Towers, Conjugacy and Coding

Steve Alpern
Department of Mathematics, London School of Economics
Houghton Street, London, WC2A 2AE, UK

V. Prasad
Department of Mathematics, University of Massachussetts Lowell,
Lowell, Massachussetts, USA

CDAM Research Report CDAM-LSE-2005-21, December 2005-12-09

ABSTRACT. We consider three theorems in ergodic theory concerning a fixed aperiodic measure preserving transformation σ of a Lebesgue probability space $(X, \mathcal{A}, \mu)$ and show that these theorems are all equivalent. Two of these results concern the existence of a partition of the space X with special properties. The third theorem asserts that the conjugates of σ are dense in the uniform topology on the space of automorphisms. The first partition result is Alpern's generalization of the Rokhlin Lemma, the so-called Multiple Rokhlin Tower theorem stating that the space can be partitioned into denumerably many columns and the measures of the columns can be prescribed in advance; the second partition result is a coding result which asserts that any mixing Markov chain can be represented by σ and some partition of the space indexed by the state space of the Markov chain.

TOWERS, CONJUGACY AND CODING

S. ALPERN AND V.S. PRASAD

We dedicate this paper to the memory of Shizuo Kakutani. His kind manner and strong presence will be missed.

Abstract. We consider three theorems in ergodic theory concerning a fixed aperiodic measure preserving transformation σ of a Lebesgue probability space $(X, \mathcal{A}, \mu)$ and show that these theorems are all equivalent. Two of these results concern the existence of a partition of the space X with special properties. The third theorem asserts that the conjugates of σ are dense in the uniform topology on the space of automorphisms. The first partition result is Alpern's generalization of the Rokhlin Lemma, the so-called Multiple Rokhlin Tower theorem stating that the space can be partitioned into denumerably many columns and the measures of the columns can be prescribed in advance; the second partition result is a coding result which asserts that any mixing Markov chain can be represented by σ and some partition of the space indexed by the state space of the Markov chain.

1. Introduction

Suppose $(X, \mathcal{A}, \mu)$ is a Lebesgue probability space with σ an invertible measure preserving bijection of the measure space X onto itself, with sigma-algebra $\mathcal{A}$ and probability measure μ (we will say σ is an automorphism of $(X, \mathcal{A}, \mu)$). We assume that σ is aperiodic; i.e., the measure of the points which are periodic for σ is a μ-null set. In this paper we consider three theorems about such an automorphism σ and show that all three theorems are equivalent: two of these results assert the existence of a partition of X with special properties under the action of σ, and the other result states that the conjugates of σ form a dense class in the space of automorphisms in a certain topology on the automorphisms.

The two partition results are: the Multiple Rokhlin Tower decomposition for σ [3], a generalization of Rokhlin's Lemma, (the latter is one of the basic constructions in ergodic theory — see for example Kornfeld's survey [21]); the Coding theorem [6] shows that given any mixing Markov chain $\mathbf{P}$, there is a partition of the space X so that σ moves the partition elements according to the transitions prescribed by the Markov chain $\mathbf{P}$ — this result is a reformulation of a coding question of Kiefer that asks if a stationary stochastic process can be coded to have prescribed marginal distributions. The third theorem, the Conjugacy theorem of the title, is Alpern's Approximate Conjugacy Theorem [3] for σ, which states that except for a previously prescribed set of small measure, there is some conjugate of σ that agrees pointwise with a given weak mixing automorphism.

Date: Draft March 19, 2005.

1991 *Mathematics Subject Classification.* 37.

Key words and phrases. Rokhlin towers, conjugacy, weak mixing, stationary, aperiodic irreducible markov chain.

To show the equivalence of these theorems, there are three implications to prove. Only one of these implications in the equivalence proof is new. Because of recent activity involving these results (see the next section) we feel it might be useful to point out why and how these theorems are connected. A good reference for these theorems is our book *Typical dynamics of volume preserving homeomorphisms* [8, Appendix 1]: it contains proofs of each of the three Theorems (as well as the first two implications in Section 3).

We give a formal statement of these results and describe some applications of these results. This is followed by a proof of the main result that the three theorems are logically equivalent.

