MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

On Cutting and Stacking Transformations

verfasst von / submitted by

Kathrin Peticzka, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2020 / Vienna, 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 821
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Mathematik
Betreut von / Supervisor:	Univ.-Prof. Henk Bruin, PhD

Acknowledgements

I would like to thank my thesis adviser Univ.-Prof. Henk Bruin, PhD of the Faculty of Mathematics at the University of Vienna. He supported me in all possible ways and still gave me enough freedom to develop my thesis on my own. He kindly answered all of my questions, doesn't matter how silly they may have been.

I would like to thank my family, especially my parents, without whom I would not have been able to study in the way I did. They provided me with financial and moral support throughout my years of study and through the process of researching and writing this thesis.

This accomplishment would not have been possible without them. Thank you.

Abstract

This paper discusses the use of the Cutting-and-Stacking procedure of certain rank one transformations. Throughout the chapters we establish some important properties of such transformations, starting with basic ergodic properties as invariance and (uniquely) ergodicity and going further with more advanced ones like mixing properties. This whole Theory was kicked-off by R. V. Chacón in 1969 and S. Kakutani in 1972. Both published independently an explicit construction of rank one transformations which, in case of Chacón, is in fact weakly mixing. For the construction of a strongly mixing rank one transformation, T. M. Adams proved in 1998 an assumtion by M. Smorodinsky that every staircase transformation is strongly mixing.

In Chapter 1 we recall some basic ergodic theory and already introduce the two main Cutting-and-Stacking constructions of S. Kakutani [16] and R. V. Chacón [10]. We also study some first mixing properties.

In Chapter 2 we establish Bratteli diagrams [9], which are a useful tool in the field of Cantor minimal systems. It was shown in [15] that every minimal homeomorphism is topologically conjugate to a Vershik map of an ordered Bratteli diagram. The main theorem (Theorem 2.3.5) in this Chapter will show that result.

In Chapter 3 we discuss ergodic invariant measures on Bratteli diagrams. We will give an answer to the question of the number of ergodic invariant measures for a given Bratteli diagram. Here we follow mainly a quite recent Paper of J. Kwiatkowski, S. Bezuglyi and O. Karpel [4]. We define the class of exact finite rank Bratteli diagrams which implies finite rank Bratteli diagrams since it adds the requirement of having only towers of positive measure to the definition of finite rank Bratteli diagrams. It will turn out that every exact finite rank Bratteli diagram has a unique ergodic measure.

In Chapter 4 we introduce a new type of rank one transformations, the so called Staircase transformations. T. M. Adams had shown in [1] and [2] that this type of transformation is in fact strongly mixing.

Zusammenfassung

Diese Arbeit diskutiert die Verwendung eines sogenannten "Cutting-and-Stacking"-Prozess von bestimmten Abbildungen, die Rang-1 besitzen. Im Laufe der Kapitel, behandeln wir die wichtigsten ergodischen Eigenschaften solcher Abbildungen und gehen speziell auf die Konstruktion dieser ein. Den Beginn machten die zwei Mathematiker R. V. Chacón, 1969 und S. Kakutani, 1972, als sie unabhängig von einander die explizite Darstellung von Rang-1 Transformationen entwickelten. Ziel war es, eigenschaften wie Mischung an solchen Abbildungen nachzuweisen. Chacón beweiste mit seiner Konstruktion die Existenz von schwach mischenden Rang-1 Abbildungen und etwas später in 1998 gelang es T. M. Adams zu beweisen, dass sogenannte (Rang-1-) Stufentransformationen stark mischend sind.

In Kapitel 1, wiederholen wir einige wichtige Konzepte der Ergodentheorie und liefern die ersten beiden Beispiele des "Cutting-and-Stacking"-Prozesses.

In Kapitel 2 definieren wir Bratteli-Diagramme. Es wird sich herausstellen, dass jeder minimale Homeomorphismus, topologisch gesehen, isomorph zu sogenannten Vershik-Abbildungen ist, die auf Bratteli-Diagrammen leben.

Kapitel 3 handelt von (eindeutigen) ergodischen, invarianten Maßen auf Bratteli-Diagrammen und gibt Aufschluss über die Anzahl an ergodischen, invarianten Maßen, die ein bestimmtes Bratteli-Diagram besitzt.

In Kapitel 4 führen wir schlussendlich die Stufentransformation ein, und beweisen nach dem Text von T. M. Adams, dass diese stark mischend sind.

Contents

1 Introduction

1.1 Prerequisites in Ergodic Theory

In this section, we are mainly following lecture notes on the course "Ergodic Theory I and II" of Professor Zweimueller, held in 2018/2019 at the University of Vienna [24, 25]. We will mention the important definitions and theorems, we will need in order to prove later statements.

Definition 1.1.1. *Null- and measure-preserving transformation.*
Let $(X, \mathcal{A}, T)$ be a measurable dynamical system (short: measurable DS), where X is a measurable space, $\mathcal{A}$ a σ-Algebra and T a measurable map such that $T : (X, \mathcal{A}, \lambda) \to (X, \mathcal{A}, \lambda)$ with λ some "natural" reference measure. Then

(i) T is **null-preserving** if $\lambda \circ T^{-1} \ll \lambda$, i.e. $\forall A \in \mathcal{A} : \lambda(A) = 0 \implies \lambda(T^{-1}A) = 0$

(ii) T is **measure-preserving** if $\lambda \circ T^{-1} = \lambda$

Remark 1.1.2. (a) Every measure-preserving transformation is also null- preserving.

(b) Any null-preserving transformation induces two operators:

(i) The **Pullback (or Koopman) operator** $\mathbb{U}_T f$ is defined such that T pulls back any function f in $\mathcal{L}_\infty$, $f : f \mapsto f \circ T$. This operator is linear, positive and has norm 1.

(ii) The **Push-forward (or Transfer) operator** $\hat{T}$ is defined such that T pushes forward measures or densities; $\hat{T} : \mathcal{L}_1(X, \mathcal{A}, \lambda)$ with $\hat{T}u := \frac{d(u \odot \lambda) \circ T^{-1}}{d\lambda}$, where $u \odot \lambda$ means the density u with respect to the measure λ. This operator is measurable and defined for all $u \geq 0$.

Proposition 1.1.3. *Duality of the operators.*
$\mathbb{U}_T$ is the dual of $\hat{T}$, in other words,

$$\forall u \in L_1(\lambda), \forall f \in L_\infty(\lambda) : \int \underbrace{(f \circ T)}_{\mathbb{U}_T f} u \, d\lambda = \int f \, \hat{T} u \, d\lambda.$$

Proof. For f use the indicator function $f = \mathbb{1}_A$ for a set $A \in \mathcal{A}$ and fix a density $u \in \mathcal{L}_1(\lambda)$. We get

$$\int f \, \hat{T} u \, d\lambda \overset{Def}{=} \int f \, d((u \odot \lambda) \circ T^{-1}) \overset{f=\mathbb{1}_A}{=} u \odot \lambda(T^{-1}A)$$
$$= \int_{T^{-1}A} u \, d\lambda = \int (\mathbb{1}_A \circ T) u \, d\lambda.$$

The last equality holds since $\mathbb{1}_{T^{-1}A} = \mathbb{1}_A \circ T$. □

Definition 1.1.4. *Recurrent sets.*
Let $\mathcal{S} = (X, \mathcal{A}, \mu, T)$ be a measure-preserving dynamical system and let μ be σ-finite. Define $A \in \mathcal{A}$ with $\mu(A) > 0$ to be our reference set. Then
A is **recurrent** for $\mathcal{S}$ if almost every $x \in X$ returns to A and the **first hitting time** of A is defined as $\varphi_A : X \to \{1, 2, \ldots, \infty\}$ with $\varphi_A(x) := \inf\{n \geq 1 : T^n x \in A\}$.

Remark 1.1.5. $\mathcal{S}$ is recurrent $\Leftrightarrow$ $\mathcal{S}$ is conservative i.e. if $W \in \mathcal{A}$ is a wandering set, then $\lambda(W) = 0$.
Note: W is wandering if $\lambda(W \cap \bigcup_{n\geq 1} T^{-n}) = 0$ or equivalently $(T^{-n}W)_{n\geq 0}$ are pairwise disjoint (mod λ).

Theorem 1.1.6. *Poincaré recurrence theorem.*
If $\mathcal{S} = (X, \mathcal{A}, \mu, T)$ a measure-preserving dynamical system and $\mu(X) < \infty$, then $\mathcal{S}$ is recurrent.

Proof. Let $W \in \mathcal{A}$ be wandering. It follows that $(T^{-n})_{n\geq 0}$ are pairwise disjoint (mod μ). Hence

$$\mu\Big(\bigcup_{n=0}^{N} T^{-n}W\Big) = \sum_{n=0}^{N} \mu\big(T^{-n}W\big) \overset{m.p.}{=} N \cdot \mu(W)$$

and

$$\mu(X) \geq \mu\Big(\bigcup_{n=0}^{N} T^{-n}W\Big) \Rightarrow \forall N \geq 1 : \mu(W) \leq \frac{\overbrace{\mu(X)}^{<\infty}}{N} \Rightarrow \mu(W) = 0$$

□

Definition 1.1.7. *Ergodicity.*
Let $\mathcal{S} = (X, \mathcal{A}, \mu, T)$ be a measure-preserving dynamical system. $\mathcal{S}$ is called **ergodic**, if $T^{-1}A = A \Rightarrow 0 \in \{\mu(A), \mu(A^c)\}$. In other words, X cannot be decomposed in any non trivial way.

Proposition 1.1.8. *Ergodicity of constant functions.*
$\mathcal{S}$ is ergodic if and only if $\forall f \in \mathcal{L}_1(\mu) : f \circ T = f \Longrightarrow f =$ constant (mod μ)

Proof. ($\Rightarrow$) Suppose $f : X \to \mathbb{R}$ such that $f \circ T = f$ $\mu-$almost everywhere and assume f is not constant. Hence there exists $a \in \mathbb{R}$ such that $A := f^{-1}((-\infty, a])$ and $A^c := f^{-1}((a, \infty))$ have positive measure. Since T is invariant, it holds that $T^{-1}A = A$ (mod μ) which contradicts the assumption on ergodicity.
($\Leftarrow$) Suppose there exists a set $A \in \mathcal{A}$ such that $\mu(A) > 0$ and $T^{-1}A = A$. Define f to be the indicator function of A. Then $f = \mathbb{1}_A$ is still T-invariant since A is. By assumption, f is constant $\mu-$almost everywhere. For $x \in A^c$ we get $f(x) = 0$, therefore $\mu(A^c) = 0$. □

Theorem 1.1.9. *$\mathcal{L}_2$-Ergodic-Theorem.*
Let $\mathcal{S} = (X, \mathcal{A}, \mu, T)$ be a measure-preserving dynamical system. Then

$$\forall f \in \mathcal{L}_2(\mu) : \mathbb{A}_n(f) := \frac{1}{n} S_n(f) = \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k \to \Pi_{\mathbb{U}}(f) \text{ in } \mathcal{L}_2(\mu),$$

where $\Pi_{\mathbb{U}}(f)$ is the orthogonal projection of f onto the subspace of T-invariant functions in $\mathcal{L}_2(\mu)$.
If additionally, $\mu(X) = 1$, i.e. μ is a probability measure, then it holds that $\Pi_{\mathbb{U}}(f) = \mathbb{E}[f|\mathcal{I}]$, where $\mathcal{I}$ is the smallest σ-algebra, containing all invariant sets $A \in \mathcal{A}$.

The proof is left to the reader. For literature take any book on basic ergodic theory as for example [24] or [22].

Theorem 1.1.10. *Birkhoff-Ergodic-Theorem.*
Let $\mathcal{S} = (X, \mathcal{A}, \mu, T)$ be a measure-preserving dynamical system. Then

$$\forall f \in \mathcal{L}_1(\mu) : \mathbb{A}_n(f) \to \bar{f} \text{ a.e., where } \bar{f} \in \mathcal{L}_1(\mu) \text{ and } \bar{f} = f \circ T \text{ a.e..}$$

If additionally $\mu(X) = 1$, then $\bar{f} = \mathbb{E}[f|\mathcal{I}]$ and if $\mathcal{S}$ is ergodic $\bar{f} = \mathbb{E}[f]$. Summerized we get for measure-preserving, ergodic systems with $\mu(X) = 1$,

$$\forall f \in \mathcal{L}_1(\mu) : \underbrace{\mathbb{A}_n(f)}_{\text{time average}} \to \underbrace{\int f d\mu}_{\text{space average}} \quad \text{a.e..}$$

The proof is again left to the reader. For literature take any book on basic ergodic theory as for example [25] or [22].

An important characterization of ergodicity is the following Propostition.

Proposition 1.1.11. *Ergodicity vs. mixing properties.*
Let $\mathcal{S} = (X, \mathcal{A}, \mu, T)$ be a dynamical system with $\mu(X) = 1$. Then

$$\mathcal{S} \text{ is ergodic} \iff \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(T^{-k}A \cap B) = \mu(A)\mu(B) \quad \forall A, B \in \mathcal{A}.$$

Proof. ($\Rightarrow$) Suppose $\mathcal{S}$ is ergodic then by Proposition 1.1.8 it follows that $\bar{f}$ is constant a.e. and $\bar{f} = \int_X f d\mu$. Now suppose $f = \chi_A$, where A is $\mathcal{A}$-measurable. We get by Theorem 1.1.10,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_A(T^k x)\chi_B(x) = \mu(A)\chi_B(x) \text{ for } \mu \text{ a.e. } x \in X \text{ and for all } B \in \mathcal{A}.$$

Integrating with respect to μ and using dominated convergence, we get

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

($\Leftarrow$) Suppose we have $\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\mu(T^{-k}A\cap B) = \mu(A)\mu(B) \quad \forall A, B \in \mathcal{A}$. Let $T^{-1}(C) = C$ and set $A = B = C$. Consequently

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\mu(T^{-k}C\cap C) = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\mu(C) = \mu(C) = (\mu(C))^2.$$

Hence $\mu(C) \in \{0,1\}$ and $\mathcal{S}$ is ergodic. □

Definition 1.1.12. *Mixing.*
A measure-preserving system $\mathcal{S} = (X, \mathcal{A}, \mu, T)$ with $\mu(X) = 1$ is

(i) **weakly mixing** if $\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}|\mu(A\cap T^{-k}B) - \mu(A)\mu(B)| = 0 \quad \forall A, B \in \mathcal{A}$

(ii) **strongly mixing** if $\lim_{n\to\infty}\mu(A\cap T^{-k}B) = \mu(A)\mu(B)$.

In other words, a system is mixing, if $\{X_0 \in A\}$ and $\{X_n \in B\}$ are asymptotically independent as $n \to \infty$.

The definition of a weakly mixing transformation is a naturally stronger requirement on the transformation than the ergodicity property [23]. By Proposition 1.1.11 it follows that every weakly mixing transformation is also ergodic. It is also worth to mention, that there are more than this types of mixing, such as *mild mixing*, *light mixing* or *partial mixing*.

1.2 Mixing Properties

Concerning Definition 1.1.12, we will now investigate mixing and especially weakly mixing properties. As done in [22], we give a Characterization of weakly mixing transformations in terms of the Cartesian products.

First, we focus on another connection with ergodicity and mixing, namely the Cesàro convergence. If we write for two measurable sets A and B,

$$a_i = a_i(A, B) = \mu(T^{-i}(A)\cap B),$$

then ergodicity can be expressed by

$$\frac{1}{n}\sum_{i=0}^{n-1}a_i(A, B) \stackrel{n\to\infty}{\longrightarrow} \mu(A)\mu(B)$$

and mixing can be expressed by

$$a_i(A,B) \stackrel{i\to\infty}{\longrightarrow} \mu(A)\mu(B).$$

Both expressions using the convergence of sequences, hence the Cesàro convergence

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

and the strong Cesàro convergence

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

respectively. Using this convergence, it is easy to show the following relations for a probability-preserving transformation T:

$$\text{T mixing} \;\Rightarrow\; \text{T weakly mixing} \;\Rightarrow\; \text{T ergodic.} \tag{1.1}$$

Also important is the notion of density-zero sets.

Definition 1.2.1. (1) A set D in $\mathbb{N}$ is if **density-zero** if

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

Conversely, if

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{1}_D(i) > 0,$$

then D is of **positive density**.
(2) A sequence $\{a_i\}$ in $\mathbb{R}$ **converge in density** to a if there exists a density-zero set $D \subset \mathbb{N}$ such that

$$\forall \varepsilon > 0\; \exists N \in \mathbb{Z} \;:\; \forall i > N,\; i \notin D,\; |a_i - a| < \varepsilon.$$

We write $\lim_{i\to\infty,\, i\notin D} a_i = a$.

Lemma 1.2.2. *Let $\{b_i\}$ be a bounded sequence in $\mathbb{R}^+$. Then $\{b_i\}$ converges Cesàro to 0 if and only if $\{b_i\}$ converges in density to 0.*

Proof. The result follows by applying the definition of Cesàro convergence. □

A direct consequence from this, is the next Proposition.

Proposition 1.2.3. *Let T be measure-preserving on the probability space $(X, \mathcal{A}, \mu)$. Then the following are equivalent:*

(a) *T is weakly mixing.*

(b) *Let (A, B) be a pair of measurable sets. Then there exists a density-zero set $D = D(A, B)$ such that*

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

(c) *Let (A, B) be a pair of measurable sets, then*

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

Proof. Applying Lemma 1.2.2 to $b_i = |\mu(T^{-i}(A) \cap B) - \mu(A)\mu(B)|$ gives the statement. □

Definition 1.2.4. $\mathcal{S}$ is said to be a **sufficient semiring** on X if

$$A \in \mathcal{A} \implies \mu(A) = \inf_{A \subset \bigcup\limits_{n=1}^{\infty} S_n} \sum_{n=1}^{\infty} \mu(S_n), \quad S_n \in \mathcal{S}.$$

Having this definition, the next Lemma underline Proposition 1.2.3(b) in a stronger way.

Lemma 1.2.5. *Let $(X, \mathcal{A}, \mu)$ be a probability space. If there exists a countable sufficient semiring $\mathcal{S}$ and a sequence $\{n_k\}$ such that for all elements $I, J \in \mathcal{S}$*

$$\lim_{n\to\infty} \mu(T^{-n_k}(I) \cap J) = \mu(I)\mu(J),$$

then for all $A, B \in \mathcal{A}$,

$$\lim \mu(T^{-n_k}(A) \cap B) = \mu(A)\mu(B). \tag{1.2}$$

Note that if the sequence $\{n_k\}$ satisfy (1.2)*, the sequence is called* ***mixing sequence****.*

The proof is left to the reader. For more details see for example Chapter 6 of C. E. Silva's book, [22].

Now we are able to give the characterization theorem of weakly mixing transformations in terms of the Cartesian products.

Theorem 1.2.6. *Assume T to be a measure-preserving transformation on a probability space $(X, \mathcal{A}, \mu)$. Then the following are equivalent:*

(a) *T is weakly mixing.*

(b) $T \times T$ is weakly mixing.

(c) $T \times T$ is ergodic.

Proof. (a) $\Rightarrow$ (b) Suppose T is weakly mixing and $A, B, C, D \in \mathcal{A}$. There exist density-zero sets D_1 and D_2 such that for all $n \notin D_1 \cup D_2$

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

If we define μ_2 to be the product measure $\mu \times \mu$ on $X \times X$, then for all $n \notin D_1 \cup D_2$ and for all elements of a sufficient semiring, it holds:

$$\lim_{n\to\infty} \mu_2\Big[(T \times T)^{-n}(A \times C) \cap (B \times D)\Big] = \mu_2(A \times C)\mu_2(B \times D) \tag{1.3}$$

From Lemma 1.2.5, it follows that (1.3) holds for all sets in $\mathcal{A}$. Finally by Proposition 1.2.3 we conclude that $T \times T$ is weakly mixing.
(b) $\Rightarrow$ (c) Follows directly by (1.1).
(c) $\Rightarrow$ (a) We assume that $T \times T$ is ergodic and show that Proposition 1.2.3(c) is satisfied. First of all, by expanding the coefficients, we get

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \Big(\mu(T^{-i}(A) \cap B) - \mu(A)\mu(B)\Big)^2 \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \Big(\mu(T^{-i}(A) \cap B)\Big)^2 \\ &-2 \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \Big(\mu(T^{-i}(A) \cap B)\Big)\mu(A)\mu(B) + \Big(\mu(A)\mu(B)\Big)^2. \end{aligned} \tag{1.4}$$

Computing the terms on the right hand side, we get our result: Since we assume $T \times T$ to be ergodic, T is also ergodic (we let this result unproven since it can be easily checked by just applying the basic definitions from the introduction part). Thus by Proposition 1.1.11,

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

For $T \times T$ and the sets $A \times A$ and $B \times B$, the same Proposition gives

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu_2\Big[(T \times T)^{-i}(A \times A) \cap (B \times B)\Big] = \mu_2(A \times A)\mu_2(B \times B)$$

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

Hence the left hand side of (1.4) equals

$$\mu(A)^2\mu(B)^2 - 2\mu(A)\mu(B)\cdot\mu(A)\mu(B) + \mu(A)^2\mu(B)^2 = 0,$$

which is the condition we needed to prove. □

We close this section with more equivalent properties to weakly mixing. Those equivalences are important to characterize weakly mixing transformations. We introduce a new property:

Definition 1.2.7. *Doubly ergodic transformation.*
A measure-preserving transformation is **doubly ergodic**, if $\forall A, B \in \mathcal{A}$ such that $\mu(A) > 0$ and $\mu(B) > 0$ then there exists $n \in \mathbb{N}$ such that $\mu(T^{-n}(A)\cap A) > 0$ and $\mu(T^{-n}(A)\cap B) > 0$.
Note: T is doubly ergodic $\Longrightarrow T$ is recurrent and ergodic.

