THE NOTION OF MIXING AND RANK ONE EXAMPLES

WENYU CHEN

Abstract. This paper discusses the relationships among mixing, lightly mixing, weakly mixing and ergodicity. Besides proving the hierarchy of these notions, we construct examples of rank one transformations that exhibit each of these behaviors to demonstrate the differences.

Contents

1. Preliminaries 1
1.1. Convergence 1
1.2. Sets of Density Zero 2
1.3. Dynamical Systems 2
2. Ergodicity 3
3. Mixing and Weakly Mixing 5
4. Mixing and Lightly Mixing 9
5. Rank One Maps 10
5.1. The Dyadic Odometer 11
5.2. Chacón’s Transformation 15
5.3. Just Lightly Mixing Map 19
6. Discussion 24
Acknowledgments 25
References 26

1. Preliminaries

We assume knowledge of measure theory (for example, chapters 1 and 2 of [5] would suffice). We use the same notation as in standard measure theory.

1.1. Convergence.

Definition 1.1. A bounded sequence \( \{a_i\} \) may converge to a number \( a \) in at least three ways.

(1) Convergence:

\[
\lim_{i \to \infty} a_i = a.
\]

(2) Strong Cesàro convergence:

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} |a_i - a| = 0.
\]
(3) Cesàro convergence:
\[
\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} a_i = a.
\]

**Lemma 1.2.** Let \( \{a_i\} \) be a bounded sequence. Then convergence implies Strong Cesàro convergence, and Strong Cesàro convergence implies Cesàro convergence.

**Proof.** The result follows easily by applying the \( \epsilon - \delta \) definition of limits. \( \square \)

1.2. Sets of Density Zero.

**Definition 1.3.** A set \( D \) of nonnegative integers is said to be of zero density if
\[
\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} I_D(i) = 0,
\]
where \( I_D \) is the characteristic function defined by \( I_D(x) = 1 \) if \( x \in D \) and \( I_D(x) = 0 \) if \( x \notin D \).

**Definition 1.4.** Let \( a \in \mathbb{R} \). We say a sequence \( \{a_i\} \) of real numbers converges in density to \( a \) if there exists a zero density set \( D \subset \mathbb{N} \) such that for any \( \epsilon > 0 \), there is an integer \( N \) such that if \( i > N \) and \( i \notin D \), then \( |a_i - a| < \epsilon \).

**Lemma 1.5.** If \( \{b_i\} \) is a bounded sequence of nonnegative real numbers, then \( \{b_i\} \) converges Cesàro to 0 if and only if \( \{b_i\} \) converges in density to 0.

**Proof.** The result follows easily by applying the definitions of convergences. \( \square \)

1.3. Dynamical Systems.

**Definition 1.6** (Dynamical System). A measurable dynamical system is a quadruple \((X, S, \mu, T)\), where \( X \) is a space, \( S \) a \( \sigma \)-algebra of subsets of \( X \), \( \mu \) a measure on \( S \) and \( T : X \to X \) a measurable transformation with respect to \( \mu \). Specifically, \((X, S, \mu, T)\) is a Lebesgue measure-preserving dynamical system if \((X, S, \mu)\) is a Lebesgue measurable space and \( T \) is a measure-preserving transformation \( T : X \to X \). We will refer to such system simply as measure-preserving dynamical system in this paper.

Though we are not going to study the “sameness” of dynamical systems, we define the notion of factor for technical reasons. A factor of a dynamical system carries some of its dynamical properties. Later we show if a dynamical system has an ergodic transformation, then so does its factor.

**Definition 1.7.** Let \((X, S, \mu, T)\) be a measurable dynamical system. A dynamical system \((X', S', \mu', T')\) is a factor of \((X, S, \mu, T)\) if there exist measurable sets \( X_0 \subset X \) and \( X_0' \subset X' \) of full measure with \( T(X_0) \subset X \), \( T'(X_0') \subset X' \), and a factor map \( \phi : X_0 \to X_0' \) such that:

1. \( \phi \) is onto
2. \( \phi^{-1}(A) \in S(X_0) \) for all \( A \in X_0' \)
3. \( \mu(\phi^{-1}(A)) = \mu'(A) \) for all \( A \in S'(X_0') \)
4. \( \phi(T(x)) = T'(\phi(x)) \).
2. Ergodicity

**Definition 2.1** (Ergodicity). A measure-preserving transformation $T$ is **ergodic** if whenever $A$ is a measurable set that is strictly $T$-invariant (i.e., $T^{-1}(A) = A$), then $\mu(A)$ is either 0 or 1.

**Theorem 2.2** (Birkhoff’s ergodic theorem). Let $(X, S, \mu)$ be a probability space and let $T$ be a measure-preserving transformation on $(X, S, \mu)$. If $f : X \to \mathbb{R}$ is an integrable function, then

1. $\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x))$ exists for all $x \in X \setminus N$, for some null set $N$ depending on $f$. Denote this limit by $\tilde{f}$.
2. $\tilde{f}(Tx) = \tilde{f}(x)$ a.e.
3. For any measurable set $A$ that is $T$-invariant,
   \[ \int_A f \, d\mu = \int_A \tilde{f} \, d\mu. \]

In particular, if $T$ is ergodic, then

\[ \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)) = \int f \, d\mu \text{ a.e.} \]

For the proof, see chapter 5 of [9].

**Theorem 2.3.** Let $T$ be a finite measure-preserving transformation on a probability space $(X, S, \mu)$. $T$ is ergodic if and only if for all measurable sets $A$ and $B$,

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

**Proof.** First we prove the forward direction. Suppose $T$ is ergodic and $A$, $B$ measurable. Apply the Birkhoff ergodic theorem on $T$, $A$, $B$, and the characteristic map $\mathbb{1}_A \mathbb{1}_B$ to get

\[ \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{1}_A(T^i(x)) = \int \mathbb{1}_A \, d\mu = \mu(A) \text{ a.e.} \]

Thus,

\[ \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{1}_A(T^i(x)) \mathbb{1}_B(x) = \mu(A)\mathbb{1}_B(x) \text{ a.e.} \]

Observe that $|\mathbb{1}_A(T^i(x))\mathbb{1}_B(x)| \leq 1$ a.e., which enables us to apply the dominated convergence theorem. This gives

\[ \lim_{n \to \infty} \int \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{1}_A(T^i(x)) \mathbb{1}_B(x) \, d\mu(x) = \int \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{1}_A(T^i(x)) \mathbb{1}_B(x) \, d\mu(x) = \int \mu(A) \mathbb{1}_B(x) \, d\mu(x) = \mu(A)\mu(B). \]
Also, we know
\[\int \frac{1}{n} \sum_{i=0}^{n-1} I_A(T^i(x)) I_B(x) d\mu(x) = \frac{1}{n} \sum_{i=0}^{n-1} \int I_A(T^i(x)) I_B(x) d\mu(x) = \frac{1}{n} \sum_{i=0}^{n-1} \mu(T^{-i}(A) \cap B).\]

We take the limit of both sides and obtain what we want to show.

Now we prove the reverse direction. Suppose (2.4) holds for all sets $A$ and $B$ in $X$. Then it also holds for $A = B$. Let $A$ be $T$-invariant. By the definition of $T$-invariant, \(\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(T^{-i}(A) \cap A) = \mu(A)\). Let $A = B$. We get \(\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(T^{-i}(A) \cap A) = \mu(A)^2\). This gives $\mu(A) = \mu(A)^2$, which implies $\mu(A)$ is 0 or 1. Thus $T$ is ergodic.

**Proposition 2.5.** Let $T$ be defined on a probability space $(X, \mathcal{S}(X), \mu)$, $T \times T$ defined on $(X \times X, \mathcal{S}(X) \times \mathcal{S}(X), \mu \times \mu)$. If $T \times T$ is an ergodic transformation, then $T$ is also ergodic.

**Proof.** Let $A$ be a strictly $T$-invariant measurable set. Then $(T \times T)^{-1}(A \times A) = T^{-1}(A) \times T^{-1}(A) = A \times A$, so $A \times A$ is $(T \times T)$-invariant. The property of product measure shows

\[(\mu \times \mu)(A \times A) = \mu(A)\mu(A).\]

Since $T \times T$ is ergodic, $(\mu \times \mu)(A \times A)$ is 1 or 0. Therefore $\mu(A)$ is either 1 or 0. Hence $T$ is ergodic. \(\square\)

**Theorem 2.6.** Let $(Y, \mathcal{S}'(Y), \nu, S)$ be a factor of $(X, \mathcal{S}(X), \mu, T)$. If $T$ is ergodic, so is $S$.

**Proof.** Let $\phi : X \to Y$ be a factor map. If $A$ is a strictly $S$-invariant set, then

\[T^{-1}(\phi^{-1}(A)) = \phi^{-1}(S^{-1}(A)) = \phi^{-1}(A).\]

(The first equality holds by Definition 1.7 and the second holds because $A$ is $S$-invariant.) Thus $\phi^{-1}(A)$ is a $T$-invariant set. Also by Definition 1.7, $\mu(\phi^{-1}(A)) = \nu(A)$. Since $T$ is ergodic, $\nu(A)$ is 0 or 1, i.e., $S$ is ergodic. \(\square\)

**Definition 2.7** (Eigenvalue and Eigenfunction). Let $(X, \mathcal{S}, \mu)$ be a probability space and let $T : X \to X$ be a measure-preserving transformation. We say a number $\mu \in \mathbb{C}$ is an **eigenvalue** of $T$ if there exists a function $f \in L^2(X, \mathcal{S}, \mu)$ which is nonzero a.e. such that

\[f(T(x)) = \lambda f(x) \text{ } \mu\text{-a.e.}\]

The function $f$ is called an **eigenfunction**.

