\beta = E_0 \backslash F_\alpha$. Consider the partition $X = C(F_\alpha) \cup C(F_\beta)$. We define the triangular n-partition $\mathbf{R}$ as follows. On $C(F_\alpha)$ we define $\mathbf{R}$ identically with $\mathbf{Q}$, i.e.,

$$R^\alpha_{k,i} = R_{k,i} \cap C(F_\alpha) = Q_{k,i} \cap C(F_\alpha), \qquad k = 1, ..., n; \qquad i = 1, ..., k.$$

We partition F_β into n sets F^k_β, $k = 1, ..., n$ and label the set $C(F^k_\beta)$ with

$\mathbf{R}$-labels $(k, 1), \ldots, (k, k)$ in equal measures and no illegal transitions, using the Labelling Lemma. By this construction we ensure that

$$\begin{aligned} b(\mathbf{R}) &= \mu\left(\bigcup_{k=1}^{n} R_{k,1}\right) \\ &= \mu\left(\bigcup_{k=1}^{n} R_{k,1} \cap C(F_\alpha)\right) + \mu\left(\bigcup_{k=1}^{n} R_{k,1} \cap C(F_\beta)\right) \\ &\geqslant \alpha b(\mathbf{Q}) + (1-\alpha)/n. \end{aligned}$$

It follows (derived later) from the choice of α that for $k = 1, \ldots, n$,

$$\frac{\mu(R_{k,1}^{\alpha})}{b(\mathbf{R})} = \frac{\alpha\mu(Q_{k,1})}{b(\mathbf{R})} \leqslant \pi_k^n. \tag{1}$$

Hence by appropriately choosing $\mu(F_\beta^k)$ we can ensure that

$$\frac{\mu(R_{k,1})}{b(\mathbf{R})} = \frac{\alpha\mu(Q_{k,1}) + (1/k)\,\mu(C(F_\beta^k))}{\alpha b(\mathbf{Q}) + \sum_{j=1}^{n} (1/j)\,\mu(C(F_\beta^j))} = \pi_k^n.$$

Therefore, $d_b(\mathbf{R}) = \pi^n$, $W(\mathbf{R}) \subset T(\{E_i\})$, and

$$\begin{aligned} \|\mathbf{R} - \mathbf{Q}\| &\leqslant 2(1-\alpha) = 2\alpha n\,|d_b(\mathbf{Q}) - \pi^n|/s_n \\ &\leqslant 2n\,|d_b(\mathbf{Q}) - \pi^n|/s_n \end{aligned}$$

as required.

The derivation of (1) is as follows:

Let $q_k = \mu(Q_{k,1})/b(\mathbf{Q})$, so $d_b(\mathbf{Q}) = (q_1, q_2, \ldots, q_n, 0, \ldots)$. For any $k = 1, \ldots, n$ with $q_k > \pi_k^n$ we have $q_k - \pi_k^n \leqslant |d_b(\mathbf{Q}) - \pi^n|$, and so since $\pi_k^n \geqslant s_n$, we have the following inequalities:

$$\frac{\alpha}{1-\alpha} \leqslant \frac{\pi_k^n}{(q_k - \pi_k^n)\,nb(\mathbf{Q})}$$

$$\alpha(q_k - \pi_k^n)\,b(\mathbf{Q}) \leqslant \frac{1-\alpha}{n}\,\pi_k^n$$

$$\alpha q_k b(\mathbf{Q}) \leqslant \alpha\pi_k^n b(\mathbf{Q}) + \frac{1-\alpha}{n}\,\pi_k^n$$

$$\alpha q_k b(\mathbf{Q}) \leqslant \left[\alpha b(\mathbf{Q}) + \frac{1-\alpha}{n}\right]\pi_k^n$$

$$\alpha q_k b(\mathbf{Q}) \leqslant b(\mathbf{R})\,\pi_k^n$$

$$\frac{\alpha\mu(Q_{k,1})}{b(\mathbf{R})} \leqslant \pi_k^n.$$

This establishes (1) for k with $q_k > \pi_k^n$. For k with $q_k \leqslant \pi_k^n$, (1) is obvious.

