INSTITUTO DE MATEMÁTICA

Universidade Federal do Rio de Janeiro

UFRJ

Bernoulli systems, Ornstein Theory and Partially Hyperbolic Diffeomorphisms

Eric Alberto Cabezas Bonilla

Rio de Janeiro, Brasil

9 de março de 2020

Bernoulli systems, Ornstein Theory and Partially Hyperbolic Diffeomorphisms

Eric Alberto Cabezas Bonilla

Dissertação de mestrado apresentada ao Programa de Pós-graduação em Matemática do Instituto de Matemática da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Matemática

Universidade Federal do Rio de Janeiro

Instituto de Matemática

Programa de Pós-Graduação em Matemática

Supervisor: Alexander Eduardo Arbieto Mendoza

Rio de Janeiro, Brasil

9 de março de 2020

CIP - Catalogação na Publicação

M962t

Mengano, Fulano
Titulo / Fulano Mengano. -- Rio de Janeiro, 2019.
101 f.

Orientador: Fulanaorientadora Menganaorientadora.
Coorientador: Fulanacoorientadora Menganacoorientadora.
Tese (doutorado) - Universidade Federal do Rio de Janeiro, Instituto de Matemática, Programa de Pós Graduação em Matemática, 2019.

1. análise. 2. álgebra. 3. geometria. I. Menganaorientadora, Fulanaorientadora, orient. II. Menganacoorientadora, Fulanacoorientadora, coorient. III. Título.

Elaborado pelo Sistema de Geração Automática da UFRJ com os dados fornecidos pelo(a) autor(a), sob a responsabilidade de Miguel Romeu Amorim Neto - CRB-7/6283.

Eric Alberto Cabezas Bonilla

Bernoulli systems, Ornstein Theory and Partially Hyperbolic Diffeomorphisms

Dissertação de mestrado apresentada ao Programa de Pós-graduação em Matemática do Instituto de Matemática da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Matemática

Trabalho aprovado por

Alexander Eduardo Arbieto Mendoza
Orientador

Sergio Augusto Romaña Ibarra

José Regis Varão Filho

Rio de Janeiro, Brasil
9 de março de 2020

Aos meus pais.

Agradecimentos

Agradeço ao meu orientador Alexander por todo seu apoio, incentivo e paciência.

Á minha familia pelo apoio e carinho incondicional.

Á meus amigos da salinha do IM pelo belo clima e pelas boas conversas sobre matemática.

Agradeço finalmente mas não menos importante á CAPES e ao CNpQ pelo apoio financeiro.

Resumo

O deslocamento de Bernoulli é um dos exemplos mais importantes em sistemas dinâmicos, a classe de automorfismos isomórficos ao deslocamento é chamada de automorfismo de Bernoulli, possuindo as mesmas propriedades que a ergodicidade, a mistura. Mas Kolmogorov define uma classe de automorfismo entre os Mixing e os Bernoulli, chamados automorfismos de Kolmogorov. Tais sistemas também eram caóticos e uma das questões em que todo o automorfismo de Kolmogorov era Bernoulli.

Neste trabalho, mostraremos um exemplo dado por Ornstein de um automorfismo de Kolmogorov que não é Bernoulli, como as técnicas para encontrar partições geradoras de Bernoulli muito fraco para obter Bernoullicidade.

Terminando com um caso parcialmente hiperbólico quando um automorfismo de Kolmogorov implica Bernoulli.

Palavras-chave: Kolmogorov automorfismo, sistema Bernoulli, parcialmente hiperbolico

Abstract

The Bernoulli shift is one of the more important examples in dynamical systems, the class of automorphisms isomorphic to the shift are called the Bernoulli automorphism, having the same properties as ergodicity, mixing. But Kolmogorov define a class of automorphism between the Mixing and the Bernoulli called the Kolmogorov automorphisms. Such systems were chaotic as well and one of the question where if all the kolmogorov automorphism were Bernoulli.

In this work we will show an example given by Ornstein of a Kolmogorov automorphism that is not Bernoulli, as the techniques to find generating partitions very weak Bernoulli to obtain Bernoullicity.

Finishing with a partially hyperbolic case when a Kolmogorov automorphism implies Bernoulli.

Keywords: Kolmogorov automorphism, Bernoulli property, Partially hyperbolic diffeomorphism.

List of Figures

Contents

1 Introduction

The main subject in dynamical system is to analyze what happened in the future of such system or how much "chaotic" an orbit is.

For example we can analyze the behavior of the roll of a dice in the future. That kind of system is the motivation for the Bernoulli shift, one of the more important systems in dynamical systems because of the chaotic properties.

In 1958-59 ([7], [8]) Kolmogorov introduced the concept of entropy in dynamical systems, which is an invariant under isomorphisms. Because of that a natural question was if two systems with the same entropy could be isomorphic or which condition is needed to obtain such characteristic.

In 1970 Donald Ornstein [11] proved that in the case of the Bernoulli shift the equality of entropies determines if these systems are isomorphic or not.

Theorem 1.0.1. *Let $T : X \longrightarrow X$ and $S : Y \longrightarrow Y$ two Bernoulli shifts. If they have the same entropy they would be isomorphic.*

Such analysis brought us a powerful technique to determine if system is a candidate to be isomorphic to a Bernoulli shift or not.

Knowing that entropy, ergodicity and mixing properties are preserved under isomorphism we can have some candidates that could be isomorphic with the Bernoulli shift. Such systems isomorphic with the Bernoulli shift are called Bernoulli automorphisms.

However there is a class of systems between the Bernoulli automorphisms and such with the mixing property called the Kolmogorov automorphisms.

It can be proved that Bernoulli automorphism are Kolmogorov automorphism, but in 1973 Ornstein [12] give and example of a Kolmogorov automorphism that is not Bernoulli.

So, there remains a question: what are the conditions to ensure that a Kolmogorov automorphism would be isomorphic to a Bernoulli shift.

Thanks to Ornstein Theory, we can answer the question for some systems as mixing Markov chain, automorphisms of $\mathbb{T}^n$, Volume preserving Anosov Diffeomorphism and many others.

The technique used to prove that an automorphism in the bi torus is Bernoulli relies heavily on the fact that such automorphisms are Anosov. Motivated by this example, Pesin [17] extend the result to Anosov Diffeomorphisms in a compact Riemannian Manifold.

The following question would be how to attack in the Partially Hyperbolic case.

In 2018, Ponce et al. [20] proves that under some conditions in the center foliations in an derived from an Anosov Diffeomorphism f. If f is Kolmogorov implies that f is a Bernoulli automorphism. More exactly:

Theorem 1.0.2. *Let $f : \mathbb{T}^3 \to \mathbb{T}^3$ be a C^2 volume preserving derived from Anosov diffeomorphism with linearization $A : \mathbb{T}^3 \to \mathbb{T}^3$. Assume that f is Kolmogorov and one of the following occurs:*

1. *$\lambda_A^c < 0$ and $\mathcal{F}^{cs}$ is absolutely continuous, or*
2. *$\lambda_A^c > 0$ and $\mathcal{F}^{cu}$ is absolutely continuous.*

Then f is Bernoulli

So in this work, we would show the isomorphism theorem given by Ornstein [11], present the technique given in [12]. Using it to show an example of a Bernoulli automorphism in a smooth field.

Ending with a brief comment about the proof of Ponce et al [20].

2 Preliminares

2.1 Ergodic Theory

In this chapter, we will remember some definitions and results of ergodic theory that we will use throughout the following chapters such as partitions, entropy and gadgets. We begin by remembering some definitions about measure theory.

Let X be a set. A *σ-algebra* of subsets of X is a collection $\mathcal{B}$ of subsets of X satisfying the following three conditions:

- $X \in \mathcal{B}$,
- if $A \in \mathcal{B}$ then $X \setminus A \in \mathcal{B}$,
- if $B_n \in \mathcal{B}$ for $n \geq 1$ then $\bigcup_{n=1}^{\infty} B_n \in \mathcal{B}$.

We call the pair $(X, \mathcal{B})$ a measurable space.

Definition 2.1.1. A sub σ-algebra $\mathscr{K}$ of $\mathcal{B}$ is a σ-algebra that is contained in $\mathcal{B}$, we will denote it as $\mathscr{K} \subset \mathcal{B}$

We would like to transmit the sense of measure volume for this abstract set. A *measure* on $(X, \mathcal{B})$ is a function $\mu : \mathcal{B} \longrightarrow \mathbb{R}^+$ satisfying

- $\mu(\emptyset) = 0$
- $\mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n)$

whenever (B_n) , $n \geq 1$ is a sequence of measurable sets pairwise disjoint subsets of X. A *measure space* is a triple $(X, \mathcal{B}, \mu)$ where $(X, \mathcal{B})$ is a measurable space. When $\mu(X) = 1$, we say that $(X, \mathcal{B}, \mu)$ is a probability space.

Definition 2.1.2. If $\mathscr{C}, \mathscr{D}$ are sub-σ-algebras of $\mathcal{B}$ we write $C \doteq D$ if for every $C \in \mathscr{C}$ there exists $D \in \mathscr{D}$ with $\mu(D \triangle C) = 0$ and if for every $D \in \mathscr{D}$ there exists $C \in \mathscr{C}$ with $\mu(D \triangle C) = 0$

Definition 2.1.3. Suppose $(X, \mathcal{A}, \mu)$, $(Y, \mathcal{B}, \nu)$ are probability spaces.

1. A transformation $T : X \longrightarrow Y$ is *measurable* if $T^{-1}(B) \in \mathcal{A}$ where B belongs to $\mathcal{B}$.

2. A transformation $T : X \longrightarrow Y$ is *measure-preserving* if T is measurable and $\mu(T^{-1}(A)) = \nu(A)$ where A belongs to $\mathcal{B}$.

3. we say that $T : X \longrightarrow Y$ is an *invertible measure-preserving* transformation if T is measure-preserving, bijective and T^{-1} is also measure-preserving.

With the motivation to work on a set that is preserved over iterations, we define the following:

Definition 2.1.4. Let $(X, \mathcal{B}, \mu)$ be a probability space and $T : X \longrightarrow X$ be a measure preserving transformation. We say that B measurable set is T-*invariant* if $T^{-1}B = B$.

Definition 2.1.5. Let $(X, \mathcal{B}, \mu)$ be a probability space. A measure-preserving transformation T of $(X, \mathcal{B}, \mu)$ is called *ergodic* if the only members B of $\mathcal{B}$ T-invariant satisfy $\mu(B) = 0$ or $\mu(B) = 1$.

There are several way to characterize the ergodicity condition and we present some of them in the next theorem.

Theorem 2.1.6. *If $T : X \longrightarrow X$ is a measure-preserving transformation of the probability space $(X, \mathcal{B}, m)$ then the following statements are equivalent:*

1. *T is ergodic.*

2. *The only members B of $\mathcal{B}$ with $m(T^{-1}(B)\Delta B) = 0$ are those with $m(B) = 0$ or $m(B) = 1$.*

3. *For every $A \in \mathcal{B}$ with $m(A) > 0$ we have $m(\bigcup_{n=1}^{\infty} T^{-n}A) = 1$.*

4. *Whenever f is measurable and $(f \circ T)(x) = f(x)$ almost everywhere then f is constant almost everywhere.*

Observation: We say that X has the P property *almost everywhere*(a.e.) if $X \setminus A$ has the P property where $\mu(A) = 0$.

Theorem 2.1.7. *(Birkhoff's Ergodic Theorem) Suppose $T : (X, \mathcal{B}, m) \longrightarrow (X, \mathcal{B}, m)$ is measure-preserving and $f \in L^1(m)$. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)) = f^* \ \ a.e.$$

where $f^* \in L^1(m)$. Also $f^* \circ T = f^*$ a.e. and if $m(X) < \infty$, then $\int f^* dm = \int f dm$.

The solutions of these theorem can be found on [28].

Definition 2.1.8. Let T be a measure-preserving transformation of a probability space $(X, \mathcal{B}, m)$. We say that T is *strong mixing* (or have the strong mixing property) if for every A, B in $\mathcal{B}$

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

2.2 Shifts and partitions

One of the classic and more powerful examples in ergodic theory is the Bernoulli shift, is defined as follows:

Let $\pi = (p_1, ..., p_n)$ be the *probability vector* such that $p_i > 0$ and $\sum p_i = 1$. Define the alphabet $\{1, 2, .., k\}$ as a set of symbols. We will consider Σ_k as the set of all doubly infinite sequence of the symbols $1, 2, ..., k$ that is:

$$\Sigma_k = \{(x_n) : x_n \in \{1, ..., k\}, n \in \mathbb{Z}\}$$

The σ-algebra $\mathscr{B}$ defined on Σ_k it would be generated by finite unions of cylinder sets where a *cylinder set* is a subset of Σ_k determined by a finite number of values, such that:

$$A = \{(x_n) : x_i = t_i, -m \leq i \leq n\}$$

where t_i is any fixed symbol in the alphabet. Therefore there exist a unique measure μ such that if A is a cylinder set, then $\mu(A) = \prod_{i=-m}^{n} p_{t_i}$.

The transformation T defined by

$$(Tx)n = x_{n+1}, n \in \mathbb{Z},$$

is an invertible μ-preserving transformation called Bernoulli shift.

The measure space $(\Sigma_k, \mathscr{B}, \mu, T_\pi)$, will be called the *Bernoulli shift space with distribution* π.

There are many ways to determine the space Σ_k i.e. it has many isomorphisms, so a given Benoulli shift can be described in many other ways.

For example, for the case when $\pi = (1/2, 1/2)$. For convenience we shall use the indexing $\{0, 1\}$ instead of $\{1, 2\}$; that is Σ_2 will be the set of all doubly infinite sequences of zeros and ones.

An *isomorphism* ϕ of $(X, \mathcal{A}, \mu)$ onto $(Y, \mathcal{B}, \nu)$ is a mapping $\phi : X \longrightarrow Y$ such that

$$\phi(\mathcal{A}) \subset \mathcal{B},\ \phi^{-1}(\mathcal{B}) \subset \mathcal{A}$$

$$\nu(\phi(A)) = \mu(A), A \in \mathcal{A}\ ;\ \mu(\phi^{-1}(B)) = \nu(B), B \in \mathcal{B}$$

also ϕ is one-to-one and onto (mod0), that is, there are sets $\tilde{X} \subset X$ and $\tilde{Y} \subset Y$ such that $\mu(\tilde{X}) = \mu(X)$, $\nu(\tilde{Y}) = \nu(Y)$ and ϕ is a one-to-one map of $\tilde{X}$ onto $\tilde{Y}$.

We say that X is a *Lebesgue space* if X is isomorphic to the unit interval with Lebesgue measure.

Some systems that are Lebesgue space are the Bernoulli shift, the irrational rotation, tent function, etc.

A collection $\xi \subset \mathcal{B}$ separates X if there is a set $E \in \mathcal{B}$, $\mu(E) = 0$, such that if $x, y \notin E$, there is a set $A \in \xi$ such that $x \in \mathcal{A}$, $y \notin A$ or $x \notin \mathcal{A}$, $y \in A$

There is a class of automorphism isomorphic to a shift Bernoulli called Bernoulli automorphism (or with the Bernoulli property) and in this dissertation we will talk a lot about it.

If π and $\overline{\pi}$ are given distributions, when will T_π and $T_{\overline{\pi}}$ are isomorphic? The Kolmogorov-Ornstein isomorphism theorem answer this question.

Two Bernoulli shifts T_π, $T_{\overline{\pi}}$ are isomorphic if and only if

$$\sum_{i=1}^{k} p_i \log p_i = \sum_{i=1}^{k} \overline{p}_i \log \overline{p}_i \tag{2.1}$$

where $\pi = (p_1, p_2, ..., p_k)$, $\overline{\pi} = (\overline{p}_1, \overline{p}_2, ..., \overline{p}_k)$.

The necessity of the condition (2.1) was established by Kolmogorov, while Ornstein prove its sufficiency. In this section, we explain the ideas behind Ornstein's result.

A basic tool that will help us to describe Bernoulli shifts are the partitions.

Let $(X, \mathscr{B}, \mu, T)$ be a dynamical system, we say that $\mathcal{P}$ is a *partition* of X if it is an ordered disjoint collection of measurable sets(called the *atoms* of $\mathcal{P}$) whose union is X. The partitions could be enumerable but in this work we assume finite partitions. We will denote the number of atoms in $\mathcal{P}$ by $|\mathcal{P}|$.

If $\mathcal{P}$ and $\mathcal{Q}$ are partitions of X, then $\mathcal{P}$ *refines* $\mathcal{Q}$ if each atom in $\mathcal{P}$ is in some atom in $\mathcal{Q}$, from that we can see that each atom $\mathcal{Q}$ could be seen as a union of atoms in $\mathcal{P}$. If $\mathcal{P}$ refines $\mathcal{Q}$, we write $\mathcal{Q} \prec \mathcal{P}$.

If $\mathcal{P}$ and $\mathcal{Q}$ are partitions, their join is

$$\mathcal{P} \vee \mathcal{Q} = \{P_i \cap Q_j : P_i \in \mathcal{P}, Q_j \in \mathcal{Q}\}$$

with lexicographic ordering. Clearly $\mathcal{P} \vee \mathcal{Q}$ is the least partition that refines $\mathcal{P}$ and $\mathcal{Q}$.

For sequence of partitions $\mathcal{P}_i$, $1 \leq i \leq n$, we use the notation

$$\bigvee_{i=1}^{n} \mathcal{P}_i = \mathcal{P}_1 \vee \mathcal{P}_2 \vee ... \vee \mathcal{P}_n$$

We can characterize the path of every point through the partition in the following way. Let $P = \{P^1, P^2, ..., P^k\}$ a partition of X. The *P-name of* x will denote the sequence $\{\alpha_i\}_{-\infty}^{\infty}$; where $\alpha_i = j$ if $T^i x \in P^j$

The *P-n-name of* x will be $\{\alpha_i\}_0^{n-1}$

The *P-name of an atom* $A \in \bigvee_0^n T^{-i}P$ will be the P-n-name of the $x \in A$

A partition $\mathcal{P}$ determines a σ-algebra $\mathscr{B}(\mathcal{P})$ which is just the set of all unions of members of $\mathcal{P}$. Note that

$$\mathcal{P} \succ \mathcal{Q} \text{ iff } \mathscr{B}(\mathcal{P}) \supset \mathscr{B}(\mathcal{Q}),$$

The distribution of P or $\operatorname{dist}(P)$ will be the vector $(\mu(P^1), \mu(P^2), ..., \mu(P^k))$. If E is a measurable set, $\operatorname{dist}(P/E)$ will denote $\dfrac{\mu(P^1 \cap E)}{\mu(E)}, ..., \dfrac{\mu(P^k \cap E)}{\mu(E)}$.

Let T be a transformation and $\mathcal{P}$ is a partition, then $T\mathcal{P} = \{TP^i : P^i \in \mathcal{P}\}$, for example:

$$\bigvee_0^n T^i\mathcal{P} = \mathcal{P} \vee T\mathcal{P} \vee ... \vee T^n\mathcal{P}.$$

We say that $\mathcal{P}$ is a *generator* for T, if $\bigvee_{-\infty}^{\infty} T^i\mathcal{P} = \mathscr{B}$, where $\bigvee_{-\infty}^{\infty} T^i\mathcal{P}$ is the σ-algebra generated by the atoms in $\bigvee_{-k}^{k} T^i\mathcal{P}$ for every $k \in \mathbb{N}$.

Example

If a $\mathcal{A}$ is a sub-σ-algebra of (X, σ, m) and P a partition of (X, σ, m) we will say that $\mathcal{A} \supset^{\epsilon} P$ if there is a $\widehat{P}$, measurable with respect to $\mathcal{A}$, and $|P - \widehat{P}| < \epsilon$. If P is measurable with respect to $\mathcal{A}$, we will write $\mathcal{A} \supset P$

We will write $P \subset^{\epsilon} \bigvee_{-K}^{K} T^iQ$ assuming that $\bigvee_{-K}^{K} T^iQ$ is denoted as the algebra of sets generated by the partition $\bigvee_{-K}^{K} T^iQ$.

Lemma 2.2.1. *A partition P is a generator for T if and only if, for each partition Q and each $\epsilon > 0$, there is an n such hat $Q \subset^{\epsilon} \bigvee_{-n}^{n} T^iP$.*

We say that two partitions $\mathcal{P}$ and $\mathcal{Q}$ are independent if

$$\mu(P^i \cap Q^j) = \mu(P^i)\mu(Q^j),\ \ P^i \in \mathcal{P}, Q^j \in \mathcal{Q}$$

This say that $\mathcal{P}$ partitions each atom in $\mathcal{Q}$ in the same proportion as it partitions the entire space.

We say that the sequence of partitions $(\mathcal{P}_n)$, $n \in \mathbb{N}$, is an *independent sequence*, if for each $n > 1$, $\mathcal{P}_n$ and $\bigvee_1^{n-1} \mathcal{P}_i$ are independent.

We can characterize a Bernoulli shift with some especial partition

Theorem 2.2.2. *A transformation T is isomorphic to a Bernoulli shift T_π with distribution $\pi = (p_1, p_2, ..., p_k)$ if and only if there is a partition $\mathcal{P} = \{P_1, P_2, ..., P_k\}$ such that*

1. $\operatorname{dist} \mathcal{P} = \pi$,
2. *$\mathcal{P}$ is a generator for T,*
3. *$T^n\mathcal{P}$, $n \geq 1$ is an independent sequence.*

These remarks can be generalized to yield the following theorem

Theorem 2.2.3. *The transformations T and $\overline{T}$ are isomorphic if and only if there are partitions $\mathcal{P}$ and $\overline{\mathcal{P}}$ which are generators for T and $\overline{T}$, respectively, such that*

$$\operatorname{dist} \bigvee_0^n T^i\mathcal{P} = \operatorname{dist} \bigvee_0^n \overline{T}^i\overline{\mathcal{P}}, n = 0, 1, 2, ...$$

The proofs of Theorems 2.2.2 and 2.2.3 appears in [26].

2.3 Stack

The key to an understanding of Ornstein's proof of the isomorphism theorem and to a number of other results is a simple geometric representation of a transformation.

Theorem 2.3.1. *(Rohlin [6]) If T is ergodic, n is a positive integer, and ϵ a positive number, then there is a set F such that $F, TF, T^2F, ..., T^{n-1}F$ is a disjoint sequence, such that, $\mu\left(\bigcup_{i=0}^{n-1} T^iF\right) \geq 1 - \epsilon$*

We will show the proof for $n = 2$.

Sketch of the proof: Take B a small measurable set with positive measure ($\mu(B) < \epsilon$). Since T is ergodic exists points in B such that its image does not belong in B.

$$\tilde{B} = \{x \in B : Tx \notin B\}. \tag{2.2}$$

We can picture B and $T\tilde{B}$ as sets with $T\tilde{B}$ above $\tilde{B}$.

Now put $B_1 = T\tilde{B}$ and let $\tilde{B}_1 = \{x ß B_1 : Tx \notin B\}$. Note that if $x \in \tilde{B}_1$, $Tx \notin B_1$, hence we can picture $T\tilde{B}_1$ as a set above $\tilde{B}_1$. By continuing this process by iteration we obtain that T maps B_i directly upwards into B_{i+1} or into B.

Since T is ergodic, the set $B \bigcup_{i=1}^{\infty} B_i = X(\mathrm{mod} 0)$, it follows from he invariance of the union and can not have measure 0.

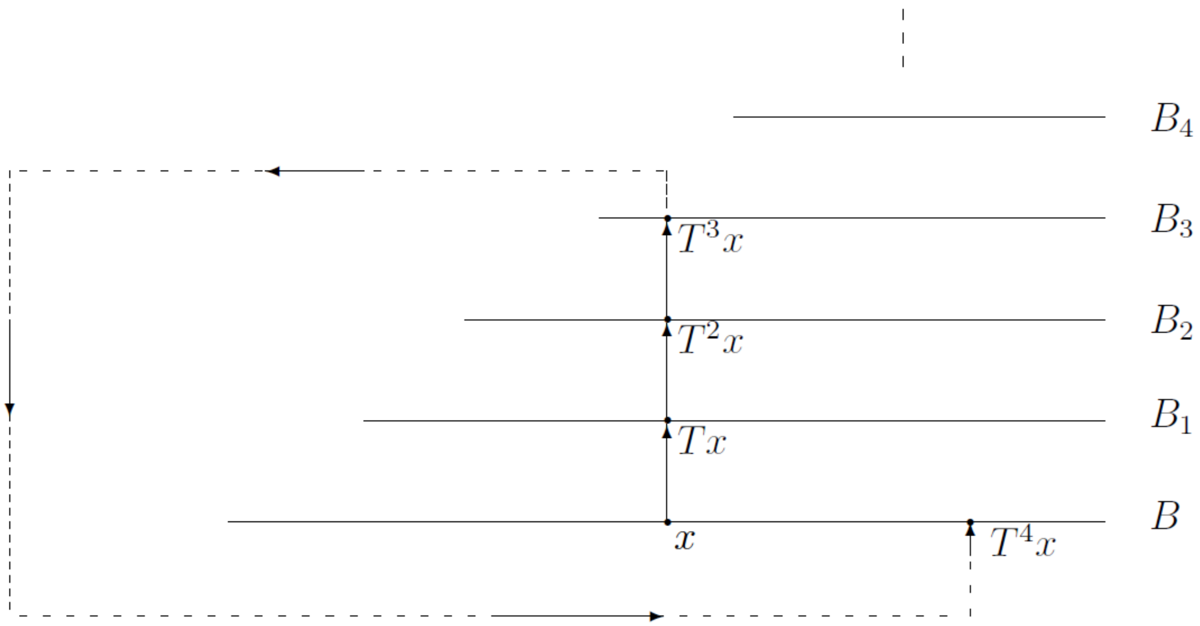

Figure 1 – Return of T^4x to B

We now define

$$F = T^{-1}B_1 \cup T^{-1}B_3 \cup T^{-1}B_5 \cup ...$$

Thus, F is made of pieces of $B, B_2, B_4, ...$ and TF is made up of $B_1, B_3, B_5, ...$ so clearly F and TF are disjoint.

The complement of them is formed by the union of $C_0 = V - T^{-1}B_1$ and $C_i = B_{2i} - T^{-1}B_{2i+1}$ for $i \geq 1$.

It follows that:

$$\mu\left(\bigcup_{i=0}^{\infty} C_i\right) = \sum_{i=0}^{\infty}\mu(C_i) = \sum_{i=0}^{\infty}\mu(D_i) \tag{2.3}$$

where $D_i = T^{-2i}C_i$. Since the D_i are disjoint and contained in B, we have

$$\mu(F \cup TF) \geq 1 - \mu(B). \tag{2.4}$$

Observation: For $n > 2$ just depend of the choice of F.

Theorem 2.3.2. *[Rohlin(Strong form)] If T is ergodic, n is a positive integer, ϵ a positive number, and P a partition, then there is a set F such that $F, TF, T^2F, ..., T^{n-1}F$ is a disjoint sequence, $\mu\left(\bigcup_{i=0}^{n-1} T^iF\right) \geq 1-\epsilon$ such that*

$$\mathrm{dist}(P/F) = \mathrm{dist}(P)$$

2.4 Gadgets

We introduce the terminology used in [11]. A *gadget* is a quadruple (T, F, n, P), where T is a transformation, F a set such that $F, Tf, ..., T^{n-1}F$ is a disjoint sequence, and a partition P of $\bigcup\limits_{0}^{n-1} T^iF$. We will call F the base of F and $T^n(F)$ the top of the gadget.

We shall say that (T, F, n, P) is *isomorphic* to $(\overline{T}, \overline{F}, n, \overline{P})$ if

$$\mathrm{dist}\left(\bigvee_{0}^{n-1} T^{-i}P/F\right) = \mathrm{dist}\left(\bigvee_{0}^{n-1} \overline{T}^{-i}\overline{P}/\overline{F}\right)$$

that is, P-n-names partition F in the same proportions as corresponding $\overline{P}$-n-names partition $\overline{F}$. It is implicit in this definition that the two gadgets have the same height, and that P and $\overline{P}$ have the same number of sets.

It is easy to see that (T, F, n, P) is isomorphic to $(\overline{T}, \overline{F}, n, \overline{P})$ if and only if there is an invertible map $S : \overline{F} \longrightarrow F$ such that, for all measurable set $\overline{A} \subset \overline{F}$ and $A \subset F$, we have

$$\overline{\mu}(\overline{A})/\overline{\mu}(\overline{F}) = \mu(S\overline{A})/\mu(F)$$
$$\overline{\mu}(S^{-1}A)/\overline{\mu}(\overline{F}) = \mu(A)/\mu(F)$$

and for $x \in \overline{F}$, the $\overline{P}$-n-name of x and the P-n-name of Sx are the same. In other words, except for a possible change of scale, two gadgets are isomorphic if one cannot distinguish between them by examining their column structures.

The statement that two gadgets are isomorphic says very little about their respective transformations, for Rohlin's Theorem and a simple construction combine to give the following result.

Lemma 2.4.1. *If (T, F, n, P) is any gadget and $\overline{T}$ is any ergodic transformation, then for any $\epsilon > 0$, there is a set $\overline{F}$ and a partition $\overline{P}$ such that $(\overline{T}, \overline{F}, n, \overline{P})$ is a gadget isomorphic to (T, F, n, P) and $\mu(\bigcup_{i=0}^{n-1} \overline{T}^i \overline{F}) \geq 1 - \epsilon$.*

Lemma 2.4.2. *Suppose (T, F, n, P) is isomorphic to $(\overline{T}, \overline{F}, n, \overline{P})$ and Q is a partition of $\bigcup_0^{n-1} T^i F$. Then there is a partition $\overline{Q}$ of $\bigcup_0^{n-1} \overline{T}^i \overline{F}$ such that $(T, F, n, P \vee Q)$ is isomorphic to $(\overline{T}, \overline{F}, n, \overline{P} \vee \overline{Q})$.*

2.5 Metrics in partitions

The first of these is the *distribution distance* given by

$$|\operatorname{dist}(P) - \operatorname{dist}(Q)| = \sum_{i=1}^{k} |\mu(P_i) - \mu(Q_i)|. \tag{2.5}$$

where the partitions P and Q each have k atoms. Note that is not required that the partitions belongs the same space.

In the case that $|\operatorname{dist}(P) - dist(Q)| = 0$, P and Q have the same distribution.

A stronger form of closeness is the *partition distance*

$$|P - Q| = \sum_{i=1}^{k} |\mu(P_i \Delta Q_i)|, \tag{2.6}$$

where P and Q have k atoms and are in the same space.

In the case when $|P - Q| = 0$ we have $\mu(P_i \Delta Q_i) = 0$, $1 \leq i \leq k$, that is $P \doteq Q$, P and Q agree except on a set of measure zero.

If P and Q are in the same partition, then

$$|\operatorname{dist}(P) - \operatorname{dist}(Q)| \leq |P - Q|.$$

Let T and $\overline{T}$ be transformations defined X, $\overline{X}$ respectively, and with respective partitions P and $\overline{P}$. The *process distance* is defined by

$$\overline{d}((T,P),(\overline{T},\overline{P})) = \sup_n \inf_{S\in\Gamma} \frac{1}{n+1}\sum_{i=0}^{n} |T^iP - S\overline{T}^i\overline{P}|, \tag{2.7}$$

where Γ is the class of isomorphisms of X onto $\overline{X}$.

We will mention some properties of the process metric.