**Lemma 2.8.** If $\lambda$ is an eigenvalue for a measure-preserving transformation $T$, then $|\lambda| = 1$.

**Proof.** Let $T$ be a measure-preserving transformation and $f$ be its eigenfunction. Let $\lambda$ be its eigenvalue. Then $|\lambda|^2 \int |f|^2 d\mu = \int |\lambda|^2 |f|^2 d\mu = \int |f| f \circ T \circ f d\mu = \int |f|^2 d\mu = \int |f| d\mu$. \(\square\)

**Lemma 2.9.** If $T$ is ergodic and $f$ is an eigenfunction, then $|f|$ is constant a.e.
Proof. Let \( f \) be an eigenfunction of \( T \) with eigenvalue \( \lambda \), i.e., \( f \circ T = \lambda f \) a.e. By Lemma 2.8, we know \( |\lambda| = 1 \). Then \( |f \circ T| = |\lambda f| = |\lambda||f| = |f| \). Since \( T \) ergodic, \( |f| \) is constant a.e. \( \square \)

Corollary 2.10. If \( T \) is ergodic and \( f \) is its eigenfunction, then \( \frac{f}{|f|} \) is also an eigenfunction. Moreover, \( \frac{f}{|f|} \) has absolute value 1. Thus we may choose an eigenfunction \( f \) of \( T \) with \( |f| = 1 \) a.e. (i.e., the values of \( f \) lie in the unit circle).

Proof. Note that \( |f| \neq 0 \) a.e. since \( f \) is non-zero a.e. \( \square \)

Definition 2.11 (Continuous Spectrum). A measure-preserving transformation \( T \) is said to have \textbf{continuous spectrum} if \( \lambda = 1 \) is its only eigenvalue and \( T \) is simple; equivalently, \( T \) is ergodic and \( \lambda = 1 \) is its only eigenvalue.

3. Mixing and Weakly Mixing

The notion of mixing is an abstract concept originated from the study of physics. We formalize this notion to describe the behavior of transformations in dynamical systems. In order to study a transformation \( T \) on a probability space, we study \( T^n(A) \cap B \) for measurable sets \( A \) and \( B \). Intuitively, the ideally “mixed” state is that the size of \( A \) in \( B \) is proportional to the measure of \( A \) and \( B \), i.e., after applying \( T \) for some \( n \) times, \( \mu(T^n(A) \cap B) = \mu(A)\mu(B) \). For example, if we mix 1 ounce of Vodka with 1 ounce of Gin, we want any subset of the new drink to be half-Vodka-half-Gin. Transformations that make the system converge to this state is defined as “mixing” in Definition 3.1.

However, we can also consider a weaker mode of convergence to the ideal state. We define it as weakly mixing. Next chapter we will discuss yet another notion of mixing that only requires the limit inferior of \( T^n(A) \cap B \) to be positive, not caring whether is it proportional to the measure of \( A \) and \( B \).

Now we explore the relationships between mixing, weakly mixing and ergodicity. For finite measure-preserving transformations, we see that

1. mixing \( \implies \) weakly mixing
2. weakly mixing \( \implies \) ergodic
3. weakly mixing \( \iff \) doubly ergodic.

Definition 3.1 (Mixing). A measure-preserving transformation \( T \) on a probability space \( (X, S, \mu) \) is \textbf{mixing} if for all measurable sets \( A \) and \( B \),

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

Definition 3.2 (Weakly mixing). A measure-preserving transformation \( T \) on a probability space \( (X, S, \mu) \) is \textbf{weakly mixing} if for all measurable sets \( A \) and \( B \),

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

Definition 3.3 (Doubly ergodic). A measure-preserving transformation \( T \) on a probability space \( (X, S, \mu) \) is \textbf{doubly ergodic} if for all measurable sets \( A \) and \( B \), there exists an integer \( n > 0 \) such that

\[
\mu(T^{-n}(A) \cap A) > 0 \text{ and } \mu(T^{-n}(A) \cap B) > 0.
\]
Remark 3.4. Construct a sequence \( \{a_i\} \) with \( a_i(A,B) = \mu(T^{-i}(A) \cap B) \). Notice that (2.4) shows that ergodicity is equivalent to Cesàro convergence of \( \{a_i\} \) to \( \mu(A)\mu(B) \). From the definitions above, we also know weakly mixing is equivalent to strong Cesàro convergence of \( \{a_i\} \) to \( \mu(A)\mu(B) \), and mixing is equivalent to convergence of \( \{a_i\} \) to \( \mu(A)\mu(B) \).

Theorem 3.5. Let \( T \) be a probability-preserving transformation.

(1) If \( T \) is weakly mixing, then it is ergodic.

(2) If \( T \) is mixing, then it is weakly mixing.

Proof. This result follows from Remark 3.4 and Lemma 1.2. \( \square \)

The next lemma provides sufficient conditions for ergodicity, mixing, and weakly mixing that will be helpful in later proofs. The conditions rely on the notion of a sufficient semi-ring.

Definition 3.6 (Semi-Ring). A semi-ring on a nonempty set \( X \) is a collection \( \mathcal{R} \) of subsets of \( X \) such that

1. \( \mathcal{R} \) is nonempty;
2. if \( A, B \in \mathcal{R} \), then \( A \cap B \in \mathcal{R} \);
3. if \( A, B \in \mathcal{R} \), then
   \[
   A \setminus B = \bigcup_{j=1}^{n} E_j,
   \]
   where \( E_j \in \mathcal{R} \) are disjoint.

Definition 3.7. Let \( (X, \mathcal{S}, \mu) \) be a measure space. A semi-ring \( \mathcal{C} \) of measurable subsets of \( X \) of finite measure is said to be a sufficient semi-ring for \( (X, \mathcal{S}, \mu) \) if it satisfies: for every \( A \subset \mathcal{S} \),

\[
\mu(A) = \inf \left\{ \sum_{j=1}^{\infty} \mu(I_j) : A \subset \bigcup_{j=1}^{\infty} I_j \text{ and } I_j \in \mathcal{C} \text{ for } j \geq 1 \right\}.
\]

Lemma 3.8. Let \( T \) be a measure-preserving transformation on a probability space \( (X, \mathcal{S}, \mu) \) with a sufficient semi-ring \( \mathcal{C} \).

1. If for all \( I, J \in \mathcal{C} \), \( \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(T^{-i}(I) \cap J) = \mu(I)\mu(J) \), then \( T \) is ergodic.

2. If for all \( I, J \in \mathcal{C} \), \( \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} |\mu(T^{-i}(I) \cap J) - \mu(I)\mu(J)| = 0 \), then \( T \) is weakly mixing.

3. If for all \( I, J \in \mathcal{C} \), \( \lim_{n \to \infty} \mu(T^{-i}(I) \cap J) = \mu(I)\mu(J) \), then \( T \) is mixing.

Proof. We provide a brief sketch of the proof. For a complete proof, see chapter 6.3 of [9] for details. Let \( A, B \) be two sets of positive measures in \( \mathcal{S} \). The proof is done by constructing \( E \) and \( F \) which are both finite disjoint unions of finite sequences of sets in \( \mathcal{C} \) such that the measure of set difference between \( A \) and \( E \), \( B \) and \( F \) are within \( \epsilon \). Then we can show by computation that if the limits in concern are bounded by \( \epsilon \) for such \( E \) and \( F \), then they are also bounded for \( A \) and \( B \). \( \square \)

Remark 3.9. We can now study a sufficient semi-ring of a measure space and extend the result to all measurable sets in the space. One typical sufficient semi-rings we utilize is intervals with dyadic rational endpoints. Another example: for \( (X, \mathcal{S}(X), \mu) \) and \( (Y, \mathcal{S}(Y), \nu) \), define \( \mu \times \nu \) on \( \mathcal{S}(X) \times \mathcal{S}(Y) \) by \( (\mu \times \nu)(A \times B) = \)
Then the semi-ring of measurable rectangles is a sufficient semi-ring for the extension measure \( \mu \times \nu \).

Now we explore the relations between ergodicity, mixing, and weakly mixing.

**Proposition 3.10.** Let \( T \) be a measure-preserving transformation on a probability space \((X, S, \mu)\). Then 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
   \[
   \lim_{i \to \infty, i \not\in D} \mu(T^{-i}(A) \cap B) = \mu(A)\mu(B).
   \]
3. For each pair of measurable sets \( A \) and \( B \),
   \[
   \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} (\mu(T^{-i}(A) \cap B) - \mu(A)\mu(B))^2 = 0.
   \]

**Proof.** (1) \( \Leftrightarrow \) (2): Let \( b_i = |\mu(T^{-i}(A) \cap B) - \mu(A)\mu(B)| \).

Then \( \{b_i\} \) is a sequence of bounded, non-negative real numbers. Notice that (1) is the condition that \( \{b_i\} \) converges Ces\'aro to 0 and (2) is the condition that \( \{b_i\} \) converges in density to 0. Hence, by Lemma 1.2, (1) and (2) are equivalent.