Remark 1.2.8. A measure-preserving transformation T has a **continuous spectrum** if their induced operators have a trivial point spectrum, equivalently if T is ergodic and $\lambda = 1$ is its only eigenvalue.

Theorem 1.2.9. *Assume T to be an invertible measure-preserving transformation on a Lebesgue probability space $(X, \mathcal{A}, \mu)$. Then the following are equivalent:*

(a) T is weakly mixing.

(b) T has continuous spectrum.

(c) T is doubly ergodic.

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

Proof. (a) ⇒ (b): By assumption T is weakly mixing. Then by Theorem 1.2.6, T and $T\times T$ are ergodic. For f eigenfuction of T and λ its eigenvalue such that $f(T(x)) = \lambda f(x)$, we get $|\lambda| = 1$ by

$$|\lambda|^2\int|f|^2d\mu = \int|\lambda|^2|f|^2d\mu = \int|f\circ T|^2d\mu = \int|f|^2\circ Td\mu = \int|f|^2d\mu$$

and $|f|$ constanct and without loss of generality $|f| = 1$ by

$$|f|\circ T = |f\circ T| = |\lambda f| = |\lambda||f| \overset{|\lambda|}{=} |f|$$

ergodicity of T and normalizing. For ahowing that $\lambda = 1$ is the only eigenvalue, we define

$$g : X\times X \longrightarrow \mathbb{C} \text{ with } g(x, y) = f(x)\bar{f}(y).$$

Then

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

Thus g is an eigenfunction of $T\times T$ and therefore constant a.e., hence f is constant a.e.

(a) $\Rightarrow$ (c):Suppose T is weakly mixing and $A,B \in \mathcal{A}$ with $\mu(A) > 0$ and $\mu(B) > 0$. Then there are two density-zero sets D_1 and D_2 such that by definition of weakly mixing, $\lim_{i\to\infty, i\notin D_1} \mu(T^{-i}(A)\cap B) = \mu(A)\mu(B)$ and $\lim_{i\to\infty, i\notin D_2} \mu(T^{-i}(A)\cap A) = \mu(A)\mu(A)$. Since A and B have positive measure, there exists $i \in \mathbb{N}$ such that $\mu(T^{-i}(A)\cap A) > 0$ and $\mu(T^{-i}(A)\cap B) > 0$. Hence T is doubly ergodic.

(b) $\Rightarrow$ (d): Showing this, needs more functional analysis then we want to present here. Since it is not our aim, we refer to [20, p.67-71] for the complete proof.

(c) $\Rightarrow$ (b): By assumption T is doubly ergodic. This implies that T is ergodic. Again there exists an eigenfunction $f \in L^2$ and an eigenvalue λ such that $|\lambda| = 1$, $|f| = 1$ and $f(T(x)) = \lambda f(x)$ a.e. As before, we need to show that $\lambda = 1$ is the only eigenvalue. Let us assume that there exist different eigenvalues on the unit circle. Parametrization gives

$$\lambda = e^{2\pi i\alpha} \text{ and } f(x) = e^{2\pi i g(x)} \text{ for } \alpha \in [0,1) \text{ and } g: X \to [0,1) \text{ measurable.}$$

Now for $R: [0,1) \to [0,1)$ with $R(t) = t+\alpha$, we have $g\circ T = R\circ g$. We construct a factor of the probability space by defining ν to be the measure on $[0,1)$ such that $\nu(A) = \mu(g^{-1}(A))$. Then g is a factor map from T to R. Since T is ergodic and R a factor of T, R is ergodic. We can now identify two cases for R: On the one hand, if α is rational, ν has to be atomic and concentrated on finitely many points. This implies that R is not doubly ergodic (take A,B two sets containing only one point). On the other hand, if α ir irrational, ν has to be the Lebegue measure. This implies that R is not doubly ergodic (take $A = [0,\frac{1}{8}), B = [\frac{1}{2},\frac{3}{4})$, then $\mu(A) > 0$ and $\mu(B) > 0$ but for every $n \in \mathbb{N} : R^n(A)\cap B \neq 0$ and $R^n(A)\cap A = 0$). Hence our assumption of different eigenvalues is false.

(d) $\Rightarrow$ (a) Apply Theorem 1.2.6 to $T = S$. □

1.3 Two prominent Examples of Cutting and Stacking constructions

The concept of Cutting and Stacking describes a construction of a measure preserving transformation, to guarantee certain properties as mixing properties and (unique) ergodicity. Kakutani and Chacón introduced this method independently, which are now famous examples of constructable transformations which satisfy one or more of those properties.

Cutting and Stacking is based on the Kakutani-Rokhlin-Lemma [3], which states that every aperiodic, measure preserving dynamical system can be written as a tower of arbitrary height plus a small remainder of arbitrary small measure, in the following called spacer.

To begin with, we give two examples of rank one transformations, one construction which is neither weakly nor strongly mixing but invariant and ergodic, the other which is weakly mixing but still not strongly mixing. Rank one transformations describe transformations which are constructed with one stack. More general, in this sense the rank of a transformation will be defined as the number of necessary stacks or towers to construct the transformation. In [23], Yassawi shows that there exists also rank one transformations which are strongly mixing. But for now we start with two examples we actually can construct.

1.3.1 Kakutani, Odometer, Dyadic adding machine

This rank one transformation is named after Shizuo Kakutani, a Japanese- American mathematician, born 1911 in Osaka and died in August 2004. This section is following the Paper of Kakutani [16], published in 1972, as well as a book of C.E. Silva [22].

Example 1.3.1. *Kakutani or Dyadic adding machine.*
Let us look at the unit interval $I = [0, 1)$. Kakutani uses the idea of replication to construct an invariant transformation by cutting and stacking intervals with endpoints of the form $\frac{k}{2^n}$. Because of this reason, this construction is also called the dyadic adding machine or Odometer. However let us get more precise:
To get the first stack S_1 we cut the interval I into two intervals, $I_1 = [0, \frac{1}{2}), I_2 = [\frac{1}{2}, 1)$, of same length. Then take the second interval and stack it above the first (see Figure 1.1).

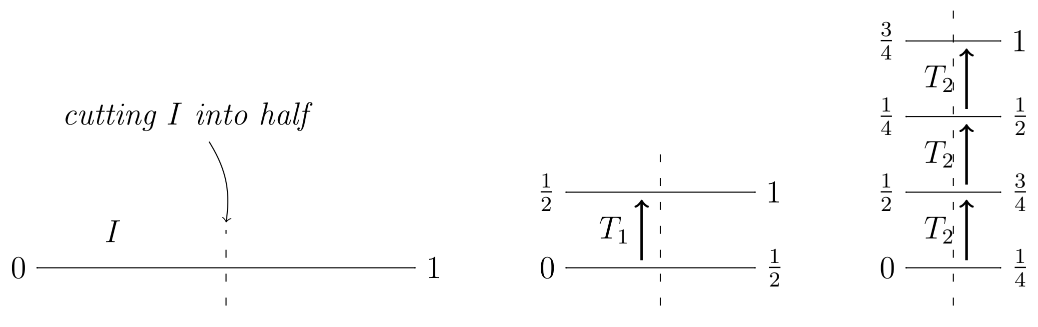


Figure 1.1: Construction of Kakutani's Example

In this first step, we got a tower of two levels. Let us define T_1 the transformation, represented by S_1.

To get S_2 act inductively and cut S_1 into half. Again stack the right half of the intervals above the left to get a new transformation T_2 represented by S_2 with now four levels (again see Figure 1.1). Continue like this to construct S_n for $n \in \mathbb{N}$ and T_n, the transformation described by S_n. Since we are cutting again and again into half, it is not surprising, that T_n is linked to the dyadic numbers, numbers of the form $\frac{k}{2^n}$, where k and n are integers. In this sense, we define $\mathcal{D}$ to be the set of left-closed, right-open dyadic intervals, i.e. intervals in $[0,1)$ with dyadic numbers as their endpoints. Then for fixed $n \in \mathbb{N}$, $\mathcal{D}^n \subseteq \mathcal{D}$ is the set of intervals in $[0,1)$ with endpoints of the form $\frac{k}{2^n}$.

Having now constructed this transformation, we are interested in the properties. Spoiler: it will turn out, that this transformation is invariant and ergodic, however not weakly mixing.

Definition 1.3.2. *The dyadic adding machine.*
By the previous construction, we define the dyadic adding machine by the piecewise function $\mathcal{K}$ on $[0,1)$:

$$\mathcal{K}(x) = \begin{cases} T_1(x): & x \in [0, \frac{1}{2}) \\ T_2(x): & x \in [\frac{1}{2}, \frac{3}{4}) \\ T_3(x): & x \in [\frac{3}{4}, \frac{7}{8}) \\ \dots \end{cases}$$

where T_n, $n \in \mathbb{N}$ are the transformations, defined by the construction.
Note: for given $n \in \mathbb{N}$, $\mathcal{K}$ coincides with T_n on the domain of the latter.

For showing that this transformation is invariant and ergodic on $[0,1)$, we need to work a little bit more and show the following lemma.

Lemma 1.3.3. *The dyadic adding machine $\mathcal{K}$ is a measure-preserving transformation of intervals in $\mathcal{D}$.*

Proof. The fact, that $\mathcal{K}$ is measure-preserving, follows directly from the construction of the transformation, since it is measure preserving in every step of the construction. Going further it also holds for infinitely many steps. For more details see for example [17]. □

We are not only interested in showing this on $\mathcal{D}$, even more we want to show this statement on all of the domain of $\mathcal{K}$, i.e. on $[0,1)$.

Theorem 1.3.4. *Let $(X, \mathcal{A}, \mu)$ be a finite measure space and $\mathcal{S}$ a sufficient semiring on X. Suppose $T : X \longrightarrow X$ such that $\mu(T^{-1}(S)) = \mu(S) \;\; \forall S \in \mathcal{S}$. Then T is an invariant transformation on all of X.*

The proof is done in [17, p.10-12] by using some "approximation techniques".

Lemma 1.3.5. *$\mathcal{D}$ is a sufficient semiring on $[0,1)$.*

Proof. First of all, looking on $\mathcal{D} := \{[a,b) : a, b$ are dyadic numbers in $[0,1)\}$, we see

1) $\forall\, D_1, D_2 \in \mathcal{D} : D_1 \cap D_2 \in \mathcal{D}$

2) $\exists\, C_1, \ldots, C_n \in \mathcal{D} : D_1 \backslash D_2 = \bigcup_{i=1}^{n} C_i$

$\Rightarrow \mathcal{D}$ is a semiring.
To show $\mathcal{D}$ is also a sufficient semiring, let us show that it is possible to approximate any open interval in the unit interval $[0,1)$ by a union of finitely many intervals in $\mathcal{D}$. Let us assume that $I = (a,b)$ is an arbitrary interval in $[0,1)$. For $\varepsilon > 0$, choose $n \in \mathbb{N}$ such that $\frac{1}{2^n} < \frac{\varepsilon}{2}$. Note that all numbers of the form $\frac{k}{2^n}$, $k = 0, 1, \ldots, 2^n - 1$ partition the interval $[0,1)$ into smaller intervals of length $\frac{1}{2^n}$. Hence there exists $l \in \mathbb{N}_0$ s.t. $\frac{l}{2^n} \le a \le \frac{l+1}{2^n}$ (for $a = 0$, define without loss of generality a as the left end of the interval) as well as there exists $m \in \mathbb{N}_0$ s.t. $\frac{m-1}{2^n} \le b \le \frac{m}{2^n}$ (for $b = 1$,define without loss of generality b as the right end of the interval).
Define now $\mathcal{D}_I := D^n \cap [\frac{l}{2^n}, \frac{m}{2^n})$. It follows

$$I \subset \bigcup_{D \in \mathcal{D}_I} \text{ and } \mu(I) \le \bigcup_{D \in \mathcal{D}_I} \mu(D) \le \mu(I) + \underbrace{\frac{1}{2^n}}_{<\frac{\varepsilon}{2}} + \underbrace{\frac{1}{2^n}}_{<\frac{\varepsilon}{2}} < \mu(I) + \varepsilon.$$

Note that for measurable sets A there exist open intervals I_n in $[0,1)$ s.t. $\mu(A) \le \sum_{n=1}^{\infty} \mu(I_n) < \mu(A) + \varepsilon$.
As long as ε is arbitrary, $\mathcal{D}_n := \mathcal{D}_{I_n}$ and it holds that

$$\mu(I_n) \le \bigcup_{D \in \mathcal{D}_n} \mu(D) < \mu(I_n) + (\mu(A) + \varepsilon - \sum_{m=1}^{\infty} \mu(I_m)) \cdot \frac{1}{2^n}$$

$$\Rightarrow \sum_{n=1}^{\infty} \sum_{D \in \mathcal{D}_n} \mu(D) < \mu(A) + \varepsilon \tag{1.5}$$

Now define $\mu_*(A) := \inf_{A \subset \bigcup_{n=1}^{\infty} D_n} \sum_{n=1}^{\infty} \mu(D_n), \quad D_n \in \mathcal{D}$. By definition $\mu(A) \le \mu_*(A)$ and by (1.5) we get that $\mu_*(A) \le \mu(A) + \varepsilon$. Therefore $\mu_*(A) \le \mu(A)$. Hence $\mathcal{D}$ is a sufficient semiring on $[0,1)$. □

Combining the above three statements, we conclude that $\mathcal{K}$ is invariant. For ergodicity we need again some additional tools.

Theorem 1.3.6. *The dyadic adding machine $\mathcal{K}$ is ergodic on $[0,1)$.*

Remark 1.3.7. For proving this theorem, let $A \in \mathcal{A}$ be a $\mathcal{K}$-invariant set of positive measure. Let $D \in \mathcal{D}_n$, where $\mathcal{D}_n$ is the set of all dyadic intervals in $[0,1)$ of length 2^{-n}. Hence $[0,1)$ is partitioned into $|\mathcal{D}_n|$ intervals of length 2^{-n} or in other words $|\mathcal{D}_n|$ is the number of dyadic intervals of length 2^{-n}, $[0,1)$ is cut into. Then

$$\mu(A) = \sum_{D \in \mathcal{D}_n} \mu(A \cap D) = |\mathcal{D}_n| \cdot \mu(A \cap D) \text{ and}$$
$$1 = \mu(I) = \mu([0,1)) = \sum_{D \in \mathcal{D}_n} \mu(D) = |\mathcal{D}_n| \mu(D)$$
$$\Rightarrow \mu(A) = \frac{\mu(A)}{\mu([0,1))} = \frac{\cancel{|\mathcal{D}_n|}\mu(A \cap D)}{\cancel{|\mathcal{D}_n|}\mu(D)} = \frac{\mu(A \cap D)}{\mu(D)}.$$

Hence the measure of A is equal to the proportion of A in D. Showing $\mathcal{K}$ being a ergodic transformation means, there exists D^δ for arbitrary given $\delta \in (0,1)$ s.t. $\mu(D^\delta) \geq \mu(A \cap D^\delta) > (1-\delta)\mu(D^\delta)$, since it follows that $\mu(A) = 1$.

To show that such a D^δ actually exists, we prove the next two lemmas.

Lemma 1.3.8. *Let $(X, \mathcal{A}, \mu)$ be a finite measure space and $\mathcal{S}$ a sufficient semiring on X. Then for any set $A \in \mathcal{A}$ and any $\varepsilon > 0$, there exists $G_\varepsilon^A = \bigcup_{i=1}^n S_i, \ \ S_i \in \mathcal{S}$ s.t. $\mu(A \triangle G_\varepsilon^A) < \varepsilon$, where $A \triangle B$ is the symmetric difference of A and B.*

Proof. Choose $\delta := \frac{\varepsilon}{2}$ and $H^\delta = \bigcup_{i=1}^\infty S_i$, where $S_i \in \mathcal{S}$ s.t. $A \subset H^\delta$ and $\mu(H^\delta \backslash A) < \varepsilon$. The existence of H^δ follows directly by the definition of the sufficient semiring $\mathcal{S}$. Let $N < M$ then

$$\bigcup_{i=1}^N S_i \subset \bigcup_{i=1}^M S_i \Rightarrow \lim_{n\to\infty} \mu\Big(\bigcup_{i=1}^n S_i\Big) = \mu\Big(\underbrace{\bigcup_{i=1}^\infty S_i}_{=H^\delta}\Big) \Rightarrow \exists n \in \mathbb{N} \text{ s.t. } \mu\Big(H^\delta \backslash \bigcup_{i=1}^n S_i\Big) < \varepsilon.$$

Define now $G_\varepsilon^A := \bigcup_{i=1}^n S_i$. Then it follows that

$$\mu(A \triangle G_\varepsilon^A) = \mu(\underbrace{A}_{\subset H^\delta} \backslash G_\varepsilon^A) + \mu(\underbrace{G_\varepsilon^A}_{\subset H^\delta} \backslash A) \leq \underbrace{\mu(H^\delta \backslash G_\varepsilon^A)}_{<\varepsilon} + \underbrace{\mu(H^\delta \backslash A)}_{<\varepsilon} < 2\varepsilon.$$

□

Lemma 1.3.9. *For all $A \in \mathcal{A}$ with $\mu(A) > 0$ and $0 < \delta < 1$ arbitrary, there exists $S^\delta \in \mathcal{S}$ s.t. $\mu(A \cap S^\delta) > (1-\delta)\mu(S^\delta)$, where $\mathcal{S}$ again is a sufficient semiring on a finite measure space.*

Proof. Choose ε between 0 and $\frac{\delta}{2-\delta}$ and $\beta := \varepsilon \cdot \mu(A)$. Define $G := G_\beta^A = \bigcap_{i=1}^{N} S_i$. Then note that $G = ((G \cap A) \cup (G \backslash A)) \subset (A \cup (A \backslash G) \cup (G \backslash A))(= A \cup G)$. Hence

$$\mu(G) < \mu(G \backslash A) \leq \mu(A \cup (A \triangle G)) \leq \mu(A)(1+\varepsilon)$$

where the second inequality follows by Lemma 1.3.8 and the definition of β. Same for $A = ((A \cap G) \cup (A \backslash G)) \subset ((A \cap G) \cup (A \triangle G))$. Hence

$$\mu(A) < \mu(A \cap G) + \mu(A) \cdot \varepsilon \Longleftrightarrow \mu(A \cap G) > (1-\varepsilon)\mu(A)$$

again by the last lemma and the definition of β.
Contrary to the claim, assume there is no $S^\delta \in \mathcal{S}$ s.t. $\mu(A \cap S^\delta) > (1-\delta)\mu(S^\delta)$, i.e. there is no $S_i \in \mathcal{S}$ s.t. $\mu(A \cap G) \leq (1-\delta)\mu(G) < (1-\delta)(1+\varepsilon)\mu(A)$ and since

$$(1-\varepsilon)\mu(A) < \mu(A \cap G) \Rightarrow (1-\varepsilon) < (1-\delta)(1+\varepsilon) \Leftrightarrow \frac{1-\varepsilon}{1+\varepsilon} < 1-\delta. \qquad (1.6)$$

But by definition of ε:

$$1+\varepsilon < \frac{2}{2-\delta} \text{ and } (1-\varepsilon) > \frac{2-2\delta}{2-\delta} = \frac{2(1-\delta)}{2-\delta} \Rightarrow \frac{1-\varepsilon}{1+\varepsilon} > 1-\delta,$$

which is a contradiction to (1.6). □

Now we can prove the Theorem and show that the dyadic adding machine is ergodic.