Note that the supremum in (2.7) is actually a limit. This follows from the fact that, if $\inf_{S\in\Gamma} 1/n\sum_{i=0}^{n-1}|T^iP - S\overline{T}^i\overline{P}| = \alpha$, hen, for all r, $\inf_{S\in\Gamma} 1/nr\sum_{i=0}^{nr-1}|T^iP - S\overline{T}^i\overline{P}| \geq \alpha$. We also observe that

Proposition 2.5.1. *Let P and $\overline{P}$ are generators for T and $\overline{T}$. If $\overline{d}((T,P),(\overline{T},\overline{P})) = 0$, then T is isomorphic to $\overline{T}$.*

The proof is as follows: The condition that $\overline{d}((T,P),(\overline{T},\overline{P})) = 0$ implies that, for each n,

$$\inf_{S\in\Gamma} \frac{1}{n}\sum_{i=0}^{n-1} |T^iP - S\overline{T}^i\overline{P}| = 0,$$

and hence that dist $\bigvee_0^{n-1} T^iP =$ dist $\bigvee_0^{n-1}\overline{T}^i\overline{P}$, $n = 1, 2, \ldots$. Then, using Theorem 2.2.3, the proposition follows.

There is another equivalence in the process metric that we shall use and it is when we helps us for another space that would be as intermediary.

Let Y be a fixed nonatomic probability space, and let Γ denote the class of all isomorphisms of the T-space onto Y, and let $\overline{\Gamma}$ denote the class of all isomorphisms of the $\overline{T}$-space onto Y. We then have

$$\overline{d}((T,P),(\overline{T},\overline{P})) = \sup_n \inf_{S\in\Gamma,\overline{S}\in\overline{\Gamma}} \frac{1}{n}\sum_{i=0}^{n-1} |ST^iP - \overline{S}\overline{T}^i\overline{P}|. \tag{2.8}$$

Lemma 2.5.2. *If $\overline{d}((T,P),(\overline{T},\overline{P})) < \epsilon^2$, then for each n, there is an isomorphism S_n from $\overline{X}$ to X such that the set of points $x \in X$ for which the P-n-name of x and the P-n-name of $S_n^{-1}x$ differ in more than ϵn places has measure less than ϵ.*

2.6 independence and ϵ-independence

Recall, two partitions P and Q are independent if

$$\mu(P^i \cap Q^j) = \mu(P^i)\mu(Q^j),\ P^i \in P,\ Q^j \in Q$$

we can redefined it as $\mathrm{dist}(P/Q^j) = \mathrm{dist}(P)$, $Q^j \in Q$ which tell us that P partitions each set in Q in the same proportions that P partitions X.

The definition of approximate independence we shall use merely asserts that P partitions most sets in Q in almost the same way that P partitions X. To be precise, we say that P is *ϵ-independent of Q* or $P \perp^{\epsilon} Q$ if there is a collection C of atoms Q such that

- $m(C) > 1 - \epsilon$,
- if $A \in C$ then $|\mathrm{dist}P/A - \mathrm{dist}P| < \epsilon$.

2.7 Entropy

Let $P = \{P^1, P^2, ..., P^k\}$ a partition of X. We say that $H(P)$ is *The entropy of T relative to P* is

$$H(P) = -\sum_i \mu(P^i) \log \mu(P^i). \tag{2.9}$$

Let P and Q finite partitions. The *conditional entropy* of P and Q is defined as

$$H(P|Q) = H(P \vee Q) - H(Q)$$

An easy calculation establish that

$$H(P|Q) = -\sum_j \mu(Q^j) \sum_i \frac{\mu(P^i \cap Q^j)}{\mu(Q^j)} \log \frac{\mu(P^i \cap Q^j)}{\mu(Q^j)} \tag{2.10}$$

The *entropy of T relative to P* is

$$H(T, P) = \lim_n \frac{1}{n} H(\bigvee_1^n T^i P). \tag{2.11}$$

Theorem 2.7.1. *(Shannon-McMillan-Breiman) If T is ergodic and P a finite partition , then given ϵ we can find an N such that if $n > N$, then there is a collection $\mathcal{C}$ of atoms of $\bigvee_0^{n-1} T^i P$ such that $m(\mathcal{C}) > 1 - \epsilon$, and if $A \in \mathcal{C}$, then the measure of A is between $\frac{1}{2}^{(H(P,T)+\epsilon)n}$ and $\frac{1}{2}^{(H(P,T)-\epsilon)n}$.*

The proof of this theorem can be found in Billingsley [1].

The number $H(T,P)$ depends upon the partition P. To obtain a invariant for T, we define *the entropy of* T as

$$H(T) = \sup\{H(T,P) : P \text{ is a finite partition}\}. \tag{2.12}$$

Is not difficult to see that the entropy is invariant upon isomorphism i.e. if S is isomorphic to T, then $H(S) = H(T)$.

We end up this subsection with a mean for calculating T given by Kolmogorov and Sinai, a proof of these theorem appears in [10] or in [28].

Theorem 2.7.2. *(Kolmogorov-Sinai) If P is a generator for T and Q is any partition, then $H(T,P) \geq H(T,Q)$. In particular $H(T) = H(T,P)$ for any generator P.*

We will say that (P,T) is finitely determined if given $\epsilon > 0$ there is an $\delta > 0$ and an integer n such that if $\overline{P}, \overline{T}$ satisfies

1. $|\overline{P}| = |P|$
2. $|H(\overline{P}, \overline{T}) - H(P,T)| < \delta$
3. $|\bigvee_0^n T^i P - \text{dist} \bigvee_0^n \overline{T}^i \overline{P}| < \delta$

Then $\overline{d}((P,T),(\overline{P},\overline{T})) < \epsilon$

3 Ornstein Theory

3.1 Isomorphism Theorem

Part of the importance of the shift resides in that some diffeomorphism have in his dynamics, transformations of this type i.e. a diffeomorphism f under a manifold contents a subset Λ compact invariant such that f restricted to Λ is "equivalent" to a certain shift.

It follows from the independence of two partitions P and Q of a probability space that $H(P \bigvee Q) = H(P) + H(Q)$. The following lemma states a relation between ϵ-independence with this property.

Lemma 3.1.1. *Given $\epsilon > 0$ and an integer k, there is a $\delta(\epsilon, k) > 0$ such that if P has k atoms and if $|H(P \vee Q) - H(P) - H(Q)| < \delta(\epsilon, k)$, then $P \perp^{\epsilon} Q$*

Proof. If P is not $\epsilon-$independent of Q, then for every collection C of atoms Q there is an $Q^i \in C$ such that $|\text{dist}P/Q_i - \text{dist}P| \geq \epsilon$.

There is an atom P' of P such that $|\dfrac{m(P' \cap Q^i)}{m(Q^i)} - m(P')| \geq |\dfrac{m(P^j \cap Q^i)}{m(Q^i)} - m(P^j)|$ for every $j = 1, ..., k$

Then

$$k\left|\frac{m(P' \cap Q^i)}{m(Q^i)} - m(P')\right| \geq |\text{dist}P/Q_i - \text{dist}P| \geq \epsilon.$$

Let $\overline{Q}$ a partition with three atoms:

- Complement of union of the Q^i in C
- The Q^i in C, $m(P' \cap Q^i)/m(Q^i) - m(P') > 0$
- The Q^i in C, $m(P' \cap Q^i)/m(Q^i) - m(P') < 0$

Denote $\overline{Q} = \{\overline{P}, P_1, P_2\}$.

If $P \perp^{\epsilon/(2k)} \overline{Q}$ then $P \perp^{\epsilon/(k)} \overline{Q}$.

Let $\overline{C}$ a collection of atoms of $\overline{Q}$. Clearly $\overline{C} \neq \overline{P}$ because $m(\overline{P}) < 1 - \epsilon/k$.

If $P_i \in \overline{C}$ then $r_i\epsilon/k < |\text{dist}P/Q_i - \text{dist}P| \geq \epsilon/k$ where r_i would be the number of atoms de Q in P_i , $i = \{1, 2\}$.

Remember that $H(P \vee \overline{Q}) - H(\overline{Q})$ could be seen as an application with $3k$ variables (Because is determinated by the measure of the partitions and $P \vee \overline{Q}$ has $3k$ atoms).

$F(\epsilon, k) = \{(m(P^i \cup Q^j))_{\substack{i\in\{1,\ldots,k\}\\ j\in\{1,2,3\}}}, P$ is not $\epsilon/(2k)$independent of$\overline{Q}\}$ is compact in $\mathbb{R}^{3k}$ because is limited and closed because the definition of $P \perp^{\epsilon/2k} \overline{Q}$.

$|H(P \vee \overline{Q}) - H(\overline{Q}) - H(P)| \neq 0$ in $F(\epsilon, k)$ because P and Q are not independent.

Therefore there is an $\delta_{(\epsilon,k)} > 0$ such that $|H(P \vee \overline{Q}) - H(\overline{Q}) - H(P)| > \delta_{(\epsilon,k)}$ (because is the minimum of that function in $F(\epsilon, k)$).

By using that $H(P \vee Q) - H(Q) \leq H(P)$ and Q refines $\overline{Q}$ then $|H(P \vee \overline{Q}) - H(\overline{Q}) - H(P)| \leq |H(P \vee Q) - H(Q) - H(P)|$.

We get that P is not $\epsilon-$independent of Q then $|H(P \vee Q) - H(Q) - H(P)| > \delta_{(\epsilon,k)}$

$\square$

We would like to know under which properties the process (Q, S) would be close in the $\overline{d}$ metric with a process (P, T) such that $T^i P$ are independent. The next lemma answer this question.

Lemma 3.1.2. *If $T^i P$ are independent then let $\epsilon > 0$, exist $\delta > 0$ such that:*

- $|Q| = |P|$
- $|H(Q, S) - H(P, T)| < \delta$
- $|distP - \ distQ| < \delta$

  then $\overline{d}((P, T), (Q, S)) < \epsilon$

Proof. Let $|P| = k$. Using Lemma 3.1.1, if $|H(Q, S) - H(Q)| < \delta_{(\epsilon,k)}$ then $S^n Q \perp^{\epsilon} \bigvee_0^{n-1} S^j Q$ for all n.

Because $\lim_{n\to\infty} |H(\bigvee_0^{n} S^i Q) - H(\bigvee_0^{n-1} S^i Q) - H(S^n Q)| = H(Q, S) - H(Q) < \delta_(\epsilon, k)$.

If $|distP - distQ| < \epsilon$, $T^i P$ are independent and $S^n Q \perp^{\epsilon} \bigvee_0^{n-1} S^j Q$ then $\overline{d}((Q, S), (P, T)) < 3\epsilon$

We will prove it for each n find two sequences of partitions $\{P_i\}_1^n$, $\{Q_i\}_1^n$ of the same space Z such that: dist $\bigvee_1^n Q_i =$ dist $\bigvee_1^n S^i Q$ and $\bigvee_1^n P_i =$ dist$\bigvee_1^n T^i P$.

To prove it we will do it by induction. Suppose we have already P_i and Q_i, $1 \leq i \leq n-1$. Take P_n and Q_n as follows:

- Take P_n such that dist$(P_n/A) =$ dist$P$, for each atom $A \in \bigvee_1^{n-1} Q_i \vee P_i$. This will ensure that dist$\bigvee_1^n P_i =$ dist$\bigvee_1^n T^i P$ because $T^i P$ are independent.

- Take Q_n such that if $A \in \bigvee_1^{n-1} P_i \vee Q_i$, $\overline{A} \in \bigvee_1^{n-1} Q_i$ that contains A and $\tilde{A} \in \bigvee_1^{n-1} S^i Q$ the atoms corresponding to $\overline{A}$, then $\text{dist}(Q_n/A) = \text{dist}(S^n Q/\tilde{A})$. This will insure that $\text{dist}\bigvee_1^n Q_i = \text{dist}\bigvee_1^n S^i Q$

Because of $S^n Q \perp^\epsilon \bigvee_1^{n-1} S^i Q$ we have that there is a collection of atoms C of $\bigvee_1^{n-1} S^i Q$ such that $m(C) < \epsilon$ and if $\tilde{A} \notin C$ then $|\ \text{dist}S^n Q/\tilde{A} - \text{dist}Q| < \epsilon$ because the definition of $\epsilon-$independent and S preserves measure.

If Q_n is chosen as above, there is a collection C_1 of atoms of $\bigvee_1^{n-1} P_i \vee Q_i$ such that $m(C_1) < \epsilon$ if $A \notin C_1$ then $|\ \text{dist}Q_n/A - \text{dist}Q| < \epsilon$, because of the definition of C and Q_n.

Because $|\ \text{dist}P - \text{dist}Q| < \epsilon$, we get that if $A \notin C_1$, then $|\ \text{dist}Q_n/A - \text{dist}P| < 2\epsilon$.

Hence A is not in C_1 we can choose P_n and Q_n so that on A, P_n and Q_n differ on a set of measure $< 2\epsilon m(A)$.

This implies that we can choose P_n and Q_n such that $|P_n - Q_n| < 3\epsilon$. Once founded $P_i, Q_i \subset Z\ /\ |P_i - Q_i| < 3\epsilon$.

We can form φ and ψ such that $P_i = \varphi(T^i P)$, $Q_i = \psi(S^i Q)$ where:

$|P_i - Q_i| = |\varphi(T^i P) - \psi(S^i Q)| < 3\epsilon$ then $\sum_{i=1}^n |\varphi(T^i P) - \psi(S^i Q)| < 3n\epsilon$.

Because the definition of $d_{\varphi,\psi}$ and taking limit, we obtain that $\overline{d}((P,T),(Q,S)) \leq 3\epsilon$

□

Corollary 3.1.3. *If $T^i P$ are independent then P, T is finitely determined*

Lemma 3.1.4. *Let R, S be finitely determined and $H(R,S) \leq H(T)$. Then given $\delta > 0$, there exist a partition P' such that*

$$\overline{d}((R,S),(P',T)) < \delta \tag{3.1}$$

Proof. Since R, S is finitely determined, given $\overline{\delta}$ and u we need to find P' such that

$$\left|\text{dist}\bigvee_0^u T^i P' - \text{dist}\bigvee_0^u S^i R\right| < \overline{\delta}, \tag{3.2}$$

and

$$|H(P',T) - H(R,S)| < \overline{\delta}. \tag{3.3}$$

Choose Q such that $H(R,S) - H(Q,T) = \beta > 0$ where $\beta < \dfrac{\overline{\delta}}{100}$.

Because the definition of entropy we can choose $\gamma < \overline{\delta}$ such that $|Q' - Q| < \gamma$ implies that $H(Q',T) \geq H(Q,T) - \dfrac{\overline{\gamma}}{100}$.

We will call an atom in $\bigvee_0^n T^{-i}Q$ a good Q-n-atom if its measure is between $\left(\dfrac{1}{2}\right)^{[H(Q,T)\pm\beta/100]\cdot n}$.

We will call an atom in $\bigvee_0^n S^{-i}R$ a good R-n-atom if its measure is between $\left(\dfrac{1}{2}\right)^{[H(R,S)\pm\beta/100]\cdot n}$, and if its n-name has the property that the distribution of names of length u in it is within $\overline{\delta}$ of the distribution of R-u-names under S(the atoms in $\bigvee_0^u S^{-i}R$).

Because of the Shannon-McMillan-Breinman theorem and the Birkhoff ergodic theorem we can choose n large enough such that the measure of the union of atoms in $\bigvee_0^n T^{-i}Q$ and $\bigvee_0^u S^{-i}R$ that are not good is less than $\gamma/100$.

We can also choose n so large such that

$$\frac{u}{n} < \frac{\gamma}{100} \text{ and } -\left[\frac{1}{n}\log\frac{1}{n} + \left(1-\frac{1}{n}\right)\log\left(1-\frac{1}{n}\right)\right] < \frac{\gamma}{100}. \tag{3.4}$$

We can see that there are more good R-n-atoms than there are good Q-n-atoms. Therefore, we can assign to each good Q-n-atom a good R-n-atom.

We can now apply the R-K theorem to obtain a set F such the T^iF $0 \leq i \leq n$ are disjoint,

$$m\left(\bigcup_0^n T^iF\right) > 1 - \frac{\gamma}{100} \text{ and } \operatorname{dist}\bigvee_0^n T^{-i}Q/F = \operatorname{dist}\bigvee_0^n T^{-i}Q. \tag{3.5}$$

Call the atoms in $\bigvee_0^n T^{-i}Q/F$ that are contained in good Q-n-atoms, good Q-n-F-atoms, we will do the same thing to the atoms in R-n-atoms and call it Q-n-F-atoms.

We are ready to define P', and we will do it over $\bigcup_0^n T^i\overline{F}$ where $\overline{F}$ is the union of the good Q-n-F-atoms.

If g is a good Q-n-F-atom, then T^jg will lie in $(P')^i$(the i^{th} atom in P') where i is the j^{th} term in the R-n-name of the R-n-atom assigned to g.

Thus if g is a good Q-n-F-atom, then the P'-n-name of $x \in g$ will be the same as the R-n-name of the R-n-atom assigned to g.

It does not matter how to define P' in the rest of X.

Let $\tilde{F}$ be the partition consisting of F and its complement.

We have that $\bigvee_{-n}^n T^i(P' \vee \tilde{F})$, restricted to $\bigcup_0^n T^i\overline{F}$, refines Q.

By the definition of Q-n-F-atoms we have that if g is a Q-n-F atom and $0 \leq j \leq n$, then $T^j g$ is entirely contained in one atom of Q. Because of each Q-n-F-atom in $\overline{F}$ was assigned a different R-n-atom we have that $\bigvee_0^n T^{-i} P'$, restricted to $\overline{F}$, is exactly the partition of $\overline{F}$ into Q-n-F-atoms. Therefore $\bigvee_{-n}^n T^i(P' \vee \tilde{F})$ contains all the $T^i g$, $0 \leq i \leq n$, g being Q-n-F-atom in $\overline{F}$.

Because the measure of the union of good Q-n-atoms is bigger than $1 - \gamma/100$, $m(\bigcup_0^n T^i F) > 1 - \gamma/100$, because of our choice of γ, and because $H(\tilde{F}) < \gamma/100$ (because (3.4)) we get that $H(P', T) > H(R, S) - \overline{\gamma}$.

Since the names of the good R-n-atoms have the property that the u-names in them are distributed well we get

$$|\operatorname{dist} \bigvee_0^u T^i P' - \operatorname{dist} \bigvee_0^u S^i R| < \delta.$$

□

The following lemma and its corollary will be used to get the full generality of Lemma 3.1.7 , one of the main lemmas of this section.

Lemma 3.1.5. *Assume R, S and P, T are such that $\overline{d}((P,T),(R,S)) < \gamma$. Then given u and δ we can find $\overline{P}$ such that: $|\overline{P}| = |P|$, $|P - \overline{P}| < 2\gamma$ and*

$$|dist(\bigvee_0^u T^i \overline{P}) - dist(\bigvee_0^u T^i P)| < \delta$$

Proof. We can assume that $\delta < \gamma$

Pick n such that $\dfrac{u}{n} < \dfrac{1}{100}\delta$

Apply the R-K theorem, because T is ergodic, to find F so that:

1. $T^i F, 0 \leq i < n$ are disjoint.
2. $m(\bigcup_{i=0}^{n-1} T^i F) > 1 - \dfrac{1}{100}\delta$.
3. $\operatorname{dist} \bigvee_0^n T^{-i} P/F = \operatorname{dist} \bigvee_0^n T^{-i} P$

Because $\overline{d}((P,T),(R,S)) < \gamma$, and the definition of $\overline{d}$, we can find partitions R_i of F such that:

$$\operatorname{dist} \bigvee_0^{n-1} R_i/F = \operatorname{dist} \bigvee_0^{n-1} S^{-i} R \tag{3.6}$$

$$\frac{1}{n}\sum |R_i/F - T^{-i}P/F| < \gamma \tag{3.7}$$

Define $\overline{P}$ such that on T^iF, $\overline{P}$ and T^iR_i agree, it does not matter how to define it in the rest of the space because the definition of the gadget.

Because of (2) from the definition of the gadget and (3.7) we get $|\overline{P} - P| < 2\gamma$, since $m(X \setminus \bigcup T^{-i}F) < \dfrac{\delta}{100}$.

It is easy to see, by (3.6) and (2) from the R-K theorem, that if $0 \le i \le n-u$ then:

$$\operatorname{dist} \bigvee_0^u T^{-j}\overline{P}/T^iF = \operatorname{dist} \bigvee_0^u S^{-j}R. \tag{3.8}$$

Because of the choose of n, the definition of F and (3.8) imply the conclusions of our lemma.

□

Corollary 3.1.6. *Let R, S and P, T satisfy: $|R| = |P|$, R, S is finitely determined and $\overline{d}((P,T),(R,S)) < \gamma$. Then given δ we can find $\overline{P}$ such that $|\overline{P}| = |P|$, $|P - \overline{P}| < 2\gamma$ and either $H(\overline{P},T) < H(R,S)$ or $\overline{d}((\overline{P},T),(R,S)) < \delta$*

Proof. Because R, S is finitely determined, there exists u_1 and δ_1 so that if $|\overline{P}| = |P|$,

$$\left|\operatorname{dist} \bigvee_0^{u_1} T^i\overline{P} - \operatorname{dist} \bigvee_0^{u_1} S^iR\right| < \delta_1 \tag{3.9}$$

and $|H(R,S) - H(\overline{P},T)| < \delta_1$ then $\overline{d}((\overline{P},T),(R,S)) < \delta$.

Afterwards, let k be an integer, such that $u = ku_1$, if we have $\widehat{\delta} > 0$ such that $\widehat{\delta} < \delta_1$, then if $\overline{P}$ satisfied

$$\left|\operatorname{dist} \bigvee_0^{u} T^i\overline{P} - \operatorname{dist} \bigvee_0^{u} S^iR\right| < \widehat{\delta} \tag{3.10}$$

then $\overline{P}$ would also satisfy (3.9).

Since,

$$\left|\operatorname{dist} \bigvee_0^{u_1} T^i\overline{P} - \operatorname{dist} \bigvee_0^{u_1} S^iR\right| \le \left|\operatorname{dist} \bigvee_0^{u} T^i\overline{P} - \operatorname{dist} \bigvee_0^{u} S^iR\right| < \widehat{\delta} < \delta_1 \tag{3.11}$$

Because the definition of $H(R,S)$ we can choose appropriately u (big enough) and $\widehat{\delta}$ so that (3.10) implies $H(\overline{P},T) - H(R,S) < \delta_1$.

Applying the previous lemma to get $\overline{P}$ such that $|P| = |\overline{P}|$, $|P - \overline{P}| < 2\delta$ and satisfies (3.10).

Then if does not satisfy $H(\overline{P}, T) < H(R, S)$ it will satisfy $\overline{d}((\overline{P}, T), (R, S)) < \delta$

$\square$

Lemma 3.1.7. *Let R, S and P, T satisfy the following: $|R| = |P|$, R, S is finitely determined , $H(R, S) \leq H(T)$ and $\overline{d}((P, T), (R, S)) < (\frac{\epsilon}{100})^2$. Then given $\delta > 0$ we can find P' such that*

$$\overline{d}((P', T), (R, S)) < \delta \tag{3.12}$$

and $|P - P'| < \epsilon$

Proof. By the previous corollary we can assume that $H(P, T) < H(R, S)$, and $\overline{d}((R, S), (P, T)) < \frac{\epsilon^2}{100}$ (otherwise our problem is solved). Because S, R is finitely determined, it will be enough to show that given $\overline{\delta}$ and u we can find P' satisfying:

$$|\text{dist} \bigvee_0^u T^i P' - \text{dist} \bigvee_0^u S^i R| < \overline{\delta}, |H(P', T) - H(R, S)| < \overline{\delta} \text{ and } |P - P'| < \epsilon$$

Since $H(P, T) < H(R, S)$ we can choose a refinement Q of P such that $H(R, S) - H(Q, T) = \beta > 0$ where $\beta < \frac{\overline{\delta}}{100}$.

Because the nature of the function $-x \log x$ we can deform Q short enough to get a partition Q' such that the difference between its entropies is less than $\frac{\overline{\delta}}{100}$ i.e. we can choose $\gamma \leq \min(\overline{\delta}, \epsilon)$ such that $|Q' - Q| < \epsilon$ implies that $H(Q', T) \geq H(Q, T) - \frac{\overline{\delta}}{100}$.

We denote an atom in $\bigvee_0^n T^{-i} Q$ as *good* (or *good Q-n atoms*) if its measure is between $(\frac{1}{2})^{[H(Q,T) \pm \beta/100]n}$.

Denote an atom in $\bigvee_0^n S^{-i} R$ as *good* (or *good R-n atoms*) if its measure is between $(\frac{1}{2})^{[H(R,S) \pm \beta/100]n}$ and if its n-name have the property that the distribution of u-names in it is within $\overline{\delta}$ of dist $\bigvee_0^u S^{-i} R$ i.e. for each good $R - n$ atom we have:

$$\left| \text{dist} \bigvee_0^u S^{-i} R/A - \text{dist} \bigvee_0^u S^{-i} R \right| < \delta.$$

Because of the Shannon-McMillan-Breiman theorem and the ergodic theorem we can choose n so large that the measure of the union of atoms in $\bigvee_0^n T^{-i} Q$ and $\bigvee_0^n S^{-i} R$ that are not good is less than $\frac{\gamma}{100}$. We can also choose n so large that: $\frac{u}{n} < \frac{\overline{\delta}}{100}$, $n\beta > 100$ and $m(A) < \frac{1}{n}$ implies that $-[m(A) \log m(A) + (1 - m(A)) \log(1 - m(A))] < \frac{\overline{\delta}}{100}$.

Because R, S and P, T are close in the $\overline{d}$ metric we can choose partitions $\overline{R}_i$, and $0 \leq i \leq n$, such that dist $\bigvee_0^n \overline{R}_i =$ dist $\bigvee_0^n S^{-i}R$, and $\frac{1}{n+1}\sum_{i=0}^n |\overline{R}_i - T^{-i}P| < \frac{\epsilon^2}{100}$.

The good atoms in $\bigvee_0^n \overline{R}_i$ will be those that correspond to the good atoms in $\bigvee_0^n S^{-i}R$ and we will call these as *good* $\overline{R}$*-n-atoms*.

The $\overline{R}$-n-name of x will be defined with respect to $\bigvee_0^n \overline{R}_i$ (the i^{th} term being the atom of $\overline{R}_i$ that contains x).

We will pick out a subset of the good Q-n-atoms which we will call *very good* Q*-*n*-atoms*.

Let E be the set of x whose $\overline{R}$-n-name is good and differs from its P-n-name in $< \epsilon n$ places. We will call g a very good Q-n-atoms, if g is good and more than half of g ($>\frac{1}{2}$ measure of g) lies in E.

Because $\frac{1}{n+1}\sum_{i=0}^n |\overline{R}_i - T^{-i}P| < \frac{\epsilon^2}{100}$, we get that $m(E) > 1 - \epsilon$ and hence the measure of the union of very good $Q - n$-atoms is $> 1 - \frac{\epsilon}{2}$ since the definition of the very good $Q - n$-atoms.

Because the measure of a good Q-n-atom is greater than two times the measure of a good $\overline{R}$-n-atom, we have that any ℓ very good Q-n-atoms intersect at least ℓ good $\overline{R}$-n-atoms in E.

Applying the marriage lemma we can assign to each every good Q-n-atom a good $\overline{R}$-n-atom which intersects it in E. Assign to each very good Q-n-atom, g, a good $\overline{R}$-n-atom whose $\overline{R}$-n-name agrees with the P-n-name of any point in g, in more than $(1 - \frac{\epsilon}{n})n$ places.

Because there are more good $\overline{R}$-n-atoms than there are good Q-n-atoms, we can extend the above assignment so that each good Q-n-atom is assigned a good $\overline{R}$-n-atom.

Apply the R-K theorem to obtain a set F such that T^iF, $0 \leq i \leq n$ are disjoint and $m(\cup_0^n T^iF) > 1 - \frac{\gamma}{100}$ and dist $\bigvee_0^n T^{-i}Q/F =$ dist $\bigvee_0^n T^{-i}Q$.

Call the atoms in $\bigvee_0^n T^{-i}Q/F$ that are contained in good and very good Q-n-atoms, good and very good Q-n-F-atoms, we can therefore carry over our assignment of good $\overline{R}$-n-atoms to Q-n-F-atoms.

Define P' on $\cup_0^n T^i\overline{F}$ where $\overline{F}$ is the union of the good Q-n-F-atoms such as if g is a good Q-n-F-atom, then T^jg will lie in $(P')^i$ where i is the j^{th} term in the $\overline{R}$-n-atom assigned to g. where $(P')^i$ is the i^{th} atom of P', in the rest of the space we can define it in any way we want .

From the definition of P' we can see that every name in any atom A on P' is the same as the $\overline{R}$-name of the $\overline{R}$-n-atom assigned to A.

It does not matter how to define P' on the rest of X.

Let $\tilde{F}$ be the partition consisting of F and its complement. Because each Q-n-F-atom in $\overline{F}$ was assigned a different R-n-atom, we have that $\bigvee_{-n}^{n} T^i(P' \vee \tilde{F})$ restricted to $\cup_0^n T^i \overline{F}$ refines Q.

Because the measure of the union of good Q-n-atoms is $> 1 - \dfrac{\gamma}{100}$, $m(\cup_0^n T^i F) > \dfrac{\gamma}{100}$, because of our choice of γ and $H(\overline{F}) < \overline{\delta}/100$, we get that $H(P', T) > H(R, S) - \overline{\delta}$.

Because the names of the good $\overline{R}$-n-atoms have the property that the u-names in them are distributed well, we get

$$\left| \text{dist} \bigvee_0^u T^i P' - \text{dist} \bigvee_0^u S^i R \right| < \overline{\delta}$$

from the definition of P'.

Finally, since each every good Q-n-atom, g, has the property that the P-n-name of any point in g and the $\overline{R}$-n-name of the $\overline{R}$-n-atom assigned to g differ in less than $\dfrac{\epsilon}{2} n$-places, and since the measure of the union of the very good Q-n-atoms is $> 1 - \dfrac{\epsilon}{2}$ we have $|P' - P| < \epsilon$. (Note that the $\overline{P}$-n-name of x in $g \cap F$ is the $\overline{R}$-n-name of the $\overline{R}$-n-atom assigned to g).

□

The following two propositions are consequences from the previous lemma that will be vital to our main subject, it will provided us a partition close enough to a process finitely determined.

Proposition 3.1.8. *Let R, S and P, T satisfy the following: $|R| = |P|$, R, S is finitely determined $H(R, S) \leq H(T)$ and $\overline{d}((R, S), (P, T)) < (\dfrac{\epsilon}{100})^2$, then we can find a P' such that $|P' - P| < \epsilon$ and $P', T \sim R, S$.*

It follows immediately, by applying Lemma 3.1.7 a countable number of times.