(2) \( \Leftrightarrow \) (3): By the fact that \( \{b_i\} \) converges to 0 if and only if \( \{b_i^2\} \) converges to 0, (2) is equivalent to

\[
\lim_{i \to \infty, i \not\in D} (\mu(T^{-i}(A) \cap B) - \mu(A)\mu(B))^2 = 0.
\]

By Lemma 1.5, (3.11) is equivalent to (3). This completes the proof. \(\square\)

**Theorem 3.12.** Let \( T \) be a measure-preserving transformation on a probability space \((X, S, \mu)\). Then the following are equivalent.

1. \( T \) is weakly mixing.
2. \( T \times T \) is weakly mixing.
3. \( T \times T \) is ergodic.

**Proof.** (1) \( \Rightarrow \) (2): Assume \( T \) is weakly mixing. Let \( A, B, C, D \) be measurable subsets of \( X \). By Proposition 3.10, there exist \( D_1 = D_1(A, B) \) and \( D_2 = D_2(C, D) \) such that for \( n \notin D_1 \cup D_2 \),

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

\[
\lim_{n \to \infty} \mu(T^{-n}(C) \cap D) = \mu(C)\mu(D).
\]

Then for all \( n \notin D_1 \cup D_2 \),

\[
\lim_{n \to \infty} (\mu \times \mu)[(T \times T)^{-n}(A \times C) \cap (B \times D)] = (\mu \times \mu)(A \times C)(\mu \times \mu)(B \times D).
\]

By the observation in Remark 3.9, (3.13) holds on a sufficient semi-ring. Then by Lemma 3.8, it holds for all measurable sets of \( X \). Then by Proposition 3.10, (3.13) implies \( T \times T \) is weakly mixing.

(2) \( \Rightarrow \) (3): This follows from Theorem 3.5.
(3)⇒(1): Assume $T \times T$ is ergodic. We will show the condition in Proposition 3.10(3) holds. Expand the coefficients in that condition:

\[
\frac{1}{n} \sum_{i=0}^{n-1} (\mu(T^{-i}(A) \cap B) - \mu(A)\mu(B))^2
\]

(3.14) \[
= \frac{1}{n} \sum_{i=0}^{n-1} (\mu(T^{-i}(A) \cap B))^2
\]

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

We compute each of the three terms on the right hand side. Since $T \times T$ is ergodic, $T$ is ergodic by Proposition 2.5. Thus apply Theorem 2.3 on $T$, $A$, $B$ to get

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

Apply Theorem 2.3 on $T \times T$, $A \times A$, $B \times B$ to get

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

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

Thus, (3.14) equals $\mu(A)^2\mu(B)^2 - 2\mu(A)^2\mu(B)^2 + \mu(A)^2\mu(B)^2 = 0$. By Proposition 3.10, $T$ is weakly mixing.

The next theorem gives the equivalent condition for weakly mixing and doubly ergodic transformations.

**Theorem 3.15.** Let $T$ be an invertible measure-preserving transformation on a Lebesgue probability space $(X, S, \mu)$. Then the following are equivalent.

1. $T$ is weakly mixing.
2. $T$ has continuous spectrum. (Recall Definition 2.11.)
3. $T$ is doubly ergodic.
4. $T \times S$ is ergodic for any ergodic, finite measure-preserving transformation $S$.

**Proof.**

(1)⇒(2): Assume $T$ is weakly mixing. By Theorem 3.12, $T$ and $T \times T$ are ergodic. By Lemma 2.8 and Corollary 2.10, there exist an eigenvalue $\lambda$ and an eigenfunction $f \in L^2$, with $|\lambda| = 1$ and $|f| = 1$ such that $f(T(x)) = \lambda f(x)$ a.e. To show $T$ has continuous spectrum, we need to show $\lambda = 1$ is the only eigenvalue. Define $g : X \times X \to \mathbb{C}$ by $g(x, y) = f(x)\bar{f}(y)$. Since $\lambda \bar{\lambda} = |\lambda|^2 = 1$,

\[
g(T(x), T(y)) = f(T(x))\bar{f}(T(y)) = \lambda f(x)\bar{\lambda}\bar{f}(y) = f(x)\bar{f}(y) = g(x, y).
\]

This shows $g$ is the eigenfunction for $T \times T$. Since $T \times T$ is ergodic, $g$ is constant a.e.; thus $\bar{f}$ is constant a.e.

(1)⇒(3): Assume $T$ is weakly mixing, and $A$ and $B$ are measurable sets of positive measures. Then by Proposition 3.10 there exist $D_1 = D_1(A, B), D_2 = \ldots$
$D_2(A, A)$ such that for $n \notin (D_1 \cup D_2)$,
\[
\lim_{n \to \infty} \mu(T^{-n}(A) \cap B) = \mu(A)\mu(B),
\]
\[
\lim_{n \to \infty} \mu(T^{-n}(A) \cap A) = \mu(A)\mu(A).
\]
Since $\mathbb{Z}_{>0} - (D_1 \cup D_2)$ has density 1, there exists a nonnegative integer $n$ such that $\mu(T^{-n}(A) \cap A) > 0$ and $\mu(T^{-n}(A) \cap B) > 0$, as required.

(2)⇒(4): This proof involves a lot of functional analysis, which is not a focus of this paper, so we only present a outline. For the complete proof of this theorem, see [8, page 67-71]. We define fiber sets $A_x = \{y \in Y : (x, y) \in A\}$. If $T \times S$ is not ergodic, then there would be an invariant set $(T \times S)(A) = A$. We then know $S(A_x) = A_{T(x)}$ for a.e. $x$. So $S$ takes the fiber at $x$ to the fiber at $T(x)$. Define the metric $d(x, x') = \mu(A_x \triangle A_{x'})$. Identify $x \sim x'$ if and only if $A_x = A_x'$ a.e. Then $X/ \sim$ with $d$ is a metric space. We claim that $T$ acts on $X/ \sim$ as an isometry and this space is compact. Then $T$ on $X/ \sim$ is isomorphic to a rotation on a compact map, which can be shown to admit a non-constant eigenfunction. This is a contradiction.

(3)⇒(2): Assume $T$ is doubly ergodic. By definition, $T$ is also ergodic. Similar to the previous proof, there exist an eigenvalue $\lambda$ and an eigenfunction $f \in L^2$ with $|\lambda| = 1$, $|f| = 1$ such that $f(T(x)) = \lambda f(x)$ a.e.

We aim to show $\lambda = 1$ is the only eigenvalue. We will proceed by contradiction. Suppose there are multiple eigenvalues on unit circle. Parametrize this unit circle. Write $\lambda = e^{2\pi i \alpha}$ and $f(x) = e^{2\pi i g(x)}$ for some $\alpha \in [0, 1)$ and measurable $g : X \to [0, 1)$. Define $R : [0, 1) \to [0, 1)$ by $R(t) = \alpha + t \pmod{1}$. Then $g \circ T = R \circ g$.

Now we construct a factor of the probability space as below: Let $\nu$ be a measure on $[0, 1)$ with $\nu(A) = \mu(g^{-1}(A))$. Then $g$ is a factor map from $T$ to $R$. By Theorem 2.6, $R$ is ergodic because $T$ is. There are two cases for $R$: (a) If $\alpha$ is rational, then $\nu$ is atomic and concentrated on finitely many points, thus not doubly ergodic. (This can be shown by taking $A$ and $B$ to be sets with singleton points.) (b) If $\alpha$ is irrational, then $\nu$ is a Lebesgue measure. Lebesgue measure is not doubly ergodic for irrational rotation by the following argument: Let $A = [0, \frac{1}{2})$ and $B = [\frac{1}{2}, \frac{3}{4})$. Then for any integer $n$ such that $R^n(A) \cap B \neq 0$, $R^n(A) \cap A = 0$. Since in both cases $R$ is not doubly ergodic, we get a contradiction. Thus $\lambda = 1$ is the only eigenvalue, i.e., $T$ has continuous spectrum.

(4)⇒(1): Let $S = T$; the result follows directly from Theorem 3.12.

\[\square\]

Remark 3.16. (1) and (3) are proved to be equivalent for finite measure-preserving transformations, but the result may not always hold for infinite measures. In fact, [2] shows that for infinite measure-preserving transformations, doubly ergodic implies weakly mixing but there exists a weakly mixing map that is not doubly ergodic.

4. Mixing and Lightly Mixing

As we introduced in the previous chapter, lightly mixing is a notion that is stronger than weakly mixing but weaker than mixing. Now we study its relationship with the other notions.

Definition 4.1 (Lightly mixing). A finite measure-preserving transformation $T$ on a probability space $(X, S, \mu)$ is \textbf{lightly mixing} if for all measurable sets $A, B$ of
positive measure,
\[ \liminf_{n \to \infty} \mu(T^{-n}(A) \cap B) > 0. \]

**Lemma 4.2.** \( T \) is lightly mixing if and only if for each set \( A \) of positive measure, there exists \( N \) such that \( \mu(T^{-n}(A) \cap A) > 0 \) for all \( n \geq N \).

**Proof.** The forward direction follows by letting \( B = A \) in the definition of lightly mixing. For the other direction, we prove the contrapositive. Suppose \( T \) is not lightly mixing. Then there exists a set \( E \) of positive measure such that \( \liminf_{n \to \infty} \mu(T^{-n}(E) \cap E) = 0 \). Then we can always pick \( n_k \) large enough such that \( E_k = T^{-n_k}(E) \cap E \) satisfies \( \mu(E_k) < \frac{\mu(E)}{3} \).

