Antonio SASAKI

ENSAE 3rd year

The 3rd year internship
Academic year: 2023-2024

An extension of Kac's lemma for Furstenberg's ergodic multiple recurrences

Instituto de Matemática
Rio de Janeiro, Brazil

Main supervisor: **Jaqueline SIQUEIRA**
from May 27th to September 30th in 2024

An extension of Kac's lemma for
Furstenberg's ergodic multiple recurrences

Antonio Sasaki

Universidade Federal do Rio de Janeiro

A one-page summary

The Poincaré recurrence theorem [P90], a cornerstone of ergodic theory, posits that under a finite invariant measure μ of a function $T\colon X \to X$, the probability of any point in space recurring is one. In simple terms, a positive measure set $\Omega \in X$ will be recurrent when, for μ-almost all points $x \in \Omega$, the sequence of T-iterations returns close to its starting point x infinite times. This result was conjectured by Poincaré and proved by Carathéodory [C19].

This recurrence property shown is intriguing in its elementariness, requiring only that the transformation be μ-invariant, which means the pair T and μ must satisfy $\mu(T^{-1}(\Omega)) = \mu(\Omega)$, for every measurable subset $\Omega \in X$. A pertinent question emerges: is it possible to generalize Poincaré's result while preserving its fundamental assumptions? Furstenberg [F82] showed that recurrence occurs not only for single returns, but also for multiple returns that form arbitrarily long arithmetic progressions. Previously, Kac [K47] had published a significant refinement by not only proving the existence of singular recurrences but also calculating their average time.

Given a positive measure subset $\Omega \in X$, Kac establishes that the expected return time to Ω is inversely proportional to its measure, $\mu(\Omega)$. Our goal is to extend this to multiple recurrences in arithmetic progressions, as introduced by Furstenberg. To achieve this, we focused on discrete-time dynamical systems, particularly Markov chains, a dense branch of stochastic theory.

Original work results and proposals

1. We first defined the concept of Furstenberg recurrence time, which is merely an intuitive generalization of conventional recurrence time for a sequence with k multiple returns. Furthermore, we created a computational algorithm designed to calculate the time of these consecutive recurrences, through combinatorial and probabilistic reasoning.
2. Using Monte Carlo estimators and hypothesis tests, we were able to observe a reliable convergence of average multiple recurrence times and this data empowered us to postulate an extension of the Kac's lemma tailored to short Furstenberg recurrences. These results would be similar to those found by Bergelson et al. [B05] for Khintchine recurrences.
3. As a bonus, while investigating compact systems, we identified a dynamical-topological proof of Euler's formula [E48]. Specifically, we derived this equation from the existence of a *lift* associated with rotation by a right angle. This perspective naturally suggests a direct generalization to a wider class of algebras beyond the complex numbers.

Une extension du lemme de Kac pour les récurrences ergodiques multiples de Furstenberg

Antonio Sasaki

Universidade Federal do Rio de Janeiro

La note de synthèse

Le théorème de récurrence de Poincaré [P90], pilier de la théorie ergodique, affirme que sous une mesure invariante finie μ d'une fonction $T\colon X \to X$, la probabilité qu'un point revienne est de un. En termes simples, un ensemble de mesure positive $\Omega \in X$ sera récurrent si, pour μ-presque tous les points $x \in \Omega$, la suite des itérations de T repasse infiniment près de x. Ce résultat fut conjecturé par Poincaré et prouvé par Carathéodory [C19].

Cette récurrence, fascinante par sa simplicité, exige seulement que la transformation soit μ-invariante, avec $\mu(T^{-1}(\Omega)) = \mu(\Omega)$ pour tout sous-ensemble $\Omega \in X$. Peut-on alors généraliser ce résultat en conservant ces hypothèses ? Furstenberg [F82] a montré que la récurrence concerne aussi des retours multiples formant des progressions arithmétiques. Kac [K47], quant à lui, avait raffiné ces travaux en calculant le temps moyen de telles récurrences singulières.

Kac a montré que le temps de retour moyen vers un sous-ensemble de mesure positive $\Omega \in X$ est inversement proportionnel à sa mesure, $\mu(\Omega)$. Nous visons à étendre ce résultat aux récurrences multiples en progressions arithmétiques, introduites par Furstenberg. Pour cela, nous nous sommes concentrés sur les systèmes dynamiques en temps discret, notamment les chaînes de Markov, une branche clé de la théorie des processus stochastiques.

Les résultats et propositions originales

1. Nous avons d'abord défini le concept de temps de récurrence de Furstenberg, qui est simplement une généralisation intuitive du temps de récurrence conventionnel pour k retours multiples. De plus, nous avons créé un algorithme de calcul conçu pour calculer le temps de ces récurrences consécutives, en utilisant un raisonnement combinatoire et probabiliste.
2. En utilisant des estimateurs de Monte-Carlo et des tests d'hypothèses, nous avons observé une convergence fiable des temps moyens de récurrence multiple, suggérant une généralisation du lemme de Kac aux courtes récurrences de Furstenberg. Ces résultats seraient similaires à ceux de Bergelson et al. [B05] pour les récurrences de Khintchine.
3. En complément, lors de l'étude des systèmes compacts, nous avons identifié une preuve dynamico-topologique de la formule d'Euler [E48]. Plus précisément, nous avons déduit cette équation de l'existence d'un *lift* associé à la rotation d'un angle droit. Cette approche suggère naturellement une généralisation directe à une classe plus large d'algèbres au-delà des nombres complexes.

Table of Contents

About the author

Antonio Sasaki is a double-degree student at the *École nationale de la statistique et de l'administration économique* and the *Universidade Federal do Rio de Janeiro*. Between 2022 and 2024, he was an Eiffel Excellence Scholar from the Government of France. In October, Sasaki will begin a doctoral contract in applied mathematics (non-convex optimization) at the *École nationale supérieure des mines de Paris*, member of *Université Paris Sciences et Lettres*.

Academic and personal acknowledgments

I am deeply grateful to God for my life and to my parents, Luciano Sasaki and Tatiana Tadros, without whom none of this would have been possible. I also thank ENSAE Paris for the opportunity to pursue a graduate degree, and extend my appreciation to the professors, administrative staff and internship team, who were essential throughout my journey. I am particularly thankful to Stefanella Boatto and Tania Castro, whose support was fundamental in my mobility and adaptation process. I also express my gratitude to the French government for its financial support over two years. Finally, a heartfelt thank you to my advisors, Jaqueline Siqueira and Alexander Arbieto, for their unwavering guidance during my time at UFRJ.

1 Brief introduction to ergodic theory

In the late 19th century, the French mathematician Henri Poincaré began investigating the motion of planets and comets [P90]. The dynamics of these celestial bodies were governed by systems of equations derived from Newton's law of gravitation [N87]. Poincaré realized that when dealing with three or more celestial bodies, these equations could not be fully solved using traditional analytical methods. Instead of seeking explicit solutions, he introduced a novel approach by focusing on the qualitative behavior of the system in x over time, represented as a sequence of states $\{T^n(x)\}_{n\in\mathbb{N}}$, which bypassed the need for closed-form formulas.

Poincaré further observed that, for almost all initial conditions in the n-body problem [N13], the system's evolution would eventually return arbitrarily close to its starting configuration after a sufficiently long period. This led to the development of the concept of recurrence, a foundational idea in what would later become ergodic theory. Poincaré formalized these ideas in his groundbreaking work on dynamical systems and chaos theory [P92]. He noted that this recurrence property was not limited to celestial mechanics but was also present in a wide range of dynamical systems. But specifically what kind of systems are we considering?

1.1 The assumptions we made in our systems

One of the key assumptions in our dynamics is that $T\colon X \to X$ is μ-invariant. This means that T preserves the measure μ for any subset $\Omega \subseteq X$, such that $\mu(T^{-1}(\Omega)) = \mu(\Omega)$. In other words, no matter how T rearranges the points in X, the measure of any set Ω remains unchanged. This property is crucial because it ensures that T does not alter the size or weight of subsets within the space, maintaining consistency across iterations of the transformation.

It is important to highlight that we use the preimage $T^{-1}(\Omega)$ in the definition of invariance rather than the image $T(\Omega)$ because T is assumed to be measurable. The measurability condition guarantees that if Ω is a measurable subset of X, then $T^{-1}(\Omega)$ will also be a measurable subset, allowing the measure μ to be consistently applied. This definition is made so that T is well-defined, allowing us to analyze its behavior in terms of recurrence and ergodic properties.

We are also considering systems with finite measures, $\mu(X) < \infty$. When $\mu(X) = 1$, we know that μ corresponds to a probability measure. However, for any other finite μ, we can normalize it to a probability measure η by defining $\eta(\Omega) = \mu(\Omega)/\mu(X)$, for any subset $\Omega \subseteq X$. Note that this change clearly maps $\eta(\Omega)$ to the interval $[0, 1]$ as in Kolmogorov's axioms [K50]. Thus, assuming a finite measure means that our analysis is, in effect, a probabilistic study.

1.2 Explaining the concept of Poincaré recurrence

The Poincaré recurrence theorem asserts that for a dynamical system with a finite invariant measure, the probability of any point in the space being recurrent is one. In other words, for μ-almost every initial point $x \in X$, after a sufficiently large number of iterations, $T^{n_k}(x)$ will return arbitrarily close to its starting position x. Therefore, if $T\colon X \to X$ is a measure-preserving transformation, a set $\Omega \subseteq X$ is said to be recurrent if, for μ-almost all points $x \in \Omega$, there exists a sequence of times $\{n_k\}_{k=1}^{\infty}$ such that $\mu(\Omega \cap T^{-n_k}(\Omega)) > 0$ as $k \to \infty$.

Theorem 1.1 (Poincaré recurrence theorem)**.** *Let $T : X \to X$ be μ-invariant and $\Omega \subseteq X$ with $\mu(\Omega) > 0$. Then, for μ-almost every point $x \in \Omega$ there exists $n(x)$ such that $n(x) \geq 1$ and $T^{n(x)}(x) \in \Omega$. Thus, there exists $N \geq 1$ so that we have the equation $\mu(\Omega \cap T^{-N}(\Omega)) > 0$.*

Proof. We can see that $(A_i)_{i\in\mathbb{N}} = \left\{T^{-i}\left(\Omega \cap \bigcap_{n=1}^{\infty} T^{-n}(X \setminus \Omega)\right)\right\}_{i=1}^{\infty} \subseteq X$ are 2-2 disjoint. The trick is to use the μ-invariance hypothesis which implies that $\mu(A_i) = \mu(T^i \circ A_i), \forall i \in \mathbb{N}$

$$\mu\left(\biguplus_{i=1}^{\infty} T^{-i}\left(\Omega \cap \bigcap_{n=1}^{\infty} T^{-n}(X \setminus \Omega)\right)\right) = \sum_{i=1}^{\infty} \mu\left(T^{-i}\left(\Omega \cap \bigcap_{n=1}^{\infty} T^{-n}(X \setminus \Omega)\right)\right)$$

$$\sum_{i=1}^{\infty} \mu\left(\Omega \cap \bigcap_{n=1}^{\infty} T^{-n}(X \setminus \Omega)\right) < \infty \implies \mu\left(\Omega \cap \bigcap_{n=1}^{\infty} T^{-n}(X \setminus \Omega)\right) = 0$$

□

A natural question to consider is whether this statement can be further strengthened while maintaining its minimal assumptions. Essentially, there are two main approaches to extending Poincaré's recurrence theorem. The first, demonstrated by Furstenberg [F82], involves exploring the statistical properties of dynamical systems and establishes the existence of recurrences that form arbitrarily long arithmetic progressions. The second approach, introduced by Khintchine [K34], goes beyond merely showing the existence of recurrences. It proves that the intersection of the moving set $T^{-N}(\Omega)$ with its original set $\Omega \subseteq X$ not only has a positive measure $\mu(\Omega \cap T^{-N}(\Omega)) > 0$, but that this measure is bounded below by a positive limit. Specifically, Khintchine demonstrated that for any $\epsilon > 0$, we have $\mu(\Omega \cap T^{-N}(\Omega)) > \mu(\Omega)^2 - \epsilon$.

