Ergodic theory, hitting time statistics

Ronan Memin

ENS Rennes

Internship report - June/July 2016

Aknowledgments

I would like to give my warmest thanks to Mike Todd who allowed me to live my first research experience at the University of St Andrews. I particularly thank Mike for his time, his constantly open door, his answers to my questions and his patience. I am extremely grateful for the efforts Mike and all the people of the mathematics department made in order to make me feel good at the University.
I also thank Jyoti who always brung some fun at the office.
Merci !

Contents

Introduction

When we try to describe physical phenomenons such as the weather, the evolution of the solar system or the repartition of a gas in a room, it is convenient to work in the formalism of dynamical systems. This theory traduces intuitive notions such as stability, recurrence and mixing. In particular, the case of ergodic systems - in which the points explore (almost) all the system - has been highly studied and gives information about the long time behaviour of points. The goal of hitting time statistics theory in ergodic systems is to estimate how fast the points enter regions of the system, refining the estimations provided by ergodic theory.
This document aims at presenting several ideas and basic constructions from ergodic theory. It is mostly divided in two parts:
• Sections 1 to 7 deal with general results and definitions of ergodic theory
• The rest of the document is dedicated in proving two hitting/return time statitics theorems.

1 Notations

If X is a set, for $T : X \to X$ we will denote by T^n the n^{th} iteration of T under composition. For $A \subset X$, $T^{-n}A$ is the inverse image of A under T^n.
Let $(X, \mathscr{F}, \mu)$ be a measure space. If $T : X \to X$ is measurable, we denote by μ_T the image measure of μ by T, defined as $\mu_T(A) := \mu(T^{-1}A)$, for $A \in \mathscr{F}$. μ_T is the unique measure on $(X, \mathscr{F})$ which satisfies the equation:

$$\text{For all } f \in L^1(\mu), \ \int_X f \circ T d\mu = \int_X f d\mu_T.$$

If $x \in X$ we will often denote by Tx the image of x by T. ($Tx = T(x)$).
We will often not distinguish elements of L^p and elements of $\mathscr{L}^p$.
λ will denote the Lebesgue measure.

2 Measure preserving systems

In this section we introduce the notions of **dynamical system** and **measure preserving transformation**. Then we prove Poincaré Recurrence Theorem, which gives a rough idea of the behaviour of the orbit of points under iterations of the tranformation.

2.1 Definitions

Let $(X, \mathscr{F}, \mu)$ be a measure space.
Let $T : X \to X$ be a measurable tranformation.
$(X, \mathscr{F}, \mu, T)$ is called a **dynamical system**.
We say that T is **measure preserving** if for any measurable set A, one has $\mu(T^{-1}A) = \mu(A)$.
$(X, \mathscr{F}, \mu, T)$ is then called a measure preserving system.

To verify that some map is measure preserving, it suffices to prove it for a big enough familly of sets:

Lemma 2.1. *Let $(X, \mathscr{F}, \mu)$ be a measure space with finite measure ($\mu(X) < +\infty$). Let $T : X \to X$ be a measurable map.*
Suppose that $\mathscr{E} \subset \mathscr{F}$ is closed under finite intersections and generates $\mathscr{F}$ (i.e the smallest σ-algebra

containing $\mathscr{E}$ is $\mathscr{F}$).
If T preserves the measure on $\mathscr{E}$, then T is measure preserving, that is:

$$\text{For all } E \in \mathscr{E},\ \mu(T^{-1}E) = \mu(E) \Rightarrow \text{For all } A \in \mathscr{F},\ \mu(T^{-1}A) = \mu(A).$$

Proof. Suppose T is measure preserving on $\mathscr{E}$. Let

$$\mathscr{M} = \left\{A \in \mathscr{F} |\ \mu(T^{-1}A) = \mu(A)\right\}.$$

$\mathscr{M}$ is a monotone class:
- Since $T^{-1}(X) = X$, we have $X \in \mathscr{M}$.
- If $A, B \in \mathscr{M}$ with $A \supset B$, then

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

hence $A \backslash B \in \mathscr{M}$.
- Let $(A_n)_n$ be an increasing sequence of elements of $\mathscr{M}$.

$$\mu(T^{-1}(\cup_n A_n)) = \mu(\cup_n T^{-1}A_n) = \lim_{n\to+\infty} \mu(T^{-1}A_n) = \lim_{n\to+\infty} \mu(A_n) = \mu(\cup_n A_n),$$

and $\cup_n A_n \in \mathscr{M}$.
By hypothesis we have $\mathscr{E} \subset \mathscr{M}$ and by the monotone class lemma, we get

$$\mathscr{F} = \sigma(\mathscr{E}) = m(\mathscr{E}) \subset \mathscr{M} \subset \mathscr{F},$$

where $\sigma(\mathscr{E})$ is the σ-algebra generated by $\mathscr{E}$ and $m(\mathscr{E})$ is the monotone class generated by $\mathscr{E}$. □

Let us give some simple examples:
Let $X = [0, 1]$ endowed with its Borel sigma algebra and the Lebesgue measure λ.
- $T : X \to X$ defined by $T(x) = 2x \mod (1)$ (called the **doubling map**)
- More generally, any $T : X \to X$ defined by $T(x) = mx \mod (1)$, $m \in \mathbb{N}^*$
- $T : X \to X$ defined by $T(x) = (x + \alpha) \mod (1)$, $\alpha \in [0, 1]$ (it is the **rotation of angle** α on the circle)

All these maps are λ-preserving.
Let us prove it for the doubling map (the proof is very similar for the other maps). Thanks to Lemma 2.1 we can verify that T is measure preserving only on the sets of the form $[a, b)$, $a < b \in [0, 1]$. For $a < b \in [0, 1]$,

$$T^{-1}([a, b)) = \left[\frac{a}{2}, \frac{b}{2}\right) \cup \left[\frac{a}{2} + \frac{1}{2}, \frac{b}{2} + \frac{1}{2}\right),$$

and $\lambda(T^{-1}([a, b))) = b - a = \lambda([a, b))$.

For $x \in X$ we might want to look at the set of the iterates of x under T, $\{T^n(x);\ n \in \mathbb{N}\}$. We call that set the **orbit** of x.

2.2 Poincaré Recurrence Theorem

Around 1880, Poincaré focussed on the three body problem:

Given a system of three bodies (let's say planets) interacting through the law of gravitation, can we predict the evolution of the system as time goes by?

Given the difficulty of the problem, Poincaré had the idea of describing the evolution of the system in a qualitative point of view. One of his results was his now famous recurrence theorem.

One can think of X as an evolving physical system, where time is represented by $\mathbb{N}$. Under that formalism the orbit of a point represents the different states it will explore. In the case where μ is a finite measure, Poincaré says that almost every point x is going to come back "near" its initial state. More precisely,

Theorem 2.1. *(Poincaré Recurrence Theorem)*
Let $(X, \mathscr{F}, \mu, T)$ be a measure preserving system with finite measure.
Let A be a measurable set with $\mu(A) > 0$. Then almost every point of A comes back to A infinitely often: For almost every $x \in A$, for infinitely many n, $T^n(x) \in A$.

Lemma 2.2. *Under the previous conditions, let $B \in \mathscr{F}$ with $\mu(B) > 0$. Then there exists $n \in \mathbb{N}^*$ such that $\mu(B \cap T^{-n}B) > 0$.*

Proof of Lemma 2.2. Assume that for all $i,\ j \in \mathbb{N}$ with $i \neq j$ we have $\mu(T^{-i}B \cap T^{-j}B) = 0$.
Then for all $n \in \mathbb{N}$ we have:

$$\mu(X) \geq \mu\left(\bigcup_{k=0}^{n} T^{-k}B\right) = \sum_{k=0}^{n} \mu(T^{-k}B),$$

And since T is measure preserving,

$$\mu(X) \geq (n+1)\mu(B),$$

which is a contradiction for n big enough. □

Proof of theorem 2.1.
Let A be a fixed event with $\mu(A) > 0$. Let E_1 be the set of points wich never come back to A:

$$E_1 = \{x \in A | \forall n \geq 1,\ T^n(x) \in A^c\} = \bigcap_{n \geq 1} (T^{-n}A^c).$$

Since $E_1 \subset A$ one has $E_1 \cap T^{-n}E_1 = \emptyset$ for all $n \geq 1$. Therefore Lemma 2.2 implies that $\mu(E_1) = 0$. Applying the same argument to T^n for each $n \geq 1$ shows that every $E_n = \{x \in A | \forall m \geq 1,\ T^{mn}(x) \in A^c\}$ has measure zero.
We finally get that

$$\mu\left(\bigcup_{n \geq 1} E_n\right) = 0,$$

And since $F := \{x \in A | x \text{ doesn't come back to A infinitely often}\}$ is contained in $\bigcup_{n \geq 1} E_n$, we get

$$\mu(A \backslash F) = \mu(A),$$

that is: Almost every point of A returns infinitely often. □

In the case where X is rich enough we can state that almost every point will come back in any accuracy of its starting location:

Corollary 2.1. *Let (X, d) be a separable metric space endowed with its Borel sigma algebra $\mathscr{F}$ and a finite Borel measure μ which gives positive measure to balls.*
Let $T : X \to X$ be a measure preserving transformation.
Then for almost every $x \in X$ there exists an increasing sequence $(n_k)_{k \geq 0}$ such that $T^{n_k}(x) \to x$ as k goes to infinity.

Proof. Let D be a dense countable subset of X. Let $\mathscr{B} = \{B(x,r);\ x \in D,\ r \in \mathbb{Q}_+^*\}$.
For each $B \in \mathscr{B}$ there exists $N_B \subset B$ with measure zero such that every point of $B \backslash N_B$ comes back infinitely often to B. Then, $N := \bigcup_{B \in \mathscr{B}} N_B$ has measure zero. For $x \in X \backslash N$, let $(B_k)_k$ be a sequence of balls centred in x with rational radii converging to zero. Since x has to come back infinitely often in any B_k, one can construct a sequence $(n_k)_k$ such that for all k, $T^{n_k}x \in B_k$. Consequently, $T^{n_k}(x) \to x$. □

Back to the physical problem that motivated Poincaré, we can look at systems for which the total energy is preserved as time goes by. It is the case for a system of planets interacting via the law of gravitation and for the movement of gas particules in a box, among others. For a system of N bodies, we work in the phase space $X = \mathbb{R}^{6N}$ of all possible states of the sytem: If $y \in X$, $y = \begin{pmatrix} \mathbf{x} \\ \mathbf{v} \end{pmatrix}$, where $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_N \end{pmatrix}$, with $\mathbf{x}_n$ denoting the position vector of the n^{th} body. Likewise, $\mathbf{v} \in \mathbb{R}^{3N}$ is the velocity vector.
Hamiltonian systems theory shows that under reasonable regularity assumptions the flow

$$\begin{array}{rccl} \varphi_t : & X & \longrightarrow & X \\ & y & \longmapsto & \varphi_t(y) = \text{the value of y after t units of time} \end{array}$$

preserves the Lebesgue measure in the phase space X, that is

$$\text{For all } t \geq 0 \text{ and } A \subset \mathbb{R}^{6N} \text{ measurable, we have } \lambda_{6N}(\varphi_t^{-1}A) = \lambda_{6N}(A),$$

where λ_{6N} is the $6N$-dimensional Lebesgue measure.
Then, discretizing the system setting $T = \varphi_1$ provides a Lebesgue preserving transformation in the phase space. We thus can apply Poincaré recurrence theorem to the system (X, T, λ_{6N}) to get the following funny fact:

If you put a glass filled with gas in an empty room, then after some time all the gas particules will come back in the glass.

More details about Hamiltonian mechanics can be found in [2].

Poincaré's theorem also implies that under the rotation of angle α, Lebesgue-almost every point of the circle is going to pass near its initial state. Note that there are two different cases:
- If α is rational, then the orbit of every point is **periodic** (that is, for all $x \in [0,1]$ there exists $p \in \mathbb{N}^*$ such that $T^p x = x$). This implies that the rotation acts in an "isolated" way: It is impossible for a point to explore all the circle.
- If α is irrational, the sytem is more complicated. One can show that the orbit of every point is dense in $[0,1]$, it is a consequence of the fact that the rotation of an irrational angle is an ergodic tranformation. This leads us to the next section.

3 Ergodicity

3.1 Definition and properties

Let $\mathscr{S} = (X, \mathscr{F}, \mu, T)$ be a measure preserving system. We say that $\mathscr{S}$ is **ergodic** if it has no proper subsystems: For all $A \in \mathscr{F}$, $T^{-1}A = A \ \Rightarrow\ \mu(A) = 0$ or $\mu(X \backslash A) = 0$.

Remark 3.1. *Depending on the context, we will either say that T or μ is ergodic.*

In the case where μ is finite, ergodicity has many equivalent definitions (we denote by $A\Delta B$ the symetric difference of A and B):

Proposition 3.1. *Let $(X, \mathscr{F}, \mu, T)$ be a measure preserving system with finite measure. The following are equivalent:*
(i) T is ergodic.
(ii) For all $A \in \mathscr{F}$, $\mu(T^{-1}A\Delta A) = 0 \;\Rightarrow\; \mu(A) = 0$ or $\mu(A) = \mu(X)$.
(iii) For all $A \in \mathscr{F}$ with $\mu(A) > 0$, $\mu\left(\bigcup_{n\geq 1} T^{-n}A\right) = \mu(X)$.
(iv) For all $A, B \in \mathscr{F}$ with $\mu(A), \mu(B) > 0$, there exists $n \geq 1$ such that $\mu(A \cap T^{-n}B) > 0$.

We are going to use the following elementary results.

Lemma 3.1.
- *For $A, B, C \in \mathscr{F}$ we have $\mu(A\Delta C) \leq \mu(A\Delta B) + \mu(B\Delta C)$.*
- *For $A, B \in \mathscr{F}$ we have $\mu(A\Delta B) = 0 \Rightarrow \mu(A) = \mu(B)$.*
- *For $A, B \in \mathscr{F}$ and $n \in \mathbb{N}$ we have $T^{-n}(A\Delta B) = (T^{-n}A)\Delta(T^{-n}B)$.*

Proof of Lemma 3.1.
- It is easier to get the first point thanks to a drawing.
- Since for $A, B \in \mathscr{F}$ we have

$$\mu(A) = \mu\big(A\backslash(A\cap B)\big) + \mu(A\cap B) \text{ and } \mu(B) = \mu\big(B\backslash(A\cap B)\big) + \mu(A\cap B),$$

we get

$$\mu(A) - \mu(B) = \mu\big(A\backslash(A\cap B)\big) - \mu\big(B\backslash(A\cap B)\big) \leq \mu(A\Delta B),$$

hence if $\mu(A\Delta B) = 0$, we have $\mu(A) \leq \mu(B)$. In the same way, we get $\mu(B) \leq \mu(A)$.
- Comes from the fact that the inverse image is compatible with intersections and unions. $\square$

Proof of proposition 3.1.
$(i) \Rightarrow (ii)$:
Let $A \in \mathscr{F}$ with $\mu(T^{-1}A\Delta A) = 0$ (we say that A is almost invariant). The idea is to construct a set B wich is invariant and satifies $\mu(A\Delta B) = 0$. Ergodicity and Lemma 3.1 then ensure that $\mu(A) = \mu(B) \in \{0, \mu(X)\}$.
Let $B = \limsup_n T^{-n}A = \bigcap_{n\geq 0}\bigcup_{k\geq n} T^{-k}A$.
Then B is invariant, and satifies $\mu(A\Delta B) = 0$:

$$\begin{aligned}
\mu(A\Delta B) &= \mu\left(\left(\bigcap_{n\geq 0}\bigcup_{k\geq n} T^{-k}A\right)\cap A^c\right) + \mu\left(\left(\bigcup_{n\geq 0}\bigcap_{k\geq n} T^{-k}A^c\right)\cap A\right)\\
&\leq \mu\left(\left(\bigcup_{n\geq 0} T^{-n}A\right)\cap A^c\right) + \mu\left(\left(\bigcup_{n\geq 0} T^{-n}A^c\right)\cap A\right)\\
&\leq \sum_{n\geq 0}\mu(T^{-n}A\Delta A).
\end{aligned}$$

The first point of Lemma 3.1 gives for $n \in \mathbb{N}$:

$$\mu(T^{-n}A\Delta A) \leq \mu(T^{-n}A\Delta T^{-(n-1)}A) + \cdots + \mu(T^{-1}A\Delta A).$$

The third point of Lemma 3.1 and the invariance of μ then give $\mu(T^{-n}A\Delta A) = 0$, we conclude that $\mu(A\Delta B) = 0$.

$\underline{(ii) \Rightarrow (iii)}$:

Let $A \in \mathscr{F}$ such that $\mu(A) > 0$. Let $B = \bigcup_{n\geq 1} T^{-n}A$. Then $T^{-1}B \subset B$ and

$$\mu(T^{-1}B\Delta B) = \mu(B\backslash T^{-1}B) = \mu(B) - \mu(T^{-1}B) = 0,$$

And since $\mu(B) \geq \mu(A) > 0$ we have $\mu(B) = \mu(X)$.

$\underline{(iii) \Rightarrow (iv)}$:

Let $A, B \in \mathscr{F}$ such that $\mu(A), \mu(B) > 0$. Then,

$$\mu(A) = \mu\Big(A \cap \bigcup_{n\geq 1} T^{-n}B\Big) = \mu\Big(\bigcup_{n\geq 1} T^{-n}B \cap A\Big).$$

Consequently, there exists $n \geq 1$ such that $\mu(T^{-n}B \cap A) > 0$.

$\underline{(iv) \Rightarrow (i)}$:

Let $A \in \mathscr{F}$ such that $T^{-1}A = A$. Assume that $\mu(A) > 0$.