Now let \( F = E - \bigcup_{k=1}^{\infty} E_k \) so that \( F \) is the part of the set that “does not come back.” Then we have
\[
\mu(F) \geq \mu(E) - \sum_{k=1}^{\infty} \mu(E_k) \quad \text{by countable additivity of } \mu
\]
\[
> \mu(E) - \frac{\mu(E)}{2} = \frac{\mu(E)}{2} \quad \text{by sum of geometric series.}
\]

Then this shows \( F \) has positive measure. Notice \( T^{-n_k}(F) \cap F \) is empty for all \( k \), since \( T^{-n_k}(A) \subset T^{-n_k}(E) \) but \( T^{-n_k}(E) \cap A = \emptyset \). \( \square \)

**Proposition 4.3.** Let \( T \) be a finite measure-preserving transformation on a probability space \((X, \mathcal{S}, \mu)\). If \( T \) is lightly mixing, then \( T \) is weakly mixing.

**Proof.** Apply Lemma 4.2: if \( T \) is lightly mixing, then for each set \( A \) of positive measure, there exists \( N \) such that \( \mu(T^{-n}(A) \cap A) > 0 \) for all \( n \geq N \). Fix this \( N \).

There exists an integer \( m > N \) such that \( \mu(T^{-m}(A) \cap B) > 0 \) by property of limit inferior. We have \( \mu(T^{-m}(A) \cap A) > 0 \) and \( \mu(T^{-m}(A) \cap B) > 0 \), so \( T \) is doubly ergodic. For finite measure this means \( T \) is weakly mixing by Theorem 3.15. \( \square \)

5. **Rank One Maps**

In Theorem 3.5 we showed that mixing implies weakly mixing, and that weakly mixing implies ergodic. In this section we show by constructing examples that the converses do not hold. To do this, we construct a map that is ergodic but not weakly mixing in Section 5.1, a map that is weakly mixing but not lightly mixing in Section 5.2 and a map that is lightly mixing but not mixing in Section 5.3. These constructions employ a technique called “cutting and stacking.”

These transformations are also called “rank one transformations” because they are constructed in a particularly simple way. (See Remark 5.19.) We will give a rigorous definition of rank one after we show the process of construction.

For this chapter, \( \mu \) denotes the Lebesgue measure.

**Construction 5.1** (Shifting Map for Two Intervals). Given two intervals \( I = [a, b] \) and \( J = [c, d] \) of the same length, define \( T_{I,J} : I \to J \) by \( T_{I,J}(x) = x + c - a \). It is easy to check that:

1. \( T_{I,J} \) is determined by \( I \) and \( J \) and is one-to-one.
2. for any measurable set \( A \subset I \), \( T_{I,J}(A) \subset J \) is measurable and \( \mu(T_{I,J}(A)) = \mu(A) \).
3. if \( I' \) and \( J' \) are both dyadic subintervals of \( I \) and \( J \) of the same order (i.e., the end points of \( I \) and \( J \) are dyadic rationals of the same order), then \( T_{I,J} \) agrees with \( T_{I',J'} \) on \( I' \). (See Figure 1a.)
Definition 5.2 (Column). Let a column be a finite sequence of disjoint intervals of the same length. Each interval is called a level and the number of intervals in a column is called the height of the column. (See Figure 1b.)

Definition 5.3 (Shifting Map for a Column). If C is a column, we define $T_C$, the shifting map for C, as follows: for each level $I$ of C except for the top level, $T_C$ maps $I$ to the level directly above it, via the shifting map for two intervals. (See Construction 5.1.)

For example, suppose $C$ has levels $I_1, \ldots, I_n$ with $I_1$ the bottom and $I_n$ the top. Then $T_C$ is defined to be $T_{I_1, I_2}$ on $I_1$, $T_{I_2, I_3}$ on $I_2$, ..., and $T_{I_{n-1}, I_n}$ on $I_{n-1}$, and not defined on $I_n$. (Notice that the domain of $T_C$ is all levels in C except for the top one.)

The three maps in this chapter are all constructed by applying this shifting map on different types of columns and taking the “limit map.”

5.1. The Dyadic Odometer.

Construction 5.4. The construction is basically “cutting column into two, and stacking the right above the left”. We inductively construct the columns $C_0$, $C_1$, $C_2$, ...

Base construction: Consider the interval $[0, 1)$. Let $C_0$ denote the first column. Then,

$$C_0 = ([0, 1)), \ h_0 = 1.$$ 

Construction of $C_1$: Divide $C_0$ into two disjoint subintervals $[0, \frac{1}{2})$ and $[\frac{1}{2}, 1)$. Stack the right above the left. (See Figure 2a.) Then,

$$C_1 = ([0, \frac{1}{2}), [\frac{1}{2}, 1)), \ h_1 = 2.$$ 

Induction: To obtain $C_{n+1}$ from $C_n = (I_{n,0}, I_{n,1}, \ldots, I_{n,h_n-1})$, divide each level into two disjoint subintervals of the same length as above. In this way the column $C_n$ is divided into two subcolumns $C_n[0]$ and $C_n[1]$, representing the left and the right subcolumns. Stack $C_n[1]$ above $C_n[0]$. (See Figure 2c.)

Since each level is cut into two, we have $h_{n+1} = 2h_n$ and $h_n = 2^n$. 

\[ \text{Figure 1. shifting map and columns} \]
Notation 5.5 (Naming the subintervals). Let $C_n$ denote the $n$th column, $h_n$ denote the height of $C_n$, and $I_{n,m}$ denote the $m$th level in $C_n$, counting from the bottom ($m = 0, 1, \ldots, h_n - 1$).

For dyadic columns, name the leftmost sublevel of a level $I$ in $C_n$ as $I_{0}$ and the rightmost $I_{1}$. For $a_0, \ldots, a_k \in \{0, 1\}$, define $I[a_0a_1 \ldots a_k]$ recursively to represent each level in $C_n$ in this way: if $I[a_0a_1 \ldots a_{k-1}]$ is defined, then $I[a_0a_1 \ldots a_{k-1}0]$ is the left sublevel of $I[a_0a_1 \ldots a_{k-1}]$ and $I[a_0a_1 \ldots a_{k-1}1]$ is the right sublevel.

Construction 5.6 (Sequence of Partial Maps $T_{C_n}$). Now we apply the shifting map on the sequence of columns we constructed. Recall Definition 5.3. For each $n$, define $T_{C_n}$. It is easy to see that $T_{C_{n+1}}$ agrees with $T_{C_n}$ whenever $T_{C_n}$ is defined.

Notice that $T_{C_n}$ is undefined on only the top level of $C_n$, of measure $\frac{1}{2^n}$.

Remark 5.7. Figure 2c is an intuitive way of representing the columns. In $C_n$, each level is mapped into the level directly above it. Another way of writing the column is to consider $C_n$ as an ordered tuple. Then $C_{n+1} = (I_{n,0}[0], \ldots, I_{n, h_n-1}[0], I_{n,0}[1], \ldots, I_{n, h_n-1}[1])$. Rename the subintervals as $C_{n+1} = (I_{n+1,0}, \ldots, I_{n+1, h_n+1-1})$. Each of the elements in the set is mapped into the element immediately after it. It is easy to see that $T_{C_n}$ is not defined only on the last element of this set. Also notice that by our construction, $I_{n,i}[0]$ is in fact $I_{n+1,i}$ and $I_{n,i}[1]$ is $I_{n+1, i+h_n+i}$.

Definition 5.8 (Dyadic Odometer). Let $T : [0, 1) \rightarrow (0, 1]$ be the map $T(x) = \lim_{n \rightarrow \infty} T_{C_n}(x)$. This map is called the dyadic odometer.

Remark 5.9 (Invertibility). This map is well defined since for each $x \in [0, 1)$, there is some $n > 0$ such that $x$ is in some level of $C_n$ that is not the top. $T$ is one-to-one with a similar argument. Also it is easy to check that $T^{-1}$ is defined for all $x \in (0, 1]$, i.e., $T$ is invertible. The invertibility is very important because it allows us to apply Theorem 3.15.
Remark 5.10. In this chapter, each of the maps we construct is invertible. Also, the columns we construct are all $T$-invariant for $T$ the limit of shifting map. Since sets that are positive invariant are strictly $T$-invariant mod $\mu$ for $T$ invertible and measure-preserving, we can study the forward image of $T$ and conclude the same result as studying the pre-image of $T$. That is, instead of studying the behavior of $\mu(T^{-1}(A) \cap B)$, we can just study the behavior of $\mu(T(A) \cap B)$, which is more intuitive in specific examples.

Now we show some properties of this map.

Lemma 5.11. Let $T$ be the dyadic odometer. Then:

(1) for all $n > 0$, $T(I_{n,h_n-1}) = I_{n,0}$.
(2) for all $n > 0$, $i = 0, \ldots, h_n - 1$, $T^{h_k}(I_{n,i}) = I_{n,i}$ for all $k \geq n$.

