CUTTING AND STACKING IN ERGODIC THEORY

By

Carole Agyeman–Prempeh

A thesis
submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Science in the Department of Mathematics

Chapel Hill
2010

Approved by:

Dr. Idris Assani

Dr. Sue Goodman

Dr. Joseph Plante

Abstract

Carole Agyeman-Prempeh : Cutting and Stacking in Ergodic Theory
(Under the direction of Dr. Idris Assani)

This thesis looks at constructing transformations using cutting and stacking methods. It focuses mainly on the construction of Chacon's transformation. This transformation provides an example of a measure-preserving transformation which is weakly mixing and not strongly mixing.

Acknowledgments

My deepest gratitude and appreciation goes to everyone who has helped and encouraged me during my postgraduate study. Your prayers, support and kind words are what kept me going during the tough times.

I am very thankful to my thesis adviser, Professor Idris Assani, for his guidance, support and patience. Under his instruction I have gained a deeper understanding of mathematics during my graduate career at UNC Chapel Hill. I was introduced to Ergodic theory by him and have amassed a lot of knowledge in cutting and stacking methods during my thesis writing.

I would like to thank members of my masters committee, Dr. Sue Goodman and Dr. Joseph Plante, for agreeing to be on my committee and for their insightful comments and suggestions.

Above all, I thank God for the opportunity given to me to study at UNC and the strength to complete my graduate study successfully.

Carole Agyeman-Prempeh

Table Of Contents

List Of Figures

Chapter 1

Introduction

1.1 Preliminaries

The goal is to define and exhibit some properties of an important measure-preserving system. This system was created in 1969 by R.V. Chacon [1]. The transformation T defined on this system is called Chacon's transformation. The Chacon example and its contstruction has become one of the fundamental examples in Ergodic theory. First we give some preliminaries, the definition and simple examples of measure-preserving systems.

We consider first a set X then define on it a σ- algebra, a measure and a measure-preserving transformation.

Definition 1.1.1. A collection of subsets of $(X, \mathcal{B})$ is defined as a σ-algebra if

1. $\emptyset \in \mathcal{B}$
2. $\mathcal{B}$ is closed under countable union and intersections.
3. $\mathcal{B}$ is closed under complements.

Definition 1.1.2. A set A is measurable if A belongs to the σ-algebra.

The pair $(X, \mathcal{B})$ forms a measurable space. In order to define a measure we first need to state the set of interest X and the σ-algebra $\mathcal{B}$. The set of interest here will be the unit interval $[0, 1]$ and the σ-algebra will consist of the Lebesgue measurable subsets of $[0, 1]$.

The measure which is defined on the elements of the σ-algebra is the Lebesgue measure μ. The triple $([0,1], \mathcal{B}, \mu)$ now makes up the measure space of interest.

Definition 1.1.3. A transformation $T : X \to X$ is said to be measure-preserving if for all $A \in \mathcal{B}$

$$\mu(T^{-1}A) = \mu(A).$$

An example of measure-preserving transformation defined on $[0,1]$ is the translation map $T(x) = x + \alpha \pmod 1$ with $\alpha \in (0,1)$. Figure 1.1.1 shows this graphically.

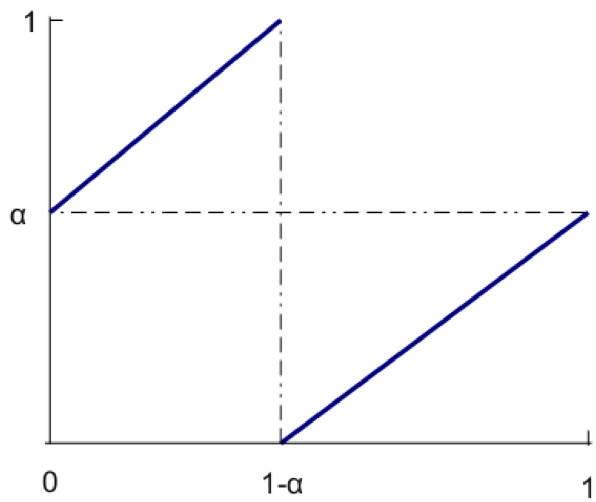

Fig. 1.1.1: Graph of the translation $T(x) = x + \alpha$ with $\alpha \in (0,1)$

Another measure-preserving transformation defined $[0,1]$ is $T : [0,1] \to [0,1]$ such that $Tx = 2x \pmod 1$ or more generally $Tx = nx(\text{mod}1)$. $Tx = 2x \pmod 1$ is seen graphically in Figure 1.1.2.

Proposition 1.1.1. *$T(x) = nx$ (mod 1) is measure preserving.*

Proof. Step 1: We first show T is measure-preserving on open intervals.

Given an interval $A \in \mathcal{B}$,

$$\begin{aligned} T^{-1}A &= \bigsqcup_{j=1}^{n} A_j \quad \text{disjoint union} \\ \mu(T^{-1}A) &= \sum_{j=1}^{n} \mu(A_j) \quad \text{because } \mu \text{ is } \sigma\text{-additive.} \end{aligned}$$

Using the fact that the slope of the line, s, is given by

$$
\begin{aligned}
s = \frac{\Delta y}{\Delta x} &= \frac{1}{\frac{1}{n}} = n \\
\Rightarrow \ \Delta y &= n\Delta x \\
\text{and thus } \mu(A) &= n\ \mu(A_j)\ \forall j.
\end{aligned}
$$

From this we see that each A_j has measure $\frac{1}{n}\mu(A)$ and therefore

$$
\begin{aligned}
\mu(T^{-1}A) &= \sum_{j=1}^{n} \frac{1}{n}\mu(A) \\
&= \mu(A).
\end{aligned}
$$

We have thus shown that the transformation T is measure-preserving on open intervals.

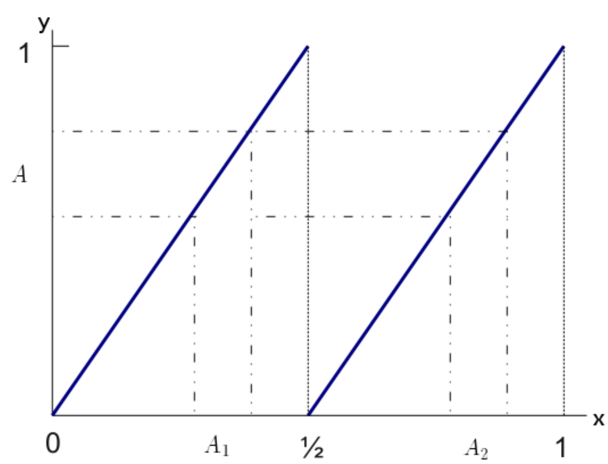

Fig. 1.1.2: Graph of the double map $T(x) = 2x \pmod 1$

Step 2: The result above can be extended to open sets then in turn extended to the set of Lebesgue measurable sets (Lebesgue measurable σ-algebra). The proof of this result is given below for open sets.

Observe that for an open set, O,

$$O = \bigcup (a_n, b_n), \text{ open, countable, and disjoint intervals}$$

$$
\begin{aligned}
T^{-1}O &= T^{-1}(\cup(a_n,b_n)) \\
\mu(T^{-1}O) &= \mu(T^{-1}(\cup(a_n,b_n))) \\
&= \mu(\cup T^{-1}(a_n,b_n)) \text{ since the intervals are disjoint} \\
&= \sum^{n} \mu T^{-1}(a_n,b_n) \\
&= \sum \mu(a_n,b_n) \\
&= \mu(O).
\end{aligned}
$$

Hence T is measure-preserving on open sets.

Step 3: To move to Lebesgue measurable sets, notice that for any $A \in \mathcal{B}$ and for all $\varepsilon > 0$, $\exists$ O an open set such that $A \subset O$ and $\mu(O \backslash A) < \varepsilon$. Pick O_1 and O_2 such that $A \subset O_1$, $A \subset O_2$. Then $A \subset (O_1 \cap O_2)$. Let $\bar{O_2}$ denote $O_1 \cap O_2$. Again, pick O_3 such that $A \subset \bar{O_3}$ where $\bar{O_3} = O_1 \cap O_2 \cap O_3$. Continuing this process generates a decreasing sequence of open sets such that $A \subset \bar{O_n}$. Let $\mathbb{I}$ denote the characteristic function on some set. Then

$$
\begin{aligned}
\mathbb{I}_{\bar{O_n}} &\longrightarrow \mathbb{I}_A \text{ a.e thus} \\
\mu(\bar{O_n}) &\longrightarrow \mu(A). \text{ Also since} \\
T^{-1}(A) &\subset T^{-1}(\bar{O_n}) \text{ then} \\
\lim_{n\to\infty} \mu(\bar{O_n}) = \lim_{n\to\infty} \int \mathbb{I}_{\bar{O_n}} d\mu &= \int \lim_{n\to\infty} \mathbb{I}_{\bar{O_n}} \text{ by the MCT} \\
&= \mu(A).
\end{aligned}
$$

Hence we have shown that $\lim_{n\to\infty} \mu(\bar{O_n}) = \mu(A)$. Note also that

$$
\begin{aligned}
\bigcap_{n=1}^{\infty} \bar{O_n} &= A \text{ and that} \\
\bigcap T^{-1}(\bar{O_n}) &= T^{-1}(\bigcap_{n=1}^{\infty} \bar{O_n}) \\
\mu(\bigcap T^{-1}(\bar{O_n})) &= \mu(T^{-1}(\bigcap_{n=1}^{\infty} \bar{O_n})) \\
\lim_{n\to\infty} \mu(T^{-1}(\bar{O_n})) &= \lim_{n\to\infty} \mu(\bar{O_n}) \\
&= \mu(A).
\end{aligned}
$$

Therefore T is measure-preserving on Lebesgue measurable subets as well. $\square$

We continue by giving more preliminaries that will be needed to prove one of the properties of Chacon's transformation.

