

---

# ERGODIC THEORY

---

## MATH 470 Report

**Samy Lahlou Kamal**

Montreal, Canada

Summer 2024

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Some Useful Definitions</b>	<b>3</b>
2.1	Measure Theory . . . . .	3
2.2	Basics of Ergodic Theory . . . . .	4
<b>3</b>	<b>Ergodic Theorems</b>	<b>7</b>
3.1	Birkhoff Ergodic Theorem . . . . .	7
3.2	Backward Ergodic Theorems . . . . .	10
3.3	Group Actions Ergodic Theorems . . . . .	12

# 1 Introduction

In this report, I will summarize the topics I have covered with Professor Anush Tserunyan and Kaining Jia during the summer 2024. The main topic was Ergodic Theory, but we focused on the Birkhoff Ergodic Theorem, some variations and generalizations (using group actions) as well as new proof methods including the tilings and boundary actions.

I will not write everything I learned during the summer in this report because I also had to learn about Measure Theory and did a lot of exercises on the Riemann Integral and Functional Analysis.

## 2 Some Useful Definitions

### 2.1 Measure Theory

It would be hard to read this report without some background in measure theory. However, I will still recall some useful definitions since the main structures we will deal with here are probability measure spaces.

**Definition 2.1** (Measure space and Probability Measure Space). Let  $X$  be a set,  $\beta \subset 2^X$  be a  $\sigma$ -algebra on  $X$  and  $\mu : \beta \rightarrow \mathbb{R}$  be a function, then the triplet  $(X, \beta, \mu)$  is a *measure space* if :

- $\mu(\emptyset) = 0$
- $\mu(\bigcup E_n) = \sum_{n=0}^{\infty} \mu(E_n)$  for any pairwise disjoint collection  $\{E_n\}_{n \in \mathbb{N}} \subset \beta$

Moreover, we say that  $(X, \beta, \mu)$  is a *probability measure space* if

$$\mu(X) = 1$$

We refer to the elements of  $\beta$  as *measurable sets*.

Sometimes, we can simply refer to a probability measure space as a *probability space*. Moreover, if the  $\sigma$ -algebra is clear from the context, we can simply refer to  $(X, \beta, \mu)$  as  $(X, \mu)$ .

Some standard probability measure space are  $([0, 1], \lambda)$ , where  $\lambda$  is the Lebesgue Measure, or  $(2^{\mathbb{N}}, \{\frac{1}{3}, \frac{2}{3}\}^{\mathbb{N}})$ .

Another key tool that will be used very often here are the measurable functions and transformations.

**Definition 2.2.** Let  $(X, \beta_1, \mu_1)$  and  $(Y, \beta_2, \mu_2)$  be measure spaces and  $f : X \rightarrow Y$  be a function. Then  $f$  is  $(\beta_1, \beta_2)$ -*measurable* if

$$f^{-1}B \in \beta_1$$

for all  $B \in \beta_2$ . If the  $\sigma$ -algebras are clear from the context, then we simply say that  $f$  is *measurable*.

Again, as an example, the baker's map  $b : [0, 1) \rightarrow [0, 1)$  is a measurable transformation if the measure is the Lebesgue measure  $\lambda$ . Another example would be the shift  $\sigma : 2^{\mathbb{N}} \rightarrow 2^{\mathbb{N}}$  defined by  $(x_n) \mapsto (x_{n+1})$  where the measure is  $\{\frac{1}{2}, \frac{1}{2}\}^{\mathbb{N}}$ .

## 2.2 Basics of Ergodic Theory

In this subsection, fix a probability space  $(X, \mu)$ . Here, I will define all notions that will let us define and understand the notion of ergodicity, which is the main notion of the report. Ergodicity can be defined for any measurable transformation  $T$  but we will only focus here on probability measure preserving transformations since the main theorems only apply to these transformations.

**Definition 2.3** (Probability Measure Preserving). A measurable transformation  $T : X \rightarrow X$  is *probability measure preserving* if

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

for any measurable set  $B$ . It will be easier to refer to this property as *pmp*.

This definition is not hard to understand and is self-contained in its name. A pmp transformation preserves the measure of a set. A very basic example would be a translation or rotation with the Lebesgue measure on  $\mathbb{R}^2$ . The next definition is similar.