Proof. (1): Let $I$ and $J$ be the top and bottom level of $C_n$. That is, $I = I_{n,h_n-1}$ and $J = I_{n,0}$. We prove by induction. Base step: $T(I[0]) = J[1]$ by definition of $T_{C_n}$.
Induction step: if $T(I[1\ldots 10]) = J[0\ldots 01]$ by $T_{C_{k+1}}$. So $T(I[1\ldots 10]) = J[0\ldots 01] \forall k$. Thus $I[1] = \bigcup_{k>0} I[1\ldots 10]$, so $T(I) = T(I[0]) + T(I[1]) = J[1] + \sum_{k>0} T(I[1\ldots 10]) = J$ as required.

(2): Since both (1) and $T_{C_n}(I_{n,i}) = I_{n,i+1}$ hold and $T$ agrees with $T_{C_n}$, we have $T^{h_n}(I_{n,i}) = I_{n,i}$ for all $0 \leq i < h_n$. Thus, also $T^{h_n}(T^{h_n}(I_{n,i})) = I_{n,i}$. (2) can be shown by induction.

![Figure 3. $T(I_{n,h_n-1}) = I_{n,0}$](image)

Definition 5.12. Given measurable sets $A$ and $I$ and $0 < \alpha \leq 1$, we say $I$ is $\alpha$-full of $A$ if $\mu(A \cap I) > \alpha \mu(I)$.

Lemma 5.13. Let $(X, \mathcal{L}, \mu)$ be a nonatomic measure space with a sufficient semi-ring $\mathcal{C}$. If $A \in \mathcal{L}$ is of finite positive measure, then for any $1 > \alpha > 0$, there exists $I \in \mathcal{C}$ such that $I$ is $\alpha$-full of $A$.

Proof. The proof is easily done in measure theory by applying the properties of semi-rings.
The notion of “α-full of” is very important in showing the ergodicity of cutting and stacking constructions. It provides a notion of “large proportion” of sets. Lemma 5.13 makes it possible for any measurable set in the σ-algebra to have a set in the semi-ring that is a “large proportion” of it. Thus, this lemma enables us to pick levels in the dyadic odometers to approximate any set in [0, 1).

Lemma 5.14. Let A be a set of positive measure and I be a dyadic interval that is $\frac{3}{4}$-full of A. Let $I_0$ and $I_1$ be the left and right half of I. Then one of $I_0$ and $I_1$ is $\frac{1}{2}$-full of A and both $I_0$ and $I_1$ are $\frac{1}{2}$-full of A.

Proof. Prove both statements by their contrapositives: if none of the two intervals is $\frac{1}{2}$-full of A, then their union is not. If either is not $\frac{1}{2}$-full of A, the union cannot be $\frac{1}{2}$-full of A even if the other interval is full of A. □

Theorem 5.15. If $T$ is the dyadic odometer, then $T$ is invertible mod $\mu$ on [0, 1). Also $T$ is measure-preserving and ergodic.

Proof. Recall that the dyadic intervals forms a sufficient semi-ring. We know for a measure space $(X, S, \mu)$ with a sufficient semi-ring $C$, if $T^{-1}(I)$ is measurable and $\mu(T^{-1}(I)) = \mu(I)$ for all $I$ in $C$, then $T$ is measure-preserving. (For the proof, see chapter 3.4 of [9].) Then the dyadic odometer is measurable and measure-preserving. The argument in Remark 5.9 shows that $T$ is invertible mod $\mu$.

Now we show $T$ is ergodic: Let $A_1, B_1$ be two sets of positive measures in [0, 1). There exist dyadic intervals $I$ and $J$ that are $\frac{1}{4}$-full of $A_1$ and $B_1$. If $I$ and $J$ have different measures, say, $\mu(I) < \mu(J)$, then since Lemma 5.14 implies at least half of $J$ is $\frac{3}{4}$-full of $B_1$, we can divide $J$ and rename the half that is $\frac{3}{4}$-full of $B_1$ as $J$. Continue this process until $I$ and $J$ have the same measure (which will eventually happen since they are dyadic). Then they are both levels of the same column (which follows from construction). Name this column $C_{n-1}$. By Lemma 5.14, each half of $J$ (in $C_n$) is $\frac{1}{2}$-full of $B_1$, and the same for $I$ and $A_1$. This means we are able to find $I$ and $J$ each $\frac{1}{4}$-full of $A_1$ and $B_1$ and $I$ above $J$ (i.e., $T^\ell(I) = J$ for some $\ell$). Let $A = A_1 \cap I$ and $B = B_1 \cap J$. We have

$$
\mu(T^\ell(A_1) \cap B_1) \geq \mu(T^\ell(A) \cap B) \\
\geq \mu(T^\ell(I) \cap J) - \mu(I \setminus A) - \mu(J \setminus B) \quad \text{by set theory} \\
> \mu(J) - \frac{1}{2} \mu(I) - \frac{1}{2} \mu(J) = 0.
$$

Pick any set $E$ and its complement $E^c$ for $A$ and $B$. Then $\mu(T^\ell(E) \cap E^c) > 0$, a contradiction. So $E$ is cannot have positive measure, i.e., $\mu(E)$ is either 0 or 1. □

Definition 5.16. Let $n > \ell$. Suppose $I$ is an interval in $C_\ell$. Then there exist intervals $I_k$ in $C_n$ such that $I = \bigcup I_k$. We say that the copies of $I$ in $C_n$ are the intervals $I_k$. (Each $I_k$ is a copy of $I$.)

We also define copies of columns. For $n > \ell$, a $C_\ell$-copy in $C_n$ is a group of consecutive intervals $J_0, \ldots, J_{n-1}$ in $C_n$ such that $J_k$ is a copy of $I_k$ for each $k$.

For this dyadic odometer example, any level $I$ in $C_\ell$ has two copies in $C_{\ell+1}$, each of length $\frac{1}{2} |I|$, and has 4 copies in $C_{\ell+2}$, each of length $\frac{1}{4} |I|$. Each $C_{n+1}$ has exactly two copies of $C_n$. (See Figure 4.)
Definition 5.17. We say levels $I$ and $J$ in $C_n$ are $|i - j|$ apart if $J$ is the $j$th level in $C_n$ and $I$ is the $i$th level of $C_n$.

Proposition 5.18. $T$ is not weakly mixing.

Proof. Since the dyadic odometer is finite-measurable, to show $T$ is not weakly mixing we only need to show it is not doubly ergodic (by Theorem 3.15). To test doubly ergodicity, simply pick $A$ and $B$ to be the top and bottom levels of $C_n$. Then $A$ is $h_n - 1$ levels above $B$. The copies of $A$ are $h_n$ levels apart, and the copies of $B$ are also $h_n$ levels apart. So if $\mu(T^m(A) \cap B) > 0$, then $m = rh_n + 1$ for some $r \in \mathbb{N}$. If $\mu(T^m(A) \cap A) > 0$, then $m = rh_n$ for some $r \in \mathbb{N}$. But there are no such $r$ and $s$ that satisfy $rh_n = sh_n + 1$, so $\mu(T^m(A) \cap A)$ and $\mu(T^m(A) \cap B)$ cannot be both positive for any $m$. □

Remark 5.19 (rank one transformation). The dyadic odometer and the two maps we construct next are all rank one transformation. However, though it is easy to give examples of rank one, it is hard to define the notion concisely. Seven mainstream definitions are given in section 1 of [4]. In this paper, we use “rank one” to describe transformations constructed with a single cutting and stacking at each step.

Since the dyadic odometer is ergodic but not weakly mixing, we have shown that ergodicity and weakly mixing are not equivalent.

5.2. Chacón’s Transformation. This transformation from R. V. Chacón’s 1969 Paper [3] is an example that is weakly mixing but not mixing. Moreover, we show that it is not lightly mixing.

Construction 5.20 (Chacón’s Transformation). We use the same notations of column, level and height. The construction is basically “cutting the column into three, putting a spacer above the middle column, and stacking from left to right.” We inductively construct the columns $C_0, C_1, C_2, \ldots$.

Base construction: Let $C_0$ denote the column consisting of a single interval $[0, 2/3)$. Then $h_0 = 1$.

Construction of $C_1$: Divide $C_0$ into three disjoint subintervals $[0, 2/9), [2/9, 4/9)$ and $[4/9, 2/3)$. Put a spacer $S_0$ (a new interval) above the middle subinterval and stack from the right on top to the left on bottom. $S_0$ is chosen to abut the current column and be of the same length of the middle subinterval. In this case, the spacer is $[2/9, 4/9)$. The union of all the intervals is $[0, 2/9)$. Notice $C_1$ has four levels (three from dividing $C_0$ and one spacer), so $h_1 = 4$. (See Figure 5a.)

Induction: We obtain $C_{n+1}$ from $C_n$. (This is a generalization of the construction of $C_1$ from $C_0$ in the previous paragraph.) First, note that the column $C_n$ has

![Figure 4](image-url)
\[ \sum_{i=0}^{n} 3^i = \frac{1}{2}(3^{n+1} - 1) \] levels, each of length \(2 \cdot 3^{-n-1}\) and that the union of these levels is \([0, 1 - 3^{-n-1})\). To obtain \(C_{n+1}\) from \(C_n\), we divide each level into three disjoint subintervals of equal length, i.e., of length \(2 \cdot 3^{-n-2}\). We define the spacer \(S_n\) to also have length \(2 \cdot 3^{-n-2}\) and to have left endpoint at \(1 - 3^{-n-1}\). (In other words, the new level is chosen to abut the current interval.) We stack the middle subcolumn on top of the left subcolumn, then we stack the spacer on top of the middle subcolumn, and then we stack the right subcolumn on top of the spacer. (See Figure 5b.) A level \(I\) in \(C_\ell\) has three copies in \(C_{\ell+1}\), each of length \(\frac{1}{3} |I|\). Notice that \(C_{n+1}\) has height \(h_{n+1} = 3h_n + 1\).