Definition 1.1.4. A set of nonnegative integers D is said to be of density zero if

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{I}_D(i)=0.$$

If this is not the case, then the $\limsup_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{I}_D(i)$ is strictly positive and the set D is described as having positive upper density.

Definition 1.1.5. Let $\{a_i\}$ be a bounded sequence of real numbers which may converge to a number a. Given are three notions of convergence. We say

1. The sequence $\{a_i\}$ converges to a if

$$\lim_{i\to\infty} a_i - a = 0.$$

2. $\{a_i\}$ has strong Cesaro convergence to a if

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}|a_i-a|=0 \text{ and}$$

3. $\{a_i\}$ is described to have Cesaro convergence to a if

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}a_i-a=0$$

Stated without proof, note that if $\{a_i\}$ is a bounded sequence, convergence to a implies strong Cesaro convergence which in turn implies Cesaro convergence. Also take note that the converse implications do not hold.

Definition 1.1.6. A sequence $\{a_i\}$ of real numbers is said to converge in density to a point a if there exists a set of density zero, D, such that for every $\varepsilon > 0$ there is an integer N such that whenever $i > N$ and $i \notin D$, $|a_i - a| < \varepsilon$. This way of looking at convergence will be used in the subsequent chapter to prove the weakly mixing property of the Chacon transformation.

We now define what it means for a transformation T to be ergodic.

Definition 1.1.7. Let $(X, \mathcal{B}, \mu)$ be a probability measure space. A measure-preserving transformation $T : X \to X$ is said to be ergodic if and only if for all A, T-invariant sets, $\mu(A) = 0$ or $\mu(A^c) = 0$.

An important result in the study of ergodic theory is the Mean ergodic theorem which is stated as follows:

Theorem 1.1.1. *Let T be a measure-preserving transformation on a probability space $(X, \mathcal{B}, \mu)$ and let $f \in L^1$. Then there exists $f^* \in L^1$ with $f^* \circ T = f^*$ a.e and*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k = f^*$$

in L^1 norm. Also $\int f^* d\mu = \int f d\mu$.

There are several ways of stating the ergodicity property. Given below are a few.

Lemma 1.1.1. *The following are equivalent:*

1. *T is ergodic.*
2. *For every $A, B \in \mathcal{B}$*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) = \mu(A)\mu(B).$$

3. *For all $A, B \in \mathcal{B}$ of positive measure, there exists some integer k such that $\mu(T^{-k}A \cap B) > 0$.*

Proof. (3) implies (1). Let A be a T-invariant set of positive measure i.e, for all $k > 0$, $T^{-k}(A) = A$. Set $B = A^c$. Then if B has positive measure there would exist some integer k so that $\mu(T^{-k}A \cap B) = \mu(T^{-k}A \cap A^c) > 0$. This is never true and hence we have a contradiction. Therefore $\mu(A^c) = 0$ which makes T ergodic.

(1) implies (2). The function $f^* \in L^1$ is given by $f^* = \mathbb{E}^{\mathcal{I}}(f)$, the expectation of f. If T is ergodic, then $\mathcal{I}$ is trivial and thus $\mathcal{I} = \{\emptyset, X\}$. Thus $f^* = \int f d\mu$. Suppose T is ergodic.

Note that with $f = \mathbb{I}_A \times \mathbb{I}_B$, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \left(\mathbb{I}_A \circ T^k \cdot \mathbb{I}_B \right) = \lim_{n\to\infty} \left(\frac{1}{n} \sum_{k=0}^{n-1} \mathbb{I}_A \circ T^k \right) \mathbb{I}_B.$$

Then by the ergodic theorem

$$\begin{aligned} \lim_{n\to\infty} \int \left(\frac{1}{n} \sum_{k=0}^{n-1} \mathbb{I}_A \circ T^k \right) \mathbb{I}_B &= \int \mathbb{I}_A d\mu \int \mathbb{I}_B d\mu \\ &= \mu(A)\mu(B). \end{aligned}$$

(2) implies (3). Suppose (2) is true. For the limit to exist, there should be infinitely many integers k such that

$$\frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}(A) \cap B) > 0$$

which would imply that there are infinitely many k such that $\mu(T^{-k}(A) \cap B) > 0$. □

Stated without proof is a Bochner-Herglotz theorem [2]. This will be needed in the next chapter to help prove the weakly-mixing property of a transformation.

Theorem 1.1.2. *Let $\{\sigma_n\}$ be a positive definite sequence. Then there exists unique non-negative measure μ on $[0, 1]$, such that*

$$\sigma(n) = \int e^{2\pi i n x} d\mu.$$

Chapter 2

Chacon's Transformations

In this chapter we look at two different ways of constructing Chacon's transformation. The first, and more common, method of construction is studied first. The transformation from this method, called the canonical Chacon's transformation, is as studied by Cesar Silva [3]. The second method of construction explored, which is actually the original method, is that seen in Chacon's paper [1].

2.1 Construction of The Canonical Chacon Transformation

We begin stage 1 by dividing the unit interval $[0, 1]$ into two disjoint pieces $\left[0, \frac{2}{3}\right)$ and $\left[\frac{2}{3}, 1\right]$. The second piece will serve as the spacer or reservoir. The first piece, $\left[0, \frac{2}{3}\right)$, is cut into three equal parts and an interval of equal length to each part, $\frac{2}{9}$, is cut from the right side of the spacer. Figure 2.1.1 shows this step. This completes the 'cutting'process.

The order of stacking of the intervals is this: First the middle piece is stacked on top of the left. The interval of equal length from the spacer is next and last in the stacking process, to form tower-1, is the rightmost piece. Figure 2.1.2 gives an illustration.

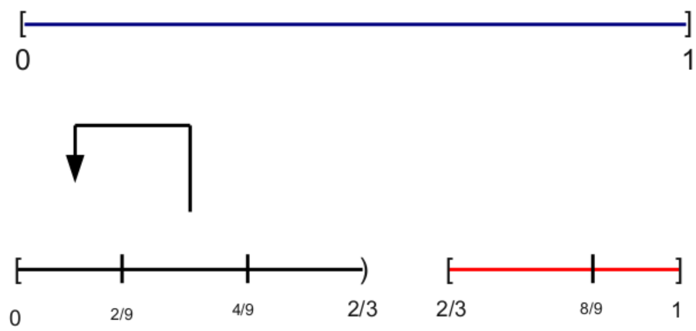

Fig. 2.1.1: Constructing Chacon's transformation at Stage 1

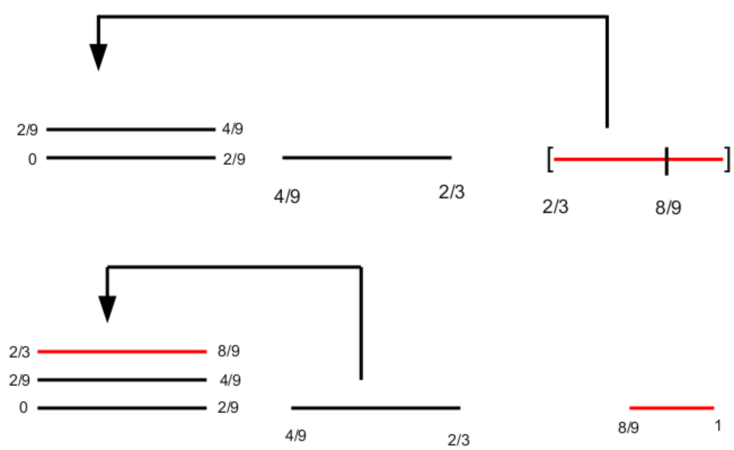

Fig. 2.1.2: Intermediate step of Chacon's process at Stage 1

Fig. 2.1.3: End of Stage 1

Thus at the end of the first stage we have a tower consisting of four levels and the remaining spacer. See Figure 2.1.3. The transformation T at this stage maps the base of the tower linearly to the interval just above it and that interval to the one just above it and so on.

We note that under T the topmost level and the remaining spacer have no images.

The process of cutting and stacking is repeated again in Stage 2. We start off first by slicing the four-tier tower from the end of stage 1, tower-1 for short, into three equal pieces.

An interval of same length as each of the newly cut parts, now $\frac{1}{3} \cdot \frac{2}{9} = \frac{2}{27}$, is separated from the right-hand side of the remaining spacer. Again the order of stacking is that the

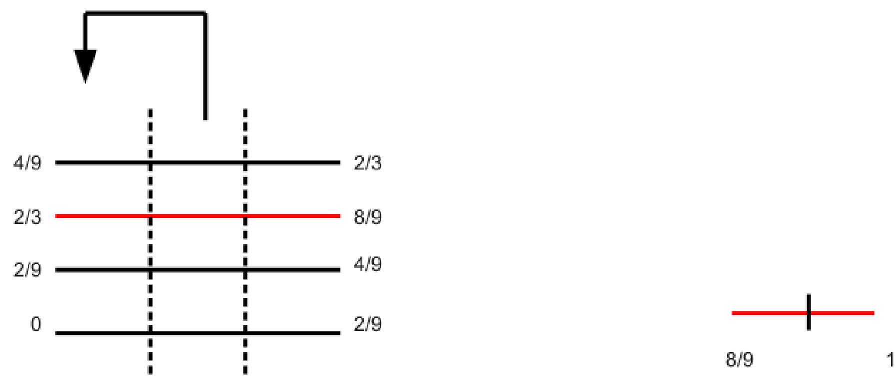

Fig. 2.1.4: Subdividing tower-1 and stacking middle portion at Stage 2

middle portion of the sliced tower is stacked on the left portion, then the spacer and last is the rightmost portion of the tower.