Furthermore, Kac [K47] provided a generalization of Poincaré's theorem by calculating the expected time of return to a given set Ω. It states that the average return time to Ω is inversely proportional to its measure. While Kac's lemma is well-established for single Poincaré recurrences, there is currently no analogous result for Furstenberg recurrences. However, there have been efforts to extend Khintchine's theorem to short multiple recurrences, as shown by Bergelson, Host and Kra [B05]. These results, while significant, leave open the question of finding an exact formula for the average return time in the context of arithmetic progressions.

1.3 Generalizing Poincaré from Khintchine's work

First, let us clarify the notations we will use. We consider the space of square-integrable functions, denoted by $L^2(X, \mathcal{B}, \mu)$, where $\mathcal{B}$ is a σ-algebra of measurable subsets of X. This function space can be decomposed into two orthogonal subspaces: $\boldsymbol{H}_{\text{inv}} \oplus \boldsymbol{H}_{\text{erg}}$. Here, $\boldsymbol{H}_{\text{inv}}$ represents the space of all invariant functions, which remain unchanged under the transformation T, while $\boldsymbol{H}_{\text{erg}}$ is the closure of all coboundary functions, essentially, the complement of $\boldsymbol{H}_{\text{inv}}$ in L^2. We will refer to C_T as the composition operator, or Koopman operator [K31], defined by $C_T(\varphi) = \varphi \circ T$. Note that if $\varphi \in L^2(X, \mathcal{B}, \mu)$ is an invariant function, then $C_T(\varphi) = \varphi$.

Next, consider the indicator function $\chi_\Omega \colon X \to \{0,1\}$ for a subset $\Omega \subseteq X$. This function equals 1 for points in Ω and 0 for points outside of Ω. We denote by $(\chi_\Omega)_{\text{inv}}$ the projection of χ_Ω onto $\boldsymbol{H}_{\text{inv}}$, the subspace of almost everywhere invariant functions in $L^2(X, \mathcal{B}, \mu)$. This projection, $(\chi_\Omega)_{\text{inv}}$, represents the component of the indicator function that remains unchanged under the transformation T, which is fundamental for analyzing the invariant properties of Ω.

Definition 1.2 (Syndetic set)**.** *A subset $A \subseteq \mathbb{N}$ is called syndetic if $\mathbb{N}$ can be covered by finitely many translates of A. That is, if there exists a set $\{n_1, \ldots, n_k\} \subseteq \mathbb{N}$ such that $\mathbb{N} = \bigcup_i (A - n_i)$.*

Theorem 1.3 (Khintchine recurrence theorem)**.** *Consider a measure-preserving transformation $T \colon X \to X$ and $\Omega \subseteq X$ with $\mu(\Omega) > 0$. Hence, there exists $N \geq 1$ such that $\forall \epsilon > 0$ we have $\mu\left(\Omega \cap T^{-N}(\Omega)\right) > \mu(\Omega)^2 - \epsilon$. Therefore, $\{n \in \mathbb{N} \colon \mu\left(\Omega \cap T^{-n}(\Omega)\right) > \mu(\Omega)^2 - \epsilon\}$ is syndetic.*

Proof. We will use the mean ergodic theorem [N32] in the indicator function $\chi_\Omega \colon X \to \{0,1\}$, followed by applying the operator $\langle \cdot, \chi_\Omega \rangle$ and using Cauchy–Schwarz inequality afterwards.

$$\left\langle (\chi_\Omega)_{inv}, \chi_\Omega \right\rangle = \left\langle \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} {C_T}^n (\chi_\Omega), \chi_\Omega \right\rangle = \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left\langle {C_T}^n (\chi_\Omega), \chi_\Omega \right\rangle$$

$$= \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \chi_\Omega \circ T^n(x)\, \chi_\Omega(x)\, d\mu = \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu\left(\Omega \cap T^{-n}\Omega\right) \geq \mu(\Omega)^2$$

□

As highlighted in [B05], Furstenberg's theorem, although often seen as a broad extension of Poincaré's recurrence, does not fully generalize Khintchine's result for arithmetic progressions. Instead, the article establishes that it is possible to extend Khintchine's theorem only for short multiple recurrences. More formally, for a measure-preserving $T \colon X \to X$ and a set $\Omega \subseteq X$ of positive measure, the subsets defined by $\left\{n \in \mathbb{N} \colon \mu\left(\bigcap_{l=0}^{k-1} T^{-ln}(\Omega)\right) > \mu(\Omega)^k - \epsilon\right\}$ are syndetic for $k \leq 4$. But what happens when we have longer arithmetic progressions?

Bergelson et al. [B05] also emphasizes that for longer arithmetic progressions, specifically when $k \geq 5$, this extension no longer holds in general. There exists a counterexample, demonstrating that the syndetic property fails for these longer sequences. This distinction between short and long multiple recurrences reveals a surprising behavior: shorter recurrences adhere to a more predictable structure, while longer ones exhibit anomalies that defy typical expectations.

This result is intriguing because it mirrors a well-known phenomenon in Galois theory [G30a]: while polynomial equations of degree four or lower have explicit analytical formulas for their solutions, polynomials of degree five and higher generally do not [G30b]. This theory rigorously explains this limitation by showing that the symmetry groups associated with higher-degree polynomials often lack the necessary structure for constructing such formulas. Similarly, the difference in behavior between both multiple recurrences suggests an inherent complexity in the underlying dynamics that prevents a straightforward generalization.

1.4 Calculating the average Poincaré recurrence time

Let $T : X \to X$ be a measure-preserving transformation of a finite measure space $(X, \mathcal{B}, \mu)$, and let $\Omega \subset X$ be a measurable subset with $\mu(\Omega) > 0$. According to the Poincaré recurrence theorem [P90], for almost every point $x \in \Omega$, the system will return to Ω after some finite time. The recurrence time, denoted by $\rho_\Omega(x) = \inf\{n \in \mathbb{N} : T^n(x) \in \Omega\} < \infty$, is well-defined for almost every x in Ω. However, due to the inherent stochasticity of the dynamics, it is typically impossible to calculate the exact recurrence time deterministically for each point. Nonetheless, we can compute the expected return time, offering a statistical measure of how long it takes, on average, for the system to revisit Ω. This is precisely where Kac [K47] comes into play.

Lemma 1.4 (Kac's lemma)**.** *Let $T : X \to X$ be an ergodic measure-preserving transformation of the finite measure space $(X, \mathcal{B}, \mu)$. Let $\Omega \in \mathcal{B}$ be such that $\mu(\Omega) > 0$, then $\int_\Omega \rho_\Omega \, d\mu = \mu(X)$.*

Proof. For each $n \in \mathbb{N}$ let us define $\Omega_n \subseteq X$ as the set of points of Ω that return to Ω for the first time at the moment $n \in \mathbb{N}$ and let us consider $\tilde{\Omega}_n \subseteq X$ as the set of points that is not in Ω and enter in Ω for the first time also at $n \in \mathbb{N}$. Now, expressing in mathematical terms what we have just conceptualized, $\tilde{\Omega}_n = \{x \in X : x \notin \Omega \ldots, T^{n-1}(x) \notin \Omega \ \wedge \ T^n(x) \in \Omega\} \subseteq X$ and

$$\Omega_n = \left\{x \in \Omega : T(x) \notin \Omega, \ldots, T^{n-1}(x) \notin \Omega \ \wedge \ T^n(x) \in \Omega\right\} = \{x \in \Omega : \rho_\Omega(x) = n\} \subseteq X.$$

These sets are measurable therefore the function ρ_Ω is also measurable. Furthermore, for $n \in \mathbb{N}$ the sets $\Omega_n, \tilde{\Omega}_n \in X$ are pairwise disjoint and the union is X. Note that, as by hypothesis we are working with an ergodic $T : X \to X$, the measure $\mu(\tilde{\Omega}_0)$ is zero. We can conclude that

$$\mu(X) = \sum_{n=1}^{\infty} \left(\mu\left(\Omega_n\right) + \mu(\tilde{\Omega}_n)\right) = \sum_{n=1}^{\infty} \sum_{i=n}^{\infty} \mu\left(\Omega_i\right) = \sum_{n=1}^{\infty} n \, \mu\left(\Omega_n\right) = \int_\Omega \rho_\Omega \, d\mu = \mu(\Omega) \, \mathbb{E}[\rho_\Omega]$$

□

The term $\int_\Omega \rho_\Omega \, d\mu$ represents the integral of the recurrence time function ρ_Ω over the set Ω, capturing the total recurrence time for all points in Ω. To compute the average recurrence time, we normalize this total by the measure of Ω, yielding $\int_\Omega \rho_\Omega / \mu(\Omega) \, d\mu$. This expression gives the expected recurrence time, as $\mu(\Omega)$ plays the role of the weight for the integral, ensuring that we average the recurrence times across all points in Ω according to their measure under μ.

The reason we **cannot** directly calculate the average time of a Furstenberg recurrence using this lemma is that it only accounts for the first return of points to a given subset Ω, not the structured sequence of returns required to complete an arithmetic progression of length $k \geq 3$. Even if we divided the average Furstenberg recurrence time by k, this would not yield the time given by Kac's formula, because during the process of forming the arithmetic progression, there are likely many earlier individual returns to Ω that do not fulfill the stricter requirements of the progression. Thus, a lower bound for the average Furstenberg recurrence time is $k/\mu(\Omega)$.

2 The Furstenberg multiple recurrence theorem

The Furstenberg recurrence can be seen as a profound generalization of the classical Poincaré recurrence [P90]. While Poincaré's theorem guarantees that, for almost every point in a measure-preserving system, there will be a return to a subset after a finite time, Furstenberg's theorem extends this to sequences of returns. Specifically, it ensures that for any positive integer, a point not only returns but forms an entire arithmetic progression of returns to the subset. This deepens our understanding of recurrences by introducing structured, sequential returns rather than isolated events, which has powerful implications for the study of this type of happening.