Proof. (*of Theorem 1.3.6*) By Lemma 1.3.5, we know that $\mathcal{D}$ is a sufficient semiring on $[0,1)$. Additionally, by the last lemma, there exists $D^\delta \in \mathcal{D}$ s.t. $\frac{\mu(A \cap D^\delta)}{\mu(D^\delta)} > (1-\delta)$ and hence by Remark 1.3.7, $\mu(A) = 1$. □

Having now an invariant and ergodic transformation $\mathcal{K}$, the missing statement is the one regarding its mixing properties. As we mentioned in the beginning of this section, this transformation is not weakly mixing.

Remark 1.3.10. From Theorem 1.2.9 we know that weakly mixing and doubly ergodicity are two equivalent properties.

Theorem 1.3.11. *The dyadic adding machine is not weakly mixing, i.e. is not doubly ergodic by Remark 1.3.10.*

Proof. We show that $\mathcal{K}$ is not doubly ergodic. To prove this, we pick two levels A and B in C_n which are precise the top and the bottom levels of C_n respectively. Then A is clearly $h_n - 1$ levels below B. Let us define the set I_k to be the copies of a level I in C_n, if for $I_k \in C_l$, $l > n$ it holds that $I = \bigcup_{k \in \mathbb{N}} I_k$. Suppose we have our copies of the levels A and B, then each copy of both, A and B are h_n levels apart. Looking at the illustration in Figure 1.2 this statement is anyway quite obvious.

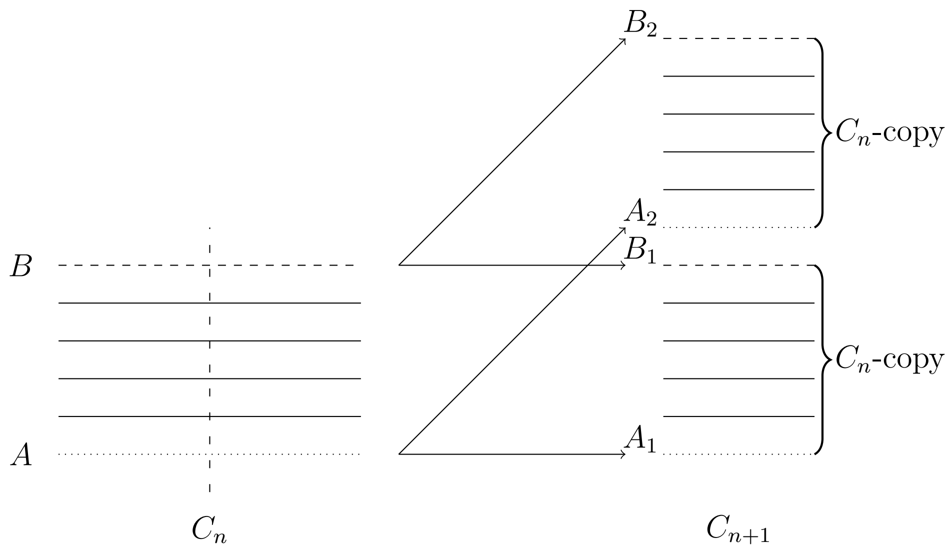


Figure 1.2: Copies of two levels A and B

Moreover, if $\mu(T^m(A) \cap B) > 0$, then $m = rh_n - 1$ for $r \in \mathbb{N}$ and if $\mu(T^m(A) \cap A) > 0$, then $m = rh_n$ for $r \in \mathbb{N}$. Since there exists no such r and s such that $rh_n = sh_n + 1$, we can conclude that $\mu(T^m(A) \cap A)$ and $\mu(T^m(A) \cap B)$ cannot be both positive for $m \in \mathbb{N}$. Hence by the definition of doubly ergodic, T is not doubly ergodic consequently not weakly mixing. □

1.3.2 Chacón's transformation

This transformation is an example for a rank one transformation which is weakly mixing but not strongly mixing. It is named after R.V. Chacón, who published this constructive example in his 1969's Paper [10].

Example 1.3.12. *Chacón.*
Again, as in Example 1.3.1, Chacón is working with the unit interval $I = [0, 1]$. The construction auf Chacón's transformation is somehow comparable to Kakutani, in the sense that again it uses cutting and stacking of some smaller intervals contained in the initial interval I.

- **Cutting step 1:** Cut the interval I into two pieces $[0, \frac{2}{3})$ and $[\frac{2}{3}, 1]$. The second interval will now be the so called "spacer or reservoir" [22]. The first interval will again be cut into three smaller intervals of same length $\frac{2}{9}$. To get the correct length of the spacer, cut the spacer at $\frac{2}{9}$ from the left of the spacer (see Figure 1.3).

Figure 1.3: Cutting Step 1

- **Stacking step 1:** Now stack these intervals in the following fashion: First, take the second most left part and stack it over the most left part. Second, take the interval of same length, starting left, from the spacer and stack it on top. Last, take the right interval and stack it on top. Together we get a tower of 4 levels plus a remaining spacer $[\frac{8}{9}, 1]$ (see Figure 1.4). Since the most upper level and the remaining spacer have no image, we need another iterate of this two construction steps.

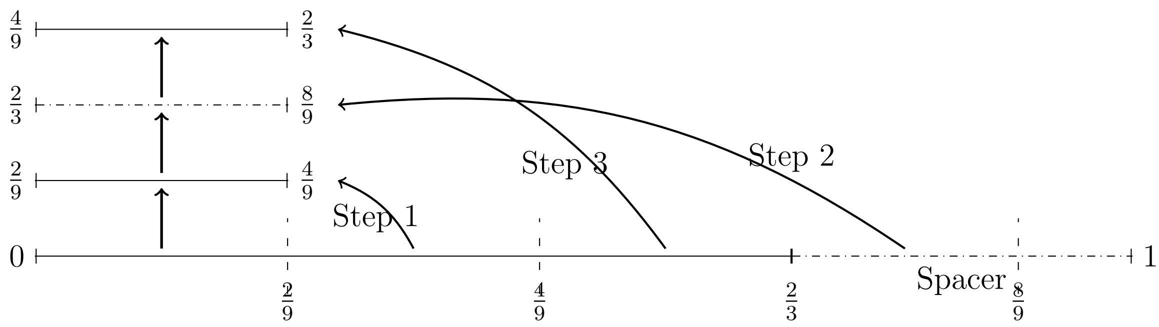


Figure 1.4: Stacking Step 1

- **Iteration of Cutting and Stacking:** The further construction, is illustrated in Figure 1.5. We cut the tower in Figure 1.4 again into three intervals with same length, this time, $\frac{2}{27}$ and the remaining spacer is again cut into an interval of same length $\frac{2}{27}$ starting on the left end of the spacer-interval. We get a new Tower with 13 levels and a remaining spacer $(\frac{26}{27}, 1]$.

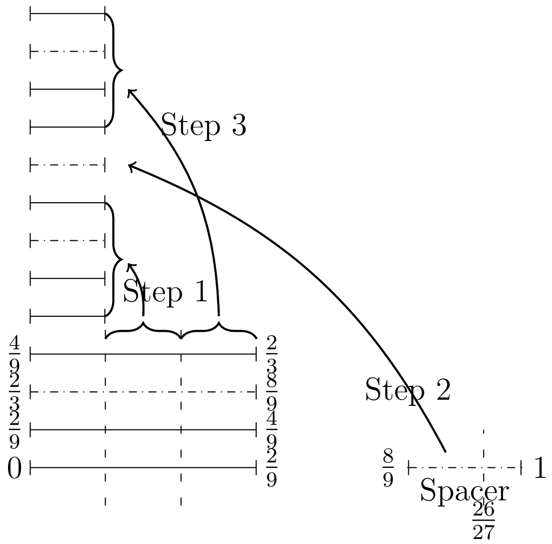


Figure 1.5: Cutting and Stacking; Step 2

Repeat this n times, $n \in \mathbb{N}$, to get a transformation T which is defined on all levels except for the upper most one and the remaining spacer. Hence the transformation is defined on a subset of $[0,1]$ of measure $1 - \frac{1}{3^n} - 2(\frac{1}{3})^{n+1} = 1 - \frac{5}{3^{n+1}}$, where the first subtrahend equals the length of the remaining spacer and the second subtrahend equals the length of the top interval. Hence the domain of the transformation T increases by increasing the number of steps, $n \to \infty$. Therefore the limiting transformation is defined on all of $[0,1]$ with measure $\mu([0,1]) = 1$.

Theorem 1.3.13. *The Chacón transformation is not strongly mixing.*

Proof. Remember, T is strongly mixing $\Leftrightarrow \lim_{n\to\infty} \mu(T^{-n}A \cap B) = \mu(A)\mu(B)$. Equivalently, T is not strongly mixing $\Leftrightarrow \exists A \in \mathcal{A},\ \mu(A) > 0 : \limsup_{n\to\infty} \mu(T^{-n}A \cap A) \neq (\mu(A))^2$. Note that

$$\mu(T^nA \cap A) \overset{\text{m.p.}}{=} \mu(T^{-n}(T^nA \cap A)) \overset{\text{m.p.}}{=} \mu(T^{-n}(T^nA) \cap T^{-n}A) \overset{\text{T inv.}}{=} \mu(A \cap T^{-n}A).$$

Hence we show that there exists $A \in \mathcal{A},\ \mu(A) > 0 : \limsup_{n\to\infty} \mu(T^nA \cap A) \neq (\mu(A))^2$. For proving this, fix a step k_0 in the construction of Chacón's transformation and define the transformation, which is representing the stack S_{k_0} in this step k_0, to be T_{k_0}. For A we choose the base of the stack and the height of T_{k_0} should be h_{k_0}. Now we claim

$$\mu(T^{h_n}A \cap A) \geq \frac{1}{3}\mu(A). \tag{1.7}$$

Then for the strong mixing property, we need $(\mu(A))^2 \geq \frac{1}{3}\mu(A) \Rightarrow \mu(A) \geq \frac{1}{3}$. But since by construction each level has length $\leq \frac{2}{9}\mu(A)$, it follows $(\mu(A))^2 \leq \frac{4}{81}\mu(A)$, hence we get a contradiction. Proving (1.7), we choose again A to be the base of the stack S_{k_0}. In step k_0+1, one more time we cut A into three subintervals A_1, A_2, A_3 and apply the stacking process to get the transformation T_{k_0+1} representing the stack S_{k_0+1}. Then for step k_0+1:

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

Here only $(T^{h_{k_0}} A_1 \cap A)$ is a non-empty set, i.e. $\mu(T^{h_{k_0}} A_1 \cap A) = \frac{1}{3}\mu(A)$ but $\mu(T^{h_{k_0}} A_2 \cap A) = \mu(\varnothing) = 0$. We cannot find an estimate for $\mu(T^{h_{k_0}} A_3 \cap A)$ so we stay with $\mu(T^{h_{k_0}} A_1 \cap A) \geq \frac{1}{3}\mu(A)$. For step k_0+2: A_1 will be cut into three parts and again only one will intersect A_1 under $T^{h_{k_0+1}}$. Similarly for A_2 and A_3. Hence we get

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

Now by induction follows for each step $n \geq k$:
We have 3^{n-k_0} parts of A, each with length $\frac{1}{3^{n-k_0+1}}\mu(A)$. Hence $\mu(T^{h_n} A \cap A) \geq 3^{n-k_0}\frac{1}{3^{n-k_0+1}}\mu(A) = \frac{1}{3}\mu(A)$. ☐

Following [11], the next Lemma shows a stronger statement of (1.7).

Lemma 1.3.14. *Let $n > 0$. Then is holds that*

(a) $\forall k \geq n: \ \mu(T^{h_k} I \cap I) \geq \frac{1}{3}\mu(I)$

(b) $\forall l \geq 0: \ \exists H = H(n, l)$ *such that* $\mu(T^H I \cap J) \geq (\frac{1}{3})^l \mu(J)$

where I and J are two levels of the stack S_k which are at most $l \geq 0$ apart, I is stacked above J and h_k is the height of S_k.

Proof. (a) is done in the proof of Theorem 1.3.13 (1.7).
For (b) we use the fact that $T^{h_n}(I)$ intersects $T^{-1}(I)$ in measure at least $\frac{1}{3}$ times $\lambda(T^{-1}(I)) = \lambda(I)$. Additionally, $T^{h_n}(I)$ intersects I. Now let $I_0, \ I_1, \ I_2$ be the copies of the level I in C_n. Then it holds that

$$T^{h_n}(I_0) = I_1 \text{ and } T^{h_n}(I_1) = T^{-1}(I_2).$$

From this and since $T^{h_n}(I)$ contains two (full) levels in the next tower C_{n+1}, we get that $T^{h_n+h_{n+1}}(I)$ intersects I and $T^{-1}(I)$ and $T^{-2}(I)$ in measure at least $(\frac{1}{3})^2$ times the measure of each of the two intervals. By using induction, it follows that the stated $H = \sum_{i=0}^{l-1} h_{n+i}$ satisfies the Lemma. □

As we mentioned before, weakly mixing and doubly ergodic are two equivalent properties, which we will now use, to prove that Chacón's transformation is weakly mixing.

Theorem 1.3.15. *The Chacón transformation is doubly ergodic.*

Proof. Choose A_1, B_1 to be two sets of positive measure and levels I_1, I_2 in some stack S_n such that $\mu(I_1 \cap A_1) > \frac{2}{3}\mu(I_1)$ and $\mu(J_1 \cap B_1) > \frac{2}{3}\mu(J_1)$, I_1 lies above J_1 and I_1 is l apart from J_1, $0 \leq l \leq h_n - 1$. Define $\delta := (\frac{1}{3})^l$. Now we can take two subintervals I and J of I_1 and J_1 respectively such that

$$\mu(I \cap A_1) > \big(1 - \frac{\delta}{3}\big)\mu(I) \text{ and } \mu(J \cap B_1) > \big(1 - \frac{\delta}{3}\big)\mu(J).$$

Let $A = I \cap A_1$ and $B = J \cap B_1$. Then by Lemma 1.3.14 it holds that

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

Hence

$$\mu(T^H A \cap B) \geq \underbrace{\mu(T^H I \cap J)}_{\geq \delta\mu(J)} - \underbrace{\mu(I \backslash A)}_{\leq \frac{\delta}{3}\mu(I)} - \underbrace{\mu(J \backslash B)}_{\leq \frac{\delta}{3}\mu(J)} > 0 \text{ and}$$
$$\mu(T^H A \cap A) \geq \mu(T^H I \cap I) - \mu(I \backslash A) - \mu(I \backslash B) > 0.$$

□

Note: Since ergodicity follows from the weakly mixing propertie, Chacón's transformation is certainly ergodic.

2 Bratteli diagrams

Remembering the construction of such (unit) interval maps, we can as well see those transformations as mappings on the Cantor set $\mathcal{C}$.

Recall:

- A Cantor set $\mathcal{C}$ is a zero-dimensional compact metric space with no isolated points.
- A homeomorphism T is aperiodic if every T-orbit is infinite.
- A homeomorphism $T : \mathcal{C} \longrightarrow \mathcal{C}$ is minimal if every T-orbit is dense in $\mathcal{C}$.

Over the last years, minimal homeomorphisms of Cantor sets have been studied in plenty Papers. In the next part of this text we are dealing with so called Bratteli diagrams, introduced in 1972 by O. Bratteli in his Paper [9]. They are a useful tool in the field of Cantor minimal systems. It was shown by R. H. Herman, I. F. Putnam and C. F. Skau in [15] that every minimal homeomorphism is topologically conjugate to the Vershik map of an ordered simple Bratteli diagram (see Theorem 2.3.5). Bratteli diagrams of its simplest form are stationary Bratteli diagrams and it was shown for example in [14] that this class is generated by the minimal substitutional system and Odometers (see Remark 2.1.6(d)).

2.1 Basic Notation

Definition 2.1.1. *Ordered Bratteli diagram.*
A **Brateli diagram** $B = (V, E)$ is a infinite graph such that $V = \bigcup_{i \geq 0} V_i$ is the Vertex-set and $E = \bigcup_{i \geq 1} E_i$ is the Edge-set, where V and E are partitioned into disjoint subsets V_i and E_i satisfying the following:

(1) $V_0 = \{v_o\}$ is a single point,

(2) V_i and E_i are finite,

(3) there exists t (target map) and s (source map) from E to V such that $t(E_i) = V_i$, $s(E_i) = V_{i-1}$ and $s^{-1}(v) \neq \emptyset \ \forall v \in V$ and $t^{-1}(v') \neq \emptyset \ \forall v' \in V \backslash V_0$ and

(4) for each $v \in \bigcup_{i \geq 1} V_i$ there is a total order $<$ between the incoming edges to v (see Definition 2.2.1 and Definition 2.2.2.

For convenience we will give some additional notation

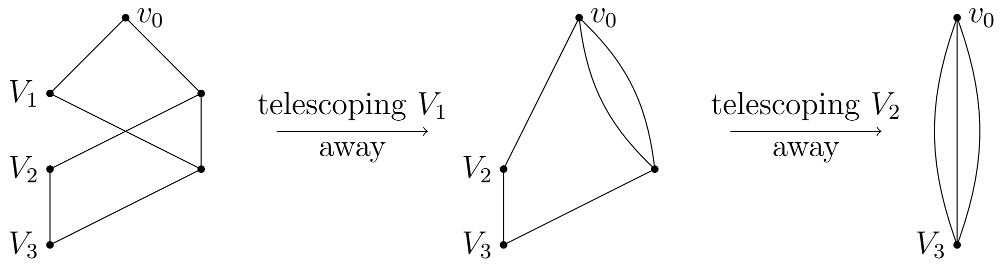
- (V_i, E_i) or just V_i is called the i-th level of B.
- $(e_i : e_i \in E_i)$ such that $t(e_i) = s(e_{i+1})$ is called the path of B.
- $e(v, v')$ is the path going from v to v'.
- X_B is the set of infinite paths, starting in v_0. X_B is equipped with the topology, generated by cylinder-sets $U(\underbrace{e_1, \cdots, e_n}_{\text{finite path in } B}) = \{x \in X_B : x_i = e_i,\ i = 1, \cdots, n\}$.

Definition 2.1.2. *Incidence matrix of a Bratteli diagram.*
To each Edge-set E_i we identify a **incidence matrix** $M(i) = (m_{v,w}(i))_{v \in V_{i-1}, w \in V_i}$, a $|V_{i-1}| \times |V_i|$ matrix whose entries $m_{v,w}(i)$ are equal to the number of edges between the vertices $v \in V_{i-1}$ and $w \in V_i$, i.e. $m_{v,w}(i) = |\{e \in E_i : s(e) = v,\ t(e) = w\}|$.

Definition 2.1.3. *Height vector of a Bratteli diagram.*
Let $B = (V, E)$ be a Bratteli diagram. The number of n-paths from v_0 to $v \in V_n$ is defined by the **height vector** $(h_v(n))_{v \in V_n}$ with $h_v(n) = |\{x_1, \cdots, x_n : s(x_1) = v_0,\ t(x_n) = v\}|$.

So far we considered the collection of edges E_i separately. If we take all paths from V_{i-1} to V_j for some $j \geq i$ instead, we define $E_{i,j}$ the collection of paths from V_{i-1} to V_j. Observe that the product matrix $M(i) \cdots M(j)$ is the incidence matrix $M(i, j)$ of the edge set $E_{i,j}$. This process, where we put vertices and levels together, is called **telescoping**. The reversed process, where we add more levels of vertices and edges, is called **microscoping**.

Example 2.1.4. *Telescoping a Bratteli diagram.*
Let B be the Bratteli diagram illustrated below. Telescoping the first and second level away corresponds to the following computation from Figure 2.1:



$$M(1) = \begin{pmatrix} 1 & 1 \end{pmatrix} \qquad M(1,2) = \begin{pmatrix} 1 & 2 \end{pmatrix}$$

$$M(2) = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \qquad M(1,3) = \begin{pmatrix} 3 \end{pmatrix}$$

$$M(3) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad M(3) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Figure 2.1: Telescoping a Bratteli diagram

Through telescoping (and/or microscoping), we transform Bratteli diagrams into equivalent diagrams. It produces isomorphic Bratteli-Vershik systems. It is used to transform the incidence matrices in such a way, that all their entries are strictly positive. This is only possible, if the Bratteli diagram is simple (see Remark 2.1.6(a)).

Definition 2.1.5. *Simple Bratteli diagram.*
We call a Bratteli diagram $B = (V, E)$ **simple**, if the following two conditions are satisfied:

(1) The set of vertices and edges (V, E) is non-trivial and

(2) B can be transformed via telescoping to a diagram, such that it has full connectivity between any two consecutive levels.