Fig. 2.1.5: Intermediate steps in Stage 2

Thus at the end of stage 2, we have a new tower, tower-2, consisting of 13 levels and the remaining spacer.

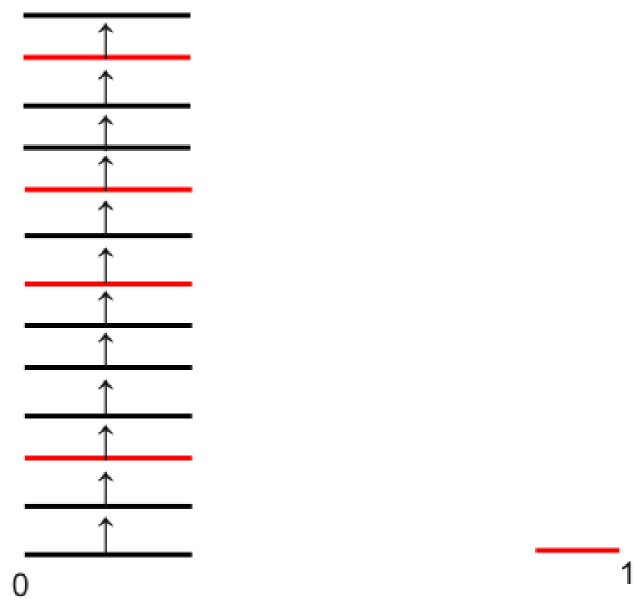

Fig. 2.1.6: End of Stage 2 of the Chacon process

This is an iterative process and at the $(n+1)$th stage of construction the nth tower is again sliced into 3 equal pieces. An interval of equal length to each level of the tower is cut off from the remaining spacer.

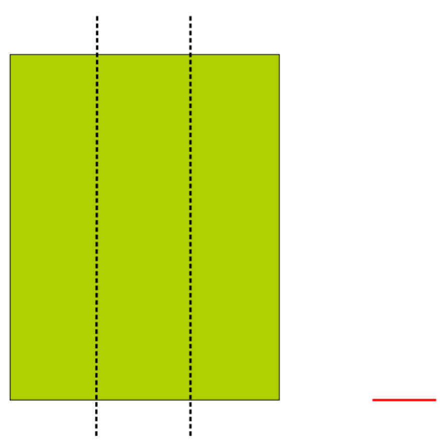

Fig. 2.1.7: Slicing and stacking the nth tower

Stacking first the middle $\frac{1}{3}$ of the tower, then the equal length spacer, then the rightmost portion of the tower we are again presented with a new tower at the end of the nth stage.

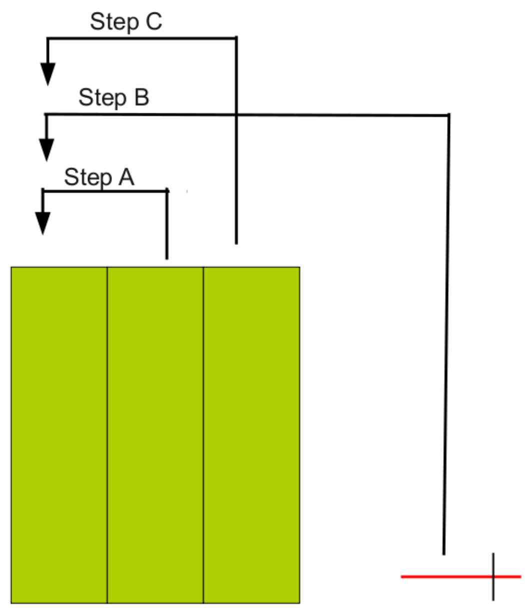

Fig. 2.1.8: Intermediate steps of $(n+1)$th stage

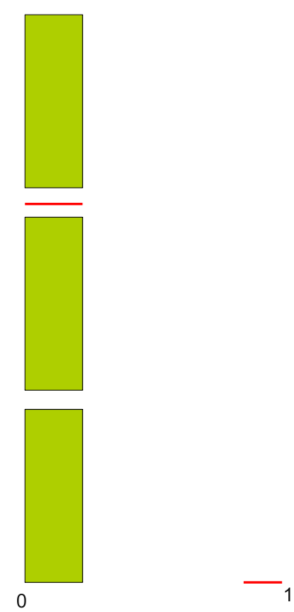

Fig. 2.1.9: End of the $(n+1)$th stage

2.1.1 Some calculations

We now look at some calculations that can be deduced about the Chacon transformation at the **beginning** of the stages.

At the beginning of Stage 1:

† Height of tower = 1.

† Length of spacer $= \frac{1}{3}$.

† Length of each interval (excluding spacer) after slicing $= \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9} = \frac{2}{3^2}$.

At the beginning of Stage 2:

† Height of tower $= (3 \times 1) + 1 = 4$.

† Length of spacer $= \frac{1}{3} - \frac{2}{9} = \frac{1}{9} = \frac{1}{3^2}$.

† Length of each interval after slicing $= \frac{1}{3} \cdot \frac{2}{9} = \frac{2}{27} = \frac{2}{3^3}$

Therefore by induction we observe that at the beginning of the nth stage:

† The height of tower is given by $h_n = 3h_{n-1} + 1$.

† The length of the spacer is given by $\frac{1}{3^n}$ and

† The length of each interval after cutting is $2\left(\frac{1}{3}\right)^{n+1}$.

Thus at the end of the nth stage the transformation T is defined on all the levels except the top level and the remaining spacer. Explicitly, the transformation is defined on a subset of $[0, 1]$ of measure $1 - \frac{1}{3^n} - \frac{2}{3^{n+1}} = 1 - \frac{5}{3^{n+1}}$. We note that the process of cutting and stacking increases the domain of where the transformation is defined and as n gets larger and larger, the limiting transformation T is defined on a subset of $[0, 1]$ which has measure 1.

2.2 Some Properties of Chacon's Transformation

2.2.1 Measure-Preserving Property

The goal is to prove that $\mu(T^{-1}A) = \mu(A)$ for all measurable $A \subset [0,1]$. Since A will be broken up into pieces, at the nth stage of the Chacon process $A = (A \bigcap \text{spacers}) \bigcup (A \bigcap \text{top levels}) \bigcup (A \bigcap \text{middle levels})$. Bear in mind that on the nth stage on the tower, T is not defined on the top level if we move in the forward direction or T is not defined on the bottom levels if we move in the backward direction. Let $\mathscr{T}_n$ denote the tower at the nth stage and $I_{n,k}$ be the levels of the tower at the nth stage and in kth position from the base of the tower.

Lemma 2.2.1. *Given $X_n \subset [0,1]$ and that $\mu(X_n) \to \mu(X) = 1$. We claim that $\mu(B \cap X_n) \to \mu(B)$ for any $B \subset [0,1]$.*

Proof.

$$\begin{aligned} \mu(B) &= \mu\left(B \cap X_n\right) + \mu\left(B \backslash X_n\right) \\ \mu(B) - \mu\left(B \cap X_n\right) &= \mu\left(B \backslash X_n\right) \\ \text{Note that } \mu(B \backslash X_n) &\leq \mu\left([0,1] \backslash X_n\right) \\ &= 1 - \mu\left(X_n\right) \to 0 \text{ as } n \to \infty \text{ since } \mu(X_n) \to 1. \end{aligned}$$

And therefore $\mu(B) = \lim_{n \to \infty} \mu(B \cap X_n)$. □

Proof of measure-preserving property. Let A be a measurable subset of $[0,1]$.

A can be written as $A = (A \bigcap I_{n,1}) \bigcup \left(\bigsqcup_{k=2}^{h_n} A \bigcap I_{n,k} \right) \bigcup (A \bigcap \text{remaining spacer at } n\text{th stage})$.

Thus

$$\mu(A) = \mu\left(A \bigcap I_{n,1}\right) + \mu\left(\bigsqcup_{k=2}^{h_n} \left(A \bigcap I_{n,k}\right)\right) + \mu(A \bigcap \text{ remaining spacer at } n\text{th stage}).$$

Notice that $\mu\left(A \bigcap I_{n,1}\right) \to 0$ as $n \to \infty$ because $\mu(I_{n,1}) = \frac{2}{3^{n+1}} \to 0$ as $n \to \infty$. Similarly

$\mu(A \bigcap$ remaining spacer at nth stage$) \to 0$ as $n \to \infty$. Therefore as $n \to \infty$ we have that

$$\mu(A) = \lim \mu \left(\bigsqcup_{k=2}^{h_n} \left(A \bigcap I_{n,k} \right) \right)$$

From the preceding lemma we know this holds since $\mu \left(\bigsqcup_{k=2}^{h_n} I_{n,k} \right) \to 1$ as $n \to \infty$.