LEMMA 3 (Adding $(n+1)$-labels). *Let $\{E_i\}$ be a skyscraper and let $\mathbf{R}$ be a triangular n-partition with $W(\mathbf{R}) \subset T(\{E_i\})$ and $d_b(\mathbf{R}) = \pi^n$. Then there is a triangular $n+1$-partition $\mathbf{S}$ with $W(\mathbf{S}) \subset T(\{E_i\})$ and $d_b(\mathbf{S}) = \pi^{n+1}$, and*

$$\|\mathbf{S} - \mathbf{R}\| \leqslant 2(n+1)\, \pi_{n+1}/\pi_N.$$

Furthermore there is a contraction factor α_{n+1} such that for any k, i with $k \leqslant n$ and $i \leqslant k$, $\mu(S_{k,i}) = \alpha_{n+1}\mu(R_{k,i})$, with $1 - \alpha_{n+1} \leqslant (n+1)\, \pi_{n+1}/\pi_N$.

Proof. The basic idea is the same as in the previous lemma. We "erase" the labels from a set $C(F_\beta)$, with $\mu(C(F_\beta)) = \beta$ and then label this set with equal amounts of labels $(n+1, 1), ..., (n+1, n+1)$. Thus for $\alpha = \alpha_{n+1} = 1 - \beta$, this ensures that the new partition $\mathbf{S}$ satisfies $\mu(S_{k,i}) = \alpha\mu(R_{k,i})$ when $k \leqslant n$, and $\mu(S_{n+1,i}) = \beta/(n+1)$ for $i = 1, ..., n+1$. Since we add only "$n+1$-labels," it follows that, $b(\mathbf{S}) = \alpha b(\mathbf{R}) + \beta/(n+1)$. To ensure that $d(\mathbf{S}) = \pi^{n+1}$ it is sufficient to choose β so that

$$\frac{\mu(S_{n+1,1})}{b(\mathbf{S})} = \frac{\beta/(n+1)}{(1-\beta)\, b(\mathbf{R}) + \beta/(n+1)} = \pi_{n+1}^{n+1}$$

or equivalently

$$1 - \alpha = \beta = \frac{b(\mathbf{R})(n+1)\, \pi_{n+1}}{b(\mathbf{R})(n+1)\, \pi_{n+1} + c_n}.$$

Since $\beta \leqslant (n+1)\, \pi_{n+1}/\pi_N$, it follows that

$$\|\mathbf{S} - \mathbf{R}\| \leqslant 2\mu(C(F_\beta)) = 2\beta \leqslant 2(n+1)\, \pi_{n+1}/\pi_N.$$

We combine the first two lemmas of this section to obtain the following:

LEMMA 4. *Given $n \geqslant N$ and $\varepsilon > 0$ there is a $\delta > 0$ such that: Given any skyscraper $\{E_i\}$ and triangular n-partition $\mathbf{P}$ with $\mu(W(\mathbf{P})) < \delta$ and $d_b(\mathbf{P}) = \pi^n$ there is a triangular n-partition $\mathbf{R}$ with $W(\mathbf{R}) \subset T(\{E_i\})$, $d_b(\mathbf{R}) = \pi^n$ and $\|\mathbf{R} - \mathbf{P}\| < \varepsilon$.*

Proof of Theorem 1. Let ε_n, $n \geqslant N$, be any summable sequence of positive numbers, and let δ_n, $n \geqslant N$, be the associated δ's of Lemma 4 (with respect to n and ε_n). We may assume $\delta_n \to 0$.

We obtain the required partition $\mathbf{P}$ as the limit in the $\|-\|$ metric of triangular partitions $\mathbf{P}^n$, $n \geqslant N$, satisfying

(1) $\mathbf{P}^n$ is an n-partition and $d_b(\mathbf{P}^n) = \pi^n$

(2) $\mu(W(\mathbf{P}^n)) < \delta_{n+1}$

(3) $\|\mathbf{P}^n - \mathbf{P}^{n+1}\| < \varepsilon_{n+1} + 2(n+1)\,\pi_{n+1}/\pi_N$, $n > N$.

(4) $\mu(P^n_{n,i}) = \mu(P^n_{n,1})$, $i = 1, ..., n$, and $\mu(P^N_{k,i}) = \mu(P^N_{k,1})$, $k = 1, ..., N$, $i = 1, ..., k$.