The lengths of spacers added is a geometric sequence with total length
\[
\frac{2}{3} + \frac{2}{27} + \ldots = \frac{1}{3}
\]

Let \(T_{C_n}\) be defined on the column \(C_n\) as in Definition 5.3. Taking \(n\) to infinity gives a measure-preserving transformation on \([0, 1)\). The resulting transformation \(T\) agrees to each transformation \(T_n\), as in the case of dyadic odometer. This new map is called the **canonical Chacón’s Transformation**.

![Figure 5. the canonical Chacón’s Transformation](image)

We observe that each level in \(C_n\) has 3 copies in \(C_{n+1}\). Pick a level, say level \(I_m\), of \(C_n\). We can see from Figure 6 that:

1. \(T^{h_n}\) maps the leftmost copy of \(I_m\) back to level \(I_m\) itself. (Recall Definition 5.16.)
2. \(T^{h_n}\) maps the middle copy of \(I_m\) level to \(T^{-1}(I_m)\) for \(m > 0\). For \(m = 0\), it simply maps \(I_0\) to the spacer \(S_n\) and we know \(S_n \subset T^{-1}(I_0)\).

The next lemma is based on this observation.
\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig6.png}
\caption{$T^{h_n}(I)$ intersects with itself and the level below it for at least $\frac{1}{3}\mu(I)$}
\end{figure}

**Lemma 5.21.** Let $n > 0$. Then:

1. For all $k \geq n$, $\mu(T^{h_k}(I) \cap I) \geq \frac{1}{3}\mu(I)$.
2. For each $\ell \geq 0$, there exists an integer $H = H(n, \ell)$ such that if $I$ and $J$ are at most $\ell \geq 0$ apart, with $I$ above $J$, then $\mu(T^H(I) \cap J) \geq (\frac{1}{3})^\ell \mu(J)$.
3. If $I$ is the top level of $C_n$, then $T^{h_n}I \subset I \cup T^{-1}I$.

**Proof.** (1): First consider $k = n$. For any level $K$ in $C_n$, name the three copies of $K$ in $C_{n+1}$ by $K[0], K[1]$ and $K[2]$, from left to right respectively. We have

$$T^{h_n}(K[0]) = K[1] \text{ and } T^{h_n}(K[1]) = (T^{-1}(K))[2].$$

Thus, $T^{h_n}(I)$ intersects both $I$ and $T^{-1}(I)$ in at least $\frac{1}{3}$ of the measure of $I$. This means

$$\mu(T^{h_n}(I) \cap I) \geq \frac{1}{3}\mu(I)$$

(5.23)

$$\mu(T^{h_n}(I) \cap T^{-1}(I)) \geq \frac{1}{3}\mu(I).$$

(5.24)

For $k \geq n$, each $I$ in $C_n$ has $3^{k-n}$ copies in $C_k$. Applying (5.23) and (5.24) to each copy $I'$ of $I$, we get $\mu(T^{h_n}(I') \cap I') \geq \frac{1}{3}\mu(I')$. Then we add the copies together to get (1).

(2): We claim that we can take $H(n, \ell) = \sum_{i=0}^{\ell-1} h_{n+i}$. (In particular, $H(n, 0) = 0$.) We prove this result via induction on $\ell$. Fix $n$ and let $H_\ell = H(n, \ell)$

The base case $\ell = 0$ is trivial. For the induction step, suppose the result works for $\ell \leq k$. This means $T^{H_k}$ intersects with $T^{-i}(I)$ in measure at least $\left(\frac{1}{3}\right)^k$ times for all $0 \leq i \leq k$. By (5.22), we know that if $K$ contains a full level in $C_n$, then $T^{h_n}$ contains two full levels in $C_{n+1}$. Our choice of $H_k$ makes it so that $T^{H_k}$ has two full levels in $C_{n+m}$ for all $m > 0$. Then we know $T^{H_k+h_{n+k}}(I)$ intersects with $T^{-i}(I)$ in measure at least $\left(\frac{1}{3}\right)^{k+1}$ for all $0 \leq i \leq k$ by (5.23). The case $i = k + 1$ holds by (5.24). So by induction, $H(n, \ell) = \sum_{i=0}^{\ell-1} h_{n+i}$ for all $\ell$.

(3): Let $I$ be the top level of $C_n$. We already know $T^{h_n}(I[0]) = I[1] \subset I$ and $T^{h_n}(I[1]) = T^{-1}(I[2]) \subset T^{-1}I$. Now consider $T^{h_n}(I[2])$. Observe that $T^{h_n}(I[20]) = I[11] \subset I$ and $T^{h_n}(I[21]) = T^{-1}(I[12]) \subset T^{-1}I$. Induction shows that $T^{h_n}(I[2\ldots20]) \subset I$ and $T^{h_n}(I[2\ldots21]) \subset T^{-1}I$. Since $I[2\ldots2]$ converges to the single point $\{1\}$ and $T$ is defined on $[0,1)$, the induction shows that (3) holds for all $n$. \hfill $\Box$

**Theorem 5.25.** The canonical Chacón’s transformation is a measure-preserving transformation on a probability Lebesgue space that has continuous spectrum.
By triangle inequalities, \( |f(x) - c| < \epsilon \) has positive measure. (If not, we would contradict countable additivity.) Now by Lemma 5.13, there exists a level \( I \) in some column \( C_n \) such that \( \mu(I \cap A) > \frac{2}{3} \mu(I) \).

(5.22) gives \( T^{h_n}(I[0]) = I[1] \) and \( T^{h_n+1}(I[1]) = I[2] \). Since \( I \) is \( \frac{2}{3} \)-full of \( A \) and \( T \) is measure-preserving, there must be a point \( x \in A \cap I \) such that \( T^{h_n}(x) \in A \cap I \) and \( T^{h_n+1}(T^{h_n}(x)) \in A \cap I \). This gives

\[
|f(x) - c| < \epsilon, \quad |\lambda^h f(x) - c| < \epsilon, \quad |\lambda^{2h_n+1} f(x) - c| < \epsilon.
\]

By triangle inequalities, \( |f(x)||\lambda^{h_n} - 1| < 2\epsilon \) and \( |f(x)||\lambda^{2h_n+1} - 1| < 2\epsilon \). Since \( |f| = 1 \), we get \( \lambda = 1 \). Thus, \( T \) has continuous spectrum. \( \square \)

Remark 5.26. The proof above is as presented in Chacón’s original paper. Another way to show the weakly mixing property, in the next theorem, follows the same track of Lemma 5.14.

Theorem 5.27. The canonical Chacón’s transformation \( T \) is not doubly ergodic.

Proof. Pick any two sets \( A_1 \) and \( B_1 \) of positive measure. We are able to choose levels \( I_1 \) and \( J_1 \) in some column \( C_n \) that are \( \frac{2}{3} \)-full of \( A_1 \) and \( B_1 \) respectively. We can choose \( I_1 \) and \( J_1 \) so that \( I_1 \) is above \( J_1 \) and they are \( \ell \) apart with \( 0 \leq \ell \leq h_n \).

Let \( \delta = \left( \frac{1}{3} \right)^\ell \). Following the same approximation method in Theorem 5.15, we are able to pick \( I \) and \( J \) subintervals of \( I_1 \) and \( J_1 \) that are \( (1 - \frac{\delta}{3}) \)-full of \( I_1 \) and \( J_1 \) respectively. Let \( A = A_1 \cap I \), \( B = B_1 \cap J \).

Let \( H = \sum_{i=1}^{\ell+1} h_{n+i} \). Lemma 5.21(2) gives \( \mu(T^H(I) \cap J) \geq \left( \frac{1}{3} \right)^\ell \mu(J) \) and \( \mu(T^H(I) \cap I) \geq \left( \frac{1}{3} \right)^\ell \mu(I) \).

Then

\[
\mu(T^H(A) \cap B) \geq \mu(T^H(I) \cap J) - \mu(I \setminus A) - \mu(J \setminus B)
\]

\[
\geq \delta \mu(J) - \frac{\delta}{3} \mu(I) - \frac{\delta}{3} \mu(J) > 0
\]

and

\[
\mu(T^H(A) \cap A) \geq \mu(T^H(I) \cap I) - \mu(I \setminus A) - \mu(I \setminus A) > 0.
\]

\( \square \)

Proposition 5.28. The canonical Chacón’s transformation \( T \) is weakly mixing.

Proof. It follows directly from Theorem 3.15. \( \square \)

Theorem 5.29. The canonical Chacón’s transformation is not mixing.

Proof. Let \( n > 0 \) be such that if \( I \) is a level in \( C_n \), then \( \mu(I) < \frac{1}{3} \). Then by Lemma 5.21, for all \( k \geq n \),

\[
\mu(T^{h_k}(I) \cap I) \geq \frac{1}{3} \mu(I) > \mu(I) \mu(I) + \frac{1}{3} \mu(I) > \mu(I) \mu(I) + \frac{1}{27}.
\]

Thus, \( T \) is not mixing. \( \square \)

Theorem 5.30. The canonical Chacón’s transformation is not locally mixing.
Lemma 5.34. For a spacer $I$ denote the left sublevel of $C$ as in Definition 5.2. Similar to Definition 5.3, denote the right sublevel of $I$ as in Figure 8. For example, $I_{n,i}$ explicitly, $T$ is measure-preserving as in Remark 5.9. We will call this map the FK map.