Also note that $\bigsqcup_{k=2}^{h_n} \left(A \bigcap I_{n,k} \right) = A \bigcap \left(\bigsqcup_{k=2}^{h_n} I_{n,k} \right)$ and that

$$\begin{aligned} \mu \left(A \bigcap \bigsqcup_{k=2}^{h_n} I_{n,k} \right) &= \mu \left(T^{-1} \left(A \bigcap \bigsqcup_{k=2}^{h_n} I_{n,k}) \right) \right) \text{ as T is m.p on levels of the tower} \\ &= \mu \left(T^{-1}(A) \bigcap T^{-1} \left(\bigsqcup_{k=2}^{h_n} I_{n,k} \right) \right) \\ &= \mu \left(T^{-1}(A) \bigcap \bigsqcup_{k=1}^{h_n - 1} (I_{n,k}) \right) \\ &\to \mu \left(T^{-1}(A) \right) \end{aligned}$$

Again this result is from the preceding lemma and from fact that $\mu(\bigsqcup_{k=1}^{h_n-1} I_{n,k}) \to 1$ as $n \to \infty$. So for the same expression we arrive at two conclusions and therefore $\mu(A) = \mu(T^{-1}A)$. □

2.2.2 The Chacon Transformation is Not Strongly Mixing

Definition 2.2.1. A measure-preserving transformation T is said to be strongly mixing if $\forall A \in \mathcal{B}$ and $\forall B \in \mathcal{B}$

$$\lim_{n \to \infty} \mu(T^{-n}(A) \cap B) = \mu(A)\mu(B).$$

Proposition 2.2.1. *Chacon's transformation is not strongly mixing.*

Proof. We will be able to conclude Chacon's transformation is not strongly mixing if we can find at least one $A \in \mathcal{B}$ of positive measure such that

$$\limsup_{n \to \infty} \mu(T^{-n}(A) \cap A) \neq (\mu(A))^2.$$

Note that since T is invertible almost everywhere.

$$\begin{aligned}\mu(T^n A \cap A) &= \mu\left(T^{-n}(T^n A \cap A)\right) \text{ because T is measure-preserving}\\ &= \mu\left(T^{-n}(T^n A) \cap T^{-n} A\right)\\ &= \mu(A \cap T^{-n} A).\end{aligned}$$

With this result if we can show instead that $\limsup_{n\to\infty} \mu(T^n(A) \cap A) \neq (\mu(A))^2$ we will still have established that T is not strongly mixing. We fix the stage k_0 and denote the tower at this stage by $\mathscr{T}_{k_0}$. As a candidate for A, we choose one of the levels of the tower. The most convenient choice for A is the base of the tower of $\mathscr{T}_{k_0}$. The height of this tower is h_{k_0}.

We note that $\mu(T^{h_n} A \cap A) \geq \frac{1}{3}\mu(A)$. To establish this, first observe that A is split into three intervals, A_1, A_2 and A_3 in the (k_0+1)th stage under the cutting and stacking process. Using some analysis of the measures of the intervals and by induction we will obtain the inequality above and use it to prove why the Chacon tranformation is not strongly mixing.

For stage k_0+1 and on tower $\mathscr{T}_{k_0+1}$,

$$\begin{aligned}T^{h_{k_0}}(A) \cap A &= T^{h_{k_0}}(A_1 \cup A_2 \cup A_3) \cap A\\ &= \left(T^{h_{k_0}}(A_1) \cap A\right) \cup \left(T^{h_{k_0}}(A_2) \cap A\right) \cup \left(T^{h_{k_0}}(A_3) \cap A\right)\end{aligned}$$

Observe that on this tower only $T^{h_{k_0}} A_1$ hits A again. In particular

$$\begin{aligned}\mu(T^{h_{k_0}}(A_1) \cap A) &= \mu(A_2 \cap A) = \frac{1}{3}\mu(A) \text{ and}\\ \mu\left(T^{h_{k_0}}(A_2) \cap A\right) &= \mu(\emptyset) = 0.\end{aligned}$$

We are unable to make a direct estimate for $\mu(T^{h_{k_0}}(A_3) \cap A)$ in the (k_0+1) stage so we end the (k_0+1) stage with the result that,

$$\mu\left(T^{h_{k_0}}(A) \cap A\right) \geq \frac{1}{3}\mu(A).$$

Moving to the (k_0+2)th stage by cutting and stacking, A_1 intersects with itself at least once under T^{h_0+1}. The same happens for A_2 and A_3. Therefore

$$\begin{aligned} \mu(T^{h_{k_0+1}}(A_1) \cap A_1) &\geq \frac{1}{3} \cdot \frac{1}{3}\mu(A) = \frac{1}{9}\mu(A). \\ \text{Likewise } \mu\left(T^{h_{k_0+1}}(A_2) \cap A_2\right) &\geq \frac{1}{9}\mu(A) \text{ and} \\ \mu(T^{h_{k_0+1}}(A_3) \cap A_3) &\geq \frac{1}{9}\mu(A). \end{aligned}$$

Therefore at the end of the $k_0 + 2$ stage we arrive at the inequality

$$\mu(T^{h_{k_0+1}}(A) \cap A) \geq 3 \cdot \frac{1}{9}\mu(A) = \frac{1}{3}\mu(A).$$

By induction we see that at stage $n \geq k$ we have 3^{n-k_0} pieces of A with each contributing $\frac{1}{3^{n-k_0+1}}\mu(A)$. Thus

$$\mu(T^{h_n}(A) \cap A) \geq 3^{n-k_0} \cdot \frac{1}{3^{n-k_0+1}}\mu(A) = \frac{1}{3}\mu(A).$$

With this inequality holding for all stages of the Chacon transformation, then for the Chacon transformation to be strongly mixing it has to satisfy that $(\mu(A))^2 \geq \frac{1}{3}\mu(A) \Rightarrow \mu(A) \geq \frac{1}{3}$. However this fails for the Chacon process since each level of the transformation has length of at most $\frac{2}{9}\mu(A)$. Therefore $(\mu(A))^2$ is at most $\frac{4}{81}\mu(A)$. Hence this proves that Chacon's transformation is not strongly mixing. □

2.2.3 Weakly Mixing Property

The main theorem which will be used to prove that the Chacon transformation is weakly mixing is the following:

Theorem 2.2.1. *Let T be an invertible measure-preserving transformation on a Lebesgue probability space. Then the following are equivalent.*

1. *T is weakly mixing.*
2. *T is doubly ergodic.*
3. *T has continuous spectrum.*

4. $T \times S$ is ergodic for any ergodic, finite measure-preserving transformation S.

The following defintions, proposition and lemma will be needed to prove the implication (1) $\Rightarrow$ (2).

Definition 2.2.2. Let $(X, \mathcal{B}, \mu)$ be a measure-preserving space and T a measure-preserving transformation. T is said to be weakly mixing if $\forall\ A, B \in \mathcal{B}$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \left|\mu(T^{-k}A \cap B) - \mu(A)\mu(B)\right| = 0$$

Recall the definition of what it means for a sequence to have *convergence* to a limit, to have *strong Cesaro convergence* and to have *Cesaro convergene* to a limit. Denoting $a_k(A, B) = \mu(T^{-k}(A) \cap B)$ and $a = \mu(A)\mu(B)$ we note that the property of T being weakly mixing is equivalent to the sequence a_k converging strong Cesaro to a.

Definition 2.2.3. A transformation T is defined to be doubly ergodic if $\forall\ A, B \in \mathcal{B}$ of positive measure, there exists an integer n such that

$$\mu(T^{-n}A \cap A) > 0 \text{ and } \mu(T^{-n}A \cap B) > 0.$$

Lemma 2.2.2. *Let $\{a_k\}$ be a bounded nonnegative sequence. Then $\frac{1}{N}\sum_{k=1}^{N} a_k$ converges to 0 if and only if $\{a_k\}$ converges to 0 in density.*

Proof. Suppose $\{a_k\}$ converges to 0 in density, that is, there exists a zero density set D such that outside D, $\lim_{k\to\infty, k\notin D} a_k = 0$. Let $\{a_k\}$ be bounded by a. Then

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} a_k &= \lim_{n\to\infty}\left[\frac{1}{n}\sum_{k=0, k\in D}^{n-1} a_k + \frac{1}{n}\sum_{k=0, k\notin D}^{n-1} b_k\right] \\
&= \limsup_{n\to\infty}\left[\frac{1}{n}\sum_{k=0, k\in D}^{n-1} a_k + \frac{1}{n}\sum_{k=0, k\notin D}^{n-1} a_k\right] \\
&\leq \limsup_{n\to\infty} \frac{1}{n}\sum_{k=0, k\in D} a + \limsup_{n\to\infty}\frac{1}{n}\sum_{k=0, k\notin D} a_k \\
&= a \limsup_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} \mathbb{I}_{D(k)} + \limsup_{n\to\infty}\frac{1}{n}\sum_{k=0, k\notin D}^{n-1} a_k
\end{aligned}$$

$$= a \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{I}_{D(k)} + \lim_{n\to\infty} \frac{1}{n} \sum_{k=0,k\notin D}^{n-1} a_k = 0.$$

On the other hand, suppose that $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} a_k = 0$. For each $m = 1, 2, ...$, let

$$D_m = \{k : \frac{1}{m} < a_k\}.$$

Note that $D_1 \subset D_2 \subset ....$ Each D_m has density 0 as for a fixed m,

$$\frac{1}{n} \sum_{k=0}^{n-1} \mathbb{I}_{D_m}(k) < m \cdot \frac{1}{n} \sum_{k=0}^{n-1} a_k \to 0$$

as $n \to \infty$. Also recall that $k \in D_m \Rightarrow a_k > \frac{1}{m}$. Choose $i_1 < i_2 < ...$ such that for each $n \geq i_m$, $\frac{1}{n} \sum_{k=0}^{n-1} \mathbb{I}_{D_m}(k) < \frac{1}{m}$. Define

$$D = \bigsqcup_{m=1}^{\infty} D_m \bigcap (i_m, i_{m+1}).$$

It is clear that $\lim_{k\to\infty, k\notin D} a_k = 0$ since as off the set D, a_k , a nonnegative sequence, decreases as m tends to infinity.