Theorem 2.1 (Furstenberg recurrence theorem)**.** *Let $T : X \to X$ be a measure-preserving transformation and $\Omega \subseteq X$ with $\mu(\Omega) > 0$. Then, for μ-almost every point $x \in \Omega$, there exists $n(x)$ such that $n(x) \geq 1$ and $T^{n(x)}(x), \ldots, T^{(k-1)n(x)}(x) \in \Omega$ for all integer $k \geq 1$. Therefore, for any integer $k \geq 1$, there exists $N \geq 1$ such that $\mu(\Omega \cap T^{-N}(\Omega) \cap \cdots \cap T^{-(k-1)N}(\Omega)) > 0$.*

Furstenberg's work has had monumental importance in mathematics, particularly in his innovative proof of Szemerédi's theorem [F77], which was originally proved using advanced graph-theoretical methods. This result asserts that any subset of integers with positive density contains arbitrarily long arithmetic progressions [S75]. Furstenberg's approach offered a simpler, more conceptual demonstration by applying the multiple recurrence theory to **compact** and **weakly mixing** systems. This breakthrough illustrated how deep connections between ergodic theory and combinatorics could offer elegant solutions to previously complex problems.

Briefly, compact systems are those associated with topological groups that are sequentially compact, these groups have specific topological properties that are important in ergodic theory. The weakly mixing systems, on the other hand, are dynamical systems where different parts of the system become asymptotically independent under the transformation T. The idea in studying these two cases is that there is a result, called the Furstenberg-Zimmer theorem [Z76], which states that an ergodic action of a compact group on a probability measure space can be decomposed into two components: a weakly mixing component and a compact component.

2.1 Multiple recurrences in quasi-periodic dynamics

In this subsection we will focus on proving Furstenberg's theorem for the simplest scenario: compact systems. We choose to demonstrate the theorem in this setting because the case of weakly mixing systems involves more intricate arguments. Compact systems, in essence, are those dynamical systems associated with topological groups that are sequentially compact. These systems exhibit quasi-periodic behavior, meaning they are a slight extension of purely periodic dynamics, which are relatively trivial and not particularly interesting from an analytical standpoint. One notable example of quasi-periodic dynamics, which we will explore later with our proof of Euler's formula, includes certain types of rotations on the unit circle $\mathbb{S}^1$.

Definition 2.2 (Quasi-periodic systems)**.** *A measure-preserving system $(X, \mathcal{B}, \mu, T)$ is called compact system if $\forall \varphi \in L^2(X, \mathcal{B}, \mu)$ the closure $\overline{\{T^n \varphi\}_{n=0}^{\infty}} \subseteq L^2(X, \mathcal{B}, \mu)$ is a compact set.*

Proposition 2.3. *Let $(X, \mathcal{B},\mu, T)$ be a compact system and $k \in \mathbb{N}$ be any positive integer. Then, for all non-zero functions almost everywhere $\varphi \in L^\infty(X, \mathcal{B}, \mu)$ and assuming that $\varphi \geq 0$,*

$$\liminf_{N\to\infty} \left(\frac{1}{N} \sum_{n=1}^{N} \int \left(\varphi \cdot T^n \, \varphi \cdot T^{2n} \, \varphi \cdots \, T^{(k-2)n} \, \varphi \cdot T^{(k-1)n} \, \varphi \right) d\mu \right) > 0$$

Proof. Firstly, the notation $T^n\varphi$ in the definition is known as push-forward, which means that $T^n\varphi = \varphi(T^{-n})$. Let $\delta_+ = \displaystyle\int \prod_{i=0}^{k-1} \varphi \, d\mu > 0$, and let $\Psi_0, \Psi_1, \ldots, \Psi_{k-1}$ be measurable functions such that $\|\Psi_i - \varphi\|_{L^\infty(X)} < \epsilon$ is essentially bounded for $i = 0, 1, \ldots, k-1 \in \mathbb{N}$. Therefore,

$$\left\| \int \prod_{i=0}^{k-1} \Psi_i \, d\mu - \int \prod_{i=0}^{k-1} \varphi \, d\mu \right\|_{L^\infty(X)} = \left\| \int \left(\sum_{j=0}^{k-1} \prod_{i=0}^{j-1} \Psi_i \, (\Psi_j - \varphi) \prod_{i=j+1}^{k-1} \varphi \right) d\mu \right\|_{L^\infty(X)}$$

$$\left\| \int \prod_{i=0}^{k-1} \Psi_i \, d\mu - \int \prod_{i=0}^{k-1} \varphi \, d\mu \right\|_{L^\infty(X)} \leq \sum_{j=0}^{k-1} \int \prod_{i=0}^{j-1} \Psi_i \, \|\Psi_j - \varphi\|_{L^\infty(X)} \prod_{i=j+1}^{k-1} \varphi \, d\mu$$

Suppose $0 \leq \varphi \leq 1$, $0 \leq \Psi_i \leq 1$ for $i = 0, \ldots, k-1 \in \mathbb{N}$ and $\epsilon \leq \delta_+/k$. Then, we have

$$\left\| \int \prod_{i=0}^{k-1} \Psi_i \, d\mu - \int \prod_{i=0}^{k-1} \varphi \, d\mu \right\| < \int \prod_{i=0}^{k-1} \varphi \, d\mu \implies \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \prod_{i=0}^{k-1} \Psi_{i,n} \, d\mu > 0$$

Due to the $\overline{\{T^n\varphi\}_{n=0}^{\infty}} \subseteq L^2(X, \mathcal{B}, \mu)$ compactness, the set $\{T^n\varphi\}_{n=0}^{\infty}$ is totally bounded. Thus, it is easy to show that $\|T^n\varphi - \varphi\|_{L^\infty(X)} < \dfrac{\epsilon}{k}$ is essentially bounded for a subset $A \ni n$ of $\mathbb{N}$ with positive lower asymptotic density. By the μ-invariance hypothesis, it follows that $\|T^{(j+1)n}\varphi - T^{jn}\varphi\|_{L^\infty(X)} < \dfrac{\epsilon}{k}$ and, after applying k times the triangular inequality it is finally possible to conclude that $\|T^{in}\varphi - \varphi\|_{L^\infty(X)} < \epsilon$ for $i = 0, 1, \ldots, k-1$ as desired. Accordingly,

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \prod_{i=0}^{k-1} T^{in}\varphi \, d\mu = \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \left(\varphi \cdot T^n\varphi \cdots \, T^{(k-2)n}\varphi \cdot T^{(k-1)n}\varphi \right) d\mu > 0$$

□

This topological compactness of transformed function sets $\{T^n\varphi\}_{n=0}^{\infty}$ is pivotal to proving the Furstenberg theorem for these specific systems. The compactness ensures that $\{T^n\varphi\}_{n=0}^{\infty}$ remains bounded and well-behaved, allowing for the approximation of functions with measurable maps while maintaining convergence properties. Without this quasi-periodicity, it would be challenging to control the iterated compositions, making this proof significantly less trivial. Finally, we just need to choose $\varphi = \chi_\Omega$, for any $\mu(\Omega) > 0$, and the result is completely proven.

Theorem 2.4 (Furstenberg recurrence theorem for compact systems)**.** *Consider a compact measure-preserving system $(X, \mathcal{B}, \mu, T)$ and $\Omega \subseteq X$ with $\mu(\Omega) > 0$. Then, for μ-almost every $x \in \Omega$, there exists $n(x)$ such that $n(x) \geq 1$ and $T^{n(x)}(x), \ldots, T^{(k-1)n(x)}(x) \in \Omega$, for all $k \geq 1$. Hence, for any $k \geq 1$, there exists $N \geq 1$ such that $\mu(\Omega \cap T^{-N}(\Omega) \cap \cdots \cap T^{-(k-1)N}(\Omega)) > 0$.*

Proof. By the proposition shown, it is known that if $(X, \mathcal{B}, \mu, T)$ is compact, then for a characteristic function $\chi_\Omega \in L^\infty(X, \mathcal{B}, \mu)$, the following equations are valid for every integer $k \geq 1$.

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \left(\chi_\Omega \cdot T^n \chi_\Omega \cdot T^{2n} \chi_\Omega \cdots T^{(k-2)n} \chi_\Omega \cdot T^{(k-1)n} \chi_\Omega \right) d\mu > 0$$

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int \left(\chi_\Omega \cdot \chi_{T^{-n}(\Omega)} \cdot \chi_{T^{-2n}(\Omega)} \cdots \chi_{T^{-(k-2)n}(\Omega)} \cdot \chi_{T^{-(k-1)n}(\Omega)} \right) d\mu > 0$$

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu\Big(\Omega \cap \cdots \cap T^{-(k-1)n}(\Omega) \Big) = \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu\left(\bigcap_{i=0}^{k-1} T^{-in}(\Omega) \right) > 0$$

□

2.2 An introduction to rotations on the unit circle

Firstly, it is important to recognize that in a purely periodic dynamical system $(X, \mathcal{B}, \mu, T)$, multiple recurrence is trivial. This is because, by definition, we have $T^n = T$ for some integer $n \in \mathbb{Z}$, making recurrences evident. To make periodic systems more interesting, we change this concept to quasi-periodicity. One classical example is a rotation on the circle [A09], where $\mathbb{S}^1 \cong \mathbb{R}/\mathbb{Z}$ is our quotient space and $T(x) = x + \alpha \pmod 1$ denotes a rotation by α on $\mathbb{S}^1$. Let $\pi \colon \mathbb{R} \to \mathbb{S}^1$ be the unit circle map given by $\pi(x) = \cos(2\pi x) + i \sin(2\pi x)$, we can take μ as the Lebesgue measure $\mathrm{Leb}_{\mathbb{S}^1}(\Omega) = \mathrm{Leb}_{\mathbb{R}}(\pi^{-1}(\Omega) \cap [0, 1])$, which is usually known as an angle.