We show there is a partition $\mathbf{P}^N$ satisfying (1), (2), and (4). Let $\{E_i\}$ be a skyscraper with $\mu(T(\{E_i\})) < \delta_{N+1}$. Let E_0 be written as the disjoint union of sets $F_1, F_2, ..., F_N$ with $\mu(C(F_k)) = k\pi^N_k(\sum_{i=1}^N i\pi^N_i)^{-1}$. The existence of such a partition follows from the Erasing Lemma. Fix any k and apply the Labelling Lemma to k and $F = F_k$. Call the resulting partition of $C(F_k)$, $\{P^N_{k,i}, i = 1, ..., k\}$. Observe that $\mu(P^N_{k,i}) = \mu(P^N_{k,1})$, which is the strong form of condition (4) required when $n = N$. Condition (2) follows from the choice of the skyscraper. Condition (1) is a consequence of $\mu(P^N_{k,1}) = \mu(C(F_k))/k = \pi^N_k(\sum_{i=1}^N i\pi^N_i)^{-1}$, and hence $b(\mathbf{P}^N) = (\sum_{i=1}^N i\pi^N_i)^{-1}$.

Suppose $\mathbf{P}^n$ satisfies conditions (1) and (2). Here is the way to construct $\mathbf{P}^{n+1}$ satisfying (1)–(4): Let $\{E_i\}$ be a skyscraper with $\mu(T(\{E_i\})) < \delta_{n+2}$. Apply Lemma 4 to the partition $\mathbf{P}^n$ with $\varepsilon = \varepsilon_{n+1}$ and $\delta = \delta_{n+1}$ to obtain a partition $\mathbf{R} = \mathbf{R}^n$ with $d_b(\mathbf{R}^n) = \pi^n$, $W(\mathbf{R}^n) \subset T(\{E_i\})$ and $\|\mathbf{R}^n - \mathbf{P}^n\| < \varepsilon_{n+1}$. Next apply Lemma 3 to $\mathbf{R} = \mathbf{R}^n$ which yields a partition $\mathbf{S} = \mathbf{P}^{n+1}$ which satisfies conditions (1)–(4).

Since ε_n and $n\pi_n$ are summable, condition (3) ensures the convergence of $\mathbf{P}^n$ to a limit partition $\mathbf{P}$ by the completeness of the $\|-\|$ partition metric. Since $\mu(W(\cdot))$ is continuous in this metric, condition (2) ensures that we have $\mu(W(\mathbf{P})) = 0$. Similarly, $d_b(\mathbf{P}) = \lim_n d_b(\mathbf{P}^n) = \lim_n \pi^n = \pi$. Thus $\mathbf{P}$ satisfies conditions (1) and (2) of Theorem 2 and hence its base $B = \bigcup_{k=1}^\infty P_{k,1}$ is the required set of Theorem 1.

4. Proof of Theorem 2

The proof of Theorem 2, parts (3), (4), (5) is based on a modification of the technique we used for Theorem 1. The idea is that if N is chosen sufficiently large and the ε_n small, then the partition $\mathbf{P} = \lim \mathbf{P}^n$ obtained in the proof of Theorem 1 also satisfies (3), (4), (5). We also need to make the following observation about the construction in the proof of Theorem 1:

Proposition 1. *In the construction above in the proof of Theorem* 1, *the following inequality holds for any* $k = 1, 2, ..., i = 1, ..., k$, *and* $n \geqslant \max(N, k)$, *where* α_{n+1} *is the "contraction" factor of Lemma* 3:

$$\alpha_{n+1}(\mu(P^n_{k,i}) - \varepsilon_{n+1}) \leqslant \mu(P^n_{k,i}) \leqslant \alpha_{n+1}(\mu(P^n_{k,i}) + \varepsilon_{n+1}).$$

Furthermore, for $j = 1, 2, \ldots$, we have

$$(\alpha_{k+1} \cdots \alpha_{k+j})\, \mu(P_{k,i}^k) - \sum_{l=k+1}^{k+j} \varepsilon_l \leqslant \mu(P_{k,i}^{k+j})$$
$$\leqslant (\alpha_{k+1} \cdots \alpha_{k+j})\, \mu(P_{k,i}^k) + \sum_{l=k+1}^{k+j} \varepsilon_l.$$

Proof. Recall the inductive step in the proof of Theorem 1, which starts with $\mathbf{P}^n$, uses Lemma 4 to produce $\mathbf{R}^n$, and then uses Lemma 3 to modify $\mathbf{R}^n$ to $\mathbf{P}^{n+1}$.