What is left to prove is that D has density 0. Fix some large $n \in \mathbb{N}$. We can find an m with $i_m \leq n \leq i_{m+1}$. Then

$$\begin{aligned}
\frac{1}{n} \sum_{k=0}^{n-1} \mathbb{I}_D(k) &= \frac{1}{n} \sum_{k=0}^{i_{m-1}} \mathbb{I}_D(k) + \frac{1}{n} \sum_{k=0}^{i_m} \mathbb{I}_D(k) \\
&\leq \frac{1}{n} \sum_{k=0}^{i_{m-1}} \mathbb{I}_{D_{m-1}}(k) + \frac{1}{n} \sum_{k=0}^{i_m} \mathbb{I}_{D_m}(k) \\
&\leq \frac{1}{m-1} + \frac{1}{m}.
\end{aligned}$$

Given any $\varepsilon > 0$ we can choose n so that $n \geq i_m \Rightarrow \frac{1}{m-1} < \frac{\varepsilon}{2}$. □

Lemma 2.2.3. *For T, a measure-preserving transformation on a probability space $(X, \mathcal{B}, \mu)$, the following are equivalent:*

1. *T is weakly mixing.*

2. *For each pair of measurable sets A and B, there is a zero density set $D = D(A, B)$ such that*

$$\lim_{k\to\infty, k\notin D} \mu(T^{-k}(A) \cap B) = \mu(A)\mu(B).$$

3. *For each pair of A, $B \in \mathcal{B}$*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \Big[\mu(T^{-k} \cap B) - \mu(A)\mu(B)\Big]^2 = 0$$

The second statement in words implies that T is weakly mixing if there is a set D such that off this set T is strongly mixing.

Proof. (1) implies (2) Suppose T is weakly mixing. By definition

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} |\mu(T^{-n}A \cap B) - \mu(A)\mu(B)| = 0$$

Denoting $b_k = |\mu(T^{-k}A \cap B) - \mu(A)\mu(B)|$ and noting that b_k is bounded by 1, then

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} b_k &= \lim_{n\to\infty} \left[\frac{1}{n} \sum_{k=0, k\in D}^{n-1} b_k + \frac{1}{n} \sum_{k=0, k\notin D}^{n-1} b_k \right] \\
&= \limsup_{n\to\infty} \left[\frac{1}{n} \sum_{k=0, k\in D}^{n-1} b_k + \frac{1}{n} \sum_{k=0, k\notin D}^{n-1} b_k \right] \\
&\le \limsup_{n\to\infty} \frac{1}{n} \sum_{k=0, k\in D} 1 + \limsup_{n\to\infty} \frac{1}{n} \sum_{k=0, k\notin D} b_k \\
&= 1 \limsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{I}_{D(k)} + \limsup_{n\to\infty} \frac{1}{n} \sum_{k=0, k\notin D}^{n-1} b_k \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{I}_{D(k)} + \lim_{n\to\infty} \frac{1}{n} \sum_{k=0, k\notin D}^{n-1} b_k = 0
\end{aligned}$$

by the fact that D is a set of density zero and by Lemma 2.2.2.

(2) implies (3) Suppose (2) is true. Then

$$\begin{aligned}
&\lim_{k\to\infty, k\notin D} \mu(T^{-k}(A) \cap B) - \mu(A)\mu(B) = 0 \\
\Rightarrow &\lim_{k\to\infty, k\notin D} \Big[\mu(T^{-k}(A) \cap B) - \mu(A)\mu(B)\Big]^2 = 0.
\end{aligned}$$

Using Lemma 2.2.2 again then this implies that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \Big[\mu(T^{-k} \cap B) - \mu(A)\mu(B)\Big]^2 = 0.$$

(3) implies (1) If (3) holds, then clearly,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\mu(T^{-k} \cap B) - \mu(A)\mu(B)| = 0$$

which is the definition of weakly mixing and hence T is weakly mixing. □

Theorem 2.2.1 (1) ⇒ (2) [*T is weakly mixing implies T is doubly ergodic.*]

Proof. Let A and B be sets with positive measure. Then by Lemma 2.2.3 there exist sets of density zero $D_1 = D(A, B)$ and $D_2 = D(A, B)$ such that

$$\lim_{k\to\infty, k\notin D_1} \mu(T^{-k}(A) \cap B) = \mu(A)\mu(B)$$

and

$$\lim_{k\to\infty, k\notin D_2} \mu(T^{-k}(A) \cap A) = \mu(A)\mu(A)$$

Therefore there is some integer $k > 0$, $k \in (D_1 \cup D_2)^c$ such that $\mu(T^{-k}(A) \cap B) > 0$ and $\mu(T^{-k}(A) \cap A) > 0$. Hence T is doubly ergodic. □

To prove the implication that $2 \Rightarrow 3$ of Theorem 2.2.1, recall what it means for a transformation to be ergodic. We will have need for the following lemma.

Lemma 2.2.4. *Let T be a probability-preserving transformation. If T is weakly mixing, then it is ergodic.*

Proof. If T is weakly mixing, recall that for all measurable sets A, $B \in \mathcal{B}$ then $a_k(A, B) = \mu(T^{-k}(A) \cap B)$ converges strong Cesaro to $\mu(A)\mu(B)$. Then for A, a T-invariant set, $a_k(A, A^c) = 0$ but we also know that it converges to $\mu(A)\mu(A^c)$. Therefore either $\mu(A) = 0$ or $\mu(A^c) = 0$ which is the requirement for a transformation to be ergodic for T-invariant sets. Thus T is ergodic. □

Definition 2.2.4. A dynamical system $(Y, \mathcal{C}, \nu, S)$ is called the factor of $(X, \mathcal{B}, \mu, T)$ if there are measurable sets X_0 and Y_0 of full measure such that

$$X_0 \subset X,\ T(X_0) \subset X_0 \ \text{ and } \ Y_0 \subset Y,\ S(Y_0) \subset Y_0,$$

and an onto map (not necessarily one-to-one a.e) $\phi : X_0 \to Y_0$ such that for all $A \in \mathcal{C}(Y_0)$, where $\mathcal{C}(Y_0)$ is the trace of $\mathcal{C}$ onto Y_0, that is, $\mathcal{C}(Y_0) = \mathcal{C} \cap Y_0 = \{A \cap Y_0 : A \in \mathcal{C}\}$,

1. $A \in \mathcal{C}(Y_0)$ if and only if $\phi^{-1}(A) \in \mathcal{B}(X_0)$
2. $\mu(\phi^{-1}(A)) = \nu(A)$ for all $A \in \mathcal{C}(Y_0)$.

Next we give some material on a small portion of spectral theory on transformations.

Let T be a measure-preserving transformation and $f \in L^2$ i.e $\int |f|^2 d\mu < \infty$. Set $\hat{\sigma}(n) = \langle T^n f, f\rangle = \langle f \circ T^n, f\rangle = \int f \circ T^n \cdot \bar{f} d\mu$. So

$$\hat{\sigma}(0) = \int f \cdot \bar{f} d\mu,\ \hat{\sigma}(1) = \int f \circ T \cdot \bar{f} d\mu,\ \hat{\sigma}(2) = \int f \circ T^2 \cdot \bar{f} d\mu ...$$

The sequence $\{\hat{\sigma}(n)\}$, $n \geq 0$ is positive definite, that is, for each system of complex numbers $z_0, z_1, ... z_n$, $\sum_{j,k=0}^{n} \hat{\sigma}(j+k) z_j \bar{z_k} \geq 0$. By the Bochner-Herglotz theorem, there exists σ_f, a positive measure defined on $\mathcal{B}[(0,1)]$,

$$\sigma_f : \mathcal{B}[(0,1)] \to \mathbb{R}^+.$$

such that

$$\hat{\sigma}_f(n) = \hat{\sigma}(n) = \int_0^1 e^{-2\pi int} d\sigma_f(t).$$

Each spectral measure, σ_f, associated with f can be broken up as follows:

$$\begin{aligned} \sigma_f &= \underbrace{\sigma_f^1}_{} + \underbrace{\sigma_f^2 + \sigma_f^3}_{} \quad \text{where} \\ &= \sigma_f^d + \quad \sigma_f^c \end{aligned}$$

$\sigma_f^1 = \sigma_f^d$ is the discrete part of σ_f. The measures σ_f^2 and σ_f^3 constitute the continuous part of σ_f. By *continuous* we mean that $\sigma\{t\} = 0 \ \ \forall\ t \in [0,1]$. The measure σ_f^3 called the Lebegue component of σ_f is absolutely continuous with respect to the Lebesgue measure while σ_f^2, also continuous, is singular with respect to the Lebesgue measure.

Let $(X, \mathcal{B}, \mu)$ be a probability space and T a measure-preserving transformation. λ is called an *eigenvalue* of T if there exists a nonzero a.e complex-valued function $f : X \to \mathbb{C}$, $f \in L^1$, such that $f \circ T = \lambda f$ a.e. The function f is the corresponding *eigenfunction* to λ.