**Definition 2.4** (T-invariance for sets). For a measurable transformation  $T : X \rightarrow X$ , a set  $B \subset X$  is *T-invariant* if

$$T^{-1}B = B$$

A set is *T-invariant* when its elements simply get *shuffled* by  $T$ . An easy example would be any rotation applied to any circle centered at zero. We can define the notion of *T-invariance* for functions as well.

**Definition 2.5** (T-invariance for functions). For a measurable transformation  $T : X \rightarrow X$ , a function  $f : X \rightarrow \mathbb{R}$  is *T-invariant* if

$$f = f \circ T$$

Equivalently, it is not hard to prove that  $f$  is *T-invariant* if  $f^{-1}(A)$  is a *T-invariant* set for all  $A \subset \mathbb{R}$ .

With all of these definitions in hand, we are now able to define the notion of ergodicity.

**Definition 2.6** (Ergodicity). Let  $T : X \rightarrow X$  be a measurable transformation, then  $T$  is *ergodic* if every *T-invariant* measurable set is null or conull, i.e,  $\mu(B) \in \{0, 1\}$  for all *T-invariant* measurable set  $B$

A transformation is ergodic if we cannot decompose the space  $X$  into positive measure invariant subsets. As for prime numbers or a lot of other notions in mathematics, ergodicity has something to do with some elements being decomposable or not.

Since invariance is important when defining ergodicity and invariance is defined for both sets and functions, we can deduce an ergodicity criterion using invariance of functions.

**Theorem 2.1** (Ergodicity Criterion). *A measurable transformation  $T : X \rightarrow X$  is ergodic iff every  $T$ -invariant measurable function  $f : X \rightarrow \mathbb{R}$  is constant a.e on  $X$ .*

*Proof.* ( $\Rightarrow$ ) Suppose that  $T$  is ergodic and let  $f \in L^1(X, \mu)$  be  $T$ -invariant, let's show that  $f$  is constant a.e on  $X$ . Consider the sets  $A, B \subset \mathbb{R}$  defined as

$$\begin{aligned} A &:= \{a \in \mathbb{R} : \mu(f^{-1}(a, \infty)) = 0\} \\ B &:= \{b \in \mathbb{R} : \mu(f^{-1}(-\infty, b]) = 0\} \end{aligned}$$

First, let's show that  $A$  and  $B$  are disjoint. By contradiction, suppose that there exists an  $x \in \mathbb{R}$  such that  $x \in A \cap B$ , then

$$\begin{aligned} \mu(f^{-1}(x, \infty)) &= 0 \\ \mu(f^{-1}(-\infty, x]) &= 0 \end{aligned}$$

So we get

$$\begin{aligned} \mu(X) &= \mu(f^{-1}(\mathbb{R})) \\ &= \mu(f^{-1}((-\infty, x] \cup (x, \infty))) \\ &= \mu(f^{-1}(-\infty, x]) \cup f^{-1}(x, \infty) \\ &= \mu(f^{-1}(-\infty, x]) + \mu(f^{-1}(x, \infty)) \\ &= 0 + 0 \\ &= 0 \end{aligned}$$

A contradiction, so  $A \cap B = \emptyset$ . Let's now prove that  $\mathbb{R} = A \cup B$ . Let  $x \in \mathbb{R} \setminus A$ , then  $\mu(f^{-1}(-\infty, x]) \neq 0$ . By ergodicity of  $T$  and by the  $T$ -invariance of  $f$ ,  $f^{-1}(-\infty, x]$  is  $T$ -invariant and measurable, so we must have  $\mu(f^{-1}(-\infty, x]) = 1$ . Thus,

$$\begin{aligned} \mu(X) &= \mu(f^{-1}(-\infty, x]) + \mu(f^{-1}(x, \infty)) \\ \implies 1 &= 1 + \mu(f^{-1}(x, \infty)) \\ \implies \mu(f^{-1}(x, \infty)) &= 0 \\ \implies x &\in B \end{aligned}$$