Remark 2.1.6. Let $B = (V, E)$ be a Bratteli diagram.

(a) B is called **simple** if $\forall n \; \exists m > n : e(v, w) \neq \emptyset \; \forall v \in V_n$ and $\forall w \in V_m$.

(b) If $M(n) = M(1) \; \forall n \geq 2$, in other words, the diagram repeats itself after each level, then B is **stationary**.

(c) If $|V_n| \leq d \; \forall n \geq 1$ then B is said to be a **finite rank** diagram. Let d be the smallest number such that there exists infinitely many levels in B with d vertices, then the **rank** of B is d.

(d) If B has rank one, i.e. $|V_n| = 1 \; \forall n$ then B is called a **minimal Cantor dynamical system**. Note that every Odometer is a minimal Cantor system.

2.2 Ordering of Bratteli-Vershik Systems

As already mentioned in Definition 2.1.1(4), B is totally ordered on the set of N-paths ending in $v \in V_N$. Showing this, we take two finite edge-labeled paths $v = v_1 \cdots v_N$ and $w = w_1 \cdots w_N$, $v_i, w_i \in E_i$. Suppose v and w have the same endpoint in V_{N+1}, we can find $m < N$ such that m is the largest index where v_m and w_m differ, i.e. $t(v_m) = t(w_m)$. The ordering follows by defining $v < w$ if $v_m < w_m$. This construction is also a partial order on the set of arbitrary N-paths.

Definition 2.2.1. *Minimal and maximal path.*
We call a path $x = (x_1, x_2, \ldots, x_i, \ldots)$ **minimal**, respectively **maximal**, if every x_i is minimal (maximal) amongst all elements from $t^{-1}(t(x_i))$ (according to the introduced lexicographic order above). For every vertex $v \in V_i,\ i \geq 0$ there exists a unique minimal, maximal respectively, path in $E(v_0, v)$. The collection of (infinite) minimal, respectively maximal, paths in X_B is denoted by $X_{\min},\ X_{\max}$ respectively.

Definition 2.2.2. *Ordered Bratteli diagram.*
Let $B = (V, E)$ be a Bratteli diagram. B together with a partial order $\leq$ on E is called an **ordered Bratteli diagram** $B = (V, E, \leq)$.

One can see, that there are infinite long paths $x' \in X_{BV}$ such that the initial N-path $x = (x_1, \ldots, x_n)$ is minimal among all N-paths with the same endpoint. The same holds again for maximal N-paths. Additionally, every $v \in V_N$ can just have one minimal incoming path x with v being the endpoint. If the set of all (infinite) minimal paths $X_{\min}$ is a single element, we call $x^{\min}$ the unique minimal path in X_B. Again same is true for the unique (infinite) maximal path $x^{\max}$. The next Example, Example 2.2.3, by [18] shows an ordered Bratteli diagram with two (infinite) minimal and one (infinite) maximal paths.

Example 2.2.3. *Minimal and maximal paths of a Bratteli diagram.*[18]
Let $B = (V, E, \leq)$ be an ordered Bratteli diagram as constructed below. (Note that by telescoping the levels 2, 4, ... one could transform B into a simple Bratteli diagram B' where each vertex from one level is connected to each vertex from the next level, see right hand side of Figure 2.2.) For B, we can see (left hand side of Figure 2.2), that $X_{\min}$ consists of two paths, whereas $X_{\max}$ consists of one unique path. The two minimal paths are the dashed lines and the maximal path the dotted line. For B' there are again two minimal paths and one unique maximal path.

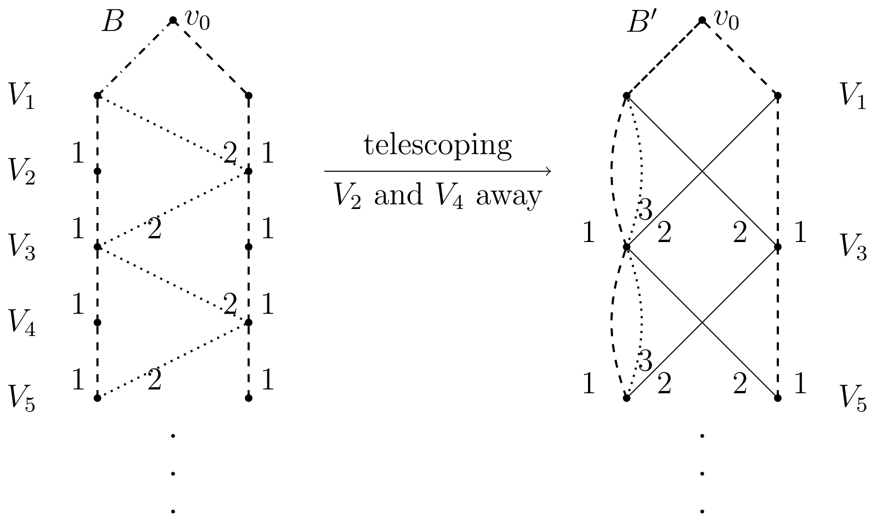


Figure 2.2: Minimal and Maximal Paths of a ordered Bratteli diagram

Definition 2.2.4. *Properly ordered Bratteli diagram.*
A Bratteli diagram $B = (V, E, \leq)$ is called **properly ordered** if

(1) (V, E) is a simple diagram and

(2) $X_{\min} = \{x^{\min}\}$ and $X_{\max} = \{x^{\max}\}$, i.e. each $X_{\min}$ and $X_{\max}$ contains only one unique path.

Since we have now a propertly ordered Bratteli diagram B, we define a minimal homeomorphism $\mathcal{V} : X_B \longrightarrow X_B$ on it. This homeomorphism will be called the **Vershik map**. The resulting system $(X_B, \mathcal{V})$ is then a Cantor minimal system and called **Bratteli-Vershik system** (compare to the beginning of this chapter). The Vershik map $\mathcal{V}$ is defined as follows:

- $\mathcal{V}(x^{\max}) = x^{\min}$.
- If $x = (e_1, e_2, \dots) \neq x^{\max}$, choose i to be the smallest number such that e_i is not a maximal edge.
- Let f_i be the successor of e_i, i.e. $t(e_i) = t(f_i)$.
- Then $\mathcal{V}(x) = y = (f_1, \dots, f_{i-1}, f_i, e_{i+1}, e_{i+2}, \dots)$, where $(f_1, \dots f_{i-1})$ is the minimal path connection v_0 with $s(f_i)$.

This map is also called the **Vershik adic transformation**.

2.3 Kakutani-Rokhlin partitions and the Bratteli-Vershik model theorem

The fact, that every minimal homeomorphism is pointedly isomorphic to a Bratteli-Vershik system, is now presented below. First we introduce the concept of Kakutani-Rokhlin partitions by R. Herman, F. Putnam and F. Skau in [15].

Definition 2.3.1. *Kakutani-Rokhlin partition.*
Let X be a Cantor set and T a continuous map on X. A **Kakutani-Rokhlin partition** is a partition

$$\mathcal{P} = \{T^j(B_i) | i \in A \text{ and } 0 \leq j < h_i \text{ for } A \text{ a finite set and } h_i \in \mathbb{N}\}$$

of X into clopen sets (i.e. open sets with open complement) that are pairwise disjoint and cover X. The base of a Kakutani-Rokhlin partition is $B = \bigcup_{i \in A} B_i$ and h_i are the heights.

Remark 2.3.2. Cutting and Stacking systems as we discussed in the first chapter can now be represented as such Kakutani-Rokhlin partitions by taking the stacks constructed in the n-th step of the Cutting and Stacking process. In the sense of Kakutani-Rokhlin partitions, this stacks are represented by the sets $\{T^j(B_i(n))\}_{j=0}^{h_i(n)-1}$.

Definition 2.3.3. *Essentially minimal dynamical system.*
A pointed dynamical system (X, T, x) is called **essentially minimal** if the following holds:

$$\bigcup_{n \in \mathbb{N}} T^n(U) = X, \text{ for every open set } U, x \in U.$$

Remark 2.3.4. Minimality of a dynamical system implies essentially minimality of the dynamical system. Unlike for minimal systems, essentially minimal systems are also such systems containing a fixed point. For an example see the Image on the left of Figure 2.4.

The next Theorem will now show, that we can in fact interpret every Cutting and Stacking system as a Bratteli-Vershik system.

Theorem 2.3.5. *The Bratteli-Vershik model theorem ([15, Theorem 4.7]).*
For every (pointed) Cantor minimal system (X, T, x), where X is the Cantor set, T an essentially minimal homeomorphism and x an element in X, there exists a propertly ordered Bratteli diagram $B = (V, E, \leq)$ such that (X, T, x) is pointedly conjugate to $(X_B, \mathcal{V}, x^{min})$, with $\mathcal{V}$ the Vershik adic transformation and x^{min} the unique minimal path in X_B.

Proof. Let $(\mathcal{P}_n)_{n\geq 0}$ be a sequence of Kakutani-Rokhlin partitions, where

$$\mathcal{P}_n = \{T^j B_{n,i} | i \in A_n,\ A_n \text{ a sequence of finite sets and } 0 \leq j < h_{n,i}\},$$

$\mathcal{P}_0 = \{X\}$ and with base $B_n = \bigcup_{i \in A_n} B_{n,i}$. Further let $(\mathcal{P}_n)_{n\geq 0}$ satisfy the following conditions:

(KR1) $\mathcal{P}_n$ is nested, i.e. $B_{n+1} \subset B_n$.

(KR2) $\mathcal{P}_{n+1}$ refines $\mathcal{P}_n$, i.e. $\mathcal{P}_{n+1} \succeq \mathcal{P}_n$.

(KR3) $\bigcap_{n\geq 0} B_n$ is a single point.

(KR4) $\{A \in \mathcal{P}_n : n \in \mathbb{N}$ spans the topology of $X\}$.

Ad (KR1): Take an ordered Bratteli diagram $B = (V, E, \leq)$ and identify it with the nested sequence $(\mathcal{P}_n)_{n\geq 0}$ in the following way:

1) There is a 1-1 correspondence between the number of towers $|A_n|$ in $\mathcal{P}_n$ and V_n, the set of vertices at level n, i.e. each vertex $v_{n,i} \in V_n$ corresponds to a tower $S_{n,i} = \{T^j B_{n,i} | 0 \leq j \leq h_{n,i}\}$ by Remark 2.3.2.

2) $S_{n,i}$ will cover towers in $\mathcal{P}_{n-1}$. This will happen in a certain order, lets say $S_{n-1,k_1}, \ldots, S_{n-1,k_m}$, i.e. $S_{n,i}$ is partitioned into m different parts. We link the m edges, ordered such that $e_{1,i} < e_{2,i} < \cdots < e_{m,i}$ with $t(e_{j,i}) = v_{n,i}$ and $s(e_{j,i}) = v_{n-1,k_j}$. E_n is then the disjoint union over $i \in A_n$ of the edges with endpoint $v_{n,i}$.

The two constructed systems (X, T, x) and $(X_B, \mathcal{V}, x^{\min})$ are pointedly isomorphic for $B = (V, E, \geq)$ a properly orderes Bratteli diagram with unique minimal path $x^{\min}$. More precisely, we can define the isomorphism $F : X \longrightarrow X_B$ in such a way, that for $y \in X$, the path $F(y)$ goes through the vertex in V_n which corresponds to the tower $\mathcal{P}_n$, where y is located. Suppose $F(y)$ goes through $w \in V_{n-1}$ and $v \in V_n$, each corresponds to the tower $S_w \in \mathcal{P}_{n-1}$, $S_v \in \mathcal{P}_n$ respectively. Hence $F(y)(n)$ is the k-th edge among the ordered edges e with $t(e) = v$ and the k-th time, S_v covers one of the towers at level $n-1$,where y is being "picked up". Note that S_v will cover S_w the k-th time here. From this we see that $F(x) = x^{\min}$. Now choose a nested sequence $(B_n)_{n\in\mathbb{N}}$ of clopen sets, shrinking to x, i.e. $B_{n+1} \subset B_n$ and $\bigcap_n B_n = \{x\}$, i.e. a single point. Thus we get $\mathcal{P}_n$ by building towers over B_n by considering the return map (or hitting map, see Definition 1.1.4) to B_n. □

This proof is as well done in more detail in [15].

Example 2.3.6. *Bratteli diagram of the dyadic Odometer.*
Looking back to the first chapter, where we constructed the Kakutani-transformation

or dyadic Odometer according to the Cutting and Stacking process, we can identify this transformation by Theorem 2.3.5 with a Bratteli diagram of rank one.

Remember the dyadic Odometer or dyadic adding machine from Definition 1.3.2. Suppose $(X, \mathcal{K})$ is this transformation. We already now the Cutting and Stacking process from Example 1.3.1. Let us as well suppose that we have a properly ordered Bratteli diagram $B = (V, E, \leq)$.
Then we define the map $\gamma : E \longrightarrow \{0, 1\}$ such that $\forall e \in E$:

$$\gamma(e) = \begin{cases} 0 & \text{if } e \text{ is a minimal edge} \\ 1 & \text{if } e \text{ is a maximal edge.} \end{cases}$$

Hence we get a conjugacy between γ and $\mathcal{V}$, the Vershik map, such that

$$\mathcal{V}((e_1, e_2, \dots)) = (\gamma(e_1), \gamma(e_2), \dots).$$

In other words, we identify the dyadic Odometer (or dyadic adding machine) with the Bratteli diagram shown in Figure 2.3.

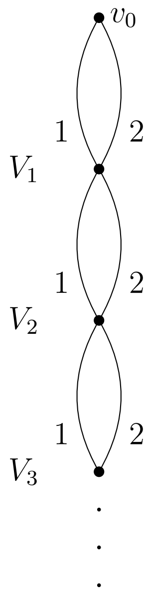


Figure 2.3: Bratteli diagram of the Odometer

Example 2.3.7. *Bratteli diagram of Chacón's transformation.*
Not only the Kakutani example, also the Chacón-transformation can be identified with Bratteli diagrams. Even though it is much harder. This is done in Park's Paper [19]. The reason why it is harder, is by the fact, that Chacón is using a spacer. He showed there is a isomorphism between the Bratteli-Vershik system $(X_B, \mathcal{V}_B)$ and (X_C, T_C) for an ordered Bratteli diagram B and Chacón's

transformation T_C ([19, Theorem 5]). The easiest way of imaging Chacón's transformation (see Figure 2.4, left hand side) is in fact not properly ordered, since it has two minimal paths $x = SSS\cdots$ and $y = 000\cdots$ and two maximal paths x and $z = 333\cdots$. Taking $T(x) = x$ and $T(z) = y$, we see that x is a fixed point but by Remark 2.3.4, (X_C, T_C, x) is at least essential minimal. Still T cannot be extended continuously to $X_{\max}$. Park found a way to fix this problem by finding a Bratteli diagram with unique minimal and maximal paths (see right part of Figure 2.4). To conclude, we will give a nudge by illustrating what Park did in his Paper.

Chacón's transformation is a substitution system, defined by

$$\begin{cases} 0 \longrightarrow & 0010 \\ 1 \longrightarrow & 1. \end{cases}$$

Alternatively we may interpret the Chacón's transformation by successive blocks B_k where

$$B_0 = 0,\ B_1 = 0010,\ B_{k+1} = B_k B_k 1 B_k.$$

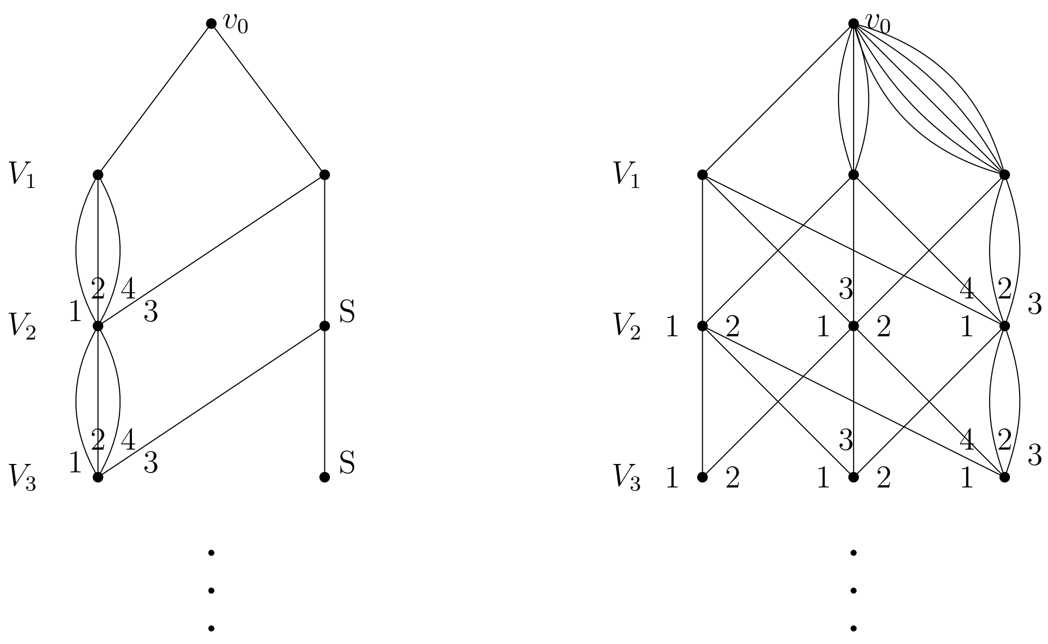


Figure 2.4: Bratteli diagram of Chacón's transformation

2.4 Cutting-and-Stacking versus Bratteli-Vershik systems

We already saw how to transfer certain Cutting-and-stacking transformations into Bratteli diagrams (Figure 2.3 and Figure 2.4). In this section, we are giving an algorithm how to turn Bratteli diagrams into cutting-and-stacking.

Suppose we have an orderes Bratteli diagram $B = (V, E, \leq)$. Beginning at level $V_0 = \{v_0\}$ of the Bratteli diagram, we declare the first stack S_0 of the cutting-and-stacking transformation as single stack of height 1. The ordering of B will be transferred to a coding at each stack. For V_0, we code S_0 with the empty code ε. In the following, we assume the stacks of each cutting-and-stacking step to be S_v for $v \in V_{n-1}$.

1) For each level V_n of the Bratteli diagram, cut the stack S_n for $v \in V_{n-1}$ into $\#\{e \in E_n : s(e) = v\}$ pieces $S_{v,e}$.

2) Coding each piece $S_{v,e}$ means copying the symbol of each edge e to $S_{v,e}$. If in the first step, $i = 1$ with $M(1) = \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix}$, in other words there exists just one edge $e \in E_1$ with $t(e) = w$, $w \in V_1$, we again take the empty code ε for slice $S_{w,e}$.

3) If there exists, take now the direct successor e', $e' \in V_{n-1}$ and $s(e') = v$, of e among all edges in E_n such that they have the same target vertex $v \in V_n$ to get $S_{v',e'}$. To conclude the first run, put $S_{v,e'}$ on top of $S_{v,e}$.

Example 2.4.1. Recall Figure 2.2, the Bratteli diagram B with two minimal and one maximal path. Applying the algorithm from above, we get a cutting-and-stacking transformation where two minimal paths corresponds to the bottom of each stack and the maximal path corresponds to the top of one part of the split stack. Again minimal and maximal paths are illustrated in dashed and dotted respectively in Figure 2.5.

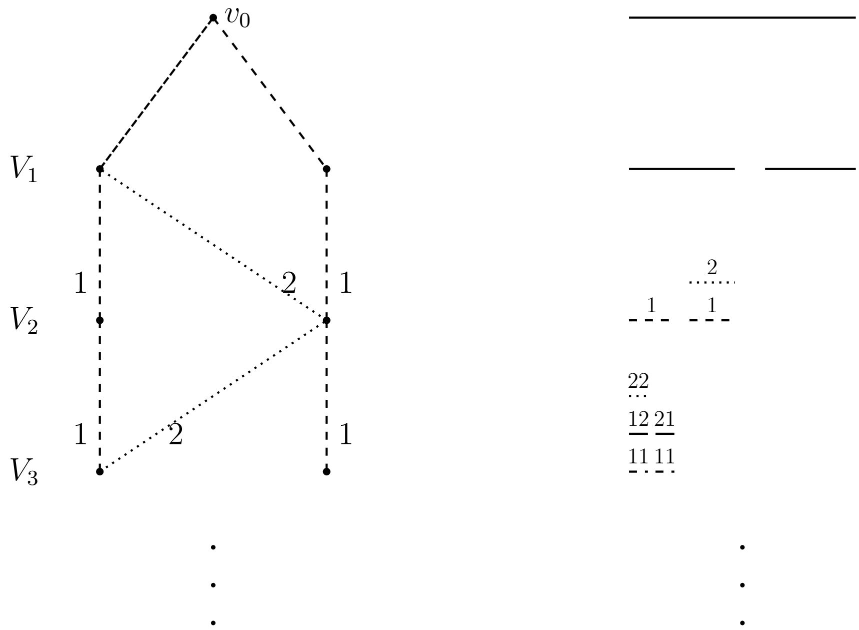


Figure 2.5: Bratteli diagram with equivalent cutting-and-stacking

For ordered Bratteli diagrams, this algorithm produces stacks, which represent the minimal and maximal path of the Bratteli diagram at each step n. The bottom level of the stacks show the minimal n-paths whereas the top level of the stacks show the maximal paths. From this, it follows that the number of minimal and maximal paths is bounden by the smallest number of pieces a stack is cut into, hence by the smallest number of vertices at one particular level such that there exists infinitely many levels with this many vertices, hence bounded by the rank of B (by Remark 2.1.6).