Lemma 2.2.5. *Let T be a measure-preserving transformation on $(X, \mathcal{B}, \mu)$. T is ergodic if and only if for all measurable functions $f : X \to \mathbb{R}$ whenever $f(x) = f(T(x))$ a.e. then f is constant a.e.*

Proof. We prove that forward implication by contradiction. Suppose T is ergodic and $f \circ T = f$ a.e. Suppose for the sake of contradiction that f is not constant a.e. Then there exists some number t such that the sets $\{x : f(x) < t\}$ and $\{x : f(x) > t\}$ both have positive measure. Observe that these sets are disjoint and are T- invariant. Since T is ergodic they cannot both have positive measure so f must be constant a.e. $\square$

Note that for T we have the following holding. T is weakly mixing which makes it ergodic. If f is an eigenfuction of T then f is constant a.e. Also observe that T is measure-preserving and because constant functions are always eigenfunctions, $\lambda = 1$ is always an eigenvalue. Since we are working on a probability space, if λ is an eigenvalue, then $|\lambda| = 1$ i.e, λ lies on the unit circle in $\mathbb{C}$. We are able to confirm this by observing that λ being an eigenvalue of f it implies that

$$f \circ T = \lambda f.$$

From a pointwise perspective

$$\begin{aligned} |f \circ T(x)| &= |\lambda||f(x)| \\ \int |f \circ T(x)| d\mu &= |\lambda| \int |f(x)| d\mu \text{ and since } T \text{ is measure-preserving} \\ \int |f(x)| d\mu &= |\lambda| \int |f(x)| d\mu \end{aligned}$$

Since $\int |f(x)| d\mu \neq 0$ as we are looking at a nonzero eigenfunction f, then this implies $|\lambda| = 1$.

Definition 2.2.5. A measure-preserving transformation is defined to have continuous spectrum if $\lambda = 1$ is its only eigenvalue and in addition it is simple.

Theorem 2.2.1 (2) $\Rightarrow$ (3). [*T is doubly ergodic implies T has continuous spectrum*]

Proof. Suppose T is doubly ergodic, then we have for free that T is ergodic. Also suppose λ is an eigenvalue of T. Then by definition, there exists a measurable function $f : X \to \mathbb{C}$ such that $f(T(x)) = \lambda f(x)$ almost everywhere and in addition, $|\lambda| = 1$ and f is constant a.e. Without loss of generality let $|f| = 1$.

Expressing λ and f in Fourier, let $\lambda = e^{2\pi i\alpha}$ and $f(x) = e^{2\pi i g(x)}$ for some $\alpha \in [0, 1)$ and a measurable function $g : X \to [0, 1)$.

Let $S : [0, 1) \to [0, 1)$ such that $S(t) = t+\alpha$. Note from figure 2.2.1 below that $g\circ T = S\circ g$. We define a measure ν on $[0, 1)$ by $\nu(A) = \mu(g^{-1}(A))$.

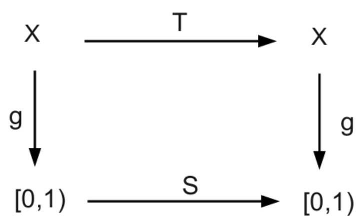

Fig. 2.2.1: factor map

Then by previous defintion of what a factor map is, we note that g is a factor map from T to S and since S is a factor of T, it inherits all properties of T and thus it implies that S is also ergodic.

We have some scenarios arising from the nature of α.

Case 1: Refering back to Figure 1.1.1 from Chapter 1, note that there are two partitions of $[0, 1)$. One consists of the disjoint subintervals $[0, 1 - \alpha)$ and $[1 - \alpha, 1)$. The other consists of disjoint intervals $[0, \alpha)$ and $[\alpha, 1)$. In addition, note that an interval in the first partition is sent to an interval of the same length in the second partition by S. In particular, an interval in $[0, 1-\alpha)$ is mapped to an interval in $[\alpha, 1)$ under the S and an interval which is a subset of $[1 - \alpha)$ has its image as an interval in $[0, \alpha)$. If α is rational, then ν is atomic and concentrated on only a finite number of points. Let I and J be intervals of equal measure in $[0, 1)$ that are sufficiently apart. Observe that $S^n(I)$ is an interval thus for any integer n, if $S^n(I) \cap J \neq \emptyset$ we would expect that $S^n(I) \cap I = \emptyset$. This being the case, then S is not doubly ergodic which is a contradiction to our assumption.

Case 2: If α is irrational, then ν must be the Lebesgue measure and in this case also it implies that S is not doubly ergodic. Again we have a contradiction to our initial assumption.

Hence $\lambda = 1$ is the only eigenvalue and this implies that T has continuous spectrum. □

Lemma 2.2.6. *(Bochner's Theorem) Let μ be a measure on $\mathcal{B}([0,1])$. The measure μ is continuous if and only if*

$$\frac{1}{2n+1}\sum_{k=-n}^{n} |\hat{\mu}(k)| \longrightarrow 0 \text{ as } n \to \infty.$$

Remark:

1. Note that if $\frac{1}{2n+1}\sum_{k=-n}^{n} |\hat{\mu}(k)| \longrightarrow 0$ then

$$\frac{1}{n}\sum_{k=0}^{n} |\hat{\mu}(k)| \longrightarrow 0 \text{ as } n \to \infty.$$

2. T is ergodic on a probability measure space if and only if

$$\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k(x) \longrightarrow \int f d\mu \text{ in } L^2 \text{ norm.}$$

Lemma 2.2.7. *Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving transformation on a finite measure space and σ_f be the spectral measure associated with f. If σ_f is continuous, then*

$$\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k \longrightarrow 0$$

in L^2 norm as $n \to \infty$.

This follows from the fact that $\mathbb{E}^{\mathcal{I}}(f) = 0$ or we can observe this by direct calculation.

$$\begin{aligned}\int |\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k|^2 d\mu &= \frac{1}{n^2}\int \left(\sum_{k=0}^{n-1} f \circ T^k\right)^2 d\mu \\ &= \frac{1}{n^2}\int \left(\sum_{k=0}^{n-1} f \circ T^k\right)\left(\sum_{j=0}^{n-1} f \,\bar{\circ}\, T^j\right) d\mu\end{aligned}$$

$$= \frac{1}{n^2} \sum_{j,k=0}^{n-1} \int \left(f \circ T^k\right) \left(f \circ \bar{T}^j\right) d\mu$$

$$= \frac{1}{n^2} \sum_{j,k=0}^{n-1} \int \left(f \circ T^{k-j} \cdot \bar{f}\right) d\mu \text{ because T is measure-preserving.}$$

$$\leq \frac{1}{n^2} \sum_{j,k=0}^{n-1} |\hat{\sigma_f}(n-j)|$$

$$\leq \frac{2}{n} \sum_{k=0}^{n-1} |\hat{\sigma}(k)| \rightarrow 0$$

by the Bochner theorem since σ_f is continuous.

Theorem 2.2.1 (3) $\Rightarrow$ (4). [*T has continuous spectrum implies that $T \times S$ is ergodic for any ergodic, finite measure-preserving transformation S.*]

Proof. Suppose T has continuous spectrum and S is ergodic. Set $U = T \times S$. The goal is to show that U is ergodic. Observe that U is ergodic if and only if for all bounded F on $X \times Y$

$$\frac{1}{N} \sum_{k=0}^{N-1} F \circ U^k \longrightarrow \int F d(\mu \times \nu).$$

If we are able to prove this on a dense set in $X \times Y$ then it holds on the whole set. Let $F = \mathbb{I}_A \mathbb{I}_B$ where $F(x,y) = \mathbb{I}_A(x) \mathbb{I}_B(y)$. Thus to prove that U is ergodic we need to prove that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \left[\left(\mathbb{I}_A \circ T^k\right) \left(\mathbb{I}_B \circ S^k\right)\right] \longrightarrow \int (\mathbb{I}_A \mathbb{I}_B)\, d\mu\, d\nu.$$

Using the fact that T has continous spectrum, we can choose $f_A = \mathbb{I}_A - \mu(A)$. Then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \left[(f_A \circ T^k)(\mathbb{I}_B \circ S^k)\right] =$$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \left[\mathbb{I}_A(T^k(x))\ \mathbb{I}_B(S^k(y))\right] - \mu(A) \lim_{n\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathbb{I}_B(S^k(y)).$$

Because S is ergodic,

$$\lim_{n\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathbb{I}_B(S^k(y)) = \nu(B) \text{ a.e.}$$

Hence for U to be ergodic is suffices to show that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \Big[(f_A \circ T^k)(\mathbb{I}_B \circ S^k) \Big] = 0.$$

To prove this we compute the (real) Fourier coefficents of $f_A \times \mathbb{I}_B$.

$$\langle f_A \times \mathbb{I}_B, (f_A \times \mathbb{I}_B) \circ U^k \rangle = |\int (f_A \times \mathbb{I}_B)(f_A \times \mathbb{I}_B) \circ U^k d\mu d\nu|.$$

Then in average,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} |\int\int (f_A \times \mathbb{I}_B)(f_A \times \mathbb{I}_B) \circ U^k d\mu d\nu| =$$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} |\int f_A \cdot f_A \circ T^k d\mu| \; |\int \mathbb{I}_B \cdot \mathbb{I}_B \circ S^k d\nu|.$$

Note that $|\int \mathbb{I}_B \cdot \mathbb{I}_B \circ S^k d\nu| \leq 1$. Therefore

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} |\int\int (f_A \times \mathbb{I}_B)(f_A \times \mathbb{I}_B) \circ U^k d\mu d\nu| \leq \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} |\int f_A \cdot f_A \circ T^k d\mu|.$$

But since σ_{f_A} is continuous then by Lemma 2.2.6,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |\hat{\sigma}_{f_A}(n)| = 0.$$

Therefore

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} |\int\int (f_A \times \mathbb{I}_B)(f_A \times \mathbb{I}_B) \circ U^k d\mu d\nu| = 0$$

which implies that $\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \Big[\mathbb{I}_A(T^k(x)) \; \mathbb{I}_B(S^k(y)) \Big] = \mu(A)\nu(B).$ □

Recall Lemma 2.2.3. To prove the implication (4) implies (1) we will show that $T \times S$ satisfies the third equivalent statement of weakly mixing in Lemma 2.2.3.