Fix k and i. Since $\|\mathbf{P}^n - \mathbf{R}^n\| \leqslant \varepsilon_{n+1}$, then $|\mu(P_{k,i}^n) - \mu(R_{k,i}^n)| < \varepsilon_{n+1}$. Furthermore, $\alpha_{n+1}\mu(R_{k,i}^n) = \mu(P_{k,i}^n)$. These two facts imply the first inequality. The second inequality follows by induction on j.

Proof of Theorem 2. Conditions (3), (4), (5) follow from the construction $\mathbf{P} = \lim \mathbf{P}^n$ given in the proof of Theorem 1, if N is chosen sufficiently large and ε_n are chosen sufficiently small. Specifically, choose N so that $\pi_N > 0$ and

(1) $\{k : 1 \leqslant k \leqslant N, \text{ and } \pi_k > 0\}$ is relatively prime,

(2) $c_N = \pi_1 + \pi_2 + \cdots + \pi_N > \frac{1}{2}$, and

(3) $c_N^{-1} \sum_{j=N+1}^{\infty} j\pi_j < \frac{1}{2}$.

Observe that condition (3) above guarantees that the product of the contraction factors α_l, $l > N+1$, exceeds $\frac{1}{2}$:

$$\begin{aligned} \alpha_l \alpha_{l+1} \cdots &\geqslant \alpha_{N+1}\alpha_{N+2} \cdots \\ &\geqslant 1 - \sum_{j=N+1}^{\infty} (1 - \alpha_j) \\ &\geqslant 1 - \sum_{j=N+1}^{\infty} \frac{j\pi_j}{c_N} \qquad \text{by Lemma 3} \\ &\geqslant 1/2 \text{ by (3)}. \end{aligned} \tag{2}$$

Next, choose the summable sequence ε_l, $l > N$, so that

$$\sum_{l=N+1}^{\infty} \varepsilon_l < \pi_N \pi^*/6 \tag{3}$$

(recall $\pi^* = (\sum_k k\pi_k)^{-1}$) and

$$\sum_{l=\max(N,k)}^{\infty} \varepsilon_l < \frac{\pi^* \pi_k \gamma_k}{6c_N(2+\gamma_k)}, \qquad k = 1, 2, \ldots. \tag{4}$$

These two conditions may be satisfied because each successive ε_l has only a finite number of constraints. Formula (4) is needed because it implies

$$\frac{(\pi_k \pi^*/6c_N) + \sum_{l=\max(N,k)}^{\infty} \varepsilon_l}{(\pi_k \pi^*/6c_N) - \sum_{l=\max(N,k)}^{\infty} \varepsilon_l} < 1 + \gamma_k, \qquad k = 1, 2, ... \tag{5}$$

We first establish a lower bound for the measure of the sets $P_{N,1}^{N+j}$ for $j = 0, 1, ...$ For $j = 0$ the construction of $\mathbf{P}^N$ ensured that $\mu(P_{N,1}^N) = \pi_N^N b(\mathbf{P}^N)$ and that

$$b(\mathbf{P}^N) = (\pi_1^N + 2\pi_2^N + \cdots + N\pi_N^N)^{-1} \geqslant \pi^*.$$

Hence $\mu(P_{N,1}^N) \geqslant \pi_N \pi^*$. For any positive integer j the Proposition and (2) and (3) guarantee that

$$\begin{aligned} \mu(P_{N,1}^{N+j}) &\geqslant \alpha_{N+1}\alpha_{N+2}\cdots\alpha_{N+j}\mu(P_{N,1}^N) - \sum_{l=N+1}^{N+j} \varepsilon_l \\ &\geqslant \frac{\pi_N \pi^*}{2} - \frac{\pi_N \pi^*}{6} = \frac{\pi_N \pi^*}{3}. \end{aligned} \tag{6}$$