As seen in the beginning of this text, cutting-and-stacking procedures are used for constructing interval maps $T : [0,1] \longrightarrow [0,1]$. If we have given a Bratteli diagram, we cannot say something about the width of each stack at first. For this, we need a Bratteli-Vershik system $(X_{BV}, \mathcal{V})$ equipped with a $\mathcal{V}$-invariant measure μ. Even having a ordered Bratteli-Vershik system is not enough. We need an order $\leq_s$ of each outgoing edge $e \in E_n$ such that $s(e) = v$, $v \in V_{n-1}$ for all $n \in \mathbb{N}$. With this we get a total lexicographical order $\prec$ on X_{BV} (unlike the partial order $\leq$ of incoming edges). For $x \neq x' \in X_{BV}$, we can find the smallest $m \in \mathbb{N}$ such that $x_m \neq x'_m$ but $s(x_m) = s(x'_m)$ and set $x \prec x' \Longleftrightarrow x_m <_s x'_m$.

We want to prove the following Proposition:

Proposition 2.4.2. $(X_{BV}, \mu, \mathcal{V})$ *is isomorphic to the cutting-and-stacking system* $([0,1], \lambda, T)$*, where* λ *is the Lebegue-measure.*

First we define a weight function $\omega : \bigsqcup_n E_n \longrightarrow [0,1]$, where

$$\omega(e) := \frac{\mu([x_1 \cdots x_n e])}{\mu([x_1 \cdots x_n])} \text{ for an arbitrary path } x_1 \cdots x_n \text{ such that } t(x_n) = s(e) \in V_n.$$

By our assumption on $\mathcal{V}$-invariance of μ, every such path $x_1 \cdots x_n$ with $t(x_n) = v \in V_n$ and any other path with target vertex $v \in V_n$ has equal mass. We also notice that

$$\sum_{s(e)=v} \omega(e) = 1 \; \forall n = 0, 1, 2, \ldots, v \in V_n.$$

Then we can define $\varphi : X_{BV} \longrightarrow [0,1]$ such that

$$\varphi(x) = \sum_{s(e)=v_0, e<_s x_1} \omega(e) + \sum_{n=2}^{\infty} \mu\Big([x_1 \cdots x_{n-1}]\Big) \times \sum_{s(e)=s(x_n), e<_s x_n} \omega(e).$$

Using this function, we define a semi-metric on X_{BV}. For x and x' two paths in X_{BV}:

$$d(x, x') = \big|\varphi(x) - \varphi(x')\big|.$$

Now we can define the transformation $T : [0,1) \longrightarrow [0,1)$ by

$$T(y) = \varphi \circ \mathcal{V} \circ \varphi^{-1}(y).$$

Proof. (*of Proposition 2.4.2*) Let $x = x_1 \cdots x_n$ be a n-path starting in v_0 in X_{BV}. Let $x^{\min}$ and $x^{\max}$ be the minimal respective maximal path in the cylinder $[x]$ corresponding to the lexicographical order $\prec$. Then

$$\varphi(x^{\max}) - \varphi(x^{\min}) = \mu([x]). \tag{2.1}$$

Now suppose $\tilde{x}$ is another n-path starting in v_0 succeeding x with respect to $\prec$, then

$$\varphi(x^{max}) = \varphi(\tilde{x}^{\min}). \tag{2.2}$$

Hence φ is not injective, but identifies pairs of point and turns the Cantor set X_{BV} into the unit interval $\varphi(X_{BV}) = [0,1)$. Since there are just countable many pairs $x^{\max}$ and $\tilde{x}^{\min}$, they have total μ-measure zero and without these points, φ is injective. From (2.1) and (2.2) it follows that φ is an order-preserving isomorphism from $(X_{BV}, \mu, \prec)$ to the unit interval $([0,1), \lambda, <)$ with λ being the Lebegue measure and $<$ the standard ordering. Thus also $T = \varphi \circ \mathcal{V} \circ \varphi^{-1}$ is injective but for countably many points where we have an one-to-two correspondence as in (2.2):

Suppose we have $x = x_1 \cdots x_n$ and $\tilde{x} = \tilde{x}_1 \cdots \tilde{x}_n$ with $\varphi(x^{\max}) = \varphi(\tilde{x}^{\min}) = y$ as before. If there exists $m < n$ such that $x_i = \tilde{x}_j \; \forall j \leq m$ and x_m is not a maximal incoming edge, then

$$\mathcal{V}(x^{\max}) = v^{\max} \text{ and } \mathcal{V}(\tilde{x}^{\max}) = \tilde{v}^{\max}$$

for two consecutive n-paths $v = \mathcal{V}(x)$ and $\tilde{v} = \mathcal{V}(\tilde{x})$. For this paths, it follows that

$$T(y) = \varphi(v^{\max}) = \varphi(\tilde{v}^{\min}),$$

hence T is well-defined, and even more, T is continuous at y. From this it follows that there are at most $\sum_{i,j} m_{ij}(n)$ pairs $x^{\max}$ and $\tilde{x}^{\min}$ for every n (they correspond to the total number of slices after the $n-1$-st cutting step to create the n-th level stack C_n) such that T can be discontinuous at $\varphi(x^{\max}) = \varphi(\tilde{x}^{\min}) = y$. Except of those points, T acts as a local isometry, in other words, exactly as a cutting-and-stacking map. □

3 Unique ergodicity

Since we have discussed Bratteli diagrams in the last chapter, we now focus on ergodic invariant measures for Bratteli diagrams. We are trying to answer the question: what is the exact number of ergodic probability measures on B such that they are invariant with respect to the tail equivalence relation $\mathcal{E}$ (see Definition 3.0.1), for a given Bratteli diagram B. This is done in [4]. In this section, we will follow this Paper by Bezuglyi, Karpel and Kwiatkowski. We study the set $\mathcal{M}_T(X)$, the set of all T-invariant Borel probability measures where X is a compact space and $T : X \longrightarrow X$ a continuous mapping. By the Kakutani-Markov fixed-point theorem, this set is always non-empty and each extreme point is a ergodic measure.

To start with, we need to specify some notation and definitions. First of all, as in [4], we are working with invariant probability measures with respect to the tail equivalence relation $\mathcal{E}$ on the path space X_B of a Bratteli diagram B, denoted by $\mathcal{M}_1(B)$.

Definition 3.0.1. *Tail equivalence relation $\mathcal{E}$.*
To paths in X_B are **tail equivalent** (i.e. $(x, y) \in \mathcal{E}$) if $\exists k \in \mathbb{N}$ such that for $x = (x_n), y = (y_n)$ it holds that $x_i = y_i \; \forall i \geq k$.

Remember Definition 2.1.5 and observe that a Bratteli diagram is simple if $\mathcal{E}$ is minimal. With this in mind, we should mention that telescoping preserves X_B, $\mathcal{E}$ and $\mathcal{M}_1(B)$. Secondly, we combine the information of the incidence matrix (Definition 2.1.2) and the height vector (Definition 2.1.3) to get the sequence $F(n)$ of stochastic matrices. Those will be needed to work with measures. This **stochastic version of the incidence matrix** of B is defined by:

$$F(n) = (f_{v,w}(n))_{v \in V_{n-1}, w \in V_n}, \quad f_{v,w}(n) = m_{v,w}(n)\frac{h_v(n-1)}{h_w(n)}. \tag{3.1}$$

3.1 Bratteli diagrams and their invariant measures

Let μ be a measure in $\mathcal{M}_1(B)$, which means μ is a Borel probability non-atomic $\mathcal{E}$-invariant measure on X_B. If we suppose that $e,\ e'$ are two edges in $E(v_0, v)$ for $v \in V_n,\ n \geq 1$, then μ is an $\mathcal{E}$-invariant measure if and only if

$\mu([e]) = \mu([e'])$, where $[e], [e']$ are the cylindersets of the infinite paths e and e'.

Let us determine the measure by the vector

$$p(n) = \big(p_v(n) : v \in V_n\big),\ n \geq 1 \text{ such that } p_v(n) = \mu([e(v_0, v)]).$$

In other words, $p_v(n)$ is the measure of the cylinder set, containing the finite path $e(v_0, v)$. Keeping this in mind, we see for a vertex $v \in V$, that the cylinder set $[e(v_0, v)]$ is equal to the union of $[e(v_0, v)]$ and all other possible paths going from v to any vertex w in the next level, hence

$$[e(v_0, v)] = \bigcup_{e(v,w), w \in V_{n+1}} [e(v_0, v), e(v, w)]. \tag{3.2}$$

If we pick one particular $w \in V_{n+1}$, we obtain that

$$[e(v_0, w)] \subset [e(v_0, v)]. \tag{3.3}$$

Thus we get the vector $p(n)$ by the formula

$$p_w(n) = \frac{\mu\big(X_w(n)\big)}{h_w(n)}, \text{ where } X_w(n) = \bigcup_{e \in E(v_0, w)} [e], \ w \in V_n. \tag{3.4}$$

By [15], given a vertex w, $X_w(n)$ is the corresponding tower in the Kakutani-Rokhlin partition. We can define the measure of the tower by

$$\mu\big(X_w(n)\big) = h_w(n) p_w(n) =: q_w(n).$$

Note that so far, $p(n)$ was not a probability vector in general, but $q(n)$ is.

By (3.2) and (3.3) we obtain that

$$p(n-1) = M(n) p(n), \tag{3.5}$$

since $M(n)$ gives us the incidence matrix going from level $n-1$ to n. Trying to get the analogous result for $q(n)$ we see,

$$\begin{aligned} q_v(n-1) = h_v(n-1) p_v(n-1) &\overset{(3.5)}{=} h_v(n-1) \sum_{w \in V_n} m_{v,w}(n) p_w(n) \\ &\overset{(3.4)}{=} \sum_{w \in V_n} h_v(n-1) m_{v,w}(n) \frac{q_w(n)}{h_w(n)} \overset{(3.1)}{=} \sum_{w \in V_n} f_{v,w}(n) q_w(n). \end{aligned}$$

Hence

$$q(n-1) = F(n) q(n). \tag{3.6}$$

Lemma 3.1.1. *([6, Theorem 2.9]) Every $F(n)$ is a column-probability matrix, i.e. the entries in a column add to 1.*

Proof. By the calculations above, we get

$$\begin{aligned} \sum_{v \in V_{n-1}} f_{v,w}(n) = \sum_{v \in V_{n-1}} m_{v,w}(n) \frac{h_v(n-1)}{h_w(n)} &= \frac{1}{h_w(n)} \underbrace{\sum_{v \in V_{n-1}} m_{v,w}(n) h_v(n-1)}_{\# \{\text{paths from } v_0 \text{ to } w\}} \\ &= \frac{1}{h_w(n)} h_w(n) = 1. \end{aligned}$$

□

Next, we give an important definition:

Definition 3.1.2. *Standard simplex.*
$\triangle^{(n)} := \{(x_w^{(n)})_{w\in V_n}; : \sum_{w\in V_n} x_w^{(n)} = 1 \text{ and } x_w^{(n)} \geq 0, w \in V_n\}$ are the **standard simplices** in $\mathbb{R}^{|V_n|}$ with $|V_n|$ vertices $\{e^{(n)}(w) = (0,\dots,0,1,0,\dots,0) : w \in V_n \text{ and } e_u^{(n)}(w) = 1 \Leftrightarrow u = w\}$

The next Proposition is for example done in [6] as well as in [5].

Proposition 3.1.3. *The number of ergodic invariant measures of a Bratteli-Vershik system is bounded by its rank.*

Proof. Suppose μ is a $\mathcal{E}$-invariant probability measure on X_B with values $q_w(n) = \mu\big([e(v_0,v)]\big)$ for $v \in V_n, n \in \mathbb{N}$ which, by the Kolmogorov Extension Theorem, determine the measure completely and satisfy (3.6). Thus we can interpret (3.6) as infinite decreasing sequence of linear maps between $\mathbb{R}^{|V_n|}$:

$$\mathbb{R} \xleftarrow{F(1)} \mathbb{R}^{|V_1|} \xleftarrow{F(2)} \mathbb{R}^{|V_2|} \xleftarrow{F(3)} \mathbb{R}^{|V_3|} \xleftarrow{F(4)} \cdots .$$

As we mentioned in the beginning of this chapter, the ergodic measures are exactly the extreme points of the set

$$\triangle_\infty^{(n)} := \bigcap_{j\geq 1} \underbrace{F(n+1)F(n+2)\cdots F(n+j)(\triangle^{(n)})}_{=:\triangle_j^{(n)}} .$$

If we now suppose that the Bratteli-Vershik system is of rank r, i.e. $|V_n| = r$ infinitely often, then $\triangle_\infty^{(n)}$ cannot have more than r extreme points for each n, hence (X,T) has not more than r ergodic measures. □

3.2 Condition for unique ergodicity of Bratteli diagrams

In this part, we are following [4] to give a proof for a condition of unique ergodicity, i.e. for a Bratteli diagram with just one invariant measure or more precise we are talking about the situation where $\mathcal{M}_1(B)$ consists of just one element.

Theorem 3.2.1. *Unique ergodicity of Bratteli diagrams([4, Theorem 3.1]).*
Suppose $B = (V,E)$ is a Bratteli diagram and (X_B,V) the corresponding Bratteli-Vershik system. Then (X_B,V) is uniquely ergodic if and only if, after appropriate telescoping,

$$\lim_{n\to\infty} \max_{w\neq w'\in V_n} \Big(\sum_{v\in V_{n-1}} |f_{v,w}(n) - f_{v,w'}(n)| \Big) = 0. \tag{3.7}$$

Remember $f_{v,w}(n)$ denotes the entries of the stochastic incidence matrix $F(n)$, defined by B.

Before we are able to prove this theorem, we may talk about those extreme points of convex sets which corresponds to the ergodic measures and how we can interpret these convex sets as decreasing sequences. First define

$$G(n, n+m) = F(n) \cdot F(n+1) \cdots F(n+m),\ m \geq 0, n \geq 1.$$

For $u \in V_{n-1}$ and $v \in V_{n+m}$ the elements of $G(n, n+m)$ are defined by

$$g_{u,v}(n, n+m) = \sum_{(v_1, v_2, \ldots, v_m) \in V_n \times \cdots \times V_{n+m-1}} f_{u,v_1}(n) \cdots f_{v_{m-1},v_m}(n+m-1) \cdot f_{v_m,v}(n+m)$$

As defined above, the set $\triangle_m^{(n)} = F(n+1)F(n+2) \cdots F(n+m)(\triangle^{(n)})$ form a decreasing sequence of convex polytopes in $\triangle^{(n)}$. The vertices of $\triangle_j^{(n)}$ are elements of $\{g_v(n, n+m) : v \in V_{n+m}\}$, where

$$g_v(n, n+m) = \big(g_{u,v}(n, n+m)\big)_{u \in V_{n-1}} = \sum_{u \in V_{n-1}} g_{u,v}(n, n+m) e^{(n)}(u) \qquad (3.8)$$

By definition, the simplex $\triangle_m^{(n)}$ is the convex hull of the vectors $\{g_v(n, n+m)\}_{v \in V_{n+m}}$.

Remark 3.2.2. We define the metric d^* by

$$d^*(x, y) = \sum_{i=1}^{k} |x_i - y_i| \text{ for } x = (x_1, \ldots, x_k)^T, y = (y_1, \ldots, y_k)^T \in \mathbb{R}^k.$$

We notice that d^* is equivalent to the euclidean norm $\| \, . \, \|$ on $\mathbb{R}^k$.

Proof. (*of Theorem 3.2.1.*)
First of all, we notice, that being uniquely ergodic means $\triangle_\infty^{(n)}$ is a singleton for all n.

($\Leftarrow$) It is sufficient to show that $\operatorname{diam}(\triangle_m^{(n)}) \overset{m \to \infty}{\longrightarrow} 0$. Let $v, v' \in V_{n+m}, u \in V_{n-1}$

and use (3.8) to calculate:

$$d^*\Big(g_v(n,n+m), g_{v'}(n,n+m)\Big) = \sum_{u\in V_{n-1}} \Big| g_{u,v}(n,n+m) - g_{u,v'}(n,n+m)\Big|$$

$$= \sum_{u\in V_{n-1}} \Bigg| \underbrace{\sum_{(v_1,\dots,v_m)\in V_n\times \dots V_{n+m-1}} f_{u,v_1}(n)\cdots f_{v_{m-1},v_m}(n+m-1)\cdot f_{v_m,v}(n+m)}_{\sum_{w\in V_{n+m-1}} g_{u,w}(n,n+m-1)\cdot f_{w,v}(n+m)}$$

$$- \underbrace{\sum_{(v'_1,\dots,v'_m)\in V_n\times \dots V_{n+m-1}} f_{u,v'_1}(n)\cdots f_{v'_{m-1},v'_m}(n+m-1)\cdot f_{v'_m,v'}(n+m)}_{\sum_{w\in V_{n+m-1}} g_{u,w}(n,n+m-1)\cdot f_{w,v'}(n+m)} \Bigg|$$

$$= \sum_{u\in V_{n-1}} \Bigg| \sum_{w\in V_{n+m-1}} g_{u,w}(n,n+m-1)\cdot \Big(f_{w,v}(n+m) - f_{w,v'}(n+m)\Big)\Bigg|$$

$$\le \sum_{w\in V_{n+m-1}} \Big| f_{w,v}(n+m) - f_{w,v'}(n+m)\Big| \underbrace{\sum_{u\in V_{n-1}} g_{u,w}(n,n+m-1)}_{=1 \text{ by definition of } g}$$

$$= d^*\Big(f_v(n+m) - f_{v'}(n+m)\Big) \overset{m\to\infty}{\longrightarrow} 0.$$

Thus we get $\operatorname{diam}(\triangle_m^{(n)}) \overset{m\to\infty}{\longrightarrow} 0$ and therefore $\triangle_\infty^{(n)}$ contains one single element. Hence B is uniquely ergodic.

($\Rightarrow$) Suppose B is uniquely ergodic, thus $\triangle_\infty^{(n)}$ is a singleton. It holds for $n\in\mathbb{N}$:

$$\max_{v,v'\in V_{n+m}} d^*\Big(g_v(n,n+m), g_{v'}(n,n+m)\Big) \le \operatorname{diam}(\triangle_m^{(n)}) \overset{m\to\infty}{\longrightarrow} 0,$$

by assumption. Hence we can find a convergent sequence $(\varepsilon_n)_{n\in\mathbb{N}} \overset{n\to\infty}{\longrightarrow} 0$ and for every n we find a m large enough such that

$$\max_{v,v'\in V_{n+m}} \sum_{u\in V_{n-1}} \Big| g_{u,v}(n,n+m) - g_{u,v'}(n,n+m)\Big| \le \varepsilon_n.$$

It follows that there are sequences $(n_i)_{i\ge1}$ and $(m_i)_{i\ge1}$ with entries in $\mathbb{N}$ such that

$$\max_{v,v'\in V_{n_i}} \sum_{u\in V_{n-1}} \Big| g_{u,v}(n_i,n_i+m_i) - g_{u,v'}(n_i,n_i+m_i)\Big| \le \varepsilon_{n_i} \tag{3.9}$$

with $1\le n_1 < n_1+m_1 = n_2 < n_2+m_2 = n_3 < n_3+m_3 = n_4 < \dots$

We have now an ordering on the levelset of our Bratteli diagram. If we telescope B with respect to this order $n_1 < n_2 < n_3 < n_4 < \dots$, we get a new Bratteli diagram $B' = (V', E')$. The corresponding stochastic incidence matrices are $G(n_i, n_i+m_i)$, $i\in\mathbb{N}$. Thus we get the result since it follows directly from (3.9). □

The authors in [4] showed the importance of finding an appropriate telescoping and the use of stochastic matrices by the following two examples:

Example 3.2.3. Suppose $B = (V, E)$ is a Bratteli diagram with incidence matrix $M(n) = \begin{pmatrix} n & 1 \\ 1 & n \end{pmatrix}$. For finding the corresponding stochastic incidence matrix $F(n)$, we observe that $M(n)$ has Equal Row Sums (short 'ERS') $r(n)$ (this property is introduced in [4]). Obviously the entries of $F(n)$ are obtained by $f_{v,w}(n) = \frac{m_{v,w}(n)}{r(n)}$. Hence

$$F(n) = \begin{pmatrix} 1 - \frac{1}{n+1} & \frac{1}{n+1} \\ \frac{1}{n+1} & 1 - \frac{1}{n+1} \end{pmatrix}.$$

Calculating the limit (3.7), to get a result about unique ergodicity, we get a limit converging to 2. At this moment B would not be uniquely ergodic.