Theorem 2.2.1 (4) ⇒ (1) [*If $T \times S$ is ergodic for any ergodic finite measure-preserving transformation S then T is weakly mixing.*]

Proof. Suppose $T \times S$ is ergodic. In addition suppose S is also ergodic thus T is ergodic and therefore by definition satisfies

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\mu(T^{-k}(A\cap B))=\mu(A)\mu(B).$$

Let $(X,\mathcal{B},\mu,T)$ and $(Y,\mathcal{C},\nu,S)$ be measure-preserving systems. Since $T\times S$ is ergodic, then for all $A,\ B\in\mathcal{B}$ and for all $C,\ D\in\mathcal{C}$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}(\mu\times\nu)\left[(T\times S)^{-k}\,(A\times C)\cap(B\times D)\right]=(\mu\times\nu)(A\times C)\ (\mu\times\nu)(B\times D).$$

Observe that this is the same as

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\left[\Big(\mu(T^{-k}(A\cap B)\Big)\times\nu(S^{-k}(C\cap D))\right]=\mu(A)\mu(B)\nu(C)\nu(D).$$

Also note that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\left[(\mu\times\nu)\Big((T\times S)^{-k}(A\times C)\cap(B\times D)\Big)\right]^2=$$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\left[\mu(T^kA\cap B)^2\nu(S^kC\cap D)^2\right]$$

$$=\mu(A)^2\mu(B)^2\nu(C)^2\nu(D)^2.$$

Therefore

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\left[(\mu\times\nu)(T\times S)^{-k}(A\times C)\cap(B\times D)-\mu(A)\mu(B)\nu(C)\nu(D)\right]^2$$

$$=\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\left[\Big(\mu(T^{-k}(A)\cap B)\Big)^2\Big(\nu(S^{-k}(C)\cap D)\Big)\right]-$$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\left[2\mu(T^{-k}(A)\cap B)(\nu(S^{-k}(C)\cap D)\mu(A)\mu(B)\nu(C)\nu(D)+(\mu(A)\mu(B)\nu(C)\nu(D))^2\right]$$

$$=2\mu(A)^2\mu(B)^2\nu(C)^2\nu(D)^2-2\mu(A)^2\mu(B)^2\nu(C)^2\nu(D)^2=0$$

Hence T is weakly mixing. □

This completes the proof of Theorem 2.2.1.

Theorem 2.2.2. *The Chacon transformation is doubly ergodic.*

By showing this and with the help of the just proved theorem for a measure-preserving transformations in general, we can conclude that Chacon's transformation is weakly mixing.

Proof. First we fix a stage k. Let $\mathscr{T}_k$ be the tower at this stage. We pick two levels of the tower and denote them by I and J. Let I and J be separated by l levels with J above I, that is, J is the $(i+l)$th level in $\mathscr{T}_k$ when I is the ith level. In proving that the Chacon transformation was not strongly mixing we were able to establish that for any level of the tower A, $\mu(T^n A \cap A) \geq \frac{1}{3}\mu(A)$. Without loss of generality this holds that for any stage k. Thus we have the inequality that $\mu(T^k I \cap I) \geq \frac{1}{3}\mu(I)$. We claim that there exists an H such that $\mu(I \cap T^H J) = \mu(T^H J \cap I) \geq \frac{1}{3^l}\mu(J)$. If we are able to establish that the same H holds for these two inequalities at the same stage, then we can conclude that the Chacon transformation is doubly ergodic.

To establish this observe that at stage k the tower $\mathscr{T}_k$ has height h_k. Moving to the $(k+1)$th stage, all the levels of the tower are split up into three subintervals each. In particular, the level J is divided into three subintervals J_1, J_2, J_3. Further observe that $T^{h_k} J_1 = J_2$ and $T^{h_k} J_2 = T^{-1} J_3$. This is because of the spacer introduced in $\mathscr{T}_{k+1}$ during the cutting and stacking process. Thus we can say that $\mu(T^{h_k} J_2 \cap T^{-1}(J_3)) = \mu(T^{h_k} J \cap T^{-1}(J)) \geq \frac{1}{3}\mu(J)$. Keeping track of only one of these subintervals of J, set $T^{h_k} J_2$ which is a level in the $(k+1)$th stage to L. In moving from the $(k+1)$th stage to the $(k+2)$th stage let L_1, L_2, L_3 denote the subintervals of L. Again because the introduction of a new spacer we can see that $\mu(T^{h_{k+1}} L_2 \cap T^{-1}(L_3)) \geq \frac{1}{3}\mu(L)$. In fact, this helps us establish that $\mu(T^{h_{k+1}} L \cap T^{-1}(L)) \geq \frac{1}{3}\mu(L)$. Which means that

$$\mu(T^{h_k}(T^{h_{k+1}} J) \cap T^{-2}(J))$$

$$= \mu(T^{h_k + h_{k+1}} J \cap T^{-2} J) \geq \frac{1}{3^2}\mu(J).$$

In tracking any level we notice that the tower is split into three pieces and a spacer is introduces when we move to the next stage. In studying the middle and right portions of a split tower, the level always misses intersecting with itself because of the introduced spacer.

Then by induction we are able to establish that $(k+l)$th stage

$$\mu(T^{h_k+h_{k+1}+...+h_{k+l-1}}J \cap T^{-l}(J)) \geq \frac{1}{3^l}\mu(J).$$

But observe that $T^{-l}J = I$ hence

$$\mu(T^H(J) \cap I) \geq \frac{1}{3^l}\mu(J)$$

with $H = h_k + h_{k+1} + ... + h_{k+l-1} = \sum_{i=0}^{l-1} h_{k+i}$.

We claim that the same H will make the inequality $\mu(T^H(I) \cap I) > 0$ true. This time we use the left and middle portions of the tower to attain desired result. Recall that we have already proven that at any $(k+1)$th stage $\mu(T^{h_k}(I) \cap I) \geq \frac{1}{3}\mu(I)$. Note that this applies to any stage in the construction of Chacon transformation.

If we look at $T^{h_k}(I) \cap I$ as an interval, then in the $(k+2)$th stage, we know that

$$\begin{aligned}\mu\left(T^{k+1}(T^{h_k}(I) \cap I) \cap (T^{h_k}(I) \cap I)\right) &= \mu\left(T^{h_{k+1}+h_k}(I) \cap T^{h_{k+1}}(I) \cap T^{h_k}(I) \cap J)\right) \\ &\geq \frac{1}{3^2}\mu(I).\end{aligned}$$

Now observe that $T^{h_{k+1}+h_k}(I) \cap T^{h_{k+1}}(I) \cap T^{h_k}(I) \cap J$ is a subset of $T^{h_{k+1}+h_k}(I) \cap J$. Thus $\mu(T^{h_{k+1}+h_k}(I) \cap I) \geq \frac{1}{3^2}\mu(I)$.

Then by induction we can say that

$$\mu(T^{h_{k+l-1}+h_{k+l}+...+h_{k+1}+h_k}(I) \cap I) \geq \frac{1}{3^l}\mu(I)$$

with $h_{k+l-1} + h_{k+l} + ... + h_{k+1} + h_k = H$.

Let A, $B \subset [0,1]$ be sets of positive measure. We claim that there exists I, a level in a tower $\mathscr{T}_k$ and a level J in $\mathscr{T}_{k'}$ such that $\mu(A \cap I) \geq \frac{2}{3}\mu(I)$ and $\mu(B \cap J) \geq \frac{2}{3}\mu(J)$. By proving this we will be able to conclude that Chacon's transformation is doubly ergodic. We state without proof the Increasing Martingale theorem which will be needed.