Proposition 3.1.9. *Let R, S be finitely determined then given ϵ there is a δ and u such that if P, T satisfies $H(R, S) \leq H(T)$, $|P| = |R|$,*

$$|dist(\bigvee_0^u T^i P) - dist(\bigvee_0^u S^i R)| < \delta \tag{3.13}$$

and

$$|H(P, T) - H(R, S)| < \delta, \tag{3.14}$$

then there exists a P' such that $|P - P'| < \epsilon$ and $P', T \sim R, S$

Proof. Because of the definition of finitely determined we can find δ and u so that $|P| = |R|$, (3.13) and (3.14) imply $\bar{d}[(R,S),(P,T)] < (\frac{\epsilon}{100})^2$. We can now apply Proposition 3.1.8. $\square$

Lemma 3.1.10. *Let T be a transformation and P and Q partitions such that P generates $H(P,T) = H(Q,T)$ and P,T and Q,T are both finitely determined Then given ϵ we can find $\overline{Q}$ such that $\overline{Q},T \sim Q,T$, $|Q - \overline{Q}| < \epsilon$, and $\bigvee_{-\infty}^{\infty} T^i\overline{Q} \supset^{\epsilon} P$.*

Proof. **Claim:**

There is a $\overline{P}$ such that

1. $\overline{P} \subset \bigvee_{-\infty}^{\infty} T^i Q$.
2. $\overline{P},T \sim P,T$.
3. There is a K and operator L such that $\left|L\left(\bigvee_{-K}^{K} T^i P\right) - Q\right| < \frac{\epsilon}{2}$ and $\left|L\left(\bigvee_{-K}^{K} T^i \overline{P}\right) - Q\right| < \frac{\epsilon}{2}$

Because P generates, we can approximate Q long enough by a partition made by union of atoms such that its refinement is some $\bigvee_{-K}^{K} T^i P$ i.e. there is a K and L such that

$$\left|L\left(\bigvee_{-K}^{K} T^i P\right) - Q\right| < \frac{\epsilon}{100} \tag{3.15}$$

Applying Proposition 3.1.9 to T acting on $X_Q = (X, \bigvee_{-\infty}^{\infty} T^i Q, m)$, we get u and δ such that if P' satisfies $|P'| = |P|$, $|\operatorname{dist} \bigvee_0^u T^i P' - \operatorname{dist} \bigvee_0^u T^i P| < \delta$ and $|H(P',T) - H(P,T)| < \delta$ ($H(T) > H(P,T)$ is automatically by definition), then there is a $\overline{P}$ such that $\overline{P} \subset \bigvee_{-\infty}^{\infty} T^i Q$, $|\overline{P} - P'| < \epsilon/(300K)$ and $\overline{P},T \sim P,T$.

Choose $\bar{\epsilon}$ such that $\bar{\epsilon} < \delta, \bar{\epsilon} < \epsilon$, and if $|Q^*| = |Q|$ and $|Q^* - Q| < \bar{\epsilon}$, then $|H(Q^*,T) - H(Q,T)| < \delta/100$. Thus if P' is any partition such that $\bigvee_{-\infty}^{\infty} T^i P' \supset^{\bar{\epsilon}} Q$, there is a partition $Q' \subset \bigvee_{-\infty}^{\infty} T^i P'$ such that $|Q' - Q| < \bar{\epsilon}$, since there exist a refinement of $Q' \prec \bigvee_{-S}^{S} T^i P'$ it follows that $|H(P',T) - H(Q,T)| < \delta$.

We now choose K_1 such that ($K_1 > K$ and $K_1 > u$), Then

$$\bigvee_{-K_1}^{K_1} T^i P \supset^{\bar{\epsilon}/2} Q. \tag{3.16}$$

Choose N so that $K_1 < \frac{\bar{\epsilon}}{100} N$.

Apply the R-K theorem to find a set F such that $T^i F$, $0 \le i \le N$ are disjoint, and $m(\bigcup_0^N T^i F) > 1 - \bar{\epsilon}/100$, and $F \subset \bigvee_{-\infty}^{\infty} T^i Q$.

Let G_1 be the gadget $\bigcup_0^N T^iF, P \vee Q, T$. Choose P' such that $P' \subset \bigvee_{-\infty}^{\infty} T^iQ$ and $G_2 = \bigcup_0^N T^iF, P' \vee Q, T$ is isomorphic to G_1 because of Lemma 2.4.2 (extend the definition of P' to the rest of X in some arbitrary way).

Since $G_1 \sim G_2$, we get (from (3.15) and (3.16)).

$$\left| L\left(\bigvee_{-K}^{K} T^iP'\right) - Q\right| < \frac{2\epsilon}{100} \tag{3.17}$$

and

$$\bigvee_{-K_1}^{K_1} T^iP' \supset^{\bar{\epsilon}} Q. \tag{3.18}$$

We get (3.17) and (3.18) as follows:

$|L(\bigvee_{-K}^{K} T^iP) - Q|$ is the measure of the set of x such that x is in i^{th} atom of Q, and its $-K$ to K name $\{\alpha_i\}_{-K}^{K}$ lies in the j^{th} atom of L and $i \neq j$. (Here $\alpha_i = k$ if T^ix lies in the k^{th} atom of P).

Because $G_1 \sim G_2$, the contribution to $|L(\bigvee_{-K}^{k} T^iP') - Q|$ coming from $\bigcup_k^{N-K} T^iF$ is less than $\epsilon/100$.

Since $\bigcup_k^{N-K} T^iF$ is most of the space we get (3.17). The same argument gives (3.18).

It follows from $G_1 \sim G_2$ that $|\mathrm{dist} \bigvee_0^u T^iP' - \mathrm{dist} \bigvee_0^u T^iP| \leq \bar{\epsilon} \leq \delta$. It follows from 3.18 and the choice of $\bar{\epsilon}$, that $|H(P', T) - H(Q, T)| < \delta$.

From the definition of δ and u we get:

$$\overline{P} \subset \bigvee_{-\infty}^{\infty} T^iQ, \tag{3.19}$$

$$\overline{P}, T \sim P, T \tag{3.20}$$

and

$$|\overline{P} - P'| < \frac{\epsilon}{100K}. \tag{3.21}$$

From (3.21) and (3.17) we get

$$\left| L\left(\bigvee_{-K}^{K} T^i\overline{P}\right) - Q\right| < \epsilon \tag{3.22}$$

as we want it.

Because of (3.19), (3.20) and (3.22) we get our Claim.

From now on $u, \delta, \overline{\epsilon}, N, F, G_1$ and G_2 will have different values but we will still use the same notation, the symbols ϵ, K and L remains with the same meaning.

Due to $\overline{P}$ is a subset from $\bigvee_{-\infty}^{\infty} T^i Q$, we can find M such that

$$\bigvee_{-M}^{M} T^i Q \supset^{\overline{\frac{\epsilon}{100}}} \overline{P}. \tag{3.23}$$

As before, apply Proposition 3.1.9 to get a δ and u, such that if Q' satisfies $|Q'| = |Q|$, $|\operatorname{dist} \bigvee_0^u T^i Q' - \operatorname{dist} \bigvee_0^u T^i Q| < \delta$ and $|H(Q',T) - H(Q,T)| < \delta$, then there is a $\overline{Q}$ satisfying $|\overline{Q} - Q'| < \epsilon/(300M)$, and $\overline{Q}, T \sim Q, T$.

Choose $\overline{\epsilon}$ so that if $\bigvee_{-\infty}^{\infty} T^i Q' \supset^{\overline{\epsilon}} P$, then $|H(Q',T) - H(P,T)| < \delta$.

Now choose M_1 so that $(M_1 > M, M_1 > u)$

$$\bigvee_{-M_1}^{M_1} T^i Q \supset^{\overline{\epsilon}/2} \overline{P}. \tag{3.24}$$

Choose N so that $M_1 < (\epsilon N)/100$.

We now apply the R-K Theorem to find F such that $T^i F$ are disjoint $0 \le i \le N$, $m\left(\bigcup_0^N T^i F\right) > 1 - \overline{\epsilon}/100$ and

$$\operatorname{dist}\left(\bigvee_0^N T^i(\overline{P} \vee Q \vee F)\right) = \operatorname{dist}\left(\bigvee_0^N T^i(\overline{P} \vee Q \vee P)\right). \tag{3.25}$$

Let G_1 be the gadget, $\bigcup_0^N T^i F, \overline{P} \vee Q$. Define Q' so that the gadget $G_2 = \bigcup_0^N T^i F, P \vee Q'$ is isomorphic to G_1. (This can be done because the gadgets $\bigcup_0^N T^i F, P$ and $\bigcup_0^N T^i F, \overline{P}$ are isomorphic by (3.25), and the fact that $\overline{P}, T \sim P, T$).

Since $G_1 \sim G_2$, (3.23) and (3.24) imply

$$\bigvee_{-M}^{M} T^i Q' \supset^{\epsilon/5} P \tag{3.26}$$

and

$$\bigvee_{-M_1}^{M_1} T^i Q' \supset^{\overline{\epsilon}} P. \tag{3.27}$$

Because $G_1 \sim G_2$, $|\operatorname{dist} \bigvee_0^u T^i Q' - \operatorname{dist} \bigvee_0^u T^i Q| < \delta$

because of (3.27) and our choice of $\overline{\epsilon}$, $|H(Q',T) - H(P,T)| < \delta$.

Because of our choice of δ and u we can find a $\overline{Q}$ satisfying

$$|\overline{Q} - Q'| < \frac{\epsilon}{30M} \tag{3.28}$$

and $\overline{Q}, T \sim Q, T$.

Because of (3.26) and (3.28) we get:

$$\bigvee_{-M}^{M} T^i\overline{Q} \supset^{\epsilon} P. \tag{3.29}$$

Since $G_1 \sim G_2$ and $|L\left(\bigvee_{-K}^{K} T^i\overline{P}\right) - Q| < \epsilon/100$ (from our Claim), we get that $|L\left(\bigvee_{-K}^{K} T^iP\right) - Q| < (2\epsilon)/100$, and because $|Q' - \overline{Q}| < \epsilon/(30M)$, we get

$$\left|L\left(_{-K}^{K} T^iP\right) - \overline{Q}\right| < \frac{\epsilon}{2} \tag{3.30}$$

However

$$\left|L\left(\bigvee_{-K}^{K} T^iP\right) - Q\right| < \frac{\epsilon}{2} \tag{3.31}$$

(see the Claim)

Because of (3.30) and (3.31) we have:

$$|\overline{Q} - Q| < \epsilon. \tag{3.32}$$

Our lemma follows from the fact that $\overline{Q}, T \sim Q, T$,(3.29) and (3.32).

□

Proposition 3.1.11. *Let T be a transformation and P and Q partitions such that P generates, $H(P,T) = H(Q,T)$ and (P,T) and (Q,T) are finitely determined Then given ϵ there is a $\overline{Q}$ such that $|Q - \overline{Q}| < \epsilon$, $\overline{Q}$ generates and $\overline{Q}, T \sim Q, T$*

Proof. It follows from the previous lemma. Form Q_1 by the previous lemma so that: $Q_1, T \sim Q, T$, $|Q_1 - Q| < \epsilon_1$ and $\bigvee_{-N_1}^{N_1} T^iQ_1 \supset^{\epsilon_1} P$. We apply the lemma again to obtain Q_2, where $|Q_2 - Q_1| < \epsilon_2$, and ϵ_2 is so small hat $\bigvee_{-N_1}^{N_1} T^iQ_2 \supset^{(1+\frac{1}{2})\epsilon_1} P$ and $\bigvee_{-N_2}^{N_2} T^iQ_2 \supset^{\epsilon_2} P$. Applying the lemma once more, we get Q_3 such that $\bigvee_{-N_1}^{N_1} T^iQ_3 \supset^{(1+\frac{1}{2}+\frac{1}{2^2})\epsilon_1} P$, $\bigvee_{-N_2}^{N_2} T^iQ_3 \supset^{(1+\frac{1}{2})\epsilon_2} P$ and $\bigvee_{-N_3}^{N_3} T^iQ_3 \supset^{\epsilon_3} P$.

Continuing in this way we get that if ϵ_i are chosen as above and $\epsilon_i < \dfrac{1}{2^i}$, then Q_i converge to $\overline{Q}$ and $\overline{Q}, T \sim Q, T$. Furthermore:

$$\bigvee_{-N_1}^{N_1} T^i\overline{Q} \supset^{2\epsilon_1} P, \bigvee_{-N_2}^{N_2} T^i\overline{Q} \supset^{2\epsilon_2} P, \ldots$$

Therefore, $\overline{Q}$ will generate. □

Theorem 3.1.12. *If $H(T) = H(S)$ and if T and S have finite partitions which generate and are finitely determined, then T is isomorphic to S*

Proof. Let Q be such partition that generates and Q, S are finitely determined.Given $\epsilon > 0$, since $H(T) = H(Q, S)$ we have from the lemma 3.1.4 a partition P' such that $\bar{d}((Q, S), (P', T)) < (\epsilon/100)^2$.

It follows from Proposition 3.1.8 that we can find a partition P_1 such that $|P_1 - P'| < \epsilon$ and

$$P_1, T \sim Q, S \tag{3.33}$$

Therefore, because of (3.33) and the fact that Q, S is finitely determined it follows that P_1, T is finitely determined.

Let P be such partition that generates and P, T are finitely determined. We have that $H(P_1, T) = H(Q, S)$ because of (3.33) then using Proposition 3.1.11 there is a partition $\overline{P}$ that generates and

$$\overline{P}, T \sim P_1, T \tag{3.34}$$

It follows from (3.33) and (3.34) that $\overline{P}, T \sim Q, S$. Thus, because of Theorem 2.2.3 S and T are isomorphic in the measure sense.

□

Theorem 3.1.13. *Two Beroulli shifts with the same entropy are isomorphic*

Proof. This is a particular case of the previous theorem because of Corollary 3.1.3 and since the definition of Bernoulli shift the spaces involved are Lebesgue spaces. □

The Isomorphism theorem gives us information about Bernoulli shift like:

Corollary 3.1.14. *Bernoulli shifts have roots of all orders.*

Proof. Let T be a Bernoulli shift. Let T_1 be a Bernoulli shift such that $H(T_1) = \frac{1}{n}H(T)$. It is easy to see by the definition of entropy that $H(T_1^k) = kH(T_1)$.

Hence $T_1^k \sim T$. Hence T has a k^{th} root. □

3.2 Factors of Shift Bernoulli

In this section we will show that if T is Bernoulli, then for any partition P, P, T is finitely determined, then acting in $\bigvee_{-\infty}^{\infty} T^i P$ is Bernoulli shift. This will imply that factors of a Bernoulli shift are Bernoulli shifts. Additionally roots of Bernoulli shift will be Bernoulli.

We say that P, T is a *factor of Bernoulli shift*, when P is a partition of the space which is supported T and T is a Bernoulli shift.

Lemma 3.2.1. *Let T be a Bernoulli shift and P a finite partition such that $H(P, T) = H(T)$. Then given ϵ, there is a δ and n such that if P, T satisfies, $\overline{T}$ is ergodic, $H(\overline{T}) \geq H(T)$, $|P| = |\overline{P}|$*

$$|H(\overline{P}, \overline{T}) - H(P, T)| < \delta, \tag{3.35}$$

and

$$\left| \operatorname{dist} \bigvee_0^n T^i P - \operatorname{dist} \bigvee_0^n \overline{T}^i \overline{P} \right| < \delta. \tag{3.36}$$

Then there is a $\hat{P}$ such that $|\hat{P} - P| < \epsilon$ and $\hat{P}, \overline{T} \sim P, T$

Proof. Let B be a finite partition such that $T^i B$ are independent and generate.

Since B generate, because of Lemma 2.2.1 we can pick K so the $\bigvee_{-K}^{K} T^i B \supset^{\epsilon/10} P$. Let L be as in the above definition and such that $\left| L \left(\bigvee_{-K}^{K} T^i B \right) - P \right| < \epsilon/10$.

Choose δ' (using corollary 3.1.3 and proposition 3.1.9) such that if $|B'| = |B|$, $|H(B', \overline{T}) - H(B, T)| < \delta'$, and $|\operatorname{dist} B' - \operatorname{dist} B| < \delta'$, then there is a $\overline{B}$ such that $|\overline{B} - B'| < \epsilon/(100K)$, $\operatorname{dist} \overline{B} = \operatorname{dist} B$. Since B is independent and they have the same distribution we will get that $\overline{T}^i \overline{B}$ are independent.

Now take $\gamma, \gamma < \epsilon/100$, and $\gamma < \delta'/100$ such that if $|P_1| = |\overline{P}|$ and $|P_1 - \overline{P}| < \gamma$, then $|H(P_1, \overline{T}) - H(\overline{P}, \overline{T})| < \delta'/10$.

Select $K' > K$, so that $\bigvee_{-K'}^{K'} T^i B \supset^{\gamma/10} P$,and n such that $\frac{K'}{n} < \frac{\gamma}{100}$ and

$$\left| \frac{1}{n} \log \frac{1}{n} + (1 - \frac{1}{n}) \log(1 - \frac{1}{n}) \right| < \frac{\delta'}{10} \tag{3.37}$$

Note that so far our choice of n and γ depend only on T, P and B. We now choose δ so that $\delta < \frac{\omega}{100}$. Having chosen n and δ we will now assume that we have $\overline{P}, \overline{T}$ satisfying hypothesis $|\overline{P}| = |P|$,(3.35) and (3.36).

Now apply the R-K-Theorem to find a set F such that $T^i F, 0 \le i \le n-1$ are disjoint, and $m\left(X \setminus \bigcup_{i=0}^{n-1} T^i F\right) < \gamma/100$ (X is the space on which T acts) and

$$\operatorname{dist} \bigvee_0^{n-1} T^{-i}(P \vee B)/F = \operatorname{dist} \bigvee_0^{n-1} T^{-i}(P \vee B). \tag{3.38}$$

Applying the Rochlin's theorem again, we can find a set $\overline{F}$ such that $\overline{T}^i\overline{F}, 0 \le i \le n-1$, are disjoint $m\left(\overline{X} \setminus \bigcup_{i=0}^{n-1} \overline{T}^i \overline{F}\right) < \gamma/100$, and

$$\operatorname{dist} \bigvee_0^{n-1} \overline{T}^{-i}\overline{P}/\overline{F} = \operatorname{dist} \bigvee_0^{n-1} \overline{T}^{-i}\overline{P}. \tag{3.39}$$

Now because of our hypotheses (3.36) and our choice of n and δ, we could (by removing a set with measure less than $\delta m(\overline{F})$ from $\overline{F}$) assume that we have instead of (3.39) the following:

$$\operatorname{dist} \bigvee_0^{n-1} \overline{T}^{-i}\overline{P}/\overline{F} = \operatorname{dist} \bigvee_0^{n-1} T^{-i}P. \tag{3.40}$$

If we let G be the gadget formed by partitioning T^iF by P, and $\overline{G}$ the gadget formed by partitioning the $\overline{T}^i\overline{F}$ by $\overline{P}$, then because of (3.38) and (3.40) it follows that G and $\overline{G}$ are isomorphic.

Let G' be the gadget formed by partitioning T^iF by $P \vee B$.

Since G and $\overline{G}$ are isomorphic using Theorem ****, we can take a partition B' of $\bigcup_{i=0}^{n-1} \overline{T}^i\overline{F}$ such that if we form the gadget $\overline{G}'$ by partitioning the $\overline{T}^i\overline{F}$ by $\overline{P} \vee B'$, then $\overline{G}'$ is isomorphic to G'.(Now extend B' to the rest of the space in any way).

Because $\overline{G}'$ and G' are isomorphic and fill up must of the space, because $\frac{K}{n} < \frac{\gamma}{10}$, and because $|L(\bigvee_{-K}^{K} T^iB) - P| < \epsilon/10$ we get that

$$\left| L\left(\bigvee_{-K}^{K} \overline{T}^i B'\right) - \overline{P} \right| < \frac{2\epsilon}{10} \tag{3.41}$$

We also have $\bigvee_{-K'}^{K'} T^iB' \supset^{(2\gamma)/10} \overline{P}$, and hence by our choice of γ.

$$|H(B', \overline{T}) - H(\overline{P}, \overline{T})| < \frac{\delta'}{10} \tag{3.42}$$

Furthermore, since $\overline{G}'$ and G' are isomorphic and fill up must of the space, we get that

$$|\,\mathrm{dist}\, B' - \mathrm{dist}\, B| < \delta' \tag{3.43}$$

Since B is generator and $H(T) = H(P,T) = H(B,T)$(because Kolmogorov-Sinai Theorem) using hypothesis (3.35) we have that $|H(\overline{P},\overline{T}) - H(B,T)| < \delta < \delta'$

Using this fact, (3.43) , (3.42) and our choice of δ' imply that there is a $\overline{B}$ such that:

$$|\overline{B} - B'| < \frac{\epsilon}{100K} \tag{3.44}$$

and

$$\overline{B}, \overline{T} \sim B, T \tag{3.45}$$

Because of (3.41) and (3.44), we get that

$$\left| L\left(\bigvee_{-K}^{K} \overline{T}^i \overline{B} \right) - \overline{P} \right| < \frac{2\epsilon}{10} \tag{3.46}$$

Because of (3.45) and the Theorem 3.1.12, $\overline{T}$ acting on $\bigvee_{-\infty}^{\infty} \overline{T}^i \overline{B}$ is isomorphic to T under an isomorphism φ, of the measure algebras which takes B on $\overline{B}$. Let $\widehat{P}$ be $\varphi(P)$.

Then $P, T \sim \widehat{P}, \overline{T}$. Again, because of the isomorphism, and because L was defined so that $|L(\bigvee_{-K}^{K} T^i B) - P| < \epsilon/10$ we get that

$$\left| L\left(\bigvee_{-K}^{K} \overline{T}^i \overline{B} - \widehat{P} \right) \right| < \frac{\epsilon}{10}. \tag{3.47}$$

(3.46) and (3.47) imply $|\overline{P} - \widehat{P}| < \epsilon$. □

We will now modify the proof of the previous lemma to cover the case where $H(P,T) \leq H(T)$.

Theorem 3.2.2. *Let T be a Bernoulli shift, and let P be a finite partition. Then (P,T) is finitely determined*

Proof. We will show that P, T is finitely determined by proving the following:

Claim: Let $\widehat{T}$ be a Bernoulli shift acting on $\widehat{X}$ satisfying $H(\widehat{T}) > H(T)$, then given ϵ there is a δ and an n such that if $\overline{T}$ is ergodic and acts on $\overline{X}$, and if $|\overline{P}| = |P|$,

$$|H(\overline{P},\overline{T}) - H(P,T)| < \delta, \tag{3.48}$$

$$|\operatorname{dist} \bigvee_0^n \overline{T}^i\overline{P} - \operatorname{dist} \bigvee_0^n T^iP| < \delta \tag{3.49}$$

then there is a partition $\tilde{P}$ of $\overline{X}\times\widehat{X}$ such that $(\tilde{P},\overline{T}\times\widehat{T}) \sim P,T$ and $|\widehat{P}-(\overline{P}\times\widehat{X})| < \epsilon$. (See that $\overline{P}\times\widehat{X}$ is the partition of $\overline{X}\times\widehat{X}$ formed by partitioning the $\overline{X}$ coordinate by $\overline{P}$. Thus $\overline{P}\times\widehat{X},\overline{T}\times\hat{T} \sim \overline{P},\overline{T}$).

We will now prove it.

Since T is a Bernoulli shift there exist B an independent generator for T. Pick K so that $\bigvee_{-K}^K T^iB \supset^{\epsilon/10} P$. Define L so that $|L(\bigvee_{-K}^K T^iB) - P| < \epsilon/10$.

Choose δ_1 (using corollary 3.1.3 and proposition 3.1.9) so that if T_1 is ergodic, $H(T_1) \geq H(T)$ and $|B_1| = |B|, |H(B_1,T_1) - H(B,T)| < \delta_1$, and $|\operatorname{dist} B_1 - \operatorname{dist} B| < \delta_1$, then there is $\overline{B}$ such that $\overline{B},T_1 \sim B,T$ and $|\overline{B} - B_1| < \epsilon/(100K)$.

Take γ such that $\gamma < \epsilon/100$ and $\gamma < \delta_1/100$, and if P_2 and P_3 are any two partitions of the same space such that $|P_2| \leq |\overline{P}||B|$, $|P_3| \leq |\overline{P}||B|$ and $|P_2 - P_3| < \gamma$, then there is an R such that $H(R) < \delta_1/100$ and $P_2 \vee R \supset P_3$. Select K' such that $\bigvee_{-K'}^{K'} T^iB \supset^{\gamma/10} P$.

Let $\widehat{T}$ be a Bernoulli shift such that $H(\widehat{T}) > H(T)$. Choose $\widehat{Q}$ an independent partition under $\widehat{T}$ such that $|\widehat{Q}| \leq |B|$, and

$$H(T) - \frac{\delta_1}{10} < H(\widehat{Q}) + H(P,T) < H(T) - \frac{\delta_1}{100}. \tag{3.50}$$

If n is large enough, then Shannon-McMillan-Brieman theorem implies that there is a collection C_1 of atoms of $\bigvee_{-n+1}^0 \widehat{T}^i\widehat{Q}$, a collection C_2 of atoms of $\bigvee_{-n+1}^0 T^i(P\vee B)$ and a collection $\overline{C_2}$ of atoms of $\bigvee_{-n+1}^0 T^iP$ such that

- $m(C_1) > 1 - \frac{\gamma}{100}, m(C_2) > 1 - \frac{\gamma}{100}, m(\overline{C_2}) > 1 - \frac{\gamma}{100}.$
-
  - If $\overline{A_2} \in \overline{C_2}$, then $m(\overline{A_2}) > \frac{1}{2}^{\left[H(P,T)+\frac{\delta_1}{1000}\right]\cdot n}$
  - If $A_1 \in C_1$, then $m(A_1) > \frac{1}{2}^{\left[H(\widehat{T})+\frac{\delta_1}{1000}\right]\cdot n}$,
  - If $A_2 \in C_2$, then $m(A_2) < \frac{1}{2}^{\left[H(T)-\frac{\delta_1}{1000}\right]\cdot n}$,

Clearly any atom in C_2 is contained in an atom in $\overline{C_2}$. Thus if n is large enough we will have

$$\frac{\gamma}{100} m(A_1) m(\overline{A_2}) > m(A_2). \tag{3.51}$$

Pick an n so large, that in addition to the above conditions we have $\frac{K'}{n} < \frac{\gamma}{100}$ and

$$\left| \frac{1}{n} \log \frac{1}{n} + \left(1 - \frac{1}{n}\right) \log \left(1 - \frac{1}{n}\right) \right| < \frac{\gamma_1}{100}. \tag{3.52}$$

Choose $\delta < \frac{\gamma}{100}$. We have now chosen δ and n. Let us now assume we have a $\overline{P}, \overline{T}$ satisfying the conditions in the Claim. Apply the R-K theorem to obtain an F such that $T^i F, 0 \le i \le n-1$ are disjoint and fill up X to within $\frac{\gamma}{100}$ and

$$\operatorname{dist} \bigvee_{-n+1}^{0} T^i (P \vee B) = \operatorname{dist} \bigvee_{-n+1}^{0} T^i (P \vee B)/F \tag{3.53}$$

Applying the R-K theorem again and using (3.49) we get $\overline{F}$ such that $\overline{T}^i \overline{F}, 0 \le i \le n-1$ are disjoint and fill up $\overline{X}$ to within $\frac{\gamma}{100}$ and

$$\operatorname{dist} \bigvee_{-n+1}^{0} \overline{T}^i \overline{P}/\overline{F} = \operatorname{dist} \bigvee_{-n+1}^{0} T^i P \tag{3.54}$$

Let $P_1 = \overline{P} \times \widehat{X}$, $F_1 = \overline{F} \times \widehat{X}$ and let $T_1 = \overline{T} \times \widehat{T}$, $X_1 = \overline{X} \times \widehat{X}$ and $Q_1 = \overline{X} \times \widehat{Q}$. We thus have: $T_1^i F_1, 0 \le i \le n-1$ are disjoint and fill up X_1 to within $\frac{\gamma}{100}$ and

$$\operatorname{dist} \bigvee_{-n+1}^{0} T_1^i P_1/F_1 = \operatorname{dist} \bigvee_{-n+1}^{0} T^i P. \tag{3.55}$$

Let G be the gadget $\bigcup_{i=0}^{n-1} T^i F$ partitioned by P and $\overline{G}$, and $\bigcup_{i=0}^{n-1} T^i F$ partitioned by $P \vee B$. Let G_1 be the gadget $\bigvee_0^{n-1} T_1^i F_1$ partitioned by P_1.

Since $G \sim G_1$,because of (3.53) and (3.55), we can find B_1 such that if $\overline{G_1}$ is formed by $\bigcup_0^{n-1} T_1^i F_1$ partitioned by $P_1 \vee B_1$, then $\overline{G} \sim \overline{G_1}$,due the Theorem ***. All that is involved in checking that $\overline{G} \sim \overline{G_1}$ is the measure of atoms of $\bigvee_{-n+1}^{0} T_1^i (B_1 \vee P_1)/F_1$. We are still free to arrange the atoms of $\bigvee_{-n+1}^{0} T_1^i (B_1 \vee P_1)/F_1$ within the atoms of $\bigvee_{-n+1}^{0} T_1^i (P_1)/F_1$ in any way we want.

We will now show that B_1 could have been chosen with the additional property that

$$\bigvee_{-n}^{n} T_1^i (B_1 \vee P_1 \vee F_1) \supset^{\gamma/10} Q_1. \tag{3.56}$$

Here F_1 denotes the partition consisting of F_1 and its complement. Because of the definitions of C_2 and $\overline{C}_2$ we have that if A_1 is in $\bigvee_{-n+1}^{0} T_1^i Q_1/F_1$ and corresponds to an atom in C_1, and A_2 is an atom in $\bigvee_{-n+1}^{0} T_1^i(P_1 \vee B_1)/F_1$ and corresponds to an atom in C_2 and $\overline{A_2}$, the atom of $\bigvee_{-n+1}^{0} T_1^i P_1/F_1$ containing A_2, then

$$\frac{\gamma}{100} m(A_1 \cap \overline{A_2}) > m(A_2). \tag{3.57}$$

This means that we can change B_1 so that all but $\dfrac{\gamma}{50}$ of the atoms of $\bigvee_{-n+1}^{0} T_1^i(P_1 \vee B_1)/F_1$ that correspond to atoms in C_2 lie inside an atom of $\bigvee_{-n+1}^{0} T_1^i Q_1/F_1$.

Because $m(C_2) > 1 - \dfrac{\gamma}{100}$ we get that

$$\bigvee_{-n}^{n} T_1^i(B_1 \vee P_1 \vee F_1) \supset^{\gamma/10} Q_1. \tag{3.58}$$

Since $\overline{G} \sim \overline{G_1}$ we could define B_1 off $\bigcup_0^{n-1} T_1^i F_1$ so that dist $B =$ dist B_1.

We will now show that $|H(B_1, T_1) - H(B, T)| < \gamma_1$. (Since $H(B, T) = H(B) = H(B_1) \geq H(B_1, T_1)$, it is enough to show that $H(B_1, T_1) \geq H(B, T) - \gamma_1$).

Because $\overline{G_1} \sim \overline{G}$, $\bigvee_{-K'}^{K'} T^i B \supset^{\gamma/10} P$, and $\overline{G}$ and G_1 fill up most of X and X_1, we get that $\bigvee_{-K'}^{K'} T_1^i B_1 \supset^{\gamma} P_1$.

Because of our choice of γ, there is a partition R, $H(R) < \dfrac{\gamma_1}{100}$, such that $\bigvee_{-K'}^{K'} T_1^i(B_1 \vee R) \supset P_1$.

Thus $\bigvee_{-n}^{n} T_1^i(B_1 \vee P_1 \vee F_1) \supset^{\gamma/10} Q_1$, we have $\bigvee_{-\infty}^{\infty} T_1^i(B_1 \vee R \vee F_1) \supset P_1 \vee Q_1$.

It follows from (3.50) that

$$H(B_1 \vee R \vee F_1, T_1) \geq H(\widehat{Q}, \widehat{T}) + H(\overline{P}, \overline{T}) - \frac{\delta_1}{10} \geq H(T) - \frac{2\delta_1}{10}. \tag{3.59}$$

Hence $H(B_1, T_1) \geq H(T) - \delta_1$.