Hence,  $\mathbb{R} = A \cup B$ . Let's now prove that both  $A$  and  $B$  are nonempty. Suppose that  $B = \emptyset$ , then  $A = \mathbb{R}$ . In this case  $n \in A$  for all  $n \in \mathbb{N}$ :

$$\begin{aligned}
\mu(X) &= \mu(f^{-1}(\mathbb{R})) \\
&= \mu\left(f^{-1}\left(\bigcup_{n=0}^{\infty}(-\infty, n]\right)\right) \\
&= \mu\left(\bigcup_{n=0}^{\infty}(f^{-1}(-\infty, n])\right) \\
&\leq \sum_{n=0}^{\infty} \mu(f^{-1}(-\infty, n]) \\
&= \sum_{n=0}^{\infty} 0 \\
&= 0
\end{aligned}$$

Thus,  $B$  is nonempty. Similarly,  $A$  is nonempty as well. Finally, let's show that if  $a_0 \in A$  and  $a \leq a_0$ , then  $a \in A$ :

$$\begin{aligned}
a &\leq a_0 \\
\implies (-\infty, a] &\subset (-\infty, a_0] \\
\implies f^{-1}(-\infty, a] &\subset f^{-1}(-\infty, a_0] \\
\implies \mu(f^{-1}(-\infty, a]) &\leq \mu(f^{-1}(-\infty, a_0]) = 0 \\
\implies \mu(f^{-1}(-\infty, a]) &= 0 \\
\implies a &\in A
\end{aligned}$$

Similarly, if  $b_0 \in B$  and  $b_0 \leq b$ , then  $b \in B$ . Using these properties of  $A$  and  $B$ , it follows by Completeness of  $\mathbb{R}$  that there exists a  $c \in \mathbb{R}$  such that  $(-\infty, c) \subset A$  and  $(c, \infty) \subset B$  ( $c = \sup(A) = \inf(B)$ ). Now, it remains to show that  $f(x) = c$  for a.e  $x \in X$ , to do this, we will show equivalently that  $\mu(f^{-1}(\mathbb{R} \setminus \{c\})) = 0$ :

$$\begin{aligned}
\mu(f^{-1}(\mathbb{R} \setminus \{c\})) &= \mu(f^{-1}(-\infty, c) \cup f^{-1}(c, \infty)) \\
&= \mu(f^{-1}(-\infty, c)) + \mu(f^{-1}(c, \infty)) \\
&= \mu\left(f^{-1}\bigcup_{n=1}^{\infty}(-\infty, c - \frac{1}{n}]\right) + \mu\left(f^{-1}\bigcup_{n=1}^{\infty}(c + \frac{1}{n}, \infty)\right) \\
&\leq \sum_{n=1}^{\infty} \mu(f^{-1}(-\infty, c - \frac{1}{n}]) + \sum_{n=1}^{\infty} \mu(f^{-1}(c + \frac{1}{n}, \infty)) \\
&= \sum_{n=1}^{\infty} 0 + \sum_{n=1}^{\infty} 0 \\
&= 0
\end{aligned}$$

This concludes the first part of the proof.

( $\Leftarrow$ ) Suppose now that every  $T$ -invariant measurable function  $f : X \rightarrow \mathbb{R}$  is constant a.e on  $X$ . Let's show that  $T$  is ergodic. Let  $B \subset X$  be a  $T$ -invariant measurable set and consider

the function  $\chi_B : X \rightarrow \mathbb{R}$ . Since  $B$  is measurable, then  $\chi_B$  is measurable as well. Let  $x \in X$ , since  $T^{-1}B = B$ , we must have  $\chi_B(x) = \chi_B(T(x))$ , so  $\chi_B$  is  $T$ -invariant. By our assumption on  $T$ -invariant measurable functions,  $\chi_B$  is a.e constant. If  $\chi_B(x) = 1$  for a.e  $x \in X$ , then  $B$  is conull. If  $\chi_B(x) = 0$  for a.e  $x \in X$ , then  $B$  is null. Therefore,  $T$  is ergodic.  $\square$

This criterion will be very useful when proving the Birkhoff Ergodic Theorem.

### 3 Ergodic Theorems

#### 3.1 Birkhoff Ergodic Theorem

Let  $(X, \mu)$  be a probability space with a pmp transformation  $T : X \rightarrow X$ .