If $\mu(A^c) > 0$, then there exists $n \geq 1$ such that $0 < \mu(A \cap T^{-n}A^c) = \mu(A \cap A^c)$, contradiction.

Hence $\mu(A) = \mu(X)$. □

Proposition 3.2. *Under the conditions of Proposition 3.1, T is ergodic if and only if for every $f : X \to \mathbb{R}$ measurable, $f \circ T = f$ almost everywhere implies that f is constant almost everywhere.*

Proof.

• Assume that T is ergodic and let $f : X \to \mathbb{R}$ measurable such that $f \circ T = f$ almost everywhere. For $k \in \mathbb{Z}$ and $n \in \mathbb{N}^*$, set

$$A_n^k = \left\{\frac{k}{n} \leq f < \frac{k+1}{n}\right\}.$$

Then $T^{-1}A_n^k\Delta A_n^k = \left\{\frac{k}{n} \leq f\circ T < \frac{k+1}{n}\right\} \Delta \left\{\frac{k}{n} \leq f < \frac{k+1}{n}\right\}$ is contained in $\{f \circ T \neq f\}$, hence $\mu(T^{-1}A_n^k\Delta A_n^k) = 0$.

By ergodicity, each A_n^k has measure 0 or $\mu(X)$.

Since for all $n \in \mathbb{N}^*$ we have $X = \bigsqcup_{k\in\mathbb{Z}} A_n^k$, there exists a unique $k(n)$ such that $\mu(A_n^k) = \mu(X)$.

By construction, $\left(\frac{k(n)}{n}\right)_n$ is a non decreasing and bounded sequence, hence $\lim_n \frac{k(n)}{n}$ exists. Let $Y = \bigcap_{n\in\mathbb{N}^*} A_n^{k(n)}$. Then $\mu(Y) = \mu(X)$ and for $x \in Y$, $f(x) = \lim_n \frac{k(n)}{n}$, hence f is constant almost everywhere.

• Suppose now that $f \circ T = f$ almost everywhere implies f is constant almost everywhere, and let $A \in \mathscr{F}$ such that $T^{-1}A = A$. Then $\mathbf{1}_A$ is measurable and satifies $\mathbf{1}_A \circ T = \mathbf{1}_{T^{-1}A} = \mathbf{1}_A$. By hypothesis, $\mathbf{1}_A$ is constant to 0 or 1 almost everywhere, which means that $\mu(A) = 0$ or $\mu(A) = \mu(X)$. □

Remark 3.2. *To prove that T is ergodic, it suffices to prove that for **one** $p \in [1, +\infty]$, for all $f \in L^p(\mu)$, $f \circ T = f$ implies f is constant (Since it suffices to prove it only for $\mathbf{1}_A$, for all $A \in \mathscr{F}$).*

Example 3.1. *Let $T : [0, 1] \to [0, 1]$ be the rotation of angle α, where α is irrational. Let $f \in L^2(\lambda)$ such that $f \circ T = f$.*

For almost every $x \in [0, 1]$ we have $f(x) = \sum_{n\in\mathbb{Z}} c_n e^{2i\pi nx}$ so that for almost every x,

$$\sum_{n\in\mathbb{Z}} c_n e^{2i\pi n(x+\alpha)} = f(Tx) = f(x) = \sum_{n\in\mathbb{Z}} c_n e^{2i\pi nx},$$

whence for almost every $x \in [0,1]$

$$\sum_{n\in\mathbb{Z}} c_n(1-e^{2in\pi\alpha})e^{2i\pi nx} = 0.$$

The uniqueness of the Fourier coefficients and the fact that $1-e^{2in\pi\alpha} \neq 0$ for $n \neq 0$ then gives: for all $n \neq 0$, $c_n = 0$, and $f = c_0$ almost everywhere.

3.2 Ergodic Theorems

In 1871 Boltzmann introduced his *Ergodic Hypothesis*, which stated that at an equilibirum state of a physical dynamical system, for every integrable function the mean time value and the mean space value are the same, that is, for all $f : X \to \mathbb{R}$ integrable,

$$\lim_n \frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k = \frac{1}{\mu(X)}\int_X f d\mu.$$

This is false in the general case but it is true to a certain extent (depending on what type of convergence we are looking at) in the ergodic case.
More precisely,

Theorem 3.1. *(Von Neumann's ergodic theorem)*
Let $(X, \mathscr{F}, \mu, T)$ be a measure preserving system, with possibly infinite measure. Then, for all $f \in L^2(\mu)$, there exists $\tilde{f}$ in $L^2(\mu)$ such that $\tilde{f}$ is invariant ($\tilde{f}\circ T = \tilde{f}$) and $\frac{1}{n}S_n f \to \tilde{f}$ in L^2 as n goes to infinity, where $S_n f := \sum_{k=0}^{n-1} f\circ T^k$.
If T is ergodic and μ is finite, then $\tilde{f} = \frac{1}{\mu(X)}\int_X f d\mu$ in L^2.

We might want the convergence to be pointwise, which will be a powerful tool to establish "almost everywhere" results. The following theorem will help.

Theorem 3.2. *(Birkhoff's ergodic theorem)*
Let $(X, \mathscr{F}, \mu, T)$ be a measure preserving system with finite measure. Let f be in $L^1(\mu)$. Then there exists an invariant function $\tilde{f}$ in $L^1(\mu)$ such that $\frac{1}{n}S_n f = \frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k \to \tilde{f}$ almost everywhere and in L^1.
If T is ergodic, we have $\tilde{f} = \frac{1}{\mu(X)}\int_X f d\mu$.

3.2.1 Proof of Von Neumann's theorem

Von Neumann's theorem uses a general result in Hilbert spaces:

Proposition 3.3. *Let $\mathscr{H}$ be a Hilbert space and $U \in L(\mathscr{H})$ a continuous operator with $||U|| \leq 1$. Then for all $x \in \mathscr{H}$, $\frac{1}{n}\sum_{k=0}^{n-1} U^k(x) \to \pi(x)$, where $\pi : \mathscr{H} \to \ker(I-U)$ is the orthogonal projection.*

Remark 3.3. *In other words, the Cesàro sums $\frac{1}{n}\sum_{k=0}^{n-1} U^k(x)$ tend to be invariant under the composition with U.*

Proof of proposition 3.3.
The proof consists in splitting $\mathscr{H}$ into spaces where the theorem is obvious: We are going to show that $\mathscr{H} = \ker(I-U) \oplus \overline{\mathrm{Im}(I-U)}$.

Let's begin by noticing that if x is in $\ker(I-U)$, then for all n, $\frac{1}{n}\sum_{k=0}^{n-1} U^k(x) = x$, and if $x = y - Uy$ is in $\mathrm{Im}(I-U)$ then $\frac{1}{n}\sum_{k=0}^{n-1} U^k(x) = \dfrac{1}{n}(y - U^n y)$, and since $||y - U^n y|| \leq 2\,||y||$, $\frac{1}{n}\sum_{k=0}^{n-1} U^k(x)$ tends to zero. Notice that if x is only a limit point of $\mathrm{Im}(I-U)$ the result still works:
Let x be in $\overline{\mathrm{Im}(I-U)}$, and let (x_n) be a sequence of $\mathrm{Im}(I-U)$ converging to x. Let $\epsilon > 0$. There exists m such that $||x - x_m|| < \frac{\epsilon}{2}$, and there exists $N > 0$ such that for all $n \geq N$ $\left|\left|\frac{1}{n}\sum_{k=0}^{n-1} U^k(x_m)\right|\right| < \frac{\epsilon}{2}$.
Then, for $n \geq N$,

$$\left|\left|\frac{1}{n}\sum_{k=0}^{n-1} U^k(x)\right|\right| \leq \frac{1}{n}\sum_{k=0}^{n-1} \left|\left|U^k(x - x_m)\right|\right| + \left|\left|\frac{1}{n}\sum_{k=0}^{n-1} U^k(x_m)\right|\right|,$$

and since $||U|| \leq 1$ we get

$$\left|\left|\frac{1}{n}\sum_{k=0}^{n-1} U^k(x)\right|\right| \leq \frac{1}{n}\sum_{k=0}^{n-1} \frac{\epsilon}{2} + \left|\left|\frac{1}{n}\sum_{k=0}^{n-1} U^k(x_m)\right|\right|,$$

hence $\left|\left|\frac{1}{n}\sum_{k=0}^{n-1} U^k(x)\right|\right| < \epsilon$,
And we proved that

$$\forall \epsilon > 0,\ \exists N \in \mathbb{N}^*,\ \forall\ n \geq N,\ \left|\left|\frac{1}{n}\sum_{k=0}^{n-1} U^k(x)\right|\right| < \epsilon.$$

We now show that the cases we looked at are essentially the unique ones to consider, that is:

$$\mathscr{H} = \ker(I-U) \oplus \overline{\mathrm{Im}(I-U)}.$$

<u>We prove that $\mathrm{Im}(I-U)^\perp \subset \ker(I-U)$.</u>
Let $x \in \mathrm{Im}(I-U)^\perp$. Then,

$$||(I-U)x||^2 = ||x||^2 + ||Ux||^2 - 2\langle x, Ux\rangle \leq 2(||x||^2 - \langle x, Ux\rangle)$$

and since $\langle x, Ux\rangle = \langle x, Ux - x\rangle + ||x||^2 = ||x||^2$, we get $||(I-U)x|| = 0$.

<u>Let's prove the other inclusion.</u>
We first prove that $\ker(I-U) = \ker(I-U^*)$.
If $x \in \ker(I-U^*)$,

$$||(I-U)x||^2 = ||x||^2 + ||Ux||^2 - 2\langle x, Ux\rangle \leq 2||x||^2 - 2\langle x, Ux\rangle$$

and $\langle x, Ux\rangle = \langle x, (U-I)x\rangle + ||x||^2 = \langle (U-I)^* x, x\rangle + ||x||^2 = ||x||^2$, hence $||(I-U)x||^2 = 0$. We thus have $\ker(I-U^*) \subset \ker(I-U)$.
Applying the same argument to U^* gives the converse inclusion.

Back to the proof of the inclusion $\text{Im}(I-U)^{\perp} \supset \ker(I-U)$.
Let $x \in \ker(I-U)$ and $y \in \mathscr{H}$, we have

$$\langle x, (I-U)y\rangle = \langle (I-U)^* x, y\rangle = 0,$$

since $\ker(I-U^*) = \ker(I-U)$.
Finally

$$\mathscr{H} = \ker(I-U) \oplus \ker(I-U)^{\perp} = \ker(I-U) \oplus (\text{Im}(I-U)^{\perp})^{\perp},$$

Which is the desired decomposition of $\mathscr{H}$ and the theorem is proved. $\square$

To establish Von Neumann's ergodic theorem, let $U : L^2 \to L^2$ be the Koopman operator, defined by $U(f) = f \circ T$. U is a well defined unitary operator, since we have

$$\int_X |f \circ T|^2 d\mu = \int_X |f|^2 d\mu_T = \int_X |f|^2 d\mu,$$

because T is measure preverving.
Applying Proposition 3.3 to U with $\mathscr{H} = L^2$ gives the first part of the theorem.
For the second part, since for f in L^2 we have $\pi(f) \circ T = \pi(f)$, ergodicity of the system provides that $\pi(f)$ is constant in L^2.
Integrating $\pi(f)$ gives:

$$\int_X \pi(f) d\mu = \mu(X)\pi(f).$$

Now using the fact that for all $n \geq 1$ we have

$$\int_X \frac{1}{n} S_n f d\mu = \int_X f d\mu,$$

we get by the Dominated Convergence Theorem (see remark 3.4)

$$\int_X f d\mu = \int_X \pi(f) d\mu.$$

Finally, we have (in L^2)

$$\pi(f) = \frac{1}{\mu(X)} \int_X f d\mu.$$

Remark 3.4. *It is not straightforward. Note that if f is bounded the result holds, and if not, apply the previous result with the bounded functions $\min(f, L)$ for $L \geq 0$. A dominated convergence allows us to conclude.*

3.2.2 Proof of Birkhoff's ergodic theorem

We will need the following result.

Lemma 3.2. *Let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be non negative sequences. Suppose that there exists an integer $M > 0$ such that for all $n \in \mathbb{N}$ there exists $1 \leq m \leq M$ verifying*

$$\sum_{k=n}^{m+n-1} a_k \geq \sum_{k=n}^{m+n-1} b_k.$$

Then, for all $N > M$, we have

$$\sum_{k=0}^{N-1} a_k \geq \sum_{k=0}^{N-M-1} b_k.$$

Proof of lemma 3.2.
Let $N > M$. By hypothesis, one can construct an increasing finite sequence of integers $(m_i)_{0\leq i\leq n}$ such that one has $m_n \leq N$, $M + m_n \geq N$, $m_0 = 0$ and for all $0 \leq i \leq n$

$$\sum_{k=m_i}^{m_{i+1}-1} a_k \geq \sum_{k=m_i}^{m_{i+1}-1} b_k.$$

Adding all these inequalities and using the fact that the a_k and b_k are non negative lead to

$$\sum_{k=0}^{N-1} a_k \geq \sum_{k=0}^{m_n-1} a_k \geq \sum_{k=0}^{m_n-1} b_k \geq \sum_{k=0}^{N-M-1} b_k.$$

□

Proof of theorem 3.1.
We prove the theorem for non negative functions (we then get the general result by writing for any measurable function $f = f^+ - f^-$).
Let $f : X \to \mathbb{R}_+$ measurable. For $x \in X$ define $\underline{f}(x) = \liminf_n S_n f(x)$ and $\overline{f}(x) = \limsup_n S_n f(x)$. $\underline{f}$ and $\overline{f}$ are T-invariant. We are going to show that those two functions coincide almost everywhere and that we have $\int_X f d\mu = \int_X \underline{f} d\mu = \int_X \overline{f} d\mu$. To prove it we just have to show that

$$\int_X \underline{f} d\mu \geq \int_X f d\mu \geq \int_X \overline{f} d\mu,$$

since we know that $\underline{f} \leq \overline{f}$.
<u>Let us prove that $\int_X f d\mu \geq \int_X \overline{f} d\mu$.</u>

Let $L > 0$ and $0 < \epsilon < 1$. By definition of the upper limit, for all x there exists $m \geq 1$ such that
$$\frac{S_m f(x)}{m} \geq \overline{f}(x)(1-\epsilon) \geq \min(\overline{f}(x), L)(1-\epsilon).$$
By an increasing continuity argument we get that for all $\delta > 0$ there exists $M > 0$ such that the set

$$X_0 := \left\{ x \in X | \ \exists\ 1 \leq m \leq M, \ \frac{S_m f(x)}{m} \geq \min(\overline{f}(x), L)(1-\epsilon) \right\}$$

has measure greater than $\mu(X) - \delta$. Define $F : X \to \mathbb{R}_+$ by $F(x) = f(x)$ if $x \in X_0$ and $F(x) = L$ if $x \in (X_0)^c$. Then, $f \leq F$: it is clear on X_0, and if $x \in (X_0)^c$, then $f(x) = S_1 f(x) \leq \min(\overline{f}(x), L)(1-\epsilon) \leq L$.
Now, for $x \in X$, let $a_n = F(T^n x)$ and $b_n = \min(\overline{f}(x), L)(1-\epsilon)$ (independant from n). (a_n) and (b_n) satisfy the conditions of Lemma 3.2
Let $n \geq 0$.
• If $T^n x \in X_0$, there exists $1 \leq m \leq M$ such that

$$\begin{aligned} F(T^n x) + \cdots + F(T^{n+m-1}x) &\geq f(T^n x) + \cdots + f(T^{n+m-1}x) \\ &\geq m \min(\overline{f}(x), L)(1-\epsilon) \\ &= b_n + \cdots + b_{n+m-1}. \end{aligned}$$

• If $T^n x \notin X_0$, take $m = 1$ since

$$a_n = L \geq \min(\overline{f}(x), L)(1-\epsilon) = b_n.$$

By lemma 3.2, for all $N > M$,

$$\sum_{k=0}^{N-1} f(T^k x) \geq (N-M)\min(\overline{f}(x), L)(1-\epsilon).$$

Since M and N are independant from $x \in X_0$, integrating both parts of this inequality and using the invariance of μ gives, for all $N > M$,

$$N \int_X F d\mu \geq (N-M) \int_X \min(\overline{f}(x), L)(1-\epsilon) d\mu.$$

We thus have

$$\begin{aligned}
\int_X f d\mu &\geq \int_{X_0} f d\mu \\
&= \int_X F d\mu - L\mu(X \backslash X_0) \\
&\geq \frac{N-M}{N} \int_X \min(\overline{f}(x), L)(1-\epsilon) d\mu - L\delta.
\end{aligned}$$

Now, letting N tend to infinity and δ to zero we have

$$\int_X f d\mu \geq \int_X \min(\overline{f}(x), L)(1-\epsilon) d\mu.$$

We now let ϵ tend to zero and L to infinity. The Monotone Convergence Theorem allows us to write

$$\int_X f d\mu \geq \int_X \overline{f} d\mu.$$

Thanks to the same kind of arguments we show that

$$\int_X \underline{f} d\mu \geq \int_X f d\mu,$$

and the result is proved. □

3.3 Remark about the Ergodic hypothesis

Ergodic theory provides a generalisation of the notion of i.i.d (independent and identically distributed) processes. Indeed, Birkhoff's theorem is a generalisation of the strong law of large numbers:

Theorem 3.3. *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, and let $(X_i)_i$ be an i.i.d sequence of real valued random variables, with $X_1 \in L^1(\mathbb{P})$.*
Then the means $\frac{1}{n}\sum_{k=0}^{n-1} X_k$ converge almost surely to $\mathbb{E}[X_1]$.