The term quasi-periodic is motivated by the fact that T is truly periodic if, and only if, α is a rational number. When α is irrational, however, the dynamics of T can still be considered almost periodic because α can be approximated by rational numbers, causing the system to exhibit a behavior that is close to periodic. Given a measurable subset Ω with $\mu(\Omega) > 0$, it can be observed that $\int \chi_\Omega(x + y)\, d\mu(x)$ is continuous in y. This continuity implies that for any $\varepsilon > 0$, there exists a δ such that if $|y| < \delta$, then $\mu\left(\Omega \cap (\Omega - y)\right) > \mu(\Omega) - \varepsilon$. Consequently,

$$\mu(\Omega \cap (\Omega - y) \cap (\Omega - 2y) \cap \cdots \cap (\Omega - (k-1)y)) > \mu(\Omega) - k\,\varepsilon$$

This equation above already shows the existence of the desired multiple recurrence. Note that, by choosing $\varepsilon < \mu(\Omega)/k$ and fixing the corresponding δ, we can conveniently define the set $\Sigma_\delta = \{n \geq 1 \colon n\alpha \in (-\delta, \delta) \pmod 1\}$. To conclude, if $n \in \Sigma_\delta$, we obtain the inequality

$$\mu\left(\Omega \cap T^{-n}(\Omega) \cap T^{-2n}(\Omega) \cap \cdots \cap T^{-(k-1)n}(\Omega)\right) > \mu(\Omega) - k\,\varepsilon > 0$$

2.3 A dynamical-topological proof of Euler's formula

Historically, Leonhard Euler is credited with first proving this very famous formula [E48], but his classic work involved power series expansions, since it can easily be deduced just using Taylor's methods and complex numbers. Is it possible to show this demonstration in one line

$$\exp ix = \sum_{n=0}^{\infty} \frac{ix^n}{n!} = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = \cos x + i \sin x, \text{ where } i = \sqrt{-1}.$$

The idea of this subsection is completely different, we will prove the formula using dynamical systems theory. Let $\mathbb{S}^1 \cong \mathbb{R}/2\pi\mathbb{Z}$ be the unit circle, which is a metric space, and $f : \mathbb{S}^1 \to \mathbb{S}^1$ a continuous map. The iteration of f defines an action of a semi-group $\mathbb{Z}^+$ on $\mathbb{S}^1$ and this can be expressed as $f^k = f \circ \cdots \circ f$ for k times. In this context, it is extremely intuitive to realize that the simplest orbit comes from fixed points, $\operatorname{Fix}(f) := \{x \in \mathbb{S}^1 : f(x) = x\}$.

It is important to emphasize that it is now more advantageous to consider $\mathbb{S}^1 \cong \mathbb{R}/2\pi\mathbb{Z}$ rather than $\mathbb{S}^1 \cong \mathbb{R}/\mathbb{Z}$ as in the previous example. This is because $\exp ix = \cos x + i \sin x$ involves angles in radians, and using the quotient $\mathbb{R}/\mathbb{Z}$ would result in $\cos 2\pi x + i \sin 2\pi x$, requiring an additional scaling factor to align with the standard trigonometric representation. By choosing $\mathbb{R}/2\pi\mathbb{Z}$, we ensure that x directly corresponds to the angle in radians, preserving the form of Euler's expression without further adjustments. This normalization also impacts the definition of μ on $\mathbb{S}^1$. Instead of the measure being $\operatorname{Leb}_{\mathbb{S}^1}(\Omega) = \operatorname{Leb}_{\mathbb{R}}(\pi^{-1}(\Omega) \cap [0, 1])$, it now becomes $\operatorname{Leb}_{\mathbb{S}^1}(\Omega) = \operatorname{Leb}_{\mathbb{R}}(\pi^{-1}(\Omega) \cap [0, 2\pi])$, where $\pi \colon \mathbb{R} \to \mathbb{S}^1$ is the canonical projection.

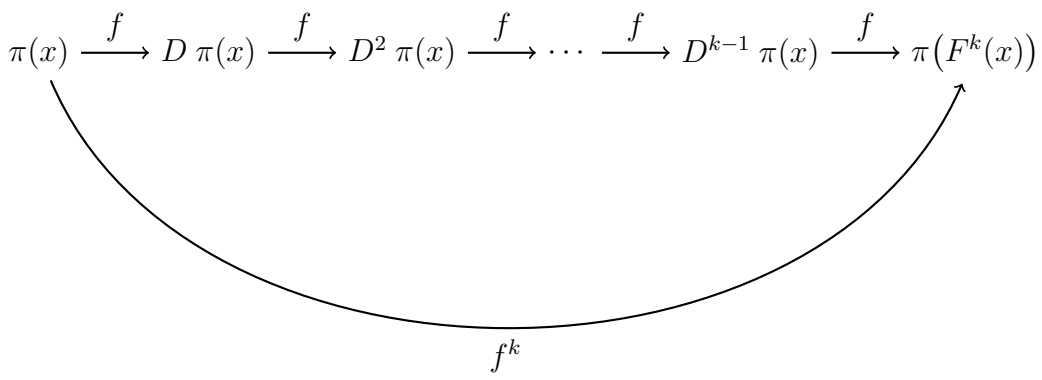
Additionally, by normalizing the quotient to $\mathbb{R}/2\pi\mathbb{Z}$, the relationship between the rotation angle α and periodicity in our dynamics is essentially inverted. Previously, we observed that rotations were periodic if, and only if, the rotation angle α was a rational number, while irrational rotations generated quasi-periodic dynamics. With the new quotient, however, irrational rotations become trivial and rational rotations now exhibit more structured dynamics.

Definition 2.5. *Let $f : \mathbb{S}^1 \to \mathbb{S}^1$ be a continuous transformation of the circle $\mathbb{S}^1 \cong \mathbb{R}/2\pi\mathbb{Z}$. Then a continuous function $F : \mathbb{R} \to \mathbb{R}$ is called a lift of f to $\mathbb{R}$ if we have $\pi \circ F = f \circ \pi$.*

Theorem 2.6 (Euler's formula)**.** *For every $x \in \mathbb{R}$, $\exp ix = \cos x + i \sin x$, where $i = \sqrt{-1}$.*

Proof. Our plan is to build $f^k \in \operatorname{Homeo}(\mathbb{S}^1)$ such that we have $\operatorname{card} \operatorname{Fix}(f) = 0$. Let us define $F^k \colon \mathbb{R} \to \mathbb{R}$ and $f^k \colon \mathbb{S}^1 \to \mathbb{S}^1$ as the rotation by the right angle $F^k(x) = x + k\pi/2 \pmod{2\pi}$ and as the derivative operator $f^k(x) = D^k \operatorname{Re} \pi(x) + iD^k \operatorname{Im} \pi(x)$, for the same fixed $k \in \mathbb{N}$.

If we take the canonical projection $\pi(x) = \cos x + i \sin x$, for $x \in \mathbb{R}$, we obtain the space $\mathbb{S}^1$. Note that starting from $\pi \colon \mathbb{R} \to \mathbb{S}^1$ and using the homeomorphism $f^k \colon \mathbb{S}^1 \to \mathbb{S}^1$, we construct, for any $k \in \mathbb{N}$ previously chosen, a *lift* of f^k given by $F^k \colon \mathbb{R} \to \mathbb{R}$. This geometrically means that we are rotating our point $\pi(x) \in \mathbb{S}^1$ to $\pi\left(F^k(x)\right) \in \mathbb{S}^1$ on the unit circle space.



The existence of this *lift* is precisely what guarantees us the formula we want to prove. As a consequence, using basic differential and integral calculus techniques to explore the equation $\pi \circ F^k = f^k \circ \pi$, and particularly $\pi \circ F = f \circ \pi$ through the choice of $k = 1$, we have

$$D^k\,\pi(x) = \cos\left(x + \frac{k\pi}{2}\ (\mathrm{mod}\ 2\pi)\right) + i\sin\left(x + \frac{k\pi}{2}\ (\mathrm{mod}\ 2\pi)\right) = \big(\cos x + i\sin x\big)\prod_{n=1}^{k} i$$

$$\frac{D\,\pi(x)}{\pi(x)}\,dx = \int_0^x \frac{D^k\,\pi(x)}{\cos x + i\sin x}\,dx = \int_0^x \frac{D^k\,(\cos x + i\sin x)}{\cos x + i\sin x}\,dx = \int_0^x dx\,\prod_{n=1}^{k} i = ix$$

□

Note that this proof works because there is an equivalence in complex space between rotating a unit vector by a right angle or simply taking its orthogonal complement. From the literature on homeomorphisms on $\mathbb{S}^1$, we know that this notion of rotation can be generalized to other functions through a concept called *rotation number* [D18, V16]. Obviously, as we will show below, the rotation number $\rho(f)$ of a transformation f which rotates α must be the same α.

Remark 2.7 (Rotation number of the derivative)**.** *Let $f : \mathbb{S}^1 \to \mathbb{S}^1$ be the previously defined map from the circle to itself. The rotation number $\rho(f)$ of f is a real number that characterizes the average angular displacement of points under iteration of $F : \mathbb{R} \to \mathbb{R}$, which is their lift.*

$$\rho(f^k) = \rho\left(D^k\,\mathrm{Re}\,\pi(x) + iD^k\,\mathrm{Im}\,\pi(x)\right) \coloneqq \lim_{n\to\infty} n^{-1}\left(F^{kn}(x) - x\right) = k\pi/2$$

To conclude, this new dynamical-topological proof has a quality that needs to be mentioned. By replacing the imaginary unit i with algebraic generators J in a general algebra $\mathcal{A}$, we introduce functions $\pi_J(x)$ that analogously encode rotations. The derivatives of these functions naturally yield rotations through the action of J, and integrating them recovers the generalized exponential map. This approach generalizes Euler's identity to arbitrary algebras containing elements whose squares equal $-I$, thereby offering a unified perspective that encompasses Clifford algebras, quantum mechanics (via Pauli matrices), and other related structures.

3 The ergodic theory used to study gas kinetics

Before diving deeper into the mathematical details, we will present a physical motivation for our theory based on gas dynamics. This example, rooted in classical mechanics, provides an intuitive foundation for understanding the recurrence phenomena that Poincaré and later Furstenberg formalized. The behavior of gas particles in a confined space serves as an ideal scenario to visualize why, under certain conditions, systems tend to return to states arbitrarily close to their initial configuration. We will explore this concept in detail using the positions and momenta of gas particles, which evolve according to Hamilton's equations [H33].

Imagine a box filled with gas, made of N identical molecules. Classical mechanics tells us that if we know the positions $q_i = (q_i^1, q_i^2, q_i^3)$ and momenta $p_i = (p_i^1, p_i^2, p_i^3)$ of the i-th molecule for all $i = 1, \dots, N$, then we can determine the positions and momenta of each particle at any given time t deterministically by solving $p_i' = -\partial H/\partial q_i$ and $q_i' = \partial H/\partial p_i$, where $H = H(q_1, \dots, q_N, p_1, \dots, p_N)$ represents the total energy of the system [S09].

In this physical context, it is natural to call $(q, p) = (q_1, \dots, q_N, p_1, \dots, p_N)$ the state of the system. Let X denote the collection of all possible states. If we assume that the total energy is bounded above, then for many reasonable choices of the Hamiltonian H, this is an open bounded subset of $\mathbb{R}^{6N}$. Let $T_t : (q, p) \mapsto (q(t), p(t))$ denote the map that gives the solution of with initial condition $(q(0), p(0))$. If H is sufficiently regular, then our problem has a unique solution for all t and every initial condition. This uniqueness of the solution implies that T_t is a flow and the law of conservation of energy implies that $x \in X \Rightarrow T_t(x) \in X$ for all t.

Suppose the system starts at a certain state $(q(0), p(0))$, will it eventually return to a state close to $(q(0), p(0))$? Given a total energy H, the question seems intractable because N is huge, if $N = 10^{24}$, then $\mathbb{R}^{6N} \approx \mathbb{R}^{10^{25}}$. However, as has been shown previously, Poincaré surprisingly answered this question using his recurrence theorem showing that the answer is yes, from a qualitative and not analytical study [P90]. Recurrence theorems are impressive because they turns a problem that looks intractable into a triviality by simply looking at it from a different viewpoint. The only thing required is the existence of a invariant measure μ on X.