Now suppose more generally, $F_n = \begin{pmatrix} a_n & b_n \\ b_n & a_n \end{pmatrix}$, a ERS-matrix. Similar to the proof of Theorem 3.2.1, we take $G(n, n+m) = \Big(g_{u,v}(n, n+m)\Big)$ and by induction,

$$S(n,m) := \sum_{u \in V_{n-1}} \Big| g_{u,v}(n, n+m) - g_{u,v'}(n, n+m) \Big| = 2 \prod_{i=0}^{m} \Big| (a_{n+i} - b_{n+i}) \Big|.$$

Applying this to B, we get $S(n,m) = 2 \prod_{i=0}^{m} \left(1 - \frac{2}{n+i+1}\right) \overset{m \to \infty}{\longrightarrow} 0$. To get the correct order of the telescoping, again as in the proof, construct for $\varepsilon_n \overset{n \to \infty}{\longrightarrow} 0$:

For ε_1, let $n = n_1$ then $\exists m_1$ such that $S(n_1, m_1) < \varepsilon_1$ and define $n_2 = n_1 + m_1$.
For ε_2, $\exists m_2$ such that $S(n_2, m_2) < \varepsilon_2$ and define $n_3 = n_2 + m_2$.

$\Longrightarrow$ Telescoping with respect to (n_k) gives that B is uniquely ergodic.

The same does not hold for diagrams where $S(n,m)$ does not converge to 0. Hence there is no telescoping such that (3.7) is satisfied. For example, take the Bratteli diagram $\bar{B} = (\bar{V}, \bar{V})$ with incidence matrix $\bar{M}(n) = \begin{pmatrix} n^2 & 1 \\ 1 & n^2 \end{pmatrix}$.

Example 3.2.4. Suppose C has incidence matrix $M(n) = \begin{pmatrix} 4 & 0 \\ 1 & 3 \end{pmatrix}$. Without using stochastic incidence matrices, we would get that C is not a uniquely ergodic matrix by calculating (3.7) for the entries of $M(n)$. Nevertheless, C is uniquely ergodic since for the corresponding stochastic incidence matrix $F(n) = \begin{pmatrix} \frac{4}{5} & 0 \\ \frac{1}{5} & 1 \end{pmatrix}$, (3.7) is satisfied.

If we now want to close the circle to our two prominent Kakutani- and Chacón transformations, we obtain the next to examples rather easy by applying Theorem 3.2.1:

Example 3.2.5. Looking back to the Kakutani transformation, we see, that we already had an example for a unique ergodic transformation. In Theorem Theorem 1.3.6, we proved, that the dyadic adding machine $\mathcal{K}$ is ergodic. Including what we know about Bratteli diagrams, $\mathcal{K}$ has an isomorphic Bratteli diagram $B = (V, E)$ with incidence matrix $M(n) = \begin{pmatrix} 2 \end{pmatrix}$. And since the stochastic version of $M(n)$ satisfies (3.7), we obtain that the dyadic adding machine is uniquely ergodic.

Example 3.2.6. As constructed in Section 1.3.2, Chacón's transformation is as well uniquely ergodic. We already have seen that this transformation is stationary with incidence matrix $M(n) = \begin{pmatrix} 3 & 0 \\ 1 & 1 \end{pmatrix}$ (compare to the left hand side of Figure 2.4). Hence we apply Theorem 3.2.1 and conclude that the stochastic version of $M(n)$ satisfies again (3.7).

3.3 Exact finite rank diagrams

Even more general, it is shown in [5] that every transformation of exact finite rank (see Definition 3.3.1) is uniquely ergodic. This fact is not only used for Bratteli diagrams and finite rank transformations, it is actually used in some other areas as well. Giving an example, it is used as a version of Boshernitzan's theorem [8] in the field of symbolic dynamics and recently to prove the uniform convergence in the multiplicative ergodic theorem (for both, one can read for example a paper by David Damanik and Daniel Lenz [13]), which in turn has applications in quantum mechanics as for example the spectral properties of Schrödingers operators. Following the Paper of S. Bezuglyi, J. Kwiatkowski, K. Medynets and B. Solomyak [5], we will give an introduction to what it means being an exact finite rank Bratteli diagram.

Definition 3.3.1. *Exact finite rank Bratteli diagram.*
A Bratteli diagram B has **exact finite rank** if there exists a finite invariant measure μ and a constant $\delta > 0$ such that, after appropriate telescoping,

$$\mu(X_v(n)) \geq \delta \text{ for all } n \in \mathbb{N} \text{ and } v \in V_n.$$

$X_v(n)$ denotes the set of all n-paths going through $v \in V_n$. In other words, exact finite rank adds the requirement of having only towers of positive measure to the definition of finite rank diagrams.

Recall the definition of finite rank diagrams, we can transfer them into a canonical block-triangular from, such that X_B is decompositioned into a finite number of tail-invariant subsets. The proof of the next Theorem can be found for example in [5, Theorem 2.6] or [4].

Theorem 3.3.2. *Let $B = (V, E)$ be a Bratteli diagram of finite rank. Then B can be isomorphically transformed to a diagram with incidence matrices $\{F(n)\}_{n\geq 1}$ such that*

$$F(n) = \begin{pmatrix} F_1(n) & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & F_2(n) & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \cdots & \cdots & \vdots \\ 0 & 0 & \cdots & F_s(n) & \cdots & \cdots & \vdots \\ X_{s+1,1}(n) & X_{s+1,2}(n) & \cdots & X_{s+1,s}(n) & F_{s+1}(n) & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \cdots & \cdots & \vdots \\ X_{m,1}(n) & X_{m,2}(n) & \cdots & X_{m,s}(n) & X_{m,s+1}(n) & \cdots & F_m(n) \end{pmatrix} \tag{3.10}$$

Each matrix $F_i(n),\ i = 1, \ldots, s,\ n \geq 1$ has strictly positive entries and each matrix $F_i(n),\ i = s+1, \ldots, m,\ n \geq 1$ has either strictly positive or zero entries. Also there is at least one matrix $X_{j,k}(n),\ k = 1, \ldots, j-1,\ j = s+1, \ldots, m-1$ which is non-zero.

In the next part, we want to show that we can derive ergodic measures from extending measures of simple, pairwise disjoint subdiagram. For this, we introduce Bratteli subdiagrams and give an overview how to extend measures from a subdiagram.

Definition 3.3.3. *Subdiagram of a Bratteli diagram.*
Let $B = (V, E)$ be a Bratteli diagram. Then $S = (W, R)$ is a **subdiagram** of B if it is constructed by taking a subset W of V and then considering all the edges of B, which connect the vertices in W. This set of edges, called R, is then a subset of E.

Let us mention here that a generated subdiagram is in general not invariant under the telescoping since one can add some edges to the subdiagram S which are not contained in the diagram B.

If we take the set Y, the set of all infinite paths of a certain subdiagram S of B, it is natural to say Y is a subset of X_B. Now we denote the set of paths $x \in X_B$ that are $\mathcal{E}$-equivalent to a path $y \in Y$ with X_S. Hence X_S is $\mathcal{E}$-invariant. We can also find an extension of the measure μ to X_S if we consider $\hat{\mu}$ on X_S such that $\hat{\mu}$ induced on Y coincides with μ. Obviously $\hat{\mu}$ is then $\mathcal{E}$-invariant as well. For extending the measure, we take a finite path $\bar{e} \in E(v_0, v)$ from the edge set that belongs to the subdiagram S and from the vertex v_0 to $v \in V_n$. Let us now define the cylinderset $[\bar{e}]_S$ to be the set of all paths in Y that coincide with $\bar{e}$ in the first n edges. Then we identify each finite path $\bar{e}'$ from the diagram B where $t(\bar{e}') = t(\bar{e})$ with the same measure, i.e. $\hat{\mu}([\bar{e}']) = \mu([\bar{e}]_S)$. Defining it like this, $\hat{\mu}$ is extended to the σ-algebra of Borel subsets of X_B generated by $[\bar{x}]$, where $\bar{x}$ is a finite path, ending in a vertex of S. The support of the measure $\hat{\mu}$ is the set of all paths which are cofinal to paths in Y. In fact we get a formula to compute $\hat{\mu}(X_S)$.

Define $W_n = W \cap V_n$ and $X_S(n) = \{x = (x_i) \in X_B \ : \ t(x_i) \in W_i, \ \forall i \geq n\}$. Then it holds

(1) $X_S(n) \subset X_S(n+1)$ and

(2) $\hat{\mu}(X_S) = \lim_{n\to\infty} \mu(X_S(n)) = \lim_{n\to\infty} \sum_{w\in W_n} \hat{h}_w(n)\mu\Big(\big[e_S(v_0, w)\big]\Big)$

where $\hat{h}_w(n)$ is the height of the tower $X_w(n)$ in the diagram B and $e_S(v_0, w)$ is the path in S which goes from v_0 to w. This measure $\hat{\mu}(X_S)$ may be either finite or infinite.

Let us clarify that from now on, finite rank Bratteli diagrams B have the structure of (3.10). We define a subset on $\{1, \dots, m\}$, denoted by Λ, such that the corresponding incidence matrices $F(n)$, $n \in \Lambda$ are non-zero. Then take $\alpha \in \Lambda$ and define B_α as the subdiagram of B which has incidence matrices $\{F_\alpha(n)\}_n$. Since $F_\alpha(n)$ is strictly positive we can imply that B_α is simple. Let the subdiagram B_α have a path space Y_α and $X_\alpha = \mathcal{E}(Y_\alpha)$. The next theorem describes the structure of ergodic invariant measures (with respect to some subdiagrams).

Theorem 3.3.4. *Let $B = (V, E)$ be a Bratteli diagram of finite rank m. Then it holds*

(1) Each finite ergodic measure on Y_α extends to an ergodic measure on X_α. The measure extension can be finite or infinite.

(2) Each ergodic measure (not necessarily finite) on X_α is obtained as an extension of a finite ergodic measure from Y_α.

(3) The number of ergodic measures is not greater than m.

(4) B can be telescoped in such a way, that for every probability ergodic measure μ there exists a subset W_μ of vertices from $\{1, \dots, m\}$ such that the support of μ consists of all infinite paths that eventually go along the vertices of W_μ only.

(4-i) Let ν be a second ergodic measure. Then $W_\mu \cap W_\nu = \emptyset$.

(4-ii) Let μ be a probability ergodic measure. Then there exists $\delta > 0$ such that for any $v \in W_\mu$ and any level n

$$\mu\Big(X_v(n)\Big) \geq \delta$$

for $X_v(n)$ the set of all paths going through the vertex v at level n.

(4-iii) *The subdiagram generated by W_μ is simple and uniquely ergodic. The unique ergodic measure on the path space of the subdiagram is precisely the restriction of the measure μ.*

Proof. Since we want to mention a nice consequence of the statements of part (4), we will refer for (1) and (2) to [5], Theorem 3.3 and for (3) to Proposition 3.1.3 from the beginning of this chapter. For proving part (4), we define the probability ergodic measures on X_B by $\mu_1, \ldots, \mu_m$. Without loss of generality, we may assume that each μ_i is restricted to a simple subdiagram B_{α_i}. In the following, we only construct W_1 for μ_1, since we find the corresponding sets $W_2, \ldots, W_m$ for $\mu_2, \ldots, \mu_m$ by repeating the same arguments as for W_1 and μ_1. For μ_1,

$$\sum_{v\in V} \limsup_{n\to\infty} \mu_1\Big(X_v(n)\Big) \geq \limsup_{n\to\infty} \sum_{v\in V} \mu_1\Big(X_v(n)\Big) = 1.$$

Thus, there exists a vertex v_1 such that

$$\limsup_{n\to\infty} \mu_1\Big(X_{v_1}(n)\Big) = \delta_1 > 0.$$

Then we can telescope the diagram such that $\mu_1\Big(X_{v_1}(n)\Big) > \frac{\delta_1}{2}$ for all levels n. Taking now a new vertex $V_2 \neq v_1$ it follows for $\delta_2 > 0$

$$\limsup_{n\to\infty} \mu_1\Big(X_{v_2}(n)\Big) = \delta_2 > 0.$$

Again we telescope the diagram such that $\mu_1\Big(X_{v_2}(n)\Big) > \frac{\delta_2}{2}$. Going along like this, we get a set of vertices W_1 such that

$$\mu_1\Big(X_v(n)\Big) > \delta > 0,\ \forall n \geq 0 \text{ and } v \in W_1.$$

Note that $\delta = \frac{1}{2}\min_i \delta_i$ and $\limsup_n \mu_1\Big(X_v(n)\Big) = 0$, for all $v \notin W_1$. Let $\biguplus$ define the disjoint union, we telescope the diagram once more, to get

$$\sum_{k=n}^{\infty} \mu_1\Big(\biguplus_{v\notin W_1} X_v(k)\Big) < \frac{1}{n},\ \text{for all } n.$$

If we denote the set of all paths, that eventually go along the vertices of W_μ only, by S_1, then S_1 is the support of μ_1 since for a set $R_1 := X_B \setminus S_1 = \bigcap_{n\geq 1}\bigcup_{k\geq n}\biguplus_{v\notin W_1} X_v(k)$,

$$\mu_1(R_1) = \lim_{n\to\infty} \mu_1\Big(\bigcup_{k\geq n}\biguplus_{v\notin W_1} X_v(k)\Big) \leq \lim_{n\to\infty}\sum_{k=n}^{\infty} \mu_1\Big(\biguplus_{v\notin W_1} X_v(k)\Big) = 0.$$

As said before, the sets $W_2, \dots, W_m$ will be constructed the same way. Now we claim that for two sets W_i, W_j with $i \neq j$, their intersection is empty. We prove by contradiction that if we have two probability ergodic measures μ and ν and a vertex w it does not hold

$$\mu\Big(X_w(n)\Big) \geq \gamma \text{ and } \nu\Big(X_w(n)\Big) \geq \gamma, \forall n \in \mathbb{N} \text{ and } \gamma = \frac{1}{2}\min\Big(\delta(\mu), \delta(\nu)\Big) > 0.$$

By ergodicity follows for a set $C = \bigcap_{k\geq 1}\bigcup_{n\geq k} X_w(n)$ that $\mu(C) = \nu(C) = 1$. Since by assumption μ and ν are distinct ergodic measures, we apply the Radon-Nikodym derivative which satisfies

$$\frac{d\mu}{d(\mu+\nu)}(x) \equiv 0 \text{ for } \nu\text{-a.e. } x \in X_B.$$

Then for $h_{v_n(x)}(n)$, the number of n-paths from vertex $v_n(x) \in V_n$, such that $x \in X_B$ goes through this vertex, to the top vertex, we calculate

$$\begin{aligned} 0 &= \frac{d\mu}{d(\mu+\nu)}(x) \\ &= \lim_{n\to\infty} \frac{\mu\big([x]_n\big)}{(\mu+\nu)\big([x]_n\big)} \\ &= \lim_{n\to\infty} \frac{h_{v_n(x)}(n)\mu\big([x]_n\big)}{h_{v_n(x)}(n)(\mu+\nu)\big([x]_n\big)} \\ &= \lim_{n:v_n(x)=w} \frac{\mu\big(X_w(n)\big)}{\mu\big(X_w(n)\big)+\nu\big(X_w(n)\big)} \geq \frac{\gamma}{2} \geq 0. \end{aligned}$$

Hence we get the contradiction and proved (4-i) and (4-ii).

For (4-iii) we denote the subdiagram generated by the vertices of W_μ by B_μ. Since B_μ is again a subdiagram of corresponding simple diagram B_{α_i}, we can telescope the original diagram such that for consecutive levels, there exists at least one edge between any pair of vertices. Thus B_μ is simple and by the same arguments as before, B_μ cannot admit any other probability ergodic measure. □

Example 3.3.5. Let us assume the Bratteli diagram B with incidence matrix

$$F(n) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 2 \end{pmatrix}.$$

Then the incidence matrix is of the form of (3.10), containing two subdiagrams B_1 and B_2 with incidence matrices $F_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ and $F_2 = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix}$ respectively.

For $\alpha \in \{1,2\}$, each B_α, has a path space Y_α and $X_\alpha = \mathcal{E}(Y_\alpha)$. On the path space Y_1, the subdiagram is not invariant since a path can possibly reach Y_1 from the right subdiagram Y_2. On the other side, we have invariance on the path space Y_2. This is illustrated in Figure 3.1 (dotted lines for Y_1 and dashed lines for Y_2). The measures of both subdiagrams are finite since for a certain vertex v in B_1 we can choose the measure of each path, ending in v in such a way, that they sum up to 1. Doing this in every step, it also holds for the limit. Analogously for B_2. Extending the measure from Y_1 gives an invariant ergodic measure on the path space X_B of the original diagram B.

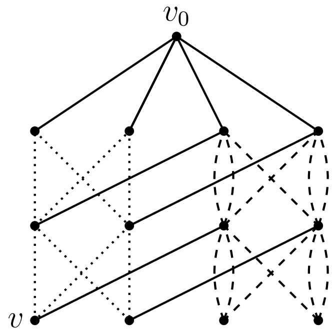


Figure 3.1: Bratteli diagram B with its subdiagrams Y_α

As a direct consequence of Theorem 3.3.4 (4-iii), we get a version of Boshernitzan's theorem [8].

Corollary 3.3.6. *A Bratteli diagram B of finite rank is uniquely ergodic if it is of exact finite rank.*

Note that we can apply the definition of exact finite rank only to Bratteli diagrams, not to the Bratteli-Vershik systems since there exists Bratteli diagrams with different Vershik-maps, one with exact finite rank and one without.

Example 3.3.7. The definition of exact finite rank diagrams, suggest that transformations constructed by the cutting and stacking procedure without adding spacers with corresponding Rokhlin towers which do not asymptotically vanish, i.e. the measure of ever tower is always bigger than some $a > 0$ are of this class. Hence for example the Kakutani transformation or Odometer is of exact finite rank.

Still the notion of exact finite rank is rather poor. Although it is a class of transformation which is automatically uniquely ergodic, the converse is not true. It also excludes strongly mixing. A. Rosenthal showed in 1984 that ergodic transformations of exacts finite rank are partially rigid which means by Definition 4.1.1 that they cannot satisfy the mixing property.

4 Construction of strong mixing rank one transformations

In this chapter, we are mainly refer to [1] a Paper of Adams, proving that many staircase transformations are strongly mixing. Note that in the reference literature, we will only read mixing instead of strongly mixing. Adams proved in his 1998's Paper the theorem which was originally stated by M. Smorodinsky. A little bit earlier, in 1992, Adams and N. Friedman [2] gave an algorithm including cutting-and-stacking constructions, with which (strongly) mixing rank one transformations can be produced. We are also following a Paper of D. Creutz and C. E. Silva [12], who showed mixing for polynomial staircase transformations.

Staircase transformations are a simple class of rank one transformations, in fact the simplest class which is strongly mixing. They are basically cutting-and-stacking transformations such that at each step n, the stack is cut into r_n pieces for a given sequence $\{r_n\}$ of non-negative integers. Before the stacking step, the spacers are put above each subcolumn (i.e. the resulting slices after cutting) increasingly by 1. While proving that Staircase transformations are mixing, Adams assumed that $\frac{r_n^2}{h_n}\longrightarrow 0$ for $n \to \infty$, where h_n is the height of the n-th stack. This result implies the finite measure-preserving property for the mixing staircase transformation and is called "restricted growth". Asking for the somehow other implication, Ryzhikov proved in [21] that all staircase transformations are mixing hence every finite measure-preserving is.