Because of our choice of δ_1 there is a $\overline{B}$ such that $|\overline{B} - B_1| < \dfrac{\epsilon}{100K}$ and $\overline{B}, T_1 \sim B, T$. Let φ be the isomorphism between T_1 acting on $\bigvee_{-\infty}^{\infty} T^i \overline{B}$ and T acting on $\bigvee_{-\infty}^{\infty} T^i B$ that takes B onto $\overline{B}$. Let $\tilde{P} = \varphi(P)$. It follows that $\tilde{P}, T_1 \sim P, T$.

Since $\overline{G} \sim \overline{G_1}$, and $\left|L\left(\bigvee_{-K}^{K} T^i B\right) - P\right| < \dfrac{\epsilon}{10}$, we get $\left|L\left(\bigvee_{-K}^{K} T_1^i B_1\right) - P_1\right| < \dfrac{2\epsilon}{10}$ and because $|B_1 - \overline{B}| < \dfrac{\epsilon}{100K}$, we get $\left|L\left(\bigvee_{-K}^{K} T_1^i \overline{B}\right) - P_1\right| < \dfrac{3\epsilon}{10}$.

On the other hand,

$$\left|L\left(\bigvee_{-K}^{K} T_1^i \overline{B}\right) - \tilde{P}\right| < \frac{\epsilon}{10}$$

(since $\varphi(B) = \overline{B}$ and $\varphi(P) = \tilde{P}$). Thus $|P_1 - \tilde{P}| < \epsilon$ and the Claim follows (recall that $P_1 = \overline{P} \times \widehat{X}$ and $T_1 = \overline{T} \times \widehat{T}$).

It follows directly from $|\widehat{P}-(\overline{P}\times\widehat{X})|<\epsilon$, $\overline{P}\times\widehat{X},\overline{T}\times\widehat{T}\sim\overline{P},\overline{T}$ and $\overline{P}\times\widehat{X},\overline{T}\times\widehat{T}\sim\overline{P},\overline{T}$ that $\overline{d}((P,T),(\overline{P},\overline{T}))<\epsilon$.

□

A direct consequence of this theorem is the following corollary.

Corollary 3.2.3. *Factors of Bernoulli shifts are Bernoulli shifts*

We can restated as: Let T be a Bernoulli shift such that $H(T)$ is finite. Let P be a finite partition. Then there is a finite partition B in $\bigvee_{-\infty}^{\infty}T^iP$ such that T^iB are independent and $\bigvee_{-\infty}^{\infty}T^iB=\bigvee_{-\infty}^{\infty}T^iP$.

As a application of the finitely determined characterization of arbitrary partitions under a Bernoulli shift we have:

Theorem 3.2.4. *If T^k is a Bernoulli shift then T is a Bernoulli shift*

Proof. Let P be a finite generator for T^k, and hence for T.

We must show that given ϵ we can find δ and u such that if $|\overline{P}|=|P|$,

$$|\mathrm{dist}(\bigvee_{1}^{kn}\overline{T}^i\overline{P})-\mathrm{dist}(\bigvee_{1}^{kn}T^iP)|<\delta \tag{3.60}$$

and

$$|H(P,T)-H(\overline{P},\overline{T})|<\delta, \tag{3.61}$$

then $\overline{d}[(P,T),(\overline{P},\overline{T})]<\epsilon$.

We have

$$|\mathrm{dist}(\bigvee_{1}^{n}(\overline{T}^k)^i(\bigvee_{1}^{k}\overline{T}^j\overline{P}))-\mathrm{dist}(\bigvee_{1}^{n}(T^k)^i(\bigvee_{1}^{k}T^jP))|=|\mathrm{dist}(\bigvee_{1}^{kn}\overline{T}^i\overline{P})-\mathrm{dist}(\bigvee_{1}^{kn}T^iP)|<\delta \tag{3.62}$$

and (3.61) implies

$$|H((\bigvee_{1}^{k}(\overline{T}^j\overline{P}),\overline{T}^k)-H((\bigvee_{1}^{k}T^jP),T^k)|<k\delta \tag{3.63}$$

Because T^k is Bernoulli, we have $\bigvee_1^kT^iP,T^k$ is finitely determined Thus δ and u can be chosen so that if $\overline{P},\overline{T}$ is such that (3.62) and (3.63) are satisfied (and $|\overline{P}|=|P|$)then

$$\overline{d}((\bigvee_{1}^{k}T^iP,T^k),(\bigvee_{1}^{k}\overline{T}^i\overline{P},\overline{T}^k))<\epsilon. \tag{3.64}$$

This implies $\overline{d}[(P,T),(\overline{P},\overline{T})] < \epsilon$

□

Definition 3.2.5. Let T be a Bernoulli shift and P an arbitrary partition. We say that the process (P,T) is a Bernoulli process or B-process.

The following Theorem show us that the Bernoulli property is preserved under approximations in the $\overline{d}$ metric over a sequence of Bernoulli process.

Theorem 3.2.6. *Let P_i, T_i and P, T processes such that: T_i are Bernoulli shifts,* $\lim_{i\to\infty} \overline{d}((P_i,T_i),(P,T)) = 0$. *Then acting on $\bigvee_{-\infty}^{\infty} T^i P$ is a Bernoulli shift.*

Proof. By Theorem 3.2.2, we get that P_i, T_i are finitely determined. (We can also assume that $|P_i| = |P|$.)

Let $\overline{T}$ be a Bernoulli shift such that $H(\overline{T}) \geq \sup H(P_i, T_i)$.

The proof will consist of showing that there is a $\overline{P}$ such that $\overline{P}, \overline{T} \sim P, T$.

By taking a subsequence of the P_i, T_i, we can assume that $\sum_1^{\infty} [\overline{d}((P_i,T_i),(P,T))]^{1/2} < \infty$

Because P_1, T_1 is finitely determined and $H(\overline{T}) \geq H(P_1, T_1)$, there is (by Lemma 3.1.8 and Proposition 3.1.4) a partition $\overline{P_1}$ such that $\overline{P_1}, \overline{T} \sim P_1, T_1$. Since P_2, T_2 is finitely determined, we can apply proposition 3.1.8 to get $\overline{P_2}$ such that

$$|\overline{P_1} - \overline{P_2}| < 100[\overline{d}((P_1,T_1),(P_2,T_2))]^{1/2}, \text{ and } \overline{P_2}, \overline{T} \sim P_2, T_2. \tag{3.65}$$

In general, if we have $\overline{P_i}$ such that $\overline{P_i}, \overline{T} \sim P_i, T_i$, then we can find $\overline{P_{i+1}}$ such that

$$|\overline{P_i} - \overline{P_{i+1}}| < 100[\overline{d}((P_i,T_i),(P_{i+1},T_{i+1}))]^{1/2}, \text{ and } \overline{P}_{i+1}, \overline{T} \sim P_{i+1}, T_{i+1}. \tag{3.66}$$

Since $\lim_{i\to\infty} \overline{d}((P_i,T_i),(P,T)) = 0$. Given $\epsilon > 0$ we can find n_0 such that $\overline{d}((P_i,T_i),(P,T)) < \epsilon$ for $i \geq n_0$.

It follows:

$$\overline{d}((\overline{P}_i,\overline{T}_i),(P,T)) \leq \overline{d}((\overline{P}_i,\overline{T}_i),(P_i,T_i)) + \overline{d}((P_i,T_i),(P,T)) < \epsilon \tag{3.67}$$

It is now easy to see that the $\overline{P_i}$ will converge to a partition $\overline{P}$ such tha $\overline{P}, \overline{T} \sim P, T$ (i.e. $\overline{d}((\overline{P},\overline{T}),(P,T)) = 0$).

□

The next section introduce us a new class of systems derived of the Bernoulli property that was introduced by Ornstein in [13]. This property is one of the most powerful methods to find whether a automorphism is Bernoulli or not.

3.3 Very Weak Bernoulli

The next subsection introduce us a new class of systems derived of the Bernoulli property that Ornstein used as a tool to know some cases when a system have the Bernoulli property.

Definition 3.3.1. P, T is very weak Bernoulli. If given ϵ, then: There is an n such that for all m we can find a collection C_m of atoms of $\bigvee_{-m}^{0} T^i P$ such that $m(C_m) > 1 - \epsilon$ and if $A \in C_m$, then $\overline{d}(\{T^i P\}_1^n, \{T^i P \setminus A\}_1^n) < \epsilon$.

The following lemma show us that if a process is sufficiently close in measure and entropy to a very weak Bernoulli process then it will be preserved the very weak Bernoulli property.

Lemma 3.3.2. *Assume P, T is very weak Bernoulli Then given ϵ, there is a u and δ such that if $\overline{P}, \overline{T}$ satisfies: $|\overline{P}| = |P|$,*

$$|dist(\bigvee_0^u T^i P) - dist(\bigvee_0^u \overline{T}^i \overline{P})| < \delta \tag{3.68}$$

and

$$|H(P,T) - H(\overline{P}, \overline{T})| < \delta, \tag{3.69}$$

then: There is an n and for all m there is a collection C_m of atoms of $\bigvee_{-m}^{0} \overline{T}^i \overline{P}$, $m(C_m) > 1 - \epsilon$, and if $A \subset C_m$, then $\overline{d}(\{\overline{T}^i \overline{P}\}_1^n, \overline{T}^i \overline{P} \setminus A\}_1^n) < \epsilon$.

Proof. Pick n from the very weak Bernoulli definition holds for (P, T).

By Lemma 3.1.1 we can pick δ_1 such that if $|Q| = |\bigvee_1^n T^i P|$, and if $|H(Q/Q_1) - H(Q)| < \delta_1$ then $Q \perp^{\epsilon^2} Q_1$.

Pick ℓ so that

$$\left| H\left(\bigvee_1^n T^i P / \bigvee_{-\ell}^{0} T^i P \right) - nH(P,T) \right| < \frac{\delta_1}{2}. \tag{3.70}$$

Let $u = n + \ell$.

If δ is small enough, then (3.70) and the hypotheses $|\overline{P}| = |P|$, (3.68) and (3.69) imply

$$\left|H\left(\bigvee_1^n \overline{T}^i\overline{P}/\bigvee_{-\ell}^0 \overline{T}^i\overline{P}\right) - nH(\overline{P},\overline{T})\right| \le \left|H\left(\bigvee_1^n \overline{T}^i\overline{P}/\bigvee_{-\ell}^0 \overline{T}^i\overline{P}\right) - H\left(\bigvee_1^n T^iP/\bigvee_{-\ell}^0 T^iP\right)\right|$$
$$+\left|H\left(\bigvee_1^n T^iP/\bigvee_{-\ell}^0 T^iP\right) - nH(P,T)\right|$$
$$+\left|nH(T,P) - nH(\overline{P},\overline{T})\right|$$

such that $\left|H\left(\bigvee_1^n \overline{T}^i\overline{P}/\bigvee_{-\ell}^0 \overline{T}^i\overline{P}\right) - nH(\overline{P},\overline{T})\right| < \delta_1$. This inequality can be restated as: for all $\overline{\ell} > 0$

$$\left|H\left(\bigvee_1^n \overline{T}^i\overline{P}/\bigvee_{-\ell}^0 \overline{T}^i\overline{P}\right) - H\left(\bigvee_1^n \overline{T}^i\overline{P}/\bigvee_{-\ell-\overline{\ell}}^0 \overline{T}^i\overline{P}\right)\right| < \delta_1. \tag{3.71}$$

Pick δ so small that the hypotheses of very weak Bernoulli, P and $\overline{P}$ have the same number of atoms and (3.68) imply there is a collection C_ℓ of atoms of $\vee_{-\ell}^0 \overline{T}^i\overline{P}$ such that $m(C_\ell) > 1 - 2\epsilon$, and if $A \in C_\ell$, then

$$\overline{d}(\{\overline{T}^i\overline{P}\}_1^n, \{\overline{T}^i\overline{P}/A\}_1^n) < 2\epsilon. \tag{3.72}$$

Because of the definition of δ_1 and (3.71), we have that for all $\overline{\ell} > 0$ there is a collection $\mathscr{C}_{\overline{\ell}}$ of atoms of $\bigvee_{-\ell}^0 \overline{T}^i\overline{P}$, $m(\mathscr{C}_{\overline{\ell}}) < \epsilon$, and if $A \subset \mathscr{C}_{\overline{\ell}}$, then $\bigvee_1^n \overline{T}^i P$ restricted to A. The lemma follows from this and (3.72).

□

Proposition 3.3.3. *If P, T is very weak Bernoulli then P, T is finitely determined.*

Proof. We must show that given ϵ there is a u and a δ such that if

1. $|\overline{P}| = |P|$.
2. $|\operatorname{dist} \bigvee_0^u T^iP - \operatorname{dist} \bigvee_0^u \overline{T}^i\overline{P}| < \delta$.
3. $|H(P,T) - H(\overline{P},\overline{T})| < \delta$

then $\overline{d}((P,T),(\overline{P},\overline{T})) < \epsilon$.

Choose u and δ by Lemma 3.1.1 so that the hypotheses (1),(2) and (3) imply that there is an n such for all m there is a collection C_m of atoms of $\bigvee_{-m}^0 \overline{T}^i\overline{P}$, $m(C_m) > 1-\epsilon/10$, and if $A \subset C_m$, then

$$\overline{d}(\{\overline{T}^i\overline{P}\}_1^n, \{\overline{T}^i\overline{P}/A\}_1^n) < \epsilon/10. \tag{3.73}$$

Because of these we can assume that the hypotheses (1),(2) and (3) imply

$$\overline{d}(\{T^iP\}_1^n, \{\overline{T}^i\overline{P}\}_1^n) < \frac{\epsilon}{10} \tag{3.74}$$

To show that $\overline{d}((P,T),(\overline{P},\overline{T})) < \epsilon$, we must, for each k, find sequences $\{P_i\}_1^k, \{\overline{P}_i\}_1^k$ of partitions of the same space such that

$$\text{dist} \bigvee_1^k P_i = \text{dist} \bigvee_1^k T^iP, \text{dist} \bigvee_1^k \overline{P_i} = \text{dist} \bigvee_1^k \overline{T}^i\overline{P} \text{ and } \frac{1}{k}\sum |P_i - \overline{P_i}| < \epsilon. \tag{3.75}$$

We will prove it inductively. Assume we have found the first k partitions. We will now define $\{P_i\}_{k+1}^{k+n}, \{\overline{P_i}\}_{k+1}^{k+n}$.

Let A be an atom of $\bigvee_1^k P_i$, $\overline{A}$ an atom of $\bigvee_1^k \overline{P_i}$, and let A_1 and $\overline{A_1}$ be the atoms in $\bigvee_1^k T^iP$ and $\bigvee_1^k \overline{T}^i\overline{P}$ corresponding to A and $\overline{A}$.

We will define $\{P_i\}_{k+1^{k+n}}$ on $A \cap \overline{A}$ so that the distribution of $\bigvee_{k+1}^{k+n} P_i$ on $A \cap 0\overline{A}$ is the same distribution of $\bigvee_{k+a}^{k+n} T^iP$ on A. This will insure that

$$\text{dist} \bigvee_1^{k+n} P_i = \text{dist} \bigvee_1^{k+n} T^iP.$$

We will define $\{\overline{P_i}\}_{k+1}^{k+n} T^iP$ on $A \cap \overline{A}$ in the analogous way.

Because of the definition of u and δ we have that there is a collection $\mathscr{C}$ of atoms of $\bigvee_1^k (P_i \vee \overline{P_i}), m(\mathscr{C}) > 1 - 2\epsilon/10$ and if $A \cap \overline{A} \subset \mathscr{C}$, then

$$\overline{d}(\{P_i\}_{k+1}^{k+n}, \{P_i/(A\cap\overline{A})\}_{k+1}^{k+n}) < \frac{\epsilon}{10} \text{ and } \overline{d}(\{\overline{P_i}\}_{k+1}^{k+n}, \{\overline{P_i}/(A\cap\overline{A})\}_{k+1}^{k+n}) < \frac{\epsilon}{10} \tag{3.76}$$

Thus by (3.73) we have $\overline{d}(\{P_i/(A\cap\overline{A})\}_{k+1}^{k+n}, \{\overline{P_i}/(A\cap\overline{A})\}_{k+1}^{k+n}) < \dfrac{3\epsilon}{10}$.

This last inequality implies, that on $A \cap \overline{A}$, $\{P_i\}_{k+1}^{k+n}$ and $\{\overline{P_i}\}_{k+1}^{k+n}$ can be chosen so that

$$\frac{1}{n}\sum_{i=k+1}^{k+n} |P_i/(A\cap\overline{A}) - \overline{P_i}/(A\cap\overline{A})| < \frac{3\epsilon}{10} \tag{3.77}$$

This in turn implies $\dfrac{1}{n}\displaystyle\sum_{i=k+1}^{k+n} |P_i - \overline{P_i}| < \epsilon$.

Which proves (3.75), and the proposition follows. □

3.4 Some criterion for Bernoulli shifts

The main object of this section is to proof that an increasing sequence of Bernoulli process implies the Bernoulli property in the whole space.

Lemma 3.4.1. *Let T be a Bernoulli shift of finite entropy, and P a finite partition such that $H(P,T) < H(T)$, and T^iP are independent. Then given ϵ there are P' and R' such that $P',T \sim P,T$, $|P - P'| < \epsilon$, $P' \vee R'$ generates and the σ-algebras $\bigvee_{-\infty}^{\infty} T^iP'$ and $\bigvee_{-\infty}^{\infty} T^iR'$ are independent and T^iR' are independent.*

Remark R' can be chosen to have any distribution whose entropy equals $H(T) - H(P,T)$. This follows from the isomorphism theorem for Bernoulli shifts.

Proof. Let $\alpha = H(T) - H(P,T)$. Because of the isomorphism theorem we can find $\overline{T}$ a Bernoulli shift with independent generator $\overline{P} \vee \overline{R}$ where $\overline{P} \perp \overline{R}$ and $\operatorname{dist} P = \operatorname{dist} \overline{P}$ and $H(\overline{R}) = \alpha$. Then $H(\overline{T}) = H(T)$.

Since T^iR is independent, because of corollary 3.1.3, P,T is finitely determined. Using proposition 3.1.10 and proposition 3.1.11 , there is a δ such that if R is a partition satisfying $|R| = |\overline{R}|$

$$|H(R \vee P, T) - H(\overline{R} \vee \overline{P}, \overline{T})| < \delta \tag{3.78}$$

and

$$|\operatorname{dist} R \vee P_1 - \operatorname{dist} \overline{R} \vee \overline{P}| < \delta \tag{3.79}$$

then there are partitions R_1 and P_1 such that $|R_1 \vee P_1 - R \vee P| < \dfrac{\epsilon}{2}$, $R_1 \vee P_1, T \sim \overline{R} \vee \overline{P}, \overline{T}$, and $R_1 \vee P_1$ generate.

We can find a $\delta_1 < \delta$ such that if

$$|H(R \vee P, T) - H(\overline{R} \vee \overline{P}, \overline{T})| < \delta_1 \tag{3.80}$$

and

$$|\operatorname{dist} R - \operatorname{dist} \overline{R}| < \delta_1 \tag{3.81}$$

then

$$\operatorname{dist} R \vee P - \operatorname{dist} \overline{R} \vee \overline{P}| < 2\delta. \tag{3.82}$$

We can see this as follows: Using (3.80) we get $H(R \vee P) \geq H(R \vee P, T) \geq H(T) - \delta_1 \geq H(P) + H(\overline{R}) - \delta_1$.

Th equation (3.81) implies that if δ_1 is small enough, $H(R \vee P)$ is so close to $H(P) + H(R)$ that R must be $\dfrac{\delta}{10}$ independent of P (because of Lemma 3.1.1). This and (3.81) implies $|\operatorname{dist} R \vee P - \operatorname{dist} \overline{R} \vee \overline{P}| < 2\delta$

The rest of the proof will be devoted to finding an R satisfying $|R| = |\overline{R}|$,(3.80) and (3.81).

Let Q be a refinement of P such that:

$$H(T) - \frac{\gamma_1}{100} < H(Q, T) < H(T). \tag{3.83}$$

Define $\beta = H(T) - H(Q, T)$.

Pick $\gamma < \delta_1$ so that if $|Q_1 - Q| < \gamma$, then

$$|H(Q_1, T) - H(Q, T)| < \frac{\delta_1}{100}. \tag{3.84}$$

Call an atom of $\bigvee_{-n+1}^{0} T^i P$ a good P-n-atom if its measure is between $\dfrac{1}{2}^{(H(P,T)-\beta/100)n}$. Call an atom of $\bigvee_{-n+1}^{0} \overline{T}^i \overline{R}$ a good $\overline{R}$-n-atom if its measure is between $\dfrac{1}{2}^{(H(\overline{R},\overline{T})\pm\beta/100)n}$, and if the distribution of its $\overline{R}$-n-name is within $\dfrac{\gamma_1}{100}$ of the distribution of $\overline{R}$. (That is: each atom of $\bigvee_{-n+1}^{0} T^i \overline{R} = \bigcap_{-n+1}^{0} T^i \overline{R} f^{(i)}$ when the $\overline{R} f^{(i)}$ are atoms of $\overline{R}$. We are requiring that the number of $f^{(i)}$ that equal j divided by n be close to the measure of $\overline{R}^j$).

Call a good P-n-atom very good if more than $1 - \dfrac{\gamma}{100}$ of it is covered by good Q-n-atoms.

Call the good Q-n-atoms in a very good P-n-atom very good. The Shannon-McMillan-Brieman theorem implies that if n is large enough, then: All but $1 - \dfrac{\gamma}{100}$ of the Q-n-atoms are very good. It also implies that:

Each very good P-n-atom contains fewer very good Q-n-atoms than there are good $\overline{R}$ atoms (We can see this, since: If n is large enough, the number of good $\overline{R}$-n-atoms are greater than $2^{(H(\overline{R},\overline{T})-\beta/100)n}$ which is equal to $2^{H(T)-(H(P,T)+\beta/100)n}$. The number of very good Q-n-atoms in a very good P-n-atom is smaller than $2^{(E(Q,T)-H(P,T)+(2\beta)/100)n}$).

Because of that we can assign to each very good Q-n-atom in a fixed very good P-n-atom a good $\overline{R}$-n-atom (so that no 2 Q-n-atoms in the same P-n-atom are assigned the same $\overline{R}$-n-atom).

Using the R-K theorem we can find a set F such that $T^i F, 0 \leq i \leq n-1$ are disjoint and $m\left(\bigcup_{i=0}^{n-1} T^i F\right) \geq 1 - \dfrac{\gamma}{100}$, and such that

$$\operatorname{dist}\left(\bigvee_{-n+1}^{0} T^i Q/F\right) = \operatorname{dist} \bigvee_{-n+1}^{0} T^i Q$$

Call the part of Q-n-atom that lies in F a Q-n-F-atom. Make the same definition for good and very good Q-n-atoms and P-n-atoms.

We will now define R as follows: Let A be a very good Q-n-F-atom.

Let $\bigcap_{-n+1}^{0} \overline{T}^i \overline{R} f^{(i)}$ be the good $\overline{R}$-n-atom assigned to A. Then $T^i A$ will be in $\overline{R} f^{(i)}$. (Thus the R-n-name of points in A will be the same as the $\overline{R}$-n-name of points in the $\overline{R}$-n-atoms assigned to A).

Let F_1 be the union of the very good Q-n-F-atoms. We have defined R on $\bigcup_0^{n-1} T^i F_1$. Because of the definition of the very good Q-n-atoms we have that

$$m(F_1) > (1 - \frac{\gamma}{100}) m(F).$$

Afterwards, R satisfies that $|R| = |\overline{R}|$ and (3.81) i.e. has good distribution, because the distribution of the R-n-name of each point in F_1 will itself have a distribution within $\frac{\delta_1}{100}$ of the dist $\overline{R}$. This shows that

$$|\operatorname{dist} R - \operatorname{dist} \overline{R}| < \delta_1, \tag{3.85}$$

because $m\left(\bigcup_0^{n-1} T^i F\right) > 1 - \frac{\gamma}{100}$ and $m(F_1) > \left(1 - \frac{\gamma}{100}\right) m(F)$.

Also $R \vee P$ satisfies (3.80) because $\bigvee_{-n}^{n} T^i(R \vee P \vee F)$ refines Q on $\bigcap_0^{n-1} T^i F_1$ (here F denotes the partition consisting on F and its complement).

Thus $\bigvee_{-n}^{n} T^i(R \vee P \vee F) \subset^{\gamma} Q$. By the choice of γ and (3.83)

$$H(R \vee P \vee F, T) > H(Q, T) - \frac{\delta_1}{100} \geq H(T) - \frac{2\delta_1}{100}.$$

Since $H(F) < \frac{\delta_1}{100}$(by the choice of n), and the definition of $\overline{P} \vee \overline{R}$ we get that

$$H(R \vee P, T) \geq H(T) - \delta_1 = H(\overline{R} \vee \overline{P}, \overline{T}) - \delta_1,$$

and hence using (3.85) and the independence of $\overline{T}^i(\overline{R} \vee \overline{P})$,

$$|H(R \vee P, T) - H(\overline{R} \vee \overline{P}, \overline{T})| < \delta_1. \tag{3.86}$$

Because of (3.82), we can use the first part of the proof to find R' and P' such that

$$R' \vee P', T \sim \overline{R} \vee \overline{P}, \overline{T} \tag{3.87}$$

$$|R' \vee P' - R \vee P| < \frac{\epsilon}{2} \tag{3.88}$$

and

$$R' \vee P' \text{ generate.} \tag{3.89}$$

Finally because of (3.87), (3.88) and (3.89) together with the properties of $\overline{P}$, $\overline{R}$ and $\overline{T}$ imply that P' and R' satisfy the conclusions of the lemma.

□

Definition 3.4.2. We say that T is an increasing union of Bernoulli shifts if there is an increasing sequence of invariant σ-algebras σ_i such that T restricted to each σ_i is a Bernoulli shift of finite entropy, and the smallest σ-algebra containing all the σ_i is the entire σ-algebra.

Lemma 3.4.3. *Let T and $\overline{T}$ be increasing unions of Bernoulli shifts with invariant σ-algebras σ_i and $\overline{\sigma}_i$ respectively. We will assume that the entropy of T restricted to σ_i is the same as the entropy of $\overline{T}$ restricted to $\overline{\sigma}_i$. Then T and $\overline{T}$ are isomorphic(in the measure algebra sense, and if T and $\overline{T}$ act on Lebesgue space and $\bigcup \sigma_i$ and $\bigcup \overline{\sigma}_i$ separate points, then T and $\overline{T}$ are isomorphic in the pointwise sense.)*

Proof. Because of the isomorphism theorem for Bernoulli shifts we can pick $P_{1,1}$ and $\overline{P}_{1,1}$ such that $P_{1,1}, T \sim \overline{P}_{1,1}, \overline{T}$, $P_{1,1}$ generates σ_1, and $\overline{P}_{1,1}$ generates $\overline{\sigma_1}$.

We now apply the previous lemma to σ_2 and $\overline{\sigma_2}$ to get:

Given ϵ_1 there exist $P_{1,2}, P_{2,2}$ and $\overline{P}_{1,2}, \overline{P}_{2,2}$ satisfying

$$|P_{1,2} - P_{1,1}| < \epsilon_1 \text{ and } |\overline{P}_{1,2} - \overline{P}_{1,1}| < \epsilon_1.$$

$$P_{1,2} \vee P_{2,2} \text{ generates } \sigma_2 \text{ and } \overline{P}_{1,2} \vee \overline{P}_{2,2} \text{ generates } \overline{\sigma_2}.$$

We will also have: $\bigvee_{-\infty}^{\infty} T^i(P_{1,2}) \perp \bigvee_{-\infty}^{\infty} T^i(P_{2,2})$, $\bigvee_{-\infty}^{\infty} \overline{T}^i(\overline{P}_{1,2}) \perp \bigvee_{-\infty}^{\infty} \overline{T}^i(\overline{P}_{2,2})$ and $P_{1,2}, T \sim \overline{P}_{1,2}, \overline{T}$, $P_{1,1}, T \sim \overline{P}_{1,1}, \overline{T}$, $P_{2,2}, T \sim \overline{P}_{2,2}, \overline{T}$. Thus we get $(P_{1,2} \vee P_{2,2}), T \sim (\overline{P}_{1,2} \vee \overline{P}_{2,2}), \overline{T}$.

If we repeat this process once more we get $P_{1,3}, P_{2,3}, P_{3,3}$ and $\overline{P}_{1,3}, \overline{P}_{2,3}, \overline{P}_{3,3}$ satisfying

$$|P_{1,3} \vee P_{2,3} - P_{1,2} \vee P_{2,2}| < \epsilon_2$$

$$\left(\bigvee_{i=1}^{3} P_{i,3}\right) \text{ generate } \sigma_3 \text{ and } \left(\bigvee_{i=1}^{3} \overline{P}_{i,3}\right) \text{ generate } \overline{\sigma_3}$$

and

$$\left(\bigvee_{i=1}^{3} P_{i,3}\right), T \sim \left(\bigvee_{i=1}^{3} \overline{P}_{i,3}\right), \overline{T}.$$

Continuing in this way we get $P_{i,n}$ and $\overline{P}_{i,n}$, $1 \le i \le n$ satisfying

$$\left|\left(\bigvee_{i=1}^{n-1} P_{i,n-1}\right) - \left(\bigvee_{i=1}^{n-1} P_{i,n}\right)\right| < \epsilon_{n-1}, \left|\left(\bigvee_{i=1}^{n-1} \overline{P}_{i,n-1}\right) - \left(\bigvee_{i=1}^{n-1} P_{i,n}\right)\right| < \epsilon_{n-1} \tag{3.90}$$

.

$$\bigvee_{i=1}^{n} P_{i,n} \text{ generate } \sigma_n \text{ and } \bigvee_{i=1}^{n} \overline{P}_{i,n} \text{ generate } \overline{\sigma_n}$$

$$\left(\bigvee_{i=1}^{n} P_{i,n}\right), T \sim \left(\bigvee_{i=1}^{n} \overline{P}_{i,n}\right), \overline{T}. \tag{3.91}$$

If the ϵ_i are chosen small enough, $P_{i,n}$ converge (as $n \longrightarrow \infty$, i fixed) to P_i, and it is clear that $\left(\bigvee_{i=1}^{n} P_i\right), T \sim \left(\bigvee_{i=1}^{n} \overline{P}_i\right), \overline{T}$.

We will be finished if we show that $\bigvee_1^\infty P_i$ and $\bigvee_1^\infty \overline{P}_i$ generate.

We will do this as follows: Let Q_n be a generator for σ_n such that Q_{n+1} refines Q_n.

Thus $\bigvee_{-\infty}^{\infty} T^i \left(\bigvee_{j=1}^{n} P_{j,n}\right) \supset Q_n$ and for some K_n $\bigvee_{-K_n}^{K_n} T^i \left(\bigvee_{j=1}^{n} P_{j,n}\right) \supset Q_n$.

If we choose $\epsilon_n, \epsilon_{n+1}$, etc. Properly, we get that $\bigvee_{-K_n}^{K_n} T^i \left(\bigvee_{j=1}^{n} P_j\right) \supset^{1/(2^{n-1})} Q_n$.

This shows that $\bigvee_1^\infty P_i$ generates, because of Lemma 2.2.1. The same argument works for $\overline{T}$.