2. Three theorems from ergodic theory

Suppose $(X, \mathcal{A}, \mu)$ is Lebesgue probability space and σ an automorphism of the measure space X onto itself which is aperiodic. We consider in detail the three results mentioned in the previous section.

2.1. **Towers.**

S. Alpern has proven the following theorem which extends Rokhlin's Lemma (see [21], for a survey and history of this "theorem" of Rokhlin). We will refer to this Mutiple Rokhlin Tower theorem as MRT.

Theorem 2.1 (MRT)**.** *Let σ be an aperiodic μ-preserving automorphism of a Lebesgue probability space $(X, \mathcal{A}, \mu)$ onto itself. Let $\pi = (\pi_1, \pi_2, \ldots)$ be a probability distribution such that $\{k : \pi_k > 0\}$ is a relatively prime set of positive integers. Then there is a partition $\mathcal{P} = \{P_{k,i} : k = 1, 2, \ldots, \infty; i = 1, 2, \ldots, k\}$ of X so that for each integer $k = 1, 2, \ldots, \infty$*

1. *$P_{k,i} = \sigma^{i-1}(P_{k,1})$ for $i = 1, 2, \ldots, k$.*
2. *$\mu(\cup_{i=1}^{k} P_{k,i}) = \pi_k$.*

We call $\mathcal{P}$ a multitower partition for σ.

The Rokhlin Lemma asserts that for σ as above, and any positive integer n and any positive number $\epsilon > 0$, there is a subset $R \in \mathcal{A}$, so that the first n iterates of R are disjoint and $\mu(\cup_{i=0}^{n-1} \sigma^i(R)) = 1 - \epsilon$. The Multiple Rokhlin Tower theorem applied to the probability vector π with $\pi_1 = \epsilon$ and $\pi_n = 1 - \epsilon$ yields a set $P_{n,1}$, which, if we take to be R, gives the Rokhlin Lemma . We note that the Multiple Rokhlin Tower theorem states that more generally (than the Rokhlin Lemma) that an aperiodic automorphism can be represented by (denumerably many) columns $(\cup_{i=1}^{k} P_{k,i})$ of given heights (k) and given measures (π_k) as long as the heights are relatively prime. The MRT asserts the existence of a set $P = \cup_k P_{k,1}$ whose relative distribution of first return times to P is given by π. Note that if d divides all k for which $\pi_k > 0$ then for any set P whose relative distribution of return times to P is given by π and for any m which is not a multiple of d, we would have $\mu(P \cap \sigma^m(P)) = 0$. If σ is mixing for example, such a set cannot exist.

For finite dimensional distributions π, Theorem 2.1 was obtained by Alpern [2] (see [12] for another proof which uses Kakutani's proof of the Rokhlin Lemma) and the general case is proved in [3]. See also [25] for another proof by Prikhod′ko and Ryzhikov. An extension of Theorem 2.1 for $\mathbb{Z}^d$-actions [24], has also been obtained by Prikhod′ko. We also note for finite dimensional distributions, that Grillenberger and Krengel [13] previously considered a deep generalization of Theorem 2.1 where the partition of X is required to generate the sigma algebra. Extensions of the

MRT to automorphisms of an infinite measure spaces can be found in [5] and [11]; for aperiodic nonsingular automorphisms of a Lebesgue space see [7].

The MRT is one of the main tools used by Alpern and Alpern-Prasad to prove that any measure theoretic property which is generic for automorphisms of a Lebesgue measure space is a generic property for measure preserving homeomorphisms of manifolds. This forms the main topic of our book [8]. In recent work on topolgical (and measurable) orbit equivalence, the MRT has been used by N. Ormes [23] and Kornfeld-Ormes [22] to prove strong orbit equivalence theorems and generalizations of the Jewett-Krieger theorem. A nice survey of the use of MRT and associated Bratteli diagrams to obtain Vershik adic-maps in topological orbit equivalence theory can be found in [21].

2.2. **Conjugacy.** We refer to the following Approximate Conjugacy Theorem as ACT (note in our book [8], the ACT is referred to as the Pointwise Conjugacy Theorem); the ACT is used to prove the denseness of the conjugacy class of the automorphism σ.