First, let us recall some ergodic properties:

Remark 4.0.1. (1) By Theorem 1.1.10 $\Longrightarrow$

$$\begin{aligned}\Big|\frac{1}{n}\sum_{k=0}^{n-1}\mu(T^{-k}A\cap B)-\mu(A)\mu(B)\Big| &= \Big|\frac{1}{n}\sum_{k=0}^{n-1}\mu(A\cap T^{k}B)-\mu(A)\mu(B)\Big|\\ &= \int_A\Big|\frac{1}{n}\sum_{k=0}^{n-1}\chi_B(T^{-k}x)-\mu(B)\Big|d\mu\\ &= \Big|\Big|\frac{1}{n}\sum_{k=0}^{n-1}\chi_B(T^{-k}x)-\mu(B)\Big|\Big|_1\longrightarrow 0\end{aligned}$$

by ergodicity of T.
(2) An **ergodic sequence** $\{a_n\}$ is defined by

$$\lim_{n\to\infty}\int\Big|\frac{1}{n}\sum_{j=0}^{n-1}\chi_B(T^{-a_j}x)-\mu(B)\Big|d\mu=0\quad\forall B\in\mathcal{B},$$

where $\mathcal{B}$ is the Borel σ- algebra.
(3) A transformation is **power ergodic** if all powers of the transformation are ergodic, i.e. the ergodic averages $\mathbb{A}_n$ converge uniformly. In other words if $\forall B \in \mathcal{B}$:

$$\lim_{n\to\infty} \sup_{l\in\mathbb{N}} \int \left| \frac{1}{n} \sum_{k=0}^{n-1} \chi_B(T^{-kl}x) - \mu(B) \right| d\mu = 0.$$

This notion comes from the property of being uniform Cesàro (as used in [1] and [2]).

Talking about mixing sequences as well as ergodicity on sequences, we should mention the theorem of Blum-Hanson [7].

Theorem 4.0.2. *The Blum-Hansen theorem.*
Assume T to be an ergodic transformation. T is mixing if and only if every strictly increasing sequence of integers is ergodic with respect to T.

4.1 Staircase transformations

Using the notation from [1], we define (and illustrate in Figure 4.1) for the staircase transformation T:

- $\{s_{n,j}\}_{\{r_n\}}$ the **spacer sequence**, where n and j describe the step n and subcolumn j respectively. The classical staircase transformation is defined by the spacer sequence $s_{n,j} = j$ for $0 \leq j \leq r_n - 1$.

- $\{r_n\}$ the **cut sequence**. As in [1], we may assume from now on that $\lim r_n = \infty$, since otherwise the transformation automatically cannot be mixing. This is because of the partial rigidity of a transformation satisfying $\liminf r_n < \infty$:

  Definition 4.1.1. *Partially rigid transformations.*
  A transformation T is **partially rigid**, if there exists a constant $c > 0$ such that
  $$\limsup_{n\to\infty} \mu(T^n A \cap A) \geq c\mu(A)$$
  for all sets A of finite measure.

  Every transformation with $\liminf r_n < \infty$ is partially rigid and hence cannot be mixing.

- $\{h_n\}$ the **height sequence**, corresponding to the number of levels in each stack: $h_0 = 1,\ h_{n+1} = r_n h_n + \sum_{j=0}^{r_n-1} s_{n,j}$.

- $I_{n,i}$ the i-th level in the n-th stack with $I_{n,0}$ the bottom level and

$$C_n = \bigcup_{i=0}^{h_n-1} I_{n,i}.$$

- S_n the set of spacers added in the n-th step, i.e.

$$S_n = C_{n+1} \setminus C_n,\ S_n = \bigcup_{j=0}^{r_n-1} s_{n,j}.$$

- $I_{n,i}^{[j]}$ the j-th sublevel of the i-th level of the n-th stack.

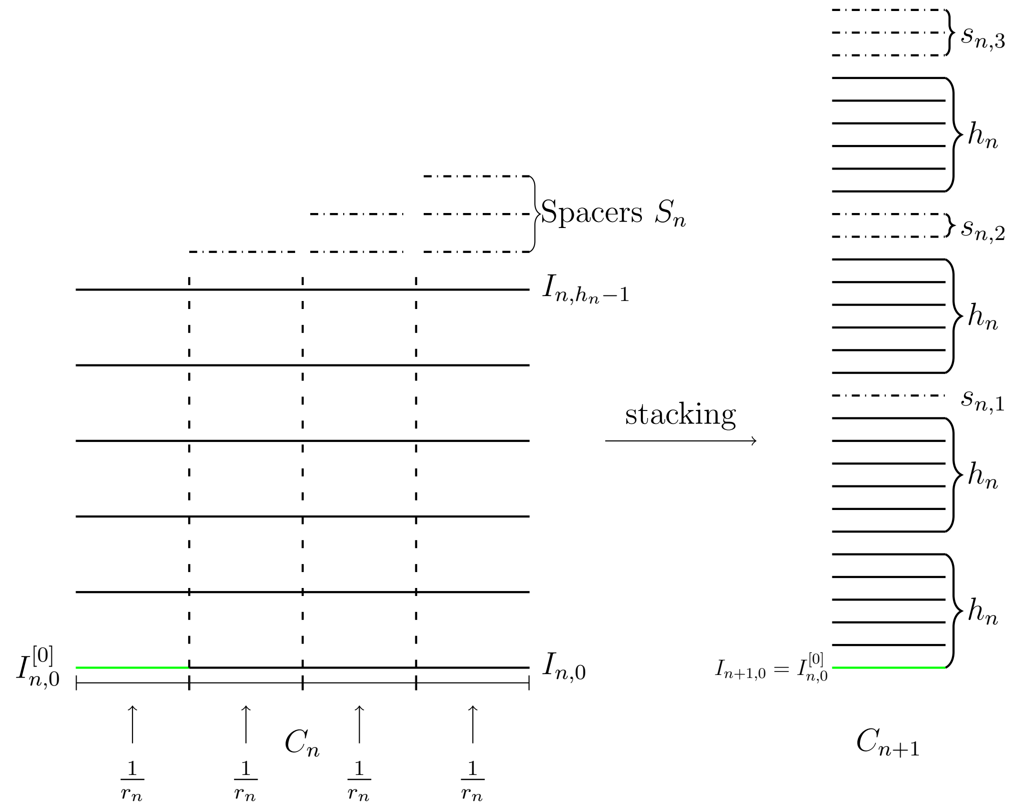


Figure 4.1: One construction step of a staircase transformation

Linking this to what we have seen in the previous chapters, we note that T is only defined on a finite measure space if and only if $\sum_{n=0}^{\infty} \mu(S_n) < \infty$. As we working on the interval $[0, 1)$ in general, we need this assumption to guarantee that T is isomorphic to the transformation defined on the unit interval constructed through cutting-and-stacking, with the only difference that $C_0 = [0, \frac{1}{L})$ with L the measure of the domain of the original T.

4.2 Staircase transformations are strongly mixing

For the moment, we investigate the idea of mixing by looking at the construction of the staircase transformation. Suppose we have a staircase transformation T in the known cutting-and-stacking fashion presented above. Let us take a level I of the stack C_n which is $i \geq r_n$ above the bottom level. If we now apply h_n times our transformation to the j-th sublevel of I, say $I_{n,i}^{[j]}$, it will be mapped to the next subcolumn $j+1$ and j levels below the $j+1$-st sublevel of I. This is because of the j added spacers above the subcolumn j (see Figure 4.2).

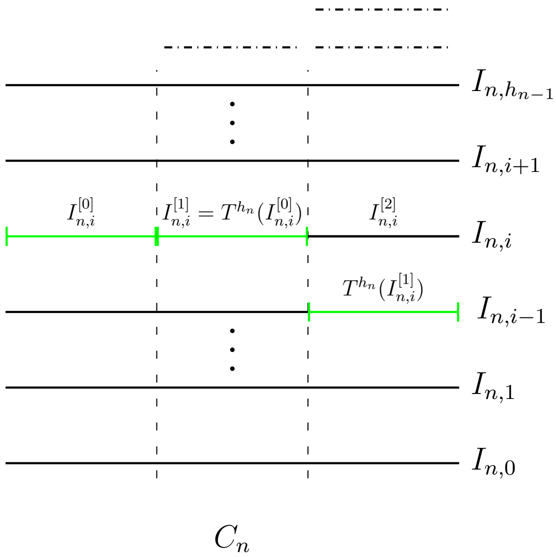


Figure 4.2: Idea of mixing on staircase transformations

Each length of such $I_{n,i}^{[j]}$ is exact $\frac{1}{r_n}$ and I contains r_n sublevels. Hence T^{h_n} maps the level I to a sequence of r_n succeeding smaller levels of length $\frac{1}{r_n}$, which means that the characteristic function of $T^{h_n}(I)$ is of the form

$$\frac{1}{r_n} \sum_{k=0}^{r_n-1} \Big(\chi_I(T^{-k}x) - \mu(I) \Big). \tag{4.1}$$

By ergodicity of T this term converges to 0, thus the sequence $\{h_n\}$ will be mixing. Making this whole construction more concrete, we obtain the next lemma.

Lemma 4.2.1. *Let $n \in \mathbb{N}$, $i \in \{0, 1, \ldots, h_n\}$, $j \in \{0, 1, \ldots, r_n\}$ where h_n and r_n are positive integers and B a union of levels in C_n. Then*

(a) $I_{n,i} = \bigcup_{j=0}^{r_n-1} I_{n,i}^{[j]}$,

(b) $T^{kh_n+jk+\frac{1}{2}k(k-1)}I_{n,i}^{[j]} = I_{n,i}^{[j+k]}$ *and*

(c) $\mu(I_{n,i}^{[j]} \cap B) = \frac{1}{r_n}\mu(I_{n,i} \cap B)$.

Proof. (a) is obvious by the construction, see Figure 4.2.
(b) follows by applying k times $T^{h_n+j}I_{n,i}^{[j]} = I_{n,i}^{[j+1]}$, hence by adding the new spacers in each step.
(c) Since B is a union of levels in C_n, it follows that $I_{n,i} \subseteq B$ or $I_{n,i} \cap B = \emptyset$. □

Our next goal is to show that the ergodic sum in (4.1) actually converges to zero. The next Lemma, called the Block Lemma [12], shows that an ergodic type average as in (4.1) is dominated by an ergodic type average with proportional fewer terms.

Lemma 4.2.2. *Block Lemma.*
Let T be a measure-preserving transformation and $B \in \mathcal{B}$. Then for any positive integers R, L and q,

$$\int \Big| \frac{1}{R} \sum_{i=0}^{R-1} \chi_B(T^{-i}x) - \mu(B) \Big| d\mu(x) \leq \int \Big| \frac{1}{L} \sum_{i=0}^{L-1} \chi_B(T^{-iq}x) - \mu(B) \Big| d\mu(x) + \frac{qL}{R}.$$

Proof. This is a basic statement about measure-preserving transformations. The name of the Lemma comes from the idea to split the sum into blocks of size $L \cdot p$ and each of these blocks again into L blocks. Applying the measure-preserving property, we can combine the terms. □

Now we set up a bounded sequence ρ_n and prove in the next Proposition, that ρ_n is a mixing sequence.

Proposition 4.2.3. *[1] Suppose T is a staircase transformation and let B be a union of levels in some stack C_n. Then for all sequences ρ_n of positive integers there exists a sequence l_n of positive integers, both converging to infinity with $h_n \leq \rho_n \leq 2h_n$, such that*

$$\lim_{n\to\infty} \int \Big| \frac{1}{l_n} \sum_{i=0}^{l_n-1} \chi_B(T^{-i\rho_n}x) - \mu(B) \Big| d\mu(x) = 0.$$

Before we proof this Proposition, we shall need the next Lemma by Blum-Hansen [7].

Lemma 4.2.4. *Let $\{C_{i,j},\ i,j \geq 1\}$ be a bounded sequence of numbers such that $\lim_{|i-j|\to\infty} C_{i,j} = 0$. Then*

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{i,j=1}^{n} C_{i,j} = 0.$$

Proof. We define the supremum of the sequence $C_{i,j}$ by $C = \sup_{i,j} |C_{i,j}|$. For $\varepsilon > 0$, we take $M \in \mathbb{N}$ such that $|C_{i,j}| \leq \varepsilon$ for $|i-j| > M$. Then for $n > M$ it holds

$$\frac{1}{n^2}\sum_{i,j=1}^{n} |C_{i,j}| = \frac{1}{n^2}\sum_{i,j=1,|i-j|\leq M} |C_{i,j}| + \underbrace{\frac{1}{n^2}\sum_{i,j=1,|i-j|>M} |C_{i,j}|}_{<\varepsilon}$$

$$\leq \frac{(2M+1)nC}{n^2} + \varepsilon.$$

If we now choose $N > \frac{(2M+1)C}{\varepsilon}$, it follows for $n > N$,

$$\frac{1}{n^2}\sum_{i,j=1}^{n} |C_{i,j}| \leq 2\varepsilon.$$

□

Proof. (of Proposition 4.2.3.) Let $\varepsilon > 0$. Take $j \in \mathbb{N}$ such that

$$i\rho_n = jh_n + t,\ 0 \leq t \leq h_n \text{ and } 1 \leq j \leq r_n.$$

Then define D_1 to be the union of the top t levels of C_n and D_2 be the union of the lower $h_n - t$ levels. Then obviously $C_n = D_1 \uplus D_2$.

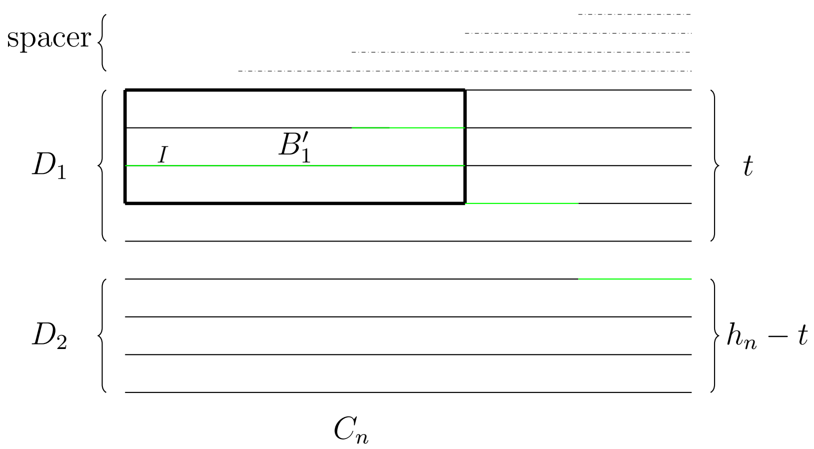


Figure 4.3: Partition of C_n into D_1 and D_2

Define further:
(i) $B_1 = B \cap D_1$ and $B_2 = B \cap D_2$ where B is by assumption a union of levels in some stack.
(ii) $B_1' =$

$$B_1 \setminus \Big\{\{\text{bottom } (j+1)\cdot r_n \text{ levels of } D_1\} \cup \{(j+1) \text{ rightmost subcolumns of } C_n\}\Big\}.$$

Then the symmetric difference of B_1 and B_1' satisfy

$$\mu(B_1' \triangle B_1) \leq \frac{(j+1)r_n}{h_n} + \frac{j+1}{r_n} \overset{n\to\infty}{\longrightarrow} 0.$$

(iii) Let I be a level in B_1', which is only a $\frac{r_n-j-1}{r_n}$ part of a level in C_n.

Then by applying $T^{i\rho_n}$ to I, I will be mapped $(j+1)$ times through the staircase on C_n by the construction of B_1' in (ii). Hence $T^{i\rho_n}I$ is intersecting with $(r_n-(j+1))$ levels of C_n. Even more, we see that $T^{i\rho_n}I$ intersects the levels of C_n in an arithmetic progression $\{k+i(j+1)\}_{i=0}^{r_n-j-2}$ as we see in Figure 4.3.

Now suppose that I^* is the highest of such levels and define $B_1^* = \bigcup_{I'\subset B_1'} I^*$. It follows that

$$\mu(T^{i\rho_n}B_1' \cap B) = \frac{1}{r_n}\sum_{k=0}^{r_n-j-2} \mu(T^{-k(j+1)}B_1^* \cap B). \tag{4.2}$$

Analogously we define $B_2' \subset B_2$ and B_2^* such that

$$\mu(T^{i\rho_n}B_2' \cap B) = \frac{1}{r_n}\sum_{k=0}^{r_n-j-2} \mu(T^{-kj}B_2^* \cap B). \tag{4.3}$$

By ergodicity of T^{-j} and $T^{-(j+1)}$, we can apply Proposition 1.1.11 to get

$$\mu(T^{i\rho_n}B \cap B) \longrightarrow \mu(B)^2 \text{ for } n \to \infty.$$

Therefore for different $i \neq j$, we combine (4.2) and (4.3) (this is okay since B_1' and B_2' are disjoint), and get

$$\mu(T^{i\rho_n}B \cap T^{j\rho_n}B) \longrightarrow \mu(B)^2 \text{ for } n \to \infty.$$

Now by taking an integer L and applying Cauchy-Schwarz and T-invariance of μ, we get

$$\begin{aligned}
&\Big(\int \Big|\frac{1}{L}\sum_{i=0}^{L-1} \chi_B(T^{-i\rho_n}x) - \mu(B)\Big| d\mu\Big)^2 \\
&\leq \int \Big|\frac{1}{L}\sum_{i=0}^{L-1} \chi_B(T^{-i\rho_n}x) - \mu(B)\Big|^2 d\mu \\
&= \int \Big|\frac{1}{L^2}\sum_{i,j=0}^{L-1} \big(\chi_B(T^{-i\rho_n}x) - \mu(B)\big)\big(\chi_B(T^{-j\rho_n}x) - \mu(B)\big)\Big| d\mu \\
&= \int \Big|\frac{1}{L^2}\sum_{i,j=0}^{L-1} \chi_B(T^{-i\rho_n}x)\chi_B(T^{-j\rho_n}x) - \mu(B)^2\Big| d\mu.
\end{aligned}$$

For the last line, we can now apply Lemma 4.2.4 if we identify

$$C_{i,j} = \chi_B(T^{-i\rho_n}x)\chi_B(T^{-j\rho_n}x) - \mu(B)^2.$$

Thus for $\varepsilon > 0$ and $\rho_n \to \infty$ we can take L such that

$$\int \Big| \frac{1}{L} \sum_{i=0}^{L-1} \chi_B(T^{-i\rho_n}x) - \mu(B) \Big| d\mu(x) < \varepsilon.$$

Hence for $\rho_n \to \infty$ fast enough, there exists $l_n \to \infty$ such that

$$\limsup_{n\to\infty} \int \Big| \frac{1}{l_n} \sum_{i=0}^{l_n-1} \chi_B(T^{-i\rho_n}x) - \mu(B) \Big| d\mu(x) < \varepsilon.$$

□

As we mentioned in the beginning of this chapter, Adams observed that for any staircase transformation, $\frac{r_n^2}{h_n} \longrightarrow 0$ if $n \to \infty$. Before, we are observing that

$$\frac{h_{n-1}^2}{h_n} = \frac{r_{n-1}h_{n-1}^2}{r_{n-1}h_n} = \underbrace{\frac{r_{n-1}h_{n-1}}{h_n}}_{\longrightarrow 1} \cdot \underbrace{\frac{h_{n-1}}{r_{n-1}}}_{\longrightarrow \infty} \longrightarrow \infty.$$

Now we are ready to prove the main theorem (see [1, Theorem 2.3]):

Theorem 4.2.5. *Let r_n be the cutting sequence such that $\lim r_n = \infty$. Then the staircase transformation T is mixing if $\lim_{n\to\infty} \frac{r_n^2}{h_n} = 0$, where h_n is the height sequence of each stack C_n.*

Proof. Remembering the construction once more, we start with a single interval C_0 and a certain number of spacer. For the induction step, we assume C_n to be the tower of height h_n. Then we cut C_n into r_n subcolumns and place $i-1$ layers of spacers at each subcolumn i. To get the new tower C_{n+1} we stack all the subcolumns i onto $i-1$ for $2 \leq i \leq r_n$. We obtain for the height of tower C_{n+1}

$$h_{n+1} = r_n h_n + \frac{1}{2} r_n(r_n - 1).$$

The "restricted growth" condition $\frac{r_n^2}{h_n} \to 0$ implies that

$$\frac{1}{2} r_n(r_n - 1) < \frac{r_n}{10} \text{ for large enough } n.$$

In the limit, we get a measure-preserving and invertible (up to sets of measure zero) transformation $T : X \to X$.