□

Theorem 3.4.4. *If $H(T) < \infty$ and T is an increasing union of Bernoulli shifts, then T is a Bernoulli shift(i.e. has a finite independent generator).*

Observation: The above theorem allows us to prove that a transformation is Bernoulli without first finding a generator.

Proof. The theorem follows from our previous lemma and the following:

1. If T is a Bernoulli shift and if r_i are numbers such that $r_{i+1} > r_i$ and $\sup r_i = H(T)$, then we can find invariant σ-algebra σ_i such that T restricted to σ_i is a Bernoulli shift of entropy r_i.

   We will get (1) immediately from the theorem that factors of Bernoulli shifts are Bernoulli shifts and from:

2. Let T be a Bernoulli shift and let σ_1 be an invariant σ-algebra such that the entropy of T restricted to σ_1 (denoted by $H(\sigma_1)$) is strictly less than $H(T)$, then given r, $H(\sigma_1) < r < H(T)$ we can find an invariant σ-algebra σ_2 such that $\sigma_2 \supset \sigma_1$ and $H(\sigma_2) = r$.

We get (2) as follows: Let P_1 be a generator for T acting on σ_1 and P_2 a generator for T. Deform P_2 continuously, through P^t, into a partition with only one atom. Then $H(P_1 \vee P^t, T)$ will go continuously from $H(T)$ to $H(\sigma_1)$.

□

3.5 Estimating distances between partition

Let $(X, \mathcal{B}, \mu)$ a probability space, $\varphi : X \longrightarrow X$ an invertible measure-preserving transformation. Let $\alpha = \{A_1, A_2..., A_a\}$ and $\beta = \{B_1, B_2, ..., B_b\}$ partitions of X. Remembering the metric defined in the preliminaries 2.5 we will use the metric

$$d(\alpha, \beta) = \sum_i \mu(A_i \Delta B_i).$$

In the case when the partitions $\alpha_i = \{A_1^{(i)}, ..., A_{a_i}^{(i)}\}$ and $\beta_i = \{B_1^{(i)}, ..., B_{b_i}^{(i)}\}$, $1 \leq i \leq n$ are on different spaces we could compare their distributions, we will write $\{\alpha_i\}_1^n \sim \{\beta_i\}_1^n$, if for all k_i, $1 \leq i \leq n$

$$\mu(\cap_1^n A_{k_i}^{(i)}) = \nu(\cap_1^n B_{k_i}^{(i)})$$

where μ and ν are the measures on the spaces X and Y respectively.

With this we can reformulate our metric $\overline{d}$ between partitions α_i in X and β_i in Y.

We will write $\overline{d}((\alpha_i)_1^n, (\beta_i)_1^n) \leq \epsilon$ if there are partitions $\overline{\alpha}_i$, $\overline{\beta}_i$ partitions on the same space such that $(\overline{\alpha}_i)_1^n \sim (\alpha_i)_1^n$ and $(\overline{\beta}_i)_1^n \sim (\beta_i)_1^n$ and

$$\frac{1}{n} \sum_1^n d(\overline{\alpha}_i, \overline{\beta}_i) \leq \epsilon.$$

In the last two sections we showed two strong methods: the Corollary **??** and Theorem 3.4.4 to determine if an automorphism is Bernoulli. We can reformulate it in the following way.

Theorem 3.5.1. *If α is very weak Bernoulli, then $(X, \bigvee_{-\infty}^{\infty} \varphi^n \alpha, \mu\varphi)$ is a B-shift.*

Theorem 3.5.2. *If $A_1 \subset A_2 \subset ...$ are an increasing sequence of φ-invariant σ-algebras $\bigvee_1^\infty A_n = B$ and for each n (X, A_n, μ, φ) is a B-shift, then (X, B, μ, φ) is a B-shift.*

To apply theorem 3.5.1 we need a method for showing that partitions are close in the $\overline{d}$-metric, and we devote the rest of the section to this point.

The $\{\alpha_i\}_1^n$-name of x is the sequence $\ell_i = \ell_i(x)$ determined by $x \in A_{\ell_i}^{(i)}, \alpha_i = \{A_1^i, A_2^i, ...A_{a_i}^i\}$.

Definition 3.5.3. We say that a transformation $\theta : (X, \mu) \longrightarrow (Y, \nu)$ between Lebesgue spaces *ϵ-preserves measure* if there is a set $E \subset X$ such that $\mu(E) \leq \epsilon$ and for every measurable set $A \subset X \setminus E$,

$$\left| \frac{\nu(\theta(A))}{\mu(A)} - 1 \right| \leq \epsilon.$$

Define a function $e : \mathbb{N} \longrightarrow \{0, 1\}$ defined by $e(0) = 0$ and $e(n) = 1$ when $n > 0$.

Lemma 3.5.4. *Let α and β two finite partitions defined in X with the names $\ell(x)$ and $m(x)$ respectively. Consider $f : X \longrightarrow \{0, 1\}$ defined by:*

$$f(x) = 1, if\ \ell(x) \neq m(x)$$
$$f(x) = 0, if\ \ell(x) = m(x).$$

Then

$$d(\alpha, \beta) = 2 \int f(x) d\mu(x).$$

Proof. Since the metric of the partitions α and β are defined as

$$d(\alpha, \beta) = \sum_{j=1}^{k} \mu(A_j \Delta B_j)$$

where k is the number of atoms in α and β.

Dividing $A_j \Delta B_j$ we get:

- If $x \in A_j \cap B_j$ then $f(x) = 0$, since $\ell(x) = j = m(x)$,
- If $x \in A_j \setminus B_j$ then $f(x) = 1$, since $\ell(x) = j \neq m(x)$,
- If $x \in B_j \setminus A_j$ then $f(x) = 1$, since $\ell(x) \neq j = m(x)$,

Furthermore:

$$\begin{aligned}
\int_{A_j} f(x)d\mu + \int_{B_j} f(x)d\mu &= \int_{A_j \setminus B_j} f(x)d\mu + 2 \cdot \int_{A_j \cap B_j} f(x)d\mu + \int_{B_j \setminus A_j} f(x)d\mu \\
&= \int_{A_j \setminus B_j} 1d\mu + 2 \cdot 0 + \int_{B_j \setminus A_j} 1d\mu \\
&= \mu(A_j \setminus B_j) + u(B_j \setminus A_j) \\
&= \mu(A_j \Delta B_j).
\end{aligned}$$

Then,

$$\begin{aligned}
d(\alpha, \beta) = \sum_{j=1}^{k} \mu(A_j \Delta B_j) &= \sum_{j=1}^{k} \int_{B_j} f(x)d\mu + \sum_{j=1}^{k} \int_{A_j} f(x)d\mu \\
&= \int_{\bigcup_{j=1}^{n} A_j} f(x)d\mu + \int_{\bigcup_{j=1}^{n} B_j} f(x)d\mu = 2 \int f(x)d\mu
\end{aligned}$$

□

Lemma 3.5.5. *Let $\{\alpha_i\}_1^n$ be partitions of X with name functions $\ell_i(x)$ and $\{\beta_i\}_1^n$ partitions of Y with name functions $m_i(y)$. If there is a measure-preserving mapping $\theta : X \to Y$ such that*

1. $\mu(E) \leq \epsilon$ *and*
2. $\dfrac{1}{n}\sum_{i=1}^{n} e(\ell_i(x) - m_i(\theta x)) \leq \epsilon$, $x \in X \setminus E$

then

$$\overline{d}(\{\alpha_i\}_1^n, \{\beta_i\}_1^n) \leq 4\epsilon.$$

Proof. Let $\overline{\alpha}_i := \theta\alpha_i$ and $\overline{\beta}_i = \beta_i$ partitions of Y, since θ is measure-preserving, we have

$$\{\theta\alpha_i\}_1^n \sim \{\alpha_i\}_1^n,$$

then

$$\{\overline{\alpha}_i\}_1^n \sim \{\alpha_i\}_1^n.$$

Define:

$$f_i(y) := e(\ell_i(\theta^{-1}(y)) - m_i(y)). \tag{3.92}$$

If the $\overline{\alpha}_i$-name of y is different from $\overline{\beta}_i$-name of y then $f_i(y) = 1$. Otherwise $f_i(y) = 1$. Because of the previous lemma, we have

$$d(\overline{\alpha}_i, \overline{\beta}_i) = 2 \cdot \int e(\ell_i(\theta^{-1}(y)) - m_i(y))d\mu$$

This gives

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} d(\overline{\alpha}_i, \overline{\beta}_i) &= \frac{2}{n}\sum_{i=1}^{n}\int e(\ell_i(\theta^{-1}(y)) - m_i(y))d\mu \\
&= 2\int_{X\setminus E}\frac{1}{n}\sum_{i=1}^{n} e(\ell_i(\theta^{-1}(y)) - m_i(y))d\mu + \int_E \frac{1}{n}\sum_{i=1}^{n} e(\ell_i(\theta^{-1}(y)) - m_i(y))d\mu \\
&\leq 2\int_{X\setminus E}\epsilon d\mu + 2\int_E 1 d\mu \\
&\leq 2\cdot\epsilon + 2\epsilon = 4\cdot\epsilon.
\end{aligned}$$

Furthermore, we conclude:

$$\overline{d}(\{\alpha_i\}_1^n, \{\beta_i\}_1^n) \leq 4\epsilon.$$

□

Lemma 3.5.6. *If $\theta : X \to Y$ is ϵ-measure-preserving and $\alpha = \{A_1, ..., A_k\}$ is a partition of X, there is a mapping $\overline{\theta} : X \to Y$ such that*

$$\{\overline{\theta}\alpha\} \sim \{\alpha\} \tag{3.93}$$

and

$$\mu(\{x : \theta x \neq \overline{\theta}x\}) \leq 8 \cdot \epsilon. \tag{3.94}$$

Proof. Since θ is ϵ-measure-preserving, there exist an E_2 such that for every measurable set $A \subset X$, denote $\overline{A}$ such that $\overline{A} = A \setminus E_2$,

$$\left|\frac{\nu(\theta(\overline{A}))}{\mu(\overline{A})} - 1\right| \leq \epsilon.$$

It follows,

$$\sum_j \nu(\theta(\overline{A}_j)) < 1 + \epsilon, \tag{3.95}$$

and

$$\nu(\bigcup_j \theta(\overline{A}_j)) > 1 - 2\epsilon. \tag{3.96}$$

From (3.95) and (3.96) we have

$$\sum_j \nu(\theta(\overline{A}_j)) - \nu(\bigcup_j \theta(\overline{A}_j)) < 1 + \epsilon - 1 + e \cdot \epsilon = 3 \cdot \epsilon. \tag{3.97}$$

Define $B := \theta^{-1}(\bigcup_{i\neq j}(\theta(\overline{A}_i) \cap \theta(\overline{A}_j)))$ as the points which have the same images in different atoms.

Note that

$$\nu(\bigcup_{i\neq j}(\theta(\overline{A}_i) \cap \theta(\overline{A}_j))) = \sum_j \nu(\theta(\overline{A}_j)) - \nu(\bigcup_j \theta(\overline{A}_j)),$$

which concludes

$$\nu(\bigcup_{i\neq j}(\theta(\overline{A}_i) \cap \theta(\overline{A}_j))) < 3 \cdot \epsilon.$$

Since θ is ϵ-measure-preserving:

$$\mu(\overline{B}) < 3 \cdot \epsilon \cdot (1 + \epsilon) < 6 \cdot \epsilon$$

Now, consider $\tilde{A}_j = A_j \setminus (B \cup E_2)$. Since $\tilde{A}_j \cap B = \emptyset$ we have $\theta(\tilde{A}_i) \cap \theta(\tilde{A}_j) = \emptyset$, for every $i \neq j$. Thus,

$$\mu(\bigcup_j \tilde{A}_j) = \mu((\bigcup_j A_j) \setminus (B \cap E_2)) = 1 - \mu(B \cup E_2) > 1 - 7 \cdot \epsilon. \tag{3.98}$$

We will construct $\overline{\theta}$ in the following way: we will modify θ in a little portion of $\tilde{A}_j$ in order to get the same measure between the image of $\overline{A}_j$ and itself, there are no common points between the image of $\tilde{A}_i$ and the image of $\tilde{A}_j$, such that $i \neq j$. Finally we will map $B \cup E_2$ in the rest of Y preserving measure.

First Step: If $\nu(\theta(\tilde{A}_j)) > \mu(\tilde{A}_j)$, consider $G_j \subset \tilde{A}_j$ such that

$$\nu(\theta(G_j)) = \mu(\tilde{A}_j). \tag{3.99}$$

we can do that since the space is non-atomic. It follows:

$$\mu(G_j) > (1-\epsilon)\nu(\theta(G_j)) = (1-\epsilon)\mu(\tilde{A}_j). \tag{3.100}$$

Take any set of null measure as $P_j \subset Y$ and choose an arbitrary point $p_j \in Y$. Define $\overline{\theta}(x) := \theta(x)$ for every $x \in G_j$ and define $\overline{\theta(x)} := p_j$ for every $x \in \tilde{A}_j \setminus G_j$.

Ir follows from the definition of $\overline{\theta}$ and (3.99) that $\overline{\theta}$ preserve the measure of $\tilde{A}_j$ and is different from θ in a subset of $\tilde{A}_j$ with relative measure less than ϵ.

Now, in the case when $\nu(\theta(\tilde{A}_j)) < \mu(\tilde{A}_j)$, consider a set $G_j \subset \tilde{A}_j$ such that its measure is bigger than $(1-\epsilon) \cdot \mu(\tilde{A}_j)$. Define $\overline{\theta}(x) := \theta(x)$ when $x \in G_j$ and such way that $\overline{\theta}(\tilde{A}_j \setminus G_j) = \mu(\tilde{A}_j) - \nu(\theta(G_j))$ and dose not intersect any image of other $\tilde{A}_j$'s.

Second Step: All that it rest to define is $\overline{\theta}$ in $B \cup E_2$. From the first step, we have

$$\nu\left(\overline{\theta}\left(\bigcup A_j \setminus (B \cup E_2)\right)\right) = \mu\left(\bigcup A_j \setminus (B \cup E_2)\right),$$

because $\overline{\theta}$ preserves the measure of n any set $\tilde{A}_j$, $1 \leq j \leq k$ and the images of them are mutually disjoint.

Because of (3.5) all that it rest in Y to use is a subset of the same measure as $B \cup E_2$. Hence, we can define $\overline{\theta}$ in $B \cup E_2$ in such a way that it would be injective and send $B \cup E_2$ to the rest set of Y.

Furthermore, it follows from the definition of $\overline{\theta}$ and (3.5) that $\{\overline{\theta}\alpha\} \sim \{\alpha\}$

It follows from the definition of G_j and using (3.100) we get:

$$\mu(\tilde{A}_j \setminus G_j) < \epsilon \cdot \mu(\tilde{A}_j)$$

for every $1 \leq j \leq k$.

This and (3.98) give $\mu(\{x : \theta x \neq \overline{\theta} x\}) \leq 8 \cdot \epsilon$.

$\square$

Note that in general, $\overline{\theta}$, can't be chosen to satisfy (3.93) and (3.94) for "all" α simultaneously. Now for the strengthening of lemma 3.5.5.

Lemma 3.5.7 (Ornstein-Weiss). *Let $\{\alpha_i\}_1^n$ be partitions of X with name functions $\ell_i(x)$ and $\{\beta_i\}_1^n$ partitions of Y with name functions $m_i(y)$. If there is an ϵ-measure-preserving mapping $\theta : X \to Y$ such that*

$$\begin{aligned} &\frac{1}{n}\sum_{i=1}^{n} e(\ell_i(x) - m_i(\theta x)) \leq \epsilon,\ x \in X \setminus E \\ &\mu(E) \leq \epsilon \end{aligned} \tag{3.101}$$

where E is some subset of X then $\overline{d}(\{\alpha_i\}_1^n, \{\beta_i\}_1^n) \leq 36 \cdot \epsilon$.

Proof. Take $\alpha := \bigvee_1^n \alpha_i$ and using Lemma 3.5.6 we get

$$\{\overline{\theta}\alpha_i\} \sim \{\alpha_i\}$$

and

$$\mu(\{x : \theta x \neq \overline{\theta} x\}) \leq 8 \cdot \epsilon.$$

It follows from (3.101) and (3.5) that $\mu(E') \leq 9 \cdot \epsilon$ and

$$\frac{1}{n}\sum_{i=1}^{n} e(\ell_i(x) - m_i(\overline{\theta} x)) \leq 9 \cdot \epsilon,\ x \in X \setminus E' \tag{3.102}$$

where E' is the union of E with the set formed by the points which θ and $\overline{\theta}$ differs. Using Lemma 3.5.5, we have

$$\overline{d}(\{\alpha_i\}_1^n, \{\beta_i\}_1^n) \leq 36 \cdot \epsilon$$

$\square$

4 Kolmogorov automorphisms

Remembering the first chapter, we have that Bernoulli automorphisms have strong mixing property. A natural question appears: Between the Bernoulli property and the strong mixing property is there another class of transformations? Introduced by Andrei Kolmogorov in his article "A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces". That class is formed by *the Kolmogorov automorphisms* or K-systems.

Definition 4.0.1. An invertible measure preserving transformation T of a probability space $(X, \mathcal{B}, m)$ is a Kolmogorov automorphism (K-automorphism) if there exists a sub σ-algebra $\mathscr{K}$ of $\mathcal{B}$ such that:

- $\mathscr{K} \subset T\mathscr{K}$.
- $\vee_{n=0}^{\infty} T^n \mathscr{K} \doteq \mathcal{B}$.
- $\cap_{n=0}^{\infty} T^{-n} \mathscr{K} \doteq \mathcal{N} = \{X, \emptyset\}$.

Obs: If a system is a K-automorphism we say that the system has the K property

This definition tell us that $\mathscr{K}$ have a good behavior under iterations of the dynamic, that the collection of its iterations generates $\mathcal{B}$ and that our system is very chaotic, they have a massive loss of memory.

Theorem 4.0.2. *Let $(X, \mathcal{B}, m)$ a probability space and $T : (X, m) \longrightarrow (X, m)$ be a measure preserving transformation. T is a K-automorphism iff there is a finite partition P such that:*

1. *$\{T^i P\}_0^{\infty}$ generates $\mathcal{B}$.*
2. *Given K and ϵ there is an $N > 0$ such that $\bigvee_{-K}^{K} T^i P$ is ϵ-independent of $\bigvee_{N}^{N+m} T^i P$ for all $m > 0$.*

From the definition of Kolmogorov automorphism we have that such automorphisms have the strong mixing property. The proof of this claim appears in [28].

There is another characterization of the Kolmogorov property with the entropy of the system, we just stated the statement of the theorem for further reading could see [22]

Theorem 4.0.3. *(Rohlin and Sinai) Let $(X, \mathcal{B}, m)$ be a Lebesgue space and let $T : X \longrightarrow X$ be an invertible measure-preserving transformation.*

Then T is a Kolmogorov automorphism iff $H(T,\mathcal{A}) > 0$ for all finite non trivial partition $\mathcal{A}$.

We can see from Theorem 4.0.2 that Bernoulli automorphism are Kolmogorov, since taking the cylinders $[i]$, $i \in \{1,...,n\}$ as the partition we have that for an N big enough $\bigvee_{-K}^{K} T^i P$ is independent of $\bigvee_{N}^{N+m} T^i P$ for all $m > 0$.

Another way to proof that is with Rohlin and Sinai's theorem, because every non-trivial partition give us a positive entropy.

Therefore, we have:

Corollary 4.0.4. *Every Bernoulli automorphism is a Kolmogorov automorphism.*

Remark: It is quite obvious that the K property is an isomorphism invariant. Just need to take $\psi(\mathscr{K})$ as the sub σ-algebra generated from the conjugated function ψ and the sub σ-algebra $\mathscr{K}$ from the K property.

The following results show that the class of Kolmogorov automorphisms does not share all the properties the class of Bernoulli automorphisms enjoys. The proofs are given in the reference cited.

Theorem 4.0.5.

1. *There are uncountably many non-conjugate Kolmogorov automorphisms with the same entropy(Ornstein and Shields [16]).*
2. *There is a Kolmogorov automorphism T not conjugate to its inverse T^{-1} (Ornstein and Shields [16]).*
3. *There is a Kolmogorov automorphism which has no n-th roots for any $n \geq 2$ (Clark [4]).*
4. *There are non-conjugate Kolmogorov automorphisms T, S with $T^2 = S^2$ (Rudolf [25]).*
5. *There are two non-conjugate Kolmogorov automorphisms each of which is a factor of the other (Polit [19] and Rudolf [25]).*

In the next section we will give an example of a Bernoulli automorphism in a manifold, in order to do that, we state the following theorem

Theorem 4.0.6. *(Rokhlin [23], Yuzvinskii [29]) Let G be a compact topological group and $\varphi : G \longrightarrow G$ an ergodic automorphism, then φ is a Kolmogorov automorphism.*

Corollary 4.0.7. *Ergodic automorphism of the bi-torus are Kolmogorov*

Proof. Since the ergodic automorphisms of $\mathbb{T}^2$ are induced by matrices 2×2 with integer entries and no unitary eigenvalues. Because of the previous theorem this corollary follows. $\square$

4.1 Ergodic automorphisms of $\mathbb{T}^2$ are Bernoulli

Let f be an ergodic automorphism induced by the integer matrix A with non unitary eigenvalues.

Define λ_1 and λ_2 as the eigenvalues of A such that

$$|\lambda_1| < 1 < |\lambda_2|$$

Let F^s and F^u(in $\mathbb{R}^2$) be the lines associated to λ_1 and λ_2 respectively.

Since $|\lambda_1| < 1$, F^s have a contraction behavior , this is given $x, y \in F^s$, we have:

$$d(Ax, Ay) = |\lambda_1| d(x, y).$$

The same thing happens with F^s having an expansive behavior with a $|\lambda_2|$ factor.

When $x \in \mathbb{R}^2$ we can define the lines $F^s(x)$ and $F^u(x)$ as:

$$F^s(x) = x + F^s,$$
$$F^u(x) = x + F^u.$$

Let $\pi : \mathbb{R}^2 \longrightarrow \mathbb{T}^2$ the natural projection, define $\mathcal{F}^s(x)$ and $\mathcal{F}^u(x)$ as the projections of $F^s(x)$ and $F^u(x)$ respectively in the 2-torus, this is:

$$\mathcal{F}^s(x) = \pi \circ F^s(x),$$
$$\mathcal{F}^u(x) = \pi \circ F^u(x).$$

Notation: Given a set J, denote for $\mathcal{F}^s(x) \cap J$ a connected component of the intersection $\mathcal{F}^s(x) \cap J$ which have the point x. Similar for $\mathcal{F}^u(x) \cap J$.

Definition 4.1.1. A set Π will be called a *parallelogram* if satisfies the following conditions:

- Π it is connected,
- $\overline{\Pi} = \overline{\mathrm{Int}(\Pi)}$,

- for every $x, y \in \Pi$ we get

$$(\mathcal{F}^s(x) \cap \Pi) \cap (\mathcal{F}^u(y) \cap \Pi) = z \in \Pi.$$

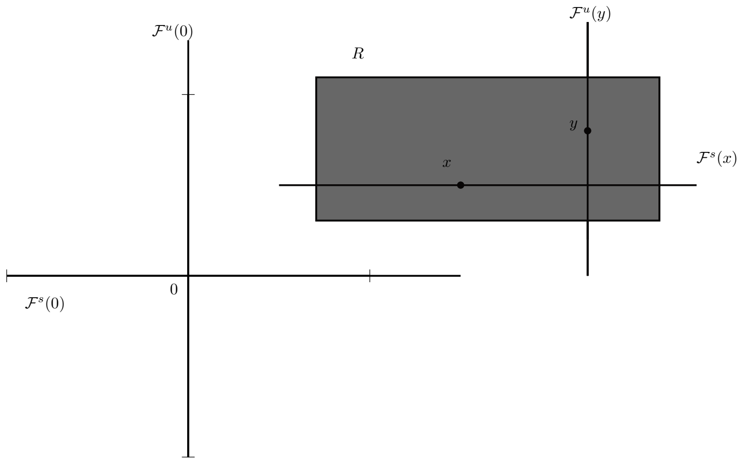

Figure 2 – Sketch of a parallelogram set

Definition 4.1.2. Let $\alpha = \{A^1, A^2, ..., A^k\}$ a partition in $\mathbb{T}^2$. We say that α is *piecewise smooth* if

- The frontier of each atom A^i is a union of finite number of smooth curves,
- $\overline{A^i} = \overline{\mathrm{Int}(A^i)}$.

Definition 4.1.3. Let Π be a parallelogram. Given a set $E \subset \mathbb{T}^2$, we say that E *intersects* Π *in a u-tubular* if for every $x \in \Pi \cap E$ we have $\mathcal{F}^u(x) \cap \Pi \subset E \cap \Pi$.

Lemma 4.1.4. *Suppose that α is a piecewise smooth partition, Π is a parallelogram, and $\delta > 0$ is given. There exists N_1 such that for any $N' > N \geq N_1$ and $\delta-$almost every atom $A \in \bigvee_N^{N'} f^k\alpha$ there is a subset $E \subset A$ with*

$$\frac{m(E)}{m(A)} > 1 - \delta, \tag{4.1}$$

and

$$E \textit{ intersects } \Pi \textit{ in a u-tubular subset.} \tag{4.2}$$

Proof. Notice that since $\mathcal{F}^u$ is invariant for f the property u-tubular for sets is not affected by iterations of f.

Let the set G_k be the intersection not tubular of Π with $f^k(A)$:

$$G_k = \{y \in \Pi \cap f^k(A) : \mathcal{F}^u(y) \cap \Pi \not\subset \Pi \cap f^k(A)\}.$$

To estimate the size of G_k, it is convenient to apply f^{-k}.

We want to proof that G_k have small measure. The next lemma show that the bad elements(elements in G_k) are close from the frontier

Lemma: For a certain constant $C > 0$, given any $y \in f^{-k}(G_k)$ we have

$$d(y, \partial A) \leq C \cdot |\lambda_2|^{-k}.$$

Proof: Consider $y \in f^{-k}(G_k)$. Since y is not an intersection u-tubular, we have $\mathcal{F}^u(y) \cap f^{-k}(\Pi) \not\subset f^{-k}(\Pi) \cap A$.

As A is connected and y belongs to A it follows that $\mathcal{F}^u(y)$ intersects ∂A in a point that we called z, it also intersects $f^{-k}(\Pi)$ in a point, that we called w such that

$$d(y, w) = d(y, z) + d(z, w)$$

because the intersection is not tubular and the foliations are lines. Since Π is a parallelogram, and the foliation $\mathcal{F}^u$ contracts exponentially with negative powers of f we have

$$d(y, w) \leq d(f^k(y), f^k(w)) \cdot |\lambda_2|^{-k} \leq \text{diam}(\Pi) \cdot |\lambda_2|^{-k}.$$

Taking $C = \text{diam}(\Pi)$, with this we conclude the proof of our lemma.

Since the boundary of A is smooth it follows that

$$m(G_k) \leq C \cdot |\lambda_2|^{-k} \tag{4.3}$$

Choosing N_1 big enough, in order to

$$m(G) \leq \sum_{k=N_1}^{\infty} m(G_k) \leq \delta^2,$$

where $G := \bigcup_{k=N_1}^{\infty} G_k$. Consider

$$\Omega := \left\{ B \in \bigvee_{N}^{N'} f^k \alpha : \frac{m(B \cap G)}{m(B)} > \delta \right\}$$

and take

Lemma(Bad atoms with small measure in proportion): Given Ω as stated before. Define:

$$L := \bigcup_{B\in\Omega} B.$$

Then $m(L) < \delta$.

Proof. If $m(L) > \delta$. Since L is a disjoint union of sets

$$m(L) = \sum_{B\in\Omega} m(B).$$

Furthermore

$$\delta^2 > m(G) \geq m(G\cap L) = \sum_{B\in\Omega} m(G\cap B) > \sum_{B\in\Omega} \delta\cdot m(B) = \delta\cdot m(L) > \delta^2,$$

which give us a contradiction. □

This means δ-almost every atom $B \in \bigvee_N^{N'} f^k\alpha$ intersects G in a set of relative measure at most δ.

For each atom $A \in \bigvee_N^{N'} f^k\alpha$ that does not belong to Ω, take $E = A\cap G^c$. Clearly E is tubular, since every intersection not tubular is in G, also

$$\frac{m(E)}{m(A)} > 1-\delta$$

because $A \notin \Omega$, thus we have

$$\frac{m(A\cap G)}{m(A)} \leq \delta \Rightarrow \frac{m(A\cap G^c)}{m(A)} > 1-\delta$$

□

The following lemma is key since is here where we construct the map θ that is fundamental in the Ornstein technique in order to proof that our system is Bernoulli. The construction of this map is based initially in build an bijective map between an stable foliation in an u-tubular set and all of it in a parallelogram, such map will be normalized measure preserving over the two foliations. Finally obtaining the desired map composing the previous map with parallel projections.

Lemma 4.1.5. *Given $\delta_1 > 0$ there is a $\delta_2 > 0$ such that if Π is a parallelogram of diameter less than δ_2, and E is an u-tubular subset of Π, there is a one-to-one, onto mapping $\theta : E \to \Pi$ such that*

1. *θ is measure preserving*
2. *$d(f^k\theta x, f^k x) < \delta_1$,*

for all $k > 0$, $x \in E$

Proof. If δ_2 is small enough, the second item will be satisfied provided that $\theta(x) \in \mathcal{F}^s \cap \Pi$, since distances contract as f operates on $\mathcal{F}^s$. Fix x_0 an interior point of Π.

Consider the segments of lines $E \cap \mathcal{F}^s(x) \cap \Pi$ and $\mathcal{F}^s(x_0) \cap \Pi$.

Define $\theta_0 : E \cap \mathcal{F}^s(x) \cap \Pi \longrightarrow \mathcal{F}^s(x_0) \cap \Pi$ a bijective linear map between these lines, taking extremes into extremes of the lines.

As it is defined, this map preserve the normalized Lebesgue measure. We will define θ from θ_0.

Let $x \in \Pi$ define $\pi_{x_0,x} : \mathcal{F}^s(x_0) \cap \Pi \longrightarrow \mathcal{F}^s(x) \cap \Pi$ as follows

$$\pi_{x_0,x}(y) = (\mathcal{F}^u(y) \cap \Pi) \cap (\mathcal{F}^s(x_0) \cap \Pi).$$

This transformation is just a parallel projection, that preserves Lebesgue measure. Now, define θ in the following way: for each $x \in E$ and given $y \in \mathcal{F}^s(x) \cap \Pi$,define

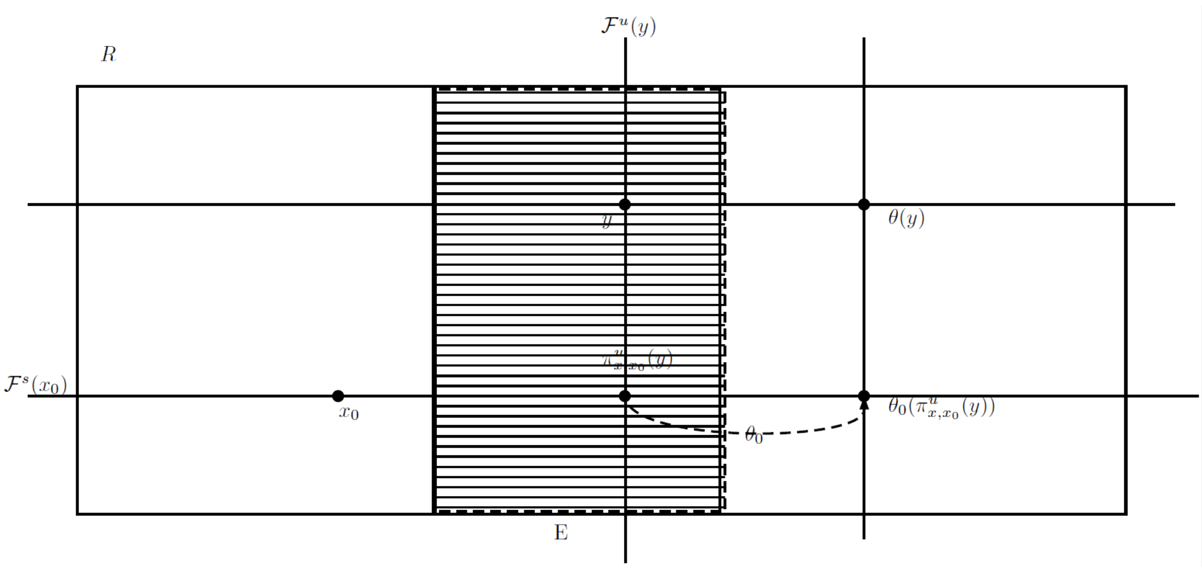

Figure 3 – Construction of the map θ.

$$\theta(y) := \pi_{x0,x} \cdot \theta_0 \cdot \pi^{-1}_{x0,x}(y). \tag{4.4}$$

This definition is possible, since E is an u-tubular set in Π. By Fubini's Theorem the mapping θ defined in $E \cap \Pi$ is measure preserving.