Construction 5.31. This construction is basically “cutting the column into two, adding a spacer on the right, and stacking from left to right.” We inductively construct the columns $C_0, C_1, C_2, \ldots$

Base construction: For convenience, let $C_0$ be empty. Then $h_0 = 0$.

Construction of $C_1$: Stack spacer $s_1 = [0, \frac{1}{2})$. So $C_1 = ([0, \frac{1}{2})$ and $h_1 = 1$.

Construction of $C_2$: Cut $C_1$ into two subcolumns and stack $s_2 = [\frac{1}{2}, \frac{3}{4})$ above $C_{1,1}$. So $C_2 = ([0, \frac{1}{2}), ([\frac{1}{2}, \frac{1}{2}), ([\frac{1}{2}, \frac{3}{4})$ and $h_2 = 3$.

Induction: To obtain $C_{n+1}$ from $C_n$, divide $C_n$ into two subcolumns, $C_{n,0}$ and $C_{n,1}$, representing the left half and the right half respectively. Let $s_{n+1} = [1 - \frac{4}{2^n}, 1 - \frac{1}{2^{n+1}}]$. We stack $C_{n,1}$ above $C_{n,0}$ and $s_{n+1}$ above $C_{n,1}$. Notice that $h_{n+1} = 2h_n + 1$.

Let $T_{C_n}$ be defined on the column $C_n$ as in Definition 5.3. Taking $n$ to infinity gives a measure-preserving transformation on $[0, 1)$. The resulting transformation $T$ agrees with each transformation $T_n$. Call it the FK map.

Remark 5.32. It is easy to check that $T$ is invertible mod $\mu$ on $[0, 1)$. Also, $T$ is measure-preserving as in Remark 5.9.

![Figure 7. construction of $C_2$ and $C_3$](image)

Notation 5.33. Similar to Definition 5.2, denote the left sublevel of $I_n$ as $I_n[0]$ and the right of $I_n$ as $I_n[1]$. For $1 \leq i \leq h_n$, let $I_{n,i}$ denote the $i$th level in column $C_n$, starting from the top (as in Figure 8). For example, $I_{n,i}[0]$ is the left sublevel of $i$th level in $C_n$, $I_{n,i}[0] = I_{n+1,h_n+i+1}$ and $I_{n,i}[1] = I_{n+1,i+1}$.

A lemma similar to Lemma 5.11 can be proved:

Lemma 5.34. For a spacer $S_n$, $T^{h_n}(S_n) = S_{n+h_n} \cup (\bigcup_{i=1}^{h_n} (T^{h_n}(S_n) \cap I_{n,i}))$

Proof. By the construction of $T_{C_{n+1}}$, we get $T^{h_n}(S_n[0]) = S_{n+1}[1] = I_{n+1}[1]$. Similarly, $T_{C_{n+2}}$ gives us $T^{h_n}(S_n[1]) = I_{n+2}[0]$. Induction shows us $T^{h_n}(S_n[1]) = I_{n,k+1}[0, \ldots, 0]1$ for $0 \leq k \leq h_n - 1$. The union of all such intervals is $\bigcup_{i=1}^{h_n} (T^{h_n}(S_n) \cap...
$I_{n,i}$, as shown in the bold lines in Figure 8. Also notice $S_n[1...1]$, the rightmost sublevel of $S_n$, is sent to $S_{n+h_n}$.

This lemma shows that each spacer $S_n$ is sent to the union of a new spacer $S_{n+h_n}$ and $h_n$ intervals, one on each level of $C_n$, with length decreasing by a factor of $\frac{1}{2}$.

Observe that $\mu(T^{h_n}(S_n) \cap I_{n,i}) = \frac{\mu(S_n)}{2^i}$. Call the configuration of these intervals a crescent.

\[
\begin{array}{c}
\cdots \\
\vdots \\
S_{n+3} \\
S_{n+2} \\
S_{n+1} \\
S_n = I_{n,1} \\
C_n \\
\vdots \\
I_{n,h_n} \\
\end{array}
\]

**Figure 8.** $T^{h_n}(S_n)$ is the union of the spacer $S_{n+h_n}$ with the bolded lines with decreasing lengths. Call the configuration of the latter a crescent.

**Remark 5.35.** Recall the notion of copy in Definition 5.16. For each level in $C_n$, there are 2 copies in $C_{n+1}$ and $C_{n+1}$ has 2 $C_n$-copies. For $m \geq k$, $C_m$ appears in $C_k$ as $2^{m-k}$ disjoint $C_k$-copies.

**Notation 5.36.** A $C_k$-copy in $C_n$ is **preceded by** $u$ spacers if there are $u$ spacers between this copy and the closest copy below it. For example, if $n = k + 1$, one of the two $C_k$-copies is preceded by $u = 0$ spacers. If $n = k + 2$, two of the four $C_k$-copies are preceded by $u = 0$ spacers, and one is preceded by $u = 1$ spacer.
Remark 5.37. Induction shows that of the \(2^{n-k}\) \(C_k\)-copies in \(C_n\), \(2^{n-k-u-1}\) copies are preceded by \(u\) spacers, for \(0 \leq u < n-k\). Thus, the fraction of \(C_k\)-copies in \(C_n\) that is preceded by at most \(u\) spacers is \(\frac{1}{2} + \frac{1}{4} + \ldots + (\frac{1}{2})^{u+1} = 1 - (\frac{1}{2})^{n+1}\).

Remarks 5.38. Consider any level \(I_{n,i}\) in \(C_n\), similar to Lemma 5.34. The copies of \(T^n(I_{n,i})\) form a crescent in \(C_n\) from \(I_{n,i}\) to bottom, and the remainder is in \(S_{n+h_n-i+1}\).

Lemma 5.39. If \(I\) and \(J\) are both levels in \(C_n\) and \(J\) is \(r-1\) levels below \(I\), then
\[
T^n(I) \cap J \subset J[0...0]_r.
\]

Proof. Similar to Lemma 5.34, this proof is also done by induction. \(\square\)

This lemma shows that the intersection \(T^n(I) \cap J\) is always subset of the leftmost subinterval of \(J\) with length \(\frac{\mu(I)}{2^r}\).

Theorem 5.41. \(T\) is not mixing.

This proof is similar to Proposition 5.18 in that we keep track of the copies of the top level of some \(C_n\). The difference is, however, that in the dyadic odometer case, the distance between copies of top levels and copies of bottom levels is always \(rh_n + 1\) for some \(r \in \mathbb{N}\). Here, because of the existence of spacers, the copies of the top level “mixes” into the levels below in the form of a crescent. (For this reason, FK’s map is doubly ergodic, as opposed to dyadic odometer).

Proof. Fix \(k\) so large that \(\frac{h_n}{2^k} < \frac{1}{4}\). Let \(A\) be the top level of \(C_k\) and \(B\) be the bottom of \(C_k\). Then \(B\) is \(h_k - 1\) levels below \(A\). For \(n > k\), by the observation in Remark 5.35, \(A\) appears in \(2^{n-k}\) copies of \(C_k\) in \(C_n\). Name them \(I_j\) for \(1 \leq j \leq 2^{n-k}\). We have \(I_j \subset A\) for all \(j\). Now we do the same for \(B\): there exist \(2^{n-k}\) copies \(J_j\) for \(1 \leq j \leq 2^{n-k}\). We have \(J_j \subset B\) for all \(j\). For each \(j\), the level \(J_j\) is \(h_k - 1\) levels below \(I_j\). For those \(I_p\) that are below \(J_j\), \(T^n(I_p) \cap J_j = \emptyset\) by the same argument as in Proposition 5.18. For those \(I_p\) that are above \(J_j\), we have \(T^n(I_p) \cap J_j \subset J_j[0...0]_h\).

We know the leftmost subinterval of \(J_i\) has measure less than \(\frac{\mu(I_i)}{2^{k+1}}\), by Lemma 5.39.
This is a proof by computation. Let $m$ be a measure of the union of the levels in the column. Then $\lim_{n \to \infty} \mu(T^m(A) \cap B) \leq 2^{-h_k} \mu(B)$.

We know $\mu(A) = \mu(S_k) = \frac{1}{2^{kn}}$ and $h_k = 2^k - 1$, so $2h_k \mu(A) < 1$. Then it is easy to see that $\mu(T^m(A) \cap B) < 2^{-h_k}(2h_k \mu(A)) \mu(B) < \frac{1}{2} \mu(A) \mu(B)$. □

**Remark 5.42.** A transformation is called **partially mixing** if $\liminf \mu(T^m(A) \cap B) \geq \alpha \mu(A) \mu(B)$ for some positive $\alpha$. It is easy to check that mixing implies partially mixing. Instead of requiring $\frac{h_k}{2^{kn}} < \frac{1}{4}$, we can require $\frac{h_k}{2^{kn}} < \frac{\alpha}{4}$ in the proof above and conclude a stronger result that $\liminf \mu(T^m(A) \cap B) < \frac{1}{2} \alpha \mu(A) \mu(B)$ for all positive $\alpha$. So $T$ is not partially mixing.