In Birkhoff's Theorem, the X_i are defined as $X_i = f \circ (T^i)$. The X_i are identically distributed: for all $A \in \mathscr{F}$,

$$\mu(X_{i+1} \in A) = \int_X \mathbf{1}_{(f\circ T^{i+1})^{-1}A} d\mu = \int_X \mathbf{1}_{(f\circ T^{i})^{-1}A} d\mu_T = \int_X \mathbf{1}_{(f\circ T^{i})^{-1}A} d\mu = \mu(X_i \in A)$$

by invariance of μ; but the process $(X_i)_i$ may not be independent.

In section 9 we will see that for "chaotic enough" dynamical systems, some other properties of random dynamical systems persist.

These results help answering the general question: To what extent can a deterministic dynamical system behave like a random system as time goes by?

3.4 Mixing systems

We define here the notion of mixing dynamical systems, which is a particular case of ergodic systems. Heuristically, it introduces the idea of "asymptotic independence", that we will be looking for in order to compare mixing enough dynamical systems to random systems.

Definition 3.1. *Let $(X, \mathscr{F}, \mu, T)$ be a measure preserving system with finite measure. We say that T (or μ, depending on what we are looking at) is* ***mixing*** *if we have for all $A, B \in \mathscr{F}$*

$$\mu(A \cap T^{-n}B) \to \frac{\mu(A)\mu(B)}{\mu(X)}.$$

Proposition 3.4. *If $(X, \mathscr{F}, \mu, T)$ is a mixing system, then it is ergodic.*

Proof. Suppose T is mixing. Then for $A, B \in \mathscr{F}$,

$$\mu(A \cap T^{-n}B) - \frac{\mu(A)\mu(B)}{\mu(X)} \to 0.$$

Let A be an invariant set. Then

$$\mu(A \cap T^{-n}A) - \frac{\mu(A)\mu(T^{-n}A)}{\mu(X)} \to 0$$

and by invariance we get

$$\frac{\mu(A)}{\mu(X)} = \left(\frac{\mu(A)}{\mu(X)}\right)^2,$$

which implies that $\mu(A) = 0$ or $\mu(X)$. □

We now introduce some notions we will use in the rest of the document.

4 Expanding maps, Markov partitions and Gibbs Measures

In this section, $X = [0, 1]$ and $\mathscr{B}$ is the Borel sigma algebra on X.

Let $\alpha \in (0, 1]$.

Definition 4.1. *($\mathscr{C}^{1+\alpha}$ maps, epanding maps)*
$T : X \to X$ is said to be a ***piecewise $\mathscr{C}^{1+\alpha}$ map*** *if T is a piecewise (with finite subdivision) $\mathscr{C}^1$ map with, in every piece where T' is defined, T' is α-Hölder (that is, there exists a constant c such that for all x and y, $d(Tx, Ty) \leq cd(x, y)^\alpha$).*
For $\beta > 1$, we say that T is β-expanding if (everywhere where T' is defined) $|T'| \geq \beta$.

Definition 4.2. *(Markov partition)*
Let $T : X \to X$ be a piecewise $\mathscr{C}^{1+\alpha}$ map.
Let $\mathscr{J} = \{J_n;\ 0 \leq n \leq N\}$ be a finite family of closed intervals where every $J_n \subset X$. We say that $\mathscr{J}$ is a Markov partition according to T if we have the following properties:
- *For all i, T is a $\mathscr{C}^{1+\alpha}$ diffeomorphism from $int(J_i)$ onto its image*
- *$X = \bigcup_{i=0}^{N} J_i$ and if $i \neq j$ we have $int(J_i) \cap int(J_j) = \emptyset$*
- *If $T(int(J_i)) \cap int(J_j) \neq \emptyset$, we have $J_j \subset T(J_i)$.*

Remark 4.1. *One can see a Markov partition as a minimal decomposition of the space X which is compatible with the regularity of T.*

Example 4.1. *Let T be the doubling map. Then the family $\mathscr{J} = \left\{ \left[0, \frac{1}{2}\right], \left[\frac{1}{2}, 1\right] \right\}$ is a natural Markov partition.*

If $\mathscr{J}$ is a Markov partition according to T, we note $\partial\mathscr{J} := \bigcup_{J \in \mathscr{J}} \partial J$ and we define S as

$$S := \bigcup_{n \in \mathbb{N}} T^{-n}(\partial\mathscr{J}).$$

Note that the set where T' is not defined is contained is S. For x not in S, define:

- $I_n(x)$ is the element of the partition containing $T^n x$.
- For $n \geq 1$, $\mathscr{J}_n(x) = \{y \in X | \text{ for all } 0 \leq k \leq n-1,\ T^n y \in I_n(x)\}$. $\mathscr{J}_n(x)$ is called the n-**cylinder** of x.

We can now define the class of Gibbs measures. Let $T : X \to X$ be a piecewise $\mathscr{C}^{1+\alpha}$ map. Assume that there exists a Markov partition $\mathscr{J}$ compatible with T.

Recall that $S_n\varphi(x) = \sum_{k=0}^{n-1} \varphi \circ T^k(x)$.

Definition 4.3. *Let $\varphi : X \backslash S \to \mathbb{R}$ be a Hölder function and μ be a T-invariant probability meaure. We say that μ is a* ***Gibbs measure*** *for the potential φ if $\mu(S) = 0$ and if there exists a* ***pressure*** *$P(\varphi) \in \mathbb{R}$ such that for some $\kappa \geq 1$, for all $n \geq 0$ and $x \in X \backslash S$ the following holds:*

$$\frac{1}{\kappa} \leq \frac{\mu(\mathscr{J}_n(x))}{e^{S_n\varphi(x) - nP(\varphi)}} \leq \kappa.$$

Example 4.2. *Section 9.1 will provide a family of examples. For instance, if T is the doubling map and $\mathscr{J}$ is the Markov partition of example 4.1, then Lebesgue measure is a Gibbs measure for $\varphi = -\log T'$ under the potential $P(\varphi) = 0$.*

Definition 4.4. *We say that $T : X \to X$ is topologically mixing if*

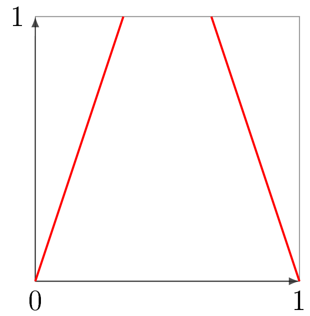
For all A, B open sets of X, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $A \cap T^{-n}B \neq \emptyset$.

The next section motivates the introduction of fractal dimension for the study of dynamical systems.

5 How dimension theory appears in dynamical systems

The following example shows how dynamical systems and Fractal theory can be related.
Let $X = [0, 1]$, $H = \left(\frac{1}{3}, \frac{2}{3}\right)$ and

$$\begin{array}{rccl} T: & X \backslash H & \longrightarrow & X \\ & x & \longmapsto & \begin{cases} 3x \text{ if } x \in \left[0, \frac{1}{3}\right] \\ 3(1-x) \text{ if } x \in \left[\frac{2}{3}, 1\right]. \end{cases} \end{array}$$



For $x \in X \backslash H$, the orbit of x is defined by the set $\left\{T^n x;\ n \text{ such that for all } 1 \leq k < n,\ T^k x \in X \backslash H\right\}$.
We say that $(X, \mathscr{F}_{X \backslash H}, \lambda, T)$ is a dynamical system with a hole (H is called the hole).
We can now look at the set A of points which never fall into the hole.
One can see that A is the Cantor set, defined as follows:
For $I = [a, b]$ define $f(I) := I \backslash \left(a + \frac{b-a}{3}, a + \frac{2(b-a)}{3}\right)$ as the set I from which we removed the middle third.
Let $C_0 = X$.
If $C_n = \bigcup_j I_j^n$ is constructed, where each I_j^n is a closed interval, define $C_{n+1} = \bigcup_j f(I_j^n)$.
Finally, define the Cantor set as the decreasing intersection $C = \bigcap_n C_n$.

Figure 1: The first iterations in the construction of the Cantor set

This is an example of fractal set which naturally appears when studying some dynamical systems. In this case, we might be interested by finding an appropriate measure - an invariant ergodic measure that would lie in the fractal set, called the attractor.
The attractor can be seen as the "limit set" in which lies the dynamical system. Let us introduce another example. Let

$$\begin{array}{rccc} T: & X & \longrightarrow & X \\ & x & \longmapsto & 2x(1-x) \end{array}$$

be the **logistic map of parameter** $\frac{1}{2}$. Note that T has two fixed points: 0 and $\frac{1}{2}$. One can show that Lebesgue-almost every point is attracted by $\frac{1}{2}$. Hence the "limit set" (according to Lebesgue measure) is reduced to a point; to study this dynamical system the Dirac measure $\delta_{\frac{1}{2}}$ is therefore an appropriate measure. Note that the attractor is often X, for example for the doubling map, wich is already ergodic according to Lebesgue measure.
In order to restrict our attention on the attractor, we thus search for an appropriate measure - an invariant measure that sees the attractor (that is, a measure with small enough dimension).

As a consequence, the fractal dimension appears in the study of dynamical systems. We define it in Appendix A.
Hausdorff dimension is a useful tool if we want to study sets such as the Cantor set, for which Lebesgue measure is zero. In the same way, the notions of Hausdorff dimension and pointwise dimension of a measure will help.
Let μ be a measure on a space $(X, \mathscr{F})$, where $X \subset \mathbb{R}^n$. We define the **Hausdorff dimension** of μ as

$$\dim_H(\mu) = \inf \left\{\dim_H(A);\ A \text{ such that } \mu(A) > 0\right\},$$

that is, $\dim_H(\mu)$ gives the size of the smallest sets that have positive measure.

Remark 5.1. *A Dirac measure has dimension zero, the d-dimensional Lebesgue measure has dimension d.*

The **lower and upper pointwise dimensions** are defined by $\underline{d}_\mu(x) = \liminf_{r\to 0} \dfrac{\log \mu(B(x,r))}{\log r}$ and $\overline{d}_\mu(x) = \limsup_{r\to 0} \dfrac{\log \mu(B(x,r))}{\log r}$.

When the limit exists, we define the **pointwise dimention** as $d_\mu(x) = \lim_{r\to 0} \dfrac{\log \mu(B(x,r))}{\log r}$.

Intuitively, the local dimension gives the power of r one has to calculate to have a rough idea of the measure of a small ball of radius r.

As an illustration of what ergodic theory provides, the next section shows that the local dimension of an ergodic measure is almost everywhere equal to its Hausdorff dimension, by introducing two quantities: Entropy and Lyapunov exponant.

6 Entropy, Lyapunov exponent and dimension

Let $(X, \mathscr{F}, \mu, T)$ be a probability measure preserving system.

6.1 Entropy

Let ξ be a finite partition of X such that every element of ξ is measurable and has positive measure. For $n \geq 1$ and $x \in X$ denote by $\xi_n(x)$ the n**-cylinder centred in** x:

$$\xi_n(x) = \left\{ y \in X |\ y \in X_0, \ldots, T^{n-1}y \in X_{n-1} \right\},$$

where X_i is the element of ξ containing x.
For $n \geq 1$ we define

$$H_n(\xi) = -\sum_{A \in \Xi_n} \log\left(\mu(A)\right)\mu(A),$$

where Ξ_n is the set of all n cylinders. The sequence $(H_n(\xi))$ is subadditive, that is:

$$\text{For all } m, n \geq 1, \text{ we have } H_{n+m}(\xi) \leq H_n(\xi) + H_m(\xi).$$

Indeed note that if α and β are finite partitions of X, putting $\alpha \vee \beta := \{A \cap B,\ A \in \alpha,\ B \in \beta\}$ we have $H(\alpha \vee \beta) \leq H(\alpha) + H(\beta)$, where $H(\gamma)$ is defined for a measurable partition γ as

$$H(\gamma) = -\sum_{A \in \gamma} \log\left(\mu(A)\right)\mu(A).$$

This can be shown in the following way:

$$\begin{aligned}
H(\alpha \vee \beta) &= -\sum_{A\in\alpha}\sum_{B\in\beta} \log(\mu(A\cap B))\mu(A\cap B) \\
&= -\sum_{A\in\alpha}\sum_{B\in\beta} \log\left(\frac{\mu(A\cap B)}{\mu(A)}\right)\mu(A\cap B) - \sum_{A\in\alpha}\sum_{B\in\beta}\log(\mu(A))\mu(A\cap B) \\
&= -\sum_{A\in\alpha}\sum_{B\in\beta} \log\left(\frac{\mu(A\cap B)}{\mu(A)}\right)\mu(A\cap B) + H(\alpha).
\end{aligned}$$

Now note that $f : x \mapsto -x\log(x)$ is concave, hence by Jensen's inequality

$$\begin{aligned}
-\sum_{B\in\beta}\sum_{A\in\alpha} \log\left(\frac{\mu(A\cap B)}{\mu(A)}\right)\mu(A\cap B) &= \sum_{B\in\beta}\sum_{A\in\alpha} \mu(A) f\left(\frac{\mu(A\cap B)}{\mu(A)}\right) \\
&\leq \sum_{B\in\beta} f\left(\mu(A\cap B)\right) \\
&= H(\beta).
\end{aligned}$$

To conclude on the subadditivity of $H_n(\xi)$ remark that for $n \geq 1$ we have $\Xi_n = \bigvee_{i=0}^{n-1} T^{-i}\xi$, and use the fact that T is measure preserving to get

$$\sum_{A \in \Xi_m} \log(\mu(A))\,\mu(A) = \sum_{A \in \bigvee_{i=n}^{n+m-1} T^{-i}\xi} \log(\mu(A))\,\mu(A).$$

The entropy of T according to ξ is then defined as the limit

$$h_{\mu,\xi}(T) = \lim_n \frac{1}{n} H_n(\xi).$$

The limit exists by subadditivity.
We now define the **entropy** of T (according to μ) as

$$h_\mu(T) = \sup_{\xi \text{ partition of X}} h_{\mu,\xi}(T),$$

where the partitions are taken to be finite and measurable.
The following theorem is a consequence of the Birkhoff Ergodic Theorem. We state it without proof. For more details, see [3].

Theorem 6.1. *(Shannon-McMillan-Breiman)*
Let $(X, \mathscr{F}, \mu, T)$ be an ergodic dynamical system with $\mu(X) = 1$, and let ξ be a finite measurable partition of X. Then, for almost every $x \in X$, we have

$$\lim \frac{1}{n} \mu(\xi_n(x)) = h_{\mu,\xi}(T).$$

6.2 Lyapunov exponent

We define here the Lyapunov exponent in the case where $X = [0,1]$ and T is differentiable with $|T'| > 0$.
The Lyapunov exponent of T at x is defined as the limit, when it exists,

$$\lambda_T(x) = \lim_n \frac{1}{n} \log\left|(T^n)'\right|(x).$$

One can see $\lambda_T(x)$ as the mean expansion value of the iterates of T in x. The following theorem is an immediate consequence of Birkhoff's ergodic theorem.

Theorem 6.2. *Suppose that $T : [0,1] \to [0,1]$ is a $\mathscr{C}^1$ ergodic transformation with respect to μ, with T' measurable and $|T'| > \epsilon$, where $\epsilon > 0$. Then the Lyapunov exponent of T is defined almost everywhere and for almost every $x \in [0,1]$ we have*

$$\lambda_T(x) = \int_{[0,1]} \log\left|(T^n)'\right|(x) d\mu(x) =: \lambda_T(\mu)$$

Proof. By hypothesis $\log|T'|$ is bounded (and measurable). It is therefore in L^1.
Since for $n \geq 1$ we have for almost every x

$$\frac{1}{n} \log\left|(T^n)'\right|(x) = \frac{1}{n} \sum_{k=0}^{n-1} \log(T' \circ T^k(x)),$$

Theorem 3.1 allows us to conclude. □

6.3 Entropy, Lyapunov exponent and local dimension

The following theorem links the previous quantities to the local dimension in ergodic systems. A proof can be found in [8]

Theorem 6.3. *Let $T : [0,1] \to [0,1]$ be a $\mathscr{C}^{1+\alpha}$, expanding, μ-preserving transformation, where μ is a Gibbs measure. Then the local dimension d_μ is constant almost everywhere and for almost every $x \in [0,1]$ we have*

$$d_\mu(x) = \frac{h_\mu(T)}{\lambda_T(\mu)} = \dim_H(\mu).$$

7 Induced transformation, Kac's lemma

Let $(X, \mathscr{F}, \mu, T)$ be a measure preserving system with finite measure.
For A measurable with $\mu(A) > 0$, and for $x \in A$ define

$$\begin{array}{rccl} \tau_A : & A & \longrightarrow & \mathbb{N} \cup \{+\infty\} \\ & x & \longmapsto & \inf\{n \geq 1 | T^n x \in A\}. \end{array}$$

τ_A is measurable. Indeed, for $n \in \mathbb{N}$,

$$\{\tau_A \leq n\} = \bigcup_{k=1}^{n} T^{-k} A \in \mathscr{F} \text{ and } \{\tau_A = +\infty\} = \bigcap_{n \geq 1} T^{-n} A^c \in \mathscr{F}.$$

Thanks to the Poincaré Recurrence Theorem, we know that τ_A is finite almost everywhere. τ_A is called the **return time map**.