$\square$

Proposition 4.1.6. *Let $\epsilon, \epsilon' > 0$ and α a partition piecewise smooth. Then there is an N, such that for all $N' \geq N$, and for ϵ-almost every atom $A \in \bigvee_N^{N'} f^k\alpha$ there is a set $E \subset A$, and a one-to-one mapping θ of E onto $\mathbb{T}^2$ such that*

1. $\dfrac{m(E)}{m(A)} > 1 - \epsilon$,
2. *θ is $D \cdot \epsilon$-measure preserving,*
3. *$d(f^k\theta x, f^k x) < \epsilon'$, for any $k \geq 0$, $x \in E$.*

Proof. Given $\epsilon' > 0$ fixed, applying the previous lemma (for $\delta_1 := \epsilon'$), there exists $\delta_2 > 0$ such that Π is a parallelogram with diameter less than δ_2 and E is a subset u-tubular of Π, there is an bijective map $\theta : E \longrightarrow \Pi$ that satisfies (1) and (2) from the previous lemma. We can take a partition $\beta = \{\Pi_0, ..., \Pi_b\}$ of $\mathbb{T}^2$ where $\Pi_1, ..., \Pi_b$ are parallelograms with diameter not bigger than δ_2 and $m(\Pi_0) < \epsilon/10$, where Π_0 is the bad set from the lemma 4.1.4.

Consider:

$$\gamma := \epsilon \cdot b^{-1} \cdot \min\{m(\Pi_i) : 1 \leq i \leq b\}.$$

Because lemma 4.1.4, for each $1 \leq i \leq b$ we can find $N_1^i > 0$ big enough such that for every $N > N' \geq N_1^i$ and for every γ-almost every atom A of $\bigvee_N^{N'} f^k\alpha$, there exist a set $E \subset A$ such that

$$\frac{m(E)}{m(A)} > 1 - \gamma > 1 - \epsilon$$

such that E intersects Π_i in u-tubular way. Taking $N_1 = \max\{N_1^i : 1 \leq i \leq b\}$ we have that for ϵ-almost every atom A of $\bigvee_N^{N'} f^k\alpha$, for every $1 \leq i \leq b$ there exist a set $E_i \subset A$ such that

$$\frac{m(E_i)}{m(A)} > 1 - \gamma > 1 - \epsilon$$

and E_i intersects Π_i in u-tubular way for every $1 \leq i \leq b$. Define

$$E = \bigcup (E_i \cap \Pi_i) \subset A. \tag{4.5}$$

It is clear that E intersects each Π_i for $1 \leq i \leq b$ since E_i intersects Π_i in a μ-tubular way and each Π_i are disjoint, since E is defined for $i \geq 1$ does not have points in common with Π_0 and since

$$\frac{m(E_i^c \cap A)}{m(A)} < \gamma$$

we have that

$$\frac{m(E_i^c \cap A \cap \Pi_i)}{m(A)} < \frac{\gamma}{b}$$

Furthermore

$$\frac{m(E)}{m(A)} > 1 - \gamma > 1 - \epsilon.$$

Also, since $E \subset A$, we have:

$$\begin{aligned}\left|\frac{m(E \cap \Pi_i))}{m(E)} - m(\Pi_i)\right| &= \left|\left[\frac{m(A \cap \Pi_i)}{m(A)} - m(\Pi_i)\right]\frac{m(A)}{m(E)} + m(\Pi_i)\left(\frac{m(A)}{m(E)} - 1\right)\right| \\ &< \frac{1}{1-\epsilon}\left|\frac{m(A \cap \Pi_i)m(\Pi_i)}{m(A)} - m(\Pi_i)\right| + \frac{1}{1-\epsilon} - 1.\end{aligned}$$

Now, because the theorem 4.0.6, we know that f is Kolmogorov. Moreover, it follows from lemma 4.0.3 that each finite partition, in particular α is Kolmogorov.

Because the definition of Kolmogorov partition, it follows: for each $1 \leq i \leq b$, given $\xi_i > 0$ there exist $N_0^i > 0$ such that for every $N' > N \geq N_0^i$ and ξ-almost every atom $A \in \bigvee_{k=N}^{N'} f^k\alpha$ we have

$$\left|\frac{m(A \cap \Pi_i)}{m(A)} - m(\Pi_i)\right| \leq \xi_i. \tag{4.6}$$

Let $C > 0$, be a constant such that:

$$C(1-\epsilon) \cdot \min\{m(\Pi_i) : 1 \leq i \leq b\} > 1$$

Take:

$$\xi_i < \epsilon \cdot (C(1-\epsilon) \cdot \min\{m(\Pi_i) : 1 \leq i \leq b\} - 1), 1 \leq i \leq b$$

and $N_0 = \max\{N_0^i : 1 \leq i \leq b\}$. Taking ξ_i small enough, we have that for each $N' > N \geq N_0$ and γ-almost every atom $A \in \bigvee_{k=N}^{N'} f^k\alpha$ we have a set $E \subset A$ intersecting each Π_i in a u-tubular way, who satisfies

$$\frac{m(E)}{m(A)} > 1 - \gamma > 1 - \epsilon$$

and, for each $1 \leq i \leq b$ we have

$$\begin{aligned}
\left| \frac{m(E \cap \Pi_i)}{m(E)} - m(\Pi_i) \right| &< \frac{1}{1-\epsilon} \cdot \xi_i + \frac{1}{1-\epsilon} - 1 \\
&< \frac{1}{1-\epsilon} (\epsilon \cdot (C(1-\epsilon) \cdot \min\{m(\Pi_i) : 1 \leq i \leq b\} - 1) + 1) - 1 \\
&= C \cdot \epsilon \min\{m(\Pi_i) : 1 \leq i \leq b\} \leq C \cdot \epsilon \cdot m(\Pi_i).
\end{aligned}$$

In particular we have:

$$\left| \frac{m(E)m(\Pi_i)}{m(E_i \cap \Pi_i)} - 1 \right| \leq \frac{C \cdot \epsilon}{1 - C \cdot \epsilon}. \tag{4.7}$$

For each $1 \leq i \leq b$ define $E_i := E \cap \Pi_i$. It follows from the previous lemma that for each $1 \leq i \leq b$ we can define a bijective map that is measure-preserving

$$\theta_i : E_i \longrightarrow \mathbb{T}^2$$

satisfying

$$d(f^k(\theta_i(x)), f^k(x)) < \epsilon \tag{4.8}$$

for each $x \in E_i$, $k \geq 0$.

Define $\theta : E \longrightarrow \mathbb{T}^2$ such that

$$\theta(x) = \theta_i(x) \text{ if } x \in E_i.$$

The item (3) is clearly satisfied because of (4.8). Item (1) is a consequence of the definition of E.

Take $B \subset E$ a measurable set. Since $\theta_i : E_i \longrightarrow \Pi_i$ is measure-preserving, we have:

$$\frac{m(\theta(B) \cap \Pi_i)}{m(\Pi_i)} = \frac{m(B \cap \Pi_i)}{m(E \cap \Pi_i)}$$

Also, we have:

$$\left|\frac{m(\theta_i(B_i))m(E)}{m(B_i)} - 1\right| = \left|\frac{m(\theta_i(B_i))m(E_i)}{m(B_i)m(\Pi_i)} - 1 + \frac{m(\theta_i(B_i))m(E_i)}{m(B_i)m(\Pi_i)}\left(\frac{m(E)m(\Pi_i)}{m(E_i)} - 1\right)\right|$$
$$= \left|\frac{m(E)m(\Pi_i)}{m(E_i)} - 1\right|.$$

where $B_i : B \cap \Pi_i$.

Because of 4.7 and the previous equality we have:

$$\begin{aligned}\left|\frac{m(\theta(B))m(E)}{m(B)} - 1\right| &= \left|\frac{\sum_{i=1}^{b} m(\theta_i(B_i))m(E)}{m(B)} - 1\right| \\ &= \left|\sum_{i=1}^{b}\left[\frac{m(\theta_i(B_i))m(E)}{m(B_i)} - 1\right]\cdot\frac{m(B_i)}{m(B)}\right| \\ &\leq \sum_{i=1}^{b}\frac{m(B_i)}{m(B)}\cdot\left|\frac{m(E)m(\Pi_i)}{m(E_i\cap\Pi_i)} - 1\right| \\ &\leq \sum_{i=1}^{b}\frac{m(B_i)}{m(B)}\cdot\frac{C\cdot\epsilon}{1 - C\cdot\epsilon} \\ &= \frac{C\cdot\epsilon}{1 - C\cdot\epsilon} < D\cdot\epsilon,\end{aligned}$$

for a certain constant $D > 0$. Thus, θ is $D\cdot\epsilon$-measure preserving as we wanted.

□

By Theorems 3.5.1 and 3.5.2 to prove that an ergodic automorphism on the bitorus is Bernoulli it is sufficient to find a sequence of increasing of very weak Bernoulli partitions $\alpha_1 < \alpha_2 < ...$ such that $\bigvee_{i=1}^{\infty}\alpha_i = \mathcal{B}$where $\mathcal{B}$ is the Borel σ-algebra. For example we can choose $(\alpha_i)_1^n$ a sequence of partition of $[0,1]^2$ by 2^{2i} rectangles whose sizes have length $1/2^i$, then $(\alpha_i)_1^n$ is a decreasing sequence of partitions such that generates $\mathcal{B}$.

Then it is sufficient to prove that such partitions are very weak Bernoulli, in fact Ornstein proves a more general statement. Any finite piecewise smooth partition of $\mathbb{T}^2$ is very weak Bernoulli for f.

Lemma 4.1.7. *Given a partition α with the property piecewise smooth and an $\epsilon > 0$, there is an N such that for all $N' \geq N$ and ϵ-a.e. atom A of $\bigvee_N^{N'} f^k\alpha$,*

$$\overline{d}(\{f^{-i}\alpha|A\}_1^n, \{f^{-i}\alpha\}_1^n) \leq \epsilon$$

for all $n \geq 1$. In other words, any partition piecewise smooth is very weak Bernoulli.

Proof. Fix $\epsilon > 0$ such that $\epsilon > \epsilon' > 0$ and choose N_2 as N in the previous lemma, let Y the exception set from the previous lemma where θ is not defined, this is:

$$m(Y) \leq D \cdot \epsilon, Y = \text{ union of atoms in } \bigvee_{k=N}^{N'} f^k \alpha,$$

where $N' > N \geq N_2$. Take an atom $A \in \bigvee_{k=N}^{N'} f^k \alpha$. Consider $E \subset A$ and $\theta : E \longrightarrow \mathbb{T}^2$ as we constructed in the previous lemma.

Take a point y_0 in the atom A and define $\overline{\theta} : A \longrightarrow \mathbb{T}^2$ as

$$\overline{\theta}(x) = \theta(x) \text{ if } x \in E$$
$$\overline{\theta}(x) = y_0 \text{ if } x \notin E$$

for every $k \leq 0$. Remember that $m(A \setminus E) < 2\epsilon$. If $x \in E$, then from the previous lemma we have

$$d(f^k \theta x, f^k x) < \epsilon' \tag{4.9}$$

Afterwards, if $e(l_i(x) - m_i(\overline{\theta}(x))) = 1$ then $f^i(x)$ and $f^i(\overline{\theta}(x))$ are in different atoms of α, which implies

$$f^i(x) \in B_{\epsilon'}(\partial A_{l_i(x)}),$$

where $B_{\epsilon'}(\partial A_{l_i(x)})$ is an ϵ'-neighborhood of $\partial A_{l_i(x)})$.

Define

$$B_{\epsilon'} := \bigcup_{i=1}^{k} B_{\epsilon'}(\partial A_i).$$

From the definition of the function e, if $x \in E$ then

$$\frac{1}{n} \sum_{i=1}^{n} e(l_i(x) - m_i(\overline{\theta}(x))) \leq \frac{1}{n} \sum_{i=1}^{n} \chi_{B_{\epsilon'}}(f^i(x)).$$

Since f is ergodic, it follows from the Birkhoff's Ergodic Theorem

$$\frac{1}{n} \sum_{i=1}^{n} e(l_i(x) - m_i(\overline{\theta}(x))) \leq \frac{1}{n} \sum_{i=1}^{n} \chi_{B_{\epsilon'}}(f^i(x)) \to m(B_{\epsilon'}).$$

in almost every point $x \in E$.

By the *piecewise smooth* property of α we know that

$$\lim_{\epsilon' \to 0} m(B_{\epsilon'}) = 0 \tag{4.10}$$

Taking ϵ' arbitrarily small we can assure that in almost every point $x \in E$ that

$$\frac{1}{n}\sum_{i=1}^{n} e(l_i(x) - m_i(\overline{\theta}(x))) \leq d \cdot \epsilon$$

for some constant $d > 0$ From lemma 3.5.7 we get

$$\overline{d}(\{f^{-i}\alpha|A\}_1^n, \{f^{-i}\alpha\}_1^n) \leq d \cdot \epsilon$$

Since ϵ is arbitrary we get that α is very weak Bernoulli. Using Theorems 3.5.1 and 3.5.2 we have that $(\mathbb{T}^2, m)$ is a Bernoulli automorphism.

□

4.2 A Kolmogorov automorphism that is not Bernoulli

Since 1958 when Kolmogorov introduced the definition of Kolmogorov automorphisms in [7], a natural question appears: The Kolmogorov automorphisms are the same as the Bernoulli automorphisms? It was until 1973 that Ornstein [14] construct an example (not natural) were the system were Kolmogorov automorphism an not Bernoulli.

In this subsection we will present such example.

Before that, the following corollary is a direct consequence of Lemma 3.2.1 and we will use it in order to proof that next example is not Bernoulli.

Corollary 4.2.1. *Let T acting on $\bigvee_{-\infty}^{\infty} T^i P$ is not a Bernoulli transformation if we can find an $\epsilon_1 > 0$ and a sequence of ergodic transformations T_i and partitions P_i such that:*

1. $H(T_i) \geq H(T)$;
2. $\lim_{i\to\infty} |H(P_i, T_i) - H(P, T)| = 0$;
3. $\lim_{i\to\infty} |\operatorname{dist}(\vee_{j=0}^{u_i} T^j P) - \operatorname{dist}(\vee_{j=0}^{u_i} T_i^j P_i)| = 0$ *and* $\lim_{i\to\infty} u_i = \infty$;
4. *There are arbitrarily large i such that it is impossible to find a partition P^i such that* $\operatorname{dist}(\vee_{j=0}^{n} T_i^j P^i) = \operatorname{dist}(\vee_{j=0}^{n} T^j P)$ *for all n and* $|P^i - P_i| < \epsilon_1$.

4.2.1 Definition of (T,P)

Let $(X, \mathcal{B}, m)$ where X is a measurable set in the real line, $\mathcal{B}$ the σ-algebra of Borel and m the Lebesgue measure.

We define our transformation F and partition P in a constructive way trough gadgets i.e. T will be defined in stages and at each stage we will extend the definition of T to a larger part of the measure space. Moreover P will have four sets: P_0, P_e, P_f, P_s.

Notation:

At stage n we will have the following situation: We will have a set F_n, T^i, $0 \leq i \leq h(n) - 1$ will be defined on F, and the $T^i F$, will all be disjoint. P will be defined on $\bigcup_{i=0}^{h(n)-1} T^i F_n$. We thus have a gadget.

If the gadget $G = \bigcup_{i=0}^{r-1} T^i B$, then we will call r the height of G and denote it by $h(G)$. Thus $h(G_n) = h(n)$. If x is in a gadget G, then only some elements of the name will be defined and we will call these the name of x in G. Remembering that a name of a point x is a sequence $(\alpha_i)_i$ where α_i is in $0, e, f$ or s according to whether $T^i x$ is in P_0, P_e, P_f or P_s.

If G is a gadget, we will define a *slice* as follows:

Partition the base B according to the name (in G) of the points in B. If J is an atom in this partition, we will call $\bigcup_{i=0}^{h(G)-1} T^i J$ partitioned by P a slice.

4.2.1.1 Construction of (T,P)

- If the gadget at stage $2n$ is defined, we will get the stage $2n+1$ as follows:

  Divide F_{2n} into $f(2n) - 1$ ($f(2n)$ will equal $2n$) disjoint sets of equal measure $F_{2n,i}$, $1 \leq i \leq f(2n) - 1$ such that the gadget $\bigcup_{j=0}^{h(2n)-1} T^j F_{2n,i}$, partitioned by P (which will be defined along the way) is isomorphic to $\bigcup_{j=0}^{h(2n)-1} T^j F_{2n}$ partitioned by P.

  For each $F_{2n,i}$ pick $f(2n)$ disjoint sets $F_{2n,i,j}$, $1 \leq j \leq f(2n)$ not in $\bigcup_{j=0}^{h(2n)-1} T^j F_{2n}$ and define T on these (except for $F_{2n,i,f(2n)}$) so that

$$T(F_{2n,i,i}) = F_{2n,i}\,, \qquad T(T^{h(2n)-1} F_{2n,i}) = F_{2n,i,i+1}$$

  and

$$T(F_{2n,i,j}) = F_{2n,i,j+1} \quad \text{if } j \neq i, j \neq f(2n)$$

  Extend the definition of P so that $F_{2n,i,j}$, $j \geq i$ are in P_f and $F_{2n,i,j}$, $j > i$ are in P_e.

  This defines the gadget at stage $2n+1$, $F_{2n+1} = \bigcup_{i=1}^{f(2n)} F_{2n,i,1}$.

  To obtain the height of the gadget G_{2n+1} we will sum the previous gadget G_{2n} and the number of sets added f(2n). In other words:

$$h(2n+1) = h(2n) + f(2n). \tag{4.11}$$

  Observation: We will begin our construction with G_2 of height $h(2)$. ($h(2)$ will be determined later). The gadget G_2 is a subset of P_0 (in fact G_2 will turn out to equal P_0).

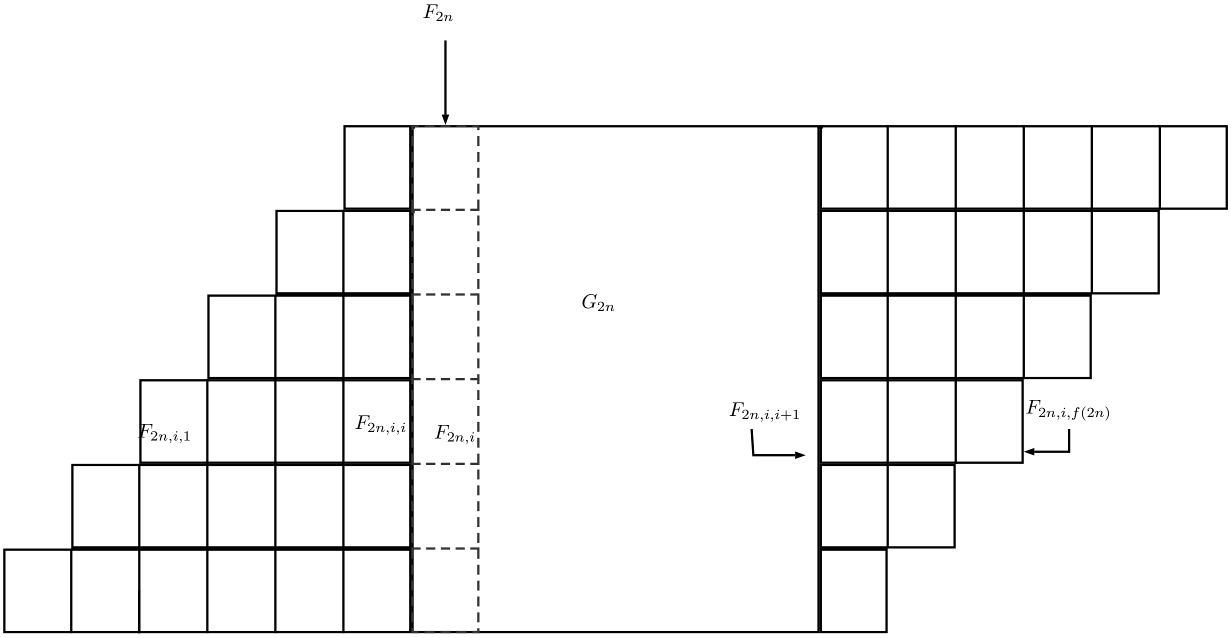

Figure 4 – Construction of the gadget G_{2n+1}.

Before continuing with our construction, we will define the "joint" of two gadgets.

Suppose we have two gadgets G_1 and G_2 where G_i, $i = 1, 2$, is the union of $T^j J_i$, $0 \le j \le r_i$, partitioned by some partition Q. We also assume that the measure of J_1 is the same as the measure of J_2 and $G_1$and G_2 are disjoint.

We will now define T from $T^{r_1} J_1$ onto J_2 (this will give us a new gadget $G_1 * G_2$, consisting of $\bigcup_0^{r_1+r_2} T^j J_1$ partitioned by P).

Partition J_1 according to the r_1-name of its points. Let I_ℓ be an atom in this partition and let $I'_\ell = T^{r_1} I_\ell$ ($\bigcup_{j=0}^{r_1} T^j I_\ell$ is a slice). Partition J_2 into sets E_ℓ in such a way that for each E_ℓ the gadget $\bigcup_{j=0}^{r_2} T^j E_\ell$ is isomorphic to G_2 and $m(E_\ell) = m(I'_\ell)$. Define $T(I'_\ell) = E_\ell$.

- We will now use the above construction to define the gadget at stage $2n+2$, given the gadget at stage $2n+1$ partition F_{2n+1} into 2^{2n+1} disjoint sets of equal measure such that each of them is the base of a gadget isomorphic to the gadget at stage $2n+1$. We call these gadgets $G^1, G^2, ..., G^{2^{2n+1}}$. For each G^i pick a collection of $i \cdot s(2n+1)$ (let $s(2n+1)$ be a number to be determined later) disjoint sets which will be in P_s (not in the gadget at stage $2n+1$!) all having the same measure as the base of G^i.

  Define T so that it maps the first of these sets onto the second, the second onto the third, etc., and the last onto the base of G^i. Call the resulting gadget $\overline{G}^i$. Form

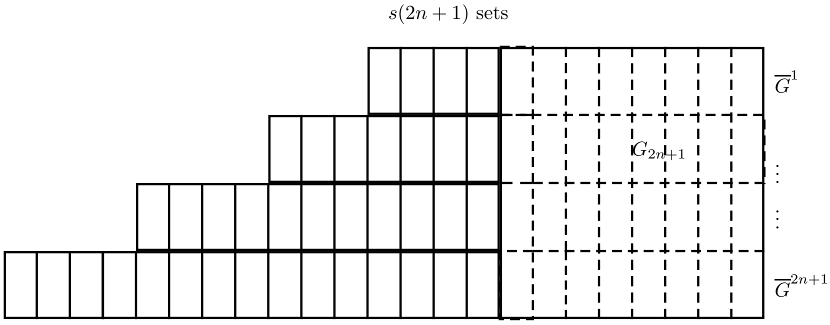

Figure 5 – Construction of the gadget $\overline{G}^i$.

$((...((\overline{G}^1 * \overline{G}^2) * \overline{G}^3)...) * \overline{G}^{2n+1})$. Now take $s(2n+1)$ intervals and map the top of $\overline{G}^{2n+1}$ onto the first of these, the first onto the second, etc. This resulting gadget will be G_{2n+2}.

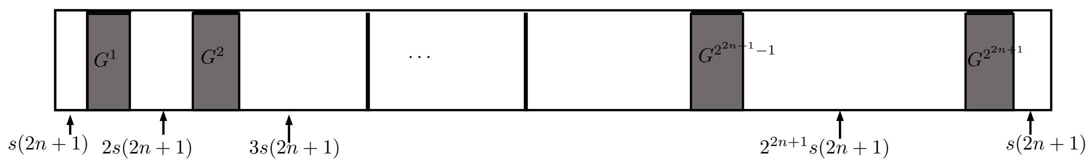

Figure 6 – Construction of the gadget G_{2n+2}.

It follows from the construction that the height of G_{2n+2} is obtained by the addition of 2^{2n+1} times the height of G_{2n+1} and the number of sets add in the construction $s(2n+1)$, $2s(2n+1)$,...,$2^{2n+1}s(2n+1)$. Therefore:

$$h(2n+2) = s(2n+1)(1 + 2^{2n}(2^{2n+1}+1)) + h(2n+1)2^{2n+1} \tag{4.12}$$

4.2.2 Choice of $f(n)$ and $s(n)$

Let $f(n) = n$ and let $s(n) = 100n^3$ (we will only use $f(n)$ for n even and $s(n)$ for n odd).

we will have:

1. $s(n) > 100 \sum_{i=1}^{n} f(i)$.
2. $s(n) < \frac{1}{2}^{10n} h(n)$

The first item is obvious. For the second one we just need to prove

$$s(n) < \left(\frac{1}{2}\right)^{10n+2} h(n) \tag{4.13}$$

for n even.

We can take $h(2)$ big enough such that (4.13) holds for $n \leq 20$ Assuming that (4.13) holds for n, we will prove it for $n+2$.

Choose ϵ, $0 < \epsilon < 1$ such that $s(n+1) < s(n)(1+\epsilon)$ in the following way:

Since $s(n) = 100n^3$ it follows

$$\frac{s(n+2)}{s(n)} = 1 + \left(\frac{6}{n} + \frac{12}{n^2} + \frac{8}{n^3}\right) < 1+\epsilon \tag{4.14}$$

for $n > 20$ we can take an ϵ that satisfies (4.13)

It follows that

$$s(n+2) < s(n)(1+\epsilon) < \left(\frac{1}{2}\right)^{10n+2}(1+\epsilon)h(n) < \left(\frac{1}{2}\right)^{10n+2}(2)h(n) < \left(\frac{1}{2}\right)^{10n+2}(2)h(n)\left(\frac{1}{2}\right)^{20}2^n$$

This follows from the fact that $n > 20$. This gives:

$$s(n+2) < \left(\frac{1}{2}\right)^{10(n+2)+2} 2^{n+1}h(n) < \left(\frac{1}{2}\right)^{10(n+2)+2} h(n+2) \tag{4.15}$$

which follows from the definition of $h(n+2)$:

$$h(n+2) = s(n+1)(1 + 2^n(2^{n+1}+1)) + h(n)2^{n+1} + n2^{n+1}.$$

For n odd we do the exact same thing, and the item (2) is proved.

We can see from this that $\sup\{m(G_n)\}$ is finite, thus T is supported in a set of finite measure, we can assume the total measure as 1. Since at stage n, when n is odd, the measure of the P_s that we add is less than

$$\frac{m(G_n)\cdot s(n)}{h(n)} \cdot \left(\frac{2^n+1}{2} + \frac{1}{2^n}\right) < \frac{m(G_n)\cdot s(n)}{h(n)} \cdot 2^{2n+2} < m(G_n)\left(\frac{1}{2}\right)^{8n}$$

Because $f(n) < s(n)$ the measure of the $P_j \cup P_e$ added at stage n, n even is less than

$$\frac{f(n)}{h(n)} \cdot m(G_n) < \frac{s(n)}{h(n)} \cdot m(G_n) < \left(\frac{1}{2}\right)^{10n} m(G_n)$$

Because of that we can assume that our space is a probability space where the partition is formed by $P = \{P_0, P_s, P_e, P_f\}$ defined in an iterative way.

We will define T_n, P_n in the same way as T, P except that at stage $2n+1$ we will form G_{2n+2} by taking

$((...((\overline{G}^{2n+1} * \overline{G}^{2n}) * \overline{G}^{2n-1})...) * \overline{G}^{1})$ instead of $((...((\overline{G}^{1} * \overline{G}^{2}) * \overline{G}^{3})...) * \overline{G}^{2n+1})$ i.e. we put the P_s's in the reversed order.

Figure 7 – Rearrangement of the gadget G_{2n+2}.

Lemma 4.2.2. *If we are given the name of x we can, for each F_n, determine which $T^i x$ are in F_n.*

Proof. We will do it by induction and showing that if we are given more than $h(n)$ consecutive terms in the name of x, we can tell which of the $T^i x$'s are in F_n.

It is obvious for $n = 2$ since we know that each $T^i F_2$ is in P_0, $0 \leq i \leq h(2)$ we just need to take the first term in each group of $h(2)$ consecutive 0's.

If the lemma holds for n even, it obviously holds for $n+1$, we just see the position of the first of the f's in front of each term in F_n such that the numbers of f's and e's that we have after $h(n)$ positions is the same.

If it holds for n odd, we get for $n+1$ as follows:

Find the term in F_n with exactly $s(n) \cdot s$-terms in front of it. The first of these s's will be in F_{n+1}. $\square$

Definition 4.2.3. Let G be a gadget with base J. We define a *rectangle* R in G as follows:

Let E be a set in some $T^i J$ and let $k > i$ be an integer smaller that the height of G. Then $\bigcup_{l=0}^{k-i} T^l E$ will be called a rectangle, and E its base.

We will say that R is pure below if each of the sets $T^l E$, $-i \leq l \leq 0$ is contained in some atom of P.

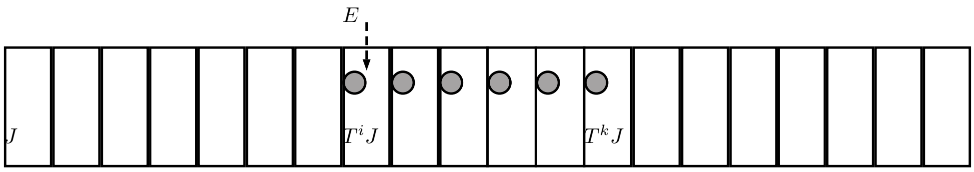

Figure 8 – Definition of rectangle.

Lemma 4.2.4. *If n is less than m, then G_m contains disjoint rectangles such that*

1. *Each of these rectangles is isomorphic (as a gadget partitioned by P) to G_n.*

2. *Each rectangle is pure below.*

3. *The union of the rectangles in G_m is equal to $G_n \cap G_m = G_n$ and the union of bases of the rectangles in G_m is equal to $F_n \cap G_m = F_n$.*

Proof. We will fix n and induct on m. We know that in the case of $m = n$ the lemma holds automatically.