In this subsection, we will focus on the Birkhoff Ergodic Theorem and prove it. To prove this theorem, we will first define a few functions and prove two important lemmas : the local-global bridge and the  $T$ -invariance of  $\bar{f}$  and  $\underline{f}$  (which will be defined later).

**Theorem 3.1** (Birkhoff Ergodic Theorem).  *$T$  is ergodic iff for each  $f \in L^1(X, \mu)$  and a.e  $x \in X$ ,*

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{i=0}^n (f \circ T^i)(x) = \int_X f d\mu$$

What this theorem says is that ergodicity implies that for almost every point  $x$  in the space  $X$ , the *forward orbit* of  $x$ ,  $\{x, Tx, T^2x, \dots\} = \{T^n x : n \in \mathbb{N}\}$  is enough to sample every part of the space. Moreover,  $T$  being pmp makes this sampling of  $X$  uniform enough to give us arbitrarily precise approximations of  $\int_X f d\mu$  for any  $f \in L^1(X, \mu)$ .

**Definition 3.1.** For a function  $f \in L^1(X, \mu)$  and  $n \in \mathbb{N}$ , define  $A_n f : X \rightarrow \mathbb{R}$  by

$$A_n f(x) = \frac{1}{n+1} \sum_{i=0}^n (f \circ T^i)(x)$$

for  $x \in X$ . For a fixed  $x \in X$ ,  $A_n f(x)$  is the average of  $f$  over  $x, Tx, T^2x, \dots, T^n x$ .

**Lemma 3.2** (Local-global bridge). *For all  $f \in L^1(X, \mu)$  and any natural number  $n \geq 1$ ,*

$$\int_X f d\mu = \int_X A_n f d\mu$$

*Proof.* Let  $f \in L^1(X, \mu)$  and  $n \geq 1$ . Let's first show that

$$\int_X (f \circ T) d\mu = \int_X f d\mu$$

. Since  $T$  is pmp, for any measurable set  $B \in \beta$ :

$$(T_*\mu)(B) = \mu(T^{-1}B) = \mu(B)$$

so  $T_*\mu = \mu$ . Thus, using the change-of-variable formula :

$$\int_X (f \circ T) d\mu = \int_X f d(T_*\mu) = \int_X f d\mu$$

Hence, by induction on  $i \in \mathbb{N}$ :

$$\int_X (f \circ T^i) d\mu = \int_X f d\mu$$

Therefore,

$$\begin{aligned} \int_X A_n f d\mu &= \int_X \frac{1}{n+1} \sum_{i=0}^n (f \circ T^i) d\mu \\ &= \frac{1}{n+1} \sum_{i=0}^n \int_X (f \circ T^i) d\mu \\ &= \frac{1}{n+1} \sum_{i=0}^n \int_X f d\mu \\ &= \frac{n+1}{n+1} \int_X f d\mu \\ &= \int_X f d\mu \end{aligned}$$

□

The statement of the Ergodic Theorem deals with the function  $x \mapsto \lim_{n \rightarrow \infty} A_n f(x)$ , hence, it is clear to see the role that the next definition will play in the proof of the theorem.

**Definition 3.2.** For a function  $f \in L^1(X, \mu)$ , define the functions  $\bar{f}$  and  $\underline{f}$  by

$$\bar{f}(x) = \limsup_n A_n f(x)$$

and

$$\underline{f}(x) = \liminf_n A_n f(x)$$

for  $x \in X$ .

**Lemma 3.3** (Invariance of  $\bar{f}$  and  $\underline{f}$ ). *For all  $f \in L^1(X, \mu)$ , the functions  $\bar{f}$  and  $\underline{f}$  are  $T$ -invariant.*

*Proof.* Let  $x \in X$  :

$$\begin{aligned} (\bar{f} \circ T)(x) &= \limsup_n A_n f(Tx) \\ &= \limsup_n \frac{1}{n+1} \sum_{i=0}^n (f \circ T^{i+1})(x) \\ &= \limsup_n \frac{1}{n+1} \left( \sum_{i=0}^{n+1} (f \circ T^i)(x) - f(x) \right) \\ &= \limsup_n \left( \frac{n+2}{n+1} A_{n+1} f(x) - \frac{1}{n+1} f(x) \right) \\ &= \limsup_n A_{n+1} f(x) \\ &= \limsup_n A_n f(x) \\ &= \bar{f}(x) \end{aligned}$$



So  $\bar{f}$  is  $T$ -invariant. Similarly,  $\underline{f}$  is  $T$ -invariant as well. □

Let's now prove the Ergodic Theorem.