Now fix any k, i with $1 < i \leqslant k$. Let $M = \max(N, k)$ be the stage at which label k is introduced. By condition (4) from the proof of Theorem 1, there is a common value $w = \mu(P_{k,1}^M) = \mu(P_{k,i}^M)$. Since $d(\mathbf{P}^M) = \pi^M$ we have $\mu(P_{k,1}^M)/\mu(P_{N,1}^M) = \pi_k/\pi_N$ and so by (6)

$$w = \frac{\pi_k \mu(P_{N,1}^M)}{\pi_N} \geqslant \frac{\pi_k \pi^*}{3}. \tag{7}$$

From the second inequality of the Proposition and using (5), we have for all $j \geqslant 1$

$$\begin{aligned} \frac{\mu(P_{k,i}^{M+j})}{\mu(P_{k,1}^{M+j})} &\leqslant \frac{\alpha_{N+1}\cdots\alpha_{N+j}w + \sum_{l=M+1}^{M+j} \varepsilon_l}{\alpha_{N+1}\cdots\alpha_{N+j}w - \sum_{l=M+1}^{M+j} \varepsilon_l} \\ &\leqslant \frac{(1/2)(\pi_k \pi^*/3) + \sum_{l=M+1}^{M+j} \varepsilon_l}{(1/2)(\pi_k \pi^*/3) - \sum_{l=M+1}^{M+j} \varepsilon_l} \\ &\leqslant 1 + \gamma_k. \end{aligned} \tag{8}$$

Since inequality (8) applies equally when i and 1 are interchanged, conclusion (3) of Theorem 2 is established by letting $j \to \infty$.

To establish conclusion (4) of the Theorem, observe $1 = \sum_{k=1}^{\infty} \sum_{i=1}^{k} \mu(P_{k,i})$. Hence it follows from conclusion (3), with $\gamma = \gamma_k$ that

$$\sum_{k=1}^{\infty} \frac{k\mu(P_{k,1})}{1+\gamma} \leqslant 1 \leqslant \sum_{k=1}^{\infty} k\mu(P_{k,1})(1+\gamma). \tag{9}$$

But by conclusion (2), $\mu(P_{k,1}) = \pi_k b(\mathbf{P}) = \pi_k \mu(B)$, so that (9) implies

$$\frac{\mu(B)}{1+\gamma} \sum_{k=1}^{\infty} k\pi_k \leqslant 1 \leqslant \mu(B)(1+\gamma) \sum_{k=1}^{\infty} k\pi_k. \tag{10}$$

Since $\pi^* = (\sum_{k=1}^{\infty} k\pi_k)^{-1}$, (10) implies conclusion (4). Conclusions (3) (with $\gamma_k = \gamma$) and (4) imply conclusion (5):

$$\begin{aligned} \mu(P_{k,i}) &\leqslant (1+\gamma)\,\mu(P_{k,1}) \\ &= (1+\gamma)\,\pi_k \mu(B) \\ &\leqslant (1+\gamma)\,\pi_k (1+\gamma)\,\pi^*. \end{aligned}$$

Since the lower bound follows similarly, conclusion (5) is established.

Proof of Theorem 3. Let $\phi = (\phi_1, ..., \phi_k, ...)$ be the given probability vector satisfying the premises of Theorem 3. The proof of Theorem 3 follows the demonstration of Theorem 1 in that starting with a partition of a skyscraper we modify it to another partition of the same skyscraper, except that now the modifications are done so that the new partition is "closer" to the partition required for Theorem 3. Since it is the column distribution d_c, instead of the base distribution d_b, on the space of triangular partitions that provides our measure of "closeness," Lemmas 1–3 used in proving Therem 1 must be correspondingly altered. We outline only the changes required in proving the modifications to these lemmas. Lemma 1 is unchanged. Lemma 2 is changed to the following:

LEMMA 2*. *Let* $\{E_i\}$ *be any skyscraper and let* $\mathbf{Q}$ *be an n-partition with* $W(\mathbf{Q}) \subset T(\{E_i\})$. *Then there is a triangular n-partition* $\mathbf{R}$ *with* $W(\mathbf{R}) \subset T(\{E_i\})$, $d_c(\mathbf{R}) = \phi^n$, *and*

$$\|\mathbf{R} - \mathbf{Q}\| \leqslant 2\,|d_c(\mathbf{Q}) - \pi^n|/s_n,$$

where $s_n = \min\{\pi_k : 1 \leqslant k \leqslant n,\ \pi_k \neq 0\}$.