If m is greater than n, we have:

If m is even and the lemma holds for it, we can take $\cup_{j=1}^{h(m)} F_{2n,i}^{j}$, $1 \leq i \leq h(h(n))$ as the rectangles we need in G_{m+1}. Clearly this rectangles holds the items of the lemma.

In the case that m is odd, we must see if G^1 and G^2 are gadgets that fulfil the consequences of the lemma, taking the union of its rectangles $G^1 * G^2$ also holds.

Then noting that taking G^i as the rectangle of $\overline{G}^i$, $\overline{G}^i$ holds. Consequently $(...(\overline{G}^1 * \overline{G}^2) * ... * \overline{G}^{2n+1})$ holds too.

□

The iteration of this map is similar to the translation map, such that the only invariant measurable set will have zero measure or total measure.

In order to proof the ergodicity of the map we suppose that exist an invariant measurable set with positive measure different from 1.

We can find a gadget big enough such that intersects E almost completely so that the rectangles which intersect E will have the same relative measure, reaching to a contradiction with the invariance, because in order to E be invariant With positive measure the gadget G_n should be getting bigger which contradict the fixed measure of E.

Lemma 4.2.5. *T is ergodic.*

Proof. Assume there is a measurable set E, $\mu(E) = \alpha$, $0 < \alpha < 1$ and $T(E) = E$.

From lemma 4.2.2, we can get G_n big enough such that(except for a set whose measure is less than ϵ) E intersects each level in a slice in G_n such that

$$\frac{\mu(E \cap L_i)}{\mu(L_i)} > 1 - \epsilon \tag{4.16}$$

or

$$\frac{\mu(E \cap L_i)}{\mu(L_i)} < \epsilon \tag{4.17}$$

in each level of a slice in G_n.

Since E is T-invariant, the measure of the intersection of E with any two levels in the same slice is the same.

$$\mu(E \cap T^j J) = \mu(E \cap T^i J). \tag{4.18}$$

Affirmation: Let L_1 and L_2 be levels of different slices in G_n. Then

$$\frac{\mu(E \cap L_1)}{\mu(L_1)} = \frac{\mu(E \cap L_2)}{\mu(L_2)} \tag{4.19}$$

Given ϵ we can find a K and E' such that $E' \subset \bigvee_{-K}^{K} T^i P$ and $|E - E'| < \epsilon$. For each m Lemma 4.2.4 implies that G_m is the union of rectangles, R_i (in G_m), each of which is isomorphic to G_n and pure below.

Therefore $\ell > h(n) + K$, $T^\ell E'$ will intersect each slice of R_i in the same proportion, except in those R_i such that $T^{-\ell} R_i$ that are not defined in G_m. If we take m large enough, we can assume that the union of exceptional R_i is arbitrarily small and that $T^\ell E'$ intersects each slice in G_n in almost the same proportion.

Since we can take E' sufficiently close as we want from E, $T^\ell E'$ is arbitrarily close to E.

Since we have (4.18) for different slices we have that

$$\frac{\mu(E \cap G_n)}{\mu(G_n)} = \sum_{i.j} \frac{\mu(E \cap T^j B_i)}{\mu(T^j B_i)} \geq h(n)(1 - \epsilon) \tag{4.20}$$

Then $\mu(E) = \sum_n \frac{\mu(E \cap G_n)}{\mu(G_n)} \geq h(n)(1-\epsilon)$ that derived us in a contradiction. □

Lemma 4.2.6. *Given an integer ℓ and $\epsilon > 0$, there exists N such that all $m > N$ (N even) have the following property:*

Let $\{J_i\}$ be the collection of sets of the form $T^j F_m$, and $0 \leq j \leq h(m)$. Except for a collection C of J_i the measure of whose union is less than ϵ we have that the distribution of ℓ-names of points in $\bigcup_{j=0}^{f(m)} T^j J_i$ is within ϵ of the distribution of ℓ-names of points in X.

Proof. Since T is ergodic, by the Birkhoff's Ergodic Theorem given $\bar{\epsilon} > 0$ there is an M such that $f(m) > M$, then all x, except for x in a set of measure less than $\bar{\epsilon}$, have that the distribution of ℓ-names in $\bigcup_0^{f(m)} T^i(x)$ is within $\bar{\epsilon}$ of the distribution of ℓ-names in X. Since $\bar{\epsilon}$ can be taken less than $(\epsilon/100)^2$, which proves our lemma. □

Lemma 4.2.7. *Given an integer ℓ and $\epsilon > 0$, there exists N such that if $n > N$ and n is odd we have the following: Let $\{J_i\}$ be the collection $T^i F_n$. Except for a collection of J_i the measure of whose union is less than ϵ, the distribution of the ℓ-names of points in J_i is within ϵ of the distribution of ℓ-names of points in X.*

Proof. Pick m, even, $m > N$, as in the previous lemma. Let $n = m+1$. We can see from the construction of the gadget G_{m+1} that the base F_{m+1} will be the union of disjoint sets $K_1, K_2, ..., K_{f(m)}$ such that $T^{f(m)-i}K_i = {}_iK$ is the base of a gadget isomorphic to G_m. If $f(m) + \ell < r < h(m+1) - f(m) - \ell$, then

$$T^r F_{m+1} = \bigcup_{i=1}^{f(m)} T^r K_i = \bigcup_{i=1}^{f(m)} T^{r-f(m)+i}{}_iK \tag{4.21}$$

Therefore the distribution of the ℓ-names of points in $T^r F_{m+1}$ is the same as the distribution of ℓ-names of points in $\bigcup_{i=1}^{f(m)} T^i(T^{r-f(m)})F_m$ since the distribution of K_i is the same as the distribution of F_m and hence from the previous lemma is within ϵ of the distribution of ℓ-names in X unless $T^{r-f(m)}F_m$ is in the exceptional set C for the lemma 4.2.6. Since ℓ is fixed and since $\dfrac{f(m)}{f(m+1)} \longrightarrow 0$ and the lemma follows. □

Theorem 4.2.8. *The automorphism T is K-automorphism.*

Proof. We must show the following: Given ℓ, there is an N such that $\bigvee_0^{\ell} T^iP$ is ϵ-independent of $\bigvee_N^{N+r} T^iP$ for all $r > 0$.

From lemma 4.2.7 we get that $\bigvee_0^{\ell} T^iP$ is ϵ-independent of Q, when Q is the partition of G_n into levels.

For each $m > n$, we will define Q_m in the following way:

We can take a collection of disjoint rectangles R_i in G_m by lemma 4.2.4, isomorphic to G_n. If B_i is the base of R_i and T^jB_i, $0 \leq j \leq h(n)$, the levels of R_i.

The partition Q_m will be formed by the levels of the rectangles R_i and $X - \cup R_i$, by the lemma 4.2.7, $\bigvee_0^{l} T^iP$ is 2ϵ-independent of Q_m.

Take $N = h(n)$. We have that Q_m refines $\bigvee_N^{r+N} T^iP$, since R_i are pure below and restricted for each $R_i \cap \bigcup_{i=r+N}^{h(m)} T^iF_m$. taking n large enough to make $X - G_m$ sufficiently small and $h(m)$ large enough compared to r, we get that $\bigvee_0^{l} T^iP$ is $\sqrt{3\epsilon}$-independent of $\bigvee_N^{N+r} T^iP$

□

Because of lemma 4.2.2 we know for each point in what time intersect the base of G_n, with this in mind, given the moment these points intersect with F_n, would like to know how close they are to each other by measuring the time it took to get there, in order to do that we must define formally sections of the P-names that represents all he path through G_n and how to measure those sections for different points.

Definition 4.2.9. The n-block in the P-name $(\alpha_i)_i$ of x, is by definition a sequence of $h(n)$ consecutive terms, the first of which is in the base F_n. Lemma 4.2.2 shows that each n-block is uniquely determined by the P-name of x.

If $m > n$, we will define the m-order of an n-block as follows:

Each n-block a is contained in a unique m-block b. The m-order of a will be i if a is the i-th n-block b (note that if n is odd, then the $n+1$ order of an n-block is determined by the number of s in front of it).

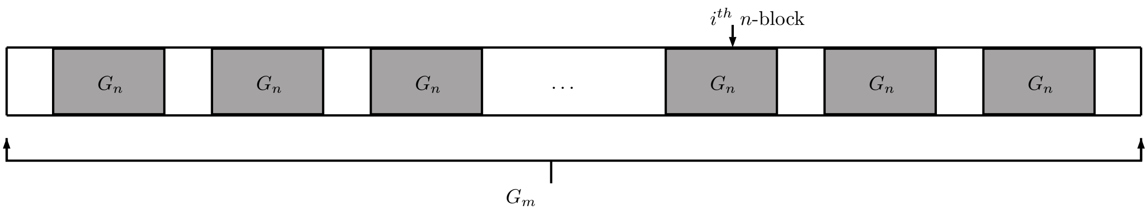

Figure 9 – The m-order of a n-block.

The next definitions will concern two points x and y and their respective P-names $(\alpha_i)_i$ and $(\beta_i)_i$.

Let a be an n-block in $(\alpha_i)_i$ and b an n-block in $(\beta_i)_i$. Let $|a-b|$ denote the absolute value of the difference of the indices of the first terms of a and b. We will say that a and b are close if $|a-b| < \sum_{k<n} f(k)$ (note that if a is also close to an n-block b', then $b' = b$. This follows from the choice of $s(n)$ and $f(n)$ which imply $\sum_{k<n} f(k) < \dfrac{1}{2}h(n)$).

Let a be an n-block in $(\alpha_i)_i$. we will call α_j in a *bad* if $\alpha_j = 0$ and $\beta_j \neq 0$

In the case when n is odd if two n-blocks of two different P-names of different points with the same $n+1$-order it is almost natural, because the construction of the gadget G_{n+1}, that all the n-blocks inside the $n+1$-block will be closed to some other n-block with the same $n+1$-order.

Lemma 4.2.10. *Let $(\alpha_i)_i$ and $(\beta_i)_i$ be the P-name of x and y respectively. Let a and b be the n-blocks in $(\alpha_i)_i$ and $(\beta_i)_i$, n odd. Assume a and b are close and have the same $n+1$ order. Let a' and b' be the $n+1$ blocks containing a and b. Then a' is close to b', and every n-block in a' is close to an n-block in b' of the same $n+1$-order.*

Proof. Because of the construction of G_{n+1} the blocks in a' are separated by a number of s-terms which depends only on the $n+1$-order of such block,the same thing happens in b', because of that every n-blocks in $(\alpha_i)_i$ and $(\beta_i)_i$ with the same $n+1$-order are close each other, it follows that a' and b' are also close, since the firsts n-blocks in them are close.

$\square$

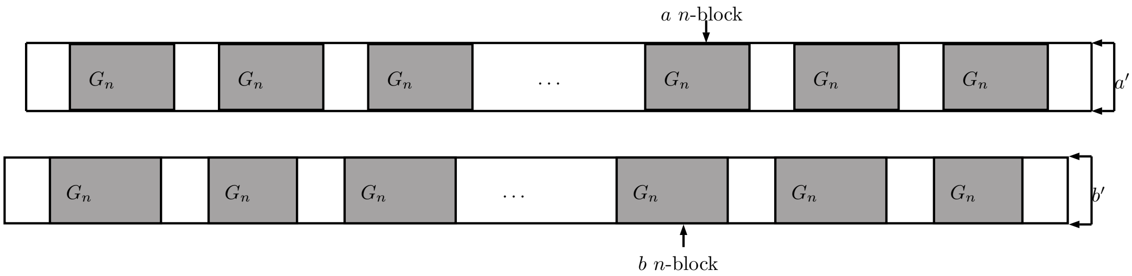

Figure 10 – Relation between $n+1$-blocks.

In the case when n is even losing that the $n-1$-blocks have the same n-order we can not warranty that the rest of the $n-1$-blocks will be closed to some other $n-1$-block with the same n-order as the previous lemma, because of the construction of the gadget G_n and the size of the added sets to it.

Just when these $n-1$-blocks are closed to each other we can warranty that there is at most another $n-1$-block closed to another $n-1$-block in the P-name.

Lemma 4.2.11. *Let $(\alpha_i)_i$ and $(\beta_i)_i$ be the P-names of x and y. Let a be an n-block in $(\alpha_i)_i$, n even. Let a' be an $n-1$-block in a, and assume that a' is close to an $n-1$-block b' whose n-order is different from the n-order of a'. Then there is at most one other $n-1$-block a'' in a such that a'' is close to an $n-1$ block in $\{\beta_i\}$.*

Proof. It would be enough to show that if b is an n-block that containing b', then there is no other $n-1$-block in a b that is close to an $n-1$-block in b.

To see that we must check that there is at most one $n-1$-block $\bar{b}$different from b which contains an $(n-1)$-block $\bar{b}'$ close to an $n-1$-block in a, $\bar{a}'$. From lemma 4.2.2 $\bar{a}'$ and $\bar{b}'$ must have different n-orders.

Let a'_i and b'_i be the $n-1$-blocks in a and b respectively whose n-order differ from the n-order of a' and b' by 1. That is a'_1 would be the next $n-1$-block after a',a'_2 would be the second $n-1$-block after a' and so on.

Because of that and the fact that a' and b' have different n-orders we get

$$|a'_i - b'_i| \geq i \cdot s(n-1) - \sum_{k<n-1} f(k)$$

Since $s(n) > 100 \sum_1^n f(k)$, a'_i and b'_i are not close.

Affirmation:

$$|a'_i - b'_i| \leq 2^n \cdot 2^n s(n-1) + \sum_{k<n-1} f(k) \leq \frac{1}{2} h(n-1). \tag{4.22}$$

We can see in $|a'_1 - b'_1|$ depend mostly of the difference of n-order between a' and b', because of that it increases by at most $2^n s(n-1)$, the maximum length of a string of s. since $i \leq 2^n$ we get our first inequality, the second one follows directly from the relation of $s(n)$ and $f(n)$ with $h(n)$ defined in "choice of f(n) and s(n)".

□

Lemma 4.2.12. *There exist a sequence $\epsilon_n > \overline{\epsilon} > 0$ such that if $(\alpha_i)_i$ and $(\beta_i)_i$ are the P-names of x and y, and if a is an n-block in $(\alpha_i)_i$, then either*

1. *There is an n-block in $(\beta_i)_i$ close to a or*
2. *There are more then $\epsilon_n h'(n)$, α_i in a are very bad. ($h'(n)$ is the number of 0's in an n-block).*

Proof. From the definition of G_2, for $n = 2$ is true taking $\epsilon_2 = (h(2))^{-1}$.

We now proceed by induction. Assume the statement is true for $n-1$. If n is odd, by the induction hypothesis a',the $n-1$-block in a, must be close to some $n-1$-block b'. Therefore a most be close to the n-block containing b'.

If n is even, we can assume that there is some $n-1$-block a'' in a that is close to an $n-1$-block b''. If a'' and b'' have the same n-order the lemma 4.2.10 implies that a is close to the n-block containing b''. If a'' and b'' have different orders, they are the only $n-1$-blocks that are close by lemma 4.2.11. Therefore because our induction hypothesis this implies that each of these $n-1$-blocks have more than $\epsilon_{n-1}h'(n-1)$ α_i's bad in a. Hence

$$\epsilon_n h'(n) > (\epsilon_{n-1})h'(n-1)(2^{n-1}-1)$$

Since $h'(n) = 2^{n-1}h'(n-1)$ we have

$$\epsilon_n > (\epsilon_{n-1})\left(1 - \left(\frac{1}{2}\right)^{n-1}\right)$$

□

Lemma 4.2.13. *Let $(\alpha_i)_i$ be the P-name of x and $(\beta_i)_i$ the P_n-name of y under T_n. Let a be a $2n+2$-block in $(\alpha_i)_i$. Then a contains at most four $2n+1$-blocks which are close to $2n+1$-blocks in $(\beta_i)_i$.*

Since the gadget in G_{2n+2} have the same number of initial intervals in the beginning such as the end of it. If the $2n+1$-blocks have the same $2n+2$-order, the blocks at the beginning and the end will be close too. For the intermediary case, it follows from the same analysis in the Lemma 4.2.11.

In the case where the $2n+1$-blocks are close it follows the same argument as the Lemma 4.2.11.

Lemma 4.2.14. *$H(P_n, T_n) = H(P, T)$.*

Proof. Let ${}_mQ = \bigvee_0^{m-1} T^i P$ and ${}_mQ_n = \bigvee_0^{m-1} T_n^i P_n$.

Define ${}_m\overline{Q}$ as follows:

For each x, if $x \in G_{2n+2}$, let $i > 0$ be the smallest integer such to $T^{-i}(x) \in F_{2n+2}$, otherwise let $i = 0$.

If $T^{m-1}(x) \in G_{2n+2}$, let j be the smallest integer such that $j \geq m-1$ and $T^j(x) \in T^{h(2n+2)} F_{2n+2}$, otherwise, let $j = m-1$. Take the P-name of x from $T^{-i}x$ to $T^j x$. Partition the x according to these names and call tat partition ${}_m\overline{Q}$.

Define ${}_m\overline{Q}_n$ in the same way, for P_n, T_n.

Define ${}_m\tilde{Q}$ as follows:

For each x, if $x \in G_{2n+2}$, let $i > 0$ be the smallest integer such that $T^i(x) \in T^{h(2n+2)} F_{2n+2}$ otherwise, let $i = 0$. If $T^{m-1}x \in G_{2n+2}$, let j be the largest integer such that $j \leq m-1$ and $T^j(x) \in F_{2n+2}$. Otherwise let $j = m-1$. Take the name of x from $T^i x$ to $T^j x$ and let ${}_m\tilde{Q}$ be the partition of the x according to these names.

Define ${}_m\tilde{Q}_n$ in an analogous way for P_n, T_n.

Since the construction of P_n, T_n is very similar to P, T, there is a 1-1 correspondence between the atoms of ${}_m\overline{Q}$ and ${}_m\overline{Q}_n$ that preserves their measure. The same holds for ${}_m\tilde{Q}$ and ${}_m\tilde{Q}_n$.

We can see that $H({}_m\tilde{Q}) \leq H({}_mQ) \leq H({}_m\overline{Q})$ and $H({}_m\tilde{Q}_n) \leq H({}_mQ_n) \leq H({}_m\overline{Q}_n)$.

For each fixed ϵ and m large enough $\bigvee_0^{(1+\epsilon)m} T^i P$ refines ${}_m\overline{Q}$ and $\bigvee_0^{(1-\epsilon)m} T^i P$ is refined by ${}_m\tilde{Q}$, with this applying the sandwich theorem, we have:

$$\lim_{m\to\infty} \frac{1}{m} |H({}_m\overline{Q}) - H({}_m\tilde{Q})| = 0 \text{ and } \lim_{m\to\infty} \frac{1}{m} |H({}_m\overline{Q}_n) - H({}_m\tilde{Q}_n)| = 0. \tag{4.23}$$

Since there is a 1-1 correspondence between the atoms in ${}_m\overline{Q}$ and ${}_m\overline{Q_n}$ we have that $H(P_n, Tn) = H(P, T)$.

□

Theorem 4.2.15. *The automorphism T is not a Bernoulli automorphism.*

Proof. From the definition of P_i, T_i and the previous lemma P_i, T_i satisfy the hypotheses $((1)-(3))$. We finish if we can show that is impossible to find a partition P^i such that $\text{dist}\left(\bigvee_0^n T_i^j P^i\right) = \text{dist}\left(\bigvee_0^n T^i P\right)$ for all n and $|P^i - P_i| < \dfrac{1}{8}\bar{\epsilon} \cdot m(P_0)$.

This will follow from:

Claim: Fixed i. If K is large enough, then the P-name of x of length K differs from the P_i-name of y of length K in more than $\dfrac{1}{4}\bar{\epsilon} \cdot m(P_0)$, for all x in a set of measure more than a half and all y.

Because of Lemmas 4.2.12 and 4.2.13 each $2n+2$-block in the P-name of x contains more than $\dfrac{1}{2}\bar{\epsilon}h'(2n+2)$ bad terms.

Therefore, the fraction of bad terms in each $2n+2$-block is more than $\dfrac{1}{2}\bar{\epsilon}m(P_0)$. Taking K large enough, the most x will have the property that the fraction of terms in its P-name, of length K, belonging to a $2n+2$-block is greater than $\dfrac{1}{2}$, the claim follows.

Because of that there is not a partition which every P^i-n-T_i-atom have the same distribution as the P-n-T-atom and at the same time $|P^i - P_i| < \dfrac{1}{8}\bar{\epsilon}m(P_0)$

$\square$

5 Partially Hyperbolic Diffeomorphisms

Motivated by the example given in the first section of Chapter 4 proved by Ornstein [15], Pesin [17] prove that every volume preserving $C^{1+\alpha}$ Anosov diffeomorphism over a compact Riemannian Manifold M is a Bernoulli automorphism.

In [20], Ponce et al. show that for a weaker version of hyperbolicity (derived from Anosov diffeomorphism) preserves the result that Kolmogorov automorphisms have the Bernoulli property under some hypothesis over the center manifold.

Theorem 5.0.1. *(Ponce-Tahzibi-Varão [20]) Let $f : \mathbb{T}^3 \to \mathbb{T}^3$ be a C^2 volume preserving derived from Anosov diffeomorphism with linearization $A : \mathbb{T}^3 \to \mathbb{T}^3$. Assume that f is Kolmogorov and one of the following occurs:*

1. *$\lambda_A^c < 0$ and $\mathcal{F}^{cs}$ is absolutely continuous, or*
2. *$\lambda_A^c > 0$ and $\mathcal{F}^{cu}$ is absolutely continuous.*

Then f is Bernoulli

5.1 Preliminaries

We now introduce the class of diffeomorphism, base of this chapter, which will present in some direction no hyperbolicity.

Definition 5.1.1. Given a smooth compact Riemannian manifold M. A diffeomorphism $f : M \to M$ is called partially hyperbolic if the tangent bundle of the ambient manifold admits an invariant decomposition $TM = E^s \oplus E^c \oplus E^u$, such that all unit vectors $v^\sigma \in E_x^\sigma, \sigma \in \{s, c, u\}$ for any $x, y, z \in M$

$$||D_x f v^s|| < ||D_y f v^c|| < ||D_z f v^u||$$

and $||D_x f v^s|| < 1 < ||D_z f v^u||$ where v^s, v^c and v^u belong respectively to E_x^s, E_y^c and E_z^u.

Let $f : \mathbb{T}^n \to \mathbb{T}^n$ partially hyperbolic diffeomorphism. Consider $f_* : \mathbb{Z}^n \to \mathbb{Z}^n$ the action of f on the fundamental group of $\mathbb{T}^n$. The function f_* can be extended to $\mathbb{R}^n$ and the extension is the lift of a unique linear automorphism $A : \mathbb{T}^n \to \mathbb{T}^n$

Definition 5.1.2. Let $f : \mathbb{T}^n \to \mathbb{T}^n$ partially hyperbolic diffeomorphism. The unique linear automorphism $A : \mathbb{T}^n \to \mathbb{T}^n$ with lift $f_* : \mathbb{R}^n \to \mathbb{R}^n$ as constructed in the previous paragraph, is called the linearization of f.

Definition 5.1.3. We say that $f : \mathbb{T}^n \to \mathbb{T}^n$ is a derived from Anosov diffeomorphism or just a DA diffeomorphism if it is partially hyperbolic and its linearization is a hyperbolic automorphism (no eigenvalue of norm one).

Theorem 5.1.4. *(Oseledets Theorem) Let $f : M \longrightarrow M$ be a C^1 diffeomorphism defined on a compact Riemannian manifold M. The set of points $x \in M$ which satisfies:*

- *There exists a splitting*

$$T_xM = E_{1(x)} \oplus ... \oplus E_{k(x)};$$

- $Df_xE_i(x) = E_i(f(x)), k(x) = k(f(x))$;
- *The exist the limits*

$$\lambda_i(x) := \lim_{n\to\infty} \frac{1}{n} \log ||Df^n(x) \cdot v|| = \lim_{n\to-\infty} \frac{1}{n} \log ||Df^n(x) \cdot v||$$

  for all $v \in E_i(x) \setminus \{0\}$.

is a set of full measure for any f-invariant probability measure.

Definition 5.1.5. We will call such $\lambda_i(x)$ from the Oseledets Theorem as the *Lyapunov exponents* of f at x.

Observe from the definition of the Lyapunov exponent that such limit tries to measure the exponential growth of the function in the directions of the subspace E_{i_x}. Remember that in the Anosov case if we consider the direction on the stable direction it turns out that $\lambda_{i(x)} < 0$ even for a time n sufficiently large it preserves some hyperbolicity but we can not conclude anything when the Lyapunov exponent is zero.

Let $(M, \mu, \mathcal{B})$ be a probability space, where M is a compact metric space, μ a probability measure and $\mathcal{B}$ the Borelian σ-algebra. Given a partition $\mathcal{P}$ of M by measurable sets, we associate the probability space $(\mathcal{P}, \tilde{\mu}, \tilde{\mathcal{B}})$ by the following way. Let $\pi : M \to \mathcal{P}$ be the canonical projection, that is π associates to a point x of M the partition element of $\mathcal{P}$ that contains it. Then we define $\tilde{\mu} := \pi_*\mu$ and $\tilde{\mathcal{B}} := \pi_*\mathcal{B}$

Definition 5.1.6. Given a partition $\mathcal{P}$. A family $\{\mu_P\}_{P\in\mathcal{P}}$ is a system of conditional measures for μ (with respect of $\mathcal{P}$) if:

1. given $\phi \in C^0(M)$, then $P \mapsto \int \phi\mu_P$ is measurable.

2. $\mu_P(P) = 1$ $\tilde{\mu}$-a.e.

3. if $\phi \in C^0(M)$, then $\int_M \phi d\mu = \int_P (\int_P \phi d\mu_P) d\tilde{\mu}$

When it is clear which partition we are referring to, we say that the family $\{\mu_P\}$ disintegrates the measure μ.

Proposition 5.1.7. *Given a partition $\mathcal{P}$, if $\{\mu_P\}$ and $\{\nu_P\}$ are conditional measures that disintegrate μ on $\mathcal{P}$, then $\mu_P = \nu_P$, $\tilde{\mu}$-a.e.*

Corollary 5.1.8. *If $T : M \to M$ preserves a probability μ and the partition $\mathcal{P}$, then $T_{*\mu_P} = \mu_{T(P)}$, $\tilde{\mu}$-a.e.*

Definition 5.1.9. We say that a partition $\mathcal{P}$ is measurable (or countably generated) with respect to μ if there exist a measurable family $\{A_i\}_{i\in\mathbb{N}}$ and and a measurable set F of full measure such that if $B \in \mathcal{P}$, then there exists a sequence $\{B_i\}$, where $B_i \in \{A_i, A_i^c\}$ such that $B \cap F = \bigcap_i B_i \cap F$.

Theorem 5.1.10 (Rokhlin's disintegration)**.** *Let P be a measurable partition of a compact metric space M and μ a Borelian probability. Then there exists a disintegration by conditional measures for μ*

Definition 5.1.11. We say that a foliation $\mathcal{F}$ is *absolutely continuous* if for any foliated box, the disintegration of volume on the segment leaves have conditional measures equivalent to the Lebesgue measure on the leaf.

We know from the previous chapter when we show that ergodic automorphism on the bitorus are Bernoulli that we depend on the application of Fubini's Theorem to the unstable manifolds of the automorphism. Unfortunately the Fubini's Theorem is not true for every continuous foliation of a given manifold, therefore to generalize this particular step in the context of this chapter we rely on the absolute continuity property on the $\mathcal{F}^{cs}$.

Observation: It is known that with our hypothesis the $\mathcal{F}^u$ and the $\mathcal{F}^s$ are absolutely continuous [21].

Definition 5.1.12. We say that a foliation $\mathcal{F}$ has *atomic disintegration* with respect to a measure μ , or that μ has atomic disintegration along $\mathcal{F}$, if for any foliated box, the disintegration of μ on the segment leaves have conditional measures which are finite sums of Dirac measures.

Given a derived from Anosov diffeomorphism, then by results of Franks [5] and Manning [9] there is a semi-conjugacy $h : \mathbb{T}^3 \longrightarrow \mathbb{T}^3$ which we will call the Franks-Manning semi-conjugacy, between f and its linearization A, that is, h is a continuous surjection satisfying

$$A \circ h = h \circ f$$

Moreover, this semi-conjugacy has the property that there exists a constant $K \in \mathbb{R}$ such that if $\tilde{h} : \mathbb{R}^3 \to \mathbb{R}^3$ denotes the lift of h to $\mathbb{R}^3$ we have $||\tilde{h}(x) - x|| \leq K$ for all $x \in \mathbb{R}^3$, and given two points $a, b \in \mathbb{R}^3$, there exists a constant $\Omega > 0$ with

$$\tilde{h}(a) = \tilde{h}(b) \Leftrightarrow ||\tilde{f}^n(a) - \tilde{f}^n(b)|| < \Omega, \text{ for all } n \in \mathbb{Z}.$$

R.Ures [27] proved that h takes center leaves of f onto center leaves of A, that is,

$$\mathcal{F}^c_A(h(x)) = h(\mathcal{F}^c_f(x))$$

Given a point $x \in \mathbb{T}^3$ define the set $c(x) \subset \mathcal{F}^c(x)$ by:

$$c(x) := h^{-1}(\{h(x)\})$$

Take:

$$\mathcal{C} := \bigcup_{y \in \{x \in \mathcal{T}^3 | c(x) \neq \{x\}\}} c(y)$$

We can see that $f(\mathcal{C}) = \mathcal{C}$, for if $h(a) = h(b)$ then by the definition of semiconjugacy we have:

$$h(f(a)) = A(h(a)) = A(h(b)) = h(f(b))$$

and if $h(a) \neq h(b)$, we have:

$$h(f(a)) = A(h(a)) \neq A(h(b)) = h(f(b))$$

Definition 5.1.13. An f-invariant measure μ is called *virtually hyperbolic* if there exists a full measurable invariant subset Z such that Z intersects each center leaf in at most one point.

If μ is virtually hyperbolic , then the central foliation is measurable with respect to μ and conditional measures along center leaves are (mono-atomic) Dirac measures. Indeed the partition into central leaves is equivalent to the partition into points.

The main difficulty in the partially hyperbolic case is the lack of hyperbolic behavior along the center directions. When we restricted ourselves to the derived from Anosov diffeomorphisms we have the advantage that the center direction in some sense carries some information from the center foliation of its linearization. If the linearization of a derived from Anosov diffeomorphism f has negative center exponent for example, then the center foliation of f has globally the same behavior as the center foliation for the linearization.

The idea to tackle the problem derived from Anosov diffeomorphisms is to treat the center foliation as a contracting(or expansive) foliation in "as many points as possible". Given a derived from Anosov diffeomorphism f with linearization A, we have the semiconjugacy h between f and A. Assume that the center Lyapunov exponent of A is negative. Ponce et al. proved [20] that if a pair of points $(a, b) \in M \times M$ is such that $h(a) \neq h(b)$ and $b \in \mathcal{F}^{cs}(a)$ then their orbits by f behaves, for most of the time, as if they were in a same contracting foliation, that is, the distance between $f^n(a)$ and $f^n(b)$ is very small for most of the natural numbers n.

Therefore, if we restrict our analysis to the set of points $x \in M$ for which $h(x) \neq h(y)$ for all $y \in M \setminus \{x\}$. We can "treat" the center foliation as if it was a contracting foliation. Although, we do not know how big this set is.