What Poincaré's Theorem doesn't tell is the time for a given point to come back to A. Kac's lemma goes in that way. In order to prove it we introduce the **induced dynamical system** on A:

Let $\mathscr{F}_{|A} := \{A \cap B;\ B \in \mathscr{F}\}$ be the trace sigma algebra on A.
Define μ_A on $\mathscr{F}_{|A}$ by $\mu_A(B) = \frac{\mu(A \cap B)}{\mu(A)}$, for all $B \in \mathscr{F}$.
Poincaré Recurrence Theorem insures the existence of a set $N \subset A$ such that $\mu(N) = 0$ and for all $x \in A \backslash N$, there exists $n \geq 1$ such that $T^n x \in A$.
Now, let $M = \{x \in A \backslash N |\ \exists n \geq 1\ T^n x \in N\}$. Since M is contained in $\bigcup\limits_{n \geq 1} T^{-n} N$, and

$$\mu(\bigcup_{n \geq 1} T^{-n} N) \leq \sum_{n \geq 1} \mu(T^{-n} N) = 0$$

by T-invariance of μ, we have $\mu(M) = 0$.
For all $x \in A \backslash M$, $\tau_A(x) < +\infty$.
Now, let $T_A : A \backslash M \to A$ be defined by $T_A(x) = T^{\tau_A(x)} x$.
T_A and all its iterates are well defined measurable functions according to $\mathscr{F}_{|A}$, by definition of M. We call T_A the **induced transformation** of T on A, and $(A, \mathscr{F}_{|A}, \mu_A, T_A)$ is the **induced dynamical system**.
The following property holds:

Proposition 7.1. *If $(X, \mathscr{F}, \mu, T)$ is a finite measure preserving system, and $A \in \mathscr{F}$ is a set of positive measure, then $(A, \mathscr{F}_{|A}, \mu_A, T_A)$ is a probability measure preserving system.*
If in addition T is ergodic, T_A is ergodic.

Proof.
• Let's prove that T_A is μ_A preserving.
Let $B \in \mathscr{F}_{|A}$. If we show that $\mu(T_A^{-1}B) = \mu(T^{-1}B)$, the result is proved by T-invariance of μ.
For $k \geq 2$ let $C_k = \{x \in X \backslash A |\ Tx \in X\backslash A, \dots, T^{k-1}x \in X\backslash A,\ T^k x \in A\}$ and $C_1 = (X\backslash A) \cap T^{-1}A$.

Using the fact that $\mu(A) = \mu\left(\dot\bigcup_{k\geq 1} \{\tau_A = k\}\right)$, we get

$$\begin{aligned}\mu(T_A^{-1}B) &= \mu\left(\dot\bigcup_{k\geq 1} \{\tau_A = k\} \cap T_A^{-1}B)\right) \\ &= \mu\left(\dot\bigcup_{k\geq 1} \{\tau_A = k\} \cap T^{-k}B\right) = \sum_{n=1}^{+\infty} \mu(\{\tau_A = k\} \cap T^{-k}B),\end{aligned}$$

and since we have $T^{-1}A = \{\tau_A = 1\} \cup C_1$,

$$\begin{aligned}\mu(T^{-1}B) &= \mu(T^{-1}B \cap T^{-1}A) \\ &= \mu\left((T^{-1}B) \cap \{\tau_A = 1\}\right) + \mu\left((T^{-1}B) \cap C_1\right) \\ &= \mu\left((T^{-1}B) \cap \{\tau_A = 1\}\right) + \mu\left((T^{-2}B) \cap T^{-1}C_1\right),\end{aligned}$$

using T-invariance of μ.
Note that for $k \geq 1$, $T^{-1}C_k = \{\tau_A = k+1\} \cup C_{k+1}$. Hence for all $n \geq 1$,

$$\begin{aligned}\mu(T^{-1}B) &= \mu\left((T^{-1}B) \cap \{\tau_A = 1\}\right) + \mu\left((T^{-2}B) \cap \{\tau_A = 2\}\right) + \mu\left((T^{-2}B) \cap C_2\right) \\ &= \mu\left((T^{-1}B) \cap \{\tau_A = 1\}\right) + \mu\left((T^{-2}B) \cap \{\tau_A = 2\}\right) + \mu\left((T^{-3}B) \cap T^{-1}C_2\right) \\ &= \mu\left((T^{-1}B) \cap \{\tau_A = 1\}\right) + \mu\left((T^{-2}B) \cap \{\tau_A = 2\}\right) + \mu\left((T^{-3}B) \cap \{\tau_A = 3\}\right) + \mu\left((T^{-3}B) \cap C_3\right) \\ &= \dots \\ &= \left(\sum_{k=1}^{n} \mu\left((T^{-k}B) \cap \{\tau_A = k\}\right)\right) + \mu\left((T^{-n}B) \cap C_n\right).\end{aligned}$$

Now, since the C_n are disjoint,

$$\mu(X) \geq \mu\left(\bigcup_{n\geq 1} (T^{-n}B) \cap C_n\right) = \sum_{n=0}^{+\infty} \mu\left((T^{-n}B) \cap C_n\right).$$

Consequently, $\mu\left((T^{-n}B) \cap C_n\right) \to 0$ as n goes to infinity, and we get that $\mu(T_A^{-1}B) = \mu(T^{-1}B)$.
• Suppose that T is ergodic. Let $f : A \to \mathbb{R}$ be a measurable function such that f is μ_A-almost everywhere T_A invariant ($f \circ T_A = f$ μ_A-almost everywhere) . Since $\mu(A) > 0$, by ergodicity of T, for almost every $x \in X$ there exists $n \geq 1$ such that $T^n x \in A$. Hence

$$\begin{array}{rccc} r_A & : & X & \to & \mathbb{N} \\ & & x & \mapsto & \min\{n \geq 1 |\ T^n x \in A\} \end{array}$$

and

$$\begin{array}{rccc} F & : & X & \to & \mathbb{R} \\ & & x & \mapsto & f(T^{r_A(x)}x) \end{array}$$

are almost everywhere well defined.
In addition, F is T-invariant: To see it, let $x \in X$ such that $F(x)$ is well defined. We have to

distinguish two cases.
(i) If $r_A(x) \geq 2$ then $r_A(Tx) = r_A(x) - 1$. In this case, $F(Tx) = f(T^{r_A(Tx)}(Tx)) = f(T^{r_A(x)-1}(Tx)) = f(T^{r_A(x)}x) = F(x)$.
(ii) If $r_A(x) = 1$ then $Tx \in A$. We thus have $F(Tx) = f(T^{r_A(Tx)}(Tx)) = f(T_A Tx) = f(T_A T^{r_A(x)}x) = f(T^{r_A(x)}x) = F(x)$ by T_A invariance of f. By ergodicity of T, F is constant almost everywhere, and since $F_{|A} = f$ we get that f is constant almost everywhere. □

Theorem 7.1. *(Kac's Lemma)*
Let $(X, \mathscr{F}, \mu, T)$ be an ergodic system with finite measure, and let $A \in \mathscr{F}$ with $\mu(A) > 0$. Then,

$$\int_A \tau_A d\mu_A = \frac{\mu(X)}{\mu(A)}.$$

Proof. Ergodicity of T_A allows us to apply Birkhoff's theorem with respect to μ_A: (The sets of measure 0 contained in A are the same for μ_A and μ) For almost every $x \in A$,

$$\frac{1}{n}\sum_{k=0}^{n-1} \tau_A \circ T_A^k x \to \frac{1}{\mu(X)} \int_A \tau_A d\mu_A. \tag{1}$$

Note that for $n \geq 1$ and for $x \in A \backslash M$, the sum $\sum_{k=0}^{n-1} \tau_A \circ T_A^k x =: N_n(x)$ denotes the number of iterations we have to wait before x comes back to A exactly n times. Consequently,

$$\sum_{k=0}^{N_n(x)-1} \mathbf{1}_A(T^k x) = n.$$

Furthermore $N_n(x) \to +\infty$ as n goes to infinity. Applying Birkhoff's theorem to μ gives: for almost every x in A,

$$\frac{1}{N_n(x)} \sum_{k=0}^{N_n(x)-1} \mathbf{1}_A(T^k x) \to \frac{\mu(A)}{\mu(X)}, \tag{2}$$

as n goes to infinity.
Equations (1) and (2) then give:
For almost all $x \in A$, $\frac{N_n(x)}{n} \to \frac{\mu(X)}{\mu(A)}$ and $\frac{N_n(x)}{n} \to \frac{1}{\mu(X)} \int_A \tau_A d\mu_A$. Now pick one x such that the two previous convergences hold and the result is proved. □

Remark 7.1. *We could have verified in an easy way that the result is true for i.i.d sequences. We develop this heuristic argument in the next section.*

8 Hitting time statistics - A naive approach

Suppose that (X, d) is a metric space. Let $\mathscr{B}$ be the borel sigma algebra over X, μ a probability measure and T an ergodic transformation of X. Given $x_0 \in X$ and $r > 0$, for $x \in X$, let $\tau_{B(x_0,r)}(x)$ denote the first time x hits the ball $B(x_0, r)$:

$$\tau_{B(x_0,r)}(x) = \inf\left\{n \geq 1 |\ T^n x \in B(x_0, r)\right\}.$$

If A is a set of positive measure, Kac's lemma gives an idea of the return time. We might want to have more accurate estimates according to return/hitting times. More precisely, we would like to know the behavior of the quantities

$$G_r(t) = \mu\left(\left\{x \in B(x_0, r) |\ \tau_B(x_0, r)(x) > \frac{t}{\mu(B(x_0, r))}\right\}\right),$$

for all $t > 0$, when r is going to zero.
For example, does the limit $\lim_{r\to 0} G_r(t)$ exist?
In section 9 we will see that for mixing enough dynamical systems, $\mu(B(x_0,r))\tau_B(x_0,r)(.)$ converges in law to an exponential of parameter one. In particular, for all $t \geq 0$, $G_r(t) \to e^{-t}$.
The following heuristic argument gives an explanation of this result:
Assume that $X := [0,1]$ and $\mu = \lambda$ is the Lebesgue measure. We make the hypothesis that the orbit of every element is an $i.i.d$ process.
Pick an i.i.d sequence $Y(x) = (y_i)_{i\geq 0}$, where each y_i is a X-valued random variable with distribution λ.
Assume that $(y_i)_{i\geq 0}$ is the orbit of y_0. Let $T_r = \lambda\big(B(x_0,r)\big)\tau_{B(x_0,r)}(.)$.
Then for $t \in \mathbb{R}$,

$$\phi_{T_r}(t) = \mathbb{E}[e^{itT_r}] = \int_{\mathbb{R}} e^{2itrx} d\mu_{\tau_{B(x_0,r)}}(x), \tag{3}$$

And since for $n \geq 1$ we have

$$\begin{aligned}\lambda(\{\tau_{B(x_0,r)} = n\}) &= \lambda(\{x \in X | y_1, \dots, y_{n-1} \in B(x_0,r)^c \text{ and } y_n \in B(x_0,r)\}) \\ &= \Big(1 - \lambda\big(B(x_0,r)\big)\Big)^{n-1} \lambda\big(B(x_0,r)\big) \\ &= (1-2r)^{n-1}2r,\end{aligned}$$

(3) gives us:

$$\phi_{T_r}(t) = \sum_{n\geq 1} e^{2itrn}(1-2r)^{n-1}2r = \frac{2r}{2r - (1 - e^{-2itr})}$$

Hence for all $t \in \mathbb{R}$, $\phi_{T_r}(t) \to \frac{1}{1-it}$ as r goes to zero, which means that T_r converges in distribution to an exponential law of parameter one.

9 Hitting/Return time statistics for Gibbs measures

In all the section, $X = [0,1]$ endowed with its Borel sigma algebra $\mathscr{B}$.
For $A \in \mathscr{B}$, τ_A is the hitting time in A: $\tau_A : X \to \mathbb{N} \cup \{+\infty\}$.

9.1 The case of expanding maps

Expanding maps which preserve an absolutely continuous (according to Lebesgue) measure give an example of dynamical systems for which Theorem 9.1 is going to apply.
We first prove a general estimate for the size of cylinders.
Let $T : X \to X$ be a topologically mixing, piecewise $\mathscr{C}^{1+\alpha}$, β-expanding map, compatible with a Markov partition $\mathscr{J}$. Define $S = \bigcup_{n\geq 0} T^{-n}\partial\mathscr{J}$.

Proposition 9.1. *There exists $C, D > 0$ such that for any x not in S, we have*

$$C \leq |(T^n)'(x)|\, diam\,(int\mathscr{J}_n(x)) \leq D.$$

Lemma 9.1. *Let $\psi : X \to \mathbb{R}$ be an α-Hölder function. Then, for any x, y in the same n-cylinder, with x, y not in S, we have*

$$|S_n\psi(x) - S_n\psi(y)| \leq \frac{|\psi|_\alpha}{\beta^\alpha - 1},$$

where $|\psi|_\alpha$ is the α-Hölder modulus of ψ : $|\psi|_\alpha = \sup_{x\neq y} \dfrac{d(\psi(x),\psi(y))}{d(x,y)^\alpha}$.

Proof of Lemma 9.1.
Note that if x, y are in the same element of the Markov partition (without being in S), the Mean Value Theorem ensures that for all $x, y \in X$ with $x \neq y$, we have $\frac{d(Tx,Ty)}{d(x,y)} \geq \beta$. We deduce that if $n \geq k$ and x, y are in the same n-cylinder, we have

$$\frac{d(T^n x, T^n y)}{d(T^k x, T^k y)} = \frac{d(T^n x, T^n y)}{d(T^{n-1} x, T^{n-1} y)} \cdots \frac{d(T^{k+1} x, T^{k+1} y)}{d(T^k x, T^k y)} \geq \beta^{n-k}.$$

Now, for x, y in the same n-cylinder, we get

$$\begin{aligned}
|S_n\psi(x) - S_n\psi(y)| &\leq \sum_{k=0}^{n-1} d(\psi(T^k x), \psi(T^k y)) \\
&\leq \sum_{k=0}^{n-1} |\psi|_\alpha d(T^k x, T^k y)^\alpha \\
&\leq |\psi|_\alpha \sum_{k=0}^{n-1} \left(\beta^{k-n} d(T^n x, T^n y)\right)^\alpha \\
&\leq |\psi|_\alpha \sum_{k=0}^{n-1} \left(\beta^{k-n}\right)^\alpha \leq \frac{|\psi|_\alpha}{\beta^\alpha - 1}.
\end{aligned}$$

□

Proof of proposition 9.1.
Note that $\psi = \log |T'|$ is α-Hölder (as a composition of Lispschitz functions with an α-Hölder function, because T' is bounded by Hölder continuity). Applying the chain rule shows that for x, y in the same n-cylinder,

$$S_n\psi(x) - S_n\psi(y) = \log \left| \frac{(T^n)'x}{(T^n)'y} \right|.$$

The lemma gives a constant $D \geq 1$ such that

$$\left| \frac{(T^n)'x}{(T^n)'y} \right| \leq D. \tag{4}$$

Let $x \in X \backslash S$. For $n \geq 0$, thre exists $i_0, \ldots, i_{n-1}$ such that

$$\mathscr{J}_n(x) = J_{i_0} \cap T^{-1} J_{i_1} \cap \cdots \cap T^{-(n-1)} J_{i_{n-1}},$$

So that

$$\mathrm{int}\, \mathscr{J}_n(x) = \mathrm{int}(J_{i_0}) \cap T^{-1} \mathrm{int}(J_{i_1}) \cap \cdots \cap T^{-(n-1)} \mathrm{int}(J_{i_{n-1}}), \tag{5}$$

because the interior of a finite intersection is the intersection of the interiors and that since the J_j are closed intervals, we have $T^{-i}\mathrm{int}(J_j) = \mathrm{int} T^{-i}(J_j)$. (Inclusion $\subset$ is easy. We get the other one writing $J_j = \mathrm{int}(J_j) \cup \partial J_j$)
From equation (5), we conclude that $\mathrm{int}\, \mathscr{J}_n(x)$ is an open interval and that T^n is a diffeomorphism from $\mathrm{int}\, \mathscr{J}_n(x)$ onto its image (which is consequently an open interval).
If we write $(a, b) := \mathrm{int}\, \mathscr{J}_n(x)$, we have $T^n(a, b) = (T^n a, T^n b)$ or $(T^n b, T^n a)$, since T is increasing or decreasing on every $\mathrm{int}(J_i)$.
The mean value theorem gives a $y \in \mathrm{int}\, \mathscr{J}_n(x)$ such that

$$\mathrm{diam}(T^n \mathrm{int}\, \mathscr{J}_n(x)) = |(T^n)'y| \mathrm{diam}\left(\mathrm{int}\, \mathscr{J}_n(x)\right).$$

(In fact T^n may not be continuous on the boundary of $\mathrm{int}\,\mathscr{J}_n(x)$, but applying the Mean Value Theorem to the continuous extension of T^n on $[a,b]$ gives the result)
Since y is in the n-cylinder of x, inequality (4) is true. Now,

$$|(T^n)'x|\mathrm{diam}\,(\mathrm{int}\,\mathscr{J}_n(x)) \le D|(T^n)'y|\mathrm{diam}\,(\mathrm{int}\,\mathscr{J}_n(x)) = D\mathrm{diam}(T^n\mathrm{int}\,\mathscr{J}_n(x)) \le D,$$

since $\mathrm{diam}(T^n\,\mathscr{J}_n(x)) \le \mathrm{diam}(X) = 1$.
Let $\rho = \min_{J\in\mathscr{J}} \mathrm{diam}J > 0$. Since by (4) we have

$$\left|\frac{(T^n)'y}{(T^n)'x}\right| \le D,$$

we get

$$\begin{aligned}|(T^n)'x|\mathrm{diam}\,(\mathrm{int}\,\mathscr{J}_n(x)) \ge \frac{1}{D}|(T^n)'y|\mathrm{diam}\,(\mathrm{int}\,\mathscr{J}_n(x)) &= \frac{\mathrm{diam}(T^n\mathrm{int}\,\mathscr{J}_n(x))}{D}\\ &= \frac{\mathrm{diam}\,(\mathrm{int}T^n\,\mathscr{J}_n(x))}{D}\\ &\ge \frac{\rho}{D},\end{aligned}$$

since $T^n\,\mathscr{J}_n(x)$ is a union of some of the J_i. □

Corollary 9.1. *In addition to the previous conditions, suppose that T preserves an absolutely continuous measure μ with respect to Lebesgue measure λ, such that there exists $0 < d_1 \le d_2$ with $d_1\lambda \le \mu \le d_2\lambda$.*
μ is a Gibbs measure under the potential $\varphi = -\log|T'|$.