3.1 A mechanical model of gas particle diffusion

The Ehrenfest model [E07] provides an illustrative framework for studying the dynamics of gas particles in two hermetically sealed containers, where no particles can escape the system. Consider a setup with N gas particles distributed between two containers, labeled A and B, such that the total number of particles remains constant throughout the experiment. At each time step, a particle is randomly selected and moved from its current container to the other.

This setup is analogous to observing how gas particles diffuse between two compartments of a closed system. Let $X(t)$ represent the number of gas particles in container A at time t. Suppose the experiment starts with all $2N$ particles in container A. We want to know if, at some point, the system will return to this initial configuration. In other words, Will all $2N$ particles eventually return to container A? If so, **how many steps** would it take on average?

Intuitively, one might expect the system to eventually reach an equilibrium state with approximately N particles in each container, fluctuating slightly around this distribution. It might seem highly unlikely for the system to return to its original state, where all $2N$ particles are in container A. However, Poincaré's theorem [P90] assures us that, given sufficient time, this return is almost certain to occur. To understand this better, we model the problem as a Markov process and define a shift map on an alphabet of $2N$ symbols. Let x_j denote the number of particles in container A after j rounds, and thus the sequence $(x_j)_{j=0}^{\infty}$ keeps track of the number of particles in container A at each time step. Since the number of particles in container A changes by ± 1 at each stage, these sequences form our shift dynamics of finite type, denoted by Σ_{2N}

$$\Sigma_{2N} = \left\{ (x_j)_{j=0}^{\infty} : x_j \in \{0, \dots, 2N\} \ \wedge \ \|x_j - x_{j+1}\| = 1 \ \wedge \ j \in \mathbb{N} \right\}$$

This process can be described by a Markov transition matrix $M = (m_{ij})$, where $m_{i,i+1} = m_{i+1,i} = 1$ for $i = 0, \dots, 2N-1$ and $m_{ij} = 0$ otherwise. Let p_i represent the probability of having i particles in container A. This probability remains constant over time and is given by $p_i = 2^{-2N}\binom{2N}{i}$ for each $i \in \{0, \dots, 2N\}$. If we have i particles in container A, then at the next stage, we either have $i-1$ or $i+1$ particles in container A. The number of particles decreases to $i-1$ if a particle in A is selected, which happens with probability $\frac{i}{2N}$. Similarly, the number of particles increases to $i+1$ if a particle in B is selected, which occurs with probability $\frac{2N-i}{2N}$.

$$\mathbb{P} = \begin{bmatrix} 0 & \frac{2N}{2N} & 0 & 0 & 0 & \cdots \\ \frac{1}{2N} & 0 & \frac{2N-1}{2N} & 0 & 0 & \cdots \\ 0 & \frac{2}{2N} & 0 & \frac{2N-2}{2N} & 0 & \cdots \\ 0 & 0 & \frac{3}{2N} & 0 & \frac{2N-3}{2N} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \cdots \\ \frac{1}{2N} & 0 & 1-\frac{1}{2N} & 0 & 0 & \cdots \\ 0 & \frac{2}{2N} & 0 & 1-\frac{2}{2N} & 0 & \cdots \\ 0 & 0 & \frac{3}{2N} & 0 & 1-\frac{3}{2N} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Note that $\mathbb{P}_{ij} \neq 0$ if and only if $m_{ij} = 1$. Thus, $\mathbb{P}$ and M are compatible, allowing us to define a probability measure μ on Σ_{2N}. The matrix M is irreducible, meaning that the system is ergodic. Let us consider the cylinder set $\Omega = [2N]$, representing the state where all $2N$ particles are in container A. By Poincaré's recurrence theorem, if we start in Ω, we will return to Ω infinitely often. Moreover, by Kac's lemma [K47], the expected recurrence time to Ω is given by $\mathbb{E}[\rho_\Omega] = 1/\mu(\Omega) = 2^{2N}$. Therefore, this model demonstrates that, even though returning to the initial configuration Ω is μ-almost certain, the expected time $\mathbb{E}[\rho_\Omega]$ for this to occur grows exponentially with the number of particles $2N$ considered in the experiment.

Using Stirling's approximation, we find that if the system starts at equilibrium, where both containers have an equal amount, the expected time to return to this state is approximately $\sqrt{\pi N}$. However, if the system begins from an extreme configuration, such as all elements in one box, the expected recurrence time becomes staggeringly large [K79]. For example, if $2N = 100$, the system would only take around 13 seconds to return to equilibrium if particles change containers at a rate of one per second. Yet, to revert back to its original state, with all of them on the same side again, the dynamics would require an estimated 4×10^{22} years. Our goal from now on is to study this system by **creating an algorithm** that calculates the average time of multiple recurrences to the initial state. For this to be computationally viable, we will only consider Furstenberg recurrences that form arithmetic progressions of length $k \leq 3$.

3.2 A computational algorithm for Furstenberg recurrences

To begin this subsection, we will separately estimate the time to Poincaré recurrences, which are a particular case of Furstenberg recurrences for $k = 2$. They are obviously the simplest to compute because no very deep idea is needed, but we should not undervalue the importance of the algorithm for these recurrences as they will be reused in calculations for lengths $k \geq 3$.

Algorithm 1 POINCARÉ RECURRENCE TIME (n)

Container A $\leftarrow$ [ ] ▷ First, let us define the Container A which will start full.
for every $i \in \{1, \ldots, n\}$ **do**
    Container A .append(i) ▷ The model has $n \in \mathbb{N}$ particles, such that $n \geq 1$.
end for
Container B $\leftarrow$ [ ] ▷ Let us define the Container B which will start empty.
$\{1, \ldots, n\} \ni x \leftarrow$ random
Container A .remove(x) ▷ Moving $x \in$ Container A from Container A to Container B.
Container B .append(x)
Time counter $\leftarrow 1$ ▷ Defining the counter to measure the number of steps.
while |Container B| > 0 **do**
    $\{1, \ldots, n\} \ni y \leftarrow$ random ▷ Continuing this process until Container B is empty again.
    if $y \in$ Container A **then**
        Container B .append(y) ▷ Moving $y \in$ Container A from Container A to Container B.
        Container A .remove(y)
    else ▷ Or if the particle y has already been moved previously.
        Container A .append(y)
        Container B.remove(y) ▷ Moving $y \in$ Container B from Container B to Container A.
    end if
    Time counter $\leftarrow$ Time counter $+ 1$ ▷ Defining the increment to calculate the recurrence time.
end while
return Time counter ▷ Returning the time (number of steps) until |Container B| $= 0$.

Now that we are able to construct a finite sequence of random returns for our system, let us consider the specific characteristics which lead these returns to become multiple recurrences of length $k = 3$. The simplest case occurs if in a sequence of recurrences two consecutive returns have the same time. On the other hand, given a sequence $(a_k)_{k\in\{1,\ldots,n\}}$ of Poincaré's return times, we form a multiple recurrence if the accumulated time up to a point in this sequence is equal to the accumulated time in the remaining part of it. As a consequence, we need to create an auxiliary algorithm to identify this non-trivial type of Furstenberg recurrences with $k = 3$.

Algorithm 2 SEQUENCE CHECKER $\left((a_k)_{k\in\{1,\ldots,n\}}\right)$

$n \leftarrow$ length of a given $(a_k)_{k\in\{1,\ldots,n\}}$ ▷ Let us define the length of the imputed time sequence.
for $i \in \{1, \ldots, n\}$ **do**
    First sum $\leftarrow \sum_{j=1}^{i} (a_k)_{k\in\{1,\ldots,n\}}\ [j]$ ▷ We want the two accumulated times to be equal to check.
    Last sum $\leftarrow \sum_{j=i+1}^{n} (a_k)_{k\in\{1,\ldots,n\}}\ [j]$
    if First sum $=$ Last sum **then** ▷ If this happens, then our returns form a multiple recurrence.
        return True
    end if ▷ Otherwise, our returns do not form a multiple recurrence like Furstenberg.
end for
return False ▷ The algorithm returns in a binary manner whether or not multiple recurrence exists.

Our task moving forward will simply involve combining both algorithms and incorporating the trivial case where two consecutive returns occur with the same time, although the latter is exceedingly unlikely given the chaotic behavior of our system. The program will work as follows: we will define a control variable for multiple returns that will pause our looping structure when it detects a Furstenberg recurrence. Unfortunately, this may take a long time because, until this identification happens, the stochastic process will continue evolving indefinitely.

Algorithm 3 FURSTENBERG RECURRENCE TIME (n)

Container A $\leftarrow$ [ ] ▷ First, let us define the Container A which will start full.
for $i \in \{1, \ldots, n\}$ **do**
    Container A .append(i) ▷ The model has $n \in \mathbb{Z}_+$ particles, such that $n \geq 1$.
end for
Container B $\leftarrow$ [ ] ▷ Let us define the Container B which will start empty.
$\{1, \ldots, n\} \ni x \leftarrow$ random
Container B .append(x) ▷ Moving $x \in$ Container A from Container A to Container B.
Container A .remove(x)
Recurrence time $\leftarrow$ [ ] ▷ In this set we will put the time of each recurrence.
Furstenberg recurrence checker $\leftarrow 0$
while Furstenberg recurrence checker $= 0$ **do** ▷ We will stop when we have a multiple recurrence.
    Time counter $\leftarrow 1$
    $\{1, \ldots, n\} \ni y \leftarrow$ random ▷ Defining the counter to measure the number of steps.
    while |Container B| > 0 **do**
        $\{1, \ldots, n\} \ni y \leftarrow$ random ▷ Continuing this process until Container B is empty again.
        if $x \in$ Container A **then**
            Container B .append(y) ▷ Moving $y \in$ Container A from Container A to Container B.
            Container A .remove(y)
        else ▷ Or if the particle y has already been moved previously.
            Container A .append(y)
            Container B .remove(y) ▷ Moving $y \in$ Container B from Container B to Container A.
        end if
        Time counter $\leftarrow$ Time counter $+ 1$ ▷ Defining the increment to calculate the recurrence time.
    end while
    Recurrence time .append(Time counter) ▷ Creating the set of each recurrence time to check.
    if |Recurrence time| ≥ 1 **then**
        if Time counter $=$ Recurrence time $[|\text{Recurrence time}| - 1]$ **then** ▷ Cases of consecutive returns.
            Furstenberg recurrence checker $\leftarrow 1$
        else if SEQUENCE CHECKER (Recurrence time) **then** ▷ Cases checked by the previous algorithm.
            Furstenberg recurrence checker $\leftarrow 1$
        end if ▷ If one of these cases happens, the process stops.
    end if
end while ▷ Ending the repetition structure of our code.
Furstenberg time $\leftarrow \sum_{i=1}^{|\text{Recurrence time}|}$ Recurrence time$[i]$
return Furstenberg time ▷ Returning the desired cumulative recurrence time.