This is proved by taking $1 - \alpha = 2\,|d_c(\mathbf{Q}) - \phi^n|/s_n$. The skyscraper is partitioned into the distinct $\mathbf{Q}$-names. For a set corresponding to a $\mathbf{Q}$-m name, $F \subset E_0$ (and $C(F)$ the corresponding column of the skyscraper), use the Erasing Lemma to find a subset F_α so that the relative distribution of $\mathbf{Q}$ on $C(F_\alpha)$ is the same as that on $C(F)$, and so that $C(F_\alpha)$ has α of the measure of $C(F)$. The labels on $C(F_\alpha)$ are kept and the other labels on $C(F)$ are "erased." If this is done for each distinct $\mathbf{Q}$-name, the resulting partition $\bar{\mathbf{R}}$ has a labeled part with labels (k, i) with $i \leqslant k$, and a (temporarily) unlabeled part of measure $\beta = 1 - \alpha$. For $k = 1, ..., n$ and $\bar{R}_k$ the set consisting of all points with first label k, we have by our choice of β that

$\mu(\bar{R}_k) \leqslant \phi_k^n$. Partition the unlabeled part of $C(F_\beta)$ of each m-name $C(F)$ (where $F_\beta = F \backslash F_\alpha$) into n sets F_β^k, $k = 1, ..., n$. For each $k = 1, ..., n$ starting at F_β^k attatch the labels $(k, 1), (k, 2), ..., (k, k), (k, 1), ...$ until the top level of $C(F_\beta^k)$ is reached. By appropriately choosing $\mu(C(F_\beta^k))$ (so that $\mu(\bar{R}_k) + \sum_F \mu(C(F_\beta^k)) = \phi_k^n$) we obtain a partition $\mathbf{R}$ with $W(\mathbf{R}) \subset T(\{E_i\})$, $d_c(\mathbf{R}) = \phi^n$, and $\|\mathbf{R} - \mathbf{Q}\| \leqslant 2\beta = 2\, |d_c(\mathbf{Q}) - \pi^n|/s_n$.

When Lemma 3 is changed by replacing d_b with d_c, and π by ϕ, the distance between the resulting partition $\mathbf{S}$ and the given $\mathbf{R}$ then becomes $\|\mathbf{S} - \mathbf{R}\| \leqslant 2\phi^{n+1}/\phi^N$. The last statement of Lemma 3 on the contraction factor is not needed for Theorem 3. The proof that Lemma 3 can be so changed, uses the same ideas employed in the previous paragraph, and is not repeated.

Theorem 3 then follows from these modifications of the lemmas in the same way that Theorem 1 follows from Lemmas 1, 2, and 3.

References

1. S. Alpern, Generic properties for measure preserving homeomorphisms, *in* "Ergodic Theory, Proceedings Oberwolfach 1978," pp. 16–27, Lecture Notes in Mathematics, Vol. 729, Springer-Verlag, Berlin/New York, 1979.
2. S. Alpern, Return times and conjugates of an antiperiodic transformation, *Ergodic Theory Dynamical Systems* **1** (1981), 135–143.
3. S. Alpern and V. S. Prasad, Coding a stationary process to one with prescribed marginals, *Ann. Probab.* **17** (1989), 1658–1663.
4. R. Chacon and N. Friedman, Approximation and invariant measure, *Zeit. Wahr. Geb.* **3** (1965), 286–295.
5. J. R. Choksi and S. Kakutani, Residuality of ergodic measurable transformations and of ergodic transformations which preserve an infinite measure, *Indiana Univ. Math. J.* **28** (1979), 453–469.
6. N. Friedman, "Introduction to Ergodic Theory," Van Nostrand–Reinhold, New York, 1970.
7. C. Grillenberger and U. Krengel, On marginal distributions and isomorphisms of stationary processes, *Math. Z.* **149** (1976), 131–154.
8. J. C. Kieffer, On coding a stationary process to achieve a given marginal distribution, *Ann. Probab.* **8** (1980), 131–141.
9. J. Lindenstrauss, A short proof of Liapunoff's convexity theorem, *J. Math. Mech.* **15** (1966), 971–972.