*Proof.* ( $\Rightarrow$ ) Suppose that  $T$  is ergodic and let's show using a tiling argument that

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{i=0}^n (f \circ T^i)(x) = \int_X f d\mu$$

for all  $f \in L^1(X, \mu)$  and a.e  $x \in X$ . Let  $f \in L^1(X, \mu)$ , then using Lemma 3.3 and the Ergodicity Criterion 2.1, both  $\bar{f}$  and  $\underline{f}$  are constant a.e on  $X$ . WLOG, we will suppose that

$$\int_X f d\mu = 0$$

By way of contradiction, suppose that  $\bar{f} > 0$  (a.e), then there must be a  $\Delta > 0$  s.t

$$\bar{f} > \Delta > 0$$

. Hence, we can define the function  $x \mapsto n_x$  which maps  $x$  to the least positive integer  $n_x$  satisfying  $A_{n_x} f(x) > \Delta$ . We will assume here that  $x \mapsto n_x$  is bounded by some  $l > 0$  and that  $f$  is bounded as well. Let  $\epsilon > 0$  and define  $N \in \mathbb{N}$  such that  $l/(N+1) < \epsilon$ , then for all  $x \in X$ ,

$$\begin{aligned} A_N f(x) &= \frac{1}{N+1} (f(x) + f(Tx) + \dots + f(T^N x)) \\ &> \frac{1}{N+1} (\Delta n_x + \Delta n_{T^{n_x+1}x} + \dots - l\|f\|_\infty) \\ &> \frac{1}{N+1} (l\Delta + l\Delta + \dots - l\|f\|_\infty) \\ &\geq \frac{1}{N+1} \left( \frac{N+1}{l} l\Delta - l\|f\|_\infty \right) \\ &= \Delta - \frac{l}{N+1} \|f\|_\infty \\ &> \Delta - \epsilon \|f\|_\infty \end{aligned}$$

Thus,

$$\begin{aligned} \int_X f d\mu &= \int_X A_N f d\mu > \Delta - \epsilon \|f\|_\infty \\ \Rightarrow \epsilon \|f\|_\infty + \int_X f d\mu &> \Delta \end{aligned}$$

But since it holds for all  $\epsilon > 0$ , by letting  $\epsilon \rightarrow 0$  we get

$$\int_X f d\mu \geq \Delta > 0$$

which contradicts our assumption that  $\int_X f d\mu = 0$ . Thus,  $\bar{f} \leq 0$  a.e on  $X$ , similarly,  $\underline{f} \geq 0$

a.e on  $X$ . But we know that  $\underline{f} \leq \bar{f}$  on  $X$ , so for a.e  $x \in X$ :

$$\begin{aligned} \lim_{n \rightarrow \infty} A_n f(x) &= \underline{f}(x) = \bar{f}(x) = \int_X f d\mu \\ \implies \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{i=0}^n (f \circ T^i)(x) &= \int_X f d\mu \end{aligned}$$

( $\Leftarrow$ ) Suppose that

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{i=0}^n (f \circ T^i)(x) = \int_X f d\mu$$

for all  $f \in L^1(X, \mu)$  and a.e  $x \in X$ , let's show that  $T$  is ergodic using the Ergodicity Criterion 2.1. Let  $f \in L^1(X, \mu)$  be a  $T$ -invariant function, let's show that  $f$  is constant a.e on  $X$ . By our assumptions :

$$\begin{aligned} \int_X f d\mu &= \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{i=0}^n (f \circ T^i)(x) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{i=0}^n f(x) \\ &= f(x) \end{aligned}$$

So if we let  $c := \int_X f d\mu$ , then  $f(x) = c$  for a.e  $x \in X$ . Therefore, by the ergodicity criterion,  $T$  is ergodic.  $\square$

### 3.2 Backward Ergodic Theorems

Again, in this section, fix a probability space  $(X, \mu)$ .

The Backward Ergodic Theorems are a variation on the Birkhoff Ergodic Theorem. To visualize these theorems, recall the notion of forward orbit of an element  $x \in X$ .