Suppose A, B are subsets of X. Then we have to show that

$$\lim_{m\to\infty} |\mu(T^m(A)\cap B) - \mu(A)\mu(B)| \to 0. \tag{4.4}$$

For $\varepsilon > 0$, we can take n large enough and A', B' union of full levels of C_{n_0} such that $\mu(A \triangle A') < \varepsilon$ and $\mu(B \triangle B') < \varepsilon$. Hence if (4.4) holds for A', B' then it holds for A, B. Therefore we can assume without loss of generality, that A, B are two unions of full levels of C_n for every $n \geq n_0$.

Similar to what we have done in the last proof, we choose a positive integers m such that

$$h_n \leq m = k_n h_n + t_n < h_{n+1} \text{ where } 1 \leq k_n \leq r_n \text{ and } 0 \leq t_n < h_n.$$

Again we can assume that m is large enough to satisfy $n \geq n_0$.

We will split our stack C_n in three disjoint parts and investigate the mixing property for each part independently. The parts are chosen in the following way (see Figure 4.4 with $r_n = 12,\ k_n = 3$ and $t_n = 4$): First cut C_n in r_n subcolumns, each subcolumn has length $\frac{1}{r_n}$. Then define

(i) D_1 to be the set of the $(k_n + 1)$ rightmost subcolumns of C_n,

(ii) $D_2 = \{\text{top } t_n \text{ levels of } C_n\} \setminus D_1$ and

(iii) $D_3 = \{\text{bottom } h_n - t_n \text{ levels of } C_n\} \setminus D_1$.

Obviously C_n is partitioned into three disjoint sets.

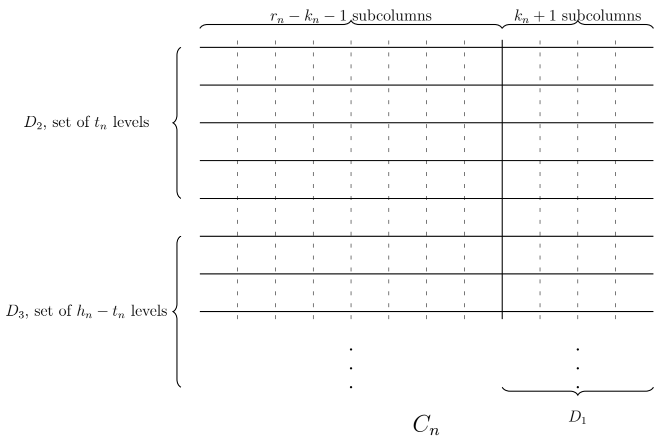


Figure 4.4: Partition of C_n with $r_n = 12,\ k_n = 3$ and $t_n = 4$

Part 1 (Mixing on D_1): By placing the layers of spacer above the tower C_n, D_1 occupies the top $(k_n + 1)h_n + (r_n - 1) + (r_n - 2) + \cdots + (r_n - k_n - 1) =$

$(k_n + 1)h_n + \frac{1}{2}(k_n + 1)(2r_n - k_n - 2)$ levels of C_{n+1}. Define further a subset of C_{n+1} by

$$\bar{D}_1 = D_1 \setminus \Big(\{\text{bottom } h_n + r_n \text{ levels of } C_{n+1}\} \cup \{\text{right-most subcolumn of } C_{n+1}\} \Big).$$

Let $\bar{A}_1 = A \cap \bar{D}_1$. Then $\bar{A}_1$ consists of levels of $C_{n+1} \setminus \{\text{right-most subcolumn of } C_{n+1}\}$ and $\mu((A \cap D_1) \triangle \bar{A}_1) < \frac{1}{r_{n+1}} \to 0$ as $n \to \infty$. Assume I to be one of those levels in $\bar{A}_1$. If we apply T^m, I gets pushed through the roof of the stack C_{n+1}, where it splits up into $r_{n+1} - 1$ intervals that intersect $r_{n+1} - 1$ consecutive levels of C_{n+1} (for illustration see Figure 4.5, the gray part in the right sketch). Each of this intersections has length $\frac{\mu(I)}{r_{n+1}-1}$.

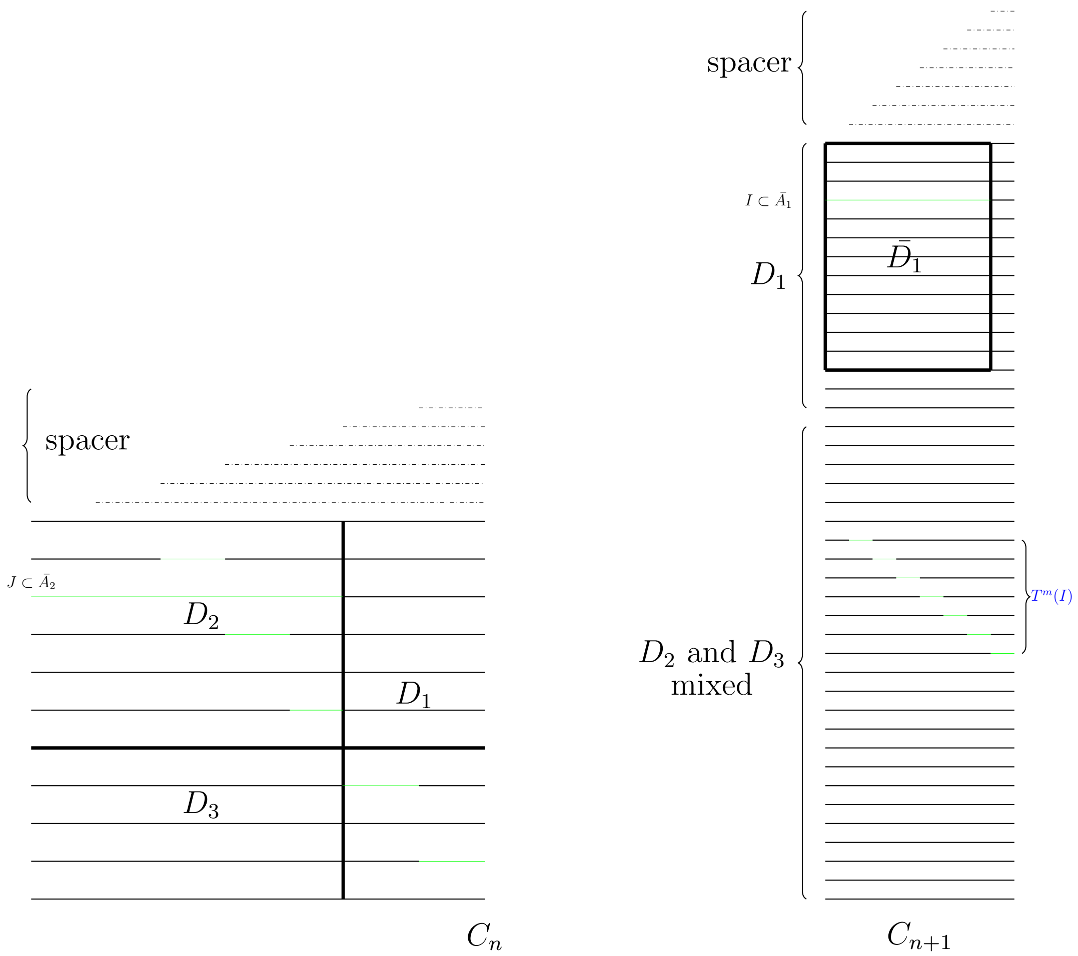


Figure 4.5: Staircase-construction for C_n (left) and C_{n+1} (right)

Now we define I^* to be the highest level which $T^m(I)$ intersects. Then

$$\mu(T^m(I) \cap B) = \frac{1}{r_{n+1} - 1} \sum_{i=0}^{r_{n+1}-2} \mu(T^{-i}(I^* \cap B),$$

and $\mu(I^*) = \frac{r_{n+1}}{r_{n+1}-1}\mu(I)$. Let $A_1^* = \bigcup_{I \subset \bar{A}_1} I^*$. We obtain by summing over all $I \subset \bar{A}_1$:

$$|\mu(T^m(\bar{A}_1) \cap B) - \mu(\bar{A}_1)\mu(B)| =$$
$$\Big| \underbrace{\frac{1}{r_{n+1}-1} \sum_{i=0}^{r_{n+1}-2} \mu(T^{-i}(A_1^*) \cap B) - \mu(A_1^*)\mu(B)}_{\to 0 \text{ by ergodicity (see Proposition 1.1.11).}} + \underbrace{\frac{1}{r_{n+1}-1}\mu(A_1^*)\mu(B)}_{\to 0 \text{ for } n \to \infty} \Big| \to 0$$

This proves (4.4) for the set $A_1 \subset A$.

Part 2 (Mixing on D_2): Define

$$\bar{A}_2 = A \cap D_2 \setminus \{\text{bottom } r_n^2 \text{ levels of } D_2 \text{ in } C_n\}.$$

It follows that $\mu((A \cap D_2) \triangle \bar{A}_2) \leq \frac{r_n^2}{h_n} \to 0$ for $m \to \infty$ (and hence $n \to \infty$ by definition of m). Let J be a level in $\bar{A}_2$. Obviously, J has length $\frac{r_n - k_n - 1}{r_n}$ in C_n. Applying T^m to J, it pushes it through the roof $k_n + 1$ times, dividing into $r_n - k_n - 1$ smaller intervals which intersect $r_n - k_n - 1$ levels on C_n. As we see in the left part of Figure 4.5, the intervals, shown in blue, intersect in an arithmetic progression $\{j + i(k_n+1)\}_{i=0}^{r_n - k_n - 2}$. As for Part 1, define J^* to be the highest level which $T^m(J)$ intersects. Then

$$\mu(T^m(J) \cap B) = \frac{1}{r_n - k_n - 1} \sum_{i=0}^{r_n - k_n - 2} \mu(T^{-i(k_n+1)}(J^*) \cap B),$$

and $\mu(\bar{A}_2) = \frac{r_n - k_n - 1}{r_n}\mu(A_2^*)$, for $A_2^* = \bigcup_{J \subset \bar{A}_2} J^*$. From this, we can assume that $\frac{k_n+1}{r_n} < 1 - \varepsilon$, since otherwise $|\mu(T^m(\bar{A}_2) \cap B) - \mu(\bar{A}_2)\mu(B)| < \varepsilon$ immediately. Summing up over all $J \subset \bar{A}_2$ gives

$$|\mu(T^m(\bar{A}_2) \cap B) - \mu(\bar{A}_2)\mu(B)| =$$
$$\frac{r_n - k_n - 1}{r_n} \cdot \Big| \frac{1}{r_n - k_n - 1} \sum_{i=0}^{r_n - k_n - 2} \mu(T^{-i(k_n+1)}(A_2^*) \cap B) - \mu(A_2^*)\mu(B) \Big|.$$

By assumption, T is measure-preserving and invertible (mod μ). Hence we can estimate the factor in the absolute value bars in the following way:

$$\begin{aligned}
&\frac{1}{r_n-k_n-1}\sum_{i=0}^{r_n-k_n-2}\mu(T^{-i(k_n+1)}(A_2^*)\cap B)-\mu(A_2^*)\mu(B)\\
&=\frac{1}{r_n-k_n-1}\sum_{i=0}^{r_n-k_n-2}\mu(A_2^*\cap T^{i(k_n+1)}(B))-\mu(A_2^*)\mu(B)\\
&=\frac{1}{r_n-k_n-1}\sum_{i=0}^{r_n-k_n-2}\int_{A_2^*}\mathbb{1}_B\circ T^{-i(k_n+1)}-\mu(B)d\mu\\
&\leq\int\Big|\frac{1}{r_n-k_n-1}\sum_{i=0}^{r_n-k_n-2}\mathbb{1}_B\circ T^{-i(k_n+1)}-\mu(B)\Big|d\mu. \qquad (4.5)
\end{aligned}$$

If we are able to show that (4.5) tends to 0 as $n\to\infty$, we get the required result for D_2.

To show this, we take $p\in\mathbb{N}$ such that $h_{p-1}\leq k_n+1<h_p$. Then we have

$$\begin{aligned}
\frac{(r_n-k_n-1)(k_n+1)}{h_p}&\geq\frac{\varepsilon r_n(k_n+1)}{h_p}\geq\varepsilon\frac{h_{p-1}^2}{h_p}\\
&=\varepsilon\frac{r_{p-1}h_{p-1}}{h_p}\frac{h_{p-1}}{r_{p-1}}\to\infty.
\end{aligned}$$

Now we choose $k'_n\geq 1$ minimal such that $k'_n(k_n+1)\geq h_p$. Then clearly $\frac{r_n-k_n-1}{k'_n}\to\infty$ as well. Applying Proposition 4.2.3 we choose a sequence $l_n\to\infty$ in such a way, that $a_n:=\lfloor\frac{r_n-k_n-1}{l_nk'_n}\rfloor\to\infty$, but

$$\int\Big|\frac{1}{l_n}\sum_{i=0}^{l_n-1}\underbrace{\mathbb{1}_B\circ T^{-ik'_n(k_n+1)}-\mu(B)}_{=:g\circ S^{-ik'_n}}\Big|d\mu\to 0\text{ for }n\to\infty. \qquad (4.6)$$

For simplicity, we abbreviate S by T^{k_n+1} and g by $\mathbb{1}_B-\mu(B)$. By S-invariance of μ,

$$\int\Big|\frac{1}{l_n}\sum_{i=0}^{l_n-1}g\circ S^{-ik'_n}\Big|d\mu=\int\Big|\frac{1}{l_n}\sum_{i=0}^{l_n-1}g\circ S^{-ik'_n+j}\Big|d\mu$$

for $0\leq j<k'_n$. Taking the average of this expression over j, we get

$$\begin{aligned}
\int\Big|\frac{1}{l_n}\sum_{i=0}^{l_n-1}g\circ S^{-ik'_n}\Big|d\mu&=\int\Big|\frac{1}{k'_n}\sum_{j=0}^{k'_n-1}\frac{1}{l_n}\sum_{i=0}^{l_n-1}g\circ S^{-ik'_n+j}\Big|d\mu\\
&=\int\Big|\frac{1}{k'_nl_n}\sum_{i=0}^{k'_nl_n-1}g\circ S^{-i}\Big|d\mu. \qquad (4.7)
\end{aligned}$$

Remembering how we defined the sequence a_n, we have $r_n - k_n - 1 = a_n k'_n l_n + b_n$ where b_n is a sequence of integers such that $0 \leq b_n < k'_n l_n$. Therefore

$$\begin{aligned}
(4.5) &= \int \Big| \frac{1}{r_n - k_n - 1} \sum_{i=0}^{r_n - k_n - 2} g \circ S^{-i} \Big| d\mu = \int \Big| \frac{1}{a_n k'_n l_n + b_n} \sum_{i=0}^{a_n k'_n l_n + b_n - 1} g \circ S^{-i} \Big| d\mu \\
&= \int \Big| \frac{a_n k'_n l_n}{a_n k'_n l_n + b_n} \frac{1}{a_n k'_n l_n} \Big(\sum_{i=0}^{a_n k'_n l_n - 1} g \circ S^{-i} + \sum_{i=a_n k'_n l_n}^{a_n k'_n l_n + b_n - 1} g \circ S^{-i} \Big) \Big| d\mu \\
&\leq \int \Big| \frac{1}{a_n k'_n l_n} \sum_{i=0}^{a_n k'_n l_n} g \circ S^{-i} \Big| d\mu + \frac{2 b_n \parallel g \parallel_\infty}{a_n k'_n l_n + b_n} \\
&\leq \int \Big| \frac{1}{a_n} \sum_{j=0}^{a_n - 1} \frac{1}{k'_n l_n} \sum_{i=0}^{k'_n l_n - 1} g \circ S^{-(i + j k'_n l_n)} \Big| d\mu + \frac{2 \parallel g \parallel_\infty}{a_n} \\
&\leq \int \Big| \frac{1}{k'_n l_n} \sum_{i=0}^{k'_n l_n - 1} g \circ S^{-i} \Big| d\mu + \frac{2 \parallel g \parallel_\infty}{a_n}.
\end{aligned}$$

The last inequality uses the other direction of (4.7). Hence by combining the arguments from (4.6) and (4.7), we can conclude the convergence for (4.5), in other words, (4.5) $\to 0$ for $n \to \infty$ and we proved the mixing property on D_2.

Part 3 (Mixing on D_3): Here we can again use the same argumentation as in Part 2 if we adjust in such a way that $T^m(J)$ goes through the roof only k_n times.

□

Bibliography

[1] T. M. Adams. "Smorodinsky's Conjecture on Rank-One Mixings". In: *Proceedings of the American Mathematical Society* (1998), pp. 739–744.

[2] T. M. Adams and N. Friedman. *Staircase Mixing*. Preprint. 1992.

[3] S. Bezuglyi, A. Dooley, and K. Medynets. "The Rokhlin lemma for homeomorphisms of a Cantor set". In: *Proceedings of the American Mathematical Society* 133 (Oct. 2004).

[4] S. Bezuglyi, O. Karpel, and J. Kwiatkowski. "Exact number of ergodic invariant measures for Bratteli diagrams". In: *Ergodic Theory and Dynamical Systems* (Feb. 2015).

[5] S. Bezuglyi et al. "Finite Rank Bratteli Diagrams: Structure of Invariant Measures". In: *Transactions of the American Mathematical Society* 365.5 (2013), pp. 2637–2679.

[6] S. Bezuglyi et al. "Invariant measures on stationary Bratteli diagrams". In: *Ergodic Theory and Dynamical Systems* 30(4) (2010), pp. 973–1007.

[7] J. R. Blum and D. L. Hansen. "On the mean ergodic theorem for subsequences". In: *Bull. Amer. Math. Soc. 66* (1960), pp. 308–311.

[8] M. Boshernitzan. "A condition of unique ergodicity of minimal symbolic flows". In: *Ergodic Theory and Dynamical Systems* 12 (1992), pp. 425–428.

[9] O. Bratteli. "Inductive limits of finite-dimensional C*-algebras". In: *Transactions of the American Mathematical Society* 171 (1972), pp. 195–234.

[10] R. V. Chacón. "Weakly mixing transformations which are not strongly mixing". In: *Proceedings of the American Mathematical Society* 22 (Jan. 1969), pp. 559–562.

[11] W. Chen. *The Notion of Mixing and Rank One Examples*. 2015. URL: `http://math.uchicago.edu/~may/REU2015/REUPapers/Chen,Wenyu.pdf`.

[12] D. Creutz and C. E. Silva. "Mixing on a class of rank-one transformations". In: *Ergodic Theory and Dynamical Systems* (2004), pp. 407–440.

[13] D. Damanik and D. Lenz. "A Condition of Boshernitzan and Uniform Convergence in the Multiplicative Ergodic Theorem". In: *Duke Math. J.* 133 (2006), pp. 95–123.

[14] F. Durand, B. Host, and C. F. Skau. "Substitutional dynamical systems, Bratteli diagramsand dimension groups". In: *Ergodic Theory and Dynamical Systems, Cambridge University Press* (1999), pp. 953–993.

[15] R. H. Herman, I. F. Putnam, and C. F. Skau. "Orderes Bratteli diagrams, Dimension Groups and topological Dynamics". In: *International Journal of Mathematics* (1992), pp. 827–864.

[16] S. Kakutani. "Examples of Egodic Measure Preserving Transformations which are Weakly mixing but not Strongly Mixing". In: *Springer Verlag, Lecture Notes in Mathematics* 318 (1972), pp. 143–150.

[17] J. Liang. *Measure-preserving dynamical systems and approximation techniques*. 2014. URL: http://math.uchicago.edu/~may/REU2014/REUPapers/Liang,Jason.pdf.

[18] P. Nyland. *Bratteli Diagrams: Modeling AF-algebras and Cantor Minimal Systems Using Infinite Graphs*. 2016. URL: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2391566.

[19] S. Park. "Bratteli Diagram Isomorphic to Chacon Homeomorphism". In: *Bulletin of the Korean Mathematical Society* (2000), pp. 519–536.

[20] Karl E. Petersen. *Ergodic Theory*. Cambridge University Press, 1983.

[21] V. Ryzhikov. *On Mixing of Staircase Transformations*. 2011. arXiv: 1108.3522.

[22] C. E. Silva. *Invitation in Ergodic Theory*. American Mathematical Society, 2007.

[23] R. A. Yassawi. *Rank One Transformations and Mixing: a Survey*. 1993. URL: http://digitool.library.mcgill.ca/webclient/StreamGate?folder_id=0&dvs=1568995670650~390.

[24] R. Zweimüller. *Ergodic Theory I*. 2019.

[25] R. Zweimüller. *Ergodic Theory II*. 2019.