Proof. Since we have $0 < d_1 \le \mu \le d_2$ and $\mathrm{diam}\,(\mathrm{int}\,\mathscr{J}_n(x)) = \lambda(\mathscr{J}_n(x))$, we get

$$Cd_1 \le |(T^n)'(x)|\mu(\mathscr{J}_n(x)) \le Dd_2.$$

Now set $\varphi = -\log|T'|$ and $\kappa = \max(1+Dd_2, 1+\frac{1}{Cd_1})$. The chain rule gives for all x not in S:

$$e^{S_n\varphi(x)} = \frac{1}{|(T^n)'x|}$$

so that for all x not in S:

$$\frac{1}{\kappa} \le \frac{\mu(\mathscr{J}_n(x))}{e^{S_n\varphi(x)}} \le \kappa.$$

Hence μ is a Gibbs measure for the pressure $P(\varphi) = 0$. □

9.2 Statement of Theorem 9.1

We now state the return/hitting time theorem.

Theorem 9.1. *Let $T : X \to X$ be a piecewise topologically mixing $\mathscr{C}^{1+\alpha}$ expanding ($|T'| \ge \beta > 1$ for some β) map. Suppose that T preserves a Gibbs measure of positive dimension μ with α-Hölder potential φ and pressure zero, such that there exists c verifying $\varphi < c < 0$. Then for almost all $x_0 \in X$, the measurable functions $\mu(B(x_0,r))\tau_B(x_0,r)(.)$ converge in law with respect to μ and $\mu_{B(x_0,r)}$ to an exponential of parameter one as r decreases to zero.*

Example 9.1. *The doubling map $x \mapsto 2x \mod (1)$ is the simplest example of such maps. It is topologically mixing: Since we are dealing with open sets, it suffices to prove it for dyadic intervals. Let I be an interval of X and J be a dyadic interval of length $\frac{1}{2^N}$. Then, for all $n \ge N$, $T^{-n}I \cap J$ is not empty.*

Remark 9.1. *We actually know from a general theorem (Ruelle-Perron-Frobenius) that such a measure exists and is unique. See [8].*

We follow the proof of [8].

9.3 Idea of the proof of Theorem 9.1

Lemma 9.2 doesn't use the hypothesis of theorem 9.1.

Lemma 9.2. *Let $A \in \mathscr{B}$ with $\mu(A) > 0$. Set $\delta(A) = \sup_{n \in \mathbb{N}} |\mu(\tau_A > n) - \mu_A(\tau_A > n)|$. Then, for all $n \in \mathbb{N}$, we have*

$$|\mu(\tau_A > n) - (1 - \mu(A))^n| \leq \delta(A).$$

Proof. The proof if based on the fact that for $k \in \mathbb{N}$,

$$T^{-1}\{\tau_A > k\} = \Big\{x \in X |\ T^2 x \in X \backslash A, \ldots, T^k x \in X \backslash A, T^{k+1} x \in A\Big\} = \{\tau_A > k+1\} \cup \Big(T^{-1} A \cap \{\tau_A \circ T > k\}\Big),$$

hence

$$T^{-1}\{\tau_A > k\} = \{\tau_A > k+1\} \cup T^{-1}(A \cap \{\tau_A > k\}).$$

Since T is measure preserving, we get

$$\mu(\{\tau_A > k\}) = \mu(\{\tau_A > k+1\}) + \mu(A)\mu_A(\{\tau_A > k\}). \tag{6}$$

Rewriting (6) gives

$$\mu(\{\tau_A > k+1\}) = [1 - \mu(A)]\,\mu(\{\tau_A > k\}) + \mu(A)\,[\mu(\{\tau_A > k\}) - \mu_A(\{\tau_A > k\})],$$

and thus passing to the supremum

$$|\mu(\{\tau_A > k+1\}) - [1 - \mu(A)]\,\mu(\{\tau_A > k\})| \leq \mu(A)\delta(A). \tag{7}$$

Now, for $n \in \mathbb{N}$, using repeatedly (7) in a chain of triangle inequalities gives

$$\begin{aligned} |\mu(\tau_A > n) - (1 - \mu(A))^n| &\leq \sum_{k=0}^{n-1} (1 - \mu(A))^k \delta(A)\mu(A) \\ &\leq \frac{1}{\mu(A)} \delta(A)\mu(A) = \delta(A). \end{aligned}$$

□

Assume that for almost every $x_0 \in X$ we have $\lim_{r \to 0} \delta(B(x_0, r)) = 0$. Then we can prove Theorem 9.1.

Proof. (of Theorem 9.1 having $\lim_{r \to 0} \delta(B(x_0, r)) = 0$ for almost every x_0.)
Let $x_0 \in X$ such that $\lim_{r \to 0} \delta(B(x_0, r)) = 0$. This concerns almost all $x_0 \in X$. Let $A = B(x_0, r)$ for $r > 0$.
Let $t > 0$ and $n = \left\lfloor \frac{t}{\mu(A)} \right\rfloor$. Since we have $\mu(\{\mu(A)\tau_A > t\}) = \mu(\{\tau_A > n\})$, we can apply lemma (9.2) to get

$$\left|\mu(\{\mu(A)\tau_A > t\}) - e^{-t}\right| \leq \delta(A) + \left|(1 - \mu(A))^n - e^{-t}\right|,$$

for all $n \in \mathbb{N}$. Since $\delta(A) \to 0$ as r goes to zero, we control the second term.

$$\left|(1 - \mu(A))^n - e^{-t}\right| \leq \left|\left(1 - \frac{t}{n}\right)^n - e^{-t}\right| + \left|(1 - \frac{t}{n})^n - (1 - \mu(A))^n\right|.$$

The first term goes to zero as n goes to $+\infty$ and

$$\left|\left(1-\frac{t}{n}\right)^n-(1-\mu(A))^n\right|\leq n\left|\mu(A)-\frac{t}{n}\right|$$

by the mean value theorem.
Finally, basic inequalities show that

$$n\left|\mu(A)-\frac{t}{n}\right|\leq\frac{t}{n-1}.$$

Now fix $\epsilon>0$. Taking n big enough gives that for r small enough,

$$\left|\mu(\{\mu(A)\tau_A>t\})-e^{-t}\right|\leq\epsilon,$$

and the result is proved for μ distributions (hitting times).
Finally, since

$$\left|\mu_A(\{\mu(A)\tau_A>t\})-e^{-t}\right|\leq\left|\mu(\{\mu(A)\tau_A>t\})-e^{-t}\right|+\delta(A)$$

the statement is still true for μ_A distributions (return times). □

We thus just have to prove that under the conditions of theorem 9.1 the following result holds.

Proposition 9.2. *For μ-almost every $x_0\in X$ we have $\delta(B(x_0,r))\to 0$ as r goes to zero.*

9.4 Proof of Proposition 9.2

Let $T:X\to X$ verifying the hypothesis of the theorem.
We admit the following two theorems. The proofs can be found in [8].

Theorem 9.2. *μ has the following mixing property:*
There exists $c>0$ and $\theta\in(0,1)$ such that for all $f:X\to X$ Lipschitz and $g:X\to X$ $\mathscr{L}^\infty$ we have for all $n\in\mathbb{N}$

$$\left|\int_X fg\circ T^n d\mu-\int_X fd\mu\int_X gd\mu\right|\leq c\theta^n|f|_{Lip}||g||_\infty,$$

where $|f|_{Lip}$ is defined as: $|f|_{Lip}=\sup_{x\neq y}\dfrac{d(f(x),f(y))}{d(x,y)}$.

Theorem 9.3. *Under the conditions of theorem 9.1, for almost all $x\in X$, the limit*

$$R(x):=\lim_{r\to 0}\frac{\log\tau_{B(x,r)}(x)}{-\log r}$$

exists and is equal to the dimension of μ.

Also recall that from theorem 6.3 we know that the pointwise dimension d_μ is constant almost everywhere and is equal to the dimension of μ.

We now state Lebesgue's density theorem. We won't prove it. Note that it is a consequence of a more general result: Lebesgue differentiation theorem, which is proved in [6].

Theorem 9.4. *(Lebesgue's density theorem)*
*Let $\emptyset\neq A\in\mathscr{B}$. For $x\in A$ and $r>0$ define $d_r(x)=\frac{\mu(A\cap B(x,r))}{\mu(B(x,r))}$. Then for almost every $x\in A$, $d_r(x)\to 1$ as r goes to zero. We call such points **density points of** A.*

Lemma 9.3. *For μ-almost every $x_0 \in X$ and for all $d \in (0, \dim_H(\mu))$, we have*

$$\mu_{B(x_0,r)}\left(\left\{\tau_{B(x,r)} \leq \frac{1}{r^d}\right\}\right) \to 0$$

as r goes to zero.

We call such x_0 a ***non sticky point****.*

Proof. Let $d \in (0, \dim_H(\mu))$. For $r_0 > 0$, let $L_{r_0} = \{x \in X | \ \forall r < r_0, \ \tau_{B(x,2r)}(x) > \frac{1}{r^d}\}$.

• $L := \bigcup_{r_0>0} L_{r_0}$ has full measure (note that for $r_1 < r_0$, $L_{r_0} \subset L_{r_1}$, so $\bigcup_{r_0>0} L_{r_0} = \bigcup_{n\geq 1} L_{\frac{1}{n}}$ is measurable):

Using theorem 9.3, almost every $x \in X$ satisfies $R(x) = \dim_H(\mu)$. Let $x \in X$ be such that the previous equality is true.

Suppose that $\tau_{B(x,2r_n)}(x) \leq \frac{1}{r_n^d}$ for a sequence (r_n) of radii decreasing to zero. Then,

$$\begin{aligned}\dim_H(\mu) = R(x) &= \lim_{n\to+\infty} \frac{\log \tau_{B(x,2r_n)}(x)}{-\log 2r_n} \\ &\leq \lim_{n\to+\infty} \frac{d \log r_n}{\log 2r_n} = d,\end{aligned}$$

which is a contradiction. Hence there exists $r_0 > 0$ such that for all $r < r_0$ we have $\tau_{B(x,2r)}(x) > \frac{1}{r^d}$. We thus know that there exists an r_0 such that $x \in L_{r_0}$.

This proves that L has full measure.

• Let $r_0 > 0$ and $x_0 \in L_{r_0}$. Almost such x_0 is a density point of L_{r_0} by Lebesgue's Density Theorem - assume that x_0 is a density point. Let $r < r_0$ and $x \in B(x_0, r)$ such that $\tau_{B(x_0,r)}(x) \leq \frac{1}{r^d}$. Then $\tau_{B(x_0,2r)}(x) \leq \frac{1}{r^d}$ and $x \in X \backslash L_{r_0}$. Whence

$$\mu_{B(x_0,r)}\left(\left\{\tau_{B(x_0,r)} \leq \frac{1}{r^d}\right\}\right) \leq \mu_{B(x_0,r)}(X \backslash L_{r_0}) \to 0$$

as r goes to zero, since x_0 is a density point of L_{r_0}. □

Lemma 9.4. *For μ-almost every $x_0 \in X$ and for all $d \in (0, \dim_H(\mu))$, we have*

$$\mu\left(\left\{\tau_{B(x_0,r)}(x) \leq \frac{1}{r^d}\right\}\right) \to 0$$

as r goes to zero.

Proof. For A measurable and $k \geq 1$, we have

$$\mu(\{\tau_A \leq n\}) = \mu\bigg(\bigcup_{k=1}^{n} T^{-k} A\bigg).$$

By invariance of μ we get

$$\mu(\{\tau_A \leq n\}) \leq n\mu(A). \tag{8}$$

Now, let $x_0 \in X$ such that $d_\mu(x) \left(= \lim_{r\to 0} \frac{\log \mu(B(x_0,r))}{\log r}\right) = \dim_H(\mu)$. This concerns almost every point by Theorem 6.3. Then (8) gives

$$\mu\left(\left\{\tau_{B(x_0,r)} \leq \frac{1}{r^d}\right\}\right) \leq r^d \left\lfloor \frac{1}{r^d} \right\rfloor \frac{\mu(B(x_0,r))}{r^d},$$

and

$$\log\left(\frac{\mu(B(x_0,r))}{r^d}\right) = \log(r)\left[\frac{\log(\mu(B(x_0,r)))}{\log r} - \frac{\log(r^d)}{\log r}\right]$$
$$= \log(r)\left[\frac{\log(\mu(B(x_0,r)))}{\log r} - d\right] \to -\infty, \qquad \sim_{r\to 0} \log(r)\left[\dim_H(\mu) - d\right]$$

since

$$\frac{\log(\mu(B(x_0,r)))}{\log r} \to \dim_H(\mu) > d.$$

Therefore $\frac{\mu(B(x_0,r))}{r^d} \to 0$ as r goes to zero. □

Lemma 9.5. *There exists γ and $\delta > 0$ such that for any $x_0 \in X$, for any $0 < r$ small enough, we have for all $0 \leq \rho < r$*

$$\mu\big(B(x_0,r)\backslash B(x_0,r-\rho)\big) \leq \gamma\rho^\delta.$$

Proof. Since for x not in S we have $\varphi \leq c < 0$, by the Gibbs property we get for x not in S and for $k \in \mathbb{N}$

$$\mu(\mathscr{I}_k(x)) \leq \kappa e^{S_k\varphi(x)} \leq \kappa e^{kc},$$

so there exists $a > 0$ such that for k big enough,

$$\mu(\mathscr{I}_k(x)) \leq e^{S_k\varphi(x)} \leq e^{-ka}. \tag{9}$$

Besides, there exists $b > 0$ such that for k big enough,

$$\mathrm{diam}(\mathrm{int}\,\mathscr{I}_k(x)) \geq e^{-kb}. \tag{10}$$

To see it, use Lemma 9.1: Since we have

$$C \leq \left|(T^n)'(x)\right| \mathrm{diam}\left(\mathrm{int}\,\mathscr{I}_n(x)\right),$$

we get by the chain rule

$$\mathrm{diam}\left(\mathrm{int}\,\mathscr{I}_n(x)\right) \geq \frac{C}{\left(\sup_I |T'|\right)^n}.$$

Now, if I is a small enough interval of X there exists $k \in \mathbb{N}$ such that $e^{-(k+1)b} \leq \mathrm{diam}(I) \leq e^{-kb}$. Equation (10) and the fact that $\mathrm{diam}(I) \leq e^{-kb}$ then implies that I can't overlap with 3 or more k-cylinders. Hence by equation (9) and the inequality $e^{-(k+1)b} \leq \mathrm{diam}(I)$ we get that

$$\mu(I) \leq 2e^{-ak} = 2(e^{-bk})^{\frac{a}{b}} \leq K\mathrm{diam}(I)^{\frac{a}{b}},$$

where $K = 2e^a$.
We thus get the desired result for $\gamma = 2K$ and $\delta = \frac{a}{b}$. □

Lemma 9.6. *For μ-almost every $x_0 \in X$ we have $\delta(B(x_0,r)) \to 0$ as r goes to zero.*

Proof. Let $d \in (0, \dim_H(\mu))$ and $x_0 \in X$ such that the results of lemmas 9.3, 9.4 and 9.5 are true in x_0.
For $r > 0$ and $\rho \in (0,r)$, write $A := B(x_0,r)$, $A' := B(x_0, r-\rho)$ and $E_n := \{\tau_A \geq n\}$. We prove several inequalities. Then we will chose an appropriate ρ. For $n \in \mathbb{N}$ and $g \leq n$ integer,

$$A \cap E_n \subset A \cap T^{-g}E_{n-g} \subset \big(A \cap \{\tau_A \leq g\}\big) \cup \big(A \cap T^{-g}E_{n-g}\big)$$

and

$$A \cap T^{-g}E_{n-g} = \big(A \cap E_n\big) \cup \big(A \cap \{\tau_A \leq g\} \cap T^{-g}E_{n-g}\big) \subset \big(A \cap \{\tau_A \leq g\}\big) \cup \big(A \cap E_n\big),$$

therefore

$$|\mu(A \cap E_n) - \mu(A \cap T^{-g}E_{n-g})| \leq \mu(A \cap \{\tau_A \leq g\}). \tag{11}$$

In the same way we get

$$|\mu(E_n) - \mu(T^{-g}E_{n-g})| \leq \mu(\{\tau_A \leq g\}). \tag{12}$$

For $\rho \in (0, r)$, Set $\phi(x) = \max(0, 1 - \frac{d(x,A')}{\rho})$. One can show that ϕ is $\frac{1}{\rho}$-Lipschitz by looking at all the possible different cases. In addition, we have $\mathbf{1}_{A'} \leq \phi \leq \mathbf{1}_A$. Hence

$$\mu(A) - \int_X \phi d\mu = \left|\mu(A) - \int_X \phi d\mu\right| \leq \mu(A \backslash A'). \tag{13}$$

Besides, theorem 9.2 implies that

$$\left|\int_X \phi \left(\mathbf{1}_{E_{n-g}} \circ T^n\right) d\mu - \mu(E_{n-g}) \int_X \phi d\mu\right| \leq c\theta^n \frac{1}{\rho}, \tag{14}$$

where $c > 0$ and $\theta \in (0, 1)$.
Finally, since we have

$$\mathbf{1}_{A' \cap T^{-g}E_{n-g}} \leq \phi \mathbf{1}_{E_{n-g}} \circ T^g \leq \mathbf{1}_{A \cap T^{-g}E_{n-g}} \circ T^g,$$