Lemma 2.2.8. *(Increasing Martingale theorem) [2] If $\{\mathcal{B}_n\}$ is an increasing sequence if sub-σ-algebras with $\mathcal{B}_n \uparrow \mathcal{B}_\infty$ (i.e $\cup_n \mathcal{B}_n$ generates $\mathcal{B}_\infty$), then the conditional expectation of $f \in L^1(X)$ with respect to $\mathcal{B}_n$, $\mathbb{E}(f|\mathcal{B}_n) \to \mathbb{E}(f|B_\infty)$ a.e and in $L^1(X)$.*

Continuing the proof of the doubly ergodic property of Chacon's transformation, let $\mathcal{B}_k$ be the increasing sequence of σ-algebras generated by the levels of $\mathscr{T}_k$ and the spacer which converges to B (the σ-algebra of Lebesgue measurable subsets of $[0, 1]$.) The conditional expectation of $\mathbb{I}_A$ with respect to $\mathcal{B}_k$ is given by

$$\mathbb{E}(\mathbb{I}_A|\mathcal{B}_k)(x) = \sum_{j=1}^{h_k-1} \frac{\int_{I_j} \mathbb{I}_A d\mu}{\mu(I_j)} \cdot \mathbb{I}_{I_j}(x) + \frac{\int_{S_k} \mathbb{I}_A d\mu}{\mu(S_k)} \cdot \mathbb{I}_{S_k}(x).$$

where S_k represents the remaining spacer at the kth stage. Then by the Martingale theorem,

$$\mathbb{E}(\mathbb{I}_A|\mathcal{B}_k)(x) \longrightarrow \mathbb{I}_A(x) \text{ a.e.}$$

After eliminating a set of measure 0 in A, pick a point $c \in A$ then

$$\mathbb{E}(\mathbb{I}_A|\mathcal{B}_k)(c) \longrightarrow \mathbb{I}_A(c) = 1$$

since the point c will be in only one set. Then for all k there exists a j_k such that

$$\frac{\int_{I_{j_k}} \mathbb{I}_A d\mu}{\mu(I_{j_k})} = \frac{\mu(A \cap I_{j_k})}{\mu(I_{j_k})} \longrightarrow 1.$$

Thus we can conclude that there exists a k_0 such that for all $k \geq k_0$, $\dfrac{\mu(A \cap I_k)}{\mu(I_k)} \geq \dfrac{2}{3}\mu(I)$.

The same argument is applied to B to establish the corresonding inequality that there exists a $k' > k_0$ such that $\dfrac{\mu(B \cap J_{k'})}{\mu(J_{k'})} \geq \dfrac{2}{3}\mu(J)$.

Now suppose $k < k'$. Pick I in $\mathscr{T}_k$ and J in $\mathscr{T}_{k'}$. Let I_1, I_2, I_3 denote the subintervals of I in the $(k+1)$th stage then $\mu(A \cap I) = \sum_{i=1}^{3} \mu(A \cap I_i)$. Since $\mu(A \cap I) \geq \frac{2}{3}\mu(A)$ at k we claim that there exists at least one subinterval of I such that $\mu(A \cap I') \geq \frac{2}{3}\mu(I')$ at stage $(k+1)$. This result holds for all subsequent stages. Then by induction we can find levels I and J , l levels apart, in the same tower $\mathscr{T}_n$ that are $\frac{2}{3}$-full of A and B respectively such that $\mu(A \cap I) \geq \frac{2}{3}\mu(I)$ and $\mu(B \cap J) \geq \frac{2}{3}\mu(J)$. $\square$

2.3 The Original Chacon Transformation

The construction of the Chacon transformation in [1] is slightly different from that of the canonical Chacon transformation descibed earlier. The difference between these two methods of constructing the transformation is given next. In the first stage the unit interval in divided into three equal and disjoint intervals I_1^1, I_2^1 and R_1. Just as in the canonical construction, the third portion $R_1 = \left[\frac{2}{3}, 1\right]$ is set aside as the spacer or reservoir.The stacking process begins immediately after this. I_1^2 is stacked on top of I_1^1. This ends stage 1. The transformation T constructed at the end of this stage maps I_1^1 linearly onto I_1^2 and is undefined on R_1.

At the beginning of stage 2, the tower of two levels is cut in half. An interval of equal length to the length of I_1^1 is separated from the right side of R_1. This piece is put on top of the right-side sliced tower and then everything is stacked on top of the left piece. Thus at the end of the second stage we have a tower consisting of five levels and the remaining spacer, R_2. The intervals at this stage satisfy the condition that $h_2 \cdot \mu(I_2^1) + \mu(R_2) = 1$ where h_2 is the height of the tower at this stage and is five.

This inductive process continues and at the nth stage we have the intervals $I_n^1, I_n^2, ..., I_n^{h_n}, R_n$, $h_n + 1$ partitions of the unit interval where $h_n \cdot \mu(I_n^1) + R_n = 1$ with $I_n^1, I_n^2, ...I_n^{h_n}$ all having equal length. The map constructed at this stage is defined on $\cup_{k=1}^{h_n - 1} I_n^k$ and is undefined on the topmost level $I_n^{h_n}$ and the remaining spacer R_n. To form the $(n+1)$th tower, we first create a new partition of the unit interval by splitting the $(n-1)$th tower into two thus creating a new partition of the unit interval $I_{n+1}^1, I_{n+1}^2, ...I_{n+1}^{h_{n+1}}, R_{n+1}$. Note that $I_{n+1}^1, I_{n+1}^2, ...I_{n+1}^{h_{n+1}}$ have equal length. An interval of the same length as one of these pieces is cut off from the R_n, placed on top of the right part of the sliced tower and then everything is stacked on the left part of the sliced tower. The spacer that remains is what forms R_{n+1}. Therefore at the end of the nth stage we have formed a new tower which has height $h_{n+1} = 2h_n + 1$.

Remark: Note that the transformation constructed using this method is the same as the resulting transformation using the previously described method, i.e, splitting intervals into three. It therefore has the same properties, in that, it is measure-preserving, invertible with respect to μ a.e. and it is an example of a weakly mixing transformation which is not strongly mixing. The main advantage of contructing the transformation by splitting the towers into three instead of two is realized when proving that the Chacon transformation is doubly ergodic. Recall that to prove that $\mu(T^H(I) \cap I) > 0$ we studied the left and middle portions of the split tower while the middle and right portions of the tower were used to prove that $\mu(T^H(I) \cap I) > 0$. In constructing the transformation using the original method there is no clear cut division and thus proving the double ergodic property is quite tedious.

Chapter 3

Some Results for Cutting and Stacking Methods

There are several interesting results for cutting and stacking methods. We state without proof results obtained by Arnoux et al [4].

Definition 3.0.1. [5] Let $(X, \mathcal{B}, \mu)$ be a Lebesgue space and T an invertible measure-preserving transformation of $(X, \mathcal{B}, \mu)$. T is said to be *aperiodic* if

$$\mu\left(\bigcup_{k\in\mathbb{Z}, k\neq 0} \{x \in X : T^k(x) = x\}\right) = 0.$$

Definition 3.0.2. An interval exchange transformation (I.E.T) is a kind of dynamical system which is a generalized idea of a circle rotation. I.E.T is a piecewise linear map of $[0, 1]$ into itself. It has the properties of being one-to-one and continuous except for a finite set of points and also preserves Lebesgue measure. An interval exchange transformation acts by cutting the unit interval into several subintervals, and then permuting these subintervals.

Mathematically, let $n > 0$ and let $\mathcal{S}$ be a vector of real numbers $\mathcal{S} = \{s_1, s_2, ..., s_n\}$ such that $\sum_n s_n = 1$. Set $a_0 = 0$, $a_i = \sum_{j=1}^{i} s_j$ and $\lambda_i = [a_{i-1}, a_i)$. Let $\mathcal{P}$ be a permutation of $\{1, 2, ..., n\}$ and set $\mathcal{S}^{\mathcal{P}} = \{s_{\mathcal{P}^{-1}(1)}, s_{\mathcal{P}^{-1}(2)}, ..., s_{\mathcal{P}^{-1}(n)}\}$. The map $T_{\mathcal{P},\mathcal{S}} : [0, 1] \to [0, 1]$ is an interval exchange transformation to the pair $(\mathcal{P}, \mathcal{S})$ defined by

$$T_{\mathcal{P},\mathcal{S}}(x) = x - a_{i-1} + a^{\mathcal{P}}_{\mathcal{S}(i)-1} \text{ for } x \in \lambda_i,\ 0 \leq i \leq n.$$

Theorem 3.0.1. *[4] Any aperiodic measure-preserving transformation on a probability space is isomorphic to a interval exchange transformation* $T : [0,1) \to [0,1)$ *(obtained from cutting and stacking) which satisfies the following properties:*

1. *There is a strictly increasing sequence* $\{t_n\} \subset [0,1)$ *and a sequence* $\{a_n\} \subset \mathbb{R}$ *such that* $t_0 = 0$, $\lim_{n\to\infty} t_n = 1$ *and* $T(x) = x + a_n$ *for all* $x \in [t_{n-1}, t_n)$.

2. $T(I_n) \subset [0,1)$ *for all* $n \in \mathbb{N}$.

3. T *is bijective.*

Remark: This is a very useful and strong theorem in that if given an interval exchange transformation we know we can construct it inductively using a cutting and stacking method. Also there exists an aperiodic transformation which is isomorphic to it. This allows us to move between these two areas and obtain several other interesting results which have been established for aperiodic transformations.

Bibliography

[1] R. V. Chacon. Weakly mixing transformations which are not strongly mixing. *Proc. Amer. Math. Soc.*, 22:559–562, 1969. ISSN 0002-9939.

[2] William Parry. *Topics in ergodic theory*, volume 75 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2004. ISBN 0-521-60490-7. Reprint of the 1981 original.

[3] C. E. Silva. *Invitation to ergodic theory*, volume 42 of *Student Mathematical Library*. American Mathematical Society, Providence, RI, 2008. ISBN 978-0-8218-4420-5.

[4] Pierre Arnoux, Donald S. Ornstein, and Benjamin Weiss. Cutting and stacking, interval exchanges and geometric models. *Israel J. Math.*, 50(1-2):160–168, 1985. ISSN 0021-2172. doi: 10.1007/BF02761122. URL `http://dx.doi.org.libproxy.lib.unc.edu/10.1007/BF02761122`.

[5] Peter Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1982. ISBN 0-387-90599-5.

[6] Nathaniel A. Friedman. *Introduction to ergodic theory*. Van Nostrand Reinhold Co., New York, 1970. Van Nostrand Reinhold Mathematical Studies, No. 29.