**Definition 3.3** (T-Forward-orbit). Let  $T : X \rightarrow X$  be a transformation and  $x$  an element of  $X$ , then the  $T$ -forward-orbit of  $x$  (or simply *forward-orbit* of  $x$ ) is the set

$$\{x, Tx, T^2x, T^3x, \dots\} = \{T^n x : n \in \mathbb{N}\}$$

In general, we can define the general orbit of  $x$  as follows.

**Definition 3.4** (T-Orbit). Let  $T : X \rightarrow X$  be a transformation and  $x$  an element of  $X$ . Define now the equivalence relation  $E_T$  on  $X$  as follows

$$xE_T y : \iff T^n x = T^m y \text{ for some } n, m \in \mathbb{N}$$

for  $x, y \in X$ .

Then the  $T$ -orbit of  $x$  (or simply *orbit* or  $x$ ) is the  $E_T$ -equivalence class of  $x$ . Without mentioning the equivalence relation  $E_T$ , we can define the orbit of  $x$  as follows :

$$[x]_T := \bigcup_{n \in \mathbb{N}} \bigcup_{m \in \mathbb{N}} T^{-m} T^n \{x\}$$

Using these definitions, we can visualise  $T : X \rightarrow X$  as in Figure 1.

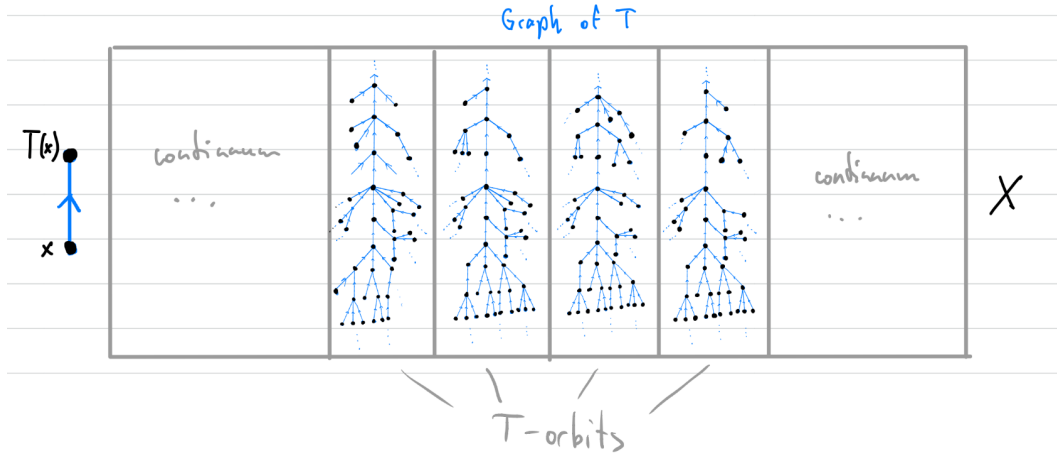


Figure 1: T-orbits

For example, if we focus on a fixed  $x \in X$ , then the Birkhoff Ergodic Theorem involves taking the averages on the red-line in Figure 2 which represents the forward-orbit of  $x$

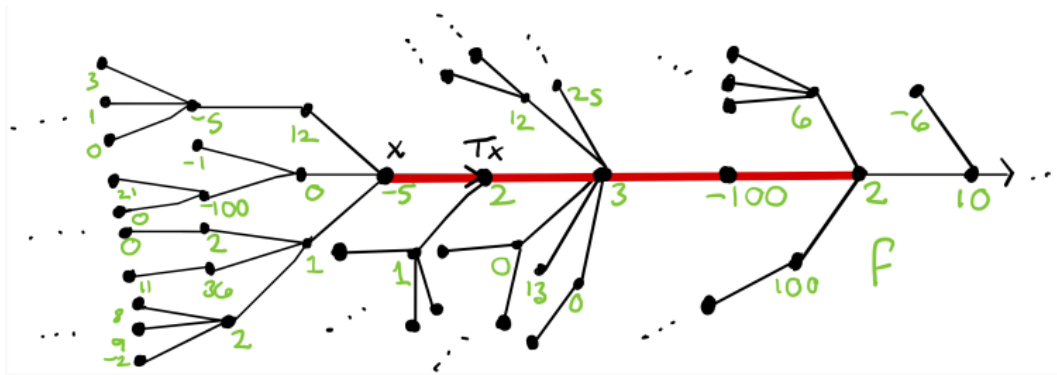


Figure 2: Forward-orbit of  $x$

The Backward Triangle Theorem involves however taking the *weighted* average on backward-orbits of  $x$ .