Integrating every terms of the inequality, we get

$$\mu(A' \cap T^{-g}E_{n-g}) \leq \int_X \phi \mathbf{1}_{E_{n-g}} \circ T^g d\mu \leq \mu(A \cap T^{-g}E_{n-g}),$$

thus

$$\left|\int_X \phi \left(\mathbf{1}_{E_{n-g}} \circ T^g\right) d\mu - \mu(A \cap T^{-g}E_{n-g})\right| \leq \mu\left((A \backslash A') \cap T^{-g}E_{n-g}\right) \leq \mu(A \backslash A'). \tag{15}$$

Now,

$$\begin{aligned}
|\mu_A(E_n) - \mu(E_n)| &= \frac{1}{\mu(A)} \left|\mu(A \cap E_n) - \mu(A)\mu(E_n)\right| \\
&\leq \frac{1}{\mu(A)} \bigg(\left|\mu(A \cap E_n) - \mu(A \cap T^{-g}E_{n-g})\right| + \left|\mu(A \cap T^{-g}E_{n-g}) - \int_X \phi \mathbf{1}_{E_{n-g}} \circ T^g d\mu\right| + \\
&+ \left|\int_X \phi \mathbf{1}_{E_{n-g}} \circ T^g d\mu - \mu(E_{n-g}) \int_X \phi d\mu\right| + \left|\mu(E_{n-g}) \int_X \phi d\mu - \mu(A)\mu(E_n)\right| \bigg) \\
&\leq \mu_A(\{\tau_A \leq g\}) + \frac{\mu(A \backslash A')}{\mu(A)} + \frac{c\theta^g}{\mu(A)\rho} + \left|\mu(E_{n-g}) \int_X \phi d\mu - \mu(A)\mu(E_n)\right|,
\end{aligned}$$

thanks to equations (11), (14) and (15).
We also have

$$\begin{aligned}
\left|\mu(E_{n-g}) \int_X \phi d\mu - \mu(A)\mu(E_n)\right| &= \left|\mu(E_{n-g})\Big(\int_X \phi d\mu - \mu(A)\Big) + \Big(\mu(E_{n-g}) - \mu(E_n)\Big)\mu(A)\right| \\
&\leq \mu(E_{n-g}\mu(A \backslash A') + \mu(A)\mu(\{\tau_A \leq g\}) \\
&\leq \mu(A \backslash A') + \mu(A)\mu(\{\tau_A \leq g\}),
\end{aligned}$$

by equations (12) and (13).
Finally,

$$|\mu_A(E_n) - \mu(E_n)| \leq \mu_A(\{\tau_A \leq g\}) + 2\frac{\mu(A\backslash A')}{\mu(A)} + \frac{c\theta^g}{\mu(A)\rho} + \mu(\{\tau_A \leq g\}). \tag{16}$$

Note that for $g \geq n$ we have

$$\begin{aligned}|\mu_A(\{\tau_A \leq n\}) - \mu(\{\tau_A \leq n\})| &\leq \mu_A(\{\tau_A \leq n\}) + \mu(\{\tau_A \leq n\}) \\ &\leq \mu_A(\{\tau_A \leq g\}) + \mu(\{\tau_A \leq g\}).\end{aligned}$$

Consequently (16) is true for all $g \geq 1$.
We thus have, passing to the supremum over n,

$$\delta(A) \leq \mu_A(\{\tau_A \leq g\}) + 2\frac{\mu(A\backslash A')}{\mu(A)} + \frac{c\theta^g}{\mu(A)\rho} + \mu(\{\tau_A \leq g\}).$$

We now take $g = \left\lfloor \frac{1}{r^d} \right\rfloor$ and $\rho = \theta^{\frac{g}{2}}$. One can check that $\rho < r$ by setting $f(r) := \theta^{\frac{1}{2}(\frac{1}{r^d}-1)} - r$ and showing that f is always negative (f is decreasing and $f(r) \to 0$ as $r > 0$ goes to zero).
Now,

$$\mu_A(\{\tau_A \leq g\}) \to 0$$

since x_0 is a non sticky point by lemma 9.3,
and

$$\mu(\{\tau_A \leq g\}) \to 0$$

by lemma 9.4.
Lemma 9.5 shows that $\mu(A\backslash A') \leq \gamma\theta^{\frac{ag}{2}}$. As in lemma 9.4, the existence of pointwise dimension ensures that the two middle terms converge to zero as r goes to zero. □

10 Symbolic dynamics

Theorem 9.1 deals with hitting/return time statistics according to balls in a metric space. It is not always easy to do so: Sometimes it is easier to find estimates on cylinders. We thus code the system in order to work with symbolic dynamics. Before defining the coding of a dynamical system, let us give an example.

10.1 Coding of the doubling map

Let $T : [0,1] \to [0,1]$ be the doubling map: $Tx = 2x \mod (1)$.
For $x \in [0,1]$ let $(x_n)_{n\geq 0} \in \{0,1\}^{\mathbb{N}}$ such that for $n \geq 0$, $x_n = 0$ if $T^n x \in \left[0, \frac{1}{2}\right]$ and $x_n = 1$ otherwise (note that $\mathscr{J} := \left\{\left[0, \frac{1}{2}\right], \left[\frac{1}{2}, 1\right]\right\}$ is a Markov partition according to T). We now work with the set of sequences $X := \{0,1\}^{\mathbb{N}}$ endowed with the product sigma-algebra (see Appendix B) and the map

$$\begin{array}{rccc}\sigma : & X & \longrightarrow & X \\ & (x_n)_{n\geq 0} & \longmapsto & (x_{n+1})_{n\geq 0}.\end{array}$$

σ is called the **left shift** over the alphabet $\{0,1\}$. It is measurable and is measure preserving according to the Bernoulli probability $\mathbb{P}$, defined as

$$\mathbb{P}(\{0,1\}^{\mathbb{N}}) = 1$$

and for $n \geq 1$

$$\mathbb{P}\left(\{a_{i_0}\} \times \{a_{i_1}\} \times \cdots \times \{a_{i_{n-1}}\} \times \prod_{k\in\mathbb{N}\backslash\{i_0,\ldots,i_{n-1}\}} \{0,1\}\right) = \frac{1}{2^n},$$

for each finite family of a_i with $a_i \in \{0,1\}$ (we call such a set a n-**cylinder**). Theorem $B.1$ ensures that $\mathbb{P}$ exists and is uniquely defined on X. To see that σ is $\mathbb{P}$-preserving it suffices to prove it for the class of cylinders - for which the result is straightforward- which is closed under intersection and generates the product sigma-algebra.
Note that if λ is the Lebesgue measure on $[0,1]$, we have for $x \in [0,1] \setminus S$ and $n \geq 1$

$$\lambda(\mathscr{J}_n(x)) = \mathbb{P}\Bigg(\{x_0\} \times \{x_1\} \times \cdots \times \{x_{n-1}\} \times \prod_{k \geq n} \{0,1\} \Bigg),$$

where $S = \bigcup_{n \geq 0} T^{-n} \partial \mathscr{J}$. The product space $(X, \mathbb{P})$ is thus a natural coding of the doubling map.
Remark that in addition, the symbolic system is ergodic (We would expect it since the map T is ergodic):
Let $A \subset X$ be a cylinder. Then for $n \geq 1$, $\sigma^{-n}A$ is in the sigma-algebra generated by the projections π_i, $i \geq n$, which we denote by $\mathscr{F}^n$. As a consequence for all $A \subset X$ measurable, $\sigma^{-n}A \in \mathscr{F}^n$. Now, let A be an invariant subset. Then for all $n \geq 0$, $A \in \mathscr{F}^n$. Hence $A \in \bigcap_{n \in \mathbb{N}} \mathscr{F}^n$.
Note that by definition of $\mathbb{P}$, the π_i are independent. The Kolmogorov Zero-One law then ensures that $\mathbb{P}(A) \in \{0,1\}$.

10.2 Symbolic dynamics for expanding maps

The previous construction can be generalized to expanding maps of $[0,1]$:
Let $T : [0,1] \to [0,1]$ be a piecewise $\mathscr{C}^1$ β-expanding map. Suppose that there exists a Markov partition $\mathscr{J} = \{J_0, \ldots, J_p\}$ compatible with T (the α-Hölder property is not required here).
Let $\mathscr{A} = \{0, \ldots, p\}$, and $A = (a_{i,j})$ be a $(p+1) \times (p+1)$ matrix defined as

$$a_{i,j} = \begin{cases} 1 & \text{if } T_j \subset J_i \\ 0 & \text{otherwise.} \end{cases}$$

A is called the transition matrix of T. Now set

$$X := \Big\{ (x_n)_{n \in \mathbb{N}} \in \mathscr{A}^{\mathbb{N}} | \text{ For all } n \in \mathbb{N},\ a_{x_n, x_{n+1}} = 1 \Big\}.$$

Remark 10.1. *In the case where T is the doubling map, A is the 2×2 matrix filled with ones, whence $X = \{0,1\}^{\mathbb{N}}$.*

For $x \in X$, let $\varphi(x) := \bigcap_{n \in \mathbb{N}} T^{-n}(\mathrm{int} J_{x_n})$. Note that $\varphi(x) \subset [0,1]$ and that the expension property of T ensures that $\varphi(x) = \{a\}$, where $a \in [0,1]$:
If a and b are elements of $\varphi(x)$ for some $x \in X$, then for all $n \in \mathbb{N}$ $T^n a$ and $T^n b$ are in the same element of the partition. We thus can apply the Mean Value Theorem to get for all $n \geq 1$

$$d(a,b) \leq \frac{1}{\beta} d(Ta, Tb) \leq \frac{1}{\beta^2} d(T^2 a, T^2 b) \leq \cdots \leq \frac{1}{\beta^n} d(T^n a, T^n b) \leq \frac{1}{\beta^n} \to 0,$$

therefore $d(a,b) = 0$.
For $x \in X$ we define $\chi(x) := a$, where a is the unique element of $\varphi(x)$.
$\chi : X \to [0,1] \setminus S$ is bijective by construction, where $S = \bigcup_{n \geq 0} T^{-n} \partial \mathscr{J}$. We sum up the coding in the following diagram. σ is the left shift.

$$\begin{array}{ccc} X & \xrightarrow{\sigma} & X \\ \Big\downarrow \chi & & \Big\downarrow \chi \\ [0,1] \setminus S & \xrightarrow{T} & [0,1] \setminus S \end{array}$$

Let $\mathscr{X}$ be the product sigma-algebra restricted to X. Suppose that μ is an invariant measure over $[0,1]$ for which S has measure zero. For $A \in \mathscr{X}$ define $\nu(A) = \mu(\chi A)$.
Note that by construction, the coded system $(X, \mathscr{X}, \nu)$ is ergodic if and only if the system $([0,1], \mathscr{B}([0,1]), \mu)$ is ergodic.

The next section introduces randomness in dynamical systems. Since the hitting/return time statistics results are harder to establish than in the deterministic case, we forget the balls and look at cylinders in symbolic dynamics.

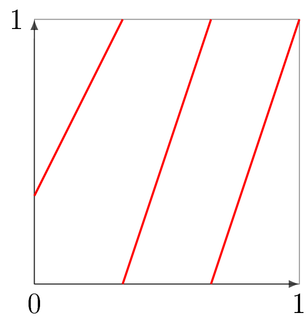
11 Hitting time statistics for random dynamics

In this section we state and prove a hitting time statistics law under strong enough mixing assumptions. The next paragraph gives an example of random dynamical system.

11.1 An example

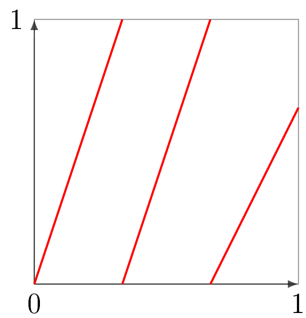
Let $T_1, T_2 : [0,1] \to [0,1]$ defined as

$$T_1 x = \begin{cases} 2x + \frac{1}{3} \text{ if } x \in \left[0, \frac{1}{3}\right) \\ 3x \bmod 1 \text{ otherwise} \end{cases}$$



and

$$T_2 x = \begin{cases} 3x \bmod 1 \text{ if } x \in \left[0, \frac{2}{3}\right) \\ 2x - \frac{4}{3} \text{ otherwise.} \end{cases}$$



The sets $J_1 = \left[0, \frac{1}{3}\right]$, $J_2 = \left[\frac{1}{3}, \frac{2}{3}\right]$ and $J_3 = \left[\frac{2}{3}, 1\right]$ form a Markov partition according to T_1 and T_2. The transition matrices are

$$A_1 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \text{ and } A_2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

We introduce randomness by flipping a coin at each step. Depending on the result we apply T_1 or T_2.

More precisely, let $\Omega = \{1,2\}^{\mathbb{N}}$, $\theta : \Omega \to \Omega$ be the left shift and $\mathbb{P}$ be the $(\frac{1}{2}\ \frac{1}{2})$ Bernoulli measure on Ω. Pick a sequence $\omega \in \Omega$ and define

$$T_\omega x = \begin{cases} T_1 x \text{ if } \omega_0 = 1 \\ T_2 x \text{ if } \omega_0 = 2 \end{cases}.$$

For $x \in [0,1]$ we would like to look at the random iterates

$$T_{\theta^i \omega} \circ \cdots \circ T_{\theta\omega} \circ T_\omega x.$$

It is the topic of the next section.

11.2 Statement of theorem 11.1

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and $\theta : \Omega \to \Omega$ $\mathbb{P}$-preserving.
Define a **finite** alphabet $\mathscr{A} := \{1, \ldots, m\}$ for some $m \geq 2$.
Let $X := \mathscr{A}^{\mathbb{N}}$, endowed with the product sigma algebra $\mathscr{X}$ and the left shift $\sigma : X \to X$.
For $\omega \in \Omega$, let $A(\omega) = (a_{i,j}(\omega))_{1 \leq i,j \leq m}$ where for every ω we have:
- For each $1 \leq i,j \leq m$, $a_{i,j} : \Omega \to \{0,1\}$ is $\mathscr{F}$-measurable
- $A(\omega)$ has at least one non zero entry in each row and each column.

We then define, for $\omega \in \Omega$,

$$X_\omega := \left\{ x \in X |\ \forall i \geq 0,\ a_{x_i, x_{i+1}}(\theta^i \omega) = 1 \right\}.$$

X_ω represents all the admissible paths in $\mathscr{A}$ for a given sequence $(\theta^i(\omega))_{i \geq 0}$.
We assume that there exists a family of probability measures $(\mu_\omega)_{\omega \in \Omega}$ on $(X, \mathscr{X})$ such that:
- For all $A \in \mathscr{X}$, $\omega \mapsto \mu_\omega(A)$ is measurable
- For all $\omega \in \Omega$ μ_ω lies in X_ω (that is, if $A \cap X_\omega = \emptyset$, then $\mu_\omega(A) = 0$)
- For almost every $\omega \in \Omega$, we have for all $i \geq 1$ and $A \in \mathscr{X}$: $\mu_\omega(\sigma^{-i} A) = \mu_{\theta^i \omega}(A)$.

We now define a probability measure μ on X as

$$\mu(A) = \int_\Omega \mu_\omega(A) d\mathbb{P}(\omega).$$

Recall that for $y \in X$, for $n \geq 1$, the n-cylinder centered in y is defined as

$$C_n(y) = \{x \in X |\ x_0 = y_0, \ldots x_{n-1} = y_{n-1}\}.$$

For $n \geq 1$ we let $\mathscr{X}_n$ be the sigma algebra generated by n-cylinders.
We assume that there exists $c_0, h_0 > 0$ and a summable function $\psi : \mathbb{N} \to \mathbb{R}^+$ such that for all $m, n \geq 1$, $g \geq 0$, $A \in \mathscr{X}_n$ and $B \in \mathscr{X}_m$, we have:
(I) $|\mu(A \cap \sigma^{-n-g} B) - \mu(A)\mu(B)| \leq \psi(g)$
(II) for $\mathbb{P}$-almost every $\omega \in \Omega$, for all $y \in X_\omega$, $\frac{1}{c_0} e^{-h_0 n} \leq \mu(C_n(y))$.
(III) for $\mathbb{P}$-almost every $\omega \in \Omega$,

$$\left|\mu_\omega(A \cap \sigma^{-n-g} B) - \mu_\omega(A)\mu_{\theta^{n+g}\omega}(B)\right| \leq \psi(g)\mu_\omega(A)\mu_{\theta^{n+g}\omega}(B)$$

(IV) $\sup_{\omega \in \Omega, x \in X} \mu_\omega(C_1(x)) < 1$

Remark 11.1. *Assumption (II) is impossible in the case of infinite alphabets (for $\omega \in \Omega$, $y \in X_\omega$ and $n = 1$ the fact that there exists infinitely many disjoint 1-cylinders would be a contradiction).*

We have the following lemma. We need it to state theorem 11.1.