Theorem 2.2 (ACT). *Let σ be an aperiodic μ-preserving automorphism of a Lebesgue probability space $(X, \mathcal{A}, \mu)$ onto itself and let τ be a weak mixing invertible measure preserving automorphism of a Lebesgue space $(Y, \mathcal{B}, \nu)$. Then for any $F \in \mathcal{B}$ with $\nu(F) < 1$, there is conjugate of σ which agrees with τ on F; i.e., there is an invertible measure preserving bijection*

$$\psi : (Y, \mathcal{B}, \nu) \to (X, \mathcal{A}, \mu)$$

such that

$$\begin{aligned} \hat{\sigma}(y) &= \psi^{-1}\sigma\psi(y) \\ &= \tau(y) \end{aligned}$$

for ν-a.e. $y \in F$.

This theorem asserts that the conjugates $\mathcal{C}(\sigma)$ of an aperiodic automorphism are dense in the weak mixing automorphisms of $(X, \mathcal{A}, \mu)$ if the distance between two automorphisms σ, τ is given by $\mu\{x \in X : \sigma(x) \neq \tau(x)\}$.

In approximation problems in ergodic theory, the denseness of $\mathcal{C}(\sigma)$, is used to show that if a particular aperiodic automorphism σ has a certain measure theoretic property in the space of automorphisms, then there is dense class of automorphisms possessing the same property — namely, its conjugacy class $\mathcal{C}(\sigma)$, in the group of all automorphisms with the weak topology [17]. For infinite measure spaces a version of the ACT can be found in [9], see also [8]. We further note that in their memoir [1], Akin et al define a (metric) space X to have the Rokhlin property if there is some homeomorphism of X whose conjugates (by other homeomorphisms) is dense in the space of homeomorphisms (they show that when X is the Cantor set, it has the Rokhlin property; further they obtain a relative Rokhlin property for circle homeomorphisms).

2.3. **Coding.** We refer to the next result as the CMC (for Coding Markov Chains) Theorem. Given a mixing Markov Chain $\mathbf{P}$, we ask if it is possible to "represent" $\mathbf{P}$, by the aperiodic automorphism σ and some partition $\mathcal{P}$. The theorem below explains in what sense this is possible.

Theorem 2.3 (CMC)**.** *Let σ be an aperiodic μ-preserving automorphism of a Lebesgue probability space $(X, \mathcal{A}, \mu)$ onto itself. Suppose that $\mathbf{P} = (p(i,j) : i, j \in \mathbb{N})$ is the transition probability matrix for a positive recurrent, aperiodic, irreducible Markov chain with state space $\mathbb{N}$ the set of positive integers. Let $\mathbf{p} = (p(i) : i \in \mathbb{N})$ be the unique positive invariant distribution $\mathbf{pP} = \mathbf{p}$ (i.e., $p(j) = \sum_{i \in \mathbb{N}} p(i)p(i,j)$). Then there is a partition $\mathcal{P} = \{P_i : i \in \mathbb{N}\}$ of X such that for all $i, j \in \mathbb{N}$*

$$\mu\left(P_i \cap \sigma^{-1} P_j\right) = \mu(P_i)p(i,j) = p(i)p(i,j). \tag{2.1}$$

The result above was proved by us in [6]. This extended to denumerable state spaces, J. Kieffer's finite state space result [20] where he considered the question of coding a stationary stochastic process to one with prescribed marginal distributions. After reformulation, Kieffer notes that Theorem 2.3 for finite state space Markov Chains and ergodic σ follows from a deep result of Grillenberger and Krengel [13] where they also consider the question of obtaining partitions which generate the sigma algebra $\mathcal{A}$. Kieffer also obtains a "universal" partition $\mathcal{P}$ which satisfies (2.1) for all antiperiodic automorphisms σ. Cohen [10] considers the following finite dimensional variant of Theorem 2.3, the so-called rotational representations of $\mathbf{P}$: The transition probabilities $(p(i,j) : 1 \leq i, j \leq n)$ are given and it is required to find a circle rotation σ and a circle partition $\mathcal{P}$ consisting of intervals satisfying (2.1). See also the work of Alpern [4], Haigh [14], Kalpazidou [19], and Kalpazidou-Tzouvaras [18] for related developments on rotational representations of stochastic matrices, and cycle decompositions of Markov chains.