**Definition 3.5** (T-Backward-orbit). Let  $T : X \rightarrow X$  be a transformation and  $x$  an element of  $X$ , then the  $T$ -backward-orbit of  $x$  (or simply *backward-orbit* of  $x$ ) is the set

$$\triangleright_T^n(x) := \bigcup_{n=0} T^{-n}\{x\}$$

We can now state the Backward Triangle Theorem.

**Theorem 3.4** (Backward Triangle Theorem). A countable-to-one pmp transformation  $T : X \rightarrow X$  is ergodic iff

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{y \in \triangleright_T^n(x)} f(y) w_x(y) = \int_X f d\mu$$

for all  $f \in L^1(X, \mu)$  and a.e  $x \in X$ .

In this theorem, the weight function  $w : E_T \rightarrow \mathbb{R}$  is called the *Radon-Nikodym cocycle* and satisfies the following properties :

- $w_x(y) \cdot w_y(z) = w_x(z)$  for all  $x E_T y E_T z$
- for any Borel bijection  $\gamma : A \rightarrow B$  with  $A, B \subset X$  and  $\gamma \subset E_T$ :

$$\mu(\gamma(A)) = \int_A w_x(\gamma(x)) d\mu(x)$$

It is a good way of assigning a relative weight between the elements of  $X$  that are  $E_T$ -equivalent.

However, the previous theorem can be generalized to any backward trees instead of backward triangles. From this idea, we get the Backward Trees Theorem :

**Theorem 3.5** (Backward Trees Theorem). *A countable-to-one pmp transformation  $T : X \rightarrow X$  is ergodic iff*

$$\frac{1}{w_x(\tau_x)} \sum_{y \in \tau_x} f(y) w_x(y) \rightarrow \int_X f d\mu \quad \text{as } w_x(\tau_x) \rightarrow \infty$$

for all  $f \in L^1(X, \mu)$  and a.e  $x \in X$  and where  $\tau_x$  ranges over finite subtrees in  $\triangleright_T^n(x)$ .

These two theorems (the first one is a special case of the second one) can also be proved using the Local-Bridge + Invariance + Tiling technique used previously to prove the Birkhoff Ergodic Theorem.

### 3.3 Group Actions Ergodic Theorems

In this subsection, consider a probability space  $(X, \mu)$ . Using group actions, we can generalize the Birkhoff Ergodic Theorem in many ways. First, let's define ergodicity for equivalence relations and group actions.

**Definition 3.6** (Ergodicity for equivalence relation). Let  $E$  be an equivalence relation, then we say that  $E$  is *ergodic* if each  $E$ -invariant (which must be union of  $E$ -equivalence classes) measurable set is null or conull.

**Definition 3.7** (Ergodicity for Group Actions). A group acting on  $X$  is *ergodic* if its orbits equivalence relation is ergodic.

Using these new definitions, we can now restate the Birkhoff Ergodic Theorem :

**Theorem 3.6** (Birkhoff Ergodic Theorem). *A pmp action  $\mathbb{Z}$  on  $(X, \mu)$  is ergodic iff*

$$\lim_{n \rightarrow \infty} \sum_{y \in F_n \cdot x} f(y) = \int_X f d\mu$$

for all  $f \in L^1(X, \mu)$  and a.e  $x \in X$  where  $F_n = \{0, \dots, n\}$ .

In fact, we can generalize this theorem to any pmp actions of an amenable group along tempered Følner sequences by the Lindenstrauss Theorem. The free groups are non amenable, unfortunately, it was proved by Terence Tao that the pointwise ergodic theorem fails for such groups.

## References

- [1] Bowen and Nevo (2015), Amenable Equivalence Relations and the construction of ergodic averages for group actions.
- [2] Anush Tserunyan (2024), A descriptive approach to pointwise ergodic theorems