Lemma 11.1. *Under asumptions (III) and (IV), there exists $c_1, c_2, h_1 > 0$ such that for all $y \in X$, for all $n, m \geq 1$ and for $\mathbb{P}$-almost every $\omega \in \Omega$, we have:*

$$\mu_\omega(C_n(y)) \leq c_1 e^{-h_1 n} \tag{17}$$

and

$$\sum_{k=m}^{n} \mu_\omega \left(C_n(y) \cap \sigma^{-k} C_n(y) \right) \leq c_2 e^{-h_1 m} \mu_\omega \left(C_n(y) \right). \tag{18}$$

Proof.
• Let $y \in X$ and $n \geq 1$. We can write the n-cylinder centered in y as

$$C_n(y) = \bigcap_{i=0}^{n-1} \sigma^{-i} \left(C_1(\sigma^i y) \right).$$

Then for all $1 \leq n_0 < n$,

$$C_n(y) \subset \bigcap_{i=0}^{\left\lfloor \frac{n-1}{n_0} \right\rfloor} \sigma^{-i n_0} \left(C_1(\sigma^{i n_0} y) \right).$$

Now use (III) to get for almost every ω

$$\mu_\omega \left(C_n(y) \right) \leq s \left(\left[1 + \psi(n_0 - 1) \right] s \right)^{\left\lfloor \frac{n-1}{n_0} \right\rfloor},$$

where $s = \sup_{\omega \in \Omega, x \in X} \mu_\omega(C_1(x))$. Since ψ is summable and $s < 1$, there exists n_0 such that

$$\left[1 + \psi(n_0 - 1) \right] s < 1,$$

and equation (17) is proved for n big enough. Changing the constant c_1 allows to use the inequality for all $n \geq 1$.
• For the second point we use the fact that for $1 \leq k \leq n$

$$\sigma^{-k} C_n(y) \subset \{ x \in X | \; x_n = y_{n-k}, \ldots, x_{n+k-1} = y_{n-1} \} = \sigma^{-n} A_k^n,$$

where $A_k^n := \{ x \in X | \; x_0 = y_{n-k}, \ldots, x_{k-1} = y_{n-1} \}$. Then for almost all ω

$$\mu_\omega \left(C_n(y) \cap \sigma^{-k} C_n(y) \right) \leq \mu_\omega \left(C_n(y) \cap \sigma^{-n} A_k^n \right),$$

and by (III)

$$\mu_\omega \left(C_n(y) \cap \sigma^{-k} C_n(y) \right) \leq \left[1 + \psi(0) \right] \mu_\omega \left(C_n(y) \right) \mu_{\theta^n \omega} \left(A_k^n \right).$$

Equation (17) then gives

$$\mu_\omega \left(C_n(y) \cap \sigma^{-k} C_n(y) \right) \leq \left[1 + \psi(0) \right] c_1 \mu_\omega \left(C_n(y) \right) e^{-h_1 k}.$$

Using the fact that for all $n \geq 1$ we have

$$\sum_{k=m}^{n} e^{-h_1 k} \leq \sum_{k=m}^{+\infty} e^{-h_1 k} = \frac{1}{1 - e^{-h_1}} e^{-h_1 m},$$

we get equation (18). $\square$

For $A \in \mathscr{X}$ define the hitting time $\tau_A : X \to \mathbb{N} \cup \{+\infty\}$ as

$$\tau_A(x) = \inf \{n \geq 1 | \ \sigma^n x \in A\} .$$

We can now state the hitting time statistics law for the measures μ_ω. We follow the proof of [7].

Theorem 11.1. *Suppose assumptions (I) to (IV) are true and that there exists $q > 2\frac{h_1}{h_0}$ such that $\psi(g)g^q \to 0$ as g goes to infinity. Then for all $z \in X$, and for almost every $\omega \in \Omega$,*

- *If z is a periodic point of period $p \geq 1$ and $\Theta := \lim_{n \to +\infty} \frac{\mu\left(C_n(z) \backslash C_{n+p}(z)\right)}{\mu\left(C_n(z)\right)}$ exists, then we have for all $t > 0$*

$$\lim_{n \to +\infty} \mu_\omega \left(\tau_{C_n(z)} \geq \frac{t}{\mu(C_n(z))} \right) = e^{-\Theta t}.$$

- *If z is not periodic, then for all $t > 0$*

$$\lim_{n \to +\infty} \mu_\omega \left(\tau_{C_n(z)} \geq \frac{t}{\mu(C_n(z))} \right) = e^{-t}.$$

Remark 11.2. *Note that theorem 11.1 doesn't only deal with codings as in example of section 11.1. It also aplies to symbolic systems which are not codings of "concrete" dynamics.*

11.3 Proof of Theorem 11.1

We now prove theorem 11.1. Note that the proof is simpler than the proof of theorem 9.1 in a theoretical point of view - this can be explained by the fact that we are dealing with cylinders and not balls. However, the two proofs use the same kind of arguments.
Suppose that the assumptions of theorem 11.1 are satisfied. We are going to work in the set

$$\Omega' = \left\{ \omega \in \Omega | \ \forall A \in \mathscr{X}, \ \forall i \geq 1, \ \mu_\omega(\sigma^{-i} A) = \mu_{\theta^i \omega}(A) \right\}$$

which has full probability by hypothesis.
For $z \in X$, $A, A' \in \mathscr{X}$ such that $z \in A' \subset A$, define

$$\delta_{z,\omega}(A, A') = \sup_{j \geq p} \left| \mu_\omega(A')\mu_\omega(\{\tau_A > j\}) - \mu_\omega(A \cap \{\tau_A > j\}) \right|,$$

where $p = p(z)$ is the period of z if z is periodic, and $p = 0$ otherwise.

Lemma 11.2. *For $z \in X$, $A, A' \in \mathscr{X}$ such that $z \in A'$, set $p = p(z)$. For almost every $\omega \in \Omega$, we have for $k \geq p$*

$$\left| \mu_\omega\left(\{\tau_A > k\}\right) - \mu_{\theta^{k-p}\omega}\left(\{\tau_A > p\}\right) \prod_{i=1}^{k-p} \left(1 - \mu_{\theta^i \omega}(A')\right) \right| \leq \sum_{i=1}^{k-p} \delta_{\theta^i \omega}(A, A') \prod_{j=1}^{i-1} \left(1 - \mu_{\theta^j \omega}(A')\right).$$

Proof. For $\omega \in \Omega'$, we have by the same calculus as in (6) (in Lemma 9.2), for $i \geq p$,

$$\mu_\omega(\{\tau_A > i+1\}) = \mu_{\theta\omega}(\{\tau_A > i\}) - \mu_{\theta\omega}(A \cap \{\tau_A > i\}),$$

therefore

$$\left| \mu_\omega\left(\{\tau_A > i+1\}\right) - \left(1 - \mu_{\theta\omega}(A')\right) \mu_{\theta\omega}(\{\tau_A > i\}) \right| \leq \delta_{\theta\omega}(A, A').$$

An induction of the previous inequality gives the desired result. $\square$

From now on, $t>0$**,** $z\in X$ **and** $p=p(z)$ **are fixed.** The proof of theorem 11.1 consists in showing that for $A_n=C_n(y)$, $A'_n=\begin{cases} C_n(z)\backslash C_{n+p}(z) \text{ if } z \text{ is periodic} \\ C_n(z) \text{ otherwise}\end{cases}$ and $k_n=\left\lfloor \frac{t}{\mu_\omega(C_n(y))}\right\rfloor$, the term

$$\mu_{\theta^{k_n-p}\omega}\left(\{\tau_{A_n}>p\}\right)\prod_{i=1}^{k_n-p}\left(1-\mu_{\theta^i\omega}(A'_n)\right)$$

goes $\mathbb{P}$-almost surely to $e^{-\Theta t}$ as n goes to infinity if $\Theta=\lim_n \dfrac{\mu(A'_n)}{\mu(A_n)}$ exists, and that the error term

$$\sum_{i=1}^{k_n-p}\delta_{\theta^i\omega}(A_n,A'_n)\prod_{j=1}^{i-1}\left(1-\mu_{\theta^j\omega}(A'_n)\right)\left(\leq \sum_{i=1}^{k_n-p}\delta_{\theta^i\omega}(A_n,A'_n)\right)$$

goes to zero $\mathbb{P}$-almost surely.

Note that if z is not periodic, $A_n=A'_n$ and $\Theta=1$ exists.

Let $M_n(\omega)=\displaystyle\sum_{i=1}^{k_n}\mu_{\theta^i\omega}(A'_n)$. By hypothesis, $M_n:\Omega\to\Omega$ is a random variable, with expectation $\mathbb{E}(M_n)=k_n\mu(A'_n)$ by the fact that θ is $\mathbb{P}$-preserving. Note that if Θ exists, we have

$$\lim_{n\to+\infty}\mathbb{E}(M_n)=\Theta t.$$

The next lemma shows that $M_n-\mathbb{E}(M_n)\to 0$ $\mathbb{P}$-almost surely. Lemma 11.4 will then provide

$$\mu_{\theta^{k_n-p}\omega}\left(\{\tau_{A_n}>p\}\right)\prod_{i=1}^{k_n-p}\left(1-\mu_{\theta^i\omega}(A'_n)\right)\to e^{-M_n}.$$

Lemma 11.3. *If the limit* $\Theta=\lim_n \dfrac{\mu(A'_n)}{\mu(A_n)}$ *exists, then we have*

$$M_n\to\Theta t$$

as n goes to infinity.

Proof. The proof consists in showing that $\displaystyle\sum_{n=1}^{+\infty}\mathrm{Var}(M_n)<+\infty$, where $\mathrm{Var}(M_n)=\mathbb{E}\left((M_n-\mathbb{E}(M_n))^2\right)$.
Then for $\epsilon>0$ we have by Chebyshev's inequality

$$\mathbb{P}\left(|M_n-\mathbb{E}(M_n)|\geq\epsilon\right)\leq\frac{\mathrm{Var}(M_n)}{\epsilon^2}.$$

Summing these inequalities leads to

$$\sum_{n=1}^{+\infty}\mathbb{P}\left(|M_n-\mathbb{E}(M_n)|\geq\epsilon\right)<+\infty,$$

and Borel-Cantelli lemma gives

$$M_n-\mathbb{E}(M_n)\to 0\ \ \mathbb{P}-\text{almost surely}.$$

We now show that $\displaystyle\sum_{n=1}^{+\infty}\mathrm{Var}(M_n)<+\infty$.
We estimate the second moment of M_n.

$$\mathbb{E}(M_n^2)=\sum_{i,j=1}^{k_n}\int_\Omega \mu_{\theta^i\omega}(A'_n)\mu_{\theta^j\omega}(A'_n)d\mathbb{P}(\omega).$$

Let $m \geq 0$. Equation (17) gives

$$\begin{aligned}
\sum_{|i-j|<m} \int_\Omega \mu_{\theta^i\omega}(A'_n)\mu_{\theta^j\omega}(A'_n)d\mathbb{P}(\omega) &\leq \sum_{|i-j|<m} \int_\Omega \mu_{\theta^i\omega}(A_n)\mu_{\theta^j\omega}(A_n)d\mathbb{P}(\omega) \\
&\leq \sum_{|i-j|<m} c_1 e^{-h_1 n} \int_\Omega \mu_{\theta^i\omega}(A_n)d\mathbb{P}(\omega) \\
&\leq \sum_{|i-j|<m} c_1 e^{-h_1 n}\mu(A_n),
\end{aligned}$$

by definition of μ and the fact that θ is $\mathbb{P}$-preserving.
Thus for n big enough we have $k_n \geq 1$ and

$$\sum_{|i-j|<m} \int_\Omega \mu_{\theta^i\omega}(A'_n)\mu_{\theta^j\omega}(A'_n)d\mathbb{P}(\omega) \leq 2mc_1e^{-h_1 n}\mu(A_n) \leq 2mc_1e^{-h_1 n}k_n\mu(A_n),$$

which gives

$$\sum_{|i-j|<m} \int_\Omega \mu_{\theta^i\omega}(A'_n)\mu_{\theta^j\omega}(A'_n)d\mathbb{P}(\omega) \leq 2mtc_1e^{-h_1 n}. \tag{19}$$

For the other terms of the sum, we have for $n+p \leq m \leq k_n$

$$\begin{aligned}
\sum_{|i-j|\geq m} \int_\Omega \mu_{\theta^i\omega}(A'_n)\mu_{\theta^j\omega}(A'_n)d\mathbb{P}(\omega) \leq 2\sum_{i=1}^{k_n}\sum_{j=m+i}^{k_n} \Bigg(&\int_\Omega \psi(m-n-p)\mu_{\theta^i\omega}(A'_n)\mu_{\theta^j\omega}(A'_n)d\mathbb{P} \\
&+ \int_\Omega \mu_{\theta^i\omega}(A'_n \cap \sigma^{-(j-i)}A'_n)d\mathbb{P}\Bigg),
\end{aligned}$$

thaks to equation (III) and by symmetry of the terms of the sum. (Note that we have to be careful when using (III), because $A'_n \in \mathscr{X}_{n+p}$)
We thus get

$$\begin{aligned}
\sum_{|i-j|\geq m} \int_\Omega \mu_{\theta^i\omega}(A'_n)\mu_{\theta^j\omega}(A'_n)d\mathbb{P}(\omega) &\leq 2k_n\sum_{i=1}^{k_n}\mu(A_n)\psi(m-n-p) + 2\sum_{i=1}^{k_n}\sum_{j=i+1}^{k_n}\mu(A'_n\cap\sigma^{-(j-i)}A'_n) \\
&\leq 2tk_n\psi(m-n-p) + 2\sum_{i=1}^{k_n}\sum_{j=i+1}^{k_n}\left(\psi(m-n-p)+\mu(A'_n)^2\right) \\
&\leq 2tk_n\psi(m-n-p) + k_n^2\mu(A'_n)^2 + k_n^2\psi(m-n-p),
\end{aligned}$$

using (I) and $2\sum_{i=1}^{k_n}\sum_{j=i+1}^{k_n} 1 \leq k_n^2$. Then, putting the previous estimates together with the fact that $\mathrm{Var}(M_n) = \mathbb{E}(M_n^2) - (\mathbb{E}(M_n))^2 = \mathbb{E}(M_n^2) - k_n^2\mu(A'_n)^2$ leads to

$$\mathrm{Var}(M_n) \leq 2mtc_1e^{-h_1 n} + 2tk_n\psi(m-n-p) + k_n^2\psi(m-n-p) \tag{20}$$

For $n \geq 0$ we chose $m = m_n = \left\lfloor e^{h_1 n/(1+\epsilon)}\right\rfloor$, where $\epsilon > 0$ can be chosen such that $\sum_{n\in\mathbb{N}} \mathrm{Var}(M_n) < +\infty$.
Indeed, in (20) the first term is summable by definition of m_n and setting $\gamma = 2h_0 - \frac{qh_1}{1+\epsilon}$ (we recall that q is introduced in the statement of theorem 11.1), we have $\gamma < 0$ for $\epsilon > 0$ small enough by the hypothesis on q, and for n big enough, we get by (II)

$$2tk_n\psi(m_n-n-p) \leq k_n^2\psi(m_n-n-p) \leq t^2c_0^2e^{2h_0 n}\frac{\psi(m_n-n-p)}{(m_n-n-p)^q}(m_n-n-p)^q,$$

whence for n big enough, since we have $\psi(m_n - n - p)(m_n - n - p)^q \leq 1$ and $m_n \leq 2(m_n - n - p)$, we have

$$k_n^2 \psi(m_n - n - p) \leq t^2 c_0^2 e^{2h_0 n} \frac{2^q}{m_n^q} \leq \alpha t^2 c_0^2 2^q e^{\gamma n},$$

where α is a constant. (Comes from the integer part in m_n)
Therefore $\sum_{n\in\mathbb{N}} \mathrm{Var}(M_n) < +\infty$ and the result is proved. $\square$

Lemma 11.4. *For almost every $\omega \in \Omega$, we have*

$$\lim_{n\to+\infty} \mu_{\theta^{k_n-p}\omega}(\{\tau_{A_n} > p\}) \prod_{i=1}^{k_n-p} (1 - \mu_{\theta^i\omega}(A'_n)) - e^{-M_n(\omega)} = 0.$$

We will need the following estimate:
For $n \geq 1$, $0 < \epsilon \leq \frac{1}{2}$ and $x_1, \ldots, x_n \in [0, \epsilon]$, we have

$$\exp\left(-(1+2\epsilon)\sum_{i=1}^{n} x_i\right) \leq \prod_{i=1}^{n}(1 - x_i) \leq \exp\left(-(1-2\epsilon)\sum_{i=1}^{n} x_i\right). \tag{21}$$

To prove this, one only has to check it for $n = 1$. Then, the right inequality is a consequence of the inequality $1 - x \leq e^{-x}$ and we get the left one by studying the function $x \mapsto e^{-(1+2\epsilon)x} - (1 - x)$.

Proof of Lemma 11.4. Equation (17) gives that for all $\epsilon > 0$ there exists $n \geq 0$ such that for almost every $\omega \in \Omega$, we have for all $1 \leq i \leq k_n$, $\mu_{\theta^i\omega}(A'_n) \in [0, \epsilon]$. Since by lemma 11.3 the sequence (M_n) is almost surely bounded, using (21) we get

$$\prod_{i=1}^{k_n} (1 - \mu_{\theta^i\omega}(A'_n)) - e^{-M_n(\omega)} \to 0$$

for almost every ω.
We thus just have to prove that as n goes to infinity, we have almost surely

$$\mu_{\theta^{k_n-p}\omega}(\{\tau_{A_n} > p\}) \prod_{i=1}^{k_n-p} (1 - \mu_{\theta^i\omega}(A'_n)) - \prod_{i=1}^{k_n} (1 - \mu_{\theta^i\omega}(A'_n)) \to 0$$

as n goes to infinity.
Note that

$$\left| \mu_{\theta^{k_n-p}\omega}(\{\tau_{A_n} > p\}) \prod_{i=1}^{k_n-p} (1 - \mu_{\theta^i\omega}(A'_n)) - \prod_{i=1}^{k_n}(1 - \mu_{\theta^i\omega}(A'_n)\right| \leq \left| \mu_{\theta^{k_n-p}\omega}(\{\tau_{A_n} > p\}) - \prod_{i=k_n-p+1}^{k_n} (1 - \mu_{\theta^i\omega}(A'_n)\right|,$$

and the two terms of the previous difference go to 1: For the first term, we have

$$\begin{aligned} \mu_{\theta^{k_n-p}\omega}(\{\tau_{A_n} > p\}) &= 1 - \mu_{\theta^{k_n-p}\omega}(\{\tau_{A_n} \leq p\}) \\ &\geq 1 - \sum_{i=1}^{p} \mu_{\theta^{k_n-p+i}\omega}(\sigma^{-i}A_n) \\ &\geq 1 - pc_1 e^{-h_1 n}, \end{aligned}$$

using equation (17). Finally equation (17) and the fact that $A'_n \subset A_n$ gives

$$\prod_{i=k_n-p+1}^{k_n} (1 - \mu_{\theta^i\omega}(A'_n)) \geq \left(1 - c_1 e^{-h_1 n}\right)^p,$$

and the almost everywhere convergence result is established. $\square$

So far, lemmas 11.2, 11.3 and 11.4 give the estimate

$$
\begin{aligned}
\left|\mu_\omega(\tau_{A_n} > k_n) - e^{-\Theta t}\right| \leq & \sum_{i=1}^{k_n-p} \delta_{\theta^i\omega}(A_n, A_n') \\
& + \left|\mu_{\theta^{k_n-p}\omega}(\{\tau_{A_n} > p\}) \prod_{i=1}^{k_n-p} \left(1-\mu_{\theta^i\omega}(A_n')\right) - e^{-M_n(\omega)}\right| \\
& + \left|e^{-M_n(\omega)} - e^{-\Theta t}\right|,
\end{aligned}
$$

where the two last term go $\mathbb{P}$-almost surely to zero.
We now only have to prove that the first term goes $\mathbb{P}$-almost surely to zero.