This same combinatorial reasoning of dividing the sequence into $k - 1$ parts and comparing the accumulated time can be generalized to even larger recurrences. Although it is theoretically possible to extend this to handle returns forming arbitrarily long arithmetic progressions, it becomes computationally infeasible for cases where $k \geq 4$. The exponential increase in time to detect such long sequences makes practical computation unmanageable. Thus, we will limit our simulations to the cases where $k \leq 3$ to evaluate the average recurrence time effectively.

3.3 Simulating the expected Furstenberg recurrence time

The algorithm we created allows us to precisely calculate the time of a Poincaré recurrence when $k = 2$ and to detect multiple recurrences for $k = 3$, simulating the dynamical evolution of the model. To investigate how the recurrence time behaves for different configurations, we employ the Monte Carlo method [M49]. This approach is particularly powerful in this context, as it enables us to numerically approximate the expected recurrence time by simulating numerous independent realizations of the process and observing the resulting statistical behavior.

The method is fundamentally based on the concept of repeated random sampling to estimate properties of a random variable. The effectiveness of this is underpinned by the law of large numbers, which guarantees that as the number of simulations increases, the sample mean $\hat{\rho}_\Omega$ converges almost surely to the true expected value $\mathbb{E}[\rho_\Omega]$. Additionally, the central limit theorem provides a measure of how the sample mean $\hat{\rho}_\Omega$ fluctuates around the true mean $\mathbb{E}[\rho_\Omega]$ for a finite number of simulations $n \in \mathbb{N}$. Specifically, the distribution of $\sqrt{n}(\hat{\rho}_\Omega - \mathbb{E}[\rho_\Omega])$ converges to a normal distribution with mean zero and variance σ^2, where σ^2 is the variance of ρ_Ω. Therefore, for large n, the standard error of the estimator is approximately $\sigma/\sqrt{n}$, indicating that the precision of the Monte Carlo estimate improves as n increases.

In our system, which we vary the number of particles $2N$ to modify the initial measure $\mu(\Omega)$, this method becomes invaluable in studying the convergence properties of the recurrence times for different values of $\mu(\Omega)$. By observing how $\hat{\rho}_\Omega$ changes as we alter $\mu(\Omega)$, we can empirically test whether the recurrence times for $k = 2$ and $k = 3$ follow the expected scaling behavior. Particularly, for Poincaré recurrences, Kac's lemma predicts that $\mathbb{E}[\rho_\Omega] = 1/\mu(\Omega)$. However, for $k \geq 3$ recurrences, no analogous formula exists, making the Monte Carlo simulations crucial for exploring the relationship between $\mathbb{E}[\rho_\Omega]$ and $\mu(\Omega)$ in higher dimensions.

First, we conducted a simulation to calculate the average Poincaré recurrence time, varying the number of particles in the system from 1 up to 20. As shown below, the average time increases exponentially as the number of particles grows. To verify this result, we compared it with the map $f(x) = 2^x$ over the same domain cutoff, $x \in \{1, \ldots, 20\}$. This convergence to the base-two exponential function validates our theoretical expectations, confirming that the recurrence times scale precisely as calculated before, in agreement with Kac's theoretical result.

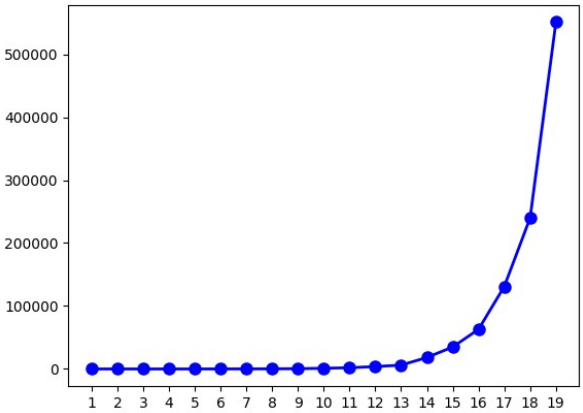


Figure 1: The average time versus number of particles

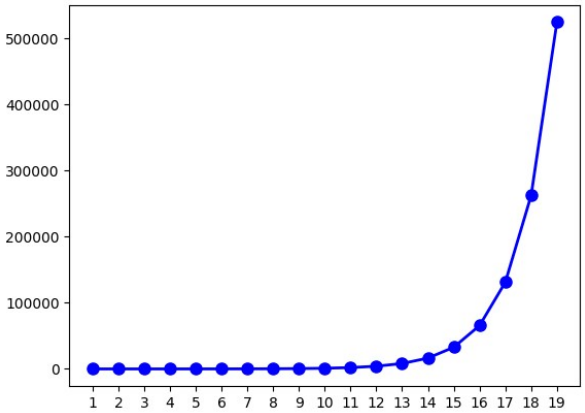


Figure 2: The exponential map $f(x) = 2^x$ versus $x \in \mathbb{N}$

This strong empirical agreement provides good evidence that our computational approach successfully captures the dynamics of Poincaré recurrences in the model. And this is excellent because it enhances the validity of a possible convergence in the case of larger sequences. However, before analyzing the multiple recurrence case, we must precisely define what we consider as time for Furstenberg recurrences of length $k \in \mathbb{N}$. In this context, we define it as the time required to go through the entire arithmetic progression. This innovative definition ensures more consistency as it reflects the complete structure of a k-term sequence of recurrences.

Definition 3.1 (Furstenberg recurrence time)**.** *Consider a subset $\Omega \subseteq X$ and a measure-preserving $T\colon X \to X$, then the Furstenberg recurrence time $\rho_{\Omega,k}$ for a k-term sequence is defined as $\rho_{\Omega,k}(x) = (k-1)\ \inf\{n \in \mathbb{N} : x, T^{-n}(x), T^{-2n}(x), \ldots, T^{-(k-1)n}(x) \in \Omega\} < \infty$.*

Technically, it does not matter whether we consider the time of the first recurrence $T^n(x)$ or the last recurrence $T^{(k-1)n}(x)$, as we have done in our definition. This is because we are assuming that the sequences are arithmetic progressions, meaning that knowing aany element of the temporal sequence allows us to determine all the other times directly. However, our definition was chosen to facilitate the calculation, as it is inspired by our algorithmic approach.

Finally, we performed a simulation to calculate the average Furstenberg recurrence time for 3-term sequences, varying the number of particles in the system from 1 up to 10. As depicted, the average time grows exponentially with the increase in the number of particles. To validate this empirical observation, we compared it against the function $f(x) = 2^{2x}$ over the same domain, $x \in \{1, \ldots, 10\}$. The resulting convergence to the squared base-two exponential function was surprisingly precise, exceeding our initial expectations. It confirmed that the average recurrence time is $\mathbb{E}[\rho_{\Omega,3}] \approx 1/\mu(\Omega)^2$, which appears to be mathematically consistent.

The visual similarity of the generated graphs can be quantitatively confirmed using hypothesis tests. As we will show, given that the Shapiro-Wilk [S65] and Anderson-Darling [A52, A54] tests indicated non-normality of the residuals, parametric tests are not suitable in this context. Instead, non-parametric ones like the Wilcoxon signed-rank [W45], which do not require normality assumptions and works for paired data, were employed to validate the convergence.

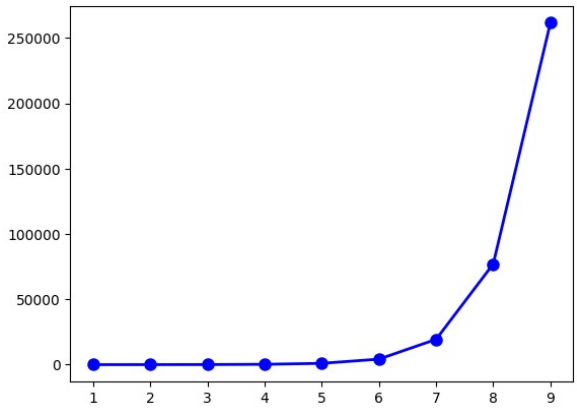


Figure 3: The average time versus number of particles

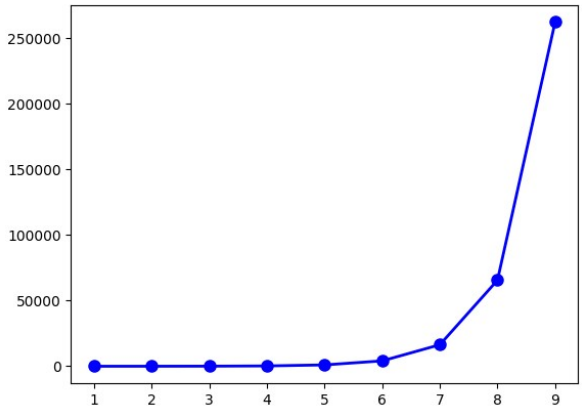


Figure 4: The exponential map $f(x) = 2^{2x}$ versus $x \in \mathbb{N}$

3.4 Analyzing residual normality before hypothesis testing

Before performing hypothesis tests to compare the calculated and conjectured values, it is essential to verify whether the residuals, defined as the difference between these two types of values, follow a normal distribution. Given the computational infeasibility of calculating Furstenberg multiple recurrence times for complex systems, we focus on sample sizes $x < 10$.

For this analysis, we use the Shapiro-Wilk and Anderson-Darling tests, as they are recommended for evaluating normality in small samples. The Shapiro-Wilk test assesses the correlation between the sample values and the expected values under normality. On the other hand, the Anderson-Darling test is more sensitive to deviations in the tails of the distribution, making it more robust than the Shapiro-Wilk test for detecting non-normality in our context.

Test name	**Result**	$\mathbb{P}\big(\geq \lvert t\rvert\big)$
Shapiro-Wilk	0.570	7.153×10^{-5}

Table 1: Shapiro-Wilk test results

Test name	**Result**	$\mathbb{P}\big(\geq \lvert t\rvert\big)$
Anderson-Darling	1.713	6.271×10^{-5}

Table 2: Anderson-Darling test results

The results of both normality tests above indicate that the residuals do not follow a normal distribution. Since our samples are paired, meaning that each value in one sample corresponds to a unique and meaningful value in the other, parametric tests such as the paired t-test would typically be used. However, the lack of normality in the residuals makes these tests inappropriate. Thus, the most suitable alternative is to use a non-parametric test, such as the Wilcoxon signed-rank test, which does not rely on the assumption of normality and is specifically designed for paired data. This test compares the differences between paired observations, making it ideal for assessing whether the medians of the two related distributions differ significantly.