The following Theorem tell us that if this set has zero measure then there exists a full measure set intersecting almost every center leaf in exactly one point. Since points in two separate center leaves have distinct images by h then we can restrict our analysis to the set of atoms and again treat, in some sense, the center foliation as a “contracting foliation”.

Theorem 5.1.14. *(Ponce-Tahzibi-Varão [20]) Let $f : \mathbb{T}^3 \to \mathbb{T}^3$ be a C^2 volume (m) preserving derived from Anosov diffeomorphism with h a semi conjugacy to linear Anosov diffeomorphism. Then, h is m-almost everywhere injective. More precisely, the following dichotomy is valid:*

- *Either the set $\mathcal{C}$ has zero volume, or*
- *$\mathcal{C}$ has full measure and (f, m) is virtually hyperbolic.*

In the latter case (f, m) is Kolmogorov.

Definition 5.1.15. We define the essential injectivity domain X of h as follows: If $m(C) = 0$ define

$$X := \mathbb{T}^3 \setminus \mathcal{C}.$$

Otherwise, define

$$X := \text{set of atoms.}$$

The following Theorems and Lemmas will be used during the proof of the Theorem 5.0.1:

Lemma 5.1.16 (Pinsker [18], Rokhlin-Sinai [24])**.** *Let f be a K-automorphism of a Lebesgue space (Y, μ), then every finite partition $\mathcal{E}$ of Y is a Kolmogorov partition.*

Theorem 5.1.17 (Burns-Wilkinson[2])**.** *Let $f \in C^2$, volume-preserving, partially hyperbolic and center bunched. If f is essentially accessible, then f is ergodic, and in fact has the Kolmogorov property.*

Corollary 5.1.18 ([2])**.** *Let f be a $C^{1+\alpha}$ volume preserving and partially hyperbolic with $dim(E^c) = 1$. If f is essentially accessible, then f is ergodic, and in fact has the Kolmogorov property.*

Proposition 5.1.19. *Let M be a compact Riemannian manifold and $f : M \to M$ a volume preserving $C^{1+\alpha}$ partially hyperbolic diffeomorphism and one dimensional center direction. Thus f is Kolmogorov if, and only if, it is essentially accessible.*

5.2 Proof of Theorem 5.0.1

Without loss of generality we can assume that the center Lyapunov exponent $\lambda^c_A < 0$(other wise we work with f^{-1} which is homotopic to A^{-1}).

From the ergodicity of f it follows that $\mathcal{C}$ has either full or zero volume.

In our case the existence of a center foliation requires a fine comprehension of the disintegration of volume on the center foliation. The Theorem 5.1.14 is the key to overcome this issue.

5.2.1 Sketch of the proof

Take $\mathcal{E}$ to be the partition of $\mathbb{T}^3$ by points. Let α be a partition of $\mathbb{T}^3$ by measurable sets such that the boundary of any element in α is piecewise smooth and such that each atom $D \in \alpha$ is an open set with boundary of zero measure. It is easy to construct such a partition on the 3-torus. We will prove that α is very weak Bernoulli. Then we will take a sequence of such partitions α_n with: $\alpha_1 \leq \alpha_2 \leq ... \leq \alpha_n \leq ...$ such that $\alpha_n \to \mathcal{E}$, concluding that f is indeed Bernoulli by Theorems 3.5.1 and 3.5.2.

In order to do so, we shall consider some specific partitions with dynamical meaning which will help us to control the behavior of $f^i\alpha, i \in \mathcal{N}$.

Given two points y and z close enough to each other, we know that $\mathcal{F}^{cs}(y)$ and $\mathcal{F}^{u}(z)$ will intersect each other and that the intersection is, locally, a single point. We denote this point by $[y,z]$. Sometimes along this section we also write $W^{cs}(y) \cap W^{u}(z)$ to mean the point $[y,z]$.

Definition 5.2.1. A measurable set Π is called δ-rectangle at a point w if $\Pi \subset B(w,\delta)$ and for any $y, z \in \Pi$ the local intersection belongs to Π, that is

$$[y,z] \in \Pi$$

Note that, by the local product structure of the rectangles, we can think of a rectangle Π as a Cartesian product of $\mathcal{F}^{u}_{x} \cap \Pi$ and $\mathcal{F}^{cs}_{z} \cap \Pi$, where $x, z \in \Pi$.

Let $f : M \to M$ be a partially hyperbolic diffeomorphism with absolutely continuous center-stable foliation. Then we can take, for a typical z,m^{u}_{z} the measure m conditioned on $\mathcal{F}^{u}_{z} \cap \Pi$ and m^{cs}_{f} the factor measure on the leaf $\mathcal{F}^{cs}_{z}$, that is as in Definition 5.1.6 the partition $\mathcal{P} = \{\mathcal{F}^{u}_{z} \cap \Pi\}_{z\in\Pi}$, $\{m^{u}_{z}\}$ is a system of conditional measures and m^{cs}_{f} is the factor measure. From the absolute continuity of the unstable and center-stable foliations it follows that for a typical z the product measure

$$m^{P}_{R} := m^{u}_{z} \times m^{cs}_{f},$$

which is defined on Π, satisfies:

$$m^{P}_{R} << m.$$

Definition 5.2.2. Given any $\epsilon > 0$, an ϵ-regular covering of M is a finite collection of disjoint rectangles $\mathcal{R} = \mathcal{R}_{\epsilon}$ such that:

1. $m(\cup_{R\in\mathcal{R}} R) > 1 - \epsilon$

2. For every $R \in \mathcal{R}$ we have

$$|\frac{m^{P}_{R}(R)}{m(R)} - 1| < \epsilon$$

   and moreover, R contains a subset, G with $m(G) > (1-\epsilon)m(R)$ which has the property that for all points in G,

$$|\frac{dm^{P}_{R}}{dm} - 1| < \epsilon$$

We will state some Lemmas in order to show briefly the ideas developed in Ponce et al.'s work. The proof of those Lemmas appears in [20].

Lemma 5.2.3. *Given any $\delta > 0$ and any $\epsilon > 0$, there exist an ϵ-regular covering of connected rectangles $\mathcal{R}_\epsilon$*

Definition 5.2.4. We say that a measurable set A intersects a rectangle Π, leafwise if

$$\mathcal{F}^u(w) \cap \Pi \subset A \cap \Pi, \text{ for any } w \in A \cap \Pi$$

Lemma 5.2.5. *If E is a set intersecting a rectangle Π leafwise then the intersection $E \cap \Pi$ is a rectangle.*

Lemma 5.2.6. *The set $\mathbb{T}^3 \setminus \mathcal{C}$ is a u-saturated. In particular, given any rectangle Π, if $X = \mathbb{T}^3 \setminus \mathcal{C}$ then X intersects Π leafwise.*

Lemma 5.2.7. *Given a rectangle Π and $\beta > 0$, one can find $N_1 > 0$ such that for any $N' \geq N \geq N_1$ and β-almost every element $A \in \bigvee_N^{N'} f^k\alpha$, there exists a subset $E \subset A$, intersecting Π leafwise, for which:*

$$\frac{m(E)}{m(A)} \geq 1 - \beta$$

The proof of this fact in our setting is exactly the same proof as for the non uniformly hyperbolic case. Therefore we refer the reader to [17] for a detailed demonstration. We point out that the proof does not need contraction of $\mathcal{F}^{cs}$, but it depends only on the expansion of $\mathcal{F}^u$,

We now proceed to the most important part of the proof: for $N' \geq N$ big enough and ϵ-almost every $A \in \bigvee_N^{N'} f^k\alpha$ we will construct a function $\theta : A \to M$ satisfying the hypothesis of Lemma 3.5.7 , in particular it should be $c\epsilon$-measure preserving, for some constant c. Such function θ should have the property that for most of the points x, the orbits of x and of $\theta(x)$ have almost the same information (asymptotically) with respect to the partition in question. In particular, if we get good control on the distance$d(f^n(x), f^n(\theta(x)))$we can expect that for most points x, $f^n(x)$and $f^n(\theta(x))$ belong to the same partition element; this will be much clearer below. In order to get this control, we need to restrict to a compact set (where his uniformly continuous) and then use the fact that the center manifold of A is uniformly contracting and h sends center manifolds to center manifolds. This is what we do in the next lemma.

Lemma 5.2.8. *Given any $\epsilon > 0$ and any compact set K in the essential injectivity domain of h(see Definition 5.1.15) there exists $n_0 \in \mathbb{N}$ such that for any two points $a \in K, b \in \mathcal{F}^{cs}(a) \cap K$, with $d(a, b) < \frac{1}{2}$we have*

$$d(f^n(a), f^n(b)) < \epsilon$$

whenever $f^n(a), f^n(b) \in K$ *for* $n \geq n_0$

Proof. We split the proof in two cases.

First case: $X = \mathbb{T}^3 \setminus \mathcal{C}$.

Consider the lifts to the universal cover $\tilde{A}, \tilde{f} : \mathbb{R}^3 \to \mathbb{R}^3$ the lift of the conjugacy such that $\tilde{h}(0) = 0$. We now that

$$\tilde{A}^n \circ \tilde{h} = \tilde{h} \circ \tilde{f}^n \tag{5.1}$$

for all n. Thus

$$(e^{\lambda^c_A})^n ||\tilde{h}(a) - \tilde{h}(b)|| \geq ||\tilde{A}^n \circ \tilde{h}(a) - \tilde{A}^n \circ \tilde{h}(b)|| = ||\tilde{h} \circ \tilde{f}^n(a) - \tilde{h} \circ \tilde{f}^n(b)||.$$

Since h s a bounded distance from the identity, then if $d(a, b) < \dfrac{1}{2}$ we have:

$$||\tilde{A}^n \circ \tilde{h}(a) - \tilde{A}^n \circ \tilde{h}(b)|| \leq (e^{\lambda^c_A})^n D$$

for a certain constant $D > 0$. By the uniform continuity of h^{-1} inside $h(K)$ we can take n_0 big enough so that $n \geq n_0$ implies:

$$d(f^n(a), f^n(b)) < \epsilon$$

Second case: X =set of atoms.

In this case, we know by Theorem 5.1.14 that each leaf has only one atom, that is, for almost every $x \in \mathbb{T}^3$ we have $X \cap \mathcal{F}^c(x) = \{a_x\}$ Thus, given any two points $a, b \in X$, a and b are not collapsed by h (since they do not belong to the same central leaf). Thus the proof of the first case works for this case as well.

□

Lemma 5.2.9. *For any $\delta > 0$, there exists $0 < \delta_1 < \delta$ with the following property. Let Π be a δ_1-rectangle and E a set intersecting Π leafwise. Then we can construct a bijective function $\theta : E \cap \Pi \to \Pi$ such that for every measurable set $F \subset E \cap \Pi$ we have*

$$\frac{m^P_\Pi(\theta(F))}{m^P_\Pi(\Pi)} = \frac{m^P_\Pi(F)}{m^P_\Pi(E \cap \Pi)}$$

and for every $x \in E \cap \Pi$

$$\theta(x) \in \mathcal{F}^{cs}(x).$$

Proof. Since E intersects Π leafwise by Lemma 5.2.5 we know that $E \cap \Pi$ is a sub-rectangle, and since the center stable foliation is absolutely continuous the intersection $\mathcal{F}^{cs}(x) \cap E \cap \Pi$ has positive leaf Lebesgue measure for almost every $x \in E \cap \Pi$. Let $z \in E \cap \Pi$ be any point such that $\mathcal{F}^{cs}(z) \cap E \cap \Pi$ has positive leaf Lebesgue measure. To construct θ we will first define it in the subset $\mathcal{F}^{cs}(z) \cap E \cap \Pi$ of the center-stable leaf $\mathcal{F}^{cs}(z)$ of z and then, using unstable holonomy and leafwise intersection property of $E \cap \Pi$, we will extend it to the whole set $E \cap \Pi$. Because $\mathcal{F}^{cs}(z) \cap E \cap \Pi$ and $\mathcal{F}^{cs}(z) \cap \Pi$ are both probability Lebesgue spaces with the normalized m_f^{cs}-measures we can construct a bijection $\theta_0 : \mathcal{F}^{cs}(z) \cap E \cap \Pi \to \mathcal{F}^{cs}(z) \cap \Pi$ preserving these normalized measures, that is, for any measurable subset $J \subset \mathcal{F}^{cs}(z) \cap E \cap \Pi$ we have

$$\frac{m_f^{cs}(\theta_0(J))}{m_f^{cs}(\Pi)} = \frac{m_f^{cs}(J)}{m_f^{cs}(\mathcal{F}^{cs}(z) \cap E \cap \Pi)} \tag{5.2}$$

Since $m_\Pi^P := m_z^u \times m_f^{cs}$ and $E \cap \Pi$ is a leafwise intersection we have that $m_\Pi^P(\Pi) = m_f^{cs}(\mathcal{F}^{cs}(z) \cap \Pi)$ and $m_\Pi^P(E \cap \Pi) = m_f^{cs}(\mathcal{F}^{cs}(z) \cap E \cap \Pi)$, then (5.2) can be rewritten as

$$\frac{m_f^{cs}(\theta_0(J))}{m_\Pi^P(\Pi)} = \frac{m_f^{cs}(J)}{m_\Pi^P(E \cap \Pi)} \tag{5.3}$$

Now, given any $y \in E \cap \Pi$ we define (using that the intersection is leafwise and inside the rectangle) $\theta(y) \in \Pi$ by

$$\theta(y) := (\pi_{y,z}^u)^{-1} \circ \theta_0 \circ \pi_{y,z}^u(y) \tag{5.4}$$

where $\pi_{y,z}^u$ denotes the local unstable holonomy from an open set of $\mathcal{F}^{cs}(y)$ to an open set of $\mathcal{F}^{cs}(z)$.

This $\theta : E \cap \Pi \to \Pi$ s well defined and $\theta(y) \in \mathcal{F}^{cs}(y) \cap \Pi$. Since m_Π^P is a product measure, given a measurable set $F \subset E \cap \Pi$ we have

$$m_\Pi^P(F) = \int m_f^{cs}(\pi_{y,z}^u(\mathcal{F}^{cs}(y) \cap F)) dm_z^u(y)$$

Thus,

$$
\begin{aligned}
m_\Pi^P(\theta(F)) &= \int m_f^{cs}(\pi_{y,z}^u(\mathcal{F}^{cs}(y)\cap\theta(F)))dm_z^u(y)\\
&= \int m_f^{cs}(\pi_{y,z}^u\circ\theta(\mathcal{F}^{cs}(y)\cap F))dm_z^u(y)\\
&= \int m_f^{cs}(\theta_0\circ\pi_{y,z}^u(\mathcal{F}^{cs}(y)\cap F))dm_z^u(y)
\end{aligned}
$$

Substituting (5.3) we have

$$
\begin{aligned}
m_\Pi^P(\theta(F)) &= \int m_f^{cs}(\theta_0\circ\pi_{y,z}^u(\mathcal{F}^{cs}(y)\cap F))dm_z^u(y)\\
&= \frac{m_\Pi^P(\Pi)}{m_\Pi^P(E\cap\Pi)}\int m_f^{cs}(\pi_{y,z}^u(\mathcal{F}^{cs}(y)\cap F))dm_z^u(y)\\
&= \frac{m_\Pi^P(\Pi)}{m_\Pi^P(E\cap\Pi)} m_\Pi^P(F)
\end{aligned}
$$

as we wanted to show. □

The following Lemma together with Theorems 3.5.1 and 3.5.2 concludes the proof of Theorem 5.0.1.

Lemma 5.2.10. *Let α be a finite partition with the property that each atom of α has piecewise smooth boundary. Then α is very weak Bernoulli.*

The proof of this Lemma closely follows from the arguments already used by Pesin [17] and Chernov–Haskell [3] with the technical difference that, by Lemma 5.2.9, the function θ, constructed from a tubular intersection to a rectangle containing this intersection, does not preserve stable manifolds as in [17] and [3]. Instead, θ preserves center-stable manifolds. Points belonging to the same center-stable manifold do not have the property of getting exponentially close to each other as we iterate the dynamics, thus we can not directly say that given $\epsilon >$ a large set of pairs of points on the same center-stable manifold will asymptotically visit the same atoms.

To overcome this difficulty we use Theorem 5.1.14. It says that either $m(C) = 0$ and then we have defined $X = \mathbb{T}^3 \setminus \mathcal{C}$, or $m(C) = 1$ and then we can take a full measure set $X \subset \mathbb{T}^3$ intersecting almost every center-leaf in exactly one point. By Lemma 5.2.8, we can choose arbitrarily large compact sets K such that any pair of point $x, y \in K \subset X$ have the property that if $y \in \mathcal{F}^{cs}(x)$ then the distance between f^n (x) and $f^n(y)$ is very small every time both of them visits the set K simultaneously. The point is that since we can take K arbitrarily large and f is ergodic, the set of natural numbers $\{n : f^n(x), f^n(y) \in K\}$ has arbitrarily large density, this will allow us to conclude that indeed for a large set of points $x \in X$ the Cesaro-mean of Lemma 3.5.7 are indeed arbitrarily small.

Proof. As explained on the last paragraph of last section, the technical difference of our case is that we have to restrict ourselves to large compact sets where points on the same center-stable leaves behave well. Therefore, we will keep notations similar to the notation used by Chernov–Haskell so that we can omit some calculations, which are equal to the (nonuniformly) hyperbolic case, and refer the reader to [3] for the detailed estimates.

Let $\alpha = \{A_1, A_2, ...A_b\}$ be a finite partition of M where each element has piecewise smooth boundaries. Thus we can take a D_0 such that for any $\epsilon > 0$ and any $i = 1, ..., b$, the ϵ-neighborhood of ∂A_i , denoted by $O_\epsilon(A_i)$, has volume measure less than $D_0\epsilon$.

Let $\epsilon > 0$ and, as in [3], we take $\delta = \epsilon^4$. Let $\{R_1, ..., R_k\}$ be a δ-regular covering of M and define the partition $\pi = \{R_0, R_1, ..., R_k\}$ of M by taking

$$R_0 := M \setminus \bigcup_{i=1}^{k} R_i.$$

By the definition of δ-regular covering we have $m(R_0) < \delta$ and, for each $1 \leq i \leq k$, we can take a set $G_i \subset R_i$ which satisfies condition (2) of Definition 5.2.2. Since f has the K-property there exists a natural number N such that for all $N_1 > N_0 \geq N$ and for δ-almost every atom $A \in_{N_0}^{N_1} f^i\alpha$ has the property that for all $R \in \pi$

$$\left|\frac{m(R \cap A)}{m(R)m(A)} - 1\right| < \delta$$

that is

$$\left|\frac{m(R/A)}{m(R)} - 1\right| < \delta \tag{5.5}$$

where $m(./A)$ denotes the measure m conditioned on A.

Now, fix natural numbers $N_1 > N_0 > N$ and $n > 0$. We want to prove that for a certain constant $D > 0$, $D\epsilon$-almost every atom of $\bigvee_{N_0}^{N_0} f^i\alpha$ satisfies

$$\overline{d}(\{f^{-i}\alpha\}_1^n, \{f^{-i}\alpha|A\}_1^n) \leq D\epsilon$$

As in [3], the first step is to identify the set of "bad" elements which will have measure at most $D\epsilon$.

First bad set B_1: Let B_1 be the union of all atoms of $\bigvee_{N_0}^{N_1} f^i\alpha$ which do not satisfy (5.5).

Second bad set B_2: Denote $F_2 = \bigcup_{i=1}^{k} R_i \setminus G_i$. Take B_2 to be the union of atoms $A \in \bigvee_{N_0}^{N_1} f^i\alpha$ such that either

$$m(F_2/A) > \delta^{1/2}$$

or

$$\sum_{i=1}^{k} \frac{m_{R_i}^P(A \cap F_2)}{m(A)} > \delta^{1/2}$$

It follows that $m(B_2) < c_1 \cdot \delta^{1/2}$ for some constant $c_1 > 0$.

Third bad set B_3: By Lemma 5.2.7, given a rectangle Π and any $\beta > 0$ we can find $\tilde{N}_1$ such that for all $\tilde{N}' \geq \tilde{N} \geq \tilde{N}_1$ and β-almost every atom $A \in \bigvee_{\tilde{N}}^{\tilde{N}'} f^i\alpha$, there exists a subset $E \subset A$, intersecting the rectangle Π leafwise for which

$$\frac{m(E)}{m(A)} \geq 1 - \beta.$$

Taking β small enough and N big enough, the set F_3 of all points $x \in M \setminus R_0$ which lies in a non-leafwise intersection (with respect to the rectangle of π containing x) satisfies

$$m(F_3) < \delta.$$

Let B_3 denote the union of all atoms A of $\bigvee_{N_0}^{N_1} f^i\alpha$ for which

$$m(F_3/A) > \delta^{1/2}.$$

Then it follows that $m(B_3) < c_1 \cdot \delta^{1/2}$.

It follows from the estimates of the bad sets that the union of all atoms on the complement of $B_1 \cup B_2 \cup B_3$ is at least $1 - c_1\epsilon$ for a certain constant $c_1 > 0$.

Let A be an atom of $\bigvee_{N_0}^{N_1} f^i\alpha$ which is in the complement of B_1, B_2 and B_3. By Lemma 5.2.9, for each $1 \leq i \leq k$ with $A \cap R_i \neq \emptyset$, we can construct a bijective function $\theta_i : A \cap R_i \cap F_3^c \longrightarrow R_i$ satisfying

$$\frac{m_{R_i}^P(\theta_i(B))}{m_{R_i}^P(R_i)} = \frac{m_{R_i}^P(B)}{m_{R_i}^P(A \cap R_i \cap F_3^c)}.$$

for every measurable set $B \subset A \cap R_i \cap F_3^c$ and

$$\theta_i(x) \in \mathcal{F}^{cs}(x),$$

for every $x \in A \cap R_i \cap F_3^c$. If we denote $\mu_{R_i}^P := m_{R_i}^P(\cdot/R_i)$ we can rewrite the previous equality as

$$\mu^P_{R_i}(B/A \cap F_3^c) = \mu^P_{R_i}(\theta_i(B)),$$

for every measurable set $B \subset A \cap R_i \cap F_3^c$.

Lemma: There exists a constant $c_2 > 0$ and a measurable set $E_2 \subset \mathbb{T}^3$ with

$$m(E_2) < c_2 \cdot \epsilon,$$

and such that for each $1 \leq i \leq k$ with $A \cap R_i \neq \emptyset$ the function θ_i satisfies

$$\left| \frac{m(B/A \cap R_i)}{m(\theta_i(B)/R_i)} - 1 \right| < c_2 \cdot \epsilon,$$

for any $B \subset A \cap R_i \cap E_2^c$.

In [3] pag. 24 and 25 is a detailed proof of this Lemma.

Now define the function $\theta : A \longrightarrow \mathbb{T}^3$ as $\theta(x) = \theta_i(x)$ if $x \in A \cap R_i \cap F_3^c$ for some $1 \leq i \leq k$, and $\theta(x) = x$ otherwise.

Lemma: $\theta : A \longrightarrow \mathbb{T}^3$ is $c_2 \cdot \epsilon$-measure preserving.

Proof. See [3] pag. 26. □

We are almost ready to conclude the proof that α is very weak Bernoulli. Remember that the function $\theta : A \longrightarrow \mathbb{T}^3$ from above has the property of being $c_2 \cdot \epsilon$-measure preserving and

$$\theta(x) \in \mathcal{F}^{cs}(x) \cap R_i$$

for any $x \in A \cap R_i$, $1 \leq i \leq k$. To use Lemma 3.5.7 and finish the proof we still need to prove that the Cesaro sum that appears in Lemma 3.5.7 is small for a large set of points x. Here is where we need to restrict ourselves to a large compact set in order to use the Lemma 5.2.8 for the current pair of points in this set.

Take an arbitrary $\zeta < 1$ and consider $K \subset E_2^c \cap X$ (see Definition 5.1.15) a compact set with

$$m(K) > \zeta \cdot m(E_2^c).$$

Take κ the set of points of K such that the past and future Birkhoff averages coincide and converge to $m(K)$, that is: $x \in \kappa$ if $x \in K$ and

$$\lim_{n\to-\infty}\frac{1}{n}\sum_{j=0}^{n-1}\chi_K(f^j(x))=\lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}\chi_K(f^j(x))=m(K).$$

By Birkhoff's Theorem we know that $m(\kappa)=m(K)$. Take $P:=\kappa\cap\theta^{-1}(\kappa)$. Since θ is $c_2\epsilon$-measure preserving, taking ζ close enough to one we have that

$$m(E_2^c\setminus P)\leq 2c_2\cdot\epsilon.$$

Observe that by Lemma 5.2.8 we can take $n_0\geq 0$ such that for any $x\in\kappa$ we have

$$d(f^n(x),f^n(\theta(x)))<\epsilon,$$

for all $n\geq n_0$ with $f^n(x),f^n(\theta(x))\in\kappa$.

Let $\ell_i(x)$ be the name of x with respect to the sequence of partitions $\xi_i:=f^{-i}\alpha|A$ and $m_i(x)$ the name of x with respect to the partitions $\eta_i=f^{-i}\alpha$. If $x\in\kappa$, $i\geq n_0$ and $e(\ell_i(x)-m_i(\theta(x)))=1$ then either:

- $f^i(x)\notin\kappa$ or $f^i(\theta(x))\notin\kappa$ or
- $d(f^i(x),f^i(\theta(x)))<\epsilon$ and then

$$d(f^i(x),\partial A_{\ell_i(x)})<\epsilon\Rightarrow f^i(x)\in B_\epsilon(A_{\partial\ell_i(x)})).$$

Take

$$B_\epsilon:=\bigcup_{i=1}^{k}B_\epsilon(\partial A_i)$$

and consider

$$J^x=\{j\in\mathbb{N}\text{ such that }f^j(x)\notin\kappa\text{ or }f^j(\theta(x))\notin\kappa\},$$

$$J_n^x:=J^x\cap[1,n].$$

By the definition of the function e we have:

$$\frac{1}{n}\sum_{i=1}^{n}e(\ell_i(x)-m_i(\theta(x)))\leq\frac{1}{n}\sum_{i=1}^{n}\chi_{B_\epsilon}(f^j(x))+\frac{1}{n}\#J_n^x$$

By ergodicity the right side converges to $[m(B_\epsilon)+\text{dens}(J^x)]$ for almost every x.

It follows from the fact that we can take K with arbitrarily large measure and $m(B_\epsilon) < D_0 \cdot \epsilon$ that there exist a set $\widehat{P} \subset P$ with measure $m(\widehat{P}) > 1 - c_3\epsilon$ such that for all $x \in \widehat{P}$

$$\frac{1}{n}\sum_{i=1}^{n} e(\ell_i(x) - m_i(\theta(x))) \leq c_3 \cdot \epsilon,$$

for a certain constant $c_3 > 0$ and n large. Applying Lemma 3.5.7 we conclude that

$$\overline{d}(\{\xi_i\}_1^n, \{\eta_i\}_1^n) \leq c_4\epsilon, \tag{5.6}$$

for some constant $c_4 > 0$ which does not depend on ϵ. Since $\epsilon > 0$ is arbitrary it follows that α is very weak Bernoulli as we wanted to show.

$\square$

Bibliography

[1] Patrick Billingsley. Ergodic theory and information. Technical report, 1965.

[2] Keith Burns and Amie Wilkinson. On the ergodicity of partially hyperbolic systems. *Annals of Mathematics*, pages 451–489, 2010.

[3] NI Chernov and Cymra Haskell. Nonuniformly hyperbolic k-systems are bernoulli. *Ergodic theory and dynamical systems*, 16(1):19–44, 1996.

[4] Jack Elman Clark. *A Kolmogorov shift with no roots.* Department of Mathematics, Stanford University., 1972.

[5] J Franks. Anosov diffeomorphisms. 1970 global analysis (proc. sympos. pure math., vol. xiv, berkeley, calif., 1968) pp. 61–93 amer. *Math. Soc., Providence, RI.*

[6] Paul R Halmos. *Lectures on ergodic theory.* Courier Dover Publications, 2017.

[7] A. N. Kolmogorov. New metric invariant of transitive dynamical systems and automorphisms of lebesgue spaces. 1958.

[8] Andrei Nikolaevitch Kolmogorov. Entropy per unit time as a metric invariant of automorphisms. In *Dokl. Akad. Nauk SSSR*, volume 124, pages 754–755, 1959.

[9] Anthony Manning. There are no new anosov diffeomorphisms on tori. *American Journal of Mathematics*, 96(3):422–429, 1974.

[10] Krerley Oliveira and Marcelo Viana. Fundamentos da teoria ergódica. *IMPA, Brazil*, pages 3–12, 2014.

[11] Donald Ornstein. Bernoulli shifts with the same entropy are isomorphic. *Advances in Mathematics*, 4(3):337–352, 1970.

[12] Donald Ornstein and Benjamin Weiss. Geodesic flows are bernoullian. *Israel Journal of Mathematics*, 14(2):184–198, 1973.

[13] Donald S Ornstein. Imbedding bernoulli shifts in flows. In *Contributions to ergodic theory and probability*, pages 178–218. Springer, 1970.

[14] Donald S Ornstein. An example of a kolmogorov automorphism that is not a bernoulli shift. *Advances in Mathematics*, 10(1):49–62, 1973.

[15] Donald S Ornstein. Ergodic theory, randomness, and" chaos". *Science*, 243(4888):182–187, 1989.

[16] Donald S Ornstein and Paul C Shields. An uncountable family of k-automorphisms. *Advances in Mathematics*, 10(1):63–88, 1973.

[17] Yakov Borisovich Pesin. Characteristic lyapunov exponents and smooth ergodic theory. *Uspekhi Matematicheskikh Nauk*, 32(4):55–112, 1977.

[18] Mark Semenovich Pinsker. Dynamical systems with completely positive or zero entropy. In *Doklady Akademii Nauk*, volume 133, pages 1025–1026. Russian Academy of Sciences, 1960.

[19] S Polit. *Weakly isomorphic maps need not be isomorphic*. PhD thesis, PhD thesis, Stanford, 1974.

[20] Gabriel Ponce, Ali Tahzibi, and Régis Varão. On the bernoulli property for certain partially hyperbolic diffeomorphisms. *Advances in Mathematics*, 329:329–360, 2018.

[21] Charles Pugh and Michael Shub. Ergodicity of anosov actions. *Inventiones mathematicae*, 15(1):1–23, 1972.

[22] Vladimir Abramovich Rokhlin. On the fundamental ideas of measure theory. *Matematicheskii Sbornik*, 67(1):107–150, 1949.

[23] Vladimir Abramovich Rokhlin. Metric properties of endomorphisms of compact commutative groups. *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya*, 28(4):867–874, 1964.

[24] VA Roklin and Ja G Sinai. Construction and properties of invariant measurable partitions. In *Dokl. Akad. Nauk SSSR*, volume 141, pages 1038–1041, 1961.

[25] Daniel J Rudolph. Two non isomorphic k-automorphisms with isomorphic squares. *Israel Journal of Mathematics*, 23(3-4):274–287, 1976.

[26] Paul C Shields. *The theory of Bernoulli shifts*. University of Chicago Press Chicago, 1973.

[27] Raúl Ures. Intrinsic ergodicity of partially hyperbolic diffeomorphisms with a hyperbolic linear part. *Proceedings of the American Mathematical Society*, 140(6):1973–1985, 2012.

[28] Peter Walters. *An introduction to ergodic theory*, volume 79. Springer Science & Business Media, 2000.

[29] Sergei Aronovich Yuzvinskii. Metric properties of the endomorphisms of compact groups. *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya*, 29(6):1295–1328, 1965.