3. Proofs of Equivalencies

We show that the three results above are equivalent.

Theorem 3.1. *Let σ be an aperiodic automorphism of a Lebesgue probability space $(X, \mathcal{A}, \mu)$ onto itself. Then the three theorems for σ, the Multiple Rokhlin Tower Theorem (MRT, Theorem 2.1), the Approximate Conjugacy Theorem (ACT, Theorem 2.2), and the Coding Markov Chain Theorem (CMC, Theorem 2.3), are all equivalent.*

The theorem will be proved by showing the following three implications. Again we note that the first two implications are in Appendix 1 of [8], but in order to make this paper self contained, we include the proofs of all of these implications.

(1) MRT $\Rightarrow$ ACT
(2) ACT $\Rightarrow$ CMC
(3) CMC $\Rightarrow$ MRT

Proof: MRT $\Rightarrow$ ACT

Let $\tau : (Y, \mathcal{B}, \nu) \to (Y, \mathcal{B}, \nu)$ be a weak mixing automorphism of the Lebesgue space $(Y, \mathcal{B}, \nu)$ and let $F \in \mathcal{B}$ be any set with $\nu(F) < 1$. Since τ is ergodic the τ-orbit of every point $y \in F$ eventually exits F. For each integer $k = 2, 3, \ldots, \infty$, set $R_{k,1} \subset F \setminus \tau(F)$ to be the set of points in $F \setminus \tau(F)$ whose τ-orbit first leaves F on the $(k-1)$th iterate. Setting $R_{k,i} = \tau^{i-1}(R_{k,1})$ for $i = 1, 2, \ldots, k$, we note that

$$F \cup \tau(F) = \cup_{k=2}^{\infty} \cup_{i=1}^{k} R_{k,i}$$

Set $R_{1,1} = Y \setminus (F \cup \tau(F))$ and let $\mathcal{R} = \{R_{k,i} : k = 1, 2, \ldots, \infty; i = 1, 2, \ldots, k\}$. Then $\mathcal{R}$ is a multitower partition of Y for τ, with column distribution $\pi = (\pi_1, \pi_2, \ldots)$, given by $\pi_k = \nu(\cup_{i=1}^{k} R_{k,i})$.

We note that the $\gcd\{k : \pi_k > 0\} = 1$: if $\nu(R_{1,1}) > 0$, then $1 \in \{k : \pi_k > 0\}$ and $\{k : \pi_k > 0\}$ is a relatively prime set of integers; on the other hand, if $\nu(R_{1,1}) = 0$, then $Y = F \cup \tau(F)$ and if $\gcd\{k : \pi_k > 0\} = p$, then all column heights are multiples of p. Then for any k with $\pi_k > 0$, the set $D = R_{k,1}$ satisfies $\nu(D \cap \tau^{np+1}(D)) = 0$ for all n. But then this contradicts the assumption that τ is weak mixing. Thus the distribution π satisfies the hypotheses of the MRT, Theorem 2.1.

Let $\mathcal{P} = \{P_{k,i}\}$ be the multitower partition for σ on $(X, \mathcal{A}, \mu)$ with distribution π. Since $\nu(R_{k,1}) = \mu(P_{k,1})$ for each k, there is a measure preserving bijection $\psi : \cup_k R_{k,1} \to \cup_k P_{k,1}$. Extend $\psi : (Y, \mathcal{B}, \nu) \to (X, \mathcal{A}, \mu)$ by setting for $y \in R_{k,i}$, $\psi(y) = \sigma^{i-1}\psi\tau^{1-i}(y)$ for all k with $\pi_k > 0$ and $i = 1, 2, \ldots k$. Since τ and $\psi^{-1}\sigma\psi$ differ only on the "top" of the tower (which is a subset of $Y \setminus F$), it follows that $\tau(y) = \psi^{-1}\sigma\psi(y)$ for all $y \in F$.