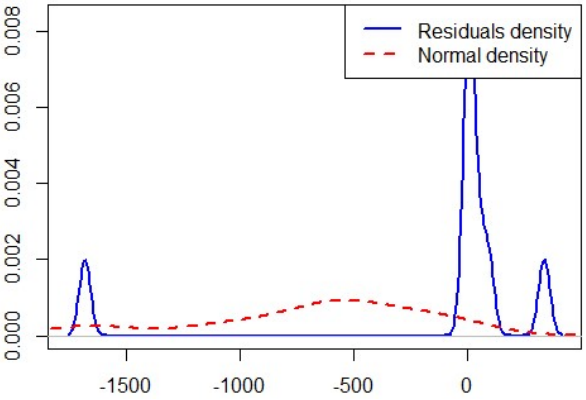


Figure 5: The density plot of our residuals

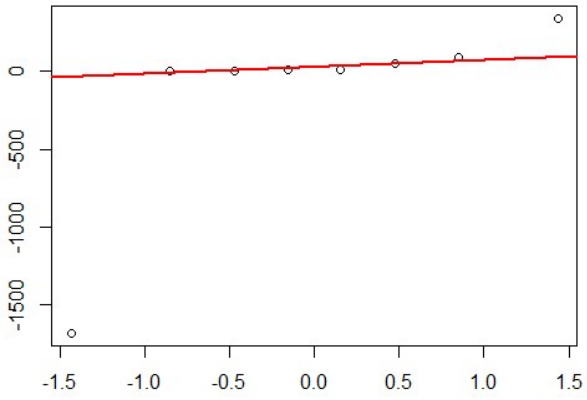


Figure 6: The Q-Q plot of our residuals

The plots were generated to visually confirm the non-normality of the residuals, as indicated by the Shapiro-Wilk and Anderson-Darling test results. The density plot reveals a multimodal distribution with peaks that deviate significantly from the smooth, bell-shaped curve of a Gaussian function, suggesting that the residuals are not centered around zero. Similarly, the Q-Q plot shows a clear departure from the straight reference line, with several points falling far from the line, especially at the tails. These deviations reinforce the test results and are expected, as the imprecision of the estimation tends to increase as our dynamics becomes more complex.

3.5 Creating the hypothesis test for our conjecture

As we said previously, the Wilcoxon test works by calculating the differences between each pair of observations. It then ranks the absolute values of these differences and assigns signs, positive or negative, based on the direction of the difference. The test sums the ranks of positive differences and compares them to the ranks of negative differences. If the sum of the positive ranks is significantly different from the sum of the negative ranks, it suggests that the median of the differences is not zero. This approach ensures that the test is sensitive to shifts in the distribution's median while remaining unaffected by the non-normality of the data.

The result of the test shows a p-value of $0.2719 \gg 0.05$. This clearly indicates that there is insufficient evidence to reject the null hypothesis that the median of the differences between the Monte Carlo estimators and the conjectured values is equal to zero. In other words, there is no significant difference between the samples, suggesting that any discrepancy observed between the empirical and theoretical values can be attributed to chance rather than a systematic change. This finding confirms that, although the residuals are not normal, the Wilcoxon test did not detect a significant shift in the location of the differences between the two sets of values.

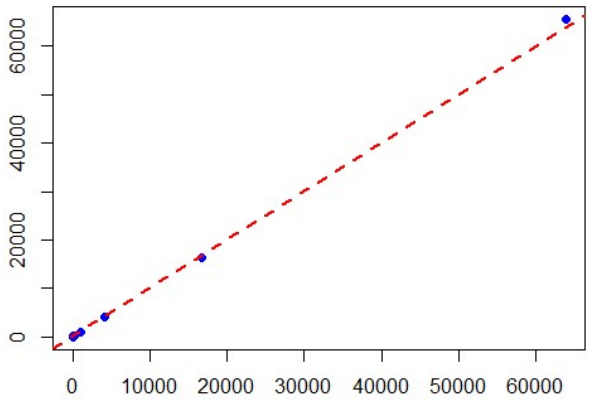


Figure 7: The scatter plot of paired samples

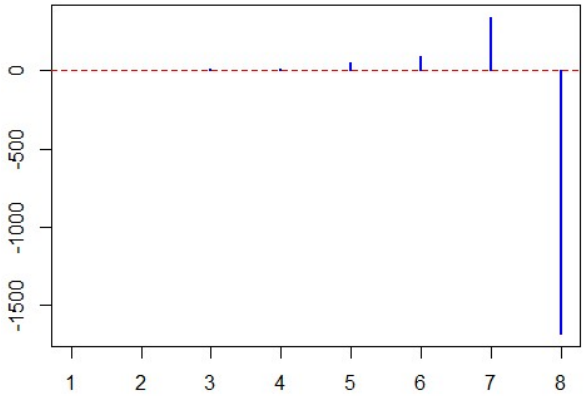


Figure 8: The paired differences plot

The left figure above is comparing the estimated values to the expected values based on the expression we want to prove. The data points align almost perfectly along the reference red line, indicating that the estimates closely match the theoretical values, suggesting a strong linear relationship with minimal deviation. In a similar way, the figure on the right shows the differences between the estimates and the expected values for each pair. Most of the differences are centered around zero, and only one observation, corresponding to the eighth pair, displays a substantial negative deviation, which stands out as a potential outlier. Overall, the distribution does not show a systematic trend favoring one set over the other. Both graphs support the test result and confirm that there is no significant difference between the two sample groups.

The results of the hypothesis tests were **highly satisfactory**, providing strong evidence that our empirical observations are statistically significant and reliable. With this validation, we are now better equipped to formalize our conjecture regarding a potential formula for the average return time of Furstenberg recurrences. This foundation allows us to proceed with more confidence, as the consistency of our findings reinforces the plausibility of extending the classical recurrence theory to encompass multiple recurrences in a rigorous and generalized manner.

4 An extension of Kac's lemma for Furstenberg recurrences

We found it more convenient to divide our conjecture into three levels of weakness, so that each weaker proposal is contained within the stronger one. This structure allows us to provide a clear progression in the formulation of our hypothesis, given that our ability to produce computational evidence is unfortunately restricted. The constraints arise from the limitations of our algorithm, which only allows us to simulate specific types of dynamics and short sequences. By carefully delineating these conjectures, we are able to present our findings in a structured manner while acknowledging the practical boundaries we faced. This organization also serves to emphasize the potential for broader extensions for longer sequences or more complex systems.

4.1 Hierarchical conjecture framework for Furstenberg recurrences

As mentioned, we propose three conjectures of varying strength regarding the distribution of mean Furstenberg recurrence times. The weak conjecture states that the average recurrence time for a 3-term sequence should be $1/\mu^2(\Omega)$. This conjecture is supported by our empirical findings. Building upon this, we present a strong conjecture stating that the average time for any k-term sequence, with $k \leq 4$, is given by $1/\mu^{k-1}(\Omega)$. This is motivated by the observation from Bergelson et al. [B05] that short arithmetic sequences tend to exhibit more structured behavior, suggesting that their recurrence times scale predictably with $\mu(\Omega)$. Finally, our very strong conjecture extends the relationship to all k-term sequences, asserting that the average recurrence time is $1/\mu^{k-1}(\Omega)$ for every $k \in \mathbb{N}$. This conjecture, if true, would provide a full generalization of Kac's lemma to multiple recurrences in any measure-preserving dynamics.

Conjecture 4.1 (Weak conjecture)**.** *Let $T : X \to X$ be an ergodic measure-preserving transformation of the probability space. If we consider $\mu(\Omega) > 0$, then the average Furstenberg recurrence time related to Ω for a k-term sequence will be $1/\mu^{k-1}(\Omega)$ for $k \in \{2, 3\}$.*

Conjecture 4.2 (Strong conjecture)**.** *Let $T : X \to X$ be an ergodic measure-preserving transformation of the probability space. If we consider $\mu(\Omega) > 0$, then the average Furstenberg recurrence time related to Ω for a k-term sequence will be $1/\mu^{k-1}(\Omega)$ for $k \in \{2, 3, 4\}$.*

Conjecture 4.3 (Strongest conjecture)**.** *Let $T : X \to X$ be an ergodic measure-preserving transformation of the probability space. If we consider $\mu(\Omega) > 0$, then the average Furstenberg recurrence time related to Ω for a k-term sequence will be $1/\mu^{k-1}(\Omega)$ for every $k \in \mathbb{N}$.*

Up to this point, our arguments have been primarily statistical, relying on empirical evidence obtained through simulations and hypothesis testing. At the end of this work, we will revisit the definition of weakly mixing systems mentioned earlier in the section on Furstenberg's recurrences, where we highlighted their complementary nature to compact systems. We will show that, for this class of problems, our strongest conjecture formula provides an accurate approximation for the average recurrence time. In mathematical terms, $\mathbb{E}[\rho_{\Omega,k}] \approx 1/\mu^{k-1}(\Omega)$.

4.2 A probabilistic analysis for weakly mixing systems

The conjecture that the expected return time for an arithmetic progression of length $k \in \mathbb{N}$ is proportional to the inverse of $\mu(\Omega)$ raised to the power of k can be understood through a stochastic framework. To begin, recall that for a measure-preserving transformation $T\colon X \to X$ on a probability space and a measurable subset $\Omega \subseteq X$ with positive measure, Kac's lemma asserts that the expected return time to Ω is given by $\mathbb{E}[\rho_{\Omega,2}] = 1/\mu(\Omega)$, by our own notation.

When extending this concept to k-term arithmetic progressions, where returns occur in the steps $\{n, 2n, 3n, \ldots, (k-1)n\}$, the structured dependencies complicate the situation because these events are not independent, as they are influenced by the deterministic T-dynamics. To analyze this, we will use the characteristic function $\chi_\Omega\colon X \to \{0,1\}$. The transformation T is **weakly mixing** if and only if for all $\Psi, \varphi \in L^2(X, \mathcal{B}, \mu)$, and particularly for χ_Ω, the following limit converges to zero. Let us remember that $T^n\varphi = \varphi(T^{-n})$ in our push-forward notation.

$$\lim_{N\to\infty} \sum_{n=1}^{N} \frac{\left(\int \Psi\, T^n\varphi\, d\mu - \int \Psi d\mu \int \varphi d\mu\right)^2}{N} = \lim_{N\to\infty} \sum_{n=1}^{N} \frac{\left(\int \chi_\Omega\, T^n\chi_\Omega\, d\mu - \left(\int \chi_\Omega d\mu\right)^2\right)^2}{N} = 0$$

This means that the correlation between $\chi_\Omega(x)$ and $\chi_\Omega(T^{-n}(x))$ diminishes as n increases, making the events $T^{-n}(x)$ and $x \in \Omega$ asymptotically independent as $n \to \infty$. Now consider k-term recurrences, where we are interested in returns to Ω at times $\{n, 2n, 3n, \ldots, (k-1)n\}$. Let us define the characteristic function for these k-returns as $\chi_{\Omega,k}(x) = \prod_{i=1}^{k-1} \chi_\Omega(T^{-in}(x))$. For $x \in \Omega$, this function takes the value one if $T^{-n}(x), T^{-2n}(x), \ldots, T^{-(k-1)n}(x) \in \Omega$, and zero otherwise. Note that the measure of observing a k-term recurrence at time $(k-1)\,n$ is

$$\mu\left(\bigcap_{i=0}^{k-1} T^{-in}(\Omega)\right) = \int \chi_\Omega(x)\, \chi_\Omega(T^{-n}(x))\, \chi_\Omega(T^{-2n}(x)) \cdots \chi_\Omega(T^{-(k-1)n}(x))\, d\mu > 0$$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left(\int \chi_\Omega(T^{-n}(x))\, \chi_\Omega(T^{-2n}(x)) \cdots \chi_\Omega(T^{-(k-1)n}(x))\, d\mu - \left(\int \chi_\Omega\, d\mu \right)^{k-1} \right)^2 = 0$$

For weakly mixing systems, the joint correlation between these characteristic functions also decays as n increases, leading to the limit above. This implies that for sufficiently large n, the joint probability of observing all k-returns converges to $\mu(\Omega)^k$. While the returns are not strictly independent, the characteristic functions $\chi_\Omega(x), \chi_\Omega(T^n x), \ldots, \chi_\Omega(T^{(k-1)n}x)$ behave as though they were approximately independent, making the product of their expectations a good approximation for the joint expectation, mainly because the $n \in \mathbb{N}$ needed to obtain multiple recurrence are always naturally very large due to the chaotic nature of our ergodic systems.