**Theorem 5.43.** $T$ is lightly mixing.

**Remark 5.44.** The following proof is long, but the idea is simple. We want to show $\mu(T^m(A) \cap A) > 0$ for large $m$. Fix a set $A$, say, the top level of $C_k$. For any large $m$, choose $n$ such that $h_n \leq m < h_{n+1}$.

We consider two sets:

1. the union of all “good” copies of $C_k$ in $C_n$ that are not far from each other.
2. the union of all crescents created by applying $T$ $m$ times on $A_i$, the copies of $A$ in $C_n$.

Our choice of $k$ would make the two sets intersect in a set of positive measure, and our choice of $n$ would make most of $A_i$ have the property that $T^m(A_i)$ is a crescent or a pair of crescents for all $m$. So there would have to be a copy of $A$ in $C_n$ whose $m$th image, $T^m(A_i)$, intersects with the top level of $C_k$-copy, for all $m$.

**Proof.** This is a proof by computation. Let $A$ be a set of positive measure. By Lemma 4.2, we need to find an integer $M$ such that

$$\mu(T^m(A) \cap A) > 0 \text{ for all } m \geq M. \quad (5.45)$$

We know $T$ is defined on $[0,1)$. We define the measure of a column as the measure of the union of the levels in the column. Then $\lim_{n \to \infty} \mu(C_n) = 1$.

Choose $k$ large enough so that $\mu(C_k) > 0.9$ and some level $J$ in $C_k$ satisfies $\mu(I \cap A) > 0.99 \mu(I)$ (i.e., $I$ is 0.99-full of $A$). Without loss of generality, apply $T$ on $A$ several times to send $A$ to the top level of $C_k$ and then shrink $A$. So now we assume $A \subset S_k$ where $S_k$ is the top level of $C_k$, and also $\frac{\mu(A)}{\mu(S_k)} > 0.99$.

For $n \geq k$, the level $S_k$ appears as $2^{n-k} C_n$-copies $I_i$ for $1 \leq i \leq 2^{n-k}$ by Remark 5.35. Also observe that each $I_i$ is the top level of a $C_k$-copy in $C_n$.

Fix $1 \geq \alpha > 0$. We say a level $n$ in $C_n$ is **accurate** if it is $\alpha$-full of $A$. Say a $C_k$-copy in $C_n$ is **good** if its top level is accurate and the top level of the $C_k$-copy below it is accurate, and is preceded by at most $u = 10$ spacers. Now fix $N$ sufficiently large so that for all $n \geq N$,

$$\#\{\text{accurate } C_n\text{-levels } I\} > \left(\frac{\mu(A)}{\mu(S_k)} - 0.01\right) \cdot 2^{n-k} > 0.98 \cdot 2^{n-k}. \quad (5.46)$$

---

\(^1\)This manipulation is without losing of generality because if (5.45) holds on a subset of $A$, it also holds on $A$. 
The second inequality holds for sufficiently large $n$ because $\frac{\mu(A)}{\mu(S_k)}$ increases as $n$ increases.

By this inequality we know the fraction of $C_k$-copies whose top level is not accurate is dominated by $0.02$. Also, Remark 5.37 shows that the proportion of $C_k$ copies preceded by more than 10 spacers is $1 - \left(\frac{1}{2}\right)^{10+1}$. Let $G_n$ be the union of all good $C_k$-copies in $C_n$. We then have

\begin{equation}
\mu(G_n) > (1 - (\frac{1}{2})^{10+1} - 0.02 \cdot 2)\mu(C_k) > 0.
\end{equation}

Now we do the computation. Set the $M$ in (5.45) to be $h_N$. For an $m \geq M$, let $n \geq N$ denote the value for which

$$h_n \leq m < h_{n+1} = 2h_n + 1.$$  

Though the choice of $n$ depends on $m$, it suffices to show $\mu(T^m A \cap A) > 0$ for all $m$ in $[h_n, h_{n+1})$.

Let $S_{n,k} = \bigcup_{i=1}^{h_k} S_{n+i}$, as in Figure 8. Suppose $I$ is an accurate interval in $C_n$ whose image $T^m I$ is disjoint from $S_{n,k}$. Then $T^m I$ appears in $C_n$ as either a single or double crescents, as in Figure 11.

![Figure 11. $T^m I$ is either single or a pair of crescents](image)

Recall Remarks 5.38. We check the configuration of $T^m I$ for all values of $m$ in $[h_n, h_{n+1})$. When $m = h_n$, $T^m I$ is a single crescent as in Figure 10. When $m$ increases, $T^m I$ moves up. After the top level of $T^m I$ intersects with $S_{n,k}$, say, when $m = p$, this $I$ becomes invalid. But after $m = p + h_k$, $I$ becomes disjoint with $S_{n,k}$ again. Now the set $T^m I$ becomes a pair of crescents.
Figure 12. the pattern of $T^m I$ as $m$ increases.

Notice that the accurate levels are at least $h_k$ from each other and $S_{n,k}$ has $h_k$ levels. So for any $m$, there can only be at most one accurate level that is invalid (as in Figure 12c).

Let $I^*$ be the union of $h_k$ many $C_n$-levels that contain the top $h_k$ levels of $T^m I$ as in Figure 11. Thus $I^*$ has the measure of a $C_k$-copy in $C_n$, which gives $\mu(I^*) = \frac{\mu(C_k)}{2^{n-k}}$. Let $G^*_n$ be the union of all $I^*$. There is at most one $I$ whose image fails to be disjoint from $S_{n,k}$. Thus,

$$\mu(G^*_n) > (\#\text{[accurate } C_n\text{-levels]} - 1)\frac{\mu(C_k)}{2^{n-k}} > 0.8 \text{ by (5.46).}$$

This inequality, together with (5.47), shows $\mu(G_n \cap G^*_n) > 0$. Since $\mu(G_n \cap G^*_n) > 0$, there is an $I^*$ and a good $C_k$-copy, call it $D$, such that the two intersect in a set of positive measure. There are two cases: either (i) the top level of $I^*$ is below that of $D$, or (ii) the top level of $I^*$ is the same or above that of $D$.

For case (i), let $J$ denote the top level of the $C_k$-copy below $D$. For case (ii), let $J$ denote the top level of $D$. In both cases, since $I^*$ has height $h_k$ and we required $D$ to be preceded by at most 10 spacers, we know $J$ is fewer than $h_k + 10$ levels away from the top level of $I^*$. Since $I$ and $J$ are both accurate,

$$\mu(T^m(A \cap A)) \geq \mu(T^m(A \cap I) \cap (A \cap J))$$
$$\geq \mu(T^m I \cap J) - (1 - \alpha)\mu(I) - (1 - \alpha)\mu(J)$$
$$\geq \mu(J) \cdot (\frac{1}{2})^{h_k + 10} - 2(1 - \alpha)$$

This shows such $m$ satisfies (5.45) \(\square\)

6. Discussion

We have discussed three different constructions of rank one maps. In fact, we see that as spacers take more importance in the construction of the columns, the map tends to mix “better.”
First, we notice the hierarchy

1. weakly mixing
2. mildly mixing
3. lightly mixing
4. partially mixing
5. mixing.

There are rank one constructions of maps that satisfy \( i \) but not \( i+1 \) for \( i = 1, 2, 3, 4 \).

Chacón’s map is actually (2) but not (3). In fact, there are even weaker maps than Chacón’s map, that is (2) but not (1), constructed by cutting each column into many subcolumns and stacking together without adding spacers. The concept of mildly mixing involves the definition of rigidity, thus is not mentioned in this paper.

The main idea of Chacón’s map is that (i) \( T^k \) maps a fixed proportion of any level in \( C_n \) both to itself and to one level lower; (ii) the rightmost subinterval of the top-right level of \( C_n \) stays on the top-right after applying \( T^k \) for all \( m > n \) because no spacer is added above the top-right. So there is no crescent created. In fact, it is easy to show that there are maps similar to Chacón’s transformation that are only weakly mixing; for example, if we cut each column into 5 instead of 3 and put one spacer in the middle.

A crescent is created when a spacer is added above the top-right. Maps similar to FK’s map would be expected if we cut each column into multiple subcolumns and add spacers on the top-right.

FK’s map is (3) but not (4). We can construct even stronger rank one maps, but with more spacers. It is even possible to construct a rank one mixing map. In fact, Ornstein proved in [7] that a rank one map with a very large amount of spacers allocated randomly can be mixing. Apart from this stochastic construction, there is also the Smorodinsky Conjecture that states the existence of such a mixing map. M. Smorodinsky conjectured that by adding staircases whose heights increase consecutively by one, the resulting transformation (see Figure 13.) is mixing if \( \lim_{n \to \infty} \frac{r_n^2}{h_n} = 0 \). This conjecture is proved by Terrence Adams with a rank one example called the “staircases map,” in [1].

\[
\begin{array}{c}
\text{Figure 13. the staircase map}
\end{array}
\]

Acknowledgments. I would like to thank my mentor, Kathryn Lindsey, for introducing me to ergodic theory, and for leading my exploration of this material. I would also thank my project mentor Alan Chang for guidance and support throughout the program. Without his suggestions this paper would not have been possible.
In addition, I would like to thank Professor Peter May for organizing this wonderful REU program.

REFERENCES