Proof: ACT $\Rightarrow$ CMC

We represent the Markov chain $\mathbf{P} = (p(i,j) : i, j \in \mathbb{N})$ as a shift transformation τ on $Y = \mathbb{N}^\infty$, the two-sided space of infinite sequences $y = (\ldots, y_{-1}, y_0, y_1, \ldots)$ with all of the $y_i \in \mathbb{N}$, the symbol space of the Markov chain. Denote by ν, the shift invariant measure, defined on the cylinders in Y via the invariant $\mathbf{p}$ and $\mathbf{P}$. The assumptions on $\mathbf{P}$ ensure that the automorphism τ is mixing. Let $Q_i = \{y \in \mathbb{N}^\infty : y_0 = i\}$ be the time 0 partition of Y. Observe that $\frac{\nu(\tau(Q_i) \cap Q_j)}{\nu(Q_i)} = p(i,j)$. We assume that $p(0) > 0$, so that $\nu(Q_0) > 0$.

Apply the Approximate Conjugacy Theorem to $F = Y \setminus \tau^{-1}\{y \in Y : y_1 \neq 0\}$ to get a measure preserving bijection

$$\psi : (Y, \mathcal{B}, \nu) \to (X, \mathcal{A}, \mu)$$

Setting $\mathcal{P} = \{P_i : i \in \mathbb{N}\}$ where $P_i = \psi^{-1}(Q_i)$ gives the required partition.

Proof: CMC $\Rightarrow$ MRT

Given a multitower configuration for σ given by the probability distribution $\pi = (\pi_1, \pi_2, \ldots)$ we model the multitower as a Markov Chain whose state space consists of the levels of the multitower. Define the denumerable (multitower) state space $\mathcal{T}$ by

$$\mathcal{T} = \{(k,i) : k \text{ is postive integer with } \pi_k > 0 \text{ and for these } k,\ i = 1, 2, \ldots, k\},$$

on which we put an initial probability distribution $\mathbf{p} = \{p((k,i)) : (k,i) \in \mathcal{T}\}$, with $p((k,i)) = \pi_k/k$ for each $(k,i) \in \mathcal{T}$. We define a Markov Chain on $\mathcal{T}$ by defining a transition $(k,i) \to (k',i')$ to be legal in $\mathcal{T}$, if $k = k'$ and $i' = i+1$, or $i = k$ and $i' = 1$. Setting $p = \sum_k \pi_k/k$, then the nonzero transition probabilities $\mathbf{P}$ are

$$P((k,i),(k,i+1)) = 1 \quad \text{if } i < k$$

and

$$P((k,k),(k',1)) = p((k',1))/p.$$

The transition probabilities on $\mathcal{T}$ with the stationary invariant probability $\mathbf{p}$ defines a "mixing" Markov chain on $\mathcal{T}$. The partition from the CMC Theorem is our multitower partition $\mathcal{P} = \{P_{(k,i)} : (k,i) \in \mathcal{T}\}$, having column distribution π.

References

[1] Akin, E; Hurley, M.; Kennedy J. , *Dynamics of topologically generic homeomorphisms*, **Memoirs of the American Mathematical Society, 164** (2003), no. 783, viii+130pp, **MR: 2004j:37024)**.

[2] Alpern, S. , Generic properties for measure preserving homeomorphisms in *Ergodic Theory, Proceedings Oberwolfach 1978*, *Lecture Notes in Mathematics* **729**, 16–27 (1979), **MR: 80m: 28019**

[3] Alpern, S. , Return times and conjugates of an antiperiodic transformation *Ergodic Theory and Dynamical Systems* **1**, 135–143 (1981), **MR: 83i:28020**

[4] Alpern, S., Rotational representations of stochastic matrices, *Ann. Probab.*, **11, (1983), 3**, 789–794, **MR: 85a:15022**.

[5] Alpern, S.; Choksi, J. R.; Prasad, V. S., Conjugates of infinite measure preserving transformations. *Canad. J. Math.* **40 (1988), no. 3**, 742–749. **MR:89m:28026**

[6] Alpern, S.; Prasad, V. S., Coding a stationary process to one with prescribed marginals. *Ann. Probab.* **17 (1989), no. 4,** 1658–1663. **MR:91f:28007**

[7] Alpern, S.; Prasad, V. S., Return times for nonsingular measurable transformations. *J. Math. Anal. Appl.* **152 (1990), no. 2**, 470–487. **MR:91m:28023**