Let $k \geq p$ and $g \leq k$ be integers. For all $A', A \in \mathscr{X}$ with $A' \subset A$, for $\omega \in \Omega$, let

$$
\begin{aligned}
G_{k,g}(\omega) &= \sum_{i=1}^{k} \mu_{\theta^i\omega}\left(A' \cap \{\tau_A \leq g\}\right) \\
H_{k,g}(\omega) &= \sum_{i=1}^{k} \sup_{j \geq p} \left|\mu_{\theta^i\omega}\left(A' \cap \sigma^{-g}\{\tau_A > j\}\right) - \mu_{\theta^i\omega}(A')\mu_{\theta^{i+g}\omega}(\{\tau_A > j\})\right| \\
K_{k,g}(\omega) &= \sum_{i=1}^{k} \mu_{\theta^i\omega}(A')\mu_{\theta^i\omega}(\{\tau_A \leq g\}).
\end{aligned}
$$

The following lemma shows that we only need to worry about the previous quantities to establish the convergence result.

Lemma 11.5. *With the previous notations, for $g \geq n+p$, we have*

$$
\sum_{i=1}^{k-p} \delta_{\theta^i\omega}(A, A') \leq G_{k,g}(\omega) + H_{k,g}(\omega) + K_{k,g}(\omega).
$$

A proof of lemma 11.5 can be found in [7].

For $n \in \mathbb{N}$ we take $A = A_n$, $A' = A_n'$, $k = k_n$ and $g = e^{h_1 n/2}$. We now write for $\omega \in \Omega$ $G_n(\omega) = G_{k,g}(\omega)$, $H_n(\omega) = H_{k,g}(\omega)$ and $K_n(\omega) = K_{k,g}(\omega)$.

Lemma 11.6. *For $\mathbb{P}$-almost every ω, we have $\lim_{n\to+\infty} H_n(\omega) = \lim_{n\to+\infty} K_n(\omega) = 0$.*

Proof. Since $A_n' \in \mathscr{X}_{n+p}$, (III) gives for almost every ω

$$
H_n(\omega) \leq \sum_{i=1}^{k_n} \psi(g_n - n - p)\mu_{\theta^i\omega}(A_n') \leq \psi(g_n - n - p)M_n(\omega)
$$

and since (M_n) is almost surely bounded, $H_n(\omega) \to 0$ almost surely.
Now, using equation (17),

$$
\mu_\omega(\{\tau_{A_n} \leq g_n\}) = \mu_\omega\left(\bigcup_{i=1}^{g_n} \sigma^{-i} A_n\right) \leq \sum_{i=1}^{g_n} \mu_{\theta^i\omega}(A_n) \leq g_n c_1 e^{-h1n},
$$

thus

$$
K_n(\omega) \leq g_n c_1 e^{-h1n} M_n(\omega) \to 0
$$

and the result is proved. □

We now show that G_n goes almost surely to zero. So far the results were independent of the possible periodicity of z. In the following lemma we dinstinguish the aperiodic case from the periodic one. We first split the expression of G_n: For $\omega \in \Omega$ and $n \geq 1$ we have

$$\begin{aligned}
G_n(\omega) &= \sum_{i=1}^{k_n} \mu_{\theta^i\omega}\left(A'_n \cap \{\tau_{A_n} \leq g_n\}\right) \\
&= \sum_{i=1}^{k_n} \mu_{\theta^i\omega}\left(\bigcup_{j=1}^{g_n} A'_n \cap \sigma^{-j} A_n\right) \\
&\leq \sum_{i=1}^{k_n}\sum_{j=1}^{n} \mu_{\theta^i\omega}\left(A'_n \cap \sigma^{-j} A_n\right) + \sum_{i=1}^{k_n}\sum_{j=n+1}^{g_n} \mu_{\theta^i\omega}\left(A'_n \cap \sigma^{-j} A_n\right).
\end{aligned}$$

If z is periodic, the first term is null since $A'_n \cap \sigma^{-j} A_n$ is empty for $1 \leq j \leq n+p-1$, by definition of A'_n.
We can therefore write

$$G_n(\omega) \leq \sum_{i=1}^{k_n}\sum_{j=1}^{n} \mu_{\theta^i\omega}\left(A'_n \cap \sigma^{-j} A_n\right) + \sum_{i=1}^{k_n}\sum_{j=n+p}^{g_n} \mu_{\theta^i\omega}\left(A'_n \cap \sigma^{-j} A_n\right),$$

where the first term is zero is z is periodic.
If z is not periodic, the next lemma shows that the first term goes to zero.

Lemma 11.7. *If z is a non periodic point, we have for almost every ω*

$$\sum_{i=1}^{k_n}\sum_{j=1}^{n} \mu_{\theta^i\omega}\left(A'_n \cap \sigma^{-j} A_n\right) \to 0$$

We are going to use the following result:

Lemma 11.8. *If z is a non periodic point, then there exists a sequence $(p_n)_{n\geq 1}$ of positive integers going to infinity such that for every point x in A_n the first return time of x in A_n at least p_n.*

Proof of Lemma 11.8.
Suppose that there exists an integer $p \geq 1$ and an increasing sequence (u_n) such that for all $n \geq 1$ there exists a point $x \in A_{u_n}$ such that there exists $1 \leq k_n \leq p$ satifying $\sigma^{k_n} x \in A_{u_n}$. Then, there exists some integer $K \leq p$ and an increasing sequence $(n_k)_{k\geq 0}$ such that for all k the set $\left\{x \in A_{u_{n_k}} \mid \tau_{A_{u_{n_k}}}(x) = K\right\}$ is not empty. Let $k_1 \in \mathbb{N}$ such that $n_{k_1} \geq K$, and let $x \in X$ such that $\tau_{A_{u_{n_{k_1}}}}(x) = K$. Since $x \in A_{u_{n_{k_1}}}$ and $\sigma^K x \in A_{u_{n_{k_1}}}$ we get $z_0 = z_K, \ldots, z_{K-1} = z_{2K-1}$. Now taking $k_2 \in \mathbb{N}$ such that $n_{k_2} \geq 2K$, the same argument leads to $(z_0, \ldots, z_{2K-1}) = (z_{K-1}, \ldots, z_{3K-1})$. Using the same argument recursively shows that z is periodic, which is a contradiction. □

Proof of Lemma 11.7. Lemma 11.8 allows us to write

$$\sum_{i=1}^{k_n}\sum_{j=1}^{n} \mu_{\theta^i\omega}\left(A'_n \cap \sigma^{-j} A_n\right) = \sum_{i=1}^{k_n}\sum_{j=p_n}^{n} \mu_{\theta^i\omega}\left(A'_n \cap \sigma^{-j} A_n\right),$$

consequently equation (18) gives

$$\begin{aligned}
\sum_{i=1}^{k_n}\sum_{j=1}^{n} \mu_{\theta^i\omega}\left(A'_n \cap \sigma^{-j} A_n\right) &\leq \sum_{i=1}^{k_n} c_2 e^{-h_1 p_n} \mu_{\theta^i\omega}(A_n) \\
&\leq c_2 e^{-h_1 p_n} M_n(\omega)
\end{aligned}$$

and since $(M_n(\omega))_n$ is almost surely bounded, we get the desired result. □

We can now conclude that (G_n) goes almost surely to zero.

Lemma 11.9. *For every $z \in X$,*

$$\sum_{i=1}^{k_n} \sum_{j=n+p}^{g_n} \mu_{\theta^i\omega}\left(A'_n \cap \sigma^{-j} A_n\right) \to 0$$

almost surely as n goes to infinity. (Recall that $p = 0$ is z is not periodic)

Proof. We use (III) and equation (17) to get

$$\begin{aligned}
\sum_{i=1}^{k_n} \sum_{j=n+p}^{g_n} \mu_{\theta^i\omega}\left(A'_n \cap \sigma^{-j} A_n\right) &\leq \sum_{i=1}^{k_n} \sum_{j=n+p}^{g_n} \left(\mu_{\theta^i\omega}(A'_n)\mu_{\theta^{i+j}\omega}(A_n) + \psi(j-n-p)\mu_{\theta^i\omega}(A'_n)\mu_{\theta^{i+j}\omega}(A_n)\right) \\
&\leq \sum_{i=1}^{k_n} \mu_{\theta^i\omega}(A'_n) \sum_{j=n+p}^{g_n} \mu_{\theta^{i+j}\omega}(A_n)\left(1 + \psi(j-n-p)\right) \\
&\leq M_n(\omega) g_n \left(1 + \|\psi\|_\infty\right) c_1 e^{-h_1 n},
\end{aligned}$$

which goes almost surely to zero as n goes to infinity. $\square$

Theorem 11.1 is thus proved.

Under the conditions of theorem 11.1, we get the following corollary by integrating over Ω and using the Dominated Convergence Theorem.

Corollary 11.1. *Suppose assumptions (I) to (IV) are true and that there exists $q > 2\frac{h_1}{h_0}$ such that $\psi(g)g^q \to 0$ as g goes to infinity. Then for all $z \in X$,*

- *If z is a periodic point of period $p \geq 1$ and $\Theta := \lim_{n\to+\infty} \dfrac{\mu\left(C_n(z) \backslash C_{n+p}(z)\right)}{\mu\left(C_n(z)\right)}$ exists, then we have for all $t > 0$*

$$\lim_{n\to+\infty} \mu\left(\tau_{C_n(z)} \geq \frac{t}{\mu(C_n(z))}\right) = e^{-\Theta t}.$$

- *If z is not periodic, then for all $t > 0$*

$$\lim_{n\to+\infty} \mu\left(\tau_{C_n(z)} \geq \frac{t}{\mu(C_n(z))}\right) = e^{-t}.$$

A Hausdorff measure, Hausdorff dimension

A.1 Hausdorff measure

Let $n \geq 1$ and $A \subset \mathbb{R}^n$. Let $\delta > 0$. We say that $(U_i)_{i\in\mathbb{N}}$ is a δ-cover of A if $A \subset \bigcup_{n\in\mathbb{N}} U_i$ and for all $i \in \mathbb{N}$ we have $\mathrm{diam}(U_i) \leq \delta$, where $\mathrm{diam}(U_i) := \sup_{x,y\in U_i} d(x,y)$.
For $s > 0$, define

$$\mathscr{H}^s_\delta(A) = \inf\left\{\sum_{i\geq 0} \mathrm{diam}(U_i)^s;\ (U_i)_i \text{ is a } \delta\text{-cover of } A\right\}.$$

If $\delta_1 < \delta_2$ note that $\mathscr{H}^s_{\delta_1}(A) \geq \mathscr{H}^s_{\delta_2}(A)$, hence by monotonicity the limit $\mathscr{H}^s(A) := \lim_{\delta\to 0} \mathscr{H}^s_\delta(A)$ exists. This limit is called the **s-dimensional Hausdorff measure** of A.

A.2 Hausdorff dimension

The definition of the Hausdorff dimension is based on the following fact: Let $A \subset \mathbb{R}^n$ and let $0 \leq s < t$. If (U_i) is a δ-cover of A then

$$\sum_{i\geq 0} \mathrm{diam}(U_i)^t = \sum_{i\geq 0} \mathrm{diam}(U_i)^{t-s}\mathrm{diam}(U_i)^s \leq \delta^{t-s}\sum_{i\geq 0} \mathrm{diam}(U_i)^s,$$

so we get for all $\delta \geq 0$

$$\mathscr{H}^t_\delta(A) \leq \delta^{t-s}\mathscr{H}^s_\delta(A).$$

We thus have: If $\mathscr{H}^s(A) < +\infty$, then for any $t > s$, $\mathscr{H}^t(A) = 0$.

We define the **dimension** of A as the unique $s \geq 0$ verifying: for all $t < s$, $\mathscr{H}^t(A) = +\infty$ and for all $t > s$, $\mathscr{H}^t(A) = 0$.
We write $s = \dim_H(A)$.

B Product σ-algebra

Let $(E_n)_{n\in\mathbb{N}}$ be a sequence of subsets of $\mathbb{N}$. For $i \in \mathbb{N}$ let

$$\begin{array}{rccc} \pi_i : & \displaystyle\prod_{n\in\mathbb{N}} E_n & \longrightarrow & E_i \\ & (x_n)_{n\geq 0} & \longmapsto & x_i. \end{array}$$

We define the **product σ-algebra** on $\prod_{n\in\mathbb{N}} E_n$, denoted by $\bigotimes_{n\in\mathbb{N}} \mathscr{E}_n$, as the smallest σ-algebra over $\prod_{n\in\mathbb{N}} E_n$ for which all the π_i are measurable. One can check that the product σ-algebra is generated by the set of cylinders

$$\left\{ \{a_{i_0}\} \times \{a_{i_1}\} \times \cdots \times \left\{a_{i_{n-1}}\right\} \times \prod_{k\in\mathbb{N}\setminus\{i_0,\ldots,i_{n-1}\}} E_k;\ n \geq 1,\ a_{i_k} \in E_{i_k} \right\}.$$

The following theorem is a powerful tool to define probability measures on product spaces. Let us first give a definition.

Definition B.1. *Let $(E_n, \mathscr{E}_n)_{n\in\mathbb{N}}$ be a sequence of measurable spaces. We say that $(\mu_n)_{n\in\mathbb{N}}$ is a* ***projective sequence of probabilities on*** $\prod E_n$ *if:*

- *For all $n \geq 0$, μ_n is a probability measure on $\left(\prod_{i=0}^{n} E_i, \bigotimes_{i=0}^{n} \mathscr{E}_i\right)$, where $\bigotimes_{i=0}^{n} \mathscr{E}_i$ is the product σ-algebra over $\prod_{i=0}^{n} E_i$.*
- *For all $n \in \mathbb{N}$ and $A \in \bigotimes_{i=0}^{n} \mathscr{E}$, we have*

$$\mu_{n+1}(A \times E_{n+1}) = \mu_n(A).$$

We don't state Daniell theorem in its whole generality but in the case where the E_n are subsets of $\mathbb{N}$.

Theorem B.1. *(Daniell)*
Let $(E_n)_{n\in\mathbb{N}}$ be a sequence of subsets of $\mathbb{N}$ endowed with their power set $\mathscr{P}(E_n)$ σ-algebra. Let $(\mu_n)_{n\in\mathbb{N}}$ be a projective sequence of probabilities on $\prod_{n\in\mathbb{N}} E_n$.
There exists a probability measure $\mathbb{P}$ on $\left(\prod_{n\in\mathbb{N}} E_n, \bigotimes_{n\in\mathbb{N}} \mathscr{E}_n\right)$ such that $\mathbb{P}$ extends all the μ_n:

$$\textit{For all } n \in \mathbb{N},\ \textit{for all } A \in \bigotimes_{i=0}^{n} \mathscr{E}_i,\ \mathbb{P}(A \times \prod_{i>n} E_i) = \mu_n(A).$$

Daniell theorem is a consequence of Caratheodory's extension theorem. The proof can be found in [1].

References

[1] Bailleul, I. (2007). *Advanced Probability*
https://perso.univ-rennes1.fr/ismael.bailleul/files/AdvancedProbability.pdf

[2] Chartier, P. (2014). *Introduction to the theory of Ordinary Differential Equations*, chapter 5
http://www.irisa.fr/ipso/fichiers/cours5.pdf

[3] Dajani, K. and Dirksin, S. (2008). *A Simple Introduction to Ergodic Theory*
http://www.staff.science.uu.nl/%7Ekraai101/lecturenotes2009.pdf

[4] Falconer, K. (2003). *Fractal Geometry*, University of St Andrews, UK: Wiley

[5] Galves, A. and Schmitt, B. (1997). Inequalities for hitting times in mixing dynamical systems, *Random & computational dynamics*, 5(4), 337-348

[6] Kronheime, P.B. (2010). *Lebesgue's differentiation theorem*
http://isites.harvard.edu/fs/docs/icb.topic913185.files/Lebesgue-differentiation.pdf

[7] Rousseau, J. and Todd, M. (2015). Hitting Times and Periodicity in Random Dynamics, *J. Stat. Phys.* 161, 131–150

[8] Saussol, B. (2009). An introduction to quantitative Poincaré recurrence in dynamical systems, *Reviews in Mathematical Physics* Vol. 21, No. 8, 949–979