It is important to emphasize that our choice to analyze weakly mixing systems was not merely a matter of convenience or ease of calculation. If that were the case, we could have selected strongly mixing systems, which exhibit convergence to the product much more robustly. To illustrate, a measure-preserving dynamical system is called **mixing** if, for any two measurable sets $\Omega_1, \Omega_2 \subseteq X$, we have convergence $\lim_{n\to\infty} \mu(T^{-n}(\Omega_1) \cap \Omega_2) = \mu(\Omega_1)\, \mu(\Omega_2)$.

While this property would simplify our analysis, mixing systems are actually quite rare in many classes of dynamical systems. They almost never occur naturally, unlike weakly mixing systems, which are far more prevalent in practice and exhibit rich behavior that makes them ideal candidates for generalizing our results. Moreover, it is possible to show that our general conjecture provides a **reliable approximation** for the average recurrence time of all weakly mixing dynamical systems by utilizing the previous asymptotic equality for $\chi_{\Omega,k}(x)$.

Proof sketch. Let A_n denote the event that $T^n(x) \in \Omega$ for a given $n \in \mathbb{N}$. Since we are in a probability space, it is intuitive to think that the probability of observing a single return is $\mathbb{P}(A_n) = \mu(\Omega)$. As we are assuming that the dynamics starts from $\Omega \subseteq X$ and therefore this happens with probability one, for a k-term arithmetic progression, we just need to analyze the joint probability of $k-1$ events observing the remaining returns at $n, 2n, 3n, \ldots, (k-1)n$, which corresponds to the intersection of events $\mathbb{P}(\bigcap_{i=1}^{k-1} A_{in}) = \mathbb{P}(A_n \cap A_{2n} \cap \ldots \cap A_{(k-1)n})$.

Because T is measure-preserving, these events are not independent, and the probability of a return at $T^{2n}(x)$ depends on whether a return occurred at $T^n(x)$. However, under certain conditions such as weakly mixing or rapid decay of correlations, these dependencies diminish, allowing us to approximate the joint probability using the product of individual probabilities

$$\mathbb{P}(A_n \cap A_{2n} \cap \ldots \cap A_{(k-1)n}) = \int \chi_{\Omega,k}(x)\, d\mu \approx \left(\int \chi_\Omega\, d\mu\right)^{k-1} = \prod_{i=1}^{k-1} \mathbb{P}(A_{in}) = \mu(\Omega)^{k-1}$$

$$\mathbb{E}[\rho_{\Omega,k}] = \frac{1}{\mathbb{P}(A_n \cap A_{2n} \cap \ldots \cap A_{(k-1)n})} = \frac{1}{\int \chi_{\Omega,k}(x)\, d\mu} \approx \frac{1}{\left(\int \chi_\Omega\, d\mu\right)^{k-1}} = \frac{1}{\mu(\Omega)^{k-1}}$$

□

Main conclusions

Overall, we introduced the concept of Furstenberg recurrence time $\rho_{\Omega,k}$ and developed an algorithm to calculate it through stochastic approaches. Using Monte Carlo methods, we were able to estimate the average times for short multiple recurrences and found that they closely matched our theoretical expectations. Statistical hypothesis tests further validated these findings, demonstrating convergence to the values predicted by our weak conjecture.

Moreover, by analyzing the asymptotic properties of weakly mixing systems, we proved that our conjectured formula is mathematically justified as a good approximation of the average Furstenberg recurrence time $\mathbb{E}[\rho_{\Omega,k}]$ for a broader class of dynamical systems. This result lends theoretical support to our strongest conjecture. In addition, our exploration of compact systems motivated us to discover a generalizable dynamical-topological proof of Euler's formula.

Bibliographical references

[C19] Carathéodory, C. (1919). *Über den Wiederkehrsatz von Poincaré*. *Sitzungsberichte der Preußischen Akademie der Wissenschaften zu Berlin, Mathematisch-physikalische Klasse*, 580–584.

[P90] Poincaré, H. (1890). *Sur le problème des trois corps et les équations de la dynamique*. *Acta Mathematica*, Vol. 13, 1–270. Also in *Œuvres*, Vol. 7, 262–478.

[N87] Newton, I. (1687). *Philosophiæ Naturalis Principia Mathematica*. *Londini, Jussu Societatis Regiæ ac Typis Josephi Streater. Prostat apud plures Bibliopolas*.

[N13] Newton, I. (1713). *The General Scholium*. In: *Principia Mathematica*, 2nd Edition. Cambridge University Press.

[P92] Poincaré, H. (1892). *Les méthodes nouvelles de la mécanique céleste*. Gauthier-Villars.

[F82] Katznelson, Y., Furstenberg, H., Ornstein, D. (1982). *The ergodic theoretical proof of Szemerédi's theorem*. *Bulletin of the American Mathematical Society*, Vol. 7, 527–552.

[K47] Kac, M. (1947). *On the notion of recurrence in discrete stochastic processes*. *Bulletin of the American Mathematical Society*, Vol. 53(10), 1002–1010.

[E48] Euler, L. (1748). *Introductio in analysin infinitorum*. Vol. 1. *Lausanne: Marcum-Michaelem Bousquet*.

[K50] Kolmogorov, A. (1950). *Foundations of the theory of probability*. New York, US: Chelsea Publishing Company.

[K31] Koopman, B. O. (1931). *Hamiltonian Systems and Transformation in Hilbert Space*. *Proceedings of the National Academy of Sciences*, Vol. 17(5), 315–318.

[B31] Birkhoff, G. D. (1931). *Proof of the ergodic theorem*. *Proceedings of the National Academy of Sciences*, Vol. 17(12), 656–660.

[N32] von Neumann, J. (1932). *Proof of the Quasi-ergodic Hypothesis*. *Proceedings of the National Academy of Sciences*, Vol. 18(1), 70–82.

[G30a] Galois, É. (1830). *Analyse d'un Mémoire sur la résolution algébrique des équations*. *Bulletin des Sciences Mathématiques*, Vol. XIII, 271.

[G30b] Galois, É. (1830). *Note sur la résolution des équations numériques*. *Bulletin des Sciences Mathématiques*, Vol. XIII, 413.

[F77] Furstenberg, H. (1977). *Ergodic behaviour of diagonal measures and a theorem of Szemerédi on arithmetic progressions*. *Journal of Analysis and Mathematics*, Vol. 31, 204–256.

[Z76] Zimmer, R. J. (1976). *Extension of ergodic group actions*. *Illinois Journal of Mathematics*, Vol. 20(3), 373–409.

[F14] Furstenberg, H. (2014). *Recurrence in ergodic theory and combinatorial number theory*. Vol. 14, Princeton University Press. ISBN: 978-0-691-16414-7.

[A09] Arbieto, A., Matheus, C., Moreira, C. (2009). *The remarkable effectiveness of ergodic theory in number theory Part I: Green-Tao theorem*. *Ensaios Matemáticos*, Vol. 17, 1–71.

[P98] Pollicott, M., Yuri, M. (1998). *Dynamical systems and ergodic theory* (Vol. 40). Cambridge University Press.

[K34] Khintchine, A. (1934). *Eine Verschärfung des Poincaréschen "Wiederkehr-Satzes"*. *Compositio Mathematica*, Vol. 1, 177–179.

[H33] Hamilton, W. R. (1833). *On a general method of expressing the paths of light, & of the planets, by the coefficients of a characteristic function*. Printed by P.D. Hardy.

[S09] Sarig, O. (2009). *Lecture notes on ergodic theory*. Lecture Notes, Penn State University.

[P89] Petersen, K. (1989). *Ergodic Theory*. Vol. 2, Cambridge University Press.

[E07] Ehrenfest, P., Ehrenfest, T. (1907). *Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem*. *Physikalische Zeitschrift*, Vol. 8, 311–314.

[K79] Kelly, F. P. (1979). *The Ehrenfest model*, in *Reversibility and Stochastic Networks*. Wiley, Chichester. ISBN 0-471-27601-4, pp. 17–20.

[S75] Szemerédi, S. (1975). *On sets of integers containing no k elements in arithmetic progression*. *Acta Arithmetica*, Vol. 27, 199–245.

[B05] Bergelson, V., Host, B., Kra, B., Ruzsa, I. (2005). *Multiple recurrence and nilsequences*. *Inventiones mathematicae*, Vol. 160, 261–303.

[W82] Walters, P. (1982). *An Introduction to Ergodic Theory*. *Graduate Texts in Mathematics* (Vol. 79). Springer, New York, NY. ISBN: 978-0-387-95152-2.

[D18] Devaney, R. (2018). *An introduction to chaotic dynamical systems*. CRC Press. ISBN: 978-1-138-60114-1.

[V16] Viana, M., Oliveira, K. (2016). *Foundations of ergodic theory*. Cambridge University Press. ISBN: 978-1-107-10568-1.

[M87] Mañé, R. (1987). *Ergodic Theory and Differentiable Dynamics*. *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics* (Vol. 8). Springer Berlin, Heidelberg. ISBN: 978-3-642-70337-9, 978-3-642-70335-5 (eBook).

[M49] Metropolis, N., Ulam, S. (1949). *The Monte Carlo Method*. *Journal of the American Statistical Association*, Vol. 44(247), 335–341.

[S65] Shapiro, S. S., Wilk, M. B. (1965). *An analysis of variance test for normality (complete samples)*. *Biometrika*, Vol. 52(3–4), 591–611. JSTOR 2333709. MR 205384.

[W45] Wilcoxon, F. (1945). *Individual comparisons by ranking methods*. *Biometrics Bulletin*, Vol. 1(6), 80–83. JSTOR 3001968.

[A52] Anderson, T. W., Darling, D. A. (1952). *Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes*. *Annals of Mathematical Statistics*, Vol. 23(2), 193–212.

[A54] Anderson, T. W., Darling, D. A. (1954). *A Test of Goodness-of-Fit*. *Journal of the American Statistical Association*, Vol. 49(268), 765–769. JSTOR 2281537.