[8] Alpern, S.; Prasad, V. S., *Typical dynamics of volume preserving homeomorphisms.* **Cambridge Tracts in Mathematics, 139.** Cambridge University Press, Cambridge, 2001. xx+216 pp. ISBN 0-521-58287-3. **MR:2002i:37006**

[9] Choksi, J. R.; Kakutani, S., Residuality of ergodic measurable transformations and of ergodic transformations which preserve an infinite measure. *Indiana Univ. Math. J.* **28 (1979), no. 3**, 453–469. **MR:80d:28042**

[10] Cohen, J. E., A geometric representation of a stochastic matrix: theorem and conjecture, *The Annals of Probability*, **9, 1981, 5,** 899–901 **MR: 82j:15018**.

[11] Eigen, S.J.; Hajian, A.B.; Prasad, V.S., Universal skyscraper templates for infinite measure preserving transformations, *preprint.*

[12] Eigen, S. J.; Prasad, V. S., Multiple Rokhlin tower theorem: a simple proof *New York J. Math.* **3A (1997/98)**, *Proceedings of the New York Journal of Mathematics Conference, June 9–13, 1997*, 11–14 (electronic: http://nyjm.albany.edu/8000/j/1997/3A-2.pdf). **MR:99h:28032**

[13] Grillenberger, C.; Krengel, U., On marginal distributions and isomorphisms of stationary processes, *Math. Z.*, **149,(1976)**, 131–154. **MR:53 #11020**.

[14] Haigh, J., Rotational representations of stochastic matrices, Ann. Probab.,**13, (1985), 3**, 1024–1027, **MR: 87a:15031**.

[15] Hajian, A.B.; Ito, Y., Transformations that do not accept a finite invariant measure, *Bull. Amer. Math. Soc.* **84 (1978), no. 3**, 417–427. **MR: 57:#6365**

[16] Hajian, A.B.; Kakutani, S., An example of an ergodic m.p.t. defined on an infinite measure space, *Lecture Notes in Math.* **160, (1970)**, *Springer Verlag*, 45-52.

[17] Halmos, P. R. *Lectures on ergodic theory*, Chelsea Publishing Co., New York, **1960**, vii+101, **MR0111817 (22 #2677)**, originally published by Publications of the Mathematical Society of Japan, **no. 3**, The Mathematical Society of Japan, **1956**, vii+99, **MR0097489 (20 #3958)**.

[18] Kalpazidou, S. and Tzouvaras, L., The forward and backward rotational decompositions of Markov chains, *Stochastic Anal. Appl.*, **19, 2001, 3**, 399–412, **MR: 2002e:60110**.

[19] Kalpazidou, S.,Cycle representations of Markov processes: an application to rotational partitions, *Stochastic processes and related topics*, in *Trends Math.*,253–273, Birkhauser Boston, Boston, MA, **1998 99i:60138**.

[20] Kieffer, J. C., On coding a stationary process to achieve a given marginal distribution, *Ann. Probab.*, **8 (1980)**, 131–141. **MR:81d:28021**.

[21] Kornfeld, I. Some old and new Rokhlin towers *Contemporary Mathematics*, **356 (2004)**, 145–169.

[22] Kornfeld, I.; Ormes, N., Topological realizations of families of automorphisms, multitowers and orbit equivalence, *preprint.*

[23] Ormes, N., Strong orbit realization for minimal homeomorphisms., *J. Anal. Math.*, **71 (1997)**, 103–133. **MR:99c:28053**.

[24] Prikhod′ko, A. A., Partitions of the phase space of a measure-preserving $\mathbf{Z}^{\mathbf{d}}$-action into towers, **Mat. Zametki, 65, (1999), 5**, 712–725, **MR: 2001a:37007**.

[25] Prikhod′ko, A. A.; Ryzhikov, V. V., The Rokhlin-Halmos-Alpern maximal lemma. (Russian. Russian summary) *Vestnik Moskov. Univ. Ser. I Mat. Mekh.*, **no. 3, 92**, 37–41, (1996) ; translation in *Moscow Univ. Math. Bull.*, **51, no. 3**, 31–34, (1996). **MR:98k:28026**.

London School of Economics
E-mail address: s.alpern@lse.ac.uk

University of Massachusetts Lowell
E-mail address: vidhu_prasad@uml.edu